

NNT: 2016SACLS129

THÈSE DE DOCTORAT
de l'Université Paris-Saclay
préparée au CEA-Saclay, Irfu/SPP.

ÉCOLE DOCTORALE N° 576

Particules, Hadrons, Énergie, Noyau, Instrumentation, Imagerie, Cosmos et Simulation
(PHENIICS)

Spécialité du doctorat: Physique des particules

par

Matthias Saimpert

**Mesure de la section efficace de production de
paires de photons isolés dans l'expérience
ATLAS au LHC et étude des couplages à
quatre photons**

Soutenue le 27 Juin 2016 à Orsay devant la Commission d'examen:

M.	Laurent FAVART	(Directeur de recherche, Université libre de Bruxelles, Rapporteur)
M.	Michael KLASSEN	(<i>Professor Doktor</i> , Université de Münster, Rapporteur)
M.	Etienne AUGÉ	(Professeur, Université Paris-Sud, Président)
M.	Eckhard ELSÉN	(Professeur, CERN, Examineur)
M.	David D'ENTERRIA	(Professeur, CERN, Examineur)
M.	Grégory SOYEZ	(Chargé de recherche, CEA-Saclay IPhT, Examineur)
M.	Christophe ROYON	(<i>Distinguished Professor</i> , Université du Kansas, Co-directeur de thèse)
M.	Laurent SCHOEFFEL	(Ingénieur chercheur, CEA-Saclay Irfu/SPP, Directeur de thèse)
M.	Bruno LENZI	(Docteur, CERN, Invité)

“On se croit naturellement bien plus capable d’arriver au centre des choses que d’embrasser leur circonférence. L’étendue visible du monde nous surpasse visiblement; mais comme c’est nous qui surpassons les petites choses, nous nous croyons plus capables de les posséder, et cependant il ne faut pas moins de capacité pour aller jusqu’au néant que jusqu’au tout; il la faut infinie pour l’un et l’autre; et il me semble que qui aurait compris les derniers principes des choses pourrait aussi arriver jusqu’à connaître l’infini. L’un dépend de l’autre et l’un conduit à l’autre. [...]”

“Connaissons donc notre portée; nous sommes quelque chose, et nous ne sommes pas tout; ce que nous avons d’être nous dérobe la connaissance des premiers principes, qui naissent du néant; et le peu que nous avons d’être nous cache la vue de l’infini.”

Blaise Pascal (1623-1662)



Thèse préparée à
**l’Institut de recherche sur les lois fondamentales
de l’Univers**
Service de Physique des Particules, Bât. 141
CEA-Saclay
91 191 Gif-sur-Yvette CEDEX

Résumé

Le LHC, qui est le collisionneur proton-proton le plus puissant du monde situé au CERN (Suisse), donne une occasion unique de pouvoir tester nos connaissances des interactions fondamentales à l'échelle du TeV. Au cours de cette thèse, deux projets ont été accomplis dans ce but. Tout d'abord, la mesure de la section efficace de production d'au moins deux photons dans l'état final ($pp \rightarrow \gamma\gamma + X$) a été réalisée à l'aide des données prises avec une énergie dans le centre de masse de 8 TeV par le détecteur ATLAS, qui est l'un des détecteurs polyvalents installé autour du LHC. Les photons sont des particules intéressantes pour tester la théorie de l'interaction forte (chromodynamique quantique ou QCD) car ils sont couplés aux quarks sans pour autant s'hadroniser, ce qui permet de pouvoir les mesurer avec une très bonne résolution. La mesure de leur taux de production au LHC permet de tester la QCD à la fois dans le domaine perturbatif et non-perturbatif. Elle est également sensible à l'émission de particules à basse énergie dans l'état initial, qui est un phénomène délicat à décrire d'un point de vue théorique à cause des divergences molles et collinéaires découlant de la QCD. Les incertitudes expérimentales ont été divisées par un facteur 2 ou plus comparé aux mesures réalisées précédemment au LHC ou au Tevatron (Fermilab, États-Unis) et la très bonne statistique des données d'ATLAS à 8 TeV a permis une augmentation significative à la fois de la résolution et de la portée de la mesure. En général, un bon accord est observé avec les prédictions théoriques. Le second projet réalisé au cours de la thèse est consacré à l'évaluation du potentiel de découverte relié à la mesure de la diffusion $\gamma\gamma$ au LHC ($\gamma\gamma \rightarrow \gamma\gamma$). La diffusion $\gamma\gamma$ est un processus singulier car prédit uniquement via des fluctuations quantiques qui n'a jamais été observé directement. Il implique un terme de couplage à quatre photons, qui montre une sensibilité significative à une large gamme de modèles d'extension du modèle standard à haute énergie, tels que ceux prédisant l'existence de dimensions supplémentaires de l'espace pour résoudre le problème de hiérarchie. En tirant profit du flux important de photons venant des protons au LHC, je montre qu'il est possible de découvrir des couplages anomaux à quatre photons avec une sensibilité permettant de rivaliser avec certaines recherches directes. Enfin, j'ai eu l'occasion de tester la nouvelle puce SAMPIC qui vise à réaliser des mesures de temps de vol avec une précision de l'ordre de quelques picosecondes grâce à un échantillonnage rapide des signaux de détecteur. Les performances de SAMPIC ont été testées à partir de signaux gaussiens produits par un générateur et par des détecteurs soumis à des impulsions infrarouges. Dans ces conditions idéales, SAMPIC permet de réaliser des mesures de temps de vol avec une résolution de l'ordre de 4 (40) ps pour les signaux générés (de détecteurs).

Mots-clefs : ATLAS, LHC, QCD, photon, modèle standard, mesure de temps de vol.

MEASUREMENT OF THE ISOLATED-PHOTON PAIR CROSS-SECTION IN THE ATLAS EXPERIMENT AT THE LHC AND STUDY OF FOUR-PHOTON COUPLINGS

Abstract

The LHC, which is the most powerful proton-proton collider in the world located at CERN (Switzerland), brings unprecedented opportunities to test our knowledge of the fundamental interactions at the TeV scale. In this work, two main projects have been achieved for this purpose. First, the production cross section measurement of at least two photons in the final state ($pp \rightarrow \gamma\gamma + X$) is performed with data taken at a center-of-mass energy of 8 TeV by the ATLAS detector, which is one of the multipurpose detector installed around the LHC ring. Photons are interesting probes to test the theory of strong interactions (Quantum chromodynamics or QCD) since they couple significantly to quarks but do not hadronize and thus still allow to perform high resolution measurements. The measurement of their production rate at the LHC allows to test QCD in both the perturbative and the non-perturbative domain. It is also sensitive to the emission of soft particles in the initial-state, which is tricky to handle on the theory side due to the collinear and soft divergences arising in QCD. Experimental uncertainties have been reduced by a factor 2 or more with respect to the measurements performed previously at the LHC or at the Tevatron (Fermilab, USA) and the high statistics of the ATLAS data sample at 8 TeV allows to increase significantly both the reach and the resolution of the measurement. In general, a good agreement is observed with theoretical predictions. The second project achieved in this work is dedicated to the evaluation of the light-by-light scattering potential (LbyL, $\gamma\gamma \rightarrow \gamma\gamma$) for new physics searches. LbyL is an intriguing process arising from quantum fluctuations only that has never been observed directly. It involves four-photon couplings, which are shown to be highly sensitive to a broad range of new physics models at high energy such as the ones predicting the existence of extra spatial dimensions to solve the hierarchy problem currently affecting the standard model of particle physics. By taking benefit of the photon flux from the protons at the LHC, I show that one may discover anomalous four photon couplings with a sensitivity allowing to compete with several direct new physics searches. Finally, I had the opportunity to test the new SAMPIC chip which aims to perform time-of-flight measurements with a few picoseconds precision from fast samplings of detector signals. SAMPIC timing capabilities have been tested using Gaussian signals generated by a signal generator or by silicon detectors pulsed with an infrared laser. Under these ideal conditions, the SAMPIC chip has proven to be capable of timing resolutions down to 4 (40) ps with synthesized (silicon detector) signals.

Keywords : ATLAS, LHC, QCD, photon, standard model, timing measurement.

Remerciements

Je tiens tout d'abord à remercier l'ensemble des membres de mon jury d'avoir accepté d'examiner mon travail de thèse. J'ai déjà eu l'occasion de discuter très agréablement de physique et/ou de collaborer directement avec certains membres, je pense bien sûr à mes directeurs de thèse le Prof. Christophe Royon et le Dr. Laurent Schoeffel mais aussi au Dr. Bruno Lenzi, au Prof. David d'Enterria et au Dr. Grégory Soyez. J'aurai également l'honneur de rencontrer lors de ma soutenance le Dr. Laurent Favart et le Prof. Dr. Michael Klasen en qualité de rapporteurs ainsi que le Prof. Etienne Augé et le Prof. Eckhard Elsen. J'espère être en mesure d'apporter à l'ensemble de mes examinateurs tous les éléments nécessaires leur permettant de juger de la qualité de mon travail.

Je remercie également l'Institut de recherche sur les lois fondamentales de l'Univers de m'avoir hébergé au cours de ces trois ans et au CEA-Saclay d'avoir financé mon travail. Je pense avoir appris plus de choses qu'un étudiant n'est en mesure d'espérer au cours d'une thèse. Merci également à l'ensemble des permanents et non-permanents du service de physique des particules et en particulier à Camilla Maiani, Jean-François Laporte, Vanina Ruhlmann-Kleider et Laurent Schoeffel pour avoir été présents lorsque j'ai eu besoin de leur aide, et plus généralement à l'ensemble des personnes avec qui j'ai eu le plaisir de collaborer.

Je n'oublie évidemment pas les actions déterminantes d'Elisabeth Vangioni, Olivier Drapier, Jean-Philippe Uzan, Robi Peschanski, Noël Camarcat et d'autres au cours de mon parcours scolaire qui m'ont guidé sur la voie de la physique fondamentale.

Je pense également à mes amis, en particulier à ceux me posant de nombreuses questions sur la physique étudiée au LHC. Cela s'avère très bénéfique pour moi et me force à prendre un peu de recul. Je pense aussi à ma famille, en particulier à la nouvelle venue de ce printemps.

J'aimerais finalement remercier tout particulièrement Christophe Royon pour sa confiance, son suivi et son soutien tout au long de mes années de thèse et Bruno Lenzi pour son avis d'expert et tout ses bons conseils, qui ont grandement contribué à améliorer la qualité de mon travail.

Merci aussi à toi, Charline, pour ton réconfort journalier qui m'a permis de tenir les derniers mois, et à ma mère pour sa relecture attentive du manuscrit et tous ses encouragements au cours de ces trois ans.

Table des matières

Introduction	15
I Le modèle standard de la physique des particules et ses limites	19
1 Considérations générales	20
1.1 Le modèle standard : une théorie quantique des champs	20
1.2 Symétries et invariances	21
1.3 Lien avec la physique expérimentale	23
2 Électromagnétisme, interaction faible et brisure de symétrie	25
2.1 Le champ électromagnétique et nucléaire faible	25
2.2 Brisure de symétrie électrofaible	28
2.3 Succès et limites de la théorie électrofaible	30
2.4 Les couplages à quatre photons, une nouvelle sonde à haut potentiel pour des découvertes au LHC	35
3 L'interaction nucléaire forte et la structure des hadrons	42
3.1 La chromodynamique quantique perturbative et non-perturbative	42
3.2 Hadronisation des partons et autres effets de QCD non-perturbative ob- servés dans les collisionneurs hadroniques	44
3.3 Structure du proton et calculs de sections efficaces hadroniques au LHC	46
3.4 Définition des observables et calcul des ordres supérieurs	49
3.5 La production de photons au LHC comme test de l'interaction forte . . .	51
II Le grand collisionneur de hadrons et l'expérience ATLAS	61
4 Le grand collisionneur de hadrons	62
4.1 Description générale	62
4.2 Intérêt d'un collisionneur proton-proton	65
4.3 Bilan de la première prise de données (2010-2012)	66
4.4 État des lieux de la seconde prise de données (2015-2018) et projet à plus long terme	68

5	L'expérience ATLAS	70
5.1	Description générale	70
5.2	Système de coordonnées et observables	72
5.3	Le détecteur interne	74
5.4	Le calorimètre électromagnétique	78
5.5	Le calorimètre hadronique	81
5.6	Le spectromètre à muons et les autres détecteurs d'ATLAS	82
5.7	Le système de déclenchement et la distribution des données	84
6	Les détecteurs de protons à l'avant au LHC et la mesure de temps de vol	88
6.1	La physique à l'avant	88
6.2	Les détecteurs de protons à l'avant	89
6.3	La mesure de temps de vol	92
III	Le traitement des photons dans les données à 8 TeV de l'expérience ATLAS	99
7	Algorithmes de reconstruction	100
7.1	Reconstruction de l'énergie des dépôts électromagnétiques	100
7.2	Reconstruction des électrons et des photons	101
8	Calibration de l'énergie des électrons et des photons	104
8.1	Présentation générale de la procédure	104
8.2	Calibration Monte Carlo	105
8.3	Corrections effectuées dans les données	107
8.4	Calibration in-situ et validation finale	110
9	Critères d'identification	114
9.1	Définitions	114
9.2	Efficacités et taux de rejet	116
9.3	Mesures de l'efficacité d'identification dans les données	116
9.4	La sélection LOOSE'	118
10	Étude et correction de l'énergie d'isolation transverse	121

10.1	Présentation générale	121
10.2	L'énergie d'isolation transverse dans le détecteur interne	123
10.3	L'énergie d'isolation transverse dans le calorimètre	124
10.4	Corrections supplémentaires de l'isolation calorimétrique basées sur les données	132
11	Évolutions pour la seconde prise de données	148
11.1	État des lieux pour les premières données à 13 TeV	148
11.2	Quelques réflexions personnelles	149
IV	Mesure de la section efficace de production de paires de photons isolés à $\sqrt{s} = 8$ TeV	153
12	Introduction	156
12.1	Enjeux de la mesure	156
12.2	Définition de l'espace fiduciel	157
12.3	Observables mesurées	159
12.4	Aperçu des résultats précédents obtenus au LHC et au TEVATRON . . .	163
13	Échantillons de données et échantillons simulés employés dans l'analyse	169
13.1	Choix du système de déclenchement	169
13.2	Échantillons simulés	169
14	Sélection de la région de signal et de régions de contrôle dans les données	175
14.1	Critères d'identification utilisés dans l'analyse	175
14.2	Définition de régions de contrôle dans les données	176
14.3	Choix et comparaison des variables d'isolation entre signal et bruits de fond	177
14.4	Reconstruction du vertex associé à la paire	184
14.5	Présélection des événements et des candidats photons	184
14.6	Classement final des différents événements	187
15	Estimation du bruit de fond hadronique et électrons à partir des données	190
15.1	Considérations générales	190

15.2	Étude des corrélations entre les deux candidats	192
15.3	La méthode du <i>template fit</i> 2D	198
15.4	Test et validation de la méthode sur un jeu de pseudo-données	208
15.5	Décomposition finale du bruit de fond dans les données réelles	216
15.6	Étude des incertitudes associées	232
15.7	Comparaison des résultats avec une méthode de soustraction alternative .	268
16	Déconvolution des effets de détecteurs et corrections des inefficacités	275
16.1	Considérations générales	275
16.2	Matrices de réponses et définition de facteurs de correction catégorie par catégorie	276
16.3	Identification et analyse des différents effets conditionnant les facteurs de correction	278
16.4	Étude de l'énergie transverse d'isolation définie au niveau particule . . .	302
16.5	Méthode de déconvolution itérative	308
16.6	Incertitudes sur la mesure finale	315
17	Comparaison des mesures aux prévisions théoriques et conclusion	336
17.1	Prédictions théoriques	336
17.2	Comparaisons aux mesures expérimentales	350
17.3	Conclusion	363
V	Étude de la sensibilité des expériences ATLAS et CT-PPS aux couplages à quatre photons	369
18	Introduction	371
19	Calcul des termes de couplage à quatre photons	373
19.1	Utilisation d'un lagrangien effectif	373
19.2	Calcul complet des amplitudes dans le cas générique de boucles de parti- cules chargées	379
19.3	Calcul complet des amplitudes dans le cas générique de résonances de spin 0 et 2	380
20	Le canal di-photon exclusif au LHC avec détection des protons intacts	381
20.1	Les différents modes de production	381

20.2	Mise en évidence expérimentale	385
20.3	Les prédictions du modèle standard	389
21	Étude des sensibilités aux couplages anormaux à quatre photons à haute énergie	393
21.1	Scénarios considérés	393
21.2	Composition et simulation du bruit de fond	394
21.3	Prise en compte des effets de détecteurs	402
21.4	Sélection des pseudo-données	405
22	Sensibilités finales, comparaisons aux limites actuelles et perspectives futures	410
22.1	Estimation de la sensibilité des données	410
22.2	Résultats obtenus par la théorie effective	411
22.3	Limites génériques sur l'existence de nouvelles particules chargées	413
22.4	Limites sur l'existence de particules neutres de spin $0/2$	414
22.5	Conclusion et perspectives	416
VI	Test de la carte de lecture SAMPIC pour des mesures de temps à la picoseconde	421
23	Introduction	422
23.1	Considérations générales	422
23.2	La puce SAMPIC	423
23.3	Intégration et pilotage	426
23.4	Le traitement des données	430
24	Résultats des mesures expérimentales	434
24.1	Tests avec des signaux issus de générateurs	434
24.2	Tests avec des signaux issus de détecteurs ultra-rapides de silicium	438
25	Conclusion et perspectives	446
VII	Conclusion générale, appendices et références	451
26	Conclusion générale	452

A	Corrections supplémentaires de l'isolation calorimétrique basées sur les données	455
B	Estimations du taux de mauvaise identification des jets en photons dans les données	464
C	Étude des différences entre les effets d'empilement dans les données et dans la simulation d'un point de vue calorimétrique	465
D	Résumé des différentes mesures $pp \rightarrow \gamma\gamma + X$ réalisées à $\sqrt{s} = 8$ TeV	469
	Références	476

Introduction

En ce mois d'avril 2016, le grand collisionneur de hadrons (LHC, pour *large hadron collider*) [1] du Centre Européen pour la Recherche Nucléaire (CERN) est sur le point de redémarrer afin de débiter l'acquisition de nouvelles données de collisions proton-proton avec une énergie dans le centre de masse inégalée de $\sqrt{s} = 13$ TeV. Après les premières données prises à l'automne 2015, une statistique conséquente devrait pouvoir être accumulée à $\sqrt{s} = 13$ TeV d'ici la fin de l'année 2016. Il est évident que le redémarrage d'un collisionneur permettant de sonder un domaine en énergie encore inexploré constitue un événement majeur pour la physique des hautes énergies, qui pourrait bouleverser profondément notre compréhension des tous premiers instants de l'Univers. Une telle occasion pour approfondir nos connaissances dans ce domaine ne se reproduira probablement pas avant au moins plusieurs dizaines d'années.

En attendant les nouvelles données du LHC, il est intéressant de faire un bilan rapide de ses premières années d'exploitation. Les prises de données à $\sqrt{s} = 7$ et 8 TeV (2010-2012) ont permis la découverte du boson de Higgs [2, 3], qui était la dernière particule manquante prédite par le modèle standard de la physique des particules [4, 5]. Le modèle standard (MS) est la théorie actuelle décrivant les interactions entre les particules élémentaires qui constituent la matière. L'existence du boson de Higgs, prédite dans les années soixante et établie près de cinquante ans plus tard, constitue ainsi un de ses plus gros succès. Cependant, la balle est à nouveau dans le camp des expérimentateurs. En effet, le MS n'explique pas tout et un certain nombre de questions reste en suspens, comme nous le verrons. En particulier, le MS nécessite un réglage fin de ses nombreux paramètres libres afin d'être prédictif à partir de l'échelle de plusieurs TeV, ce qui pourrait indiquer l'existence de nouveaux phénomènes dans la gamme d'énergie sondée par le LHC. Les physiciens théoriciens ont cependant besoin d'indices expérimentaux supplémentaires afin de pouvoir discriminer entre les nombreuses extensions possibles du MS. Les données à 7 et 8 TeV ont malheureusement apporté très peu d'éléments probants à ce sujet, laissant planer le doute quant à l'échelle en énergie réelle de ces nouveaux phénomènes.

Pendant mes années de thèse, je me suis intéressé plus particulièrement à la production de photons au LHC. Les photons sont des particules intéressantes à plus d'un titre. Tout d'abord, elles constituent les seules particules connues de masse nulle directement observables expérimentalement. Cette caractéristique leur accorde un statut particulier vis-à-vis de la théorie de la relativité, en leur permettant de voyager à la vitesse c dans le vide. Elle a également des conséquences très importantes sur l'électrodynamique quantique (QED, pour *quantum electrodynamics*), avec par exemple la conservation de la charge électrique et la propriété de renormalisabilité, permettant de résoudre les problèmes de la QED reliés aux divergences ultraviolettes en garantissant la validité de la théorie à tous les ordres d'un développement perturbatif. En contrepartie, la masse nulle des photons mène à des divergences infrarouges, qui sont considérées cependant comme bien comprises en général.

À basse énergie, un certain nombre de phénomènes tels que l'effet Casimir ou la violation des inégalités de Bell n'a pu être observé qu'à l'aide de photons jusqu'à maintenant. Par ailleurs, bien que la QED soit le secteur du modèle standard ayant été testé avec la plus grande précision, la diffraction $\gamma\gamma$ ($\gamma\gamma \rightarrow \gamma\gamma$), processus purement

issu de fluctuations quantiques prédit par la QED, n'a jamais encore pu être observé directement. Il n'est donc absolument pas exclu que les photons soient un portail vers de nouveaux phénomènes très faiblement couplés tels que l'existence d'axions (basse énergie) [6] ou de dimensions supplémentaires de l'espace (haute énergie) [7].

Après avoir été l'un des canaux de découverte du boson de Higgs au LHC ($H \rightarrow \gamma\gamma$), les paires de photons ont été à nouveau sur les devants de la scène récemment à l'occasion de l'excès modéré observé à la fois par les collaborations ATLAS et CMS dans les états finals $\gamma\gamma$ autour d'une masse de 750 GeV dans les premières données à $\sqrt{s} = 13$ TeV du LHC [8,9]. Le mode de désintégration en paires de photons est intéressant en particulier dans le cadre de la recherche de nouvelles particules scalaires, pseudo-scalaires ou de spin 2. Dans un futur proche, il pourrait ainsi devenir une voie d'exploration privilégiée pour étudier un éventuel nouveau secteur scalaire/spin 2 à intégrer dans le MS. D'autre part, les photons étant couplés aux constituants du proton et produits en grande quantité, ils forment également une fenêtre d'observation intéressante pour l'étude de l'interaction forte (QCD, pour *quantum chromodynamics*), particulièrement pertinente dans le cas de collisions proton-proton comme au LHC. En effet, du point de vue expérimental, les photons produisent des signaux très propres dans les détecteurs contrairement aux constituants du proton qui sont observables uniquement par l'intermédiaire de jets de particules collimatés constitués de divers états liés de l'interaction forte (hadrons).

J'ai également eu l'occasion pendant mes années de thèse de travailler sur une nouvelle puce électronique appelée SAMPIC développée par une collaboration entre le service d'électronique des détecteurs et d'informatique de l'Irfu (Irfu/SEDI) et le service d'électronique, recherche en détecteurs et instrumentation du laboratoire de l'accélérateur linéaire d'Orsay (LAL/SERDI). SAMPIC permet de mesurer des temps d'arrivée de particules avec une résolution de l'ordre de quelques picosecondes à partir d'un signal de détecteur. Ce projet a constitué une ouverture très enrichissante pour moi, en me donnant à la fois les outils et l'envie de poursuivre une activité hardware dans le futur.

Le manuscrit de thèse suit le plan suivant : dans la partie I, une revue succincte des principes et propriétés fondamentales du modèle standard de la physique des particules est donnée. Après une courte introduction constituant la section 1, les sections 2 et 3 introduisent respectivement le secteur électrofaible et le secteur fort du MS. Ces dernières s'achèvent chacune sur une ouverture aux deux projets principaux de la thèse : l'étude des couplages à quatre photons, formant le projet phénoménologique principal de la thèse, et la mesure de la section efficace de production de paires de photons à $\sqrt{s} = 8$ TeV avec les données de collision proton-proton du détecteur ATLAS, formant la mesure expérimentale principale effectuée pendant la thèse.

Dans la partie II, le collisionneur du LHC ainsi que le détecteur ATLAS sont introduits. Une parenthèse est également faite dans la section 6 sur les détecteurs de protons à l'avant au LHC et la mesure de temps de vol. Cette section sera utile à la bonne compréhension du projet phénoménologique principal (voir partie V) ainsi que de ma tâche technique accomplie au sein de la collaboration ATLAS avec la puce SAMPIC (voir partie VI).

Dans la partie III, le traitement des photons dans les données à $\sqrt{s} = 8$ TeV d'ATLAS est décrit, incluant notamment la reconstruction, la calibration, l'identification et les critères d'isolation appliqués pour la sélection des photons. Un soin particulier est apporté à la définition de l'énergie d'isolation, dont l'étude et la correction ont

constitué le travail de performance principal accompli pendant les années de thèse (voir section 10).

Dans la partie IV, la mesure expérimentale principale effectuée pendant les années de thèse est décrite en détail. Celle-ci a consisté à mesurer la section efficace de production de paires de photons à $\sqrt{s} = 8$ TeV avec les données de collision proton-proton du détecteur ATLAS au LHC. Elle constitue entre autres un test de précision important de la théorie des interactions fortes (QCD) à très haute énergie.

Dans la partie V, le projet phénoménologique principal effectué pendant les années de thèse est décrit en détail. Celui-ci s'est porté sur l'évaluation de la sensibilité des expériences ATLAS-AFP et CT-PPS du LHC aux couplages anormaux à quatre photons à très haute énergie. Les contributions prédites dans le cadre du modèle standard sont revisitées au préalable dans la section 20.

Dans la partie VI, les résultats finals issus de ma tâche technique dans ATLAS sont présentés. Celle-ci a consisté à tester la nouvelle puce SAMPIC, qui a été conçue pour mesurer des temps de vol de particules avec une résolution de quelques picosecondes. Pour cela, SAMPIC échantillonne très rapidement des signaux de détecteur, ce qui constitue une rupture technologique importante actuellement en plein essor comparée aux systèmes existants.

Finalement, le manuscrit s'achève sur une conclusion générale dans la partie VII. Quatre appendices numérotés de A à D apportant des précisions supplémentaires utiles aux lecteurs intéressés sont disponibles par la suite ainsi que l'ensemble des références bibliographiques employées pour la rédaction.

Première partie

Le modèle standard de la physique des particules et ses limites

Sommaire

1	Considérations générales	20
1.1	Le modèle standard : une théorie quantique des champs	20
1.2	Symétries et invariances	21
1.3	Lien avec la physique expérimentale	23
2	Électromagnétisme, interaction faible et brisure de symétrie	25
2.1	Le champ électromagnétique et nucléaire faible	25
2.2	Brisure de symétrie électrofaible	28
2.3	Succès et limites de la théorie électrofaible	30
2.4	Les couplages à quatre photons, une nouvelle sonde à haut potentiel pour des découvertes au LHC	35
2.4.1	Le processus de diffusion $\gamma\gamma$	35
2.4.2	Motivations pour une mesure directe	36
2.4.3	Sensibilités au LHC	38
3	L'interaction nucléaire forte et la structure des hadrons	42
3.1	La chromodynamique quantique perturbative et non-perturbative . .	42
3.2	Hadronisation des partons et autres effets de QCD non-perturbative observés dans les collisionneurs hadroniques	44
3.3	Structure du proton et calculs de sections efficaces hadroniques au LHC	46
3.4	Définition des observables et calcul des ordres supérieurs	49
3.5	La production de photons au LHC comme test de l'interaction forte	51
3.5.1	Motivations	51
3.5.2	Les différents modes de production	52
3.5.3	Critères d'isolation	55
3.5.4	Déformations des distributions $\gamma\gamma + X$ au NLO	58

1 Considérations générales

1.1 Le modèle standard : une théorie quantique des champs

Le modèle standard de la physique des particules (MS) est la théorie actuelle décrivant les interactions entre les particules élémentaires qui constituent la matière. Il repose largement sur le concept de champ, qui est une notion fondamentale de la physique moderne. En effet, les champs, qui ne sont qu'un artefact de calcul en mécanique classique (exemple : le champ de pesanteur terrestre $\vec{g}$, $||\vec{g}|| = 9,8 \text{ m.s}^{-2}$), acquièrent une réalité physique intrinsèque et une dynamique propre avec la théorie de la relativité restreinte et le postulat d'une vitesse limite c à la propagation de l'information. Ils deviennent alors indispensables à la description de certains phénomènes. La théorie quantique des champs est le formalisme mathématique adapté à l'application de la mécanique quantique aux champs [10, 11].

Étant donné les très petites échelles de distance mises en jeu lors des interactions élémentaires, souvent associées à de très hautes densités d'énergie et donc à des vitesses relativistes, il est naturel de construire le MS comme une théorie quantique des champs relativiste. Le MS inclut la description des interactions associées à trois types de champs : le champ électrofaible, le champ de Brout-Englert-Higgs (BEH) et le champ de l'interaction forte. La gravitation en est pour le moment exclue. Environ 10^{-11} seconde après le Big-bang, la brisure spontanée de symétrie électrofaible par le mécanisme de BEH [4, 5] conduisit à la scission du champ électrofaible en champs électromagnétique et nucléaire faible (voir section 2), qui sont décrits respectivement par l'électrodynamique quantique (QED) [12–15] et le modèle de Glashow-Weinberg-Salam [16–18]. L'interaction forte est décrite par la chromodynamique quantique (QCD) [19] (voir section 3). Tous ces formalismes se basent sur une théorie quantique des champs particulière.

Les différentes particules élémentaires de matières (fermions) qui ont été identifiées au cours du temps peuvent être classées en fonction de leurs nombres quantiques et de leurs couplages par rapport aux interactions fondamentales mentionnées plus haut. Deux grandes catégories co-existent : les leptons, qui ne sont pas couplés à l'interaction forte, et les quarks qui eux sont couplés à l'interaction forte. Les quarks et les leptons se décomposent indépendamment en six saveurs, elles-mêmes regroupées en trois générations qui possèdent des propriétés similaires ; seules les masses diffèrent. L'existence de générations ou de sous-structures supplémentaires est fortement contrainte par les mesures expérimentales, effectuées principalement au grand collisionneur électron-positron (LEP) [20] et au grand collisionneur de hadrons¹ (LHC) [1] du Centre Européen pour la Recherche Nucléaire (CERN) situé à Genève (Suisse), ainsi qu'au Tevatron [21] du laboratoire national de Fermilab (États-Unis), qui est également un collisionneur de hadrons.

Par ailleurs, comme conséquence de la quantification, chacune des interactions fondamentales se voit attribuer un certain nombre de particules médiatrices associées (bosons) en fonction des symétries internes de la théorie des champs sous-jacente. L'importance des symétries dans la construction du MS est développée de manière succincte dans la section 1.2. Un tableau récapitulatif des particules élémentaires incluses dans le modèle standard est montré dans la figure 1.

¹Les hadrons sont des particules composites liées par l'interaction forte (voir section 3).

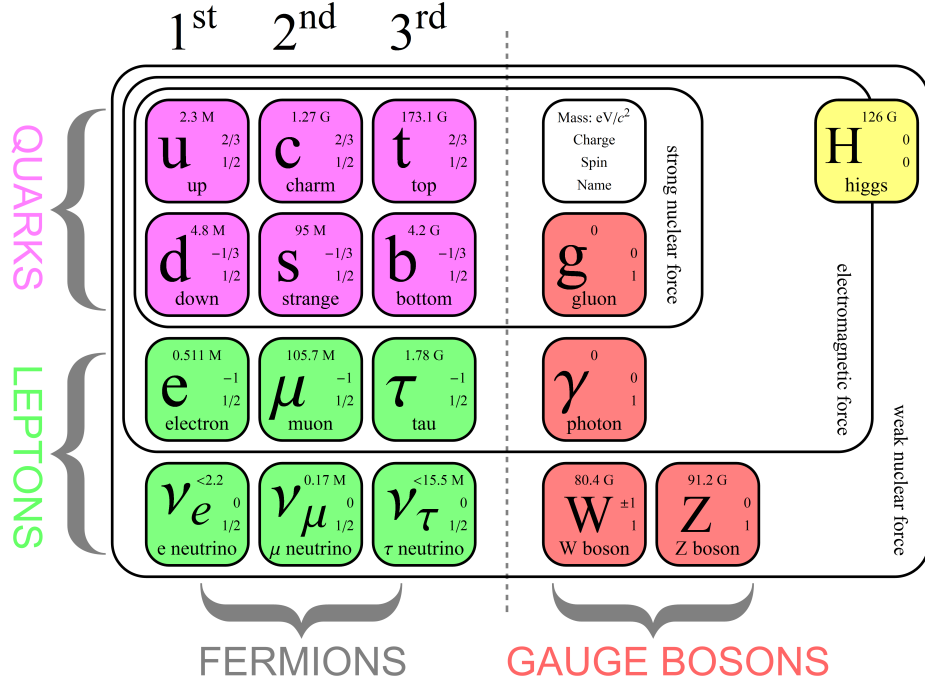


FIG. 1 – Particules élémentaires incluses dans le modèle standard de la physique des particules. Seuls les différents états de masse sont représentés [22].

1.2 Symétries et invariances

Une symétrie correspond à l'invariance d'un objet par certains types de transformations. Un des exemples le plus courant de système symétrique est la réflexion d'une image dans un miroir plan (voir figure 2). Ici, la symétrie est dite discrète car l'ensemble des transformations associées est fini, et même réduite à un élément dans ce cas particulier : $(x, y, z) \rightarrow (-x, y, z)$. Les symétries jouent un rôle prépondérant dans la structure du modèle standard, en particulier celles basées sur un ensemble de transformations infini et continu. Cela provient en grande partie du théorème formulé par Emmy Noether en 1918 [23] :

“À toute transformation continue d'un champ laissant l'action invariante est associée une grandeur qui se conserve.”

Ainsi, chaque symétrie continue d'un lagrangien a pour conséquence la conservation d'une quantité. Dans le cas d'une symétrie locale, c'est à dire lorsque l'ensemble des transformations associées est paramétré par les coordonnées d'espace-temps $x_\mu = (x^0, \mathbf{x})$, on l'appelle une invariance de jauge. Cet ensemble définit alors un groupe de jauge. Par exemple, la QED repose sur le groupe de jauge $U(1)_Q$, assurant la conservation de la charge électrique Q (voir section 2.1). Ce degré de liberté supplémentaire a également une importance particulière pour garantir le caractère renormalisable du MS, une propriété fondamentale qui lui permet de préserver sa prédictivité à tous les ordres d'un développement perturbatif [10].

Les symétries du lagrangien du modèle standard permettent de retrouver la conservation du tenseur énergie-impulsion (qui inclut la conservation de l'énergie, de l'impulsion et du moment angulaire), des charges de la QCD (couleurs), des charges de l'interaction électrofaible (hypercharge et isospin faible, et donc de la charge électrique), et du

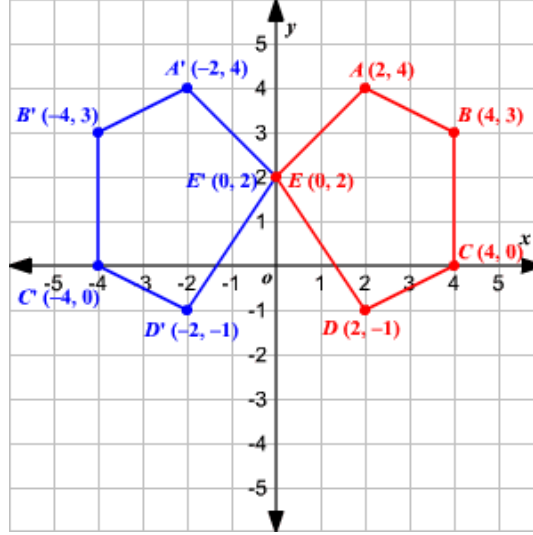


FIG. 2 – Réflexion d’une image dans un miroir plan. L’axe z est perpendiculaire au plan (x, y) . La transformation laissant la figure invariante correspond à l’ensemble réduit à l’élément $\{(x, y, z) \rightarrow (-x, y, z)\}$ [24].

nombre baryonique et leptonique global. Il existe également des symétries dites “approchées”, qui ne sont pas rigoureusement exactes mais permettent d’obtenir des résultats intéressants. Par exemple, la symétrie $SU(3)_{\text{savoir}}$ entre les quarks u , d et s est une bonne approximation permettant de décrire l’interaction forte et d’expliquer la spectroscopie des baryons et mésons plus simplement qu’avec la QCD complète [25].

Il existe également des symétries du MS qui sont discrètes, comme l’opération consistant à renverser l’écoulement du temps (T), à échanger particules et antiparticules (appelée conjugaison de charges, C), ou à changer l’orientation de l’espace (appelée parité, P). Elles ne rentrent pas dans le cadre du théorème de Noether, mais soulèvent des questions intéressantes : la gravitation, l’électromagnétisme et l’interaction forte sont symétriques par rapport à C , P et T . L’interaction faible viole C et P individuellement, mais pas CP et T . Certains processus rares impliquant des mésons K , D et B violent CP et T , mais toujours pas CPT qui pour le moment semble être une symétrie parfaite de la nature [26]. La violation de C et P individuellement par l’interaction faible est prédite de façon satisfaisante par la théorie de jauge de Glashow-Weinberg-Salam, mais l’origine des violations de CP reste à éclaircir. Ces processus rares pourraient être à l’origine de l’asymétrie matière/anti-matière observée dans l’Univers, qui constitue un des mystères de la physique moderne.

D’autre part, notre compréhension des tous premiers instants de l’Univers et les espoirs de voir émerger un jour une théorie de grande unification décrivant avec le même formalisme toutes les grandes interactions fondamentales [10, 11] reposent largement sur le concept de brisure de symétrie, qui sera discuté dans la section 2.2. Tout semble donc indiquer que les symétries de la nature, déjà indispensables à la compréhension du MS, vont continuer à jouer un rôle de première importance dans ses futurs développements.

En plus d’un cadre théorique solide, tous les modèles nécessitent des confirmations expérimentales. L’importance de la physique expérimentale dans la construction du modèle standard est rappelée dans la section 1.3.

1.3 Lien avec la physique expérimentale

Aucune théorie ne peut être complète sans un appui expérimental. Cela est particulièrement vrai dans le cas du modèle standard, qui décrit des phénomènes à l'échelle des particules élémentaires et donc non observables directement dans notre vie de tous les jours. Le tester précisément nécessite donc un investissement important de la part des physiciens expérimentateurs.

D'une manière générale, tous les modèles comportent un certain nombre de paramètres libres devant être mesurés. Au total, le MS ne compte pas moins de dix-neuf paramètres libres [27].² Ce chiffre important accorde un rôle de premier plan aux physiciens expérimentateurs, qui ont pour mission de bâtir des grandes expériences comme ATLAS au LHC (voir partie II) pour mesurer avec une grande précision ces quantités clés. Celles-ci peuvent ensuite être utilisées par les physiciens théoriciens afin de produire des prédictions fiables, qui sont à leur tour testées dans les grandes expériences.

Différentes techniques expérimentales ont été adoptées au fil des années pour développer et tester le modèle standard. En particulier, l'étude de collisions de particules à très haute énergie est l'une des méthodes les plus utilisées : grâce à la conversion de l'énergie de collision en énergie de masse, les différentes particules instables du MS peuvent être recréées artificiellement pendant un très court instant. Leurs produits de désintégration sont alors identifiés dans de grands détecteurs afin d'étudier en détail l'interaction initiale. Un résumé des principaux collisionneurs de particules construits ou en projet pour les années futures est montré dans la figure 3 en fonction de leur énergie de collision dans le centre de masse, augmentant au fur et à mesure des progrès technologiques. Les premières machines datent des années soixante et la dernière mise en service est le LHC du CERN. D'autres sont en projet, comme le collisionneur linéaire international (ILC) [28] ou le collisionneur circulaire futur (FCC, non représenté sur la figure) [29]. Elles sont classées en deux types principaux : les collisionneurs électrons-positrons et les collisionneurs de hadrons (protons, antiprotons ou ions). Des détails supplémentaires sur le LHC et l'expérience ATLAS seront donnés dans la partie II.

Cependant, des arguments à la fois théoriques (naturalité, problème de la hiérarchie, ...) et expérimentaux (matière noire, asymétrie matière/anti-matière, ...) indiquent que le modèle standard est aujourd'hui incomplet, comme cela sera développé par la suite. Par conséquent, de nombreuses extensions sont proposées par les physiciens théoriciens afin de l'améliorer, le plus souvent en postulant l'existence de nouvelles particules, comme la supersymétrie [31, 32], et/ou de nouvelles sous-structures pour des particules déjà identifiées, comme les théories de Higgs composite [33]. Notons aussi l'idée de l'existence de dimensions supplémentaires de l'espace [34]. Ces nouveaux phénomènes hypothétiques sont appelés d'une manière générale "nouvelle physique", et sont caractérisés par une échelle d'énergie typique, correspondant par exemple à la masse de nouvelles particules ou à l'énergie de liaison d'une nouvelle interaction. Ils permettent de résoudre un certain nombre de problèmes du MS actuel. Certains d'entre eux seront abordés plus en détail dans les sections suivantes.

De leur côté, les physiciens expérimentateurs cherchent à détecter des signaux de nouvelle physique, soit directement, soit indirectement. La méthode directe consiste à

²Ce chiffre est correct pour des neutrinos de masse nulle. Dans le cas contraire, sept paramètres libres supplémentaires doivent être considérés.

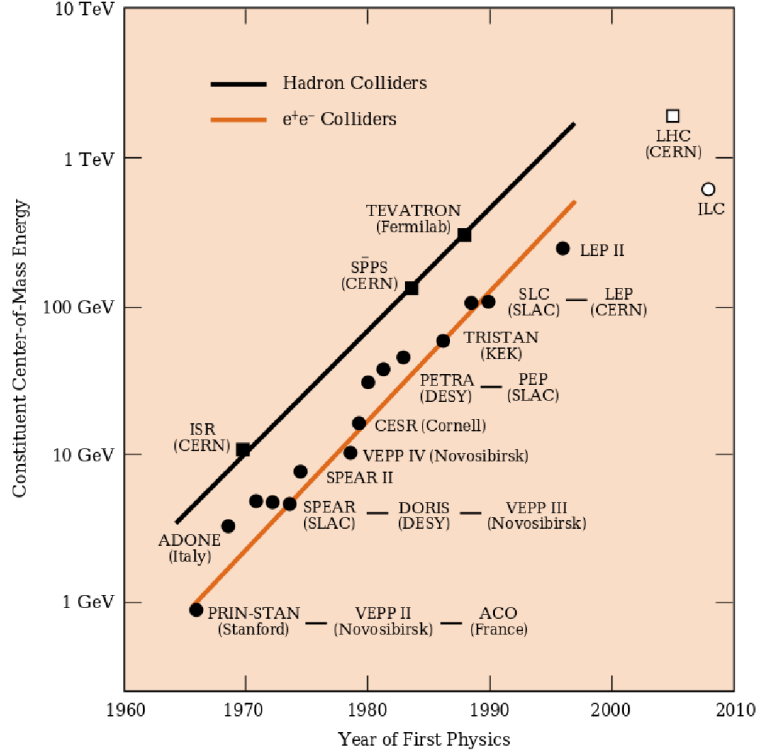


FIG. 3 – Résumé des différents collisionneurs de hadrons (carré) et électron-positron (rond) mis en service depuis les années soixante, en fonction de leur énergie de collision dans le centre de masse. Le LHC du CERN est la dernière machine à avoir été mise en service (2009), tandis que l'ILC devrait être la prochaine à être démarrée. L'énergie de collision des collisionneurs de hadrons a été divisée par un facteur allant de six à dix afin de tenir compte de la nature composite des particules qu'ils font se rencontrer [30].

sonder des domaines encore non explorés par l'homme, par exemple en faisant rentrer en collision des particules à des énergies encore jamais atteintes sur Terre, comme au LHC. La méthode indirecte se concentre sur des mesures de précision de certains paramètres très contraints du MS, comme par exemple la masse du boson W ($m_W \simeq 80.385$ GeV [35]), afin de détecter des déviations éventuelles par rapport à la théorie et de tester la cohérence générale du modèle. Elle présente l'avantage de ne pas dépendre des différents modèles de nouvelle physique. Les deux approches sont cependant très complémentaires : la première permet un test direct des différentes propositions d'extension du MS à partir de données prises dans des zones non explorées, alors que la seconde permet de détecter de légères déviations induites par une nouvelle physique peut-être encore hors de portée, mais ayant tout de même des effets à plus basse énergie grâce aux fluctuations autorisées par la mécanique quantique.

Afin de mieux comprendre les succès et limites du MS, quelques-unes de ses propriétés fondamentales dans le secteur électrofaible et nucléaire fort sont introduites dans les sections 2 et 3, respectivement. Chacune des deux sections s'achève avec une ouverture sur les travaux expérimentaux ou phénoménologiques effectués pendant les années de thèse reliés à ces deux secteurs : les couplages à quatre photons sont détaillées dans la section 2.4, tandis que l'importance des tests de précisions de la QCD au LHC à l'aide de photons est soulignée dans la section 3.5.

2 Électromagnétisme, interaction faible et brisure de symétrie

2.1 Le champ électromagnétique et nucléaire faible

Bien qu'ayant des propriétés très différentes à l'heure actuelle, les champs électromagnétique et nucléaire faible sont intimement liés dans le modèle standard. Il est possible de retrouver quelques-unes de leurs propriétés importantes en étudiant certains termes de leur lagrangien respectif.

Commençons par le cas simple de la QED. Partant des équations de Maxwell régissant l'électrodynamique classique, le lagrangien du champ électromagnétique en présence d'une charge Q décrite par le quadrivecteur courant $J^{\text{EM}} = (\rho; \vec{j})$ peut se formuler de la façon suivante :

$$\mathcal{L}_{\text{EM}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + QJ_{\mu}^{\text{EM}}A^{\mu}, \quad (1)$$

où $A = (V; \vec{A})$ est le quadrivecteur potentiel et $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ est le tenseur électromagnétique. On cherche à faire interagir une particule de matière (fermion) chargée, décrite par une fonction d'onde notée ψ , avec le champ électromagnétique. La dynamique des fermions chargés est donnée par l'équation de Dirac, qui a pour lagrangien associé :

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma_{\mu}\partial^{\mu} - m)\psi, \quad (2)$$

où γ^{μ} sont les matrices de Dirac, $\bar{\psi} = \psi^{\dagger}\gamma^0$ est l'adjoint de Dirac et m est la masse au repos de la particule. D'après l'équation de Dirac, cette particule modifie le courant de la façon suivante : $J_{\mu}^{\text{EM}} = \bar{\psi}\gamma_{\mu}\psi$. On cherche également à ce que le lagrangien de ce système soit invariant par la transformation locale :

$$A_{\mu} \rightarrow A'_{\mu} = A_{\mu} - \frac{1}{Q}\partial_{\mu}\alpha(x), \quad (3)$$

où $\alpha(x)$ est une fonction arbitraire des coordonnées de l'espace-temps. Cette transformation correspond au degré de liberté sur la définition du quadrivecteur A dans les équations de Maxwell [10, 11]. En remplaçant A_{μ} par A'_{μ} dans l'équation 1, on obtient comme condition sur la fonction d'onde ψ son invariance par la transformation suivante :

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha(x)}\psi(x). \quad (4)$$

On introduit souvent la dérivée covariante $D_{\mu} = \partial_{\mu} + iQA_{\mu}$ pour écrire le lagrangien de la QED, ce qui donne finalement :

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma_{\mu}D^{\mu} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (5)$$

On obtient ainsi la dynamique du champ vectoriel A_μ de spin 1 (le photon) avec un fermion décrit par la fonction d'onde ψ (par exemple, l'électron, spin 1/2) et leurs interactions. Afin d'avoir une théorie quantique viable qui préserve le principe de causalité, il faut également quantifier le champ de Dirac et le champ électromagnétique. Cela n'est pas abordé ici par souci de concision, mais est détaillé par exemple dans [10]. On obtient alors la notion de particules comme quanta des différents champs en présence. Notons que la condition obtenue dans l'équation 4 sur la fonction d'onde assure une symétrie locale du lagrangien de la QED correspondant au groupe de jauge $U(1)_Q$. Cela induit la conservation de la charge électrique Q grâce au théorème de Noether, mentionné dans la section 1.2. Il est également intéressant de remarquer que si le photon avait une masse non nulle, le lagrangien aurait un terme supplémentaire $\frac{1}{2}m^2 A_\mu A^\mu$ qui violerait l'invariance de jauge.

Au contraire, pour l'interaction nucléaire faible, responsable par exemple des désintégrations radioactives de type β , on observe expérimentalement des bosons de jauge très massifs ($W^\pm$, Z^0 , voir figure 4). Celle-ci peut être décrite de façon satisfaisante par une théorie de jauge basée sur le groupe $SU(2)$, mais uniquement si les bosons sont de masse nulle, en utilisant des arguments similaires à ceux de la QED. En revanche, ces hautes masses rendent les bosons très instables à cause du principe d'incertitude d'Heisenberg et donc semblent adaptées à la portée très limitée de l'interaction faible, qui est observée à l'échelle nucléaire (10^{-15} mètres) mais est absente aux échelles macroscopiques. Ce problème théorique majeur a été résolu en deux temps, d'une part grâce au mécanisme de BEH développé dans les années soixante expliquant l'acquisition des masses des différentes particules élémentaires par un processus dynamique provoqué par une brisure spontanée de symétrie [4, 5], et d'autre part par le modèle d'unification des interactions électromagnétique et nucléaire faible à haute énergie proposé par Glashow-Weinberg-Salam [16–18]. Ces derniers ont été récompensés du prix Nobel en 1979, peu après la découverte des courants neutres au CERN prédits par leur théorie. Peter Higgs et François Englert l'ont reçu en 2012, peu après l'observation au LHC d'un boson compatible avec le boson de BEH [2, 3] prédit par le mécanisme du même nom.³

L'interaction faible après brisure de symétrie électrofaible présente plusieurs particularités, en plus ou en prolongation des violations de symétrie C et P évoquées précédemment (voir section 1.2). Basée sur la charge électrique et un nombre quantique supplémentaire appelé isospin faible (ou alternativement sur l'isospin faible et l'hypercharge), elle se décompose en deux parties distinctes : les courants chargés, transportés par les bosons $W^\pm$, qui engendrent un changement de saveur et sont par exemple responsables des désintégrations radioactives de type β , et les courants neutres transportés par le boson Z^0 , qui eux ne peuvent pas changer les saveurs. Sa grande spécificité provient de sa nature chirale, donnée par certains termes de couplages aux fermions de son lagrangien explicités ci-dessous :

$$\mathcal{L}_{CC} = \frac{g}{\sqrt{2}} [J_\mu^+ W^{\mu-} + J_\mu^- W^{\mu+}], \quad (6)$$

$$\mathcal{L}_{NC}^{(1)} = \frac{g}{\cos \theta_W} J_\mu^Z Z^\mu, \quad (7)$$

³Robert Brout est malheureusement décédé en mai 2011, peu avant la découverte expérimentale du boson de BEH.

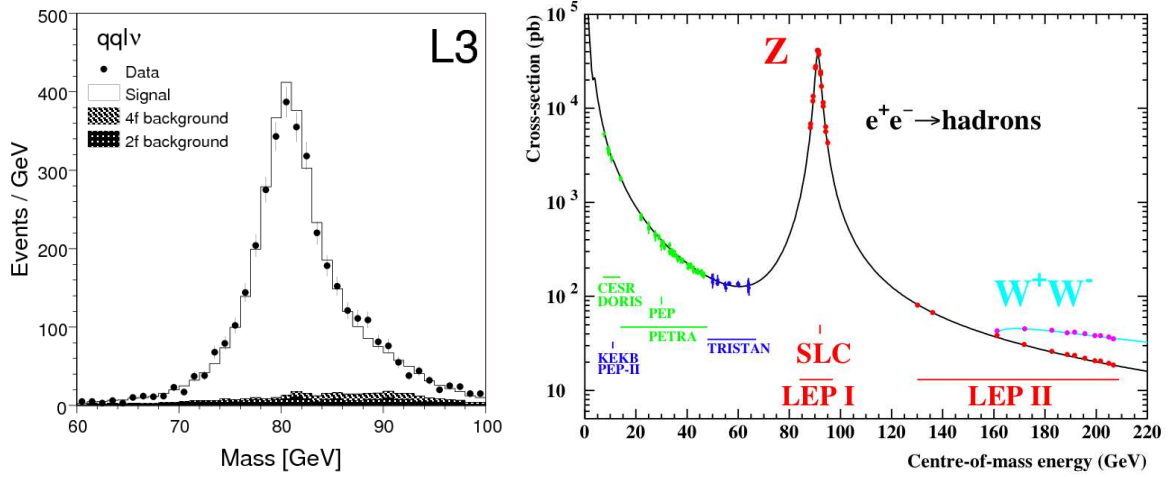


FIG. 4 – Reconstruction de la masse invariante du boson W par la collaboration L3 au LEP, CERN, à partir des états finals $W^+W^- \rightarrow q\bar{q}l\nu$, $l = e, \mu$ (gauche) [36]. Résumé des mesures de section efficace totale $e^+e^- \rightarrow \text{hadrons}$ effectuées dans plusieurs expériences, en fonction de l'énergie de collisions dans le centre de masse (droite) [37]. La résonance autour de 91 GeV correspond au boson Z^0 .

où CC et NC différencient les termes associés aux courants chargés et neutres, respectivement, tandis que ⁽¹⁾ signifie que certains termes supplémentaires du lagrangien ont été tronqués pour une meilleure compréhension. θ_W est un paramètre libre du MS définissant la charge élémentaire en fonction du couplage de l'interaction g lors de la brisure de symétrie électrofaible, $e = g \sin \theta_W$, appelé angle de mélange (voir section 2.2). Z_μ , W_μ^+ et W_μ^- sont les champs vectoriels de spin 1 associés respectivement au boson Z^0 , W^+ et W^- , et les J_μ^i sont les courants correspondants. La chiralité de l'interaction est donnée par la forme de ces courants :

$$J_\mu^+ = \sum_f \bar{f}_u \gamma_\mu \frac{1 - \gamma_5}{2} f_d, \quad J_\mu^- = \sum_f \bar{f}_d \gamma_\mu \frac{1 - \gamma_5}{2} f_u, \quad (8)$$

$$J_\mu^Z = \sum_f \bar{f} \gamma_\mu \frac{\hat{v}_f - \hat{a}_f \gamma_5}{2} f, \quad (9)$$

où les γ_μ sont les matrices de Dirac et les f , f_i sont les fonctions d'onde des différents fermions. $\frac{1 - \gamma_5}{2}$ est l'opérateur correspondant à une projection sur la composante gauche d'une fonction d'onde de fermion, habituellement décrite par un spineur à deux composantes correspondant aux deux états d'hélicité autorisés par la statistique de Fermi-Dirac. Les coefficients $\hat{v}_f$ et $\hat{a}_f$ sont fonctions de l'isospin faible et de la charge électrique des fermions considérés.

Il apparaît donc que seuls les fermions d'hélicité gauche se couplent aux bosons $W^\pm$, ce qui entraîne la violation de parité déjà mentionnée dans la section 1.2. Par invariance de jauge de l'isospin faible, qui se présente en doublet de $SU(2)_L$ pour les fermions gauches, les courants chargés entraînent nécessairement un changement de saveur des fermions, symbolisé par les indices u et d (par référence aux quarks) dans l'équation 8. L'expression du courant neutre est plus complexe, car elle fait intervenir les composantes

gauche et droite des fermions avec des poids différents, paramétrés par les coefficients $\widehat{v}_f$ et $\widehat{a}_f$.

Une difficulté supplémentaire vient du fait que les états propres de masse et de saveur des quarks ne correspondent pas exactement, ce qui rend nécessaire la construction d'une matrice de passage entre les deux, appelée matrice de Cabibbo-Kobayashi-Maskawa (CKM) [38].⁴ La propriété d'unitarité de la matrice CKM interdit les courants neutres changeant la saveur par le mécanisme de Glashow-Illiopoulos-Maiani [39].

Afin de comprendre la relation entre interaction faible, électromagnétisme et la solution aux problèmes liés à la masse des bosons de jauge, il est nécessaire de donner quelques éléments de compréhension sur un phénomène particulier ayant eu lieu au tout début de l'histoire de notre Univers : la brisure de symétrie spontanée électrofaible.

2.2 Brisure de symétrie électrofaible

D'après le modèle de Glashow-Weinberg-Salam [16–18], qui n'a jamais été mis en défaut expérimentalement, les interactions électromagnétiques et nucléaires faibles étaient au moment du Big-bang décrites par une unique théorie de jauge dite électrofaible, basée sur le groupe $SU(2)_L \times U(1)_Y$, où L est l'isospin faible et Y l'hypercharge. $SU(2)_L$ et $U(1)_Y$ sont associés respectivement aux couplages g et g' et aux champs $(W_\mu^a)_{a=0,1,2}$ et B_μ , qui ont alors des bosons vecteurs de masse nulle. En effet, comme déjà mentionné dans la section 2.1, des bosons massifs feraient apparaître des termes quadratiques dans le lagrangien de type $\frac{1}{2}m^2 X_\mu X^\mu$ qui violeraient l'invariance de jauge. Pour une raison légèrement différente reliée à la structure interne de l'isospin faible, il est également impossible d'avoir dès l'origine des fermions massifs.

Le mécanisme de BEH [4, 5] permet de faire apparaître dynamiquement des termes de masse pour les bosons vecteurs et les fermions, en postulant l'existence d'un champ scalaire complexe ϕ appelé champ de BEH et ayant un potentiel de la forme suivante :

$$U(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda}{2} (\phi^\dagger \phi)^2, \quad (10)$$

où $\mu^2 < 0$ et $\lambda > 0$ des constantes. Ce potentiel est représenté sur la figure 5. La condition $\mu^2 < 0$ impose l'existence d'une infinité de minimums répartie autour d'un cercle de rayon $v = \sqrt{-\mu^2/\lambda}$. Ce rayon correspond à la valeur moyenne du champ dans l'état fondamental, appelée *vacuum expectation value*, qui est donc non-nulle. On peut alors déjà deviner ce qui s'est passé lors du refroidissement dû à l'expansion de l'Univers : dans un Univers infiniment chaud et dense, le champ ϕ a une énergie suffisante pour monter tous les potentiels. Il a donc une valeur moyenne nulle. Cet état est représenté par la boule centrée en zéro sur la figure 5. Le potentiel au voisinage de la position d'équilibre (instable) présente alors une symétrie continue correspondant à la transformation :

$$U(\phi_0 + \sigma(x)) \rightarrow U(\phi_0 + \sigma(x) \cdot e^{i\theta}), \quad (11)$$

⁴Dans le modèle standard, il n'existe pas de phénomène similaire dans le secteur leptonique puisque les neutrinos n'ont pas de masse. Cependant, des oscillations de saveurs de neutrinos (allant donc à l'encontre de la théorie la plus simple) sont observés et ont mené à la définition d'une matrice similaire appelée Pontecorvo-Maki-Nakagawa-Sakata (PMNS).

où $\sigma(x)$ est une perturbation (complexe) du champ par rapport à sa position d'équilibre ϕ_0 , et θ est un angle quelconque. Lorsque la température atteint une valeur proche de la *vacuum expectation value*, le champ n'a plus l'énergie suffisante pour monter tous les potentiels et va donc passer dans un des ses états fondamentaux, représenté par la deuxième boule sur la figure 5. Il est impossible de prédire exactement dans quel état fondamental la transition va s'effectuer, mais cela va avoir pour effet de réduire la symétrie apparente du potentiel. En effet, au voisinage de la nouvelle position d'équilibre (stable), le potentiel n'est plus invariant par la transformation de l'équation 11 : il y a brisure spontanée de symétrie.

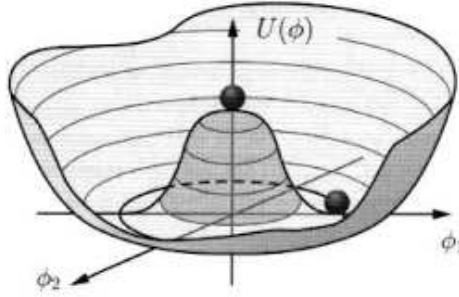


FIG. 5 – Représentation schématisée du potentiel de BEH. La première boule centrée en zéro correspond à la position d'équilibre du champ au moment du Big-bang. La deuxième boule correspond à un état fondamental pris par le champ après la brisure de symétrie spontanée électrofaible [40].

Le fait que le champ de BEH prenne désormais une valeur moyenne non nulle va générer des termes quadratiques mélangeant les champs électrofaibles $(W_\mu^a)_{a=0,1,2}$ et B_μ (voir début de la section) dans le lagrangien correspondant aux interactions de ϕ avec les champs électrofaibles. Il reste à bien choisir les invariances de jauge de ϕ afin de retrouver les propriétés du champ électromagnétique et nucléaire faible du MS. Il est possible de montrer qu'un champ Φ complexe correspondant à un doublet d'isospin faible d'hypercharge 1 permet cela [10]. On peut alors reconnaître de nouveaux états propres de masse correspondants aux champs de l'interaction nucléaire faible W_μ^+ , W_μ^- et Z_μ dans les termes quadratiques du lagrangien par identification avec la forme bien connue d'un terme d'énergie cinétique $\frac{1}{2}m^2 X_\mu X^\mu$. Un autre champ noté A_μ et mélangeant W_μ^3 et B_μ reste lui sans masse et est identifié à l'électromagnétisme.

Les combinaisons entre les champs $(W_\mu^a)_{a=0,1,2}$ et B_μ sont paramétrées par λ , v et les constantes de couplages de l'interaction électrofaible g et g' . On définit souvent un paramètre appelé angle de mélange et noté θ_W , déjà mentionné dans la section 2.1 :

$$\cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}} ; \quad \sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}. \quad (12)$$

Celui-ci permet d'écrire facilement la matrice de mélange permettant de passer de la base (W^3, B) à (Z, A) , ainsi que l'équation permettant de retrouver la charge élémentaire :

$$\begin{pmatrix} Z \\ A \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} = \begin{pmatrix} W^3 \\ B \end{pmatrix}, \quad (13)$$

$$e = g \sin \theta_W. \quad (14)$$

θ_W est un paramètre libre du modèle standard mesuré par les physiciens expérimentateurs. Le couplage g , relié au groupe $SU(2)_L$, reste inchangé entre interaction électrofaible et nucléaire faible alors que la charge électrique représente un mélange entre hypercharge et isospin faible. Les termes de mélange permettant d'obtenir W_μ^+ , W_μ^- et les termes de masse ne sont pas explicités ici mais se trouvent facilement dans les ouvrages de référence.

De façon similaire, les masses des fermions sont générées en introduisant un isodoublet du champ Φ ayant une hypercharge $Y = -1$. Les couplages de ce champ aux fermions peuvent alors être développés en utilisant des lagrangiens dit de Yukawa, du nom du premier physicien qui explicita les couplages entre scalaires et fermions [41].

Il reste à étudier la dynamique du champ de BEH lui-même après brisure de symétrie. La forme du potentiel $U(\phi)$ au voisinage de la nouvelle position d'équilibre (voir figure 5) montre que les perturbations du champ peuvent être de deux types : soit le long du cercle de rayon v , soit dans la direction perpendiculaire. Si on développe le potentiel autour de la position d'équilibre stable du champ, $\phi(x) = v + \frac{1}{\sqrt{2}}(\phi_1(x) + i\phi_2(x))$, on obtient la forme suivante :

$$U(\phi) = -\frac{\mu^4}{2\lambda} + \frac{1}{2} \cdot (-2\mu^2)\phi_1^2 + O(\phi_i^3). \quad (15)$$

Le choix de la position d'équilibre en $(v, 0)$ ne fait perdre aucune généralité car implique uniquement un choix de base approprié. L'équation 15 indique que la perturbation selon l'axe des réels ϕ_1 est un champ scalaire de masse non nulle égale à $-2\mu^2 > 0$. Sa particule associée est appelée le boson de BEH et a été observée pour la première fois en 2012 au LHC par les expériences ATLAS et CMS [2, 3]. Sa masse $-2\mu^2 > 0$ est un paramètre libre de la théorie. En revanche, les termes en ϕ_2^2 s'annulent lors du développement du potentiel, indiquant que le champ ϕ_2 est sans masse. Les champs de type ϕ_2 sont au nombre de trois pour Φ , qui compte deux degrés de liberté supplémentaires par rapport à ϕ . Les particules associées aux champs de type ϕ_2 sont appelées des bosons de Goldstone. Elles n'apparaissent pas comme des particules physiques à part entière dans la théorie mais constituent des degrés de liberté supplémentaires nécessaires à la génération des masses des trois bosons faibles.

2.3 Succès et limites de la théorie électrofaible

La brisure de symétrie électrofaible est un des exemples typiques expliquant le grand succès du modèle standard. A partir de constatations expérimentales diverses, *ie.* l'existence de deux types de courants dans l'interaction faible, et du problème théorique relié à la masse des bosons de jauge, la théorie a su fournir un mécanisme général de génération de masse fonctionnant à la fois pour les bosons et les fermions. Celui-ci résout élégamment le problème en unifiant deux interactions fondamentales à haute énergie à l'aide du postulat minimal consistant à introduire un nouveau champ scalaire ayant les bonnes propriétés. La théorie fournit des prédictions pour un grand nombre de phénomènes à des énergies variées, allant des désintégrations radioactives β ($E_e \simeq 100$ keV)

aux bosons Z^0 porteurs des courants neutres ($m_Z \simeq 91$ GeV), en passant par l'existence du boson scalaire de BEH et de la valeur de ses couplages avec les autres particules. La très grande majorité est pour le moment largement vérifiée à des précisions toujours plus hautes par les physiciens expérimentateurs.

Les figures 6 et 7 montrent deux exemples de tests de la théorie électrofaible effectués par le groupe GFITTER à l'aide de mesures de précision [42]. Dans la figure 6, un test de cohérence de la théorie électrofaible comparant les mesures directes de la masse du boson W , du quark top et du boson de BEH avec leur prédiction indirecte est représenté. Le recouvrement des barres vertes et bleues montrent que les différentes mesures sont largement compatibles avec la théorie dans les barres d'erreurs. Dans la figure 7, un résumé des mesures des observables électrofaibles est montré. Malgré quelques anomalies isolées telles que l'asymétrie avant-arrière pour les quark b , la grande majorité des mesures sont en très bon accord avec les prédictions de la théorie.

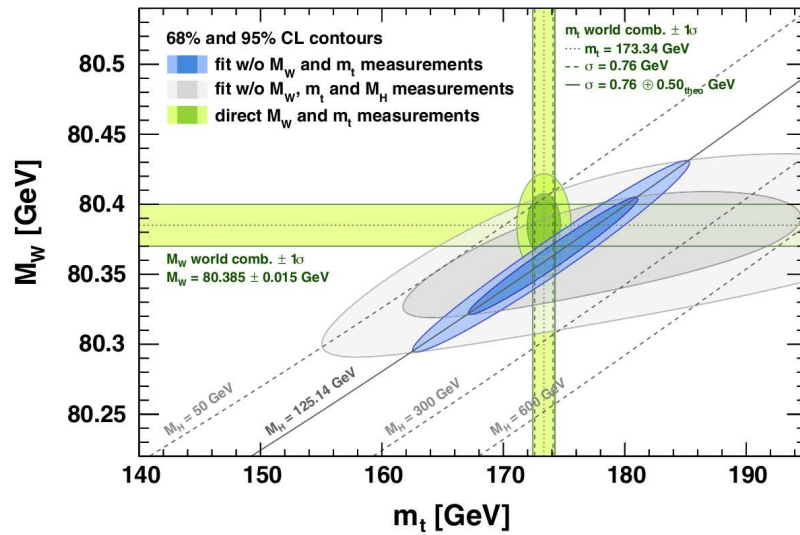


FIG. 6 – Exemple de test de cohérence de la théorie électrofaible réalisé par le groupe GFITTER à partir des masses du boson W , du quark top et du boson de Higgs [42].

Les mesures récentes reliées au boson de BEH effectuées par les collaborations ATLAS et CMS avec les données du *run 1* du LHC (voir partie II) constituent un autre exemple intéressant de test de la théorie électrofaible. En effet, une fois la masse du boson de BEH mesurée, la théorie prédit de façon très précise ses différentes sections efficaces de production en fonction du modes de production et les rapports de branchement, reliés aux modes de désintégration. Ces mesures sont résumées dans la figure 8 et montrent un très bon accord avec la théorie du MS, à la fois pour les modes de production et de désintégration [43, 44].

Cependant, la théorie reste non-satisfaisante sur un certain nombre de points. Par exemple, le mécanisme de BEH n'explique pas la hiérarchie entre les masses des trois générations de fermions, allant dans le secteur des quarks de 2.3 MeV pour le quark up à 173 GeV pour le quark top. Ces masses sont paramétrées par des couplages de Yukawa, paramètres libres du MS. Le modèle standard compte d'ailleurs un grand nombre de paramètres libres, dix-neuf au total.⁵ Malgré ce grand nombre de degrés de liberté, il

⁵Voir la note de bas de page N° 2, section 1.3.

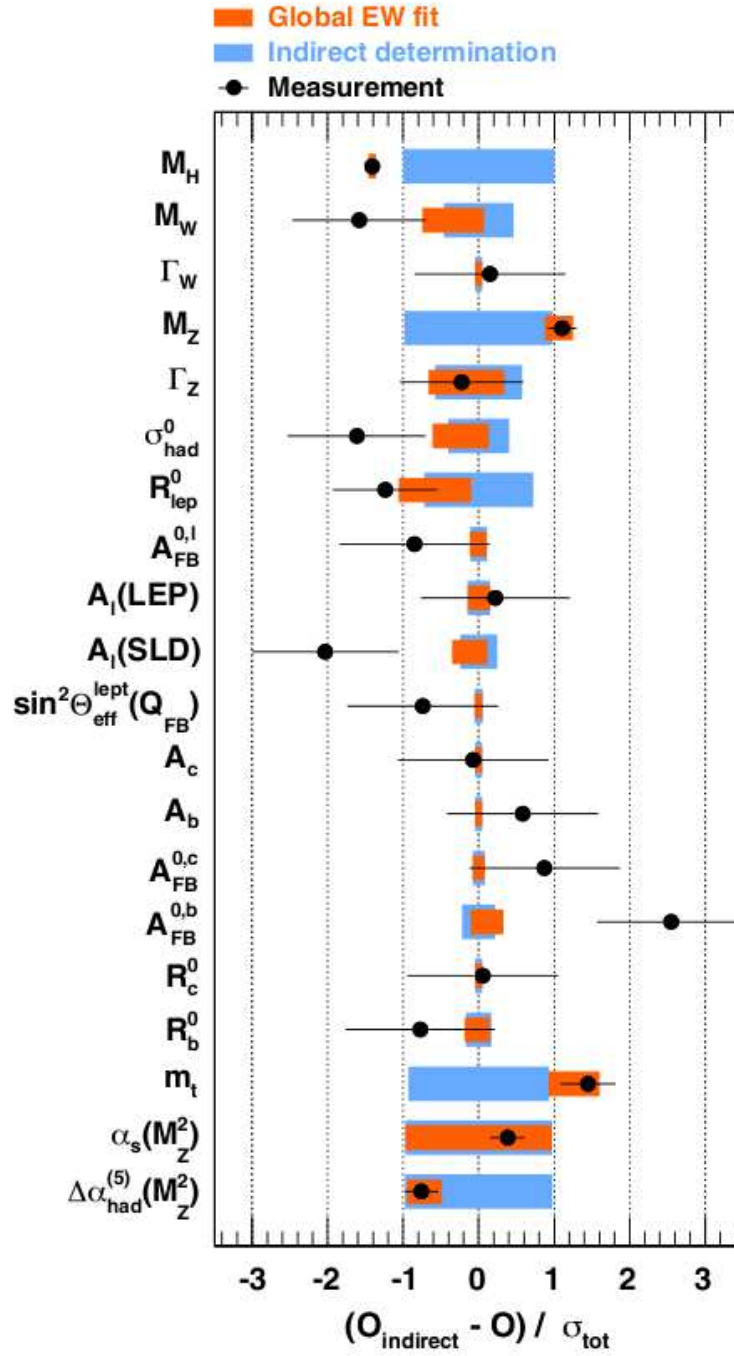


FIG. 7 – Résumé des différentes mesures des observables électrofaibles et comparaison avec les prédictions théoriques associées calculées par le groupe GFITTER. Celles-ci peuvent être obtenues par un ajustement global de toutes les mesures ou par détermination indirecte, correspondant à la prédiction d'un ajustement incluant toutes les mesures sauf celle de l'observable en question [42].

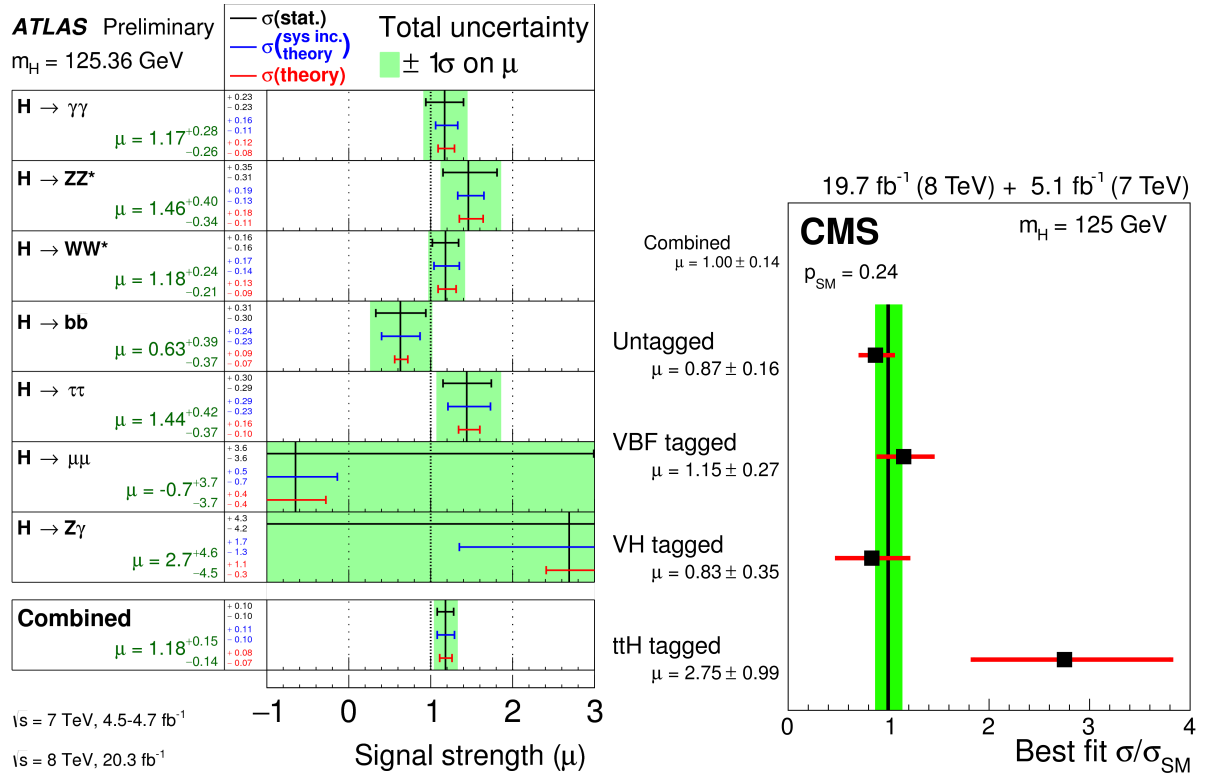


FIG. 8 – Mesures finales des modes de désintégration (gauche) [43] et de production (droite) [44] du boson de BEH effectuées respectivement par les expériences ATLAS et CMS avec les données du *run 1* du LHC. Les prédictions du modèle standard correspondent à une force de signal $\mu = 1$.

n'inclut pas la gravité, qui a une intensité 10^{17} fois plus élevée que l'interaction électro-faible.⁶ À cause de cette très grande différence d'échelle, un réglage fin des paramètres est nécessaire afin de garantir la stabilité du MS à haute énergie, ce qui constitue le problème dit de "naturalité".

D'autre part, le caractère chiral des courants chargés de l'interaction faible (voir équation 6) oblige à construire les fermions du modèle standard comme un doublet de $SU(2)_L$ et un singulet de $SU(2)_R$ afin de garantir le changement de saveur observé expérimentalement lors de l'émission ou de l'absorption d'un boson $W^\pm$, ce qui apparaît comme une construction *ad hoc*. De façon plus générale, il semble légitime de s'interroger sur la signification physique de certaines asymétries présentes à la fois dans le modèle standard et la nature, comme la différence de spin entre quanta d'interaction (bosons) et de matière (fermions),⁷ ou encore la différence de charge et de masse entre leptons ($\pm e$, $0 \text{ GeV} < m_l < 1.8 \text{ GeV}$) et quarks ($\pm \frac{e}{3}$, $\pm \frac{2e}{3}$, $0.002 \text{ GeV} < m_q < 173 \text{ GeV}$). Il est également intéressant de remarquer que certaines symétries présentes naturellement dans le MS, comme la symétrie matière/anti-matière apparaissant dans les solutions de l'équation de Dirac, sont largement violées dans la nature, ce qui semble difficile à intégrer dans le modèle de façon satisfaisante (mécanisme de violation de CP , voir section 1.2). Tous ces éléments pourraient être des indices d'une théorie plus large, dont le MS d'aujourd'hui ne serait qu'un cas limite.

D'autres aspects de la théorie restent à éclaircir comme la masse des neutrinos, plus naturellement nulle dans le modèle standard mais prouvées non-nulle dans les expériences d'oscillations, ou encore les observations en cosmologie et en astrophysique de l'effet d'une matière inconnue appelée "matière noire", qui n'a pas de particules associées dans le MS. Une liste plus complète peut être consultée dans les références [31, 32].

D'un point de vue expérimental, aucune recherche directe récente n'a permis l'observation de nouveaux phénomènes, bien que quelques anomalies restent à éclaircir comme par exemple l'excès modéré observé dans les spectres des paires de photons à la fois par les expériences ATLAS et CMS dans les premières données à $\sqrt{s} = 13 \text{ TeV}$ autour d'une masse de 750 GeV [8, 9].⁸ Cela a pour conséquence de réduire très fortement l'espace des paramètres d'un certain nombre de théories au-delà du modèle standard qui étaient favorisées jusqu'à récemment, comme le modèle supersymétrique minimum contraint (CMSSM) [46]. Il semble donc plus que jamais intéressant de rechercher des effets de nouvelle physique d'une part dans des processus encore peu étudiés ; d'autre part de façon plus robuste en s'affranchissant des dépendances aux différents scénarios de nouvelle physique. C'est ce qui est proposé dans la section 2.4, où on s'intéresse à des processus singuliers à plus d'un titre et encore peu étudiés à haute énergie : les interactions à quatre photons. En effet, ceux-ci apparaissent comme une très bonne sonde pour de nombreux modèles au-delà du MS, comme les modèles de Higgs composite [33] ou de dimensions supplémentaires de l'espace [34].

⁶Cette valeur se calcule très facilement en comparant l'échelle de Planck ($\simeq 10^{19} \text{ GeV}$) à l'échelle électrofaible ($\simeq 10^2 \text{ GeV}$).

⁷Cette aspect du modèle standard devient naturel dans le cadre de la théorie supersymétrique.

⁸Une liste plus complète des anomalies observées au LHC est disponible dans la référence [45].

2.4 Les couplages à quatre photons, une nouvelle sonde à haut potentiel pour des découvertes au LHC

2.4.1 Le processus de diffusion $\gamma\gamma$

La plupart du temps, deux faisceaux de photons rentrant en collision passent l'un à travers l'autre sans interagir. Cela peut se comprendre en étudiant les solutions des équations de Maxwell de l'électrodynamique classique, qui sont des ondes planes obéissant au principe de superposition. Cependant la QED, qui décrit avec une plus grande précision les interactions électromagnétiques grâce entre autres à la théorie quantique des champs (voir section 2.1), prédit une section efficace faible mais non nulle au processus $\gamma\gamma \rightarrow \gamma\gamma$, communément appelé diffusion $\gamma\gamma$ ou *light-by-light scattering* (LbyL) dans la littérature.

En effet, la diffusion $\gamma\gamma$ est rendue possible grâce à la mécanique quantique, qui permet à un photon de fluctuer pendant un instant très court en une paire particule/anti-particule chargée. L'une de ces particules peut alors absorber un deuxième photon. Lorsque la paire se recombine, deux photons sont à nouveau émis par l'annihilation matière/antimatière. L'interaction globale apparaît comme une diffusion élastique mettant en jeu un terme de couplage à quatre photons (4γ). L'étude de la section efficace de ce processus a donné lieu à des discussions théoriques très fructueuses tout au cours du XX^e siècle, d'abord à basse énergie (< 2 GeV) lors de la découverte progressive de la mécanique quantique et de la QED [47–51], puis à plus haute énergie pour son potentiel à contraindre les modèles de nouvelle physique [7, 52–59].

Les diagrammes de Feynman de LbyL à l'ordre de plus bas de la théorie des perturbations (LO, pour *leading order*) sont montrés dans la figure 9, gauche. Ils impliquent une boucle de fermions chargés (quarks ou e, μ, τ) ou de bosons chargés $W^\pm$, et permettent de calculer avec précision la section efficace prédite par le modèle standard aux énergies du LHC.⁹ En effet, les corrections QED et QCD à un ordre supérieur (NLO, pour *next-to-leading order*) ont été évaluées à la fois pour les boucles de bosons et de fermions [54]. Leur contribution à la section efficace est faible ($\simeq 1\%$) et sera négligée par la suite. Le processus $\gamma\gamma \rightarrow H \rightarrow \gamma\gamma$ contribue également à la section efficace LO, mais est totalement négligeable devant les autres diagrammes [60]. En revanche, le vertex $Z\gamma\gamma$ est interdit par le modèle standard par conservation de l'hélicité (théorème de Young-Landau). Les formules complètes des amplitudes invariantes du processus de diffusion $\gamma\gamma$ à une boucle (LO) incluant les différents cas limites sont disponibles dans l'appendice A de l'une des publications de thèse [57]. Les formules complètes des amplitudes invariantes à deux boucles sont disponibles dans [54].

Il est également important de mentionner l'existence d'un processus QCD menant au même état final mais initié par des gluons, représenté dans la figure 9, droite. Dans ce cas, uniquement les quarks peuvent circuler dans la boucle grâce à leur couplage à la fois à la QED et à l'interaction forte (voir section 3).

Le processus de diffusion $\gamma\gamma$ n'a jamais pu être observé directement à cause de sa section efficace très faible. En effet, la sensibilité des expériences dédiées, qui sondent majoritairement les basses énergies, est encore de nombreux ordres de grandeurs en-

⁹Pour les énergies inférieures à 8 GeV, des corrections hadroniques importantes dues aux effets de confinements et aux résonances de mésons (voir section 3) doivent être prises en compte [54].

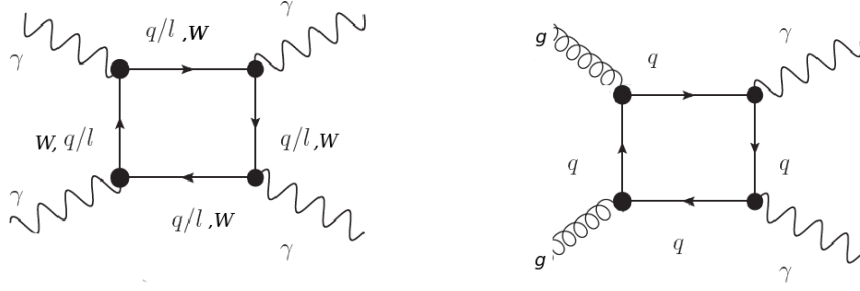


FIG. 9 – Diagrammes de Feynman correspondant au processus de diffusion $\gamma\gamma$ à l'ordre le plus bas de la théorie des perturbations prédit par la QED (gauche). Processus QCD similaire initié par des gluons (droite).

dessous des prédictions QED [61]. Cependant, son observation indirecte à basse énergie a déjà été réalisée de façon très précise par l'intermédiaire de certains tests de précision, comme la mesure du moment gyromagnétique de l'électron [62], sensible aux boucles d'électrons, ou du muon [63], sensible aux boucles de muons, d'électrons et de hadrons. Plusieurs autres expériences ont été proposées pour mesurer directement la LbyL à basse énergie, par exemple en utilisant des lasers de haute puissance [64]. Un résumé récent des différentes propositions est donné succinctement dans [64].

Par ailleurs, l'étude de deux processus proches est rapportée dans la littérature : la séparation de photons dans un champ magnétique fort [65] et la diffusion de photons dans le champ coulombien d'un noyau (diffusion de Delbrück) [66, 67]. Seul le second a été clairement observé, pour des énergies allant jusqu'à 7 GeV. Ces processus diffèrent de la LbyL classique dans le sens où deux des quatre photons de l'interaction sont issus de fluctuations quantiques (photons virtuels).

À plus haute énergie quand les diagrammes présentés dans la figure 9 (gauche) décrivent la section efficace de façon correcte, très peu de contraintes existent sur la diffusion $\gamma\gamma$. Un collisionneur photon-photon en serait l'observatoire idéal, comme le mentionnent plusieurs auteurs [52, 68]. Cependant la construction d'une telle machine est très incertaine. Il semble alors pertinent de se poser la question de la faisabilité d'une mesure au LHC.

Dans la section 2.4.2, les motivations pour une mesure directe à haute énergie de la LbyL sont énoncées brièvement et remises dans le contexte de la recherche de couplages anormaux entre bosons électrofaibles démarrée au LEP (CERN). La section 2.4.3 introduit ensuite le travail phénoménologique de la partie V, qui a consisté à estimer le potentiel d'observation du processus de LbyL dans les collisions proton-proton du LHC et à calculer sa sensibilité aux modèles de nouvelle physique. Pour cela, le LHC est utilisé effectivement comme un collisionneur photon-photon en s'appuyant sur le courant électromagnétique important induit par les protons lors de leur croisement.

2.4.2 Motivations pour une mesure directe

Tout d'abord, la mesure de la diffusion $\gamma\gamma$ à haute énergie constituerait un nouveau test de précision de la QED. Celle-ci est pour le moment très bien vérifiée par l'expérience (voir section 2.3), à quelques exceptions près comme par exemple la mesure du

moment gyromagnétique du muon déjà mentionnée au paragraphe précédent, qui bien que déterminée avec une précision impressionnante (environ 0.5 parties par million), se situe à plus de trois déviations standards des prédictions théoriques [63]. Celles-ci incluent des corrections quantiques poussées de la QED (jusqu'à quatre boucles), mais aussi des interactions faibles et nucléaires fortes, sur lesquelles repose a priori la majorité des incertitudes théoriques.

La suppression élevée du processus de LbyL dans le MS en fait une sonde particulièrement appropriée pour la recherche de nouvelle physique [7, 52–57, 59]. En effet, la section efficace tendant très vite vers zéro à haute énergie (voir partie V), la détection de quelques événements à une énergie suffisamment haute pourrait suffire à revendiquer la découverte d'un signal de physique au-delà du MS. De plus, les couplages à 4γ en général sont sensibles à une gamme très large de nouveaux phénomènes, toute particule chargée ou nouvelle résonance de spin zéro ou deux constituant une contribution supplémentaire au couplage. Notons qu'il est important dans ce cas de prendre en compte les boucles de bosons $W^\pm$ afin d'avoir une prédiction modèle standard correcte, car celles-ci dominent la section efficace à haute énergie du fait de la masse élevée du W ($m_W \simeq 80$ GeV) [57]. Tout cela sera démontré explicitement dans la partie V.

D'un point de vue historique, la recherche de couplages anormaux entre les bosons du secteur électrofaible a été démarré au LEP (CERN), et s'est poursuivie en utilisant les collisions de protons du Tevatron (Fermilab) et maintenant du LHC (CERN) [69]. De façon générale, ces recherches sont fortement motivées par le fait que les couplages entres γ , $W^\pm$ et Z^0 sont prédits de façon très précise par la théorie électrofaible (voir section 2.3). Ainsi, une légère déviation pourrait suffire à affecter significativement la dynamique de la brisure de symétrie sous-jacente (voir section 2.2), et donc à terme améliorer notre compréhension du rôle que jouent les champs scalaires dans la nature.

Techniquement, les couplages anormaux entre bosons sont le plus souvent introduits dans les équations de la façon la plus générale possible, en ajoutant à la main les premiers opérateurs de dimensions supérieures préservant les symétries du modèle standard dans le lagrangien. Les couplages de ces nouveaux termes sont alors paramétrés sous forme de réels, prenant une valeur nulle dans le cas du MS mais non-nulle pour un certain nombre de modèles de nouvelle physique. Par exemple, pour les couplages anormaux trilineaires (aTGC, *anomalous triple gauge couplings*) $WW\gamma$, deux opérateurs supplémentaires de dimensions six sont généralement introduits, ce qui modifie le lagrangien correspondant de la façon suivante [70] :

$$\mathcal{L}_{\text{eff}}^{WW\gamma} = -ie[(W_{\mu\nu}^+ W^{-\mu} - W^{+\mu} W_{\mu\nu}^-) A^\nu + (1 + \Delta\kappa_\gamma) W_\mu^+ W_\nu^- A^{\mu\nu} + \frac{\lambda_\gamma}{m_W^2} W_\mu^{+\nu} W_\nu^{-\rho} A_\rho^\mu], \quad (16)$$

où $\mathcal{L}_{\text{eff}}^{WW\gamma}$ est le lagrangien effectif total du couplage $WW\gamma$ et m_W est la masse des bosons W . Les A_μ et $W_\mu^\pm$ sont les tenseurs correspondant au photon et aux bosons $W^\pm$, respectivement, et $V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu$. Le premier terme de la somme donne le vertex $W^+W^-\gamma$ familier de la QED. $\Delta\kappa_\gamma$ et λ_γ correspondent aux couplages anormaux introduisant les deux opérateurs supplémentaires, égaux à zéro dans le MS. La dérivation du lagrangien effectif permet de calculer les sections efficaces pour différentes valeurs de couplages anormaux, qui peuvent ensuite être comparées aux données.

Il est important de souligner qu'en utilisant ce formalisme, le lagrangien modifié n'est plus qu'un lagrangien effectif, valide uniquement lorsque les effets de nouvelle physique sont suffisamment faibles pour être traités comme des petites corrections en théorie des perturbations. En effet, lorsque l'énergie de l'interaction devient trop proche de l'échelle de nouvelle physique, les effets de seuils nécessitent un modèle plus complet donnant le détail des opérateurs supplémentaires pour pouvoir produire des prédictions de sections efficaces fiables (limite de découplage). Il convient également d'évaluer la section efficace uniquement pour des valeurs de $\Delta\kappa_\gamma$ et λ_γ préservant l'unitarité de la mécanique quantique (limite d'unitarité).

Les couplages anormaux trilineaires $WW\gamma$, $Z\gamma\gamma$ ou $ZZ\gamma$ sont déjà fortement contraints par les différentes expériences de collisionneurs, comme le montre le tableau de résumé présenté dans la figure 10. Les couplages anormaux quartiques (aQGC, *anomalous quartic gauge couplings*) sont en revanche moins contraints, car plus délicats à sonder. Ainsi, seuls les couplages $\gamma\gamma WW$ et γZWW sont contraints directement par les expériences de collisions (voir figure 11), ce qui motive d'autant plus la recherche de couplages anormaux à 4γ .

Dans le cas d'une découverte de couplages anormaux à quatre photons, il pourrait être intéressant de tenter de faire un parallèle avec d'autres phénomènes mettant en jeu des photons observés en physique atomique, comme le paradoxe de Einstein-Podolsky-Rosen et la violation des inégalités de Bell, confirmée depuis 1982 par l'expérience d'Alain Aspect [71]. En effet, il est désormais établi que le phénomène d'intrication quantique implique certaines tensions avec la théorie relativiste dans sa formulation actuelle, ce qui pourrait être la porte d'entrée vers une physique au-delà du modèle standard. En particulier, l'existence de dimensions supplémentaires de l'espace pourrait nous amener à revoir la notion des distances pour certaines particules. La grande différence entre les échelles en énergies considérées en physique atomique et en physique des hautes énergies rendent néanmoins l'établissement d'un lien entre les deux domaines délicat.

Dans la section 2.4.3, l'étude phénoménologique principale effectuée pendant les années de thèse est introduite. Celle-ci consiste à estimer le potentiel d'observation de la diffusion $\gamma\gamma$ dans les collisions proton-proton du LHC et à calculer sa sensibilité aux modèles de nouvelle physique. Elle sera abordée en détail dans la partie V.

2.4.3 Sensibilités au LHC

L'étude de la sensibilité des expériences *ATLAS-Forward Proton Detector* (AFP) [72] et *CMS-TOTEM-Precision Proton Spectrometer* (CT-PPS) [73] du LHC (voir partie II) aux couplages anormaux à quatre photons représente l'un des travaux principaux effectué pendant les années de thèse, et sera abordée en détail dans la partie V. Tout d'abord, les contributions prédites par le MS sont revisitées et une mesure potentielle du processus de LbyL dans les collisions proton-proton du LHC est examinée [57]. Cela fait écho à une mesure similaire récemment proposée dans la littérature à partir des collisions Plomb-Plomb du LHC [59].

Ensuite, l'étude phénoménologique de la partie V considère les nouvelles données de collisions proton-proton du LHC actuellement en cours d'acquisition (voir partie II) afin d'évaluer le potentiel de découverte de couplages anormaux à 4γ au LHC. En particulier, elle considère le canal de production exclusif de di-photons avec détection des protons

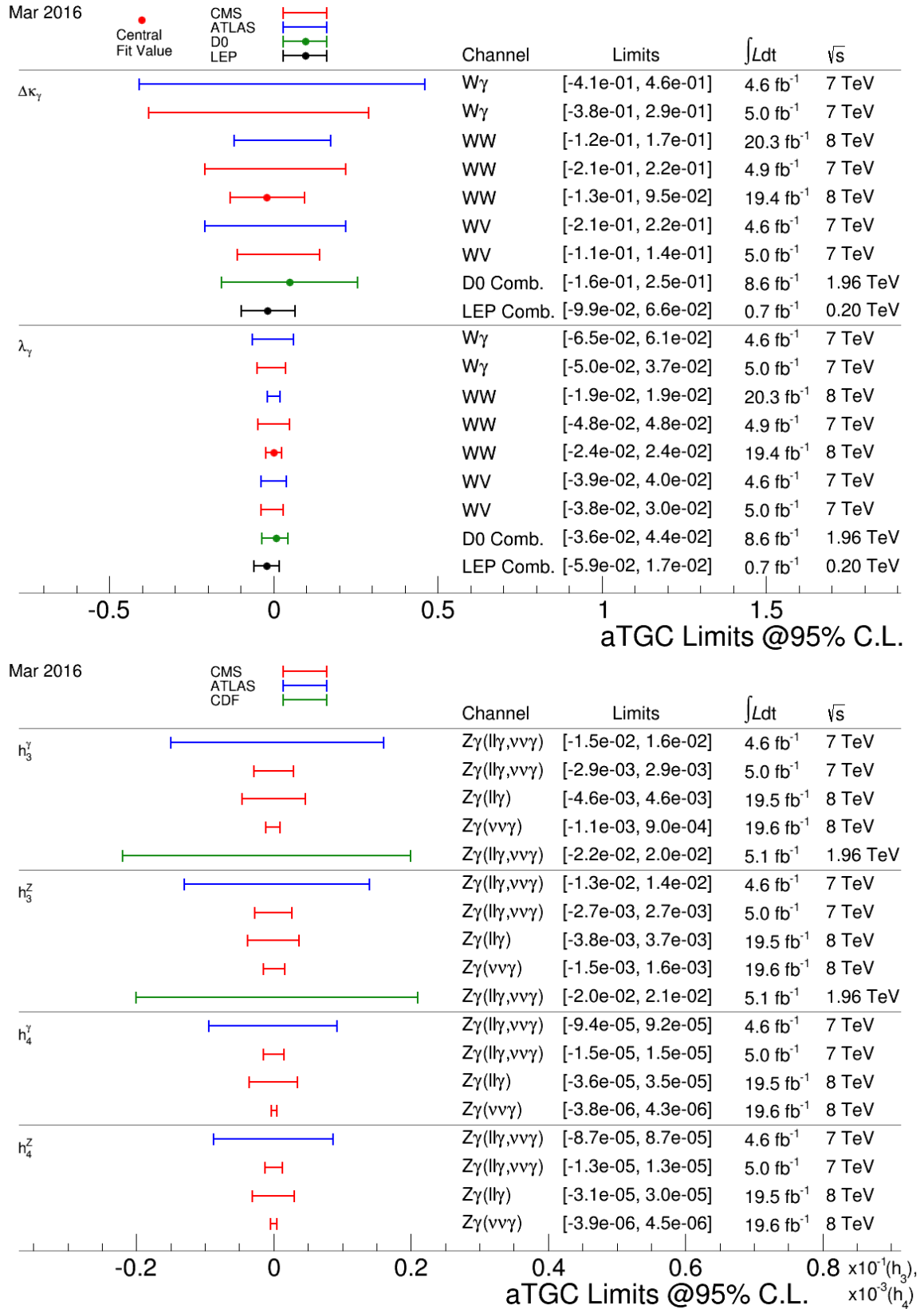


FIG. 10 – Limites actuelles sur les couplages anormaux $WW\gamma$ (haut) et $Z\gamma\gamma$, $ZZ\gamma$ (bas) en fonction de différentes expériences : CMS, ATLAS (LHC), DØ (Tevatron) et la combinaison des différents détecteurs du LEP. Les couplages anormaux sont paramétrés par différents termes sans dimension mentionnés dans la colonne la plus à gauche des tableaux [69].

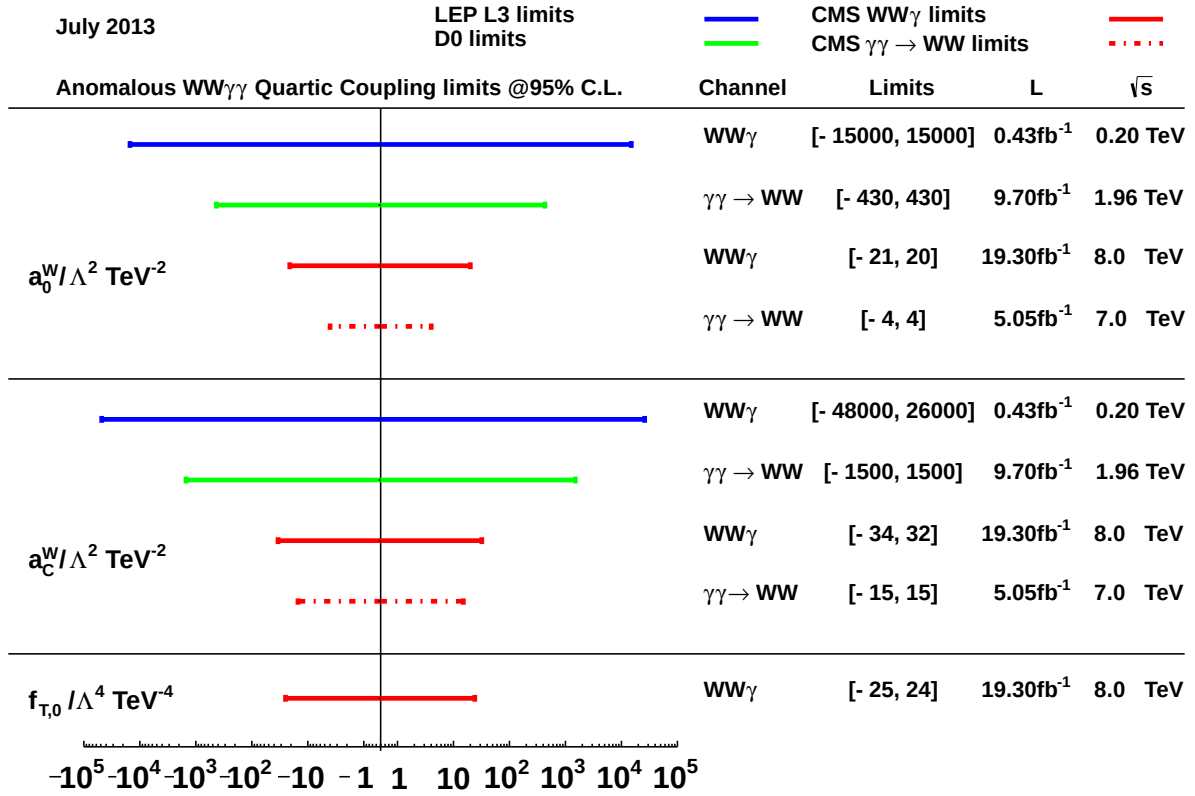


FIG. 11 – Limites actuelles sur les couplages anormaux quartiques $\gamma\gamma WW$ en fonction de différentes expériences : CMS (LHC), DØ (Tevatron) et L3 (LEP). Les couplages anormaux sont paramétrés par différents termes mentionnés dans la colonne la plus à gauche des tableaux. L'expérience ATLAS a également placé des limites en étudiant l'état final $W\gamma\gamma$ au LHC, comparable à celle de CMS. Des limites posées par CMS existent également sur γZWW et sont disponibles dans la référence [69].

intacts à l'avant, qui a l'avantage de contenir très peu de bruit de fond après coupures par rapport aux canaux de recherche classiques, généralement inclusifs. Le terme exclusif signifie que seul deux photons sont produits lors de l'interaction et rien d'autre, les protons restant intacts. Des études basées sur des canaux similaires pour la recherche des couplages anormaux entre photons et bosons $W^\pm$ ou Z^0 sont déjà présentes dans la littérature [74, 75]. Le diagramme au LO de la production exclusive de paires de photons au LHC par LbyL est représenté dans la figure 12. La détection de protons à l'avant au LHC et la physique à l'avant en général seront introduites dans la dernière section de la partie II.

Partant d'une théorie effective basée sur des opérateurs de dimensions supérieures (voir section 2.4.2) pour simuler les couplages anormaux à 4γ [56], les amplitudes complètes sont ensuite produites dans le cas générique de l'existence de nouvelles particules chargées [57] ou de nouvelles résonances neutres [58]. Ces nouvelles productions donnent lieu à de nouveaux diagrammes de LbyL, présentés dans la figure 13. Cette approche permet d'obtenir des limites largement indépendantes des différents modèles de nouvelle physique. De plus, le calcul des amplitudes complètes permet de garantir la validité des limites quelque soit l'échelle en énergie de la nouvelle physique considérée.

Le potentiel de découverte estimé est très largement compétitif avec les limites ac-

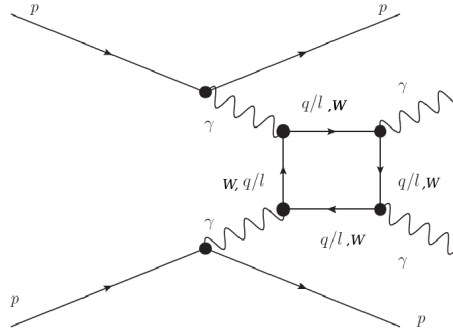


FIG. 12 – Production exclusive de paires de photon par diffusion $\gamma\gamma$. Les protons rentrant en collision sortent intact de l'interaction et peuvent être détectés par des détecteurs de protons à l'avant au LHC.

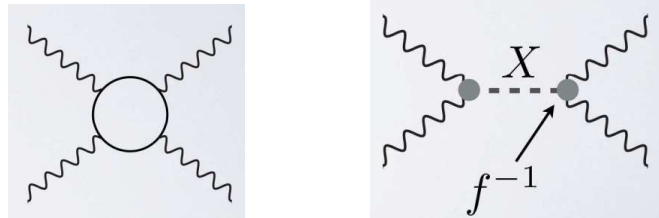


FIG. 13 – Contributions typiques de nouvelles particules chargées (gauche) et résonances neutres (droite) au processus de diffusion $\gamma\gamma$. Ces dernières font apparaître un nouveau couplage f^{-1} associé au vertex $\gamma\gamma X$.

tuelles et même à venir au LHC, en particulier pour certains modèles au-delà du MS introduisant des dimensions supplémentaires de l'espace. Des limites complémentaires aux recherches directes peuvent également être établies pour une large gamme de modèles, incluant par exemple tout nouveau phénomène impliquant l'existence de nouvelles particules chargées ou de bosons scalaires neutres. Cette étude a abouti à plusieurs publications grâce à une collaboration fructueuse avec plusieurs physiciens expérimentateurs et théoriciens [56–58].

Dans la section suivante, quelques propriétés de l'interaction nucléaire forte sont rappelées afin d'introduire le deuxième travail principal effectué pendant les années de thèse, constitué par une mesure de précision QCD avec les données du détecteur ATLAS au LHC.

3 L'interaction nucléaire forte et la structure des hadrons

3.1 La chromodynamique quantique perturbative et non-perturbative

L'interaction forte décrit les couplages entre les partons, formés par les quarks et les gluons. Elle s'appuie sur la chromodynamique quantique ou QCD, qui est une théorie quantique des champs basée sur le groupe de jauge $SU(3)_C$ [10, 11, 19]. Sa charge associée est un nombre quantique appelée couleur (C), qui peut être de trois types : R , G ou B , rappelant les initiales des trois couleurs primaires optiques en anglais *red*, *green* et *blue*, avec la propriété supplémentaire $R + G + B = 0$ (blanc). La théorie se construit avec des quarks (anti-quarks) portant une unique charge de couleur (anti-couleur) : R , G , B ($\bar{R}$, $\bar{G}$, $\bar{B}$). Le groupe $SU(3)_C$ est non-abélien, ce qui signifie que les bosons vecteurs associés, appelés gluons, portent eux-mêmes une charge de couleur. Cela leur permet d'interagir également entre eux par interaction forte.¹⁰ Le groupe $SU(3)_C$ comportant huit générateurs, huit gluons suffisent à en définir une base. Par invariance de jauge, ils portent une charge de couleur correspondant à une couleur et une anti-couleur, par exemple : $R\bar{G}$, $B\bar{G}$, etc. Ils sont aussi nécessairement de masse nulle pour la même raison (voir section 2.2).

Le caractère non-abélien de la QCD, associé au nombre élevé de gluons dû à la structure du groupe $SU(3)_C$, mène à deux propriétés très importantes car uniques dans le MS : la liberté asymptotique et le confinement. Ces deux propriétés sont des conséquences du comportement de la constante de couplage de l'interaction à haute et basse énergie. Celui-ci est obtenu par l'écriture des corrections d'ordre supérieur en théorie des perturbations. Néanmoins, il peut se sentir facilement à l'aide d'une image simple, présentée dans la figure 14.

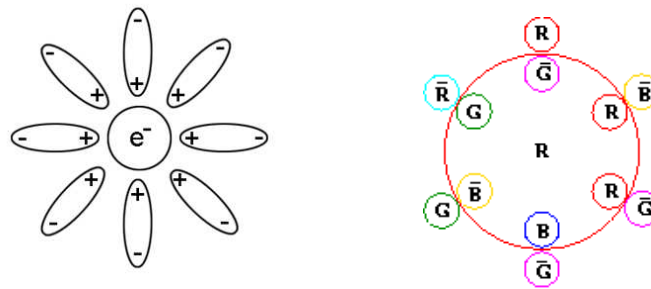


FIG. 14 – Illustration du phénomène d'écrantage en QED (gauche) [76] et anti-écrantage en QCD (droite) [77] à courte distance.

Commençons par le cas plus simple de la QED en imaginant un électron isolé. Sa charge électrique va déformer localement le champ électromagnétique aux alentours, avec la création/annihilation d'une multitude de photons pendant des instants très courts. Ces productions sont autorisées en mécanique quantique grâce au principe d'incertitude d'Heisenberg. En QED, les photons peuvent fluctuer en paires particules/anti-particules

¹⁰La théorie électrofaible est également une théorie non-abélienne, mais la QED est abélienne car le photon est une particule neutre (voir section 2).

chargées uniquement. Il est donc possible de les considérer comme des dipôles électriques qui s'orientent naturellement dans le champ créé par l'électron, comme le montre la figure 14, gauche. Maintenant considérons deux interactions, la première à longue distance (donc à basse énergie) et la seconde à plus courte distance (donc à plus haute énergie). Dans le cas d'une interaction suffisamment loin des différentes charges en présence, la charge apparente correspond à celle de l'électron. Au contraire, si l'interaction se produit avec une particule située à l'intérieur du nuage de photons, les charges positives des photons-dipôles vont avoir pour effet d'augmenter la charge apparente. C'est l'effet d'écrantage de la QED, qui a pour conséquence la (légère) croissance de la constante de couplage électromagnétique α_{em} avec l'énergie de l'interaction. Ainsi, α_{em} est environ 10% plus élevée à la masse du Z ($\simeq 91$ GeV) par rapport à sa valeur mesurée à basse énergie (limite de Thomson) [42].

Considérons maintenant une configuration similaire en QCD. Les gluons portent une charge de couleur, ce qui leur permet de fluctuer non seulement en paires quark/anti-quark, mais également en d'autres paires de gluons de différentes couleurs. Cela a pour effet de diluer davantage la charge initiale. Dans l'exemple de la figure 14 (droite), un quark rouge est entouré de gluons bicolores. Cette fois, il s'avère que les productions supplémentaires de gluons renversent la situation précédente. Ainsi, la charge rouge du quark s'entoure préférentiellement d'autres charges rouges. La charge apparente est alors diluée à courte distance, et donc augmente sur les plus longues distances.

Une compilation de mesures de la constante de couplage de l'interaction nucléaire forte α_s à différentes énergies d'interaction est montrée dans la figure 15 [78]. Comme prédit par la QCD, α_s décroît avec l'énergie de l'interaction, ce qui donne la propriété de liberté asymptotique à haute énergie permettant de faire des calculs perturbatifs en puissance de α_s . Ce domaine est appelé QCD perturbative (pQCD). En revanche, la constante de couplage devient très importante à basse énergie. En dessous d'un certain seuil appelé échelle de confinement Λ_{QCD} , les développements perturbatifs ne sont plus possibles car la série associée diverge.¹¹ C'est le domaine non-perturbatif, où les calculs théoriques deviennent très délicats et sont en général moins précis. L'échelle de confinement varie en fonction des processus, mais se situe typiquement autour de quelques centaines de MeV.

En particulier, le phénomène de confinement implique que les quarks, gluons ou toutes autres particules colorées ne peuvent pas être observés à l'état libre, car leur charge de couleur ferait diverger leur énergie. On dit alors que les états asymptotiquement stables de la QCD sont nécessairement des singulets de $SU(3)_C$, appelés hadrons. La structure du groupe $SU(3)_C$ prévoit deux types de hadrons principaux : les mésons, constitués dans leur état fondamental par une paire $q\bar{q}$, et les baryons, correspondant à qqq . Un exemple de baryon bien connu est le proton, qui peut être considéré dans son état fondamental comme un état lié par l'interaction forte des quarks uud .

Certains états liés plus exotiques, comme des boules de gluons ou des pentaquarks $q\bar{q}qqq$, ne sont pas exclus par la théorie. Des observations en accord avec l'existence de deux résonances de type pentaquarks ont d'ailleurs été reportées récemment par l'expérience LHCb au LHC [80].

¹¹La conjecture de Maldacena [79] permet cependant d'établir dans un certain nombre de cas des correspondances entre théories quantiques à champ fort et théories de type relativité générale, aussi connue sous le nom de correspondances AdS/CFT (*anti-de Sitter/conformal field theory*).

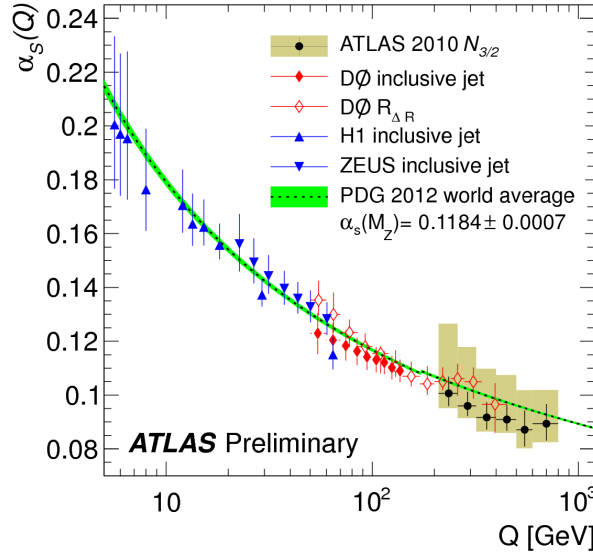


FIG. 15 – Mesure de α_s à différentes énergies dans les collisionneurs. Toutes les mesures sont en bon accord avec la prédiction de l'évolution du couplage donné par la QCD [78].

Dans les sections suivantes, certaines propriétés ou phénomènes observés en QCD utiles pour comprendre la modélisation des interactions protons-protons au LHC sont explicitées. Dans la section 3.2, le processus d'hadronisation relié à la fragmentation des partons et d'autres effets de QCD non-perturbatives jouant un rôle dans les collisionneurs hadroniques sont décrits brièvement. Dans la section 3.3, des rappels sont faits sur la structure du proton et le théorème de factorisation, largement utilisés dans le calcul des sections efficaces hadroniques à haute énergie. Le principe de ces calculs est également abordé. Dans la section 3.4, quelques problèmes importants de divergence inhérents aux calculs des ordres supérieurs en pQCD sont mentionnés.

Finalement, une ouverture sur les tests de précision de la QCD dans les collisionneurs hadroniques à l'aide de photons est présentée dans la section 3.5. Ce dernier paragraphe anticipe l'analyse de données principale effectuée pendant les années de thèse pour la mesure de la section efficace inclusive de production de di-photon dans l'expérience ATLAS au LHC. Celle-ci sera abordée plus en détail dans la partie IV.

3.2 Hadronisation des partons et autres effets de QCD non-perturbative observés dans les collisionneurs hadroniques

Dans les collisionneurs hadroniques comme le LHC, des hadrons sont envoyés les uns contre les autres à des énergies phénoménales. Les hadrons étant constitués de quarks et de gluons (voir section 3.1), il paraît légitime de se demander ce qu'il se passe lorsqu'un parton (quark ou gluon) est déconfiné par la force du choc.

Le déconfinement d'un parton nécessite une quantité importante d'énergie à cause de la valeur élevée de la constante de couplage (voir section 3.1). Il est donc raisonnable de penser que l'interaction parton-parton principale puisse être décrite dans le cadre de la pQCD. Cependant, nous avons vu dans la section précédente qu'un parton libre ne peut constituer une observable à l'état libre du fait de sa charge de couleur

non nulle et de l'augmentation de la constante de couplage de l'interaction forte avec la distance. Ainsi, une fois éjecté du hadron et au fur et à mesure de son éloignement du point d'interaction initial, l'énergie de liaison va augmenter. Le vide entre les différentes particules colorées va se polariser à cause des courants de couleurs croissants. Très vite ($\simeq 10^{-23}$ s.), ceux-ci vont devenir non-perturbatifs et suffisants pour créer des partons supplémentaires à partir du vide pour compenser les différentes charges de couleur, c'est le phénomène d'hadronisation (ou fragmentation). Expérimentalement, des jets de particules collimatés, principalement hadroniques et appelés *jets* par la suite, sont observés à la place des partons. Un exemple de collision proton-proton avec deux jets dans l'état final enregistrée par l'expérience ATLAS au LHC est montré dans la figure 16.

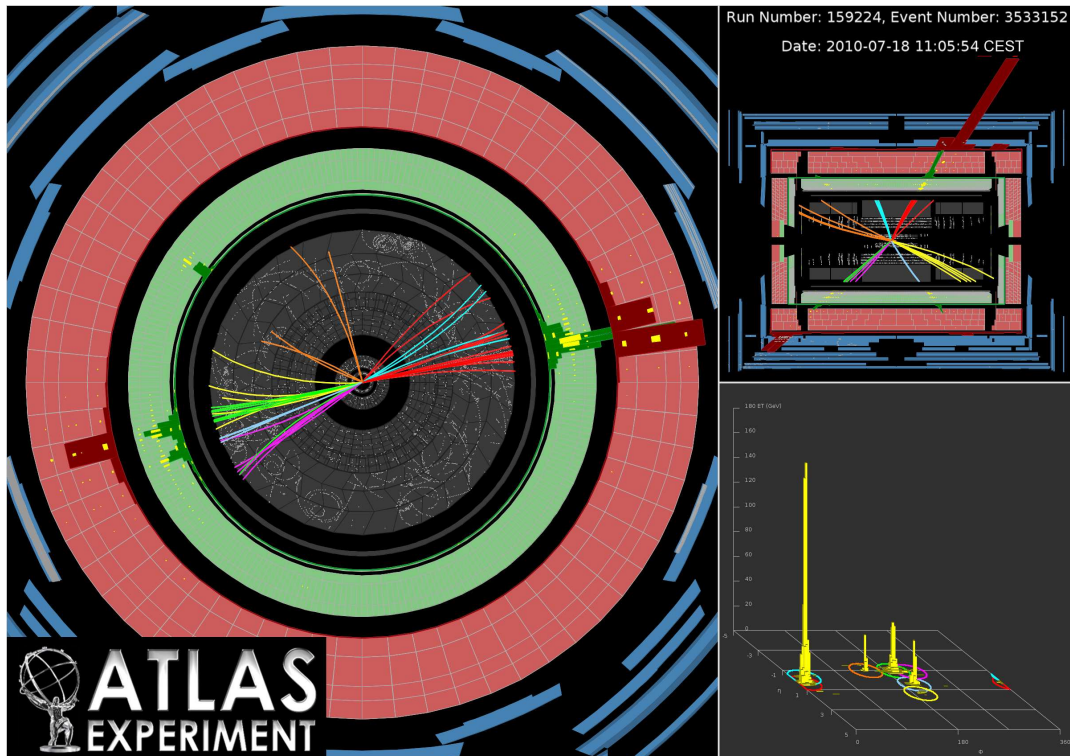


FIG. 16 – Collision proton-proton avec une énergie dans le centre de masse à 7 TeV enregistrée en 2010 par l'expérience ATLAS au LHC. Deux jets avec de hautes impulsions transverses (1.12 TeV et 480 GeV) sont visible dans l'état final [81].

Le processus d'hadronisation n'est pas totalement compris à l'heure actuelle à cause de sa nature non-perturbative, ce qui en fait souvent une source d'incertitudes conséquentes, à la fois dans les prédictions théoriques et dans les mesures. Il peut être étudié grâce à des simulations Monte Carlo (MC), où il est souvent décomposé en deux étapes :

- ◇ **Cascades partoniques** : simulation des cascades de particules secondaires de plus basses énergie formées à partir des différents partons de haute énergie jouant un rôle dans l'interaction principale. Celles-ci peuvent être électromagnétiques et/ou hadroniques. En principe, elles représentent des corrections d'ordres supérieurs par rapport au processus principale et peuvent donc être décrites par pQCD. En pratique, elles sont approximées à l'aide de générateurs de Monte Carlo (MC) et s'arrêtent à une échelle arbitraire en énergie (*cutoff*).

- ◇ **Formation des hadrons** : recombinaison des différents partons de haute et plus basse énergie en hadrons afin d'avoir un état final non coloré. Cette étape est basée sur des modèles semi-empiriques non-perturbatifs.

L'hadronisation se base sur des modèles semi-empiriques et la cascade de parton peut être formulée de plusieurs manières, à l'image du modèle de dipôle et de reconnexion de couleur de Lund utilisé par le générateur MC PYTHIA8 [82] ou des deux modèles concurrents implémentés dans les générateurs MC SHERPA [83] et HERWIG [84, 85]. Les différents algorithmes prennent en compte les mesures des fonctions de fragmentation pour régler leurs paramètres, correspondant à la probabilité pour un parton d'émettre un hadron d'un certain type avec une cinématique donnée. Ces fonctions peuvent être considérées comme universelles grâce au théorème de factorisation (voir section 3.3). Il est également nécessaire de fixer une échelle en énergie arbitraire μ_F servant à déterminer la limite entre les calculs de QCD perturbative et non-perturbative (*cutoff*).

Un autre effet de QCD non-perturbative intervenant lors de collisions proton-proton comme celles du LHC réside dans les événement sous-jacents, correspondant à des interactions additionnelles avec d'autres partons appartenant aux deux hadrons principaux rentrant en collision, en plus de l'interaction à très haute énergie que l'on vise à étudier. Les événements sous-jacents sont modélisés de façon semi-empirique dans les générateurs PYTHIA, SHERPA et HERWIG.

Un deuxième effet non directement relié à la QCD non-perturbative mais qui en dépend néanmoins fortement est l'empilement ou *pile-up*, provenant d'interactions proton-proton additionnelles à basse énergie (et donc non décrites par la pQCD) ayant lieu lors des croisements de faisceaux (voir partie II). L'empilement dépend en effet fortement à la fois de la section efficace inélastique totale proton-proton et des paramètres de collisions, incluant en particulier la luminosité instantanée (voir section 4). Les principaux effets de QCD devant être pris en compte lors de la modélisation d'une interaction proton-proton au LHC sont résumés dans la figure 17.

3.3 Structure du proton et calculs de sections efficaces hadroniques au LHC

Il paraît maintenant nécessaire de donner quelques détails supplémentaires sur l'interaction parton-parton principale, a priori décrite par la pQCD quand l'énergie d'interaction est suffisante pour déconfiner les partons. Cette interaction sera qualifiée de "dure" par la suite, en référence au transfert d'énergie important qu'elle met en oeuvre.

Le premier point concerne le choix des partons de l'interaction dure. Dans son état fondamental, le proton est constitué de trois quarks uud échangeant des gluons. Cependant, à très haute énergie sa structure évolue avec l'apparition d'une mer de quarks issue de la séparation de gluons $g \rightarrow q\bar{q}$ (*sea quark*) et d'une multitude de gluons supplémentaires par les processus de type $g \rightarrow gg$. Dans les faits, sa structure est paramétrisée par des fonctions de structure appelées PDFs (*parton density functions*) [87–89]. Celles-ci donnent la probabilité de présence d'un parton d'un certain type ($u, d, \dots, g$) en fonction de la norme de sa quadri-impulsion Q et de la fraction d'impulsion du proton qu'il porte x , de façon assez similaire aux fonctions de fragmentation mentionnées dans la

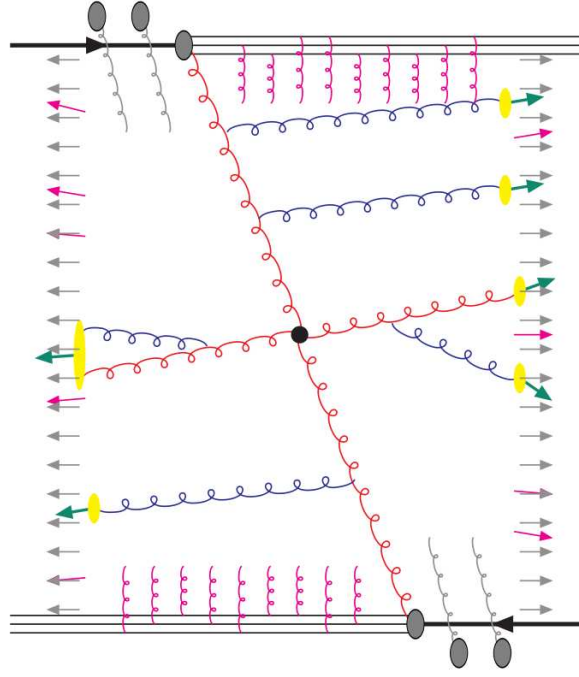


FIG. 17 – Principaux effets de QCD devant être pris en compte lors d’une interaction proton-proton au LHC, en plus de l’interaction principale entre deux partons décrit par la QCD perturbative (en rouge) : cascade de partons (bleu), hadronisation (jaune et vert), événements sous-jacent (magenta) et empilement (gris) [86].

section précédente. Les PDFs encapsulant des effets non-perturbatifs difficilement calculables, des données sont nécessaires pour les déterminer. Néanmoins, leur évolution en fonction de Q (pour Q assez haut) peut être déduite des équations de Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP), basées sur la pQCD.¹² Les PDFs sont donc finalement déterminées par des ajustements (*fits*) globaux à partir de données collectées dans différents collisionneurs, sous l’hypothèse que leur évolution en fonction de Q obéit à des équations de type DGLAP. Un exemple d’évolution de la structure du proton en fonction de x à $Q = 1$ GeV et $Q = 100$ GeV publié par la collaboration CT14 est donné dans la figure 18 [87]. À bas Q et grand x les quarks de valence uud dominent, alors qu’à plus grand Q et plus petit x les quarks de mers et surtout les gluons affectent très fortement les fonctions de structure.

Ainsi, la section efficace fiducielle totale de production d’un état final X donné à partir de deux protons peut s’écrire comme le produit entre les PDFs et les différentes sections efficaces calculées au niveau partonique ($qq \rightarrow X$, $qg \rightarrow X$, $gg \rightarrow X$) :

$$\sigma_{\text{fid}} = \sum_{i,j} \int_{Q,x} f_i(Q_1, x_1) f_j(Q_2, x_2) \int_{\Omega} \frac{d\sigma_{ij \rightarrow X}}{d\Omega} d\Omega, \quad (17)$$

où les fonctions $f(Q_k, x_k)$ représentent les PDFs des deux protons rentrant en collision et $\frac{d\sigma_{ij \rightarrow X}}{d\Omega}$ représente les sections efficaces partoniques différentielles du processus étudié, calculable à un ordre fixe donné grâce à la pQCD (LO, NLO, NNLO, ...). Notons que

¹²De façon similaire, des équations d’évolution en fonction de x ont été formulés par Balitskii–Fadin–Kuraev–Lipatov (BFKL) [90,91].

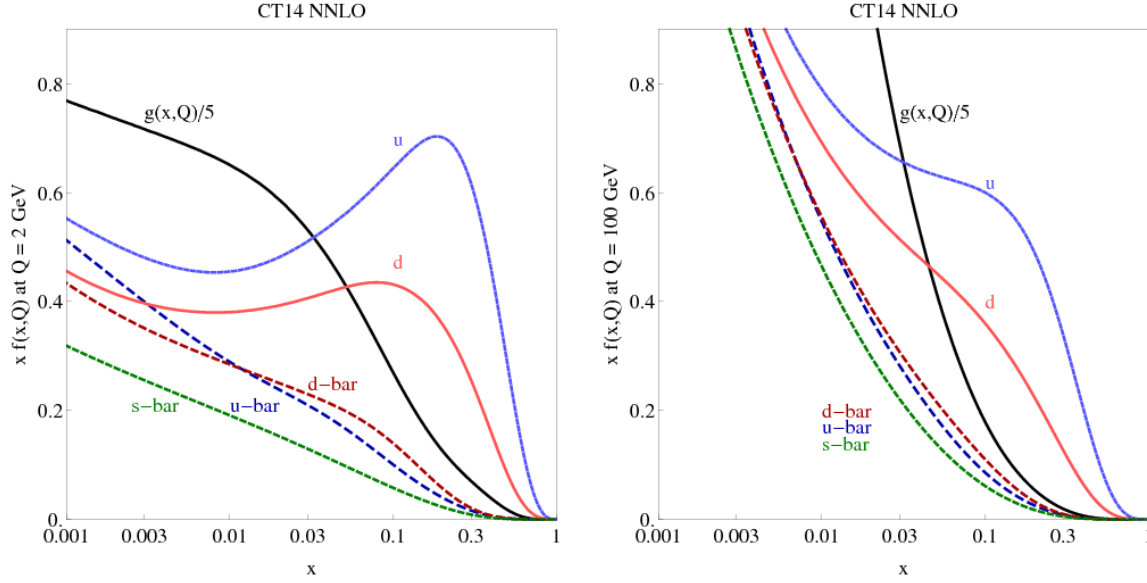


FIG. 18 – Évolution de la structure du proton à $Q = 1$ GeV (gauche) et $Q = 100$ GeV (droite) en fonction de x , d’après l’ajustement global publié par la collaboration CT14. L’évolution en Q^2 utilisée suit les équations de DGLAP calculées à l’ordre *next-to-next-to-leading order* (NNLO) [87].

par souci de cohérence et à cause de problèmes de convergence observés en QCD (voir section 3.4), les sections efficaces partoniques et l’équation d’évolution DGLAP utilisée dans la détermination des PDFs doivent être calculées au même ordre. Il est également rappelé que la formule de l’équation 17 doit être corrigée à l’aide de simulations MC afin de prendre en compte les autres effets de QCD affectant l’état final (voir section 3.2) afin de passer d’un état final “partonique” à un état final “particule” ou “vérité”.

Afin de séparer les effets perturbatifs et non-perturbatifs comme illustré dans l’équation 17, il est nécessaire de faire l’hypothèse que la physique interne régissant la structure du proton (non-perturbative) ne dépend pas de l’interaction dure (perturbative). Cela revient à considérer les PDFs comme universelles, ce qui semble raisonnable vu la différence d’échelle entre les deux processus. Ce résultat important et non trivial est donné par le théorème de factorisation [92], qui a cependant un domaine de validité en principe limité [93]. Le théorème de factorisation impose de fixer une échelle de factorisation μ_f arbitraire au-delà de laquelle les partons sont considérés comme libres (liberté asymptotique). Ce théorème est également utilisé pour factoriser les effets de QCD non-perturbative sur l’état final de l’interaction dure, reliées alors à l’échelle de fragmentation μ_F (voir section 3.2).

Le calcul de la section efficace partonique $\frac{d\sigma_{ij \rightarrow X}}{d\Omega}$ à un ordre fixe nécessite également la définition d’une autre échelle arbitraire dite de renormalisation μ_R , correspondant à la coupure en énergie lors de l’intégration de la section efficace sur l’espace des phases permettant la convergence du calcul perturbatif (divergences dans l’ultra-violet).

Finalement, l’emploi de l’adjectif “fiducielle” pour désigner la section efficace totale obtenue dans l’équation 17 n’est pas anodin. Il précise que l’expression différentielle n’a été intégrée que dans un domaine restreint de l’espace des phases Ω , appelé volume fiduciel et correspondant le plus souvent au domaine de sensibilité maximale de l’appareil

de mesure. Sa définition peut avoir des conséquences dramatiques sur la convergence des prédictions théoriques en QCD et doit donc être choisi avec soin. Quelques explications supplémentaires sur ce dernier point sont présentées dans la section 3.4.

3.4 Définition des observables et calcul des ordres supérieurs

Il est important d'apporter un soin particulier à la définition des observables en QCD, du fait de la masse nulle des gluons.¹³ En effet, cette caractéristique leur permet d'être produits à des énergies infiniment basses et à des impulsions infiniment colinéaires au parton initial. Il est assez facile de se rendre compte que cela va avoir pour conséquence de faire diverger certaines prédictions théoriques correspondant à des radiations supplémentaires (ordres supérieurs) si tous les éléments du calcul ne sont pas correctement pris en compte.

Cet effet est illustré dans la figure 19, qui présente les deux diagrammes correspondant aux premières corrections réelles QCD d'ordre supérieur (NLO) pour le processus $e^+e^- \rightarrow q\bar{q}$. La mise en équation de ces diagrammes montre une divergence logarithmique lorsque x_3 tend vers 0 (divergence molle ou infrarouge) ou lorsque l'angle entre le gluon additionnel et le quark tend vers 0 (divergence colinéaire), car il faut intégrer à partir de la masse des gluons, qui est nulle. Ces divergences posent de sérieux problèmes, mais heureusement s'annulent si on inclut dans le calcul les diagrammes correspondant aux première corrections virtuelles (NLO). l'un des ces diagrammes est représenté dans la figure 20.

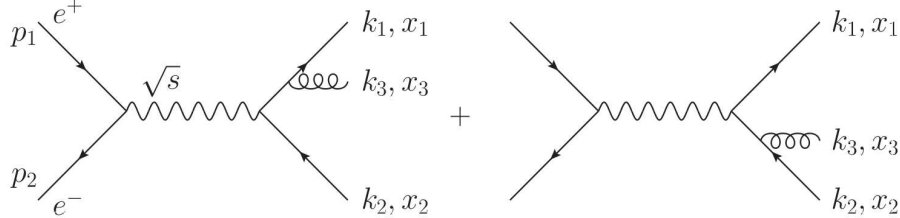


FIG. 19 – Premières corrections réelles QCD d'ordre supérieur pour le processus $e^+e^- \rightarrow q\bar{q}$. p_1, p_2 (k_1, k_2) sont les quadri-impulsions des leptons initiaux (quarks finals). $\sqrt{s}$ est l'énergie de collision dans le centre de masse des deux leptons et les x_i représentent les fractions d'énergie portées par les différents partons dans l'état final [86].

L'annulation des divergences molles et colinéaires par les corrections virtuelles est une propriété générale de la QCD (et QED) aussitôt que l'on travaille à un ordre fixe en théorie des perturbations, appelé théorème de Kinoshita-Lee-Nauenberg [94, 95] (Bloch–Nordsieck pour la QED [96]). Cependant, cette annulation est strictement vraie uniquement dans le cadre de mesures inclusives et donc ne s'applique en général pas directement aux mesures réelles, qui s'effectuent sur un volume fiduciel donné correspondant au domaine d'intégration Ω dans l'équation 17. Il revient donc aux expérimentateurs de s'assurer que leurs mesures concernent des variables *infrared-* et *collinear-safe*, c'est à dire qui préservent l'annulation des divergences nécessaire à la convergence des

¹³Ce commentaire est également valide pour les photons de la QED, mais le caractère non abélien de $SU(3)_C$ ouvrant le canal $g \rightarrow gg$ rend souvent les problèmes plus complexes en QCD.

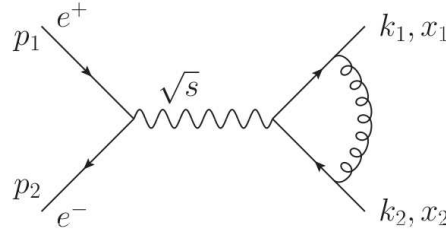


FIG. 20 – Exemple de diagramme correspondant aux premières corrections virtuelles QCD d’ordre supérieur pour le processus $e^+e^- \rightarrow q\bar{q}$. Deux autres diagrammes correspondant à l’émission et à l’absorption du gluon additionnel par le même quark sont également possibles. Ils permettent la régularisation des prédictions au premier ordre supérieur en α_s [86].

calculs de pQCD. Un exemple typique de variables non *infrared-safe* est le nombre de particules dans un jet hadronique (voir section 3.2), qui dépend naturellement fortement de l’ordre du calcul de pQCD. Au contraire, la quadri-impulsion d’un jet est une observable valide.

Il est également important de souligner qu’en général, le caractère *infrared-* et *collinear-safe* d’une variable est plus ou moins vérifié, et non un “oui ou non” strict. En effet, la taille des corrections d’ordres supérieurs en QCD peut dépendre significativement de la sensibilité aux infrarouges de l’observable considérée, et donc aux divergences abordées précédemment. On parle alors d’instabilités infrarouges. Des techniques de resommation à tous les ordres du développement perturbatif en α_s peuvent alors parfois être mises en oeuvre pour le calcul de certaines variables dans les régions sensibles aux infrarouges, à condition que la série présente des propriétés de régularité suffisantes [97–102]. Dans ce cas tous les termes de la série en puissance de α_s , divergeant logarithmiquement à cause des instabilités infrarouges, sont resommés jusqu’à une puissance de logarithme donnée. On parle alors de resommation au premier, deuxième logarithme, etc (LL, NLL pour *leading logarithm*, *next-to-leading logarithm*, etc). Ces types de calculs ont été développés en particulier pour le processus Drell-Yan ($q\bar{q} \rightarrow l^+l^-$) [97, 98, 103] afin de prédire correctement la section efficace lorsque l’impulsion transverse du propagateur Z^0/γ^* tend vers 0. Ils associent aussi souvent des ordres supérieurs en α_s dans le reste de l’espace des phases afin de décrire correctement les spectres à haute énergie, on parle alors de calculs LO+NLL, NLO+NNLL, etc.

L’utilisation d’algorithmes de cascades de partons (PS pour *parton shower*, voir section 3.2) dans les générateurs de Monte Carlo correspond en pratique à réaliser une resommation approchée, par opposition à un ordre fixe de logarithme. En effet, l’algorithme de PS étant traditionnellement exécuté indépendamment de l’ordre en α_s des éléments de matrice, des doubles comptages sont très probables, en particulier dans le cas de calculs au-delà de LO. Cependant, les générateurs de Monte Carlo de nouvelle génération (MC@NLO [104], POWHEG [105], SHERPA [83], HERWIG7 [85], ...) implémentent désormais des algorithmes d’appariement et de fusion entre les éléments de matrice et les cascades de partons, ce qui permet en principe de réaliser des resommations approchées sans double comptage tout en incluant des corrections supérieures en α_s . Ces développements sont particulièrement prometteurs pour décrire des processus sensibles à la fois aux corrections d’ordres supérieur (ex : jets additionnels dans l’état

final) et aux instabilités infrarouges (ex : émission de gluons mous dans l'état initial), comme dans le cas de la production de paires de photon au LHC (voir section 3.5 et la partie IV). Ils ont récemment permis de décrire avec succès les productions de $W^\pm$, Z et $t\bar{t} + \text{jet(s)}$ au LHC [106–108].

Les différentes remarques énoncées au cours de cette section auront une importance particulière dans la définition du signal di-photon et la compréhension du choix des variables de mesure dans l'analyse des données d'ATLAS au LHC, qui sera présentée en détail dans la partie IV. Dans la section 3.5, la production du photon au LHC est introduite.

3.5 La production de photons au LHC comme test de l'interaction forte

3.5.1 Motivations

Comme mentionné dans la section 3.1, la QCD joue un rôle particulier dans le modèle standard grâce à deux propriétés étonnantes : le confinement et la liberté asymptotique. À ce titre, il est important de tester sa véracité à la fois dans le domaine perturbatif (calculs perturbatifs à ordre fixe) et non-perturbatif (PDFs, modèles de fragmentation, ...). La question de la validité de la factorisation entre les deux domaines est également pertinente [93]. Par ailleurs, les sections efficaces de production de jets dans les collisionneurs hadroniques étant typiquement plusieurs ordres de grandeurs au-dessus des autres, les processus QCD constituent souvent des bruits de fond importants au LHC, que ce soit pour des mesures de précision (couplages du Higgs, sections efficaces du $W^\pm$ et du Z^0 , ...) ou des recherches de nouvelle physique (dimensions supplémentaires, supersymétrie, ...). Il est donc important de les déterminer avec la plus grande précision pour maximiser la sensibilité des mesures et des recherches de nouveaux phénomènes. Un résumé des mesures électrofaible du MS par la collaboration ATLAS au LHC est donné dans la figure 21. Pour un grand nombre d'entre elles, une partie significative de l'incertitude vient de la connaissance limitée du bruit de fond QCD.

Les photons sont des sondes indirectes particulièrement intéressantes de la QCD dans les collisionneurs hadroniques. En effet, ils produisent des signaux souvent plus discriminants et mesurables plus précisément par rapport aux partons qui s'hadronisent (voir section 3.2) tout en étant couplés à la QCD par l'intermédiaire des quarks, tous chargés électriquement. Par ailleurs, ils sont produits en quantités importantes avec des jets additionnels dans l'état final, correspondant à des corrections réelles en α_s . Leur mesure apparaît donc comme un banc de test idéal pour tester les prédictions pQCD à ordre fixe. Il a également été montré que les mesures des photons pourraient permettre de contraindre efficacement la PDF des gluons au LHC [109].¹⁴

Par ailleurs, dans le cas où au moins deux photons sont produits dans l'état final ($pp \rightarrow \gamma\gamma + X$), certaines observables du système $\gamma\gamma$ sont particulièrement sensibles aux instabilités infrarouges via l'émission de gluons mous dans l'état initial (voir partie IV), de façon similaire à la production Drell-Yan $q\bar{q} \rightarrow l^+l^-$ (voir section 3.4). Ainsi, les photons peuvent également être utilisés pour tester les prédictions basées sur des resom-

¹⁴Étude basée sur l'état final $\gamma + X$ uniquement.

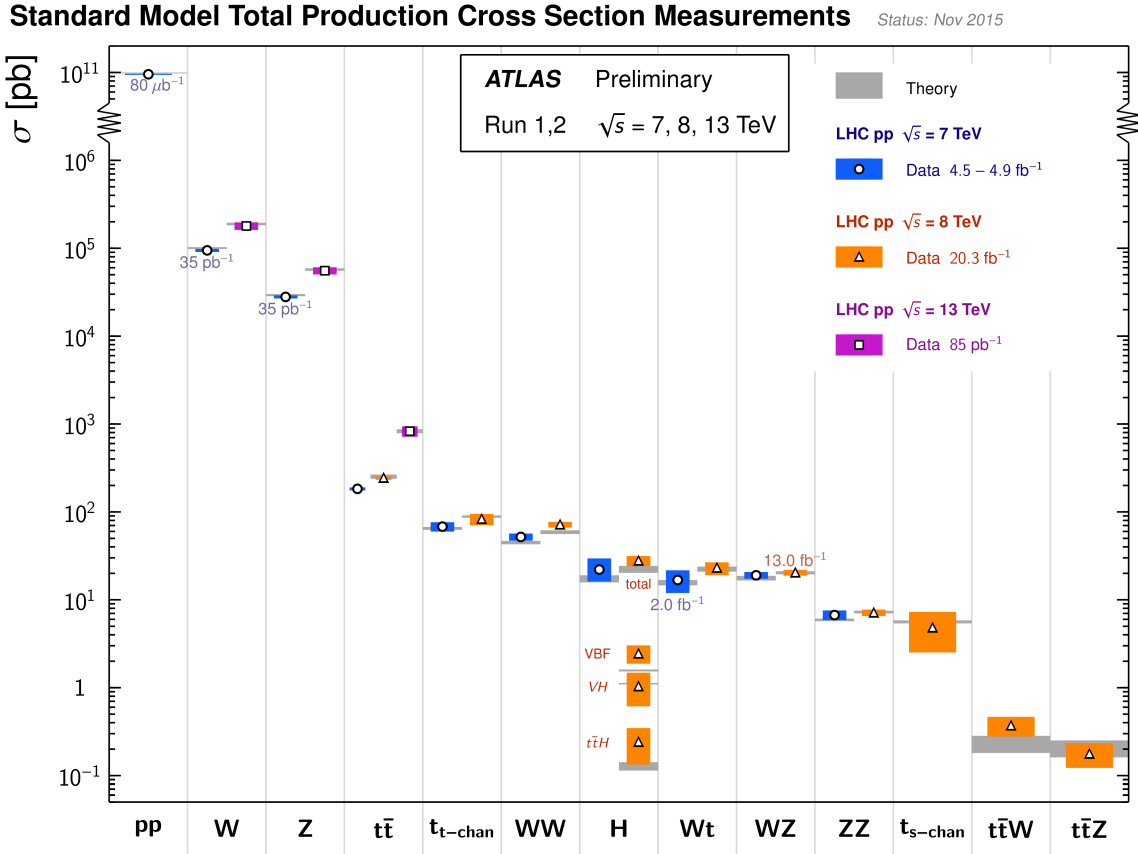


FIG. 21 – Résumé de plusieurs mesures de sections efficaces de production fiducielles et totales effectuées dans le cadre du modèle standard par la collaboration ATLAS au LHC. Les mesures sont corrigées des rapports de branchement leptoniques et comparées avec les prédictions théoriques au NLO ou à un ordre plus élevé correspondantes [110].

mations à tous les ordres en α_s , que ce soit de manière formelle à un ordre en logarithme donné en QCD (ex : calcul RESBOS [101] ou 2 γ RES [102]) ou via des générateurs de Monte Carlo de nouvelle génération implémentant des algorithmes d'appariement et de fusion entre éléments de matrice et cascades de partons permettant de tenir compte à la fois des corrections supérieures en α_s et des émissions à basse énergie (voir section 3.4).

Dans la suite, on se propose de donner quelques détails sur la production de paires de photons au LHC ($pp \rightarrow \gamma\gamma + X$), qui est la mesure principale effectuée pendant les années de thèse et sera traitée en détail dans la partie IV. Débutant par leurs différents modes de production au LHC dans la section 3.5.2, on abordera également les particularités des sections efficaces de photons dues au critère d'isolation (section 3.5.3) et quelques effets particuliers venant des corrections aux ordres supérieurs (section 3.5.4).

3.5.2 Les différents modes de production

Il existe trois modes principaux de production de paires de photons au LHC. La première représente les paires provenant directement de l'interaction partonique dure, dite composante directe. Celles-ci permettent de tester directement les prédictions de la pQCD. Leurs diagrammes de production principaux au LHC (LO, NLO) sont montrés

dans la figure 22 [111]. À LO, uniquement le processus initié par $q\bar{q}$ est réalisable à $O(\alpha_s^0)$ (diagramme a). À NLO, *ie.* $O(\alpha_s)$, les canaux gq et $g\bar{q}$ s'ouvrent avec émission d'un quark additionnel (diagramme c). Notons également qu'au NNLO, *ie.* $O(\alpha_s^2)$, le canal gg s'ouvre à son tour avec l'émission de deux photons par l'intermédiaire d'une boucle de quarks. L'ouverture de ces nouveaux canaux aux ordres supérieurs en α_s démontre l'importance des contributions des diagrammes NLO et NNLO à la section efficace totale, ce qui en fait un processus intéressant pour tester la pQCD.

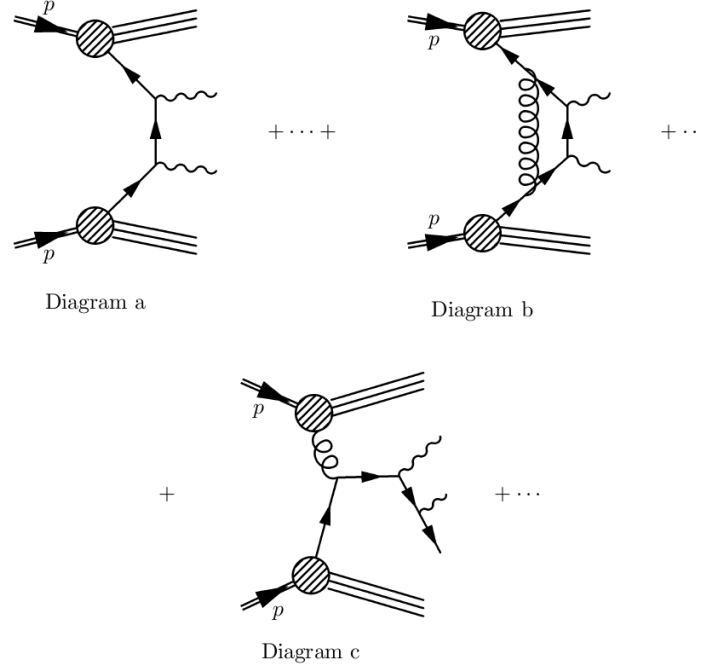


FIG. 22 – Exemple de diagrammes de production de paires de photons directs au LHC à l'ordre LO (diagramme a) et NLO (diagrammes b et c) en α_s . Les diagrammes correspondants à $q\bar{q} \rightarrow \gamma\gamma g$ et à d'autres corrections réelles et virtuelles ne sont pas représentés [111].

Le second mode de production représente les processus dans lesquels un ou deux des photons de l'état final sont issus de la fragmentation d'un parton (voir figure 23 et 24). La probabilité pour qu'un photon soit émis lors de la fragmentation d'un parton avec une fraction z de son impulsion est factorisée dans des fonctions de fragmentation $D_{\gamma/q}(z, \mu_F)$ et $D_{\gamma/g}(z, \mu_F)$, définies dans un schéma de factorisation donné et évaluées à une échelle de fragmentation arbitraire μ_F (voir section 3.2) [112].

D'un point de vue expérimental, on s'attend à ce que les photons de fragmentation soient à proximité de hadrons à cause de la dépendance angulaire de la probabilité d'émission ($\sim 1/\theta$). D'un point de vue théorique, les processus incluant des photons de fragmentation ne sont pas séparables des contributions directes dès lors que des diagrammes avec partons additionnels sont considérés. En effet, les trois composantes (directe, 1-fragmentation, 2-fragmentations) dépendent fortement du choix de l'échelle de factorisation μ_F alors que leur somme permet une annulation partielle de la dépendance du fait des problèmes de divergences colinéaires de la QCD vus en section 3.4. Cela se comprend en remarquant par exemple que le diagramme d (1-fragmentation) correspond au diagramme c (directe) dans la limite colinéaire pour le deuxième photon, illustrant

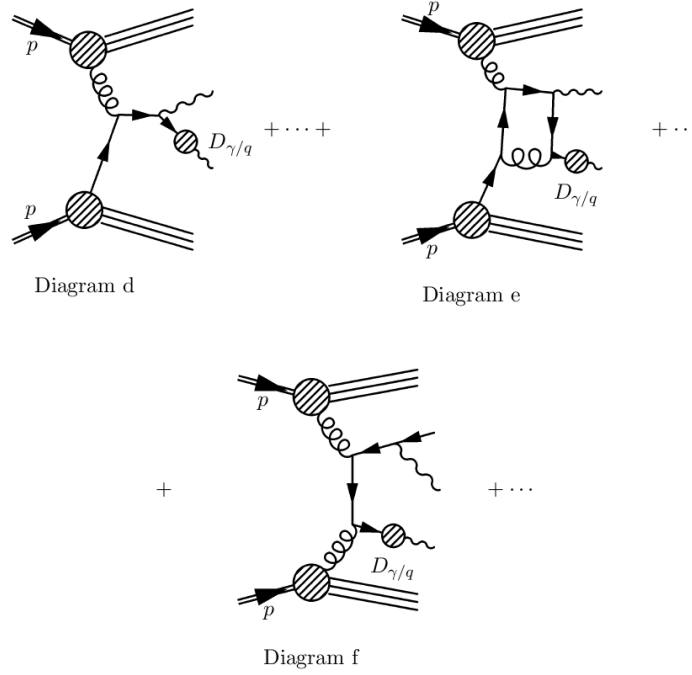


FIG. 23 – Exemple de diagrammes de production de paires de photons au LHC avec un des deux photons issus de la fragmentation d'un parton au LO (diagramme d) et NLO (diagrammes e et f) [111].

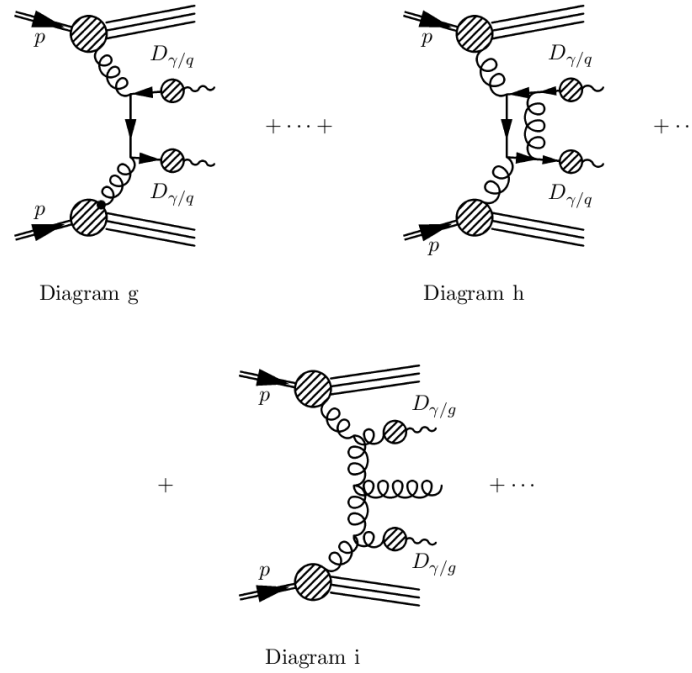


FIG. 24 – Exemple de diagrammes de production de paires de photons au LHC avec deux photons issus de la fragmentation d'un parton au LO (diagramme g) et NLO (diagrammes h et i) [111].

bien l'ambiguïté existant entre les photons directs et les photons de fragmentation à partir de NLO.

Finalement, la troisième catégorie de photons produits au LHC inclut les photons issus de la désintégration de hadrons ou de taus, l'exemple typique étant $\pi^0 \rightarrow \gamma\gamma$. Ils sont parfois appelés photons secondaires. Leur section efficace de production est gouvernée par la QCD non-perturbative, en particulier par le processus d'hadronisation, et se situe plusieurs ordres de grandeur au-dessus des autres catégories. Afin de les séparer des photons directs et de fragmentation pour garder une sensibilité à la pQCD, il est nécessaire d'introduire au niveau expérimental des critères d'isolation pour les photons, objet de la section 3.5.3.

3.5.3 Critères d'isolation

Isolation dans un cône de taille fixe

Les photons secondaires ne sont pas nécessaires à la convergence des calculs de pQCD et la taille de leur section efficace de production au LHC n'est pas calculable facilement avec une bonne précision. Par conséquent, ils sont traditionnellement exclus des mesures de photons et traités comme un bruit de fond. Au niveau expérimental, un critère d'isolation est requis sur les candidats photons afin de rejeter les photons secondaires en s'appuyant sur leur environnement hadronique important (voir partie III). Ainsi, on définit souvent l'énergie d'isolation au niveau expérimental comme la quantité d'énergie transverse se situant dans un cône d'une certaine taille autour du photon reconstruit dans le détecteur, qu'on limite à une certaine valeur :

$$E_T^{\text{iso}} < E_T^{\text{iso,max}} \text{ dans } \mathcal{C}(R), \quad (18)$$

où E_T^{iso} est l'énergie d'isolation transverse se situant dans le cône $\mathcal{C}(R)$ de rayon R centré autour du photon reconstruit et $E_T^{\text{iso,max}}$ est la coupure appliquée dans l'analyse. Des définitions plus précises des variables utilisées dans les collisionneurs hadroniques et de l'énergie d'isolation utilisée expérimentalement sont données respectivement dans les sections 5.2 et 10.

Cependant, le critère d'isolation utilisé au niveau expérimental exclut également une partie des photons de fragmentation, qui présentent aussi souvent un environnement hadronique non négligeable. Il est alors très courant (et préférable) d'appliquer en parallèle une coupure d'isolation dans les prédictions théoriques la plus proche possible de celle appliquée au niveau expérimental, afin de s'affranchir au maximum de la description théorique des photons à très hautes énergies d'isolation. En effet, la production de ces photons est mal déterminée à cause de la connaissance limitée des fonctions de fragmentations (domaine non-perturbatif). De plus, elle dépend fortement des corrections aux ordres supérieurs à cause des instabilités colinéaires (voir section 3.2). Un des critères couramment utilisé dans les prédictions théoriques est similaire à l'équation 18 en restreignant le calcul à l'énergie transverse hadronique E_T^{had} issue des partons additionnels (niveau partonique). Dans le cas d'un générateur incluant cascade de partons, hadronisation, etc, elle peut également être évaluée au niveau particule, *ie.* après modélisation complète de l'état final.

Dans le cas de calculs de sections efficaces isolées, la validité du théorème de factorisation, nécessaire à l'utilisation des fonctions de fragmentation (voir section 3.3), et l'annulation des divergences infrarouges de la QCD (voir section 3.4) ne sont a priori plus garanties. En effet, le critère d'isolation restreint l'espace des phases d'intégration, ce qui pourrait invalider les propriétés de QCD vérifiées habituellement. On parle alors de sections efficaces isolées et non plus inclusives. La validité du théorème de factorisation et l'universalité des fonctions de fragmentation pour les sections efficaces utilisant un critère d'isolation similaire à l'équation 18 ont été démontrées dans le cas de collisions e^+e^- [113] et hadroniques [114], légitimant son utilisation dans le calcul des sections efficaces.¹⁵ En revanche, des instabilités infrarouges sont attendues, en particulier dans certains domaines de l'espace des phases particulièrement sensibles à l'émission de gluons mous dans l'état initial, comme dans les topologies di-photons associées à des grandes différence d'azimuts ($\Delta\phi_{\gamma\gamma} \simeq \pi$), à des basses impulsions transverses ($p_{T,\gamma\gamma} \simeq 0$), ou plus surprenant à $p_{T,\gamma\gamma} \simeq E_T^{\text{iso,max}}$ [111]. Pour ces domaines, une resommation à tous les ordres devrait permettre d'améliorer les prédictions.

L'état de l'art concernant les prédictions théoriques $pp \rightarrow \gamma\gamma + X$ isolés dans un cône de taille fixe est le calcul DIPHOX [111], qui implémente un calcul à ordre fixe NLO en α_s pour toutes les composantes (direct, 1-fragmentation, 2-fragmentations). Le calcul RESBOS [101] implémente également des prédictions au NLO en α_s avec en plus une procédure de resommation au NNLL sensée améliorer la description des régions $p_{T,\gamma\gamma} \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$ sensibles aux infrarouges. Cependant, RESBOS traite la fragmentation de manière plus approximative en incluant uniquement la composante 1-fragmentation par une paramétrisation approchée des fonctions de fragmentation. DIPHOX et RESBOS ont été développés entre autres pour décrire les mesures effectuées au TEVATRON (voir partie IV). Finalement, les générateurs MC HERWIG 6.4 [115], PYTHIA8 [82] et SHERPA 1.4 [83] implémentent des éléments de matrice au LO avec cascade de partons dans lesquels un critère d'isolation basé sur un cône fixe peut être implémenté. Ces différents calculs seront abordés plus en détails lors de la comparaison de la mesure $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV avec les prédictions théoriques dans la partie IV.

Le calcul de la composante de fragmentation (et donc des diagrammes correspondants à γ +jet+ X et di-jet) pénalise considérablement les prédictions pQCD à ordre fixe car celle-ci doit être calculée en général à un ordre supérieur en α_s au-dessus de la composante directe afin de stabiliser la section efficace isolée. Afin de remédier à cela, un critère d'isolation basé sur un cône à rayon variable permettant de faire converger la composante directe sans la fragmentation est également possible et de plus en plus utilisé parmi les théoriciens. Celui-ci est décrit dans le paragraphe suivant.

Isolation dans un cône de taille variable

Il existe également un critère d'isolation alternatif utilisant un cône à rayon variable plutôt que fixe appelé isolation de Frixione [116]. Il est défini par la formule suivante :

$$\forall \ r < R_0, \quad E_T^{\text{had}}(r) < E_T^{\text{iso,max}}(r) = \epsilon_\gamma E_T^\gamma \left(\frac{1 - \cos r}{1 - \cos R_0} \right)^n, \quad (19)$$

¹⁵Le fait de considérer l'énergie transverse uniquement et non l'énergie totale est fondamental afin d'obtenir ce résultat important.

où E_T^{had} est l'énergie hadronique transverse se situant dans le cône $\mathcal{C}(r)$ de rayon r centré autour du photon reconstruit, et ϵ_γ , n et R_0 sont des paramètres à choisir. Un choix raisonnable est par exemple $\epsilon_\gamma = 0.5$, $n = 1$ et $R_0 = 0.4$ [117]. L'utilisation du critère de Frixione dans les calculs des sections efficaces de photons a l'immense avantage de supprimer entièrement la dépendance du calcul aux fonctions de fragmentation tout en garantissant le caractère *collinear and infrared-safe* à tous les ordres des perturbations de la QCD. La preuve formelle est donnée dans [116]. Cela permet aux théoriciens de calculer beaucoup plus facilement des prédictions aux ordres supérieurs en α_s .

En effet, l'état de l'art concernant les prédictions théoriques $pp \rightarrow \gamma\gamma + X$ isolés dans un cône de Frixione est au NNLO en α_s , avec le calcul $2\gamma\text{NNLO}$ [117] ou plus récemment MCFM [118]. un calcul NNLO en α_s avec resommation à tous les ordres jusqu'au NNLL a également été publié récemment par les mêmes auteurs que $2\gamma\text{NNLO}$ ($2\gamma\text{RES}$) [102]. Les dernières versions des générateurs MC utilisés au LHC comme SHERPA 2 [119] ou HERWIG 7 [85] implémentent également un critère de Frixione pour l'isolation des photons au niveau générateur. À nouveau, ces différents calculs seront abordées plus en détails lors de la comparaison de la mesure $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV avec les prédictions théoriques dans la partie IV.

Malheureusement, l'application d'un critère d'isolation basé sur un cône avec un rayon variant continuellement est très difficile au niveau expérimental à cause de la segmentation limitée des calorimètres (voir partie II). Cependant, les différences entre les deux critères mènent à des différences de l'ordre du pourcent uniquement pour des paramètres bien choisis [120]. Ainsi, il a été proposé récemment un accord entre théoriciens et expérimentateurs, disponible dans l'étude débutant en page 145 de la référence [121] : il est recommandé aux expérimentateurs de ne pas implémenter le critère de Frixione dans les analyses mais de définir leurs mesures en utilisant le critère de cône fixe avec un cône assez petit ($R \simeq 0.4$), une séparation angulaire minimum entre les deux photons ($\Delta R_{\gamma\gamma} \simeq 0.4$) et une énergie d'isolation maximum suffisamment basse ($E_T^{\text{iso,max}} < 5$ GeV) pour produire des résultats comparables avec les prédictions théoriques NNLO, produites uniquement avec le critère de Frixione pour le moment.

Pourtant, il est bien connu que les simulations MC utilisées dans les analyses pour corriger les mesures des effets de détecteurs (déconvolution ou *unfolding*) se basent le plus souvent sur des algorithmes de cascades de partons pour générer les photons de fragmentation (voir section 3.2), qui diffèrent significativement de la composante colinéaire générée au niveau partonique à l'aide des fonctions de fragmentation dans les prédictions à ordre fixe [122]. En effet, les algorithmes de cascades de partons sont indispensables pour obtenir une simulation réaliste de l'état final complet et ainsi évaluer la réponse de l'ensemble du détecteur. Or, l'accord énoncé plus haut a été conclu à partir d'études basées sur le programme DIPHOX [111], qui produit une composante de fragmentation colinéaire. En pratique, il est donc aujourd'hui très difficile d'estimer précisément le degré de correspondance entre les critères d'isolation appliqués par les expérimentateurs et les théoriciens. Il serait très profitable pour la communauté d'étudier plus en détails les différences entre les photons issus de la fragmentation colinéaire de partons et ceux issus des cascades de partons. Néanmoins, à l'heure actuelle, les incertitudes théoriques restent dominées par les ordres supérieurs pour tous les types de calculs, ce qui relaie le problème délicat de l'isolation à un second plan.

Dans la section 3.5.4, quelques effets particuliers sur les prédictions des distributions

cinématiques des di-photons dus aux corrections d'ordres supérieurs en α_s sont discutés succinctement.

3.5.4 Déformations des distributions $\gamma\gamma + X$ au NLO

Il est intéressant de noter quelques effets non-triviaux des corrections d'ordres supérieurs sur la forme de certaines distributions cinématiques sensibles aux productions colinéaires, comme dans la région $p_{T,\gamma\gamma} \simeq E_{T,1}^{\gamma,\min} + E_{T,2}^{\gamma,\min}$, où les $E_{T,i}^{\gamma,\min}$ correspondent aux impulsions transverses minimales des photons. Dans cette région, les productions de di-photons correspondant au diagramme avec une fragmentation de la figure 25 (gauche) s'accroissent (limite colinéaire) et déforment la distribution en $p_{T,\gamma\gamma}$ [123]. Le diagramme avec deux fragmentations correspondant est également concerné (le deuxième photon est alors issu de la fragmentation d'un gluon). La section efficace est aussi déformée par une contribution 100% directe venant de l'ouverture de l'espace des phases pour $p_{T,\gamma\gamma} > E_{T,1}^{\gamma,\min} + E_{T,2}^{\gamma,\min}$ par les contributions NLO correspondant à $\Delta\phi_{\gamma\gamma} < \pi/2$. L'altération correspondante est appelée épaule de Guillet [123] et visible dans la figure 25 (droite) grâce aux prédictions NLO du calcul DIPHOX pour les trois composantes (directe, une fragmentation, deux fragmentations).

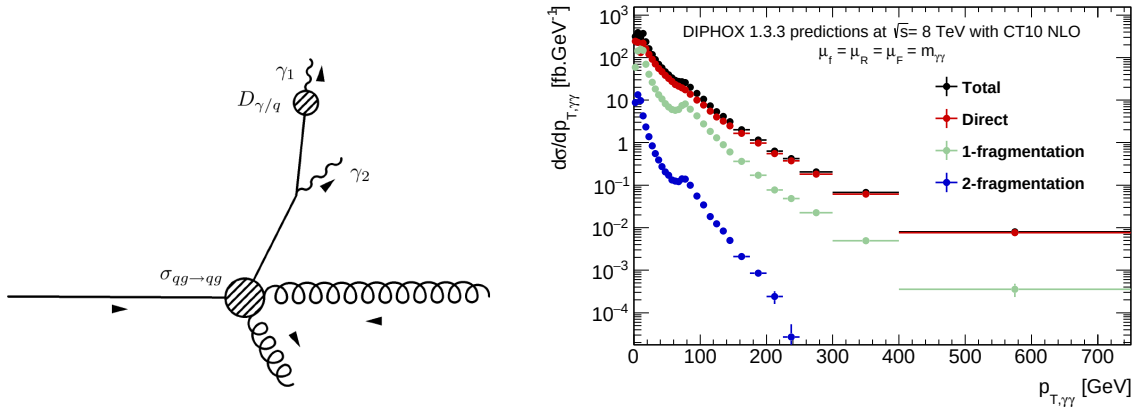


FIG. 25 – (gauche) Configuration cinématique avec un photon direct et un photon de fragmentation responsable de la déformation appelée épaule de Guillet. Une contribution additionnelle venant de l'ouverture de l'espace des phases pour $p_{T,\gamma\gamma} > E_{T,1}^{\gamma,\min} + E_{T,2}^{\gamma,\min}$ par les contributions NLO directs est également significative [123]. (droite) Prédictions NLO en α_s pour la section efficace $\gamma\gamma + X$ générées à $\sqrt{s} = 8$ TeV avec le code DIPHOX v1.3.3 [111] en utilisant les PDFs CT10 NLO [124]. Les différentes prédictions sont décomposées en terme de composantes directes, 1-fragmentation et 2-fragmentations (voir section 3.5.2). Les échelles de fragmentation, factorisation et renormalisation sont fixées à $m_{\gamma\gamma}$ et la sélection fiducielle est celle employée dans l'analyse présentée dans la partie IV.

Au cours de cette première partie, le modèle standard de la physique des particules a été introduit succinctement. Certains aspects de la théorie électrofaible et de la QCD nécessaires à la compréhension des deux travaux principaux effectués pendant la thèse, *ie.* l'étude phénoménologique des couplages à quatre photons et la mesure de la section efficace de paires de photons dans l'expérience ATLAS, ont été développés. Les travaux

principaux de thèse ont été ensuite introduits et motivés dans la section 2.4 et 3.5, préparant leur traitement détaillé qui sera effectué les parties IV et V.

Dans la partie suivante, le LHC et l'expérience ATLAS sont introduits. Les détecteurs de protons à l'avant installés ou à venir au LHC sont également présentés.

Deuxième partie

Le grand collisionneur de hadrons et l'expérience ATLAS

Sommaire

4	Le grand collisionneur de hadrons	62
4.1	Description générale	62
4.2	Interêt d'un collisionneur proton-proton	65
4.3	Bilan de la première prise de données (2010-2012)	66
4.4	État des lieux de la seconde prise de données (2015-2018) et projet à plus long terme	68
5	L'expérience ATLAS	70
5.1	Description générale	70
5.2	Système de coordonnées et observables	72
5.3	Le détecteur interne	74
5.3.1	Le détecteur à pixels de silicium	74
5.3.2	Le détecteur à micro-pistes de silicium (SCT)	75
5.3.3	Le détecteur à rayonnement de transition (TRT)	76
5.3.4	Résolution globale attendue et quantité de matière	76
5.4	Le calorimètre électromagnétique	78
5.4.1	Le calorimètre électromagnétique à argon liquide (LAr EM)	78
5.4.2	Le calorimètre électromagnétique à l'avant	80
5.5	Le calorimètre hadronique	81
5.5.1	Le calorimètre hadronique à tuiles (<i>TileCal</i>)	81
5.5.2	Les bouchons hadroniques (<i>HEC</i>)	81
5.5.3	Le calorimètre hadronique à l'avant	82
5.6	Le spectromètre à muons et les autres détecteurs d'ATLAS	82
5.6.1	Le spectromètre à muons	82
5.6.2	Les autres détecteurs d'ATLAS	83
5.7	Le système de déclenchement et la distribution des données	84
5.7.1	Le système de déclenchement	84
5.7.2	La distribution des données	86
6	Les détecteurs de protons à l'avant au LHC et la mesure de temps de vol	88
6.1	La physique à l'avant	88
6.2	Les détecteurs de protons à l'avant	89
6.2.1	État des lieux	89
6.2.2	Mesures, résolution et acceptance	90
6.3	La mesure de temps de vol	92
6.3.1	Systèmes de mesures	92
6.3.2	Le temps de vol des protons à l'avant	94

4 Le grand collisionneur de hadrons

4.1 Description générale

Le grand collisionneur de hadrons (*large hadron collider*, LHC) [1] est un accélérateur de particules circulaire situé au Centre Européen pour la Recherche Nucléaire (CERN) à la frontière franco-suisse, près de Genève. Construit dans un tunnel souterrain de près de vingt-sept kilomètres de circonférence compris entre quarante et cent quatre-vingt mètres de profondeur, il constitue à ce jour le plus puissant collisionneur de particules au monde, avec une énergie nominale de collision de 14 TeV dans le centre de masse dans le cas de faisceaux de protons.¹⁶

Au LHC, deux faisceaux de particules sont accélérés en parallèle et en sens opposé grâce à un système magnétique à deux ouvertures installé le long de l’anneau. Des collisions quasi-frontales des deux faisceaux peuvent être effectuées à quatre emplacements distincts, où sont situées les expériences ATLAS (*a toroidal LHC apparatus*) [72], CMS (*compact muon solenoid*) [125], LHCb (*LHC beauty experiment*) [126] et ALICE (*a large ion collider experiment*) [127]. D’autres expériences plus petites et dédiées à des mesures plus spécifiques sont situées à quelques centaines de mètres des différents points d’interaction : TOTEM (*total elastic and diffractive cross-section measurement*) [128], LHCf (*LHC forward experiment*) [129] et MoEDAL (*monopole and exotics detectors at the LHC*) [130]. Un schéma général de l’installation est montré dans la figure 26.

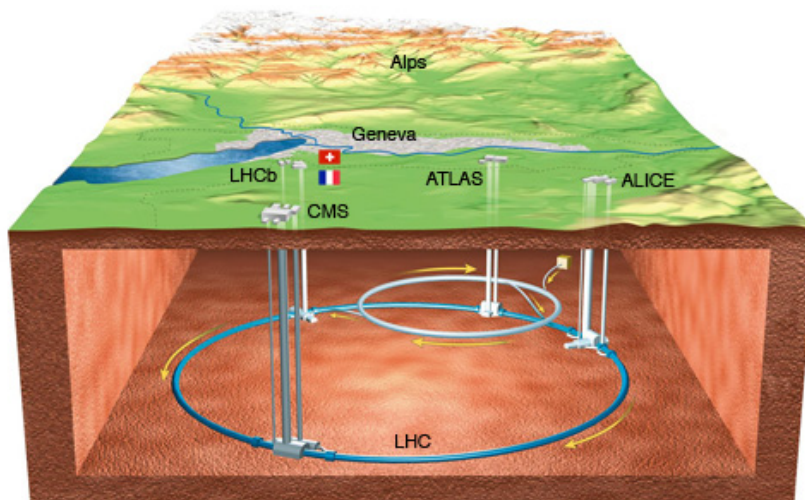


FIG. 26 – Schéma montrant le LHC et ses quatre expériences principales, situées au niveau des points de croisement des faisceaux. Trois autres expériences plus petites et dédiées à des mesures plus spécifiques sont également présentes le long de l’anneau mais non-représentées ici [131].

Le LHC a été conçu pour collisionner des hadrons chargés. Depuis sa mise en service en 2009, des collisions proton-proton, proton-plomb et plomb-plomb ont été produites. Il n’est pas exclu de produire d’autres types de collisions comme par exemple des collisions proton-oxygène [132], qui pourraient servir à mieux contraindre les modèles de

¹⁶ À cause des problèmes de connections dans le système cryogénique, l’énergie de collision maximale d’opération a été abaissée à 13 TeV.

désintégration des rayons cosmiques dans l'atmosphère. Dans la suite, le cas de collisions proton-proton est considéré, sauf indication contraire.

Les faisceaux nécessitent d'être portés à une énergie de 450 GeV avant de pouvoir être injectés dans le LHC, qui a une fréquence de fonctionnement limitée à une plage réduite (synchrocyclotron). Cela implique l'existence d'une chaîne de pré-accélérateurs, incluant un accélérateur linéaire et trois autres synchrotrons : le Linac 2, le PSB (*proton synchrotron booster*), le PS (*proton synchrotron*) et le SPS (*super proton synchrotron*). Les protons proviennent initialement d'atomes d'hydrogène gazeux ionisés dans un champ électrique et sont accélérés au fur et à mesure de leur passage dans les différentes machines, jusqu'à atteindre une énergie nominale de 7 TeV dans le LHC.¹⁷ Ils rentrent ensuite en collision aux différents points d'interaction. Une vue d'ensemble de la chaîne d'accélération est présentée dans la figure 27.

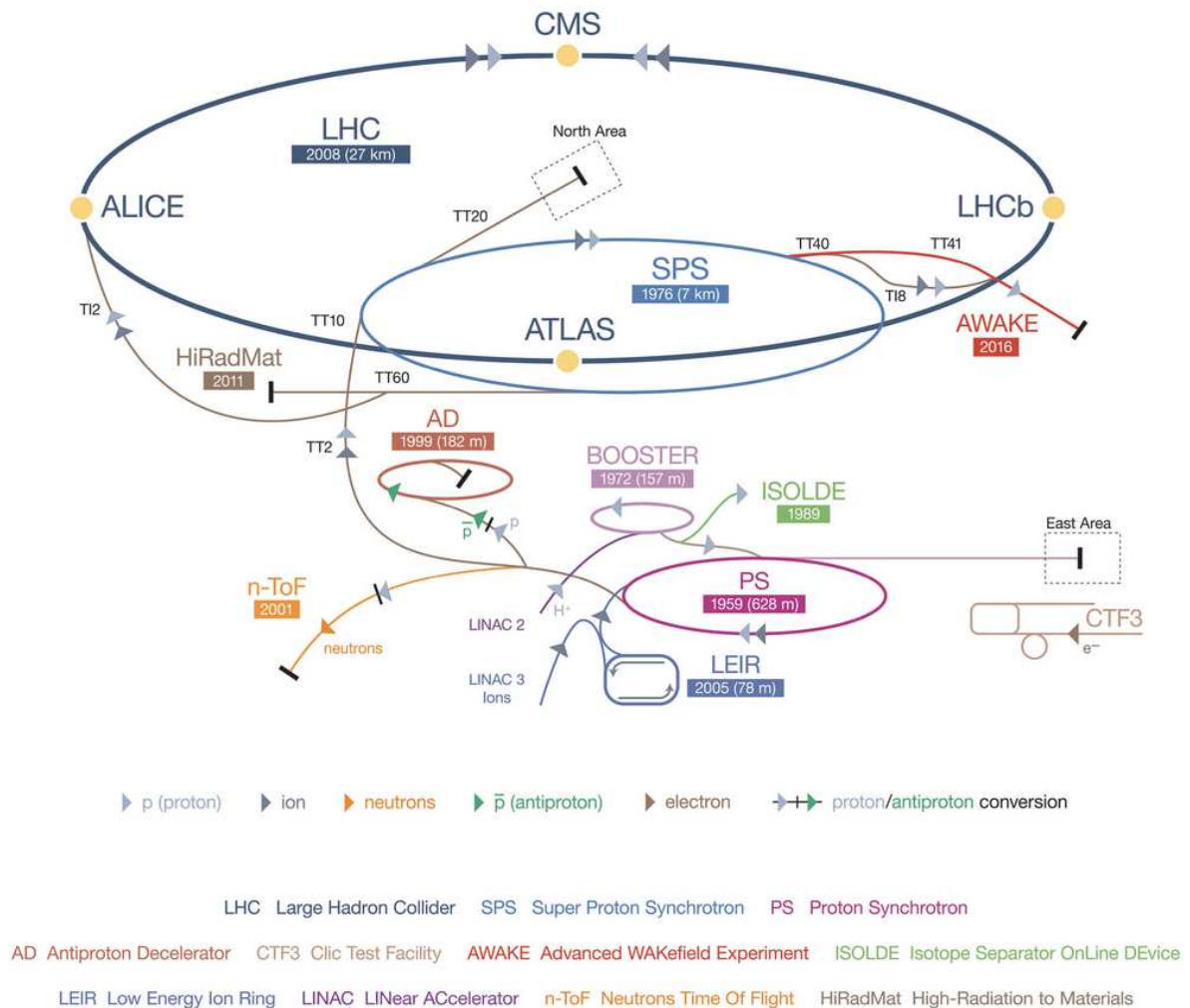


FIG. 27 – Schéma du complexe d'accélérateurs de particules en opération au CERN. Le trajet des protons rentrant en collision au LHC est marqué par une flèche grise. Les chaînes d'accélérations servant à d'autres expériences sont également représentées [133].

¹⁷En réalité, l'énergie nominale des faisceaux de proton a été abaissée à 6.5 TeV, voir note de bas de page N° 16.

Au LHC, les protons sont accélérés à chaque tour de 485 keV grâce à huit cavités radiofréquences réparties le long de l'anneau. Des cavités sont également utilisées dans les synchrotrons précédents, ce qui a pour conséquence de diviser les faisceaux en paquets de particules distincts, appelés *bunchs*. La taille des *bunchs* est paramétrisable mais est d'environ 7.55 centimètres de long pour 16.7 micromètres de large au niveau des points d'interaction. Ils contiennent initialement environ 10^{11} protons et sont espacés entre eux de 50 nanosecondes (2009-2012, 2015) ou 25 nanosecondes (2015-...) dans le cas d'une opération nominale de la machine. Lors d'un espacement de 25 nanosecondes, le LHC peut contenir au maximum environ 2800 *bunchs* par ligne de faisceau.

Le contrôle des faisceaux au LHC est assuré par plusieurs types d'aimants, la plupart supraconducteurs. Les dipôles permettent de courber les faisceaux tandis que des quadrupôles, sextupôles, octopôles, décapôles et même dodécapôles corrigent finement le champ magnétique.

Ainsi, à chaque point d'interaction, des paquets de particules à très haute énergie entrent en collision à une fréquence de 20 ou 40 mégahertz, selon leur espacement. Le nombre de protons dans chaque paquet baissant au fur et à mesure du temps, il est nécessaire de contrôler le nombre de collisions produites par unité de temps et pendant une période donnée. Il est alors pratique de définir la luminosité instantanée L , et intégrée $\mathcal{L} = \int L dt$. Ces quantités ont respectivement la dimension de l'inverse d'une section efficace par unité de temps et de l'inverse d'une section efficace. La luminosité intégrée est une quantité particulièrement importante pour les mesures de physique, car elle permet d'écrire simplement le passage entre le nombre d'événements observé et la section efficace que l'on cherche à extraire :

$$N = \mathcal{L} \cdot \epsilon \cdot \sigma, \quad (20)$$

où N est le nombre d'événements observé, $\mathcal{L}$ la luminosité intégrée, ϵ un facteur corrigeant les différents effets de détecteur (déclenchement, acceptance, reconstruction, identification, ..., voir section IV) et σ la section efficace que l'on veut mesurer. Dans la collaboration ATLAS (voir section 5), la mesure la plus précise de la luminosité est effectuée grâce à des scans de Van Der Meer. Cette technique a permis une détermination de la luminosité à 1.9% près pour les données 2012 (mesure finale) [134]. Des détails supplémentaires sur les différentes méthodes utilisées par l'expérience ATLAS pour mesurer la luminosité sont disponibles dans la référence [135].

Par ailleurs, à cause du grand nombre de protons contenu dans chaque paquet entrant en collision et de la section efficace totale inélastique proton-proton (pp) élevée ($\simeq 70$ mb) [136],¹⁸ plusieurs interactions proton-proton ont généralement lieu en même temps lors d'un croisement de faisceau. C'est l'effet d'empilement ou *pile-up*, déjà mentionné dans la section 3.2. Celui-ci peut être décomposé en deux parties : l'empilement en-temps (*in-time*), venant de collisions supplémentaires ayant lieu pendant le même croisement de faisceau, et hors-temps (*out-of-time*), qui lui vient de collisions produites lors de croisements différents. L'empilement est typiquement maximum au début d'une période de collision quand les *bunchs* ont leur taille maximale, et baisse au fur et à mesure du temps. Pour un croisement de faisceau donné, on parle souvent d'empilement

¹⁸Valeur pour une énergie de collision d'environ 14 TeV. Le barn est l'unité habituellement utilisée pour les sections efficaces. Elle est homogène à une surface avec $1 \text{ barn} = 10^{-24} \text{ cm}^2$.

moyen, noté μ . μ représente le paramètre de la distribution de Poisson donnant la probabilité d'avoir X interactions lors du croisement en question. Il s'obtient à partir de la section efficace totale inélastique pp $\sigma_{\text{inel}}^{pp}$, de la luminosité instantanée du croisement de faisceau L_{crois} , et de la fréquence de révolution dans le LHC f_r :

$$\mu = \frac{L_{\text{crois}} \cdot \sigma_{\text{inel}}^{pp}}{f_r}. \quad (21)$$

Dans la section 4.2, les avantages d'un collisionneur proton-proton par rapport aux choix concurrents sont mis en avant. Des détails supplémentaires sont ensuite apportés concernant les conditions de fonctionnement du LHC lors des différentes prises de données ayant eu lieu jusqu'à maintenant : dans la section 4.3, la première prise de données stable de longue durée appelée *run 1* (2010-2012) est abordée, tandis que la section 4.4 dresse un premier état des lieux de la seconde prise de données (*run 2*), débutée à l'été 2015.

4.2 Interêt d'un collisionneur proton-proton

Il est intéressant de se demander pourquoi il a été décidé de construire un collisionneur circulaire proton-proton plutôt qu'un autre pour le LHC. Pour cela, il est nécessaire de rappeler brièvement le contexte historique. Après l'observation des bosons $W^\pm$ et Z^0 à UA1 et UA2 dans les années quatre-vingts, il était nécessaire de confirmer leur découverte et d'étudier précisément leur couplage afin de tester la théorie électrofaible formulée par Glashow-Weinberg-Salam (voir section 2). Pour cela, le grand collisionneur circulaire électron-positron du CERN a été construit (LEP, 1989-2000). Celui-ci reste encore aujourd'hui le plus puissant collisionneur leptonique qui ait jamais fonctionné, avec une énergie dans le centre de masse atteignant jusqu'à $\sqrt{s} = 192$ GeV. À la fin des opérations du LEP, la théorie électrofaible était presque parfaitement confirmée par les données. Seule la découverte du boson de Higgs manquait au tableau, car sa masse, bien que fortement contrainte par les données du LEP, est un paramètre libre du modèle standard. Dès le début des années quatre-vingt dix, un projet de remplacement du LEP dans les années deux mille par une nouvelle machine, le LHC, fut proposé afin de déterminer de façon définitive l'existence du boson de Higgs et d'explorer la physique au téra-électron-volts, particulièrement motivée théoriquement par le problème de hiérarchie et des arguments de naturalité (voir section 2.3).

Il fallait donc concevoir une machine pouvant être installée dans le tunnel du LEP et capable de déterminer de façon définitive l'existence ou non du boson de Higgs. Dans ce contexte, un collisionneur hadronique présente deux avantages déterminants par rapport à un collisionneur électron-positron :

- ◊ **Les protons ne perdent quasiment aucune énergie par rayonnement synchrotron** en comparaison aux électrons, grâce à leur masse 1 836 fois plus élevée. Cela permet de construire une machine beaucoup plus puissante pour la même taille d'anneau.
- ◊ **La collision d'objets composites comme les protons permet de sonder une gamme d'énergie très large** grâce aux fonctions de structure (voir sec-

tion 3.3), ce qui donne une situation propice à de nouvelles découvertes, comme par exemple celle de nouvelles particules massives prédites par certaines extensions du MS.

La contrepartie du deuxième point réside dans le fait que la cinématique des collisions n'est pas contrainte dans la direction du faisceau, ce qui oblige à porter une attention particulière à la définition des observables (voir section 5.2). Il implique également la présence de débris de collision venant des protons, constituant un bruit de fond hadronique important pour le détecteur. Ces deux caractéristiques sont absentes des collisions électron-positron.

Par ailleurs, la luminosité constitue un autre paramètre fondamental lorsque l'on vise à observer des phénomènes rares. Il est alors intéressant d'imaginer une machine comportant deux systèmes magnétiques séparés afin de réaliser des collisions particule/particule comme le LHC, et non plus particule/anti-particule comme par exemple le Tevatron (Fermilab, États-Unis). En effet, la création d'anti-particule étant techniquement difficile, son taux de production limitée constitue un frein très important à l'augmentation de la luminosité dans les collisionneurs $p\bar{p}$, comme cela a été constaté au Tevatron. Le LHC a donc été conçu pour accélérer deux faisceaux de protons en sens opposé, comme décrit précédemment dans la section 4.1.

4.3 Bilan de la première prise de données (2010-2012)

En 2010 et 2011, le LHC a été configuré pour produire des collisions proton-proton avec une énergie de $\sqrt{s} = 7$ TeV dans le centre de masse. La luminosité instantanée maximale mesurée pendant cette période est de $3.65 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$, tandis que la luminosité intégrée s'élève à 5.9 fb^{-1} [135]. En 2012, l'énergie des collisions a été augmentée de 7 à 8 TeV. Le pic de luminosité instantanée a également augmenté en passant à $7.73 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$, notamment grâce à une augmentation du nombre de protons par paquet. Cela a permis de produire l'équivalent de 22.8 fb^{-1} de données en une année seulement. L'espacement des paquets de protons pendant toute la période 2010-2012 était de 50 nanosecondes. La figure 28 montre l'évolution de la luminosité intégrée produite au LHC et enregistrée par l'expérience ATLAS au cours du *run 1*. La proximité des différentes courbes montre le bon déroulement général de la première prise de données.

L'empilement est également un paramètre sensible des prises de données, qui a des conséquences importantes sur la reconstruction et l'identification des particules dans le détecteur (voir partie III). Très dépendant de la luminosité instantanée, celui-ci a augmenté considérablement entre 2011 et 2012, comme le montre la figure 29. Ainsi, l'empilement moyen $\langle \mu \rangle$ pendant les années 2011 et 2012 s'élève respectivement à 9.1 et 20.7 interactions par croisement de faisceaux. Notons que des collisions plomb-plomb (2010-2011) et proton-plomb (2013) ont également été produites au cours du *run 1*, avec une luminosité intégrée respective environ huit et six ordres de grandeur plus faible et un empilement négligeable dans les deux cas. L'analyse de données principale effectuée pendant les années de thèse (voir partie IV) utilise les collisions proton-proton de 2012 à 8 TeV présentées dans cette section.

Dans la section 4.4, un premier état des lieux de la seconde prise de données, débutée en 2015, est dressé.

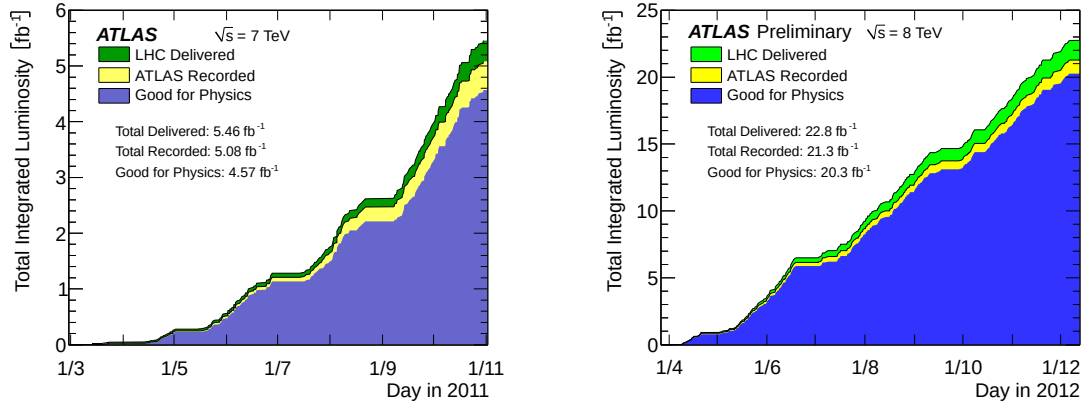


FIG. 28 – Évolution de la luminosité intégrée au cours des années 2011 (gauche, $\sqrt{s} = 7$ TeV) et 2012 (droite, $\sqrt{s} = 8$ TeV). La luminosité produite au LHC correspond à la bande verte, tandis que celle enregistrée par ATLAS est représentée en jaune. Parmi les données prises par ATLAS, celles utilisables pour la physique (détecteurs en bon état de fonctionnement) est représentée en bleu. Environ 0.48 fb^{-1} de données à $\sqrt{s} = 7$ TeV ont également été délivrées en 2010, non représenté ici [137].

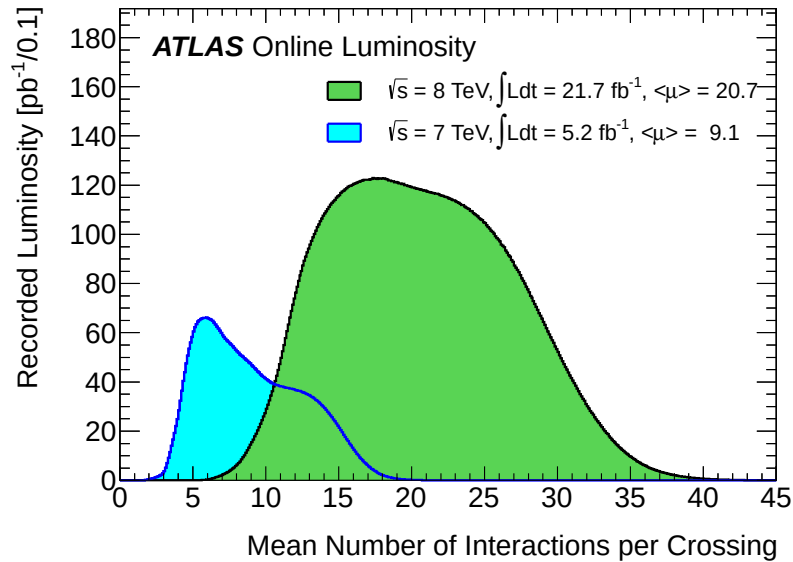


FIG. 29 – Distribution du nombre moyen d'interactions par croisement de faisceau μ en 2011 (bleu, $\sqrt{s} = 7$ TeV) et 2012 (vert, $\sqrt{s} = 8$ TeV), pondérée par la luminosité intégrée correspondante [137].

4.4 État des lieux de la seconde prise de données (2015-2018) et projet à plus long terme

Au terme d'un arrêt d'exploitation de deux ans au cours duquel de nombreuses améliorations et maintenances ont été effectuées sur la machine, le LHC a redémarré avec succès en juin 2015 avec une énergie nominale de collisions portée à 13 TeV. Environ 225 pb^{-1} de données de collisions pp ont été délivrées avec des paquets espacés de 50 nanosecondes et un empilement moyen de 20 entre juin et août 2015. Le collisionneur a fait ensuite circuler de septembre à novembre des faisceaux comportant des paquets espacés de 25 nanosecondes, ce qui a permis une augmentation conséquente de la luminosité tout en gardant l'empilement sous contrôle. Le jeu de données proton-proton complet de 2015 à 13 TeV correspond à une luminosité intégrée de 3.87 fb^{-1} , comme le montre la figure 30, gauche. Le pic de luminosité instantanée est légèrement plus faible comparé aux données de 2012 ($5.1 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$), ce qui est largement compensé par le passage à des *bunchs* espacés de 25 nanosecondes. L'empilement est également légèrement inférieur aux données 2012, avec une valeur moyenne d'interaction par croisement de faisceaux $\langle \mu \rangle$ égale à 13.7.

Notons également qu'une campagne de collisions d'ions lourds a été conduite en décembre 2015, menant à $680 \mu\text{b}^{-1}$ de données Pb-Pb à 5 TeV par nucléon enregistrées par le détecteur ATLAS. La machine est désormais arrêtée pour l'hiver, avec un redémarrage prévu pour avril 2016 et donc actuellement en cours.

Le plan d'opération à plus long terme du LHC est présenté dans la figure 31 (juin 2015). De nouvelles données proton-proton à $\sqrt{s} = 13 \text{ TeV}$ vont être enregistrées à partir de fin avril 2016. La seconde prise de données devrait s'achever fin 2018. Un arrêt long est ensuite prévu pour permettre d'augmenter à nouveau la luminosité instantanée du LHC afin de pouvoir constituer un total de 300 fb^{-1} de données proton-proton à 13 TeV d'ici 2023. Finalement, un nouvel arrêt long débutant en 2024 permettra une prise des données à ultra-haute luminosité (HL-LHC, *high luminosity LHC*), marquant le début de la seconde phase d'exploitation (Phase-2). Une mise à jour très importante de la machine et des expériences sera alors nécessaire afin de leur faire supporter entre autres des très haut taux d'empilement ($\simeq 200$). La Phase-2 vise à collecter 3000 fb^{-1} de données de collisions proton-proton à 13 TeV au terme de dix ans d'exploitation supplémentaires. Le projet HL-LHC a été approuvé officiellement par le conseil du CERN en juin 2016. Il a été défini comme la priorité absolue de la stratégie européenne pour la physique des particules lors de sa mise à jour de 2013, et il est inscrit sur la feuille de route 2016 du Forum stratégique européen sur les infrastructures de recherche (ESFRI). Le projet a aussi été considéré comme prioritaire dans le processus d'établissement des priorités pour les projets de physique des particules aux États-Unis (P5), ainsi que dans la vision stratégique du Japon pour la discipline. Cependant, en pratique, son succès semble aujourd'hui toujours conditionnée aux résultats de la Phase-1 d'exploitation.

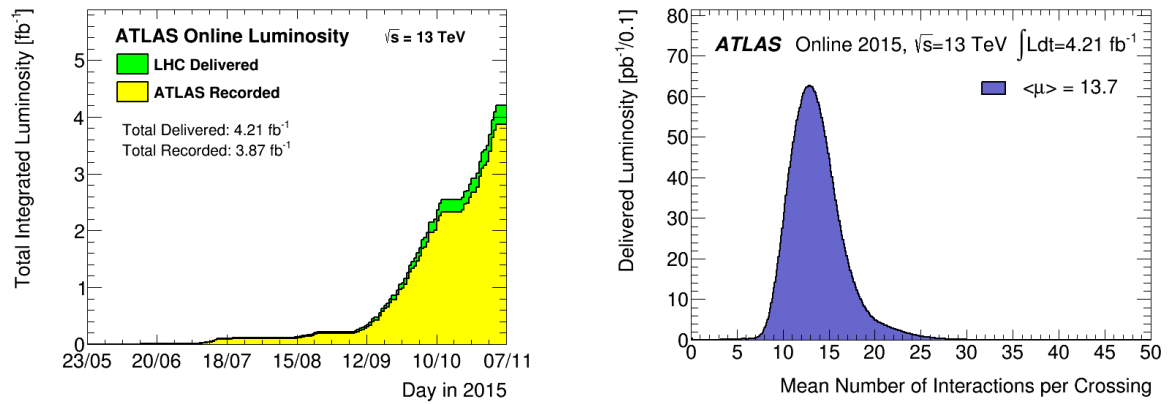


FIG. 30 – Évolution de la luminosité intégrée au cours de l'année 2015 à $\sqrt{s} = 13$ TeV (gauche). Distribution du nombre moyen d'interactions par croisement de faisceau μ pondérée par la luminosité intégrée correspondante (droite). De juin à août 2015, les données ont été délivrées avec des paquets espacés de 50 nanosecondes, tandis que l'espacement a été abaissé à 25 nanosecondes de septembre à novembre 2015 [138].

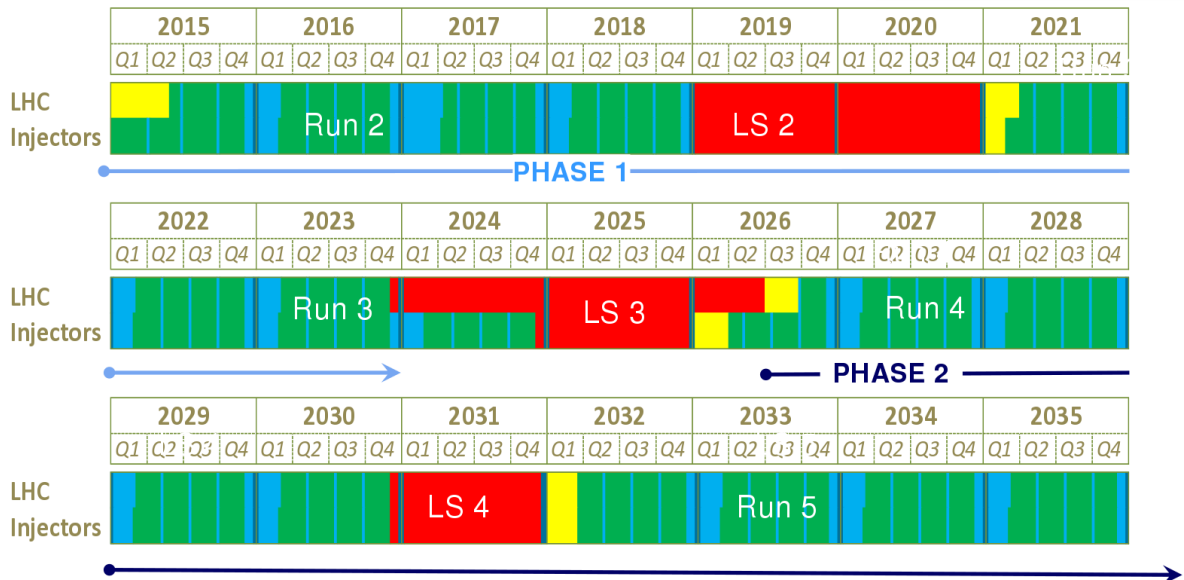


FIG. 31 – Plan à long terme pour les opérations du LHC. Les prises de données pour la physique sont représentées en vert, les arrêts longs en rouge, les arrêts techniques en bleu et les périodes de démarrage en jaune [139].

5 L'expérience ATLAS

5.1 Description générale

Le détecteur de particules ATLAS [72] est une expérience polyvalente de très grande taille, située à l'un des quatre points de collision du LHC. Afin de répondre aux conditions expérimentales imposées par le collisionneur et les objectifs du programme de physique, le détecteur a été conçu avec une granularité très fine, une très grande couverture angulaire, un système de déclenchement rapide et efficace et une bonne tenue aux radiations.

Un schéma descriptif général de l'expérience ATLAS est représenté dans la figure 32. Haut d'environ vingt-cinq mètres pour une longueur de quarante-quatre mètres, ATLAS est constitué de trois modules principaux : le module principale (environ vingt-deux mètres de long), qui est centré sur le point d'interaction et appelé tonneau (*barrel*), et deux modules cylindriques plus petits placés aux extrémités, appelés bouchons (*end-caps*). L'ensemble du détecteur pèse environ sept mille tonnes.

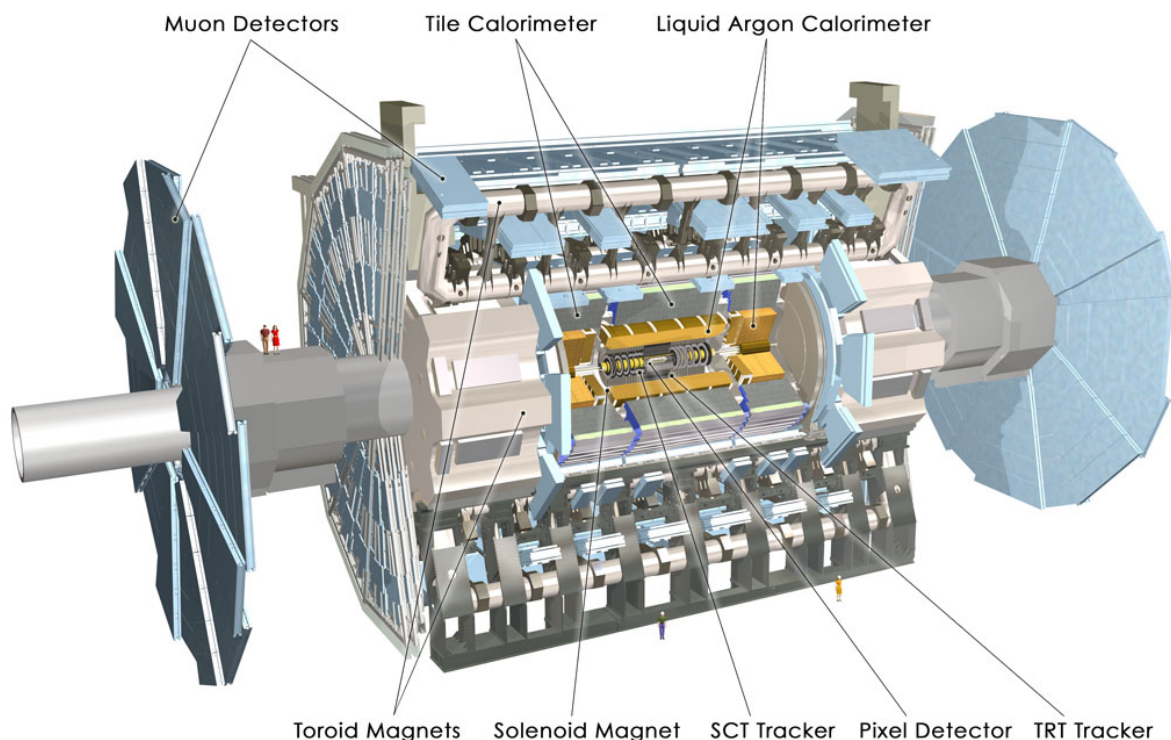


FIG. 32 – Vue d'ensemble du détecteur ATLAS et de ses principaux constituants [140].

Chacun des modules présente une structure en pelure d'oignon autour du tube faisceau, typique des détecteurs de particules polyvalents. Chaque couche intègre un sous-détecteur particulier dont la liste principale est donnée ci-dessous, de l'intérieur vers l'extérieur :

- ◊ **Détecteur interne** : trajectographe composé de trois parties distinctes (détecteur à pixels de silicium, détecteur à micro-pistes de silicium et détecteur à rayonnement de transition) mesurant la trajectoire des particules chargées.

- ◇ **Calorimètre électromagnétique** : composé d'argon liquide entrecoupé d'absorbeurs en plomb, il permet de déterminer précisément l'énergie et la position des électrons et des photons qui y sont stoppés.
- ◇ **Calorimètre hadronique** : composé d'une partie à tuiles dans la partie tonneau combinant polystyrène dopé/acier et de parties en cuivre dans les bouchons, il permet de déterminer précisément l'énergie et la position des jets qui y sont stoppés.
- ◇ **Spectromètre à muons** : trajectographe composé de différents types de chambres à muons, permettant leur identification et la mesure précise de leur impulsion.

L'expérience présente également deux systèmes magnétiques supraconducteurs. Ceux-ci courbent les trajectoires des particules chargées dans le détecteur interne et dans le spectromètre à muons afin de pouvoir en déterminer leur impulsion :

- ◇ **Solénoïde** : situé dans un cryostat entre le détecteur interne et le calorimètre électromagnétique, il plonge le détecteur interne dans un champ magnétique constant de 2 teslas orienté selon l'axe du tube faisceau. Cela a pour effet de courber la trajectoire des particules chargées dans le plan transverse au faisceau, permettant la mesure précise de leur impulsion.
- ◇ **Toroïdes** : des tores situés dans des cryostats autour du calorimètre hadronique génèrent un champ magnétique quasi-perpendiculaire à la trajectoire des particules chargées passant dans le spectromètre à muons. Le champ présente une valeur moyenne de 0.5 tesla mais varie significativement en fonction des régions. Dans la partie tonneau, le champ toroïdal est constitué de huit boucles de courant réparties autour du détecteur, tandis que les deux bouchons intègrent des tores de plus petites tailles pour refermer les lignes de champs.

Un schéma illustrant le parcours des différents types de particules traversant le détecteur est représenté dans la figure 33. Les photons et les électrons s'arrêtent dans le calorimètre électromagnétique alors que les hadrons (protons, neutrons, ...) sont stoppés dans le calorimètre hadronique. Les muons et les neutrinos sont les seules particules du modèle standard ayant la possibilité de traverser ATLAS entièrement. Les muons laissent des traces dans le spectromètre et sont ensuite arrêtés par les couches d'acier entourant le détecteur, tandis que les neutrinos s'échappent hors du détecteur sans interagir et donc ne sont pas détectés directement, du fait de leur section efficace très faible avec la matière. Cependant, la détermination partielle de leur cinématique est possible de façon indirecte, comme nous le verrons dans la section 5.2.

Dans la section suivante, le système de coordonnées utilisé dans ATLAS est introduit et certaines observables importantes pour l'étude des collisions hadroniques sont présentées. Ces notions seront largement utilisées par la suite, en particulier pour la description plus approfondie des différents systèmes d'ATLAS (sections 5.3 à 5.7), pour les détecteurs à l'avant du LHC (section 6), la reconstruction des photons dans ATLAS (partie III) et les deux travaux principaux de thèse (partie IV et V).

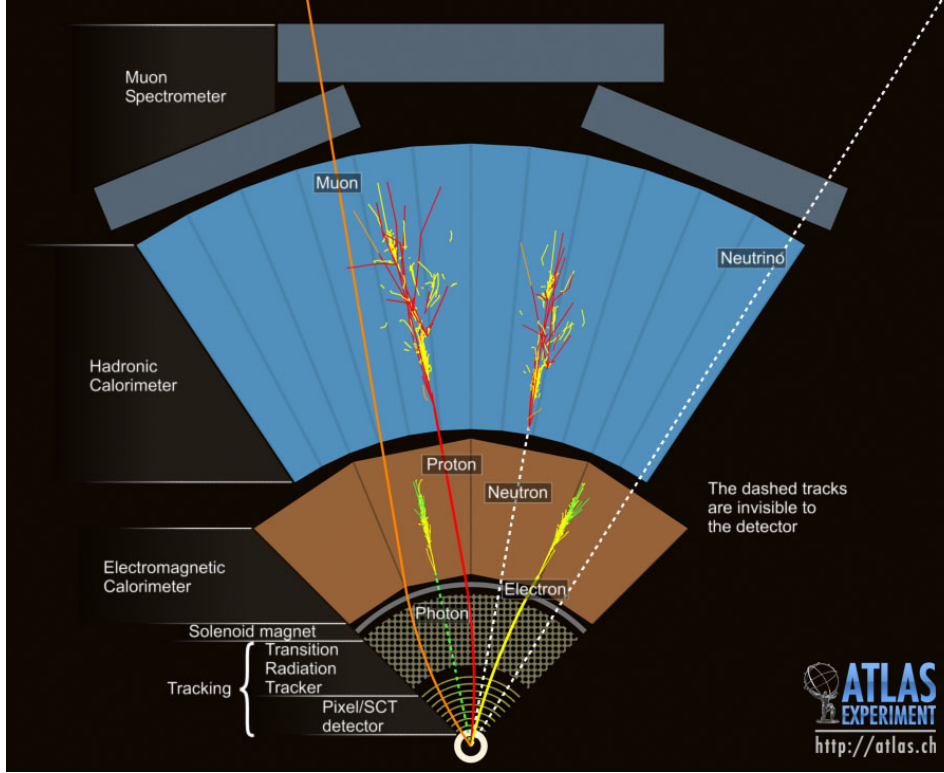


FIG. 33 – Trajets de différents types de particules dans le détecteur ATLAS [141].

5.2 Système de coordonnées et observables

Le système de coordonnées d'ATLAS est un système cartésien direct qui a pour origine le point d'interaction au centre du détecteur. L'axe z suit l'axe de collision tandis que les axes x et y définissent le plan transverse à la direction du faisceau. L'axe x pointe vers le centre de l'anneau du LHC et l'axe y pointe vers la surface. La coordonnée radiale dans le plan (x, y) est notée R , et l'axe associé r .

Ainsi, le champ magnétique solénoïdal courbe la trajectoire des particules chargées traversant le détecteur interne dans le plan (x, y) , tandis que le champ toroïdal agit sur les muons au niveau du spectromètre dans le plan (r, z) . La mesure des rayons de courbure permet alors une détermination précise des différentes impulsions.

On définit également l'angle d'azimut ϕ dans le plan transverse (x, y) , et l'angle polaire θ dans le plan (r, z) . On utilise cependant souvent la rapidité y au lieu de l'angle polaire θ :

$$y = \frac{1}{2} \cdot \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (22)$$

où E est l'énergie de la particule considérée et p_z est la projection de son impulsion selon l'axe z . Dans le cas d'une particule ultra-relativiste ($E \gg m$), la pseudo-rapidité η est une bonne approximation de la rapidité :

$$\eta = -\ln \left| \tan \frac{\theta}{2} \right|. \quad (23)$$

Cette approximation est largement vérifiée dans un collisionneur à très haute énergie comme le LHC. Une des propriétés intéressantes de y et η est qu'une différence de (pseudo-)rapidité est invariante selon n'importe quelle translation de Lorentz le long de l'axe z . Or, dans les collisionneurs hadroniques, les interactions principales de hautes énergies correspondent à des interactions parton-parton et non proton-proton, ce qui génère des centres de masse de collisions translatés en z les uns par rapport aux autres (voir section 3). Ces centres de masse sont également translatés en z par rapport au référentiel du laboratoire, qui est confondu avec le centre de masse du système proton-proton. La pseudo-rapidité permet alors de définir simplement un certain nombre d'observables indépendamment du référentiel dans le centre de masse des collisions. Par exemple, la séparation angulaire entre deux particules peut être défini comme la moyenne quadratique entre différence d'azimut et différence de rapidité :

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \quad (24)$$

Cette formulation permet l'application d'un critère similaire quelle que soit l'interaction parton-parton. Un ΔR constant autour d'une particule donnée définit alors un cône avec une ouverture croissante en pseudo-rapidité.

De plus, le flux de particules produit dans les collisions étant du même ordre de grandeur par unité de pseudo-rapidité, il est courant d'utiliser η pour définir la segmentation des détecteurs, comme nous le verrons dans les sections 5.3 à 5.7. On appelle souvent les domaines de grande rapidité ($|\eta| > 2$) les régions à l'avant, à opposer avec la région centrale qui se situe à plus faible rapidité ($|\eta| < 2$). Les couvertures en pseudo-rapidité des différents sous-détecteurs d'ATLAS mentionnés dans la section précédente sont résumées dans la table 1. Les dimensions en termes de rayons et longueurs sont également précisées.

Composant	Rayon [m]	Longueur [m]	Couverture [η]
Spectromètre à muons (tonneau)	11	26	$ \eta < 1.4$
Spectromètre à muons (bouchons)	11	2.8	$1.1 < \eta < 2.8$
Calorimètre hadronique (tonneau)	4.25	12.2	$ \eta < 1.7$
Calorimètre hadronique (bouchons)	3.0	1.8	$1.5 < \eta < 3.2$
Calorimètre électromagnétique (tonneau)	3.0	6.4	$ \eta < 1.475$
Calorimètre électromagnétique (bouchons)	3.0	0.63	$1.375 < \eta < 3.2$
Détecteur interne (tonneau + bouchons)	1.05	6.2	$ \eta < 2.4$

TAB. 1 – Dimensions et couverture en pseudo-rapidité des principaux sous-détecteurs d'ATLAS.

De manière générale, le plan transverse (x, y) joue un rôle particulier dans les collisionneurs hadroniques. Cela vient du fait que dans le référentiel du laboratoire, l'impulsion selon z d'un système formé de deux partons rentrant en collision est non-nulle et inconnue, alors qu'elle est nulle dans le plan transverse en première approximation.¹⁹ Ainsi, par conservation de l'impulsion dans le plan (x, y) , on obtient la relation suivante

¹⁹Ceci est vrai à une erreur près correspondant à l'ordre de grandeur de Λ_{QCD} (soit quelques centaines de MeV, voir section 3.1) à cause d'effets reliés à la structure interne du proton.

pour une interaction parton-parton donnée avec n particules dans l'état final :

$$\sum_{i=1}^n \vec{p}_{T,i} = \vec{0}, \quad (25)$$

où les $\vec{p}_{T,i}$ sont les impulsions transverses des particules de l'état final. Les neutrinos étant les seules particules n'interagissant pas avec le détecteur, il est courant d'utiliser la relation de l'équation 25 afin de définir une observable importante pour leur étude appelée énergie transverse manquante :

$$\vec{E}_T^{miss.} = - \sum_{i=1}^n \vec{p}_{T,i}, \quad (26)$$

où $\vec{E}_T^{miss.}$ correspond à la somme vectorielle des $\vec{p}_T$ des neutrinos émis dans l'événement. L'hermiticité du détecteur est fondamentale pour son calcul précis.

Par ailleurs, les coupures autorisées sur la cinématique d'un état final particulier, c'est-à-dire préservant l'annulation des divergences en théorie des perturbations (voir section 3.4), font souvent intervenir les impulsions transverses des particules et non leur énergie totale à cause des impulsions initiales des partons en z inconnues.

Dans les sections suivantes, des détails supplémentaires sont donnés sur les sous-détecteurs d'ATLAS, en particulier pour ceux utilisés dans le cadre des travaux de thèse.

5.3 Le détecteur interne

Le détecteur interne est le sous-détecteur d'ATLAS le plus proche du tube faisceau, et donc le premier à être traversé par les produits de collisions. C'est un trajectographe de forme cylindrique qui a pour but de mesurer l'impulsion des particules chargées dans un large domaine en pseudo-rapacité ($|\eta| < 2.5$, voir tableau 1), grâce à la courbure de leur trajectoire par le champ magnétique du solénoïde. Il permet également la reconstruction des positions des points d'interactions primaires et secondaires, appelés vertex.

Il est constitué de trois parties distinctes : le détecteur à pixels de silicium, le détecteur à micro-pistes de silicium (ou SCT pour *semi-conductor tracker*) et le détecteur à rayonnement de transition (TRT pour *transition radiation tracker*). L'ensemble est représenté dans la figure 34.

5.3.1 Le détecteur à pixels de silicium

Le détecteur à pixels comporte une très fine segmentation afin de pouvoir reconstruire les vertex des interactions. Ses modules en silicium fournissent typiquement trois points de mesure de la trajectoire à proximité immédiate du faisceau. Dans la partie tonneau, il est constitué de trois couches cylindriques concentriques de rayon respectif 50.5 mm, 85.5 mm et 122.5 mm. La couche la plus proche du faisceau, appelée couche- b (*b-layer*), est essentielle à la reconstruction précise des vertex, en particulier pour les

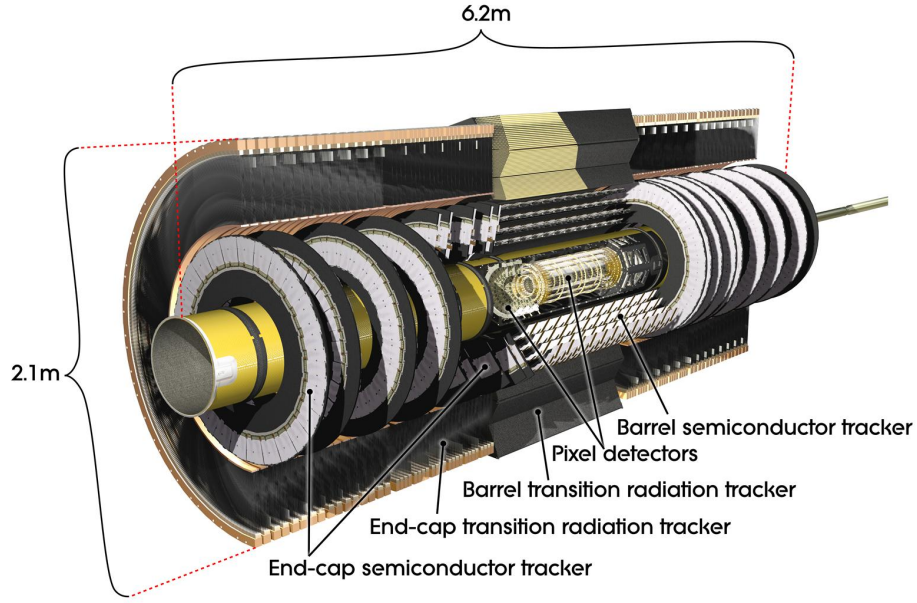


FIG. 34 – Vue d’ensemble du détecteur interne d’ATLAS et de ses principaux constituants [142].

hadrons composés de quarks b (étiquetage des jets de b) ou les taus. Elle est également utile pour différencier les électrons des photons convertis en paires électron-positron par interaction avec le milieu, les conversions ayant lieu généralement après la traversée de la couche- b . Dans chaque bouchon, une série de trois disques autour de l’axe du faisceau positionnés respectivement en $z = \pm 495$ mm, ± 580 mm et ± 650 mm permettent de couvrir la zone $1.1 < |\eta| < 2.5$.

Au total, le détecteur à pixels comporte environ 80 millions de voies de lectures. La taille d’un pixel est de $50\ \mu\text{m}$ dans le plan $R-\phi$ et $400\ \mu\text{m}$ en z . La précision attendue sur la position est de l’ordre de $14\ \mu\text{m}$ dans le plan $R-\phi$ et de $115\ \mu\text{m}$ dans la direction longitudinale (radiale) pour le tonneau (bouchons). Enfin, un refroidissement à environ $-10\ ^\circ\text{C}$ est assuré en permanence afin de limiter les dégâts dus aux fortes radiations engendrées par la proximité du faisceau.

Pour la seconde prise de données (*run 2*), une quatrième couche de pixels appelée IBL (*insertable b-layer*) a été installée afin d’améliorer l’identification des jets de b et présente de bonnes performances après prise en compte de sa déformation mécanique [143].

5.3.2 Le détecteur à micro-pistes de silicium (SCT)

Le détecteur à micro-pistes de silicium (SCT) participe à la reconstruction des trajectoires des particules chargées et des vertex, en complément du détecteur à pixels. Dans le tonneau, il est constitué de quatre couches cylindriques concentriques de rayon respectif 299 mm, 371 mm, 443 mm et 514 mm, permettant une couverture de la zone $|\eta| < 1.1$. Dans chaque bouchon, neuf disques permettent d’étendre la couverture jusqu’à $|\eta| < 2.5$. Le SCT utilise une technologie à base de silicium similaire au détecteur à pixels (voir section 5.3.1), mais avec des modules segmentés en micro-pistes verticales et horizontales alternées. Cette géométrie présente l’avantage de réduire considérablement

le nombre de canaux de lecture. Chaque couche inclut deux plaques dos-à-dos, ce qui permet de fournir en moyenne huit points de mesure additionnels pour le calcul des trajectoires.

Le SCT compte environ 6.3 millions de voies de lectures pour une surface totale de 63 m^2 . La précision attendue sur la position est de l'ordre de $17 \text{ } \mu\text{m}$ dans le plan $R\text{-}\phi$ et de $580 \text{ } \mu\text{m}$ dans la direction longitudinale (radiale) pour le tonneau (bouchons).

Pour la seconde prise de données, l'électronique de transmission du détecteur à micro-pistes a été considérablement améliorée, avec notamment une meilleure tenue à l'humidité [144].

5.3.3 Le détecteur à rayonnement de transition (TRT)

Le détecteur à rayonnement de transition (TRT) est constitué de tubes à dérives (ou pailles) de 4 mm de diamètre contenant un mélange gazeux Xe-CO₂-O₂ (70% Xe, 27% CO₂ et 3% O₂ pour le *run 1*) et un fil de tungstène, alternés avec des fibres ou films minces de polypropylène/polyéthylène. Lorsqu'une particule chargée relativiste passe du polymère à un tube, elle émet un rayonnement de transition qui ionise le mélange gazeux et délivre un signal, collecté par le fil de tungstène. La partie centrale est constituée de 53 000 tubes répartis sur trois cylindres de 160 cm de long, tandis que chaque bouchon compte 125 000 tubes répartis sur trois roues. La couverture en pseudo-rapacité du TRT est légèrement plus faible que pour le détecteur à pixels et le SCT, avec $|\eta| < 2.0$ seulement.

Le TRT permet de créer un grand nombre de points de mesure ($\simeq 40$), très utile pour la reconstruction des traces. De plus, il permet de distinguer efficacement les électrons des hadrons légers comme les pions ou les kaons grâce à l'amplitude du rayonnement de transition, qui est typiquement au-dessus de 6 keV pour les électrons et à bien plus basse valeur pour les hadrons légers ($< 0.5 \text{ keV}$), grâce à leur masse environ 250 fois plus élevée. Deux seuils en amplitude sont paramétrés sur le détecteur afin de séparer les deux contributions. La précision attendue pour la position d'une particule dans le plan $R\text{-}\phi$ est de l'ordre de $130 \text{ } \mu\text{m}$, significativement plus faible que pour les couches de pixels ou le SCT.

Pendant la première prise de données, des fuites importantes de gaz ont été constatées à cause de problèmes de corrosion. Pour la seconde prise de données, le mélange gazeux a été modifié afin de minimiser les pertes. Le système de lecture a également été amélioré afin de pouvoir utiliser le TRT avec des collisions toutes les 25 ns [145].

5.3.4 Résolution globale attendue et quantité de matière

En faisant l'hypothèse d'une trace reconstruite dans le tonneau avec trois mesures de trajectoire dans le détecteur à pixel, huit dans le SCT, et trente-six dans le TRT, la résolution attendue sur la mesure de l'impulsion transverse d'une particule chargée dans le détecteur interne est donnée par [146] :

$$\frac{\sigma_{\text{pT}}}{\text{pT}} = 0.05\% \cdot \text{pT} [\text{GeV}] \oplus 1\%, \quad (27)$$

où σ_{p_T} représente l'écart-type de la gaussienne de résolution et $\oplus$ précise une somme en quadrature. La résolution attendue est donc globalement excellente, avec une dégradation pouvant néanmoins devenir importante pour les hauts p_T ($p_T > 100$ GeV). Cependant, une très bonne résolution est absolument nécessaire surtout à basse énergie afin de garder les effets d'empilement sous contrôle (voir section 3.2).

Par ailleurs, la quantité de matière dans le détecteur interne moyenné sur ϕ en fonction de la pseudo-rapacité pour les différents sous-détecteurs est montrée dans la figure 35. Elle est donnée en terme de nombre de longueurs de radiation X/X_0 , qui a été déterminé précisément au cours de la première prise de données [147]. Une longueur de radiation correspond à la distance moyenne où un électron de haute énergie E perd une énergie $(E - 1/E)$ par interaction avec la matière environnante, ou de façon équivalente au 7/9 du parcours libre moyen d'un photon avant conversion en une paire e^+e^- .

La distribution de quantité de matière dans le détecteur interne présente des inhomogénéités importantes en fonction de η , variant d'un facteur trois à quatre entre $|\eta| < 0.6$ et $0.6 < |\eta| < 1.81$. Elle baisse à nouveau pour $1.81 < |\eta| < 2.37$. La région au-delà n'est pas couverte par le détecteur de traces et par conséquent est souvent exclue des analyses de physique. Le caractère inhomogène de cette distribution a des conséquences importantes sur la fraction de photons convertis, plus nombreux dans la fin du tonneau et les bouchons, et plus généralement sur la calibration du calorimètre électromagnétique se trouvant plus en aval, comme nous le verrons dans la partie III.

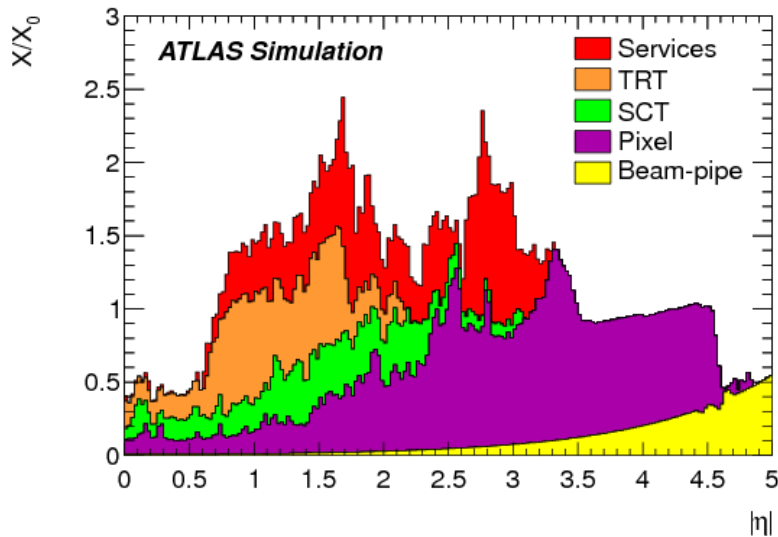


FIG. 35 – Quantité de matière dans le détecteur interne moyenné sur ϕ en fonction de la pseudo-rapacité, mesurée en termes de nombre de longueurs de radiations. Les contributions des différents sous-détecteurs sont séparées en bandes de couleurs différentes [147].

5.4 Le calorimètre électromagnétique

5.4.1 Le calorimètre électromagnétique à argon liquide (LAr EM)

Le calorimètre électromagnétique à argon liquide (LAr EM) mesure l'énergie et la direction des photons et des électrons dans la région $|\eta| < 3.2$, en disposant d'une granularité fine jusqu'à $|\eta| < 2.37$. Sa partie tonneau s'étend jusqu'à $|\eta| < 1.475$, tandis que les deux bouchons couvrent $1.375 < |\eta| < 3.2$. Leur épaisseur correspond respectivement à environ 22 et 24 longueurs de radiation X/X_0 , ce qui leur permet d'arrêter la très grande majorité des photons et les électrons. La partie tonneau est contenue dans le même cryostat que celui utilisé pour le solénoïde, alors que les bouchons sont placés dans des cryostats séparés.

Un schéma général du calorimètre électromagnétique d'ATLAS est présenté dans la figure 36. La région de transition entre tonneau et bouchons ($1.37 < |\eta| < 1.56$) contient un certain nombre de câbles et de services pour le détecteur interne, et est faiblement instrumentée. Par conséquent, elle est généralement exclue des analyses de physique. La partie tonneau est divisée en deux tonneaux identiques placés de part et d'autre de $z = 0$, tandis que chaque bouchon est divisé en deux roues, couvrant respectivement $1.375 < |\eta| < 2.5$ et $2.5 < |\eta| < 3.2$.

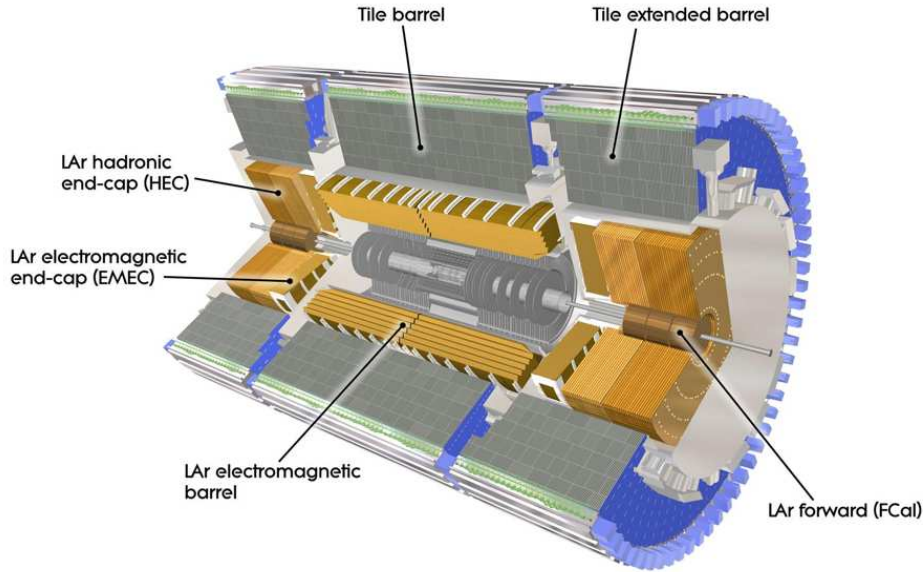


FIG. 36 – Vue d'ensemble du système calorimétrique d'ATLAS, composé de calorimètres à argon liquide (électromagnétique, hadronique) et à tuiles (hadronique) [148].

Chaque module du LAr EM est composé d'absorbeurs en plomb, initiant les cascades électromagnétiques, et d'argon liquide, qui constitue le milieu actif. Les électrodes de lecture sont composées de trois couches de cuivre entrecoupées de kapton, ce qui permet de transmettre les signaux induits par les électrons d'ionisation dans l'argon liquide par couplage capacitif, grâce à une tension de 2000 V nominale appliquée entre l'argon liquide et les couches externes des électrodes. De plus, afin d'assurer une couverture complète en ϕ et des performances similaires pour les signaux collectés à l'avant ou à l'arrière de l'électrode, absorbeurs et électrodes sont répartis selon une géométrie en

accordéon. Les couches de plomb ont une épaisseur respective d'environ 1.5 et 1.7 mm dans le tonneau et les bouchons. Les couches d'argon liquide ont une épaisseur de 2.1 mm dans le tonneau et de 0.9 à 3.8 mm dans les bouchons. La composition des électrodes et la géométrie en accordéon du calorimètre sont illustrées dans la figure 37, gauche.

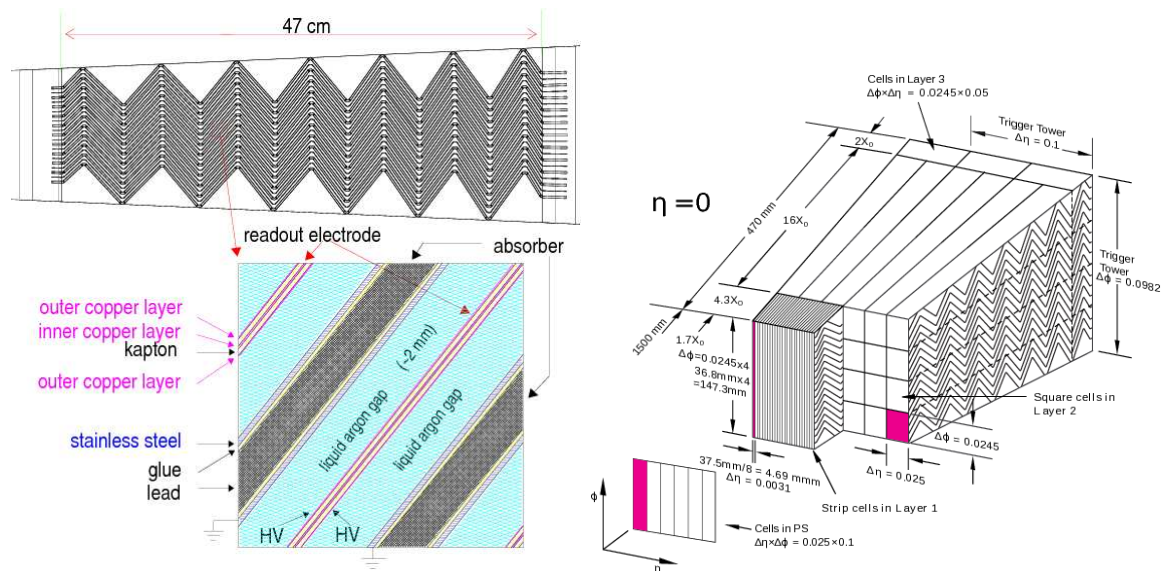


FIG. 37 – Illustration de la géométrie en accordéon dans la partie tonneau du calorimètre électromagnétique d’ATLAS (gauche) et schéma général montrant la structure en trois couches (droite) [149].

Le LAr EM est segmenté longitudinalement en trois couches, comme présenté dans la figure 37, droite. Chaque couche présente des caractéristiques et fonctions particulières :

- ◇ **La première couche** présente une profondeur d'environ $4 X_0$. Elle s'étend jusqu'à $|\eta| < 2.37$ et est formée par des bandes radiales avec une granularité très fine en η d'environ $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$ radian, dont la valeur exacte dépend de la pseudo-rapidité. Cette première couche est très importante pour séparer les gerbes spatialement proches, typiquement émises lors de désintégrations de mésons neutres légers comme $\pi^0 \rightarrow \gamma\gamma$.
- ◇ **La seconde couche** est celle qui reçoit la plus grande partie de l'énergie collectée par le calorimètre, avec une profondeur qui varie de 16 à 18 X_0 . Elle présente une granularité moins fine de $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ radian et assure avec la première couche la mesure en (η, ϕ) de la particule incidente. Celle-ci constitue la seule mesure angulaire pour les photons ne laissant pas de traces exploitables dans le détecteur interne.
- ◇ **La troisième et dernière couche** a une profondeur d'environ 2 X_0 et sert à estimer les pertes dues aux queues des gerbes non contenues dans le LAr EM. Sa granularité en $\Delta\eta \times \Delta\phi$ est de 0.05×0.025 .

De plus, la région $|\eta| < 1.8$ est instrumentée avec un pré-échantillonneur situé devant le calorimètre, qui présente une granularité $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$ radian. Celui-ci permet

d'estimer les pertes d'énergies des photons et des électrons dans la matière située en amont du calorimètre (détecteur interne, solénoïde, cryostat).

La résolution angulaire du calorimètre électromagnétique est d'environ $\frac{50[\text{mrad}]}{\sqrt{E[\text{GeV}]}}$, ce qui permet une bonne distinction entre les différentes cascades électromagnétiques. Sa résolution en énergie est donnée par la somme en quadrature de trois termes indépendants :

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c. \quad (28)$$

- ◇ $\frac{a}{\sqrt{E}}$ **est le terme d'échantillonnage.** Il décrit les fluctuations statistiques liées au développement de la gerbe électromagnétique dans l'argon liquide et le plomb. Sa valeur est de l'ordre de 10% à $|\eta| = 0$ et comprise typiquement entre 0 et 20%.
- ◇ $\frac{b}{E}$ **est le terme de bruit.** Il décrit les fluctuations venant principalement de la chaîne électronique et des événements d'empilement. Sa valeur est de l'ordre de 300 MeV, mais varie avec la luminosité instantanée et peut devenir particulièrement important à basse énergie.
- ◇ **Le terme c, indépendant de l'énergie, est le plus délicat à estimer.** Il traduit la non-uniformité de la réponse du calorimètre, venant par exemple des légers défauts des matériaux, de la qualité de l'échantillonnage, ou de la linéarité de réponse. C'est ce terme qui domine la résolution à haute énergie et donc qui nous intéressera par la suite (voir partie IV et V). La calibration finale de la première prise de données l'estime à 1.2% dans le tonneau, et 1.8% dans les bouchons pour $|\eta| < 2.47$. Au delà, il atteint 3.3% [149].

5.4.2 Le calorimètre électromagnétique à l'avant

Dans la région très à l'avant ($3.1 < |\eta| < 4.9$), les calorimètres électromagnétiques à l'avant (voir figure 36) assurent une couverture grâce à une technologie incluant une matrice de cuivre entrecoupée d'argon liquide. Leur granularité en $\Delta\eta \times \Delta\phi$ est de 0.1×0.1 radian pour un total d'environ 2000 canaux de lecture. Cette configuration permet de supporter les radiations, très fortes à l'avant des faisceaux. Cependant, la résolution en énergie de ce détecteur est significativement plus faible et de l'ordre de :

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%. \quad (29)$$

Son objectif principal est d'assurer une meilleure hermiticité de l'expérience ATLAS, utile notamment à la détermination de l'énergie transverse manquante (voir section 5.2).

Aucune amélioration significative n'a été apportée au calorimètre électromagnétique pour la seconde prise de données. Ses seuils de déclenchement ont été néanmoins adaptés aux nouvelles conditions imposées par le LHC.

5.5 Le calorimètre hadronique

Le calorimètre hadronique permet de mesurer l'énergie des jets et de calculer l'énergie transverse manquante (voir section 5.2). En effet, le calorimètre électromagnétique n'arrête en général pas les hadrons, qui ont des gerbes typiquement plus longues et nécessitent donc un détecteur supplémentaire pour leur acquisition complète. Ainsi, environ deux tiers de l'énergie des jets de haute énergie ($E > 100$ GeV) est déposée dans le LAr EM en moyenne. Le rapport entre l'énergie collectée dans le calorimètre hadronique et dans le LAr EM pour une gerbe donnée est également une observable importante utilisée pour discriminer les photons des jets, comme nous le verrons dans la partie III.

Le calorimètre hadronique est divisé en trois parties : le calorimètre à tuiles couvrant la région centrale ($|\eta| < 1.7$), les bouchons hadroniques ($1.5 < |\eta| < 3.2$), et le calorimètre hadronique à l'avant couvrant $3.1 < |\eta| < 4.9$ (voir figure 36).

5.5.1 Le calorimètre hadronique à tuiles (*TileCal*)

Le calorimètre hadronique à tuiles est situé juste après le calorimètre électromagnétique à argon liquide. Il comporte une partie tonneau ($|\eta| < 1.0$), et deux parties bouchons ($0.8 < |\eta| < 1.7$). Il s'agit d'un calorimètre à scintillateurs, comme le LAr EM. Le milieu dopé est constitué de tuiles en polystyrène dopé et les gerbes sont initiées par des couches d'acier. Les hadrons qui passent par un module excitent les molécules du milieu dopé, qui émettent de la lumière pour retourner dans leur état fondamental. La lumière est acheminée vers des photomultiplicateurs grâce à un réseau de fibres optiques et constitue le signal.

La partie tonneau comme les bouchons sont segmentés en trois couches de granularité $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$ radian pour les deux premières couches, et 0.1×0.1 radian pour la dernière. Au total, le calorimètre à tuiles est connecté à environ 7 000 voies de lectures.

5.5.2 Les bouchons hadroniques (*HEC*)

Les bouchons hadroniques couvrent la partie $1.5 < |\eta| < 3.2$. Ils sont logés dans les mêmes cryostats que les bouchons électromagnétiques et incluent des absorbeurs en cuivre avec un milieu actif formé d'argon liquide. Ils présentent une géométrie constituée de plans successifs perpendiculaires à l'axe du faisceau, regroupés en deux roues. Leur granularité typique en $\Delta\eta \times \Delta\phi$ est 0.1×0.1 radian pour $1.5 < |\eta| < 2.5$ et 0.2×0.2 radian pour $2.5 < |\eta| < 3.2$, avec environ 3 000 voies de lecture associées au total.

Les contraintes en terme de performance sont moindres pour les calorimètres hadroniques, qui souffrent de la grande variabilité des gerbes hadroniques. Cela est compensé partiellement par l'excellente résolution du calorimètre électromagnétique à argon liquide, qui mesure finalement la plus grande partie de l'énergie des jets. La résolution en énergie des calorimètres hadroniques correspond à :

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%. \quad (30)$$

5.5.3 Le calorimètre hadronique à l'avant

Le calorimètre hadronique à l'avant permet de couvrir la région $3.1 < |\eta| < 4.9$, utile notamment pour garantir une meilleure hermiticité de l'expérience. Cela permet entre autre d'améliorer la détermination de l'énergie transverse manquante (voir section 5.2). Il permet aussi de reconstruire les jets très à l'avant, utile pour certaines analyses de physique.

Ce détecteur combine des plaques de cuivre et de tungstène avec de l'argon liquide, ce qui lui permet d'atteindre un grand nombre de longueur de radiation dans une épaisseur restreinte. Sa granularité typique en $\Delta\eta \times \Delta\phi$ est 0.2×0.2 radian. Il présente une résolution similaire au calorimètre électromagnétique à l'avant.

Aucune amélioration significative n'a été apportée au calorimètre hadronique pour la seconde prise de données. Ses seuils de déclenchement ont néanmoins été adaptés à la nouvelle configuration des collisions du LHC.

5.6 Le spectromètre à muons et les autres détecteurs d'ATLAS

5.6.1 Le spectromètre à muons

Les muons et les neutrinos sont les seules particules du MS à ne pas être stoppées par le système calorimétrique d'ATLAS. Les muons étant chargés, leur trajectoire se courbe sous l'impulsion du champ toroïdal dans le plan (r, z) (voir section 5.1) et est mesurée à nouveau par le dernier sous-détecteur d'ATLAS, le spectromètre à muons. Les données du spectromètre n'ayant pas été utilisées pour les travaux de thèse, celui-ci est détaillé de façon très succincte.

Un schéma général du système à muons d'ATLAS est représenté dans la figure 38. Les chambres à muons dans le tonneau couvrent la région $|\eta| < 1.0$ tandis que les chambres des bouchons, montées sur des roues orthogonales à l'axe du faisceau, permettent d'étendre l'acceptance à $|\eta| < 2.7$. Quatres types de chambres différentes sont utilisées, deux servant pour les mesures de précision (tubes à dérives et chambres à pistes cathodiques) et deux formant un système de déclenchement autonome vis-à-vis du reste de l'expérience (chambres à plaques résistives et à fentes minces).

Les chambres de précision sont organisées en trois couches. Les tubes à dérives couvrent la plupart du domaine en pseudo-rapacité tandis que les chambres à pistes cathodiques garantissent une granularité fine dans la première couche et les régions les plus à l'avant.

Le système de déclenchement [150] permet de décider de l'acquisition ou non d'une collision en cas de présence de muons dans une large gamme en énergie. Les chambres à plaques résistives sont utilisées dans le tonneau et les chambres à fentes minces dans les bouchons. Ces chambres fournissent également une mesure précise du temps de passage des particules (résolution de l'ordre de quelques nanosecondes), ce qui permet d'associer toutes les détections de muons avec le croisement de faisceau correct.

La combinaison des mesures du spectromètre et du détecteur interne permet une détermination très précise de l'impulsion transverse des muons. La résolution finale en

combinant les deux détecteurs varie en fonction de l'énergie et de l'emplacement du muon dans le détecteur. Elle est de l'ordre de 4% (10%) pour un muon de 50 GeV (1 TeV).

ATLAS inclut également un blindage sur sa couche la plus externe afin d'arrêter les radiations restantes, provenant majoritairement de muons.

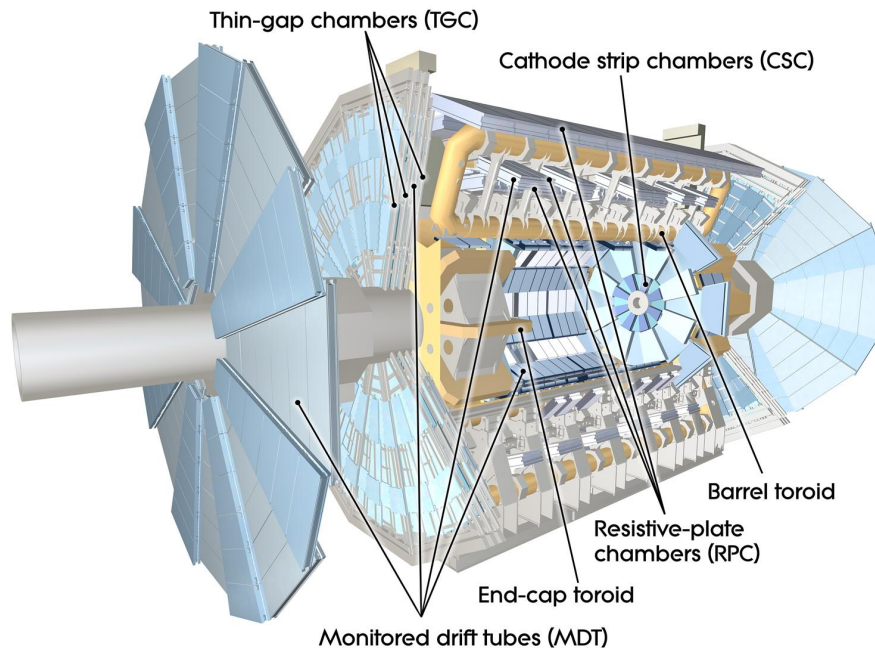


FIG. 38 – Vue d'ensemble du système à muons d'ATLAS [151].

5.6.2 Les autres détecteurs d'ATLAS

L'expérience ATLAS inclut également d'autres sous-détecteurs plus petits dans les régions très à l'avant [152] :

- ◇ **LUCID** (*luminosity measurement using Cherenkov detector*) : composé de deux modules proches du tube faisceau placés à environ dix-sept mètres de part et d'autre du point d'interaction d'ATLAS, LUCID permet de détecter le passage de particules chargées à l'aide de leur cône de Cherenkov et de photomultiplicateurs, et ainsi de surveiller la luminosité instantanée par collision au cours de la prise de données. LUCID a subi une importante mise à niveau après la première prise de données afin de pouvoir continuer à fonctionner pour le *run 2*. Il a redémarré avec succès depuis mai 2015 [153].
- ◇ **Les ZDCs** (*zero degree calorimeter*) : situés à 140 m du point d'interaction dans l'axe de collision, ces calorimètres à zéro degré permettent de détecter les particules neutres de très hautes pseudo-rapidités restant pratiquement dans l'axe des collisions ($|\eta| > 8.3$). Les particules chargées sont déviées au préalable par les aimants du LHC. Les ZDCs présentent un intérêt particulier dans le cas des collisions d'ions lourds.

- ◇ **ALFA** (*absolute luminosity for ATLAS*) : détecteurs situés à environ 240 m de part et d'autre du point d'interaction d'ATLAS pour la mesure des protons déviés à petits angles. ALFA permet de mesurer la section efficace élastique et inélastique totale proton-proton, ainsi que d'étudier des processus de physique diffractive. Quelques détails seront donnés dans la section 6.

Les zones les plus à l'avant étant exposées à de très fortes radiations, les détecteurs mentionnés plus haut peuvent être opérés uniquement lors de prise de données dédiées à plus basse luminosité avec des taux d'empilement très réduits.

5.7 Le système de déclenchement et la distribution des données

5.7.1 Le système de déclenchement

Le LHC croise les faisceaux à une fréquence de 40 (20) MHz dans le cas de paquets espacés de 25 (50) ns (voir section 4). En 2012, le taux de 1 milliard d'interactions par seconde a été atteint. ATLAS étant équipé d'une centaine de millions de canaux de lecture, il paraît clair qu'un flux de données correspondant à l'enregistrement de toutes les collisions produites est beaucoup trop important pour être correctement traité. En effet, l'enregistrement de l'intégralité du détecteur prend un temps de l'ordre de quelques millisecondes et un espace disque non négligeable. C'est de plus inutile, les physiciens ne s'intéressant qu'à certaines collisions bien particulières.

Par conséquent, l'expérience intègre un système de déclenchement (*trigger*) à trois niveaux permettant de réduire le flux de données de plusieurs dizaines de mégahertz à quelques centaines de hertz, en limitant les enregistrements aux collisions susceptibles d'être utiles aux physiciens. Ce système intègre différentes chaînes de décision plus ou moins complexes en fonction du processus étudié, appelées *menus*. Le cas du *menu* servant aux études des collisions avec au moins un électron dans l'état final est illustré dans la figure 39. Le fonctionnement général du système est résumé dans la figure 40 et décrit ci-dessous.

Le système de déclenchement de niveau 1 (L1) est entièrement basé sur l'électronique. Il utilise les informations du spectromètre à muons ainsi que des calorimètres avec une granularité réduite de $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ radian afin de définir des régions d'intérêt (RoI pour *regions of interest*). Lorsque l'énergie déposée est supérieure à un certain seuil, l'événement est transmis au niveau suivant. Certaines chaînes requièrent le rejet d'une fraction précise d'événements passant le L1 (par exemple 1 sur 10) afin de ne pas excéder le flux maximum de sortie en L1 (75 kHz). Cette fraction est appelée *prescale* et varie en fonction des chaînes. Les chaînes de déclenchement ne nécessitant pas de *prescale* sont dites non réduites ou *unprescaled*. Le temps maximal avant la prise de décision au L1 est de 2.5 μ s.

Le niveau 2 (L2) est exécuté par un programme informatique. Les données du détecteur interne deviennent accessibles et l'ensemble des informations collectées dans les RoI sont utilisées afin de vérifier certains critères de qualité et d'énergie des dépôts. Pour cela, une première reconstruction simplifiée des objets, basée uniquement sur les RoI et les traces, est effectuée. Différents *menus* sont définis en fonction du type et du nombre d'objets à étudier. Le taux de déclenchement en sortie du L2 est d'environ 3.5 kHz, et

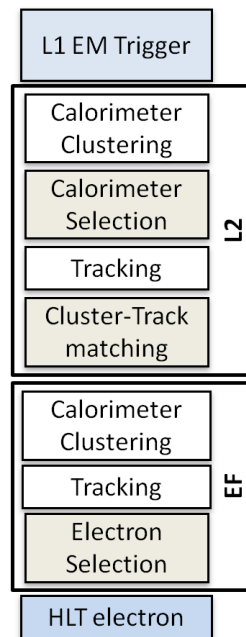


FIG. 39 – Schéma des différentes étapes du système de déclenchement d'ATLAS pour la chaîne électron au cours de la première prise de données [150].

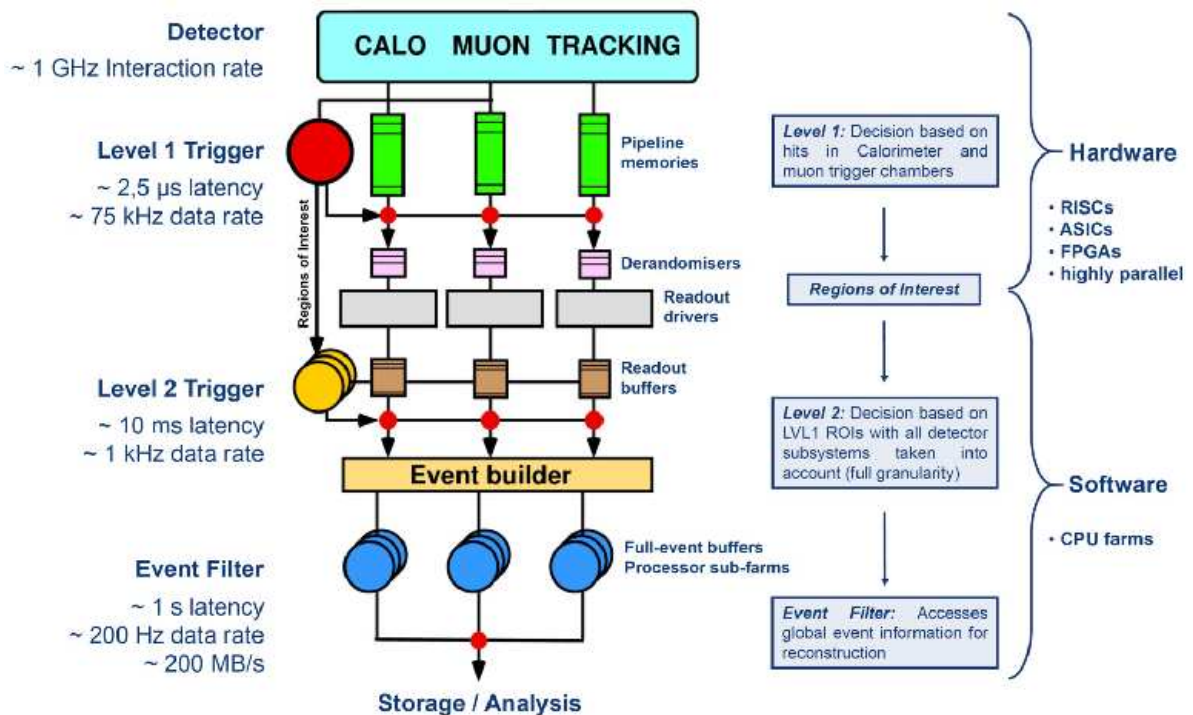


FIG. 40 – Schéma récapitulatif du système de déclenchement de l'expérience ATLAS [154].

la décision est prise dans un délai d'environ 40 ms.

La prise de décision finale est effectuée par le niveau 3 ou EF (*event filter*). Celui-ci est également basé sur un programme informatique. La reconstruction finale des objets non calibrés (non limitée aux RoI) est faite, et des critères d'identification pour les différents types de particules proches des critères finals sont appliqués. Le détail de ces critères dans le cas des photons sera donné dans la partie III. Si toutes les conditions sont satisfaites, l'ensemble du détecteur est enregistré. La fréquence finale d'enregistrement est typiquement de quelques centaines de hertz, correspondant à un débit de sortie de 200 Mo/s. Le temps de prise de décision de l'EF est typiquement d'une seconde.

Il n'y a pas eu d'amélioration significative du système de déclenchement pour la seconde prise de données. La vitesse du L2+EF a toutefois pu être multipliée par trois grâce à diverses optimisations. Il est aussi envisagé d'augmenter la fréquence de lecture des sommes analogiques du trigger du calorimètre de 40 à 80 MHz au cours du *run 2*, ce qui permettrait une meilleure identification du croisement de faisceau pour un signal donné, particulièrement dans le cas de saturation des seuils [155].

5.7.2 La distribution des données

Les différents *menus* sont regroupés en flots de données ou *stream*, qui sont divisés en plusieurs catégories en fonction de leur but (calibration, mesure de physique) et des objets qu'ils contiennent (électrons/photons, muons, jets, ...). Les acquisitions sont également découpées en fonction du temps, l'unité la plus fine étant le bloc de luminosité (*lumiblock*).

Les données sont alors prêtes pour la phase dite de *processing* [156]. Celle-ci est effectuée sur des ordinateurs du CERN appelés *tier-0*, et consiste à faire une première copie intégrale des données brutes ainsi que la première étape de reconstruction hors-ligne. Les données brutes et la première reconstruction sont ensuite distribuées à treize autres centres de calcul appelés *tiers-1*, répartis partout dans le monde. Ceux-ci traitent à nouveau les données et sont responsables de leur stockage et de leur accessibilité par les membres d'ATLAS. Ils les distribuent à leur tour à des centres de calculs secondaires appelés *tiers-2*, qui constituent des zones de stockage supplémentaires.

Les données brutes passant l'EF sont écrites au format binaire et comptent environ 1.6 Mo par événement. Elles sont ensuite converties au format ESD (*event summary data*), qui est lisible dans le logiciel dédié d'ATLAS, ATHENA. Les ESD présentent typiquement 500 ko par événement. Les données sont à nouveau converties, et passent du format ESD au format AOD (*analysis object data*), qui a une organisation orientée objet. Les AOD permettent à nouveau de réduire la taille à environ 100 ko par événement.

En parallèle des données véritables, des échantillons simulés sont produits et distribués au format AOD directement. Ils suivent ensuite une chaîne de production identique aux données. Ces échantillons incluent la génération de processus physiques particuliers présents ou non dans le modèle standard à l'aide de générateurs Monte Carlo, ainsi qu'une simulation complète du détecteur basée sur un code GEANT4 [157, 158]. Les échantillons simulés jouent un rôle très important lors de l'analyse des données, comme nous le verrons dans les parties III et IV. Plus de détails sont disponibles dans la référence [159].

Finalement, des dérivations de données spécifiques et adaptées au besoin de chaque groupe sont produites. Elles sont réduites au maximum afin d'optimiser les temps d'analyse et lisibles avec le logiciel ROOT [160] dans leur totalité. Pour le *run 1*, le format de données en bout de chaîne s'appelle DPD (*derived physics data*). Pour le *run 2*, des améliorations importantes ont été apportées dans la chaîne de traitement et les formats DPD et AOD ont été fusionnés dans un nouveau format unique, les xAOD [161]. Les dérivations distribuées aux différents groupes sont produites dans un format similaire appelé DxAOD.

Pendant mes années de thèse coïncidant avec la préparation et le début du *run 2*, j'ai été responsable de la mise en place et de la maintenance du code de dérivation des données pour le groupe de physique étudiant les photons dans le cadre du modèle standard (STDM2). En particulier, j'ai été amené à interagir avec les experts des *menus* du système de déclenchement du *run 2* et avec les différents groupes d'analyse de physique et de performance intéressés par STDM2 (photon MS, groupe de performance e/γ) afin de définir une dérivation de données adaptée aux besoins de chacun tout en gardant une taille raisonnable.

Dans la section 6, la physique à l'avant et les détecteurs des protons diffractés à petits angles au LHC sont introduits. Les détecteurs de protons à l'avant sont indispensables pour réaliser la mesure proposée dans la partie V. Par ailleurs, un rapide état de l'art est dressé concernant la détection du temps de vol, qui est une des mesures essentielles des détecteurs de protons à l'avant en cas de collisions à haute luminosité. Ce dernier point constitue une bonne introduction à la partie VI, qui abordera en détail la tâche technique effectuée durant les années de thèse, reliée aux tests de la carte de lecture SAMPIC pour la détection de temps avec une résolution de quelques picosecondes.

6 Les détecteurs de protons à l'avant au LHC et la mesure de temps de vol

6.1 La physique à l'avant

La physique à l'avant est le domaine de la physique des particules s'intéressant aux processus émettant des particules à très grande pseudo-rapacité. Les expériences du LHC incluent dans leur ensemble un programme riche de physique à l'avant. Sur l'initiative du comité du LHC (LHCC), un groupe de travail regroupant des théoriciens et des membres des différentes expériences du LHC a été constitué en 2014 (*LHC forward working group*) afin de rédiger un rapport public (*yellow report*) sur les opportunités apportées par le LHC dans le domaine et de coordonner leurs activités [132].²⁰

Pendant mes années de thèse, j'ai assumé avec les théoriciens Lucian Harland-Lang et Valery Khoze la responsabilité de la constitution et de l'édition finale du chapitre 5 de ce rapport, dédié aux productions centrales exclusives. J'y ai également contribué personnellement grâce à l'étude phénoménologique des couplages à quatre photons dans le canal avec protons intacts à l'avant, introduite rapidement dans la section 2.4 et qui sera abordée en détail dans la partie V. Les références associées sont [56] et [57]. Une autre contribution personnelle proposant l'étude de certains processus au LHC pour déterminer la structure du pomeron en terme de quarks et de gluons est présente dans le chapitre 4 du même rapport (processus de diffraction dure inclusive), extraite de mon travail de master [162] et basée sur la référence [163].

La section 6.2 donne un aperçu rapide des détecteurs de protons à l'avant capable de prendre des données à haute luminosité au LHC, installés ou en projet. Les lecteurs intéressés peuvent consulter le *yellow report* mentionné dans le paragraphe précédent et plus particulièrement le chapitre 9, consacré aux détecteurs, pour plus de détails. De tels détecteurs sont nécessaires en plus d'un détecteur central de type ATLAS ou CMS afin de réaliser la mesure proposée dans la partie V.

La section 6.3 introduit la mesure de temps de vol, particulièrement importante pour la détection des protons à l'avant à haute luminosité. Cette problématique est étroitement reliée à ma tâche technique effectuée dans ATLAS pendant la thèse, requise pour figurer sur la liste des auteurs des papiers de collaboration. Elle a consisté à développer des outils d'analyse pour tester la carte de lecture de temps à la picoseconde SAMPIC [164–166], conçue au service d'électronique de l'Irfu (Irfu/SEDI) en collaboration avec le service d'instrumentation du laboratoire de l'accélérateur linéaire d'Orsay (LAL/SERDI). Le compte-rendu détaillé des travaux et des résultats obtenus avec SAMPIC sera donné dans la partie VI.

²⁰Le rapport final est disponible en téléchargement à l'adresse suivante : http://www-d0.fnal.gov/Run2Physics/qcd/loi_atlas/fpwg-yellow_report_final.pdf

6.2 Les détecteurs de protons à l'avant

6.2.1 État des lieux

Au LHC, certaines interactions de haute ou basse énergie, comme par exemple la diffusion élastique [167], l'échange inélastique de pomeron [163], ou encore les interactions à deux photons (voir partie V), laissent un ou deux des protons intacts dans l'état final. Les protons sont alors diffractés à petits angles (ou de façon équivalente, très haute pseudo-rapacité), et restent dans le tube du faisceau après interaction. Dans ce cas, les faisceaux en sortie du détecteur central contiennent en général des protons n'ayant pas interagi, des débris de collision divers, ainsi que les protons intacts diffractés. Les différentes particules passent ensuite dans plusieurs aimants destinés à re-collimater les paquets et à éliminer les débris de collision. Les aimants agissent comme un spectromètre pour les protons diffractés ayant perdu une partie de leur énergie en les écartant légèrement du faisceau, ce qui leur permet de pouvoir être détectés par la suite dans des stations dédiées très proches du tube faisceau et situées à plusieurs centaines de mètres de leur point d'interaction.

Des détecteurs de protons à l'avant à environ deux cents mètres de part et d'autre des points d'interaction d'ATLAS et CMS sont opérationnels depuis les premières données du LHC. Il s'agit des détecteurs de la collaboration TOTEM [128] qui sont situés à proximité de CMS, et des détecteurs d'ALFA [168], partie intégrante de la collaboration ATLAS (voir section 5.6.2). Néanmoins, ces détecteurs, conçus principalement pour la mesure des sections efficaces élastiques et inélastiques totales proton-proton [136, 167, 169, 170], ne sont pas adaptés aux hautes luminosités instantanées et donc prennent des données uniquement lors de *runs* dédiés à plus basse luminosité.

Les projets AFP (*ATLAS forward proton*) [171] et CT-PPS (*CMS-TOTEM precision proton spectrometer*) [73] visent à installer des détecteurs de protons à l'avant pouvant fonctionner pendant les prises de données nominales du LHC à partir de la seconde prise de données (voir section 4.4). Les deux projets sont approuvés pour leur première phase, consistant à déployer les détecteurs au plus tard début 2017 et à prendre des données pendant les *runs* dédiés à basse luminosité, afin de démontrer la faisabilité d'une opération à plus haute luminosité. CT-PPS a pris de l'avance et devrait être en mesure de prendre des données à toutes les luminosités d'ici l'été 2016. Un schéma simplifié des détecteurs d'AFP est représenté dans la figure 41. Une configuration similaire est prévue de l'autre côté du point d'interaction d'ATLAS. Le système retenu par CT-PPS diffère légèrement mais suit la même logique et présente des caractéristiques largement similaires.

AFP inclut deux stations situées à environ 200 mètres de part et d'autre du point d'interaction d'ATLAS. Chaque station est placée à une distance du faisceau correspondant à entre dix et vingt fois sa largeur, ce qui représente un espacement de un à trois millimètres seulement. Deux détecteurs à pixels utilisant une technologie similaire à celle du détecteur interne d'ATLAS (voir section 5.3.1) assurent la mesure de la trajectoire des protons et un détecteur de temps de vol mesure leur temps d'arrivée. Dans un premier temps, la mesure du temps sera assurée par un détecteur Tcherenkov utilisant des barres de quartz. Le signal sera amplifié par une galette à microcanaux (MCP, *micro-channel plate*) couplée à un photomultiplicateur. Il sera ensuite lu par une électronique rapide classique de lecture de temps (TDC, *time-to-digital converter*). Cependant, il est

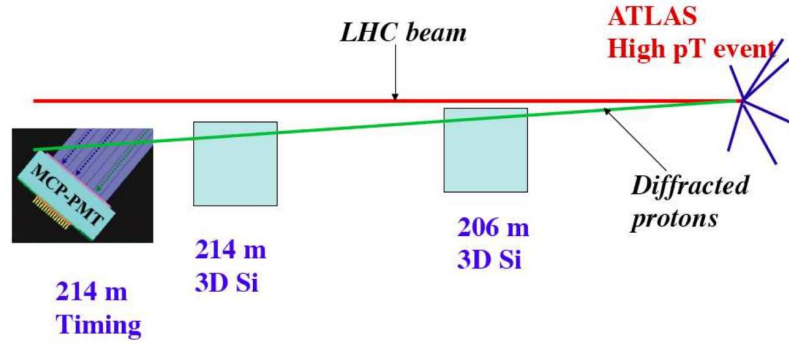


FIG. 41 – Représentation simplifiée des détecteurs d'AFP. Une configuration similaire est prévue de l'autre côté du point d'interaction d'ATLAS [172].

envisagé de passer à des détecteurs rapides à base de diamant [173] lus par la carte SAMPIC pour les prises de données à plus haute luminosité, qui comportent un très bon potentiel en terme de résolution et une pixellisation facilitée (voir partie VI). CT-PPS prévoit d'utiliser cette technologie dès la première phase du projet, installée sur les détecteurs de TOTEM depuis l'hiver 2015. Des détecteurs à silicium ultra-rapide [174] pourraient également être testés au cours de la première phase.

6.2.2 Mesures, résolution et acceptance

La mesure de la trajectoire des protons par les couches de pixels permet de reconstruire l'impulsion perdue par le proton, souvent exprimée en fraction et notée ξ . Une excellente précision allant de 5 à 10 GeV est attendue sur cette mesure grâce à la très bonne qualité des aimants du LHC (voir figure 42, gauche). La trajectoire des protons permet également de mesurer leur petite impulsion transverse avec une précision d'environ 100 MeV (voir figure 42, droite). D'autre part, la mesure du temps de vol est importante pour rejeter le bruit de fond venant des débris entourant le faisceau (halo). Elle est également fondamentale pour distinguer les protons de signal des protons venant de l'empilement lors d'une opération à haute luminosité, à condition d'avoir une résolution en temps suffisante de l'ordre de 10 à 15 ps (voir section 6.3) [172].

Pour des collisions proton-proton nominales au LHC à haute luminosité, l'acceptance de chaque station en terme de fraction d'impulsion manquante ξ est limitée et donnée approximativement par l'équation 31 :

$$0.015 < \xi < 0.15. \quad (31)$$

Les fractions d'impulsion trop basses sont indétectables car les protons diffractés restent trop proches du faisceau, tandis que les fractions d'impulsion trop hautes sont écartées par les collimateurs magnétiques. En revanche, l'acceptance en impulsion transverse p_T n'est pas limitante dans cette configuration. D'une façon générale, les acceptances en (ξ, p_T) dépendent fortement des réglages du LHC, ce qui permet d'étudier des processus différents en fonction des *runs* (*runs* nominaux, *runs* dédiés) et des détecteurs considérés (ALFA, TOTEM, AFP, CT-PPS). Par exemple, des configurations alternatives d'injection du faisceau mises en place lors de *runs* dédiés permettent d'aug-

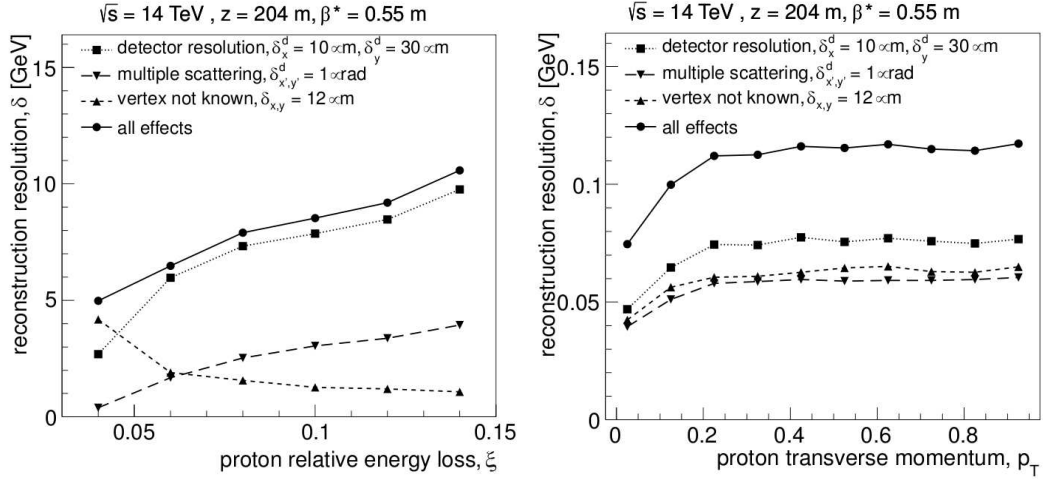


FIG. 42 – Résolutions attendues sur la mesure d'énergie perdue (gauche) et d'impulsion transverse (droite) des protons à l'avant reconstruits dans les détecteurs d'AFP, dans le cas de collisions proton-proton nominale à 14 TeV au LHC. Ces résultats ont été obtenus à l'aide d'une simulation complète des détecteurs [171].

menter l'acceptance en ξ au prix d'une réduction très importante de luminosité intégrée (voir section 20.2). Sauf mention contraire, la configuration nominale sera supposée dans la suite.

Dans le cas d'une interaction laissant les deux protons intacts dans l'état final, on s'attend à une détection en coïncidence dans les deux détecteurs d'AFP (ou de CT-PPS), c'est à dire une détection à droite et une détection à gauche du point d'interaction principal pour un même croisement de faisceaux. Il est alors possible de reconstruire la masse et la rapidité manquante du système di-proton grâce aux formules suivantes :

$$m_{pp}^{\text{miss}} = \sqrt{\xi_1 \xi_2 s} \quad ; \quad y_{pp}^{\text{miss}} = \frac{1}{2} \ln \frac{\xi_1}{\xi_2}, \quad (32)$$

où m_{pp}^{miss} et y_{pp}^{miss} sont la masse et la rapidité manquante, respectivement. $\sqrt{s}$ est l'énergie de collisions proton-proton dans le centre de masse et les $\xi_{1,2}$ sont les mesures des fractions d'impulsion manquantes des protons. L'équation 31 entraîne cependant automatiquement une acceptation limitée en masse manquante, représentée dans la figure 43. Ainsi, uniquement les masses manquantes comprises entre 200 GeV et 2 TeV peuvent être détectées pour des collisions proton-proton nominales à haute luminosité et $\sqrt{s} = 14$ TeV au LHC. Cependant, cette mesure présente une excellente résolution de l'ordre de quelques pourcents, ce qui est comparable voir meilleur qu'une mesure typique réalisée avec le détecteur central, qui dépend largement de l'état final considéré (jets, photons, $W^\pm$, ...).

La détection des protons à l'avant permet ainsi de réaliser de nouvelles mesures au LHC, en particulier pour contraindre les modèles de nouvelle physique. En effet, la baisse de section efficace engendrée par la demande de deux protons intacts dans l'acceptance des détecteurs est dans certains cas largement compensée par un bien meilleur contrôle du bruit de fond, grâce aux contraintes imposées par les protons sur la cinématique (voir partie V). Or, le bruit de fond est bien souvent un des facteurs principaux limitant la

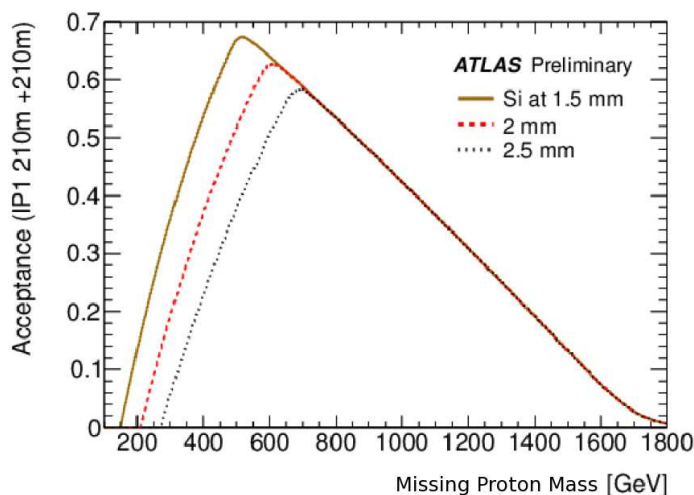


FIG. 43 – Acceptance en terme de masse manquante des détecteurs d’AFP lors d’une détection de paires de proton intacts en coïncidence. Trois espacements possibles entre le faisceau et les détecteurs à pixels sont considérés [175].

sensibilité.

Dans la section suivante, les problématiques principales liées à la mesure de temps dans le cadre des détecteurs à protons à l’avant au LHC sont introduites. Celles-ci seront utiles à la bonne compréhension de la partie VI sur les tests de la carte SAMPIC, effectuées dans le cadre de la tâche technique pour figurer sur la liste d’auteur qualifié d’ATLAS.

6.3 La mesure de temps de vol

6.3.1 Systèmes de mesures

La mesure du temps sert dans des domaines très divers : dates calendaires, épreuves sportives, systèmes de guidage, imageries médicales, ... , et également en physique des hautes énergies. Il est important de réaliser qu’un temps se mesure toujours à partir d’une référence ; par exemple pour revenir sur les cas précédents : phase de la Lune ou du Soleil, coup de feu de départ, émission de l’onde sonore (pour un sonar), début d’acquisition de l’image. Au LHC, le temps d’arrivée d’une particule se mesure le plus souvent par rapport à l’horloge de l’accélérateur lui-même, qui cadence les croisements des faisceaux.

Une différence majeure entre les différentes applications utilisant des mesures de temps vient de la précision attendue. S’il est parfaitement inutile de connaître l’heure précise à la centième de milliseconde près dans la vie de tous les jours, cela peut devenir crucial pour analyser les collisions du LHC, qui croise des faisceaux de protons quarante millions de fois par seconde. Ainsi, les chambres à muons du spectromètre incluses dans le système de déclenchement d’ATLAS (voir section 5.6.1) intègrent une mesure de temps de vol à quelques nanosecondes près afin d’associer les particules incidentes avec le croisement de faisceaux correct, tout comme les “tours” de déclenchement du calorimètre (*trigger towers*).

Il existe plusieurs techniques pour mesurer le temps de passage d'une particule à partir d'un signal analogique de détecteur. Le système le plus courant consiste à mesurer la charge et le temps par des chaînes séparées [176]. Un discriminateur basé sur l'évolution de la charge permet de détecter le passage d'une particule et déclenche l'acquisition du temps de passage. Un TDC (*time-to-digital converter*) et un ADC (*analog-to-digital converter*) assurent alors respectivement la reconstruction du temps et des charges collectées. Le schéma général d'un tel système est représenté dans la figure 44. Bien que présentant de bonnes performances générales, ces systèmes ont quelques défauts :

- ◇ La mesure est souvent limitée par le pas en temps fixe des lignes à retard intégrées dans le TDC et par la précision du discriminateur (une unique valeur de temps est sauvegardée).
- ◇ Il est nécessaire de configurer le discriminateur au préalable de l'acquisition, ce qui entraîne un manque d'adaptabilité vis-à-vis de la variabilité ou de l'évolution des signaux au cours de la prise de données (ex : *time walk*, voir partie VI).
- ◇ La détermination du temps de passage de deux particules trop proches est faussée en général (signal à double pic).
- ◇ Les composants requis pour atteindre de très bonnes précisions (< 10 ns) sont difficiles à intégrer et gourmands en énergie. Ils deviennent très chers pour des résolutions en-dessous de 100 ps.

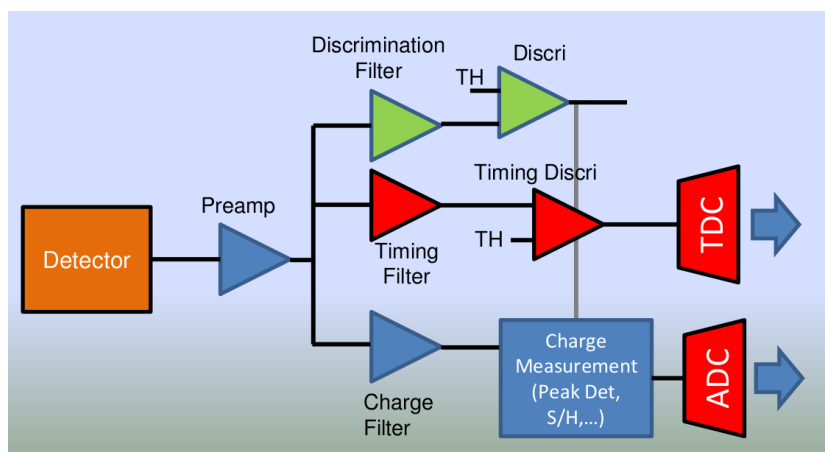


FIG. 44 – Système analogique standard pour mesurer le temps de passage d'une particule. Deux chaînes indépendantes mesurent respectivement la charge et le temps. L'acquisition est commandée par un discriminateur configuré au préalable et basé sur la variation de charge [177].

Une technique alternative comme celle adoptée dans la carte de lecture SAMPIC (voir partie VI) [178] consiste à échantillonner constamment et précisément le signal du détecteur dans une mémoire analogique rapide, comme illustré dans la figure 45. En cas d'activation du système de déclenchement, chaque point présent dans la mémoire analogique est gelé, numérisé, puis sauvegardé. Les données d'un compteur de temps sont également numérisées et sauvegardées. Le temps de passage exact de la particule

peut être alors reconstruit hors-ligne à l'aide d'algorithmes dédiés (interpolations, ...), alors que dans le cas du système précédent celui-ci était retourné directement par le TDC (et donc plus dépendant du pas en temps utilisé) sur la base des critères implémentés a priori dans le discriminateur.

Ce fonctionnement accorde une meilleure robustesse vis-à-vis des conditions expérimentales et la possibilité d'un traitement hors-ligne plus poussé. Par ailleurs, la mesure n'est plus limitée par la précision du discriminateur qui sert uniquement à la définition d'une région d'intérêt, contrairement au système précédent. Il permet également de pouvoir mesurer l'intégrale de la charge pour une région d'intérêt donnée. Cependant, l'échantillonnage très précis et ultra-rapide nécessaire pour atteindre des résolutions en-dessous de la centaine de picosecondes nécessite des composants très performants ainsi que des efforts particuliers de conception. La fabrication de cartes électroniques capables d'échantillonner un signal à une fréquence de plusieurs centaines de millions d'échantillons par seconde (MSPS, *mega-sample per second*) pour un prix raisonnable a été rendue possible entre autres grâce aux énormes progrès réalisés sur les ADCs et les mémoires analogiques depuis une vingtaine d'années. Des précisions supplémentaires sur l'architecture de SAMPIC sont données dans la partie VI.

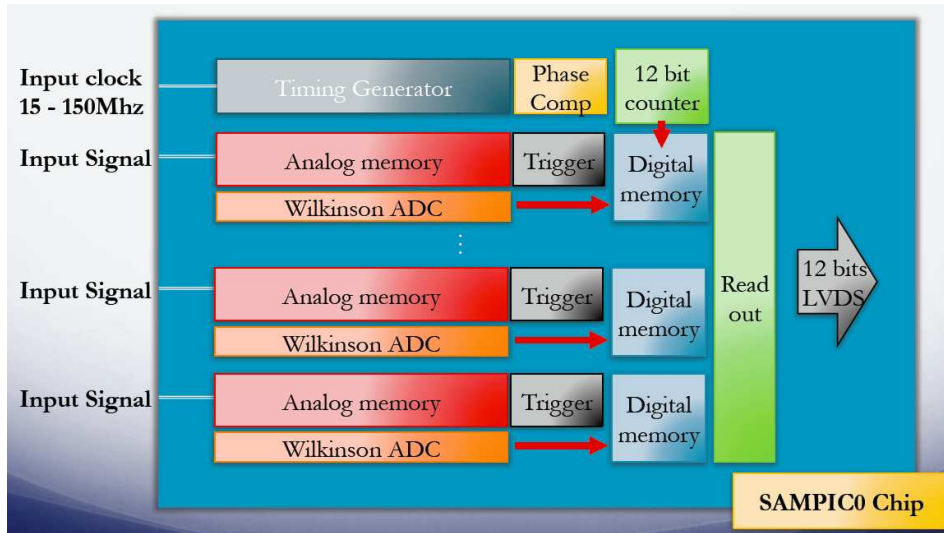


FIG. 45 – Système adopté par la carte de lecture SAMPIC pour mesurer le temps de passage d'une particule, basé sur un échantillonnage rapide de la charge [172].

6.3.2 Le temps de vol des protons à l'avant

La mesure du temps de vol des protons à l'avant par AFP et CT-PPS servira tout d'abord à rejeter le bruit de fond formé par tous les objets contenus dans le halo du faisceau n'arrivant pas en coïncidence avec une collision de haute énergie dans le détecteur central. En effet, la proximité des détecteurs à l'avant par rapport au faisceau associée aux très hautes énergies des protons laisse penser que le bruit de fond induit par le faisceau dominera si aucune détermination du temps de passage des particules n'est effectuée. Celui-ci est pour le moment difficile à estimer précisément par manque de données à 13 TeV, bien que des analyses soient en cours. Les différents retours d'expérience venant de la collaboration TOTEM et de ATLAS/ALFA acquis pendant la

première prise de données indiquent qu'une résolution en temps de 100 picosecondes sur le temps de passage sera suffisante pour obtenir un rapport signal sur bruit acceptable pendant les premières prises de données à 13 TeV (*runs* dédiés à basse luminosité). Ces premières données permettront également de mener des études dédiées sur le bruit de fond des détecteurs.

Cependant, le LHC en fonctionnement nominal produira des collisions avec un fort taux d'empilement ($\langle \mu \rangle \simeq 50$). Les interactions d'empilement ayant une forte probabilité d'émettre des protons intacts dans l'acceptance des détecteurs (voir figure 46), le bruit de fond associé risque de devenir très important. Celui-ci est illustré dans la figure 47 : les étoiles représentent les différents vertex d'interaction ayant lieu durant un même croisement de faisceaux. Dans l'événement du haut, deux particules à hautes impulsions transverses sont émises dans le détecteur central et deux protons intacts dans les détecteurs à l'avant par une unique interaction proton-proton, ce qui constitue un événement de signal. En revanche, dans l'événement du bas, deux particules sont émises dans le détecteur central sans proton intact, et des protons intacts issus d'interaction d'empilement sont détectés par les stations à l'avant, ce qui le rend a priori indistinguable d'un événement de signal.

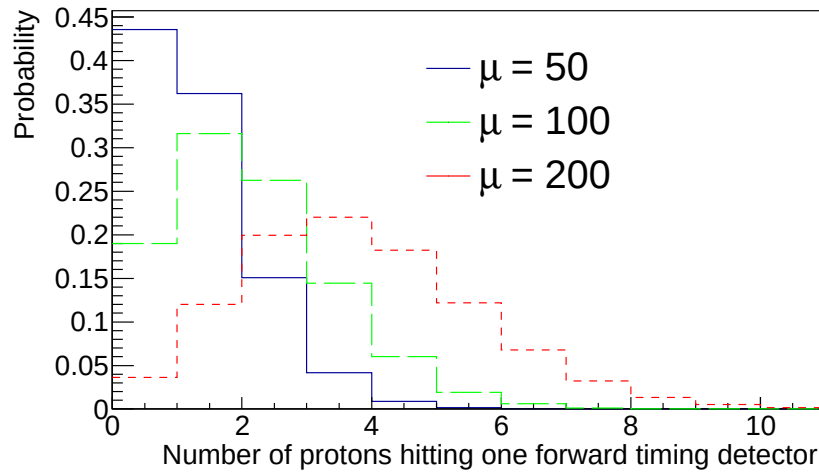


FIG. 46 – Probabilité de détection de protons intacts d'empilement par les détecteurs à l'avant pour différents scénarios d'empilement pour un croisement de faisceau donné [172].

Bien qu'il soit possible de définir des critères cinématiques ou encore d'utiliser le système de détection des traces à l'avant afin de rejeter le bruit de fond créé par l'empilement de façon conséquente (voir partie V), cela n'est pas toujours suffisant. Une solution complémentaire consiste alors à mesurer très précisément la différence entre le temps d'arrivée des protons détectés à gauche et à droite du point d'interaction, ce qui permet de reconstruire leur point d'émission (vertex) avec une incertitude minimale à partir des protons uniquement. Il est ensuite possible de vérifier la compatibilité entre la position du vertex déterminé à partir des protons et celui reconstruit à partir du détecteur central. Les différentes interactions d'empilement étant très proches spatialement (de l'ordre du millimètre), une résolution de l'ordre de dix à quinze picosecondes

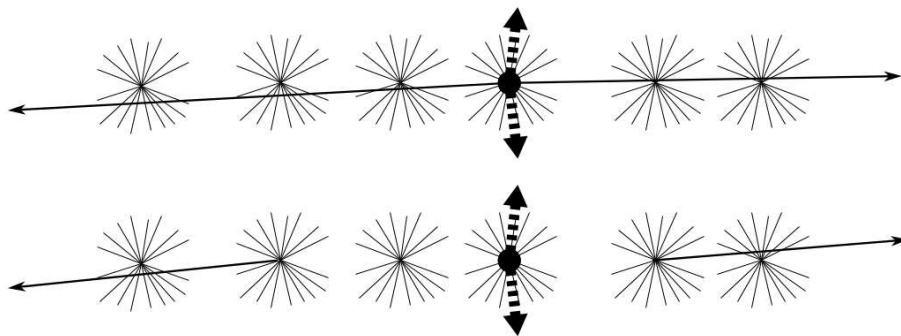


FIG. 47 – Interaction laissant les deux protons intacts (haut) et processus de bruit de fond formé par deux interactions d’empilement laissant des protons intacts et un processus dur sans protons intacts (bas) [179].

est nécessaire sur le temps de vol afin d’obtenir un pouvoir discriminant significatif sur l’emplacement des différents points d’interaction pour un événement donné. Ainsi, il a été montré qu’une résolution en temps de 10 picosecondes mènerait à une réduction du bruit de fond lié à l’empilement par un facteur quarante [172]. Il est également important de souligner qu’à haut taux d’empilement une segmentation adéquate des détecteurs (pixels + temps de vol) devient primordiale afin de minimiser les inefficacités de détections, comme mentionné dans [162, 180]. Atteindre une telle résolution de mesure dans un environnement complexe comme celui du LHC est sans précédent et pose de nombreux défis à résoudre. Comme nous le verrons dans la partie VI, plusieurs technologies de détecteurs et de méthodes de lecture semblent cependant pouvoir remplir ce critère, parmi lesquels figure la carte de lecture SAMPIC.

Au cours de cette seconde partie, la machine du LHC ainsi que ses principaux détecteurs ont commencé par être introduits (voir section 4). Des détails supplémentaires sur le détecteur ATLAS ainsi que sur les détecteurs de protons à l’avant ont été apportés dans les sections 5 et 6, respectivement. Ces sections permettent d’introduire l’ensemble des dispositifs expérimentaux employés dans les parties IV et V, qui forment les travaux principaux de thèse. Finalement, quelques explications sur la mesure de temps ont été données en section 6.3, préparant le compte-rendu sur les tests de la carte SAMPIC de la partie VI.

Dans la partie III, la reconstruction, calibration, identification et critères d’isolation des photons utilisés pour les données 2012 à 8 TeV du détecteur ATLAS sont abordés plus en détail. Les données 2012 sont celles utilisées pour l’analyse de données principale, qui sera présentée dans la partie IV. Les changements principaux apportés pour le traitement des nouvelles données du *run 2* à 13 TeV sont également évoqués.

Troisième partie

Le traitement des photons dans les données à 8 TeV de l'expérience ATLAS

Sommaire

7	Algorithmes de reconstruction	100
7.1	Reconstruction de l'énergie des dépôts électromagnétiques	100
7.2	Reconstruction des électrons et des photons	101
8	Calibration de l'énergie des électrons et des photons	104
8.1	Présentation générale de la procédure	104
8.2	Calibration Monte Carlo	105
8.3	Corrections effectuées dans les données	107
8.3.1	Inter-calibration des couches du calorimètre	107
8.3.2	Autres corrections	109
8.4	Calibration in-situ et validation finale	110
9	Critères d'identification	114
9.1	Définitions	114
9.2	Efficacités et taux de rejet	116
9.3	Mesures de l'efficacité d'identification dans les données	116
9.4	La sélection LOOSE'	118
10	Étude et correction de l'énergie d'isolation transverse	121
10.1	Présentation générale	121
10.2	L'énergie d'isolation transverse dans le détecteur interne	123
10.3	L'énergie d'isolation transverse dans le calorimètre	124
10.3.1	Définition	124
10.3.2	Corrections dues à l'empilement et à l'événement sous-jacent	126
10.3.3	Corrections des fuites à partir de la simulation	128
10.3.4	Comparaison entre données et simulation après corrections .	129
10.4	Corrections supplémentaires de l'isolation calorimétrique basées sur les données	132
10.4.1	Origines des différences observées	132
10.4.2	Méthodologie	134
10.4.3	Résultats	137
10.4.4	Interpolation des paramètres et corrections finales	144
11	Évolutions pour la seconde prise de données	148
11.1	État des lieux pour les premières données à 13 TeV	148
11.2	Quelques réflexions personnelles	149

7 Algorithmes de reconstruction

7.1 Reconstruction de l'énergie des dépôts électromagnétiques

Comme nous l'avons déjà vu dans la section 5 de la partie II, les mesures principales concernant les photons et les électrons dans le détecteur ATLAS s'effectuent au niveau du calorimètre électromagnétique à argon liquide (voir section 5.4) [147]. Tout au long de la partie III, nous nous limiterons aux photons présents dans les données à $\sqrt{s} = 8$ TeV inclus dans la région $|\eta^\gamma| < 2.37$, qui est couverte à la fois par les calorimètres et le détecteur interne (voir section 5.3). Le traitement dans les données à $\sqrt{s} = 13$ TeV est largement similaire mais présente quelques différences qui seront discutées dans la section 11.

À leur arrivée dans le calorimètre électromagnétique, les photons et les électrons développent des gerbes électromagnétiques en interagissant avec les couches d'absorbeurs en plomb. Celles-ci ionisent l'argon liquide, situé entre les couches d'absorbeurs, et donnent lieu à des courants de dérive constitués d'électrons. Ces courants sont collectés dans chaque cellule par des électrodes pour former un signal électrique proportionnel à l'énergie déposée dans le volume actif du détecteur. Le signal est alors envoyé vers l'électronique de lecture à l'aide de câbles, où il est amplifié une première fois par un pré-amplificateur de charge (*charge sensitive amplifier* ou CSA).

Afin de permettre un fonctionnement optimal sur une grande plage en amplitude tout en évitant les saturations, le signal est filtré et amplifié selon trois gains linéaires différents (faible, moyen, haut). Chaque signal ainsi obtenu est alors échantillonné en cinq points. Cela permet au système de pouvoir s'adapter à des niveaux de bruits très variables, principalement dus à l'empilement hors-temps (voir section 4) et au bruit résiduel de l'électronique. En cas de déclenchement au L1 (voir section 5.7), le choix final de niveau de gain est effectué en fonction de l'amplitude maximum obtenue pour le signal amplifié avec le gain moyen.

L'énergie totale déposée dans une cellule du calorimètre électromagnétique peut s'écrire sous la forme suivante :

$$E_{cell}[\text{MeV}] = F_{\mu A \rightarrow \text{MeV}} \times F_{DAC \rightarrow \mu A} \times \frac{1}{\frac{M_{phys}}{M_{cali}}} \times G \times \sum_{j=1}^5 a_j (s_j - p), \quad (33)$$

où les s_j sont les cinq échantillons du signal amplifié avec le gain choisi, p est le piédestal de l'électronique, mesuré pour chaque gain lors de prises de données dédiées pour la calibration, et les a_j sont des poids optimisés pour l'extraction du signal prenant en compte les composantes dues aux effets d'électronique et d'empilement. G est le gain spécifique de la cellule, mesuré au préalable en injectant un signal de calibration connu. Il est corrigé du facteur $\frac{1}{\frac{M_{phys}}{M_{cali}}}$, qui permet de tenir compte des différences entre les signaux

de physique et de calibration. Finalement, $F_{DAC \rightarrow \mu A}$ est le facteur de conversion du signal numérique en courant de dérive, déterminé à partir de l'électronique, et $F_{\mu A \rightarrow \text{MeV}}$ est le facteur de conversion du courant de dérive en énergie déposée dans la cellule, déterminé lors de tests faisceaux.

Dans la section 7.2, la définition d'un candidat électron et photon reconstruit dans le calorimètre est détaillée.

7.2 Reconstruction des électrons et des photons

Une fois l'énergie des cellules du calorimètre reconstruite pour un événement donné, il est nécessaire de définir les candidats photons et électrons proprement dits. Pour cela, les cellules des trois couches du calorimètre sont regroupées en "tours" de $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ et parcourues par un algorithme à fenêtre glissante (*sliding window*), basé sur des fenêtres de taille fixe 3×5 (soit $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$). Les énergies des cellules à l'intérieur de chaque tour sont sommées et les fenêtres d'énergie maximale contenant une énergie totale supérieure à 2.5 GeV sont appelées *clusters* électromagnétiques.

Un *cluster* électromagnétique associé à au moins une bonne trace reconstruite dans le détecteur interne pointant vers un vertex dans la direction du faisceau est classé dans un premier temps comme candidat électron. Dans le cas où aucune trace n'est associée au *cluster*, celui-ci est classé comme candidat photon non converti. Pour être associée à un *cluster*, la trace doit être contenue dans une fenêtre en $\Delta\eta \times \Delta\phi = 0.05 \times 0.10$ autour du barycentre des dépôts d'énergies contenues dans le *cluster* et son impulsion doit correspondre à au moins 10% de l'énergie du *cluster*.

Cependant, il existe également une fraction importante de photons qui se convertissent en paires électron-positron en amont du calorimètre par interaction avec le matériel mort du détecteur interne (voir la figure 35 de la section 5.3), de l'ordre de 15% dans le tonneau et 30% dans les bouchons. À ce stade, la grande majorité de ces photons est classée comme candidat électron ($\simeq 95\%$) à cause de la présence de traces associées. Néanmoins, leur classement est très dépendant du lieu de conversion. Par exemple, les conversions tardives s'effectuant dans le TRT ont moins de chance de mener à un *cluster* associé à une trace, tandis que celles se produisant à l'entrée du détecteur interne donnent souvent lieu à de bonnes traces reconstruites.

Un algorithme dédié est alors exécuté afin de récupérer les photons convertis parmi les candidats électrons. Celui-ci démarre lors de la première identification des *clusters*, en recherchant dans le détecteur interne la présence d'un vertex secondaire de conversion pour la trace associée. Ceux-ci peuvent être définis par deux traces reconstruites de signes opposés (*double track conversion*), mais également par une trace unique lorsque la conversion est très asymétrique ou si les deux traces sont très proches (*single track conversion*). La distribution radiale des vertex de conversion dans les premières données à 7 TeV d'ATLAS est présentée dans la figure 48 et comparée à la simulation, où un bon accord général est observé.

Les candidats électrons pour lesquels la trace est associée à un vertex de conversion sont alors classés comme photons convertis à une ou deux traces, selon la multiplicité du vertex de conversion. Par ailleurs, les candidats ayant une trace non associée à un vertex de conversion et telle que $p_T < 2$ GeV ou $E/p > 10$ sont reclassés comme photons non convertis. Ce critère de qualité supplémentaire sur les traces permet de récupérer environ 7% des photons non convertis totaux, identifiés comme candidats électrons à cause d'artefacts dans le détecteur interne.

La fraction de photons convertis et non convertis en fonction du nombre d'interac-

tions d'empilement dans les données à 8 TeV de 2012 est montrée dans la figure 49. Elle est stable à 1% près, avec une légère migration des photons convertis à deux traces vers les photons convertis à une trace qui augmente avec l'empilement. À l'issue de la procédure, l'efficacité de reconstruction est estimée à 99.83% pour les photons non convertis avec les simulations Monte Carlo [181]. Elle est environ 6% plus faible pour les photons convertis, principalement à cause de leur ressemblance plus forte avec les électrons (94.33%).

Les *clusters* sont ensuite agrandis en fonction de leur classification. En effet, les particules chargées étant déviées dans le plan transverse au faisceau par le champ magnétique du solénoïde lors de leur passage dans le détecteur interne (voir section 5), il est nécessaire d'augmenter la taille en ϕ des *clusters* associés à des traces afin de récupérer une plus grande partie de l'énergie des candidats. Ainsi, les candidats électrons et photons convertis sont redéfinis comme étant des *clusters* de 3×7 dans le tonneau, tandis que les *clusters* de photons non convertis restent inchangés à 3×5 . De plus, tous les *clusters* sont agrandis dans les bouchons en η à 5×5 , afin de prendre en compte les largeurs plus importantes des gerbes (matériau mort) tout en minimisant les contributions dues au bruit et à l'empilement.

Ainsi, dans tous les cas, les candidats photons et électrons sont reconstruits dans des *clusters* de taille fixe. Cela permet entre autres d'éviter des biais de reconstruction dus au bruit du calorimètre, qui inclut des composantes statistiques mais également systématiques (empilement, non-uniformités, ...). Une discussion non-exhaustive sur les avantages et inconvénients d'utiliser de *clusters* de taille fixe est proposée dans la section 11.2. Ce choix de reconstruction donne lieu à des corrections importantes d'énergie au cours de la calibration ($\simeq 20\%$), objet de la section 8.

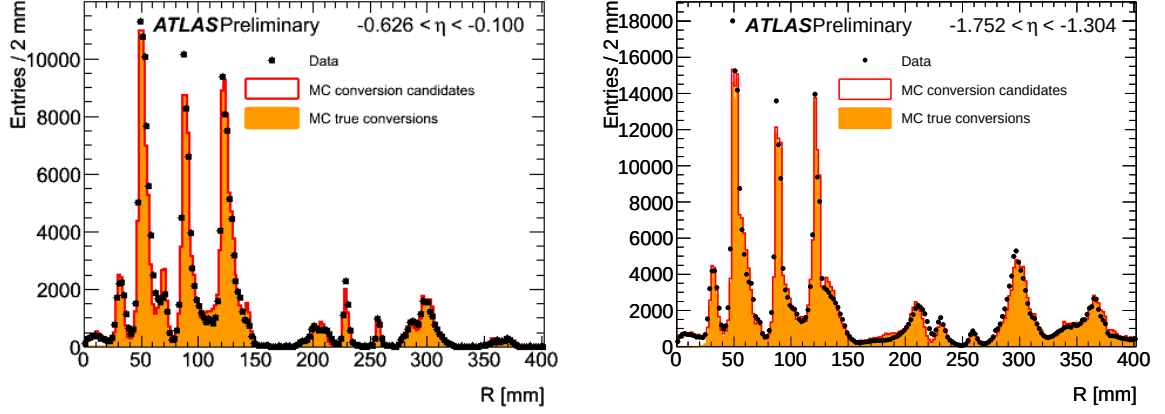


FIG. 48 – Distribution radiale des vertex de conversion reconstruits dans le tonneau (gauche) et la partie bouchon couverte par le TRT (droite). Les trois couches de pixels (50.5 mm, 85.5 mm et 122.5 mm) et les deux premières couches de SCT (300 mm, 373 mm) sont clairement visibles [182].

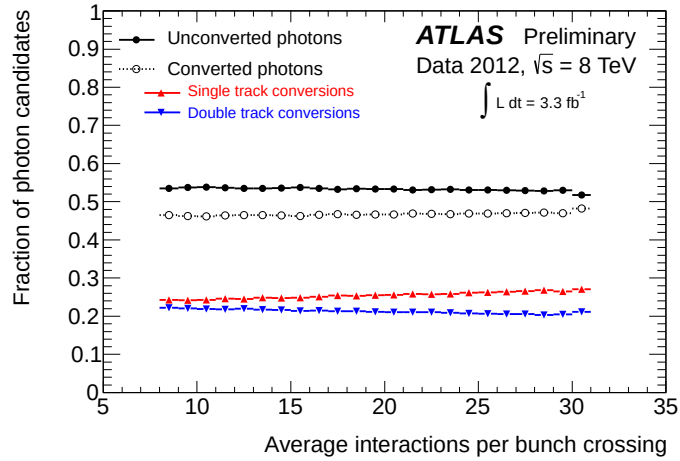


FIG. 49 – Décomposition des candidats photons reconstruits selon leur statut de conversion (non converti, converti à une trace et converti à deux traces) en fonction du nombre moyen d'interactions par croisement de faisceaux [183].

8 Calibration de l'énergie des électrons et des photons

8.1 Présentation générale de la procédure

Les différentes étapes de la calibration en énergie des *clusters* électromagnétiques de taille fixe définis dans la section précédente sont résumées dans la figure 50.

La première étape consiste à entraîner un algorithme d'apprentissage artificiel multivarié (MVA, pour *multivariate algorithm*) à corriger l'énergie des *clusters* électromagnétiques en se basant sur des échantillons simulés (module 1). Au préalable, la distribution de matière située devant le calorimètre est mesurée précisément dans les données afin d'avoir une description adéquate des interactions rayonnement/matière dans la simulation. Ces parties sont décrites succinctement dans la section 8.2.

Diverses corrections de la réponse du calorimètre dans les données, incluant l'inter-calibration entre les différentes couches du calorimètre (module 2) et des corrections d'uniformité de réponse diverses reliées à certaines inhomogénéités du détecteur (module 4), sont présentées dans la section 8.3. La détermination des constantes d'inter-calibration est nécessaire à la validation de la distribution de matière implémentée dans la simulation et doit donc être appliquée aux données avant la calibration MC (module 3), tandis que les corrections d'uniformité sont appliquées a posteriori de celle-ci.

Finalement, la réponse en énergie obtenue pour les électrons et les photons est corrigée lors d'une dernière étape concernant à la fois les données et la simulation, consistant à comparer un échantillon d'électrons autour de la masse du boson Z^0 dans les deux cas de figure. Cette procédure est appelée calibration *in-situ* (module 5) et s'effectue après calibration MC et corrections d'uniformité. La calibration finale est validée pour les électrons en utilisant les résonances $J/\psi \rightarrow ee$ et pour les photons grâce aux émissions radiatives des désintégrations leptoniques du boson Z^0 , $Z^0 \rightarrow ll\gamma$ (module 6).²¹ Cela permet entre autres de vérifier la linéarité des corrections en sondant d'autres domaines en énergie. Cette partie est traitée dans la section 8.4.

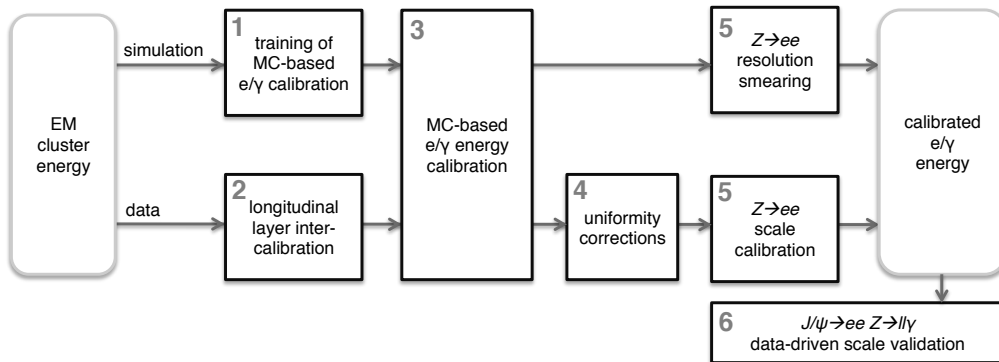


FIG. 50 – Vue schématique de la procédure utilisée pour calibrer la réponse en énergie des photons et électrons dans l'expérience ATLAS [147].

²¹Pour toutes les sélections effectuées dans les données, des critères d'identification et d'isolation sont appliqués aux candidats photons et électrons afin de maximiser la pureté des échantillons. Ceux-ci seront détaillés pour les photons dans les sections 9 et 10.

8.2 Calibration Monte Carlo

Afin d'obtenir l'énergie correcte des électrons et des photons traversant le calorimètre, il est nécessaire de corriger de plusieurs effets les énergies des *clusters* électromagnétiques (voir section 7). Une liste non exhaustive de ces effets est donnée ci-dessous :

- ◇ **Pertes d'énergie** dans la matière précédant et au-delà du calorimètre.
- ◇ **Dépôts d'énergie dans les cellules avoisinant** les *clusters*, qui ont une taille fixe en η et ϕ .
- ◇ **Variation de réponse** en fonction des points d'impact dans les cellules.

Pour cela, un algorithme de type MVA est entraîné sur des échantillons simulés d'électrons et de photons (produits seuls) de très grande statistique, couvrant tout le domaine cinématique (1 GeV à 3 TeV). Cette étape correspond au module 1 de la figure 49. L'entraînement est effectué séparément pour les électrons, photons convertis et non convertis, et catégorisé en p_T , η . Il se base sur six observables au niveau détecteur afin de produire la meilleure estimation de l'énergie réelle de la particule : l'énergie totale mesurée dans le calorimètre E_{calo} , le rapport entre E_{calo} et l'énergie mesurée dans le pré-échantillonneur, la profondeur de la gerbe, la pseudo-rapacité du barycentre en énergie du *cluster* dans le système de coordonnées d'ATLAS (η_{cl}) ainsi que la pseudo-rapacité et l'angle azimuthal du *cluster* dans le repère du calorimètre (η , ϕ). η_{cl} est nécessaire afin d'estimer correctement le matériel mort traversé, tandis que η et ϕ permettent de corriger les effets dus à la géométrie du calorimètre.

Pour les photons convertis, le rayon de conversion, le rapport entre l'impulsion transverse des traces de conversion et E_{calo} et le rapport entre la trace reconstruite de plus haut p_T et l'impulsion transverse des traces de conversion sont utilisés comme variables supplémentaires sous certaines conditions.

Les performances de l'algorithme MVA sont évaluées à partir de la simulation et présentées dans la figure 51 en fonction de η et dans plusieurs gammes de p_T , pour les photons convertis et non convertis. La très bonne linéarité de réponse est illustrée par les distributions des valeurs les plus probables (MPV, *most probable value*), plates à la fois en η et p_T . Les effets de résolution restent sous contrôle, avec une variation de 2 à 4% (4 à 8%) pour les photons non convertis (convertis). La résolution est fortement corrélée avec la quantité de matériaux précédant le calorimètre (voir figure 35 de la section 5.3, partie II), qui produit des queues en énergie significatives au niveau des gerbes.

Afin de pouvoir appliquer la calibration MC aux données réelles, diverses corrections sont nécessaires et détaillées dans la section 8.3. Celles-ci permettent entre autres de valider la distribution de matière implémentée dans la simulation et d'uniformiser la réponse réelle du calorimètre.

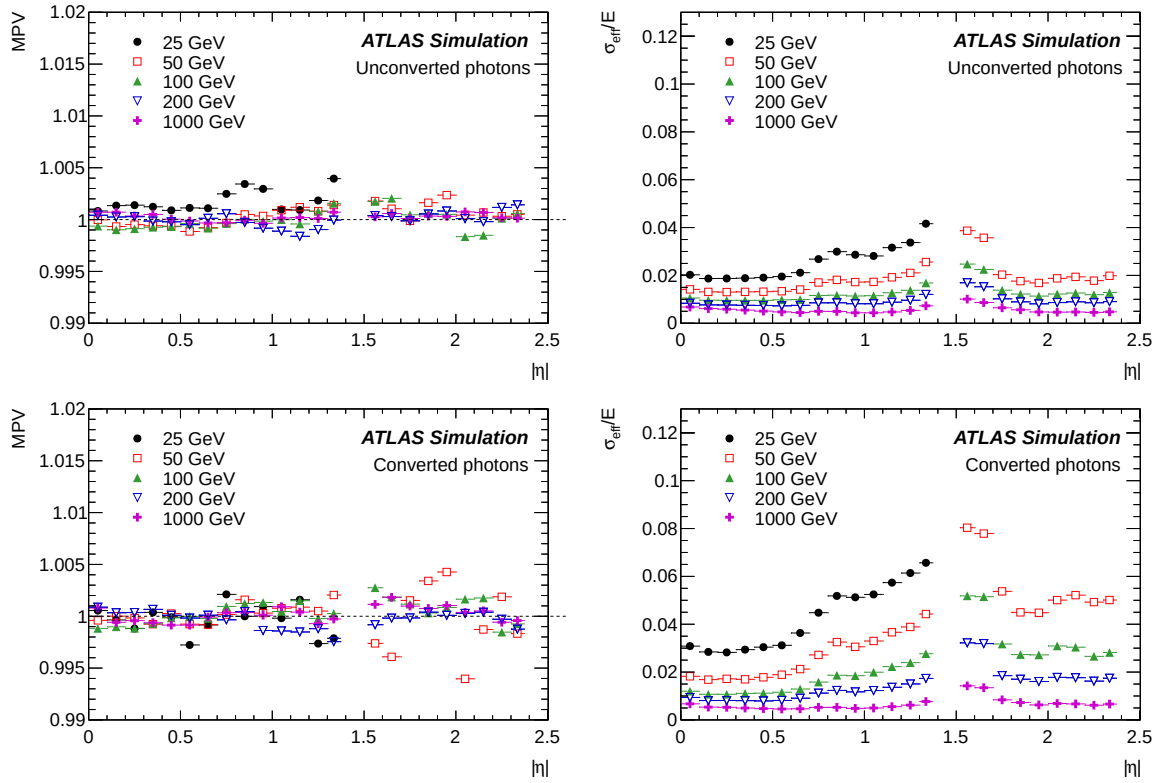


FIG. 51 – Performances de l'algorithme MVA utilisé pour la calibration Monte Carlo des photons non convertis (haut) et convertis (bas), données en terme de valeur la plus probable E/E_{true} (MPV) et de résolution (σ_{eff}/E) en fonction de l'impulsion transverse et de la pseudo-rapidity au niveau reconstruit [147].

8.3 Corrections effectuées dans les données

8.3.1 Inter-calibration des couches du calorimètre

Comme mentionné dans la section 8.2, la réponse en énergie du calorimètre dépend fortement de la distribution de matière en amont. Il est donc indispensable de valider la distribution implémentée dans la simulation au préalable de la calibration MC afin d'assurer un entraînement fiable de l'algorithme MVA. Pour cela, la différence entre le taux d'interaction des électrons, sensibles à l'intégralité du matériel qu'ils traversent, et des photons non convertis, insensibles au matériel traversé jusqu'à leur éventuel point de conversion dans le détecteur interne, est exploitée. En pratique, le rapport entre l'énergie déposée dans la première et la seconde couche du calorimètre ($E_{1/2}$) constitue une bonne observable pour cette étude. En effet, une valeur de $E_{1/2}$ plus élevée dans certains domaines de pseudo-rapacité indique une gerbe démarrant plus tôt et donc une quantité de matière traversée en amont plus importante.

Cependant, la réalisation de cette étude nécessite en premier lieu d'inter-calibrer les différentes couches du calorimètre dans les données. Pour cela, les désintégrations $Z^0 \rightarrow \mu\mu$ sont exploitées, le dépôt d'énergie des muons dans le calorimètre étant quasi-indépendant de la quantité de matière traversée et ne nécessitant pas de calibration particulière. $E_{1/2}$ est estimée à l'aide de deux méthodes différentes, l'une basée sur un ajustement et l'autre sur une simple moyenne tronquée, et comparée entre simulation et données dans plusieurs intervalles en pseudo-rapacité. Plus de détails sur les deux méthodes de mesure sont disponibles dans [147]. La constante d'inter-calibration $\alpha_{1/2}$ est alors définie comme le rapport données sur MC. Cette distribution est montrée dans la figure 52, gauche.

La valeur centrale de la constante est définie comme la moyenne des deux méthodes de mesure, tandis que la différence entre les deux est prise comme incertitude systématique ($< 2\%$). L'incertitude statistique est négligeable. La différence entre les données et la simulation est attribuée principalement aux petites inhomogénéités des cellules du calorimètre en terme de géométrie et à la variation du champ électrique entre les différentes couches et modules.

L'échelle en énergie du pré-échantillonneur E_0 est également inter-calibrée avec les deux premières couches. Cette calibration est rendue indépendante de la description du matériel en amont au premier ordre grâce à la corrélation approximativement linéaire entre E_0 et $E_{1/2}$ vis-à-vis de la quantité de matière, comme le montre la figure 52, droite. L'inter-calibration du pré-échantillonneur est ensuite validée en utilisant les électrons des bosons $W^\pm$ et Z^0 , illustrée dans la figure 53, gauche. La mesure de son échelle d'énergie comporte une incertitude finale de 2 à 3%, variant légèrement en fonction de la pseudo-rapacité.

En revanche, aucune étude dédiée n'a été menée pour inter-calibrer la troisième couche avec le reste du calorimètre. Cela est principalement dû au fait que l'énergie déposée y est négligeable dans le cas des collisions proton-proton à 7 et 8 TeV réalisées lors de la première prise de données. Les changements dans le traitement des photons pour le *run 2* seront détaillés brièvement dans la section 11.

Il est alors possible de valider la distribution du matériel mort implémentée dans la

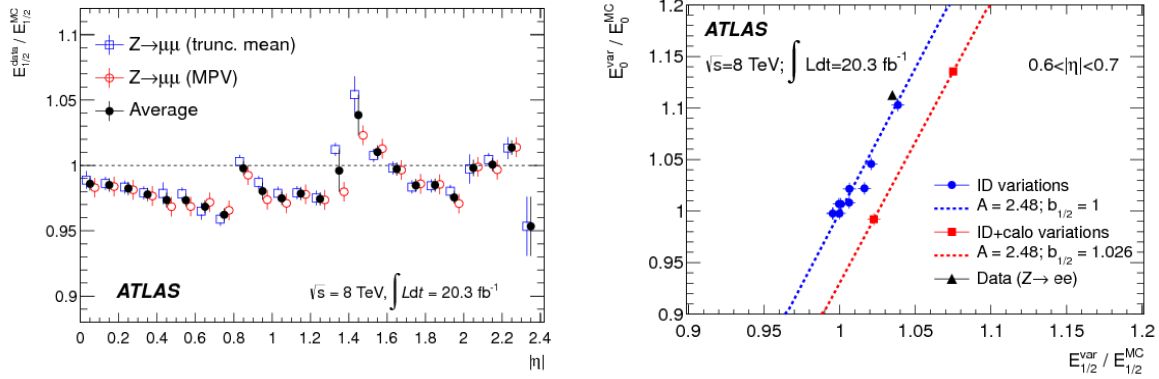


FIG. 52 – (gauche) Comparaison de l'observable $E_{1/2}$ entre données et simulation en fonction de la pseudo-rapacité, estimée à partir des muons des désintégrations $Z^0 \rightarrow \mu\mu$ et deux méthodes de mesure de l'énergie. La première méthode est basée sur l'énergie moyenne des dépôts tandis que la seconde méthode utilise les résultats issus de l'ajustement d'un modèle correspondant au produit de convolution d'une distribution de Landau (particule d'ionisation minimale) et d'une gaussienne (bruit électronique). (droite) Corrélation entre E_0 et $E_{1/2}$ pour différentes variations du matériel mort en amont du calorimètre dans la région $0.6 < |\eta| < 0.7$. Le modèle de corrélation est validé dans les données en utilisant les désintégrations électroniques du Z^0 [147].

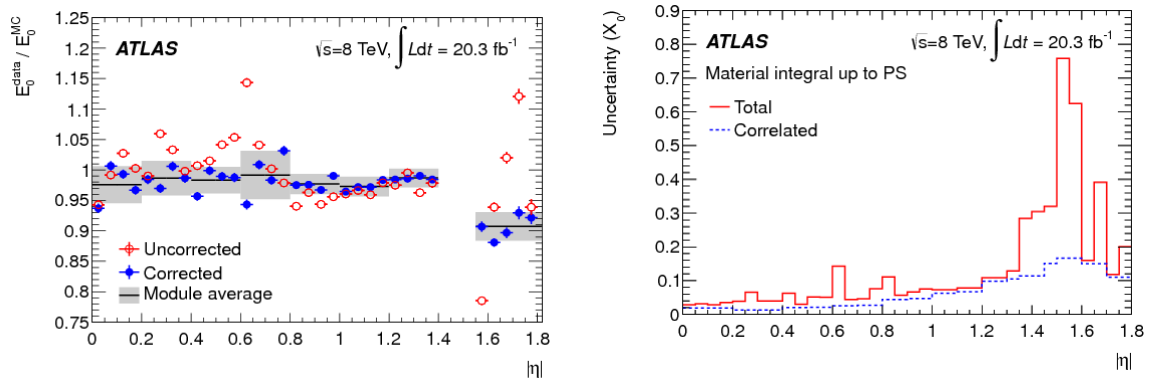


FIG. 53 – (gauche) Comparaison de l'énergie des électrons mesurés dans le pré-échantillonneur entre simulation et données avant et après inter-calibration, en fonction de η . (droite) Incertitude finale sur la distribution du matériel mort jusqu'à l'entrée du pré-échantillonneur [147].

simulation, comme mentionné au début de la section. La distribution finale de matière en amont du calorimètre déterminée lors du *run 1* a déjà été montrée dans la figure 35 de la section 5.3 (partie II) ; son incertitude finale est représentée en terme de longueur de radiation X_0 dans la figure 53, droite. La seule incertitude significative restante est localisée dans la zone de transition entre le tonneau et les bouchons, qui est faiblement instrumentée ($1.37 < \eta < 1.56$). Cette région est généralement exclue des analyses de données pour la physique. En particulier, elle est exclue de l'analyse présentée dans la partie IV.

8.3.2 Autres corrections

Une fois l'algorithme MVA correctement entraîné, la calibration MC est appliquée aux données inter-calibrées. Cependant, il est nécessaire de corriger a posteriori de la calibration MC plusieurs effets non linéaires présents uniquement dans les données afin d'obtenir une calibration satisfaisante pour les électrons et photons passant par le détecteur ATLAS.²² Ces corrections sont spécifiques à certaines régions du détecteur et difficilement reproductibles dans la simulation. Elles incluent :

- ◇ **Corrections de la réponse des couches 1 et 2 du calorimètre** dues à des inhomogénéités dans la haute tension appliquée aux bornes des électrodes dans certaines régions ($\simeq 2\text{-}3\%$ dans moins de 2% des secteurs).
- ◇ **Corrections de la réponse du pré-échantillonneur** dues aux différentes valeurs de haute tension appliquées aux bornes des électrodes dans certaines régions ($\simeq 1\%$).
- ◇ **Corrections des effets de bords en ϕ** dues à la taille des cellules du calorimètre ($\simeq 1\text{-}2\%$).
- ◇ **Corrections de la dépendance de la calibration vis-à-vis du gain** utilisé ($\simeq 1\%$).

La première correction est liée à la présence de courts-circuits dans certains secteurs du calorimètre. Ceux-ci obligent à baisser la haute tension appliquée aux bornes des électrodes dans les régions concernées, se présentant typiquement sous la forme de “tours” de $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$. Ces changements sont pris en compte dès la reconstruction des *clusters*, en utilisant la dépendance attendue de la réponse des cellules en fonction de la valeur de haute tension. Cependant, dans moins de 2% des secteurs du calorimètre, des inhomogénéités de réponse en ϕ sont encore observées après correction. Celles-ci sont corrigées de manière empirique en utilisant la masse bien connue du boson Z^0 et ses désintégrations électroniques dans les données. Un exemple de correction est fourni dans la figure 54, gauche.

²²Ces corrections sont également importantes pour l'inter-calibration des couches du calorimètre présentée dans la section 8.3.1, ce qui engendre nécessairement certaines iterations entre les étapes 1 à 4 du schéma de calibration (voir figure 50).

Au cours de la première prise de données, la haute tension appliquée dans le pré-échantillonneur a également dû être abaissée mais cette fois dans tous les secteurs, afin de réduire le bruit de l'électronique qui croît avec l'augmentation de luminosité. Comme pour la situation précédente, cela est d'abord pris en compte au niveau de la reconstruction des *clusters* et correspond à des effets allant de 8 à 20%. À nouveau, ceux-ci sont vérifiés grâce à la masse bien connue du boson Z^0 et ses désintégrations électroniques. Un biais résiduel de l'ordre de 1% en η est détecté et corrigé empiriquement afin d'uniformiser la réponse du pré-échantillonneur, abaissant le biais moyen à 0.4%. Cette correction est illustrée dans la figure 54, droite.

Notons également une dépendance résiduelle supplémentaire en ϕ de la réponse du calorimètre, visible grâce aux électrons des désintégrations leptoniques des bosons $W^\pm$ et localisée aux transitions entre les différents modules de la partie tonneau. Cet effet de 1 à 2% est interprété comme un éloignement des modules les uns des autres à cause de leur poids et est corrigé empiriquement dans les données.

Finalement, dans certaines régions spécifiques en η , une dépendance de la calibration d'environ 1% vis-à-vis du choix de gain dans les cellules lors de la reconstruction des *clusters* est observée (voir section 7) et corrigée à l'aide de la simulation.

La stabilité de réponse du calorimètre dans le temps et en fonction du nombre moyen d'interactions d'empilement est estimée dans les données à partir des désintégrations électroniques des bosons électrofaibles $W^\pm$, Z^0 après corrections d'uniformité. Elle est illustrée dans la figure 55. Une excellente stabilité de réponse est atteinte après corrections, avec des écarts de moins de 0.1% pour l'ensemble des données à $\sqrt{s} = 8$ TeV.

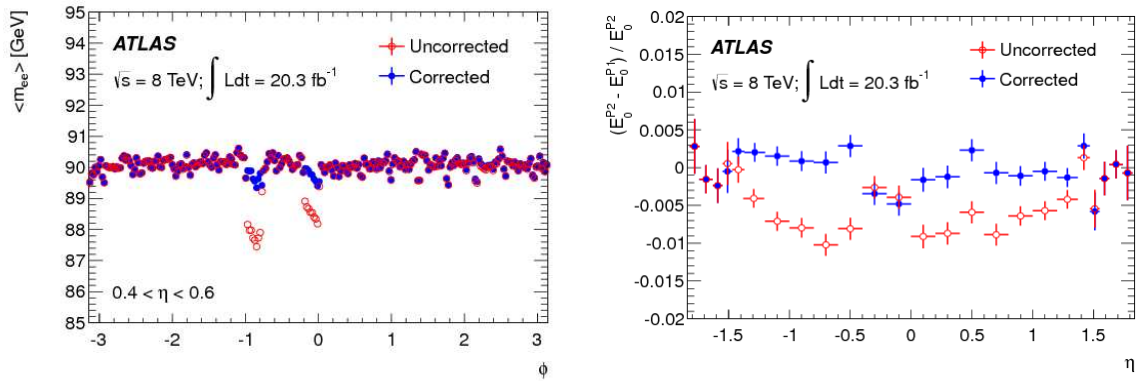


FIG. 54 – Illustration des corrections résiduelles d'inhomogénéité de réponse du calorimètre (gauche) et du pré-échantillonneur (droite) dues aux différentes valeurs de haute tension appliquées aux électrodes selon les secteurs [147].

8.4 Calibration in-situ et validation finale

Étant donné l'excellente stabilité de réponse du calorimètre au cours de la première prise de données (voir section 8.3), l'échelle absolue en énergie est catégorisée en η uniquement et intégrée en ϕ , μ , etc. En particulier, aucune catégorisation n'est effectuée pour les différentes périodes de fonctionnement. Des catégories de largeur $\Delta\eta = 0.2$ sont utilisées dans l'intervalle $-2.47 < \eta < 2.47$.

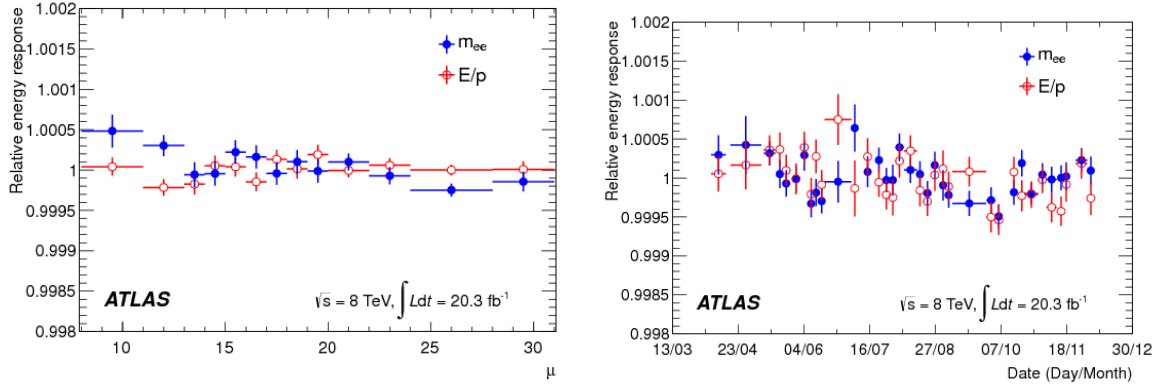


FIG. 55 – Illustration de la stabilité de réponse du calorimètre électromagnétique en fonction du nombre moyen d’interactions d’empilement (gauche) et au cours du temps lors de la prise de données à $\sqrt{s} = 8$ TeV (droite). Les barres d’erreurs représentées sont statistiques uniquement [147].

Au cours de la dernière étape de calibration appelée calibration *in-situ* (module 5), la distribution en masse invariante des paires d’électrons autour de la masse du boson Z^0 est comparée entre données et MC, après calibration et corrections. Les différences résiduelles sont corrigées en abaissant légèrement la résolution des échantillons simulés et en corrigeant légèrement la position du pic en masse dans les données. Ces deux effets sont de l’ordre de 1 à 2%. Les distributions finales sont montrées dans la figure 56. Les différences entre données et échantillons simulés sont sous contrôle et couvertes par les incertitudes issues des différentes étapes de calibration dans tout l’intervalle observé.

La linéarité de la calibration des électrons et des photons est ensuite vérifiée dans différents régimes cinématiques grâce aux désintégrations de $J/\psi \rightarrow ee$ ayant une haute impulsion transverse ($\langle E_T \rangle = 11$ GeV pour les électrons) et $Z \rightarrow ll\gamma$ ($E_T^\gamma < 60$ GeV), respectivement. Les différences observées sont toujours sous contrôle et aucune incertitude supplémentaire n’est introduite. Les incertitudes principales sur l’échelle en énergie des photons non convertis sont résumées dans la figure 57 pour la région $|\eta| < 0.6$. Les photons convertis possèdent des niveaux d’incertitudes similaires. L’erreur est dominée par la détermination de la constante d’inter-calibration de la première et la deuxième couches du calorimètre et de la correction de gain dans la seconde couche. Elle reste néanmoins inférieure à 0.4% dans l’intervalle considéré.

D’autre part, les effets de résolution sont faibles, de l’ordre de 2% à $E_T = 40$ GeV et 1% à plus haute impulsion transverse, à la fois pour les photons et les électrons. Ceux-ci sont résumés dans la figure 58. Leur incertitude est de l’ordre de 10% à bas E_T mais a tendance à augmenter à plus haute impulsion transverse et atteint 25% à $E_T = 200$ GeV. Elle est alors dominée par l’incertitude sur le terme constant de l’équation 28 (voir section 5.4), qui est déterminé après application des corrections d’uniformité (voir section 8.3.2). Des détails complémentaires sur les résultats de calibration des électrons et des photons dans les données à 8 TeV sont disponibles dans la référence [147].

Dans la section 9, les critères d’identification utilisés pour les gerbes de photons sont présentés.

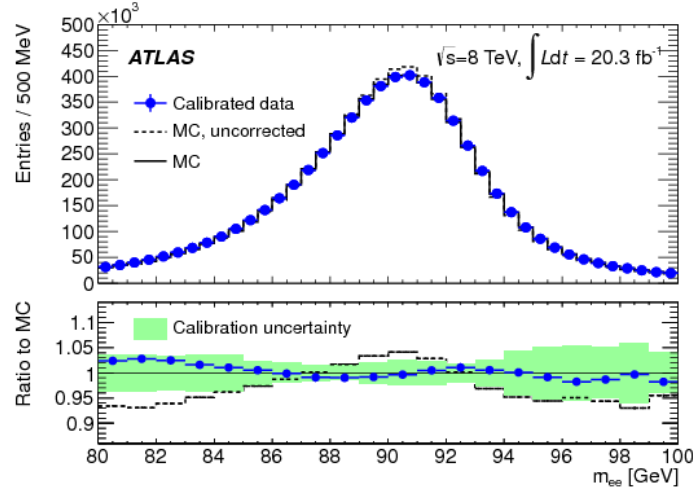


FIG. 56 – Comparaison finale de la distribution en masse invariante des paires d'électrons des désintégrations $Z \rightarrow ee$ entre données et simulations après calibration pour les données 2012. La calibration *in-situ* est appliquée dans les données tandis que son effet sur la résolution des échantillons simulés est représenté dans la figure. Les incertitudes finales de la calibration sont également données [147].

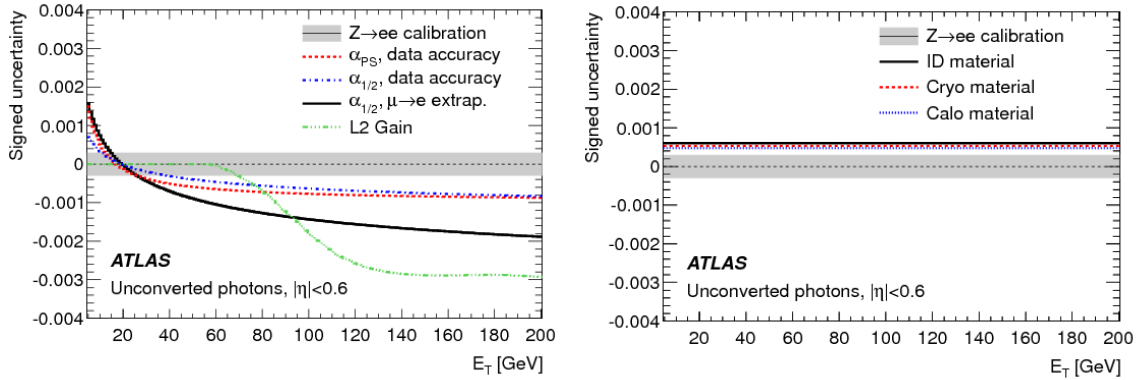


FIG. 57 – Incertitudes principales sur la détermination de l'échelle en énergie des photons non convertis en fonction de leur impulsion transverse : calibration *in-situ* ($Z \rightarrow ee$ calibration), constantes d'inter-calibration (α_{PS} , $\alpha_{1/2}$), correction de gain dans la seconde couche ($L2$ gain) et description du matériel mort dans la simulation (ID , $Cryo$ and $Calo$ material). Les photons convertis présentent des incertitudes dominantes similaires en première approximation [147].

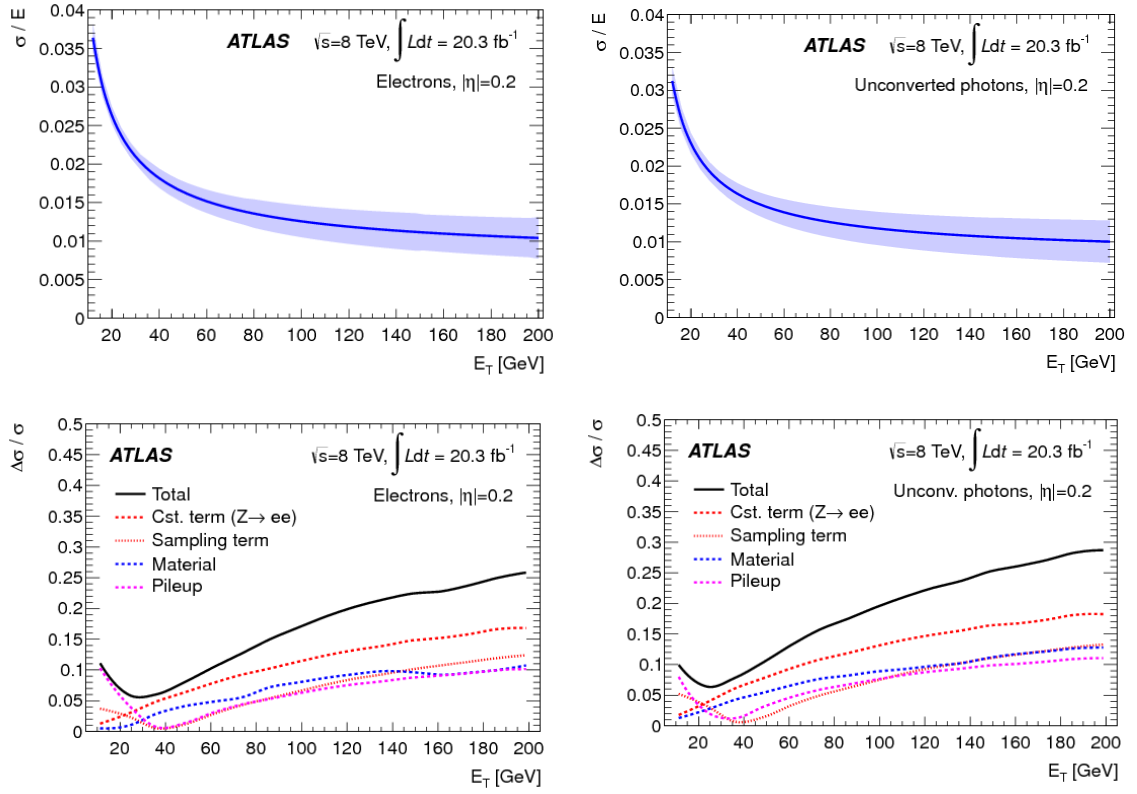


FIG. 58 – Résolution en énergie sur les électrons (gauche) et les photons non convertis (droite) en fonction de leur impulsion transverse. Les incertitudes principales sont mentionnées sur les courbes inférieures : terme constant, terme d'échantillonnage (voir équation 28, section 5.4), description du matériel mort dans la simulation et empilement [147].

9 Critères d'identification

9.1 Définitions

Il existe trois modes principaux de production de photons dans les collisions hadroniques du LHC (voir section 3.5.2). Pour la très grande majorité des analyses utilisant les données du détecteur ATLAS, seuls les photons directement issus de l'interaction à haute énergie (photons directs) et de la fragmentation de partons (photons de fragmentation) sont intéressants à étudier. Ils seront appelés photons immédiats par la suite (dits *prompt* en anglais). Cependant, la troisième catégorie, incluant les photons issus de la désintégration de hadrons, est produite dans des quantités bien plus importantes que les photons immédiats, en particulier pour les impulsions transverses inférieures à 100 GeV. Il est donc nécessaire, une fois les candidats photons reconstruits dans le détecteur et leur énergie finement calibrée (voir section 7 et 8), d'établir des critères supplémentaires dit d'identification afin de rejeter ce bruit de fond très important.

Dans la collaboration ATLAS [184, 185], l'identification des photons immédiats se fait en utilisant des coupures franches sur neuf variables, présentées dans le tableau 2. Parmi celles-ci, huit caractérisent le développement latéral et longitudinal de la gerbe électromagnétique associée déposée dans l'argon liquide, tandis que la neuvième représente la fraction d'énergie collectée dans le calorimètre hadronique. La plupart sont communes aux photons et aux électrons, cependant les électrons utilisent également un certain nombre de critères reliés à la qualité de leur trace (E/p , nombre de dépôts dans le TRT, ...), absents pour les photons. L'identification repose sur le fait que les photons immédiats produisent typiquement des gerbes plus étroites et comportent des fuites réduites dans le calorimètre hadronique par rapport aux photons du bruit de fond. En effet, les hadrons sont le plus souvent émis en jets très collimatés (voir section 3.2), ce qui a pour effet d'étaler et de prolonger les gerbes. De plus, les photons issus de la désintégration de mésons légers, typiquement $\pi^0 \rightarrow \gamma\gamma$, donnent souvent lieu à deux maxima locaux d'énergie dans la première couche très segmentée du calorimètre (voir section 5.4) à cause de la proximité angulaire des deux photons, contrairement aux photons immédiats.

Il est important de mentionner que l'empilement provoque des émissions supplémentaires à basse énergie qui a également tendance à étaler les gerbes, et donc diminue la discrimination entre photons immédiats et bruit de fond. Ainsi, les critères d'identification doivent être réévalués lors d'un changement important des conditions d'empilement. Cela a été le cas lors du *run 1* entre les données à 7 et 8 TeV (voir section 4.3).

Deux sélections d'identification sont définies pour les photons immédiats, appelées LOOSE et TIGHT. La sélection LOOSE utilise la fraction d'énergie collectée dans le calorimètre hadronique ($R_{\text{had}(1)}$) et deux variables définies à partir de la seconde couche du calorimètre LAr uniquement (R_η , R_ϕ), tandis que TIGHT utilise les neuf variables définies dans le tableau 2, et donc à la fois la première et la seconde couche du LAr. La sélection TIGHT est définie de telle sorte que les photons associés forment une sous-catégorie des photons LOOSE. Dans la section 9.2, des détails supplémentaires sur les efficacités et taux de rejet des critères d'identification sont donnés.

Catégorie	Description	Nom	LOOSE	TIGHT
Fuite dans le calorimètre hadronique	Rapport entre E_T mesuré dans la première couche du calorimètre hadronique et les <i>clusters</i> électromagnétiques. (utilisé pour $ \eta < 0.8$ et $ \eta > 1.37$)	R_{had1}	oui	oui
	Rapport entre E_T mesuré dans le calorimètre hadronique et les <i>clusters</i> électromagnétiques. (utilisé pour $0.8 < \eta < 1.37$)	R_{had}	oui	oui
Seconde couche du calorimètre LAr	Rapport entre l'énergie des cellules 3×7 et 7×7 en $\eta \times \phi$.	R_η	oui	oui
	Largeur latérale de la gerbe.	$w_{\eta 2}$	oui	oui
	Rapport entre l'énergie des cellules 3×3 et 3×7 en $\eta \times \phi$.	R_ϕ	non	oui
Première couche du calorimètre LAr	Largeur de la gerbe calculée à partir des 2 bandes avoisinant celle de plus haute énergie.	w_{s3}	non	oui
	Largeur latérale totale de la gerbe.	w_{stot}	non	oui
	Énergie située à l'extérieur des 3 bandes centrales mais à moins de 7 bandes de celles-ci, divisée par l'énergie des 3 bandes centrales.	F_{side}	non	oui
	Différence entre l'énergie associée avec le 2 ^e maximum dans la première couche et l'énergie associée dans la bande d'énergie minimale située entre le premier et le second maximum.	ΔE	non	oui
	Rapport entre la différence d'énergie associée entre le premier et le second maximum et la somme de ces 2 quantités.	E_{ratio}	non	oui

TAB. 2 – Variables discriminantes utilisées pour les identifications LOOSE et TIGHT des photons reconstruits dans le détecteur ATLAS [185].

9.2 Efficacités et taux de rejet

Deux caractéristiques importantes d'un critère d'identification sont l'efficacité, c'est-à-dire la fraction de photons immédiats passant la sélection, et le taux de rejet, soit le nombre de photons de bruit de fond rejetés pour un d'accepté. Le critère LOOSE a été conçu pour avoir une efficacité d'au moins 97% pour les photons avec $E_T > 20$ GeV. Dans les faits, celle-ci se situe à 97% pour $E_T^\gamma = 20$ GeV et est supérieure à 99% pour $E_T^\gamma > 40$ GeV. Le taux de rejet associé se situe autour de 1 000 [181,185], ce qui est le plus souvent insuffisant pour obtenir un rapport signal sur bruit supérieur à 1. La sélection TIGHT permet de rejeter davantage le bruit de fond venant des mésons neutres évoqué dans la section 9.1, grâce notamment à l'utilisation de la première couche du calorimètre, finement segmentée. Afin de maximiser son efficacité, les valeurs des coupures définissant TIGHT sont définies séparément pour les photons convertis et non-convertis. Son efficacité s'élève à 85% pour $E_T^\gamma > 40$ GeV, avec un taux de rejet d'environ 5 000 [181,185], ce qui correspond à un rapport signal sur bruit d'environ 4 pour $E_T^\gamma = 50$ GeV et supérieur à 9 pour $E_T^\gamma > 100$ GeV [186]. La sélection TIGHT, optimisée pour fournir le meilleur rapport signal sur bruit, et celle recommandée dans les analyses.

Les coupures associées aux sélections LOOSE et TIGHT sont définies séparément dans sept catégories en pseudo-rapidité afin de tenir compte de la déformation des gerbes due à la granularité variable du calorimètre et à la variation du matériel situé en amont, permettant ainsi d'optimiser leurs efficacités et taux de rejet. Les valeurs des coupures peuvent varier jusqu'à 25% en fonction des catégories. Les catégories sont définies par les intervalles en pseudo-rapidités suivants : 0-0.6, 0.6-0.8, 0.8-1.15, 1.15-1.37, 1.52-1.81, 1.81-2.01, 2.01-2.37. Les sélections d'identification ne sont pas définies pour l'intervalle 1.37-1.52, celui-ci étant généralement exclu des analyses de physique (voir section 5.4).

La sélection LOOSE est utilisée entre autres pour déclencher l'acquisition d'événements contenant des photons immédiats (voir section 5.7), tandis que la sélection TIGHT sert à la définition des échantillons finals. L'utilisation de LOOSE au niveau du déclenchement permet de maintenir une efficacité de *trigger* très proche de 100%.

La section 9.3 présente un résumé succinct des différentes mesures d'efficacité d'identification des photons mises en œuvre dans les données à 8 TeV.

9.3 Mesures de l'efficacité d'identification dans les données

Il est possible d'obtenir un échantillon très pur de photons immédiats dans les données ($\simeq 99\%$) en utilisant par exemple les désintégrations radiatives du boson Z^0 , *ie.* $Z^0 \rightarrow ll\gamma$ avec le photon émis par l'un des leptons.²³ Les distributions des variables d'identification entre simulation et données peuvent alors être directement comparées, comme illustré dans la figure 59 pour F_{side} et w_{s3} (voir tableau 2). Malgré un aspect similaire, les valeurs moyennes des distributions diffèrent significativement pour des raisons pas encore bien comprises. De nombreuses sources potentielles pouvant expliquer ces différences ont été explorées, sans succès jusqu'à maintenant. La piste privilégiée à l'heure actuelle est un problème de modélisation dans le code GEANT4 utilisé pour la

²³Pour toutes les sélections effectuées dans les données, un critère d'isolation est appliqué aux candidats photons. L'énergie d'isolation sera décrite dans la section 10.

simulation du détecteur (voir section 5.7).

Dans les faits, ces écarts sont corrigés en décalant chacune des distributions simulées pour les faire correspondre aux données (voir figure 59). Ces facteurs correctifs, appelées *fudge factors* par la suite, sont calculés séparément dans des catégories en pseudo-rapidité et impulsion transverse. Ils se basent sur un échantillon TIGHT de photons inclusifs ($\gamma + X$) dans les données afin de couvrir un domaine en pseudo-rapidité et impulsion transverse le plus large possible, tout en maintenant une excellente pureté. Ils correspondent typiquement à un décalage représentant 10% de la RMS des distributions observées dans les données et leur impact sur l'efficacité d'identification varie entre -5% et -10% pour $E_T^\gamma = 10$ GeV, tendant ensuite vers 0 pour $E_T^\gamma > 50$ GeV. Les *fudge factors* ont été évalués sans incertitudes et uniquement pour les électrons et les photons lors de la première prise de données.

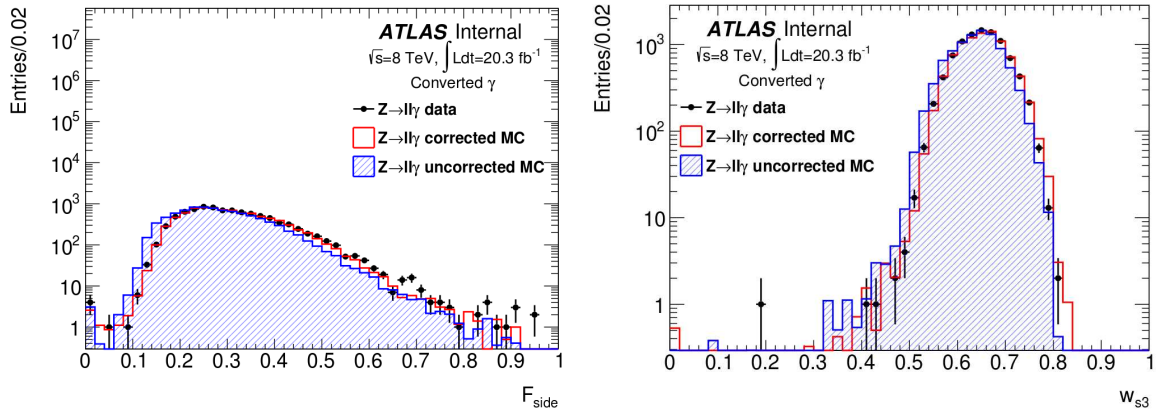


FIG. 59 – Comparaison des variables F_{side} et w_{s3} entre simulation et données à 8 TeV, basée sur les désintégrations radiatives $Z^0 \rightarrow ll\gamma$. Les distributions simulées sont corrigées à l'aide de translations simples indépendantes appelées *fudge factors* [185].

Une fois les *fudge factors* calculés et appliqués à la simulation, trois méthodes indépendantes sont utilisées afin de sélectionner des échantillons de photons immédiats purs dans les données :

- ◇ **Utilisation des désintégrations radiatives du boson Z^0** , *ie.* $Z^0 \rightarrow ll\gamma$ avec le photon émis par l'un des leptons. Elle permettent d'obtenir un échantillon de photons immédiats avec une pureté d'environ 99% dans la gamme $10 \text{ GeV} < E_T^\gamma < 80 \text{ GeV}$, y compris sans application de critères d'identification.
- ◇ **Utilisation des désintégrations électroniques du boson Z^0** et extrapolation des électrons aux photons. La similarité entre les gerbes des photons et des électrons est exploitée et permet de couvrir la gamme $30 \text{ GeV} < E_T^\gamma < 100 \text{ GeV}$ par le biais de l'échantillon très pur de di-électrons présent au voisinage de la masse du boson Z^0 . Un modèle d'extrapolation électron $\rightarrow \gamma$ est cependant nécessaire.
- ◇ **Utilisation de l'énergie d'isolation transverse mesurée dans le détecteur interne** afin de sélectionner un échantillon inclusif de photons ($\gamma + X$), avec ou sans critères d'identification (méthode de la matrice). Le caractère inclusif de l'échantillon permet de couvrir une très large gamme d'impulsions transverses, $10 \text{ GeV} < E_T^\gamma < 1.5 \text{ TeV}$.

La définition de ces échantillons permet de comparer l'efficacité d'identification TIGHT entre simulation et données dans différents domaines cinématiques, catégorisés en $|\eta^\gamma|$ et E_T^γ . De façon analogue à la définition de TIGHT, les comparaisons sont effectuées séparément pour les photons convertis et non-convertis.

Tout d'abord, l'amélioration de l'accord données/MC après application des *fudge factors* est vérifiée à nouveau sur l'ensemble de l'espace des phases grâce aux différentes méthodes. Des facteurs correctifs ainsi que leurs incertitudes sont ensuite calculés afin de tenir compte des différences résiduelles sur l'efficacité d'identification TIGHT entre simulation et données, indépendamment pour chaque méthode. Celles-ci s'élèvent à environ 10% pour $E_T^\gamma = 10$ GeV et ne représentent plus que quelques pourcents pour $E_T^\gamma > 40$ GeV. Les résultats sont en accord entre les trois méthodes dans leurs incertitudes respectives. Finalement, tous les résultats sont combinés afin d'extraire les facteurs correctifs finals, appelés *ID scale factors* dans la suite.

La combinaison des trois méthodes permet de déterminer l'efficacité d'identification TIGHT des photons immédiats à partir des données avec une incertitude très réduite. Les mesures finales sont présentées dans les figures 60 et 61 pour les photons convertis et non convertis des données à 8 TeV, séparément dans quatre catégories en $|\eta^\gamma|$. L'efficacité d'identification TIGHT est d'environ 45-50% (50-60%) pour les photons non-convertis (convertis) à $E_T^\gamma = 10$ GeV et atteint plus de 90% pour $E_T^\gamma > 40$ GeV. L'incertitude associée représente 1.5-2.5% (2-3%) pour les photons non-convertis (convertis) à $E_T^\gamma = 10$ GeV, et baisse à 0.5-1% à plus haute impulsion transverse. Les très légers sauts visibles autour de 30 et 100 GeV dans les bouchons pour les photons non-convertis sont dus à de légères tensions entre les méthodes ($\simeq 1$ -2 sigma). Dans ce cas, la différence est prise comme incertitude systématique additionnelle afin de rester conservateur sur le degré de précision atteint.

Dans la section 9.4, une sélection supplémentaire servant à l'identification des photons de bruit de fond est introduite. Celle-ci est utilisée dans la grande majorité des analyses de données d'ATLAS afin de soustraire le bruit de fond restant dans les échantillons TIGHT. En particulier, elle sera employée pour la définition de régions de contrôle dans l'analyse présentée dans la partie IV.

9.4 La sélection Loose'

Le but de la sélection LOOSE' est de définir un échantillon dominé par les photons de bruit de fond présentant des caractéristiques aussi similaires que possible avec le bruit de fond passant TIGHT. En effet, les candidats photons TIGHT sont le plus souvent encore contaminés de façon importante par les hadrons, avec une pureté d'environ 80% attendue à $E_T^\gamma = 50$ GeV [186]. Comme nous le verrons dans la partie IV, il est nécessaire de procéder à la soustraction de ce bruit de fond résiduel dans le cas d'une mesure de précision. L'échantillon défini par LOOSE' sera alors très utile.

La sélection LOOSE' est similaire au critère TIGHT concernant les coupures effectuées sur les variables de la seconde couche du calorimètre (R_η , $w_{\eta 2}$, R_ϕ) et la fuite d'énergie dans le calorimètre hadronique ($R_{\text{had}(1)}$), afin de préserver l'environnement moyen des gerbes. Cette propriété est primordiale afin d'assurer une distribution d'énergie d'isolation transverse dans le calorimètre similaire entre les photons de bruit de fond

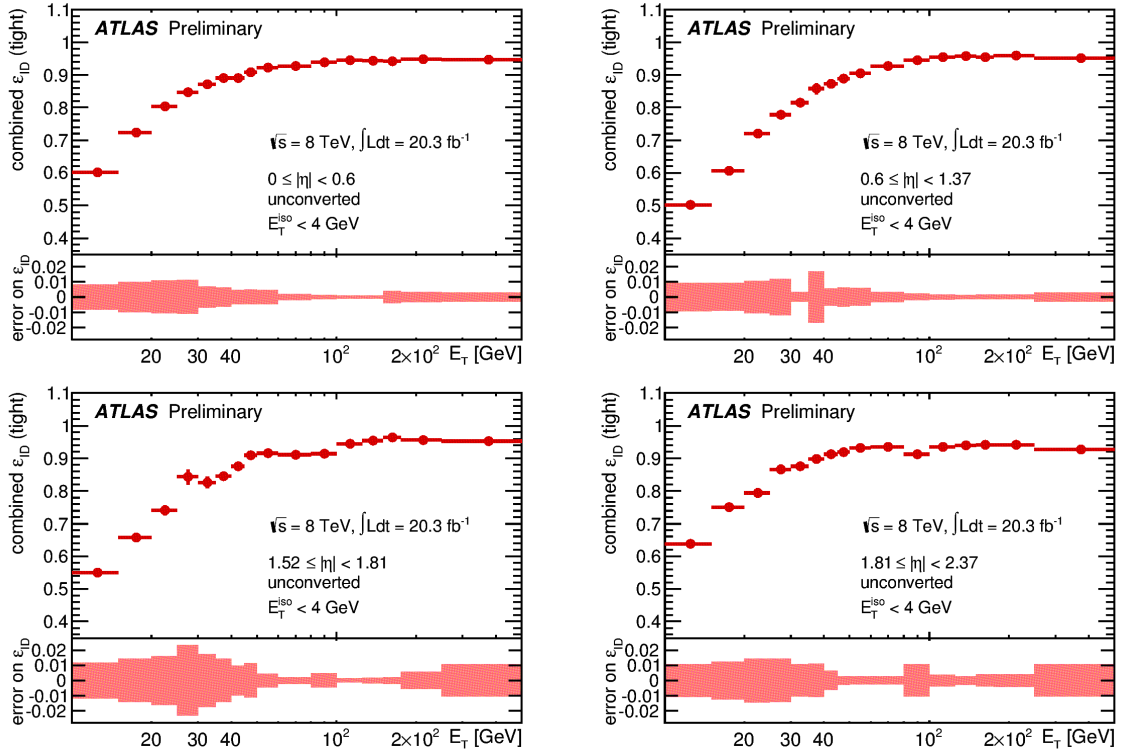


FIG. 60 – Efficacité d'identification TIGHT finale des photons non-convertis issus des données à 8 TeV, après combinaison des trois méthodes de mesures basées sur les données. Les incertitudes finales sont données dans le cadre inférieur. Un critère d'isolation $E_T^{\text{iso}} < 4 \text{ GeV}$ est appliqué aux photons (voir section 10) [187].

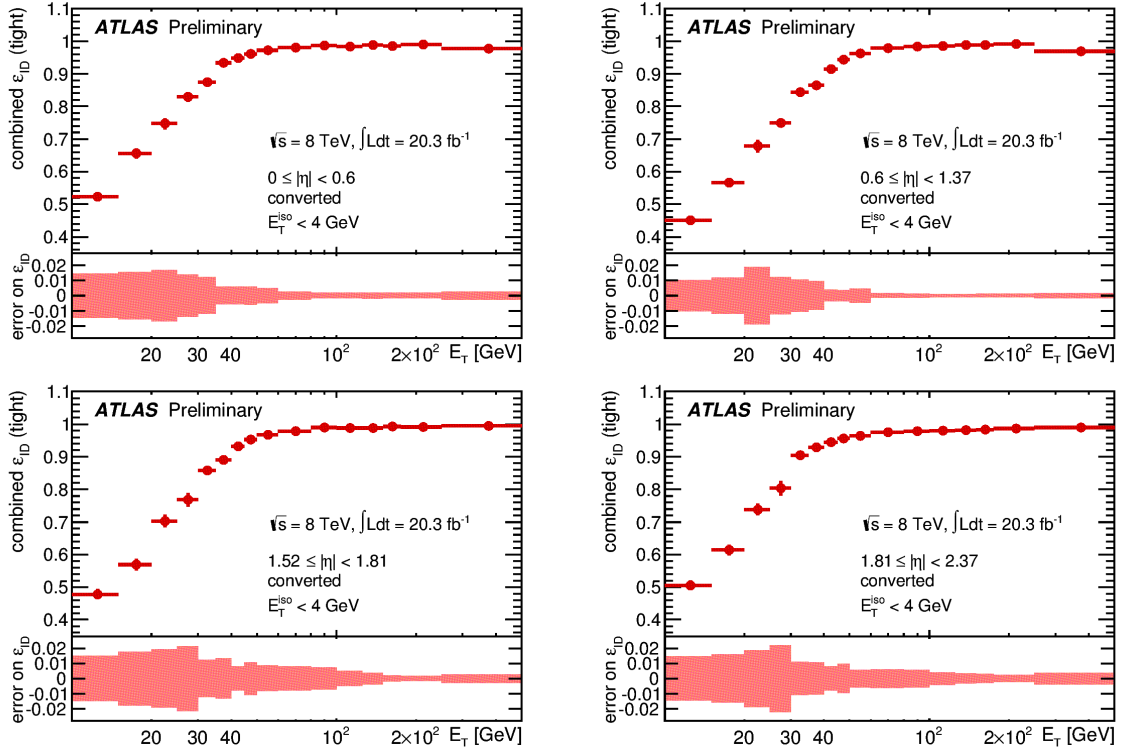


FIG. 61 – Figure similaire à la figure 60 pour les photons convertis [187].

LOOSE' et ceux contaminant de façon résiduelle l'échantillon TIGHT. En effet, la plupart de l'énergie collectée par le système calorimétrique d'ATLAS est déposée à partir de la seconde couche du calorimètre à argon liquide. Des détails supplémentaires sur l'énergie d'isolation transverse des photons seront donnés dans la section 10.

En revanche, la sélection LOOSE' n'applique pas deux à cinq (équivalent alors à la totalité) des coupures sur les variables calculées à partir de la première couche du calorimètre. Les différentes sélections ainsi définies sont notées LOOSE'X en fonction du nombre X de coupures non appliquées. Le détail de ces coupures est donné dans le tableau 3. Le critère nominal utilisé dans les analyses est en général LOOSE'4, qui est un bon compromis entre statistique et préservation de l'environnement moyen de la gerbe. Les autres critères servent à l'évaluation des erreurs systématiques dues au choix arbitraire du nombre de coupures non appliquées.

Nom	Coupures non appliquées par rapport à la sélection TIGHT
TIGHT	-
LOOSE'2	F_{side}, w_{s3}
LOOSE'3	$F_{\text{side}}, w_{s3}, \Delta E$
LOOSE'4	$F_{\text{side}}, w_{s3}, \Delta E, E_{\text{ratio}}$
LOOSE'5	$F_{\text{side}}, w_{s3}, \Delta E, E_{\text{ratio}}, w_{\text{stot}}$

TAB. 3 – Définition des différentes sélections LOOSE' par rapport à la sélection TIGHT.

Finalement, afin d'enrichir davantage l'échantillon en bruit de fond et de définir une région orthogonale à l'échantillon TIGHT, la sélection LOOSE' est appliquée parmi les photons LOOSE qui ne sont pas TIGHT. LOOSE' appliquant les mêmes coupures que TIGHT concernant les fuites dans le calorimètre hadronique et les variables définies à partir de la seconde couche du calorimètre, cela revient simplement à considérer les candidats photons passant les mêmes coupures que TIGHT pour les fuites et la seconde couche et ne passant pas au moins l'une des 5 coupures appliquées sur les variables de la première couche du calorimètre. Cette condition permet en particulier de laisser passer les photons de désintégrations de mesons légers, $\pi^0 \rightarrow \gamma\gamma$ étant la plus courante. Les photons passant LOOSE' et pas TIGHT sont alors dits NON-TIGHT. L'utilité de tels échantillons sera démontrée lors de la soustraction du bruit résiduel hadronique présentée dans la partie IV ainsi que pour les corrections d'isolation de la section 10.4.

Dans la section 10, les principales définitions de l'énergie d'isolation transverse utilisées dans ATLAS pour les photons sont introduites. Un critère d'isolation est souvent appliqué pour définir un échantillon de photons immédiats dans les données, que ce soit dans l'analyse de la partie IV ou pour la calibration et l'identification abordées dans les sections précédentes. Les corrections de fuite pour l'énergie d'isolation transverse des photons basées sur les données, travail personnel effectué au cours des années de thèse, sont présentées dans la section 10.4.2.

10 Étude et correction de l'énergie d'isolation transverse

10.1 Présentation générale

Comme mentionné dans les sections 3.5, 8 et 9, un critère d'isolation est très souvent utilisé au niveau expérimental pour améliorer la pureté des échantillons de photons immédiats définis à partir des données. L'énergie d'isolation transverse mesure, comme son nom l'indique, l'énergie transverse située autour d'un candidat photon, en excluant l'énergie de la particule en question. Appliquer une coupure sur cette quantité constitue un critère très discriminant pour séparer les photons immédiats des photons de bruit de fond, car ces derniers sont majoritairement issus de la désintégration de hadrons émis en jets collimatés et donc donnent lieu en général à une énergie d'isolation plus élevée. Dans le même temps, cela restreint l'espace fiduciel de la mesure en excluant une partie importante des photons de fragmentation. La validité du théorème de factorisation dans le cas de sections efficaces isolées pour les photons immédiats et le caractère *collinear- et infrared-safe* du critère d'isolation basé sur l'énergie transverse dans un cône de taille fixe ont été démontrés dans [114] dans le cas de collisions hadroniques (voir section 3.5).

La région à inclure dans le calcul de l'isolation est souvent délimitée par un cône de taille fixe centré autour de la particule concernée et défini à l'aide des pseudorapidités et angles azimuthaux, $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < X$. Cette formulation permet d'appliquer un critère similaire quelque soit le référentiel des partons rentrant en collision (voir section 5.2). Il a déjà été souligné dans la section 3.5 que les distributions d'isolation issues de calculs théoriques générant la composante de fragmentation (non-perturbative) en se basant sur des algorithmes de cascades de partons, dont les expérimentateurs ont absolument besoin pour déconvoluer leur mesure des effets de détecteurs (*unfolding*), diffèrent significativement de celles estimées à partir de calculs à ordre fixe utilisant les fonctions de fragmentation. Comme nous le verrons dans les sections suivantes, les efficacités des coupures d'isolation définies expérimentalement présentent également souvent des différences significatives avec leurs analogues théoriques servant à la définition de l'espace fiduciel. Il est donc nécessaire d'être particulièrement vigilant lors de la définition des coupures d'isolation expérimentales et fiducielles, afin d'éviter des corrections de déconvolution trop importantes. En effet, celles-ci engendreraient des incertitudes supplémentaires venant de la dépendance de la mesure vis-à-vis de la modélisation de l'isolation par le MC. Plus de détails sur ce point délicat seront donnés dans la partie IV.

Un autre point central donnant lieu à une des incertitudes dominantes dans toutes les analyses traitant des photons dans ATLAS concerne la corrélation entre les critères d'identification et d'isolation. Comme nous le verrons dans la partie IV, il est nécessaire de supposer les deux critères indépendants afin de soustraire le bruit de fond restant dans un échantillon donné. Or, la forme de la gerbe étant nécessairement influencée par son environnement (voir section 9), les deux critères présentent forcément une certaine corrélation. La corrélation associée est largement inconnue et estimée le plus souvent en faisant varier la définition de l'échantillon LOOSE' (voir section 9.4). Plus de détails sur ce second point délicat seront apportés dans la partie IV.

Dans le cas du détecteur ATLAS, il est possible d'estimer l'énergie d'isolation trans-

verse d'un photon à partir du détecteur interne ou du calorimètre. Les études reliées à la calibration et à l'identification (voir section 8 et 9) utilisent uniquement l'énergie d'isolation définie dans le calorimètre afin de rester le plus général possible, tandis que l'analyse du $H \rightarrow \gamma\gamma$ [188] et celle présentée dans la partie IV utilisent à la fois le détecteur interne et le calorimètre. Cela permet d'améliorer la robustesse du critère d'isolation expérimental vis-à-vis de l'empilement (voir figure 62) et d'augmenter son efficacité à haute énergie en relâchant légèrement la valeur de la coupure dans le calorimètre (voir section 10.4 et partie IV), tout en maintenant une haute pureté grâce à la coupure dans le détecteur interne.

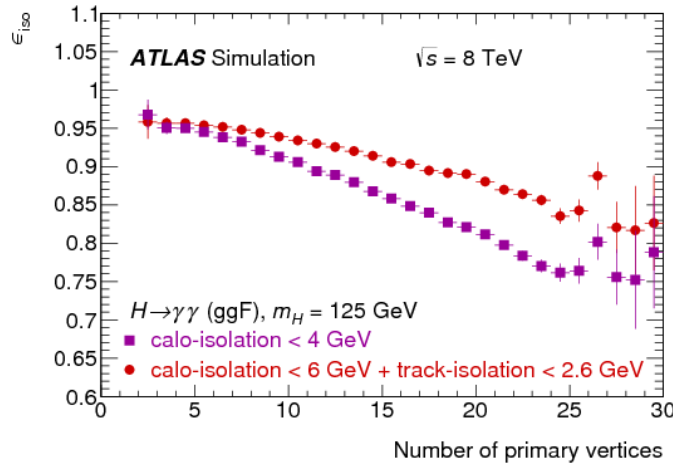


FIG. 62 – Efficacité de la coupure d'isolation en fonction du nombre de vertex reconstruits dans l'événement pour un échantillon simulé correspondant à la production du signal $H \rightarrow \gamma\gamma$ par fusion de gluons. La coupure utilisant à la fois l'isolation définie à partir du calorimètre et du détecteur interne permet une meilleure efficacité à haut taux d'empilement pour un taux de rejet similaire [189].

Pendant la première prise de données, le groupe de performance en charge des photons n'a pas réalisé de mesures d'efficacité d'isolation basées sur les données afin de quantifier l'accord données/simulation, laissant libres les groupes d'analyse dans le choix des coupures et la détermination des incertitudes systématiques associées. En revanche, des recommandations sur le calcul des variables d'isolation et leurs corrections ont été fournies, détaillées dans les sections suivantes. Des informations complémentaires sur les variables et coupures d'isolation employées dans l'analyse de données principale de la thèse seront données dans la partie IV.

Les variables d'isolation utilisant le détecteur interne et le calorimètre sont présentées respectivement dans les sections 10.2 et 10.3. L'énergie d'isolation transverse des photons dans le calorimètre nécessite le calcul de corrections particulières, introduites dans la section 10.3.2. Finalement, l'analyse des différences résiduelles observées entre données et simulations après corrections ainsi que le calcul de corrections de fuite supplémentaires basées sur les données pour l'isolation calorimétrique sont présentés dans la section 10.4. Les résultats de cette étude ont été validés par la collaboration et représentent le travail de performance principal effectué pendant les années de thèse.

10.2 L'énergie d'isolation transverse dans le détecteur interne

Une des définitions possibles de l'énergie d'isolation transverse d'un photon consiste à réaliser la somme scalaire des impulsions des traces reconstruites dans le détecteur interne situées dans un cône autour du candidat photon. On appellera cette variable l'isolation des traces dans la suite.

La quantité de traces reconstruites dans un cône donné dans le détecteur interne étant extrêmement dense à cause de la très haute luminosité des collisions du LHC (voir figure 63, gauche), il est nécessaire de sélectionner les traces à inclure dans le calcul. En pratique, uniquement les traces compatibles avec le vertex primaire du photon comme point d'émission et dont l'impulsion transverse est supérieure à 1 GeV sont considérées,²⁴ afin de minimiser les effets liés à l'empilement et à l'événement sous-jacent. Pour les mêmes raisons, un certain nombre de critères de qualité est demandé sur les traces reconstruites, comme un nombre minimum de points de mesure dans les couches de silicium (9) et un veto sur les traces passant par des cellules mortes de la première couche de pixels (voir section 5.3). Finalement, les traces associées aux photons convertis sont exclues du cône afin d'éviter une contamination de la variable par le photon lui-même. La formule finale définissant l'isolation des traces est présentée ci-dessous :

$$p_T^{\text{track.iso.}} [\text{GeV}] = \sum_{i \in \mathcal{C}(\mathbf{R})} p_T^i, \text{ pour toutes les traces } i \text{ passant la selection,} \quad (34)$$

où $p_T^{\text{track.iso.}}$ est l'isolation des traces mesurée en GeV et p_T^i est l'impulsion de la trace i située dans le cône $\mathcal{C}(\mathbf{R})$ passant la sélection évoquée dans le paragraphe précédent. La taille du cône, optimisée vis-à-vis de l'activité moyenne des événements, est généralement de 0.2. Certaines études préliminaires utilisent 0.3.

Les distributions de l'isolation des traces dans un cône de 0.3 pour les photons immédiats et le bruit de fond, obtenues à partir de candidats photons simulés passant les coupures d'identification, sont présentées dans la figure 63, droite. Elles sont composées dans les deux cas d'un pic important à 0 suivi d'une queue décroissante débutant à 1 GeV, conséquence du seuil sur l'impulsion transverse appliqué lors de la sélection des traces. Comme attendu, le signal est en moyenne bien mieux isolé que le bruit de fond. Les résultats de la simulation sont confrontés aux données dans le cadre de l'analyse de la partie IV et trouvés en bonne accord.

La très bonne résolution spatiale du détecteur interne permet ainsi d'obtenir une variable discriminante très robuste vis-à-vis de l'empilement sans corrections supplémentaires (voir figure 62). De plus, dans le cas d'un photon immédiat reconstruit avec le bon statut de conversion, il n'y a pas de problème de contamination de l'énergie d'isolation par celle du photon. Cependant, seules les particules chargées laissent des traces dans le détecteur interne, l'isolation des traces est donc nécessairement une borne inférieure de l'énergie transverse totale environnant le photon lors de son émission. Enfin, la position dans le détecteur interne et l'identification du vertex des photons convertis (1 trace, 2

²⁴En l'absence de traces associées au candidat photon (photon non-converti), le vertex primaire est identifié avec une précision de l'ordre de quelques millimètres grâce à la segmentation longitudinale du calorimètre [190].

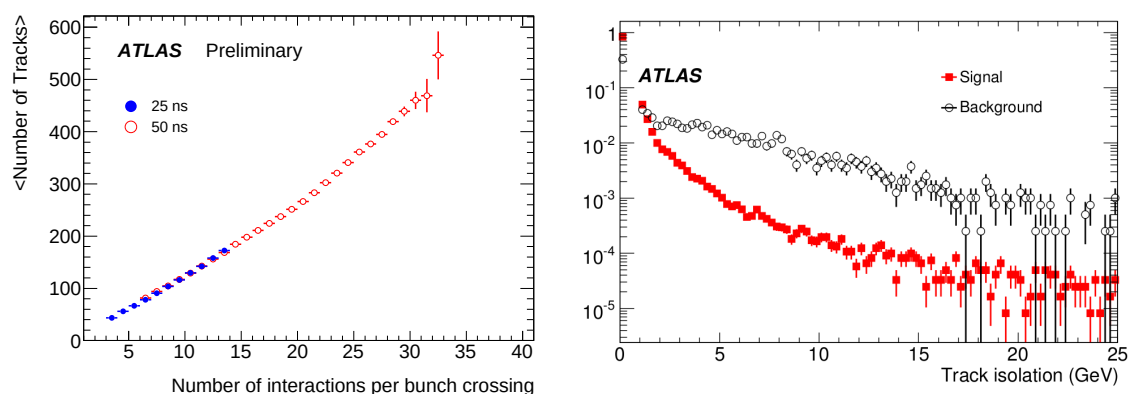


FIG. 63 – (gauche) Nombre moyen de traces reconstruites dans le détecteur interne en fonction du taux d’empilement dans les données à $\sqrt{s} = 8$ TeV pour un espacement des paquets de 50 et 25 ns [191]. (droite) Distribution normalisée de l’isolation des traces dans un cône $R = 0.3$ entre photons immédiats et bruit de fond, après application des critères d’identification. Les photons immédiats sont issus de désintégrations $H \rightarrow \gamma\gamma$ simulées, tandis que le bruit de fond est basé sur la simulation de jets hadroniques [192].

traces) et non convertis (0 trace) n’étant pas déterminées avec la même précision, des efficacités et taux de rejet différents sont attendus en fonction du statut de conversion.

L’isolation des traces est souvent utilisée en complément de l’énergie d’isolation transverse calculée à partir du calorimètre, objet de la section 10.3.

10.3 L’énergie d’isolation transverse dans le calorimètre

10.3.1 Définition

Il est également possible de définir une énergie d’isolation transverse pour les photons en se basant sur les dépôts d’énergie dans le calorimètre électromagnétique et hadronique. De façon similaire à l’isolation des traces, le but est de calculer une variable évaluant l’énergie transverse environnant le candidat photon lors de son émission. Elle est estimée le plus souvent dans un cône de taille fixe 0.4. Cette valeur provient d’un compromis permettant aux expérimentateurs de réaliser une mesure ne dépendant pas trop de l’empilement (cône pas trop grand) pouvant être comparée aux calculs théoriques (cône pas trop petit), suivant les recommandations de la référence [121]. La variable associée sera appelée isolation calorimétrique dans la suite.

Il est intéressant d’assembler les cellules environnant le *cluster* du candidat photon (voir section 7) en *clusters* particuliers à trois dimensions (η , ϕ , R), appelés *clusters* topologiques [193] pour le calcul de l’isolation calorimétrique. Les *clusters* topologiques ont la particularité de ne pas avoir de taille fixe et suppriment très efficacement les effets dus au bruit des cellules du calorimètre. L’algorithme de reconstruction associé considère d’abord uniquement les cellules dont l’énergie est supérieure à quatre fois l’énergie correspondant au bruit moyen, $4\sigma_{\text{bruit}}$. Celui-ci est constitué de la superposition

du bruit de l'électronique et du bruit dû à l'empilement :

$$\sigma_{\text{bruit}} = \sqrt{(\sigma_{\text{bruit}}^{\text{electronique}})^2 + (\sigma_{\text{bruit}}^{\text{empilement}})^2}, \quad (35)$$

où σ_{bruit} est le bruit total, $\sigma_{\text{bruit}}^{\text{electronique}}$ est le bruit de l'électronique et $\sigma_{\text{bruit}}^{\text{empilement}}$ est le bruit de l'empilement. σ_{bruit} est déterminé séparément pour chaque cellule du calorimètre et mis à jour lors de changements importants de configuration de la machine. Par exemple pour les données à 8 TeV, σ_{bruit} est estimé à environ 40 MeV à $\eta = 0$ et augmente rapidement avec la pseudorapidité à cause de l'empilement, atteignant environ 100 MeV à $\eta = 2$.

Toutes les cellules voisines des cellules initiales dont l'énergie est supérieure à $2\sigma_{\text{bruit}}$ sont ensuite ajoutées itérativement, en fusionnant les *clusters* topologiques de départ si nécessaire. À la fin de l'exécution de l'algorithme, les *clusters* présentant plusieurs extrema locaux sont à nouveau divisés. La description complète de l'algorithme de reconstruction des *clusters* topologiques est disponible dans la référence [193].

Les énergies transverses non-calibrées²⁵ et positives des *clusters* topologiques ayant leur barycentre inclus dans le cône d'isolation sont alors sommées scalairement. L'énergie transverse contenue dans le *cluster* électromagnétique étendu de taille fixe 5×7 en $\eta \times \phi$ centré sur le candidat photon (et non le *cluster* topologique associé) est soustraite afin d'exclure l'énergie du candidat dans le calcul d'isolation. Un schéma récapitulant les différentes étapes du calcul de l'isolation calorimétrique est donné dans la figure 64.

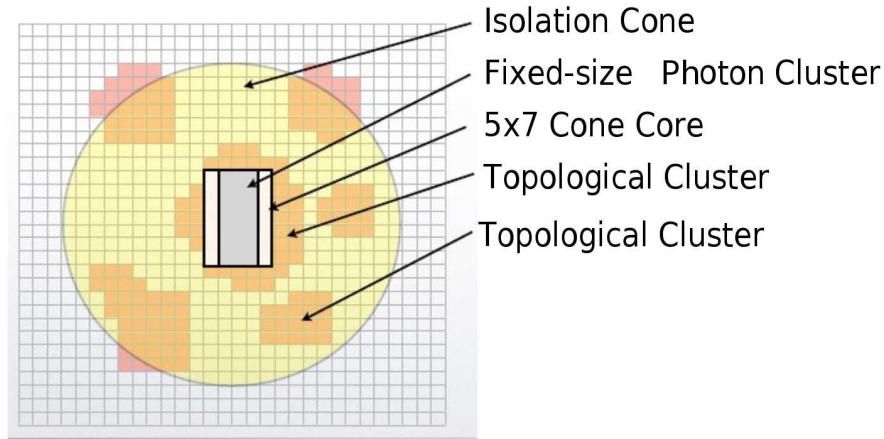


FIG. 64 – Schéma en η - ϕ illustrant le calcul de l'isolation dans le calorimètre basé sur les *clusters* topologiques. La grille représente les cellules du calorimètre, le cône d'isolation est affiché en jaune et les *clusters* topologiques en rouge. La partie blanche correspond au *cluster* étendu de 5×7 qui est soustrait afin d'exclure l'énergie du candidat dans le calcul d'isolation. Celui-ci est centré sur le candidat photon, dont le *cluster* de taille fixe apparaît en gris [194].

²⁵L'énergie des *clusters* topologiques individuels dans ATLAS est calculée à partir des énergies de chaque cellule le constituant. Celles-ci sont reconstruites grâce à l'équation 33 (voir section 7.1) et reproduisent en moyenne l'énergie d'une gerbe électromagnétique. La non-calibration des *clusters* topologiques (corrections supplémentaires dues aux variations de réponse, pertes, etc) est l'une des raisons principales pour laquelle les énergies d'isolation définies au niveau reconstruit et au niveau particule diffèrent. Une comparaison entre les deux niveaux sera menée explicitement dans la partie IV.

L'isolation calorimétrique ainsi obtenue est alors largement indépendante des fluctuations dues au bruit des cellules, mais inclut toujours des contributions non négligeables liées à l'empilement en temps et à l'événement sous-jacent. Les corrections correspondantes sont présentées dans la section 10.3.2.

Par ailleurs, on s'attend à ce qu'une partie de l'énergie du candidat photon de l'ordre du pourcent se situe hors du *cluster* de 5×7 qui est soustrait. Il est alors nécessaire de calculer des corrections supplémentaires, appelées corrections de fuite. Les corrections appliquées par défaut sont basées uniquement sur la simulation et introduites brièvement dans la section 10.3.3, tandis que les sections 10.3.4 et 10.4 présentent respectivement les différences résiduelles entre données et simulation après application des corrections de base et le calcul de corrections additionnelles de fuite basées sur les données. Cette dernière section est issue de travaux personnels réalisés pendant la thèse et validés par la collaboration.

10.3.2 Corrections dues à l'empilement et à l'événement sous-jacent

Le nombre d'interactions additionnelles pendant un croisement de faisceaux se situe en moyenne autour de vingt et un pour les données à $\sqrt{s} = 8$ TeV (voir section 4.3). Avec l'événement sous-jacent, ils contribuent à surestimer l'énergie d'isolation calorimétrique définie dans la partie précédente.

Afin de soustraire cette contribution, une correction – recalculée pour chaque événement – est appliquée en utilisant une méthode de soustraction basée sur les références [195, 196]. Cette méthode consiste à estimer l'énergie ambiante moyenne de l'événement dans des intervalles de pseudorapidité, ρ_{EA}^η , en reconstruisant les jets associés sans seuil minimum en p_T . La reconstruction des jets est basée sur les *clusters* topologiques définis dans la section précédente afin de supprimer l'effet relié au bruit des cellules du calorimètre de la même façon que pour l'énergie d'isolation. Dans chaque intervalle de pseudorapidité, la densité d'impulsion transverse moyenne des jets est alors prise comme la médiane de la distribution $p_T^{\text{jet}}/A_{\text{jet}}$, où p_T^{jet} est l'impulsion transverse non-calibrée des jets et A_{jet} est leur aire associée, définie comme la surface couverte par l'algorithme de reconstruction, en l'occurrence k_T avec un paramètre $R = 0.5$ [197].

La correction finale d'empilement et de l'événement sous-jacent est finalement donnée par la formule suivante :

$$c_{\text{EA}}^\eta = -(\pi R^2 - 5 \times 7 \times 0.025 \times \pi/128) \cdot \rho_{\text{EA}}^\eta, \quad (36)$$

où le terme entre parenthèses représente l'aire couverte par un cône d'isolation calorimétrique de taille R dans la seconde couche du calorimètre, où la plupart de l'énergie est déposée. L'aire associée au *cluster* de taille fixe est soustraite explicitement. En pratique, c_{EA}^η est déterminée pour chaque événement dans deux intervalles de rapidité, $|\eta| < 1.5$ et $1.5 < |\eta| < 3$, afin d'obtenir un nombre suffisant de jets dans chaque catégorie tout en prenant en compte la dépendance en pseudo-rapidité des effets d'empilement. La valeur typique associée à cette correction est d'environ -2 GeV et -1.5 GeV dans la partie centrale et à l'avant à $\sqrt{s} = 8$ TeV ($\langle \mu \rangle \simeq 21$), respectivement. Il est important de souligner qu'aucune extrapolation n'est faite de la simulation vers les données pour ces corrections, c_{EA}^η étant calculé séparément dans les deux cas.

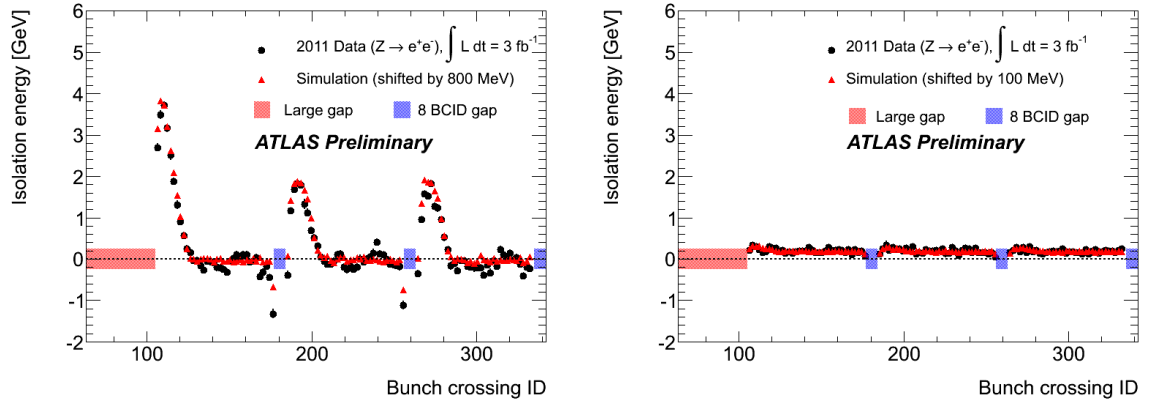


FIG. 65 – Comparaison entre l'isolation calorimétrique basée directement sur l'énergie des cellules (gauche) et celle basée sur les *clusters* topologiques (droite), en fonction du numéro de croisement de faisceaux [194].

La stabilité de la variable d'isolation calorimétrique basée sur les *clusters* topologiques après correction en fonction du croisement de faisceaux est illustrée dans la figure 65, droite. Le comportement de la variable équivalente utilisant directement la somme de l'énergie transverse des cellules est représenté dans la figure 65, gauche.²⁶ Alors que la variable d'isolation simple montre une très grande sensibilité vis-à-vis du numéro de croisement de faisceaux, qui est directement relié à la luminosité instantanée et donc à l'empilement (en temps et hors-temps), l'isolation calorimétrique topologique reste stable.

Le fait que la correction appliquée soit moyennée dans deux intervalles en η seulement a certaines conséquences importantes. Par exemple, dans le cas d'un événement comportant un photon bien isolé situé en η positif avec des dépôts hadroniques significatifs dans la région en pseudorapidité négative correspondante, il est possible que la correction devienne supérieure à l'isolation calorimétrique elle-même, donnant lieu à une valeur d'énergie d'isolation négative après correction. La situation inverse est également possible dans le cas d'un photon émis très proche d'un jet d'empilement.

Cette remarque contribue à accentuer les différences entre l'énergie d'isolation définie au niveau particule et celle reconstruite expérimentalement dans le calorimètre. Il est également probable que le champ magnétique du solénoïde dévie les particules chargées de bas p_T de façon à ce qu'un nombre non négligeable d'entre elles migre depuis leur point d'émission à l'intérieur (ou à l'extérieur) du cône d'isolation au niveau du calorimètre. Une solution possible pour améliorer l'accord entre les deux variables pourrait alors consister à faire des catégories plus petites en η centrées sur le candidat pour l'estimation de l'énergie ambiante. Plus généralement, il serait nécessaire d'imaginer une méthode de calibration des *clusters* topologiques, corrigeant entre autres les effets dus aux bas p_T . Ce dernier point ne paraît pas trivial à cause des incertitudes théoriques importantes existant sur la production de particules de basse énergie dans les collisionneurs hadroniques, rendant toute calibration potentiellement dépendantes des modèles non-perturbatifs des différents générateurs MC.

²⁶Dans les deux cas, les corrections d'empilement sont calculées à partir de jets reconstruits depuis les clusters topologiques, seul le calcul de l'énergie d'isolation elle-même change.

Dans la section 10.3.3, les corrections de fuite de l'isolation calorimétrique basées sur la simulation sont abordées.

10.3.3 Corrections des fuites à partir de la simulation

L'autre correction importante pour l'isolation calorimétrique correspond à la soustraction des fuites d'énergie du photon lui-même hors du *cluster* de 5×7 considéré lors de l'exécution de l'algorithme. Les fuites dépendent du développement latéral des gerbes électromagnétiques et devraient représenter un effet de l'ordre de quelques pourcents de l'énergie du photon. Il est donc attendu que les corrections associées montrent une forte corrélation avec l'impulsion transverse des photons. Par ailleurs, comme déjà mentionné dans les sections précédentes, la forme des gerbes dépend fortement de la quantité de matière en amont, elle-même fonction de la pseudorapidité. Il est donc nécessaire de calculer les corrections de fuite dans des catégories fines en E_T^γ , $|\eta^\gamma|$, et de traiter séparément photons convertis et non convertis.

Afin de disposer d'une statistique permettant d'effectuer une étude avec la meilleure granularité en E_T^γ , $|\eta^\gamma|$ possible, des premières corrections basées exclusivement sur la simulation et extrapolées directement dans les données sont réalisées.²⁷ Le groupe d'ATLAS chargé d'évaluer les performances des photons préconise leur utilisation dans toutes les analyses. Ces corrections utilisent des échantillons simulés de photons convertis et non convertis générés seuls afin de s'affranchir des effets autres que ceux des fuites (empilement, etc). Les candidats sont divisées en dix catégories en fonction de sa pseudorapidité, elles-mêmes subdivisées en quarante-huit catégories de E_T afin de couvrir l'intervalle [20 GeV, 500 GeV] par pas de 10 GeV. La distribution de l'énergie d'isolation brute est tracée dans chaque catégorie. Elle se présente sous la forme d'une courbe de type boule de cristal,²⁸ dont la valeur maximale correspond à l'énergie de fuite la plus probable. En effet, l'énergie déposée par le photon hors du *cluster* étendu de taille fixe 5×7 est naturellement une variable aléatoire dont la distribution est déterminée lors de la simulation MC du passage du photon à travers le détecteur. La correction associée ne peut alors constituer qu'une correction moyenne, au même titre que la correction d'empilement et de l'événement sous-jacent détaillée dans la section 10.3.2. La valeur de fuite la plus probable est collectée dans chaque catégorie en E_T^γ , $|\eta^\gamma|$ et tracée en fonction de E_T^γ dans chaque intervalle de pseudorapidité. Finalement, son comportement est modélisé par une fonction affine déterminant l'énergie à soustraire de l'isolation calorimétrique du photon en fonction de son impulsion transverse. L'étude est effectuée séparément pour les photons convertis et non convertis.

Les corrections associées varient significativement en fonction du statut de conversion et de la région du détecteur, mais sont typiquement inférieures à 1 GeV pour $E_T^\gamma = 50$ GeV et de l'ordre de 2 à 4 GeV pour $E_T^\gamma = 200$ GeV. De plus, la pente du modèle utilisé augmente sensiblement avec la pseudorapidité, passant d'environ 1.5% pour $\eta^\gamma = 0$ à 4% pour $\eta^\gamma = 2$, quelque soit le statut de conversion. En revanche, les ordonnées à l'origine restent relativement stables vis-à-vis de la pseudorapidité, de l'ordre de -0.4 GeV. Leur origine physique est encore mal comprise.

²⁷Les photons radiatifs du $Z^0 \rightarrow l\bar{l}\gamma$ sont difficilement utilisables pour cette étude à cause de leurs impulsions transverses limitées ($E_T^\gamma < 100$ GeV).

²⁸La fonction boule de cristal est définie comme la somme d'une gaussienne et d'une queue radiative décrite par une fonction puissance.

Dans la section 10.3.4, l'accord entre données et simulation après corrections d'empilement et de fuite est discuté.

10.3.4 Comparaison entre données et simulation après corrections

Après application des corrections d'empilement et des corrections de fuite, la distribution d'isolation calorimétrique des photons se présente approximativement comme une fonction de type boule de cristal avec un pic centré en zéro et une queue à plus haute énergie. La queue est peuplée à la fois de photons directs émis proches d'un jet (issu du processus dur ou de l'empilement) et de photons de fragmentation. La largeur de la partie gaussienne est dominée par les fluctuations autour de la moyenne des corrections d'empilement (de fuite) à bas (haut) E_T^γ . Des exemples de telles distributions dans la simulation et dans les données seront montrées plus loin dans cette section.

Au cours de la prise de données à $\sqrt{s} = 8$ TeV, plusieurs groupes analysant des échantillons de photons immédiats de haut E_T^γ ont voulu vérifier l'accord entre données et simulation pour l'isolation calorimétrique corrigée des deux effets présentés dans les paragraphes précédents [186, 198]. Les groupes concernés ont validé les corrections d'empilement et de l'événement sous-jacent, qui rendent effectivement l'isolation calorimétrique quasi-indépendante de l'activité de l'événement à la fois dans la simulation et les données (voir figure 66). Par ailleurs, les écarts entre données et simulation restent stable quelque soit la variable d'isolation employée.

En revanche, des différences entre données et simulation croissant avec E_T^γ ont été observées, avec notamment des distributions d'isolation dans les données décalées à plus haute énergie comparées aux simulations (voir figure 67). Une forte dépendance en pseudorapidité du désaccord a également pu être identifiée (voir figure 68). Le décalage de l'énergie d'isolation dans les données a pour conséquence directe de baisser l'efficacité d'une coupure fixe sur l'isolation calorimétrique à haut E_T^γ , particulièrement si celle-ci est fixée à une faible valeur d'énergie d'isolation, et d'augmenter l'incertitude associée, la simulation reproduisant mal ce comportement (voir figure 69). Ces effets peuvent s'avérer problématiques, particulièrement lors de recherche de nouvelle physique à haute énergie, et devraient s'aggraver avec la montée en puissance du LHC pendant la seconde prise de données si aucune amélioration n'est apportée.

Ainsi, dans chaque groupe d'analyse concerné des corrections supplémentaires adaptées aux mesures ont été calculées afin de faire correspondre données et simulation. Certains groupes ont eu besoin de considérer uniquement la dépendance en E_T^γ du désaccord (voir figure 67), tandis que d'autres analyses plus sensibles à cet effet ont dûes également faire dépendre les corrections de la pseudo-rapidité des candidats (voir figure 68).

Les mêmes questions se sont posées dans le cadre de la mesure de la section efficace de production de paires de photons à $\sqrt{s} = 8$ TeV, sujet principal de ma thèse (voir partie IV). Après réflexion, il a été décidé de calculer des corrections supplémentaires pour la simulation les plus générales possibles et basées sur les données uniquement, afin de permettre leur utilisation directe par d'autres analyses étudiant les photons à $\sqrt{s} = 8$ TeV et de tenter de comprendre davantage la physique sous-jacente au problème [201]. Le détail de cette étude, validée par le groupe de performance et dont la méthodologie a été retenue pour l'évaluation des incertitudes reliées à l'isolation pour le *run 2*, est donné dans la section 10.4.

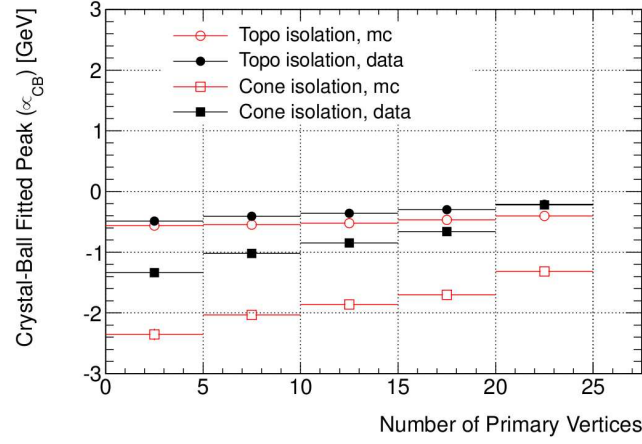


FIG. 66 – Figure issue de la note interne à la collaboration de l'analyse publiée dans la référence [198], montrant le bon accord entre données et simulation pour la variable d'isolation basée sur les *clusters* topologiques (*Topo isolation*) vis-à-vis du nombre de vertex reconstruits dans l'événement. L'accord est évalué en comparant la position du pic d'un modèle de type boule de cristal ajusté aux différentes distributions d'isolation. La variable d'isolation basée directement sur les cellules du calorimètre (*Cone isolation*) est également représentée à titre illustratif. Les échantillons simulés sont issus de la désintégration d'un graviton en deux photons [199].

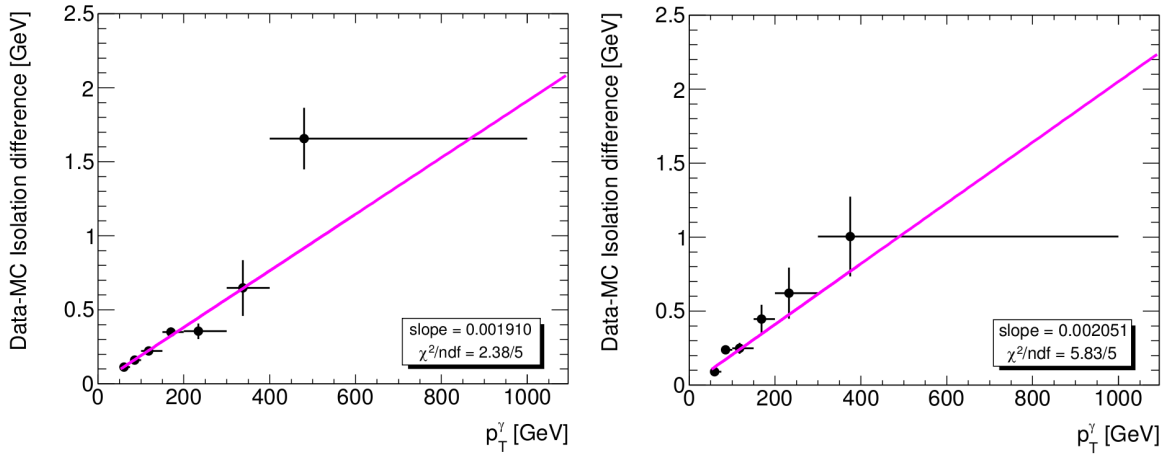


FIG. 67 – Figure issue de la même note interne que la figure 66. L'accord entre l'isolation calorimétrique des photons dans les données et la simulation est évalué en considérant la différence dans la position du pic d'un modèle de type boule de cristal ajusté aux différentes distributions. Cette variable est représentée en fonction de l'impulsion transverse des photons sélectionnés et intégrée en η . Le modèle de correction linéaire utilisé dans l'analyse pour corriger les données est également représenté [199].

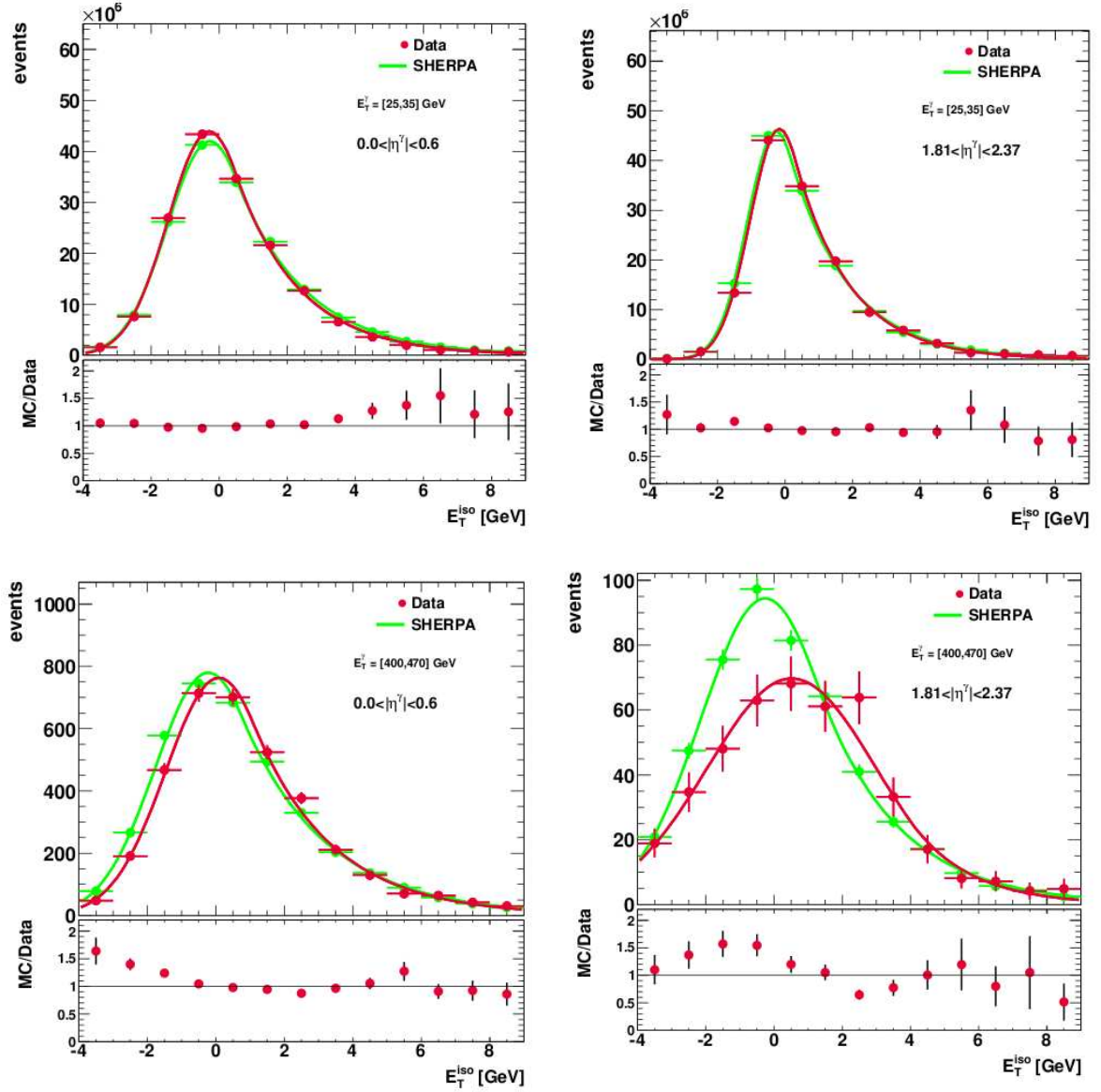


FIG. 68 – Figure issue de la note interne à la collaboration de l'analyse publiée dans la référence [186], comparant les distributions d'isolation calorimétrique entre données et simulation dans plusieurs catégories en pseudo-rapidité et impulsions transverses. Un bon accord est observé à bas E_T^{γ} uniquement. Les distributions sont modélisées par une fonction de type boule de cristal et corrigées par la suite dans le cadre de l'analyse [200].

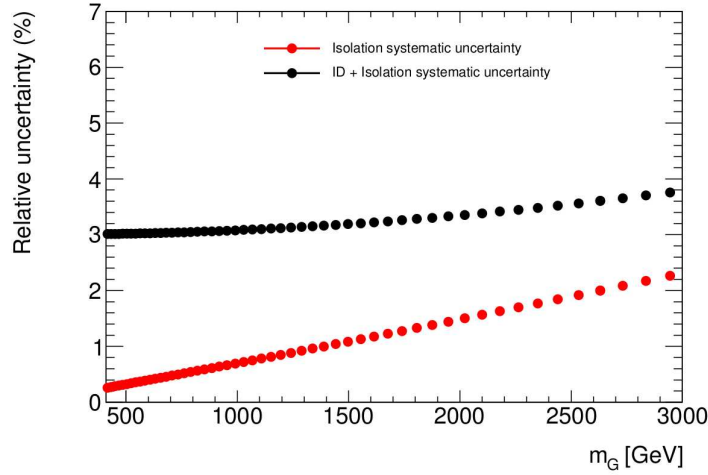


FIG. 69 – Figure issue de la même note interne que la figure 66. L’erreur systématique reliée à l’efficacité d’isolation (rouge) et l’erreur totale reliée à l’efficacité d’identification et d’isolation sont représentées en fonction de la masse de la nouvelle particule recherchée (ici un graviton). L’incertitude reliée à l’isolation devient significative pour $m_G > 2$ TeV environ [199].

10.4 Corrections supplémentaires de l’isolation calorimétrique basées sur les données

10.4.1 Origines des différences observées

Après avoir analysé avec attention les résultats des autres groupes d’analyse concernant l’accord données/simulation pour l’isolation calorimétrique des photons, une liste des sources possibles de désaccord a été établie et est rapportée ci-dessous :

- ◇ **Calibration des photons** : il est impossible d’exclure complètement qu’un problème de second ordre relié à la calibration du calorimètre soit responsable du désaccord observé. Néanmoins, la calibration a donné lieu à de nombreuses études et vérifications tout au long du *run 1*, qui ont permis d’atteindre une très bonne précision sur l’échelle et la résolution en énergie des photons (voir section 8).
- ◇ **Corrections dues à l’empilement et à l’événement sous-jacent** : ces corrections sont calculées événement par événement indépendamment pour les données et les échantillons simulés. Ainsi, aucune extrapolation n’est effectuée entre les deux. De plus, aucun groupe d’analyse n’a observé une dépendance du désaccord données/simulation en fonction d’observables sensibles à l’empilement telle que le nombre de vertex reconstruits dans l’événement (voir figure 66).
- ◇ **Corrections des fuites** : ces corrections sont basées exclusivement sur la simulation et extrapolées inchangées dans les données. Elles montrent une dépendance importante vis-à-vis de E_T^γ et $|\eta^\gamma|$, à l’image du désaccord observé. Par ailleurs, la présence des *fudge factors* dans le cadre des études d’identification démontre que la description latérale des gerbes électromagnétiques dans la simulation est imparfaite (voir section 9). Or, cette description est essentielle à la bonne détec-

mination des corrections de fuite. Il paraît donc probable que ces écarts aient un effet significatif sur les corrections de fuite décrites dans la partie précédente, qui n'utilisent pas les variables d'identification présentées dans la section 9 et donc ne peuvent pas être corrigées simplement à l'aide des *fudge factors*.

- ◊ **Génération des photons** : une partie de l'écart entre données et simulation vient très probablement de la génération des photons directement. En effet, les générateurs utilisés pour les simulations complètes du détecteur font souvent usage d'éléments de matrices limités pour reproduire la composante directe (au mieux LO avec une partie des corrections NLO), alors que les corrections supérieures et notamment celles correspondant à l'émission de jets supplémentaires jouent un rôle déterminant (voir section 3.5). La composante de fragmentation non-perturbative est également mal connue et impacte les distributions d'isolation de façon importante, les photons associés étant en général moins bien isolés. Finalement, les fluctuations autour de la contribution moyenne de l'événement sous-jacent pourrait également jouer un rôle. Cependant, le désaccord observé augmente avec E_T^γ tandis que la composante de fragmentation baisse avec l'impulsion des photons. De plus, le désaccord ne montre pas de dépendance accrue vis-à-vis de variables sensibles aux ordres supérieurs comme $\Delta\phi_{\gamma\gamma}$ pour les paires de photons. Ainsi, tout semble indiquer que les différences liées à la génération des photons constituent un problème de second ordre.

Les deux hypothèses principales retenues pour expliquer le désaccord observé sont donc la mauvaise description du développement latéral des gerbes électromagnétiques dans la simulation (privilegiée) et la génération des photons dans les simulations (secondaire). Dans le premier cas, l'origine est purement liée à un problème de modélisation du détecteur et peut donc donner lieu à un calcul de corrections générales applicables à l'ensemble des analyses comportant des photons.²⁹ En revanche, la deuxième hypothèse souligne un problème de générateur, rendant toute correction dépendante du processus physique étudié (topologie, événement sous-jacent, etc).

Le désaccord consiste dans l'observation de distributions d'isolation calorimétrique décallées à plus haute énergie dans les données par rapport aux simulations. Dans ce contexte, l'hypothèse privilégiée énoncée plus haut indiquerait la présence d'énergie de fuite résiduelle dans le cône d'isolation des données, et donc une sous-estimation du développement latéral des gerbes dans la simulation. Cela est cohérent avec le signe des corrections correspondant aux *fudge factors* introduites pour les critères d'identification, ce qui renforce d'autant plus cette hypothèse (voir section 9 et particulièrement la figure 59).

Dans la section 10.4.2, la méthode basée sur les données pour corriger les corrections de fuite imparfaites de la section 10.3.3 est présentée. Cette étude vise uniquement à corriger l'effet dû à la mauvaise description du développement latéral des gerbes électromagnétiques dans la simulation. Ainsi, une incertitude systématique importante sera calculée afin de prendre en compte le biais potentiel relié à la génération et à la topologie des événements considérés.

²⁹Un effet similaire est attendu pour les électrons mais n'a pas encore été étudié en détails.

10.4.2 Méthodologie

La méthode de correction mise au point pendant les années de thèse se base sur un ajustement de l'isolation calorimétrique des photons immédiats et du bruit de fond dans un échantillon TIGHT de photons inclusifs dans les données $(\gamma + X)$, appelée *template fit* dans la suite. Le *template fit* repose sur la modélisation adéquate de la distribution d'isolation calorimétrique des photons immédiats et du bruit de fond. Il a déjà été mis en œuvre dans plusieurs analyses de données de photons au LHC [202–204] et sera utilisé à nouveau sur un échantillon et dans un but différent dans l'analyse de la partie IV.

L'utilisation d'un échantillon de photons inclusif permet de maximiser la statistique, particulièrement à haute énergie, et de sélectionner sans distinction tous les types de topologies. L'échantillon est défini à l'aide de la même sélection que celle utilisée dans le cadre de la mesure de la section efficace de production inclusive de photons à 8 TeV par la collaboration ATLAS qui est sur le point d'être publiée [186], à l'exception de la coupure d'isolation calorimétrique qui n'est pas appliquée. Aucune coupure sur l'isolation des traces n'est appliquée afin de ne pas biaiser l'isolation calorimétrique de l'échantillon. Celui-ci contient plusieurs millions de photons avec des impulsions transverses allant de 20 GeV à 1.5 TeV.

Le modèle d'ajustement adopté se présente sous la forme suivante :³⁰

$$W_{\text{tot}} \cdot F_{\text{tot}}(E_{\text{T}}^{\text{calo.iso.}}) = W_{\gamma} \cdot F_{\gamma}(E_{\text{T}}^{\text{calo.iso.}}) + W_{\text{b}} \cdot F_{\text{b}}(E_{\text{T}}^{\text{calo.iso.}}), \quad (37)$$

où W_{tot} est le nombre total d'événements attendus dans l'échantillon et W_{γ} et W_{b} sont respectivement le nombre d'événements contenant un vrai photon immédiat et le nombre d'événements de bruit de fond, avec $W_{\text{tot}} = W_{\gamma} + W_{\text{b}}$. F_{tot} représente la fonction densité de probabilité (ou pdf pour *probability density function*) globale de l'isolation calorimétrique de l'échantillon et F_{γ} , F_{b} sont respectivement les pdfs de l'isolation des photons immédiats et du bruit de fond. Toutes les pdfs sont fonction de l'isolation calorimétrique du candidat photon de plus haute impulsion transverse présent dans l'événement uniquement, celle-ci étant corrigée des effets d'empilement et de fuite suivant les sections 10.3.2 et 10.3.3. Une fois F_{γ} et F_{b} choisies (voir paragraphe suivant), la procédure d'ajustement consiste à minimiser l'estimateur statistique de maximum de vraisemblance étendu associé au modèle sur les données à 8 TeV en variant W_{γ} et W_{b} indépendamment. La minimisation est effectuée à l'aide du code ROOFIT [205], lui-même basé sur le code MINUIT du CERN [206].

Le modèle de bruit de fond F_{b} est extrait à partir de l'échantillon NON-TIGHT, défini comme l'ensemble des candidats photons passant LOOSE'4 et ne passant pas TIGHT (voir section 9.4). Le reste de la sélection est inchangé. L'échantillon ainsi obtenu est largement dominé par le bruit de fond, avec la présence d'une légère contamination par des photons immédiats de quelques pourcents. La pdf extraite correspond alors à une somme pondérée entre la véritable pdf de bruit de fond (dominante) et la pdf de photons immédiats :

$$F_{\text{b}}(E_{\text{T}}^{\text{calo.iso.}}) = \alpha \cdot F_{\gamma}(E_{\text{T}}^{\text{calo.iso.}}) + (1 - \alpha) \cdot F_{\text{jet}}(E_{\text{T}}^{\text{calo.iso.}}), \quad (38)$$

³⁰Un bruit de fond supplémentaire très faible ($< 1\%$) composé d'électrons mal identifiés est également présent dans l'échantillon [186] mais négligé dans cette étude.

où α représente la contamination de la région NON-TIGHT par les photons immédiats. Les pdfs de l'échantillon NON-TIGHT sont identifiées à celles de l'échantillon TIGHT par propriété de la sélection LOOSE' (voir section 9.4). Il est alors aisé de montrer en combinant l'équation 37 et 38 que la contamination de la région NON-TIGHT ne change donc pas le modèle final $W_{\text{tot}} \cdot F_{\text{tot}}$ mais uniquement W_γ et W_b , W_b incluant une part des photons immédiats si aucune correction n'est effectuée. Cet effet ne sera pas pris en compte dans cette étude, qui vise uniquement à corriger l'isolation des photons immédiats dans le MC (F_γ) et non pas à obtenir des estimations précises de W_γ et W_b .

Pour le modèle de signal F_γ , la pdf résultant du produit de convolution entre la distribution de l'échantillon TIGHT issue du générateur MC SHERPA 1.4.0 [83] et une pdf gaussienne est utilisée :³¹

$$F_\gamma(E_T^{\text{calo.iso.}}, \mu, \sigma) = F_\gamma^{\text{MC}}(E_T^{\text{calo.iso.}}) * G_{\mu, \sigma}(E_T^{\text{calo.iso.}}), \quad (39)$$

où F_γ est la pdf finale des photons immédiats, F_γ^{MC} est la pdf extraite du générateur MC SHERPA et $G_{\mu, \sigma}$ est une pdf gaussienne de valeur moyenne μ et d'écart-type σ . Une telle expression est souvent utilisée pour imiter les effets de résolution d'un détecteur sur une observable. En effet, il est possible de montrer que la pdf décrivant la somme de deux variables aléatoires X et Y correspond au produit de convolution de leur pdf respective. Ainsi, l'équation 39 considère l'énergie d'isolation calorimétrique dans les données comme la somme entre l'énergie d'isolation simulée et une variable aléatoire de bruit B suivant une loi gaussienne $G_{\mu, \sigma}$:

$$E_{T, \text{data}}^{\text{calo.iso.}} = E_{T, \text{mc}}^{\text{calo.iso.}} + B, \quad (40)$$

où B représente l'énergie de fuite restante dans le cône d'isolation des données après les corrections d'empilement et les corrections de fuite basées sur la simulation (voir section 10.3.4). Elle est supposée suivre une loi gaussienne $G_{\mu, \sigma}$. Les paramètres μ et σ correspondent alors respectivement à la valeur moyenne et aux fluctuations des énergies de fuite restantes dans les données et seront laissés libres ou non pendant l'ajustement, comme nous le verrons par la suite. De plus, le modèle est implémenté de façon à ce que la gaussienne $G_{\mu, \sigma}$ devienne une distribution de Dirac de paramètre μ dans le cas où σ tend vers 0 :³²

$$\lim_{\sigma \rightarrow 0^+} G_{\mu, \sigma}(E_T^{\text{calo.iso.}}) \rightarrow \delta_\mu(E_T^{\text{calo.iso.}}). \quad (41)$$

Par propriété du produit de convolution, les équations 39 et 41 impliquent alors :

$$\lim_{\sigma \rightarrow 0^+} F_\gamma(E_T^{\text{calo.iso.}}, \mu, \sigma) \rightarrow F_\gamma^{\text{MC}}(E_T^{\text{iso}} - \mu), \quad (42)$$

signifiant que dans le cas limite où σ tend vers 0, la correction consiste uniquement à appliquer une translation constante de $+\mu$ à la distribution MC.

³¹des détails supplémentaires sur les données simulées sont disponibles dans la référence [186].

³²En pratique, la valeur minimale de σ est fixée à 0.0001 GeV.

Le modèle de bruit de fond F_b présente l'avantage d'être extrait directement des données, et donc ne souffre pas des problèmes liés à la simulation. Par ailleurs, le fait de multiplier le MC par une gaussienne pour les photons immédiats permet à la fois de corriger la valeur moyenne de fuite (μ) et de contrôler les fluctuations associées grâce à la largeur de la gaussienne (σ), tout en restant très simple à mettre en œuvre. Une des limites du modèle gaussien est probablement son caractère symétrique, impliquant nécessairement des fluctuations additionnelles symétriques si $\sigma \neq 0$. Une évolution possible pourrait considérer à convoluer la distribution MC à une fonction boule de cristal. Cependant, comme nous le verrons par la suite, σ joue un rôle minime dans les résultats finals ce qui fait de l'équation 39 un modèle bien adapté au cas d'étude considéré.

D'autre part, il est préférable de définir un modèle non-catégorisé (*unbinned*) afin de garder la procédure d'ajustement sensible aux faibles valeurs de μ et σ . En effet, dans le cas d'un modèle catégorisé (*binned*), les pdfs se présentent sous la forme de fonctions constantes par intervalle. Or, si ces intervalles ont une largeur L trop grande par rapport aux corrections attendues, $L \gg \mu, \sigma$, l'ajustement n'y sera pas sensible. Dans les faits, les pdfs d'isolation F_{γ}^{MC} et F_b ne peuvent pas être décrites simplement par une fonction paramétrique vu la précision souhaitée ($\simeq 10$ MeV). Pour résoudre ce problème, une technique de lissage basée sur des estimateurs de noyaux gaussiens (KDE, *kernel density estimator*) [207] est utilisée afin d'extraire des pdfs F_{γ}^{MC} et F_b continues (*unbinned*) à partir des données et du MC. Cette méthode sera utilisée à nouveau dans la partie IV.

Les échantillons de photons inclusif TIGHT et NON-TIGHT sont séparés en 14 catégories en pseudorapidité, elles-mêmes séparées en 14 catégories en E_T^{γ} , formant donc 196 catégories au total. Pour chaque catégorie, la procédure d'ajustement est exécutée trois fois consécutivement : la première fois en utilisant la distribution MC sans corrections additionnelles pour la pdf de photons immédiats, la deuxième fois en utilisant le modèle simplifié de l'équation 42 (translation simple μ à ajuster), et la dernière fois en laissant à la fois μ et σ libres. Cette stratégie est adoptée afin de pouvoir tester la pertinence des différents paramètres du modèle et le gain véritable apporté par les corrections μ et σ sur la qualité de l'ajustement. Les distributions F_{γ}^{MC} et F_b sont extraites séparément dans chaque catégorie afin de tenir compte des dépendances en E_T^{γ} et $|\eta^{\gamma}|$ des énergies d'isolation simulées et réelles pour les photons immédiats et le bruit de fond. L'ajustement est réalisé dans l'intervalle d'isolation calorimétrique $[-2 \text{ GeV}, 10 \text{ GeV}]$. La limite inférieure est choisie de façon à ne pas introduire trop de fluctuations statistiques, tandis que la limite supérieure est placée à une énergie d'isolation suffisamment haute pour s'assurer que les corrections ne biaisent pas les queues des distributions.

Une catégorisation de l'échantillon en fonction de l'énergie totale des photons E^{γ} aurait été préférable, les fuites dépendant a priori de E^{γ} et non de E_T^{γ} uniquement. Cependant, cela n'est pas réalisable simplement à cause des différents facteurs de *prescale* des *triggers* photons d'ATLAS (voir section 5.7 et [186]). En effet, ceux-ci dépendent de l'impulsion transverse du candidat photon et non de son énergie totale, ce qui amènerait à mélanger des événements enregistrés avec des *prescales* différentes dans le cas de catégories en E^{γ} . Des corrections d'échelle seraient alors nécessaires et augmenteraient très probablement les incertitudes. Par conséquent, les corrections sont calculées dans des catégories en E_T^{γ} , $|\eta^{\gamma}|$ incluant uniquement les données enregistrées à l'aide d'un seul *trigger*. Chaque catégorie est associée à l'intervalle en pseudorapidité considéré et à l'énergie moyenne $\langle E^{\gamma} \rangle$ des photons qu'elle contient. Étant donné la forte dépen-

dance de l'isolation vis-à-vis de la pseudorapidité, une catégorisation très fine en $|\eta^\gamma|$ est adoptée. Cela permet entre autre de réduire la dépendance des corrections vis-à-vis de la distribution en rapidité de l'échantillon inclusif et des différences cinématiques entre données et MC. Aucune distinction n'est faite entre photons convertis et non convertis par souci de simplicité. Établir des catégories en fonction du statut de conversion à l'avenir pourrait également se révéler être une évolution intéressante de la méthode. Les résultats des différents ajustements décrits au paragraphe précédent ainsi que les incertitudes systématiques associées sont discutées dans la section 10.4.3.

10.4.3 Résultats

Une comparaison entre les données et les différents modèles d'ajustement est présentée dans les figures 70, 71 et 72 pour trois catégories en $E_T^\gamma, |\eta^\gamma|$ afin d'évaluer le gain apporté par les nouvelles corrections dans différentes régions cinématiques. Dans chaque figure, les résultats de l'ajustement dans le cas du modèle sans corrections supplémentaires (gauche), avec translation simple ajustée (droite) et avec les corrections complètes μ, σ ajustées (bas), sont représentés. Les distributions des *pulls*³³ correspondants ainsi que le χ^2 par degré de liberté global de l'ajustement sont également donnés dans chaque cas. En complément, les distributions des χ^2 par degré de liberté en fonction de E_T^γ dans les trois catégories en pseudo-rapidité citées précédemment sont montrées dans la figure 73. Les distributions correspondantes pour les autres domaines en pseudo-rapidité sont disponibles dans l'appendice A.

Il est intéressant de remarquer que même si les corrections ne semblent pas indispensables pour l'ensemble de l'espace des phases, l'implémentation des deux paramètres μ et σ améliore systématiquement la description des données par le modèle d'ajustement, visible par la baisse systématique du χ^2 par degré de liberté. En particulier, les régions à bas $E_T^\gamma, |\eta^\gamma|$ ne semblent pas nécessiter de corrections particulières tandis que les fuites supplémentaires moyennes μ augmentent considérablement à plus haut $E_T^\gamma, |\eta^\gamma|$. Les corrections liées aux fluctuations des fuites résiduelles (σ) deviennent pertinentes dans les bouchons, particulièrement pour $E_T^\gamma > 125$ GeV. Cela est probablement relié au fait que la statistique (et donc la sensibilité de l'ajustement) augmente considérablement à partir de cette valeur, à cause du seuil en impulsion transverse du premier *trigger* photon sans facteur de *prescale* utilisé pour l'acquisition des données en 2012 ($E_T^\gamma > 120$ GeV).

Par ailleurs, les distributions représentant la fraction de photons immédiats (ou pureté) dans l'intervalle d'isolation $[-2 \text{ GeV}, 4 \text{ GeV}]$ après ajustement, couramment utilisées dans les analyses photons d'ATLAS, sont représentées dans la figure 74 en fonction de E_T^γ pour les trois catégories en pseudo-rapidité mentionnées précédemment. Les résultats des trois modèles d'ajustement considérés sont présentés pour comparaison. Comme pour les distributions de χ^2 , les distributions correspondant aux autres domaines en pseudo-rapidité sont disponibles dans l'appendice A. On rappelle ici que la contamination des régions de contrôle par les photons immédiats n'est pas prise en compte dans le calcul de pureté, les chiffres dévoilés ici visant uniquement à comparer les différents modèles d'ajustement et non pas à déterminer la pureté absolue des échantillons. Les corrections de fuite moyennes ont un impact non négligeable sur la pureté, de l'ordre de quelques pourcents, tandis que les corrections des fluctuations ne l'affectent pas significativement.

³³Les *pulls* sont définis dans chaque intervalle d'énergie d'isolation comme $\frac{N^{\text{data}} - N^{\text{model}}}{\sqrt{N^{\text{data}}}}$.

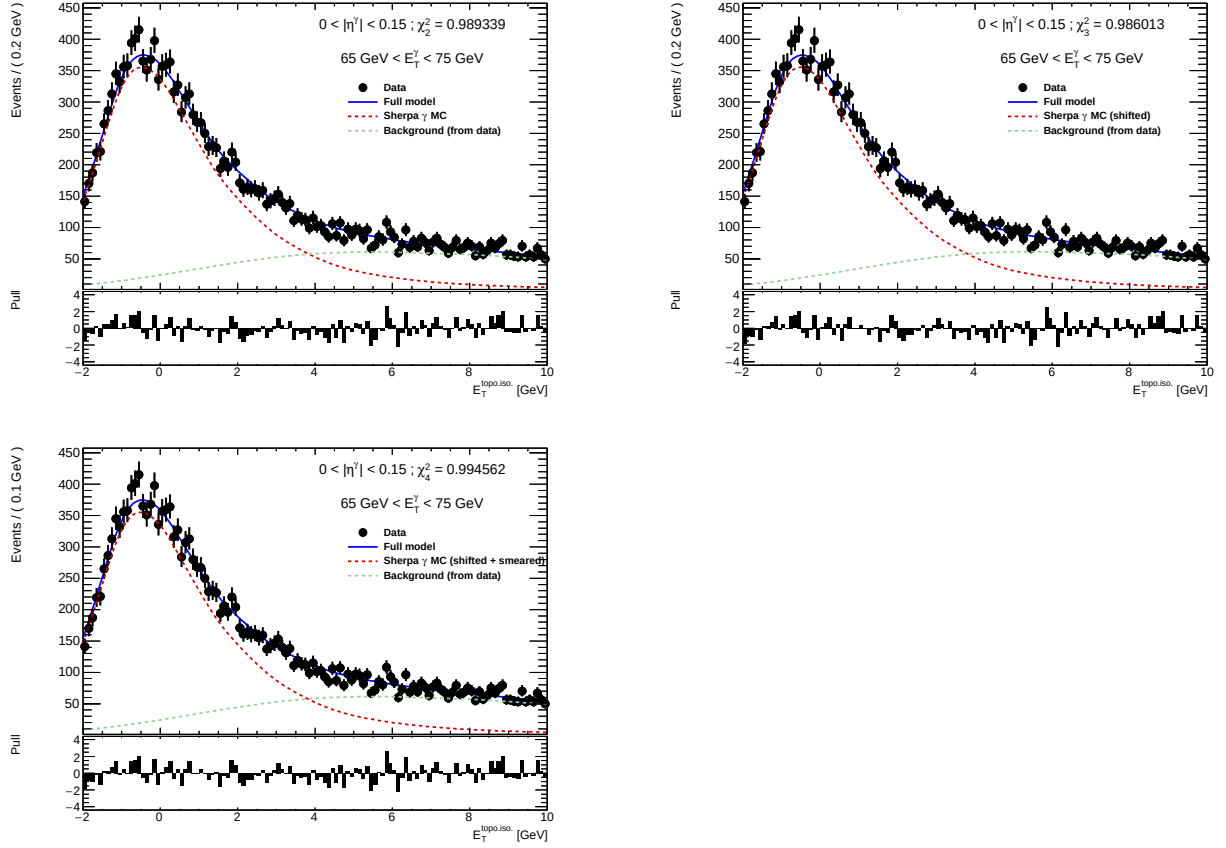


FIG. 70 – Comparaison entre données et modèle d'ajustement pour $0 < |\eta^\gamma| < 0.15$, $65 \text{ GeV} < E_T^\gamma < 75 \text{ GeV}$. Les résultats issus des modèles utilisant comme pdf photon SHERPA MC (gauche), SHERPA MC avec correction des fuites moyennes (droite) et SHERPA MC avec corrections des fuites moyennes et des fluctuations (bas) sont représentés. Le χ^2 par degré de liberté ainsi que les distributions des *pulls* correspondantes sont également représentés pour faciliter la comparaison. Dans cet exemple, les différentes corrections ne sont pas nécessaires afin d'obtenir une bonne description des données [201].

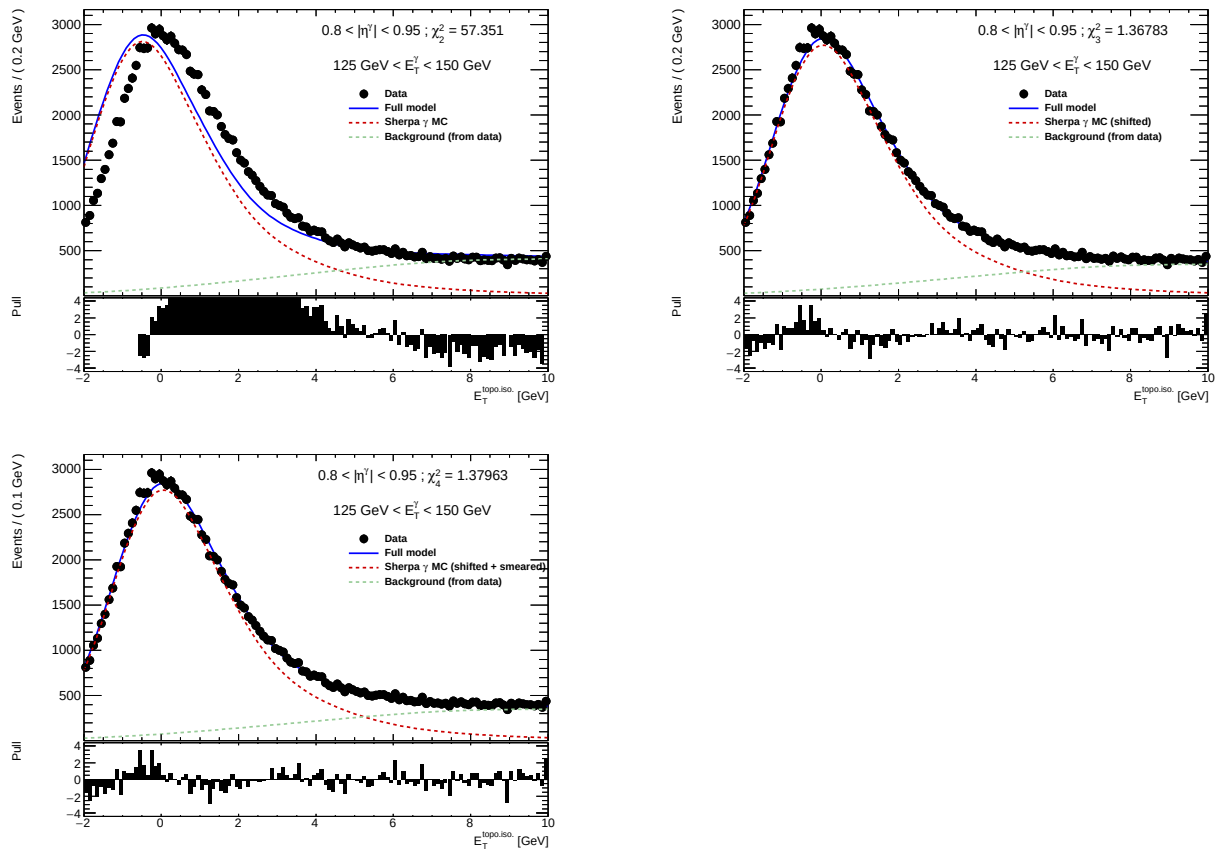


FIG. 71 – Figure similaire à la figure 70 pour $0.8 < |\eta^\gamma| < 0.95$, $125 \text{ GeV} < E_T^\gamma < 150 \text{ GeV}$. Dans cet exemple, les corrections de fuite moyennes améliorent significativement la description des données [201].

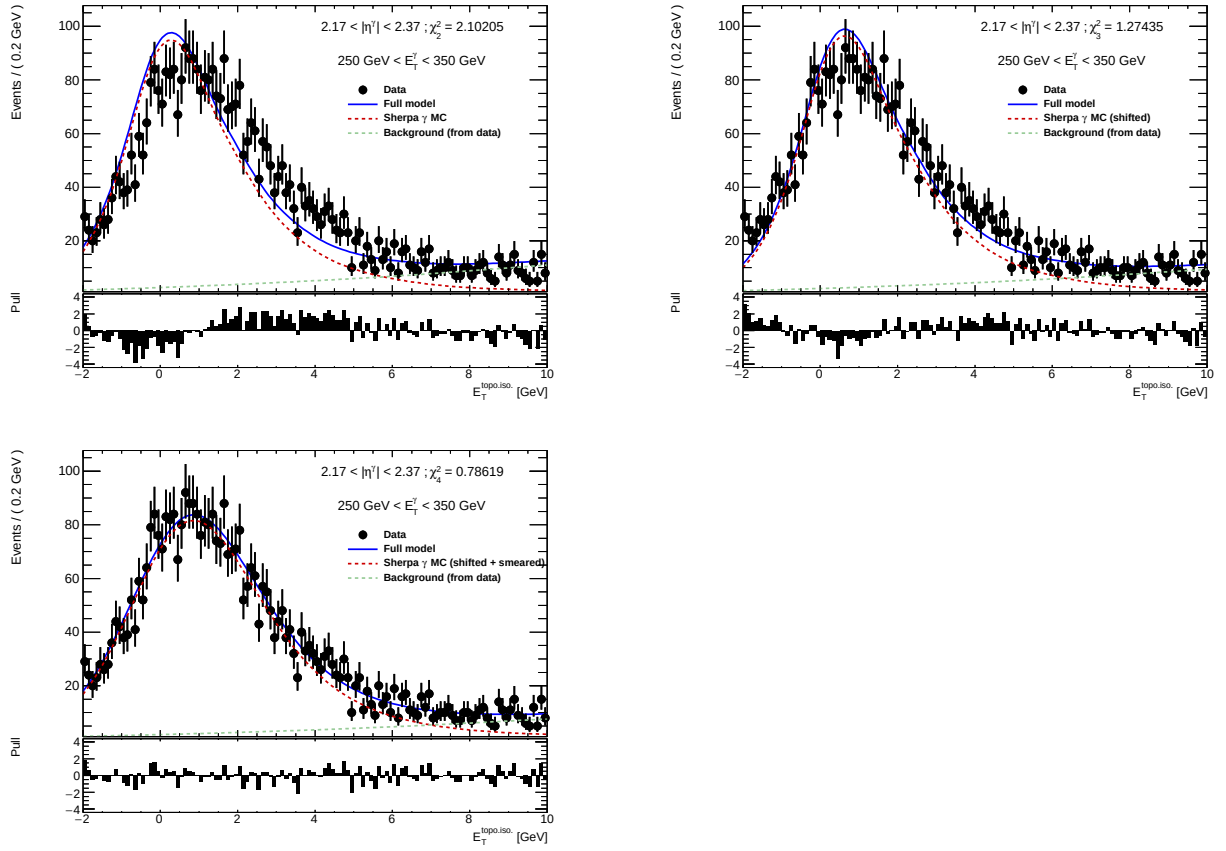


FIG. 72 – Figure similaire à la figure 70 pour $2.17 < |\eta^\gamma| < 2.37$, $250 \text{ GeV} < E_T^\gamma < 350 \text{ GeV}$. Dans cet exemple, à la fois les corrections de fuite moyennes et des fluctuations associées améliorent significativement la description des données [201].

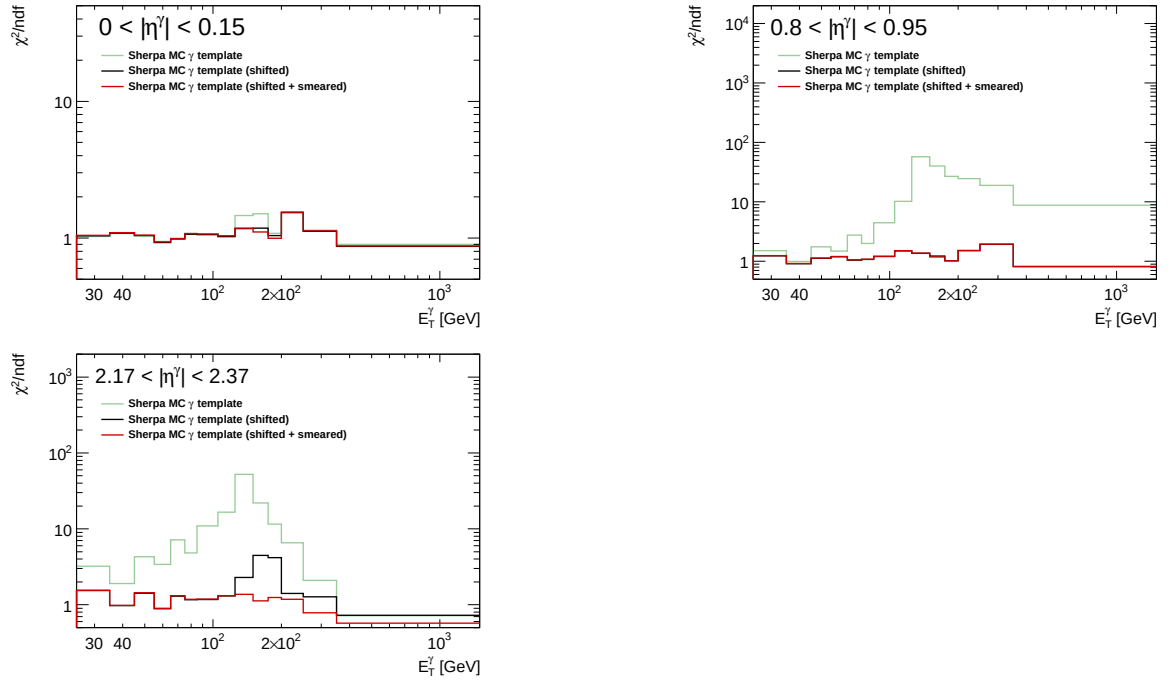


FIG. 73 – χ^2 par degré de liberté dans le cas de différentes comparaisons données/modèles d'ajustement en fonction de E_T^γ dans $0 < |\eta^\gamma| < 0.15$ (gauche), $0.8 < |\eta^\gamma| < 0.95$ (droite) et $2.17 < |\eta^\gamma| < 2.37$ (bas). Les résultats issus des modèles utilisant comme pdf photon SHERPA MC (vert), SHERPA MC avec correction des fuites moyennes (noir) et SHERPA MC avec corrections des fuites moyennes et des fluctuations (rouge) sont représentés [201].

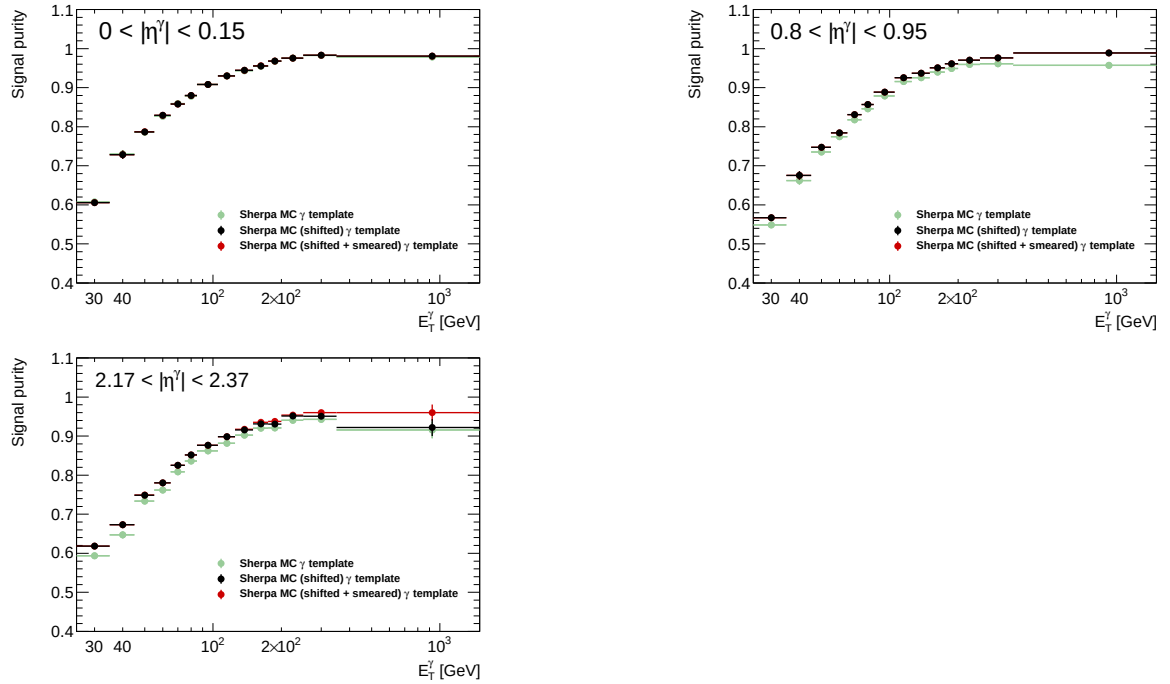


FIG. 74 – Figure similaire à la figure 73 pour la pureté des échantillons TIGHT considérés. Les barres d'erreurs sont statistiques uniquement et la contamination des régions de contrôle n'est pas prise en compte dans le calcul [201].

Ainsi, les nouvelles corrections affectent significativement la détermination de la pureté des échantillons de photons immédiats dans le cas de l'utilisation d'une méthode type *template fit* pour soustraire le bruit de fond résiduel. Une telle méthode est mise en œuvre par exemple dans l'analyse du $H \rightarrow \gamma\gamma$ [188] ou dans les recherches à plus haute masse dans le spectre des di-photons [198]. Elle est également utilisée dans l'analyse principale présentée dans la partie IV. De plus, un domaine plus élevé en E_T^γ devrait être atteint durant la seconde prise de données grâce au passage à une énergie de collision de $\sqrt{s} = 13$ TeV. Ce dernier élément renforce l'idée que de nouvelles corrections d'isolation doivent absolument être calculée pour le *run 2* afin d'obtenir une description satisfaisante de l'isolation calorimétrique dans tout l'espace des phases.

Finalement, les distributions des paramètres de corrections μ et σ sont montrées dans les figures 75 et 76 pour les trois catégories en pseudo-rapidity citées en exemple dans les paragraphes précédents. Les distributions correspondantes pour les autres domaines de pseudo-rapidity sont disponibles dans l'appendice A. Dans chaque catégorie en E_T^γ , $|\eta^\gamma|$, les μ et σ sont assignés à la valeur moyenne de l'énergie totale calibrée E^γ des photons dans la catégorie en question. L'erreur (symétrique) associée à cette simplification est prise comme l'écart-type de la distribution en E^γ dans la catégorie. Les incertitudes systématiques sur les valeurs des paramètres correctifs sont représentées dans le cas du modèle complet d'ajustement (μ et σ laissés libres). Les effets de la définition de l'échantillon NON-TIGHT, estimés en variant la définition de LOOSE' (voir section 9.4), et de la statistique limitée de ce même échantillon et du MC, estimés en faisant fluctuer les pdfs correspondantes dans l'erreur statistique, ont un impact très limité sur les corrections. Par conséquent, l'unique incertitude systématique considérée est prise comme la différence symétrisée entre les corrections obtenues en utilisant les pdfs photon F_γ^{MC} issues des générateurs PYTHIA et SHERPA (voir équation 39), afin de tenir compte de l'incertitude sur la génération évoquée dans la section 10.4.1. Les χ^2 associés aux ajustements utilisant PYTHIA étant généralement moins bons que ceux utilisant SHERPA, cette évaluation est considérée comme étant conservative.³⁴

La figure 75 montre les différences entre les paramètres μ déterminés dans le cadre du modèle complet et simplifié (voir équations 39 et 42). Les deux valeurs sont compatibles dans leurs erreurs statistiques, illustrant la robustesse de la méthode. μ montre une dépendance approximativement linéaire vis-à-vis de $\langle E^\gamma \rangle$ dans le tonneau, avec une pente dépendant de la pseudorapidity, tandis qu'un modèle parabolique semble mieux adapté dans les bouchons. La figure 76 montre que l'ajustement n'est pas sensible à un σ différent de zéro dans le tonneau, contrairement aux bouchons où une dépendance en $\sqrt{\langle E^\gamma \rangle}$ après un certain seuil est observée. Les résultats ont également été représentés dans des plans en (μ, σ) dans chaque catégorie en pseudorapidity (voir appendice A), mais aucune relation simple n'a pu en être extraite.

Afin de minimiser les effets dus aux fluctuations statistiques et au choix des catégories en E_T^γ , une interpolation des paramètres μ et σ est réalisée et présentée dans la section 10.4.4, formant les résultats finals utilisables dans les analyses de données.

³⁴De meilleurs échantillons MC pour les photons devraient être disponibles pour le *run 2*, voir partie IV.

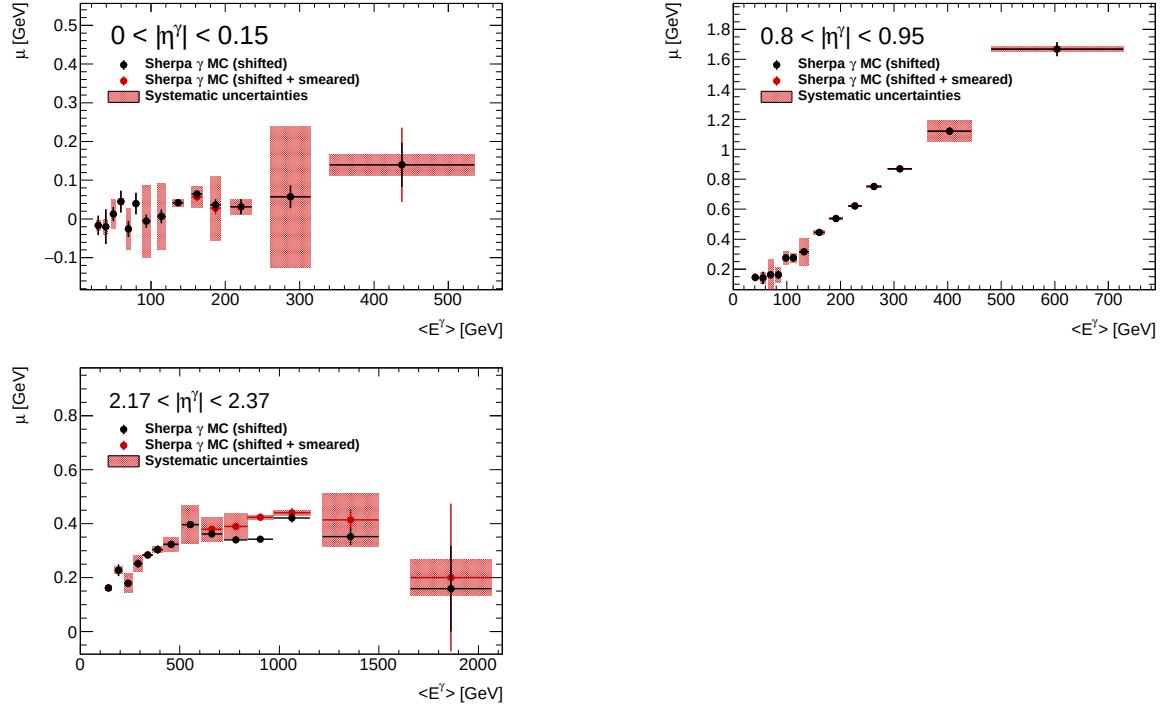


FIG. 75 – Paramètre μ déterminé par la procédure d'ajustement en fonction de l'énergie moyenne des photons dans chaque catégorie en E_T^γ dans $0 < |\eta^\gamma| < 0.15$ (gauche), $0.8 < |\eta^\gamma| < 0.95$ (droite) et $2.17 < |\eta^\gamma| < 2.37$ (bas). Les résultats des modèles utilisant comme pdf photon SHERPA MC avec corrections des fuites moyennes (noir) et des fluctuations (rouge) sont représentés. Les barres d'erreurs sont statistiques uniquement, tandis que les systématiques sont représentées pour le modèle complet par des bandes rectangulaires [201].

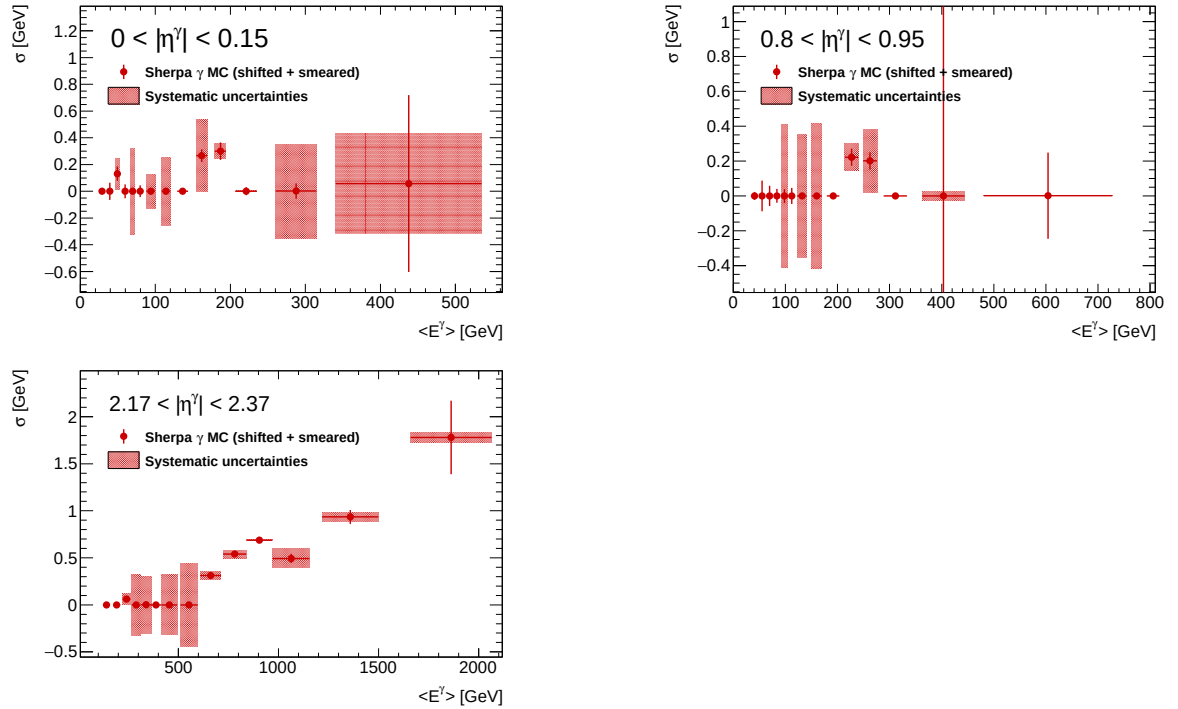


FIG. 76 – Figure similaire à la figure 75 pour le paramètre σ [201].

10.4.4 Interpolation des paramètres et corrections finales

Étant donné la faible sensibilité des données vis-à-vis du paramètre σ dans le tonneau, le modèle simplifié (voir équation 42) est adopté pour les corrections finales dans l'intervalle $|\eta^\gamma| < 1.37$ et seul μ est considéré. Dans chaque catégorie en pseudo-rapacité, la distribution de μ en fonction de l'énergie moyenne des photons est interpolée en utilisant le modèle linéaire suivant :

$$\mu_{\text{barrel}}^{|\eta^\gamma|} = a^{|\eta^\gamma|} \cdot (E^\gamma - b^{|\eta^\gamma|}), \quad (43)$$

où $a^{|\eta^\gamma|}$ et $b^{|\eta^\gamma|}$ sont les paramètres libres de l'interpolation et varient en fonction de la catégorie en pseudo-rapacité.

Au contraire, il a été montré dans la partie précédente que le paramètre σ permet d'améliorer significativement la description des données dans les bouchons. Par conséquent, le modèle complet est utilisé pour $1.56 < |\eta^\gamma| < 2.37$. Dans chaque catégorie en pseudo-rapacité, les modèles suivant sont adoptés pour μ et σ dans les bouchons :

$$\mu_{\text{end-caps}}^{|\eta^\gamma|} = c^{|\eta^\gamma|} \cdot (E^\gamma)^2 + d^{|\eta^\gamma|} \cdot E^\gamma + e^{|\eta^\gamma|}, \quad (44)$$

$$\sigma_{\text{end-caps}}^{|\eta^\gamma|} = \max(0, f^{|\eta^\gamma|} \cdot (\sqrt{E^\gamma} - g^{|\eta^\gamma|})), \quad (45)$$

où $c^{|\eta^\gamma|}$, $d^{|\eta^\gamma|}$, $e^{|\eta^\gamma|}$, $f^{|\eta^\gamma|}$ et $g^{|\eta^\gamma|}$ sont les paramètres libres de l'interpolation et varient en fonction de la catégorie en pseudo-rapacité. $g^{|\eta^\gamma|}$ correspond au seuil en $\sqrt{E_T^\gamma}$ où les corrections de fluctuations deviennent non nulles.

Les résultats d'interpolation des distributions du paramètre μ avec la bande d'erreur finale associée, incluant incertitudes statistiques et systématiques, sont montrés dans la figure 77 pour les catégories en pseudo-rapacité citées dans les paragraphes précédents. Les figures pour toutes les catégories en pseudo-rapacité sont disponibles dans l'appendice A. Les distributions équivalentes pour le paramètre σ sont montrées dans la figure 78 pour $1.56 < |\eta^\gamma| < 1.64$ (gauche), $1.72 < |\eta^\gamma| < 1.81$ (droit) et $2.17 < |\eta^\gamma| < 2.37$ (bas). Les figures correspondantes pour toutes les catégories en pseudo-rapacité sont disponibles dans l'appendice A.

Les corrections finales détaillées dans cette section doivent être appliquées en plus des corrections présentées dans les sections 10.3.2 et 10.3.3 pour les distributions MC. Les pentes du modèle d'interpolation linéaire utilisé pour μ dans le tonneau vont de 0 à 0.3%, soit environ un facteur 10 plus faible que les corrections de fuite basées sur la simulation (voir section 10.3.3). Les ordonnées à l'origine sont souvent légèrement positives mais systématiquement plus faibles que celles de la section 10.3.3. Elles montrent également des variations plus importantes et leur origine physique reste à éclaircir. Dans les bouchons, les dépendances sont plus complexes et moins bien comprises, ce qui se reflète souvent par des incertitudes plus importantes sur les paramètres μ et σ . Cela peut se comprendre en remarquant que les bouchons sont les régions recevant les photons de plus haute énergie ($E^\gamma = E_T^\gamma \cdot \cosh \eta^\gamma$). Par ailleurs, ils présentent des tailles de cellules variables en fonction de la pseudo-rapacité, ce qui pourrait expliquer

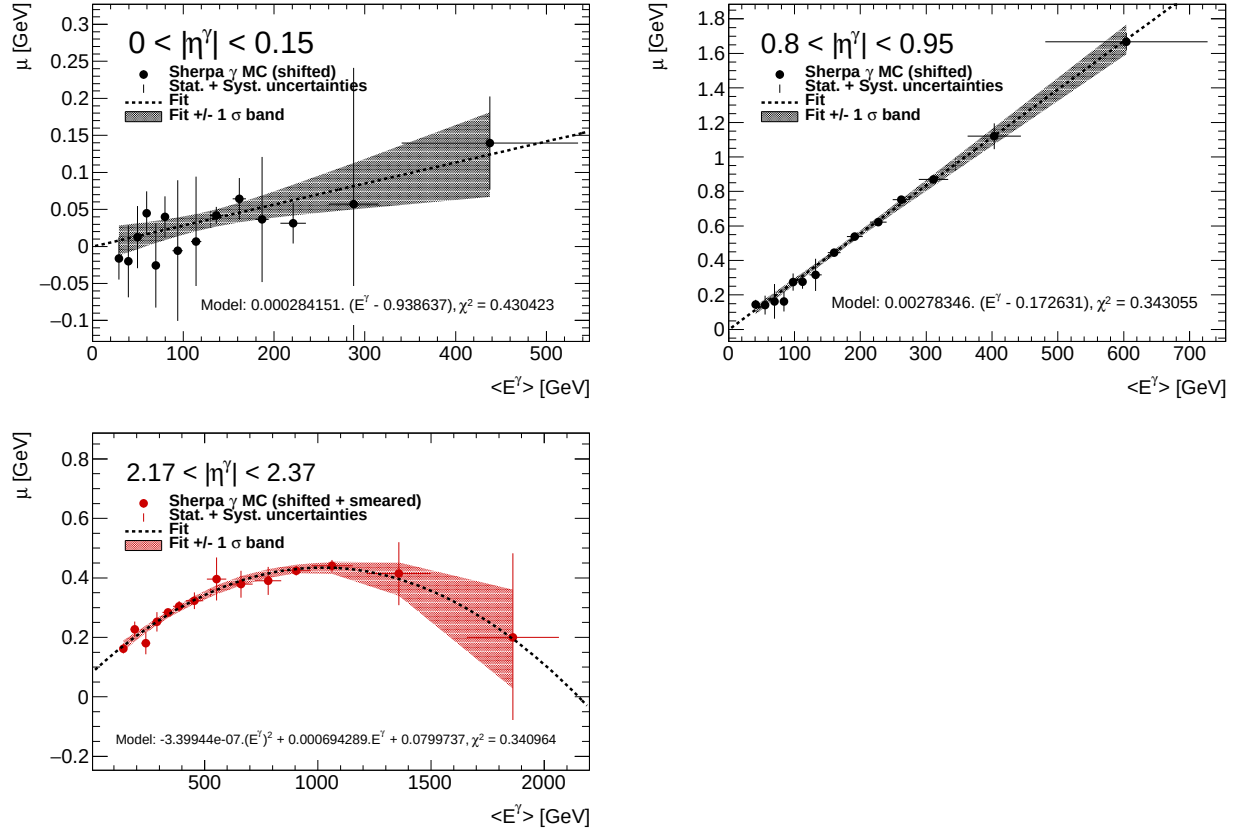


FIG. 77 – Distribution du paramètre de correction μ en fonction de l'énergie moyenne $\langle E_T^\gamma \rangle$ dans chaque catégorie en E_T^γ de $0 < |\eta^\gamma| < 0.15$ (gauche), $0.8 < |\eta^\gamma| < 0.95$ (droite) et $2.17 < |\eta^\gamma| < 2.37$ (bas), obtenues lors de l'utilisation du modèle simplifié (gauche, droite) et du modèle complet (bas). Les barres d'erreurs représentent les incertitudes finales et incluent statistique et systématique. Le χ^2 par degré de liberté associé à l'interpolation est également montré [201].

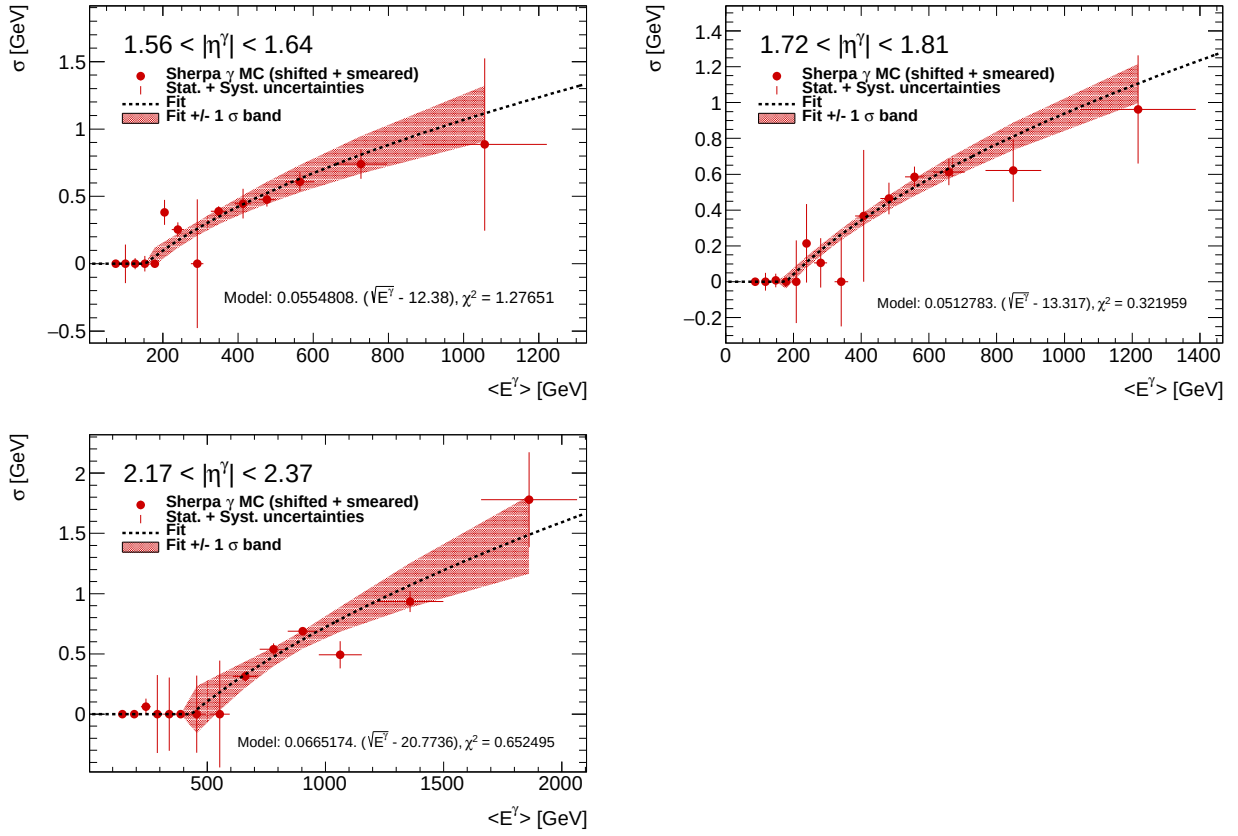


FIG. 78 – Figure similaire à la figure 77 pour σ dans $1.56 < |\eta^\gamma| < 1.64$ (gauche), $1.72 < |\eta^\gamma| < 1.81$ (droit) et $2.17 < |\eta^\gamma| < 2.37$ (bas) obtenu lors de l'utilisation du modèle complet [201].

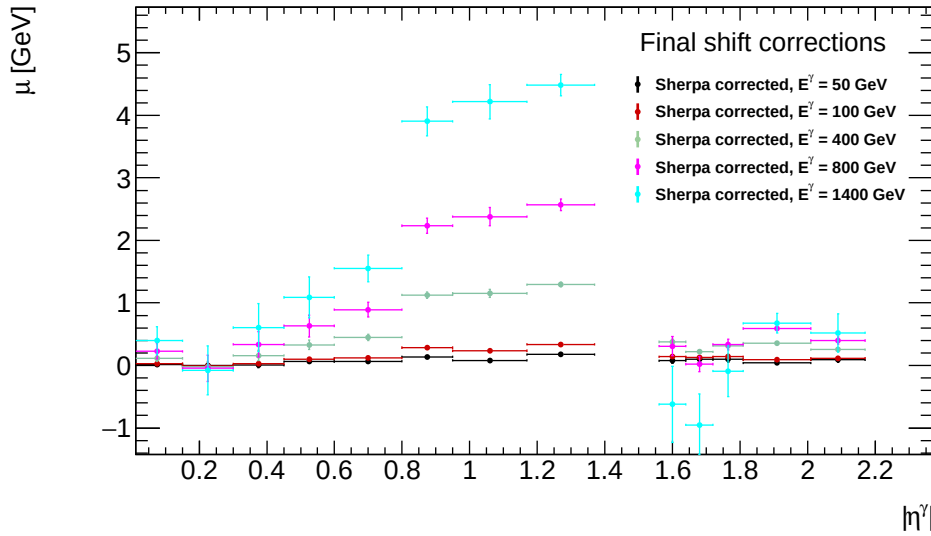


FIG. 79 – Distribution du paramètre μ en fonction de $|\eta^\gamma|$ pour des photons de différentes énergies. Les barres d'erreurs horizontales montrent les largeurs des catégories en $|\eta^\gamma|$ tandis que les barres d'erreurs verticales représentent l'incertitude finale citée pour μ . Pour $E^\gamma = 800$ et 1400 GeV, les valeurs sont parfois extrapolées à partir de mesures réalisées à plus basse énergie grâce au modèle d'interpolation (voir texte) [201].

en partie les non-linéarités observées. Les χ^2 associés aux interpolations finales sont cependant toujours aux alentours de 1, montrant le bon accord général entre les modèles paramétriques choisis et les dépendances observées.

Les résultats finals ont été implémentés dans un code C++ et mis à disposition de la collaboration. Ils permettent d'obtenir les corrections d'isolation basées sur les données pour les photons du MC en fonction de leur énergie calibrée et pseudo-rapacité. Les incertitudes associées sont également disponibles. Pour μ (tonneau + bouchons), les barres d'erreurs montrées précédemment sont utilisées (voir figure 77). Pour σ (bouchons), deux variations sont réalisées. La première consiste à rajouter à l'énergie d'isolation corrigée une variable aléatoire correspondant à une gaussienne centrée en 0 et ayant une largeur correspondant à l'incertitude totale sur σ (voir figure 78), tandis que la seconde renvoie l'énergie d'isolation obtenu dans le cadre du modèle simplifié (corrections de largeur désactivées). Cela permet d'être conservateur sur le paramètre σ , dont la physique sous-jacente est moins bien comprise que pour μ . L'utilisation de ces corrections est désormais recommandée par le groupe de performance des photons d'ATLAS pour les analyses du *run 1* à $\sqrt{s} = 8$ TeV. Elles ont été implémentées dans deux analyses actuellement en cours avec ces données ($\gamma+c$ et $\gamma+b$). La méthode de correction a par ailleurs été reproduite par le groupe de performance pour les données du *run 2* et est actuellement utilisée pour l'évaluation des incertitudes systématiques liées à l'efficacité d'isolation des photons (voir section 11).

Les catégories en pseudo-rapacité étant déjà particulièrement fines, aucune interpolation supplémentaire n'est réalisée en fonction de cette variable. Cependant, afin d'avoir une idée de la dépendance en $|\eta^\gamma|$ des corrections, le paramètre μ est représenté en fonction de $|\eta^\gamma|$ pour différentes valeurs de E^γ dans la figure 79. Le code C++ présenté dans le paragraphe ci-dessus est employé. Sur le plus long terme, ces résultats pourraient permettre de mieux comprendre les problèmes de modélisation du développement latéral des gerbes électromagnétiques dans ATLAS.

Dans la section 11, les évolutions apportées au traitement des photons pour la seconde prise de données et quelques opinions personnelles sur le sujet sont abordées.

11 Évolutions pour la seconde prise de données

11.1 État des lieux pour les premières données à 13 TeV

Le LHC a été arrêté deux ans entre 2013 et 2015 afin de subir un certain nombre de travaux de maintenance et de mises à niveau, indispensables à l'augmentation de l'énergie de collision à $\sqrt{s} = 13$ TeV (voir section 4.4). Un total de 3.87 fb^{-1} de données de collisions a pu être collecté à $\sqrt{s} = 13$ TeV en 2015. Les changements apportés aux différents paramètres de collisions (paquets espacés de 25 ns au lieu de 50 ns, etc) ainsi qu'au détecteur ATLAS (ajout de l'IBL, etc, voir section 5) ont nécessité de faire évoluer quelques éléments dans la chaîne de traitement des photons. Un aperçu des différentes évolutions est donné ci-dessous :

- ◇ **Reconstruction** : le traitement des photons convertis a changé considérablement. Désormais, l'algorithme de reconstruction classe directement chaque *cluster* électromagnétique dans trois catégories lors de son exécution (photon, électron ou ambiguïté), ce qui évite de procéder à la phase de récupération décrite brièvement dans la section 7. Par ailleurs, l'ajout de l'IBL dans le détecteur interne (voir section 5.3.1) ainsi que les changements apportés au mélange gazeux contenu dans les tubes du TRT (voir section 5.3.3) ont nécessité une modification importante des critères mis en œuvre afin d'assurer la qualité des traces et leur bonne association aux dépôts calorimétriques.
- ◇ **Calibration** : la chaîne de calibration est inchangée. Les mesures principales effectuées à 7 et 8 TeV vont être refaites à 13 TeV, incluant la mesure de quantité de matière dans le détecteur interne, modifiée par l'IBL. Des résultats préliminaires sont déjà disponibles. Il est également prévu d'étudier plus en détail l'énergie collectée dans la troisième couche du calorimètre, qui devrait recevoir une partie plus conséquente de l'énergie des photons et des électrons à $\sqrt{s} = 13$ TeV.
- ◇ **Identification** : la chaîne d'identification est inchangée. Les mesures principales effectuées à 7 et 8 TeV vont être refaites à 13 TeV afin de ré-optimiser les sélections LOOSE et TIGHT. Des résultats préliminaires sont déjà disponibles. Les *fudge factors* ont été ré-estimés et diffèrent très peu de ceux évalués à 8 TeV, qui restent appliqués par défaut pour le moment. La sélection TIGHT a nécessité plusieurs ajustements suite aux changements dans la reconstruction des photons convertis mentionnée ci-dessus. Enfin, la sélection LOOSE' employée pour la définition de régions de contrôle enrichies en bruit de fond est pour le moment inchangée par rapport à la sélection TIGHT.
- ◇ **Isolation** : un effort important d'harmonisation a été réalisé concernant les variables et les coupures d'isolation afin de faire converger les différents critères utilisés dans les analyses de données. Des outils ont été développés pour calculer et corriger l'isolation calorimétrique et l'isolation des traces des photons de la même façon pour l'ensemble des analyses. Différentes combinaisons de coupures sur les variables d'isolation sont désormais officiellement recommandées par la collaboration, pour lesquelles des mesures d'efficacité dans les données sont réalisées. Des

scale factors similaires à ceux calculés dans le cadre de l'identification sont fournis afin d'améliorer l'accord données/MC. Quelques résultats préliminaires sont déjà disponibles. Les incertitudes systématiques sur les *scale factors* sont évaluées entre autres à l'aide d'une méthode très similaire à celle présentée dans la section 10.4, développée pendant les années de thèse.

Dans la section 11.2, quelques réflexions personnelles sur la chaîne actuelle de traitement des photons dans ATLAS sont présentées.

11.2 Quelques réflexions personnelles

Le traitement des photons dans ATLAS est globalement très satisfaisant. Cette réussite est illustrée par l'incertitude systématique très faible obtenue sur la mesure de la masse du boson de Higgs dans le canal di-photon à la fin de la première prise de données ($\simeq 0.22\%$) [208]. Néanmoins, il est tentant de mentionner quelques axes de discussion intéressants afin de préparer la suite de l'exploitation des données du LHC et même d'autres accélérateurs hadroniques futurs de très haute énergie. De façon évidente, ces remarques ne sont en aucun cas exhaustives et se réduisent à des observations probablement limitées issues de ma courte expérience avec les photons dans ATLAS.

L'intégralité de la chaîne de reconstruction/calibration/identification actuelle repose sur l'utilisation de *clusters* électromagnétiques de taille fixe. Cela peut sembler discutable sachant qu'il est désormais certain que le développement latéral des gerbes est mal reproduit dans la simulation. En effet, les *clusters* de taille fixe entraînent des corrections importantes d'énergie dues aux pertes latérales, qui dépendent a priori fortement de la largeur des gerbes. Or, la calibration repose sur un algorithme MVA entraîné sur des échantillons simulés (voir section 8.2) et l'identification est presque entièrement liée à la forme des gerbes électromagnétiques (voir section 9.1). Un autre exemple très approprié ici est probablement le calcul de l'isolation calorimétrique, où les imprécisions de la simulation nécessitent le calcul de corrections supplémentaires basées sur les données afin de décrire correctement la région à haute impulsion transverse (voir section 10.4).

Les études engagées afin de pallier le problème ont principalement mené à la détermination des *fudge factors*. Ceux-ci permettent de corriger efficacement les variables discriminantes utilisées pour l'identification dans la simulation, et finalement les incertitudes sur l'identification des photons à 7 et 8 TeV, estimées directement à partir des données, restent très réduites (voir figures 60 et 61, section 9.3). Les *fudge factors* sont également appliqués dans le cadre de la calibration et les variables de l'algorithme MVA sont choisies de manière à rester largement indépendantes de la description latérale précise des gerbes. La calibration est par ailleurs validée de manière très satisfaisante à l'aide de plusieurs méthodes et l'étape de correction in-situ permet d'absorber les différences résiduelles (voir section 8.4).

Les *fudge factors* fonctionnent donc bien pour le moment, cependant ils ne fournissent pas de réponses finales au problème. En effet, ils ne permettent pas de corriger directement la forme des gerbes, ce qui pose problème par exemple dans le cas des corrections de fuite de l'isolation calorimétrique (voir section 10.4). D'autres corrections sont alors nécessaires, sans possibilité de tester leur compatibilité avec les *fudge factors* de manière entièrement satisfaisante. Les *fudge factors* simplifient également grande-

ment le problème : les huit variables d'identification sont corrigées indépendamment alors qu'elles présentent certainement des corrélations.

Une alternative (de pensée) consisterait alors à utiliser uniquement les *clusters* topologiques pour le traitement des électrons et des photons. Ceux-ci présentent l'avantage de ne pas avoir de tailles fixes, et donc sont plus sensibles aux queues de distribution en énergie. Ils pourraient permettre d'étudier plus précisément les différences de morphologie entre les gerbes réelles et simulées et pourquoi pas, de corriger directement les dépôts calorimétriques *cluster* par *cluster*. Enfin, cela permettrait d'unifier élégamment la reconstruction entre les jets hadroniques et les photons/électrons en rendant en principe possible la détermination d'une échelle et d'une résolution en énergie commune dans le calorimètre électromagnétique, quelle que soit la particule sous-jacente. Cela pourrait présenter à des résultats intéressants pour la discrimination entre photon et hadron. Il semble également raisonnable d'affirmer que les différences entre les gerbes simulées et réelles mènent probablement à des effets non-compris absorbés dans des résidus comptés finalement comme incertitude systématique, à l'exemple de la calibration in-situ. La résolution du problème engendrerait donc probablement une amélioration globale conséquente (bien que difficilement prévisible) de la chaîne de traitement des photons.

Cependant, cela demanderait un travail très important sans résultat garanti, avec probablement un grand nombre de problèmes totalement nouveaux à résoudre. Par exemple, la plus grande sensibilité des *clusters* topologiques aux queues de distribution les rendent également beaucoup plus sensibles aux fluctuations de réponses du calorimètre, statistiques mais également systématiques (gain, haute-tension, ...) et l'excellente uniformité de réponse obtenue lors du *run 1* pourrait être dégradée. Par ailleurs, il est pour le moment impossible de se passer totalement des *clusters* de taille fixe à cause du système de déclenchement au niveau L1 (voir section 5.7), basé sur un algorithme de fenêtre glissante et sur des régions d'intérêt de taille fixe.

Le calcul de l'isolation calorimétrique pourrait alors constituer un bon point de départ pour commencer l'étude des *clusters* topologiques de photons et d'électrons dans ATLAS. En effet, il semble intéressant d'essayer de soustraire du cône d'isolation l'énergie contenue dans le *cluster* topologique associé au candidat photon plutôt que le *cluster* étendu de taille fixe 5×7 et d'ensuite calculer des corrections de fuite (voir section 10.3). Si cela permet de s'affranchir totalement des corrections de fuite, ce serait un signal très encourageant afin de pousser plus loin l'utilisation de *clusters* topologiques pour les photons et les électrons.

Dans une optique similaire, des études visant à mettre au point un nouvel algorithme de reconstruction basé sur des “super-*clusters*”, inspiré de la chaîne de traitement mise en place par la collaboration CMS [209], sont actuellement en cours dans ATLAS.

Dans la partie IV, l'analyse de données principale effectuée durant la thèse, consistant à la mesure de la section efficace de production de paires de photons isolés dans les données prises à 8 TeV par le détecteur ATLAS, est décrite en détail. Elle constitue l'un des travaux majeurs de thèse et devrait être publiée dans le courant de l'année 2016. Elle a été réalisée dans son intégralité pendant la thèse et la quasi-totalité des résultats présentés dans la partie IV sont issus de travaux de recherches personnels. Bien qu'elle ne soit pas présentée dans la partie IV, je contribue également à la mesure de la section efficace de production d'un photon isolé dans les nouvelles données à 13 TeV ($\gamma + X$) toujours en cours [210], notamment au niveau de la sélection des données et de

la présentation des résultats finals. Une publication des résultats pour cette mesure est également prévue d'ici la fin de l'année.

Quatrième partie

Mesure de la section efficace de production de paires de photons isolés à $\sqrt{s} = 8$ TeV

Sommaire

12 Introduction	156
12.1 Enjeux de la mesure	156
12.2 Définition de l'espace fiduciel	157
12.3 Observables mesurées	159
12.4 Aperçu des résultats précédents obtenus au LHC et au TEVATRON	163
13 Échantillons de données et échantillons simulés employés dans l'analyse	169
13.1 Choix du système de déclenchement	169
13.2 Échantillons simulés	169
13.2.1 Considérations générales	169
13.2.2 Les échantillons simulés de signal ($\gamma\gamma + X$)	171
13.2.3 Les échantillons simulés de bruit de fond (γ -jet, $Z \rightarrow ee$)	173
14 Sélection de la région de signal et de régions de contrôle dans les données	175
14.1 Critères d'identification utilisés dans l'analyse	175
14.2 Définition de régions de contrôle dans les données	176
14.3 Choix et comparaison des variables d'isolation entre signal et bruits de fond	177
14.3.1 L'énergie d'isolation dans le détecteur interne	177
14.3.2 L'énergie d'isolation dans le calorimètre	179
14.4 Reconstruction du vertex associé à la paire	184
14.5 Présélection des événements et des candidats photons	184
14.6 Classement final des différents événements	187
15 Estimation du bruit de fond hadronique et électrons à partir des données	190
15.1 Considérations générales	190
15.2 Étude des corrélations entre les deux candidats	192
15.2.1 Rôle des corrélations dans l'analyse	192
15.2.2 Estimation des corrélations d'isolation à partir des régions de contrôle et des échantillons simulés	193
15.2.3 Analyse et conclusion pour le processus di-jet	195
15.2.4 Analyse et conclusion pour le processus di-photon	196
15.2.5 Analyse et conclusion pour les processus photon-jet et ee	197

15.3	La méthode du <i>template fit</i> 2D	198
15.3.1	Aperçu	198
15.3.2	Procédure de minimisation	199
15.3.3	Détail du modèle d'ajustement adopté	200
15.3.4	Corrections des nombres d'événements ajustés ($w_{ik}^{\text{TITI}} \rightarrow W_{ik}^{\text{TITI}}$)	202
15.3.5	Résumé de la procédure et décomposition du bruit de fond pour les différents spectres	205
15.4	Test et validation de la méthode sur un jeu de pseudo-données . . .	208
15.4.1	Création d'un jeu de pseudo-données	208
15.4.2	Validation de la méthode du <i>template fit</i> 2D en l'absence de bruit de fond électron	210
15.4.3	Validation de la méthode du <i>template fit</i> 2D en présence de bruit de fond électron	214
15.5	Décomposition finale du bruit de fond dans les données réelles	216
15.5.1	Échantillon TITI complet (cas inclusif)	216
15.5.2	Spectres obtenus pour les différentes observables de mesure .	220
15.6	Étude des incertitudes associées	232
15.6.1	Incertitudes spécifiques à la modélisation des photons	232
15.6.2	Incertitudes spécifiques à la modélisation des jets	248
15.6.3	Incertitudes spécifiques à la modélisation des électrons	254
15.6.4	Autres incertitudes	259
15.6.5	Résultats finals	266
15.7	Comparaison des résultats avec une méthode de soustraction alternative	268
15.7.1	Comparaison des chiffres bruts	268
15.7.2	Étude des corrélations entre les incertitudes systématiques . .	268
15.7.3	Comparaison finale des résultats en tenant compte des corrélations	272
16	Déconvolution des effets de détecteurs et corrections des inefficacités	275
16.1	Considérations générales	275
16.2	Matrices de réponses et définition de facteurs de correction catégorie par catégorie	276
16.3	Identification et analyse des différents effets conditionnant les facteurs de correction	278
16.3.1	Quelques définitions	278
16.3.2	Corrections d'acceptance	282
16.3.3	Efficacités des coupures de qualité	288
16.3.4	Factorisation des différents termes	292
16.3.5	Efficacité du système de déclenchement	297
16.4	Étude de l'énergie transverse d'isolation définie au niveau particule .	302
16.5	Méthode de déconvolution itérative	308
16.5.1	Aperçu	308
16.5.2	Test de la procédure et comparaison avec d'autres méthodes de déconvolution	310
16.6	Incertitudes sur la mesure finale	315
16.6.1	Classement et propagation des différentes incertitudes	315

16.6.2	Incertitudes corrélées	316
16.6.3	Incertitudes sur la soustraction du bruit de fond	320
16.6.4	Incertitudes de reconstruction	324
16.6.5	Incertitudes factorisées	329
16.6.6	Incertitudes statistiques	331
16.6.7	Bilan	333
17	Comparaison des mesures aux prévisions théoriques et conclusion	336
17.1	Prédictions théoriques	336
17.1.1	Considérations générales	336
17.1.2	SHERPA 1.4.0 (<i>ME+PS merged at LO</i>) et PYTHIA 8 (<i>LO+PS</i>)	337
17.1.3	HERWIG 7 et SHERPA 2.3 α (<i>ME+PS merged at NLO</i>) . .	338
17.1.4	DIPHOX (ordre fixe en α_s au NLO)	340
17.1.5	2γ NNLO et MCFM (ordre fixe en α_s au NNLO)	347
17.1.6	ResBos et 2γ RES (ordre fixe en α_s au NLO/NNLO avec re- sommation au NNLL)	348
17.2	Comparaisons aux mesures expérimentales	350
17.2.1	SHERPA 1.4.0 (<i>ME+PS merged at LO</i>) et PYTHIA 8 (<i>LO+PS</i>)	350
17.2.2	HERWIG 7 et SHERPA 2.3 α (<i>ME+PS merged at NLO</i>) . .	355
17.2.3	DIPHOX et ResBos (ordre fixe NLO en α_s)	358
17.2.4	2γ NNLO, MCFM et 2γ RES (ordre fixe NNLO en α_s)	361
17.3	Conclusion	363

12 Introduction

12.1 Enjeux de la mesure

La mesure de la section efficace de production de paires de photons isolés à $\sqrt{s} = 8 \text{ TeV}$ ³⁵ [211] est réalisée à partir de l'ensemble des événements³⁶ issus de collisions proton-proton enregistrés par le détecteur ATLAS pendant l'année 2012 (voir section 4.3) et menant à au moins deux photons dans l'état final ($\gamma\gamma + X$). Les trois enjeux principaux de cette mesure ont déjà été abordés dans la section 3.5 et sont rappelés brièvement ci-dessous :

◇ **Test de précision du modèle standard de l'interaction forte (QCD).**

La QCD joue un rôle particulier dans le modèle standard grâce à deux propriétés étonnantes, le confinement et la liberté asymptotique (voir section 3.1). À ce titre, il est important de tester sa véracité à la fois dans le domaine perturbatif (calculs perturbatifs à ordre fixe) et non-perturbatif (PDFs, modèles de fragmentation, ...). La question de la validité de la factorisation entre les deux domaines est également pertinente [93]. La section efficace de production de paires de photons isolés au LHC est presque exclusivement constituée de processus QCD et très fortement affectée par les corrections d'ordre supérieur en α_s (voir section 3.5), elle constitue ainsi un banc de test idéal pour la QCD perturbative. Elle inclut des photons de fragmentation et présente une certaine dépendance vis-à-vis des PDFs [109], et donc permet aussi dans une moindre mesure d'obtenir des résultats intéressants dans le domaine non-perturbatif (contraintes sur les PDFs et les fonctions de fragmentation).

◇ **Amélioration des prédictions MC pour la production QCD $\gamma\gamma + X$, bruit de fond important pour les recherches de nouvelle physique.**

À l'image de l'excitation suscitée dans la communauté scientifique par le léger excès observé à la fois par les collaborations ATLAS et CMS dans le spectre en masse des paires de photons dans les premières données à $\sqrt{s} = 13 \text{ TeV}$ du LHC [8, 9], la production de di-photons à très haute énergie est examinée très attentivement dans le cadre de nombreux modèles de nouvelle physique (Higgs composite, dimensions supplémentaires de l'espace, etc). Le bruit de fond principal et irréductible de ces recherches est constitué par les modes de production QCD, qu'il faut donc démontrer comme étant sous contrôle. Ainsi, les mesures directes de section efficaces $\gamma\gamma + X$ sont fondamentales pour comparer les prédictions MC aux données réelles et les corriger si besoin (*K-factors*, *tuning*, etc). Pour des raisons similaires, ces mesures sont également très utiles aux études de précision relatives au boson de Higgs via le canal de désintégration $H \rightarrow \gamma\gamma$.

◇ **Tests des procédés de resommation employés en QCD.** Fortement affectée à la fois par les ordres supérieurs (voir ci-dessus) et les émissions de gluons à basse énergie dans l'état initial (instabilités infrarouges), la production QCD

³⁵Analyse actuellement en cours de revue interne par la collaboration ATLAS.

³⁶Un événement est défini dans ATLAS comme l'ensemble des données enregistrées par l'ensemble des détecteurs pendant un laps de temps donné, identifié à un croisement de faisceaux particulier.

de paires de photons isolés constitue le banc de test idéal pour les prédictions combinant les corrections d'ordre supérieur en α_s et un procédé de resommation QCD à tous les ordres (voir section 3.4). Les techniques de resommation peuvent se baser exclusivement sur la pQCD (ordre en logarithme fixé) ou sur des resommations approximatives grâce aux algorithmes de combinaison entre éléments de matrice et cascades de partons implémentés dans les générateurs MC de nouvelle génération [212].

Dans la section 12.2, l'espace fiduciel de mesure est défini. Celui-ci correspond au domaine d'intégration utilisé pour le calcul des sections efficaces théoriques hadroniques (voir section 3.3, équation 17). Il constitue un ingrédient indispensable pour pouvoir comparer la mesure aux différentes prédictions (voir section 3.4). Dans la section 12.3, les observables étudiées dans le cadre de l'analyse sont explicitées et justifiées. Finalement, la section 12.4 dresse un rapide état de l'art concernant les mesures $\gamma\gamma + X$ réalisées précédemment au LHC et au TEVATRON et leur accord avec la théorie.

12.2 Définition de l'espace fiduciel

L'espace fiduciel de mesure correspond à l'ensemble des collisions proton-proton menant à la création d'au moins deux photons dans l'état final ($\gamma\gamma + X$), sans contraintes supplémentaires sur le système additionnel X qui est le plus souvent constitué de radiations diffuses à basse énergie mais peut également inclure des jets collimatés de haute énergie ($p_T \gg 1$ GeV). Dans la suite, les deux photons (niveau généré) ou candidats photons (niveau reconstruit) formant la paire sont numérotés 1 et 2 en ordre décroissant vis-à-vis de leur impulsion transverse individuel E_T^γ , *ie.* $E_{T,1} > E_{T,2}$.

Uniquement les photons immédiats (*prompts*), c'est-à-dire issus directement de l'interaction partonique dure et donc reliés aux éléments de matrice pQCD (composante directe) ou issus de la fragmentation d'un parton et donc reliés aux fonctions de fragmentation non-perturbative (composante de fragmentation), sont considérés. Par conséquent, les photons issus de la désintégration de hadrons ou de taus (l'exemple typique étant $\pi^0 \rightarrow \gamma\gamma$) sont exclus de la mesure et traités comme un bruit de fond hadronique. Ce choix est justifié dans la section 3.5.

Dans le cas où plus de deux photons immédiats sont émis dans l'état final, seuls les deux photons de plus haute énergie transverse sont considérés. Une analyse des événements avec au moins trois photons immédiats dans l'état final ($\gamma\gamma\gamma + X$) est actuellement en cours avec les données 2012 dans ATLAS [213]. La section efficace préliminaire mesurée est plus que 200 fois plus faible que pour $\gamma\gamma + X$ et donc totalement négligeable par rapport aux précisions théoriques et expérimentales (voir section 17).

Le paragraphe précédent implique *de facto* que les contributions associées à la désintégration du boson de Higgs en deux photons ($H \rightarrow \gamma\gamma$) sont incluses dans le signal. Par ailleurs, des coupures supplémentaires sont appliquées sur la cinématique de la paire de photons sélectionnée afin de restreindre la mesure au domaine de sensibilité maximale du détecteur ATLAS. Celles-ci sont décrites ci-dessous :

◇ **Pseudo-rapidités** : $|\eta^\gamma| < 2.37$, avec l'exclusion de $1.37 < |\eta^\gamma| < 1.56$ pour cha-

cun des deux photons constituant la paire. Cette coupure est fixée par la région du calorimètre électromagnétique à argon liquide couverte par la première couche d'échantillonnage, qui présente une granularité très fine indispensable à la bonne discrimination des photons vis-à-vis des jets hadroniques (voir section 5.4). La zone d'exclusion correspond à la région de transition entre tonneau et bouchons, faiblement instrumentée. Elle a été légèrement étendue par rapport à l'analyse di-photon précédente à $\sqrt{s} = 7$ TeV [203] ($1.52 \rightarrow 1.56$) afin de s'affranchir totalement des mesures des scintillateurs du calorimètre à tuiles (voir section 5.5.1), qui dégradent significativement la résolution. Cette extension constitue une recommandation officielle du groupe de performance des électrons et des photons dans ATLAS (e/γ CP pour *combined performance*).

- ◇ **Impulsions transverses** : $E_{T,1}^\gamma > 40$ GeV ($E_{T,2}^\gamma > 30$ GeV) pour le photon de plus haute (basse) énergie parmi la paire sélectionnée. Ces valeurs sont fixées par les seuils du menu de déclenchement pour les paires de photon LOOSE (voir section 9) sans facteur de réduction (*unprescaled*) et de plus basse énergie utilisés lors la prise de données à $\sqrt{s} = 8$ TeV, correspondant à la détection de deux candidats photons LOOSE avec $E_{T,1}^\gamma > 35$ GeV et $E_{T,2}^\gamma > 25$ GeV (voir section 13.1). Une marge de 5 GeV est prise sur les seuils afin d'assurer une efficacité de déclenchement très proche de 100% (voir section 16). Par ailleurs, les deux coupures sont asymétriques afin de minimiser l'effet des instabilités infrarouges dans les calculs théoriques à ordre fixe (NLO/NNLO) [214]. Les seuils ont été augmentés par rapport à l'analyse di-photon précédente (25/22 GeV) [203], principalement à cause de l'évolution des conditions d'empilement dans les données 2012 (voir section 4.3).
- ◇ **Isolations** : $E_{T,1,2}^{\text{part.iso.}} < 11$ GeV dans le cône de taille fixe $\mathcal{C}(R = 0.4)$, *ie.* chaque photon de la paire doit avoir une énergie transverse d'isolation définie au niveau particule inférieure à 11 GeV. La nécessité d'intégrer une coupure d'isolation dans la définition de l'espace fiduciel a déjà été discutée dans la section 3.5.3.
 - Pour les échantillons de signal simulés avec cascade de partons (voir section 3.2), l'énergie d'isolation utilisée pour définir le volume fiduciel est calculée au niveau particule (aussi appelé niveau vérité ou *truth*), c'est-à-dire après hadronisation. Les muons, neutrinos et le photon lui-même sont exclus du cône $\mathcal{C}(R)$ afin d'obtenir une valeur aussi proche que possible de la variable d'isolation équivalente reconstruite à partir des données du détecteur (voir section 10.3).
 - La valeur de la coupure est désormais choisie afin de minimiser les corrections de déconvolution des effets de détecteurs entre “isolé” au niveau reconstruit et “isolé” au niveau générateur, correspondant ici à $E_{T,1,2}^{\text{part.iso.}} < 11$ GeV pour le critère d'isolation adopté au niveau reconstruit dans l'analyse (voir section 14). Cette valeur relativement élevée est obtenue notamment à cause de l'absence de calibration des *clusters* topologiques (voir section 10.3). L'étude associée sera détaillée dans la section 16. Dans l'analyse précédente, la coupure était simplement fixée à la valeur de coupure appliquée au niveau reconstruit.
 - D'autre part, $E_{T,1}^{\text{part.iso.}}$ et $E_{T,2}^{\text{part.iso.}}$ sont corrigées en moyenne de l'effet de l'événement sous-jacent en utilisant la même technique que celle utilisée dans les

données pour l'empilement et l'événement sous-jacent (basée sur l'énergie ambiante moyenne de l'événement, voir section 10.3.2) [195,196], mais en utilisant les jets reconstruits au niveau particule au lieu des jets reconstruits à partir des *clusters* topologiques. L'empilement n'est pas considéré au niveau particule afin d'obtenir une mesure indépendante des conditions de fonctionnement du LHC.

- Finalement, pour les générateurs partoniques n'incluant pas d'algorithmes de cascades de partons (DIPHOX, RESBOS, 2 γ RES, 2 γ NNLO, MCFM, voir section 17), l'énergie transverse d'isolation est calculée au niveau partonique et l'espace fiduciel est corrigé des effets non-perturbatifs en utilisant le modèle d'hadronisation et d'événement sous-jacent des générateurs SHERPA 1.4.0 et PYTHIA8 (voir sections 13 et 17).
- ◇ **Séparation entre les deux photons :** $\Delta R_{\gamma\gamma} = \sqrt{\Delta\eta_{\gamma\gamma}^2 + \Delta\phi_{\gamma\gamma}^2} > 0.4$. Ce critère est nécessaire afin d'éviter un recouvrement trop important des cônes d'isolation des deux photons (voir point précédent), suivant les recommandations présentées dans [121] déjà évoquées dans la section 3.5.3. Cette définition de $\Delta R_{\gamma\gamma}$ permet d'appliquer un critère similaire quelque soit le référentiel dans le centre de masse des partons rentrant en collision (voir section 5.2). Il est important de remarquer qu'un recouvrement partiel des deux cônes d'isolation est cependant toujours autorisé. Ce critère est inchangé comparé à l'analyse précédente.

Une sélection supplémentaire correspondant à la catégorie (*bin*) considérée est également effectuée pour l'étude des différentes observables de mesure (ex : $X < m_{\gamma\gamma} < Y$). Ces dernières sont détaillées dans la section 12.3.

12.3 Observables mesurées

En plus de la section efficace de production fiducielle totale basée uniquement sur les coupures énoncées dans la section 12.2, six spectres caractéristiques des états finals à deux photons sont mesurés dans différentes sous-catégories de l'espace fiduciel (*bins*) qui seront définies dans la section 16. Les spectres considérés sont décrits ci-dessous :

- ◇ **Masse invariante**, $m_{\gamma\gamma}$: variable principale étudiée dans le cadre de recherches de résonances dans le spectre di-photon. Par conséquent, une bonne connaissance de la distribution en $m_{\gamma\gamma}$ des événements $\gamma\gamma + X$ est utile pour l'estimation du bruit de fond QCD associé à ces recherches.
- ◇ **Impulsion transverse**, $p_{T,\gamma\gamma}$: variable sensible aux instabilités infrarouges (émission de gluons mous dans l'état initial) et aux ordres supérieurs en α_s respectivement à basse ($p_{T,\gamma\gamma} \simeq 0$) et haute valeur ($p_{T,\gamma\gamma} \gg 0$). Dans la zone intermédiaire, une certaine sensibilité à la composante de fragmentation est également attendue, particulièrement au niveau de la région appelée épaule de Guillet dans la littérature (*Guillet shoulder*, voir section 3.5.3). Finalement, une certaine sensibilité aux PDFs est également attendue.
- ◇ **Différence entre les angles d'azimut des deux photons**, $\Delta\phi_{\gamma\gamma}$: variable

sensible aux instabilités infrarouges (émission de gluons mous dans l'état initial) et aux ordres supérieurs en α_s respectivement à haute ($\Delta\phi_{\gamma\gamma} \simeq \pi$) et basse valeur ($\Delta\phi_{\gamma\gamma} \ll \pi$). Une certaine sensibilité à la composante de fragmentation est également attendue via un effet similaire à l'épaule de Guillet mais est en général peu visible dans les données [123]. Finalement, une certaine sensibilité aux PDFs est également attendue.

- ◇ **La valeur absolue du cosinus de l'angle de diffusion de la paire de photons par rapport à la direction du faisceau**, $|\cos \theta_\eta^*|$: dans le cadre d'une nouvelle résonance se désintégrant en paires de photons, cette variable est très sensible au spin de la nouvelle particule. Elle a servi entre autres à mesurer le spin du boson de Higgs [215, 216]. Le plus souvent calculée dans le référentiel dit de Collins-Soper (θ_{CS}^*) [217], elle est ici et pour la première fois dans une analyse di-photon calculée dans le référentiel où les deux photons composant la paire sont dos-à-dos dans le plan transverse (θ_η^*), suivant la suggestion développée dans la référence [218]. Cela permet de s'affranchir des mesures des impulsions transverses de chaque photon ($E_{T,1}^\gamma, E_{T,2}^\gamma$), qui présentent souvent une résolution limitée à bas E_T^γ , en définissant $\cos \theta_\eta^*$ uniquement à partir des mesures de pseudo-rapidité $\eta_2^\gamma, \eta_1^\gamma$, qui bénéficient au contraire d'une excellente résolution intrinsèque (voir section 5.4 et l'équation 46 ci-dessous). La région à bas E_T^γ correspondant à $p_{T,\gamma\gamma} \simeq 0$ est intéressante en particulier pour tester les différents procédés de resommation employés en QCD (voir sections 3.4). Par ailleurs, la valeur absolue est considérée afin de maximiser la statistique, la distribution de $\cos \theta_\eta^*$ attendue étant typiquement symétrique par rapport à 0.

$$\cos \theta_\eta^* = \tanh \left(\frac{\eta_1^\gamma - \eta_2^\gamma}{2} \right) \quad (46)$$

- ◇ **Nouvelle observable**, a_T : introduit dans [219], a_T correspond à la composante transverse de $p_{T,\gamma\gamma}$ par rapport à la direction $\vec{p}_{T,1}^\gamma - \vec{p}_{T,2}^\gamma$ (voir figure 80). Elle est moins sensible aux incertitudes de mesure sur $E_{T,1}^\gamma, E_{T,2}^\gamma$ et aux effets de détecteurs (inefficacités, résolution limitée) que $p_{T,\gamma\gamma}$, tout en sondant un espace des phases proche (voir figures 81 et 82). a_T permet donc en principe un meilleur accès à la région $p_{T,\gamma\gamma} \simeq 0$ sensible aux instabilités infrarouges, dont l'étude expérimentale est traditionnellement limitée par la résolution du détecteur. a_T a déjà été utilisée comme variable discriminante dans la collaboration ATLAS pour l'étude du canal $H \rightarrow \gamma\gamma$ [189, 220], elle est mesurée ici pour la première fois dans le cadre d'une analyse $\gamma\gamma + X$.

- ◇ **Nouvelle observable**, ϕ_η^* : introduit dans [218], ϕ_η^* est défini par :

$$\phi_\eta^* = \tan \left(\frac{\pi - \Delta\phi_{\gamma\gamma}}{2} \right) \sin(\theta_\eta^*). \quad (47)$$

Comme a_T , cette variable est moins sensible aux incertitudes de mesure sur $E_{T,1}^\gamma, E_{T,2}^\gamma$ et aux effets de détecteurs (inefficacités, résolution limitée) que $p_{T,\gamma\gamma}$, tout en sondant un espace des phases proche (voir figures 81 et 82). ϕ_η^* permet donc en principe un meilleur accès à la région $p_{T,\gamma\gamma} \simeq 0$ sensible aux instabilités infrarouges, dont l'étude est traditionnellement limitée par la résolution du détecteur. Il est

possible de montrer que lorsque les photons de la paire sont dos-à-dos dans le plan transverse, $\phi_\eta^* \sim a_T/m_{\gamma\gamma}$ est vérifié. Par ailleurs, ϕ_η^* présente une forte sensibilité à la composante de fragmentation à haute valeur ($\phi_\eta^* > 2$, voir figure 83), ce qui n'avait encore jamais été souligné dans la littérature. Déjà utilisé dans ATLAS dans le cadre de la mesure de la distribution en p_T du boson $Z^0 \rightarrow l^+l^-$ [221], ϕ_η^* est étudié ici pour la première fois dans une analyse $\gamma\gamma + X$.

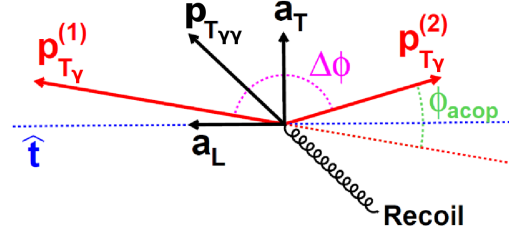


FIG. 80 – Schéma représentant a_T , $\Delta\phi_{\gamma\gamma}$ et $\phi_{\text{acop}} = \pi - \Delta\phi_{\gamma\gamma}$ dans le plan transverse à la direction des faisceaux. $\hat{t}$ est défini comme le vecteur unitaire pointant dans la direction $\vec{p}_{T,1}^\gamma - \vec{p}_{T,2}^\gamma$ [218].

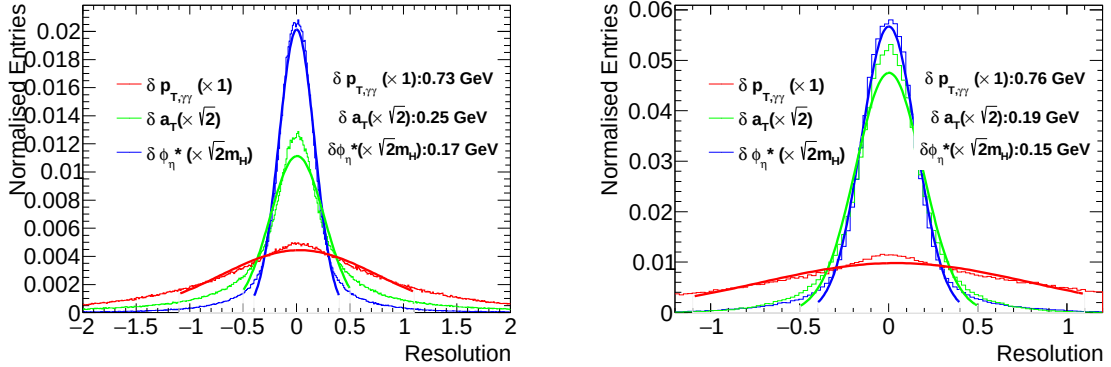


FIG. 81 – Comparaison des résolutions de $p_{T,\gamma\gamma}$, a_T et ϕ_η^* obtenues à partir de l'échantillon di-photon simulé avec le générateur SHERPA 1.4.0 (voir section 13.2) après sélection des données (gauche, voir section 14.6) et après sélection des données pour $\pi - \Delta\phi_{\gamma\gamma} < \pi/16$ (droite). Les courbes de résolution de a_T et ϕ_η^* sont multipliées respectivement par $\sqrt{2}$ et $\sqrt{2}m_H$ afin de présenter des quantités comparables, a_T représentant une composante de $p_{T,\gamma\gamma}$ et ϕ_η^* étant équivalent à $a_T/m_{\gamma\gamma}$ quand les photons sont dos-à-dos dans le plan transverse. Les résolutions de a_T et ϕ_η^* deviennent très proches pour $\pi - \Delta\phi_{\gamma\gamma} < \pi/16$ car $\phi_\eta^* \sim a_T/m_{\gamma\gamma}$ dans cette région.

Les spectres en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, $\Delta\phi_{\gamma\gamma}$ et $\cos \theta_\eta^*$ ont déjà été mesurés au TEVATRON et dans les données proton-proton à $\sqrt{s} = 7$ TeV du LHC, tandis que a_T et ϕ_η^* sont mesurés pour la première fois dans l'analyse présentée dans ce manuscrit. Dans la section 12.4, un rapide aperçu des mesures $\gamma\gamma + X$ précédentes effectuées au LHC et au TEVATRON est présenté.

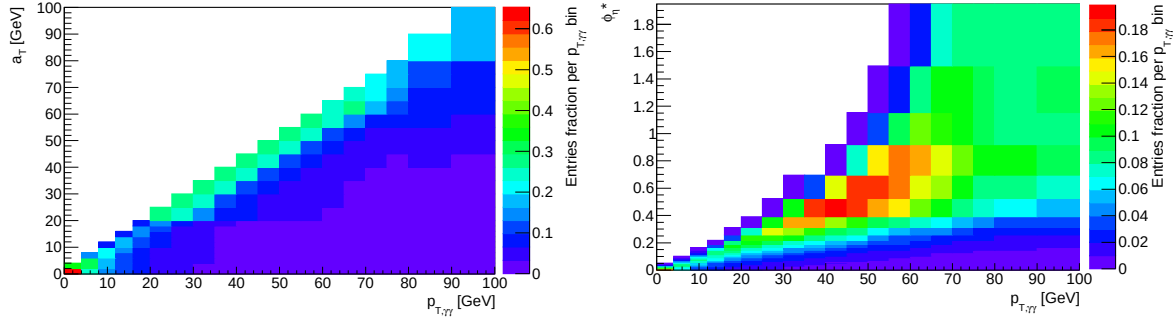


FIG. 82 – Corrélations dans le plan $p_{T,\gamma\gamma}$ - a_T et $p_{T,\gamma\gamma}$ - ϕ_η^* pour $p_{T,\gamma\gamma} \in [0, 100 \text{ GeV}]$, obtenues à partir de l'échantillon di-photon simulé avec le générateur SHERPA 1.4.0 après sélection des données (voir sections 13.2 et 14.6). Les facteurs de corrélation globaux s'élèvent respectivement à 80% et 70% environ. Le nombre d'entrées dans chaque bin 2D est normalisé par le nombre total d'événements présents dans le bin 1D en $p_{T,\gamma\gamma}$ correspondant.

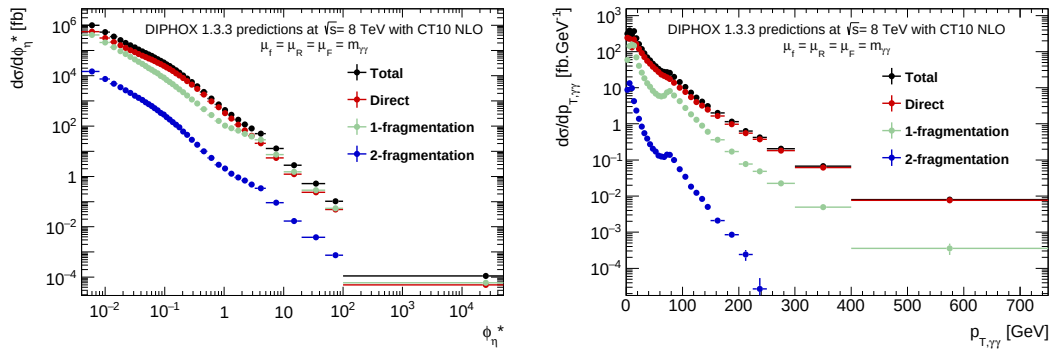


FIG. 83 – Prédications NLO en α_s pour la section efficace $\gamma\gamma + X$ générées à $\sqrt{s} = 8 \text{ TeV}$ avec le code DIPHOX v1.3.3 [111] en utilisant les PDFs CT10 NLO [124]. Les différentes prédictions sont décomposées en terme de composantes directes, 1-fragmentation et 2-fragmentations (voir section 3.5.3). Les échelles de fragmentation, factorisation et renormalisation sont fixées à $m_{\gamma\gamma}$ et la sélection fiducielle utilisée est celle présentée dans la section 12.2.

12.4 Aperçu des résultats précédents obtenus au LHC et au Tevatron

La section efficace de production $\gamma\gamma + X$ a été mesurée précédemment à haute énergie à partir de collisions proton-antiproton et proton-proton, respectivement à $\sqrt{s} = 1.96$ TeV au TEVATRON et à $\sqrt{s} = 7$ TeV au LHC par les collaborations DØ [222], CDF [223] (TEVATRON), ATLAS [203] et CMS [224] (LHC). Par ailleurs, CMS a présenté récemment une mesure préliminaire de la section efficace $\gamma\gamma + N$ jets ($N = 0, 1, 2$ et “3 ou 4”) à $\sqrt{s} = 7$ TeV [225]. Aucun résultat à $\sqrt{s} = 8$ TeV n’a été présenté pour le moment. Les mesures de DØ et CDF constituent les mesures finales du TEVATRON, tandis que celles d’ATLAS et CMS emploient l’intégralité des données prises à $\sqrt{s} = 7$ TeV en 2010-2011 (voir section 4.3). CDF utilise un jeu de données légèrement plus important que DØ (9.5 fb^{-1} vs 8.5 fb^{-1}), tandis que la statistique des données d’ATLAS et de CMS est similaire. Dans tous les cas (incluant l’analyse présentée dans ce document), la précision de la mesure est limitée par les incertitudes systématiques, sauf dans les queues de distributions où l’incertitude statistique prédomine.

Afin d’avoir une idée des différences entre les espaces des phases explorés par ces différentes mesures, les figures 84 et 86 montrent les distributions en masse invariante $m_{\gamma\gamma}$ mesurées dans les différents cas et leur comparaison avec les prédictions théoriques, extraites des références citées dans le paragraphe précédent. L’acceptance en masse est systématiquement limitée par les coupures fiducielles sur $E_{T,1}$, $E_{T,2}$, créant un pic dans les distributions en masse à environ $m_{\gamma\gamma} = 2\sqrt{E_{T,1}^{\gamma,\min} E_{T,2}^{\gamma,\min}}$.³⁷

Des informations sur les prédictions théoriques disponibles pour $\gamma\gamma + X$ ont été données dans la section 3.5, des détails complémentaires et plus exhaustifs seront apportés dans la section 17. L’énergie de collision dans le centre de masse étant bien plus élevée au LHC (1.96 TeV vs 7 TeV), les masses invariantes atteintes par les paires de photons le sont également. Cependant, la luminosité instantanée du LHC est aussi plus importante, ce qui oblige à travailler avec des seuils de déclenchement plus élevés ($E_{T,1,2}^{\gamma,\min}$). D’autre part, les PDFs diffèrent dans les deux cas, à la fois à cause de la différence d’énergie de collision (1.96 TeV vs 7 TeV) et des particules utilisées ($p\bar{p}$ vs pp) [109].

Au TEVATRON, le meilleur accord entre théorie et mesure en aval du pic cinématique est en général atteint pour les prédictions de SHERPA 1.4.0 (*ME+PS merged at LO*) [83], bien que la plupart des prédictions soient compatibles dans les incertitudes citées. L’accord de MCFM (NLO) avec les données de CDF est moins bon qu’il n’y paraît, les prédictions étant évaluées à des échelles de factorisation, fragmentation et renormalisation (voir section 3.3) correspondant à $m_{\gamma\gamma}/2$ tandis que DØ et les expériences du LHC utilisent $m_{\gamma\gamma}$ systématiquement. Au LHC, les corrections NNLO semblent nécessaires afin de décrire correctement les sections efficaces.

Par ailleurs, la région en amont du pic est dominée par des topologies particulières contenant des photons de fragmentation et/ou proches les uns des autres et reste ainsi mal décrite par les différents calculs. Cependant, les corrections NNLO permettent une nette amélioration de la situation grâce à l’apparition des processus $2 \rightarrow 4$ avec deux partons additionnels émis dans l’état final. Les distributions pour les autres observables de mesure seront examinées plus en détail dans la section 17.

³⁷ $m_{\gamma\gamma}^2 = 2E_{T,1}^{\gamma} E_{T,2}^{\gamma} [\cosh(\eta_1^{\gamma} - \eta_2^{\gamma}) - \cos(\phi_1 - \phi_2)]$.

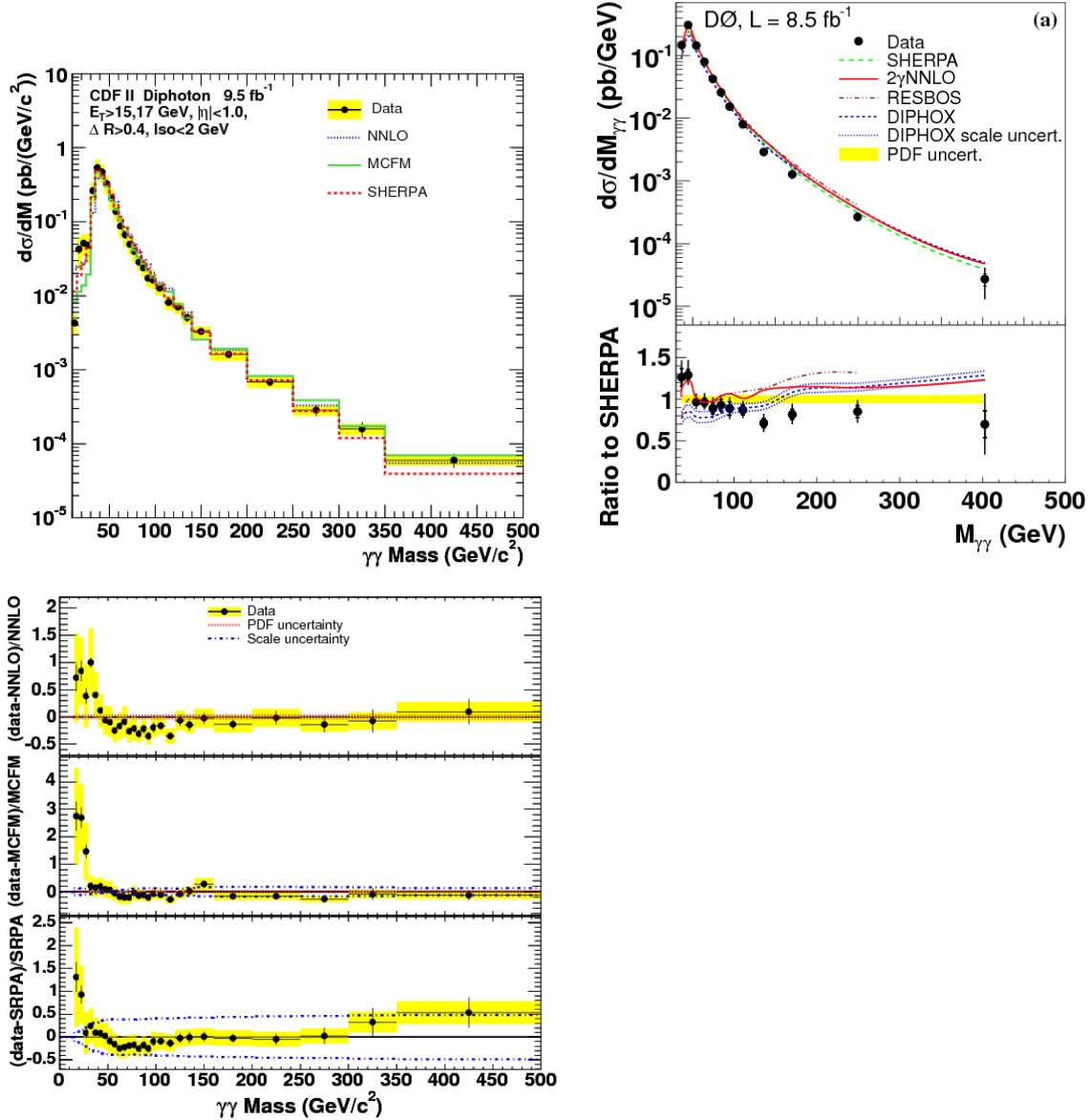


FIG. 84 – Sections efficaces différentielles de production $\gamma\gamma+X$ en fonction de $m_{\gamma\gamma}$ mesurées à partir de collisions $p\bar{p}$ à $\sqrt{s} = 1.96$ TeV au TEVATRON par les collaborations CDF (gauche, bas) et DØ (droite). L'emplacement du pic dans la distribution correspond à environ $m_{\gamma\gamma} = 2\sqrt{E_{T,1}^{\gamma,\min} E_{T,2}^{\gamma,\min}}$ et donc donne une idée de l'espace des phases exploré. Les valeurs de masse atteintes sont dans les deux cas inférieures à 500 GeV. Une comparaison des mesures avec les prédictions de SHERPA (*ME+PS merged at LO*) [83], RESBos (NLO+NNLL avec fragmentation approximée) [226], DIPHOX (NLO à ordre fixe incluant la fragmentation) [111], MCFM (NLO à ordre fixe) [227] et 2γ NNLO (NNLO à ordre fixe, isolation de Frixione) [117] est également effectuée [222, 223].

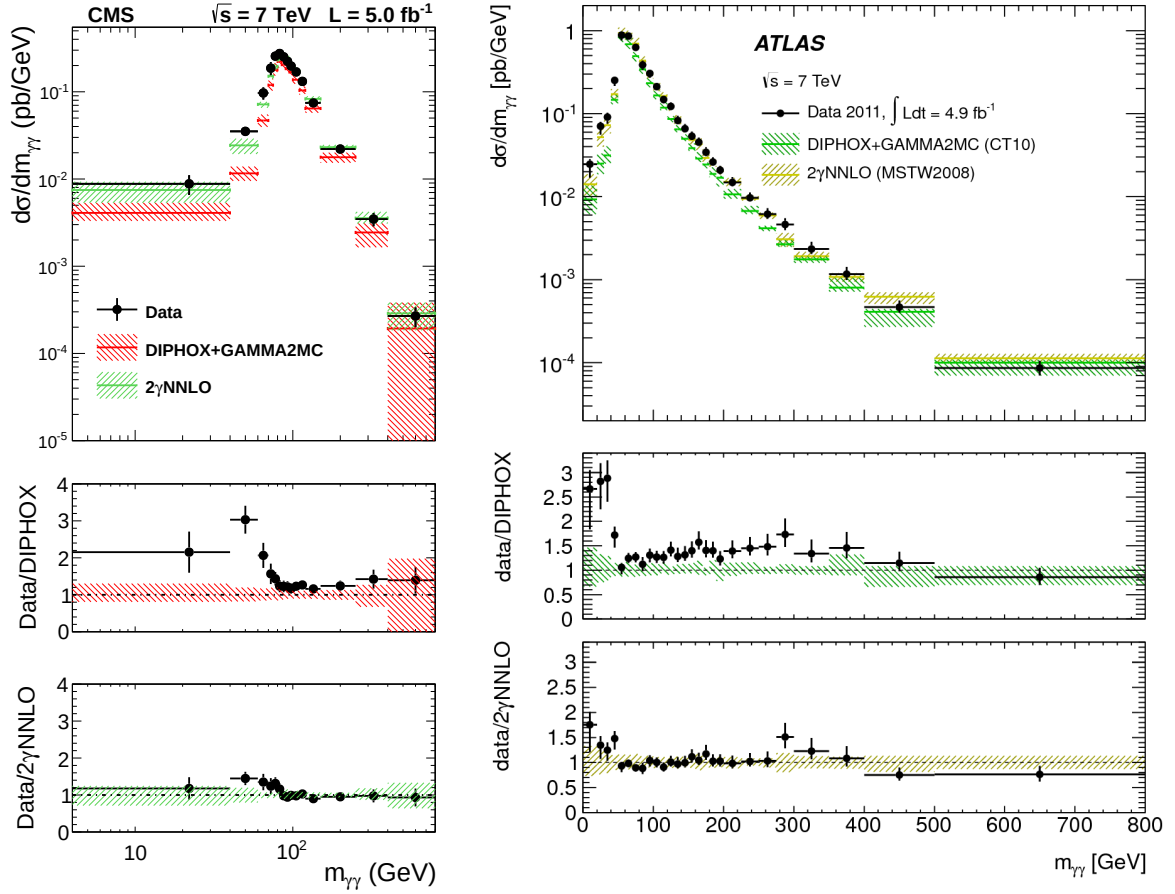


FIG. 85 – Sections efficaces différentielles de production $\gamma\gamma + X$ en fonction de $m_{\gamma\gamma}$ mesurées à partir de collisions pp à $\sqrt{s} = 7$ TeV au LHC par les collaborations CMS (gauche) et ATLAS (droite). L'emplacement du pic dans la distribution correspond à environ $m_{\gamma\gamma} = 2\sqrt{E_{T,1}^{\gamma,\min} E_{T,2}^{\gamma,\min}}$ et donc donne une idée de l'espace des phases exploré. Les valeurs de masse atteintes sont dans les deux cas inférieures à 800 GeV. Aucun excès n'est visible à 750 GeV [8, 9]. Une comparaison des mesures avec les prédictions de DIPHOX+GAMMA2MC (NLO à ordre fixe incluant la fragmentation et les premières corrections du mode de production gg) [111, 228] et 2γ NNLO (NNLO à ordre fixe, isolation de Frixione) [117] est également montrée [203, 224].

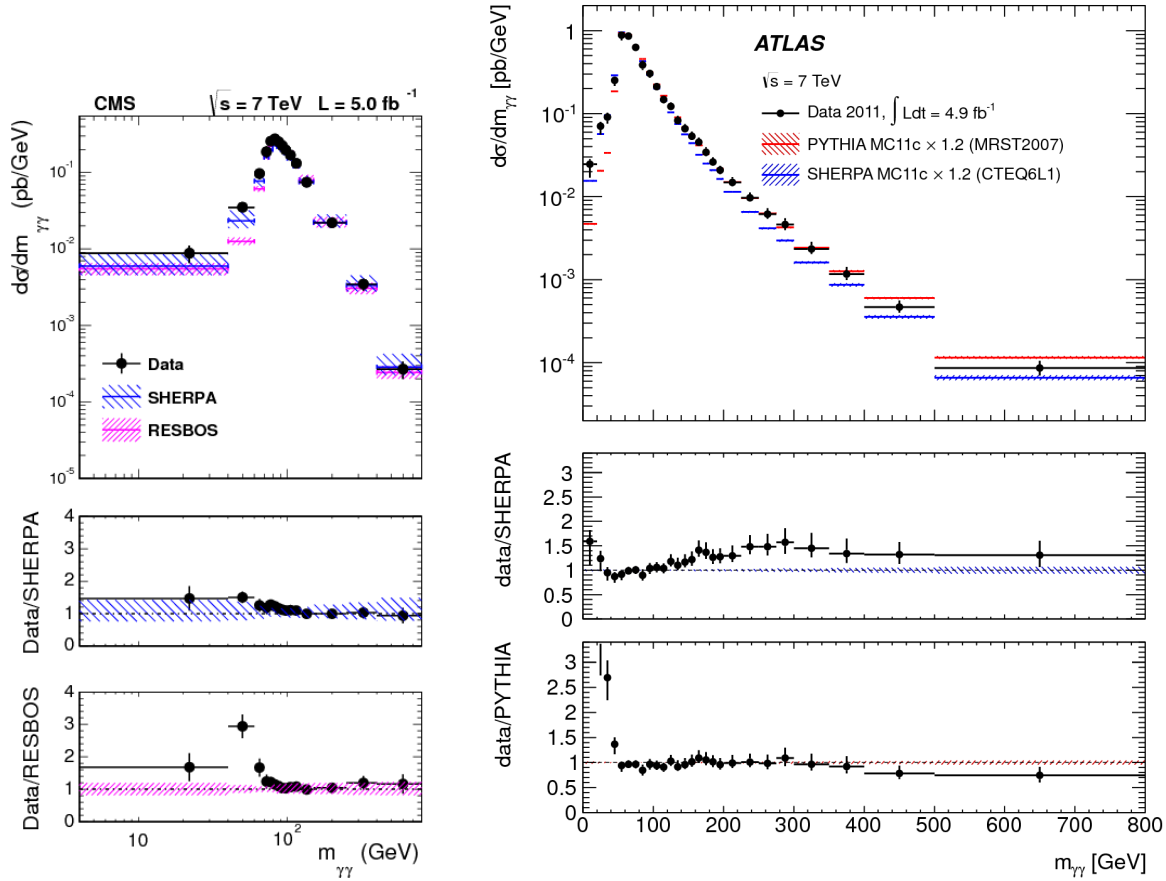


FIG. 86 – Sections efficaces différentielles de production $\gamma\gamma + X$ en fonction de $m_{\gamma\gamma}$ mesurées à partir de collisions pp au LHC par les collaborations CMS (gauche) et ATLAS (droite) à $\sqrt{s} = 7$ TeV. L'emplacement du pic dans la distribution correspond à environ $m_{\gamma\gamma} = 2\sqrt{E_{T,1}^{\gamma,\min} E_{T,2}^{\gamma,\min}}$ et donc donne une idée de l'espace des phases exploré. Les valeurs de masse atteintes sont dans les deux cas inférieures à 800 GeV. Une comparaison des mesures avec les prédictions de PYTHIA (LO $\gamma\gamma$ et γ -jet+PS) [82], RESBos (NLO+NNLL avec fragmentation approximée) [226] et SHERPA 1.4.0 ($ME+PS$ merged at LO) [83] est également montrée. Les prédictions de PYTHIA et SHERPA sont renormalisées à la section efficace fiducielle totale [203, 224].

La mesure préliminaire de CMS sélectionnant également des jets dans l'état final [225] offre des observables supplémentaires intéressantes grâce à la reconstruction des jets, qui permettent d'observer plus précisément la dynamique de l'interaction forte. Les résultats concernant deux de ces observables sont montrés en exemple dans la figure 87. Par exemple, il est intéressant d'observer que les prédictions semblent s'écarter des données à bas $\Delta R_{\gamma, j^{\text{lead}}}$ (voir figure 87) tandis que cette même variable semble décrite correctement par PYTHIA8 pour l'état final γ +jet dans ATLAS [229]. La mesure de CMS souffre néanmoins pour le moment de la statistique limitée des données à $\sqrt{s} = 7$ TeV. Une mesure similaire à $\sqrt{s} = 8$ TeV est en préparation avec les données du détecteur ATLAS [230] et devrait largement profiter de la luminosité intégrée quatre fois plus importante collectée en 2012.

La partie IV est organisée de la façon suivante : dans la section 13, l'échantillon de données de départ utilisé pour la mesure $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV ainsi que les échantillons simulés employés dans l'analyse sont décrits en détail. Les différents critères d'identification (voir section 9) et variables d'isolation (voir section 10) employés pour la sélection des données sont introduits dans la section 14. Ceux-ci permettent de définir une région enrichie en signal sur laquelle se base la mesure (région de signal) et plusieurs régions enrichies dans les différents bruits de fond présents dans la région de signal afin d'en extraire certaines caractéristiques (régions de contrôle). Dans la section 15, la soustraction du bruit de fond restant dans la région de signal et les incertitudes associées sont abordées en détail. La méthode de soustraction s'appuie au maximum sur les données afin de s'affranchir de la description incertaine de la simulation concernant la mauvaise identification des jets hadroniques et des électrons en photons. Cette partie constitue un des points les plus délicats de l'analyse, les incertitudes systématiques dues à la connaissance limitée du bruit de fond dominant l'incertitude expérimentale finale. La section 16 est dédiée aux déconvolutions des effets de détecteurs et aux corrections des inefficacités. Après avoir identifié et analysé en détail les différents effets de détecteur à l'aide d'un échantillon de signal simulé, plusieurs méthodes de déconvolution sont introduites et testées. Les distributions mesurées dans les données d'ATLAS après soustraction du bruit de fond sont finalement déconvoluées avec une méthode de déconvolution itérative [231] et les différentes incertitudes sont propagées (incertitudes issues de la soustraction du bruit de fond) et/ou ajoutées en quadrature (incertitudes issues de la déconvolution). Finalement, après avoir introduit succinctement les différents calculs théoriques actuellement disponibles pour la section efficace de production $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV, une confrontation entre résultats expérimentaux et prédictions est présentée dans la section 17. La partie s'achève sur une conclusion générale incluant un bilan des résultats et des axes d'amélioration expérimentaux possibles pour les analyses futures.

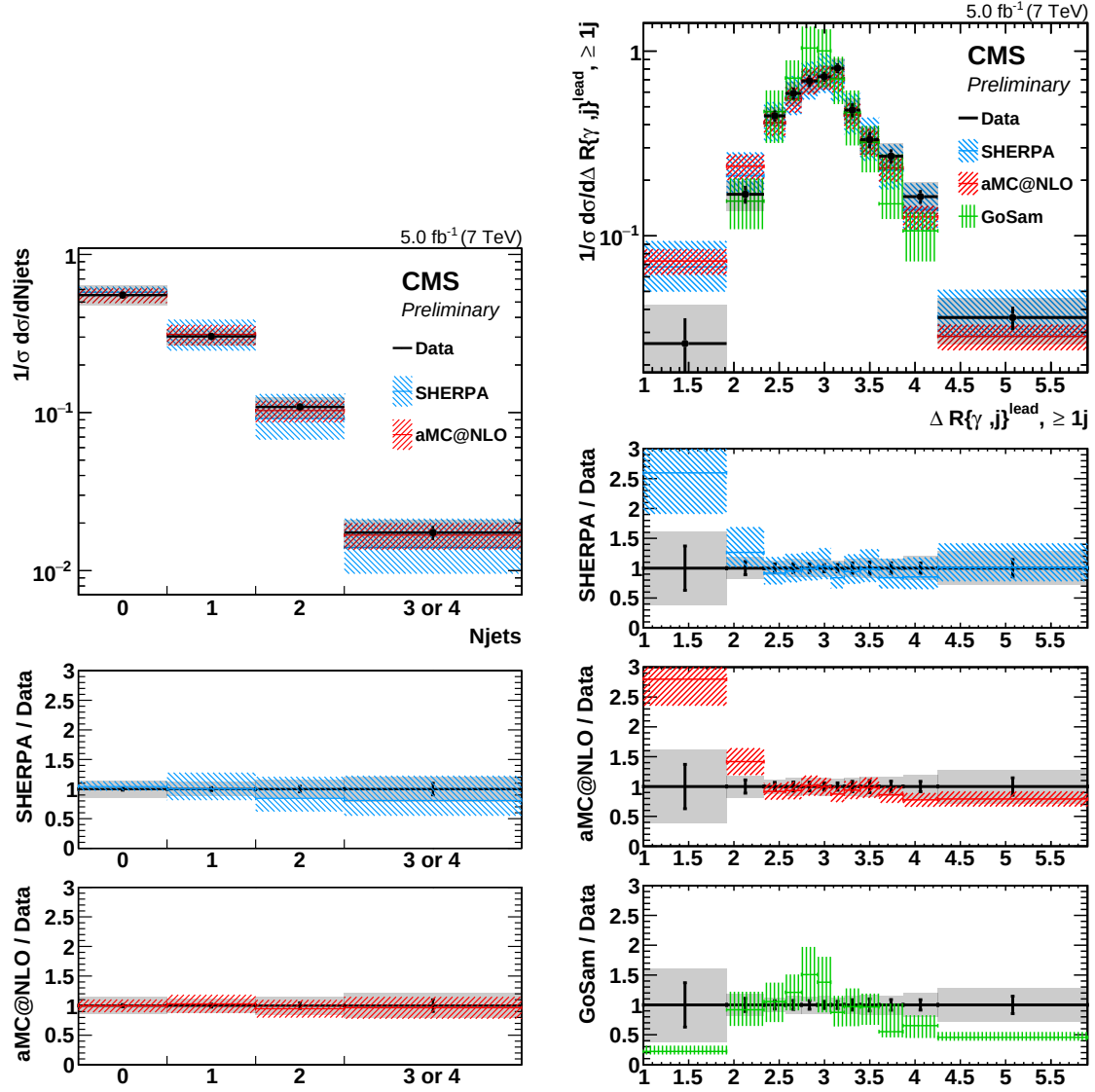


FIG. 87 – Sections efficaces différentielles de production de $\gamma\gamma + N$ jets en fonction du nombre de jets passant la sélection (gauche) et de la distance entre le photon et le jet de plus haute impulsion transverse (droite), mesurées par la collaboration CMS dans des collisions pp à $\sqrt{s} = 7$ TeV [225]. Les résultats sont comparés aux prédictions de SHERPA 1.4.0 (*ME+PS merged at LO*) [83], aMC@NLO (*NLO+PS matching*) [232] et GoSAM (NLO à ordre fixe incluant la fragmentation) [233].

13 Échantillons de données et échantillons simulés employés dans l'analyse

13.1 Choix du système de déclenchement

L'échantillon de données utilisé pour l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV a été collecté par la collaboration ATLAS en utilisant la chaîne de déclenchement appelée `g35_loose_g25_loose` pendant les collisions proton-proton produites par le LHC pendant l'année 2012 (voir section 5.7). Celle-ci n'a subi aucun facteur de réduction (*prescale*) pendant l'intégralité de la prise de données de 2012.

La chaîne de déclenchement `g35_loose_g25_loose` est basée sur les chaînes L1 appelées `EM12` et `EM16V`, demandant une région d'intérêt dans le calorimètre électromagnétique associée à une énergie transverse supérieure à 12 et 16 GeV, respectivement. La mention "V" dans `EM16V` signifie que le seuil varie légèrement en fonction de η (de 16 à 18 GeV) afin de tenir compte de la variation de quantité de matière traversée par les particules. Au niveau 3 (EF), la présence de deux objets photons (voir section 7) avec des impulsions transverses supérieures à 35 GeV et 25 GeV respectivement pour celui de plus haute et basse énergie ainsi que la sélection d'identification LOOSE satisfaite pour les deux objets est requise (voir section 9).

`g35_loose_g25_loose` constitue la chaîne de déclenchement LOOSE pour les paires de photons sans facteur de réduction avec les seuils en impulsion transverse les plus bas employée pour la prise de données de 2012. Les événements enregistrés lorsque l'acquisition était défectueuse ou les calorimètres et/ou le détecteur interne pas totalement opérationnels sont exclus. La luminosité intégrée mesurée après prise en compte du système de déclenchement et des critères de qualité s'élève à $20.24 \pm 0.38 \text{ fb}^{-1}$ (mesure finale) [134].

Dans la section 13.2, les différents échantillons simulés utilisés dans l'analyse sont introduits.

13.2 Échantillons simulés

13.2.1 Considérations générales

Comme nous l'avons vu dans la section 13 et plus généralement tout au long de la partie III, le bruit de fond principal attendu dans les données contenant des photons au LHC est issu de jets hadroniques mal identifiés en photons. Il existe également un bruit de fond plus faible constitué d'électrons mal identifiés (voir section 15). Afin d'étudier les caractéristiques des événements de signal et de bruit de fond, des échantillons Monte Carlo tentant de reproduire partiellement les données sont produits à l'aide de plusieurs générateurs. Ceux-ci incluent des interactions d'empilement afin de reproduire de façon réaliste les conditions expérimentales d'acquisition. Les particules stables, définies comme celles ayant un temps de vie supérieur à 10 ps, sont intégrées à une simulation complète de la réponse du détecteur ATLAS basée sur le programme GEANT4 [158] (voir section 5.7). Les différents événements sont ensuite reconstruits à l'aide des mêmes

algorithmes que ceux utilisés pour les véritables données de collisions. Plus d'informations sont disponibles dans la section 5.7 et la référence [159].

Les échantillons correspondant à la troisième production MC pour les données 2012 (MC12c) sont utilisés pour l'analyse. Ceux-ci reproduisent au niveau de chaque acquisition (*run*) les conditions des données véritables en terme d'empilement, d'inefficacités des différents détecteurs, etc. Une exception est faite pour les échantillons correspondant au processus Drell-Yan (DY) [103], qui sont issus de la première production MC pour les données 2012 pour des raisons pratiques (MC12a). En effet, ceux-ci servent à fournir des prédictions pour le bruit de fond correspondant au processus $Z^0 \rightarrow ee$ avec les deux électrons mal identifiés en photons dans l'analyse (voir section 15), et la différence entre données et MC concernant les taux de mauvaises identifications d'électrons en photons, a priori faible car uniquement reliée à la simulation du détecteur,³⁸ a été quantifiée précisément par le groupe performance e/γ uniquement pour la production MC12a [234]. Les résultats de l'étude sont résumés dans le tableau 4.

Pseudo-rapidité du candidat	Taux de mauvaise identification (données)	Taux de mauvaise identification (MC)	Facteur d'échelle
$-2.37 < \eta^\gamma < -2.0$	$0.0882 \pm 6 \cdot 10^{-4}$	$0.0844 \pm 0.9 \cdot 10^{-4}$	$1.04 \pm 0.01(stat) \pm 0.03(syst)$
$-2.0 < \eta^\gamma < -1.52$	$0.0978 \pm 5 \cdot 10^{-4}$	$0.0941 \pm 0.9 \cdot 10^{-4}$	$1.04 \pm 0.01(stat) \pm 0.03(syst)$
$-1.37 < \eta^\gamma < -1.0$	$0.0536 \pm 4 \cdot 10^{-4}$	$0.0495 \pm 0.6 \cdot 10^{-4}$	$1.08 \pm 0.01(stat) \pm 0.03(syst)$
$-1.0 < \eta^\gamma < -0.7$	$0.0360 \pm 3 \cdot 10^{-4}$	$0.0314 \pm 0.4 \cdot 10^{-4}$	$1.15 \pm 0.02(stat) \pm 0.04(syst)$
$-0.7 < \eta^\gamma < 0.7$	$0.0381 \pm 1 \cdot 10^{-4}$	$0.0341 \pm 0.2 \cdot 10^{-4}$	$1.12 \pm 0.01(stat) \pm 0.03(syst)$
$0.7 < \eta^\gamma < 1.0$	$0.0350 \pm 3 \cdot 10^{-4}$	$0.0315 \pm 0.4 \cdot 10^{-4}$	$1.11 \pm 0.02(stat) \pm 0.04(syst)$
$1.0 < \eta^\gamma < 1.37$	$0.0539 \pm 4 \cdot 10^{-4}$	$0.0495 \pm 0.6 \cdot 10^{-4}$	$1.09 \pm 0.01(stat) \pm 0.04(syst)$
$1.52 < \eta^\gamma < 2.0$	$0.1041 \pm 6 \cdot 10^{-4}$	$0.0951 \pm 0.8 \cdot 10^{-4}$	$1.09 \pm 0.01(stat) \pm 0.03(syst)$
$2.0 < \eta^\gamma < 2.37$	$0.0937 \pm 6 \cdot 10^{-4}$	$0.0870 \pm 1.0 \cdot 10^{-3}$	$1.08 \pm 0.01(stat) \pm 0.05(syst)$

TAB. 4 – Résumé des différences entre données et MC12a concernant les taux de mauvaises identifications d'électrons en photons, tiré de l'étude du pic du boson Z^0 détaillée dans [234]. Les *scale factors* et leurs incertitudes sont utilisés dans l'analyse $\gamma\gamma + X$ pour corriger les prédictions $\gamma\gamma$ des échantillons DY (voir section 15).

Les différences entre données et MC sont attribuées à des erreurs sur les paramètres d'efficacité de certains modules du détecteur interne dans la simulation, non corrigées dans MC12c. Elles ont été évaluées dans des catégories en η^γ uniquement, leur dépendance en E_T^γ étant négligeable dans l'intervalle de l'étude (30 GeV–60 GeV). Le taux de mauvaise identification $e \rightarrow \gamma$ à grand E_T^γ est mal connu mais relativement stable d'après la simulation. Le bruit de fond associé devrait dans tous les cas être très faible dans cette région à cause des sections efficaces hadroniques de production d'électrons très faibles en dehors du pic du Z^0 .

Les taux de mauvaise identification $e \rightarrow \gamma$ sont plus importants dans les bouchons à cause de la quantité de matière traversée plus importante (voir figure 35, section 5.3). Il est également intéressant de remarquer qu'une certaine asymétrie en η est observée dans les données. Aucun commentaire spécifique sur cette caractéristique n'est faite dans [234], bien que les asymétries soient identifiées comme provenant des candidats

³⁸La situation est différente pour le taux de mauvaise identification des jets en photons, qui dépendent significativement des modèles d'hadronisation non perturbatifs et présentent ainsi une incertitude élevée dans les échantillons simulés.

photons convertis et donc probablement issues de modules instables du détecteur interne, dont le comportement est difficilement reproductible dans la simulation.

Bien que les échantillons MC12a aient été simulés avec une ancienne version de la géométrie du détecteur, les taux de mauvaises identifications d'électrons en photons associés sont largement compatibles avec ceux obtenus à partir de la nouvelle géométrie, comme l'illustre la figure 88. Ainsi, les prédictions de bruit de fond issus des échantillons DY MC12a corrigés avec les facteurs du tableau 4 sont utilisées dans l'analyse $\gamma\gamma + X$ principale.

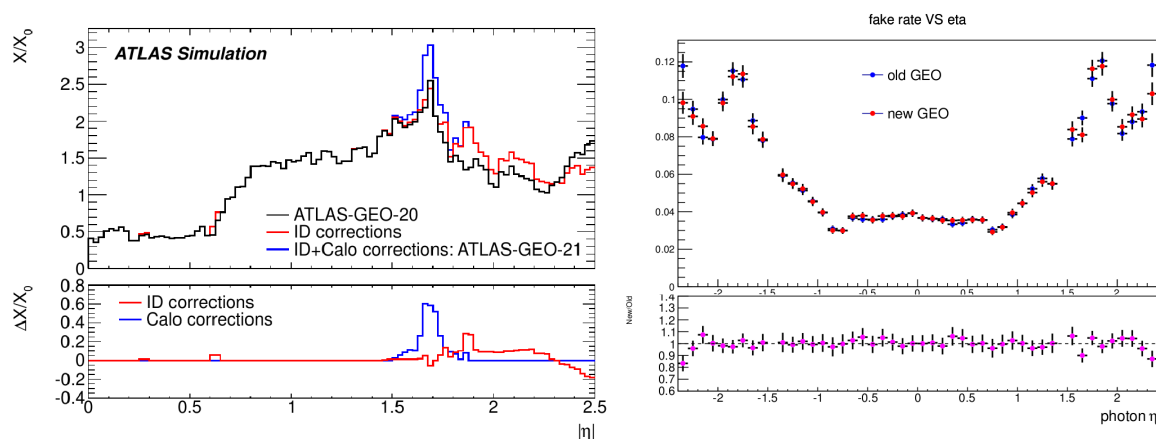


FIG. 88 – Différences entre l'ancienne et la nouvelle géométrie du détecteur ATLAS employée dans la simulation GEANT4 en terme de nombre de longueurs de radiations (gauche) et de taux de mauvaises identifications $e \rightarrow \gamma$ dans un échantillon MC $Z \rightarrow ee$ (droite). Les deux géométries sont identiques pour $|\eta| < 1.5$ et prédisent des taux de mauvaise identification compatibles quelque soit $|\eta^\gamma|$. L'ancienne (nouvelle) géométrie est utilisée pour la production MC12a (MC12c) [234].

Par ailleurs, tous les échantillons simulés incluent un modèle d'hadronisation et d'événement sous-jacent (voir section 3.2). Les paramètres des différents générateurs sont réglés selon les recommandations de la collaboration ATLAS (*tuning* AU2) [235]. Les échantillons MC sont également repondérés afin de faire correspondre leur nombre de vertex primaires d'interaction reconstruits ainsi que leur distribution en z aux données réelles.³⁹ Enfin, les variables d'identification des photons reconstruits sont corrigées à l'aide des *fudge factors* (voir section 9) et leur énergie d'isolation calorimétrique inclut les différentes corrections mentionnées dans la section 10 (corrections d'empilement/d'événement sous-jacent, corrections de fuite basées sur la simulation et corrections de fuite basées sur les données). La liste des échantillons simulés utilisés dans l'analyse est fournie dans le tableau 5 et détaillée dans les sections 13.2.2 et 13.2.3.

13.2.2 Les échantillons simulés de signal ($\gamma\gamma + X$)

Deux échantillons de signal ($\gamma\gamma + X$) sont simulés, l'un à l'aide du générateur SHERPA 1.4.0 [83] (utilisé par défaut) et l'un à l'aide du générateur PYTHIA 8.165 [82]

³⁹En pratique, les distributions du nombre de vertex primaires d'interaction reconstruits sont identifiés entre données et MC en repondérant le nombre moyen d'interactions par croisement de faisceaux $\langle \mu \rangle$ dans les MCs en considérant $\langle \mu \rangle_{\text{data}} = 1.09$. $\langle \mu \rangle_{\text{MC}}$ (voir appendice C).

Nom du générateur	Coups de génération ou filtre utilisé	Nombre d'événements générés	Section efficace (nb)	Efficacité du filtre
Signal $\gamma\gamma + X$ (MC12c)				
SHERPA 1.4.0	filtre 2DP20	9986378	0.49	0.18589
PYTHIA8	filtre 2DP20	9999576	138.09	0.00065
Bruit de fond γ -jet + X (MC12c)				
SHERPA 1.4.0	$E_T^\gamma > 15$ GeV	9999783	389.61	1.0000
SHERPA 1.4.0	$E_T^\gamma > 35$ GeV	9999183	24.475	1.0000
SHERPA 1.4.0	$E_T^\gamma > 70$ GeV	5499676	2.152	1.0000
SHERPA 1.4.0	$E_T^\gamma > 140$ GeV	2499984	0.13774	1.0000
SHERPA 1.4.0	$E_T^\gamma > 280$ GeV	999985	0.00596	1.0000
SHERPA 1.4.0	$E_T^\gamma > 500$ GeV	999976	0.00028	1.0000
SHERPA 1.4.0	$E_T^\gamma > 800$ GeV	99996	0.00001	1.0000
SHERPA 1.4.0	$E_T^\gamma > 1000$ GeV	99995	2.3821×10^{-6}	1.0000
PYTHIA8	filtre DP17	2999999	1235400	0.00023
PYTHIA8	filtre DP35	2999694	58768	0.00040
PYTHIA8	filtre DP70	8999861	3425	0.00057
PYTHIA8	filtre DP140	2998985	122.17	0.00097
PYTHIA8	filtre DP280	999989	3.3487	0.00014
PYTHIA8	filtre DP500	999877	0.11563	0.00180
PYTHIA8	filtre DP800	99997	0.00492	0.00190
PYTHIA8	filtre DP1000	99992	0.00087	0.00184
Bruit de fond $Z \rightarrow ee$ (MC12a)				
POWHEG+PYTHIA8	-	9994580	1.110	1.0000
POWHEG+PYTHIA8	$20 \text{ GeV} < m_{ee} < 60 \text{ GeV}$	4999990	0.789	1.0000
POWHEG+PYTHIA8	$120 \text{ GeV} < m_{ee} < 180 \text{ GeV}$	499999	0.009846	1.0000
POWHEG+PYTHIA8	$180 \text{ GeV} < m_{ee} < 250 \text{ GeV}$	100000	0.001571	1.0000
POWHEG+PYTHIA8	$250 \text{ GeV} < m_{ee} < 400 \text{ GeV}$	99999	0.000549	1.0000
POWHEG+PYTHIA8	$400 \text{ GeV} < m_{ee} < 600 \text{ GeV}$	99999	0.000087	1.0000
POWHEG+PYTHIA8	$600 \text{ GeV} < m_{ee} < 800 \text{ GeV}$	99998	0.000015	1.0000
POWHEG+PYTHIA8	$800 \text{ GeV} < m_{ee} < 1000 \text{ GeV}$	100000	3.75×10^{-6}	1.0000
POWHEG+PYTHIA8	$1000 \text{ GeV} < m_{ee} < 1250 \text{ GeV}$	99998	1.29×10^{-6}	1.0000
POWHEG+PYTHIA8	$1250 \text{ GeV} < m_{ee} < 1500 \text{ GeV}$	99997	3.58×10^{-7}	1.0000
POWHEG+PYTHIA8	$1500 \text{ GeV} < m_{ee} < 1750 \text{ GeV}$	99998	1.12×10^{-7}	1.0000
POWHEG+PYTHIA8	$1750 \text{ GeV} < m_{ee} < 2000 \text{ GeV}$	99998	3.84×10^{-8}	1.0000
POWHEG+PYTHIA8	$2000 \text{ GeV} < m_{ee} < 2250 \text{ GeV}$	99995	1.39×10^{-8}	1.0000
POWHEG+PYTHIA8	$2250 \text{ GeV} < m_{ee} < 2500 \text{ GeV}$	99995	5.23×10^{-9}	1.0000
POWHEG+PYTHIA8	$2500 \text{ GeV} < m_{ee} < 2750 \text{ GeV}$	99995	2.02×10^{-9}	1.0000
POWHEG+PYTHIA8	$2750 \text{ GeV} < m_{ee} < 3000 \text{ GeV}$	99896	7.89×10^{-10}	1.0000
POWHEG+PYTHIA8	$m_{ee} > 3000 \text{ GeV}$	99892	5.04×10^{-10}	1.0000

TAB. 5 – Échantillons MC utilisés dans l'analyse. Les filtres DPX et 2DPX correspondent respectivement à la présence de 1 et 2 photons immédiats dans l'état final au niveau particule avec $E_T^\gamma > X$ GeV. Des détails sur les différents paramètres de générations sont donnés dans le texte.

(utilisé pour l'évaluation des incertitudes systématiques). SHERPA fait usage des PDFs CT10 NLO [236], tandis que PYTHIA utilise les PDFs LO CTEQ6L1 [237].⁴⁰ PYTHIA contient des éléments de matrice au LO et un algorithme de cascade de partons (PS, pour *parton shower*). L'échantillon de signal est basé sur les processus durs $\gamma\gamma$ ($q\bar{q} \rightarrow \gamma\gamma$, $gg \rightarrow \gamma\gamma$) pour la composante directe et γ -jet + PS ($gg \rightarrow q\gamma$, $q\bar{q} \rightarrow g\gamma$) pour la composante 1-fragmentation. La composante 2-fragmentation est négligée (voir section 3.5.2).

SHERPA est également un générateur LO avec cascade de partons mais inclut un algorithme de fusion entre éléments de matrice et PS. Les processus durs à l'ordre le plus bas correspondant à l'état final $\gamma\gamma + 0, 1, 2$ parton(s) sont implémentés pour la composante directe et la composante de fragmentation est générée entièrement par la PS [238]. Afin d'assurer la bonne convergence des prédictions pour les processus avec 1 ou 2 partons dans l'état final, une coupure de séparation entre photons et partons $\Delta R_{\gamma, q(g)} > 0.3$ est implémentée.

Tous les événements des échantillons simulés de signal doivent passer un filtre qui vérifie la présence d'au moins 2 photons immédiats générés dans l'état final avec $E_T^\gamma > 20$ GeV, incluant les photons directs et ceux issus des radiations des quarks.⁴¹ Environ 10 000 000 événements sont générés pour chacun des échantillons. Ceux-ci sont employés en particulier pour modéliser la distribution en énergie d'isolation transverse des événements de signal (voir sections 10 et 15) et pour déterminer les corrections d'acceptance (voir section 16). Pour ce faire, les photons TIGHT sont repondérés en E_T^γ et $|\eta^\gamma|$ à l'aide des *ID scale factors*, ce qui permet d'obtenir des efficacités d'identification similaires entre MC et données (voir section 9).

13.2.3 Les échantillons simulés de bruit de fond (γ -jet, $Z \rightarrow ee$)

Des événements γ -jet sont générés en utilisant SHERPA $\gamma + 0, 1, 2$ parton(s) et PYTHIA γ -jet (voir section 13.2.2) et di-jet ($gg \rightarrow gg$, $gg \rightarrow q\bar{q}$, $qg \rightarrow qg$, $q\bar{q} \rightarrow q\bar{q}$, $q\bar{q} \rightarrow gg$). Ces processus constituent un bruit de fond important pour l'analyse $\gamma\gamma + X$ dans le cas où 1 (γ -jet) ou 2 jets (di-jet) sont mal identifiés en tant que photon(s) par le détecteur.

Les deux échantillons contiennent tous au moins 1 photon immédiat dans l'état final avec une impulsion transverse minimum $E_{T,\min}^\gamma > X$ GeV (voir tableau 5). Ce critère est implémenté via une coupure au niveau générateur (SHERPA) ou un filtre post-génération (PYTHIA).⁴² À cause du taux de mauvaise identification très bas des jets hadroniques en photons, ces échantillons ne contiennent que très peu d'événements avec deux photons reconstruits dans l'état final. D'autre part, on s'attend à ce que le taux de mauvaise identification $\text{jet} \rightarrow \gamma$ dans le MC ne soit pas fiable entre autres à cause de la dépendance importante de ce taux vis-à-vis des modèles d'hadronisation (QCD non-perturbative). Ainsi, les propriétés des bruits de fond constitués des processus γ -jet et di-jet sont déduits principalement de régions de contrôle dans les données, qui seront définies dans la section 14.6.

⁴⁰CT10 a uniquement été produit au NLO.

⁴¹Les photons issus de la radiation des quarks simulent la composante de fragmentation qui, dans les calculs partoniques à ordre fixe, est générée par la fragmentation colinéaire de partons en photons à l'aide de fonctions de fragmentation (voir section 3.5.2).

⁴²Il est également nécessaire d'exclure de l'échantillon PYTHIA les événements ayant au moins deux photons immédiats générés dans l'état final pour éviter un double comptage avec l'échantillon de signal correspondant. Les échantillons SHERPA $\gamma\gamma$ et γ -jet sont en revanche totalement disjoints.

Finalement, des événements $Z \rightarrow ee$, constituant un bruit de fond lorsque les deux électrons sont mal-identifiés en photons par le détecteur, sont générés avec POWHEG (NLO) [105] couplé à la cascade de partons de PYTHIA8 via un algorithme d'appariement. Les PDFs CT10 NLO sont utilisés. Ces échantillons sont employés pour l'étude du bruit de fond réduit constitué d'électrons mal identifiés. Afin d'assurer une statistique suffisante pour l'ensemble de l'espace des phases considéré, les échantillons sont générés dans 14 catégories en masse invariante pour les di-électrons (m_{ee}) couvrant l'intervalle 120 GeV–3000 GeV. Par ailleurs, des échantillons spécifiques sont générés pour les masses supérieures à 3000 GeV et celles comprises entre 20 et 60 GeV. Un échantillon est également produit pour toutes les masses et ensuite recoupé afin de sélectionner uniquement les événements ayant des m_{ee} générées dans l'intervalle 60 GeV–120 GeV.

Comme déjà mentionné dans la section 13.2.1, le nombre d'électrons mal identifiés en photons est corrigé dans des catégories en $|\eta^\gamma|$ via des facteurs d'échelle basés sur les données, ce qui permet d'obtenir des prédictions fiables pour le bruit de fond DY. Cependant, les corrections étant calculées uniquement selon $|\eta^\gamma|$, on ne s'attend par exemple pas une reproduction parfaite de l'énergie d'isolation transverse.

Dans la section 14, les critères d'identification et d'isolation des photons utilisés dans l'analyse $\gamma\gamma + X$ sont définis et justifiés. L'algorithme employé pour la reconstruction des vertex des di-photons est explicité. Les sélections utilisées pour définir la région de signal, qui est la région enrichie en événements $\gamma\gamma + X$ sur laquelle se base la mesure, ainsi que les régions de contrôle, qui sont les régions enrichies en jets/électrons mal identifiés utilisées pour étudier les différents bruits de fond, sont également décrites en détail.

14 Sélection de la région de signal et de régions de contrôle dans les données

14.1 Critères d'identification utilisés dans l'analyse

Les différents critères utilisés pour l'identification des photons dans ATLAS ont déjà été présentés et discutés en détail dans la section 9. Leurs caractéristiques principales ainsi que leur utilisation dans l'analyse $\gamma\gamma + X$ sont rappelées/introduites ci-dessous :

- ◇ **Loose (L)** : opère une sélection basée uniquement sur la fraction d'énergie collectée dans le calorimètre hadronique et deux variables décrivant la forme de la gerbe électromagnétique dans la seconde couche du calorimètre LAr. La sélection LOOSE présente une efficacité pour les photons immédiats d'environ 97% à $E_T^\gamma = 20$ GeV et supérieure à 99% pour $E_T^\gamma > 40$ GeV. Son taux de rejet est d'environ 1 000 pour les jets. Cette sélection est utilisée avec des valeurs de coupures légèrement plus faibles au niveau du système de déclenchement (voir section 13.1) que le critère standard défini par le groupe de performance, qui sera employé lors de la présélection (voir section 14.5).
- ◇ **Tight (T)** : opère une sélection basée sur la fraction d'énergie collectée dans le calorimètre hadronique et neuf variables décrivant la forme de la gerbe électromagnétique dans la première et la seconde couche du calorimètre LAr, ce qui lui permet d'atteindre un taux de rejet d'environ 5 000 au prix d'une efficacité légèrement plus faible que la sélection LOOSE, particulièrement à bas E_T^γ ($\simeq 80$ -90% à $E_T^\gamma = 40$ GeV, voir les figures 60 et 61 de la section 9). La sélection TIGHT est utilisée dans l'analyse afin de définir les régions les plus enrichies en photons immédiats (voir section 14.6).
- ◇ **Loose' (L')** : critère similaire à la sélection TIGHT au niveau des coupures effectuées sur les variables calculées à partir de la seconde couche du calorimètre et la fuite d'énergie collectée dans le calorimètre hadronique. Cela permet d'assurer des distributions d'énergie d'isolation transverse similaires entre le bruit de fond passant LOOSE' et celui passant TIGHT (voir section 9.4). En revanche, deux à cinq (équivalent alors à la totalité) des coupures sur les variables définies à partir de la première couche du calorimètre ne sont pas appliquées par rapport au critère TIGHT, afin de garder un maximum de bruit de fond. Les différentes sélections sont notées LOOSE'X (**L'X**) en fonction du nombre X de coupures non appliquées. Le critère LOOSE' est utilisée dans l'analyse afin de définir des régions enrichies en bruit de fond par la sélection des événements LOOSE' et non TIGHT, appelés alors NON-TIGHT (**T̃**). Le critère par défaut utilisé est **L'4**, qui est un bon compromis entre statistique et préservation de l'environnement moyen de la gerbe. Les autres critères (**L'2**, **L'3** et **L'5**) servent à l'évaluation des erreurs systématiques dues au choix arbitraire du nombre de coupures non appliquées. Plus de détails sont disponibles dans la section 9.4.

Dans la section 14.2, les régions de contrôle utilisées dans l'analyse pour étudier les bruits de fond composés d'électrons et de jets mal identifiés sont définies.

14.2 Définition de régions de contrôle dans les données

Trois régions de contrôle enrichies en jets mal identifiés sont définies à partir des critères d'identification $\mathbf{T}$ et $\mathbf{L}'$ (voir section 14.1) sur les deux photons composant la paire sélectionnée. On rappelle que le critère NON-TIGHT ($\tilde{\mathbf{T}}$) correspond aux photons LOOSE' et non TIGHT. Ces régions sont disjointes entre elles et disjointes de la région de signal. Elles sont définies par les critères d'identification suivants :

- ◇ $\mathbf{T}\tilde{\mathbf{T}}$: le photon de plus haute (basse) impulsion transverse passe la sélection TIGHT (NON-TIGHT). Région de contrôle enrichie en événements γ +jet.
- ◇ $\tilde{\mathbf{T}}\mathbf{T}$: le photon de plus haute (basse) impulsion transverse passe la sélection NON-TIGHT (TIGHT). Région de contrôle enrichie en événements jet+ γ .
- ◇ $\tilde{\mathbf{T}}\tilde{\mathbf{T}}$: les photons de plus haute et de plus basse impulsion transverse passent la sélection NON-TIGHT. Région de contrôle enrichie en événements di-jet.

Les régions $\mathbf{T}\tilde{\mathbf{T}}$, $\tilde{\mathbf{T}}\mathbf{T}$ et $\tilde{\mathbf{T}}\tilde{\mathbf{T}}$ constituent la principale source d'information permettant l'extraction des caractéristiques du bruit de fond composé de jets hadroniques mal identifiés. On attend cependant de légères différences en terme de cinématique (E_T^γ , η^γ) et d'énergie d'isolation entre les jets passant TIGHT et ceux passant NON-TIGHT à cause des dépendances cinématiques du taux de rejet de la sélection TIGHT et des corrélations résiduelles entre identification et isolation, respectivement (voir section 9). L'effet de ces différences sur la soustraction du bruit de fond restant dans l'échantillon sélectionné sera évalué dans la section 15.

Par ailleurs, deux régions de contrôle supplémentaires sont définies pour l'étude du bruit de fond issu d'électrons mal identifiés en utilisant une méthode *tag-and-probe* basée sur la désintégration électronique du boson Z^0 ($Z \rightarrow ee$) :

- ◇ **Sélection d'événements $e\gamma$ dans la fenêtre 80 GeV–100 GeV** : une sélection similaire à la sélection définissant la région de signal $\gamma\gamma + X$ est réalisée (voir section 14.6), à la différence que l'objet de plus **haute** impulsion transverse doit être reconstruit et identifié comme un électron et non un photon. Par ailleurs, la masse invariante de la paire $m_{e\gamma}$ doit être située aux alentours de la masse du boson Z^0 . Région de contrôle enrichie en électrons mal identifiés pour le photon de plus **basse** énergie.
- ◇ **Sélection d'événements γe dans la fenêtre 80 GeV–100 GeV** : idem, à la différence que l'objet de plus **basse** impulsion transverse doit être reconstruit et identifié comme un électron et non un photon. Région de contrôle enrichie en électrons mal identifiés pour le photon de plus **haute** énergie.

En effet, la section efficace ee étant très élevée dans $80 \text{ GeV} < m_{\gamma e, e\gamma} < 100 \text{ GeV}$ à cause de la résonance du boson Z^0 , on s'attend à ce que la grande majorité des photons reconstruits dans cette région soit en réalité des électrons mal identifiés. La contamination des candidats photons par de vrais photons immédiats issue par exemple de processus tels que $W\gamma$ ou $Z\gamma$ est négligeable et celle des jets est estimée comme étant

inférieure à 10% selon une étude réalisée dans le cadre de l'analyse $H \rightarrow \gamma\gamma$ [239]. Ces deux régions de contrôle pour le bruit de fond électron sont utilisées principalement pour la modélisation de l'énergie d'isolation transverse des électrons mal identifiés.

Dans la section 14.3, les variables et coupures finales d'isolation (voir sections 10.2 et 10.3) choisies dans l'analyse $\gamma\gamma + X$ sont commentées et leur pouvoir de discrimination entre signal et bruit de fond est illustré pour l'état final $\gamma\gamma + X$, notamment grâce à l'utilisation des régions de contrôle.

14.3 Choix et comparaison des variables d'isolation entre signal et bruits de fond

14.3.1 L'énergie d'isolation dans le détecteur interne

Afin de réduire le bruit de fond venant de jets et d'électrons mal identifiés dans les données de départ, une coupure sur l'isolation dans le détecteur interne (ou isolation des traces, voir section 10.2) est appliquée pour chaque photon composant la paire lors de la sélection des données (voir section 14.6). La variable utilisée, notée $p_T^{\text{track.iso.}}$ dans la suite, est définie comme la somme scalaire des impulsions transverses des traces situées dans un cône défini par $\Delta R < 0.2$ autour du candidat photon. Elle correspond à l'isolation des traces utilisée dans les analyses d'ATLAS étudiant la désintégration du boson de Higgs en paires de photons [43, 188, 208]. Afin que $p_T^{\text{track.iso.}}$ présente les meilleures performances, les traces à compter dans la somme doivent passer une sélection supplémentaire incluant les critères suivants [240] :

- ◇ **$p_T > 1 \text{ GeV}$** : ce critère permet d'exclure les traces de plus basse énergie, souvent issues d'artefacts dans le détecteur interne ou d'interactions d'empilement.
- ◇ **Au moins 9 points de mesure dans les couches en silicium** : le nombre moyen de mesures attendues étant de 3 dans le détecteur à pixels et 8 dans le SCT (voir section 5.3), cette sélection permet de limiter les traces venant d'artefacts dans le détecteur interne.
- ◇ **Pas de cellules mortes traversées dans la première couche de pixel** : un nombre réduit de cellules situées dans la première couche de pixels du détecteur interne se sont éteintes au cours de la première prise de données à cause des fortes radiations présentes dans cette région. Un veto sur les traces passant par de telles cellules est implémenté afin de réduire les artefacts de mesure.⁴³
- ◇ **$d_0 < 1.5 \text{ mm}$** : d_0 représente le paramètre d'impact dans le plan transverse aux faisceaux. La coupure associée a été optimisée afin de limiter les artefacts de mesure tout en gardant une bonne efficacité.
- ◇ **$z_0 \sin \theta < 15 \text{ mm}$** : $z_0 \sin \theta$ est appelé paramètre d'impact longitudinal modifié. Il est évalué par rapport à la position en z du vertex des di-photons, dont l'identi-

⁴³L'effet de ce critère sur l'efficacité de la coupure sur l'isolation des traces implémentée dans l'analyse est typiquement inférieure à 1%.

fication sera décrite dans la section 14.6. La multiplication de z_0 par $\sin \theta$ permet de prendre en compte la croissance de Δz_0 en fonction de $|\eta| = |\ln(\tan \frac{\theta}{2})|$. La coupure associée a été optimisée afin de limiter les artefacts de mesure tout en gardant une bonne efficacité.

La distribution de l'isolation des traces est représentée dans la figure 89 pour le candidat photon de plus haute (gauche) et basse (droite) impulsion transverse de la paire considérée, en fonction de la vraie nature du candidat. Les distributions sont normalisées à 1 afin de comparer les formes des distributions uniquement. Les photons immédiats sont extraits à partir de l'échantillon de signal SHERPA tandis que les jets et électrons mal identifiés sont issus respectivement des régions de contrôle $\tilde{\mathbf{T}}\mathbf{T}$, $\mathbf{T}\tilde{\mathbf{T}}$ et *tag-and-probe* $Z \rightarrow ee$ définies dans la section 14.2.

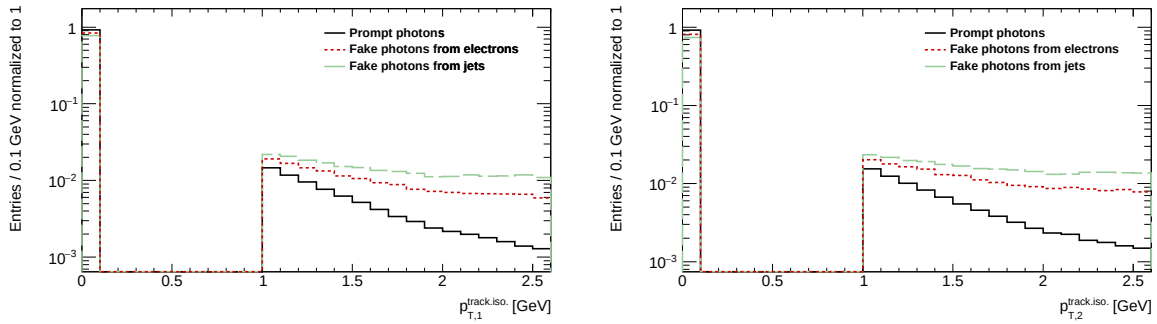


FIG. 89 – Distributions de l'énergie d'isolation dans le détecteur interne pour les candidats photons de plus haute (gauche) et basse (droite) impulsion transverse dans la paire considérée. Les différents cas sont représentés séparément en fonction de la vraie nature du candidat : photons immédiats présents dans l'échantillon de signal SHERPA (ligne pleine noire), électrons mal-identifiés issus de la région de contrôle *tag-and-probe* $Z \rightarrow ee$ dans les données (ligne rouge pointillée) et jets mal-identifiés issus des régions de contrôle $\tilde{\mathbf{T}}\mathbf{T}$ (gauche) et $\mathbf{T}\tilde{\mathbf{T}}$ (droite) dans les données. Tous les événements doivent passer l'intégralité de la sélection implémentée dans les données (voir section 14.6) à l'exception des critères d'isolation.

L'isolation des traces présente une distribution très différente pour le signal et le bruit de fond, illustrant bien le pouvoir discriminant de cette variable. Le pic très important situé à 0 (les distributions de la figure 14.2 sont représentées en échelle semi-logarithmique) montre que la plupart des candidats photons passant la sélection ne sont pas émis à proximité d'une particule chargée avec une impulsion transverse supérieure à 1 GeV. Les distributions sont ensuite vides jusqu'à 1 GeV, qui correspond au seuil de sélection en p_T des traces à inclure dans le cône d'isolation. Dans le cas des photons immédiats, une queue de distribution fortement décroissante est observée. Celle-ci correspond aux photons produits avec des jets additionnels de haute énergie (processus $2 \rightarrow 3, 4$) et/ou issus de la fragmentation d'un parton. Elle est légèrement supérieure pour le second photon, qui est émis plus proche des jets en moyenne comparé au premier photon ($\vec{p}_{T,1}^\gamma + \vec{p}_{T,2}^\gamma + \vec{p}_{T,\text{recoil}} = \vec{0}$).

Pour les jets et les électrons mal identifiés, une queue de distribution beaucoup plus plate est observée. Cela est attendu de l'activité hadronique importante du jet autour du candidat photon, qui a alors de grandes chances d'être en réalité un méson léger de

type π^0 , et du fait que la majorité des électrons mal identifiés est en réalité issue de photons de freinage (*bremsstrahlung*) émis à très haute énergie.

Dans l'analyse principale, $p_{T,1,2}^{\text{track.iso.}} < 2.6$ GeV est demandé (voir section 14.6). Cette valeur de coupure est motivée par les études réalisées précédemment dans le cadre de la sélection des candidats $H \rightarrow \gamma\gamma$ [189, 208], qui ont démontré qu'une telle coupure sur l'isolation des traces permettait de renforcer la robustesse de la sélection vis-à-vis des conditions d'empilement grâce à la quasi-indépendance de son efficacité par rapport au *pileup* (voir section 10.1, figure 62). Par ailleurs, à cause de la dépendance résiduelle en E_T^γ de l'isolation calorimétrique dans les données (voir section 10.4), couper sur l'isolation des traces est utile afin de pouvoir relâcher légèrement la coupure sur l'isolation calorimétrique de 4 GeV, qui est la valeur traditionnellement employée dans les analyses photon 2012 d'ATLAS, à 6 GeV, tout en maintenant une excellente pureté. Cela permet ainsi d'atténuer la baisse d'efficacité du critère d'isolation à haut E_T^γ (voir section 16). La coupure sur $p_T^{\text{track.iso.}}$ constitue une nouvelle sélection par rapport à l'analyse $\gamma\gamma + X$ précédente [203], qui bénéficiait du taux d'empilement plus faible des données 2011 (voir section 4.3).

Dans la section 14.3.2, la variable d'isolation calorimétrique utilisée dans le cadre de l'analyse est présentée.

14.3.2 L'énergie d'isolation dans le calorimètre

Afin de réduire encore davantage le bruit de fond, une coupure sur l'énergie d'isolation calorimétrique, notée $E_T^{\text{calo.iso.}}$ dans la suite, est implémentée dans l'analyse principale ($E_{T,1,2}^{\text{calo.iso.}} < 6$ GeV). La variable d'isolation calorimétrique considérée suit la définition standard présentée dans la section 10.3 en se basant sur les *clusters* topologiques et un cône de taille fixe défini par $\Delta R < 0.4$. Les corrections dues à l'empilement et à l'événement sous-jacent moyens (voir section 10.3.2) et les corrections de fuite moyenne basées sur la simulation (voir section 10.3.3) sont appliquées par défaut. Par ailleurs, les corrections de fuite supplémentaires basées sur les données (voir section 10.4) [201] sont appliquées aux échantillons de signal simulés SHERPA et PYTHIA. Celles-ci sont calculées séparément dans des catégories fines en E^γ et $|\eta^\gamma|$. Leur effet global sur les distributions de signal est illustré dans la figure 90, séparément pour le premier et le second photon de la paire. Ces corrections incluent des translations des distributions MC allant jusqu'à 2 GeV à haut E^γ et des corrections de largeur dans les bouchons. Elles sont attribuées principalement à la mauvaise description du développement latéral des gerbes dans la simulation (voir section 10.4).

Les distributions de $E_{T,1}^{\text{calo.iso.}}$ et $E_{T,2}^{\text{calo.iso.}}$ sont représentées pour les photons immédiats et les électrons/jets mal identifiés dans la figure 91 en se basant respectivement sur l'échantillon de signal SHERPA, les régions de contrôle *tag-and-probe* $Z \rightarrow ee$ dans les données et les régions de contrôle $\tilde{\mathbf{T}}\mathbf{I}\mathbf{I}$ (gauche) et $\mathbf{T}\tilde{\mathbf{I}}\mathbf{I}$ (droite) dans les données, où $\mathbf{I}$ signifie que le candidat doit satisfaire les critères d'isolation des traces (< 2.6 GeV) et calorimétrique (< 6 GeV), en plus du critère d'identification $\mathbf{T}$ ou $\tilde{\mathbf{T}}$.⁴⁴ Des énergies d'isolation calorimétrique négatives sont attendues et observées dans le cas où l'empile-

⁴⁴La sélection sur $p_T^{\text{track.iso.}}$ a un effet limité sur les distributions d'isolation calorimétrique pour les photons immédiats (voir section 16), en revanche elle affecte considérablement les distributions des bruits de fond ainsi que la contamination des régions de contrôle (voir section 15).

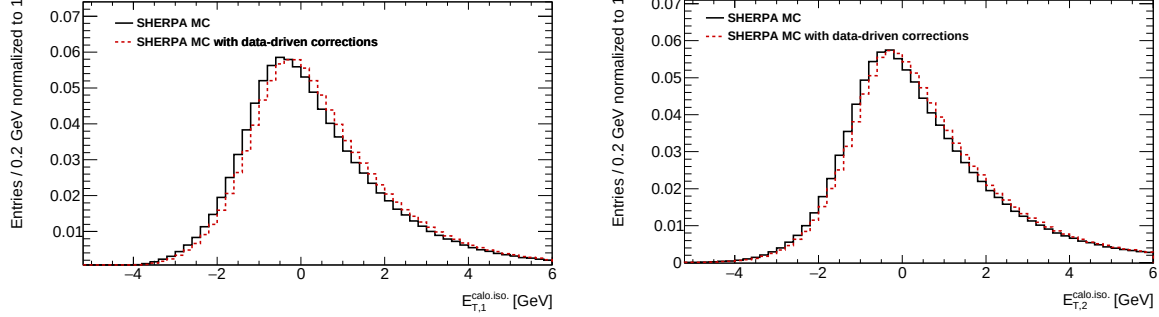


FIG. 90 – Énergie d'isolation calorimétrique pour le premier (gauche) et le deuxième (droite) candidat photon parmi les paires de photons immédiats de l'échantillon de signal SHERPA, sans (trait plein noir) et avec (trait poillé rouge) les corrections de fuites supplémentaires basées sur les données (voir section 10.4) [201]. Tous les événements doivent passer l'intégralité de la sélection implémentée dans les données (voir section 14.6).

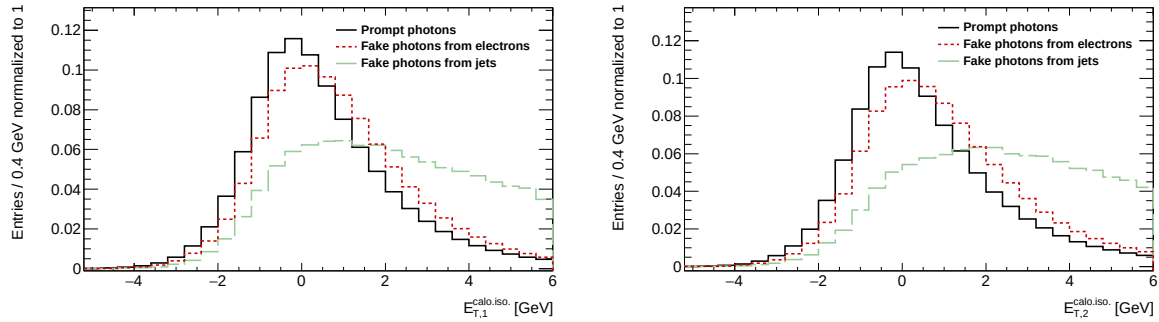


FIG. 91 – Énergie d'isolation calorimétrique pour le premier (gauche) et le deuxième (droite) candidat photon parmi les photons immédiats de l'échantillon de signal SHERPA (photons immédiats, trait plein noir), les électrons mal identifiés issus de la région de contrôle *tag-and-probe* $Z \rightarrow ee$ dans les données (ligne rouge pointillée) et les jets mal identifiés issus des régions de contrôle $\bar{\mathbf{T}}\mathbf{T}\mathbf{I}$ (gauche) et $\mathbf{T}\mathbf{T}\mathbf{I}$ (droite) dans les données. Tous les événements doivent passer l'intégralité de la sélection implémentée dans les données (voir section 14.6).

ment/événement sous-jacent et/ou l'énergie de fuite présentent une fluctuation négative vis-à-vis de leur valeur moyenne.

Dans le cas des photons immédiats, la plus grande partie des énergies d'isolation calorimétrique est répartie autour de $E_T^{\text{calo.iso.}} = 0$ GeV sur une largeur d'environ 1.5 GeV. Toutefois, une queue de distribution clairement visible du côté des énergies d'isolation positives, provenant des paires de photons produites avec des jets additionnels de haute énergie (processus $2 \rightarrow 3, 4$) et/ou issues de la fragmentation d'un parton. La queue de distribution est légèrement supérieure pour le second photon, qui est en moyenne émis plus proche des jets comparé au premier photon ($\vec{p}_{T,1}^\gamma + \vec{p}_{T,2}^\gamma + \vec{p}_{T,\text{recoil}} = \vec{0}$). Par ailleurs, à la fois le pic et la largeur des distributions de $E_T^{\text{calo.iso.}}$ montrent une certaine dépendance vis-à-vis de la pseudorapacité et de l'impulsion transverse des photons (voir figures 92 et 93, respectivement), avec en particulier une distinction nette entre les photons du tonneau ($|\eta^\gamma| < 1.37$) et des bouchons ($1.56 < |\eta^\gamma| < 2.37$), qui ont une distribution plus piquée et translatée à plus haute énergie d'environ +400 MeV. Cet effet est relié aux variations en termes de composition et de quantité de matière en amont mais aussi à l'intérieur même du calorimètre LAr.

À l'instar de l'isolation des traces, les électrons mal identifiés ont en général une énergie d'isolation calorimétrique plus élevée par rapport aux photons immédiats (voir figure 91) à cause de la présence de plusieurs photons de freinage (*bremsstrahlung*). Cette distribution ne peut pas être reproduite précisément avec le MC DY à cause de différences significatives entre les taux de mauvaise identification $e \rightarrow \gamma$ entre données et simulation, qui ne sont corrigés qu'en fonction de $|\eta^\gamma|$ par des facteurs d'échelle fournis par le groupe de performance e/γ (voir section 13.2). Il est important de souligner que photons et électrons mal identifiés ont des distributions d'isolation calorimétrique significativement différentes, élément qui a toujours été négligé lors des analyses $\gamma\gamma + X$ précédentes réalisées dans ATLAS [202, 203]. Nous reviendrons sur ce point lors de la soustraction du bruit de fond (voir section 15). Une queue de distribution très importante est observée pour les jets mal identifiés, attendue de l'activité hadronique présente autour du jet. On s'attend également à une certaine dépendance des distributions de $E_T^{\text{calo.iso.}}$ vis-à-vis de la pseudorapacité et de l'impulsion transverse des candidats dans le cas du bruit de fond, à l'image de ce qui est observé pour les photons immédiats. Cependant, la statistique limitée des régions de contrôle rend plus difficile sa mise en évidence, comme on peut le voir dans les figures 94 et 95, réalisées pour les jets avec les régions de contrôle **TITI** (gauche) et **TIT** (droite). Néanmoins, cette dépendance sera considérée lors de la soustraction du bruit de fond (voir section 15)

Comme déjà annoncé au début de cette section et dans la section 14.3.1, uniquement les candidats satisfaisant $E_{T,1}^{\text{calo.iso.}}, E_{T,2}^{\text{calo.iso.}} < 6$ GeV sont considérés dans la suite. Cela correspond à un léger assouplissement du critère d'isolation calorimétrique comparé aux analyses di-photon ATLAS précédentes ($E_{T,1}^{\text{calo.iso.}}, E_{T,2}^{\text{calo.iso.}} < 4$ GeV) [202, 203], justifié par la dépendance en E_T^γ accrue de $E_T^{\text{calo.iso.}}$ dans les données à 8 TeV due à l'accès au plus haut E_T^γ et par les résultats de l'analyse du $H \rightarrow \gamma\gamma$ concernant la robustesse de la sélection vis-à-vis de l'empilement (voir section 10.1, figure 62).

Au cours de cette section, les différentes variables d'identification et d'isolation ont été introduites dans le cadre de l'analyse $\gamma\gamma + X$. Dans la section 14.4, l'algorithme de reconstruction du vertex associé à la paire de photons est introduit succinctement.

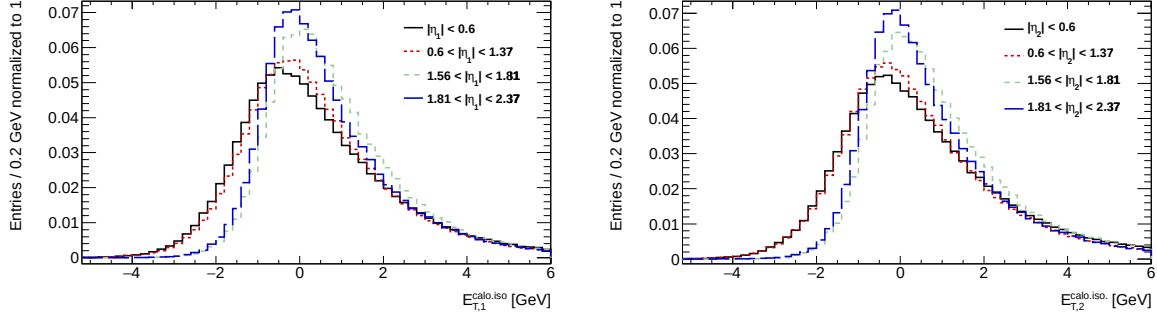


FIG. 92 – Énergie d'isolation calorimétrique pour le premier (gauche) et le deuxième (droite) photon parmi les photons immédiats de l'échantillon de signal SHERPA, dans quatre catégories en pseudo-rapidité. Tous les événements doivent passer l'intégralité de la sélection implémentée dans les données (voir section 14.6). Les corrections de fuites supplémentaires [201] sont appliquées afin de décrire de façon plus réaliste les données. Une certaine dépendance en $|\eta^\gamma|$ est observée.

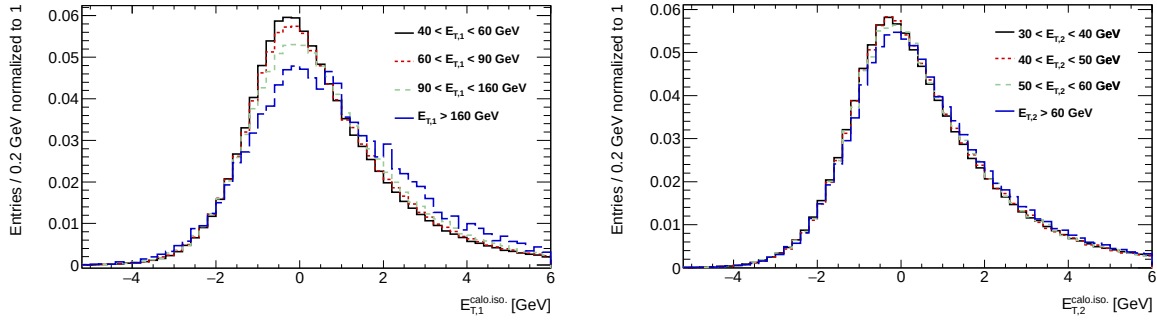


FIG. 93 – Énergie d'isolation calorimétrique pour le premier (gauche) et le deuxième (droite) photon parmi les photons immédiats de l'échantillon de signal SHERPA, dans quatre catégories en impulsion transverse. Les catégories ne sont pas identiques pour les deux photons à cause de la statistique limitée disponible à haut E_T^γ pour le deuxième photon. Tous les événements doivent passer l'intégralité de la sélection implémentée dans les données (voir section 14.6). Les corrections de fuites supplémentaires [201] sont appliquées afin de décrire de façon plus réaliste les données. Une certaine dépendance en E_T^γ est observée.

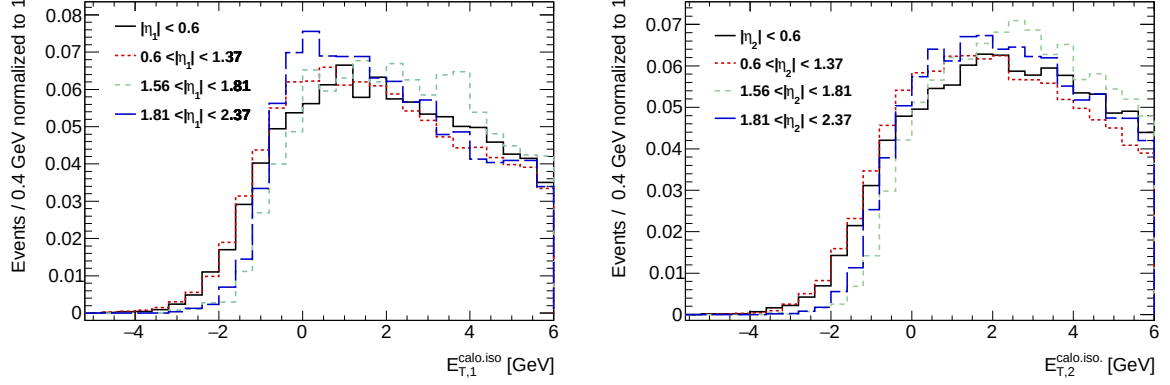


FIG. 94 – Énergie d'isolation calorimétrique pour le premier (gauche) et le deuxième (droite) candidat photon des régions de contrôle $\tilde{\mathbf{T}}\mathbf{ITI}$ (gauche) et $\mathbf{T}\mathbf{ITI}$ (droite), dans quatre catégories en pseudo-rapidité. Tous les événements doivent passer l'intégralité de la sélection implémentée dans les données (voir section 14.6).

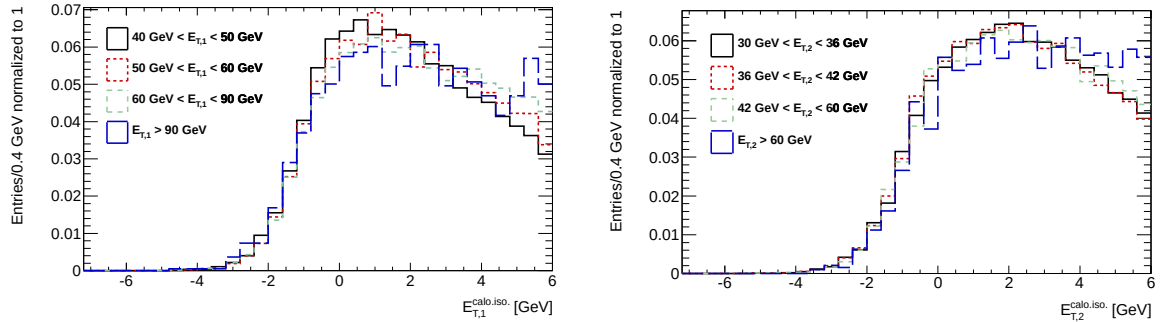


FIG. 95 – Énergie d'isolation calorimétrique pour le premier (gauche) et le deuxième (droite) candidat photon des régions de contrôle $\tilde{\mathbf{T}}\mathbf{ITI}$ (gauche) et $\mathbf{T}\mathbf{ITI}$ (droite), dans quatre catégories en impulsion transverse. Les catégories ne sont pas identiques pour les deux candidats photons à cause de la statistique limitée disponible à haut E_T pour le deuxième photon. Tous les événements doivent passer l'intégralité de la sélection implémentée dans les données (voir section 14.6).

14.4 Reconstruction du vertex associé à la paire

La zone de recouvrement des faisceaux lors d'un croisement s'étale sur une longueur non négligeable selon l'axe z ($\sigma_z = 48$ mm en 2012) de part et d'autre du point d'interaction principal d'ATLAS ($z = 0$). La distribution des interactions proton-proton selon l'axe z suit alors approximativement une distribution gaussienne de largeur σ_z centrée en 0. Ainsi, l'identification du vertex associé à la paire de photons permet de connaître plus précisément la position en z à laquelle ceux-ci ont été émis, ce qui correspond à une (petite) correction pour η_1^γ et η_2^γ logiquement centrée autour de 0. Sa RMS est de l'ordre de 3%, à la fois pour les données et les échantillons simulés.

Dans ATLAS, le vertex di-photon est identifié grâce à la méthode dite de pointage des photons (*photon pointing*) [190], qui fait usage à la fois de la trajectoire individuelle des candidats photons et de la contrainte donnée par σ_z afin de déterminer la position en z du point d'émission. Pour les photons convertis, la position du vertex de conversion est également utilisée si la ou les trace(s) associée(s) présente(nt) une quantité suffisante de points de mesure dans le silicium. Un algorithme basé sur un réseau de neurones est alors exécuté afin de sélectionner le vertex di-photon le plus vraisemblable. L'algorithme combine les informations issues du pointage des photons, incluant éventuellement des informations de conversion, avec la somme des impulsions transverses au carré $\sum (p_T^{\text{tracks}})^2$ et la somme scalaire $\sum p_T^{\text{tracks}}$ des traces associées aux différents vertex reconstruits (traces de photons convertis exclues), ainsi que la différence d'angles d'azimut $\Delta\phi$ entre $\sum \vec{p}_T^\gamma$ et $\sum \vec{p}_T^{\text{tracks}}$.

Les pseudo-rapidités $\eta_{1,2}^\gamma$ et impulsions transverses $E_{T,1,2}^\gamma$ sont alors recalculées en tenant compte de la position en z du vertex. Les corrections sont ensuite propagées aux autres observables ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, voir section 12.3). Typiquement, l'effet résultant est très faible pour les résolutions individuelles en $\eta_{1,2}^\gamma$ et $E_{T,1,2}^\gamma$ mais est plus quantifiable pour les observables propres à la paire. En effet, des études réalisées dans le cadre de l'analyse $H \rightarrow \gamma\gamma$ ont montré que l'identification du vertex $\gamma\gamma$ permettait d'améliorer la résolution en $m_{\gamma\gamma}$ d'environ 1 GeV [208]. Les performances de l'algorithme sont résumées dans la figure 96 pour un signal $H \rightarrow \gamma\gamma$ à 125 GeV [189, 241]. L'efficacité de sélection du vertex di-photon à 0.3 mm près en z est alors supérieure à 75% pour une paire de photons non-convertis. L'espace des phases considéré dans l'analyse $\gamma\gamma + X$ étant similaire à l'analyse $H \rightarrow \gamma\gamma$ en terme d'impulsion transverse minimum ($E_{T,1(2)}^\gamma > 40$ (30) GeV), les performances seront supposées comparables. Les données et la simulation sont par ailleurs en très bon accord.

Dans les sections 14.5 et 14.6, la sélection et le classement final des événements pour l'analyse $\gamma\gamma + X$ sont résumés.

14.5 Présélection des événements et des candidats photons

Partant des échantillons décrits dans la section 13, une présélection au niveau de chaque événement est tout d'abord effectuée afin de considérer uniquement ceux permettant des mesures fiables pour l'analyse. Celle-ci est parfois redondante avec la description de l'échantillon de départ fournie dans la section 13.1 ; elle est énoncée ci-dessous :

◇ **Système de déclenchement** : l'événement doit avoir été sélectionné par

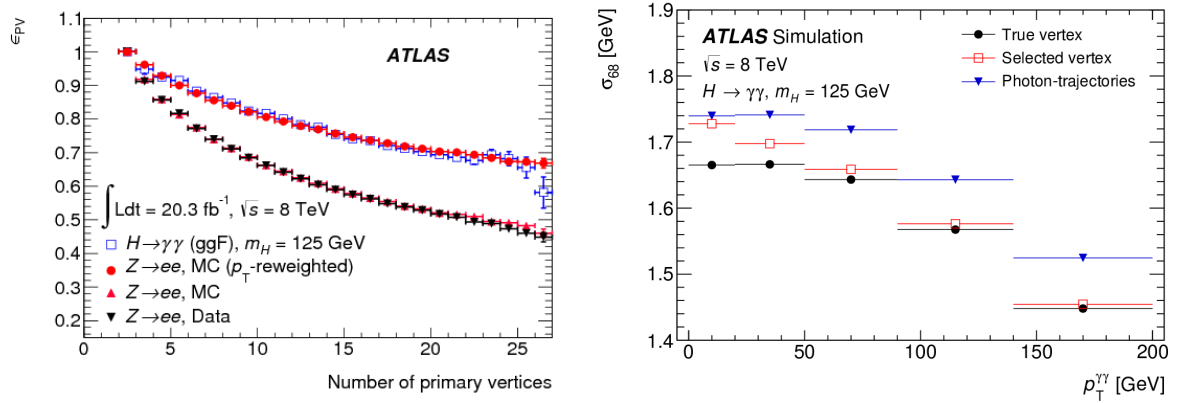


FIG. 96 – (gauche) Efficacité de sélection du vertex di-photon à 0.3 mm près en fonction du nombre de vertex reconstruits dans l'événement pour des événements $H \rightarrow \gamma\gamma$ simulés avec deux photons non-convertis (carrés bleus), $Z \rightarrow ee$ simulés et sélectionnés dans les données (triangles rouges et noirs) et les mêmes événements $Z \rightarrow ee$ simulés mais repondérés afin de reproduire le spectre en $E_{T,1,2}^\gamma$ des événements $H \rightarrow \gamma\gamma$ (ronds rouges). Les traces des électrons ne sont pas considérées pour l'identification du vertex. (droite) Résolution en masse à 68% C.L (*confidence-level*) des événements $H \rightarrow \gamma\gamma$ en fonction de $p_{T,\gamma\gamma}$ lorsque le véritable vertex (points noirs), le vertex renvoyé par l'algorithme basé sur le réseau de neurones (carrés rouges) ou celui déterminé uniquement à l'aide des trajectoires des photons (triangles bleus) sont utilisés [189, 241].

la chaîne de déclenchement `EF_g35_loose_g25_loose` décrite dans la section 13.1, qui requiert au moins deux candidats photons reconstruits au niveau 3 (*Event Filter*) passant une sélection d'identification LOOSE assouplie avec $E_{T,1(2)}^\gamma > 35$ (25) GeV.⁴⁵

- ◇ **Good Run List** : l'événement doit figurer sur la liste “*AllGood*” d'ATLAS, qui assure que celui-ci a été enregistré dans de bonnes conditions, que ce soit au niveau des faisceaux, du système de déclenchement, du détecteur interne ou des calorimètres.
- ◇ **Pas de problème pendant l'acquisition et pas de données corrompues** : il peut arriver que le calorimètre présente un niveau de bruit anormalement élevé pendant l'acquisition ou que les données soient corrompues au moment de la lecture. Les événements enregistrés pendant de telles périodes sont exclus de la mesure.
- ◇ **Vertex primaire** : l'événement doit comporter au moins un vertex primaire reconstruit comportant au moins trois traces chargées. Ce critère permet de rejeter le bruit de fond qui n'est pas issu de collisions proton-proton (voir [242] pour des détails supplémentaires). On s'attend à ce que celui-ci soit très faible dans les données 2012 grâce au chargement élevé de la machine pendant les différentes périodes d'acquisition, limitant ainsi le nombre de croisements sans collisions (voir section 4).

⁴⁵Les impulsions transverses considérées ici ne sont pas corrigées de la position du vertex di-photon ni finement calibrées.

- ◇ **Nombre de photons reconstruits dans l'événement** : l'événement doit contenir au moins deux photons reconstruits dans les données finales.

Si l'événement est présélectionné, la calibration finale en énergie des candidats photons est exécutée (voir section 8 et [147]) et les variables d'identification des échantillons simulés sont corrigées à l'aide des *fudge factors* (voir section 9 et [185]). Une deuxième présélection est alors initiée, concernant cette fois les photons reconstruits dans l'événement. Si au moins deux photons sont présélectionnés dans l'événement, celui-ci est sauvegardé. La présélection des photons vise à obtenir un échantillon avec une pureté acceptable (photons immédiats vs électrons/jets mal identifiés) tout en maintenant une efficacité (photons immédiats passant la sélection) aussi haute que possible. Les efficacités finales de sélection seront étudiées en détail dans la section 16.3. La quasi-intégralité des critères employés pour la présélection des photons constitue les recommandations de base du groupe de performance e/γ pour les analyses photons 2012 dans ATLAS ; ils sont déclinés ci-dessous :

- ◇ **Impulsion transverse** : le candidat photon doit avoir une impulsion transverse $E_{T,1(2)}^\gamma > 25$ GeV après calibration.⁴⁶
- ◇ **Pseudo-rapacité** : le candidat photon doit avoir une pseudo-rapacité η_{S2} mesurée dans la seconde couche du calorimètre telle que $|\eta_{S2}| < 2.37$, avec l'exclusion de $1.37 < |\eta_{S2}| < 1.56$, qui correspond à la région de transition entre le tonneau et les bouchons.⁴⁷
- ◇ **Résolveur d'ambiguïté électron-photon** : le candidat photon doit passer le résolveur d'ambiguïté électron-photon. Celui-ci élimine les photons convertis ayant au moins une trace associée passant par des cellules mortes de la première couche de pixels du détecteur interne. Il permet de réduire considérablement le nombre d'électrons mal identifiés en photons convertis (voir sections 5.3 et 16).
- ◇ **Qualité de l'objet et nettoyage des données optimisé pour les photons** : le candidat photon doit être à l'extérieur des régions du calorimètre comportant des cellules mortes ou des problèmes de courts-circuits (voir section 8.3.2).
- ◇ **Sélection d'identification Loose** : le candidat photon doit passer la sélection d'identification LOOSE nominale (voir sections 9 et 14.1).

Si au moins deux photons sont présélectionnés dans l'événement, celui-ci est sauvegardé et la dernière phase de sélection commence, consistant à choisir la paire de candidats photons et à en déterminer ses propriétés cinématiques finales. Cette dernière étape est décrite dans la section 14.6.

⁴⁶Les impulsions transverses considérées ici ne sont pas corrigées de la position du vertex di-photon.

⁴⁷L'utilisation de η_{S2} pour les coupures en pseudo-rapacité au niveau reconstruit suit les recommandations du groupe de performance e/γ .

14.6 Classement final des différents événements

Tout d'abord, les candidats photons présélectionnés sont classés par E_T^γ décroissant. Seuls les deux premiers candidats sont considérés dans la suite. Ainsi, tout photon supplémentaire présélectionné de plus bas E_T^γ est ignoré. Une mesure de la section efficace $\gamma\gamma + X$ est actuellement en cours dans ATLAS à $\sqrt{s} = 8$ TeV [213], les résultats préliminaires faisant état d'une section efficace plus que 100 fois plus faible que pour $\gamma\gamma + X$ et donc totalement négligeable en tenant compte des incertitudes théoriques et expérimentales (voir section 17).

La variable d'isolation calorimétrique décrite dans la section 14.3.2 étant calculée dans un cône défini par $\Delta R < 0.4$, il est important que les cônes d'isolation des deux candidats ne se recouvrent pas totalement. Ainsi, un critère de séparation similaire à celui employé pour la définition du volume fiduciel est demandé, *ie.* $\Delta R_{\gamma\gamma} > 0.4$ (voir section 12.2).⁴⁸ Celui-ci est inchangé par rapport aux analyses di-photon précédentes [202, 203].

Une fois le critère de séparation vérifié, le vertex di-photon est reconstruit (voir section 14.4) et les corrections associées en η^γ sont propagées aux différentes observables. Si les deux candidats passent la sélection d'identification **TIGHT** (**TT**), la sélection d'isolation (**II**, $p_T^{\text{track.iso.}} < 2.6$ GeV, $E_T^{\text{calo.iso.}} < 6$ GeV) et la sélection finale en impulsion transverse ($E_{T,1}^\gamma > 40$ GeV, $E_{T,2}^\gamma > 30$ GeV),⁴⁹ l'événement est classé dans la région de signal (**TITI**).

D'autre part, si les candidats passent les critères d'isolation et la sélection finale en impulsion transverse mais présentent une identification différente (**TIGHT** ou **NON-TIGHT**), ils viennent peupler les régions de contrôle pour les jets mal identifiés **TI $\tilde{\text{T}}$ I**, **$\tilde{\text{T}}$ ITI** et **$\tilde{\text{T}}\tilde{\text{T}}$ I** (voir section 14.2). Si aucun critère d'isolation particulier n'est appliqué, la mention “**I**” disparaît, ce qui donne les régions **TT**, **T $\tilde{\text{T}}$** , **$\tilde{\text{T}}\tilde{\text{T}}$** .

Le choix de la paire de photons est effectué juste après la présélection et en particulier avant la sélection d'isolation (**I**) et d'identification (**T**, **$\tilde{\text{T}}$**) afin d'avoir des cinématiques aussi proches que possible entre régions de signal et régions de contrôle. En effet, l'efficacité et les taux de rejet de ces coupures dépendent de la cinématique, ce qui pourrait introduire des biais non négligeables entre les différentes régions. Pour des raisons similaires, l'identification du vertex di-photon et les corrections associées sont appliquées après avoir choisi la paire de façon définitive. Celles-ci occasionnent dans de très rares cas ($< 1\%$) une réorganisation des deux candidats photons de la paire ($2 \rightarrow 1$ et $1 \rightarrow 2$). Un éventuel troisième candidat photon présélectionné n'est jamais considéré. L'inefficacité associée est très faible ($< 1\%$) et corrigée pendant l'étape de déconvolution (voir section 16). D'autre part, les échantillons ne sont pas différenciés en fonction du statut de conversion des candidats photons afin de garder une bonne statistique dans l'ensemble de l'espace des phases.

Les nombres d'événements après les différentes coupures appliquées pour la sélection **TITI** (région de signal) sont disponibles dans le tableau 6 pour les données 2012 et les échantillons de signal simulés SHERPA et PYTHIA. Le nombre total d'événements

⁴⁸ $\Delta R_{\gamma\gamma}$ est calculé avant reconstruction du vertex di-photon car seule la séparation des candidats dans le calorimètre importe pour limiter le recouvrement des cônes d'isolation.

⁴⁹Les impulsions transverses considérées ici sont finement calibrées et corrigées de la position du vertex di-photon.

Coupure	Données 2012	SHERPA $\gamma\gamma$	PYTHIA $\gamma\gamma$
Dans le D3PD	71 749 692	9 986 378	9 999 576
Système de déclenchement	25 701 909	2 984 908	2 961 602
GRL	24 530 463	-	-
Erreur LAr/ <i>Tile</i> , données corrompues	24 476 785	-	-
Vertex primaire	24 431 638	2 983 061	2 960 070
Préselection des photons	11 980 908	2 232 016	2 190 803
$\Delta R > 0.4$	11 791 505	2 210 262	2 188 461
LOOSE'4/LOOSE'4	2 634 216	1 990 039	1 945 181
TIGHT/TIGHT	1 122 762	1 812 545	1 769 332
TIGHT-isolé/TIGHT-isolé	528 354	1 596 075	1 496 866
$E_{T,1}^\gamma > 40$ GeV, $E_{T,2}^\gamma > 30$ GeV (TITI sample)	312 754	986 619	937 536

TAB. 6 – Nombre d'événements après chaque coupure définissant la région de signal (**TITI**) pour l'ensemble des données 2012 et les échantillons de signal simulés SHERPA et PYTHIA non-repondérés. Les chiffres de départ (*all in D3PD*) pour les données 2012 viennent de fichiers pré-filtrés pour l'analyse (voir section 5.7).

ments **TITI** dans les données 2012 est 312 754 tandis que les échantillons de signal simulés présentent une statistique environ 3 fois supérieure. Le nombre total d'événements dans les régions de contrôle enrichies en jets mal identifiés dans les données (**TITI**, **TITI** et **TITI**) est fourni dans le tableau 7 pour les différentes définitions de LOOSE' (voir section 9.4). Les régions de contrôle *tag-and-probe* $Z \rightarrow ee$ enrichies en premier et deuxième électron mal identifié en photon contiennent respectivement 153 207 et 174 017 événements.

Finalement, le nombre d'événements **TITI** enregistrés en fonction du numéro d'acquisition (*Run number*) est représenté dans la figure 97. Le nombre d'événements est normalisé par la luminosité collectée pendant l'acquisition en question de façon à ce qu'on attende un chiffre constant au cours des *runs*, ce qui est observé globalement. La figure 97 permet ainsi d'illustrer la stabilité de la prise de données du détecteur ATLAS pendant l'année 2012.

Dans la section 15, l'analyse et la soustraction du bruit de fond restant dans l'échantillon **TITI** sont abordés en détails.

Région de contrôle	Loose'2	Loose'3	Loose'4	Loose'5
$\tilde{\mathbf{T}}\mathbf{T}\mathbf{I}$	20 452	22 309	38 435	55 470
$\mathbf{T}\tilde{\mathbf{T}}\mathbf{I}$	40 636	42 741	82 616	112 039
$\tilde{\mathbf{T}}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$	3 660	4 148	14 576	28 473

TAB. 7 – Nombre d'événements dans les régions de contrôle enrichies en jets mal identifiés dans les données pour les différentes définitions de LOOSE'.

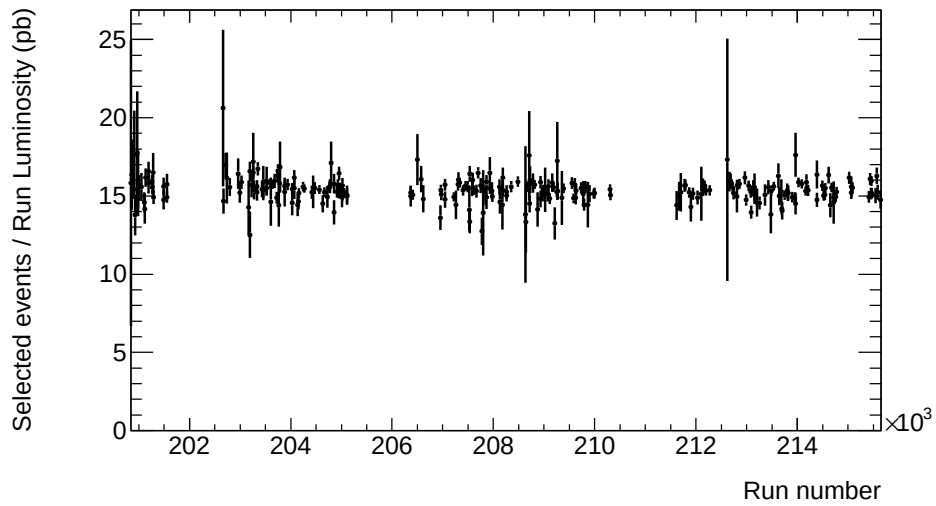


FIG. 97 – Nombre d'événements passant la sélection $\mathbf{T}\mathbf{T}\mathbf{I}$ normalisé par la luminosité intégrée et représenté en fonction du numéro d'acquisition. Les incertitudes tiennent compte uniquement de la statistique limitée.

15 Estimation du bruit de fond hadronique et électrons à partir des données

15.1 Considérations générales

Comme déjà mentionné au cours des sections précédentes, l'échantillon **TITI** défini dans la section 13 est composé essentiellement par des événements $\gamma\gamma$ (signal), γ +jet/jet+ γ /di-jet (bruit de fond hadronique) et ee (bruit de fond électrons), où les jets et les électrons sont mal identifiés par le détecteur.

On s'attend à ce que les bruits de fond γ +jet et jet+ γ soient les deux bruits de fond dominants à cause des sections efficaces de production de jets très élevées du LHC. Ces deux bruits de fond diffèrent uniquement par l'ordre en E_T^γ entre les deux candidats photons ($E_T^\gamma < E_T^{\text{fake}}$ ou vice-versa) mais sont attendus dans des proportions significativement différentes du fait de la stabilité relative du taux de mauvaise identification jet $\rightarrow \gamma$ en fonction de E_T^γ et de la baisse rapide de la section efficace de production photon-jet avec p_T^{jet} [243]. En effet, le taux de mauvaise identification jet $\rightarrow \gamma$ peut être estimé grossièrement dans les données par une méthode de type *tag-and-probe* décrite dans l'appendice B.⁵⁰ Les résultats sont représentés dans la figure 98 et montrent une dépendance significative du taux de mauvaise identification en fonction de $|\eta^\gamma|$ et une stabilité relative en fonction de E_T^γ . On s'attend donc à ce que l'échantillon **TITI** contienne significativement plus de γ +jet que de jet+ γ . Cette observation est confirmée par les mesures de pureté récentes effectuées dans l'analyse des photons inclusifs ($\gamma + X$) à $\sqrt{s} = 8$ TeV en fonction de E_T^γ [186, 200].

Les bruits de fond di-jet et ee attendus sont plus faibles parce qu'ils requièrent une mauvaise identification des deux objets afin de passer la sélection **TITI**. Par ailleurs, le bruit de fond ee est localisé autour du pôle en masse du boson Z^0 ($m_Z \simeq 91$ GeV) à cause de la très grande augmentation de la section efficace ee à cet endroit de l'espace des phases, provenant des désintégrations résonantes du $Z^0 \rightarrow ee$. Le bruit de fond di-jet devrait baisser rapidement à haut E_T^γ pour les mêmes raisons que celles évoquées dans le paragraphe précédent.

On attend également un bruit de fond très faible composé d'événements $e\gamma$ et γe .⁵¹ Ce bruit de fond est très réduit ($< 1\%$) à cause de la faible section efficace inclusive $e + \gamma$ et $\gamma + e$ au LHC, qui est formée par des états finals comme $\gamma W \rightarrow \gamma e \nu$ ou $\gamma Z \rightarrow \gamma ee$. Les événements $e\gamma$ et γe présents dans l'échantillon **TITI** seront négligés dans la soustraction nominale du bruit de fond. Une incertitude systématique dédiée sera évaluée de façon conservatrice afin de tenir compte de l'effet de cette simplification.

Afin de pouvoir mesurer précisément la section efficace $\gamma\gamma + X$, il est nécessaire d'estimer le plus rigoureusement possible le nombre d'événements de bruit de fond présents dans l'échantillon **TITI**, objet de la section 15. Cette étape constitue l'un des points les plus délicats de l'analyse, les incertitudes systématiques dues à la connaissance limitée du bruit de fond dominant l'incertitude expérimentale finale.

⁵⁰Plus précisément, cette méthode permet d'estimer la probabilité pour un jet reconstruit comme un photon LOOSE'4 de passer la sélection TIGHT.

⁵¹Le bruit de fond e +jet et jet+ e est négligeable du fait des sections efficaces associées très faibles comparées à la production de γ +jet/jet+ γ .

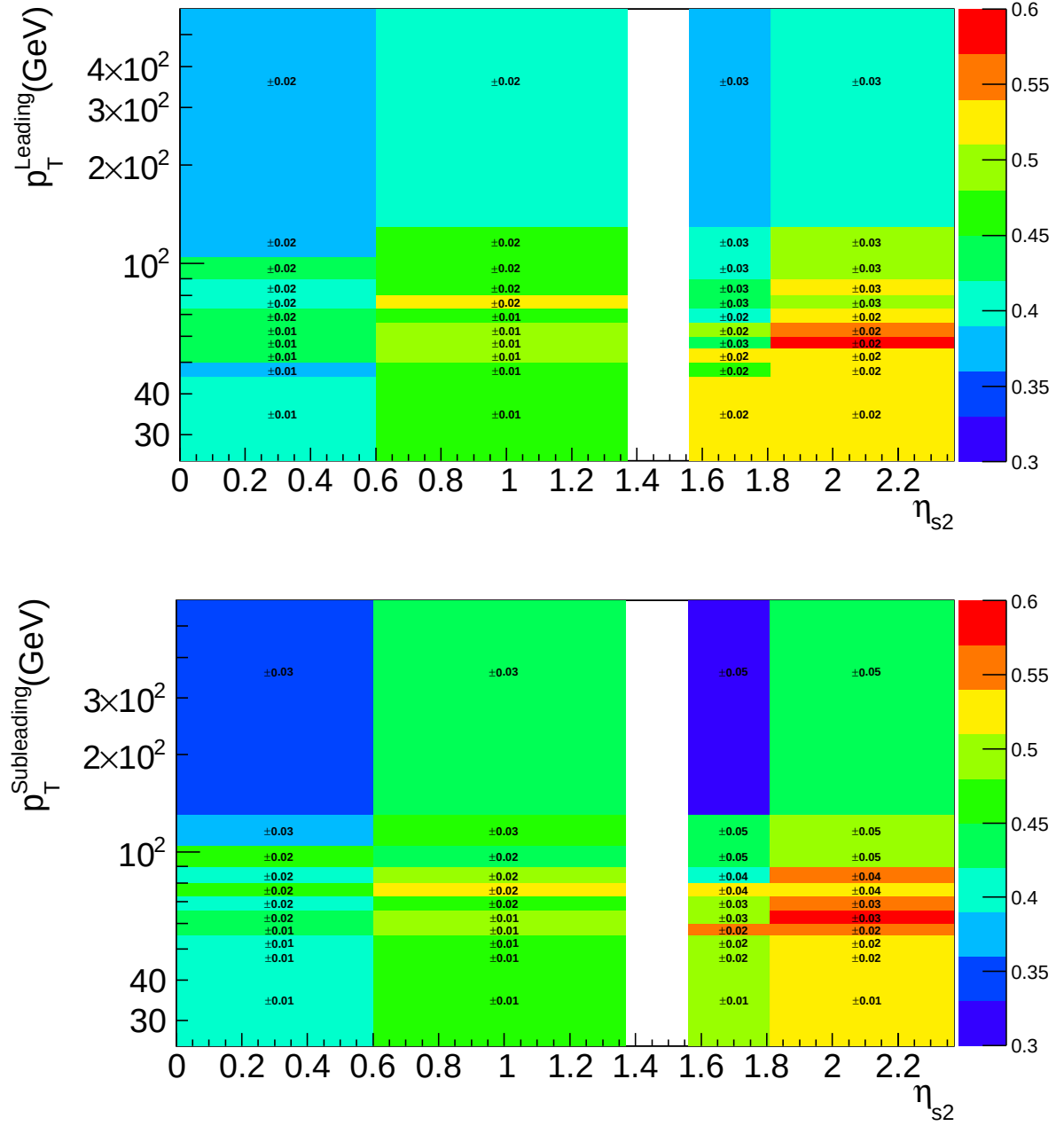


FIG. 98 – Taux de mauvaise identification $\text{jet} \rightarrow \gamma$ pour le premier (haut) et le deuxième (bas) candidat photon estimés à partir des données. La méthode de type *tag-and-probe* menant à ces résultats est décrite dans l'appendice B. Elle permet d'estimer la probabilité pour un jet reconstruit comme un photon LOOSE'4 de passer la sélection TIGHT, représentée ici. Les incertitudes mentionnées dans chaque catégorie sont statistiques uniquement.

Dans ce contexte, les échantillons simulés de bruit de fond (voir section 13.2.3) présentent un intérêt limité, particulièrement pour les bruits de fond hadroniques. En effet, ceux-ci sont générés avec une statistique très limitée et reproduisent probablement le taux de mauvaise identification $\text{jet} \rightarrow \gamma$ de façon approximative, à cause des incertitudes élevées dues aux modèles d’hadronisation non-perturbatifs. Par conséquent, il est préférable d’utiliser une méthode statistique de soustraction du fond basée au maximum sur les données, comme cela a été fait lors des analyses $\gamma\gamma + X$ précédentes [202, 203].

Les critères principaux permettant de discriminer signal et bruits de fond sont reliés aux sélections d’identification et aux variables d’isolation, désormais bien connues et déjà utilisées pour définir la région de signal (voir section 14). Ces critères permettent également la soustraction du bruit de fond à partir des données sous réserve de quelques hypothèses, comme nous le verrons. Une des particularités de l’analyse $\gamma\gamma + X$ vient des corrélations d’isolation non-négligeables entre les deux objets dans l’état final, ce qui complexifie le traitement du bruit fond par rapport à une analyse inclusive $\gamma + X$ par exemple.

Dans la section 15.2, les corrélations entre les deux objets dans l’état final sont étudiées, grâce notamment aux échantillons simulés et aux régions de contrôle définies dans les données. Dans la section 15.3, la méthode de soustraction de bruit de fond adoptée (*2D template fit*) est expliquée en détail et les différentes hypothèses conditionnant son bon fonctionnement sont mentionnées explicitement. Le *2D template fit* a déjà été utilisé dans les deux analyses $\gamma\gamma + X$ précédentes d’ATLAS [202, 203], mais également dans l’étude du canal $H \rightarrow \gamma\gamma$ [188, 189, 208]. Cependant, plusieurs améliorations y ont été apportées dans le cadre de l’analyse présentée dans ce manuscrit afin de réduire les incertitudes expérimentales. Dans la section 15.4, le *2D template fit* optimisé pour l’analyse est réalisé sur un jeu de pseudo-données créé spécialement afin de tester son bon fonctionnement et vérifier qu’aucun biais résiduel n’est présent. Les résultats de la méthode avec les données réelles sont ensuite dévoilés dans la section 15.5. La dépendance des résultats de la méthode vis-à-vis des hypothèses sous-jacentes est évaluée dans la section 15.6, menant au calcul des différentes incertitudes dues à la connaissance limitée du bruit de fond. Finalement, les résultats finals sont comparés avec une méthode de soustraction alternative dans la section 15.7, dont le principe est décrit succinctement dans l’appendice B.

15.2 Étude des corrélations entre les deux candidats

15.2.1 Rôle des corrélations dans l’analyse

Si aucune corrélation n’existait entre les variables des deux objets sélectionnés employées pour soustraire le bruit de fond (isolation, identification), celui-ci pourrait être extrait indépendamment pour les deux candidats, ce qui reviendrait à effectuer deux analyses $\gamma + X$ séparément à partir d’un échantillon di-photon. Cependant, des corrélations importantes sont observées pour les variables d’isolation. Ces corrélations varient en fonction du processus ($\gamma\gamma$, $\gamma + \text{jet}$, $\text{jet} + \gamma$, di-jet, ee) et doivent par conséquent être étudiées pour chacun d’entre eux. Il est important de bien estimer ces corrélations afin de formuler des hypothèses raisonnables qui mèneront à des incertitudes systématiques réduites lors de la soustraction du bruit de fond restant dans l’échantillon **TITI**.

En revanche, les corrélations des deux objets en terme d'identification sont faibles et arrivent comme un effet de second ordre. Cela peut se comprendre en considérant qu'au premier ordre, la forme de la gerbe dans le calorimètre est uniquement reliée à l'interaction de la particule avec le détecteur (type de particule + géométrie du détecteur), tandis que l'isolation dépend nécessairement de la physique sous-jacente : topologie de l'événement (présence de jets additionnels de haute énergie), événement sous-jacent/empilement (radiations à basse énergie), impulsion transverse et pseudo-rapidité E_T^γ et $|\eta^\gamma|$ (reliées aux fuites d'énergie hors du *cluster* de taille fixe soustrait), etc.

La bonne compréhension des corrélations est un des points clé de l'analyse, mais aussi l'un des plus difficiles à étudier. Dans la suite, nous limiterons leur étude à la région de signal **TITI**, échantillon sur lequel se base la mesure $\gamma\gamma + X$, et à l'isolation calorimétrique, qui est la variable d'isolation utilisée par le *template fit* 2D pour la discrimination du bruit de fond (voir section 15.3) et pour laquelle on peut s'attendre des corrélations importantes d'après les arguments exposés ci-dessus.

Dans la section 15.2.2, les corrélations d'isolation calorimétrique entre les deux objets sont estimées séparément pour tous les processus considérés dans l'échantillon **TITI**, directement à partir des données ou avec les échantillons simulés. Une analyse des corrélations pour chaque processus est réalisée dans les sections 15.2.3 à 15.2.5, permettant la formulation d'hypothèses qui seront utilisées par la suite pour la soustraction du bruit de fond (voir section 15.3).

15.2.2 Estimation des corrélations d'isolation à partir des régions de contrôle et des échantillons simulés

Les corrélations d'isolation calorimétrique entre les deux candidats pour les bruits de fond hadroniques (di-jet, γ +jet, jet+ γ , avec les jets identifiés en photon) sont estimées à partir des distributions 2D correspondantes dans leur région de contrôle respective, soit **TIT $\tilde{\text{T}}$** , **TIT $\tilde{\text{I}}$** et **TIT $\tilde{\text{I}}$** . Par ailleurs, la distribution 2D de l'isolation des traces ($p_{T,1}^{\text{track.iso.}}$ vs $p_{T,2}^{\text{track.iso.}}$) dans **TIT $\tilde{\text{T}}$** est également représentée à titre indicatif.

Les distributions associées au di-jet (**TIT $\tilde{\text{T}}$**) sont représentées dans la figure 99. Le facteur de corrélation entre les deux axes est respectivement de 6.2% et de 5.5% pour l'isolation calorimétrique et l'isolation des traces. La contamination de la région **TIT $\tilde{\text{T}}$** par les processus $\gamma\gamma$ et photon-jet est estimée respectivement à environ 2% et 31% avec le *template fit* 2D (voir section 15.5).⁵²

Les distributions associées à γ +jet (**TIT $\tilde{\text{I}}$**) et jet+ γ (**TIT $\tilde{\text{I}}$**) sont représentées dans la figure 100. Le facteur de corrélation entre les deux axes est respectivement de 4.1% et 4.7%. La contamination par les $\gamma\gamma$ constitue l'un des résultats du *template fit* 2D et est du même ordre de grandeur pour les deux régions (16% et 12%, respectivement). En revanche, la contamination di-jet, plus incertaine car estimable uniquement à partir du jeu de pseudo-données qui sera présenté dans la section 15.4, est sensiblement différente entre les deux (10% et 25%, respectivement) à cause du nombre plus élevé d'événements

⁵²Il est malheureusement difficile de ne pas anticiper sur quelques résultats du *template fit* 2D si l'on veut apporter certaines précisions supplémentaires utiles à la bonne compréhension de l'analyse. En pratique, celle-ci a donné lieu à un certain nombre d'itérations avant d'arriver aux différentes conclusions.

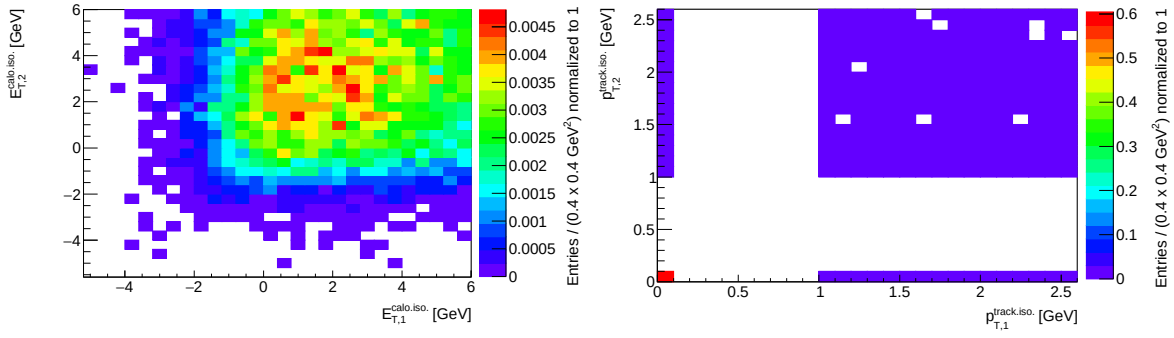


FIG. 99 – Distribution de l'isolation calorimétrique (gauche) et de l'isolation des traces (droite) des deux candidats photons dans la région de contrôle di-jet définie dans les données ($\tilde{\mathbf{T}}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$). Le facteur de corrélation entre les deux axes est respectivement de 6.2% et 5.5%. La contamination de la région par les processus $\gamma\gamma$ et photon-jet est estimée respectivement à environ 2% et 31% par le *template fit* 2D (voir section 15.5).

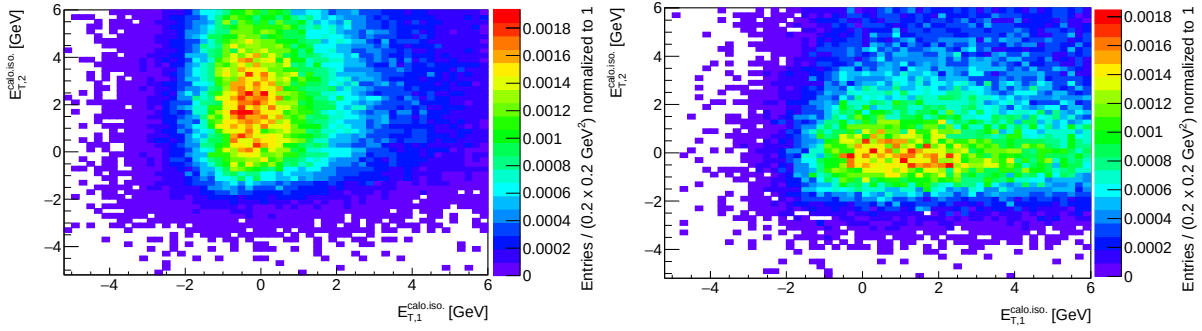


FIG. 100 – Distribution de l'isolation calorimétrique des deux candidats photons dans les régions de contrôle γ +jet ($\mathbf{T}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$, gauche) et jet+ γ ($\tilde{\mathbf{T}}\mathbf{I}\mathbf{T}\mathbf{I}$, droite) définie dans les données. Le facteur de corrélation entre les deux axes est respectivement 4.1% et 4.7%. La contamination des régions $\mathbf{T}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$ et $\tilde{\mathbf{T}}\mathbf{I}\mathbf{T}\mathbf{I}$ par le processus $\gamma\gamma$ est estimée respectivement à environ 16 et 12% par le *template fit* 2D (voir section 15.5). La contamination par les di-jets peut être estimée à partir du jeu de pseudo-données (voir section 15.4) et est de l'ordre de 10% et 25%, respectivement.

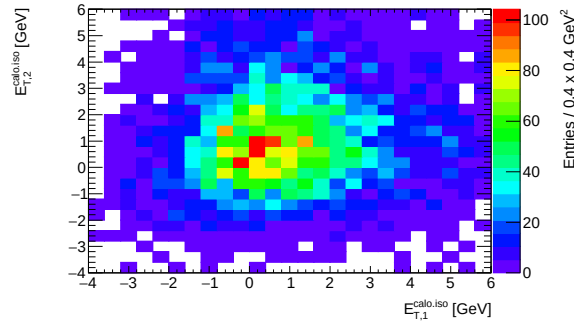


FIG. 101 – Distribution de l'isolation calorimétrique des candidats photons de l'échantillon MC DY dans la région $\mathbf{T}\mathbf{I}\mathbf{T}\mathbf{I}$. Le facteur de corrélation entre les deux axes est de 5.7%.

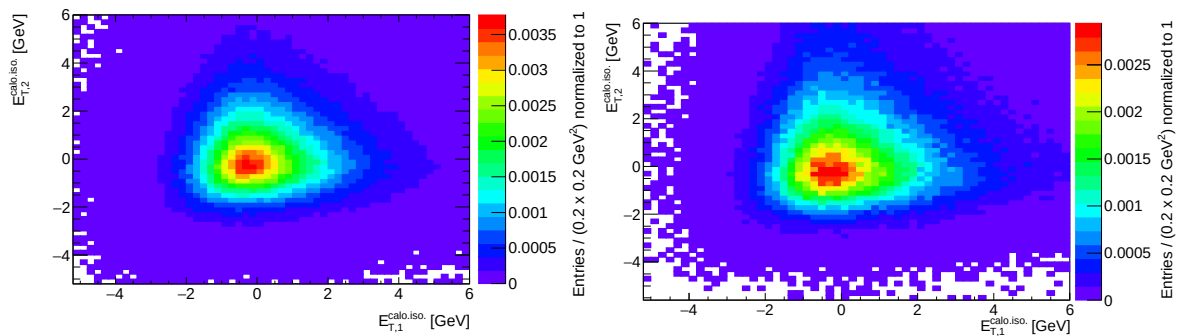


FIG. 102 – Distribution de l’isolation calorimétrique des candidats photons de l’échantillon MC de signal SHERPA dans la région **TITI** (gauche) et de l’ensemble des événements **TITI** dans les données réelles (droite). Le facteur de corrélation entre les deux axes est respectivement de 3.2% et 3.4%. D’après le *template fit* 2D, les événements $\gamma\gamma$ représentent environ 75% des événements de la région **TITI**.

γ +jet par rapport à jet+ γ (voir section 15.1). Dans tous les cas, la contamination des différentes régions de contrôle par ee est supposée négligeable, ce qui est largement vérifié dans les pseudo-données (voir section 15.4).

Les corrélations d’isolation calorimétrique entre les deux candidats pour le bruit de fond ee sont estimées à partir des échantillons DY simulés, les régions de contrôle *tag-and-probe* ne permettant pas leur sauvegarde (voir section 14.2). On ne s’attend pas à ce que les distributions d’isolation du MC reproduisent parfaitement les données, le taux de mauvaise identification $e \rightarrow \gamma$ étant corrigé uniquement en $|\eta^\gamma|$ et non en fonction de l’énergie d’isolation. Pour $\gamma\gamma$, l’échantillon MC SHERPA est utilisé. Cette fois, un bon accord est attendu grâce aux corrections supplémentaires d’isolation basées sur les données (voir section 10.4). La région **TITI** dans les données et avant soustraction du bruit de fond restant est également considérée afin d’évaluer grossièrement les corrélations d’isolation calorimétrique $\gamma\gamma$ et de comparer avec les résultats du MC.

La distribution simulée associée à ee est représentée dans la figure 101. Le facteur de corrélation entre les deux axes est de 5.7%. La simulation permet de s’affranchir de toutes les contaminations potentielles dont souffrent les régions de contrôle.

Les distributions associées à $\gamma\gamma$, incluant celle simulée avec SHERPA et celle correspondant à la région **TITI** des données, sont représentées dans la figure 102. Les facteurs de corrélation entre les deux axes sont 3.2% dans l’échantillon SHERPA et 3.4% dans l’échantillon **TITI** des données. D’après le *template fit* 2D, les événements $\gamma\gamma$ représentent environ 75% des événements de la région **TITI**.

15.2.3 Analyse et conclusion pour le processus di-jet

Les corrélations d’isolation calorimétrique dans le cas du bruit de fond di-jet sont prises en compte dans la méthode du *template fit* 2D (voir section 15.3), comme ce fut déjà le cas dans les deux analyses précédentes [202, 203]. Ce choix est motivé par le facteur de corrélation relativement élevé observé dans la région de contrôle (6.2%), qui est de plus probablement sous-estimé à cause de la forte contamination photon-jet ($\simeq 31\%$). De telles corrélations sont attribuées aux flux de couleur significatifs attendus

entre les deux partons dans l'état final, qui sont constitués majoritairement de gluons.

L'effet de négliger ces corrélations pour la soustraction du fond a été évalué avec les pseudo-données (voir section 15.4) et correspond à un biais d'environ $+0.3\%$ sur les résultats, finalement faible comparé à l'incertitude systématique totale, qui est de l'ordre de 2.6% (voir section 15.6).⁵³

15.2.4 Analyse et conclusion pour le processus di-photon

Pour les di-photons, le facteur de corrélation observé dans la simulation et les données est plus faible et en bon accord ($\simeq 3.2\%$). Dans les analyses précédentes [202, 203], son effet sur la soustraction du bruit de fond a systématiquement été négligé.

Cette hypothèse simplificatrice a été testée dans le cas présent avec les pseudo-données (voir section 15.4) et a mis en évidence l'apparition d'un biais de l'ordre de $+1.3\%$ sur les résultats de soustraction du bruit de fond, ce qui n'est pas négligeable comparé à l'incertitude systématique totale, qui est de l'ordre de 2.6% (voir section 15.6).⁵⁴ Par conséquent, cette hypothèse a été retirée pour la soustraction du bruit de fond avec le *template fit* 2D dans l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV, qui prend désormais en compte les corrélations d'isolation calorimétrique pour $\gamma\gamma$.

Une étude a été menée en parallèle afin d'essayer d'identifier plus précisément l'origine de ces corrélations avec l'échantillon de signal SHERPA. La corrélation observée dans la région **TITI** de l'échantillon MC (3.2%) reste significativement non nulle y compris lorsque l'espace des phases est restreint à un domaine où les cônes d'isolation des deux photons sont totalement disjoints. Par exemple pour $\Delta R_{\gamma\gamma} > 0.8$, le facteur de corrélation s'élève toujours à 2.8% . Plusieurs autres effets potentiels ont été testés (plus haut $\Delta R_{\gamma\gamma}$, plus bas $\langle \mu \rangle$, etc) et le critère correspondant à demander une différence importante en impulsion transverse entre les deux photons a été identifié comme celui abaissant le plus les corrélations d'isolation. Ainsi, $\Delta E_{T,1,2}^\gamma > 100$ GeV permet d'abaisser les corrélations à 0.9% .

Des E_T^γ similaires menant à des fuites d'énergie du photon dans le cône d'isolation du même ordre de grandeur, la source de corrélation principale est attribuée à l'effet suivant : l'énergie du photon fuyant hors du *cluster* de taille fixe qui est soustrait lors du calcul de l'isolation calorimétrique (voir section 10.3) peut être représentée comme une variable aléatoire distribuée selon une fonction de type boule de cristal. Cependant, les corrections de fuite associées correspondent uniquement au pic de cette distribution (valeur de fuite la plus probable), de telle façon que des corrélations résiduelles sont attendues à cause des queues de distribution asymétriques.

L'effet de l'événement sous-jacent (UE pour *underlying event*) a également été évalué grossièrement en répétant l'étude avec l'échantillon de signal PYTHIA, qui inclut un modèle d'UE différent de SHERPA. Un facteur de corrélation très proche est alors obtenu (3.14%), ce qui ne permet pas de conclure définitivement, les deux modèles d'UE pouvant dans ce cas précis donner des résultats similaires. Malheureusement, aucun échantillon de signal simulé avec UE désactivé et simulation complète du détecteur n'a été produit pour

⁵³Ces différences sont données relativement à la fraction d'événements $\gamma\gamma$ dans **TITI** (par exemple $70\% \pm 2\%$).

⁵⁴Voir note de bas de page numéro 53.

en apprendre plus sur ces corrélations. En revanche, les isolations des deux photons au niveau particule ont été étudiées avec l'échantillon de signal simulé SHERPA et montrent des corrélations même légèrement supérieures à celle du niveau reconstruit, de l'ordre de 4-5%. Cette observation semble indiquer un rôle non négligeable de l'événement sous-jacent qu'il serait bon d'étudier plus en détail dans les analyses futures.

15.2.5 Analyse et conclusion pour les processus photon-jet et ee

L'effet des corrélations d'isolation dans les événements de bruit de fond ee a été évalué avec les pseudo-données (voir section 15.4) et a un effet négligeable sur les résultats malgré le facteur de corrélation relativement important observé dans les échantillons simulés (5.7%). Cela est probablement lié au fait que les photons et les électrons ont une distribution d'isolation calorimétrique proche (bien que significativement différente, voir section 14.3.2) et que le nombre d'événements ee dans l'échantillon **TITI** est très réduit ($\simeq 3\%$, voir section 15.5). Par conséquent, les corrélations d'isolation pour ee seront négligées dans la suite de l'analyse. Cependant, le fait que les photons et les électrons aient des distributions d'isolation calorimétrique significativement différentes a des conséquences importantes. En effet, cela oblige à soustraire en même temps le bruit de fond hadronique et le bruit de fond ee des données (voir section 15.4), et donc à ajouter au modèle d'ajustement une composante ee . Dans les analyses précédentes d'ATLAS [202, 203], les distributions d'isolation entre photons et électrons étaient supposées identiques *de facto*, de telle sorte que les jets et les électrons pouvaient être soustraits les uns après les autres.

Finalement, l'effet des corrélations d'isolation calorimétrique dans les événements photon-jet ($\simeq 4\%$) reste incertain car très difficilement évaluable. En effet, comme nous le verrons dans la section 15.4, les pseudo-données négligent par construction de telles corrélations au premier ordre et ne peuvent donc par conséquent pas être utilisées pour évaluer leur impact. Par ailleurs, plusieurs arguments laissent penser qu'elles ont un impact bien plus limité sur la soustraction du fond que dans le cas $\gamma\gamma$ et di-jet :

- ◊ Contrairement aux événements di-jets, aucun flux de couleur n'est attendu entre le photon et le parton dans l'état final. Ce dernier devrait être principalement constitué de quarks légers.
- ◊ Le nombre d'événements photon-jet dans l'échantillon **TITI** des données est entre 3 et 4 fois plus faible que $\gamma\gamma$.

On s'attend par ailleurs à ce que les effets des corrélations augmentent à haute énergie d'isolation à cause du recouvrement partiel des cônes des deux candidats. La méthode du *template fit* 2D se basant exclusivement sur des régions bien isolées **II**, cela devrait lui permettre de garder ces effets sous contrôle. Cette particularité devrait également permettre à la méthode de garder sous contrôle les effets reliés aux corrélations entre identification et isolation pour un même candidat, qui augmentent également avec l'énergie d'isolation.

Dans la section 15.3, la méthode statistique utilisée pour soustraire à la fois les jets et les électrons mal identifiés de la région **TITI** dans l'analyse principale est expliquée en détail.

15.3 La méthode du *template fit* 2D

15.3.1 Aperçu

La méthode du *template fit* a déjà été introduite dans la section 10.4.2 dans le cas 1D pour le calcul des corrections de fuite pour l'isolation, également réalisé pendant la thèse. Elle est ici abordée à nouveau, dans le cas 2D.

Le *template fit* 2D est une méthode statistique se basant sur l'ajustement des distributions d'isolation calorimétrique à deux dimensions ($E_{T,1}^{\text{calo.iso.}}$, $E_{T,2}^{\text{calo.iso.}}$) des différents événements présents dans l'échantillon **TITI** ($\gamma\gamma$, γ +jet, jet+ γ , di-jet, ee) aux données correspondantes enregistrées par le détecteur ATLAS, afin d'en obtenir la composition.⁵⁵ Elle repose ainsi sur la modélisation adéquate de la distribution d'isolation calorimétrique des photons immédiats, du bruit de fond (jet et électrons) et des corrélations d'isolation entre les deux candidats dans les différents processus (voir section 15.2). Elle a déjà été mise en œuvre dans plusieurs analyses de données de photons au LHC, dont les trois analyses $\gamma\gamma + X$ à 7 TeV réalisées par ATLAS et CMS [202, 203, 224].

Les modèles d'isolation calorimétrique pour les bruits de fond sont logiquement extraits des régions de contrôle définies dans la section 14. Pour les photons immédiats, ceux-ci sont extraits de l'échantillon de signal SHERPA $\gamma\gamma$. On pourrait également songer à utiliser un échantillon de photons pur dans les données à l'image de ce qui est fait pour les mesures d'identification (voir section 9.3), consistant par exemple à sélectionner les photons radiatifs du $Z^0 \rightarrow l^+l^-\gamma$. Cela est en pratique fortement déconseillé ici pour les raisons suivantes :

- ◊ Les photons radiatifs du Z^0 présentent une gamme en $E_T^\gamma / |\eta^\gamma|$ restreinte (en particulier $E_T^\gamma < 100$ GeV) et propre au processus $Z \rightarrow l^+l^-\gamma$. Or, il a été vu dans les sections 10 et 14.3.2 que les distributions d'isolation calorimétrique des photons ont une dépendance significative à la fois en E_T^γ et en $|\eta^\gamma|$. L'utilisation des photons radiatifs du $Z \rightarrow l^+l^-\gamma$ nécessiterait alors une repondération massive des événements et une extrapolation importante à haut E_T^γ , probablement limitantes pour la mesure. L'accord données/MC pour ces photons a néanmoins été vérifié dans le cadre de plusieurs études, notamment pour l'étude de la désintégration du boson de Higgs en deux photons [189, 208]. Un bon accord général est observé.
- ◊ Il est attendu que l'isolation calorimétrique des di-photons dépende du nombre et de la distribution des jets additionnels de haute énergie émis par le processus dur, des photons de fragmentation et de l'événement sous-jacent (voir section 14.3.2). Toutes ces caractéristiques liées à la QCD et propres au signal $\gamma\gamma + X$ sont perdues (voir biaisées) si les distributions d'isolation des photons radiatifs du $Z \rightarrow l^+l^-\gamma$ sont utilisées, qui sont eux dominés par des effets QED.

La simulation a par ailleurs été corrigée afin de reproduire correctement les distributions d'isolation des photons dans les données, ce qui a donné lieu au travail de

⁵⁵Un ajustement des distributions d'isolation des traces serait peu discriminant, la statistique étant concentrée en très grande majorité dans la catégorie correspondant à $p_{T,1}^{\text{track.iso.}} = p_{T,2}^{\text{track.iso.}} = 0$ pour l'ensemble des processus (voir section 14.3.1).

performance principal effectué pendant la thèse, décrit dans la section 10.4. La quantification de cet accord sera effectuée dans le cadre de l'analyse lors de l'estimation des incertitudes systématiques associées à la soustraction du bruit de fond (voir section 15.6).

Dans la section 15.3.2, quelques informations sont apportées sur la procédure de minimisation. Tous les détails pratiques concernant l'implémentation du modèle d'ajustement sont ensuite donnés dans la section 15.3.3.

15.3.2 Procédure de minimisation

La procédure d'ajustement est effectuée dans l'intervalle d'isolation calorimétrique $[-5.2 \text{ GeV}, 6 \text{ GeV}] \times [-5.2 \text{ GeV}, 6 \text{ GeV}]$ en minimisant l'estimateur statistique de maximum de vraisemblance étendu (*extended maximum likelihood*) associé au modèle à l'aide du code ROOFIT [205], lui-même basé sur le code MINUIT du CERN [206]. Le maximum de vraisemblance est un estimateur classique employé lors de procédures d'ajustement. La borne inférieure -5.2 GeV est choisie afin d'inclure le maximum de données disponibles tout en limitant les fluctuations statistiques, qui deviennent importantes à plus basse énergie d'isolation. Cela mène à l'exclusion de 88 événements de la procédure d'ajustement, sur les 312 754 événements présents dans l'échantillon **TITI**. Cet effet minime sera négligé dans la suite.

La borne supérieure est fixée à 6 GeV de manière à ce que le *template fit* considère uniquement des régions **II**. Ce choix est motivé par le peu de données présent à plus haute énergie d'isolation du fait de la coupure sur l'isolation des traces, qui réduit fortement les queues de distribution. Par ailleurs, les corrélations d'isolation entre les deux candidats et les corrélations entre identification et isolation pour un même candidat augmentant à plus haute énergie d'isolation, il est préférable de rester autant que possible dans la région **II** (voir section 15.2). Les analyses précédentes [202, 203] étendaient l'intervalle d'ajustement au-delà de la région **II** afin d'augmenter la statistique des données, car celles-ci bénéficiaient d'une statistique moins importante et de queues de distribution chargées en bruit de fond plus conséquentes (aucune isolation des traces n'était appliquée).

L'estimateur de maximum de vraisemblance utilisé pour l'ajustement est catégorisé (*binned*) ou non (*unbinned*) en fonction de la statistique MC présente dans l'échantillon de signal. Les deux estimateurs donnent les mêmes résultats dans la limite à grande statistique. Lorsque la statistique MC est haute (typiquement au-dessus de 20 000 événements), l'ajustement est effectué dans des catégories de largeur $0.4 \text{ GeV} \times 0.4 \text{ GeV}$ dans l'intervalle $[-5.2 \text{ GeV}, 6 \text{ GeV}] \times [-5.2 \text{ GeV}, 6 \text{ GeV}]$ mentionné ci-dessus pour des raisons de performance (temps d'exécution). C'est par exemple toujours le cas pour la soustraction du fond dans la région **TITI** complète. En revanche, pour la décomposition du bruit de fond selon les différents spectres, deux cas de figures se présentent : soit la statistique MC est élevée ($> 20\,000$ événements) et l'ajustement est réalisé dans des catégories étendues de largeur $0.8 \text{ GeV} \times 0.8 \text{ GeV}$ afin de tenir compte de la baisse de statistique dans les données par rapport au cas inclusif, soit la statistique MC est plus faible et un ajustement non catégorisé est effectué.

Dans le cas d'un ajustement non catégorisé, il est nécessaire d'extraire des distributions d'isolation calorimétrique non catégorisées (soit continues ou *unbinned*) pour lancer la procédure. En pratique, les distributions jets ou électrons, basées sur les don-

nées, sont toujours extraites de façon non catégorisée et les distributions photons, issues du MC, sont extraites de façon catégorisée ou non selon le type d'ajustement. Les distributions d'isolation observées dans les données et le MC ne pouvant être paramétrées analytiquement de manière fiable pour tout l'espace des phases, une technique basée sur des estimateurs de noyaux gaussiens (KDE, *kernel density estimator*) [207] est utilisée, déjà introduite dans la section 10.4.2). Elle s'appuie sur une technique de lissage gaussien (*smoothing*) pour extraire des distributions continues (*unbinned*) à partir de distributions limitées en statistique (données ou MC). Le biais potentiel venant du lissage est largement sous contrôle et sera évalué lors du traitement des incertitudes systématiques (voir section 15.6). Un modèle *unbinned* permet d'utiliser l'estimateur du maximum de vraisemblance étendu *unbinned*, préconisé car bien plus robuste pour les régions de l'espace des phases souffrant d'une statistique limitée [244].

Dans la section 15.3.3, toutes les explications nécessaires à la compréhension détaillée du modèle d'ajustement utilisé dans l'analyse pour la soustraction du fond sont fournies.

15.3.3 Détail du modèle d'ajustement adopté

L'écriture la plus simple du modèle d'ajustement adopté dans l'analyse se présente de la façon suivante :

$$W_{\text{tot}}^{\text{TITI}} F_{\text{tot}}^{\text{TITI}} = W_{\gamma\gamma}^{\text{TITI}} F_{\gamma\gamma}^{\text{TITI}} + W_{\gamma j}^{\text{TITI}} F_{\gamma j}^{\text{TITI}} + W_{j\gamma}^{\text{TITI}} F_{j\gamma}^{\text{TITI}} + W_{jj}^{\text{TITI}} F_{jj}^{\text{TITI}} + W_{ee}^{\text{TITI}} F_{ee}^{\text{TITI}}, \quad (48)$$

où $W_{\text{tot}}^{\text{TITI}}$ est le nombre total d'événements attendus dans l'échantillon **TITI** et les W_i^{TITI} sont les nombres d'événements attendus pour les différents processus, avec $W_{\text{tot}}^{\text{TITI}} = \sum_i W_i^{\text{TITI}}$. Les W_i^{TITI} sont laissés libres pendant l'ajustement moyennant une contrainte poissonnienne sur $W_{\text{tot}}^{\text{TITI}}$ centrée sur le nombre d'événements observés dans la région **TITI** dans les données, issue de la définition de l'estimateur de maximum de vraisemblance étendu. F_{tot} représente la fonction densité de probabilité (ou pdf pour *probability density function*) globale attendue pour l'isolation calorimétrique de l'échantillon **TITI**, tandis que les F_i représentent les pdfs individuelles attendus pour les différents processus.

Par ailleurs, il a été vu précédemment que les corrélations d'isolation calorimétrique entre les candidats photons dans les processus photon-jet et ee devraient avoir un impact faible sur les résultats (voir section 15.2). Celles-ci sont donc négligées pour la soustraction du bruit de fond et les pdfs correspondants à γ +jet, jet+ γ et ee sont factorisées comme le produit de deux pdfs 1D. En raison de la dépendance en E_T^γ des distributions d'isolation relevée dans la section 14.3.2 à la fois pour les photons immédiats et le bruit de fond, une distinction est toutefois effectuée entre les pdfs 1D du premier et du deuxième candidat. En revanche, les corrélations d'isolation pour $\gamma\gamma$ et les di-jets peuvent avoir un effet important sur les résultats et les pdfs 2D complètes sont gardées pour ces deux processus. Le modèle de l'équation 48 peut alors se réécrire de la façon suivante :

$$W_{\text{tot}}^{\text{TITI}} F_{\text{tot}}^{\text{TITI}} = W_{\gamma\gamma}^{\text{TITI}} F_{\gamma\gamma}^{\text{TITI}} + W_{jj}^{\text{TITI}} F_{jj}^{\text{TITI}} + W_{\gamma j}^{\text{TITI}} F_{\gamma,1}^{\text{TITI}} F_{j,2}^{\text{TITI}} + W_{j\gamma}^{\text{TITI}} F_{j,1}^{\text{TITI}} F_{\gamma,2}^{\text{TITI}} + W_{ee}^{\text{TITI}} F_{e,1}^{\text{TITI}} F_{e,2}^{\text{TITI}}, \quad (49)$$

où $F_{\gamma,1}^{\text{TITI}}$ et $F_{\gamma,2}^{\text{TITI}}$ sont les pdfs 1D pour le premier et deuxième photon et $F_{j,1}^{\text{TITI}}$ et $F_{j,2}^{\text{TITI}}$ ($F_{e,1}^{\text{TITI}}$ et $F_{e,2}^{\text{TITI}}$) sont les pdfs 1D correspondantes pour les jets (électrons) mal identifiés.

Une complication supplémentaire vient de la contamination des régions de contrôle des jets utilisées pour l'extraction des pdfs par les autres processus (signal, autres bruits de fond jets). La contamination des régions de contrôle de jets par les électrons est négligée. Les différentes contaminations dans les régions de contrôle *tag-and-probe* utilisées pour les pdfs électrons sont également négligées. Ces hypothèses simplificatrices ont un effet nul dans les résultats finals d'après les pseudo-données (voir section 15.4). En pratique, le modèle correspond alors à :

$$W_{\text{tot}}^{\text{TITI}} F_{\text{tot}}^{\text{TITI}} = w_{\gamma\gamma}^{\text{TITI}} F_{\gamma\gamma}^{\text{TITI}} + w_{bb}^{\text{TITI}} F_{bb}^{\tilde{\text{TITI}}} + w_{\gamma b}^{\text{TITI}} F_{\gamma,1}^{\text{TITI}} F_{b,2}^{\tilde{\text{TITI}}} + w_{b\gamma}^{\text{TITI}} F_{b,1}^{\tilde{\text{TITI}}} F_{\gamma,2}^{\text{TITI}} + W_{ee}^{\text{TITI}} F_{e,1}^{\text{TITI}} F_{e,2}^{\text{TITI}}, \quad (50)$$

où $F_{b,1}^{\tilde{\text{TITI}}}$ ($F_{b,2}^{\tilde{\text{TITI}}}$) est la pdf 1D du premier (deuxième) candidat photon dans la région $\tilde{\text{TITI}}$ (TITI), extrapolée de façon inchangée à la région TITI . De la même façon, $F_{bb}^{\tilde{\text{TITI}}}$ correspond à la pdf 2D extraite depuis la région $\tilde{\text{TITI}}$. Cette extrapolation suppose la non-corrélation des critères d'identification et d'isolation,⁵⁶ l'une des hypothèses fortes de la méthode pour la modélisation du bruit de fond. Ce point sera discuté plus en détail lors de l'estimation des incertitudes systématiques (voir section 15.6).

Les w_{ik}^{TITI} ne correspondent alors plus exactement au nombre d'événements *ik* attendus pour les différents processus à cause de la contamination des régions de contrôle (par exemple, $F_{b,1}^{\text{TITI}} = (1 - \alpha_1) F_{j,1}^{\text{TITI}} + \alpha_1 F_{\gamma,1}^{\text{TITI}}$ et non $F_{j,1}^{\text{TITI}}$ directement). Ils seront appelés "nombres d'événements ajustés" dans la suite et notés avec une minuscule, tandis que les nombres réels d'événements attendus seront notés systématiquement avec une majuscule (W_{ik}^{TITI}).⁵⁷ Les formules permettant de passer de l'un à l'autre ainsi que le calculs des différents coefficients associés (α_1 , etc) seront décrits en détail dans la section 15.3.4.

Finalement, les pdfs pour les photons et les électrons sont extraites respectivement depuis l'échantillon de signal SHERPA et depuis les régions de contrôle *tag-and-probe* et ne souffrent pas de contaminations significatives pour les résultats. Cependant, les différences entre les distributions d'isolation des électrons mal identifiés et des photons (voir section 14.3.2), bien qu'assez fortes pour biaiser les résultats si l'on ne considère pas deux pdfs séparées pour $\gamma\gamma$ et ee dans le modèle (voir section 15.4), ne le sont pas assez pour les discriminer de manière satisfaisante dans la procédure d'ajustement.⁵⁸ Par

⁵⁶Ou de façon plus correcte, une corrélation similaire entre isolation et identification dans TI et $\tilde{\text{TI}}$.

⁵⁷Comme mentionné plus haut, toutes les contaminations reliées aux électrons sont négligées de sorte que $w_{ee}^{\text{TITI}} = W_{ee}^{\text{TITI}}$.

⁵⁸Cela peut se comprendre en considérant que W_{ee}^{TITI} est le plus souvent du même ordre de grandeur (cas inclusif ou région proche de m_Z) ou inférieure (ailleurs dans les spectres) à l'erreur systématique totale sur $W_{\gamma\gamma}^{\text{TITI}}$.

conséquent, l'estimateur de maximum de vraisemblance est multiplié par une contrainte gaussienne externe afin d'aider à la convergence. Celle-ci suit une distribution gaussienne fonction de $W_{ee}^{\mathbf{TITI}}$ centrée sur la prédiction de l'échantillon MC DY correspondant (voir section 13). La largeur de la gaussienne est prise comme la somme en quadrature entre l'erreur statistique MC de l'échantillon DY et l'erreur systématique sur les facteurs de correction des taux de mauvaise identification $e \rightarrow \gamma$ [234] (dominante), ce qui assure des prédictions MC fiables pour le bruit de fond ee venant du $Z \rightarrow ee$ (voir section 13).

D'autres processus sous-dominants (comme $t\bar{t}$, ZZ ou WW) pouvant générer le bruit de fond ee , on s'attend à ce que les prédictions du MC soient des bornes inférieures au nombre d'événements ee présents dans l'échantillon **TITI**. Il est enfin rappelé que les contaminations des électrons dans les autres régions de contrôle sont négligées, ce qui a un effet nul dans les résultats finals d'après les pseudo-données (voir section 15.4) et induit la relation $w_{ee}^{\mathbf{TITI}} = W_{ee}^{\mathbf{TITI}}$.

Dans la section 15.3.4, les corrections permettant de passer des nombres d'événements ajustés $w_{ik}^{\mathbf{TITI}}$ aux nombres réels d'événements attendus $W_{ik}^{\mathbf{TITI}}$ sont explicitées.

15.3.4 Corrections des nombres d'événements ajustés ($w_{ik}^{\mathbf{TITI}} \rightarrow W_{ik}^{\mathbf{TITI}}$)

Les pdfs $F_{b,1}^{\tilde{\mathbf{TITI}}}$, $F_{b,2}^{\tilde{\mathbf{TITI}}}$ et $F_{bb}^{\tilde{\mathbf{TITI}}}$ introduites dans l'équation 50 de la section 15.3.3 peuvent être exprimées en fonction des pdfs jets et photons de la façon suivante :

$$F_{b,1}^{\tilde{\mathbf{TITI}}} = (1 - \alpha_1)F_{j,1}^{\tilde{\mathbf{TITI}}} + \alpha_1 F_{\gamma,1}^{\tilde{\mathbf{TITI}}}, \quad (51)$$

$$F_{b,2}^{\mathbf{TITI}} = (1 - \alpha_2)F_{j,2}^{\mathbf{TITI}} + \alpha_2 F_{\gamma,2}^{\mathbf{TITI}}, \quad (52)$$

$$F_{bb}^{\tilde{\mathbf{TITI}}} = (1 - \alpha'_1 - \alpha'_2 - \alpha'_3)F_{jj}^{\tilde{\mathbf{TITI}}} + \alpha'_1 F_{j,1}^{\tilde{\mathbf{TITI}}} F_{\gamma,2}^{\tilde{\mathbf{TITI}}} + \alpha'_2 F_{\gamma,1}^{\tilde{\mathbf{TITI}}} F_{j,2}^{\tilde{\mathbf{TITI}}} + \alpha'_3 F_{\gamma,1}^{\tilde{\mathbf{TITI}}} F_{\gamma,2}^{\tilde{\mathbf{TITI}}}, \quad (53)$$

où α_1 (α_2) est la fraction d'événements $\gamma\gamma$ dans les régions de contrôle **TITI** (**TITI**). De la même façon, α'_1 (α'_2 , α'_3) représente la fraction d'événements jet+ γ (γ +jet, $\gamma\gamma$) dans l'échantillon **TITI**. Comme précédemment, les pdfs jets et photons des régions de contrôle sont identifiées aux pdfs correspondantes dans la région de signal **TITI**, ce qui revient à faire l'hypothèse que les corrélations entre isolation et identification dans **TI** et **TITI** sont similaires (voir note de bas de page numéro 56).

En principe, la contamination des événements di-jets dans les régions de contrôle γ +jet et jet+ γ devrait également être prise en compte dans les équations 51 à 53. Cependant, d'après les tests effectués sur les pseudo-données (voir section 15.4), les projections 1D de F_{jj} sont très similaires aux pdfs $F_{j,1}$, $F_{j,2}$ correspondantes. Par conséquent, ces contaminations seront négligées dans la suite.

Les facteurs α_i et α'_i peuvent être obtenus à partir des nombres réels d'événements attendus $W_{ik}^{\mathbf{TITI}}$ et des rapports entre les événements **TITI**/**T** dans les différents échantillons

simulés à partir des formules suivantes :

$$\alpha_1(W_{\gamma\gamma}^{\text{TITI}}) = \frac{n_{\gamma\gamma,MC}^{\tilde{\text{TITI}}}}{n_{\gamma\gamma,MC}^{\text{TITI}}} \frac{W_{\gamma\gamma}^{\text{TITI}}}{n^{\tilde{\text{TITI}}}}, \quad (54)$$

$$\alpha_2(W_{\gamma\gamma}^{\text{TITI}}) = \frac{n_{\gamma\gamma,MC}^{\text{TITI}}}{n_{\gamma\gamma,MC}^{\tilde{\text{TITI}}}} \frac{W_{\gamma\gamma}^{\text{TITI}}}{n^{\text{TITI}}}, \quad (55)$$

$$\alpha'_1(W_{j\gamma}^{\text{TITI}}) = \frac{n_{j\gamma,MC}^{\tilde{\text{TITI}}}}{n_{j\gamma,MC}^{\text{TITI}}} \frac{W_{j\gamma}^{\text{TITI}}}{n^{\tilde{\text{TITI}}}}, \quad (56)$$

$$\alpha'_2(W_{\gamma j}^{\text{TITI}}) = \frac{n_{\gamma j,MC}^{\tilde{\text{TITI}}}}{n_{\gamma j,MC}^{\text{TITI}}} \frac{W_{\gamma j}^{\text{TITI}}}{n^{\tilde{\text{TITI}}}}, \quad (57)$$

$$\alpha'_3(W_{\gamma\gamma}^{\text{TITI}}) = \frac{n_{\gamma\gamma,MC}^{\tilde{\text{TITI}}}}{n_{\gamma\gamma,MC}^{\text{TITI}}} \frac{W_{\gamma\gamma}^{\text{TITI}}}{n^{\tilde{\text{TITI}}}}, \quad (58)$$

où $n_{ik,MC}^{\text{XY}}$ est le nombre d'événements de type ik dans la région XY mesurés dans les échantillons MC correspondants et n^{XY} est le nombre total d'événements dans la région XY observés dans les données.

Afin de manipuler des quantités indépendantes des nombres réels d'événements attendus W_{ik}^{TITI} , on définit également les facteurs β_i et β'_i :

$$\beta_1 = \frac{\alpha_1(W_{\gamma\gamma}^{\text{TITI}})}{W_{\gamma\gamma}^{\text{TITI}}} = \frac{n_{\gamma\gamma,MC}^{\tilde{\text{TITI}}}}{n_{\gamma\gamma,MC}^{\text{TITI}}} \frac{1}{n^{\tilde{\text{TITI}}}}, \quad (59)$$

$$\beta_2 = \frac{\alpha_2(W_{\gamma\gamma}^{\text{TITI}})}{W_{\gamma\gamma}^{\text{TITI}}} = \frac{n_{\gamma\gamma,MC}^{\text{TITI}}}{n_{\gamma\gamma,MC}^{\tilde{\text{TITI}}}} \frac{1}{n^{\text{TITI}}}, \quad (60)$$

$$\beta'_1 = \frac{\alpha'_1(W_{j\gamma}^{\text{TITI}})}{W_{j\gamma}^{\text{TITI}}} = \frac{n_{j\gamma,MC}^{\tilde{\text{TITI}}}}{n_{j\gamma,MC}^{\text{TITI}}} \frac{1}{n^{\tilde{\text{TITI}}}}, \quad (61)$$

$$\beta'_2 = \frac{\alpha'_2(W_{\gamma j}^{\text{TITI}})}{W_{\gamma j}^{\text{TITI}}} = \frac{n_{\gamma j,MC}^{\tilde{\text{TITI}}}}{n_{\gamma j,MC}^{\text{TITI}}} \frac{1}{n^{\tilde{\text{TITI}}}}, \quad (62)$$

$$\beta'_3 = \frac{\alpha'_3(W_{\gamma\gamma}^{\text{TITI}})}{W_{\gamma\gamma}^{\text{TITI}}} = \frac{n_{\gamma\gamma,MC}^{\tilde{\text{TITI}}}}{n_{\gamma\gamma,MC}^{\text{TITI}}} \frac{1}{n^{\tilde{\text{TITI}}}}. \quad (63)$$

Finalement, il est possible d'obtenir à partir des β_i et β'_i des relations (non-linéaires) entre les nombres réels d'événements attendus W_{ik}^{TITI} et les nombres d'événements ajustés w_{ik}^{TITI} introduits dans la section 15.3.3 :

$$W_{\gamma\gamma}^{\text{TITI}} = \frac{w_{\gamma\gamma}^{\text{TITI}}}{1 - \beta_1 w_{b\gamma}^{\text{TITI}} - \beta_2 w_{\gamma b}^{\text{TITI}} - \beta'_3 w_{bb}^{\text{TITI}}}, \quad (64)$$

$$W_{\gamma j}^{\text{TITI}} = \frac{w_{\gamma b}^{\text{TITI}}[1 - \alpha_2(W_{\gamma\gamma}^{\text{TITI}})]}{1 - \beta'_2 w_{bb}^{\text{TITI}}}, \quad (65)$$

$$W_{j\gamma}^{\text{TITI}} = \frac{w_{b\gamma}^{\text{TITI}}[1 - \alpha_1(W_{\gamma\gamma}^{\text{TITI}})]}{1 - \beta'_1 w_{bb}^{\text{TITI}}}, \quad (66)$$

$$W_{jj}^{\text{TITI}} = w_{bb}^{\text{TITI}}[1 - \alpha'_1(W_{j\gamma}^{\text{TITI}}) - \alpha'_2(W_{j\gamma}^{\text{TITI}}) - \alpha'_3(W_{\gamma\gamma}^{\text{TITI}})], \quad (67)$$

$$W_{ee}^{\text{TITI}} = w_{ee}^{\text{TITI}}. \quad (68)$$

Le MC est donc utilisé pour l'extraction des pdfs photons et pour l'évaluation des rapports $\mathbf{T}/\tilde{\mathbf{T}}$ figurant dans les équations 64 à 67 dans le cadre du *template fit 2D*. Cependant, il est important de remarquer que pour $W_{\gamma\gamma}^{\text{TITI}}$, uniquement les échantillons de signal sont nécessaires, les échantillons de bruit de fond photon-jet servant uniquement à corriger $w_{b\gamma}^{\text{TITI}}$, $w_{\gamma b}^{\text{TITI}}$ et w_{bb}^{TITI} . Les rapports $\mathbf{T}/\tilde{\mathbf{T}}$, obtenus directement des échantillons simulés présentés dans la section 13, sont fournis dans le tableau 8 avec leurs erreurs statistiques pour les différentes définitions de LOOSE'.

Coefficient	Loose'2	Loose'3	Loose'4	Loose'5
$n_{\gamma\gamma,MC}^{\tilde{\mathbf{T}}\mathbf{T}}/n_{\gamma\gamma,MC}^{\mathbf{T}\mathbf{T}}$	$1.4 \cdot 10^{-2} \pm 1.4 \cdot 10^{-4}$	$1.8 \cdot 10^{-2} \pm 1.5 \cdot 10^{-4}$	$2.6 \cdot 10^{-2} \pm 1.8 \cdot 10^{-4}$	$3.4 \cdot 10^{-2} \pm 2.0 \cdot 10^{-4}$
$n_{\gamma\gamma,MC}^{\mathbf{T}\tilde{\mathbf{T}}}/n_{\gamma\gamma,MC}^{\mathbf{T}\mathbf{T}}$	$2.8 \cdot 10^{-2} \pm 1.9 \cdot 10^{-4}$	$3.1 \cdot 10^{-2} \pm 1.9 \cdot 10^{-4}$	$4.1 \cdot 10^{-2} \pm 2.1 \cdot 10^{-4}$	$5.2 \cdot 10^{-2} \pm 2.6 \cdot 10^{-4}$
$n_{j\gamma,MC}^{\tilde{\mathbf{T}}\mathbf{T}}/n_{j\gamma,MC}^{\mathbf{T}\mathbf{T}}$	$1.9 \cdot 10^{-1} \pm 1.5 \cdot 10^{-1}$	$1.9 \cdot 10^{-1} \pm 1.5 \cdot 10^{-1}$	$1.9 \cdot 10^{-1} \pm 1.5 \cdot 10^{-1}$	$1.9 \cdot 10^{-1} \pm 1.5 \cdot 10^{-1}$
$n_{\gamma j,MC}^{\tilde{\mathbf{T}}\mathbf{T}}/n_{\gamma j,MC}^{\mathbf{T}\mathbf{T}}$	$4.6 \cdot 10^{-3} \pm 3.2 \cdot 10^{-3}$	$8.0 \cdot 10^{-3} \pm 4.5 \cdot 10^{-3}$	$2.1 \cdot 10^{-2} \pm 7.5 \cdot 10^{-3}$	$3.6 \cdot 10^{-2} \pm 1.1 \cdot 10^{-2}$
$n_{\gamma\gamma,MC}^{\tilde{\mathbf{T}}\tilde{\mathbf{T}}}/n_{\gamma\gamma,MC}^{\mathbf{T}\mathbf{T}}$	$4.7 \cdot 10^{-4} \pm 2.2 \cdot 10^{-5}$	$6.4 \cdot 10^{-4} \pm 2.5 \cdot 10^{-5}$	$1.2 \cdot 10^{-3} \pm 3.6 \cdot 10^{-5}$	$1.9 \cdot 10^{-3} \pm 4.5 \cdot 10^{-5}$

TAB. 8 – Rapports entre les nombres d'événements $\gamma\gamma$, γ +jet et jet+ γ obtenus dans les régions de contrôle du bruit de fond hadronique ($\tilde{\mathbf{T}}$) et la région de signal ($\mathbf{T}$), extraits à partir des échantillons simulés $\gamma\gamma$ et photon-jet générés avec SHERPA. Ces rapports sont utilisés dans le calcul des coefficients α_i , α'_i , β_i and β'_i (voir équations 54 à 63). Les erreurs statistiques MC associées sont également mentionnées.

D'après le MC, le rapport $\mathbf{T}\tilde{\mathbf{T}}/\mathbf{T}\mathbf{T}$ pour $\gamma\gamma$ est environ deux fois plus élevé que $\tilde{\mathbf{T}}\mathbf{T}/\mathbf{T}\mathbf{T}$, très vraisemblablement à cause de la différence en E_T^γ entre les deux candidats photons et de la baisse de l'efficacité d'identification TIGHT à bas E_T^γ (voir section 9.3). Le rapport $\tilde{\mathbf{T}}\tilde{\mathbf{T}}/\mathbf{T}\mathbf{T}$ est bien de l'ordre de grandeur du produit des deux précédents, illustrant la non corrélation des identifications pour $\gamma\gamma$ (voir section 15.2). De la même façon, le rapport $\tilde{\mathbf{T}}\tilde{\mathbf{T}}/\mathbf{T}\mathbf{T}$ pour γ +jet est plus faible que celui attendu pour jet+ γ .

Les corrections associées au signal ($\gamma\gamma$) sont bien déterminées, tandis que celles associées au bruit de fond souffrent de la statistique limitée des échantillons photon-jet. Cette caractéristique a pour effet d'augmenter les incertitudes sur les compositions séparées en γ +jet, jet+ γ et di-jet, mais n'a aucun effet sur l'incertitude globale sur le bruit de fond jet global (γ jet + jet γ + di-jet) ni sur le signal ou les électrons ($\gamma\gamma$, ee) (voir équations 64 à 68). Les différentes incertitudes seront quantifiées plus précisément

dans la section 15.6. Il est rappelé que les coefficients α_i , α'_i complets, représentant les contaminations finales des régions de contrôle, ne peuvent être déterminés qu'une fois la procédure d'ajustement effectuée (voir équations 54 à 58).

Dans la section 15.3.5, la procédure du *template fit* 2D est résumée succinctement. Une parenthèse est également faite sur le traitement des différents spectres ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc), avec en particulier des détails sur les dépendances cinématiques des pdfs conservées ou non dans les différents cas.

15.3.5 Résumé de la procédure et décomposition du bruit de fond pour les différents spectres

La méthode du *template fit* 2D a été détaillée au cours de la section 15.3. Elle peut être décomposée en trois étapes principales :

- ◇ **Extraction des différentes pdfs pour le signal et les bruits de fond :** L'échantillon simulé de signal dans **TITI** et les régions de contrôle des jets **$\tilde{\text{TITI}}$** , **$\text{TIT}\tilde{\text{I}}$** et **$\tilde{\text{TIT}}\tilde{\text{I}}$** dans les données sont utilisées pour l'extraction des pdfs d'isolation calorimétrique correspondant respectivement à $\gamma\gamma$, γ +jet, jet+ γ et di-jet. La pdf pour ee est extraite des régions de contrôle *tag-and-probe* dans les données. Les corrélations d'isolation entre les deux candidats sont préservées pour $\gamma\gamma$ et di-jet uniquement.
- ◇ **Procédure d'ajustement :** l'estimateur de maximum de vraisemblance étendu associé au modèle est construit et la procédure d'ajustement est exécutée sur la région **TITI** des données à l'aide du logiciel ROOFIT [205], lui-même basé sur le code MINUIT du CERN [206]. L'estimateur peut être catégorisé (*binned*) ou non (*unbinned*) en fonction de la statistique disponible. Dans tous les cas, des estimateurs de noyaux gaussiens sont utilisés afin de lisser les différentes pdfs, ce qui permet de limiter au maximum les biais statistiques dans les régions faiblement peuplées (voir section 15.3.2). Une contrainte gaussienne supplémentaire fonction de w_{ee}^{TITI} et centrée sur la prédiction du MC DY est appliquée afin d'aider à la discrimination photons/électrons.
- ◇ **Corrections des w_{ik}^{TITI} :** Les rapports $\tilde{\text{T}}/\text{T}$ sont extraits des échantillons simulés (voir tableau 8) et utilisés pour corriger les nombres d'événements ajustés w_{ik}^{TITI} des différentes contaminations selon les équations 64 à 68. Les nombres d'événements finals W_{ik}^{TITI} sont alors obtenus pour les différents processus. La matrice de covariance de la procédure d'ajustement renvoyée par ROOFIT (procédure HESSE [206]) est employée afin de propager correctement les différentes incertitudes statistiques dans les W_{ik}^{TITI} .

Par ailleurs, afin de reconstruire les distributions correspondant aux différentes observables de mesure ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T , ϕ_η^* , $\Delta\phi_{\gamma\gamma}$, $|\cos \theta_\eta^*|$, voir section 12.3), l'échantillon **TITI** est dans un second temps catégorisé selon chacune de ces variables (*binning*) et la procédure d'ajustement est répétée dans chaque catégorie. Les catégories adoptées sont unidimensionnelles et données ci-dessous :

- ◇ $m_{\gamma\gamma}$ [GeV] : 0-30, 30-50, 50-70, 70-80, 80-90, 90-100, 100-110, 110-120, 120-130, 130-140, 140-150, 150-160, 160-170, 170-180, 180-190, 190-200, 200-225, 225-250, 250-275, 275-300, 300-350, 350-400, 400-500, 500-600, 600-700, 700-1700.
- ◇ $p_{T,\gamma\gamma}$ [GeV] : 0-4, 4-8, 8-12, 12-16, 16-20, 20-25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-55, 55-60, 60-65, 65-70, 70-75, 75-80, 80-90, 90-100, 100-110, 110-120, 120-130, 130-140, 140-150, 150-175, 175-200, 200-225, 225-250, 250-300, 300-400, 400-750.
- ◇ ϕ_η^* [sans unité] : 0-0.004, 0.004-0.008, 0.008-0.012, 0.012-0.016, 0.016-0.020, 0.020-0.024, 0.024-0.029, 0.029-0.034, 0.034-0.039, 0.039-0.045, 0.045-0.051, 0.051-0.057, 0.057-0.064, 0.064-0.072, 0.072-0.081, 0.081-0.091, 0.091-0.102, 0.102-0.114, 0.114-0.128, 0.128-0.145, 0.145-0.165, 0.165-0.189, 0.189-0.219, 0.219-0.258, 0.258-0.312, 0.312-0.391, 0.391-0.524, 0.524-0.695, 0.695-0.918, 0.918-1.153, 1.153-1.496, 1.496-1.947, 1.947-2.522, 2.522-3.277, 3.277-5.0, 5-10, 10-20, 20-50, 50-100, 100-50000.
- ◇ a_T [GeV] : 0-2, 2-4, 4-6, 6-8, 8-10, 10-12, 12-14, 14-16, 16-18, 18-20, 20-25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-55, 55-60, 60-65, 65-70, 70-75, 75-80, 80-90, 90-100, 100-110, 110-120, 120-130, 130-150, 150-200, 200-450.
- ◇ $|\cos \theta_\eta^*|$ [sans unité] : 0.00-0.04, 0.04-0.08, 0.08-0.12, 0.12-0.16, 0.16-0.20, 0.20-0.24, 0.24-0.28, 0.28-0.32, 0.32-0.36, 0.36-0.40, 0.40-0.44, 0.44-0.48, 0.48-0.52, 0.52-0.56, 0.56-0.60, 0.60-0.64, 0.64-0.68, 0.68-0.72, 0.72-0.76, 0.76-0.80, 0.80-0.84, 0.84-0.88, 0.88-0.92, 0.92-0.96, 0.96-1.00.
- ◇ $\Delta\phi_{\gamma\gamma}$ [rad.] : 0.00-0.25, 0.25-0.50, 0.50-0.75, 0.75-1.00, 1.00-1.25, 1.25-1.50, 1.50-1.62, 1.62-1.75, 1.75-1.88, 1.88-2.00, 2.00-2.12, 2.12-2.25, 2.25-2.30, 2.30-2.35, 2.35-2.40, 2.40-2.45, 2.45-2.50, 2.50-2.55, 2.55-2.60, 2.60-2.65, 2.65-2.67, 2.67-2.70, 2.70-2.73, 2.73-2.75, 2.75-2.77, 2.77-2.80, 2.80-2.83, 2.83-2.85, 2.85-2.88, 2.88-2.90, 2.90-2.92, 2.92-2.95, 2.95-2.98, 2.98-3.00, 3.00-3.02, 3.02-3.05, 3.05-3.08, 3.08-3.10, 3.10-3.12, 3.12-3.14.

En tout, 193 mesures sont donc effectuées au total, 1 correspondant à la région **TITI** totale et 192 autres dans les sous-catégories mentionnées ci-dessus. 193 procédures d'ajustement sont donc nécessaires à l'analyse complète.⁵⁹ Le choix des catégories est fait sur la base de critères statistiques et des capacités du détecteur en matière de résolution, il sera discuté plus en détail dans la section 16.

Étant donné la dépendance en pseudo-rapacité $|\eta^\gamma|$ et en impulsion transverse E_T^γ des distributions d'isolation, que ce soit pour les photons ou les jets (voir section 14.3.2), les pdfs associées ($\gamma\gamma$, γ +jet, jet+ γ , di-jet) sont extraites indépendamment selon l'équation 50 dans chaque sous-catégorie. Cependant dans les queues de distribution, la statistique est souvent limitée et certaines régions de contrôle dans les données et régions de signal dans l'échantillon MC de signal sont fusionnées. L'effet de ces fusions est très limité et sera précisé dans la section 15.6.

⁵⁹La technique de bootstrap utilisée pour l'estimation des incertitudes (voir section 15.6) implique en réalité la multiplication de ce chiffre par un facteur de quelques milliers, amenant le nombre total de procédures d'ajustement lancées pour l'analyse à environ 500 000.

Le bruit de fond électron ee est concentré dans l'espace des phases se situant à proximité de la production résonante des désintégrations du $Z \rightarrow ee$. Par conséquent, les dépendances cinématiques des distributions d'isolation sont limitées et les mêmes pdfs sont utilisées quelque soit la catégorie, extraites à partir des régions de contrôle *tag-and-probe* (voir section 15.3.3). L'effet de cette simplification est totalement négligeable sur les résultats de soustraction d'après les tests réalisées avec les pseudo-données (voir section 15.4).

Le calcul des coefficients de contamination permettant la détermination de $W_{\gamma\gamma}^{\text{TITI}}$, α_1 , α_2 et α'_3 (voir section 15.3.4, équation 64) est effectué indépendamment dans chaque sous-catégorie, à l'exception des dernières catégories dans les queues de distribution de $p_{T,\gamma\gamma}$ (5 dernières catégories), a_T (4 dernières catégories) et $m_{\gamma\gamma}$ (3 dernières catégories), où la statistique réduite des régions de contrôle (parfois inférieure à 10 événements) mène à des résultats non-physiques. Dans ce cas, les coefficients α_1 , α_2 et α'_3 sont extrapolés à partir du reste de la distribution et une incertitude systématique conservatrice est adoptée, sous-dominante par rapport aux autres sources d'erreur (voir section 15.6). Plus de détails sur cette extrapolation seront données dans la section 15.5.

Cette méthode ne peut s'appliquer à $W_{\gamma j}^{\text{TITI}}$, $W_{j\gamma}^{\text{TITI}}$ et W_{jj}^{TITI} car les échantillons simulés photon-jet n'ont pas la statistique requise afin de pouvoir calculer les coefficients α'_1 et α'_2 de manière satisfaisante. Produire la statistique requise nécessiterait une consommation intensive de CPU et n'est pas réaliste. Ces facteurs n'impactant en aucun cas les résultats pour $W_{\gamma\gamma}^{\text{TITI}}$ (voir équation 64), il a été décidé de ne pas corriger les nombres d'événements γ +jet, jet+ γ et di-jet des contaminations des différents bruit de fond hadroniques. Ainsi, dans le cas d'une décomposition du bruit de fond dans les différents spectres, la méthode du *template fit* 2D ne permet pas de citer des chiffres indépendants pour γ +jet, jet+ γ et di-jet et se contente de citer un chiffre global correspondant à la somme de ces trois processus.

Dans la section 15.4, le jeu de pseudo-données créé spécialement pour tester et valider la méthode du *template fit* 2D dans le cadre de l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV est présenté. Sa mise au point, comme la quasi-intégralité des résultats présentés dans la partie IV, représente l'un des travaux personnels effectué dans les années de thèse conjointement avec le Dr. Bruno Lenzi (CERN). C'est l'une des premières fois où une analyse $\gamma\gamma + X$ utilise un jeu de pseudo-données pour valider sa méthode de soustraction du bruit de fond dans ATLAS, ce qui a permis de réaliser de nombreuses études complémentaires et de déceler un certain nombre de biais dans le cadre du *template fit* 2D (soustraction séparée des jets et des électrons, corrélations d'isolation pour $\gamma\gamma$), mais aussi pour les méthodes de soustraction alternatives (voir section 15.7). Depuis, le jeu de pseudo-données a été ré-utilisé dans le cadre de l'analyse $\gamma + c(b) + X$ à $\sqrt{s} = 8$ TeV en cours dans ATLAS et $\gamma\gamma + X$ à 13 TeV effectuée dans le cadre de la recherche de résonances et présentée à la conférence de Moriond 2016 [245, 246].

15.4 Test et validation de la méthode sur un jeu de pseudo-données

15.4.1 Création d'un jeu de pseudo-données

Étant donné la simplicité de l'état final, il a été décidé de valider la méthode du *template fit* 2D sur un jeu de pseudo-données créé spécialement pour l'analyse $\gamma\gamma + X$. Les pseudo-données reproduisent un échantillon de di-photons LOOSE'4-LOOSE'4 à $\sqrt{s} = 8$ TeV avec toutes les coupures de sélection appliquées sauf l'identification TIGHT et les coupures d'isolation (détecteur interne + calorimètre). Elles incluent la quadri-impulsion des deux candidats sélectionnés, leur statut d'identification (TIGHT ou NON-TIGHT), leur énergie d'isolation calorimétrique et leur énergie d'isolation dans le détecteur interne.

Elles ont été constituées selon les modalités suivantes :

- ◇ **$\gamma\gamma$ et e^+e^-** : les événements $\gamma\gamma$ et ee sont directement issus des échantillons simulés de signal SHERPA et des échantillons simulés de bruit de fond DY POWHEG+PYTHIA8, respectivement (voir section 13.2). Cela permet de conserver toutes les dépendances entre isolation, identification et cinématique pour chaque objet et entre les deux candidats.
- ◇ **γ +jet et jet+ γ** : il est impossible d'utiliser les simulations photon-jet pour cette composante à cause de la statistique très limitée des échantillons correspondants. Les événements γ +jet et jet+ γ sont donc extraits respectivement des échantillons de contrôle $\mathbf{T}\tilde{\mathbf{T}}$ et $\tilde{\mathbf{T}}\mathbf{T}$ dans les données.
 - L'identification et les énergies d'isolation de la branche $\mathbf{T}$ (qui correspond le plus souvent à un photon immédiat) sont remplacées par une valeur prise de façon aléatoire dans l'échantillon de signal SHERPA parmi les candidats ayant un E_T^γ et $|\eta^\gamma|$ similaire.⁶⁰
 - Le statut d'identification de la branche $\tilde{\mathbf{T}}$ (qui correspond le plus souvent à un jet) est remplacé par une valeur aléatoire tirée du ratio $\tilde{\mathbf{T}}/\mathbf{T}$ extrait des données à partir de la méthode décrite dans l'appendice B dans des catégories en $|\eta^\gamma| - E_T^\gamma - m_{\gamma\gamma}$ afin de tenir compte des dépendances cinématiques des taux de mauvaise identification (voir section 15.1, figure 98). Les variables d'isolation sont en revanche conservées.

Cette technique permet de préserver la cinématique du système et les corrélations associées, mais néglige les corrélations supplémentaires d'isolation entre les deux candidats (événements sous-jacent, etc). Par ailleurs, la contamination par les vrais photons de la branche $\tilde{\mathbf{T}}$ n'est pas soustraite, ce qui devrait amener les jets des pseudo-données à être légèrement mieux isolés par rapport aux vraies données

⁶⁰En pratique, l'échantillon de signal SHERPA est catégorisé de la manière suivante :

- premier photon $E_{T,1}^\gamma$ [GeV] (40, 45, 50, 55, 60, 70, 85, 100, 150, 200, 400, 600, 1000).
- deuxième photon $E_{T,2}^\gamma$ [GeV] (30, 35, 40, 45, 50, 60, 70, 90, 110, 150, 250, 400, 600, 1000).
- pour $|\eta^\gamma|$, 50 (30) catégories de même largeur sont définies dans $[0, 2.37]$ pour le premier (deuxième) photon.

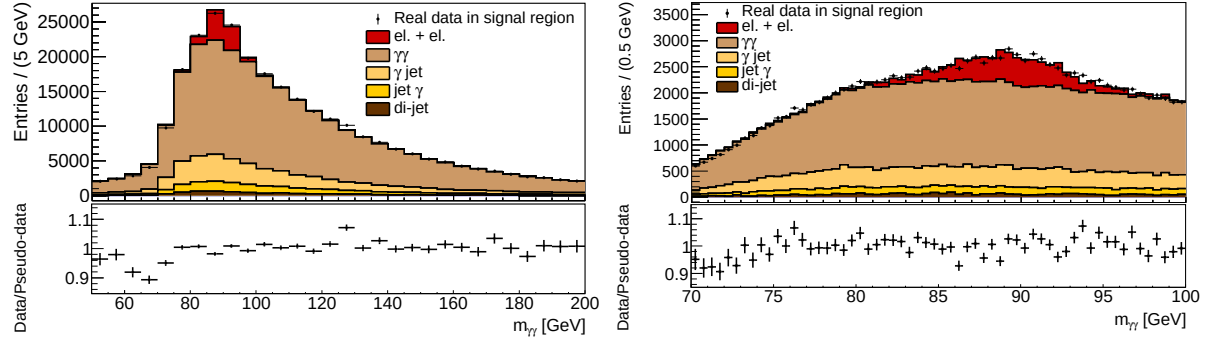


FIG. 103 – Masse invariante des candidats di-photon dans la région de signal **TITI** dans les données (points noirs) et histogrammes empilés des composantes des pseudo-données ($\gamma\gamma$, γ +jet, jet+ γ , di-jet, ee) dans deux intervalles : 50–200 GeV (gauche) et autour du pic du boson Z^0 (70–100 GeV, droite). Les barres d'erreur tiennent compte uniquement des incertitudes statistiques des vraies données.

(légère baisse du pouvoir discriminant du *template fit 2D*).

- ◇ **di-jet** : les événements sont extraits de l'échantillon de contrôle $\tilde{\mathbf{T}}\tilde{\mathbf{T}}$ dans les données en remplaçant le statut d'identification de chaque objet par une valeur aléatoire tirée du ratio $\tilde{\mathbf{T}}/\mathbf{T}$ extrait des données à partir de la même méthode que celle décrite pour γ +jet et jet+ γ (voir appendice B). Ces ratios diffèrent légèrement des précédents et sont mieux adaptés aux événements di-jet, plus d'informations sont disponibles dans l'appendice B.

Cette technique permet de préserver les corrélations cinématiques (comme pour γ +jet et jet+ γ) et d'isolation (en plus comparé à γ +jet et jet+ γ) entre les deux candidats, tandis que la corrélation entre isolation et identification pour un même objet est perdue (comme pour γ +jet et jet+ γ). De façon similaire aux processus photon-jet, les fuites de photons immédiats dans la région $\tilde{\mathbf{T}}\tilde{\mathbf{T}}$ ne sont pas soustraites, ce qui devrait amener les di-jets des pseudo-données à être légèrement mieux isolés comparés aux vraies données (légère baisse du pouvoir discriminant du *template fit 2D*).

Les événements de chaque processus sont repondérés globalement afin de correspondre approximativement aux proportions observées dans la région de signal **TITI** des vraies données, à l'exception de la composante ee qui est simplement renormalisée à la luminosité des données. De plus, la composante $\gamma\gamma$, extraite à partir de l'échantillon de signal simulé SHERPA, est repondérée en masse ($m_{\gamma\gamma}$) afin de mieux correspondre à la distribution extraite dans les vraies données (voir section 17).⁶¹ Les données et les pseudo-données complètes sont comparées en termes de masse invariante $m_{\gamma\gamma}$ après repondération dans la figure 103. Un bon accord général est observé.

Dans les sections 15.4.2 et 15.4.3, la méthode de décomposition du bruit de fond

⁶¹La repondération est basée sur la comparaison entre données et pseudo-données avant repondération et correspond à l'application des fonctions de poids suivantes :

- pour $30 \text{ GeV} < m_{\gamma\gamma} < 60 \text{ GeV}$, $w = 1.01887 - 0.00361531 \cdot m_{\gamma\gamma}$.
- pour $60 \text{ GeV} < m_{\gamma\gamma} < 70 \text{ GeV}$, $w = 0.28558 + 0.00866237 \cdot m_{\gamma\gamma}$.
- pour $m_{\gamma\gamma} > 70 \text{ GeV}$, $w = 1.47783e+00 - \exp(-3.41810e-03 \cdot (m_{\gamma\gamma} + 1.03794e+02))$.

nominale de l'analyse (*template fit* 2D, voir section 15.3) est exécutée sur le jeu de pseudo-données, tout d'abord en excluant explicitement les électrons afin de restreindre l'espace des paramètres (section 15.4.2) et finalement avec l'intégralité des composantes incluses (section 15.4.3). Ces tests se sont montrés très utiles pour évaluer l'effet des différentes hypothèses simplificatrices effectuées pour la soustraction du fond et les modifier si besoin afin de réduire les incertitudes. En particulier, l'implémentation finale du *template fit* 2D dans l'analyse $\gamma\gamma + X$ inclut désormais la prise en compte des corrélations d'isolation entre les candidats pour $\gamma\gamma$ et l'inclusion de la composante ee dans la procédure d'ajustement, aspects qui avaient été négligés lors des analyses précédentes d'ATLAS [202, 203].

15.4.2 Validation de la méthode du *template fit* 2D en l'absence de bruit de fond électron

La procédure du *template fit* 2D décrite dans la section 15.3 est tout d'abord exécutée sur le jeu de pseudo-données en excluant explicitement la composante ee , à la fois dans les pseudo-données et le modèle d'ajustement. Afin d'estimer l'effet des corrélations d'isolation dans les événements $\gamma\gamma$, la méthode est exécutée à nouveau avec une pdf $F_{\gamma\gamma}^{\text{TITI}}$ extraite à partir du produit des deux pdfs 1D correspondantes, *ie.* $F_{\gamma,1}^{\text{TITI}} \times F_{\gamma,2}^{\text{TITI}}$ (voir section 15.3.3). Les résultats sont rapportés dans le tableau 9.

Processus	Chiffres réels	<i>fit</i> 2D nominal	<i>fit</i> 2D sans corrélations d'isolation $\gamma\gamma$
$\gamma\gamma$	233 003	231 383 $\pm$ 953	236 761 $\pm$ 964
$\gamma + jet$	45 000	45 640 $\pm$ 748	40 007 $\pm$ 773
$jet + \gamma$	18 000	20 657 $\pm$ 634	16 415 $\pm$ 716
di-jet	7 000	5 247 $\pm$ 399	9 741 $\pm$ 505

TAB. 9 – Résultats de la méthode de soustraction du *template fit* 2D sur le jeu de pseudo-données avec les incertitudes statistiques associées. La composante ee est exclue explicitement à la fois des pseudo-données et du modèle d'ajustement. Deux cas de figures sont testés : pdf $\gamma\gamma$ 2D extraite à partir de l'échantillon simulé SHERPA ($F_{\gamma\gamma}^{\text{TITI}}$) et produit des deux pdfs 1D correspondantes ($F_{\gamma,1}^{\text{TITI}} \times F_{\gamma,2}^{\text{TITI}}$).

Un biais important sur le nombre d'événements $\gamma\gamma$ attendu est observé lorsque les corrélations d'isolation $\gamma\gamma$ sont négligées, correspondant à un effet de l'ordre de +1.6% en termes de nombre d'événements et 1.2% en termes de proportions d'événements de signal dans l'échantillon (ou pureté).⁶² Cela est non négligeable par rapport à l'incertitude systématique totale, qui est de l'ordre de 2.6% sur la pureté (voir section 15.6).

La prise en compte des corrélations d'isolation $\gamma\gamma$ engendre une baisse de la pureté d'environ 1.7%, une variation du même ordre de grandeur étant observée dans les véritables données de collision. Le biais résiduel final sur la pureté $\gamma\gamma$ est alors d'environ -0.5% dans les pseudo-données. L'origine du biais résiduel restant a été identifiée grâce aux pseudo-données et est attribuée aux deux raisons suivantes :

⁶²Présenter les variations des résultats $\gamma\gamma$ en termes de pureté a l'avantage de s'affranchir d'une part du nombre total d'événements dans l'échantillon et d'autre part de la différence entre nombre total d'événements attendus et observés. Les variations de pureté permettent ainsi de comparer plus directement données et pseudo-données et seront privilégiées dans la suite.

- ◇ La distribution d'isolation pour $\gamma\gamma$ dans les pseudo-données est extraite à partir de l'échantillon de signal simulé SHERPA repondéré en $m_{\gamma\gamma}$ tandis que la méthode du *template fit* 2D utilise la pdf du même échantillon sans repondération. La repondération implique des déformations de cinématique et donc des pdfs d'isolation, menant à un léger biais dans la méthode.
- ◇ Le jeu de pseudo-données a été conçu de façon à ce que la probabilité pour un jet de passer la sélection TIGHT dépende de sa cinématique, plus précisément de $|\eta^\gamma| - E_T^\gamma - m_{\gamma\gamma}$ (voir section 15.4.1). À nouveau, cela va induire des différences de cinématique entre les jets présents dans la région de signal **TITI** et ceux présents dans les différentes régions de contrôle (**$\tilde{\mathbf{T}}$**) et ainsi biaiser légèrement la méthode.

Le biais résultant des deux effets décrits ci-dessus est évalué de la façon suivante dans les pseudo-données : la pureté dans l'échantillon **TITI** complet est recalculée à partir des chiffres obtenus lors de la décomposition du bruit de fond en fonction des différentes observables de mesure ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T , etc) et comparée au chiffre obtenu en considérant uniquement la région **TITI** dans son intégralité. Le modèle d'ajustement étant extrait indépendamment dans chaque catégorie pour l'observable considéré ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T , etc) et la largeur des catégories étant en général très fine grâce à la très bonne statistique des données d'ATLAS à $\sqrt{s} = 8$ TeV et l'excellente résolution du calorimètre, on s'attend à ce que l'effet des différences cinématiques entre le modèle d'ajustement (échantillon simulé SHERPA non-repondéré, régions de contrôle basées sur **$\tilde{\mathbf{T}}$**) et les données (région **TITI**) soit bien plus faible.

Comme attendu, la pureté dans l'échantillon **TITI** total recalculée à partir de la décomposition du bruit de fond en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T , etc, est systématiquement supérieure au chiffre nominal, avec un maximum atteint pour $m_{\gamma\gamma}$ correspondant au biais observé (232 918 événements $\gamma\gamma$, soit +0.5% sur la pureté par rapport au cas nominal). Étant donné que des effets similaires sont attendus dans les vraies données, ces quantités seront également évaluées pour celles-ci et formeront une incertitude systématique distincte ajoutée en quadrature (voir section 15.6). L'effet résultant dans les données est du même ordre de grandeur que dans les pseudo-données ($\simeq 2\,000$ événements $\gamma\gamma$ de différence), illustrant à nouveau la bonne qualité des pseudo-données.⁶³

Les biais résiduels sur les composantes de bruit de fond sont plus complexes à évaluer car le *template fit* 2D ne permet pas la distinction des trois composantes indépendamment lors de la décomposition du bruit de fond dans les différentes observables (voir section 15.3.5). Par ailleurs, les corrections de fuites entre les trois composantes (γ +jet, jet+ γ , di-jet) pour la région **TITI** complète sont mal déterminées à cause de la statistique limitée des échantillons simulés photon-jet (voir section 8). Ces biais résiduels n'affectant dans tous les cas pas le nombre d'événements $\gamma\gamma$, un simple test de cohérence est effectué en répétant le *template fit* 2D avec les pdfs de jets véritables des pseudo-données au lieu de celles extraites depuis les régions de contrôle afin de s'affranchir de l'intégralité des systématiques liées à la modélisation de l'isolation des jets. L'effet des corrélations d'isolation $\gamma\gamma$ peut ainsi être testé uniquement en présence des systématiques résiduelles liées à la modélisation des photons. Les résultats sont présentés dans le tableau 10.

⁶³Bien qu'il serait tentant d'utiliser l'échantillon de signal SHERPA repondéré dans la méthode nominale de soustraction du fond sur les vraies données, cela s'avère délicat voir dangereux car une repondération en $m_{\gamma\gamma}$ dégrade la description de SHERPA pour d'autres observables telles que $p_{T,\gamma\gamma}$.

Processus	Chiffres réels	<i>fit</i> 2D avec pdfs réelles de bruit de fond + pdf $\gamma\gamma$ 2D	<i>fit</i> 2D avec pdfs réelles de bruit de fond + pdf $\gamma\gamma$ 1D×1D
$\gamma\gamma$	233 003	$231\,925 \pm 913$	$236\,971 \pm 929$
$\gamma + jet$	45 000	$45\,348 \pm 702$	$40\,454 \pm 716$
$jet + \gamma$	18 000	$18\,383 \pm 623$	$13\,764 \pm 634$
di-jet	7 000	$7\,270 \pm 474$	$11\,738 \pm 494$

TAB. 10 – Résultats similaires à ceux présentés dans le tableau 9 dans le cas où le *template fit* 2D utilise les véritables pdfs d’isolation des jets présents dans les pseudo-données au lieu des distributions observées dans les régions de contrôle. L’effet des corrélations d’isolation $\gamma\gamma$ est à nouveau testé et trouvé comme étant significatif.

Un biais similaire correspondant à une surestimation de la pureté par +1.3% est observé si les corrélations d’isolation sont négligées pour $\gamma\gamma$, qui est réduit à -0.35% si elles sont prises en compte. À nouveau, le biais résiduel de -0.35% est couvert par l’incertitude systématique décrite plus haut. Par ailleurs, les événements de bruit de fond attendus sont désormais compatibles avec leurs valeurs réelles, principalement grâce à l’annulation des incertitudes statistiques sur les contaminations de $jet+\gamma$ et $\gamma+jet$ dans $\tilde{\mathbf{T}}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$ (utilisation des pdfs de jets véritables).

En conclusion, une fois que l’on tient compte des corrélations d’isolation pour $\gamma\gamma$, de la systématique reliée aux différences de cinématique entre modèle et données et des incertitudes sur les contaminations de $jet+\gamma$ et $\gamma+jet$ dans $\tilde{\mathbf{T}}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$, toutes les composantes renvoyées par le *template fit* 2D sont en bon accord avec les chiffres réels. Il a donc été décidé d’inclure les corrélations d’isolation pour $\gamma\gamma$ dans la méthode du *template fit* 2D nominale, comme cela a déjà été annoncé précédemment dans les sections 15.2 et 15.3. Cela constitue une évolution importante par rapport aux analyses précédentes d’ATLAS [202, 203]. L’erreur systématique reliée aux différences de cinématique et les incertitudes dues aux contaminations de $jet+\gamma$ et $\gamma+jet$ dans $\tilde{\mathbf{T}}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$ avait déjà été estimées dans les analyses à $\sqrt{s} = 7$ TeV et est réévaluée à 8 TeV.

Finalement, la méthode du *template fit* 2D nominale est répétée avec le jeu de pseudo-données pour les différents spectres (voir section 15.3.5) et le nombre d’événements $\gamma\gamma$ attendus est comparé au chiffre réel pour les 192 catégories. Les résultats sont montrés dans la figure 104. Un très bon accord général est observé, le plus souvent compatible à moins d’une déviation statistique standard. Mêmes si celles-ci sont couvertes par les incertitudes statistiques, des différences importantes sont parfois observées dans les queues des distributions de $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, ϕ_{η}^* ou a_T . Cela s’explique par le fait que certaines des régions de contrôle dans les données et certaines régions de signal dans l’échantillon simulé SHERPA sont fusionnées dans ces régions (voir section 15.3.5). Dans l’analyse principale, une incertitude systématique sera évaluée afin de tenir compte du biais potentiel engendré par ces simplifications (voir section 15.6).

Comme déjà mentionné dans les sections 15.2 et 15.4.1, les pseudo-données ne permettent pas d’évaluer l’effet des corrélations d’isolation éventuelles dans les événements photon-jet. L’unique partie de ces corrélations qui est préservée lors de la construction du jeu de pseudo-données est celle induite par la cinématique et a, d’après les résultats de cette section, un impact tout à fait négligeable sur les résultats. Un autre effet potentiellement important et non reproduit par les pseudo-données est la corrélation résiduelle entre l’identification et l’isolation d’un même candidat pour les jets, inconnue

mais nécessairement négligée pour permettre l'extrapolation des pdfs d'isolation des jets depuis les régions de contrôle vers la région de signal. Cet effet est évalué dans l'analyse principale en variant la définition de la sélection LOOSE', ce qui a pour effet de déformer les pdfs des jets (voir section 15.6). Le fait d'utiliser les photons de l'échantillon SHERPA dans les pseudo-données limite également les possibilités pour tester la validité des pdfs de ce générateur pour les pdfs photons. Cela sera évaluée en détail de plusieurs manières différentes dans les vraies données, notamment en remplaçant les pdfs extraites à partir de SHERPA par celles de PYTHIA (voir section 15.6).

Dans la section 15.4.3, quelques tests complémentaires sont effectués avec les pseudo-données incluant la composante ee afin de tester son effet sur les résultats.

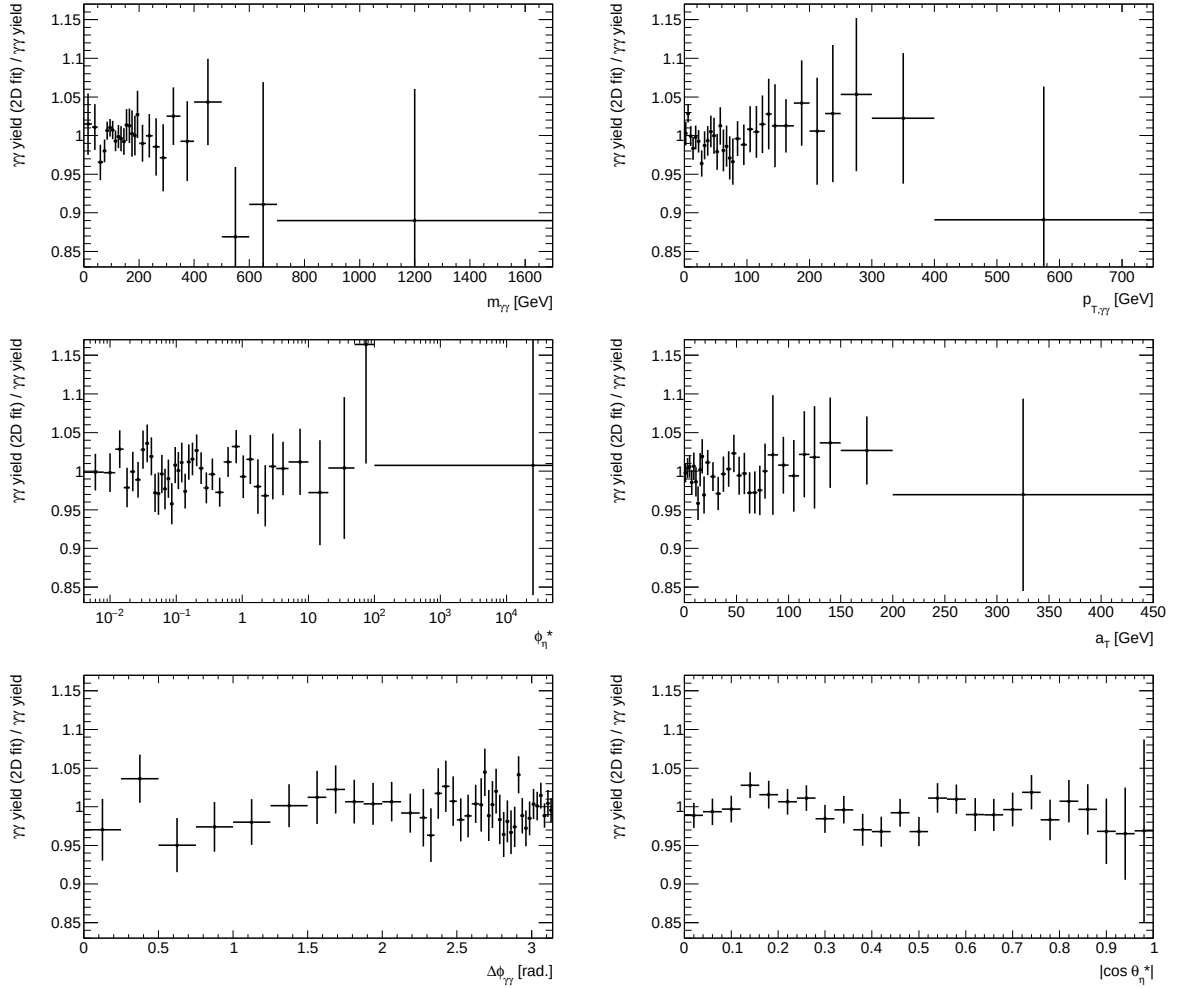


FIG. 104 – Rapport entre le nombre d'événements de signal attendus avec le *template fit* 2D et le nombre réel d'événements présents dans les pseudo-données. La composante ee est exclue à la fois du modèle d'ajustement et des pseudo-données. Les barres d'erreurs sont uniquement statistiques. Un très bon accord général est observé.

15.4.3 Validation de la méthode du *template fit* 2D en présence de bruit de fond électron

La procédure complète du *template fit* 2D (incluant la composante ee) est exécutée sur le jeu de pseudo-données complet (ee inclus) avec et sans tenir compte des corrélations d'isolation pour la composante di-électron. Les résultats sont affichés dans le tableau 11.

Processus	Chiffres réels	<i>fit</i> 2D nominal	<i>fit</i> 2D tenant compte des corrélations d'isolation pour ee
$\gamma\gamma$	233 003	$230\,923 \pm 1\,037$	$230\,518 \pm 1\,047$
$\gamma + jet$	45 000	$45\,526 \pm 774$	$45\,782 \pm 774$
$jet + \gamma$	18 000	$20\,696 \pm 667$	$20\,870 \pm 666$
di-jet	7 000	$5\,365 \pm 408$	$5\,163 \pm 409$
ee	10 582	$10\,990 \pm 639$	$11\,164 \pm 645$

TAB. 11 – Résultats de la méthode de soustraction du *template fit* 2D sur le jeu de pseudo-données avec les incertitudes statistiques associées. Toutes les composantes sont désormais incluses, à la fois dans les pseudo-données et le modèle d'ajustement. Deux cas de figures sont testés : pdf ee 2D extraite à partir de l'échantillon DY simulé (F_{ee}^{TITI}) et produit des deux pdfs 1D correspondantes ($F_{e,1}^{TITI} \times F_{e,2}^{TITI}$).

Dans les deux cas, le biais résiduel est similaire et en très bon accord avec les résultats du tableau 9, soit un effet d'environ -0.7% sur la pureté, représentant $\simeq 2\,000$ événements $\gamma\gamma$. L'écart au nombre réel d'événements pour $\gamma\gamma$ (et également ee) est à nouveau couvert par les différences de cinématique entre le modèle d'ajustement et la région de signal **TITI** des pseudo-données, tandis que celui pour $jet+\gamma$, $\gamma+jet$ et di-jet est couvert par les différences de cinématique et l'incertitude sur les contaminations des événements $jet+\gamma$ et $\gamma+jet$ dans **TITI**. Le *template fit* 2D nominal est ensuite répété pour les différents spectres et le nombre d'événements $\gamma\gamma$ attendu est comparé au nombre réel dans la figure 105. Les résultats obtenus sont très similaires à ceux de la figure 104, ce qui n'est pas surprenant vu le très faible nombre d'événements ee en général. Un très bon accord est observé pour la totalité des observables. La méthode du *template fit* 2D présentée dans la section 15.3 est donc validée pour l'ensemble de l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV.

Ainsi, les corrélations d'isolation pour le processus ee (voir section 15.2) ont un effet négligeable sur les résultats de soustraction, ce qui permet de pouvoir utiliser les pdfs électrons 1D issus des régions de contrôle *tag-and-probe* des données dans l'analyse principale. Par ailleurs, il est maintenant démontré que les contaminations des électrons dans les régions de contrôle des jets ont un impact négligeable sur les résultats (voir section 15.3.4).

Finalement, le *template fit* 2D sans la composante ee est exécutée sur le jeu de pseudo-données complet (ee inclus) et les résultats correspondants sont comparés aux chiffres réels dans le tableau 12. Si la composante ee n'est pas incluse dans le modèle d'ajustement, les événements di-électrons se répartissent entre les différentes composantes, principalement $\gamma\gamma$ mais également $\gamma+jet$ et $jet+\gamma$. Ainsi, considérer que les électrons sont absorbés en totalité dans la composante $\gamma\gamma$ amène un biais non-négligeable d'environ -2% sur la pureté " $\gamma\gamma + ee$ ". Cette hypothèse a été utilisée pour le *template fit*

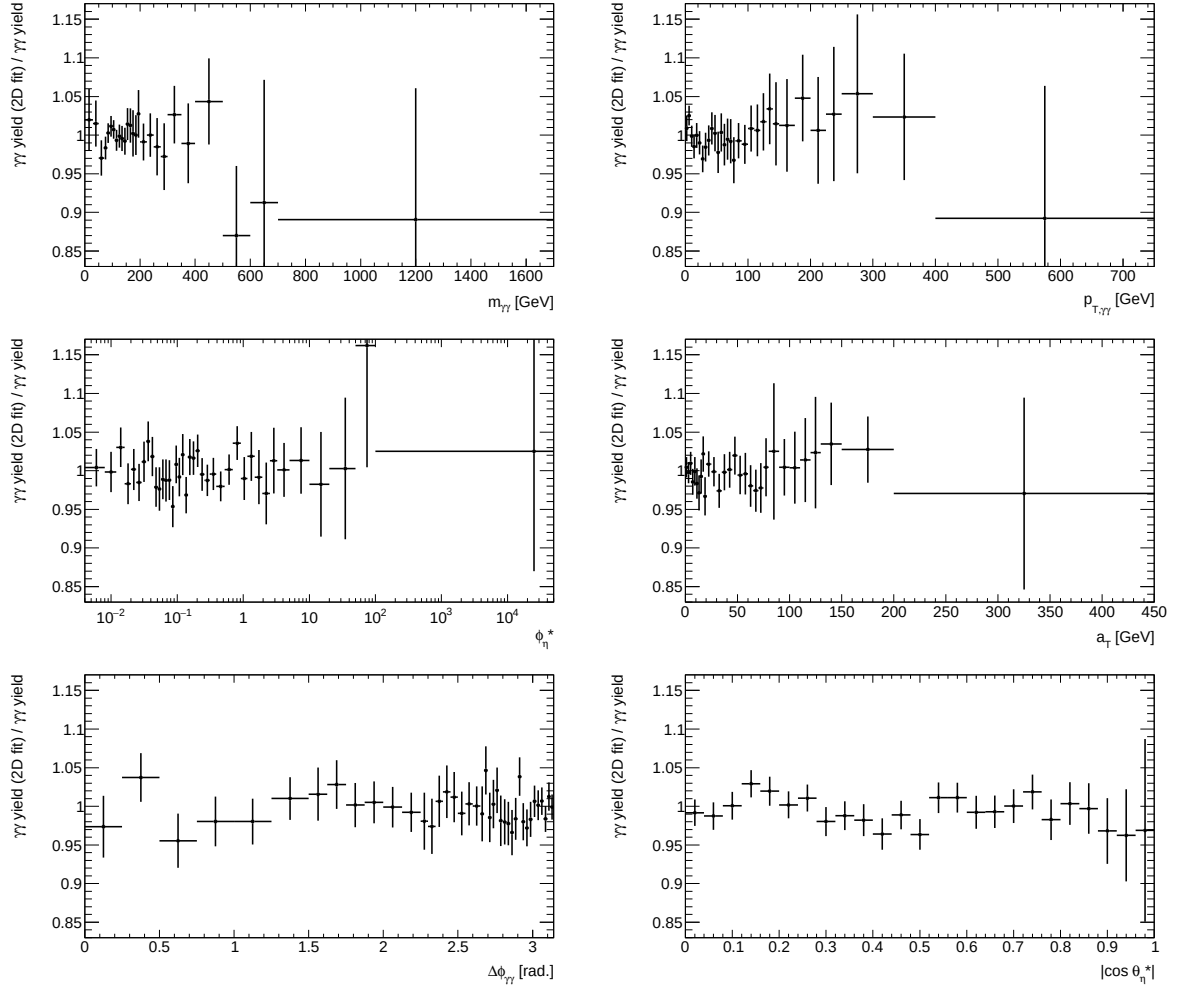


FIG. 105 – Rapport entre le nombre d'événements de signal attendus avec le *template fit* 2D et le nombre réel d'événements présents dans les pseudo-données. Toutes les composantes sont désormais incluses à la fois dans les pseudo-données et le modèle d'ajustement. Les barres d'erreurs sont uniquement statistiques. Un très bon accord général est observé.

Processus	Chiffres réels	<i>fit</i> 2D sans composant pour <i>ee</i>
$\gamma\gamma, ee$	243 582	$237\,228 \pm 971$
$\gamma + jet$	45 000	$47\,910 \pm 823$
$jet + \gamma$	18 000	$22\,979 \pm 896$
di-jet	7 000	$5\,391 \pm 277$

TAB. 12 – Résultats de la méthode de soustraction du *template fit* 2D sans composante *ee* sur le jeu de pseudo-données complet (*ee* inclus) avec les incertitudes statistiques associées. Il n'est pas correct de considérer ici que l'intégralité de la composante *ee* est absorbée dans $\gamma\gamma$.

mis en œuvre dans le cadre des deux analyses $\gamma\gamma + X$ précédentes d'ATLAS [202, 203] et également pour l'analyse de CMS [224].⁶⁴ Elle est rejetée dans l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV, où une composante ee est ajoutée au modèle d'ajustement (voir section 15.3).

La méthode du *template fit* 2D présentée dans la section 15.3 est donc validée pour l'ensemble de l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV, à la fois pour son application à la région **TITI** complète et pour la décomposition en fonction des différentes observables de l'analyse. Dans la section 15.5, les résultats de la méthode avec les données de l'échantillon **TITI** d'ATLAS sont dévoilés et analysés.

15.5 Décomposition finale du bruit de fond dans les données réelles

15.5.1 Échantillon **TITI** complet (cas inclusif)

Présentation des résultats de décomposition

Le *template fit* 2D est exécuté dans un premier temps avec 312 666 événements sur les 312 754 présents dans la région **TITI** des données (voir section 14.6), 88 événements étant exclus car au moins un de leur candidat présente une énergie d'isolation calorimétrique inférieure à -5.2 GeV, qui est la borne inférieure des intervalles utilisés pour l'ajustement $[-5.2 \text{ GeV}, 6 \text{ GeV}] \times [-5.2 \text{ GeV}, 6 \text{ GeV}]$, voir section 15.3). Cette perte est négligée dans la suite. L'ajustement consiste alors à la minimisation d'un maximum de vraisemblance étendu catégorisé dans des intervalles de largeur $0.4 \text{ GeV} \times 0.4 \text{ GeV}$ (cas inclusif, voir section 15.3.2).

Les résultats de décomposition avant et après correction des nombres d'événements ajustés ($w_{ik}^{\text{TITI}} \rightarrow W_{ik}^{\text{TITI}}$, voir section 15.3.4) sont présentés dans le tableau 13. La pureté $\gamma\gamma$ finale observée dans la région **TITI** complète est de $75.28 \pm 0.33\%$ tandis que la pureté $\gamma\gamma + ee$ est de 77.92% , significativement plus élevée que dans la mesure précédente d'ATLAS à $\sqrt{s} = 7$ TeV (67.54% pour $\gamma\gamma + ee$) [203]. Cela est tout d'abord dû aux changements dans la définition de l'espace des phases, l'analyse à $\sqrt{s} = 8$ TeV ($\sqrt{s} = 7$ TeV) demandant $E_{T,1} > 40$ GeV, $E_{T,2} > 30$ GeV ($E_{T,1} > 25$ GeV, $E_{T,2} > 22$ GeV), car le bruit de fond jet baisse rapidement avec E_T^γ (voir section 15.1). Plus d'événements γ +jet sont attendus dans l'échantillon par rapport à jet+ γ pour les mêmes raisons, comme cela a déjà été souligné dans la section 15.1. Par ailleurs, la coupure sur l'isolation des traces ($p_T^{\text{track.iso.}} < 2.6$ GeV) permet de garder une excellente pureté malgré l'assouplissement de la coupure d'isolation calorimétrique par rapport à l'analyse précédente ($E_T^{\text{calo.iso.}} < 6$ GeV vs $E_T^{\text{calo.iso.}} < 4$ GeV) [203]. Les motivations pour ces légers changements ont déjà été détaillées dans la section 14.3. Finalement, la composante ee est fortement réduite par rapport à l'analyse précédente ($2.65 \pm 0.22\%$ vs $\simeq 5.7\%$) [203], notamment grâce à la sélection plus restrictive en $E_{T,1}, E_{T,2}$, limitant l'acceptance pour $m_{\gamma\gamma} \simeq m_Z$, et la coupure sur l'isolation des traces, très discriminante pour les électrons (voir section 14.3.1).

⁶⁴Le détecteur de CMS étant totalement différent de celui d'ATLAS, il est évidemment impossible de conclure quant à la validité de cette hypothèse dans les analyses $\gamma\gamma + X$ de CMS.

Processus	Chiffres ajustés (w_{ik}^{TITI})	Chiffres finals (W_{ik}^{TITI})	% finals
$\gamma\gamma$	$226\,085 \pm 1\,074$	$235\,392 \pm 1\,022$	$75.28 \pm 0.33\%$
γ +jet	$51\,005 \pm 762$	$45\,465 \pm 672$	$14.54 \pm 0.21\%$
jet+ γ	$20\,103 \pm 693$	$18\,699 \pm 590$	$5.98 \pm 0.19\%$
di-jet	$7\,236 \pm 480$	$4\,873 \pm 353$	$1.56 \pm 0.11\%$
di-electron	$8\,275 \pm 680$	$8\,275 \pm 680$	$2.65 \pm 0.22\%$
Total	312 704	312 704	100%

TAB. 13 – Résultat de la méthode de soustraction du *template fit* 2D pour la région de signal **TITI** complète. Les incertitudes affichées sont celles renvoyées par le code MINUIT [206] et tiennent compte de la statistique limitée de la région de signal uniquement. La correction des nombres d'événements ajustés est effectuée à l'aide des équations 64 à 68 présentées dans la section 15.3.4, et les incertitudes statistiques sont propagées grâce à la matrice de covariance de l'ajustement. Les autres sources d'incertitude seront traitées en détail dans la section 15.6.

Le nombre total d'événements attendus par le *template fit 2D* (312 704) est par ailleurs en très bon accord avec le nombre d'événements observés (312 666 dans l'intervalle d'ajustement considéré). Il est rappelé que l'estimateur du maximum de vraisemblance étendu fait par définition usage d'une contrainte poissonnienne centrée sur le chiffre observé pour cette quantité (312 666). Cependant, le nombre d'électrons observé (8275 ± 680) dévie significativement de la contrainte gaussienne centrée sur la prédiction de l'échantillon DY (voir section 15.3.3), qui s'élève à $10\,580 \pm 720$ pour l'échantillon **TITI** complet. L'écart représente ainsi une déviation de $-3.2\,\sigma$ par rapport à la contrainte alors que celle-ci est considérée comme une borne inférieure du nombre d'événements ee réels dans la région **TITI**, à cause des processus autres que $Z \rightarrow ee$ pouvant jouer un rôle dans cette contamination ($t\bar{t}$, ZZ , WW , etc). Autrement dit, les données d'isolation plaident en faveur d'une contamination ee plus faible que celle prédite par le MC DY avec facteurs correctifs pour le modèle d'ajustement nominal. Bien que le bruit de fond électrons joue un rôle mineur dans l'analyse, nous reviendrons sur ce point lors de l'analyse de la qualité du modèle d'ajustement dans le paragraphe suivant et lors de l'évaluation des incertitudes systématiques dans la section 15.6.

Les différentes contaminations des événements $\gamma\gamma$ dans les différentes régions de contrôle de jets d'une part, et des événements γ +jet, jet+ γ dans **TITI** d'autre part obtenues avec le *template fit* 2D sont montrées dans le tableau 14 avec leurs incertitudes statistiques. Celles-ci tiennent compte du nombre d'événements limité dans les régions de contrôle et dans les échantillons de signal et photon-jet simulés. Les contaminations sont significativement plus élevées par rapport à l'analyse précédente d'ATLAS [203] à cause de la nouvelle coupure sur l'isolation des traces des deux candidats, qui est appliquée à la fois dans la région de signal **TITI** et dans les régions de contrôle et diminue significativement le nombre de jets tout en ayant une excellente efficacité pour les photons immédiats (voir sections 14.3.1 et 16). La contamination des événements $\gamma\gamma$ attendue dans les différentes régions de contrôle des jets est de 15.7%, 11.8% et 1.9% respectivement pour **TITI**, **TITI** et **TITI**. Dans tous les cas, elle est déterminée avec une incertitude statistique très faible. En revanche, la contamination des événements γ +jet et jet+ γ dans **TITI** souffre de la statistique limitée des échantillons photon-jet simulés (voir section 13.2). Cela n'a aucun impact sur le nombre d'événements de signal, α'_1 et α'_2 n'entrant à aucun moment dans le calcul de $W_{\gamma\gamma}^{\text{TITI}}$ (voir section 15.3.4,

Fuites de signal	Coefficient	Fraction
$\gamma\gamma$ dans $\tilde{\mathbf{T}}\mathbf{T}\mathbf{I}\mathbf{I}$	α_1	0.157 ± 0.001 (MC) ± 0.0006 (données)
$\gamma\gamma$ dans $\mathbf{T}\tilde{\mathbf{T}}\mathbf{I}\mathbf{I}$	α_2	0.118 ± 0.0006 (MC) ± 0.0005 (données)
$\gamma\gamma$ dans $\tilde{\mathbf{T}}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$	α'_3	0.019 ± 0.0006 (MC) ± 0.00008 (données)
Fuites de bruit de fond		
jet+ γ dans $\tilde{\mathbf{T}}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$	α'_1	0.242 ± 0.193 (MC) ± 0.008 (données)
γ +jet dans $\tilde{\mathbf{T}}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$	α'_2	0.065 ± 0.023 (MC) ± 0.001 (données)

TAB. 14 – Taux de contamination des événements $\gamma\gamma$ dans les différentes régions de contrôle ($\alpha_1, \alpha_2, \alpha'_3$) et des événements γ +jet, jet+ γ dans $\tilde{\mathbf{T}}\mathbf{I}\tilde{\mathbf{T}}\mathbf{I}$ (α'_1, α'_2) attendus par le *template fit* 2D. Le calcul de ces taux est effectué avec les équations 54 à 58 de la section 15.3.4. Les incertitudes issues de la statistique limitée des régions de contrôle (*data*) et des échantillons simulés de signal ($\alpha_1, \alpha_2, \alpha'_3$) et photon-jet (α'_1, α'_2) sont également fournies (*MC*).

équation 64), mais diminue la précision sur la détermination des composantes γ +jet, jet+ γ et di-jet individuelles (voir section 15.6).

Dans le paragraphe suivant, la qualité du modèle d'ajustement (*goodness-of-fit*) est évaluée grâce à la mise en place de plusieurs estimateurs statistiques post-exécution.

Analyse de la qualité du modèle d'ajustement

La méthode du *template fit* 2D fait usage d'une procédure d'ajustement dont la qualité des résultats peut être évaluée en comparant le modèle final (après ajustement) avec les données observées. Cette comparaison est montrée dans la figure 106, où données et modèle final ont été projetés sur l'isolation de chacun des deux candidats. Les différentes composantes du modèle sont également visibles. Par ailleurs, afin de mieux quantifier les différences entre le modèle d'ajustement et les données, on définit dans chaque catégorie d'isolation notée k (*bins*) un *pull* [247] de la façon suivante :

$$P_k = \frac{n_k^{\text{data}} - n_k^{\text{model}}}{\sqrt{n_k^{\text{data}}}}, \quad (69)$$

où n_k^{model} et n_k^{data} représentent respectivement le nombre d'événements attendus et observés dans la catégorie k . Les *pulls* peuvent être calculés à partir des projections (voir figure 106) mais également à partir de l'ajustement 2D complet (voir figure 107, gauche). Dans le cas d'un ajustement non biaisé avec une précision limitée par la statistique des données, l'histogramme des *pulls* (voir figure 107, droite) suit une gaussienne standard ($\mu = 0, \sigma = 1$).

La qualité de l'ajustement est dans l'ensemble très satisfaisant, même si des différences résiduelles significatives sont observées à la fois dans la distribution des *pulls* projetés (voir figure 106) et 2D (voir figure 107, gauche), ce qui se traduit par un histogramme des *pulls* suivant une distribution gaussienne centrée en 0 (ajustement non biaisé) mais avec une largeur supérieure à 1 ($\sigma \simeq 1.3$). Cela est attendu dans la mesure où la précision sur la soustraction du bruit de fond est dominée par des effet systématiques et non statistiques (voir section 15.6), par exemple sur les formes des différentes pdfs d'isolation utilisées dans le modèle d'ajustement. Cependant, il est intéressant d'étudier

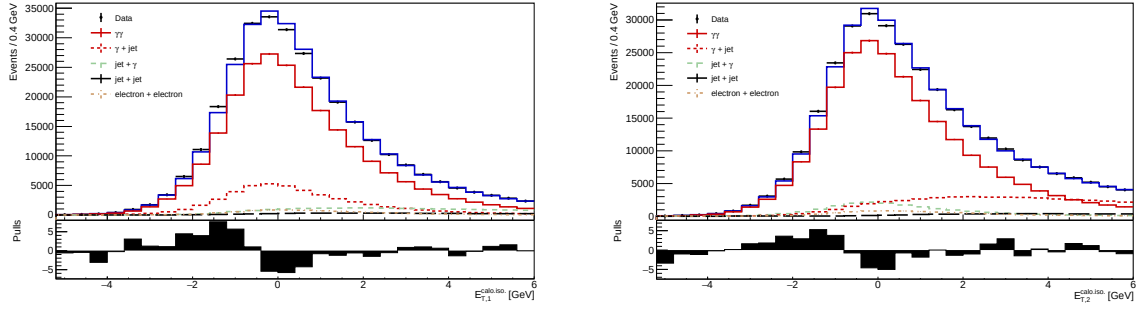


FIG. 106 – Projection des résultats finals d’ajustement du *template fit* 2D pour la région **TITI** inclusive sur l’isolation calorimétrique du premier candidat (gauche) et du second candidat (droit). Le modèle total (ligne continue bleue) est décomposé dans ses cinq composantes : $\gamma\gamma$ (ligne continue rouge), γ +jet (pointillés rouges), jet+ γ (pointillés verts), di-jet (pointillés noirs) et ee (pointillés bruns) et comparé aux données de collisions (points noirs). Les distributions des *pulls* (voir équation 15.5.1) sont également montrées pour une meilleure quantification des différences modèle/données.

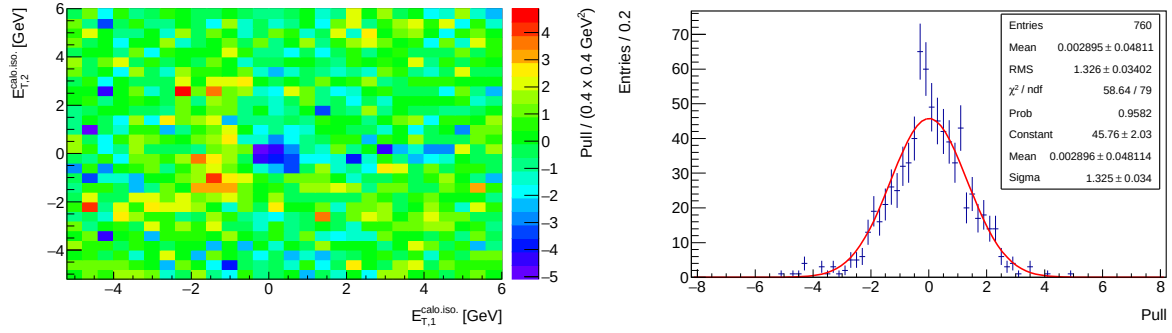


FIG. 107 – Distributions 2D des *pulls* (voir équation 15.5.1) dans le plan $E_{T,1}^{\text{calo.iso.}} - E_{T,2}^{\text{calo.iso.}}$ (gauche) et histogramme simple avec fit gaussien (droite). Pour les catégories en $E_{T,1}^{\text{calo.iso.}} - E_{T,2}^{\text{calo.iso.}}$ vides de données (24 sur 784), les *pulls* sont fixés à 0 mais non ajoutés à l’histogramme simple.

ces différences plus en détail.

D’après la figure 107, gauche, la plupart des différences significatives entre données et modèle d’ajustement se situent d’une part au niveau du pic de la distribution 2D autour de 0 (variations négatives des *pulls*) et d’autre part autour de la région $E_{T,1}^{\text{calo.iso.}}, E_{T,2}^{\text{calo.iso.}} \simeq -1.5$ GeV. La présence de déviations significatives au niveau du pic n’est pas anodine : c’est la région où la statistique dans les données est la plus élevée et donc où les données sont les plus sensibles aux déviations du modèle. Par ailleurs, vu la similarité entre les pdfs d’isolation calorimétrique des photons et des électrons dans cette région où $\gamma\gamma$ et ee dominant (voir section 14.3.2), il n’est pas impossible que ces défauts de modélisation expliquent la différence entre le nombre d’événements ee attendu avec le *template fit* 2D et l’estimation issue de la simulation $DY Z \rightarrow ee$ (voir paragraphe précédent). Plus de détails sur ce point seront fournis dans la section 15.6.

Afin de mieux quantifier l’écart observé, les sommes en quadrature des résidus entre le modèle complet et les données ($\sum_k [n_k^{\text{data}} - n_k^{\text{model}}]$) dans le cas 2D, projection 1D sur le

premier candidat et projection 1D sur le deuxième candidat sont évaluées. Elles s'élèvent respectivement à 965, 2 217 et 1 622 événements. Tous ces chiffres sont bien plus faibles que l'incertitude systématique totale assignée sur la composante $\gamma\gamma$ dans la section 15.6 (+8 070/-8 672).

Le modèle d'ajustement est donc considéré comme satisfaisant, avec des écarts entre données et modèle d'ajustement sous contrôle par rapport aux différents effets systématiques attendus. Ceux-ci seront abordés plus en détail dans la section 15.6. Dans la section 15.5.2, les résultats de décomposition pour les différentes observables de mesure sont présentés et analysés.

15.5.2 Spectres obtenus pour les différentes observables de mesure

Présentation des résultats de décomposition : histogrammes empilés

Dans un second temps, le *template fit* 2D est exécuté dans chacune des sous-catégories définies à l'aide des différentes observables de mesure présentées dans la section 15.3.5. Les pdfs d'isolation sont extraites indépendamment pour chaque sous-catégorie, sauf les pdfs d'électrons qui sont toujours issues des régions *tag-and-probe* complètes. Ce choix a déjà été justifié dans la section 15.3.5. L'ajustement consiste alors à minimiser un maximum de vraisemblance étendu dans l'intervalle $[-5.2 \text{ GeV}, 6 \text{ GeV}] \times [-5.2 \text{ GeV}, 6 \text{ GeV}]$, qui peut être catégorisé ou non selon la statistique disponible dans l'échantillon de signal simulé (voir section 15.3.2). En cas de catégorisation, des largeurs de $0.8 \text{ GeV} \times 0.8 \text{ GeV}$ sont utilisées afin de tenir compte de la baisse de statistique dans les données catégorisées par rapport au cas inclusif. La qualité du modèle d'ajustement est également estimée dans des catégories de largeur $0.8 \text{ GeV} \times 0.8 \text{ GeV}$ au lieu de $0.4 \text{ GeV} \times 0.4 \text{ GeV}$ ici, par souci de cohérence (voir paragraphe suivant).

Les résultats de décomposition sont tout d'abord montrés dans la figure 108 sous la forme d'histogrammes empilés pour les différents spectres. Uniquement la somme des trois composantes du bruit de fond jet (γ +jet, jet+ γ et di-jet) est disponible pour la décomposition selon les différentes observables, à cause de la statistique très limitée des échantillons simulés photon-jet qui ne permet pas de calculer correctement la contamination de $\tilde{\text{TIT}}\tilde{\text{T}}$ par γ +jet et jet+ γ (voir section 13.2).

L'épaule de Guillet, correspondant en fait à un coude dans les distributions d'impulsions transverses des di-photons $p_{T,\gamma\gamma}$ et a_T prédit par les calculs théoriques à partir du NLO (voir section 3.5) [123], est clairement visible pour $\gamma\gamma$ mais également pour le bruit de fond hadronique. En effet, ce coude est créé entre autres par l'augmentation brusque de la section efficace $\gamma\gamma$ grâce à l'ouverture de l'espace des phases NLO à $\Delta\phi_{\gamma\gamma} < \pi/2$ pour la composante directe.⁶⁵ Un effet similaire est attendu pour les productions γ +jet, jet+ γ et di-jet. Il est dans les faits peu observé lors des mesures de section efficaces di-jet et γ +jet à cause de la résolution expérimentale limitée sur l'impulsion transverse des jets. L'épaule de Guillet est en revanche absente pour la composante ee , car celle-ci a une section efficace de production QCD très faible en dehors de la résonance du boson Z^0 .

Comme attendu, les événements ee sont concentrés autour de $m_Z \simeq 91 \text{ GeV}$, avec

⁶⁵Il provient également d'une augmentation de la composante de fragmentation (voir section 3.5).

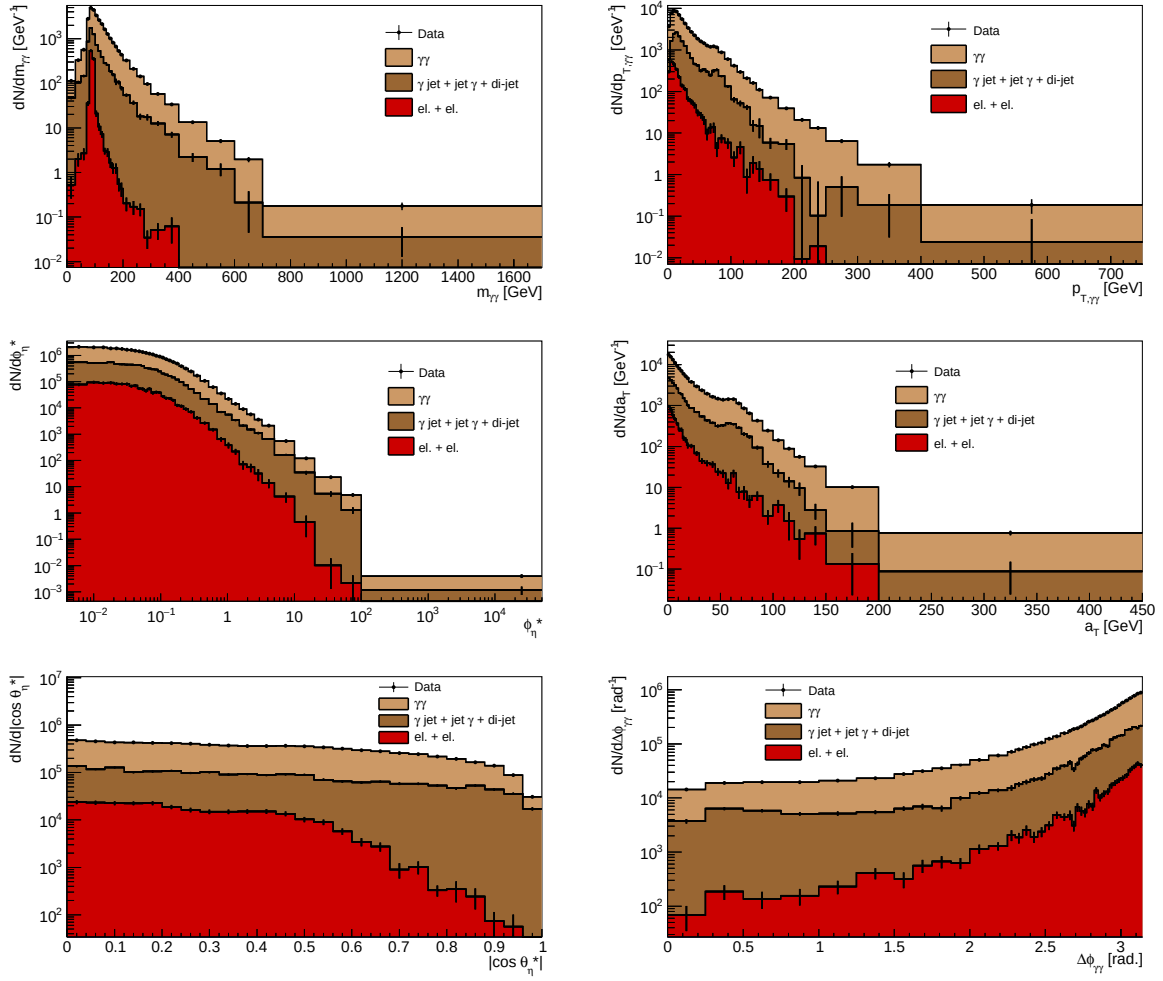


FIG. 108 – Décomposition du bruit de fond pour les différentes observables de mesure : $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T , ϕ_η^* , $|\cos \theta_\eta^*|$ et $\Delta\phi_{\gamma\gamma}$ (voir section 12.3), selon les sous-catégories (*bins*) définies dans la section 15.3.5 et obtenue avec la méthode du *template fit* 2D. Les composantes $\gamma\gamma$ (brun clair), jets (incluant γ +jet, jet+ γ et di-jet, brun foncé) et di-électron (rouge) déterminées par le *template fit* 2D sont empilées et comparées aux données observées dans les régions de signal (points noirs). Les barres d'erreurs sont uniquement issues de la statistique limitée des régions de signal dans les données.

une préférence pour $p_{T,\gamma\gamma} \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$ (production de Z^0 sans jets supplémentaires de haute énergie). Les événements attendus sont cette fois souvent en bon accord avec les prédictions de l'échantillon simulé DY, utilisé comme contrainte gaussienne pendant la procédure d'ajustement (voir section 15.3.3). Cette amélioration par rapport au cas inclusif vient très probablement de la statistique plus limitée des données, qui n'est plus suffisante pour faire dévier considérablement le nombre d'événements ee attendu de la valeur centrale de la contrainte. L'écart maximum est d'ailleurs atteint au niveau du pic du Z^0 ($90 \text{ GeV} < m_{\gamma\gamma} < 100 \text{ GeV}$), là où le rapport signal sur bruit pour les di-électrons est le plus favorable. La déviation entre le nombre d'événements ee attendu après ajustement est alors de -1.7σ par rapport à la valeur centrale de la contrainte (σ étant la largeur de la contrainte gaussienne). Les nombres totaux d'événements attendus dans **TITI** pour les différentes sous-catégories sont quant à eux à nouveau en bon accord avec les nombres observés pour l'ensemble de l'espace de mesure considéré.

La distribution en masse est considérablement déformée par la sélection en $E_{T,1}^\gamma$, $E_{T,2}$, qui ferme presque totalement l'espace des phases pour $m_{\gamma\gamma} < 70$ GeV et la limite considérablement pour $m_{\gamma\gamma} < 100$ GeV. Dans ces régions, uniquement des paires de candidats photons avec deux objets très proches en pseudo-rapacité et/ou azimut sont autorisées.⁶⁶ On y attend donc un nombre important de photons de fragmentation et d'événements avec plusieurs jets additionnels de haute énergie. D'autre part, la coupure asymétrique sur $E_{T,1}$, $E_{T,2}$ crée un pic autour de 5 GeV pour $p_{T,\gamma\gamma}$, correspondant très probablement à la transition entre une région dominée par $\gamma\gamma + 0$ jet (événements dos-à-dos dans le plan transverse et de même E_T^γ) vers $\gamma\gamma + \geq 1$ jet(s). Il est intéressant de remarquer que ce pic n'apparaît pas pour a_T .

Finalement, l'incertitude statistique est généralement sous contrôle, même si quelques fluctuations apparaissent parfois dans les queues des distributions de $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$ et a_T pour les composantes de bruit de fond. Cela a un impact très limité sur l'incertitude statistique de la composante de signal $\gamma\gamma$, qui reste très largement sous contrôle pour l'ensemble de l'espace des phases. En effet, la précision du *template fit* 2D est limitée dans la très grande majorité des sous-catégories par des effets systématiques, qui seront évalués dans la section 15.6.

Dans le paragraphe suivant, les résultats de décomposition dans **TITI** sont présentés en termes de pureté $\gamma\gamma$. Les différentes contaminations $\gamma\gamma$ dans les régions de contrôle de jets sont également examinées.

Présentation des résultats de décomposition : pureté et contamination dans les régions de contrôle

La pureté $\gamma\gamma$ ou pureté de signal est définie dans chaque sous-catégorie comme le nombre d'événements $\gamma\gamma$ attendus par le *template fit* 2D par rapport au nombre total d'événements observés. Elle correspond à la division de l'histogramme $\gamma\gamma$ de la figure 108 (brun clair) par l'histogramme des données (ligne noire) et est représentée pour les différentes observables dans la figure 109.

Les distributions montrent une pureté de signal croissante avec $p_{T,\gamma\gamma}$ et a_T . Cela est attendu de la baisse du bruit de fond jet à haut E_T^γ (voir section 12). Celles-ci semblent se stabiliser dans les queues de distribution, mais cela pourrait être dû à des fluctuations statistiques ou à la fusion des régions de signal dans l'échantillon de signal simulé SHERPA et des régions de contrôle mentionnées dans 15.3.5 pour l'exécution de la méthode. Une incertitude systématique couvrant cet effet sera évaluée dans la section 15.6 ($\simeq 2\%$). Par ailleurs, une baisse de pureté de 10% (5%) au niveau de l'épaule de Guillet pour $p_{T,\gamma\gamma}$ (a_T) est clairement visible, confirmant une hausse du bruit de fond dans cette région (voir paragraphe précédent). Deux baisses importantes de pureté sont également visibles à plus basse valeur, correspondant probablement à la région de transition $\gamma\gamma + 0 \text{ jet} \rightarrow \gamma\gamma + \geq 1 \text{ jet(s)}$ décrites précédemment.

Pour $m_{\gamma\gamma}$, la pureté est tout d'abord fortement croissante pour les mêmes raisons que $p_{T,\gamma\gamma}$ et a_T mais la tendance s'inverse pour $m_{\gamma\gamma} > 250$ GeV, avec une possible baisse de l'ordre de quelques pourcents. Cela vient du fait que les événements à très haute masse sont pour la plupart constitués d'une paire bouchon-bouchon (forte séparation en $|\eta^\gamma|$),

⁶⁶ $m_{\gamma\gamma}^2 = 2E_{T,1}^\gamma E_{T,2}^\gamma [\cosh(\eta_1^\gamma - \eta_2^\gamma) - \cos(\phi_1 - \phi_2)]$.

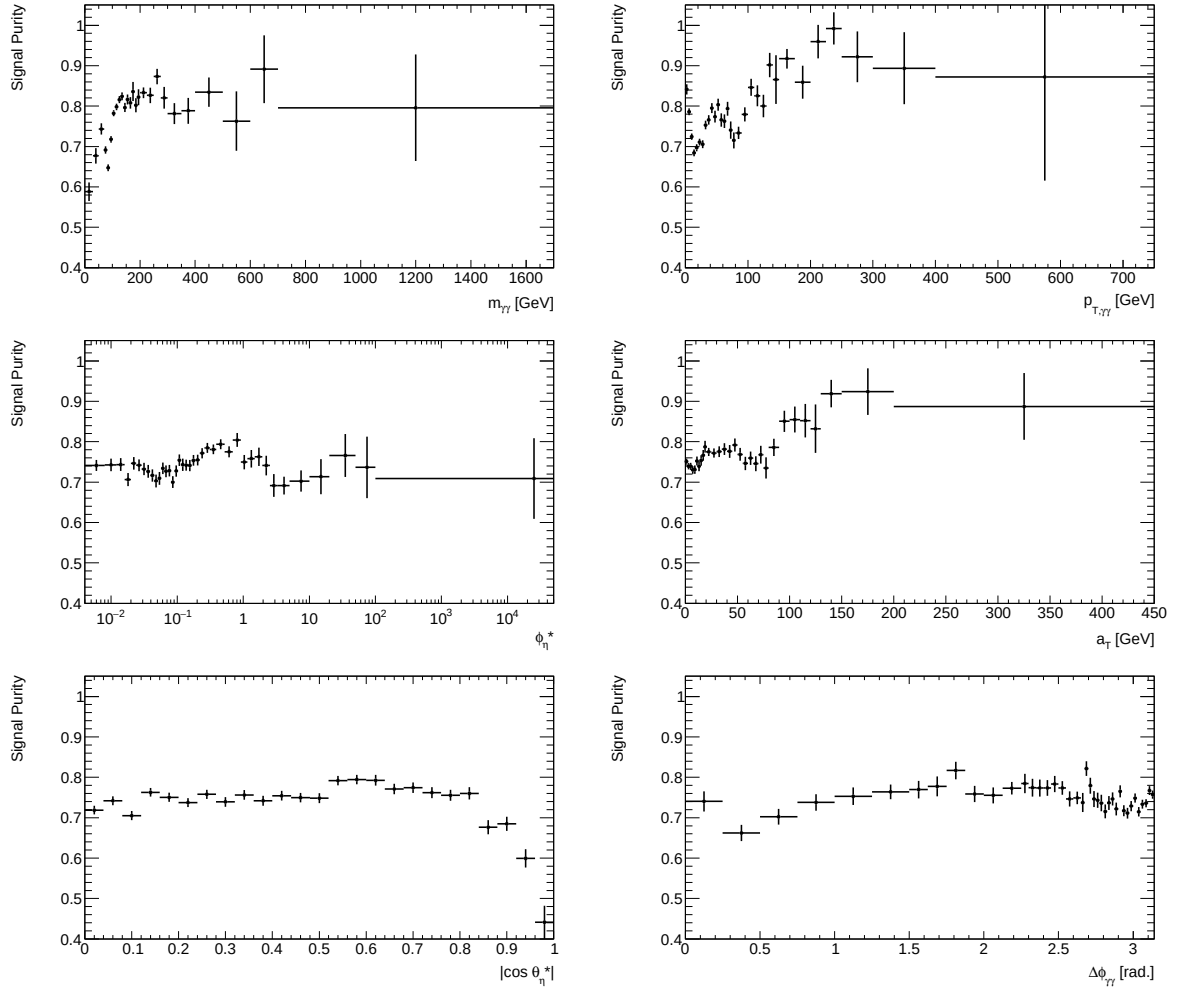


FIG. 109 – Décomposition du bruit de fond pour les différentes observables de mesure : $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T , ϕ_η^* , $|\cos \theta_\eta^*|$ et $\Delta\phi_{\gamma\gamma}$ (voir section 12.3), selon les catégories définies dans la section 15.3.5 et obtenue avec la méthode du *template fit* 2D. Les résultats sont exprimés en termes de pureté $\gamma\gamma$, correspondant à la proportion des événements $\gamma\gamma$ parmi le nombre total d'événements observés dans la catégorie. Les barres d'erreurs tiennent compte uniquement de la statistique limitée des régions de signal.

qui sont connus pour être plus affectés par le bruit de fond (voir section 15.1, figure 98). Un effet similaire est visible pour $|\cos \theta_\eta^*| > 0.8$.⁶⁷ D'autre part, une baisse de pureté pour $m_{\gamma\gamma}$ de l'ordre de 5 à 8% est visible entre 80 GeV et 100 GeV à cause de l'augmentation de la composante ee ($Z^0 \rightarrow ee$). La baisse de pureté pour $70 \text{ GeV} < m_{\gamma\gamma} < 80 \text{ GeV}$ est attribuée aux migrations importantes d'événements à plus haute masse dues à la résolution limitée sur le pic situé entre 70 et 100 GeV, particulièrement pour les candidats photons issus de jets mal identifiés. Les migrations dans le cas $\gamma\gamma$ seront étudiées plus en détail dans la section 16. Les distributions de pureté pour $\Delta\phi_{\gamma\gamma}$ et ϕ_η^* sont plus difficiles à interpréter, mais de manière générale plus plates que pour les autres observables. Une hausse de pureté de quelques pourcents dans la région $\phi_\eta^* \simeq 1$ est observée, probablement issue de la pureté croissante de $p_{T,\gamma\gamma}$ et a_T . La région à plus haut ϕ_η^* est affectée par la fragmentation (voir section 12.3), engendrant une baisse de pureté globale

⁶⁷ $\cos \theta_\eta^* = \tanh\left(\frac{\eta_1^\gamma - \eta_2^\gamma}{2}\right)$.

à cause des critères d'isolation.

Finalement, les fractions d'événements $\gamma\gamma$ dans les régions de contrôle $\tilde{\mathbf{TITI}} (\alpha_1)$, $\mathbf{TITI} (\alpha_2)$ et $\tilde{\mathbf{TITI}} (\alpha'_3)$ attendues par le *template fit* 2D sont montrées dans la figure 110. Celles-ci sont calculées grâce aux équations 54, 55 et 58 de la section 15.3.4, à l'exception des chiffres figurant dans les 3 (4, 5) dernières sous-catégories de $m_{\gamma\gamma} (a_T, p_{T,\gamma\gamma})$, qui sont extrapolées du reste de la distribution sous la forme de fonctions analytiques simples, du fait du manque de statistique dans les régions de contrôle associées (voir section 15.3.5).⁶⁸ Une incertitude conservatrice de $\pm 50\%$ sera prise sur ces derniers chiffres, couvrant l'ensemble des valeurs attendues pour α_1 , α_2 et α'_3 pour un spectre donné (voir section 15.6). Celle-ci aura un impact quasi-nul sur les résultats, les corrections dues aux contaminations $\gamma\gamma$ étant très limitées dans ces régions et la précision des mesures totalement dominée par la statistique limitée des données.

Les distributions de α_1 , α_2 et α'_3 sont difficiles à interpréter quantitativement car elles dépendent de l'évolution de l'efficacité d'identification TIGHT à la fois pour les photons et les jets mal identifiés et du rapport des sections efficaces de production. Elles sont donc montrées ici principalement à des fins illustratives. α'_3 est systématiquement inférieure à α_1 et α_2 , attendu du fait que les deux photons de la paire doivent être NON-TIGHT pour être classés $\tilde{\mathbf{TITI}}$. Il est intéressant de remarquer les régions où α_1 et α_2 diffèrent significativement (bas $|\cos \theta_\eta^*|$, grand $\Delta\phi_{\gamma\gamma}$), indiquant probablement des domaines où les distributions cinématiques des deux candidats ($|\eta^\gamma|$, E_T^γ) sont significativement différentes, que ce soit pour le signal et/ou pour le bruit de fond.

Dans le paragraphe suivante, la qualité des modèles d'ajustement employés pour les différentes observables est évaluée.

Analyse de la qualité du modèle d'ajustement

Il est intéressant d'évaluer la qualité du modèle d'ajustement, y compris lors de la décomposition du bruit de fond selon les différentes observables. En effet, cela peut permettre d'identifier des régions mieux décrites que d'autres par le modèle et ainsi préciser les axes d'amélioration. Dans cette optique, une décomposition du bruit de fond dans des sous-catégories 1D en $E_{T,1}$, $E_{T,2}$, η_1 et η_2 a été effectuée en plus des décompositions selon les observables de mesures. Les définitions des sous-catégories sont données ci-dessous :

- ◇ $E_{T,1}$ [GeV] : 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 65, 70, 75, 80, 85, 90, 95, 100, 120, 140, 160, 180, 200, 250, 325, 450, 1500.
- ◇ $E_{T,2}$ [GeV] : 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 70, 80, 100, 150, 200, 300, 800.
- ◇ η_1 [sans unité] : -2.37, -1.81, -1.56, -1.37, -0.6, 0, 0.6, 1.37, 1.56, 1.81, 2.37.
- ◇ η_2 [sans unité] : -2.37, -1.81, -1.56, -1.37, -0.6, 0, 0.6, 1.37, 1.56, 1.81, 2.37.

⁶⁸ $m_{\gamma\gamma} : \alpha_1 (\alpha_2, \alpha'_3) = 0.16 (0.12, 0.02)$.

$a_T : \alpha_1 (\alpha_2, \alpha'_3) = 0.2 + 0.01 \times X (0.14 - 0.008 \times X, 0.04 + 0.008 \times X)$, X allant de 1 à 4.

$p_{T,\gamma\gamma} : \alpha_1 (\alpha_2, \alpha'_3) = 0.25 + 0.015 \times X (0.2 + 0.015 \times X, 0.05 + 0.008 \times X)$, X allant de 1 à 5.

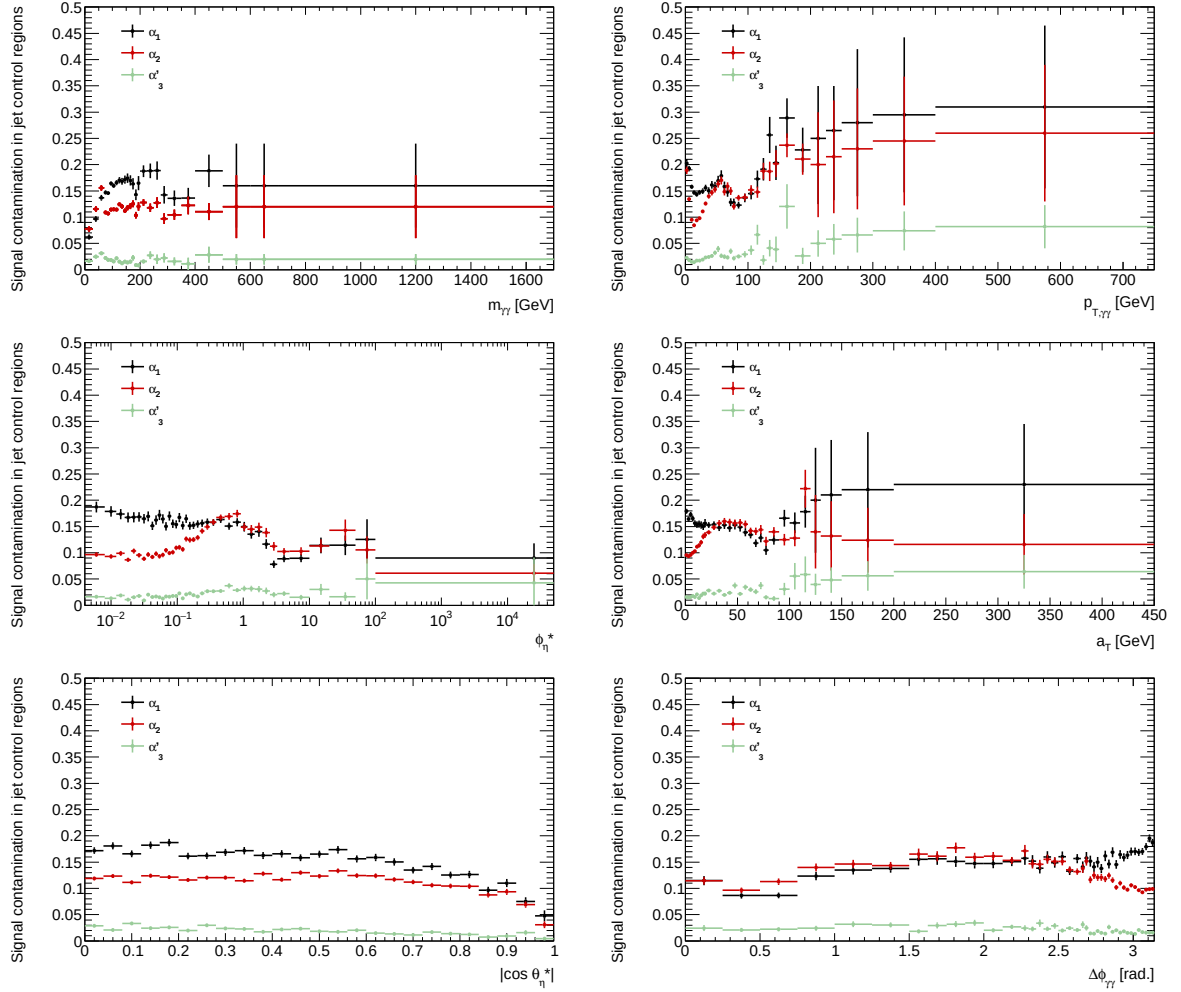


FIG. 110 – Résultats du *template fit* 2D en termes de fraction d'événements $\gamma\gamma$ dans les régions de contrôle $\tilde{\mathbf{T}}\mathbf{T}\mathbf{I}$ (α_1), $\mathbf{T}\mathbf{T}\mathbf{I}$ (α_2) et $\tilde{\mathbf{T}}\mathbf{I}\mathbf{T}$ (α'_3). Ces facteurs sont déterminées à partir du MC de signal SHERPA, du nombre d'événements observés dans les régions de contrôle et des résultats du *template fit* 2D grâce aux équations 54, 55 et 58 de la section 15.3.4, à l'exception des 3 (4, 5) dernières sous-catégories de $m_{\gamma\gamma}$ (a_T , $p_{T,\gamma\gamma}$) où elles sont extrapolées du reste de la distribution (voir texte). Une incertitude totale conservatrice de $\pm 50\%$ est considérée pour celles-ci (montrée dans la figure), tandis que pour les autres les barres d'erreur représentent uniquement l'incertitude statistique (MC et statistique des régions de contrôle).

Les sous-catégories mentionnées ci-dessus sont unidimensionnelles (1D), signifiant par exemple que la sous-catégorie $40 \text{ GeV} < E_{T,1} < 42 \text{ GeV}$ n'induit aucun critère particulier sur la cinématique du deuxième candidat. Quelques résultats sont présentés dans les figures 111 à 114, correspondant à des projections du modèle complet ajusté dans certaines sous-catégories de l'échantillon $\mathbf{T}\mathbf{T}\mathbf{I}$. La décomposition est effectuée avec et sans corrections supplémentaires d'isolation basées sur les données (voir section 10.4), afin d'en évaluer leur impact sur la qualité du modèle et vérifier leur bonne application sur un échantillon di-photon. La qualité du modèle est estimée grâce à la distribution des *pulls* 1D des projections et grâce à la RMS de l'histogramme des *pulls* affichée dans les différents cas (*2D pull RMS*).

Dans tous les cas, les corrections supplémentaires d'isolation basées sur les données

$(\gamma + X)$ mènent à une réduction de la RMS de l'histogramme des *pulls*, démontrant la pertinence de ces corrections pour améliorer la qualité de l'ajustement. Des décalages similaires à ceux observés dans l'étude de performance sont observés pour les pdfs photons si les corrections ne sont pas appliquées (voir section 10.4), pouvant mener à un biais sur la détermination de la pureté, particulièrement à haut E_T^γ (voir figure 112). Ces corrections sont fournies avec une certaine incertitude qui sera propagée dans les résultats dans la section 15.6 et mènera à un des facteurs dominants limitant la précision sur l'extraction du bruit de fond.

Une fois la pertinence des corrections établie, il est intéressant de comparer la RMS de l'histogramme des *pulls* dans les différentes sous-catégories. Il apparaît qu'un désaccord plus important entre modèle et données semble être présent pour les deux sous-catégories $1.81 < \eta_1 < 2.37$, $1.81 < \eta_2 < 2.37$ (RMS d'environ 1.4 au lieu de 1.3), qui correspondent aux parties les plus externes de l'acceptance géométrique. Le détail de la distribution des *pulls* 2D et de l'histogramme des *pulls* pour ces deux sous-catégories est donné dans la figure 115. De façon similaire au cas inclusif (voir section 15.5.1), le désaccord le plus fort est localisé au niveau du pic de la distribution d'isolation 2D autour de 0 (variations négatives des *pulls*) et autour de la région $E_{T,1}^{\text{calo.iso.}}, E_{T,2}^{\text{calo.iso.}} \simeq -1.5$ GeV. La statistique étant plus limitée que dans le cas inclusif, la segmentation des *pulls* est plus grossière ($0.8 \text{ GeV} \times 0.8 \text{ GeV}$). D'autre part, à haut E_T^γ (voir figure 112), l'accord données/modèle est très bon (RMS des *pulls* 2D $\simeq 1.0$), ce qui est probablement d'abord dû à la statistique beaucoup plus limitée dans ces régions. Partout ailleurs (petit E_T^γ , petit $|\eta^\gamma|$), l'accord est compatible avec le cas inclusif.

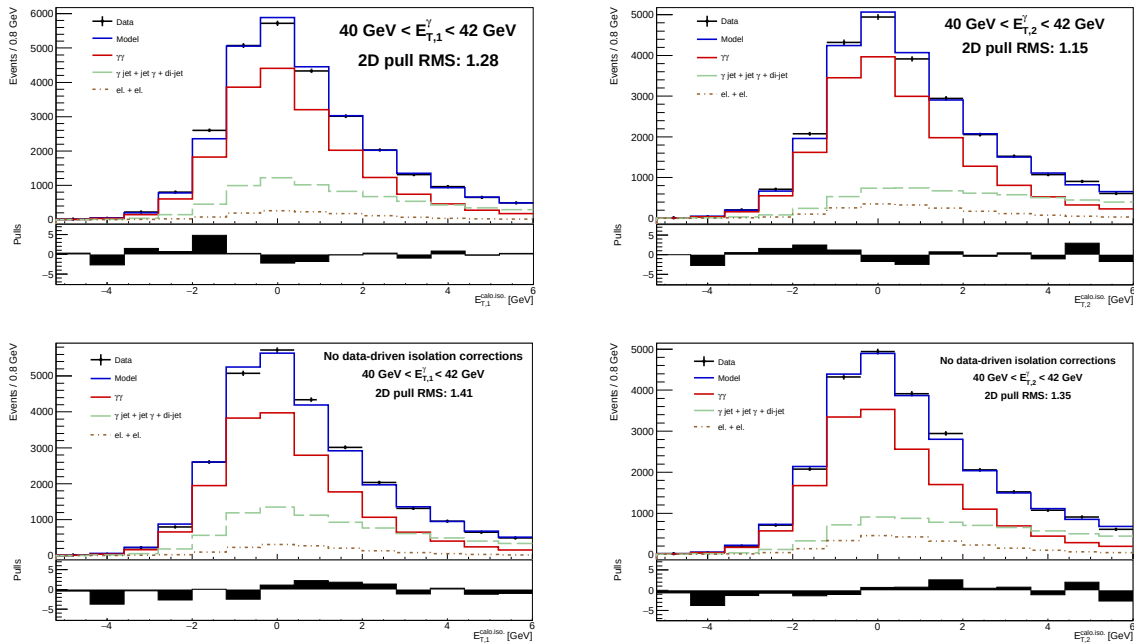


FIG. 111 – Projection du modèle d'isolation 2D après ajustement sur l'isolation calorimétrique du premier (gauche) et du deuxième (droite) candidat photon pour $40 \text{ GeV} < E_{T,1}$ (gauche), $E_{T,2}$ (droite) $< 42 \text{ GeV}$ avec (haut) et sans (bas) corrections supplémentaires d'isolation basées sur les données (voir section 10.4). Le modèle d'ajustement décrit correctement les données dans tous les cas (RMS des *pulls* 2D $\simeq 1.3$).

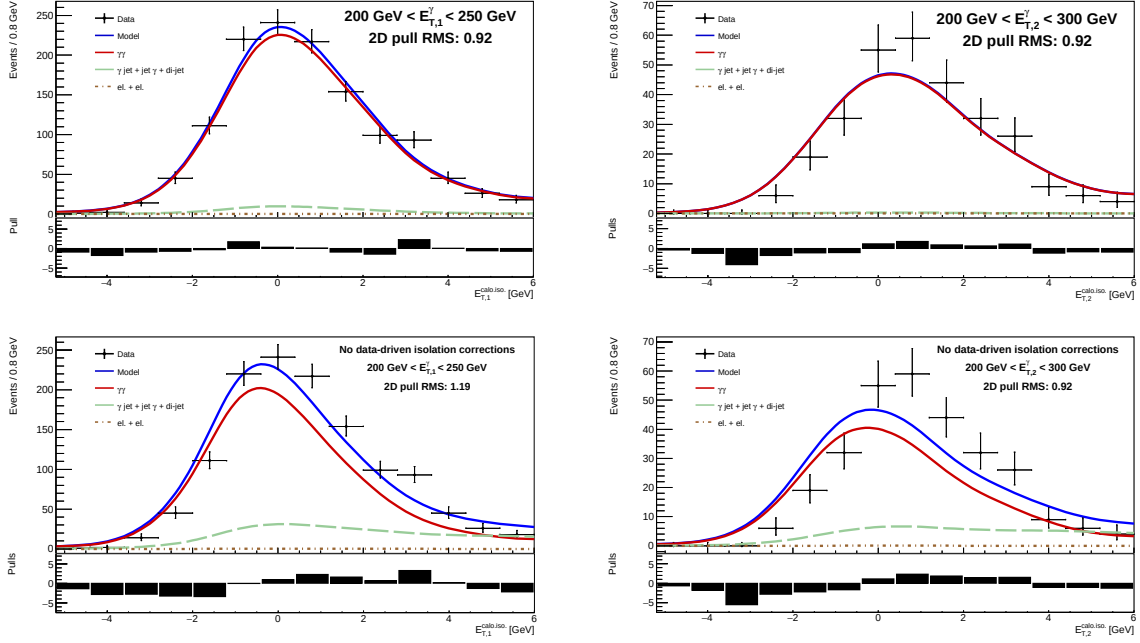


FIG. 112 – Figure similaire à la figure 111 pour $200 \text{ GeV} < E_{T,1} < 250 \text{ GeV}$ (gauche) et $200 \text{ GeV} < E_{T,2} < 300 \text{ GeV}$ (droite). La statistique étant plus faible ici, un modèle non catégorisé (*unbinned*) d'ajustement est adopté. La pureté de signal est sous-estimée si les corrections supplémentaires d'isolation ne sont pas prises en compte, avec un décalage du modèle par rapport aux données bien visible. Autrement, un très bon accord est observé dans la statistique disponible (RMS des *pulls* 2D $\simeq 1.0$).

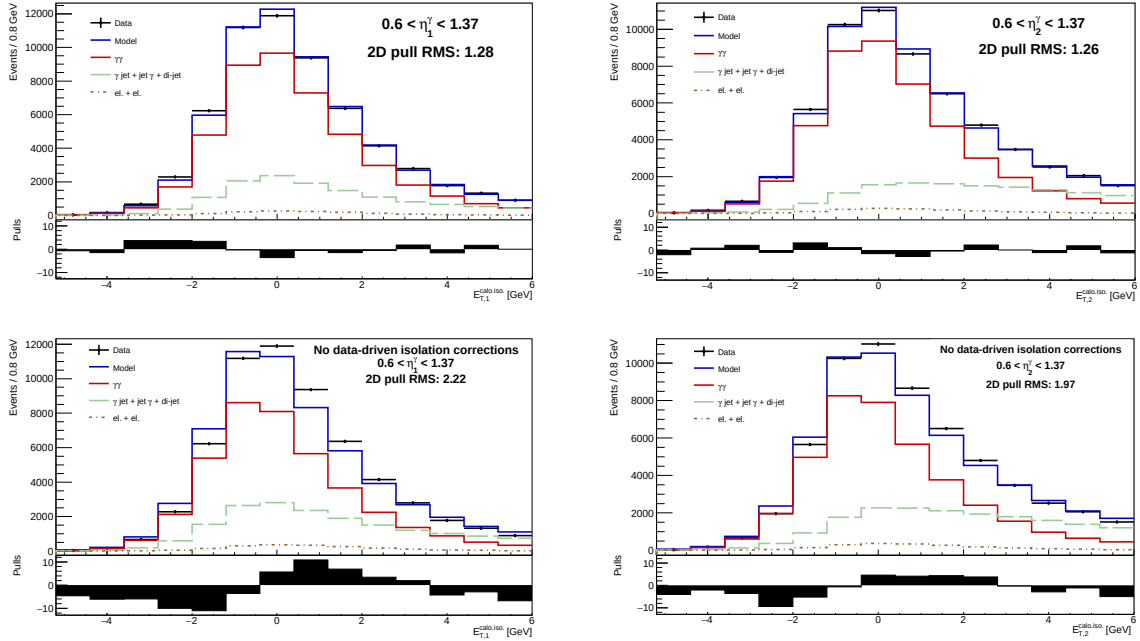


FIG. 113 – Figure similaire à la figure 111 pour $0.6 < \eta_1 < 1.37$ (gauche) et $0.6 < \eta_2 < 1.37$ (droite). Un décalage du modèle par rapport aux données est clairement visible si les corrections supplémentaires d'isolation ne sont pas prises en compte. Autrement, le modèle d'ajustement décrit correctement les données (RMS des *pulls* 2D $\simeq 1.3$).

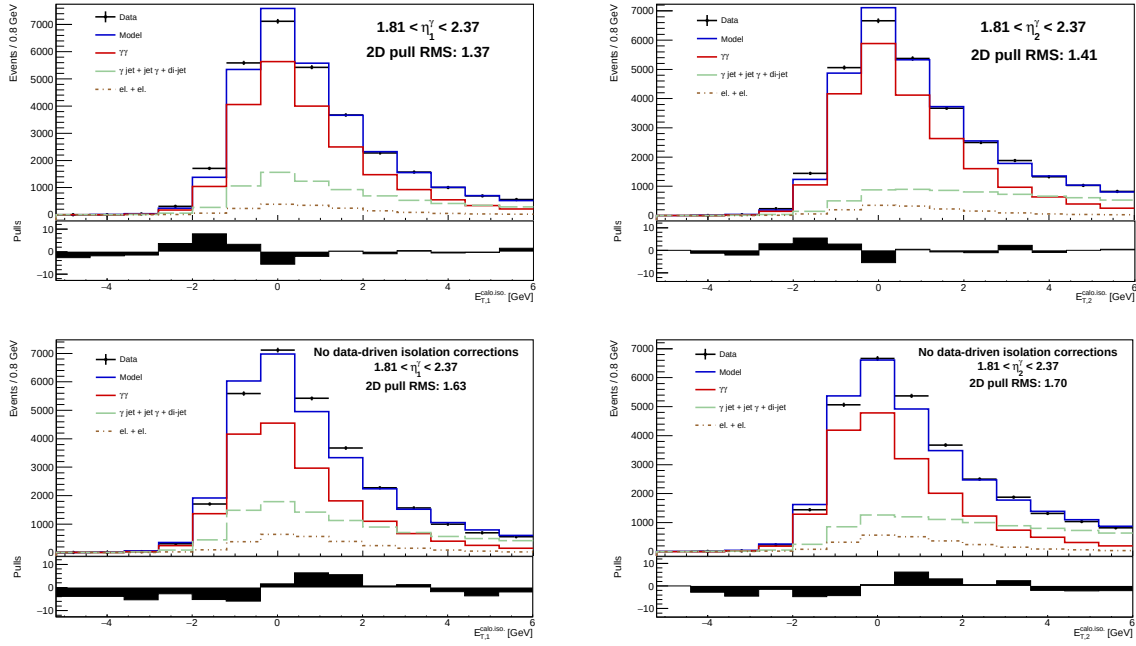


FIG. 114 – Figure similaire à la figure 111 pour $1.81 < \eta_1 < 2.37$ (gauche) et $1.81 < \eta_2 < 2.37$ (droite). Un décalage du modèle par rapport aux données est visible si les corrections supplémentaires d'isolation ne sont pas prises en compte. Autrement, un accord correct est observé, bien que significativement moins bon que pour les pseudo-rapidités centrales (RMS des *pulls* 2D $\simeq 1.4$).

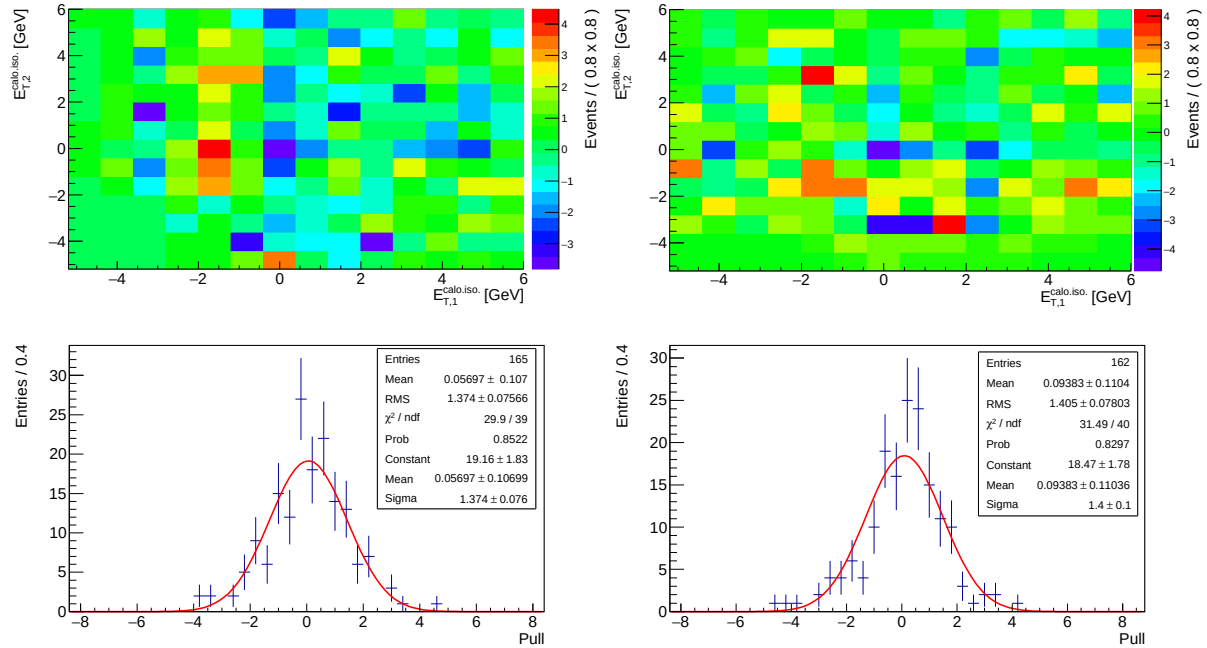


FIG. 115 – Distribution des *pulls* 2D (haut) et histogramme des *pulls* (bas) pour les sous-catégories correspondant à $1.81 < \eta_1 < 2.37$ (gauche) et $1.81 < \eta_2 < 2.37$ (droite).

Il a été tenté de localiser plus précisément l'origine des résidus dans l'espace des phases. Ainsi, l'accord données/modèle a été analysé dans les autres sous-catégories considérées dans l'analyse (en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc), quatre d'entre elles étant détaillées dans les figures 116 à 119 : $70 \text{ GeV} < m_{\gamma\gamma} < 80 \text{ GeV}$, $110 \text{ GeV} < m_{\gamma\gamma} < 120 \text{ GeV}$, $8 \text{ GeV} < p_{T,\gamma\gamma} < 12 \text{ GeV}$ et $30 \text{ GeV} < p_{T,\gamma\gamma} < 35 \text{ GeV}$. Les RMS des histogrammes des *pulls* sont toutes comprises entre 1.1 et 1.4, signe qu'aucune région plus spécifique de l'espace des phases montrant un désaccord particulier n'a pu être identifiée. En conséquence, une incertitude systématique conservatrice sera adoptée sur la modélisation de l'isolation des photons dans le *template fit* 2D, comme nous le verrons dans la section suivante.

Dans la section 15.6, les différents effets systématiques pouvant altérer les résultats du *template fit* 2D sont listés et leurs actions individuelles sur les différentes composantes sont évalués. La somme en quadrature de ces différents effets forme alors l'incertitude systématique totale reliée à la soustraction du bruit de fond. Elle constitue l'incertitude dominante de la mesure et donc celle limitant sa précision, à l'image des mesures $\gamma\gamma + X$ précédentes réalisées dans les collisionneurs hadroniques (voir section 12.4).

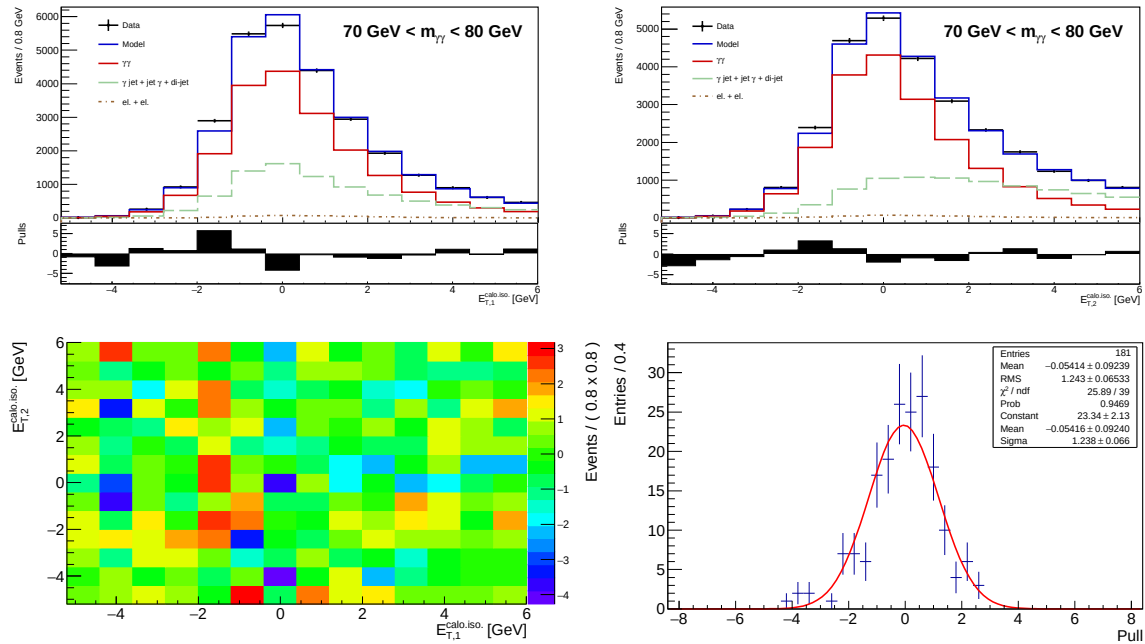
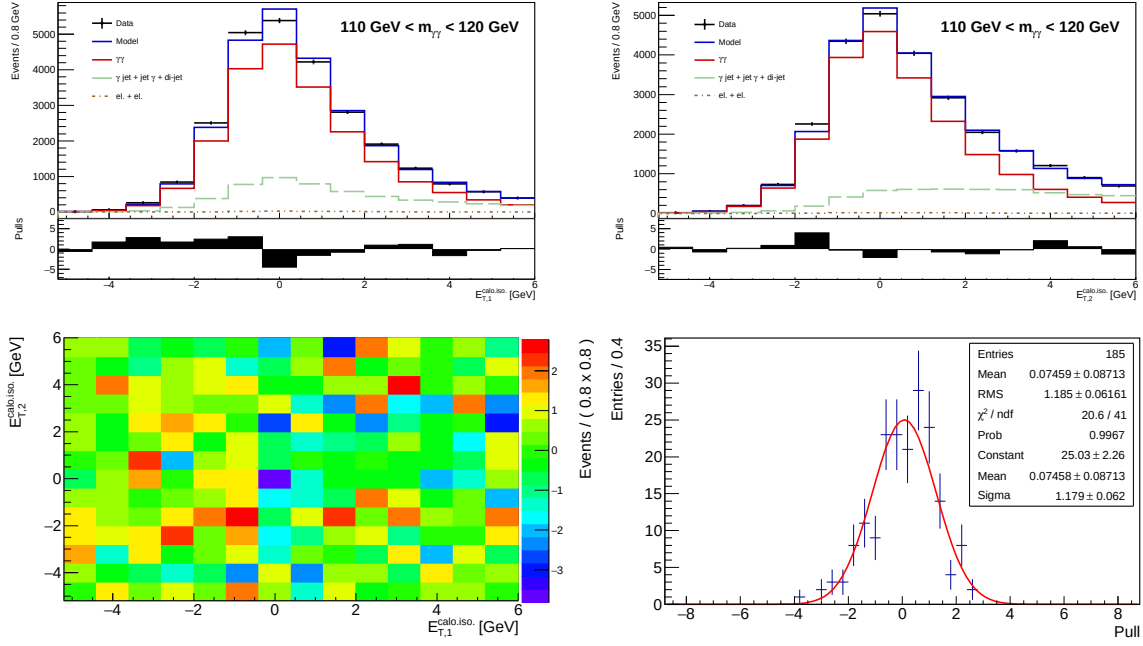
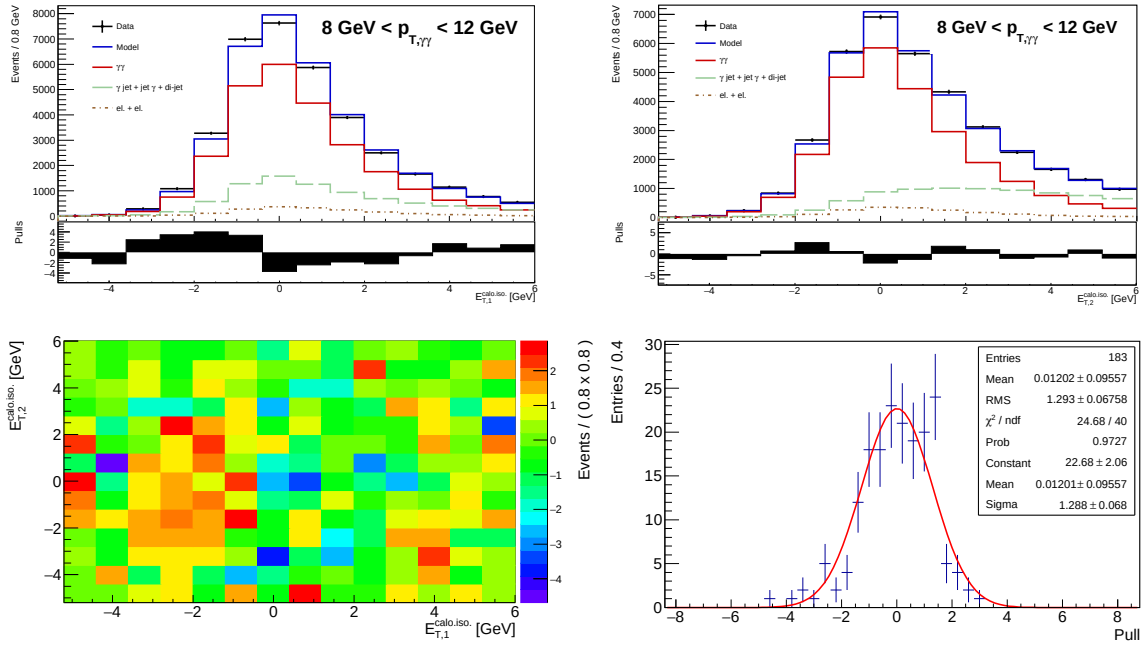
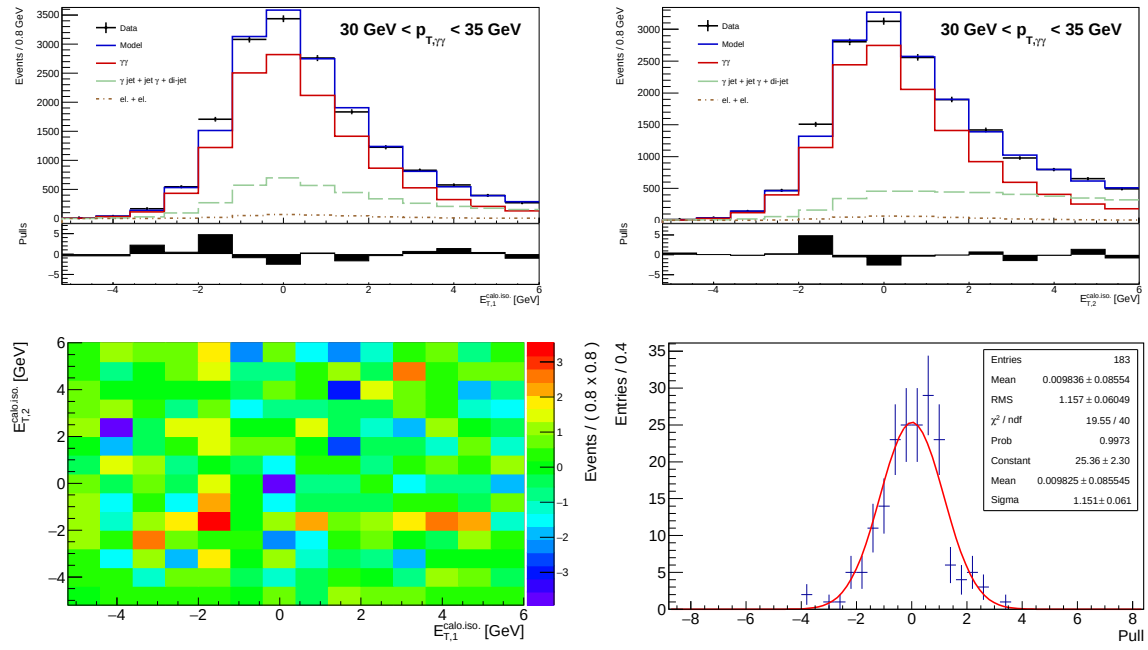


FIG. 116 – Projection du modèle d'isolation 2D après ajustement sur l'isolation calorimétrique du premier (gauche) et du deuxième (droite) candidat photons pour $70 \text{ GeV} < m_{\gamma\gamma} < 80 \text{ GeV}$ (haut) et distribution 2D et histogramme des *pulls* associés (bas).

FIG. 117 – Figure similaire à la figure 116 pour $110 \text{ GeV} < m_{\gamma\gamma} < 120 \text{ GeV}$.FIG. 118 – Figure similaire à la figure 116 pour $8 \text{ GeV} < p_{T,\gamma\gamma} < 12 \text{ GeV}$.

FIG. 119 – Figure similaire à la figure 116 pour $30 \text{ GeV} < p_{T,\gamma\gamma} < 35 \text{ GeV}$.

15.6 Étude des incertitudes associées

Dans cette section, les effets systématiques pouvant altérer les résultats du *template fit* 2D sont listés et évalués dans le cadre de l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV.

Dans les sections 15.6.1 à 15.6.3, les effets systématiques reliés à la modélisation des pdfs photons, des pdfs jets et des pdfs électrons sont évalués tour à tour. Les sources d'incertitudes ayant des conséquences sur l'ensemble du modèle, incluant les différences de cinématique entre modèle et données déjà discutées dans la section 15.4 lors des tests avec les pseudo-données et les biais potentiels venant de la méthode de lissage des différentes pdfs basées sur les estimateurs de noyaux gaussiens (voir section 15.3.2), sont traitées dans la section 15.6.4. Finalement, les résultats complets concernant les incertitudes de la méthode du *template fit* 2D sont exposés dans la section 15.6.5, incluant ceux concernant la région **TITI** complète et sa décomposition pour les différents spectres.

15.6.1 Incertitudes spécifiques à la modélisation des photons

Choix du générateur MC

(éléments de matrice, cascades de partons, événement sous-jacent, etc)

Les distributions d'isolation des di-photons dépendent significativement des $|\eta^\gamma|$, E_T^γ de chacun des photons de la paire mais également de la topologie du signal, que ce soit en terme d'éléments de matrice (nombre de jets additionnels émis à haute énergie), des paramètres des cascades de partons ou de processus de QCD non-perturbative (photons de fragmentation, événement sous-jacent, ...). Par conséquent, le choix de l'analyse a été d'utiliser un générateur MC pour modéliser les pdfs d'isolation des photons nécessaires à la réalisation du *template fit* 2D (voir section 15.3.1).

Le générateur utilisé par défaut est SHERPA 1.4.0, qui inclut les éléments de matrice LO $\gamma\gamma + 0, 1, 2$ jet(s) avec un algorithme de cascades de partons et un algorithme fusionnant les deux composantes (*ME+PS merged at LO*, voir section 13.2) [83]. L'échantillon simulé employé pour l'estimation des incertitudes systématiques est généré avec PYTHIA8, qui inclut les éléments de matrice $\gamma\gamma + 0$ jet et γ +jet avec cascade de partons (sans algorithme de fusion) pour modéliser la composante directe et la composante de fragmentation, respectivement. Ce choix est motivé par le fait que, naïvement, on s'attend à une meilleure description des données par SHERPA, qui inclut plus de jets additionnels de haute énergie dans l'état final et un algorithme de fusion. Cela étant vérifié lors de la comparaison entre la mesure finale et les différentes prédictions théoriques (voir section 17), ce choix est validé définitivement.

Cependant, on s'attend également à ce que la description de SHERPA soit limitée. En particulier, la coupure de séparation entre photons et partons appliquée au niveau générateur pour assurer la convergence des processus $\gamma\gamma + 1, 2$ jets ($\Delta R_{\gamma, q(g)} > 0.3$) biaise probablement les queues des distributions d'isolation. Ainsi, l'ensemble de la procédure du *template fit* 2D est répété avec les pdfs du générateur PYTHIA8 et la différence avec les résultats nominaux (obtenus avec les pdfs SHERPA) est prise comme erreur systématique. Le modèle ajusté utilisant les pdfs photons de PYTHIA8 est comparé avec les données dans la figure 120. Comme attendu, l'accord est moins bon qu'avec les pdfs SHERPA, avec en moyenne des pulls plus élevés en valeurs absolues et une RMS pour

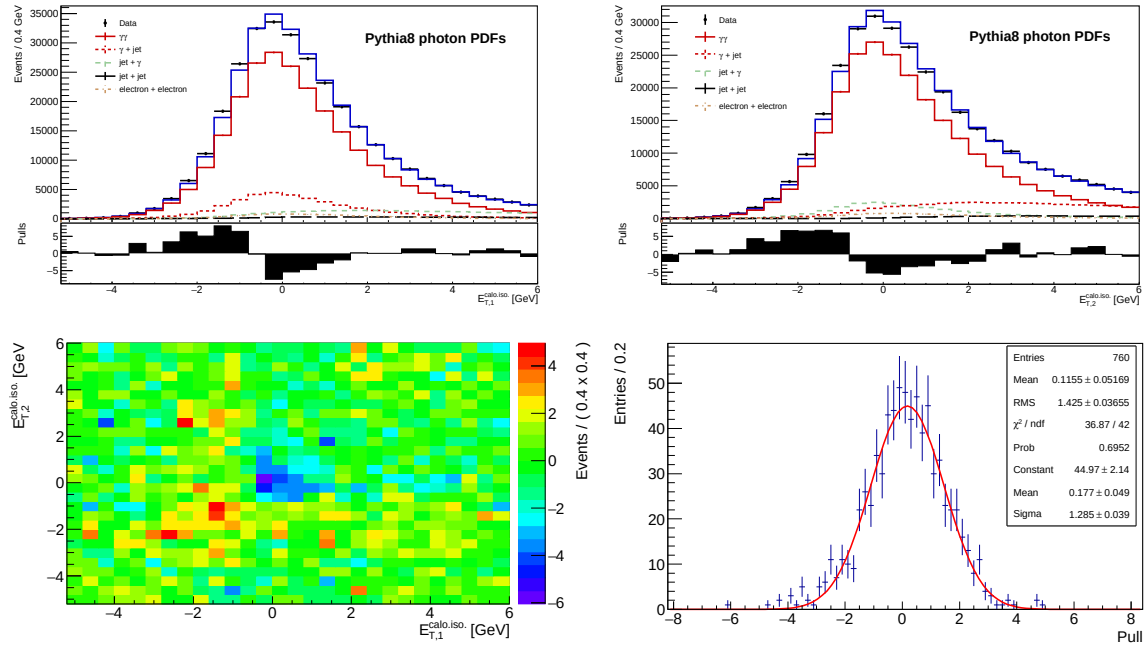


FIG. 120 – Projection des résultats finals d’ajustement du *template fit* 2D pour la région **TITI** inclusive sur l’isolation calorimétrique du premier candidat (gauche) et du second candidat (droite). Le modèle total (ligne continue bleu) est décomposé dans ses cinq composantes : $\gamma\gamma$ (ligne continue rouge), γ +jet (pointillés rouges), jet+ γ (pointillés verts), di-jet (pointillés noirs) et ee (pointillés bruns) et comparé aux données de collisions (points noirs). Les distributions des *pulls* (voir équation 15.5.1) sont également montrées pour une meilleure quantification des différences modèle/données. Les pdfs d’isolation utilisées pour les photons viennent ici de l’échantillon de signal PYTHIA8, qui est utilisé pour l’évaluation des systématiques.

l’histogramme des *pulls* supérieure à 1.4 (1.3 avec SHERPA). La distribution des *pulls* est également en légère tension avec le modèle gaussien, justement à cause des résidus négatifs plus élevés autour de $E_{T,1}^{\text{calo.iso.}}, E_{T,2}^{\text{calo.iso.}} = (0,0)$.

Comme nous le verrons dans la section 15.6.5, cette erreur systématique est l’incertitude dominante du *template fit* 2D, correspondant à un écart d’environ +1.7% sur la pureté $\gamma\gamma$ ($\simeq +2.2\%$ sur le nombre d’événements). Elle est symétrisée afin de tenir compte d’une éventuelle déviation des résultats dans le sens opposé et ainsi rester conservateur sur la validité des pdfs de SHERPA.

Cependant, des fluctuations statistiques importantes de cette erreur systématique sont observées lors de la décomposition selon les différents spectres. Par conséquent, une procédure de lissage est adoptée pour la systématique finale sur $\gamma\gamma$, implémentée de la façon suivante :

- ◊ Tout d’abord, 1 000 répliques *bootstrap*⁶⁹ sont générées de manière indépendante pour chaque région de signal. En pratique, dans le cas simple qui nous intéresse, cela revient à assigner 1 000 poids poissoniens indépendants centrés en 1 pour

⁶⁹Le *bootstrap* est une méthode de rééchantillonnage générale similaire à la méthode bien connue de *toy MC* à la différence près qu’elle s’applique facilement à des données non-catégorisées (*unbinned*).

chaque événement dans l'échantillon **TITI** des données.

- ◇ Pour chacune des répliques, la procédure de *template fit* 2D est exécutée avec SHERPA puis avec PYTHIA, ce qui permet d'extraire une valeur de systématique par réplique tout en gardant intactes les corrélations statistiques entre les résultats obtenus avec SHERPA et PYTHIA, issues de l'utilisation des mêmes données **TITI** pour l'ajustement. En revanche, les fluctuations statistiques (corrélées également) correspondant à la statistique limitée des régions de contrôle sont négligées.
- ◇ Dans chaque catégorie en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc, la RMS de la distribution des incertitudes systématiques est sommée en quadrature avec la différence entre la moyenne de la distribution des incertitudes et l'incertitude systématique nominale (*bootstrap non-closure*). Cette dernière est toujours sous-dominante. On obtient alors une première valeur pour l'incertitude sur la systématique considérée.
- ◇ Afin de tenir compte des fluctuations MC (cette fois indépendantes entre SHERPA et PYTHIA) et éviter ainsi un double comptage avec l'incertitude correspondant à la statistique limitée de l'échantillon de signal simulé (traité plus loin dans cette section), 1 000 répliques *bootstrap* des régions **TITI** des échantillons simulés SHERPA et PYTHIA sont également produites et la procédure de *template fit* 2D est exécutée à nouveau sur les données nominales avec chacune des répliques MC. Les variations obtenues avec les répliques de SHERPA et de PYTHIA sont ajoutées en quadrature à l'incertitude sur la systématique. Plus de détails sur cette étape sont disponibles dans le paragraphe de cette section intitulé *statistique limitée de l'échantillon de signal simulé* abordé plus loin.
- ◇ Pour chaque spectre ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc), les catégories sont fusionnées jusqu'à ce que les erreurs systématiques soient considérées comme significatives. Le critère de significativité standard consiste à demander une valeur centrale de l'incertitude à plus de trois déviations standard de 0, correspondant à un intervalle de confiance supérieur à 99% dans l'approximation gaussienne. Si ce critère n'est pas atteint après la fusion de trois catégories, la valeur centrale de l'incertitude et la barre d'erreur associée issue de la fusion sont tout de même considérées comme significatives.
- ◇ Une procédure de lissage (*smoothing*) à nouveau basée sur des estimateurs de noyaux gaussiens [207] est finalement utilisée afin de déterminer la distribution des systématiques finales à partir de la distribution des systématiques significatives.

Les résultats de la procédure sont résumés dans la figure 121. Il est intéressant de remarquer que la différence entre les résultats obtenus avec SHERPA et PYTHIA sont maximaux dans les régions de l'espace des phases où des configurations impliquant plusieurs jets additionnels de haute énergie et/ou des photons de fragmentation sont attendus : $m_{\gamma\gamma} < 100$ GeV, région de transition $\gamma\gamma + 0 \text{ jet} / \gamma\gamma + \geq 1 \text{ jet}$ et épaule de Guillet pour $p_{T,\gamma\gamma}$ et a_T ,⁷⁰ $\phi_\eta^* > 1$ et $\Delta\phi_{\gamma\gamma} < \pi$, démontrant que l'incertitude calculée saisit bien l'effet que l'on cherche à évaluer.

⁷⁰PYTHIA ne reproduit pas l'épaule de Guillet dans ses prédictions, voir section 17.

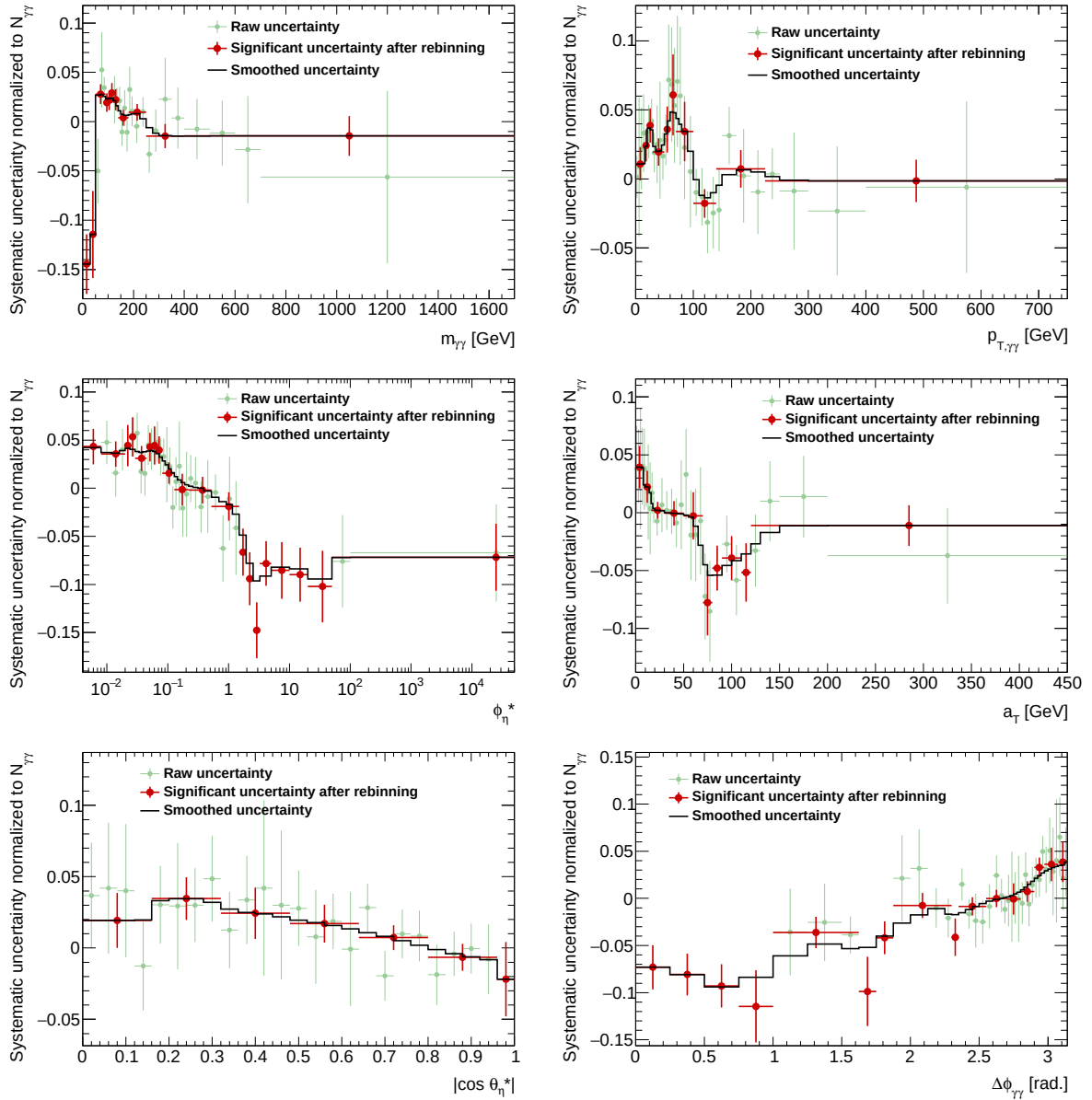


FIG. 121 – Incertitude systématique sur la soustraction du bruit de fond venant du choix du générateur MC pour la modélisation de l’isolation des photons. L’incertitude de départ (*raw uncertainty*, barres horizontales vertes) est évaluée comme la différence entre le nombre d’événements $\gamma\gamma$ obtenu avec le modèle utilisant les pdfs d’isolation photons PYTHIA et SHERPA. Son incertitude (barres verticales vertes) est évaluée par la technique de *bootstrap* décrite dans le texte. Les points rouges (*significant uncertainty*) représentent les incertitudes après le test de significativité impliquant la fusion des différentes catégories (voir texte). Finalement, la courbe noire (*smoothed uncertainty*) est la systématique finale après lissage retenue pour la mesure.

Note : les barres d’erreur verticales vertes et rouges sont légèrement sous-estimées car celles-ci ne tiennent pas compte des fluctuations dues à la statistique limitée des régions de contrôle, ce qui explique quelques écarts entre les erreurs systématiques significatives et lissées.

L'effet des gluons mous dans les régions sensibles aux instabilités infrarouges ($p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$) est également différent entre SHERPA et PYTHIA à cause des différences entre les algorithmes de cascade de partons utilisés par les deux générateurs et de l'implémentation d'un algorithme de fusion entre PS et éléments de matrice dans SHERPA.

De façon similaire au cas inclusif, la qualité de l'ajustement entre modèle et données baisse significativement avec l'usage des pdfs photons de PYTHIA lors de la décomposition dans les différents spectres. Par conséquent, l'évaluation de cette incertitude systématique est considérée comme conservatrice pour l'ensemble des sous-catégories considérées pour la mesure finale. Il a été décidé de garder cette incertitude très conservatrice afin de couvrir les résidus d'ajustement pas totalement compris dans le cas nominal déjà discutés dans la section 15.5. Ceux-ci sont suspectés de venir plutôt de la modélisation de l'empilement dans la simulation, comme cela sera discuté plus loin dans cette section.

Dans le paragraphe suivant, les incertitudes systématiques reliées aux corrections d'isolation de fuites sont décrites.

*Corrections d'isolation de fuites
(développement latéral des gerbes électromagnétiques)*

Tous les échantillons simulés de signal appliquent des corrections de fuites supplémentaires basées sur les données $\gamma + X$ pour $E_{T,1}^{\text{calo.iso.}}$ et $E_{T,2}^{\text{calo.iso.}}$ afin de tenir compte des différences de développements latéraux entre les gerbes électromagnétiques réelles et simulées. L'étude complète menant au calcul de ces corrections constitue le travail de performance principal effectué pendant les années de thèse et a été abordé en détail dans la section 10.4.

Les corrections consistent à appliquer des translations dans le tonneau et des translations et corrections de largeur dans les bouchons. Toutes les corrections sont calculées indépendamment dans des catégories fines en $|\eta^\gamma|/E^\gamma$ et fournies avec des barres d'erreur. La méthode du *template fit* 2D est alors exécutée à nouveau en faisant varier l'ensemble des corrections de translations puis l'ensemble des corrections de largeur de $+1\sigma$, -1σ . Des détails supplémentaires sur les barres d'erreurs des corrections sont disponibles dans la section 10.4 et l'appendice A. Les effets de ces corrections sur les résidus de l'ajustement dans l'analyse ont déjà été analysés dans la section 15.5.2.

Cette méthode d'estimation suppose les translations et les corrections de largeur non corrélées, ce qui est vérifié approximativement dans l'analyse de performance associée (voir section 10.4). Les corrections de largeur ayant dans tous les cas un impact négligeable sur les résultats du *template fit* 2D (voir section 15.6.5), cette hypothèse n'a aucun effet significatif sur l'incertitude finale. En revanche, la méthode d'estimation est considérée comme conservatrice dans le sens où elle fait varier l'ensemble des corrections d'un certain type quelque soit $|\eta^\gamma|/E^\gamma$, et non pas de façon individuelle dans chaque catégorie en $|\eta^\gamma|/E^\gamma$ considérée dans l'étude de performance. Cela revient donc à supposer les barres d'erreurs issues de l'étude de performance 100% corrélées alors qu'elles incluent une partie non-corrélée (erreur statistique) et une partie corrélée (différence avec PYTHIA symétrisée).

L'effet résultant constitue la seconde source d'incertitude la plus importante sur la soustraction du bruit de fond, avec un effet de l'ordre de 1.4% sur la pureté $\gamma\gamma$ pour la région **TITI** complète ($\simeq 1.9\%$ en termes de nombre d'événements), totalement dominé par les corrections de translation. Les distributions des erreurs selon $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc, seront montrées et commentées plus en détail dans la figure 126 située en fin de section.

*Rapport $\tilde{\mathbf{T}}/\mathbf{T}$ pour $\gamma\gamma$
(développement latéral des gerbes électromagnétiques)*

Certains rapports entre le nombre de photons $\tilde{\mathbf{T}}$ et $\mathbf{T}$ dans l'échantillon de signal simulé sont utilisés pour calculer les corrections de contamination dans les régions de contrôle (voir section 15.3.4). Il est donc nécessaire d'estimer l'incertitude sur ce rapport.

Les variables discriminantes utilisées pour l'identification des photons dans les échantillons de signal MC sont tout d'abord corrigées à l'aide des *fudge factors* fournis par le groupe de performance, qui permettent de faire correspondre les distributions MC avec les données (voir section 9). Ces corrections sont à nouveau rendues nécessaires par les différences entre le développement latéral des gerbes électromagnétiques de la simulation et des données réelles. Des mesures d'efficacité de la sélection d'identification TIGHT dans les données sont ensuite exploitées par le groupe de performance pour calculer des facteurs d'échelle (*ID scale factors*) corrigeant les différences résiduelles entre l'efficacité MC et celle mesurée dans les données avec une excellente précision (voir tableau 15). Les *ID scale factors* sont appliqués par défaut à tous les photons TIGHT dans les échantillons de signal.

Cependant, des mesures dans les données étant disponibles uniquement pour l'efficacité TIGHT et les sélections TIGHT et NON-TIGHT comportant une partie 100% corrélée (variables de forme de gerbe basées sur la seconde couche du calorimètre) et une partie 100% anti-corrélée (variables de forme de gerbe basées sur la première couche du calorimètre), l'évaluation de l'incertitude sur le rapport $\tilde{\mathbf{T}}/\mathbf{T}$ est délicate.

Dans la collaboration ATLAS, l'incertitude sur ce rapport pour les photons est souvent évaluée en désactivant les *fudge factors*, ce qui revient à considérer une incertitude individuelle sur TIGHT bien plus grande que celle donnée par les *ID scale factors* mais également à utiliser la corrélation entre $\mathbf{T}$ et $\tilde{\mathbf{T}}$ prédite par le MC, ce qui finalement donne une incertitude raisonnable (voir tableau 15).

Dans l'analyse $\gamma\gamma + X$, une approche plus rigoureuse et a priori très conservatrice développée pendant les années de thèse est adoptée à la place. Tout d'abord, le rapport entre les photons TIGHT et NON-TIGHT est exprimé en terme d'efficacité d'identification. L'échantillon de départ considéré est un échantillon di-photon passant toute la sélection au niveau reconstruit à l'exception de l'identification, qui est uniquement LOOSE-LOOSE (*ie.* après présélection). Les corrélations d'identification entre les deux photons sont négligées (voir section 15.2).

$$R = \frac{N_{\gamma}^{\tilde{\mathbf{T}}}}{N_{\gamma}^{\mathbf{T}}} = \frac{\epsilon_{L'} \cdot \epsilon_{\tilde{\mathbf{T}}|L'} \cdot N_{\gamma}}{\epsilon_{L'} \cdot \epsilon_{T|L'} \cdot N_{\gamma}} = \frac{1 - \epsilon_{T|L'}}{\epsilon_{T|L'}} = \frac{1}{\epsilon_{T|L'}} - 1 \quad \text{avec} \quad \epsilon_{\tilde{\mathbf{T}}|L'} = 1 - \epsilon_{T|L'}, \quad (70)$$

où $\epsilon_{T|L'}$ est l'efficacité d'identification TIGHT pour les photons passant LOOSE' (probabilité conditionnelle). L'incertitude sur ce rapport peut ensuite être calculée en utilisant une différentielle :

$$\Delta R = \frac{\Delta \epsilon_{T|L'}}{\epsilon_{T|L'}^2}. \quad (71)$$

L'unique quantité mesurée dans les données est $\epsilon_T = \epsilon_{L'} \cdot \epsilon_{T|L'}$. Son incertitude est connue grâce aux incertitudes sur les *ID scale factors*. Les coupures entre la première ($T|L'$) et la deuxième couche (L') du calorimètre étant non-corrélées ou corrélées positivement, celle-ci s'exprime de la façon suivante :⁷¹

$$\left(\frac{\Delta \epsilon_T}{\epsilon_T} \right)^2 = \left(\frac{\Delta \epsilon_{L'}}{\epsilon_{L'}} \right)^2 + \left(\frac{\Delta \epsilon_{T|L'}}{\epsilon_{T|L'}} \right)^2 + 2\rho \frac{\Delta \epsilon_{L'} \Delta \epsilon_{T|L'}}{\epsilon_{L'} \epsilon_{T|L'}} \quad \text{avec } \rho \geq 0, \quad (72)$$

soit :

$$\Delta \epsilon_{T|L'} = \epsilon_{T|L'} \cdot \sqrt{\left(\frac{\Delta \epsilon_T}{\epsilon_T} \right)^2 - \left(\frac{\Delta \epsilon_{L'}}{\epsilon_{L'}} \right)^2 - 2\rho \frac{\Delta \epsilon_{L'} \Delta \epsilon_{T|L'}}{\epsilon_{L'} \epsilon_{T|L'}}} < \epsilon_{T|L'} \cdot \frac{\Delta \epsilon_T}{\epsilon_T}. \quad (73)$$

En réinjectant le résultat de l'équation 73 dans l'équation 71, on obtient finalement :

$$\Delta R < \frac{\Delta \epsilon_T}{\epsilon_T \cdot \epsilon_{T|L'}}. \quad (74)$$

Ainsi, l'incertitude sur les rapports R_1 , R_2 (premier et deuxième photon) peut être estimée de façon conservatrice en appliquant une variation correspondant à $\pm \frac{\Delta \epsilon_{T,i}}{\epsilon_{T,i} \cdot \epsilon_{T|L',i}}$. R_1 et R_2 sont employés pour le calcul des corrections de contamination $\gamma\gamma$ dans $\tilde{\mathbf{TITI}}$ et $\mathbf{TITI}$, respectivement. Pour les contaminations $\gamma\gamma$ dans $\tilde{\mathbf{TITI}}$ ($\Delta R_{1,2}$), les identifications des deux photons ont besoin d'être variées de manière synchrone pour rester conservateur (corrélation à 100% des incertitudes), ce qui mène à la formule suivante :

$$\frac{\Delta R_{1,2}}{R_{1,2}} = \sqrt{\left(\frac{\Delta R_1}{R_1} \right)^2 + \left(\frac{\Delta R_2}{R_2} \right)^2 + 2 \cdot \frac{\Delta R_1 \cdot \Delta R_2}{R_1 \cdot R_2}} \quad (75)$$

Il est alors possible d'utiliser l'équation 71 pour obtenir une inégalité similaire à celle obtenue dans l'équation 74.

Les termes $\Delta \epsilon_{T,i}$ et $\epsilon_{T,i}$ sont évalués avec l'échantillon de signal simulé. $\Delta \epsilon_{T,i}$ est estimé en faisant varier les *ID scale factors*, qui ont des incertitudes supposées 100% corrélées entre le premier et le deuxième photon, tandis que les $\epsilon_{T,i}$ sont calculés à partir des variations d'événements MC reportés dans le tableau 8 (voir section 15.3.4) et de l'équation 70. Les variations obtenues sont rapportées dans le tableau 15.

⁷¹ $\epsilon_{L'}$ et $\epsilon_{T|L'}$ sont non-corrélés ou corrélés positivement pour les photons car $\epsilon_{L'}$ correspond principalement à des coupures sur des variables de la seconde couche tandis que $\epsilon_{T|L'}$ correspond à des coupures sur des variables de la première couche uniquement.

Coefficient	SF + 1 σ	SF - 1 σ	<i>Fudge factors</i> désactivés	Approche conservatrice
$n_{\gamma\gamma,MC}^{\tilde{\mathbf{TITI}}}/n_{\gamma\gamma,MC}^{\mathbf{TITI}} (\Delta R_1)$	-0.59%	+0.59%	$\pm 14.0\%$	$\pm 24.9\%$
$n_{\gamma\gamma,MC}^{\tilde{\mathbf{TITI}}}/n_{\gamma\gamma,MC}^{\mathbf{TITI}} (\Delta R_2)$	-0.88%	+0.90%	$\pm 29.0\%$	$\pm 23.4\%$
$n_{\gamma\gamma,MC}^{\tilde{\mathbf{TITI}}}/n_{\gamma\gamma,MC}^{\mathbf{TITI}} (\Delta R_{12})$	-1.53%	+1.56%	$\pm 39.0\%$	$\pm 48.3\%$

TAB. 15 – Variations relatives des rapports $\tilde{\mathbf{T}}/\mathbf{T}$ employées pour le calcul des corrections de contamination dans les régions de contrôle utiles à la décomposition de la région **TITI** complète (facteurs α_i et α'_i , voir section 15.3.4). Les variations systématiques correspondant à l'incertitude sur les facteurs d'échelle pour TIGHT (*SF*), la désactivation des *fudge factors* et les variations conservatrices finales adoptées pour l'analyse sont montrées séparément.

Les valeurs issues de la méthode conservatrice sont probablement très surestimées car elles supposent les incertitudes sur les *ID scale factors* 100% anti-corrélées avec NON-TIGHT. En comparant les variations avec celles obtenues lors de la désactivation des *fudge factors*, un ordre de grandeur similaire est observé, sauf pour ΔR_1 où la méthode des *fudge factors* pourrait sous-estimer l'incertitude. Par conséquent, l'incertitude conservatrice est adoptée dans l'analyse $\gamma\gamma + X$. L'effet résultant est de l'ordre de 0.9% sur la pureté $\gamma\gamma$ pour la région **TITI** complète, pas totalement négligeable comparé à l'incertitude systématique finale ($\simeq 2.6\%$). Les distributions des erreurs selon $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc, seront montrées et commentées plus en détail dans la figure 128 située en fin de section. L'approche conservatrice développée pendant les années de thèse est actuellement en cours de discussion au sein du groupe de performance e/γ d'ATLAS.

Une idée pour améliorer cette évaluation dans le futur pourrait être de mesurer la corrélation $\tilde{\mathbf{T}}/\mathbf{T}$ directement dans les données avec un échantillon de photons purs dès le critère LOOSE', par exemple issu des désintégrations radiatives $Z \rightarrow l^+l^-\gamma$.

D'autre part, les rapports servant aux corrections des contaminations γ +jet et jet+ γ dans $\tilde{\mathbf{TITI}}$ (appliqués dans le cas inclusif uniquement) font intervenir le rapport $\tilde{\mathbf{T}}/\mathbf{T}$ du MC à la fois pour les photons et les jets mal identifiés (voir section 15.3.4). Le rapport MC étant dans ce cas de toute façon très incertain à cause de la statistique limitée de l'échantillon simulé, les effets systématiques sur ces rapports MC sont négligés. Dans tous les cas, ces corrections jouent uniquement sur les proportions des trois bruits de fond hadroniques dans le cas inclusif (γ +jet, jet+ γ , di-jet) et n'ont aucun impact sur $\gamma\gamma$.

*Statistique limitée de l'échantillon de signal simulé
(statistique MC, décomposition spectrale uniquement)*

L'échantillon de signal présente globalement une excellente statistique, de l'ordre de trois à quatre fois celle des données pour la région **TITI** complète (voir section 14.6). Par conséquent, l'effet de la statistique MC sur les résultats du *template fit* est négligée dans le cas inclusif. Cependant, lorsque l'échantillon **TITI** est séparé en plusieurs sous-catégories pour la décomposition selon les différents spectres, la statistique MC est plus faible car les pdfs photons sont extraites indépendamment dans chaque région. Le MC ne reproduisant pas entièrement les données (voir section 17), certaines catégories ont un rapport statistique MC/données défavorable, le cas extrême étant la dernière catégorie

en $m_{\gamma\gamma}$ où le rapport statistique MC/données est de l'ordre de 0.5. Il devient alors nécessaire d'évaluer l'effet de la statistique du MC sur les résultats de soustraction.⁷²

Pour cela, une procédure de type *bootstrap* est à nouveau appliquée (voir le paragraphe traitant du *choix du générateur MC*), mais cette fois uniquement avec l'échantillon de signal simulé :

- ◇ 1 000 répliques *bootstrap* sont générées de manière indépendante pour chaque région de signal dans l'échantillon simulé. En pratique, 1 000 poids poissoniens indépendants centrés en 1 sont assignés pour chaque événement dans **TITI**.
- ◇ 1 000 jeux de pdfs photons sont extraits (1 par réplique) et la procédure du *template fit* 2D est exécutée à nouveau pour chacun d'entre eux sur les données de collisions nominales.
- ◇ La distribution des résultats pour chaque sous-catégorie (définie en termes de $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc) est tracée et l'incertitude finale est prise comme la somme en quadrature entre la RMS et la différence entre la moyenne de la distribution et le résultat nominal (*bootstrap non-closure*). Cette dernière est la plupart du temps très négligeable, sauf pour certaines queues de distribution où l'écart au modèle poissonien engendre de légers *non-closure*.

Les incertitudes systématiques résultantes sont de l'ordre de 1 à 2% en termes de nombre d'événements $\gamma\gamma$, avec une tendance à augmenter dans les queues de distribution (jusqu'à 5% pour la catégorie la plus haute en $m_{\gamma\gamma}$). Cependant celles-ci restent dans tous les cas sous-dominantes par rapport à l'effet dû à la statistique limitée des données dans **TITI**, qui correspond alors simplement à l'erreur statistique renvoyée par la procédure d'ajustement (voir section 15.6.5) et est par exemple de l'ordre de 18% pour la catégorie la plus haute en $m_{\gamma\gamma}$. Les distributions des erreurs selon $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc, seront montrées et commentées plus en détail dans la figure 126 située en fin de section.

*Fusion des régions **TITI** dans l'échantillon de signal simulé
(dépendances cinématiques des pdfs d'isolation, décomposition spectrale uniquement)*

Finalement, une incertitude supplémentaire spécifique à la modélisation des photons vient de la fusion de certaines régions **TITI** dans les échantillons simulés pour l'extraction des pdfs lorsque la statistique du MC devient trop faible, qui sont alors employées dans les catégories **TITI** des données non fusionnées.

Cette technique est utilisée uniquement dans les queues de distribution lors de la décomposition du bruit de fond selon les différentes observables. Elle permet de baisser l'incertitude totale liée à la statistique MC en faisant un compromis entre fluctuations statistiques des pdfs dues à la statistique limitée de l'échantillon MC (voir paragraphe précédent) et biais cinématique de la modélisation dû au non-respect de la définition de la sous-catégorie (traité ici). Par exemple, pour la dernière catégorie de $m_{\gamma\gamma}$ (700-

⁷²Afin de réduire cet effet, une nouvelle production MC dédiée pour chaque queue de distribution en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc, aurait été nécessaire. La statistique limitée du MC n'étant jamais un facteur limitant la précision de la mesure, il a été décidé de se satisfaire de l'échantillon simulé de base.

1700 GeV), les régions des trois dernières catégories sont fusionnées dans l'échantillon MC de telle sorte que le jeu de pdfs photons nominal est extrait de la région 500-1700 GeV et pourtant ajusté dans 700-1700 GeV. Cela permet de limiter l'effet des fluctuations statistiques du MC à 5% malgré un rapport statistique MC/données d'environ 0.5 dans la catégorie 700-1700 GeV (voir paragraphe précédent).

Afin d'estimer le biais cinématique résultant de cette hypothèse de régularité supplémentaire sur les pdfs des photons, la méthode du *template fit* 2D est exécutée à nouveau en fusionnant la catégorie d'intérêt dans le MC uniquement avec la ou les catégories situées sur sa gauche ou sur sa droite. Dans le cas cité précédemment, cela revient à évaluer l'incertitude systématique en utilisant le jeu de pdfs issu de 500-700 GeV uniquement, puisqu'il n'y a aucune catégorie plus à droite.

L'effet résultant est montré dans la figure 122 pour toutes les catégories où des fusions des régions **TITI** dans l'échantillon de signal simulés sont effectuées, soit principalement celles situées dans les queues de distribution de $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T , ϕ_η^* et $|\cos\theta_\eta^*|$. Les variations vont de 1% à 7% et sont maximales pour $m_{\gamma\gamma}$ (environ +7% pour 700-1700 GeV). En effet, comme cela a déjà été souligné dans la section 15.5.2, de plus en plus de paires sont émises dans les bouchons à haut $m_{\gamma\gamma}$, or les paires bouchon-bouchon mènent à des distributions d'isolation significativement différentes des parties plus centrales du détecteur (voir section 14.3.2).

Cette systématique a finalement peu d'impact sur les résultats finals, les queues de distribution où les fusions sont nécessaires étant totalement dominées par la statistique limitée de l'échantillon **TITI** des données, qui correspond alors simplement à l'erreur statistique renvoyée par la procédure d'ajustement (voir section 15.6.5). Par ailleurs, son ajout en quadrature est considéré comme conservateur car l'évaluation de cet effet souffre de la statistique limitée du MC et des données.

Dans le paragraphe suivant, l'effet systématique relié à la modélisation de l'empilement dans les échantillons simulés est discuté.

Modélisation de l'empilement dans la simulation (déformation des pdfs d'isolation à haut taux d'empilement)

Bien que l'efficacité d'isolation soit approximativement indépendante des conditions d'empilement grâce aux corrections basées sur la médiane de la densité d'énergie ambiante ρ_{UE} (voir section 10.3.2), la forme des pdfs d'isolation dépend de façon significative de μ (nombre moyen d'interactions par croisement de faisceaux, voir section 4.1). En effet, particulièrement à bas E_T^γ (corrections de fuites faibles), la largeur des pdfs d'isolation augmente avec μ à cause des fluctuations croissantes autour de la médiane de ρ_{UE} , comme cela est illustré dans la figure 123.

Les corrections d'isolation dues à l'empilement et à l'événement sous-jacent sont calculées séparément dans les données et dans la simulation afin d'éviter toute extrapolation. Cependant, il est important de veiller à ce que les distributions des densités d'énergie ambiante complètes soient similaires entre données et simulation afin de limiter les différences résiduelles dues aux fluctuations autour de la médiane de ρ_{UE} , soustraite par défaut de $E_T^{\text{calo.iso.}}$. Il est très intéressant de remarquer que μ a un effet particulièrement significatif sur la pente montante de l'isolation, qui est justement l'emplacement

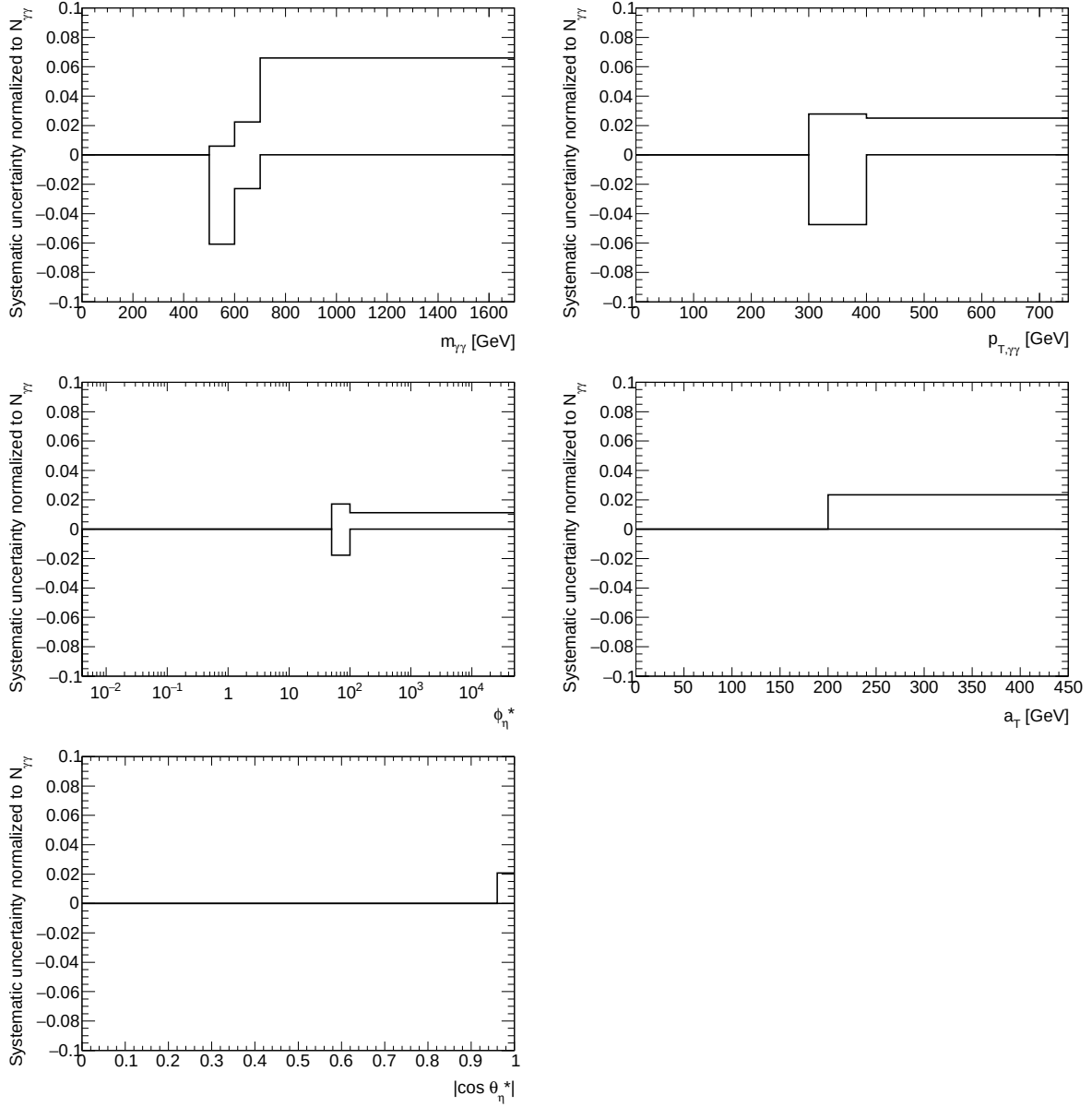


FIG. 122 – Variations des résultats de la méthode du *template fit* 2D en termes de nombre d'événements $\gamma\gamma$ dans **TITI** lors de l'évaluation de l'effet de la fusion des régions **TITI** dans l'échantillon de signal simulé. Le biais cinématique résultant sur les pdfs photons est évalué en fusionnant uniquement les régions situées à gauche, puis à droite de la catégorie considérée. Aucune fusion n'est nécessaire pour $\Delta\phi_{\gamma\gamma}$.

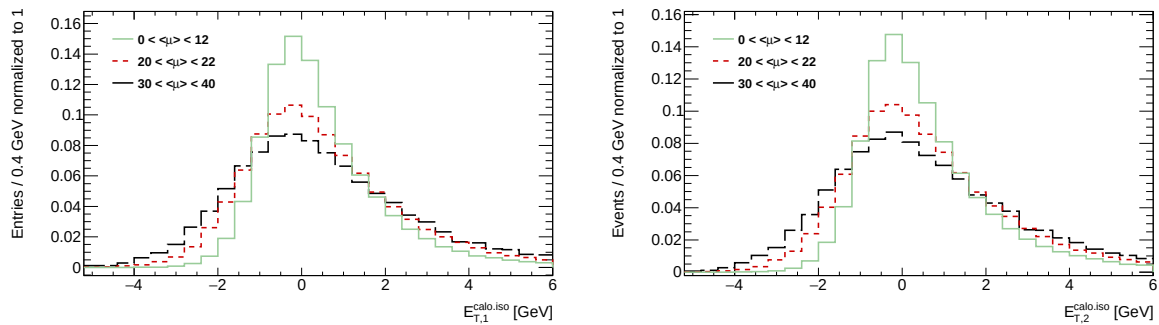


FIG. 123 – Pdfs d’isolation du premier (gauche) et du deuxième (droite) photon évaluées à partir de l’échantillon de signal simulé SHERPA dans trois différentes catégories d’empilement μ : 0-12 (trait plein vert), 20-22 (trait pointillé rouge) et 30-40 (trait pointillé noir). Aucune sélection supplémentaire que **TITI** n’est faite sur $E_{T,1}$, $E_{T,2}$, de sorte que la région est dominée par des photons ayant de faibles impulsions transverses.

des résidus les plus manifestes entre modèle d’ajustement et données (voir section 15.5). Étant donné la forme des résidus (voir figure 106), ceux-ci pourraient indiquer des fluctuations dues à l’empilement légèrement trop faibles dans la simulation.

Dans la collaboration ATLAS, la description de l’empilement par la simulation est optimisée uniquement vis-à-vis d’observables issues du détecteur interne telles que le nombre de vertex reconstruits N_{PV} (voir section 13.2.1). Le groupe de performance du détecteur interne (*tracking performance group*) recommande ainsi de repondérer les échantillons simulés à $\frac{1}{1.09} \cdot \mu_{\text{data}}$ au lieu de μ_{data} afin que la simulation reproduise la distribution en N_{PV} des données. L’incertitude sur ce facteur d’échelle est évaluée à ± 0.04 .

Le *template fit 2D* a été reproduit avec des pdfs d’isolation de photons extraits à partir de l’échantillon SHERPA repondéré à $\frac{1}{1.13} \cdot \mu_{\text{data}}$ et $\frac{1}{1.05} \cdot \mu_{\text{data}}$ afin d’estimer l’impact sur la décomposition du bruit de fond. Les résultats sont présentés dans le tableau 16. Les variations en termes de pureté $\gamma\gamma$ sont de l’ordre de $\pm 0.5\%$ ($\simeq 0.7\%$ en termes de nombres d’événements), largement sous-dominantes par rapport à la systématique reliée au choix du générateur MC ($\simeq 1.7\%$ en termes de pureté $\gamma\gamma$). Dans le cas des décompositions de l’échantillon **TITI** selon les observables de mesures, l’effet est constant à $\simeq 0.5\%$, sauf dans les régions dominées par les hauts E_T^γ où il tombe à 0. En effet, la largeur des pdfs dans les régions à haut E_T^γ est alors dominée par les fluctuations des corrections de fuites, qui ont déjà été corrigées avec les données pendant les années de thèse (voir section 10.4). Il a par ailleurs déjà été observé lors du calcul de ces corrections que les effets de largeurs de $E_T^{\text{calo.iso}}$ sur la pureté sont faibles en général et qu’ils nécessitent une bonne statistique dans les données afin de pouvoir être corrigés efficacement.

Étant donnés les différents ordres de grandeur, l’incertitude résultant de la repondération en μ des échantillons MC est négligée dans la suite. Ce choix est également motivé par le fait que l’incertitude reliée au choix du générateur MC est déjà intentionnellement conservatrice afin de couvrir les résidus les plus manifestes entre modèle d’ajustement et données (voir section 15.5).

À titre illustratif, la qualité du modèle d’ajustement utilisant les repondérations

Processus	Nombre d'événements finals (cas nominal)	Variations pour $SF_\mu = 1.05$	Variations pour $SF_\mu = 1.13$
$\gamma\gamma$	$235\,392 \pm 1\,022$ (stat) $^{+8\,070}_{-8\,672}$ (syst)	+1 527	-1 872
γ +jet	$45\,465 \pm 672$ (stat) $^{+8\,470}_{-8\,780}$ (syst)	-750	+711
jet+ γ	$18\,699 \pm 590$ (stat) $^{+4\,391}_{-4\,693}$ (syst)	-1 359	+1 272
di-jet	$4\,873 \pm 352$ (stat) $^{+2\,096}_{-1\,461}$ (syst)	-199	+191
di-electron	$8\,275 \pm 680$ (stat) $^{+2\,777}_{-1\,396}$ (syst)	+767	-301

TAB. 16 – Effet de la variation par $\pm 1\sigma$ du facteur d'échelle utilisé pour la repondération des échantillons MC au nombre moyen d'interactions par croisement de faisceaux (1.09 ± 0.04).

alternatives en μ est évaluée dans les figures 124 et 125. Malgré l'effet négligeable sur $\gamma\gamma$, une variation importante des résidus d'ajustement est observée uniquement dans la région en légère tension (0 GeV-0 GeV et énergie d'isolation négative en général), ce qui tend à confirmer le caractère conservateur de l'incertitude évaluée avec le générateur PYTHIA8 et donne de bonnes perspectives d'amélioration pour l'analyse suivante. En effet, il est possible d'imaginer des recommandations différentes de repondérations des échantillons MC pour les analyses utilisant avant tout les mesures du calorimètre électromagnétique (électrons/photons), basées par exemple sur une bonne description de la RMS de la densité d'énergie ambiante dans les données. Cela permettrait également de tenir compte de l'effet de l'empilement hors-temps sur le calorimètre, qui est un détecteur plus lent et donc plus sensible à ce type d'effet que le détecteur interne.

Je n'ai malheureusement pas eu le temps de mettre en place une méthodologie permettant une repondération en μ optimisée pour le calorimètre électromagnétique dans ATLAS, bien que je sois convaincu qu'elle permettrait de réduire significativement les résidus observés entre données et modèle d'ajustement dans la méthode du *template fit* 2D. Celle-ci pourrait ensuite permettre éventuellement de revisiter l'évaluation de la systématique due au choix du générateur MC, probablement très conservatrice à cause de l'approche simpliste de PYTHIA8 pour la modélisation du signal.

Les résidus d'ajustement baissent pour $SF_\mu = 1.05$ (-1σ), confirmant une amélioration de la qualité de l'ajustement lors d'une augmentation des fluctuations d'empilement. Dans le même temps, la composante ee augmente de +767 événements, se rapprochant ainsi de la valeur prédite par l'échantillon DY simulé avec facteurs correctifs. Bien que ces observations puissent très bien s'avérer être de simples coïncidences, elles constituent des éléments supplémentaires encourageant les recherches évoquées plus haut. Une ébauche d'étude avec quelques figures supplémentaires est disponible dans l'appendice C.

Dans le paragraphe suivant, un bilan est donné quant aux incertitudes spécifiques à la modélisation des photons dans le *template fit* 2D.

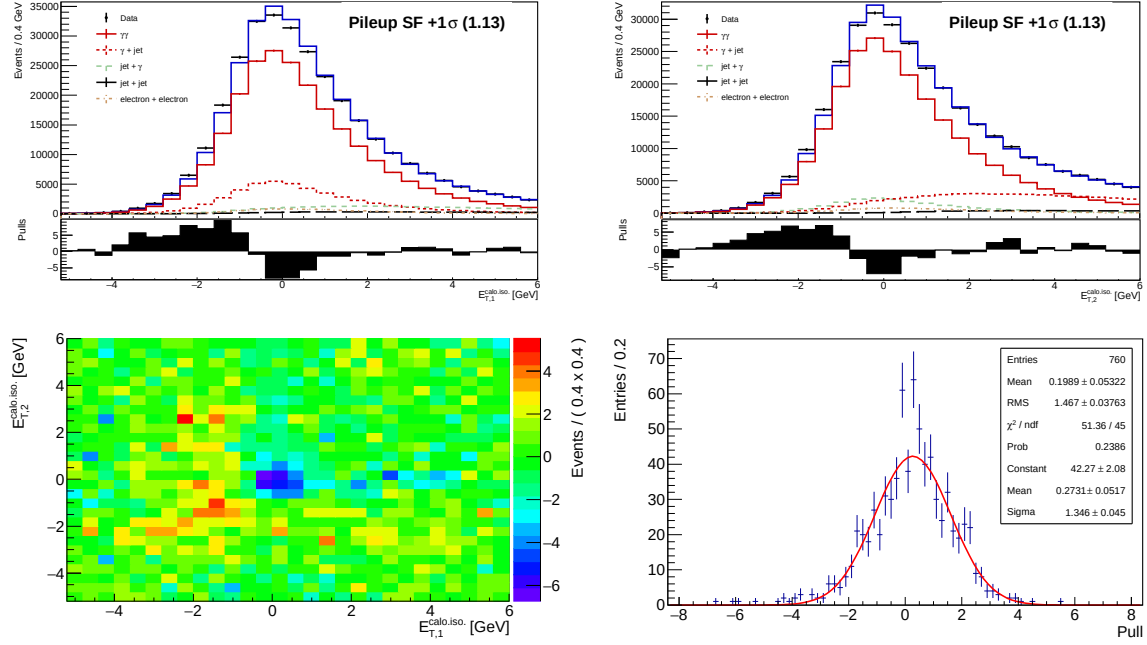


FIG. 124 – Résultats du *template fit* 2D dans la région **TITI** complète en repondérant l'échantillon de signal simulé à $\frac{1}{1.13} \cdot \mu_{\text{data}}$. Les figures représentées sont similaires à celles de la figures 120.

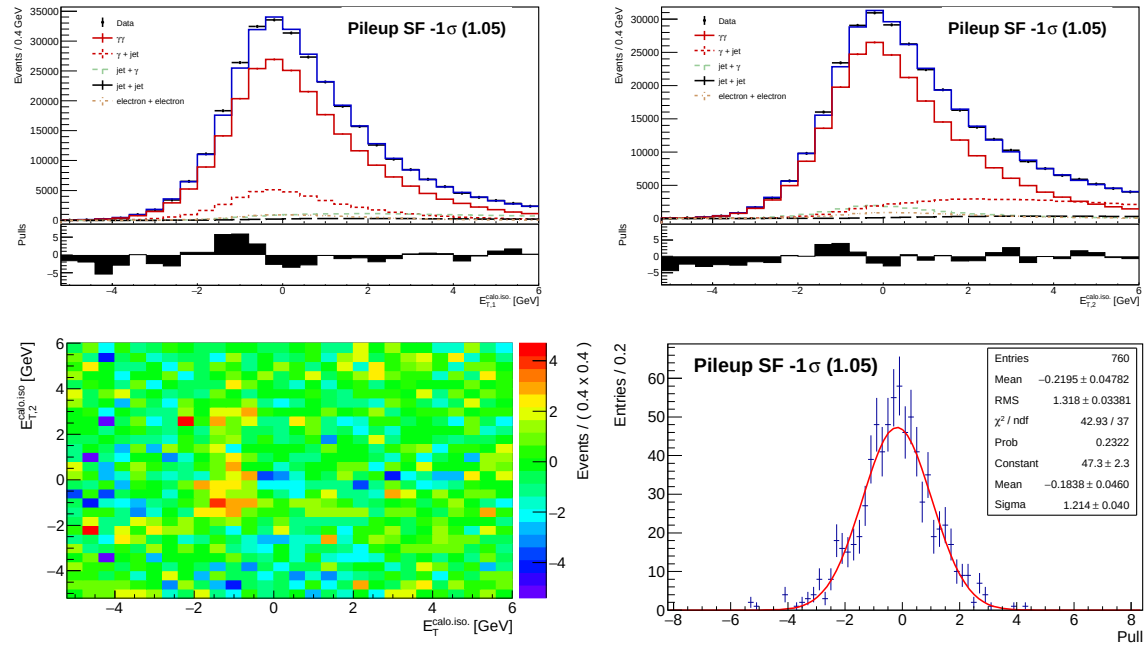


FIG. 125 – Figure similaire à la figure 124 en utilisant un échantillon de signal simulé repondéré à $\frac{1}{1.05} \cdot \mu_{\text{data}}$.

Bilan

Le résumé des différentes incertitudes spécifiques à la modélisation des photons retenues dans l'analyse dans le cas de la région **TITI** complète est présenté dans le tableau 17. Les distributions des incertitudes associées lors des décompositions en fonction des observables de mesures sont représentées dans la figure 126.

Processus	PYTHIA8 (symétrisé)	Isolation $^{+1}_{-1} \sigma$ (translation)	Isolation $^{+1}_{-1} \sigma$ (largeur)	Fuites $^{+1}_{-1} \sigma$ (syst., MC stat.)	Total
$\gamma\gamma$	+5 189	+4 493 -4 535	+53 +23	+2 187 -2 159 (± 38)	+7 204 -7 222
γ +jet	-8 037	-1 534 +1 418	-44 -135	-1 430 +1 389 (± 172)	+8 280 -8 309
jet+ γ	+3 330	-1 840 +1 737	-46 +21	-924 +895 ($\pm 1 399$)	+4 107 -4 158
di-jet	+32	+63 -45	+8 +71	+150 -141 ($\pm 1 409$)	+1 420 -1 417
ee	-509	-1 175 +1 404	+26 +19	—	+1 494 -1 281

TAB. 17 – Variation de chaque composante du modèle d'ajustement par rapport aux valeurs nominales citées dans la section 15.5.1 en fonction des différentes incertitudes spécifiques à la modélisation des photons (cas inclusif) : utilisation de pdfs d'isolation issues de PYTHIA8 (valeur symétrisée), corrections d'isolation de fuite supplémentaires variées dans leur barres d'erreur (corrections de translation et de largeur séparées) et rapport $\tilde{\mathbf{T}}/\mathbf{T}$ dans le MC.

Sans surprise, la source d'incertitude dominante concernant la modélisation des photons pour la composante $\gamma\gamma$ est liée au choix du générateur MC (ligne pleine noire) dans presque la totalité de l'espace des phases. Cependant, dans les queues de distribution ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, $|\cos \theta_\eta^*|$), les autres incertitudes jouent également un rôle important, en particulier celles liées à la statistique de l'échantillon de signal simulé (lignes pointillées rouge et bleu). Cependant, la précision des mesures dans ces régions est limitée par la statistique des données et non les effets systématiques, comme nous le verrons dans la section 15.6.5. Les corrections d'isolation principales (*photon isolation shift corrections*) ainsi que l'effet du rapport $\tilde{\mathbf{T}}/\mathbf{T}$ dans le MC de signal (voir figure 128) contribuent également dans certaines zones où PYTHIA et SHERPA mènent à des résultats similaires. En revanche, les corrections d'isolation de largeur (*photon isolation smearing corrections*) ont toujours un effet négligeable sur les résultats du *template fit* 2D.

Comme déjà mentionné au cours de la section, l'incertitude liée au choix du générateur MC est considérée comme très conservatrice à cause de la modélisation simpliste du signal par le générateur PYTHIA8. Néanmoins, il a été décidé de l'évaluer ainsi afin de couvrir l'effet des résidus mal compris observés entre données et modèle d'ajustement (voir section 15.5). Les résidus sont suspectés de provenir en grande partie de la description imparfaite de la simulation concernant les fluctuations d'isolation calorimétrique dues à l'empilement, celle-ci étant optimisée uniquement à partir d'observables du détecteur interne dans ATLAS (voir paragraphe précédent), en particulier peu sensibles à l'empilement hors-temps. Des études complémentaires sur cet effet sont disponibles dans

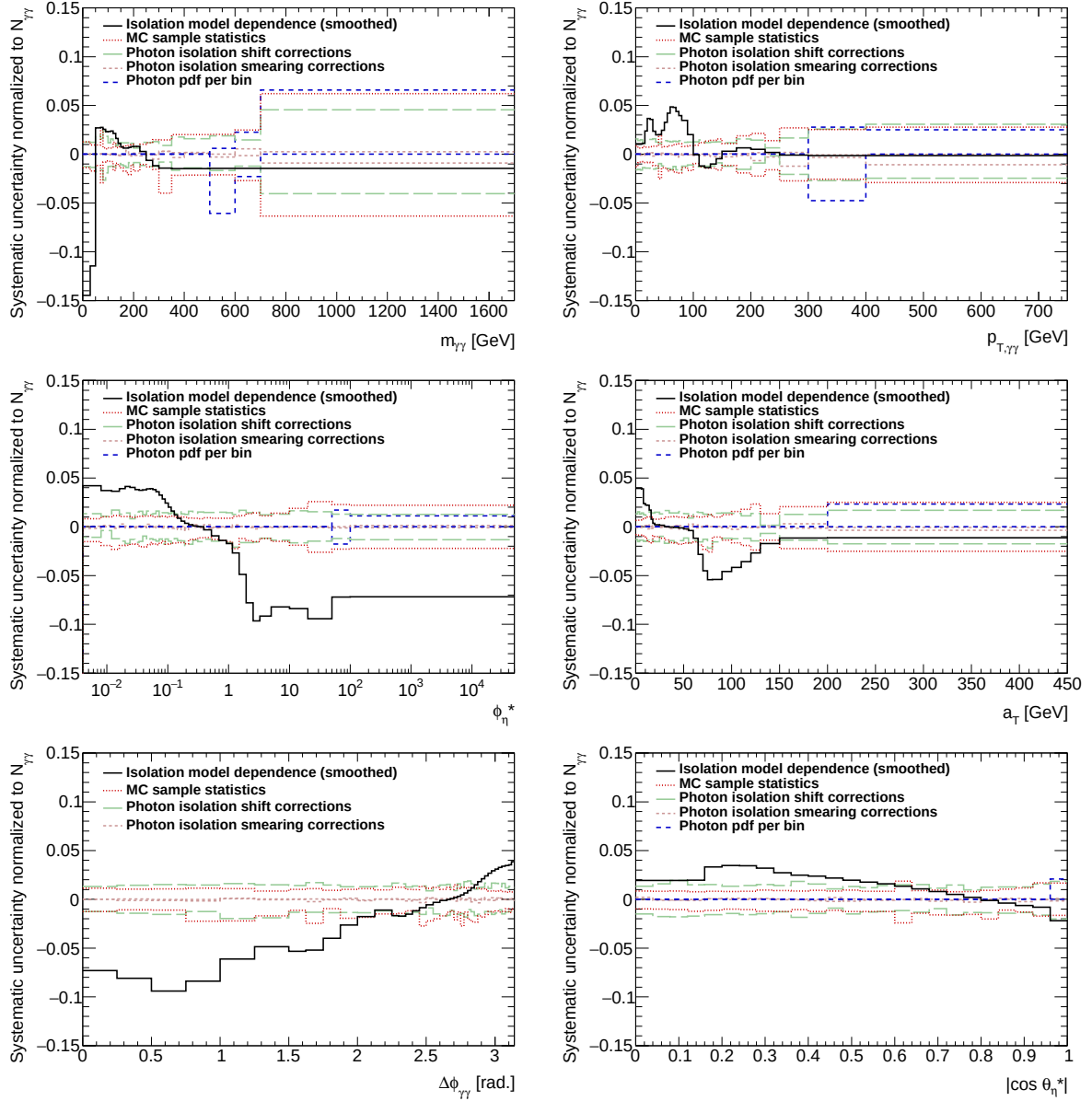


FIG. 126 – Variations relatives du nombre d'événements $\gamma\gamma$ associé aux différentes incertitudes spécifiques à la modélisation des photons : utilisation de pdfs d'isolation issues de PYTHIA8 (*isolation model dependence*), corrections d'isolation de fuite supplémentaires variées dans leur barres d'erreur (*photon isolation shift/smearing corrections*), statistique limitée du MC de signal (*MC sample statistics*) et fusion des régions **TITI** dans l'échantillon de signal (*photon pdf per bin*). L'incertitude liée au rapport $\hat{\mathbf{T}}/\mathbf{T}$ dans le MC est incluse dans la figure 128 afin d'éviter la surcharge de celle-ci.

l'appendice C mais ne permettent pas encore de fournir des recommandations générales améliorant la simulation de l'empilement pour les analyses basées sur le calorimètre électromagnétique.

La seule approche possible aujourd'hui pour estimer cet effet consiste à faire varier le facteur d'échelle actuellement recommandé pour la repondération en μ des échantillons simulés dans sa barre d'erreur (1.09 ± 0.04). L'effet résultant a été détaillé dans le paragraphe précédent et est négligeable devant les autres erreurs systématiques pour l'ensemble de l'espace des phases, validant leur robustesse.

L'effet venant de l'identification des photons pourrait également probablement être réduit par des mesures du rapport $\tilde{\mathbf{T}}/\mathbf{T}$ dans les données $Z^0 \rightarrow l^+l^-\gamma$. Malheureusement, les mesures réalisées avec les données par le groupe de performance concernent uniquement la sélection TIGHT pour le moment. Des discussions avec le groupe de performance sont en cours à ce sujet. Finalement, quelques réflexions personnelles sur le traitement des photons dans ATLAS et quelques possibilités d'évolution qui pourrait entre autres permettre l'amélioration de l'isolation calorimétrique ont été évoquées dans la section 11 (soustraction de *clusters* topologiques au lieu de *clusters* de taille fixe, etc).

Dans la section 15.6.2, les sources d'incertitude de la procédure de soustraction du fond reliées à la modélisation des jets sont examinées en détail.

15.6.2 Incertitudes spécifiques à la modélisation des jets

*Définition des régions de contrôle de jets
(corrélations résiduelles entre identification et isolation)*

L'incertitude systématique dominante reliée à la modélisation des jets vient de la définition arbitraire des différentes régions de contrôle utilisées pour l'extraction des pdfs d'isolation associées, toutes basées sur la sélection d'identification NON-TIGHT (voir sections 9.4 et 14.1).

En effet, la sélection NON-TIGHT utilisée par défaut dans le *template fit* 2D est issue de la sélection LOOSE'4, qui relâche 4 coupures sur les variables de formes de gerbes basées sur la première couche du calorimètre par rapport à TIGHT (voir section 9.4, tableau 3). Celle-ci permet une bonne statistique dans les régions de contrôle tout en limitant le biais sur les pdfs d'après les échantillons MC γ +jet simulés (voir figure 127). Cependant, les échantillons γ +jet simulés présentent une statistique très limitée et sont a priori peu fiable pour le taux de mauvaise identification $\gamma \rightarrow jet$. Le nombre de coupures relâchées apparaît alors comme un paramètre semi-arbitraire source d'incertitude.

L'effet systématique résultant de la déformation des pdfs d'isolation entre les régions TIGHT et NON-TIGHT est évalué en répétant la soustraction du bruit fond avec des pdfs jets basées sur les définitions alternatives LOOSE'2, LOOSE'3 et LOOSE'5. Dans tous les cas, les modèles basés sur LOOSE'2 et LOOSE'3 (LOOSE'5) mènent à un nombre d'événements $\gamma\gamma$ plus faible (plus élevé) que les résultats nominaux, avec de plus $W_{\gamma\gamma}^{\text{TITI}, \text{L}'3} > W_{\gamma\gamma}^{\text{TITI}, \text{L}'2}$. Ces résultats sont attendus dans le sens où plus le nombre de coupures relâchées est important, plus le biais sur l'isolation devrait être important. Ainsi, les jets LOOSE'5 sont moins bien isolés que les jets LOOSE'4, ce qui engendre une augmentation de la pureté $\gamma\gamma$ dans la région de signal.

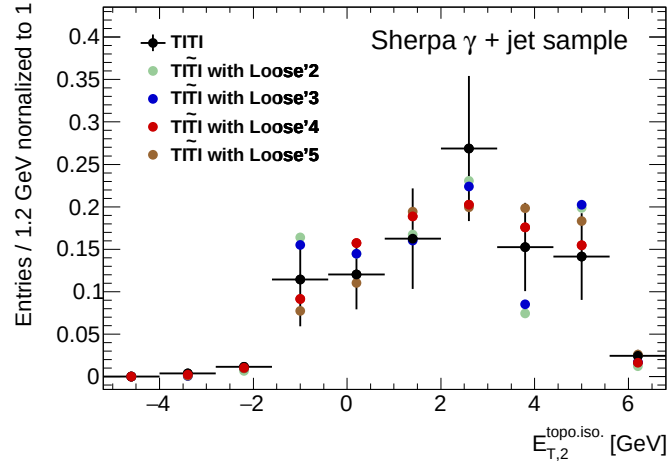


FIG. 127 – Comparaison entre les pdfs d’isolation des jets issues du processus γ +jet dans la région **TITI** et **TITI** en fonction de la définition LOOSE’ utilisée. L’échantillon simulé photon-jet SHERPA, bien que présentant une statistique très limitée, est utilisé faute de mieux. La figure correspondante pour jet+ γ (qui peut être produit à partir du même échantillon en demandant $E_T^{\text{fake}} > E_T^\gamma$) n’est pas représentée à cause de la statistique insuffisante du MC. La distribution associée à LOOSE’4 (points rouges) apparaît comme celle présentant le moins de fluctuations statistiques tout en ayant une queue d’isolation plus fidèle à **TITI** par rapport à LOOSE’5 (points bruns).

Fuites de signal	Coefficient	LOOSE’2	LOOSE’3	LOOSE’4	LOOSE’5
$\gamma\gamma$ dans TITI	α_1	0.162	0.191	0.157	0.146
$\gamma\gamma$ dans TITI	α_2	0.157	0.170	0.118	0.110
$\gamma\gamma$ dans TITI	α'_3	0.030	0.035	0.019	0.016

TAB. 18 – Fraction d’événements $\gamma\gamma$ dans les régions de contrôle de jets **TITI**, **TITI** et **TITI** calculées selon les équations 54 à 58 (voir section 15.3.4). L’incertitude statistique totale sur chacun de ces chiffres est de l’ordre de 0.001 en valeur absolue. Les incertitudes totales sont dominées par les effets systématiques.

L’incertitude systématique finale considérée est asymétrique et prise comme la différence entre les résultats nominaux et ceux issus des modèles LOOSE’2 (variation négative la plus élevée) et LOOSE’5 (variation positive) faute d’une meilleure méthode d’évaluation. Dans la région **TITI** complète, elle représente un effet de +1.1%/-1.3% en termes de pureté $\gamma\gamma$ (soit +1.5%/-1.7% en termes de nombre d’événements). Les distributions des erreurs seront montrées et commentées plus en détail dans la figure 128 située en fin de la section.

Par ailleurs, les contaminations des différentes régions de contrôle de jets par les événements $\gamma\gamma$ évaluées pendant le *template fit* 2D sont données dans le tableau 18 pour les différentes définitions de NON-TIGHT.

Ceux-ci sont stables à quelques pourcents près, bien que les régions de contrôle définies à l’aide de LOOSE’2 et LOOSE’3 soient comme attendues légèrement plus contaminées que celles définies à l’aide de LOOSE’4 et LOOSE’5. D’autre part, la contamination de LOOSE’3 est plus importante que celle de LOOSE’2 bien qu’une coupure de plus soit relâchée (ΔE , voir section 9), indiquant que celle-ci présente un plus grand taux de rejet

pour les photons NON-TIGHT que pour les jets NON-TIGHT. Aucune conclusion ne peut être donnée concernant les objets passant TIGHT à ce stade.

Dans le paragraphe suivant, l'effet venant de la statistique limitée des régions de contrôle des jets est évalué.

*Statistique limitée des régions de contrôle de jets
(impact de la statistique des régions de contrôle de jets sur les pdfs d'isolation)*

Les régions de contrôle des jets présentent une statistique limitée, particulièrement dans les queues de distribution. Les fluctuations statistiques peuvent alors avoir un effet significatif sur les formes des pdfs de jets utilisées pour la procédure du *template fit* 2D. Celles-ci étant totalement disjointes de la région de signal **TITI** et disjointes entre elles, une procédure indépendante de type *bootstrap* similaire à celle présentée dans la section 15.6.1 est utilisée afin d'évaluer l'effet associé :

- ◇ 1 000 répliques *bootstrap* de chaque région de contrôle de jets (**TIT $\tilde{\text{I}}$** , **$\tilde{\text{T}}$ ITI** et **$\tilde{\text{TIT}}\tilde{\text{I}}$**) sont générées de manière indépendante pour chaque sous-catégorie (en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc) de l'échantillon simulé. En pratique, 1 000 poids poissoniens indépendants centrés en 1 sont assignés pour chaque événement dans **TIT $\tilde{\text{I}}$** , **$\tilde{\text{T}}$ ITI** et **$\tilde{\text{TIT}}\tilde{\text{I}}$** .
- ◇ 1 000 jeux de pdfs jets sont extraits (1 par réplique) et la procédure du *template fit* 2D est exécutée à nouveau pour chacun d'entre eux sur les données **TITI** nominales.
- ◇ La distribution des résultats pour chaque sous-catégorie (définie en termes de $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc) est tracée et l'incertitude finale est prise comme la somme en quadrature entre la RMS et la différence entre la moyenne de la distribution et le résultat nominal (*bootstrap non-closure*). Cette dernière est la plupart du temps très négligeable, sauf pour certaines queues de distribution où l'écart au modèle poissonien engendre de légers *non-closure*.

Dans le cas inclusif, le *non-closure* est négligeable, de l'ordre de 20 à 40 événements pour les différents processus sur un total d'au moins plusieurs milliers ($N_{\gamma\gamma} \simeq 235\,000$, $N_{jj} \simeq 5\,000$). La RMS est également faible et représente une variation d'environ 0.11%, 0.55%, 0.93%, 2.42% et 0.79% pour le nombre d'événements $\gamma\gamma$, γ +jet, jet+ γ , di-jet et ee attendus, respectivement. Pour la décomposition selon les différents spectres, l'effet peut aller jusqu'à une variation de 4% sur le nombre d'événements $\gamma\gamma$ dans les dernières catégories en masse, bien qu'il soit le plus souvent d'environ 1%. Celui-ci reste dans tous les cas sous-dominant par rapport à la statistique limitée des données dans **TITI** ($\simeq 18\%$ pour la dernière catégorie en $m_{\gamma\gamma}$), qui correspond alors simplement à l'erreur statistique renvoyée par la procédure d'ajustement (voir section 15.6.5). Les distributions des erreurs seront montrées et commentées plus en détail dans la figure 128 située en fin de section.

Dans le paragraphe suivant, l'effet systématique venant de la fusion de certaines régions de contrôle est évalué.

*Fusion des régions de contrôle**(dépendances cinématiques des pdfs d'isolation, décomposition spectrale uniquement)*

De façon similaire aux régions **TITI** dans l'échantillon de signal simulé, une incertitude supplémentaire vient de la fusion de certaines régions de contrôle de jets (**TITI**, **TITI** et **TITI**) dans les données pour l'extraction des pdfs d'isolation lorsque la statistique devient trop faible.

Cette technique est utilisée uniquement dans les queues de distribution lors de la décomposition du bruit de fond selon les différentes observables. Elle permet de baisser l'incertitude totale reliée à la statistique des régions de contrôle en faisant un compromis entre fluctuations statistiques des pdfs de jets (voir paragraphe précédent) et biais cinématique de la modélisation dû au non-respect de la définition de la sous-catégorie (traité ici). Par exemple, pour la dernière catégorie de $m_{\gamma\gamma}$ (700-1700 GeV), les régions de contrôle des huit dernières catégories sont fusionnées dans les données de telle sorte que le jeu de pdfs jet nominal est extrait de la région 250-1700 GeV et pourtant ajusté dans 700-1700 GeV. Cela permet de limiter l'effet des fluctuations statistiques des régions de contrôle à 4% malgré le nombre d'événements très limité dans celles-ci (parfois inférieur à 10 événements).

Afin d'estimer le biais cinématique résultant de cette hypothèse de régularité supplémentaire sur les pdfs jets, la méthode du *template fit* 2D est exécutée à nouveau en fusionnant les régions de contrôle de la catégorie d'intérêt uniquement avec la ou les catégories situées sur sa gauche ou sur sa droite. Dans le cas cité précédemment, cela revient à évaluer l'incertitude systématique en utilisant le jeu de pdfs issu de 250-700 GeV uniquement, puisqu'il n'y a aucune catégorie plus à droite.

Malgré le plus grand nombre de fusions opérée par rapport à l'échantillon de signal simulé (voir section 15.6.1), l'effet résultant n'a aucun impact significatif sur les résultats à l'exception de la dernière catégorie en $|\cos \theta_\eta^*|$ ($0.96 < |\cos \theta_\eta^*| < 1.0$), où un effet de -2% sur le nombre d'événements $\gamma\gamma$ est cité. Cela est vraisemblablement dû au fait que la contamination du bruit de fond dans les queues de distribution est en général très réduite, sauf à grand $|\cos \theta_\eta^*|$ (voir section 15.5.2). Les dépendances cinématiques des pdfs de jets sont également probablement plus faibles que celle des photons (voir section 14.3.2). Par ailleurs, l'ajout en quadrature de cette systématique est considéré comme conservateur car l'évaluation de cet effet souffre de la statistique limitée des différentes régions considérées.

Dans le paragraphe suivant, un bilan est donné quant aux incertitudes spécifiques à la modélisation des jets dans le *template fit* 2D.

Bilan

Dans le tableau 19, les incertitudes systématiques reliées à la modélisation des jets sont rapportées pour chaque composante incluse dans le *template fit* 2D pour la région **TITI** complète (cas inclusif). Les distributions des erreurs associées lors des décompositions en fonction des différentes observables de mesure sont représentées dans la figure 128.

Processus	LOOSE'2	LOOSE'3 (non décompté)	LOOSE'5	Stat. des régions T (<i>non-closure</i> /RMS)	Total
$\gamma\gamma$	-4 011	-1 573	+3 587	-4 / $\pm$ 254	+3 596 / -4 019
γ +jet	+1 592	+625	-2 508	+21 / $\pm$ 246	+1 616 / -2 520
jet+ γ	+1 503	+375	-2 010	+21 / $\pm$ 174	+1 513 / -2 018
di-jet	+1 459	+1 058	+345	-39 / $\pm$ 118	+1 464 / -124
ee	-539	-485	+581	+3 / $\pm$ 65	+585 / -543

TAB. 19 – Variations de chaque composante du modèle d’ajustement par rapport aux valeurs nominales citées dans la section 15.5.1 pour la région **TITI** complète en fonction de la définition LOOSE’ utilisée pour les régions de contrôle de jets (LOOSE’4 est utilisé dans la procédure nominale) et de l’effet de la statistique limitée dans les régions de contrôle (procédure *bootstrap*, voir texte).

L’incertitude dominante reliée à la modélisation des jets est dans la quasi-totalité de l’espace des phases celle issue de la définition des régions de contrôle. Son effet est cependant fortement réduit comparé aux mesures précédentes (où elle était souvent ultra-dominante) grâce à l’enrichissement de la région de signal en $\gamma\gamma$ apportée par la coupure sur l’isolation des traces et les coupures plus hautes en E_T^γ . Par ailleurs, la méthode du *template fit* 2D implémentée dans l’analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV a été pensée afin d’avoir à considérer uniquement des régions bien isolées (**II**), ce qui a pour effet de réduire les corrélations entre isolation et identification pour un même candidat et donc baisse l’incertitude reliée à la définition des régions de contrôle. Les incertitudes reliées à la modélisation des jets sont presque toujours sous-dominantes comparées aux incertitudes sur la modélisation des photons, ces dernières présentant par ailleurs de meilleures marges d’amélioration (voir section 15.6.3).

La statistique limitée des régions de contrôle a également un certain effet sur l’incertitude finale dans les queues de distribution des différentes observables ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T), bien que la technique de fusion décrite précédemment permette de la garder largement sous contrôle sans effets systématiques supplémentaires significatifs dans la plupart des cas.

Dans la section 15.6.3, les effets systématiques reliés à la modélisation des électrons sont discutés.

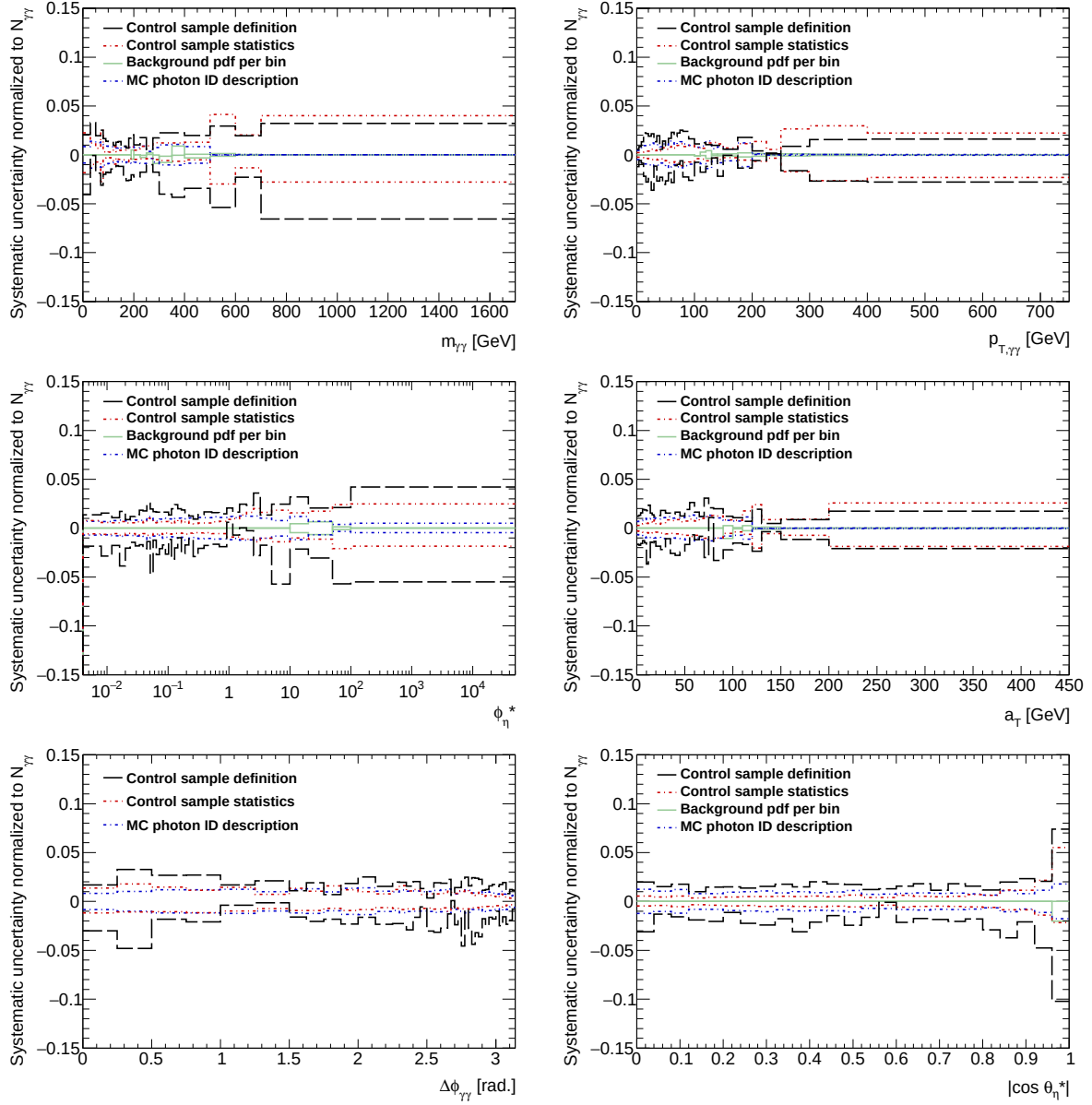


FIG. 128 – Variations relatives du nombre d'événements $\gamma\gamma$ associés aux différentes incertitudes spécifiques à la modélisation des jets : définition des régions de contrôle (*control sample definition*), statistique limitée des régions de contrôle (*control sample statistics*) et fusion des régions de contrôle (*background pdf per bin*). Par ailleurs, l'effet associé au rapport $\mathbf{T}/\tilde{\mathbf{T}}$ dans l'échantillon de signal simulé (lié à la modélisation des photons) est également représenté afin d'éviter de surcharger la figure 126.

15.6.3 Incertitudes spécifiques à la modélisation des électrons

Composante di-électron

Le *template fit* 2D inclut une composante ee dans la procédure d'ajustement. Les corrélations d'isolation associées étant négligeables (voir section 15.4), les pdfs pour le premier et le second électron sont extraites indépendamment dans les deux régions de contrôle *tag-and-probe* (voir section 14.2). Celles-ci présentent une certaine contamination composée par de vrais photons d'une part et des jets mal identifiés d'autre part, toutes deux négligeables selon une étude réalisée dans le cadre de l'analyse $H \rightarrow \gamma\gamma$ [239], mais non quantifiée plus précisément dans l'analyse. Par ailleurs, une contrainte gaussienne externe est appliquée pendant la procédure d'ajustement afin d'assurer la convergence de la composante ee . La contrainte est fonction du nombre d'événements ee attendu dans **TITI** (W_{ee}^{TITI}) et centrée sur les prédictions de l'échantillon simulé $Z^0 \rightarrow ee$, préalablement corrigées grâce à des facteurs d'échelle basés sur les données (voir section 13.2) [234]. Ces prédictions constituent une borne inférieure pour la contamination ee réelle, d'autres processus sous-dominants tels que $t\bar{t}$, ZZ et WW pouvant également contaminer l'échantillon.

Étant donné le grand nombre de sources potentielles d'incertitudes reliées à la modélisation des électrons dans le *template fit* 2D, une évaluation parallèle du nombre d'événements ee dans l'échantillon **TITI** développée par un autre groupe participant à l'analyse est utilisée pour l'estimation des incertitudes systématiques associées. L'étude se base sur le pic du $Z^0 \rightarrow ee$ pour la détermination du taux de mauvaise identification $e \rightarrow \gamma$ d'une part, et sur le nombre d'événements ee , $e\gamma$ et γe observés dans les données pour la normalisation de la composante ee dans **TITI** d'autre part.

L'estimation des taux de mauvaise identification $e \rightarrow \gamma$ dans le cadre de cette étude est très similaire à celle employée pour le calcul des facteurs d'échelle corrigeant les prédictions de l'échantillon DY (voir section 13.2) [234], ou encore à celle décrite dans l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 7$ TeV [203]. Elle devrait cependant être plus précise dans la mesure où les taux sont calculés pour les candidats photons passant la sélection exacte de l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV (en termes de E_T^γ , d'isolation, etc). D'autre part, des taux de mauvaise identification différents sont calculés pour le premier et le second candidat photons, prenant ainsi en compte leur différence de cinématique. Un facteur de corrélation entre les deux taux est également estimé à l'aide de la simulation. Les taux de mauvaise identification extraits sont en général en très bon accord avec ceux issus de l'étude détaillée dans [234] et utilisés pour corriger les prédictions DY employées comme contrainte gaussienne dans le *template fit* 2D, comme le montre la figure 129. Les légères différences visibles pour le second candidat dans les bouchons viennent de différences résiduelles de géométrie (voir section 13.2) et de la coupure sur l'isolation des traces, appliquée pour l'étude associée à l'analyse $\gamma\gamma + X$ uniquement, qui élimine plus de bruit de fond électron dans les données par rapport au MC à cause des problèmes de configuration dans la simulation discutés dans la section 13.2). L'effet est surtout visible à bas p_T (deuxième photon).

Afin de passer des taux de mauvaise identification $e \rightarrow \gamma$ aux nombres d'événements ee , $e\gamma$ et γe attendus dans l'échantillon **TITI**, les événements ee , $e\gamma$ et γe reconstruits dans les données à $\sqrt{s} = 8$ TeV sont sélectionnés et repondérés par les taux de mauvaise identification, après avoir été corrigés des efficacités de sélection (incluant reconstruction,

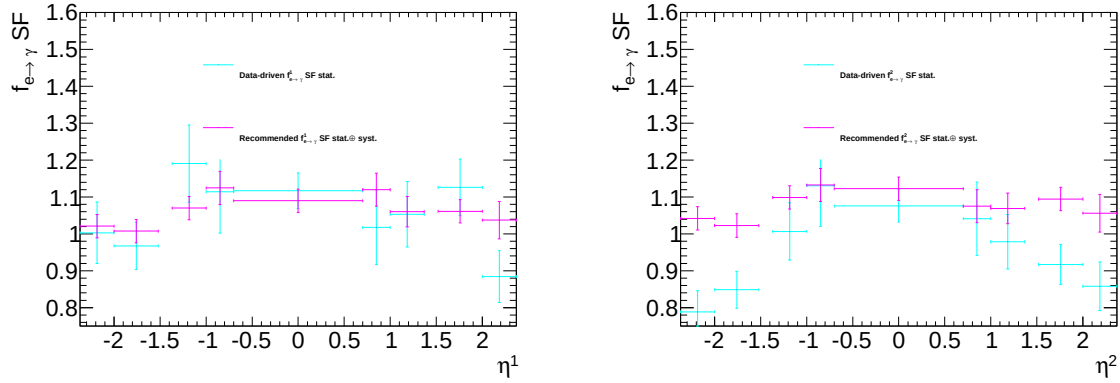


FIG. 129 – Comparaison des facteurs d'échelle sur les taux de mauvaise identification $e \rightarrow \gamma$ du MC obtenus dans l'étude de performance utilisée pour corriger les prédictions DY dans le *template fit 2D* (magenta) [234] et les résultats de l'évaluation parallèle développée par un autre groupe participant à l'analyse $\gamma\gamma + X$ (bleu) [211], représenté pour le premier (gauche) et le deuxième (droite) candidat photon en fonction de leur pseudo-rapidité respective. Les deux études se basant sur le pic du $Z^0 \rightarrow ee$, une large corrélation entre les résultats est attendue. Un très bon accord entre les différentes valeurs est observé. Les légères différences visibles pour le second candidat dans les bouchons proviennent de la coupure sur l'isolation des traces (voir texte). Les incertitudes tiennent uniquement compte de la statistique limitée de l'échantillon de données sans prise en compte des corrélations entre les deux méthodes.

Méthode	ee dans TITI
MC POWHEG+PYTHIA8 DY	$10\,580 \pm 720$
<i>template fit 2D</i>	$8\,275 \pm 680$ (stat) $^{+2\,777}_{-1\,396}$ (syst)
Étude du pic du Z^0 (pas de soustraction des jets)	10 660
Étude du pic du Z^0 (avec soustraction des jets)	9 660

TAB. 20 – Estimation du nombre d'événements ee dans la région **TITI** complète avec l'échantillon simulé DY et facteurs d'échelle correctifs (voir section 13.2), le *template fit 2D* avec contrainte gaussienne (voir section 15.3) et la méthode développée par un autre groupe participant à l'analyse basée sur l'étude du pic du Z^0 et le nombre d'événements ee , $e\gamma$ et γe reconstruit dans les données. Cette dernière peut soustraire ou non le bruit de fond jet dans les données ee , $e\gamma$ et γe . Uniquement l'incertitude sur $ee + e\gamma + \gamma e$ est quantifiée précisément dans le cadre de cette étude. Elle est de l'ordre de $\pm 1\,330$.

identification, etc). Ils sont également corrigés du bruit de fond dû aux photons mal identifiés en électrons (estimé directement à partir du MC, effet très faible), et peuvent être corrigés ou non du bruit de fond jets. Les résultats non corrigés du bruit de fond jets constituent alors une borne supérieure intéressante pour les contaminations ee , $e\gamma$ et γe dans **TITI**.

Les résultats des différentes méthodes d'estimation pour la région **TITI** complète sont résumés dans le tableau 20 et sont trouvés en bon accord. Les trois prédictions sont fortement corrélées, en particulier celle issue des échantillons DY avec facteurs d'échelle correctifs et celles basées sur le pic du Z^0 . Le *template fit* 2D utilise les prédictions de l'échantillon DY simulé comme contrainte pendant l'ajustement, qui converge néanmoins à une valeur significativement plus faible étant donnée la statistique de l'échantillon ($8\,275 \pm 680$ vs $10\,580$). Cependant, l'incertitude systématique, dominée par l'évaluation des différences de cinématiques entre le modèle d'ajustement et la région **TITI** ($+2\,218$ événements, voir section 15.6.4), couvre la différence observée. Comme mentionné dans la section 15.5, les résultats en termes de nombre d'événements ee attendus dans le cadre du *template fit* 2D sont en légère tension dans la région **TITI** complète mais en bon accord lors de la décomposition dans les différents spectres. Cela vient très probablement de la statistique limitée des données dans le second cas, qui donne *de facto* plus de poids à la contrainte gaussienne. Comme nous le verrons dans la section 15.6.4, l'effet résultant sur $\gamma\gamma$ reste sous contrôle.

Tous les chiffres du tableau 20 étant finalement compatibles dans les barres d'erreur, aucune incertitude supplémentaire spécifique à la composante ee n'est ajoutée dans le cas de la région **TITI** complète. Étant donné l'incertitude systématique évaluée dans la section 15.6.4, prendre une incertitude supplémentaire basée sur les différences de résultats entre les méthodes d'estimation mènerait dans ce cas à un double comptage.

En revanche, une incertitude spécifique est calculée lors de la décomposition du bruit de fond pour les différents spectres, prise comme la différence entre les résultats nominaux renvoyés par le *template fit* 2D (basés sur les prédictions DY corrigées, borne inférieure) et les résultats obtenus avec la composante ee fixée pendant la procédure d'ajustement à la valeur estimée par l'étude du pic du Z^0 sans soustraction des jets (borne supérieure). L'erreur résultante sur $\gamma\gamma$ est de l'ordre de 1% à 2% et sous-dominante dans l'ensemble de l'espace de mesure. Les distributions des erreurs seront montrées et commentées plus en détail dans la figure 130 située en fin de section.

Composante électron-photon

La procédure du *template fit* 2D décrite dans la section 15.3 inclut une composante ee dans la procédure d'ajustement mais néglige les événements $e\gamma$ et γe présents dans l'échantillon **TITI**, bien qu'attendus par exemple des productions (sous-dominantes) $W\gamma$ ou $Z\gamma$ (voir section 15.1). Par conséquent, il est nécessaire d'évaluer une incertitude systématique associée à cette simplification.

Les nombres d'événements $e\gamma$ et γe attendus dans l'échantillon **TITI** complet selon la méthode basée sur l'étude du pic du Z^0 (voir paragraphe précédent) sont donnés dans le tableau 21. La mesure finale $\gamma\gamma + X$ se basant uniquement sur le nombre d'événements $\gamma\gamma$ attendu dans **TITI**, l'incertitude sur la composante électron-photon est évaluée de façon conservatrice pour $\gamma\gamma$ en ajoutant en quadrature les nombres d'événements γe

Méthode	γe	$e\gamma$
Étude du pic du Z^0 (pas de soustraction des jets)	916	788
Étude du pic du Z^0 (avec soustraction des jets)	174	0

TAB. 21 – Estimation du nombre d'événements γe et $e\gamma$ dans la région **TITI** complète avec la méthode développée par un autre groupe participant à l'analyse basée sur l'étude du pic du Z^0 et le nombre d'événements ee , $e\gamma$ et γe reconstruits dans les données. Cette dernière peut soustraire ou non le bruit de fond jet dans les données ee , $e\gamma$ et γe . En cas de non-soustraction du bruit de fond hadronique, l'estimation constitue une borne supérieure au nombre d'événements attendu. Les incertitudes individuelles pour $e\gamma$ et γe ne sont pas calculées dans le cadre de cette étude, seule l'incertitude sur la valeur centrale de $ee + e\gamma + \gamma e$ est connue ($\pm 1\ 330$).

et $e\gamma$ attendus sans soustraction du bruit de fond jet (borne supérieure) à l'incertitude systématique négative totale sur $\gamma\gamma$. Cette systématique supplémentaire est implémentée pour la région **TITI** complète (cas inclusif) mais également pour la décomposition du bruit de fond selon les différents spectres.

Elle est considérée comme conservatrice d'une part car elle utilise des bornes supérieures pour les estimations électron-photon (pas de soustraction du fond jet dans les sélections ee , $e\gamma$ et γe dans les données) et d'autre part car elle considère que tous les événements γe et $e\gamma$ sont absorbés dans $\gamma\gamma$ lors de l'ajustement nominal. Or, il y a toutes les chances pour qu'une partie du bruit de fond électron-photon soit absorbée par les composantes γ +jet et jet+ γ pendant l'ajustement (et donc que celui-ci soit déjà partiellement soustrait), comme cela a déjà été constaté pour ee avec les pseudo-données lors du test d'un modèle d'ajustement sans composante ee (voir section 15.4.3).

L'effet résultant de la systématique supplémentaire dû à la composante électron-photon est de l'ordre de -0.4% sur la pureté $\gamma\gamma$ (soit -0.5% en termes de nombre d'événements) et totalement sous-dominant par rapport à la systématique totale ($\simeq \pm 2.6\%$ sur la pureté $\gamma\gamma$). Pour les différents spectres, l'effet est typiquement compris entre 0 et -1% et donc également sous-dominant.

Les distributions des erreurs sont montrées et commentées plus en détail dans le paragraphe suivant et la figure 130, où un bilan est donné concernant les incertitudes spécifiques à la modélisation des électrons dans le *template fit* 2D.

Bilan

Les distributions des incertitudes relatives à la modélisation des électrons dans le *template fit* 2D sont montrées pour les décompositions selon les différentes observables dans la figure 130. Comme décrit dans les paragraphes précédents, aucune systématique supplémentaire n'est considérée pour la composante ee dans le cas inclusif afin d'éviter un double comptage avec la première systématique décrite dans la section 15.6.4, tandis que les chiffres sans soustraction du bruit de fond jet du tableau 21 sont considérés pour $e\gamma$ et γe et ajoutés en quadrature à l'incertitude systématique négative totale sur $\gamma\gamma$.

Comme attendu, les incertitudes liées à la modélisation des électrons ($\simeq 0$ à 2%)

sont toujours totalement sous-dominantes comparées à celles reliées à la modélisation des photons et des jets (s'étalant généralement de 2% à 10%). Ainsi, l'amélioration du traitement des électrons n'est pas prioritaire pour augmenter la précision de la mesure finale.

Dans la section 15.6.4, les incertitudes ayant des conséquences sur l'ensemble du modèle sont abordées.

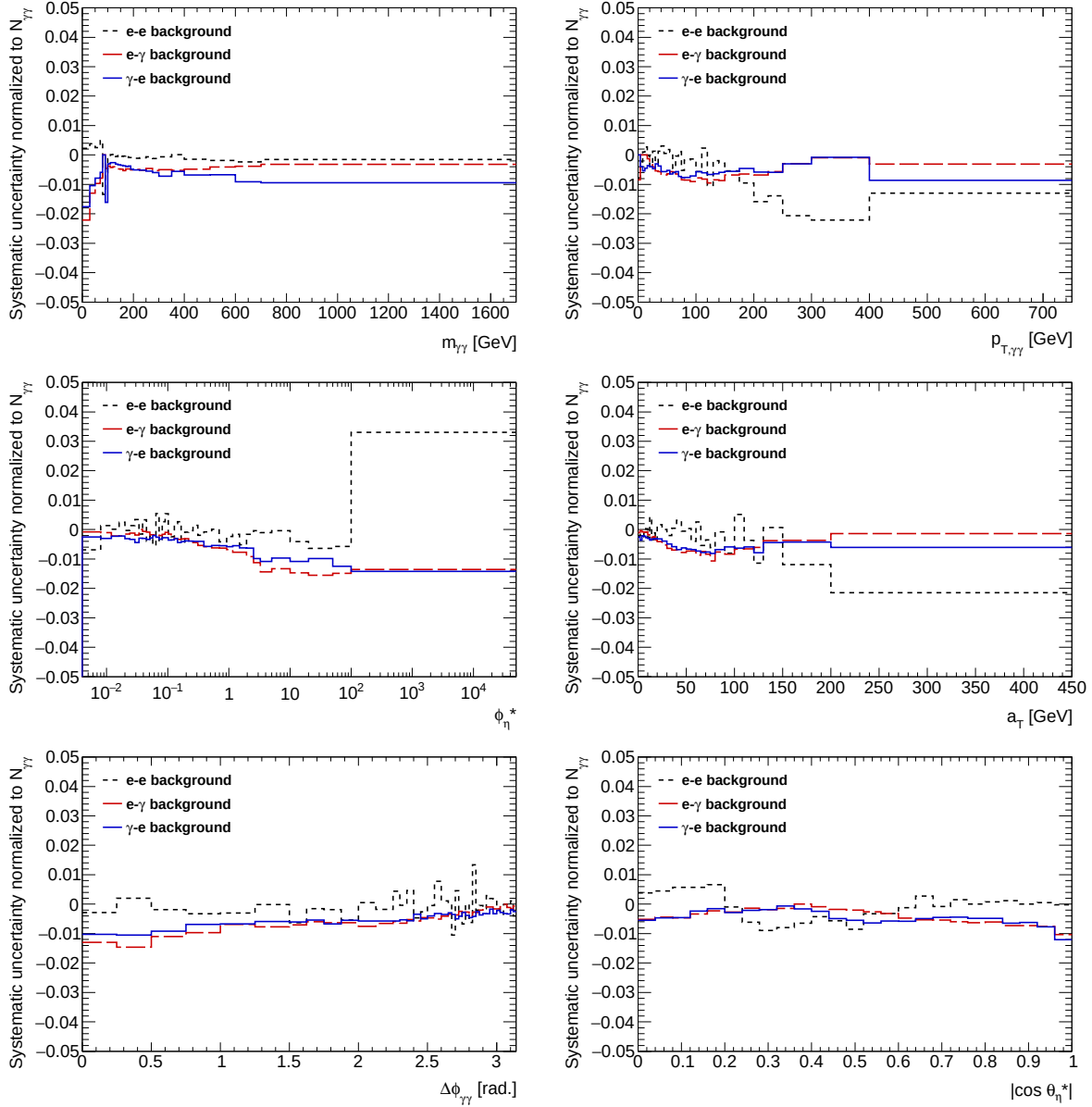


FIG. 130 – Variations relatives du nombre d'événements $\gamma\gamma$ associées aux incertitudes spécifiques à la modélisation des électrons pour les différentes observables de mesure dans le *template fit* 2D : composante ee (composante ee fixée à la valeur estimée par l'étude du pic du Z^0 dans l'ajustement), $e\gamma$ et γe (valeurs basées sur l'étude du pic du Z^0 prises comme incertitudes systématiques négatives additionnelles sur $\gamma\gamma$).

15.6.4 Autres incertitudes

*Différence de cinématique entre le modèle d'ajustement et la région **TITI** dans les données (cas inclusif uniquement)*

L'incertitude reliée aux différences de cinématique entre le modèle d'ajustement (échantillon **TITI** dans l'échantillon de signal simulé et régions de contrôle NON-TIGHT) et la région **TITI** dans les données est évaluée dans le cas inclusif en comparant les résultats $\gamma\gamma$ inclusifs avec ceux obtenus lors des différentes décompositions en termes de $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc (voir section 15.4). Les résultats associés à des décompositions 1D selon $E_{T,1}^\gamma$, $E_{T,2}^\gamma$, η_1^γ et η_2^γ sont également mentionnés à titre illustratif.⁷³ Un résumé de ces comparaisons est présenté dans le tableau 22.

Observable	Événements $\gamma\gamma$	Événements jet (γ +jet, jet+ γ , di-jet)	Événements ee	Nombre d'événements totaux retourné par l'ajustement
région TITI complète	235 392	69 037	8 275	312 704
$m_{\gamma\gamma}$	235 081	67 767	9 849	312 698
$p_{T,\gamma\gamma}$	233 780	68 584	10 328	312 692
$\Delta\phi_{\gamma\gamma}$	233 354	68 882	10 446	312 682
$\cos \theta_\eta^*$	233 593	68 807	10 301	312 700
ϕ_η^*	233 268	68 919	10 493	312 680
a_T	233 973	68 418	10 304	312 695
$E_{T,1}$	233 984	68 425	10 288	312 697
$E_{T,2}$	233 888	68 523	10 284	312 695
η_1^γ	235 375	67 486	9 820	312 694
η_2^γ	235 434	67 429	9 831	312 698
Variation maximum (+)	+0	+0	+2 218	+0
Variation minimum (-)	-2 124	-1 608	-0	-24

TAB. 22 – Estimations du nombre d'événements $\gamma\gamma$, ee et du bruit de fond hadronique total (γ +jet, jet+ γ et di-jet) dans la région **TITI**, obtenues à partir de la méthode du *template fit* 2D inclusive ou à partir de la somme des résultats obtenus dans les différentes catégories en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, $\Delta\phi_{\gamma\gamma}$, $\cos \theta_\eta^*$, ϕ_η^* , a_T , $E_{T,1}$, $E_{T,2}$, η_1^γ et η_2^γ .

Les différences maximales entre les résultats inclusifs et ceux obtenus à partir des différentes décompositions sont issues respectivement de ϕ_η^* , $\Delta\phi_{\gamma\gamma}$ et η_2^γ pour les composantes $\gamma\gamma$, ee et le bruit de fond hadronique total (γ +jet, jet+ γ et di-jet), et s'élèvent à -2 124, +2 218 et -1 608 événements. Ces valeurs sont prises comme incertitudes supplémentaires pour chacune de ces composantes dans le cas inclusif. Pour $\gamma\gamma$ (ee), cela correspond à un effet de l'ordre de -0.7% (+0.7%) sur la pureté, soit respectivement -0.9% et +26% en termes de nombre d'événements. Pour le bruit de fond hadronique, la séparation en trois composantes γ +jet, jet+ γ et di-jet n'est pas possible lors de la décomposition dans chaque sous-catégorie (en termes de $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc) à cause de la statistique très limitée de l'échantillon simulé photon-jet (voir section 15.3.5). Cependant, si on fait l'hypothèse que les proportions relatives de γ +jet, jet+ γ et di-jet sont les mêmes que dans le cas inclusif standard lorsqu'on somme les résultats obtenus dans les différentes catégories, on obtient des variations relatives de pureté respectivement de l'ordre de -0.34%, -0.14% et -0.04% pour γ +jet, jet+ γ et di-jet, prises comme incertitudes supplémentaires sur les bruits de fond hadronique individuels. Un résumé des incertitudes finales en termes de nombre d'événements pour les différentes composantes est donné dans le tableau 23.

Comme nous l'avons vu dans la section 15.4, la prise en compte de cette incertitude

⁷³Les catégories employées sont les mêmes que celles citées dans la section 15.5.2.

Processus	Variation maximum selon les observables
$\gamma\gamma$	$+0$ $-2\ 124$
$\gamma+\text{jet}$	$+0$ $-1\ 059$
$\text{jet}+\gamma$	$+0$ -435
di-jet	$+0$ -113
ee	$+2\ 218$ -0

TAB. 23 – Variations maximales de chaque composante du modèle d’ajustement par rapport aux valeurs nominales citées dans la section 15.5.1 lors de la décomposition du bruit de fond selon les différentes observables détaillées dans le tableau 22. Celles-ci sont prises comme incertitudes systématiques additionnelles dans le cas inclusif (région **TITI** complète).

est nécessaire afin de couvrir les différences résiduelles observées lors des tests effectués avec les pseudo-données, reliées alors aux différences de cinématiques entre le modèle d’ajustement (échantillon **TITI** dans l’échantillon de signal simulé et régions de contrôle NON-TIGHT) et la région **TITI** dans les données. Elle constitue finalement un effet faible comparé à l’incertitude systématique finale pour $\gamma\gamma$ (-0.7% vs $\pm 2.6\%$). L’écart est en revanche significatif pour ee , qui est en général en très bon accord avec les prédictions MC DY dans le cas des différentes décompositions mais tend à être inférieur dans le cas inclusif, lorsque les données d’isolation ont la statistique nécessaire pour contrebalancer la contrainte gaussienne centrée sur la prédiction MC. Une illustration supplémentaire de cet effet est visible dans les résultats de décomposition selon $m_{\gamma\gamma}$, η_1^γ et η_2^γ , qui contiennent les régions les plus enrichies en ee ($80\text{ GeV} < m_{\gamma\gamma} < 100\text{ GeV}$ d’une part et les bouchons d’autre part, voir section 13.2.1) et présentent des estimations ee totales comprises entre les résultats inclusifs (inférieurs aux prédictions du MC DY) et ceux des autres décompositions (en bon accord avec les prédictions MC DY).

Comme nous l’avons vu dans le paragraphe sur la modélisation de l’empilement dans la section 15.6.1, l’écart entre la composante ee inclusive et celle prédite par le MC DY est possiblement une conséquence des résidus significatifs entre données et modèle d’ajustement (voir section 15.5), eux-mêmes probablement issus de différences résiduelles entre la densité d’énergie ambiante de la simulation et des données. Un peaufinage de la simulation sur ce point constituerait un bon axe d’amélioration pour la prochaine analyse $\gamma\gamma + X$ à $\sqrt{s} = 13\text{ TeV}$ et permettrait d’améliorer la convergence de la composante ee si cette hypothèse est confirmée.

Dans le paragraphe suivant, le biais potentiel issu de la procédure de lissage des pdfs d’isolation avec les KDEs (à ne pas confondre avec la procédure utilisée pour lisser la systématique SHERPA vs PYTHIA) est évalué.

*Biais potentiel de la procédure de lissage des pdfs d'isolation
(smoothing par estimateurs de noyaux gaussiens)*

Dans la méthode du *template fit* 2D, un ajustement catégorisé (*binned*) ou non (*unbinned*) est réalisé en fonction de la statistique MC (voir section 15.3.1). Dans tous les cas, les pdfs d'isolation des jets et des électrons sont extraites des régions de contrôle sous forme continue (*unbinned*) à l'aide d'estimateurs de noyaux gaussiens (KDE, pour *kernel density estimators*) [207], qui permettent d'extraire des pdfs non-catégorisées sans faire d'hypothèse spécifique de paramétrage. En effet, les distributions d'isolation présentent des dépendances non-triviales vis-à-vis de E_T^γ , $|\eta^\gamma|$, du taux d'empilement et du processus considéré (nombre de jets additionnels, événement sous-jacent, fragmentation, ...) et par conséquent ne peuvent pas être paramétrées analytiquement simplement et de manière fiable dans tout l'espace des phases.

Dans le cas d'un ajustement catégorisé, l'utilisation de KDEs au lieu de simples histogrammes pour les pdfs de jets et d'électrons équivaut à une simple procédure de lissage, ce qui permet de limiter l'effet des fluctuations statistiques (voir section 15.6.2). La statistique du MC dans **TITI** (alors $> 20\,000$ événements par définition, voir section 15.3.1) est quant à elle suffisante pour se passer des KDEs pour les pdfs photons,⁷⁴ qui sont alors extraites à partir de simples histogrammes.

Dans le cas d'un ajustement non-catégorisé, l'ensemble des pdfs est extrait grâce aux KDEs afin de limiter l'effet des fluctuations statistiques présentes dans les régions de contrôle et dans la région **TITI** de l'échantillon de signal simulé. Le fait de disposer de pdfs non-catégorisées permet également d'effectuer une procédure d'ajustement non-catégorisée (*unbinned extended maximum likelihood*) qui exploite de façon optimale la statistique des données **TITI** comparé à l'estimateur catégorisé correspondant (*binned extended maximum likelihood*) [244].

Cependant, les KDEs incluent un paramètre libre pour l'extraction des différentes pdfs, appelée bande passante et noté h . La bande passante correspond au degré de lissage. Si le lissage est trop fort ($h \gg 1$), les pdfs obtenues seront trop lissées et manqueront la structure sous-jacente des distributions d'énergie d'isolation. En revanche, si le lissage est trop faible ($h \ll 1$), les pdfs obtenues seront affectées par les fluctuations statistiques. Cet effet est illustré dans la figure 131.

En pratique, la valeur de h est réglée à l'oeil catégorie par catégorie en fonction de la statistique de l'échantillon **TITI** considéré (inclusif, queue de distribution, etc) et est différente pour les pdfs di-photon ($h_{\gamma\gamma}$), photons (h_γ), électrons (h_e), di-jet (h_{jj}) et jets (h_j). Afin d'évaluer l'effet systématique sur les résultats dû au choix des h_i , plusieurs variations de ces paramètres sont effectuées en répétant à chaque fois la procédure d'ajustement :

- ◇ (i) Tous les h_i sont augmentés de 50%.
- ◇ (ii) Tous les h_i sont diminués de 50%.

⁷⁴L'utilisation de KDEs pour les pdfs photons (et en particulier pour la pdf 2D $\gamma\gamma$) dans les régions présentant une très haute statistique MC détériorerait significativement les performances de la procédure d'ajustement (temps d'exécution) pour un profit quasi-nul.

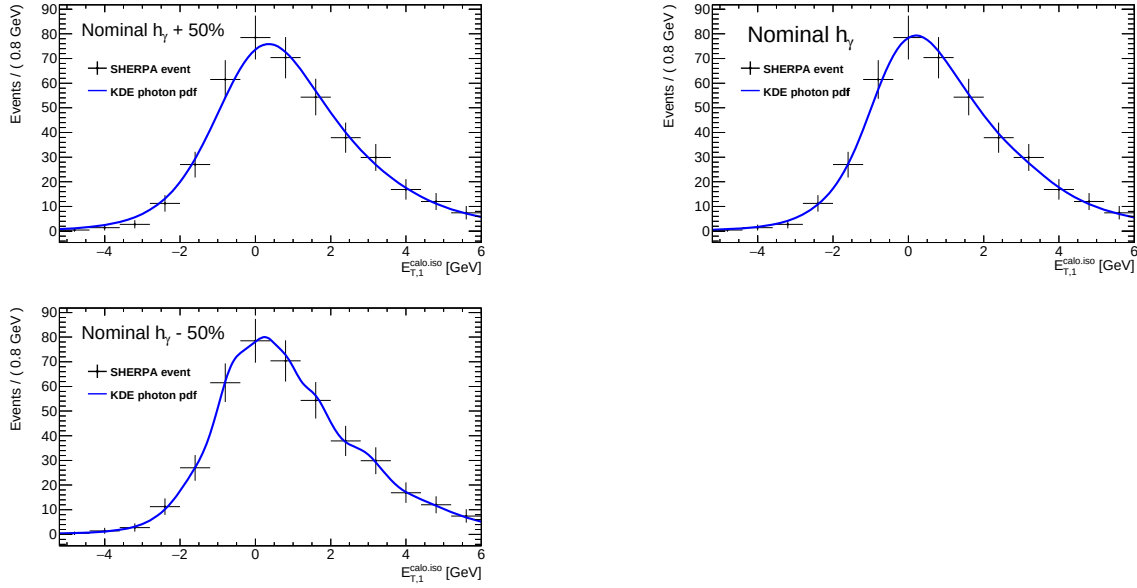


FIG. 131 – Illustration de la procédure de lissage effectuée à l’aide de KDEs pour l’extraction de la pdf 1D du premier photon ($E_{\gamma,1}^{\text{TITI}}$) dans $500 \text{ GeV} < m_{\gamma\gamma} < 1700 \text{ GeV}$ (région fusionnée, voir section 15.6.1). Les événements de l’échantillon de signal SHERPA sont utilisés. La pdf KDE est extraite avec trois bandes passantes différentes : la bande passante standard utilisée dans l’analyse ($h_\gamma = 1.5$ ici, représentée à droite) et la bande passante standard augmentée et diminuée de 50% (correspondant alors à $h_\gamma = 3$ et 0.75, et représentées à gauche). Dans le cas d’une augmentation, la pdf est trop lissée, ce qui mène à un biais sur la position du pic. Dans le cas contraire, des fluctuations apparaissent à cause de fluctuations statistiques.

- ◇ (iii) Tous les h_i reliés aux photons sont augmentés de 50%, tous les h_i reliés aux électrons et aux jets sont diminués de 50%.
- ◇ (iv) Tous les h_i reliés aux photons sont diminués de 50%, tous les h_i reliés aux électrons et aux jets sont augmentés de 50%.

Les variations positives et négatives maximales obtenues en termes de nombre d’événements (ou de pureté) $\gamma\gamma$ sont alors prises comme incertitude systématique. Dans le cas inclusif, les différentes variations obtenues sont représentées dans le tableau 24. Les variations maximales sont atteintes respectivement pour le procédé (i) et (ii) et correspondent à des variations de pureté $\gamma\gamma$ de +0.2%/-0.3%, très faibles comparées à la plupart des incertitudes évoquées dans les paragraphes précédents.

Les variations relatives obtenues en terme d’événements $\gamma\gamma$ pour les différentes observables de mesure sont montrées dans la figure 132. Les variations maximales peuvent alors être obtenues avec les procédés (i) à (iv) selon les sous-catégories considérées. Celles-ci restent en général faibles comparées aux autres incertitudes évoquées dans les sections précédentes. Elles augmentent typiquement dans les queues de distribution ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, ϕ_η^* , $\cos \theta_\eta^*$) et participent parfois au même titre que les incertitudes liées à la statistique limitée de l’échantillon de signal simulé (voir section 15.6.1) et des régions de contrôle (voir section 15.6.2). Cette incertitude montrant probablement une forte

Processus	tous les facteurs +50%	tous les facteurs -50%	photon (+) bruit de fond (-)	photon (-) bruit de fond (+)
$\gamma\gamma$	+644	-895	+192	-442
γ +jet	-397	+432	-716	+753
jet+ γ	-378	+62	-677	+363
di-jet	+248	-74	+481	-310
ee	-119	+466	+713	-360

TAB. 24 – Variations de chaque composante du modèle d’ajustement dans le cas inclusif par rapport aux valeurs nominales citées dans la section 15.5.1 lorsque les bandes passantes des KDEs sont variées ensemble de +50%, -50% et lorsque les bandes passantes reliées aux photons sont augmentées (diminuées) de 50% tandis que celles reliées aux électrons et aux jets sont diminuées (augmentées) de 50%. Les variations maximales sont prises comme incertitudes systématiques (asymétriques) supplémentaires.

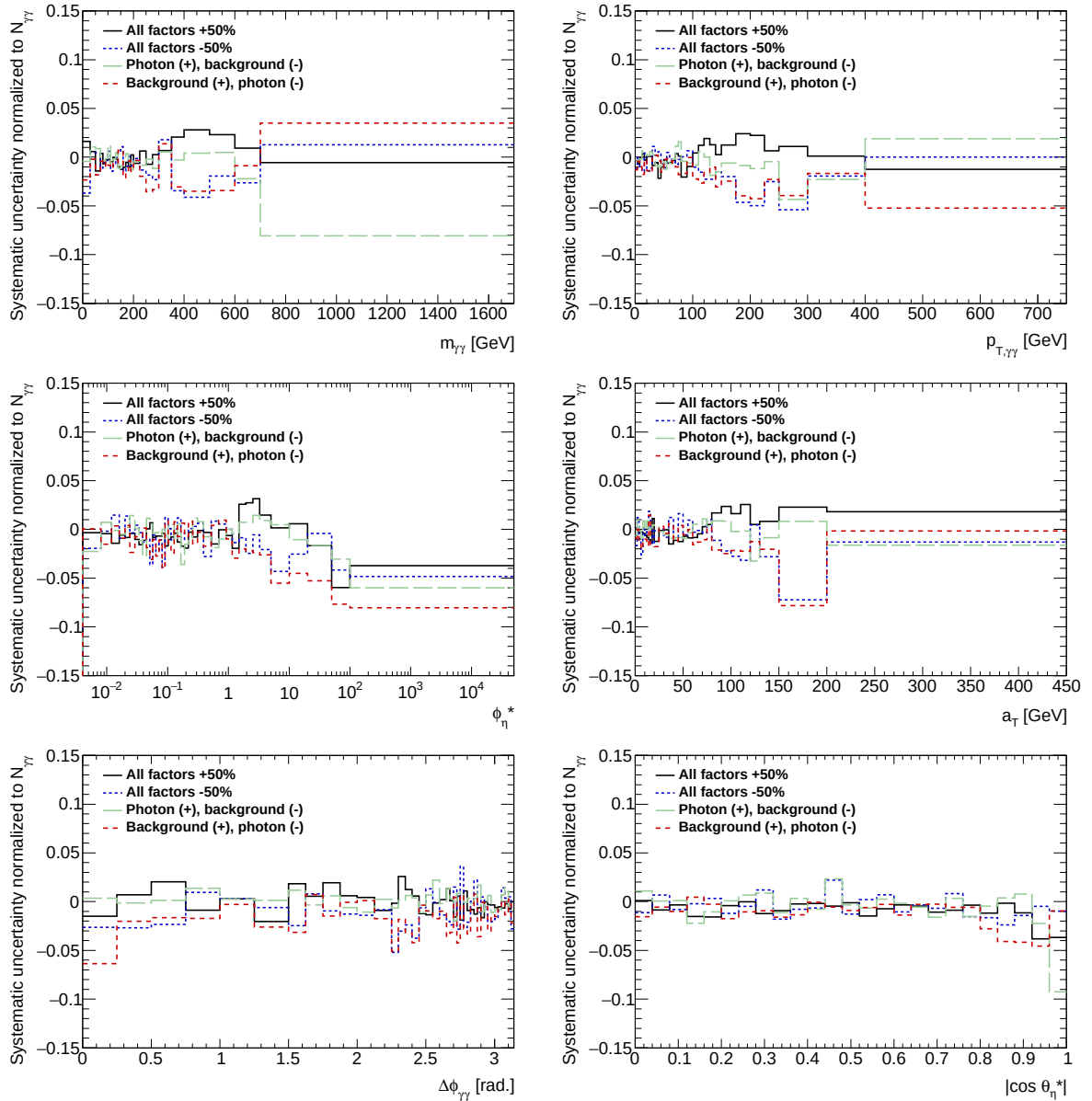


FIG. 132 – Variations relatives du nombre d’événements $\gamma\gamma$ associées au choix arbitraire des bandes passantes des KDEs pour les différentes observables de mesure (voir texte).

corrélation avec l'incertitude statistique, son ajout en quadrature est considéré comme conservateur.

Dans la paragraphe suivant, quelques précisions complémentaires sont apportées sur le traitement des incertitudes reliées aux corrections de contamination dans le *template fit 2D*.

Précisions sur les incertitudes reliées aux corrections de contamination

Dans la méthode du *template fit 2D*, les corrections dues aux contaminations dans les régions de contrôle sont calculées grâce aux équations 64 à 68 (voir section 15.3.4). La précision de ces corrections dépend de la statistique du MC et des régions de contrôle, de la précision du rapport $\tilde{\mathbf{T}}/\mathbf{T}$ dans le MC, mais également de la statistique de la région de signal et finalement de l'ensemble des effets systématiques décrits au cours de la section 15.6 puisque les nombres d'événements ajustés (w_{ik}^{TITI}) entrent dans les équations 64 à 68. Par conséquent, aucune systématique particulière n'est assignée aux corrections de contamination $w_{ik}^{\text{TITI}} \rightarrow W_{ik}^{\text{TITI}}$ directement, leur précision limitée étant automatiquement prise en compte pendant l'évaluation des différents effets systématiques décrits précédemment. En particulier, les incertitudes systématiques sur les facteurs de contamination α_i et α'_i ne sont pas calculées explicitement (voir section 15.3.4).

Une exception est faite dans les 5, 4 et 3 dernières catégories en $p_{T,\gamma\gamma}$, a_T et $m_{\gamma\gamma}$, respectivement, où la statistique réduite des régions de contrôle (parfois inférieure à 10 événements) mène à des résultats non-physiques si les équations 64 à 68 sont utilisées. Cette particularité a déjà été mentionnée dans les sections 15.3.5 et 15.5.2. Dans ce cas, les coefficients α_1 , α_2 et α'_3 sont extrapolés à partir du reste de leur distribution respective⁷⁵ et une incertitude systématique conservatrice de $\pm 50\%$ sur les valeurs obtenues est adoptée, permettant de couvrir l'intégralité des valeurs observées pour ces coefficients. Les valeurs de α_1 , α_2 et α'_3 dans ces sous-catégories sont alors fixées lors de l'évaluation des autres effets systématiques afin d'éviter un double comptage.

Les distributions complètes des coefficients α_1 , α_2 et α'_3 ont déjà été montrées dans la section 15.5.2, figure 110. L'effet de la propagation de l'incertitude systématique de $\pm 50\%$ adoptée pour les queues de distribution de $p_{T,\gamma\gamma}$, a_T et $m_{\gamma\gamma}$ sur les nombres d'événements $\gamma\gamma$ attendus est résumé dans le tableau 25. Celui-ci est compris entre 0.3% et 2.6% relativement au nombre d'événements $\gamma\gamma$ attendus de façon nominale dans l'échantillon **TITI**, ce qui est largement sous-dominant pour ces catégories situées dans les queues de distribution et autrement plus affectées par la statistique limitée des régions de contrôle et de l'échantillon MC (voir sections 15.6.1 et 15.6.2). En effet, une très haute pureté $\gamma\gamma$ est atteinte dans l'échantillon **TITI** pour ces sous-catégories (généralement $> 80\%$), limitant de fait l'influence des corrections dues aux contaminations $\gamma\gamma$ dans les régions de contrôle.

Dans la section 15.6.5, les incertitudes finales sur la soustraction du bruit de fond sont résumées et commentées.

⁷⁵Les formules suivantes, déjà énoncées dans la note de bas de page numéro 68, section 15.5.2, sont utilisées :

$m_{\gamma\gamma} : \alpha_1 (\alpha_2, \alpha'_3) = 0.16 (0.12, 0.02).$

$a_T : \alpha_1 (\alpha_2, \alpha'_3) = 0.2 + 0.01 \times X (0.14 - 0.008 \times X, 0.04 + 0.008 \times X), X$ allant de 1 à 4.

$p_{T,\gamma\gamma} : \alpha_1 (\alpha_2, \alpha'_3) = 0.25 + 0.015 \times X (0.2 + 0.015 \times X, 0.05 + 0.008 \times X), X$ allant de 1 à 5.

Observable	Catégorie	Nombre total d'événements ($\gamma\gamma$)	dont corrections de fuites $\gamma\gamma$	incertitude supplémentaires sur $\gamma\gamma$ (%)
$m_{\gamma\gamma}$	500-600 GeV	505 (385)	20	± 10 (2.6%)
$m_{\gamma\gamma}$	600-700 GeV	195 (174)	3	± 2 (1.1%)
$m_{\gamma\gamma}$	700-1700 GeV	177 (141)	6	± 3 (2.1%)
a_T	120-130 GeV	558 (462)	17	± 9 (1.9%)
a_T	130-150 GeV	647 (591)	6	± 3 (0.5%)
a_T	150-200 GeV	502 (460)	6	± 3 (0.6%)
a_T	200-450 GeV	192 (170)	5	± 3 (1.8%)
$p_{T,\gamma\gamma}$	200-225 GeV	518 (497)	6	± 3 (0.6%)
$p_{T,\gamma\gamma}$	225-250 GeV	331 (328)	1	± 1 (0.3%)
$p_{T,\gamma\gamma}$	250-300 GeV	321 (296)	9	± 5 (1.7%)
$p_{T,\gamma\gamma}$	300-400 GeV	174 (155)	7	± 4 (2.6%)
$p_{T,\gamma\gamma}$	400-750 GeV	65 (57)	2	± 1 (1.8%)

TAB. 25 – Nombre total d'événements observés, nombre final d'événements $\gamma\gamma$ attendus et nombre d'événements $\gamma\gamma$ issus des corrections dues aux contaminations dans les 5, 4 et 3 dernières catégories en $p_{T,\gamma\gamma}$, a_T et $m_{\gamma\gamma}$, respectivement. L'incertitude conservatrice de $\pm 50\%$ adoptée sur ce dernier chiffre est propagée en termes de variations pour $\gamma\gamma$ (nombre d'événements et %).

15.6.5 Résultats finals

Les résultats finals de décomposition du bruit de fond dans l'échantillon **TITI** avec la méthode du *template fit* 2D sont résumés dans le tableau 26 pour la région **TITI** complète (cas inclusif) et dans la figure 133 pour les décompositions selon les différentes observables de mesure. Les incertitudes systématiques totales citées correspondent alors à la somme en quadrature des différents effets systématiques évalués au cours de la section 15.6.

Processus	Décomposition finale (nombre d'événements)	Décomposition finale (%)
$\gamma\gamma$	$235\,392 \pm 1\,022$ (stat) $^{+8\,070}_{-8\,672}$ (syst)	$75.28\% \pm 0.33\%$ (stat) $^{+2.58\%}_{-2.77\%}$ (syst)
γ +jet	$45\,465 \pm 672$ (stat) $^{+8\,470}_{-8\,780}$ (syst)	$14.54\% \pm 0.21\%$ (stat) $^{+2.71\%}_{-2.81\%}$ (syst)
jet+ γ	$18\,699 \pm 590$ (stat) $^{+4\,391}_{-4\,693}$ (syst)	$5.98\% \pm 0.19\%$ (stat) $^{+1.40\%}_{-1.50\%}$ (syst)
di-jet	$4\,873 \pm 352$ (stat) $^{+2\,096}_{-1\,461}$ (syst)	$1.56\% \pm 0.11\%$ (stat) $^{+0.67\%}_{-0.47\%}$ (syst)
ee	$8\,275 \pm 680$ (stat) $^{+2\,777}_{-1\,396}$ (syst)	$2.65\% \pm 0.22\%$ (stat) $^{+0.89\%}_{-0.45\%}$ (syst)

TAB. 26 – Résultats finals de décomposition du bruit de fond dans l'échantillon **TITI** complet avec la méthode du *template fit* 2D. L'incertitude statistique citée tient compte uniquement de la statistique limitée de la région de signal **TITI** (chiffre issus de la procédure d'ajustement basée sur le protocole HESSE [206]), tandis que les incertitudes statistiques reliées à l'échantillon de signal simulé (voir section 15.6.1) et aux régions de contrôle (voir section 15.6.2) sont comptées dans les incertitudes systématiques.

Dans le cas inclusif, l'incertitude systématique totale s'élève à $^{+2.58\%}_{-2.77\%}$ sur la pureté $\gamma\gamma$, soit $^{+3.43\%}_{-3.68\%}$ en termes de nombre d'événements. Elle domine totalement l'incertitude reliée à la statistique limitée de l'échantillon **TITI** ($\pm 0.33\%$ sur la pureté) et constitue le facteur principal limitant la précision expérimentale. Cependant, elle a été réduite jusqu'à un facteur deux comparée à la mesure précédente réalisée par la collaboration ATLAS [203] ($^{+4.5\%}_{-7\%}$ en termes de nombre d'événements $\gamma\gamma$, environ), principalement grâce à l'enrichissement de l'échantillon **TITI** ($E_{T,1}$ et $E_{T,2}$ minimum plus élevés et coupure sur l'isolation des traces), permettant de limiter les effets reliés à la modélisation du bruit de fond, et à l'amélioration de la modélisation de l'isolation des photons, grâce notamment à la prise en compte des différentes dépendances en $E_T^\gamma/|\eta^\gamma|$ et au calcul des corrections de fuites supplémentaires basées sur les données (voir section 10.4). La mesure précédente souffrait par ailleurs de désaccords significatifs entre les diverses méthodes de soustraction du bruit de fond (non comptés dans les chiffres cités précédemment), ce qui n'est plus le cas ici (voir section 15.7).

Pour les différents spectres (voir figure 133), l'incertitude totale est également largement dominée par les effets systématiques dans la plupart des sous-catégories de mesure. L'incertitude statistique participe également significativement voire domine l'incertitude totale citée dans les queues des distributions en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T , ϕ_η^* et dans la dernière catégorie en $|\cos \theta_\eta^*|$. Les incertitudes systématiques pour les différents spectres sont

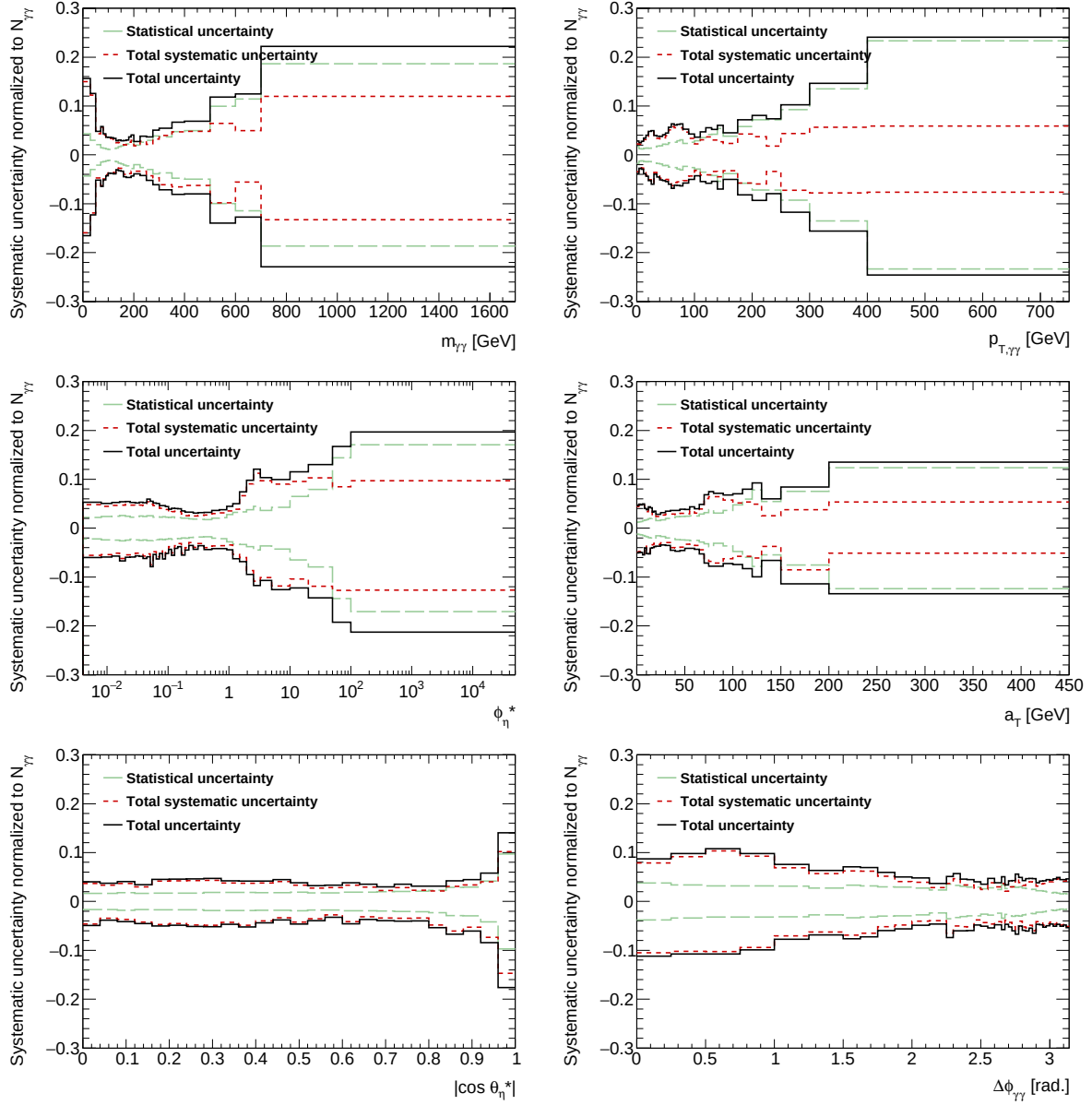


FIG. 133 – Incertitudes statistiques (tenant compte uniquement de la statistique limitée de la région de signal **TITI**), systématiques (incluant l'effet de la statistique limitée de l'échantillon de signal MC et des régions de contrôle) et totales relatives au nombre d'événements $\gamma\gamma$ dans la région de signal **TITI** décomposées en fonction des six observables de mesure : $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, ϕ_η^* , a_T , $\Delta\phi_{\gamma\gamma}$ et $|\cos \theta_\eta^*|$. La méthode du *template fit* 2D est utilisée pour la soustraction des bruits de fond (voir section 15.3).

généralement du même ordre de grandeur ou plus faibles comparées à l'analyse précédente réalisée par la collaboration ATLAS [203]. Par ailleurs, la luminosité plus élevée collectée à $\sqrt{s} = 8$ TeV a permis d'augmenter la portée en $m_{\gamma\gamma}$ et $p_{T,\gamma\gamma}$ de la mesure et de doubler la segmentation des spectres en $\Delta\phi_{\gamma\gamma}$ et $|\cos \theta_\eta^*|$.

La principale piste pour améliorer la précision du *template fit* 2D a été évoquée dans la section 15.6.1 et consiste selon moi à optimiser la description de l'empilement dans la simulation vis-à-vis d'observables basées sur le calorimètre électromagnétique telles que la densité d'énergie ambiante. En particulier, le calorimètre, plus lent, devrait être plus sensible à l'empilement hors-temps que le détecteur interne. Un début d'étude allant dans ce sens est présenté dans l'appendice C. Par ailleurs, des mesures du rapport $\tilde{\mathbf{T}}/\mathbf{T}$ pourraient être réalisées dans les données $Z^0 \rightarrow l^+l^-\gamma$ afin de mieux contrôler les corrections de contamination. Le calcul de l'isolation calorimétrique pourrait également probablement être amélioré en généralisant l'utilisation des *clusters* topologiques, comme cela a été discuté dans la section 11.

Dans la section 15.7, les résultats obtenus avec la méthode du *template fit* 2D sont comparés avec ceux produits avec une méthode de soustraction alternative.

15.7 Comparaison des résultats avec une méthode de soustraction alternative

15.7.1 Comparaison des chiffres bruts

Une méthode de soustraction alternative appelée *4x4 matrix* et décrite par exemple dans la référence [202] ou [211] a été implémentée par un autre groupe impliqué dans l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV. Les résultats de décomposition pour la région **TITI** complète sont comparés dans le tableau 27.

Tous les chiffres sont compatibles dans les incertitudes citées, incluant signal ($\gamma\gamma$) et bruits de fond. Cependant, bien que ne partageant pas la totalité de leurs hypothèses, les deux méthodes de soustraction utilisent les mêmes variables discriminantes pour séparer les différents processus présents dans l'échantillon **TITI** (*ie.* identification TIGHT/NON-TIGHT et critères d'isolation) et reposent sur l'hypothèse de non corrélation entre identification et isolation. Ainsi, des corrélations significatives sont attendues entre leurs résultats, qui devraient également être affectés par un certain nombre d'effets statistiques et systématiques communs.

Il est donc nécessaire d'étudier ces corrélations au préalable afin d'être en mesure d'effectuer une comparaison quantitative entre les résultats des deux méthodes et de démontrer leur compatibilité réelle, objet de la section 15.7.2.

15.7.2 Étude des corrélations entre les incertitudes systématiques

Après un examen attentif des différentes incertitudes affectant chacune des deux méthodes, les effets suivants ont été identifiés comme communs et donc potentiellement corrélés :

Processus	<i>template fit</i> 2D	<i>4x4 matrix</i> + étude du pic du Z^0
$\gamma\gamma$	$235\,392 \pm 1\,022$ (stat) $^{+8\,070}_{-8\,672}$ (syst)	$231\,092 \pm 901$ (stat) $^{+9\,658}_{-8\,437}$ (syst)
γ +jet	$45\,465 \pm 672$ (stat) $^{+8\,470}_{-8\,780}$ (syst)	$45\,175 \pm 529$ (stat) $^{+6\,167}_{-7\,457}$ (syst)
jet+ γ	$18\,699 \pm 590$ (stat) $^{+4\,391}_{-4\,693}$ (syst)	$17\,974 \pm 425$ (stat) $^{+1\,728}_{-1\,868}$ (syst)
di-jet	$4\,873 \pm 352$ (stat) $^{+2\,096}_{-1\,461}$ (syst)	$7\,450 \pm 228$ (stat) $^{+1\,804}_{-1\,469}$ (syst)
$ee + e\gamma + \gamma e$	$8\,275 \pm 680$ (stat) $^{+2\,777}_{-1\,396}$ (syst) (<i>ee</i> seul)	$11\,083 \pm 221$ (stat) $^{+1\,310}_{-1\,310}$ (syst) (dont 10 168 pour <i>ee</i> seul)
Total	312 704 (contrainte poissonnienne centrée en 312 666)	312 754 (fixé dans dans la méthode)

TAB. 27 – Comparaison des résultats de décomposition obtenus dans la région **TITI** complète avec la méthode du *template fit* 2D réalisée pendant la thèse et une méthode de soustraction alternative implémentée par un autre groupe impliqué dans l’analyse appelé *4x4 matrix*.

- ◊ **Statistique limitée des données** : les deux méthodes utilisent les données correspondant à l’échantillon **II**. Dans le cas du *template fit* 2D, celui-ci représente la région de signal (**TITI**) plus l’intégralité des régions de contrôle de jets (**$\tilde{\text{TITI}}$** , **$\text{TIT}\tilde{\text{I}}$** et **$\tilde{\text{T}}\text{IT}\tilde{\text{I}}$**). La méthode de la matrice utilise quant à elle l’intégralité de l’échantillon **II** plus un certain nombre de régions de contrôle supplémentaires non-isolées (voir appendice B). On s’attend donc une corrélation entre les incertitudes statistiques des deux méthodes, bien que celle-ci pourrait être plus ou moins élevée selon l’impact de la statistique des régions de contrôle non-isolés sur les résultats de la méthode de la matrice.
- ◊ **Définition des régions de contrôle de jets** : les deux méthodes utilisent la sélection NON-TIGHT basée sur LOOSE’4 pour la définition de régions de contrôle de jets et varient cette définition (LOOSE’2, LOOSE’3, LOOSE’5) pour l’évaluation des incertitudes systématiques. Une corrélation entre les incertitudes associées est donc attendue. Cependant, les régions de contrôle non-isolées supplémentaires utilisées par la méthode de la matrice pourraient dominer l’incertitude car elles sont a priori plus sensibles aux corrélations entre isolation et identification (voir section 15.2).
- ◊ **Choix du générateur MC** : La méthode de la matrice utilise l’échantillon de signal simulé afin de décontaminer les régions de contrôle non-isolées en vrais photons. Par défaut, l’échantillon PYTHIA8 est employé et l’incertitude systématique correspondante est évaluée avec l’échantillon SHERPA et symétrisée, ce qui correspond à une configuration inversée par rapport au *template fit* 2D.⁷⁶ Par conséquent, une corrélation négative (anti-corrélation) est attendue entre les deux

⁷⁶Ce choix pour la méthode de la matrice est toujours en discussion dans l’analyse.

incertitudes systématiques associées.

- ◇ **Rapport $\mathbf{T}/\tilde{\mathbf{T}}$ dans le MC :** Les deux méthodes utilisent des rapports $\mathbf{T}/\tilde{\mathbf{T}}$ dans l'échantillon de signal simulé afin de décontaminer les régions NON-TIGHT en photons immédiats. Cependant, les régions de contrôle non-isolées supplémentaires utilisées par la méthode de la matrice pourraient dominer l'incertitude car l'efficacité d'identification TIGHT est moins bien connue pour les photons non-isolés (voir section 9).

La solution la plus rigoureuse pour estimer les corrélations décrites ci-dessus consisterait à répéter l'évaluation des incertitudes pour les deux méthodes dans chaque catégorie (inclusive, sous-catégorie en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc) sur 1 000 répliques *bootstrap* communes des données et de remplir au fur et à mesure dans chaque cas (inclusif, sous-catégorie en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc) un histogramme à deux dimensions représentant l'incertitude du *template fit* 2D en fonction de celle de la méthode de la matrice. Les corrélations seraient alors connues avec une très bonne précision statistique pour chaque mesure (inclusif, sous-catégorie en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc), ce qui permettrait même en principe de pouvoir combiner les résultats des deux méthodes.

Cependant, cette solution nécessiterait la construction de répliques *bootstrap* communes et donc d'une harmonisation des variables d'entrée des deux méthodes, ainsi qu'un temps de CPU non-négligeable (de l'ordre d'une semaine, sans compter les phases de développement). Une hypothèse astucieuse permet alors de simplifier considérablement le problème si l'on cherche uniquement à déterminer l'ordre de grandeur des corrélations. En effet, si on suppose les corrélations uniformes dans l'espace des phases, les 193 soustractions de bruit de fond réalisées dans les différentes catégories de l'analyse, incluant l'échantillon **TITI** complet et les décompositions selon les différents spectres (voir section 15.3.5), permettent déjà de se faire une bonne idée de celles-ci.

Néanmoins, il est important de souligner que les données utilisées dans les 193 catégories de l'analyse ne sont pas statistiquement indépendantes. En effet, ϕ_η^* et a_T présentent par exemple une forte corrélation avec $p_{T,\gamma\gamma}$ (voir section 12.3, figure 82). Il sera donc impossible d'estimer la précision statistique sur les corrélations obtenues. Les histogrammes 2D correspondant aux corrélations entre les effets statistiques et systématiques communs entre les deux méthodes de soustraction du bruit de fond basées sur cette approche simplifiée sont présentés dans la figure 134.

Les facteurs de corrélation extraits de la figure 134 relatifs à la statistique limitée des données, au choix du générateur MC, aux définitions des régions de contrôle et au rapport $\mathbf{T}/\tilde{\mathbf{T}}$ dans le MC sont respectivement d'environ +95%, -80%, +20% (LOOSE'2), +30% (LOOSE'5) et 60%. Comme prévu, les deux méthodes ont une incertitude statistique fortement corrélée à cause de l'utilisation de données communes (région de signal + régions de contrôle **II**). Le choix du générateur MC étant inversé entre *template fit* 2D et méthode de la matrice (SHERPA est utilisé par défaut par le *template fit* alors que la matrice utilise PYTHIA8), une anti-corrélation forte est observée pour cet effet. La queue située dans la partie gauche de l'histogramme correspondant (voir figure 134, haut-droite) est reliée au fait que la méthode de la matrice est uniquement affectée par les contaminations positives de photons immédiats dans les régions non-isolées tandis que le *template fit* 2D est affecté quelque soit le sens de déformation des pdfs photons

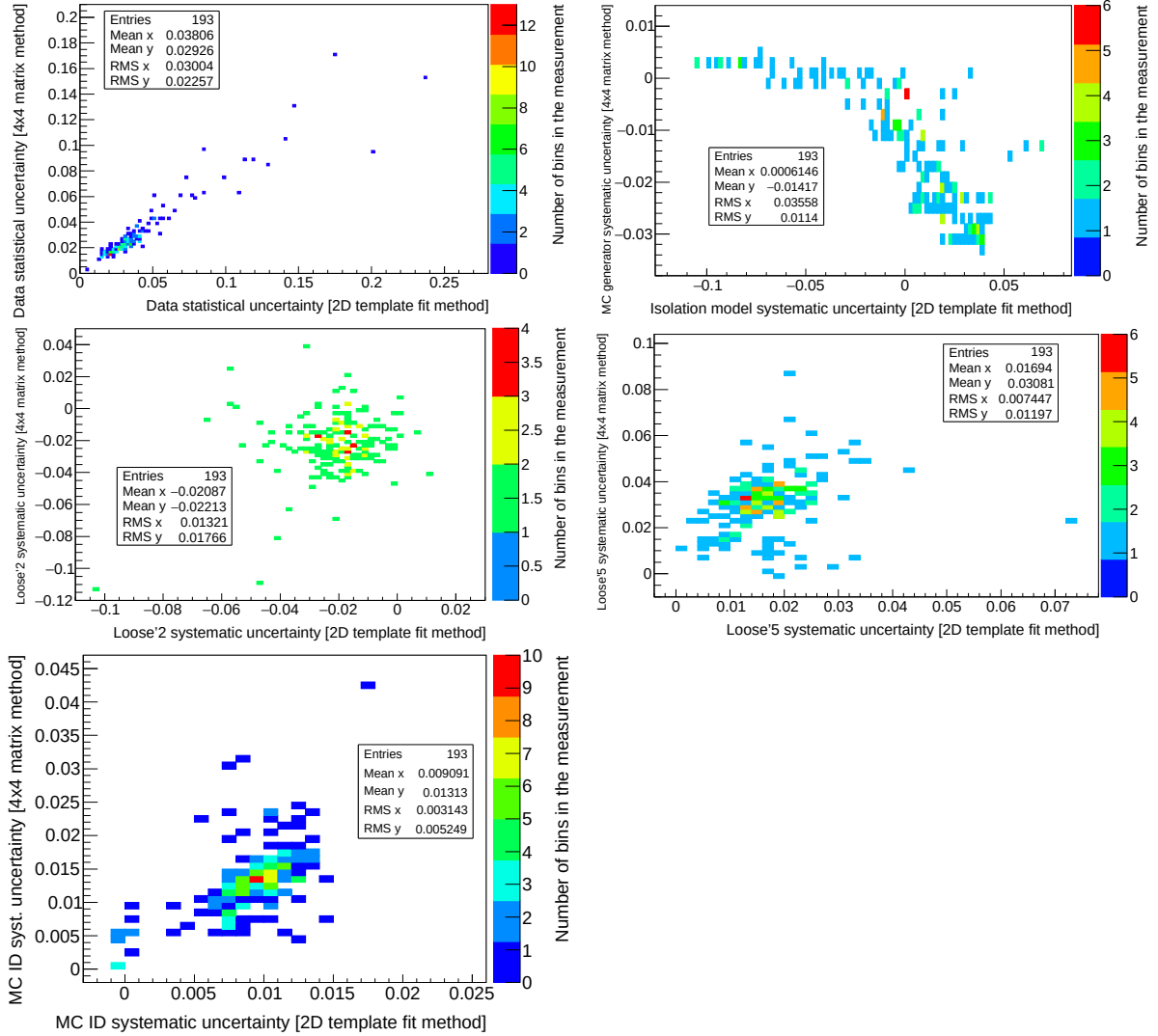


FIG. 134 – Comparaison entre les incertitudes sur le nombre d'événements $\gamma\gamma$ attendus dans **TITI** issues de la statistique limitée des données (en haut à gauche), du choix du générateur MC pour l'échantillon de signal (en haut à droite), des définitions des régions de contrôle (milieu) et du rapport $\mathbf{T}/\bar{\mathbf{T}}$ dans le MC (en bas à gauche) entre la méthode du *template fit* 2D (axe des abscisses) et de la matrice (axe des ordonnées) pour les 193 mesures réalisées dans l'analyse.

(queues de distribution plus/moins importantes).

Finalement, les incertitudes reliées au rapport $\mathbf{T}/\tilde{\mathbf{T}}$ dans le MC et à la définition des régions de contrôle de jets ne sont pas corrélées à 100% entre les deux méthodes, indiquant qu'une partie importante de l'effet systématique subi par la méthode de la matrice vient des régions de contrôle non isolées, non-employées par le *template fit* 2D. Il est intéressant de remarquer que la corrélation est plus forte pour LOOSE'5 que pour LOOSE'2. Bien qu'il ne soit pas possible d'identifier l'origine de cette différence avec certitude étant donné les éléments dont nous disposons, celle-ci pourrait être due uniquement à la statistique plus limitée des échantillons LOOSE'2, baissant la corrélation apparente.

Dans la section 15.7.3, la comparaison finale des résultats du *template fit* 2D et de la méthode de la matrice est effectuée en tenant compte des corrélations entre les différentes incertitudes.

15.7.3 Comparaison finale des résultats en tenant compte des corrélations

Étant donnée l'approche simplifiée employée pour extraire les facteurs de corrélation déterminés dans la section 15.7.2, les chiffres suivants sont considérés pour la comparaison finale entre les deux méthodes :

- ◇ **Statistique limitée des données** : corrélation positive de 100%.
- ◇ **Rapport $\mathbf{T}/\tilde{\mathbf{T}}$ dans le MC** : corrélation positive de 60%.
- ◇ **Définition des régions de contrôle de jets** : corrélation positive de 30%, identique pour LOOSE'2 et LOOSE'5.
- ◇ **Choix du générateur MC** : les valeurs centrales de la méthode de la matrice obtenues avec SHERPA sont considérées afin d'éliminer cette incertitude systématique lors de la comparaison des deux méthodes. Les autres incertitudes relatives de la méthode de la matrice sont supposées inchangées.
- ◇ **Autre incertitudes** : non-corrélées.

Les rapports entre les nombres d'événements $\gamma\gamma$ obtenus à l'aide des deux méthodes sont présentés pour la région **TITI** complète (cas inclusif) dans le tableau 28. Les incertitudes finales tiennent désormais compte des corrélations décrites dans la section 15.7.2. Le rapport (*4x4 matrix* + étude du pic du Z^0) sur (*template fit* 2D) pour $\gamma\gamma$ se situe à environ une déviation standard de 1, confirmant la bonne compatibilité des deux méthodes.

Le même exercice n'a pas été répété pour les différentes composantes de bruit de fond par manque de temps mais également car une interprétation quantitative des résultats serait plus limitée. En effet, prenons l'exemple d'une forte anti-corrélation entre les erreurs systématiques de $\gamma\gamma$ et γ +jet. Dans ce cas, une compatibilité à $+1\sigma$ (même signe) entre les deux méthodes pour ces deux types d'événements indiquerait une tension plus élevée qu'il n'y paraît entre les résultats. Afin d'aller plus loin dans cette étude, les

corrélations entre les effets systématiques des différentes composantes ($\gamma\gamma$, γ +jet, etc) devraient être étudiées pour les deux méthodes, ce qui dépasse probablement le cadre de l'analyse.

Processus	($4x4$ matrix + étude du pic du Z^0) sur (<i>template fit</i> 2D)	Incertitude non-corrélée	Statistique (corr. +100%)	$\tilde{\mathbf{T}}/\mathbf{T}$ (corr. +60%)	Définition de $\tilde{\mathbf{T}}$ (corr. +30%)	Total
$\gamma\gamma$	0.962	+0.023 -0.026	± 0.001	± 0.010	+0.031 -0.026	0.041 0.038
γ +jet	0.994					
jet+ γ	0.961					
di-jet	1.529					
ee	1.339					

TAB. 28 – Comparaison des résultats de décomposition obtenus dans la région **TITI** complète avec la méthode du *template fit* 2D (réalisée pendant la thèse) et une méthode de soustraction alternative implémentée par un autre groupe impliqué dans l'analyse (appelée *4x4 matrix*). Les corrélations entre les incertitudes sont prises en compte et propagées dans le calcul des erreurs finales. Les valeurs centrales obtenues avec l'échantillon de signal SHERPA sont considérées pour les deux méthodes.

Par ailleurs, les distributions des rapports (*4x4 matrix* + étude du pic du Z^0) sur (*template fit* 2D) pour $\gamma\gamma$ selon les différentes observables de mesures sont représentées dans la figure 135. L'accord entre les deux méthodes est en général satisfaisant avec des différences s'étalant de 0 à 2σ , à de rares exceptions près. Les résultats de la matrice en termes de nombre d'événements $\gamma\gamma$ attendus ont cependant tendance à être légèrement inférieurs à ceux du *template fit* 2D, de l'ordre de 3 à 5%. Cette tendance semble plate pour toutes les observables sauf $m_{\gamma\gamma}$, où une certaine pente est observée dans la figure 135 en haut à gauche. Des études sont toujours en cours afin d'en déterminer l'origine.

Il est tentant de vouloir combiner les différentes catégories des spectres de la figure 135 mentalement afin de faire baisser l'incertitude totale et ainsi revendiquer le fait que les deux méthodes ne sont pas compatibles. Cependant, l'incertitude est ici totalement dominée par des effets systématiques et non statistiques (l'erreur due à la statistique des données est 100% corrélée entre les deux méthodes), qui sont possiblement fortement corrélés catégorie par catégorie (*bin-by-bin correlated*). Il est donc a priori faux de penser qu'une combinaison des différentes catégories mènerait à une réduction des erreurs.

Par conséquent, les deux méthodes sont considérées comme compatibles dans les incertitudes citées et aucune erreur systématique supplémentaire n'est ajoutée. Cela constitue une validation supplémentaire importante de la méthode du *template fit* 2D et du bon traitement de ses effets systématiques (voir section 15.6), ainsi qu'une évolution majeure par rapport à la mesure précédente d'ATLAS [203], où des incertitudes supplémentaires importantes avaient dû être considérées afin de couvrir les différences observées entre les différentes méthodes de soustraction employées.

Une combinaison des deux méthodes étant très délicate du fait des corrélations non-négligeables entre les valeurs centrales et les différentes incertitudes, les résultats finals d'une seule méthode devrait figurer dans la publication d'ATLAS. Des discussions sont

toujours en cours au sein de l'équipe d'analyse à ce sujet.

Dans la section 16, la déconvolution des effets de détecteurs et les corrections des inefficacités permettant de passer de résultats cités en termes de nombre d'événements à des mesures de sections efficaces sont abordées en détail.

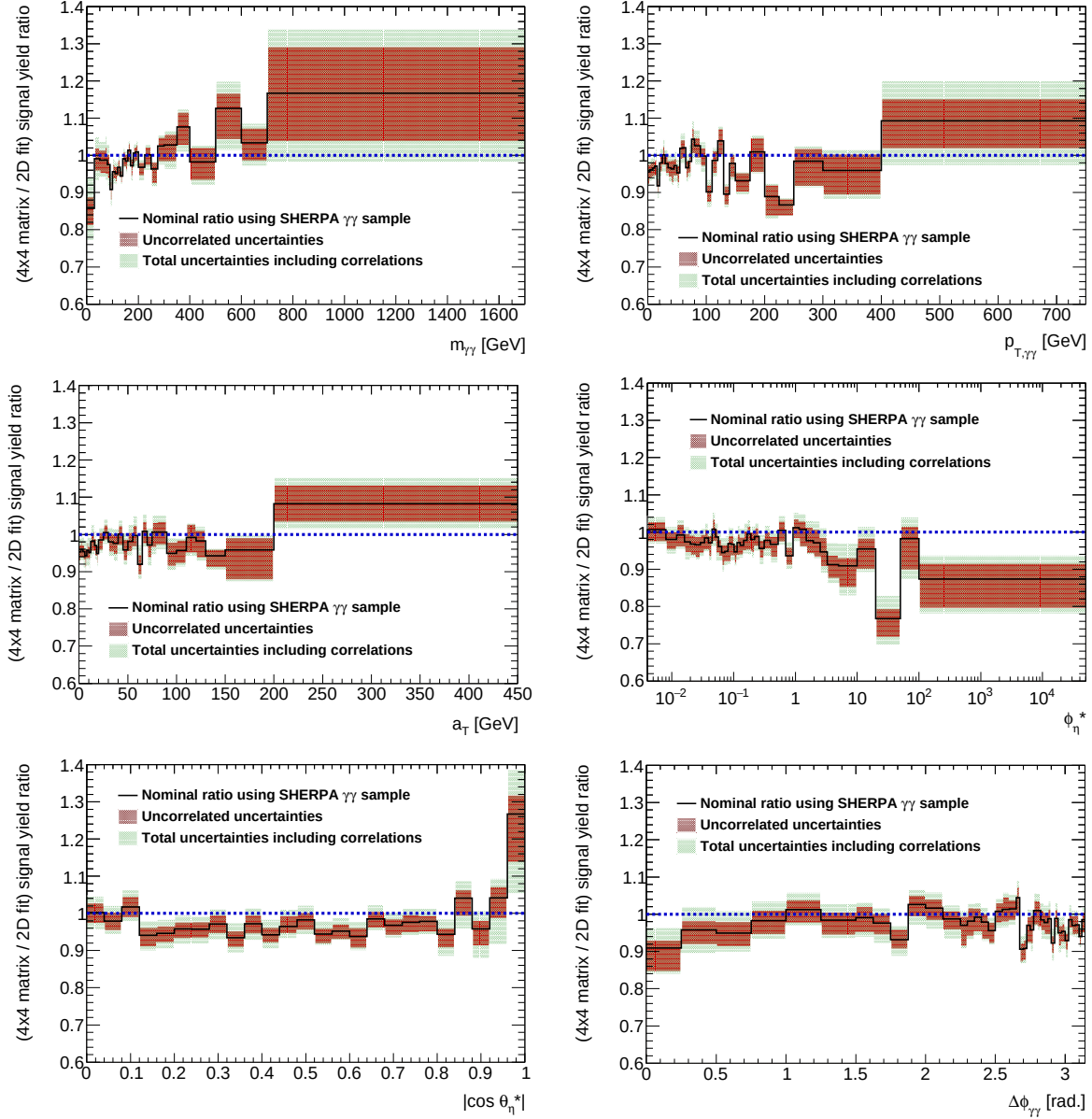


FIG. 135 – Comparaison des résultats après soustraction du bruit de fond obtenus dans la région **TITI** décomposée selon les différentes observables de mesure avec la méthode du *template fit* 2D (réalisée pendant la thèse) et une méthode de soustraction alternative implémentée par un autre groupe impliqué dans l'analyse (4×4 matrix). Les incertitudes citées sont séparées entre celles non-corrélées entre les deux méthodes (bandes rouges) et les incertitudes totales tenant compte à la fois des incertitudes non-corrélées, des corrélations statistiques, de l'usage du rapport $\mathbf{T}/\tilde{\mathbf{T}}$ dans le MC et de l'utilisation des mêmes régions de contrôle de jets dans les deux méthodes (définition de NON-TIGHT). Les valeurs centrales obtenues avec l'échantillon de signal SHERPA sont considérées pour les deux méthodes.

16 Déconvolution des effets de détecteurs et corrections des inefficacités

16.1 Considérations générales

La section 16 est dédiée à la déconvolution des effets de détecteurs, nécessaire du fait de la résolution expérimentale limitée et des inefficacités de détection (reconstruction, sélection, ...), qui altère les distributions observées dans la section 15. Cette étape permet de passer de résultats cités en terme de nombre d'événements observés dans ATLAS à des valeurs de section efficace indépendantes du détecteur considéré. Elle est obligatoire afin de pouvoir comparer les résultats de l'analyse avec des mesures réalisées dans d'autres détecteurs (CMS, D0, CDF, ...), mais également utile afin de les confronter aux prédictions théoriques les plus avancées, qui se limitent souvent à une description partonique de l'état final. La déconvolution des effets de résolution et la correction des inefficacités seront traitées en même temps dans l'analyse et appelées "étape de déconvolution" sans distinction.

Dans la section 16.2, les matrices de réponse du détecteur ATLAS vis-à-vis du signal $\gamma\gamma + X$ sont définies et estimées à l'aide de l'échantillon de signal simulé SHERPA. Des facteurs de correction catégorie par catégorie (*bin-by-bin*) sont ensuite calculés (à nouveau à l'aide de l'échantillon SHERPA) et représentés pour les différentes observables. Dans la section 16.3, ces facteurs sont étudiés en détail à l'aide de l'échantillon de signal simulé SHERPA et des données de collisions. Les différents effets de détecteur les conditionnant sont identifiés et analysés afin de s'assurer que leurs formes selon les différentes observables sont bien comprises. Les facteurs de correction catégorie par catégorie seront utilisés pour la déconvolution finale de la section efficace fiducielle totale uniquement.

Dans la section 16.4, une parenthèse est faite sur la sélection fiducielle effectuée sur l'énergie d'isolation dans l'analyse (voir section 12.2). La correspondance entre l'énergie d'isolation transverse au niveau particule et au niveau reconstruit est étudiée, permettant de déterminer à partir des coupures d'isolation au niveau reconstruit (guidées principalement par des considérations expérimentales) la sélection fiducielle optimale, sur la base de critères de migrations minimales entre le niveau généré et le niveau reconstruit.

Dans la section 16.5, une méthode de déconvolution plus complexe que la déconvolution catégorie par catégorie appelée déconvolution itérative [231] est définie et implémentée pour les sections efficaces différentielles de l'analyse. Plusieurs vérifications sont effectuées afin de s'assurer de la fiabilité des résultats, incluant la comparaison de plusieurs méthodes de déconvolution. L'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV utilise donc la méthode de déconvolution itérative (catégorie par catégorie) pour l'obtention des sections efficaces différentielles (de la section efficace fiducielle totale). Ces choix seront justifiés dans la section 16.5.

Finalement, la section 16.6 traite de la propagation des incertitudes reliées à la soustraction du bruit de fond (voir section 15) dans la procédure de déconvolution ainsi que de la quantification des incertitudes systématiques supplémentaires entrant en jeu lors de cette dernière étape. La comparaison des mesures finales avec un large éventail de prédictions théoriques sera effectuée dans la section 17.

16.2 Matrices de réponses et définition de facteurs de correction catégorie par catégorie

La matrice de réponse d'une observable X générée dans B_G catégories (*bins*) et reconstruite dans B_R catégories est définie comme la matrice de dimension $(B_G + 1) \times (B_R + 1)$ de coefficients $N_{i,j}^X$, eux-mêmes définis comme le nombre d'événements avec deux photons immédiats générés dans la catégorie i , passant la sélection fiducielle au niveau généré (voir section 12.2) et reconstruits dans l'échantillon **TITI** de la catégorie j (voir section 14.6). Par ailleurs, les deux candidats photons présélectionnés doivent correspondre grossièrement aux deux photons immédiats générés en termes d'angle d'azimut et de pseudo-rapacité afin d'éliminer les artefacts de reconstruction issus de l'empilement et de la cascade de partons présents dans l'échantillon simulé. Ce critère élimine moins de 0.1% des événements simulés. La coupure reliée au système de déclenchement n'est pas appliquée car son efficacité est évaluée séparément dans la section 16.3.5 à partir des données. Celle-ci étant très proche de 100%, elle ne devrait avoir aucun effet significatif sur les matrices de réponse.

La catégorie supplémentaire dans les catégories générées ($B_G + 1$) est nécessaire afin de tenir compte des événements di-photon ne passant pas la sélection fiducielle au niveau générée mais tout de même reconstruit dans l'échantillon **TITI** d'une catégorie au niveau reconstruit (faux événements de signal). La situation inverse requiert une catégorie supplémentaire dans les catégories reconstruites, $B_R + 1$ (inefficacités). Les catégories reconstruites incluant pour chaque observable l'ensemble des événements observés, aucune catégorie supplémentaire n'est nécessaire afin de tenir compte d'événements ayant potentiellement migrés hors de l'intervalle de mesure au niveau reconstruit.

Les matrices de réponses estimées avec l'échantillon de signal simulé SHERPA sont représentées dans la figure 136 pour les différentes observables de mesure. Elles sont toutes presque diagonales avec systématiquement moins de 30% de migrations entre les catégories générées et reconstruites correspondantes (faux événements de signal et inefficacités exclus), attendu de l'excellente résolution du calorimètre à argon liquide d'ATLAS utilisé pour la mesure des photons (voir section 5.4). Bien que la distribution de a_T semble légèrement plus affectée par les migrations, les valeurs situées sur la diagonale de sa matrice de réponse sont comparables aux autres observables. Le caractère quasi-diagonal de ces matrices constitue une première validation du choix des catégories, leur largeur étant bien au moins égale à la résolution attendue (définie généralement comme un taux de migration de 68% dans la limite gaussienne). La situation inverse engendrerait des corrections importantes de déconvolution dues à la résolution et par conséquent une augmentation sensible des incertitudes systématiques dues au choix du générateur MC pour cette étape (voir section 16.6). D'autres indicateurs permettant de valider le choix des catégories seront détaillés dans la section 16.3.

La première idée venant à l'esprit pour corriger les distributions mesurées est d'inverser les matrices de réponse de la figure 136 et de les appliquer aux résultats obtenus dans la section 15. Cependant, il est possible de montrer que cette méthode est très sensible aux fluctuations statistiques des données et donc peu robuste [244]. Ainsi, lorsque les matrices de réponse sont quasi-diagonales et que l'échantillon de signal simulé décrit raisonnablement bien les données, il est préférable de définir des facteurs de correction catégorie par catégorie pour corriger les distributions mesurées de la résolution limitée,

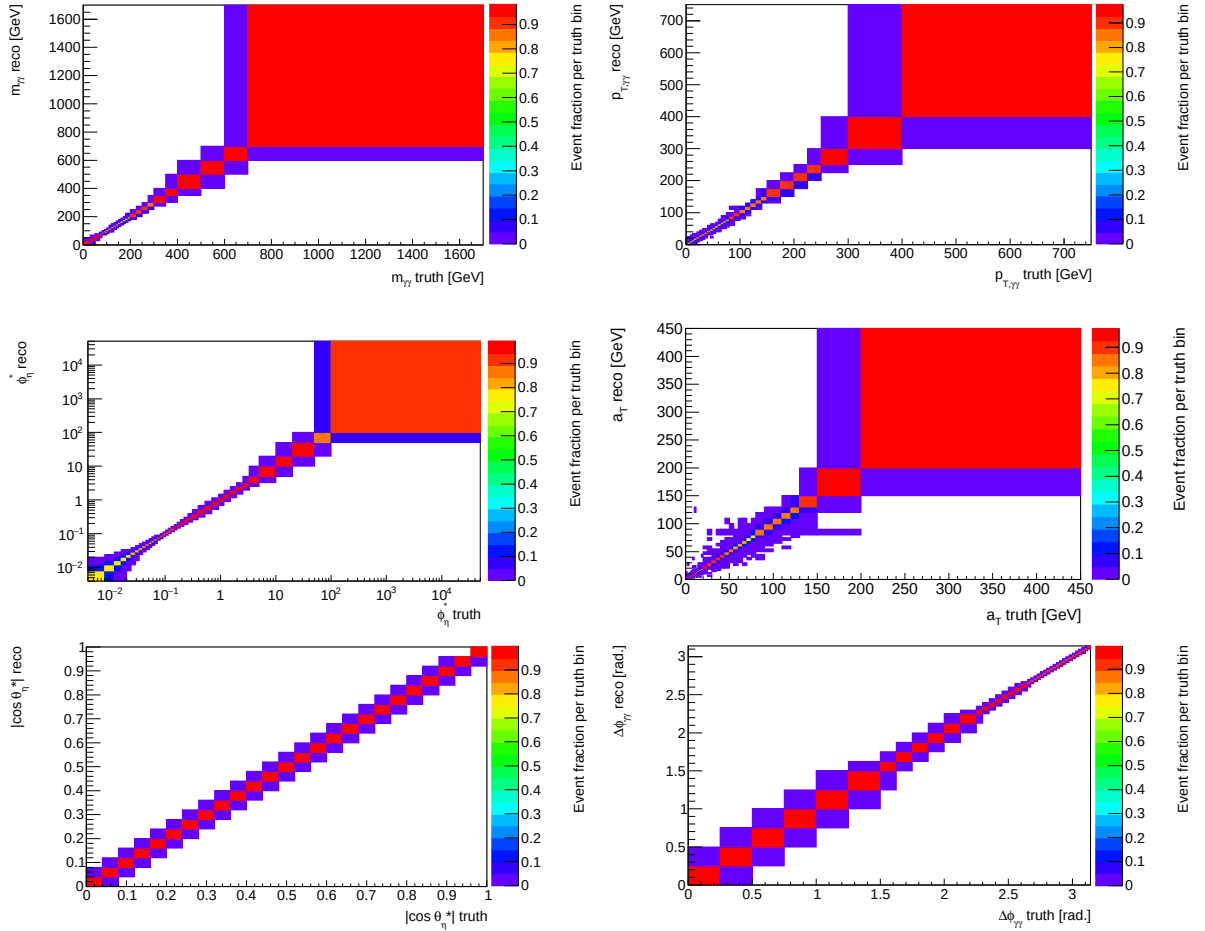


FIG. 136 – Matrices de réponse du détecteur ATLAS vis-à-vis du signal $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV estimées avec l'échantillon de signal simulé SHERPA pour les différentes observables de mesure. Les catégories ont été normalisées colonne par colonne afin d'avoir une vision plus claire des migrations au niveau reconstruit. Les catégories supplémentaires (voir texte) ont été exclues de la normalisation afin de permettre la visualisation des migrations indépendamment des inefficacités et des faux événements de signal.

des inefficacités et des faux événements de signal [244], définis de la façon suivante :

$$C_{total,i}^X = \frac{\sum_{l=1}^{B_R+1} N_{l,i}^X}{\sum_{k=1}^{B_G+1} N_{i,k}^X} = \frac{N_{i,\mathbf{TITI}}^{reco}}{N_{i,fid.(truth)}^{truth}}, \quad (76)$$

où $C_{total,i}^X$ est le facteur de correction pour la catégorie i de l'observable X , les $N_{i,j}^X$ sont les éléments de matrice de réponse et B_G (B_R) est le nombre de catégories au niveau généré (reconstruit). Le facteur de correction peut également s'exprimer comme le rapport entre le nombre d'événements reconstruits dans l'échantillon **TITI** de la catégorie i ($N_{i,\mathbf{TITI}}^{reco}$) sur le nombre d'événements générés dans l'espace fiduciel correspondant ($N_{i,fid.(truth)}^{truth}$).

La section efficace fiducielle totale et les sections efficaces différentielles peuvent alors

s'exprimer ainsi (déconvolution catégorie par catégorie) :

$$\sigma_{fid.} = \frac{N_{data}^{\mathbf{TITI}} \cdot x_{\gamma\gamma}^{\mathbf{TITI}}}{\epsilon_{trig.} \cdot C_{total} \cdot \mathcal{L}}, \quad (77)$$

$$\frac{d\sigma_i}{dX} = \frac{N_{data,i}^{\mathbf{TITI}} \cdot x_{\gamma\gamma,i}^{\mathbf{TITI}}}{\Delta X \cdot \epsilon_{trig.,i} \cdot C_{total,i} \cdot \mathcal{L}}, \quad (78)$$

où $N_{data}^{\mathbf{TITI}}$ est le nombre d'événements observés dans l'échantillon **TITI**, $x_{\gamma\gamma}^{\mathbf{TITI}}$ est la pureté $\gamma\gamma$ associée déterminée dans la section 15, $\epsilon_{trig.}$ est l'efficacité du système de déclenchement (voir section 16.3.5), C_{total} est le facteur correctif introduit précédemment et $\mathcal{L}$ est la luminosité intégrée de l'échantillon de données (voir section 13). Pour les différents spectres, les sections efficaces sont également divisées par la largeur de la catégorie ΔX considérée afin d'obtenir une quantité indépendante du choix des catégories (sections efficaces différentielles). Cependant, celles-ci sont toujours citées dans une certaine catégorie (ex : $d\sigma/dm_{\gamma\gamma}$ dans $600 \text{ GeV} < m_{\gamma\gamma} < 1700 \text{ GeV}$) afin d'éviter les problèmes liés à la répartition asymétrique de la statistique des données dans les différentes catégories, *ie.* $d\sigma/dm_{\gamma\gamma}$ à citer à 1150 GeV ou au $m_{\gamma\gamma}$ moyen des événements dans la catégorie $600 \text{ GeV} < m_{\gamma\gamma} < 1700 \text{ GeV}$ (vraisemblablement inférieur à 1150 GeV).

Les facteurs de correction catégorie par catégorie peuvent être évalués avec l'échantillon de signal simulé SHERPA. Pour celui correspondant au volume fiduciel et à l'échantillon **TITI** complet (voir équation 77), on obtient $C_{total} = 0.708$, avec une incertitude statistique MC négligeable. La méthode de déconvolution catégorie par catégorie est utilisée pour le calcul de la section efficace fiducielle totale $\gamma\gamma + X$ à $\sqrt{s} = 8 \text{ TeV}$, comme nous le verrons dans la section 16.5.

Les facteurs de correction catégorie par catégorie sont représentés pour les différentes observables de mesure dans la figure 137. Ils présentent généralement des variations significatives pour une même observable. Afin de comprendre plus en détail les formes observées, il est nécessaire de réaliser une étude dédiée basée sur l'échantillon de signal simulé SHERPA, objet de la section 16.3. La méthode de déconvolution itérative sera finalement celle utilisée pour les résultats finals concernant les sections efficaces différentielles (voir section 16.5). Les différentes incertitudes systématiques seront évaluées/propagées dans les sections 16.5 et 16.6 pour l'ensemble des résultats (inclusif, différentiels).

16.3 Identification et analyse des différents effets conditionnant les facteurs de correction

16.3.1 Quelques définitions

L'étude proposée vise à étudier et comprendre la forme et les valeurs des différents facteurs de correction présentés dans la section 16.2, entrant dans le calcul des sections efficaces finales dans le cadre d'une déconvolution catégorie par catégorie. Pour cela, l'échantillon de signal SHERPA est utilisé ainsi que les données à chaque fois que cela

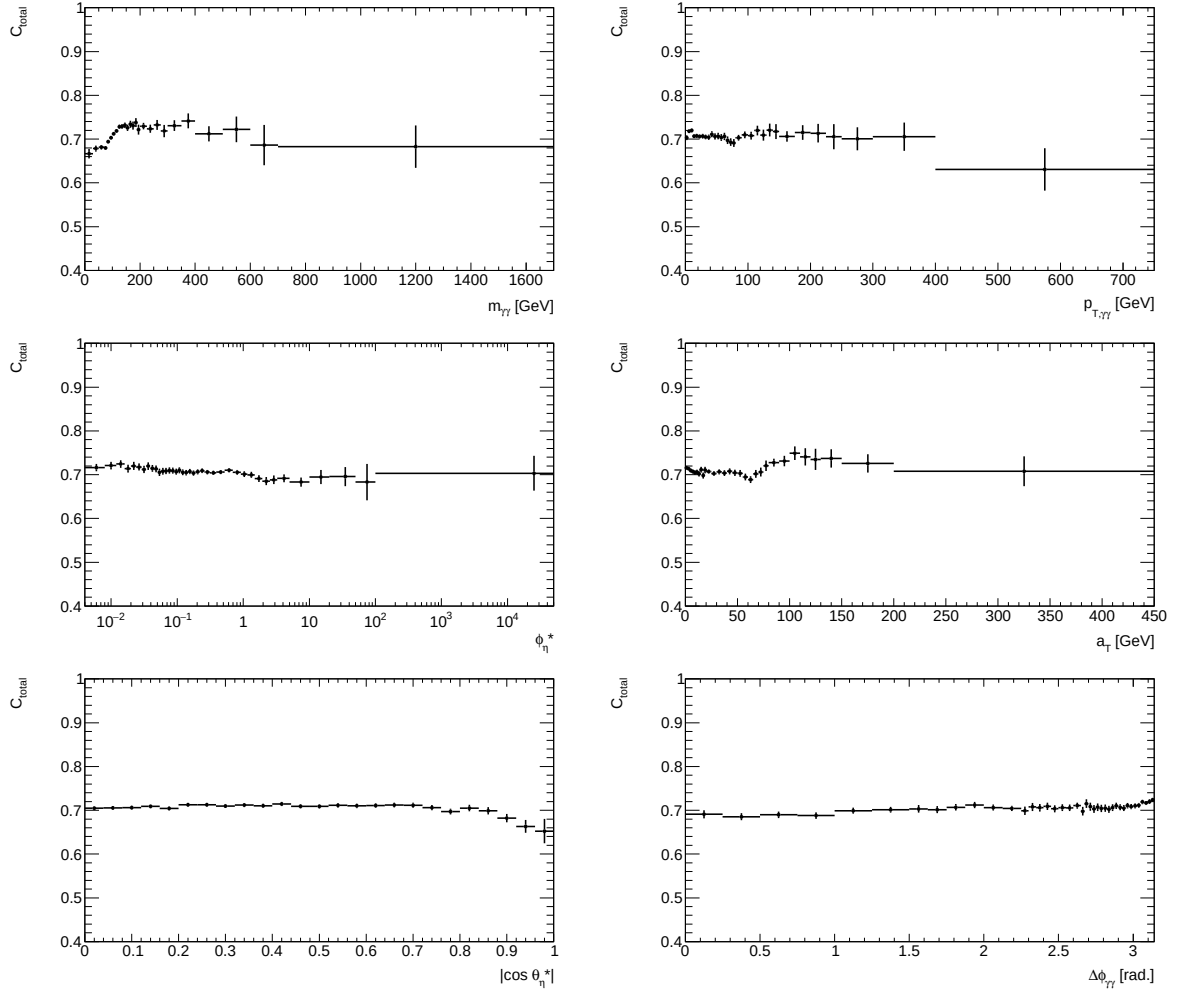


FIG. 137 – Facteurs de correction catégorie par catégorie (voir équation 76) pour les différentes observables de mesure, estimés avec l'échantillon de signal simulé SHERPA. Les barres d'erreur sont uniquement statistiques.

est possible. Au préalable, il est nécessaire d'introduire un certain nombre de définitions qui seront exploitées tout au long de l'étude. Celles-ci sont décrites ci-dessous :

◇ **Sélection fiducielle au niveau particule [*fid.(truth)*]** : détaillée dans la section 12.2, cette sélection s'applique uniquement aux quantités générées et définit le volume fiduciel de la mesure. Elle est rappelée brièvement ci-dessous :

- $|\eta_{\text{truth}}^\gamma| < 2.37$, avec l'exclusion de $1.37 < |\eta_{\text{truth}}^\gamma| < 1.56$.
- $E_{T,1}^{\gamma,\text{truth}} > 40 \text{ GeV}$, $E_{T,2}^{\gamma,\text{truth}} > 30 \text{ GeV}$.
- $E_T^{\text{part.iso.}} < 11 \text{ GeV}$ après hadronisation (niveau particule) dans un cône de taille fixe $\Delta R_{\text{truth}} = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ autour de chaque photon. La valeur de cette coupure sera discutée plus en détail dans la section 16.4.
- $\Delta R_{\gamma\gamma}^{\text{truth}} = \sqrt{\Delta\eta_{\gamma\gamma,\text{truth}}^2 + \Delta\phi_{\gamma\gamma,\text{truth}}^2} > 0.4$.
- Définition de la catégorie d'intérêt en fonction de la valeur des observables de mesure ($X < m_{\gamma\gamma} < Y$, etc) calculée au niveau particule.

◇ **Sélection fiducielle au niveau reconstruit [*fid.(reco.)*]** : la sélection fiducielle au niveau particule a été choisie de façon à être la plus proche possible de la sélection fiducielle au niveau reconstruit. Cette dernière correspond à la sélection fiducielle appliquée aux quantités reconstruites dans le détecteur.

- $|\eta^\gamma| < 2.37$, avec l'exclusion de $1.37 < |\eta^\gamma| < 1.56$.
- $E_{T,1} > 40 \text{ GeV}$, $E_{T,2} > 30 \text{ GeV}$.
- $E_T^{\text{calo.iso.}} < 6 \text{ GeV}$ dans un cône de taille fixe $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ autour de chaque photon. La valeur de cette coupure a été discutée dans les sections 10 et 14.3.2.
- $\Delta R_{\gamma\gamma} = \sqrt{\Delta\eta_{\gamma\gamma}^2 + \Delta\phi_{\gamma\gamma}^2} > 0.4$.
- Définition de la catégorie d'intérêt en fonction de la valeur des observables de mesure ($X < m_{\gamma\gamma} < Y$, etc) calculée au niveau reconstruit.

◇ **Coupsures de qualité (niveau reconstruit uniquement)**

[*quality selection*] : ces coupures supplémentaires sont appliquées au niveau reconstruit afin d'enrichir l'échantillon **TITI** en événements de signal $\gamma\gamma$ (voir section 14). Elles incluent :

- Vertex primaire : chaque événement doit inclure au moins un vertex primaire reconstruit avec trois traces chargées de $p_T > 400 \text{ MeV}$. Ce critère permet de rejeter le bruit de fond (très faible) qui n'est pas lié aux collisions (*beam-induced*) [242].
- Identification TIGHT : les deux candidats photons doivent passer la sélection d'identification TIGHT (voir section 14.1).

- Énergie d’isolation des traces : les deux candidats photons doivent avoir une énergie d’isolation des traces inférieure à 2.6 GeV (voir section 14.3.1).
 - La sélection sur le statut du système de déclenchement n’est pas appliquée pour cette étude lors de l’utilisation de l’échantillon MC SHERPA. En effet, l’efficacité de ce critère sera déterminée entièrement à partir des données dans la section 16.3.5.
- ◇ **Facteurs de correction catégorie par catégorie** : ces facteurs ont été définis dans la section 16.2 et entrent dans le calcul de la section efficace lors d’une déconvolution catégorie par catégorie (voir équation 78 et 77). Au cours de cette section, ils seront décomposés comme suit :

$$C_{total,i} = \frac{N_{i,TITI}^{reco}}{N_{i,fid.(truth)}^{truth}} = \frac{N_{fid.(reco) + quality\ selection}^{reco}}{N_{fid.(truth)}^{truth}} = \epsilon_{good\ \gamma\gamma,i} \cdot C_{fid,i} \cdot \epsilon_{TT,i} \cdot \epsilon_{II,i}, \quad (79)$$

où $\epsilon_{good\ \gamma\gamma,i}$ et $C_{fid,i}$ correspondent respectivement à l’efficacité de détection de deux bons photons dans le détecteur et aux corrections dues aux différences résiduelles entre la sélection fiducielle au niveau particule et au niveau reconstruit (voir les définitions ci-dessus), et $\epsilon_{TT,i}$ et $\epsilon_{II,i}$ sont respectivement l’efficacité d’identification TIGHT-TIGHT et isolé-isolé dans le détecteur interne pour les candidats photons de la paire. Comme nous le verrons, la sélection sur le vertex primaire d’interaction a une efficacité très proche de 100% et est par conséquent négligée dans la suite. Les études séparées de $\epsilon_{good\ \gamma\gamma,i}$, $C_{fid,i}$, $\epsilon_{TT,i}$ et $\epsilon_{II,i}$ seront détaillées dans les sections 16.3.2 et 16.3.3. Leur non-corrélation, nécessaire afin d’étudier ces quantités indépendamment, sera vérifiée dans la section 16.3.4.

- ◇ **Corrections d’acceptance** : formées de $\epsilon_{good\ \gamma\gamma,i}$ et $C_{fid,i}$, l’étude des corrections d’acceptance requiert absolument un échantillon simulé de signal afin d’avoir accès aux quantités générées. Ces corrections sont issues d’une part de diverses inefficacités de détection/identification basiques du détecteur ($\epsilon_{good\ \gamma\gamma,i}$), ayant lieu par exemple lorsque un photon passe au travers de cellules mortes du détecteur ou lorsque sa gerbe électromagnétique est trop proche de celle d’un jet (identification LOOSE non satisfaite), et d’autre part de différences résiduelles entre la sélection fiducielle au niveau particule et au niveau reconstruit ($C_{fid,i}$). En effet, à cause de la résolution limitée du détecteur, il arrive qu’un photon généré hors du volume fiduciel soit reconstruit à l’intérieur de celui-ci et vice-versa. En particulier, les sélections fiducielles d’isolation au niveau particule et au niveau reconstruit présentent des différences importantes (voir section 16.4), engendrant une faible “résolution” sur cette quantité. Les définitions exactes et l’étude complète de $\epsilon_{good\ \gamma\gamma,i}$ et $C_{fid,i}$ seront introduites dans la section 16.3.2.
- ◇ **Efficacités des coupures de qualité** : formées de $\epsilon_{TT,i}$ et $\epsilon_{II,i}$, les efficacités des coupures de qualité peuvent être évaluées à partir de la simulation mais également à partir d’échantillons de photons purs dans les données car elles nécessitent uniquement les quantités reconstruites. Pour cela, il est nécessaire de démontrer qu’elles se factorisent des effets d’acceptance et entre elles (voir section 16.3.4). Les critères TIGHT-TIGHT et isolé-isolé dans le détecteur interne permettent de

rejeter une quantité importante de bruit de fond au prix d'une efficacité légèrement abaissée (voir sections 9 et 10). Ces coupures sont difficilement applicables au niveau générateur pour des raisons techniques. La sélection sur le vertex primaire d'interaction a une efficacité très proche de 100% et est négligée dans la suite.

Dans la section 16.3.2, les corrections d'acceptance sont abordées en détail.

16.3.2 Corrections d'acceptance

Efficacité de détection de deux bons photons

L'efficacité de détection de deux bons photons dans le détecteur à partir d'événements de signal $\gamma\gamma + X$, notée $\epsilon_{good\ \gamma\gamma,i}$, est définie ci-dessous :

$$\epsilon_{good\ \gamma\gamma} = \frac{N_{fid.(truth)+2\ good\ candidates}^{truth}}{N_{fid.(truth)}^{truth}}, \quad (80)$$

où $N_{fid.(truth)}^{truth}$ est le nombre d'événements générés passant la sélection fiducielle au niveau généré. Au numérateur, deux bons candidats photons correspondant aux photons générés en termes d'angle d'azimut et de pseudo-rapacité sont demandés, un bon candidat photon étant défini comme un photon passant la présélection des photons détaillée dans la section 14.5. Celle-ci vise à obtenir un échantillon au niveau reconstruit avec une pureté acceptable tout en maintenant une efficacité aussi haute que possible. La quasi-intégralité des critères de présélection pour les photons constitue les recommandations de base du groupe de performance e/γ pour les analyses photons 2012 dans ATLAS, et donc par conséquent difficilement modifiable par les groupes d'analyses.

$\epsilon_{good\ \gamma\gamma,i}$ considère le nombre d'événements générés au numérateur et au dénominateur, ce qui lui permet de s'affranchir en grande partie des effets reliés aux différences résiduelles entre la sélection fiducielle au niveau particule et au niveau reconstruit.⁷⁷ Ainsi, $\epsilon_{good\ \gamma\gamma,i}$ est attendu < 1 et comme une mesure de l'efficacité de détection de deux photons avec des critères de qualité minimum afin d'atteindre une pureté acceptable. Pour la région **TITI** complète, $\epsilon_{good\ \gamma\gamma,i} = 0.885$ avec une erreur statistique négligeable. Les distributions selon $E_{T,1}$, $E_{T,2}$, η_1^γ et η_2^γ sont représentées dans la figure 138, tandis que celles correspondant aux différentes observables de mesure sont représentées dans la figure 139.

Ainsi, au moins 10% des événements de signal sont perdus dès la phase de présélection alors que la reconstruction et l'efficacité d'identification LOOSE après reconstruction pour un photon unique devrait se situer entre 97 et 100% (voir section 9.2). Les autres coupures entrant dans la présélection (résolveur d'ambiguïté, acceptance en pseudo-rapacité, nettoyage des données et qualité de l'objet) jouent donc un rôle non-négligeable sur la valeur des facteurs de correction catégorie par catégorie. Par ailleurs, $\epsilon_{good\ \gamma\gamma,i}$ semble baisser légèrement à haut E_T^γ , de l'ordre de 2 à 3%. Cet effet est connu

⁷⁷La sélection d'acceptance en $|\eta^\gamma|$ est néanmoins demandée pour $\epsilon_{good\ \gamma\gamma,i}$ afin d'être en mesure d'appliquer la sélection d'identification LOOSE.

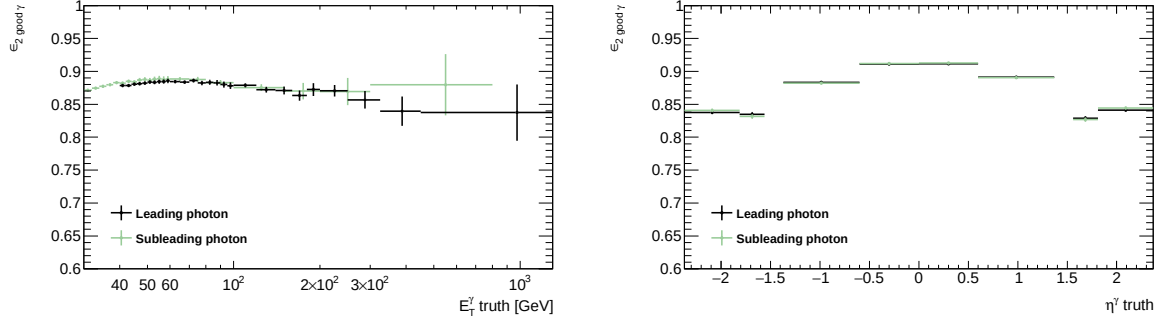


FIG. 138 – Efficacité de détection de deux bons photons dans le détecteur en fonction de l'impulsion transverse (gauche) et la pseudo-rapidité (droite) du premier (ligne continue noire) et deuxième (ligne continue verte) photon immédiat générés dans l'échantillon de signal simulé SHERPA. Les barres d'erreur représentent uniquement l'incertitude statistique du MC.

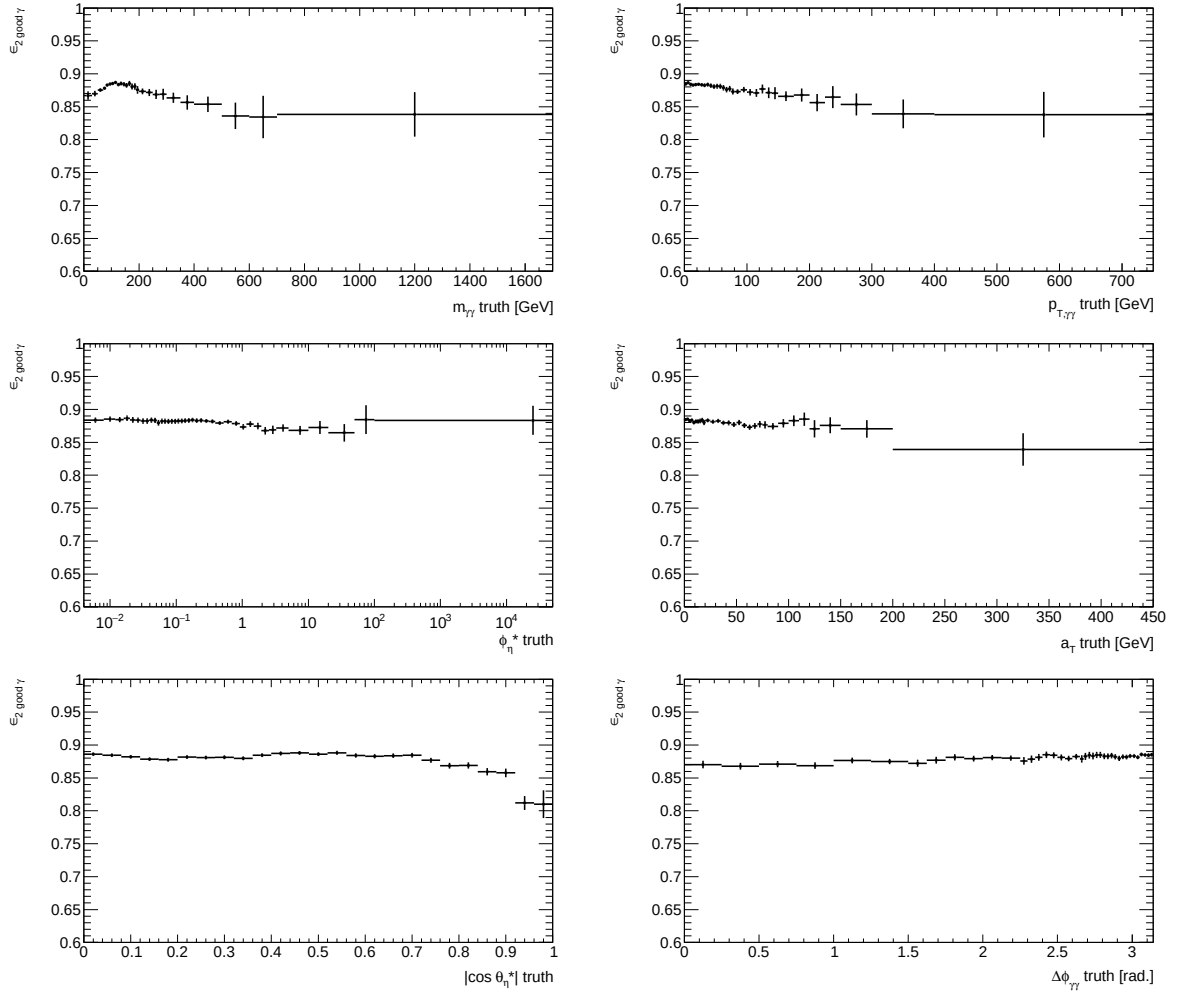


FIG. 139 – Figure similaire à la figure 138 pour les différentes observables de mesure di-photon.

dans ATLAS [200] et vient de la difficulté croissante à distinguer un photon converti d'un électron à haut E_T^γ , à cause des limitations de reconstruction du détecteur interne pour les traces très proches les unes des autres. $p_{T,\gamma\gamma}$ et a_T , dominés par les effets liés à $E_{T,1}$ et $E_{T,2}$, suivent la même tendance.

Il est également intéressant de remarquer que $m_{\gamma\gamma}$ et $|\cos \theta_\eta^*|$ baissent de plusieurs pourcents à partir de $m_{\gamma\gamma} > 200$ GeV et $|\cos \theta_\eta^*| > 0.75$, respectivement. Il est possible de démontrer que l'effet dominant est géométrique : à haute masse et $|\cos \theta_\eta^*|$, une part importante des photons est reconstruite dans les bouchons, qui présentent des performances de détection minimum significativement plus faibles que le tonneau (voir figure 138, gauche). Finalement, les distributions de ϕ_η^* et $\Delta\phi_{\gamma\gamma}$ présentent des distributions en $\epsilon_{good \gamma\gamma,i}$ plates dans la statistique considérée.

Différences résiduelles entre la sélection fiducielle au niveau particule et au niveau reconstruit

Les différences résiduelles entre la sélection fiducielle au niveau particule et au niveau reconstruit sont étudiées grâce au facteur $C_{fid,i}$, défini de la façon suivante :

$$C_{fid} = \frac{N_{2 \text{ good candidates}+fid.(reco)}^{reco}}{N_{fid.(truth)+2 \text{ good candidates}}^{truth}}, \quad (81)$$

où le dénominateur correspond au numérateur de $\epsilon_{good \gamma\gamma,i}$ présenté au paragraphe précédent et $N_{2 \text{ good candidates}+fid.(reco)}^{reco}$ est le nombre d'événements **reconstruits** ayant deux bons photons reconstruits et passant la sélection fiducielle au niveau **reconstruit**.

$C_{fid,i}$ peut être supérieur ou inférieur à 1 en fonction du bilan net des migrations (entrantes + sortantes). Pour la région **TITI** complète, $C_{fid,i} = 0.953$. Les distributions selon $E_{T,1}$, $E_{T,2}$, η_1^γ et η_2^γ sont représentées dans la figure 140, tandis que celles correspondant aux différentes observables de mesure sont représentées dans la figure 141. $C_{fid,i}$ est systématiquement inférieur à 1 principalement à cause de la mauvaise correspondance entre la coupure d'isolation appliquée au niveau particule et au niveau reconstruit (voir section 16.4), les autres migrations restant très réduites (voir section 16.2). Les distributions restent néanmoins globalement très plates, ce qui montre la très bonne résolution du détecteur et l'emploi de catégories adaptées dans l'analyse.

$C_{fid,i}$ baisse légèrement à grand E_T^γ (> 200 GeV) à cause de la translation des pics des distributions de l'isolation calorimétrique $E_T^{\text{calo.iso.}}$ à haut E_T^γ dans les données et du choix d'une coupure fixe (< 6 GeV). Cet effet a d'ailleurs nécessité le calcul de corrections supplémentaires pour que le MC reproduise ce comportement, représentant le travail de performance principal effectué pendant la thèse (voir section 10.4) [201].⁷⁸ $p_{T,\gamma\gamma}$, qui est l'observable la plus sensible à $E_{T,1}$ et $E_{T,2}$, est légèrement affectée par cet effet.

En revanche, $C_{fid,i}$ est globalement constant selon $|\eta^\gamma|$, sauf à proximité immédiate de la région de transition entre tonneau et bouchons exclue de l'analyse ($1.37 < |\eta^\gamma| < 1.56$), où une hausse de 3% est observée ($1.56 < |\eta^\gamma| < 1.81$). Cette hausse est

⁷⁸C'est bien le MC qui est corrigé afin de reproduire la dépendance en E_T^γ observée dans les données et non l'inverse.

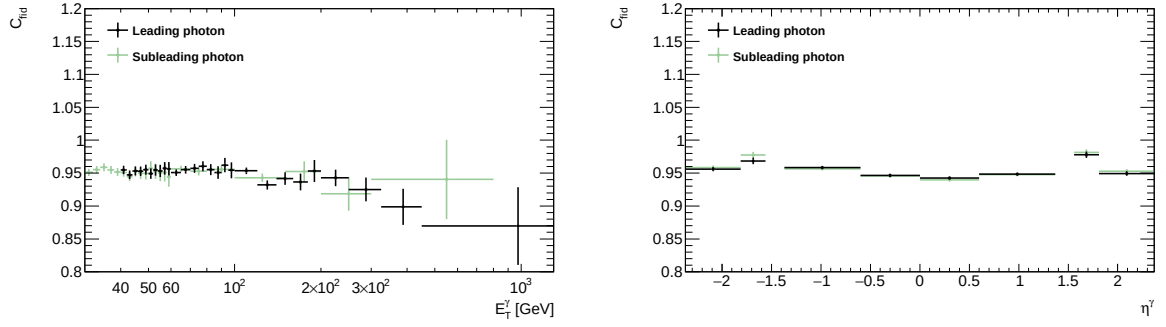


FIG. 140 – Différences résiduelles entre la sélection fiducielle au niveau particule et au niveau reconstruit (C_{fid}) en fonction de l'impulsion transverse (gauche) et la pseudo-rapidité (droite) du premier (ligne continue noire) et deuxième (ligne continue verte) photon immédiat généré dans l'échantillon de signal simulé SHERPA. Les barres d'erreur représentent uniquement l'incertitude statistique du MC.

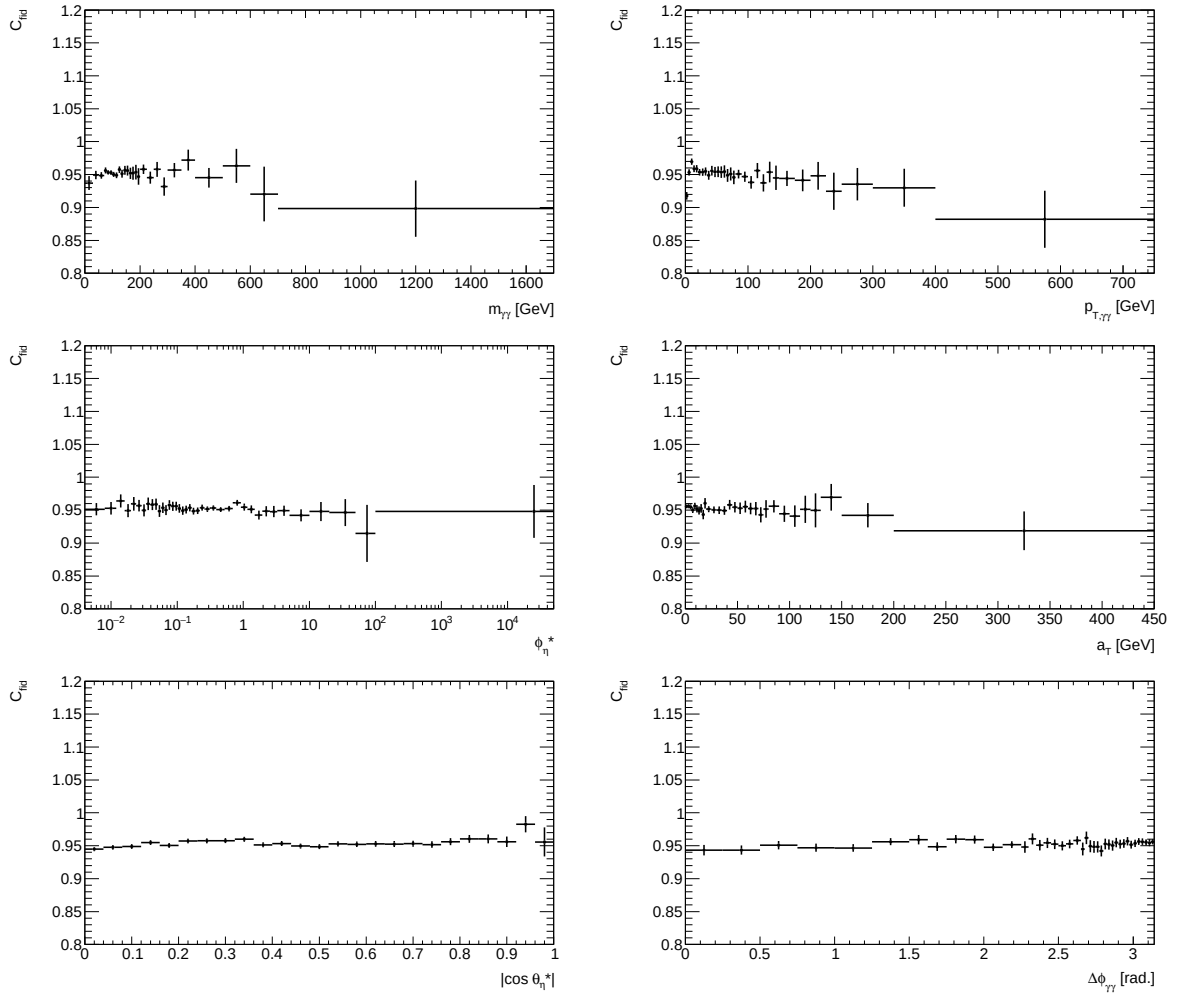


FIG. 141 – Figure similaire à la figure 140 pour les différentes observables de mesure di-photon.

également présente dans $0.6 < |\eta^\gamma| < 1.37$ mais moins visible à cause de la section efficace plus grande observée dans cette catégorie. Elle correspond à une augmentation locale des migrations dues aux événements générés dans $1.37 < |\eta^\gamma| < 1.56$ mais reconstruit dans l'acceptance par des effets de résolution.⁷⁹

Finalement les légères variations visibles à petit $m_{\gamma\gamma}$ et $p_{T,\gamma\gamma}$ proviennent de la présence d'un pic asymétrique prononcé dans les distributions correspondantes (voir par exemple la figure 108, section 15.5.2), augmentant les migrations à cause des effets de résolution. Néanmoins, toutes les fluctuations restent sous contrôle, validant à nouveau les largeurs des différentes catégories. a_T , ϕ_η^* et $\Delta\phi_{\gamma\gamma}$ restent plats, conformément aux attentes. L'effet systématique le plus significatif sur $C_{fid,i}$ est relié aux incertitudes sur les corrections supplémentaires d'isolation et sera évalué dans le cadre de l'analyse dans la section 16.6.

Afin d'avoir une idée plus complète des migrations entre les différentes catégories, il est également intéressant d'étudier leur pureté (*bin purity*, à ne pas confondre avec la pureté de signal $\gamma\gamma$), qui correspond à C_{fid} avec en plus l'obligation pour les événements reconstruits dans une certaine catégorie d'avoir été générés dans la même catégorie :

$$\begin{aligned} C_{bin\ purity} &= C_{fid} \cdot \frac{N_{2\ good\ candidates+fid.(reco)+same\ bin\ than\ generation}^{reco}}{N_{2\ good\ candidates+fid.(reco)}^{reco}} \\ &= \frac{N_{2\ good\ candidates+fid.(reco)+same\ bin\ than\ generation}^{reco}}{N_{fid.(truth)+2\ good\ candidates}^{truth}}. \end{aligned} \quad (82)$$

Ainsi, C_{fid} représente le résultat net des migrations (entrantes + sortantes) tandis que $C_{bin\ purity}$ indique la fraction d'événement n'ayant pas migré pour chacune des catégories considérées. $C_{bin\ purity}$ est donc nécessairement inférieur à 1. Pour la région **TITI** complète, $C_{bin\ purity} = 0.903$. Les distributions correspondant aux différentes observables de mesure sont représentées dans la figure 142. $C_{bin\ purity}$ est compris entre 70% et 90% pour l'ensemble des catégories considérées pour la mesure, ce qui valide définitivement le choix des catégories.

Les distributions correspondant à $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T et $\Delta\phi_{\gamma\gamma}$ utilisent des catégories dont la largeur varie par palier afin de garder une compatibilité maximum avec celles utilisées dans la mesure $\gamma\gamma + X$ précédente à $\sqrt{s} = 7$ TeV d'ATLAS [203].⁸⁰ Cela a pour effet secondaire la création de paliers décroissants ($m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T) ou constants ($\Delta\phi_{\gamma\gamma}$) pour la pureté des différentes catégories. En effet, la résolution en énergie du calorimètre est proportionnelle à un certain pourcentage de l'énergie du candidat photon approximativement constant (de l'ordre de 1 à 4%, voir section 8). Ainsi, à largeur de catégorie égale, la pureté baisse à plus haut $E_{T,1}$, $E_{T,2}$, affectant dans le même temps les observables qui leur sont corrélées ($m_{\gamma\gamma} < 200$ GeV, $p_{T,\gamma\gamma}$, a_T). En revanche, la résolution angulaire est approximativement constante en valeur absolue (et non en

⁷⁹La situation inverse, *ie.* événements générés dans l'acceptance mais reconstruits dans $1.37 < |\eta^\gamma| < 1.56$, est exclue de $C_{fid,i}$ car la sélection sur $|\eta^\gamma|$ au niveau reconstruit est appliquée dès $\epsilon_{good\ \gamma\gamma,i}$ afin d'être en mesure de demander une identification LOOSE (voir paragraphe précédent).

⁸⁰Malheureusement, aucune comparaison directe entre les deux mesures n'a été réalisée dans le cadre de l'analyse à $\sqrt{s} = 8$ TeV à cause des différents volumes fiduciels considérés dans les deux analyses (voir section 12.2).

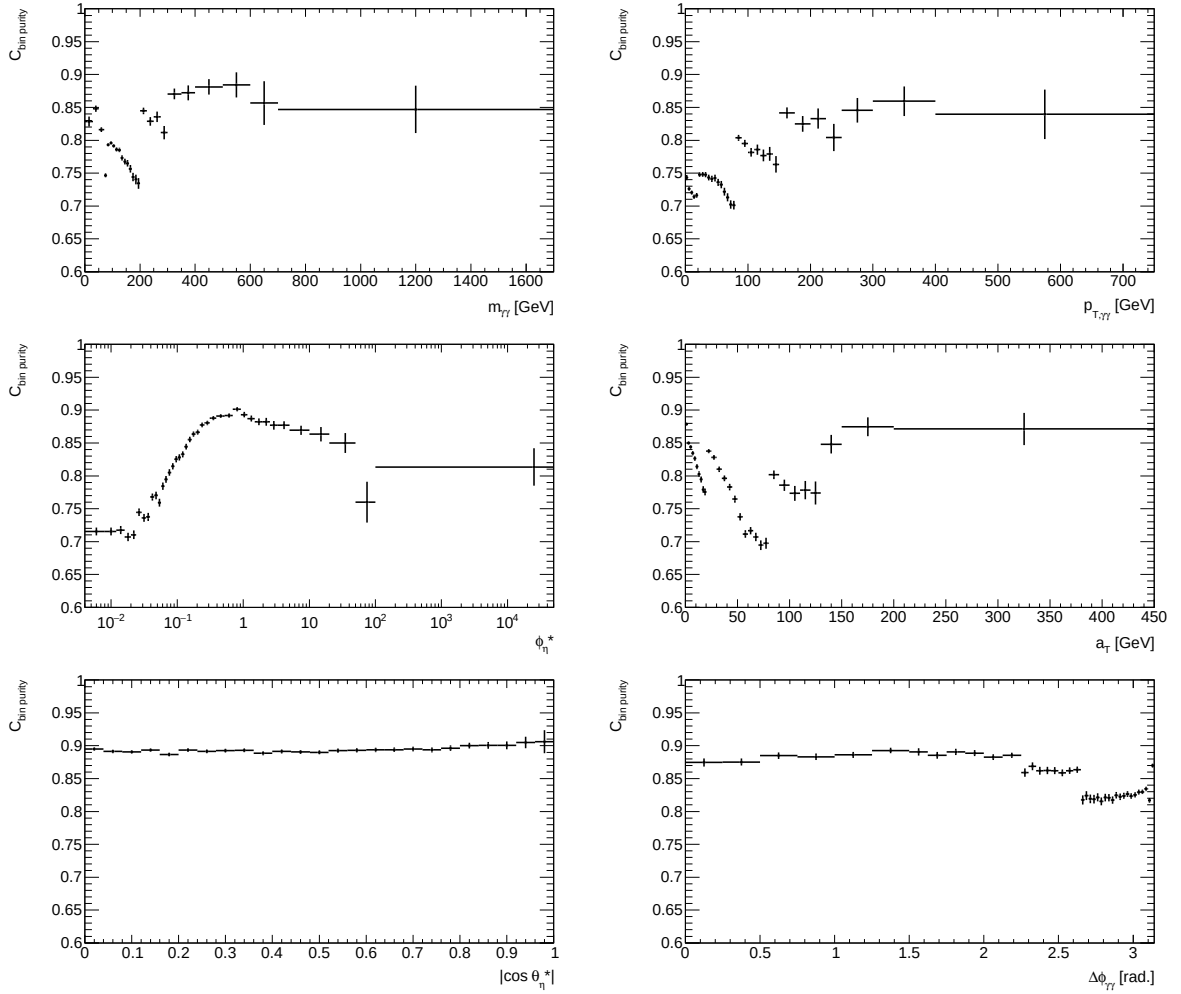


FIG. 142 – Pureté des différentes catégories (voir équation 82, à ne pas confondre avec pureté de signal $\gamma\gamma$) pour les différentes observables de mesure di-photon estimée à partir de l'échantillon de signal SHERPA simulé. Les barres d'erreur représentent uniquement l'incertitude statistique du MC.

%), ce qui mène à des paliers constants ($\Delta\phi_{\gamma\gamma}$). Dans les deux cas, lorsque la largeur de la catégorie augmente, la pureté augmente naturellement. Cet effet n'est pas visible pour ϕ_{η}^* , qui est décomposé en catégories présentant des largeurs progressives,⁸¹ ni pour $|\cos \theta_{\eta}^*|$ qui est décomposé dans des catégories de largeur fixe (résolution et statistique approximativement constantes pour l'ensemble du spectre). Pour $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, ϕ_{η}^* et a_T , l'augmentation des largeurs des catégories est donc nécessaire à la fois pour des questions de statistique et des questions de résolution. La stabilisation de la pureté de $m_{\gamma\gamma}$ pour $m_{\gamma\gamma} > 200$ GeV indique une prépondérance des effets angulaires à haute masse, déjà observée pour $\epsilon_{good\ \gamma\gamma}(m_{\gamma\gamma})$.

Finalement, la catégorie $70\text{ GeV} < m_{\gamma\gamma} < 80\text{ GeV}$ présente une pureté significative-ment plus faible que les catégories voisines car elle est placée au début du pic asymétrique de la distribution (créé par la sélection en $E_{T,1}$, $E_{T,2}$), ce qui a pour effet une augmentation significative des migrations. La bonne résolution attendue des variables a_T et ϕ_{η}^* comparées à $p_{T,\gamma\gamma}$ dans la région $p_{T,\gamma\gamma}/a_T/\phi_{\eta}^* \simeq 0$ sensible aux divergences infrarouges

⁸¹ ϕ_{η}^* n'a pas été mesuré à $\sqrt{s} = 7$ TeV.

(voir section 12.1) est confirmée. La forme de la distribution en pureté de ϕ_η^* est difficile à interpréter plus en détail à cause des catégories de largeurs progressives adoptées et de la forme logarithmique de la variable (voir section 12.3).

Dans la section 16.3.3, les efficacités des différentes coupures de qualité appliquées dans l'analyse sont abordées.

16.3.3 Efficacités des coupures de qualité

Sélection du vertex primaire de l'événement

Dans l'analyse principale, seuls les événements comportant au moins un vertex primaire reconstruit avec au moins trois traces chargées de $p_T > 400$ MeV sont considérés. Cette sélection permet de supprimer le bruit de fond qui n'est pas relié aux collisions (*beam-induced*) [242]. Celui-ci devrait cependant être très faible dans les données à $\sqrt{s} = 8$ TeV grâce au chargement élevé de la machine pendant les différentes périodes d'acquisition, limitant ainsi le nombre de croisements sans collisions (voir section 4.3). L'efficacité de cette coupure dans l'échantillon MC de signal SHERPA s'élève à 99.94%. Dans les données, elle mène au rejet de moins de 0.2% des événements (sans distinction entre les différents processus). Par conséquent, les effets de l'efficacité de cette coupure sur les facteurs de correction catégorie par catégorie sont négligés.

L'ensemble des résultats de déconvolution finals seront néanmoins corrigés de l'efficacité de cette coupure en se basant sur l'échantillon de signal SHERPA. Les différences potentielles entre données et MC seront négligées (voir section 16.5).

Sélection de photons TIGHT-TIGHT

ϵ_{TT} est défini par la formule suivante :

$$\epsilon_{TT} = \frac{N_{2 \text{ good candidates}+fid.(reco)+TT}^{reco}}{N_{2 \text{ good candidates}+fid.(reco)}^{reco}}, \quad (83)$$

Les différentes sélections d'identification pour les photons ont déjà été décrites en détail dans les sections 9 et 14.1. L'efficacité d'identification TIGHT dans les échantillons de signal simulés est corrigée à l'aide des *ID scale factors* fournis par le groupe de performance afin de correspondre aux données. Ces facteurs d'échelle sont basés sur trois mesures indépendantes de l'efficacité TIGHT dans les données.

Par conséquent, seules des vérifications élémentaires sont effectuées dans le cadre de l'analyse. Pour cela, l'efficacité d'identification TIGHT-TIGHT (par rapport à LOOSE-LOOSE, critère déjà demandé lors de la présélection) est représentée dans les figures 143 et 144 pour $E_{T,1}$, $E_{T,2}$, η_1^γ , η_2^γ et les différentes observables de mesure. Si les identifications des deux photons sont non-corrélées, l'efficacité TIGHT-TIGHT correspond au produit des efficacités individuelles. À haut E_T^γ (> 40 GeV) l'efficacité LOOSE est de 100% de telle sorte que les efficacités représentées dans les figures 143 et 144 peuvent être comparées directement aux efficacités représentées dans la section 9.2, figures 60 et 61 (qui sont évaluées par rapport à la reconstruction).

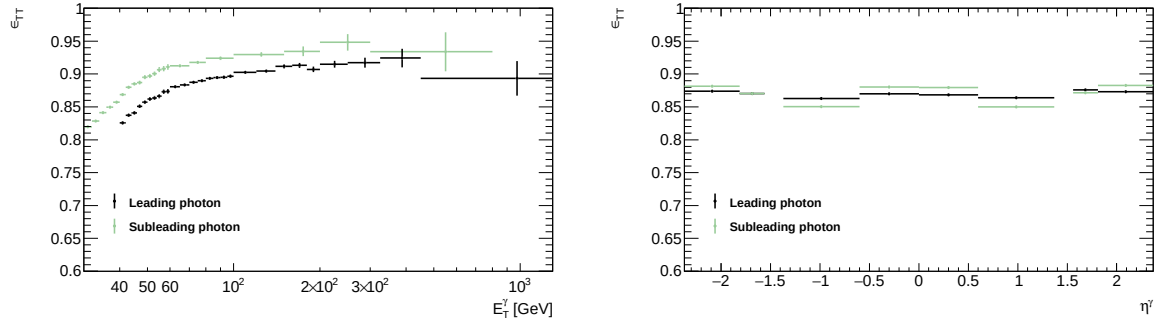


FIG. 143 – Efficacité d'identification TIGHT-TIGHT en fonction de l'impulsion transverse (gauche) et de la pseudo-rapidité (droite) du premier (ligne noire pleine) et du deuxième (ligne verte pleine) photon. L'échantillon de signal simulé SHERPA est utilisé avec les *ID scale factors*. Les barres d'erreur représentent uniquement l'incertitude statistique.

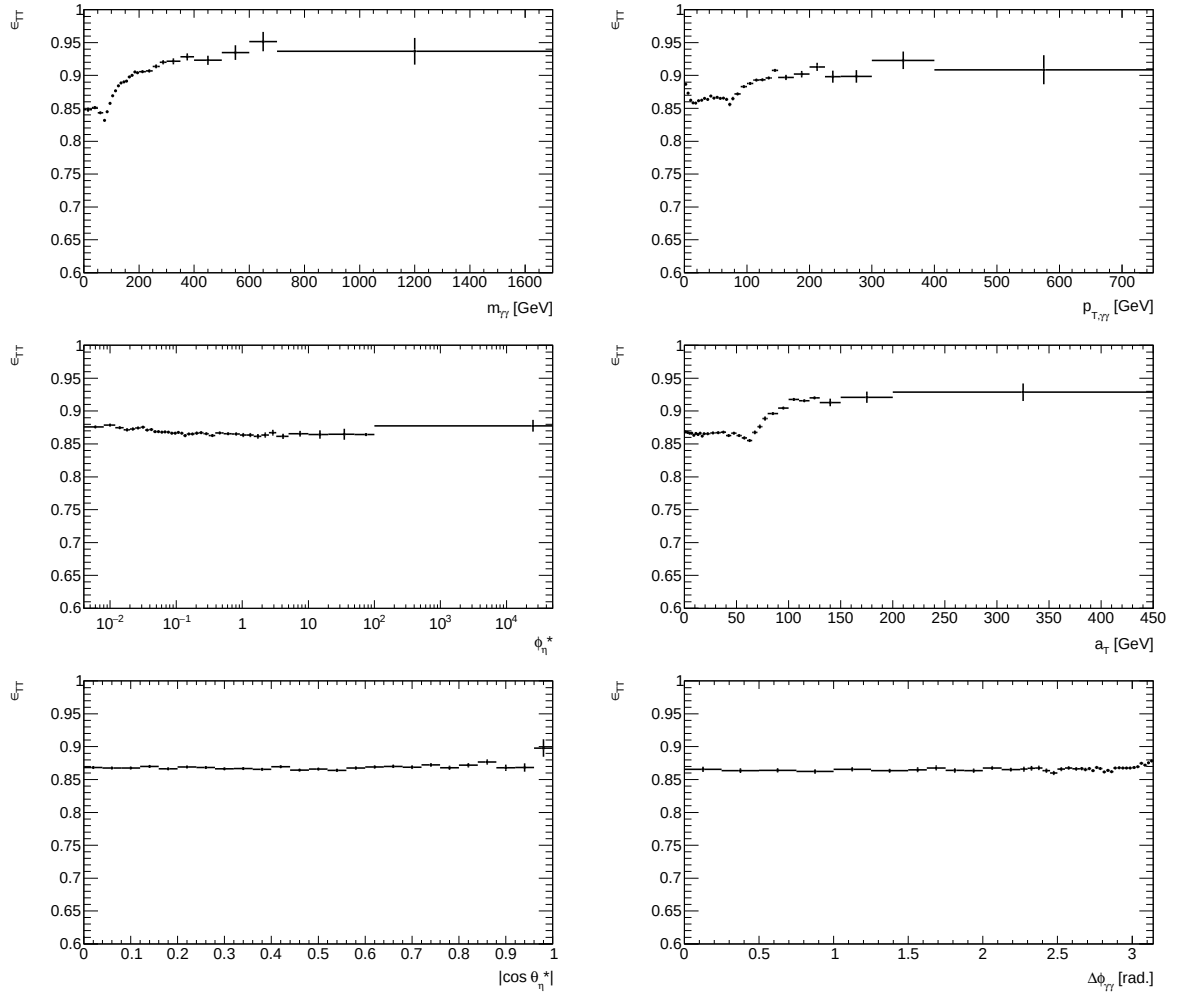


FIG. 144 – Figure similaire à la figure 143 pour les différentes observables de mesure.

Comme attendu, l'efficacité TIGHT-TIGHT est plate en $|\eta^\gamma|$, les sélections d'identification étant optimisées par le groupe de performance pour cela. Par ailleurs, le plateau en E_T^γ correspond à environ 95%, compatible avec celui présenté dans les figures 60 et 61 pour les photons uniques ($\simeq 97 - 98\%$). Finalement, la courbe correspondant à $E_{T,1}$ est plus faible que celle correspondant à $E_{T,2}$ à cause de l'intégration sur l'autre candidat, systématiquement à plus basse (haute) impulsion transverse pour $E_{T,1}$ ($E_{T,2}$).

Les distributions de ϵ_{TT} selon les différentes observables de mesure suivent des comportements légèrement différents en fonction de leur corrélation plus ou moins forte vis-à-vis de $E_{T,1}$, $E_{T,2}$, η_1^γ , η_2^γ . ϵ_{TT} atteint le plateau pour $m_{\gamma\gamma} > 200$ GeV environ et $p_{T,\gamma\gamma}$ et $a_T > 120$ GeV. ϕ_η^* , $|\cos \theta_\eta^*|$ et $\Delta\phi_{\gamma\gamma}$ sont dominés par des effets angulaires de telle sorte que ϵ_{TT} reste approximativement constant pour ces observables.

D'autre part, le court plateau observé à basse masse est une conséquence de la sélection en $E_{T,1}$, $E_{T,2}$ appliquée dans l'analyse. En effet, la plupart des événements ayant une masse inférieure à 70 GeV ont des photons ayant des impulsions transverses proches des coupures (40 GeV, 30 GeV, respectivement) et diffèrent uniquement par leur distribution angulaire dans le plan transverse ($\Delta\phi_{\gamma\gamma}$),⁸² observable présentant un ϵ_{TT} plat. a_T et $p_{T,\gamma\gamma}$ présentent un pic négatif (-1 à 2%) au niveau de l'épaule de Guillet, montrant que les photons de fragmentation ont un ϵ_{TT} significativement plus faible. Un pic similaire est également visible pour $m_{\gamma\gamma}$, qui peut également être attribué aux photons de fragmentation qui devraient être plus nombreux pour $m_{\gamma\gamma} < 70$ GeV.

Dans le paragraphe suivant, l'efficacité de la sélection correspondant à deux photons isolés dans le détecteur interne est abordée.

Sélection de photons isolé-isolé dans le détecteur interne

ϵ_{II} est défini par la formule suivante :

$$\epsilon_{II} = \frac{N_{2 \text{ good candidates} + fid.(reco) + TITI}^{reco}}{N_{2 \text{ good candidates} + fid.(reco) + TT}^{reco}}. \quad (84)$$

Contrairement aux différentes sélections d'identification, l'efficacité de sélection sur l'isolation des traces n'a pas été étudiée en détail par le groupe de performance. En particulier, aucun facteur de correction générique n'existe pour faire correspondre l'efficacité du MC aux données. Cette sélection correspond à une coupure additionnelle comparée aux analyses $\gamma\gamma + X$ précédentes d'ATLAS [202, 203], employée jusqu'à maintenant principalement dans les études des désintégrations du boson de Higgs en deux photons [43, 188, 208].

Par conséquent, en plus de l'estimation d' ϵ_{II} avec l'échantillon de signal SHERPA simulé, une seconde estimation basée sur les données est effectuée pour estimer l'incertitude du MC. Pour cela, une seconde soustraction du bruit de fond est réalisée avec la méthode du *template fit* 2D (voir section 15) sans demander de sélection particulière sur l'isolation des traces des deux photons, ce qui revient à redéfinir **I** comme isolé dans le calorimètre uniquement dans toute l'analyse. Le rapport des nombres d'événements

⁸² $m_{\gamma\gamma}^2 = 2E_{T,1}^\gamma E_{T,2}^\gamma [\cosh(\eta_1^\gamma - \eta_2^\gamma) - \cos(\phi_1 - \phi_2)]$

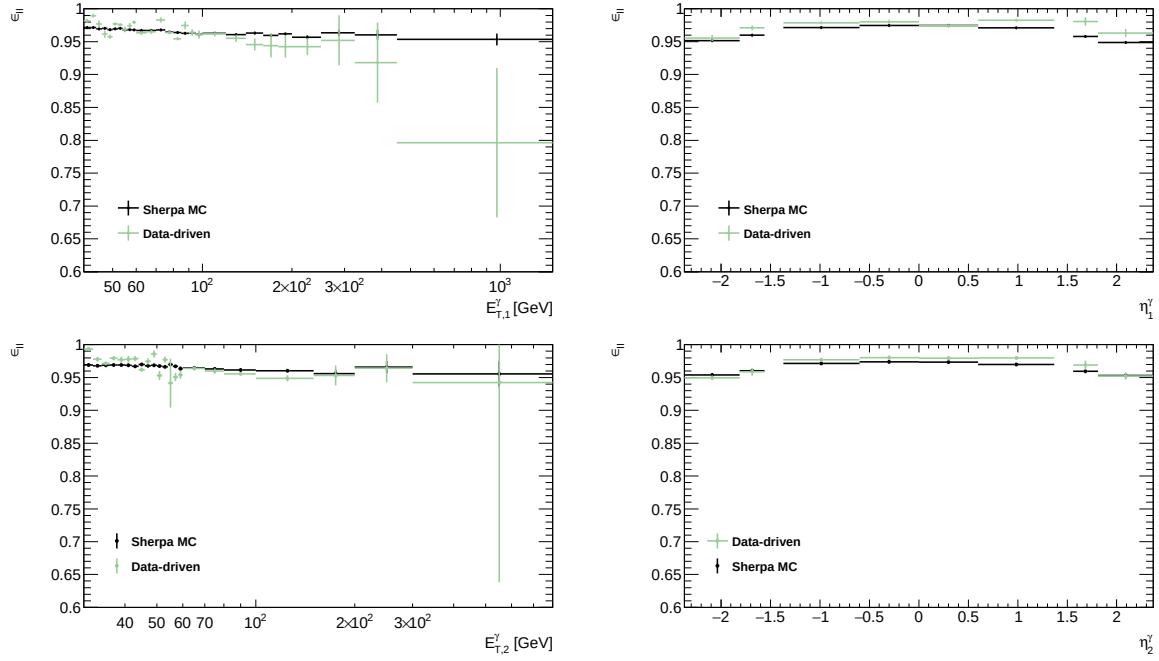


FIG. 145 – Efficacité de sélection de deux photons isolés dans le détecteur interne en fonction de l’impulsion transverse (gauche) et de la pseudo-rapacité (droite) du premier (haut) et du deuxième (droite) photon. L’échantillon de signal simulé SHERPA est utilisé (ligne noire pleine) ainsi que les données directement (ligne verte pleine). Les barres d’erreur représentent uniquement l’incertitude statistique.

$\gamma\gamma$ attendus dans les deux cas de figure (analyse nominale et analyse sans sélection sur l’isolation des traces) est alors utilisé pour fournir une estimation de ϵ_{II} basée sur les données. Dans le cas inclusif, l’efficacité isolé-isolé dans le détecteur interne est estimée respectivement à 96.6% et 96.3% avec l’échantillon SHERPA et les données, et donc en excellent accord. Les différentes estimations en fonction de $E_{T,1}^\gamma$, $E_{T,2}^\gamma$, η_1^γ , η_2^γ et selon les différentes observables de mesure sont représentées dans les figures 145 et 146. Uniquement l’incertitude statistique est représentée (erreur binomiale).

Dans l’ensemble, les deux techniques d’estimation d’ ϵ_{II} donnent des résultats compatibles dans les barres d’erreur, la précision étant généralement limitée par la statistique des données. L’incertitude sur ϵ_{II} dans le MC est estimée en considérant la somme en quadrature de l’incertitude statistique MC et de la différence entre l’estimation MC et l’estimation basée sur les données (symétrisée). Cette méthode est considérée comme conservatrice car le deuxième terme de la somme inclut très probablement une partie des fluctuations statistiques des données.

Ces fluctuations statistiques engendrant des fluctuations importantes sur l’incertitude systématique de ϵ_{II} , un lissage de l’erreur sera appliqué pour la mesure finale (voir section 16.6). La technique de lissage utilisée est très similaire à celle déjà évoquée lors de l’évaluation de l’incertitude systématique liée au choix du générateur lors de la soustraction du bruit de fond (voir section 15.6.1) et nécessite une estimation de “l’erreur sur l’erreur”. Celle-ci est estimée en faisant varier le deuxième terme de l’incertitude systématique sur ϵ_{II} (voir paragraphe précédent) en prenant la différence entre la valeur centrale du MC et la valeur à $\pm 1\sigma$ (*stat*) de l’estimation basée sur les données au lieu de sa valeur centrale. L’incertitude finale sur ϵ_{II} , de l’ordre de 0 à 3-4%, jouera finalement

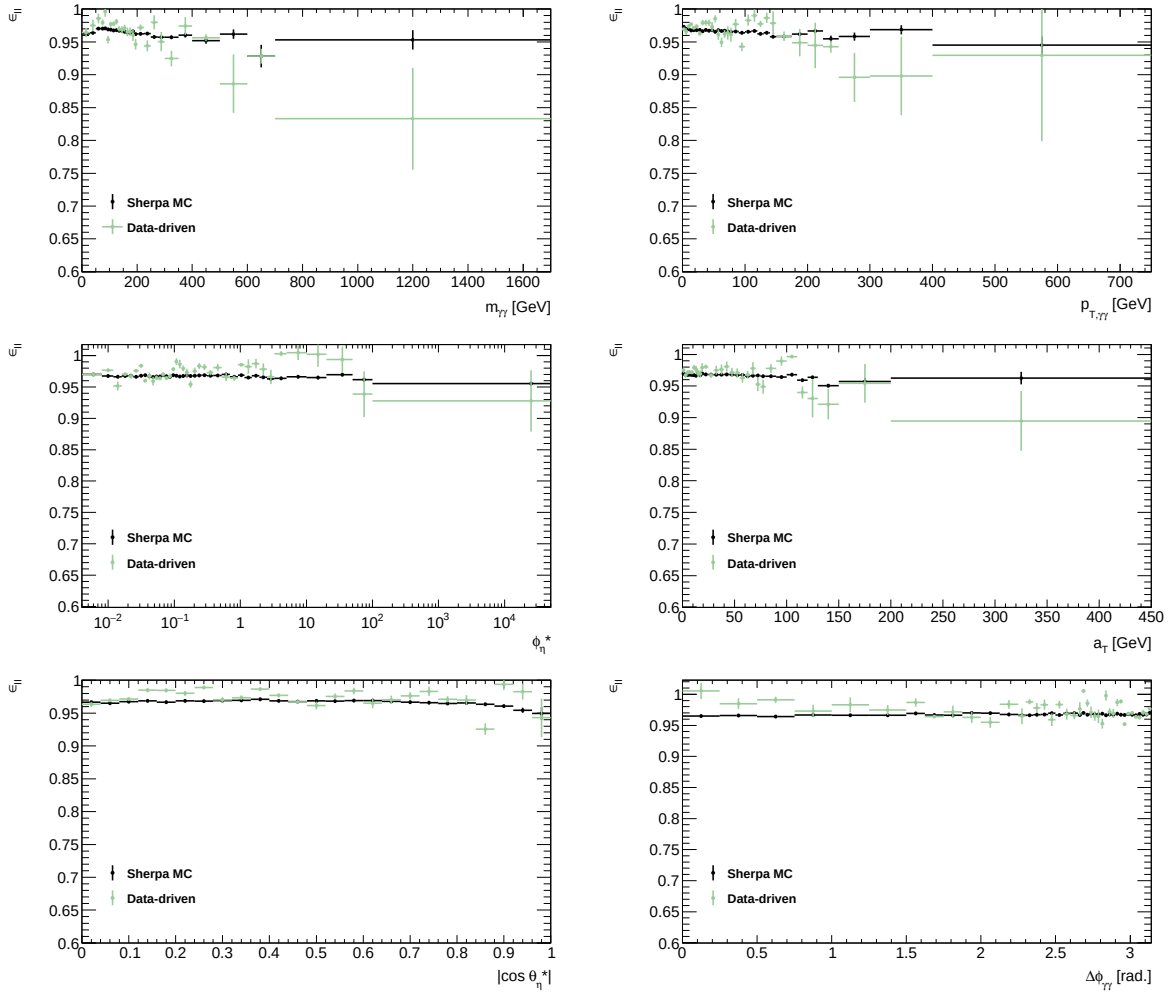


FIG. 146 – Figure similaire à la figure 145 pour les différentes observables de mesure.

un rôle très mineur sur la précision finale de la mesure. Tous les détails sur le traitement de l'incertitude systématique sur ϵ_{II} seront donnés dans la section 16.6.

Dans la section 16.3.4, les corrélations entre les différents termes employés pour décomposer les facteurs de correction catégorie par catégorie $C_{total,i}$ sont étudiées. Des corrélations faibles sont nécessaires en particulier pour valider l'évaluation indépendante des différentes systématiques, souvent effectuée dans les données à partir de plusieurs échantillons ($\gamma + X$ pour l'isolation dans le calorimètre, voir section 10.4, $Z^0 \rightarrow l^+l^- \gamma$ et $Z^0 \rightarrow e^+e^-$ avec extrapolation pour l'identification, voir section 9).

16.3.4 Factorisation des différents termes

Afin de valider l'ensemble des différentes corrections et incertitudes ayant un effet sur l'étape de déconvolution (description de l'isolation dans le calorimètre, efficacité de sélection sur l'isolation des traces, identification, etc, voir section 16.6), il est nécessaire de démontrer la faible corrélation entre les corrections d'acceptance (évaluées obligatoirement à partir de l'échantillon de signal simulé) et les efficacités des coupures de qualité (basées en grande partie sur des échantillons purs dans les données) d'une part, et entre les différentes coupures de qualité d'autre part, évaluées à partir de divers échantillons

indépendants des données $\gamma\gamma + X$.

Il est plus facile de se rendre compte de l'importance de ce critère en considérant le cas extrême d'un signal totalement dominé par des paires contenant des photons de fragmentation, basé par exemple uniquement sur le processus γ +jet (+ PS) dans l'échantillon de signal PYTHIA8 (voir section 13.2). Il est possible de sélectionner de tels événements par une redéfinition du volume fiduciel au niveau particule. Dans ce cas, les candidats photons sont très différents (*ie.* biaisés) par rapport à ceux présents dans les échantillons $\gamma + X$ ou $Z^0 \rightarrow l^+l^-\gamma$ classiques utilisés pour le calcul des différentes corrections de performance, de telle sorte que l'utilisation des *ID scale factors* ou des corrections d'isolation sans précautions supplémentaires mène très probablement à un biais significatif. Il convient alors d'évaluer et de corriger ce biais ou à défaut d'en tenir compte explicitement dans les incertitudes. Celui-ci peut être évalué directement à partir de la corrélation entre les corrections d'acceptance et les efficacités des coupures de qualité.

De façon similaire, un échantillon de photons avec une sélection sur l'isolation des traces très sévère présente probablement des gerbes biaisées par rapport à un échantillon plus classique, invalidant l'utilisation directe des *ID scale factors* du groupe de performance sans corrections et/ou incertitudes supplémentaires. Le biais associé peut alors être évalué de la corrélation entre les différentes efficacités des coupures de qualité.

Corrélation entre corrections d'acceptance et coupures de qualité

Afin d'estimer ces différents biais potentiels dans l'analyse $\gamma\gamma + X$, la démarche suivante est adoptée : tout d'abord, la corrélation entre les corrections d'acceptance ($C_{fid.}$)⁸³ et les coupures de qualité ($\epsilon_{TT} \cdot \epsilon_{II}$) est estimée. Pour cela, $C_{fid.}$ est recalculé après application des coupures de qualité au numérateur et dénominateur (voir équation 81) et comparé aux valeurs nominales obtenues dans la section 16.3.2. Les résultats sont montrés dans les figures 147 et 148. Les $C_{fid.}$ avant et après correction sont très proches pour l'ensemble des catégories utilisées dans l'analyse. Dans le cas inclusif, $C_{fid.}$ vaut 0.953 (0.958) avant (après) application des coupures de qualité. Les corrections d'acceptance et les coupures de qualité présentent donc une corrélation très faible.

Cependant, une corrélation résiduelle est clairement visible par le fait que les $C_{fid.}$ après coupures prennent de façon systématique des valeurs plus élevées que les facteurs initiaux : +0.01 environ pour les différentes catégories en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc, et +0.005 pour la région **TITI** complète. Cette différence deviendrait probablement significative si la corrélation entre les incertitudes statistiques des deux types de $C_{fid.}$ étaient prises en compte. Il est possible de montrer que cet effet vient en intégralité de la corrélation résiduelle entre la sélection sur l'isolation calorimétrique, appliquée dans le cadre de la définition du volume fiduciel, et la sélection d'identification TIGHT, incluse dans les coupures de qualité. En effet, si l'étude est répétée avec uniquement la sélection TIGHT appliquée, des distributions identiques à celles obtenus dans les figures 147 et 148 (traits pleins verts) sont obtenues. En revanche, aucune corrélation résiduelle significative entre la sélection sur l'isolation calorimétrique et la sélection sur l'isolation des traces n'est

⁸³ $\epsilon_{good \gamma\gamma,i}$ est exclu de l'étude car les différents critères utilisés lors de la présélection des photons sont appliqués dans l'ensemble des études de performance d'ATLAS, et donc se factorisent automatiquement des autres termes utilisés pour la décomposition de $C_{tot.,i}$.

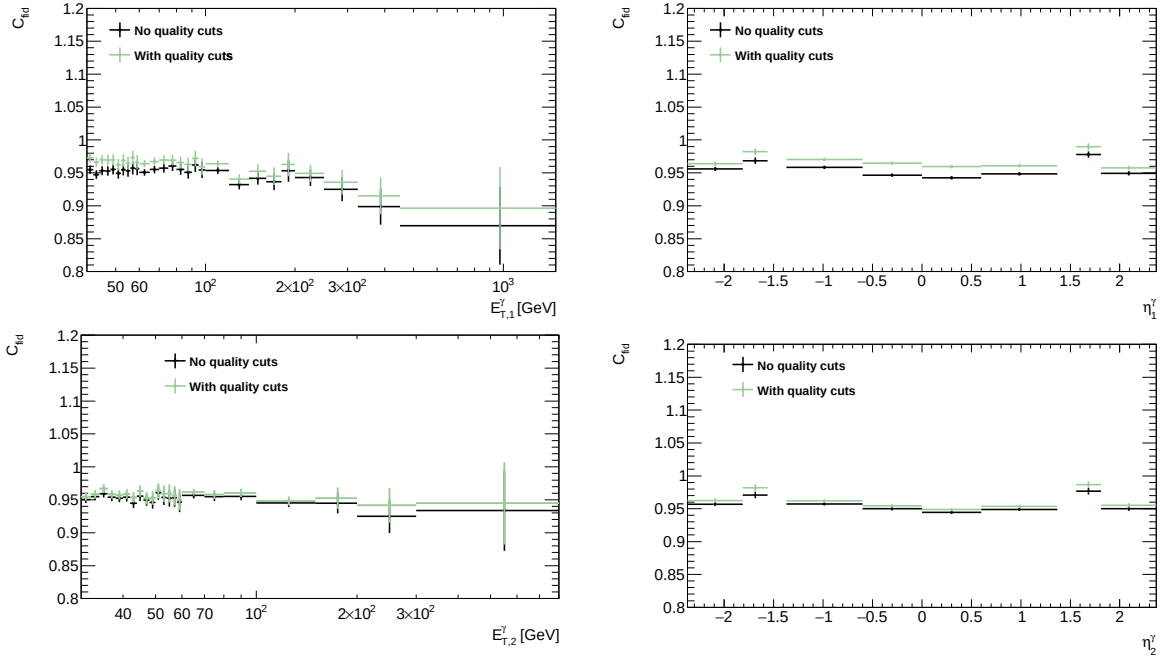


FIG. 147 – $C_{fid.}$ en fonction de l'impulsion transverse (gauche) et de la pseudo-rapidité (droite) du premier (haut) et du deuxième photon (droite) avec (traits noirs) et sans (traits verts) application des coupures de qualité. L'échantillon de signal simulé SHERPA est utilisé. Les barres d'erreur représentent uniquement l'incertitude statistique non-corrélée.

observée pour le signal $\gamma\gamma + X$, probablement parce que l'isolation des traces est très peu restrictive pour le signal (voir section 16.3.3).

Les différences entre les deux types de $C_{fid.}$ ne sont néanmoins pas prises comme incertitudes supplémentaires pour la mesure car l'effet relié à la corrélation résiduelle entre l'identification TIGHT et l'isolation calorimétrique des photons est inclus dans les incertitudes sur les *ID scale factors* fournies par le groupe de performance et mène à une systématique additionnelle du même ordre de grandeur que l'effet observé ici ($\simeq 1\%$) [186, 187]. Les figures 147 et 148 permettent ainsi une première validation de l'utilisation des *ID scale factors* et des incertitudes associées pour l'étape de déconvolution de l'analyse $\gamma\gamma + X$.

Corrélation entre la sélection d'identification TIGHT et l'isolation des traces

De façon similaire, les corrélations résiduelles entre la sélection TIGHT-TIGHT (ϵ_{TT}) et la sélection isolé-isolé dans le détecteur interne (ϵ_{II}) sont évaluées en recalculant ϵ_{II} sans demander TIGHT-TIGHT. Les résultats correspondants sont représentés dans la figure 149 pour les différentes observables de mesure. Pour l'intégralité des catégories considérées, les résultats avec et sans la sélection TIGHT-TIGHT sont en très bon accord, montrant la faible corrélation entre les deux critères pour le signal $\gamma\gamma + X$. La sélection TIGHT-TIGHT augmente néanmoins systématiquement ϵ_{II} d'environ +0.3%. À nouveau, cet effet résiduel est déjà pris en compte dans les incertitudes sur les *ID scale factors* calculés par le groupe de performance, qui incluent une incertitude systématique additionnelle spécialement dédiée aux analyses utilisant la sélection sur l'isolation des

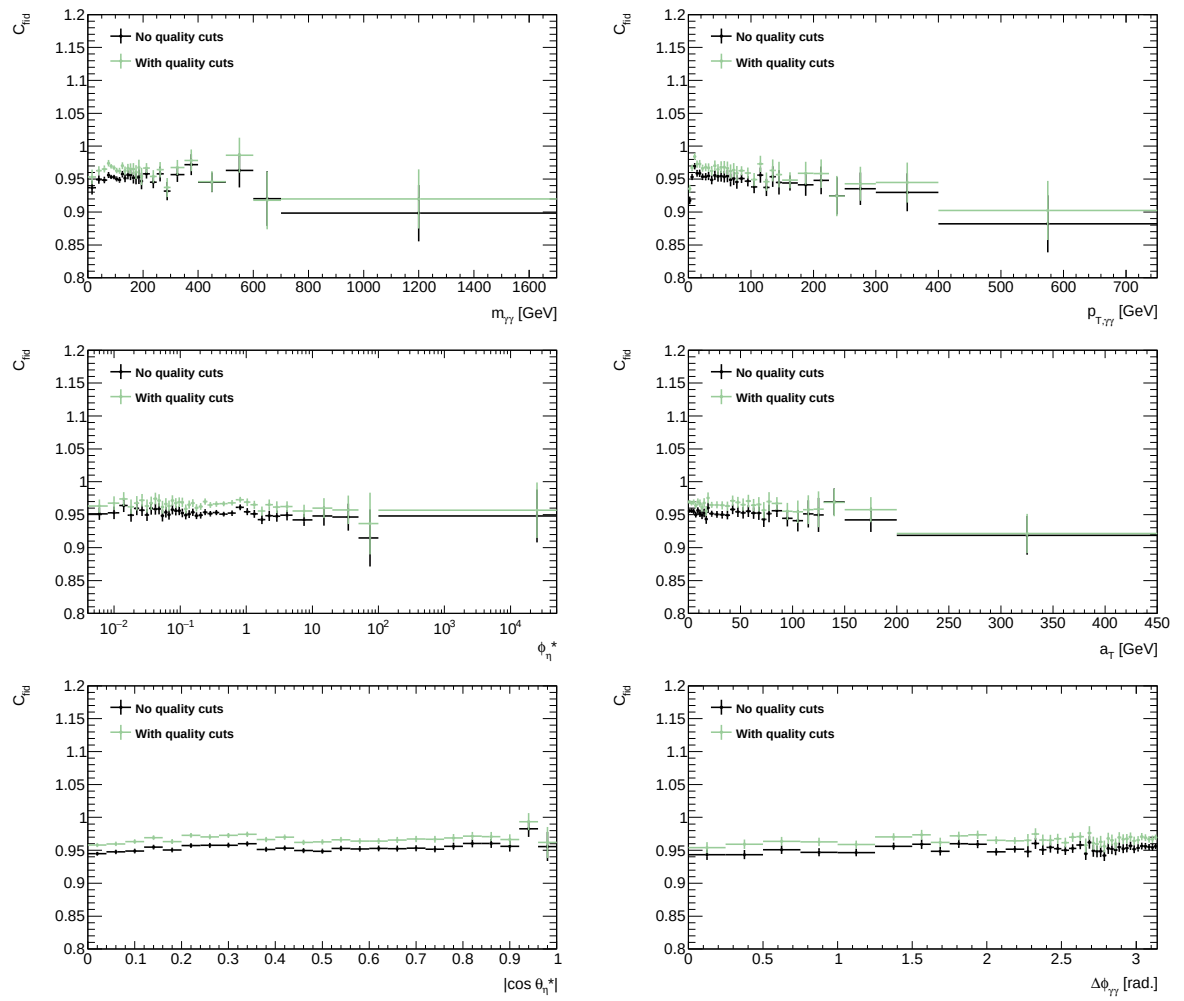


FIG. 148 – Figure similaire à la figure 147 pour les différentes observables de mesure.

traces implémenté dans l'analyse $\gamma\gamma + X$.

En conclusion, l'utilisation des *ID scale factors* et des incertitudes associées fournies par le groupe de performance permet de couvrir de façon adéquate l'effet des corrélations résiduelles entre les corrections d'acceptance et les différentes coupures de qualité. En particulier, aucune correction résiduelle significative n'a été observée entre la sélection sur l'isolation des traces et celle sur l'isolation dans le calorimètre pour le signal $\gamma\gamma + X$, probablement grâce à la très bonne efficacité de la coupure sur l'isolation des traces.⁸⁴

Dans la section 16.3.5, l'efficacité du système de déclenchement est abordée brièvement.

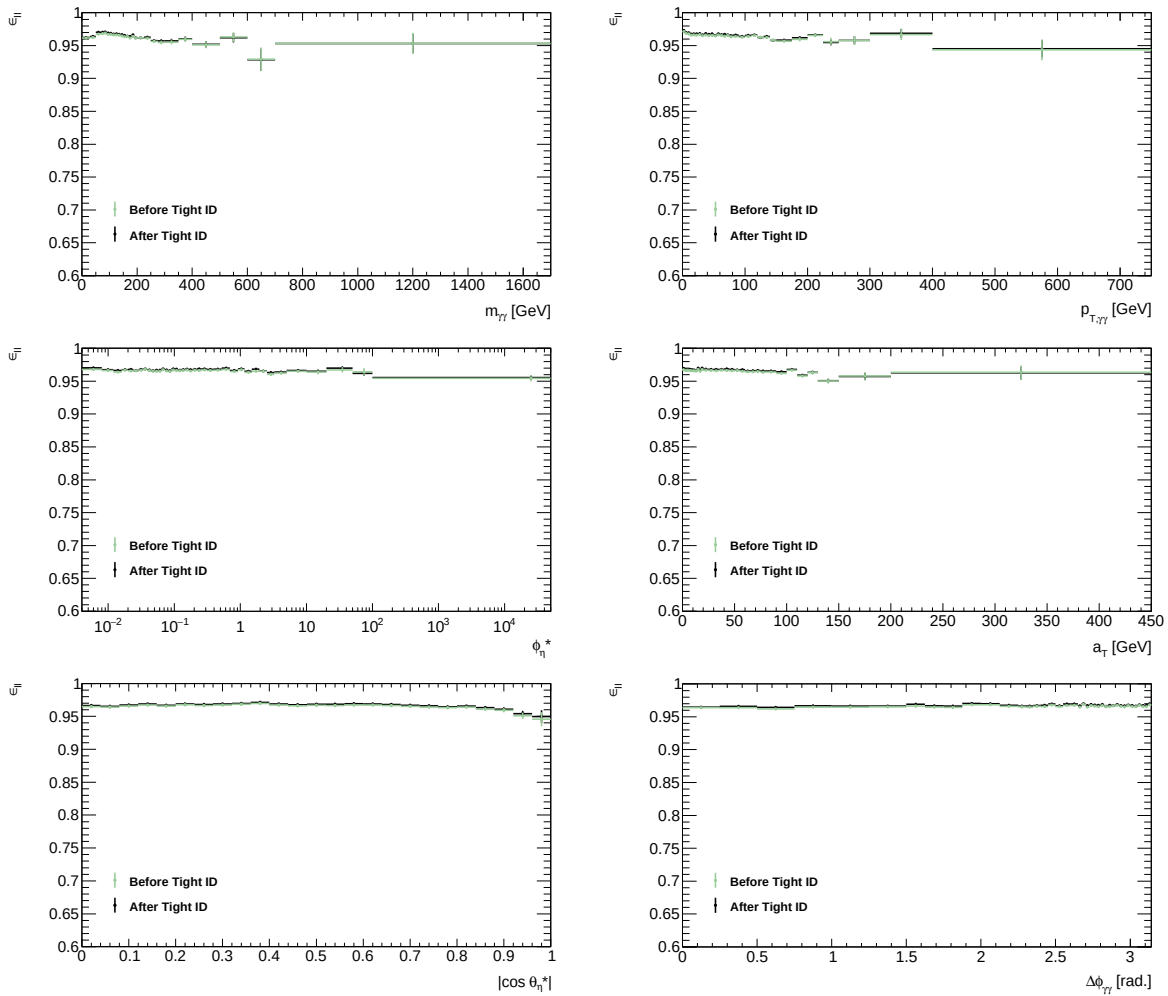


FIG. 149 – ϵ_{II} en fonction des différentes observables de mesure avec (traits noirs) et sans (traits verts) application de la sélection d'identification TIGHT-TIGHT. L'échantillon de signal simulé SHERPA est utilisé. Les barres d'erreur représentent uniquement l'incertitude statistique non-corrélée.

⁸⁴En revanche, ces deux coupures sont attendues comme fortement corrélées pour le bruit de fond.

16.3.5 Efficacité du système de déclenchement

La chaîne de déclenchement employée pour la sélection de l'échantillon $\gamma\gamma + X$ dans les données 2012 a déjà été détaillée dans la section 13.1. L'efficacité de celle-ci a été évaluée par deux mesures indépendantes réalisées avec les données dans le cadre d'une étude générale sur l'efficacité des systèmes de déclenchement pour les photons dans ATLAS [248]. Les résultats principaux sont résumés brièvement ci-dessous.

Utilisation des désintégrations $Z^0 \rightarrow l^+l^-\gamma$

La première mesure est basée sur la sélection d'événements $Z^0 \rightarrow l^+l^-\gamma$ dans les données, avec le photon émis par l'un des leptons dans l'état final (désintégrations radiatives). Comme nous l'avons déjà vu (voir section 9), une telle sélection permet d'obtenir un échantillon de photons très pur dans une gamme limitée en E_T^γ (< 100 GeV environ). Elle est appliquée avec et sans critères particuliers vis-à-vis des chaînes de déclenchement photons `g25_loose` et `g35_loose` afin de permettre l'extraction de leur efficacité respective, le déclenchement standard étant alors basé uniquement sur les leptons.

L'effet systématique dominant est relié à la présence résiduelle d'événements de bruit de fond dans l'échantillon ($< 1\%$), dominé par $Z^0 + \text{jet}$. L'effet est évalué en faisant varier les différents critères de sélection appliqués aux leptons. En revanche, la sélection des photons est fixe et correspond à peu de choses près à la sélection de l'analyse $\gamma\gamma + X$. En effet, en dehors de la zone d'exclusion en η légèrement réduite ($1.56 \rightarrow 1.52$) et du critère d'isolation limité à une coupure unique sur l'isolation calorimétrique à 4 GeV, les deux sélections sont identiques. L'effet de ces différences de sélection sur l'efficacité du système de déclenchement sera négligée dans la suite.

Les résultats en fonction de E_T^γ et $|\eta^\gamma|$ sont présentés dans la figure 150,⁸⁵ tandis que les efficacités des chaînes `g25_loose` et `g35_loose` ainsi que l'efficacité combinée de la chaîne `g35_loose_g25_loose` sont résumées dans le tableau 29.⁸⁶ Aucune corrélation significative entre les deux objets n'ayant été observée pour la chaîne `g35_loose_g25_loose`, son efficacité est calculée par un simple produit des résultats pour `g25_loose` et `g35_loose`.⁸⁷ En revanche, les incertitudes sur l'efficacité des deux chaînes `g25_loose` et `g35_loose` sont corrélées, ce qui est pris en compte dans le résultat final cité pour `g35_loose_g25_loose`.

Comme attendu, les efficacités montent très rapidement lors du dépassement du seuil de déclenchement en E_T^γ , sans dépendance significative en fonction de $|\eta^\gamma|$. Les efficacités finales pour $E_{T,1} > 40$ GeV, $E_{T,2} > 30$ GeV (*ie.* 5 GeV au-dessus des seuils) sont très proches de 100%. La précision sur l'efficacité est dominée par la statistique limitée de l'échantillon $Z^0 \rightarrow l^+l^-\gamma$ et non par les effets systématiques dus au bruit de fond résiduel.

Cependant, la gamme en E_T^γ des photons radiatifs du $Z^0 \rightarrow l^+l^-\gamma$ est limitée à $E_T^\gamma < 100$ GeV. Afin de vérifier l'efficacité de la chaîne `g35_loose_g25_loose` à plus

⁸⁵La sélection sur E_T^γ est plus faible ici afin de pouvoir visualiser la pente de montée des deux efficacités.

⁸⁶ $E_{T,1} > 40$ GeV, $E_{T,2} > 30$ GeV est considéré ici.

⁸⁷Il serait néanmoins intéressant d'étudier l'efficacité de `g25_loose` et `g35_loose` en fonction de $\Delta R_{\gamma\gamma}$, ce qui devrait être réalisé d'ici peu dans le cadre de l'analyse $\gamma\gamma + X$.

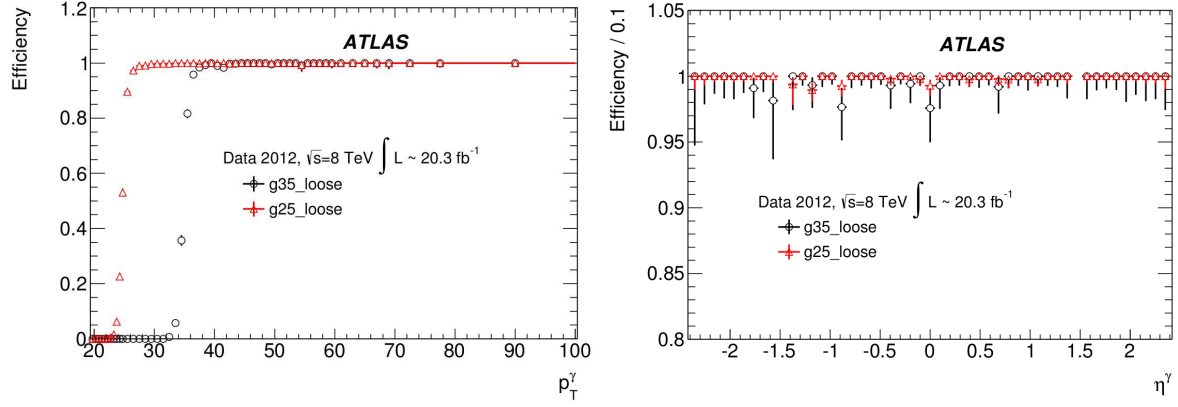


FIG. 150 – Efficacité des chaînes de déclenchement `g25_loose` (triangles rouges) et `g35_loose` (ronds noirs) en fonction de l'impulsion transverse (gauche) et de la pseudo-rapidité (droite) du candidat photon, évaluée à partir d'un échantillon très pur de photons issu des désintégrations $Z^0 \rightarrow l^+l^-\gamma$ dans les données [248].

Chaîne de déclenchement	Efficacité
<code>g25_loose</code>	$99.8^{+0.05}_{-0.06}$ (stat.) $^{+0.01}_{-0.01}$ (syst.)
<code>g35_loose</code>	$99.62^{+0.16}_{-0.21}$ (stat.) $^{+0.06}_{-0.06}$ (syst.)
<code>g35_g25_loose</code>	$99.42^{+0.22}_{-0.29}$ (tot.)

TAB. 29 – Efficacités inclusives des chaînes de déclenchement `g25_loose` et `g35_loose` évaluées à partir d'un échantillon très pur de photons issu des désintégrations $Z^0 \rightarrow l^+l^-\gamma$ dans les données [248].

haut E_T^γ , une méthode alternative appelée méthode *bootstrap* est implémentée, objet du paragraphe suivant.

Méthode *bootstrap*

La méthode du *bootstrap* permet d'évaluer l'efficacité du système de déclenchement par rapport au déclenchement à L1 uniquement (voir 5.7). Elle est basée sur l'exploitation de données $\gamma + X$ dont l'enregistrement a été déclenché à L1 lors de prises de données particulières avec le L2 et l'EF désactivés. Contrairement à la méthode précédente, l'échantillon obtenu est alors contaminé de manière non-négligeable par du bruit de fond, issu principalement de jets mal identifiés. L'efficacité de déclenchement finale peut alors être calculée en demandant le L2 et l'EF a posteriori de l'acquisition. Cependant, l'efficacité obtenue représente alors le produit entre l'efficacité relative aux photons immédiats et au bruit de fond pondéré par la pureté en photons immédiats de l'échantillon.

Le bruit de fond est estimé par une méthode basée sur les données appelée méthode des bandes latérales ou méthode ABCD, décrite par exemple dans l'analyse $\gamma + X$ à $\sqrt{s} = 8$ TeV [186, 200]. La pureté de l'échantillon *bootstrap* obtenue avec la méthode ABCD est représentée dans la figure 151 en fonction de E_T^γ . La pureté de l'échantillon est comprise entre 30 et 90% en fonction de l'impulsion transverse du candidat, en bon accord avec ce qui est observé dans le cadre de l'analyse $\gamma\gamma + X$. Les variations reliées à la définition des régions de contrôles ($\tilde{\mathbf{T}}$) et aux fuites de photons immédiats dans les régions de contrôle sont prises comme incertitudes systématiques. Bien que celles-ci soient importantes à petit E_T^γ , l'effet sur l'efficacité finale des chaînes de déclenchement pour les photons est limité grâce aux efficacités de déclenchement très similaires pour les photons immédiats et le bruit de fond.

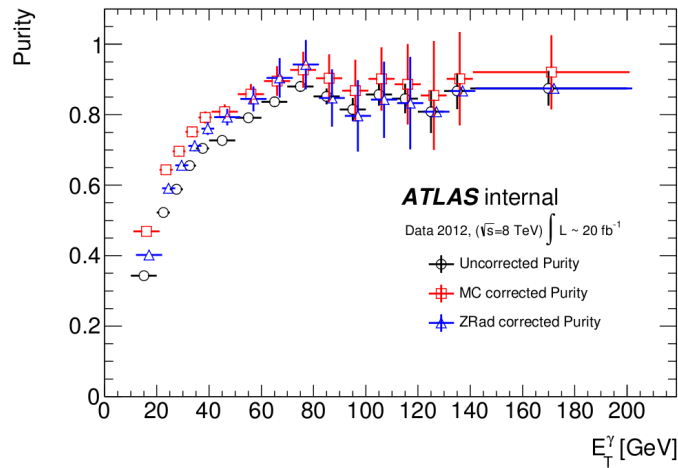


FIG. 151 – Pureté de l'échantillon *bootstrap* en fonction de l'impulsion transverse du candidat photon avec la méthode des bandes latérales ou méthode ABCD (rond noirs). Les fuites de photons immédiats dans les régions de contrôle peuvent être corrigées à l'aide du rapport $\tilde{\mathbf{T}}/\mathbf{T}$ observé dans le MC (carrés rouges) ou dans l'échantillon $Z^0 \rightarrow l^+l^-\gamma$ des données (triangles bleus). Les incertitudes représentées sont statistiques uniquement [248].

Les efficacités de déclenchement obtenues en fonction du nombre de vertex primaires reconstruits dans l'événement sont montrées dans la figure 152. Celles-ci sont totalement plates, montrant une dépendance négligeable vis-à-vis de l'empilement. Elles ont également été calculées indépendamment pour chaque période d'acquisition pendant l'année 2012 et sont toutes compatibles dans les incertitudes citées.

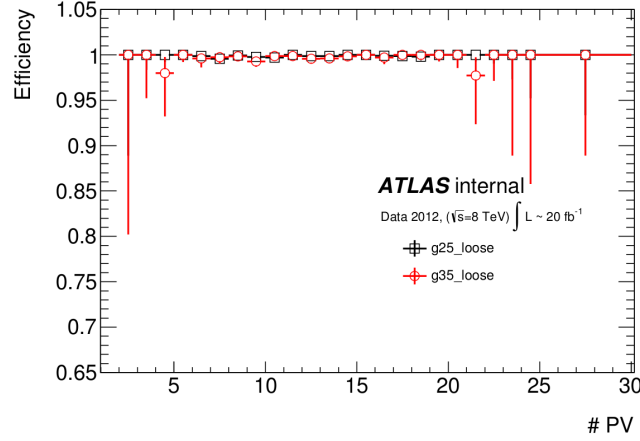


FIG. 152 – Efficacités des chaînes de déclenchement `g25_loose` et `g35_loose` en fonction du nombre de vertex primaires reconstruits dans l'événement, évaluées à partir de l'échantillon *bootstrap*. [248].

Les résultats combinant les chiffres obtenus à partir de l'échantillon *bootstrap* et l'efficacité du L1 estimée avec la méthode du $Z^0 \rightarrow l^+l^-\gamma$ ($> 99\%$) sont présentés dans le tableau 30. La précision de la méthode *bootstrap* est dominée par les effets systématiques reliés à la soustraction du bruit de fond et non plus par la statistique comme dans la méthode précédente basée sur l'échantillon $Z^0 \rightarrow l^+l^-\gamma$. L'incertitude finale est cependant très réduite et du même ordre de grandeur par rapport au paragraphe précédent. L'échantillon *bootstrap* inclut par ailleurs des photons de plus haut E_T^γ ($100 \text{ GeV} < E_T^\gamma < 200 \text{ GeV}$), inaccessibles dans les désintégrations $Z^0 \rightarrow l^+l^-\gamma$.

Finalement, les deux estimations de l'efficacité de la chaîne `g35_loose_g25_loose` (*bootstrap*, $Z^0 \rightarrow l^+l^-\gamma$) ainsi que leur combinaison finale sont présentées dans le tableau 31. Celles-ci sont en très bon accord malgré les gammes très différentes en E_T^γ des deux méthodes, démontrant la très bonne efficacité du système de déclenchement basé sur la chaîne `g35_loose_g25_loose` dans ATLAS quelques soient les impulsions transverses (et les pseudo-rapidités) considérées. Par conséquent, le chiffre de $99.44\%^{+0.14\%}_{-0.17\%}$ sera adopté dans l'analyse $\gamma\gamma + X$ pour corriger les inefficacités du système de déclenchement, que ce soit pour la région **TITI** complète ou pour les différentes catégories définies à partir des observables de mesure. L'échantillon de signal SHERPA, non utilisé dans l'analyse principale pour l'efficacité du *trigger*, donne une efficacité de 100% pour la sélection adoptée dans l'analyse $\gamma\gamma + X$ (99.96%).

Dans la section 16.4, la correspondance entre l'énergie d'isolation transverse définie au niveau particule et l'énergie d'isolation calorimétrique est étudiée grâce aux échantillons de signal simulés SHERPA et PYTHIA8. Cette étude permet, sur la base de critères de migrations minimales entre le niveau généré et le niveau reconstruit, de déterminer la sélection optimale à appliquer au niveau particule dans l'analyse $\gamma\gamma + X$ afin de limiter les incertitudes de déconvolution.

Chaîne de déclenchement	Efficacité
g25_loose	99.92 $^{+0.06}_{-0.07}$
g35_loose	99.54 $^{+0.13}_{-0.16}$
g35_loose_g25_loose	99.46 $^{+0.19}_{-0.23}$

TAB. 30 – Efficacités inclusives des chaînes de déclenchement `g25_loose` et `g35_loose` évaluées à partir de l'échantillon *bootstrap* dans les données [248].

Méthode	Chaîne de déclenchement	Efficacité
$Z^0 \rightarrow l^+ l^- \gamma$	g35_loose_g25_loose	99.42 $^{+0.22}_{-0.29}\%$
Bootstrap	g35_loose_g25_loose	99.46 $^{+0.19}_{-0.23}\%$
Combinaison	g35_loose_g25_loose	99.44 $^{+0.14}_{-0.17}\%$

TAB. 31 – Combinaison finale des efficacités inclusives des chaînes de déclenchement `g25_loose` et `g35_loose` évaluées à partir des données [248].

16.4 Étude de l'énergie transverse d'isolation définie au niveau particule

Au niveau reconstruit, le critère d'isolation final employé dans l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV est constitué à la fois d'une sélection sur l'isolation calorimétrique (< 6 GeV) et sur l'isolation des traces (< 2.6 GeV), voir section 14.3. Expérimentalement, ce critère présente l'avantage de réduire la baisse de $C_{fid.}$ à haut E_T^γ grâce à la coupure de 6 GeV sur $E_T^{calo.iso.}$ (voir section 16.3.2) tout en maintenant une haute pureté grâce à la coupure modérée sur l'isolation des traces (voir section 15). Par ailleurs, la taille des cônes utilisés ($\Delta R = 0.4$ et 0.2 , respectivement) permettent de garder la dépendance vis-à-vis des conditions d'empilement sous contrôle (voir section 10) tout en préservant la convergence des calculs QCD (voir section 3.5).

Bien qu'on attende une forte corrélation entre les énergies d'isolation définies au niveau particule et celles définies au niveau reconstruit, celles-ci ne sont en général pas identiques pour les raisons suivantes (non exhaustives) :

Énergie d'isolation calorimétrique

- ◇ Les *clusters* topologiques utilisés pour le calcul de l'isolation calorimétrique ne sont pas calibrés (voir section 10.3). Ceux-ci sont basés sur l'énergie reconstruite dans les cellules (voir section 7.1).
- ◇ Les corrections reliées aux fuites d'énergie du candidat photon et à l'empilement/événement sous-jacent sont moyennées.

Isolation des traces

- ◇ Les particules neutres ne laissent pas de dépôts dans le détecteur interne.
- ◇ Les traces entrant de le calcul de $p_T^{track.iso.}$ doivent passer une sélection de qualité stricte incluant $p_T^{track} > 1$ GeV.

Plus de détails concernant ces différents aspects sont disponibles dans la section 10.

Partant de ces considérations, il est intéressant d'étudier la correspondance entre l'énergie d'isolation au niveau reconstruit et celle au niveau particule avec les échantillons de signal simulés, dans l'optique de déterminer une coupure optimale sur $E_T^{part.iso.}$. Cependant, notre marge de manœuvre est très limitée. En effet, il est bien connu que les échantillons de signal à notre disposition pour la déconvolution (SHERPA et PYTHIA8, voir section 13.2) intègrent une approche particulièrement naïve des photons de fragmentation et des jets de haute énergie émis dans l'état final, et donc de l'énergie d'isolation. Par conséquent, les critères minimaux suivants sont adoptés dans l'étude :

- ◇ Les coupures sur $E_T^{part.iso.}$ sont prises comme identiques pour les deux photons.
- ◇ Celles-ci sont choisies uniquement de façon à minimiser les migrations entre la catégorie "isolée" au niveau particule et "isolée" au niveau reconstruit.

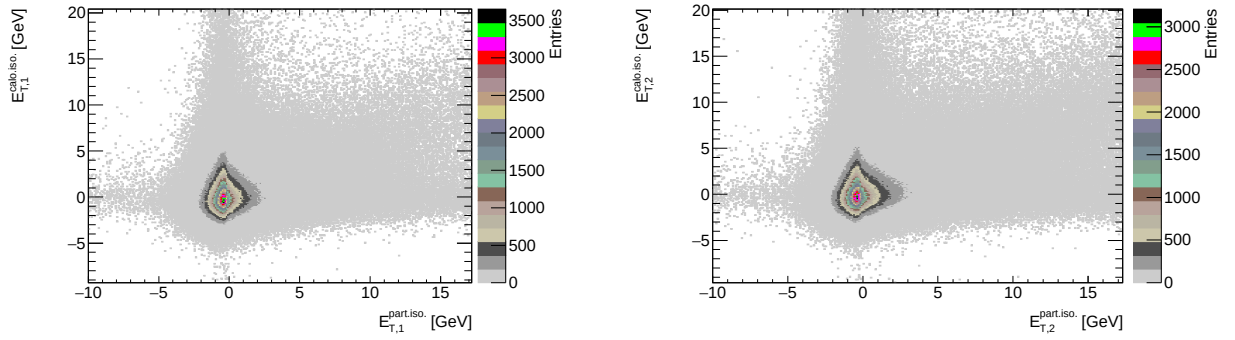


FIG. 153 – Distribution à deux dimensions de l'isolation au niveau particule (axe des abscisses) du premier (gauche) et du deuxième (droite) photon en fonction de l'isolation calorimétrique (axe des ordonnées) du candidat correspondant. L'échantillon de signal simulé SHERPA est utilisé. les événements doivent passer l'intégralité de la sélection au niveau particule et au niveau reconstruit à l'exception de la coupure sur $E_T^{\text{part.iso.}}$ et sur l'isolation calorimétrique.

Ces deux critères simples permettent de limiter au maximum les corrections de déconvolution dues à l'isolation, basées sur la description de l'isolation dans les échantillons de signal simulé et donc très incertaines.

La corrélation entre l'isolation au niveau particule et au niveau reconstruit est tout d'abord évaluée globalement dans la figure 153 avec l'échantillon de signal simulé SHERPA. Il est rappelé que l'isolation au niveau particule peut prendre des valeurs négatives à cause des fluctuations autour des corrections moyennes de l'événement sous-jacent (voir section 12.2) tandis que l'énergie d'isolation calorimétrique subit un effet similaire des corrections d'empilement et de l'événement sous-jacent et des corrections de fuites (voir section 10.3). Bien que la très grande majorité des événements soit bien isolée à la fois au niveau particule ($E_T^{\text{part.iso.}}$) et dans le calorimètre ($E_T^{\text{calo.iso.}}$), des queues de distribution importantes correspondant à des photons bien isolés au niveau particule mais peu isolés dans le calorimètre et vice-versa sont observées.

Les corrélations ainsi que leur dépendance éventuelle avec l'impulsion transverse des photons sont quantifiées plus précisément dans la figure 154, où les profils selon $E_T^{\text{part.iso.}}$, obtenus à partir de la figure précédente, sont représentés pour plusieurs catégories en E_T^γ . Une corrélation linéaire est observée pour les photons ayant $E_T^{\text{part.iso.}} > 2$ GeV. Aucune dépendance significative en fonction de E_T^γ n'est visible, indiquant qu'une coupure fixe sur $E_T^{\text{part.iso.}}$ est a priori suffisante pour limiter les migrations et satisfaire les critères mentionnés au paragraphe précédent.

Un ajustement linéaire des profils d'isolation est réalisé séparément au voisinage de $E_T^{\text{calo.iso.}} = 6$ GeV pour les deux échantillons (SHERPA, PYTHIA8) et pour le premier et le second photon (intégré en E_T^γ). Les résultats sont dévoilés dans la figure 155. Dans tous les cas une pente proche de 0.6 est obtenue, avec une valeur correspondante à $E_T^{\text{calo.iso.}} = 6$ GeV située entre 11 et 12 GeV pour $E_T^{\text{part.iso.}}$, en bon accord entre SHERPA, PYTHIA8 et pour les deux photons de la paire.

Afin de valider définitivement ces valeurs de coupure pour $E_T^{\text{part.iso.}}$ vis-à-vis des critères énoncés au début de la section, les régions R_1 , R_2 et R_3 sont définies comme indiquées dans la figure 156. Le rapport entre les différents nombres d'événements cor-

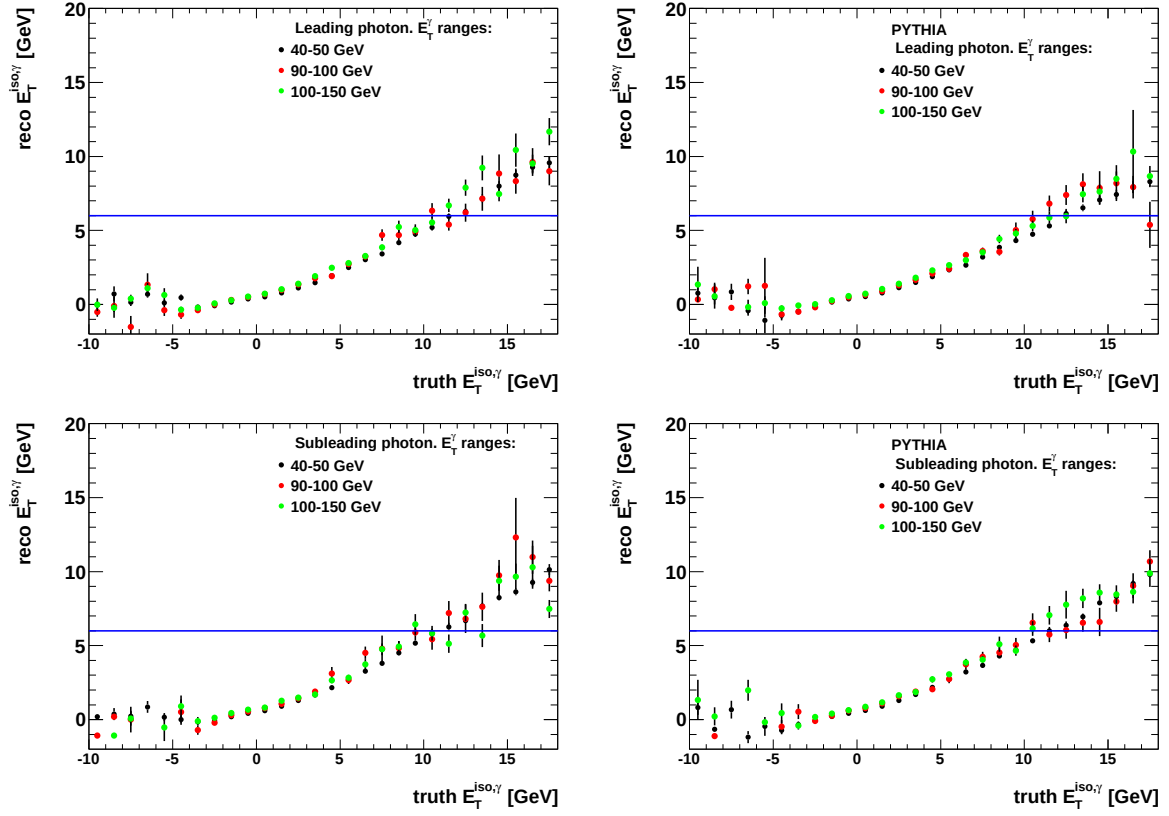


FIG. 154 – Profils des histogrammes 2D de la figure 153 pour le premier (haut) et le deuxième (bas) photon à partir de l'échantillon de signal simulé SHERPA (gauche) et PYTHIA8 (droite). La coupure finale sur l'isolation dans le calorimètre appliquée dans l'analyse $\gamma\gamma + X$, basée majoritairement sur des critères expérimentaux, est représentée par une droite horizontale (6 GeV).

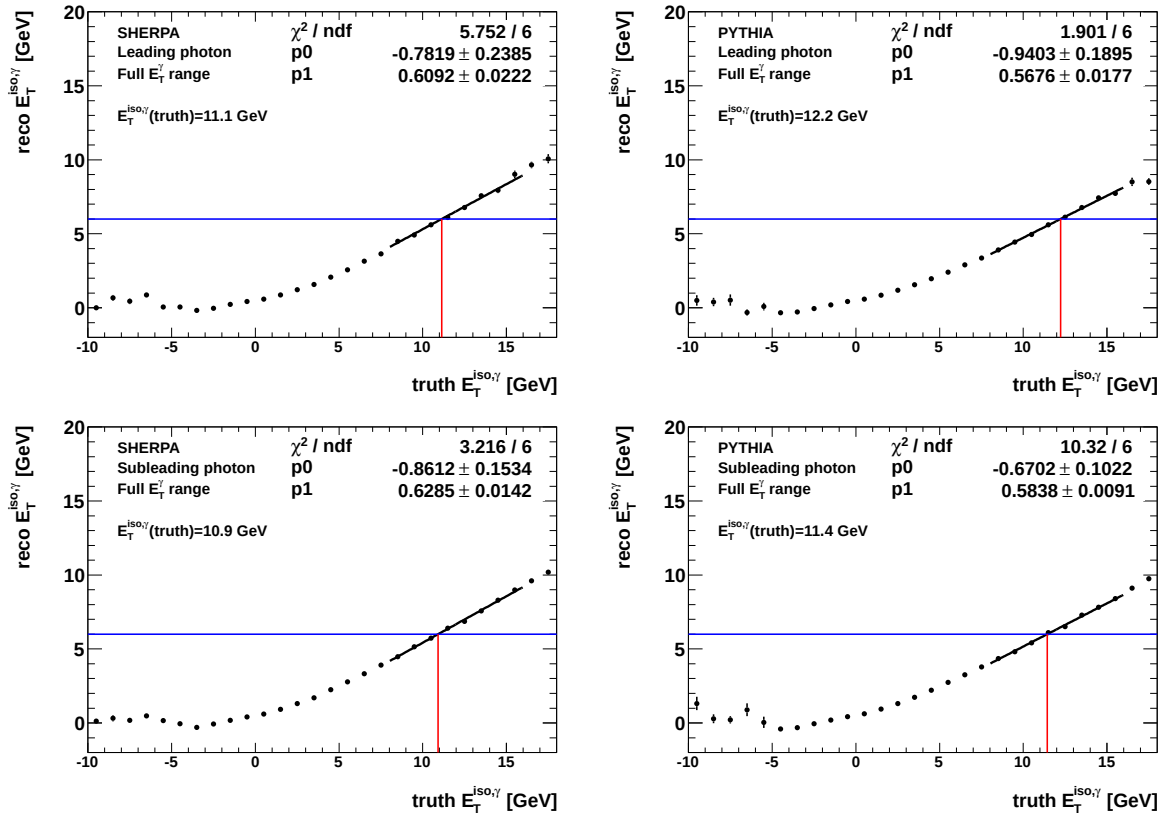


FIG. 155 – Résultat d'un ajustement linéaire des profils de la figure 154 réalisé indépendamment pour SHERPA, PYTHIA8, et pour le premier et le deuxième photon de la paire (intégré en E_T^γ). La coupure finale sur l'isolation dans le calorimètre appliquée dans l'analyse $\gamma\gamma + X$, basée majoritairement sur des critères expérimentaux, est représentée par une droite horizontale (6 GeV). La valeur correspondante au niveau particule indiquée par l'ajustement, représentée par un segment rouge, est respectivement d'environ 11 GeV pour SHERPA et entre 11.5 GeV et 12.5 GeV pour PYTHIA.

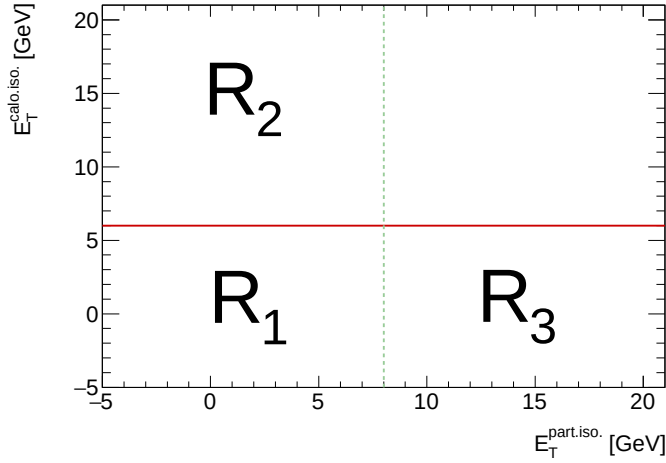


FIG. 156 – Définition des régions R_1 , R_2 et R_3 dans les échantillons de signal simulés SHERPA et PYTHIA8 pour l’estimation des taux de migration entre les catégories “isolée” et “non-isolée” entre le niveau particule et le niveau reconstruit. L’intégralité de la sélection au niveau particule et au niveau reconstruit est appliquée à l’exception de la coupure d’isolation au niveau particule ($E_T^{\text{part.iso.}}$) et dans le calorimètre ($E_T^{\text{calo.iso.}}$). La séparation entre R_2 et $R_{1,3}$ en terme d’isolation calorimétrique est fixe, tandis que la limite entre R_3 et $R_{1,2}$ est variée afin de tester plusieurs valeurs de coupure sur $E_T^{\text{part.iso.}}$.

respondant à $(R_2 + R_3)/R_1$ est ensuite calculé pour plusieurs valeurs de coupure sur $E_T^{\text{part.iso.}}$ afin d’évaluer les taux de migration entre les catégories “isolée” et “non-isolée” entre le niveau particule et le niveau reconstruit dans les différents cas.

Les résultats présentés séparément dans plusieurs catégorie en E_T^γ pour SHERPA, PYTHIA, le premier et le deuxième photon sont montrés dans la figure 157. Dans tous les cas, une coupure sur $E_T^{\text{part.iso.}}$ correspondant à $E_T^{\text{part.iso.}} < 11$ GeV permet de minimiser autant que possible les migrations entre “isolée” et “non-isolée” entre le niveau particule et le niveau reconstruit pour un photon donné. Cependant, la zone correspondant au minimum de $(R_2 + R_3)/R_1$ est très large, montrant la “résolution” médiocre sur l’isolation, attendue des figures précédentes et des remarques énoncées en début de section.

Finalement, la figure 158 montre le même rapport $(R_2 + R_3)/R_1$ lorsque les photons ne sont pas différenciés en E_T^γ , séparément pour SHERPA, PYTHIA, le premier et le deuxième photon, avec en plus le cas correspondant à l’analyse $\gamma\gamma + X$ où une coupure sur $E_T^{\text{part.iso.}}$ est appliquée aux deux photons. Les conclusions précédentes restent inchangées et la coupure symétrique correspondant à $E_{T,1,2}^{\text{part.iso.}} < 11$ GeV adoptée dans la définition de l’espace fiduciel de l’analyse $\gamma\gamma + X$ est validée définitivement (voir section 12.2). Elle mène selon les échantillons de signal simulé à un rapport $(R_2 + R_3)/R_1$ inclusif d’environ 3%.

Il serait probablement intéressant pour les analyses QCD futures incluant des photons ($\gamma + X$, $\gamma\gamma + X$, $\gamma(\gamma) + n$ jets + X , etc) de comprendre plus en détail les corrélations entre $E_T^{\text{part.iso.}}$ et $E_T^{\text{calo.iso.}}$. En principe, une amélioration significative de la “résolution” entre l’isolation au niveau particule et au niveau reconstruit pourrait permettre de mesurer des sections efficaces différentielles en fonction de $E_T^{\text{part.iso.}}$. D’après certains théo-

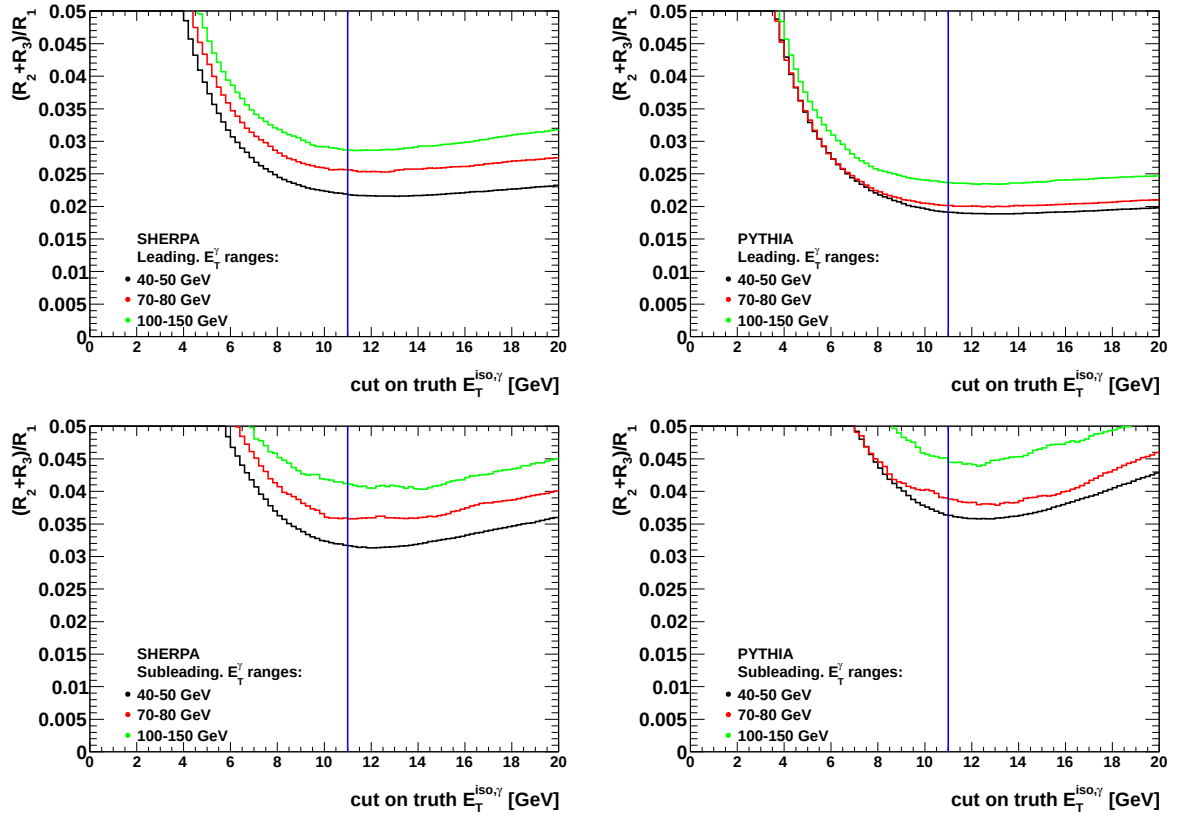


FIG. 157 – Estimation par le rapport $(R_2 + R_3)/R_1$ des taux de migration entre les catégories “isolée” et “non-isolée” entre le niveau particule ($E_T^{\text{part.iso.}}$) et le niveau reconstruit ($E_T^{\text{calo.iso.}}$) pour plusieurs valeurs de coupures sur $E_T^{\text{part.iso.}}$. Les résultats sont présentés séparément pour le premier (haut) et le deuxième (bas) photon issu de l'échantillon de signal simulé SHERPA (haut) et PYTHIA8 (bas).

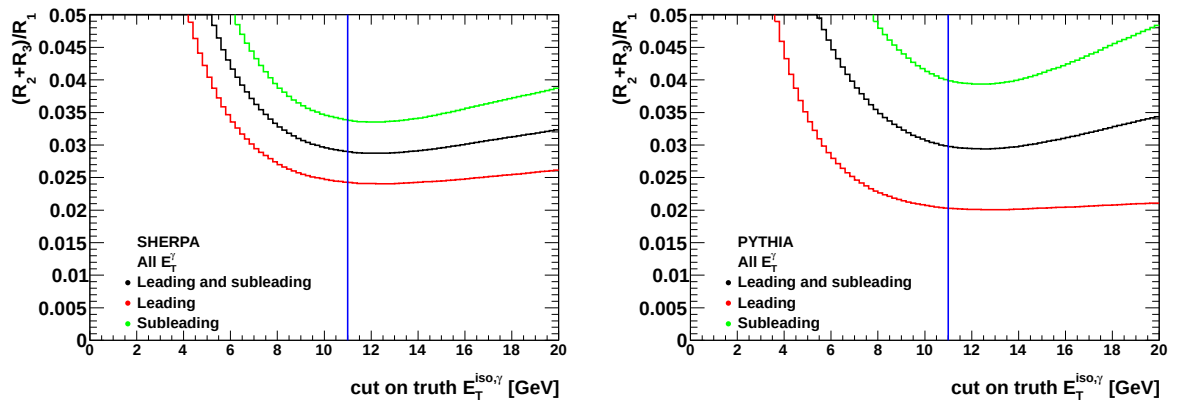


FIG. 158 – Figure similaire à la figure 157 avec les distributions intégrées en E_T^{γ} . Les cas où une coupure sur $E_T^{\text{part.iso.}}$ est effectuée sur le premier (trait plein rouge), le deuxième (trait plein vert) ou les deux photons (trait plein noir) sont représentées. Le $(R_2 + R_3)/R_1$ final est d'environ 3% pour $E_{T,1,2}^{\text{part.iso.}} < 11$ GeV.

riciens,⁸⁸ de telles mesures seraient très intéressantes pour la communauté théorique même si les composantes de fragmentation générées dans SHERPA et PYTHIA sont très différentes de la fragmentation collinéaire considérée dans les calculs QCD théoriques plus rigoureux (voir section 10). En effet, bien qu'il semble peu probable qu'un calcul partonique à ordre fixe NLO avec fonction de fragmentation soit capable de produire des prédictions fiables à cause des instabilités collinéaires déjà discutées dans la section 3.5, ces mesures pourraient permettre une adaptation des algorithmes de cascade de partons des générateurs MC et de manière plus générale apporter des informations utiles sur le processus de fragmentation.

Néanmoins, cela nécessiterait un travail important du côté expérimental, incluant par exemple la calibration des *clusters* topologiques et/ou la combinaison des isolations mesurées dans le détecteur interne et dans le calorimètre (*particle flow isolation*). CMS utilise déjà une énergie d'isolation combinée entre le calorimètre et le détecteur interne, malheureusement aucune figure montrant la correspondance entre $E_T^{\text{part.iso.}}$ et $E_T^{\text{calo.iso.}}$ n'a été rendue publique [224].

Dans la section 16.5, la méthode de déconvolution itérative employée pour la déconvolution des distributions selon les différentes observables de mesure est présentée.

16.5 Méthode de déconvolution itérative

16.5.1 Aperçu

La méthode de déconvolution itérative [231] est une méthode de déconvolution plus évoluée par rapport à la déconvolution catégorie par catégorie abordée dans la section 16.2. Elle se base sur un certain nombre d'itérations déterminé par l'utilisateur à partir de la formule de Bayes bien connue des statisticiens afin de limiter la dépendance de la procédure de déconvolution vis-à-vis de l'échantillon de signal simulé employé, relégué alors au statut de *prior*.

Cela est particulièrement utile lors de l'extraction de sections efficaces différentielles, plus affectées par les migrations d'événements d'une catégorie à l'autre entre le niveau généré et le niveau reconstruit par rapport à la section efficace fiducielle totale (voir le paragraphe concernant $C_{\text{bin purity}}$ de la section 16.3.2) [249]. Ainsi, la déconvolution itérative sera utilisée pour l'extraction nominale des sections efficaces différentielles tandis que la déconvolution catégorie par catégorie sera adoptée pour la section efficace fiducielle totale. Les deux donnent de toutes façons des résultats très proches, comme nous le verrons dans la section 16.5.2. Dans la suite de la section, la méthode itérative est décrite plus en détail.

Partant de la formule de Bayes, la probabilité conditionnelle pour qu'un événement reconstruit dans la catégorie R_j ait été généré dans la catégorie G_i peut être calculée de

⁸⁸Ce commentaire est issu en particulier de discussion avec le Professeur Joey Huston (MSU).

la manière suivante :

$$P(G_i|R_j) = \frac{P(R_j|G_i)P_0(G_i)}{\sum_{l=1}^{B_G+1} P(R_j|G_l)P_0(G_l)}, \quad (85)$$

où $P(R_j|G_i)$ est la probabilité pour qu'un événement généré dans la catégorie G_i ait été reconstruit dans la catégorie R_j , $P_0(G_i)$ est la probabilité (non conditionnelle) pour qu'un événement soit généré dans la catégorie G_i et B_G est le nombre de catégories considéré pour la mesure au niveau généré (voir section 16.2).

$P(R_j|G_i)$ correspond à l'élément de matrice $N_{i,j}^X$ de la matrice de réponse définie dans la section 16.2 normalisé par $\sum_{j=1}^{B_R+1} N_{i,j}^X$. Il peut être obtenu à partir d'un échantillon de signal simulé mais est indépendant de la forme des distributions MC générés, celle-ci étant factorisée dans le terme $P_0(G_i)$ appelé *prior* dans le langage des statistiques bayésiennes. $P(R_j|G_i)$ est donc totalement déterminé par les effets de reconstruction, obtenus à partir de la simulation complète du détecteur ATLAS.

La méthode de déconvolution itérative commence par un premier calcul des différents $P(G_i|R_j)$ en utilisant l'équation 85 et un échantillon de signal simulé comme *prior*. Les nombres d'événements générés sont ensuite recalculés dans chaque catégorie à partir de la formule suivante :

$$\hat{n}(G_i) = \sum_{j=1}^{B_R+1} n(R_j) \cdot P(G_i|R_j), \quad (86)$$

où $\hat{n}(G_i)$ est le nombre d'événements générés dans la catégorie i après une itération,⁸⁹ B_R est le nombre de catégories considéré au niveau reconstruit et $n(R_j)$ est le nombre d'événements observés dans la catégorie j dans les données (résultats de la section 15).

$P_0(G_i)$ est alors recalculé en normalisant $\hat{n}(G_i)$ par $\sum_{i=1}^{B_G+1} \hat{n}(G_i)$ et une nouvelle itération peut commencer si besoin.

Une fois les résultats stabilisés, les sections efficaces différentielles finales sont obtenues grâce à la formule suivante :

$$\frac{d\sigma_i}{dX} = \frac{1}{\Delta X \cdot \epsilon_{trig.,i} \cdot \mathcal{L}} \sum_{j=1}^{B_R+1} P_K(G_i|R_j) \cdot N_{data,j}^{\mathbf{TITI}} \cdot x_{\gamma\gamma,j}^{\mathbf{TITI}}, \quad (87)$$

où $N_{data,j}^{\mathbf{TITI}}$ est le nombre d'événements observés dans la catégorie j de l'échantillon **TITI**, $x_{\gamma\gamma,j}^{\mathbf{TITI}}$ est la pureté $\gamma\gamma$ associée déterminée dans la section 15, $\epsilon_{trig.}$ est l'efficacité du système de déclenchement (voir section 16.3.5), les $P(G_i|R_j)$ sont les facteurs issus de

⁸⁹Les corrections reliées au système de déclenchement sont exclues de la méthode de déconvolution itérative car l'estimation de l'efficacité de déclenchement est basée exclusivement sur les données (voir section 16.3.5) et donc factorisable. Celle-ci étant mesurée comme fixe et très proche de 100%, sa factorisation n'a aucun effet sur les résultats finals.

l'équation 85 après K itérations et $\mathcal{L}$ est la luminosité intégrée de l'échantillon de données (voir section 13). Pour les différents spectres, les sections efficaces sont également divisées par la largeur de la catégorie ΔX considérée afin d'obtenir une quantité indépendante du choix des catégories (sections efficaces différentielles).

L'implémentation de la procédure et les différents tests de validation associés sont présentés dans la section 16.5.2.

16.5.2 Test de la procédure et comparaison avec d'autres méthodes de déconvolution

L'implémentation de la méthode de déconvolution itérative disponible dans le code public RooUnfold [231, 250] est utilisée dans l'analyse. Dans un premier temps, les distributions de l'échantillon de signal SHERPA au niveau reconstruit sont déconvoluées afin de vérifier le bon fonctionnement du code. Après avoir constaté la bonne correspondance des distributions obtenues avec celles au niveau généré, la méthode est exécutée sur les distributions observées dans les données après soustraction du bruit de fond (voir section 15).

La variation des résultats en fonction du nombre d'itérations est alors examinée pour chaque observable de mesure, voir figure 159. La convergence est rapide pour toutes les observables et peut être considérée comme atteinte après 5 itérations, qui est le chiffre adopté dans l'analyse nominale.

L'incertitude statistique sur les résultats est également un paramètre important car celle-ci augmente significativement en fonction du nombre d'itérations, à cause des corrélations successives introduites entre les distributions déconvoluées et les matrices de réponse à chaque nouvelle itération (voir section 16.5.2). L'incertitude statistique finale est calculée directement dans le code RooUnfold en tenant compte des corrélations [251], ce qui permet de se passer de l'utilisation de répliques *bootstrap* (ou de *toy MC*) pour son évaluation. L'augmentation en fonction du nombre d'itérations est représentée pour les différentes observables de mesure dans la figure 160. Bien qu'une augmentation de l'incertitude statistique soit bien observée à chaque nouvelle itération, l'évolution la plus significative vient des deux premières itérations, qui changent également significativement les distributions finales. Par conséquent, le fait de prendre 5 itérations de façon nominale ne dégrade pas significativement la précision des mesures, l'augmentation de l'incertitude par rapport à 3 itérations étant au maximum de 10% soit $\ll 1\%$ par rapport aux valeurs centrales des sections efficaces considérées.

Finalement, les résultats de sections efficaces différentielles $\gamma\gamma + X$ obtenus avec la méthode de déconvolution itérative sont comparés avec ceux issus de la déconvolution catégorie par catégorie et une autre méthode de déconvolution appelée TUnfold [252] dans la figure 161. Tous les résultats sont obtenus à partir du code RooUnfold, qui implémente chacune de ces méthodes. Chaque procédure de déconvolution est testée au préalable avec les distributions de l'échantillon SHERPA au niveau reconstruit afin de vérifier leur bonne implémentation. Par ailleurs, les résultats de RooUnfold pour la déconvolution catégorie par catégorie ont été comparés avec une implémentation privée et trouvés en parfait accord.

Les méthodes TUnfold et itérative implémentent toutes les deux une procédure dite

de régularisation⁹⁰ et donc fournissent un meilleur traitement des migrations que la déconvolution catégorie par catégorie. En particulier, les résultats devraient être moins dépendants de la forme des distributions de l'échantillon simulé, qui dans la déconvolution catégorie par catégorie peuvent avoir un effet significatif sur les corrections via le facteur C_{fid} . (voir section 16.3.2).

Les résultats présentés dans la figure 161 sont très similaires pour les trois méthodes. Les méthodes TUnfold et itérative mènent à des résultats quasi identiques et suivent les mêmes variations de l'ordre de 0 à 3% avec la méthode de déconvolution catégorie par catégorie. Par conséquent, aucune incertitude systématique additionnelle intrinsèque au choix de la méthode de déconvolution n'est considérée. Comme annoncé précédemment, les sections efficaces différentielles $\gamma\gamma + X$ sont extraites à partir de la méthode itérative tandis que la section efficace fiducielle totale, quasi insensible aux migrations, est obtenue avec la déconvolution catégorie par catégorie.

Dans la section 16.6, les incertitudes finales reliées à la soustraction du bruit de fond (voir section 15) sont propagées dans la procédure de déconvolution et les incertitudes systématiques supplémentaires entrant en jeu lors de cette étape sont quantifiées soigneusement.

⁹⁰Celle-ci consiste à procéder à plusieurs itérations pour la déconvolution itérative. Plus d'informations sur les procédures de régularisation sont disponibles dans les références [244, 249, 252].

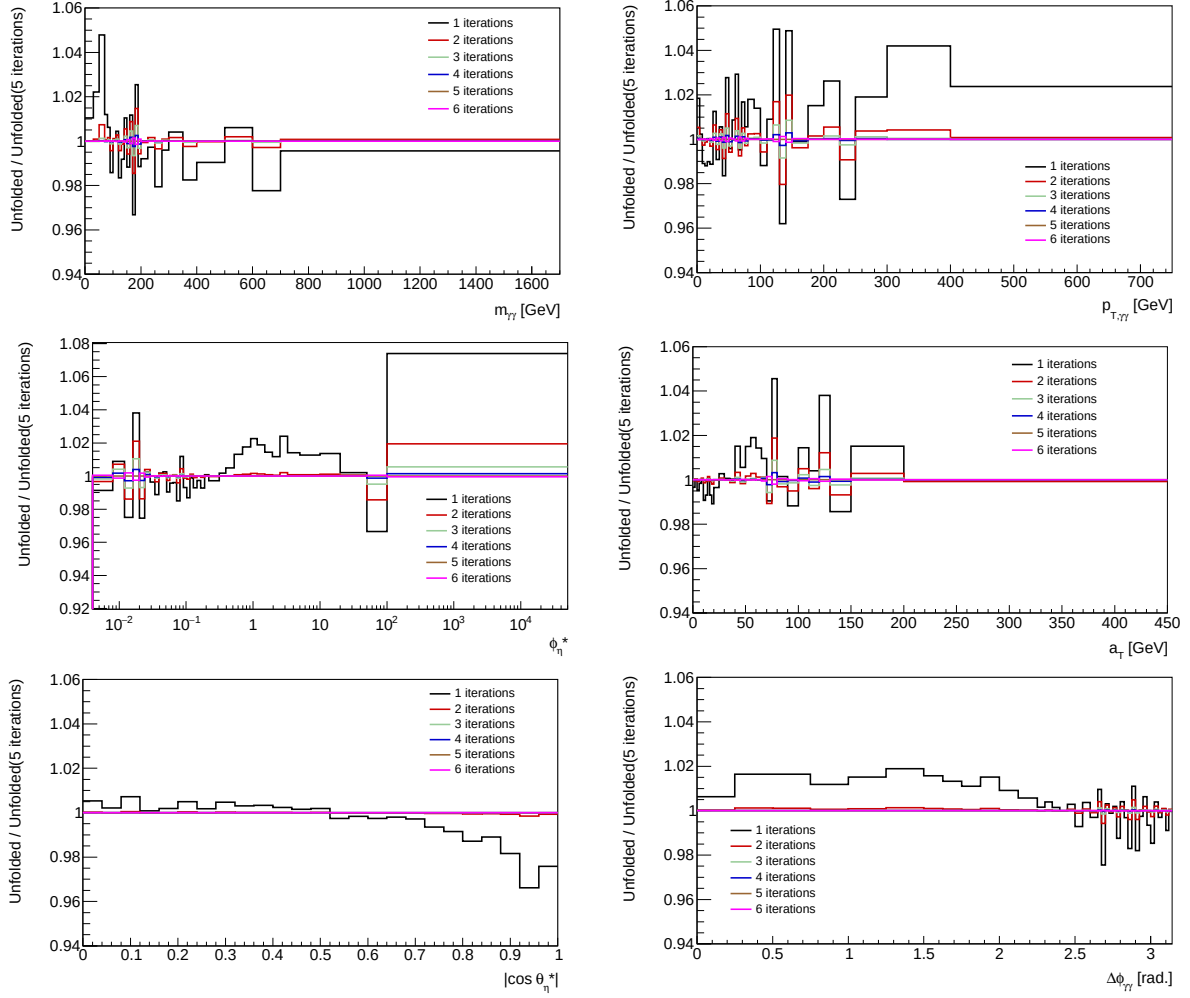


FIG. 159 – Distributions déconvoluées des observables de mesure avec la méthode itérative en fonction du nombre d'itérations, normalisés aux résultats obtenus après cinq itérations. Le code RooUnfold [250] et l'échantillon de signal simulé SHERPA sont utilisés.

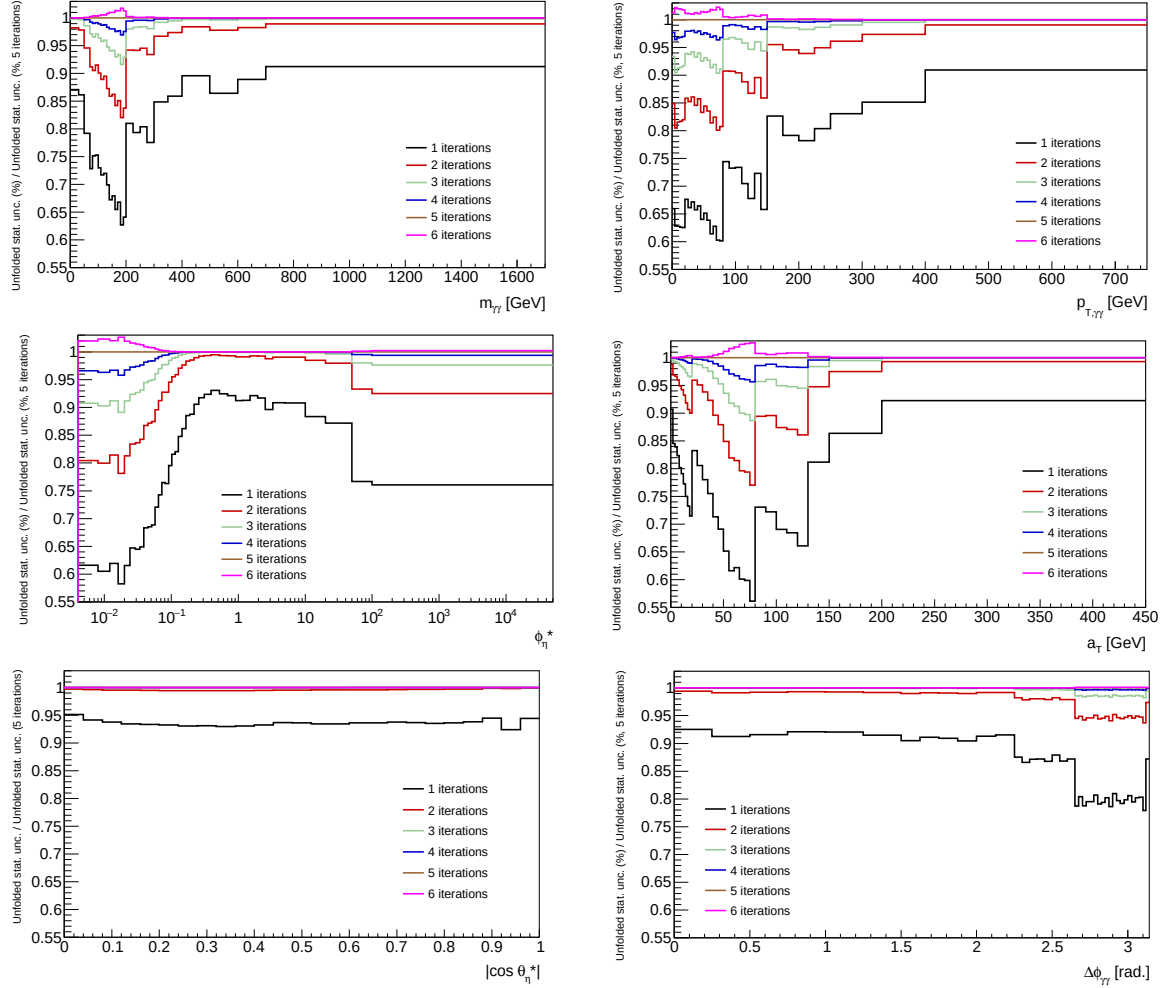


FIG. 160 – Évolution de l'incertitude statistique des distributions déconvoluées avec la méthode itérative pour les différentes observables de mesure en fonction du nombre d'itérations. Les résultats sont normalisés par les résultats obtenus après cinq itérations. Le code RooUnfold [250] et l'échantillon de signal simulé SHERPA sont utilisés.

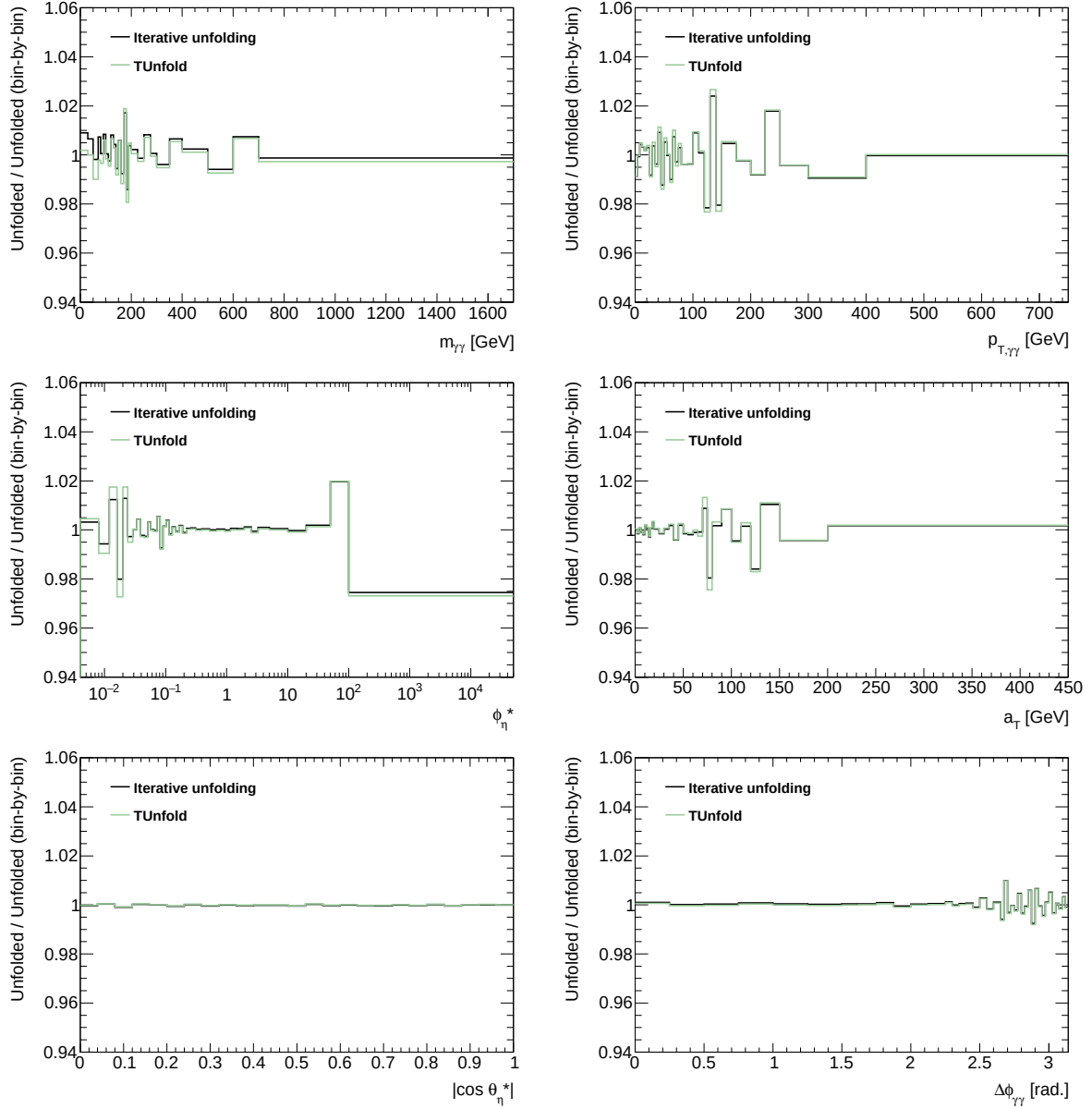


FIG. 161 – Comparaison des résultats déconvolués obtenus en utilisant la méthode de déconvolution itérative [231] (trait plein noire) et TUnfold [252] (trait plein vert) par rapport aux résultats de la déconvolution catégorie par catégorie. Le code ROOUNFOLD [250] et l'échantillon de signal simulé SHERPA sont utilisés.

16.6 Incertitudes sur la mesure finale

16.6.1 Classement et propagation des différentes incertitudes

Dans cette section, le traitement et la propagation des différentes incertitudes dans les mesures finales de section efficace sont abordés en détail. Il est rappelé que la section efficace fiducielle totale est obtenue à partir d'une déconvolution catégorie par catégorie (voir équation 77, section 16.2), tandis que les sections efficaces différentielles sont extraites par la méthode itérative afin d'améliorer le traitement des migrations (voir équation 87, section 16.5.1).

Dans les deux cas, l'étape de déconvolution implique deux entrées distinctes : les distributions observées après soustraction du bruit de fond (voir section 15) et les matrices de réponses (méthode itérative) ou de façon équivalente les facteurs de corrections catégorie par catégorie (déconvolution catégorie par catégorie, voir section 16.2). Les matrices de réponse/facteurs de correction, appelés sans distinction matrices de réponse dans la suite, incluent les effets de reconstruction et d'efficacité (alignement, calibration, identification, etc) mais également les formes des distributions MC générées, affectant les migrations (résolution). Ces deux effets sont factorisés explicitement uniquement dans le cadre de la méthode itérative.

Dans la suite, les différentes sources d'incertitudes sont classées selon cinq catégories :

- ◇ **Les effets systématiques ayant un impact significatif à la fois sur la soustraction du bruit de fond et sur les matrices de réponses**, principalement issus de la dépendance du *template fit* 2D vis-à-vis de l'échantillon de signal simulé via les pdfs de photons. Ces effets ont déjà été étudiés dans le cadre de la soustraction du fond dans la section 15.6.1 et ont été identifiés comme dominants. Ils doivent être propagés dans la mesure finale en tenant compte des déformations associées pour les matrices de réponse (corrélations). Ils représentent les sources d'incertitudes expérimentales dominantes sur les mesures finales et sont traitées dans la section 16.6.2.
- ◇ **Les effets systématiques ayant un impact significatif uniquement sur la soustraction du bruit de fond**, comme par exemple le choix de la définition de NON-TIGHT pour les régions de contrôle des jets. Ces effets ont déjà été évalués dans les sections 15.6.2 à 15.6.4 et doivent être propagés dans la mesure finale après déconvolution. Cette opération est réalisée dans la section 16.6.3.
- ◇ **Les effets systématiques ayant un impact significatif uniquement sur les matrices de réponses**. Ceux-ci incluent des effets fins de reconstruction comme par exemple la calibration des photons ou l'alignement du calorimètre. Ils sont détaillés dans la section 16.6.4.
- ◇ **Les effets systématiques factorisables**. Ceux-ci incluent les incertitudes sur l'efficacité du système de déclenchement 16.3.5 et sur la luminosité de l'échantillon de données (voir section 13.1). Ils seront abordés dans la section 16.6.5.
- ◇ **Les effets statistiques**, incluant région de signal (issus de la procédure HESSE

de MINUIT [206] pendant l'exécution du *template fit* 2D) et régions de contrôle (voir section 15.6.2). Ceux-ci ont la particularité d'être indépendants entre eux et décorrélés catégorie par catégorie,⁹¹ ce qui permet de pouvoir les propager automatiquement dans les résultats finals lors de la déconvolution grâce à la procédure implémentée dans RooUnfold [251]. Par ailleurs, la statistique limitée du MC a un impact non nul sur la soustraction du bruit de fond uniquement dans les queues de distribution, lorsque les effets statistiques sont totalement dominés par la région de signal. Par conséquent, les corrélations entre les effets statistiques du MC sur la soustraction du bruit de fond et sur les matrices de réponse sont négligées et propagées indépendamment en même temps que les effets statistiques dans les données grâce au code RooUnfold [251].⁹² Des détails supplémentaires sur cette étape sont fournis dans la section 16.6.6.

Finalement, un bilan général des incertitudes de la mesure sera présenté dans la section 16.6.7.

16.6.2 Incertitudes corrélées

Les effets systématiques ayant un impact significatif à la fois sur la soustraction du bruit de fond et sur les matrices de réponses incluent tous ceux abordés dans la section 15.6.1, à l'exception de la fusion des pdfs photons dans les queues de distribution (voir section 16.6.3) et de la statistique limitée de l'échantillon simulé (traitée avec les autres effets statistiques dans la section 16.6.6). Ils forment les incertitudes systématiques expérimentales dominantes des mesures $\gamma\gamma + X$ finales.

Étant donné que les corrélations catégorie par catégorie (*bin-by-bin*) pour ces différents effets n'ont pas été étudiées en détail, celles-ci sont supposées à 100% pour la déconvolution itérative afin de simplifier l'estimation des erreurs.⁹³ En pratique, cela revient à évaluer les différentes systématiques en répétant l'opération de déconvolution pour chaque paramètre de nuisance avec des distributions d'événements observés après soustraction du fond augmentées et abaissées d'une déviation standard pour la totalité des catégories d'un spectre donné, et ainsi évite l'usage systématique de *toy MC*. Étant donné les migrations très faibles pour toutes les observables (voir section 16.2), aucun effet significatif sur l'étape de déconvolution n'est attendu de corrélations éventuelles entre les différentes catégories.

En revanche, afin de tenir compte des corrélations entre l'effet sur les distributions d'événements et l'effet sur les matrices de réponse d'une même source d'incertitude, les matrices sont déformées à $\pm 1\sigma$ de façon synchronisée avec les distributions d'événements pour chaque paramètre de nuisance.

De façon similaire à ce qui avait été déjà observé dans la section 15.6.1, des fluctuations importantes de l'erreur systématique reliée au choix du générateur MC (SHERPA

⁹¹Les corrélations résiduelles entre les différentes catégories issues de la fusion de certaines régions de contrôle dans les données pour la soustraction du bruit de fond sont négligées pour la déconvolution. Les migrations étant faibles, cette simplification ne devrait avoir aucun effet sur les résultats finals.

⁹²La note de bas de page N° 91 s'applique également ici pour la fusion des régions **TITI** dans les échantillons simulés.

⁹³Il n'a pas été encore décidé quelles matrices de corrélation figureraient dans la publication finale.

vs PYTHIA8) sont présentes dans le cas des distributions différentielles si celle-ci est prise simplement comme la différence entre les deux résultats. Par conséquent, une procédure de lissage similaire à celle décrite dans la section 15.6.1 est adoptée pour les résultats finals après déconvolution. Les résultats de la procédure de lissage sont résumés dans la figure 162.

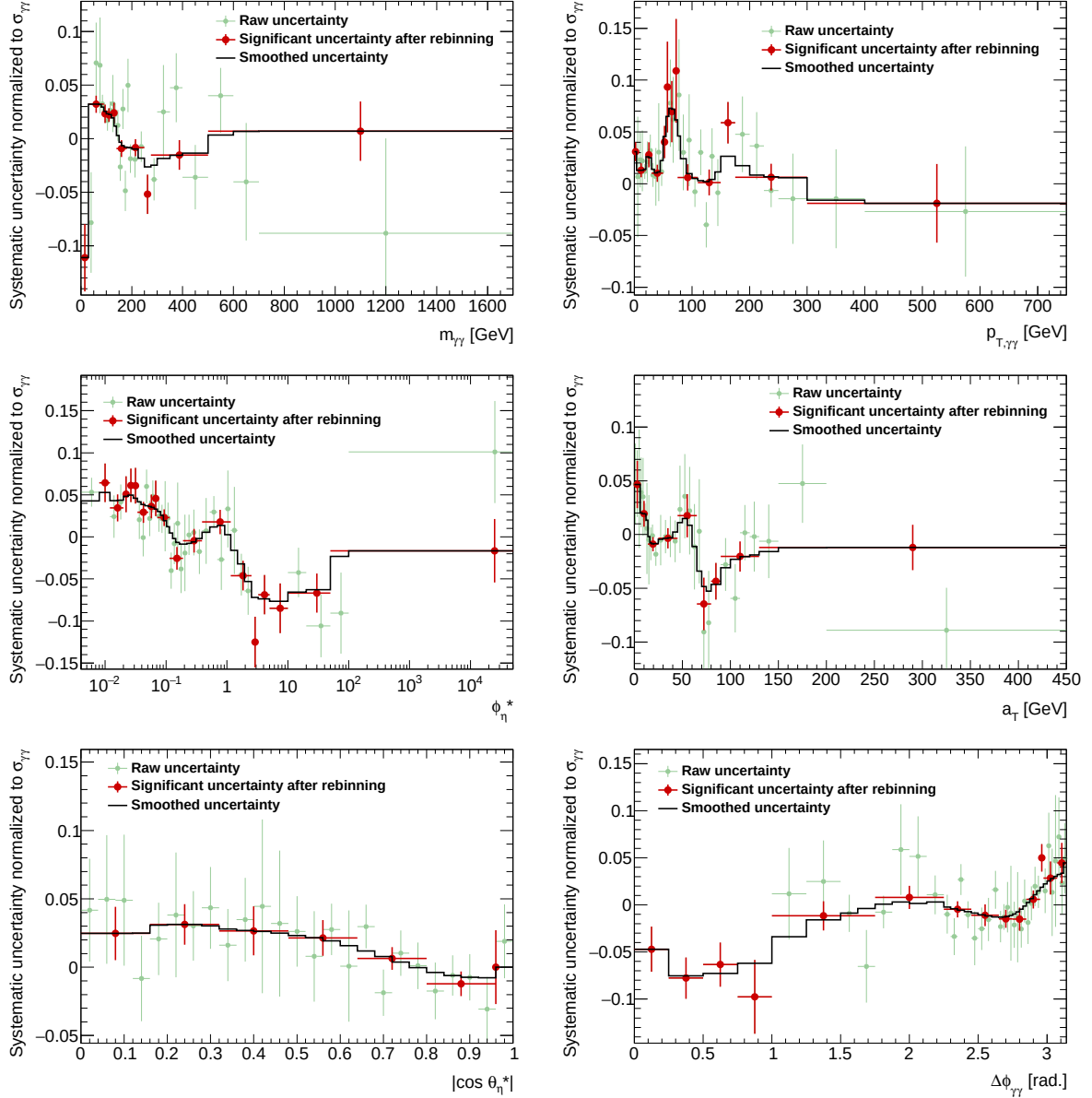


FIG. 162 – Résumé des différentes étapes de la procédure de lissage adoptée pour l'incertitude systématique finale sur les sections efficaces différentielles liée au choix du générateur MC. La soustraction du bruit de fond est effectuée avec la méthode du *template fit* 2D (voir section 15). Plus de détails sur la procédure de lissage sont disponibles dans la section 15.6.1.

Note : les barres d'erreur verticales vertes et rouges sont légèrement sous-estimées car celles-ci ne tiennent pas compte des fluctuations dues à la statistique limitée des régions de contrôle, ce qui explique quelques écarts entre les systématiques significatives et lissées.

Paramètre de nuisance	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (pb)	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (%)
Description MC de l'ID des photons	+0.465/-0.450	+2.81/-2.73
Dépendance vis-à-vis du modèle MC (Sherpa/Pythia)	+0.393/-0.336	+2.38/-2.04
Description MC de l'isolation (corr. de translation dans le calo.)	+0.324/-0.327	+1.97/-1.98
Description MC de l'isolation (corr. de largeurs dans le calo.)	+0.004/-0.000	+0.02/-0.00
Total	+0.690/-0.650	+4.18/-3.94

TAB. 32 – Incertitudes systématiques finales sur la section efficace fiducielle totale affectant à la fois la soustraction du bruit fond résiduelle dans l'échantillon **TITI** et les matrices de réponse. La soustraction du bruit de fond est effectuée avec la méthode du *template fit* 2D (voir section 15). Les effets statistiques seront traités dans la section 16.6.6. Les incertitudes dominantes de la mesure sont marquées en gras.

Les résultats pour la section efficace fiducielle totale sont résumés dans le tableau 32, tandis que ceux obtenus pour les sections efficaces différentielles sont montrés dans la figure 163 (figure 164 pour l'incertitude sur l'identification des photons).

L'effet systématique dominant vient désormais parfois de l'incertitude sur l'identification des photons, comme c'est le cas par exemple pour la région **TITI** complète (+2.81/-2.73%). En effet, l'identification des photons a un effet significatif à la fois sur les contaminations des régions de contrôle utilisées pour la soustraction du bruit de fond (voir section 15.3.4) et sur les matrices de réponse (inefficacités reliées à ϵ_{TT} , voir section 16.3.3). Par exemple, supposons que l'efficacité d'identification TIGHT dans le MC soit sous-estimée. Dans ce cas, à la fois les corrections de contamination dans les régions de contrôle et les corrections d'inefficacité sont surestimées, menant à un biais positif sur la section efficace d'une part à cause du biais positif sur la pureté $\gamma\gamma$ (corrections de contamination) mais également à cause des matrices de réponses (corrections d'inefficacité).

Les incertitudes issues du choix du générateur MC et des corrections d'isolation sont également importantes. Cependant, elles ne sont que légèrement amplifiées par les corrélations avec les matrices de réponse par rapport aux chiffres rapportés dans la section 15.6.1, indiquant que la plupart de l'effet systématique vient de la soustraction du bruit de fond et non des matrices. Finalement, la fusion des pdfs photons, affectant uniquement la soustraction du fond et non les matrices mais figurant tout de même sur la figure 163 car reliée au pdfs de photons, a un impact non nul uniquement dans les queues de distribution de $m_{\gamma\gamma}$ et $p_{\text{T},\gamma\gamma}$, dominées de toute façon par l'incertitude statistique (voir section 16.6.6).

Des pistes afin de réduire ces incertitudes dans les analyses $\gamma\gamma + X$ futures ont été évoquées dans la section 15.6, comme par exemple le fait d'optimiser la simulation de l'empilement par rapport à des variables calorimétriques. Concernant l'effet de l'identification des photons sur les matrices de réponse, des efforts considérables ont déjà été fournis par le groupe de performance avec les *ID scale factors*, qui sont considérés comme finals pour les données 2012. Cependant, des mesures du rapport $\mathbf{T}/\tilde{\mathbf{T}}$ pourraient être réalisées par exemple dans les données du $Z^0 \rightarrow l^+l^-\gamma$ afin de mieux contrôler la conta-

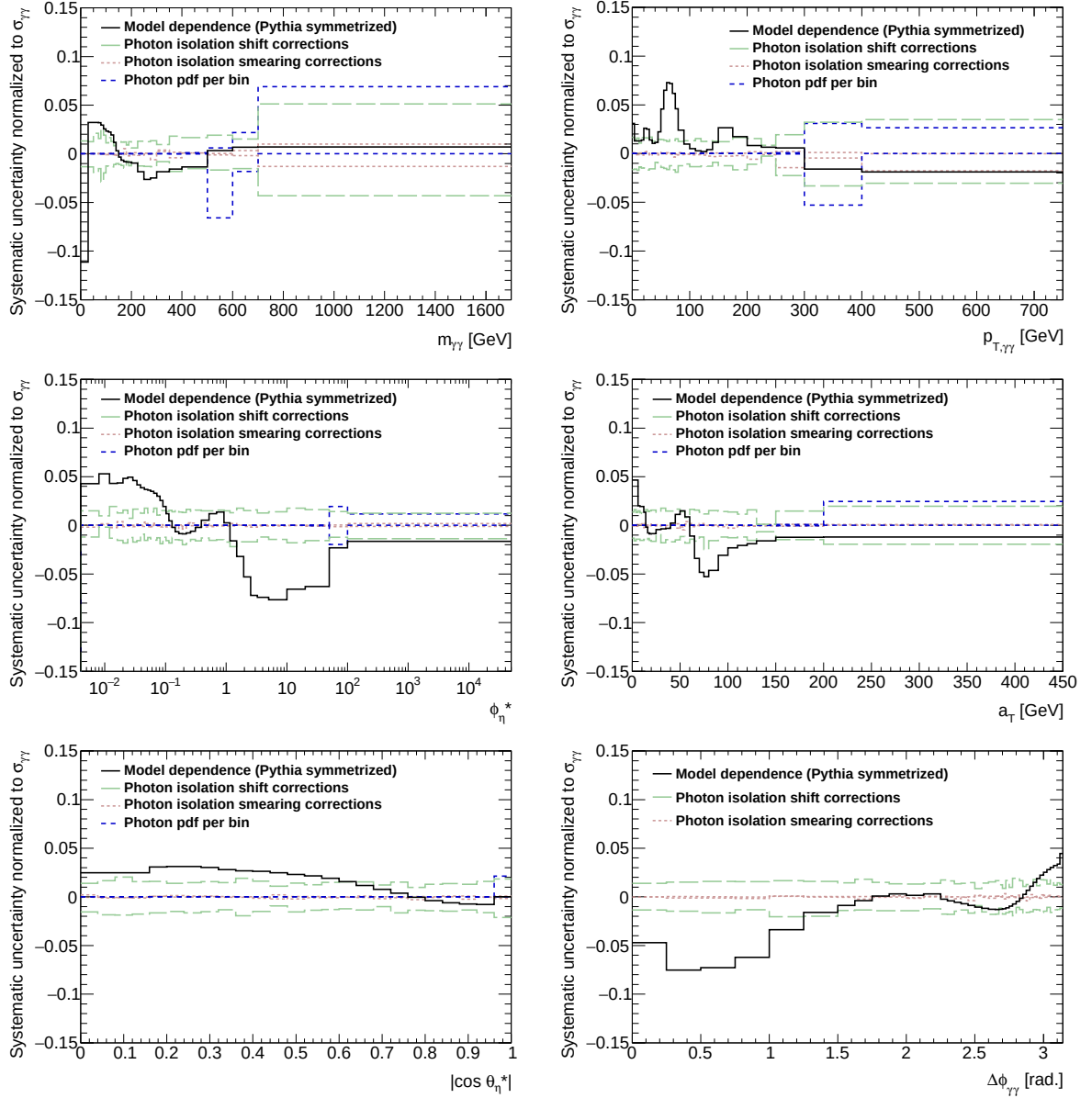


FIG. 163 – Incertitudes systématiques finales sur les différentes sections efficaces différentielles issues des paramètres de nuisance jouant à la fois sur la soustraction du bruit de fond résiduel dans l'échantillon **TITI** et sur les matrices de réponse, à l'exception de celle liée à la fusion des pdfs photons dans les queues de distribution (*Photon pdf per bin*) qui affecte uniquement la soustraction du bruit de fond. L'incertitude liée à l'identification des photons dans le MC (*MC photon ID description*) est incluse dans la figure 164 afin d'éviter la surcharge de celle-ci et ne pas mélanger les effets issus des photons et des jets. Plus d'informations sur l'origine des différentes sources sont disponibles dans la section 15.6.1. Les effets statistiques seront traités dans la section 16.6.6.

mination des régions de contrôle. Quelques réflexions personnelles plus générales sur le traitement des photons dans ATLAS ont été données dans la section 11.

Dans la section 16.6.3, les effets systématiques ayant un impact significatif uniquement sur la soustraction du bruit de fond sont discutés.

16.6.3 Incertitudes sur la soustraction du bruit de fond

Les effets systématiques ayant un impact significatif uniquement sur la soustraction du bruit de fond incluent tous ceux abordés dans les sections 15.6.2 à 15.6.4, à l'exception de la statistique limitée des régions de contrôle qui sera traitée avec les autres effets statistiques dans la section 16.6.6.⁹⁴ Ces incertitudes sont donc reliées à la modélisation du bruit de fond jets et électrons dans le *template fit* 2D (voir sections 15.6.2 et 15.6.3) ainsi qu'aux autres sources étudiées dans la section 15.6.4.

De façon similaire à la section 16.6.2, les corrélations catégorie par catégorie pour ces différents effets sont supposées à 100% pendant la déconvolution afin de simplifier le traitement des incertitudes. Étant donné les migrations très faibles (voir section 16.2), aucun effet significatif sur la déconvolution n'est attendu des corrélations éventuelles entre les différentes catégories.

Les résultats pour la section efficace fiducielle totale sont présentés dans le tableau 33, tandis que ceux pour les sections efficaces différentielles sont montrés dans les figures 164 à 166. Dans le cas inclusif comme pour la plupart des distributions selon les différentes observables de mesure, l'incertitude affectant uniquement la soustraction du bruit de fond est dominée par la définition des régions de contrôle de jets (NON-TIGHT), s'étalant de 1 à 8% avec le maximum de 8% atteint dans la dernière catégorie en $|\cos \theta_\eta^*|$.

Paramètre de nuisance	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (pb)	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (%)
Définition des régions de contrôle	+0.252/-0.281	+1.52/-1.70
Bande passante des KDEs	+0.045/-0.063	+0.27/-0.38
Bruit de fond électrons ($e\gamma$, γe)	+0.000/-0.119	+0.00/-0.72
Différence maximale avec les distributions différentielles	+0.000/-0.149	+0.00/-0.90
Total	+0.26/-0.35	+1.54/-2.09

TAB. 33 – Incertitudes systématiques finales sur la section efficace fiducielle totale reliées aux paramètres de nuisance affectant uniquement la soustraction du bruit de fond résiduel dans l'échantillon **TITI**. La source d'incertitude dominante provient de la définition des régions de contrôle de jets (mentionnée en gras). La soustraction du bruit de fond est effectuée avec la méthode du *template fit* 2D (voir section 15). Les effets statistiques seront traités dans la section 16.6.6.

⁹⁴Par ailleurs, la fusion des pdfs de photons dans les queues de distribution, qui devrait théoriquement être discutée ici, a été abordée dans la section 16.6.2 pour ne pas mélanger les effets issus des photons et des jets.

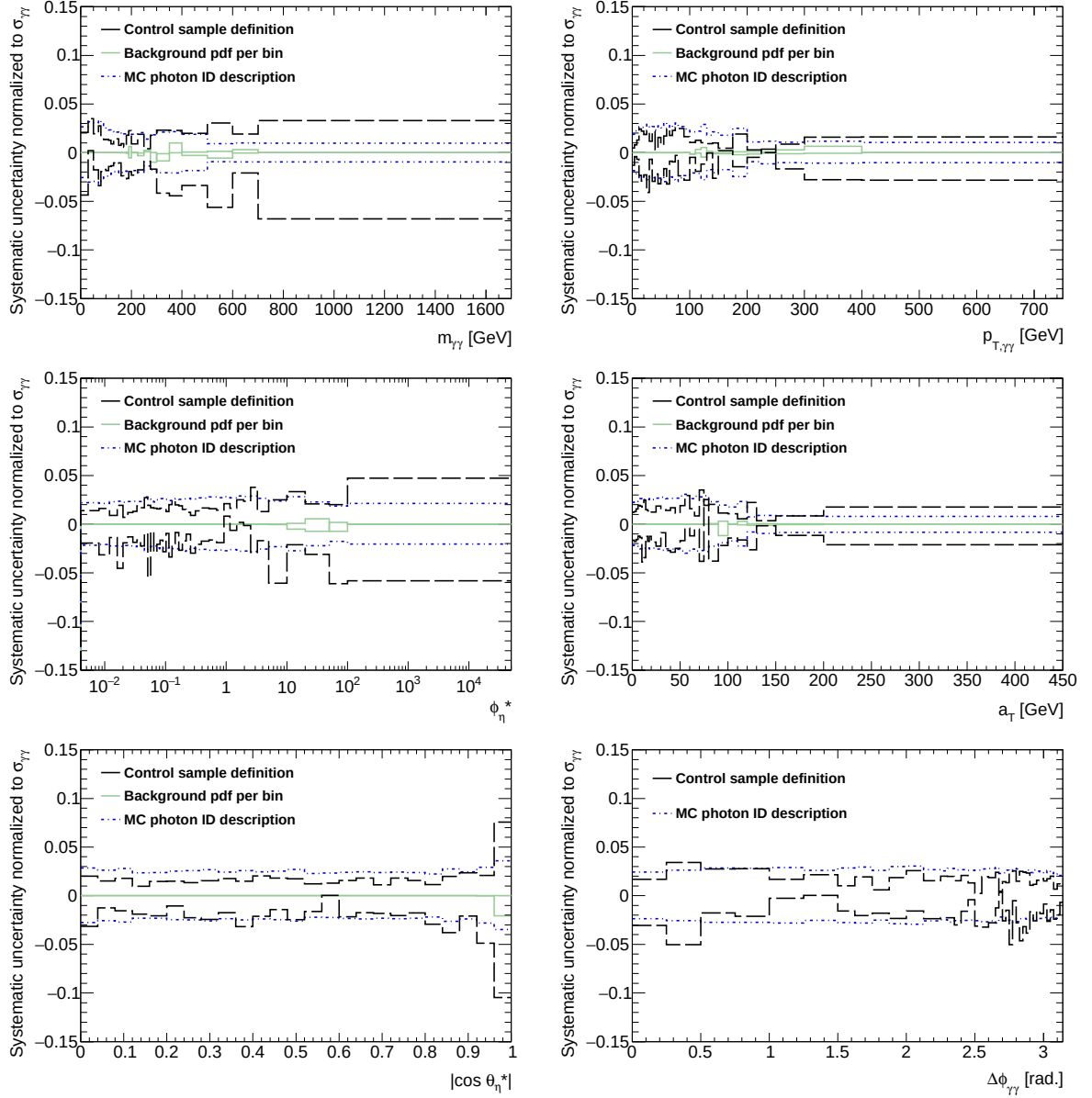


FIG. 164 – Incertitudes systématiques finales sur les différentes sections efficaces différentielles spécifiques à la modélisation des jets dans le *template fit* 2D (voir section 15.6.2) affectant uniquement la soustraction du bruit de fond résiduel dans l'échantillon **TITI**, à l'exception de la description de l'identification des photons dans le MC (*MC photon ID description*) qui affecte à la fois la soustraction du fond et les matrices de réponses. L'incertitude liée à la fusion des pdfs photons dans les queues de distribution (*Photon pdf per bin*) est incluse dans la figure 163 pour ne pas surcharger celle-ci et ne pas mélanger photons et jets. Les effets statistiques seront traités dans la section 16.6.6.

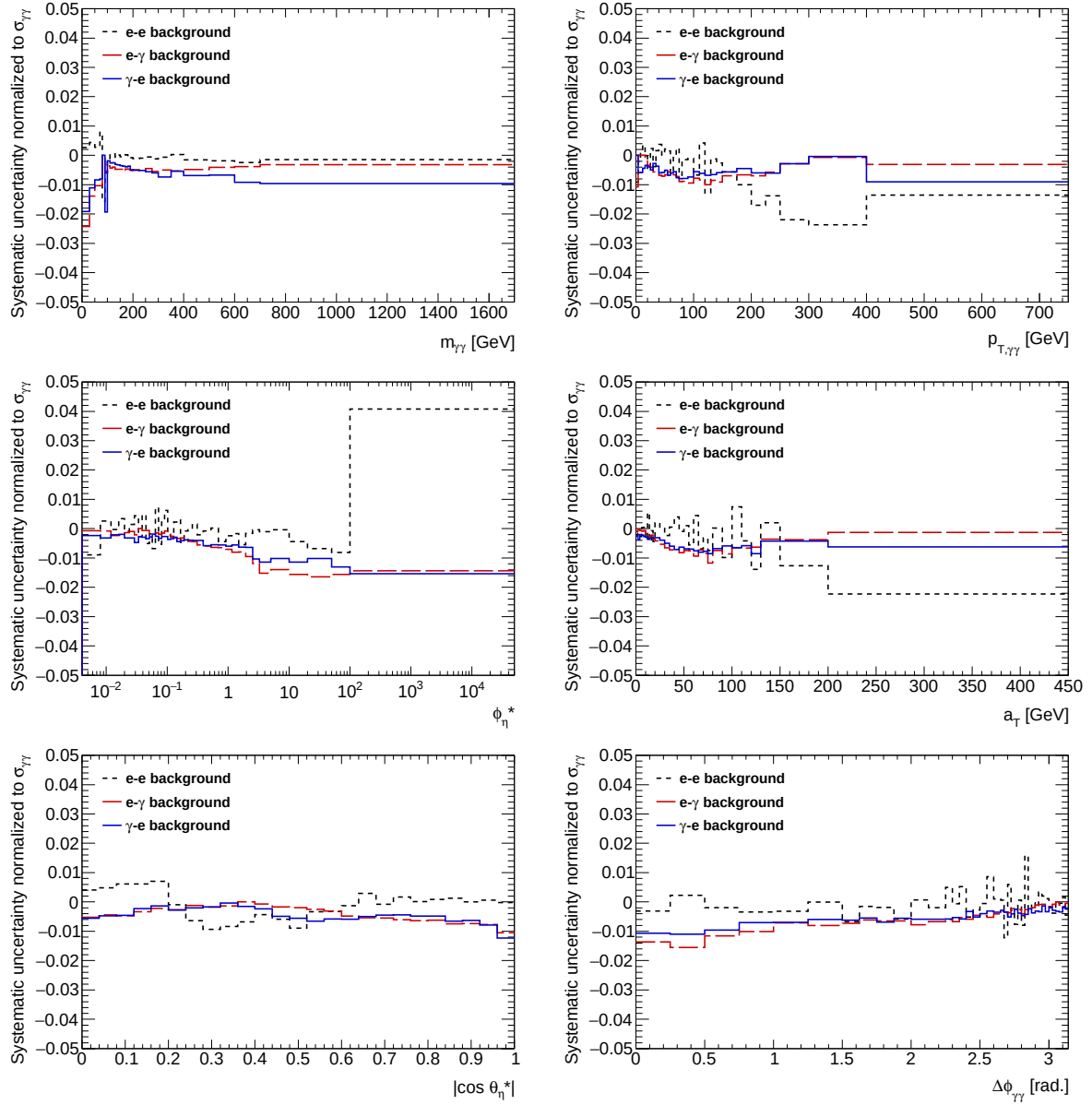


FIG. 165 – Incertitudes systématiques finales sur les différentes sections efficaces différentielles spécifiques à la modélisation des électrons dans le *template fit* 2D (voir section 15.6.3) affectant uniquement la soustraction du bruit de fond résiduel dans l'échantillon **T1T1**.

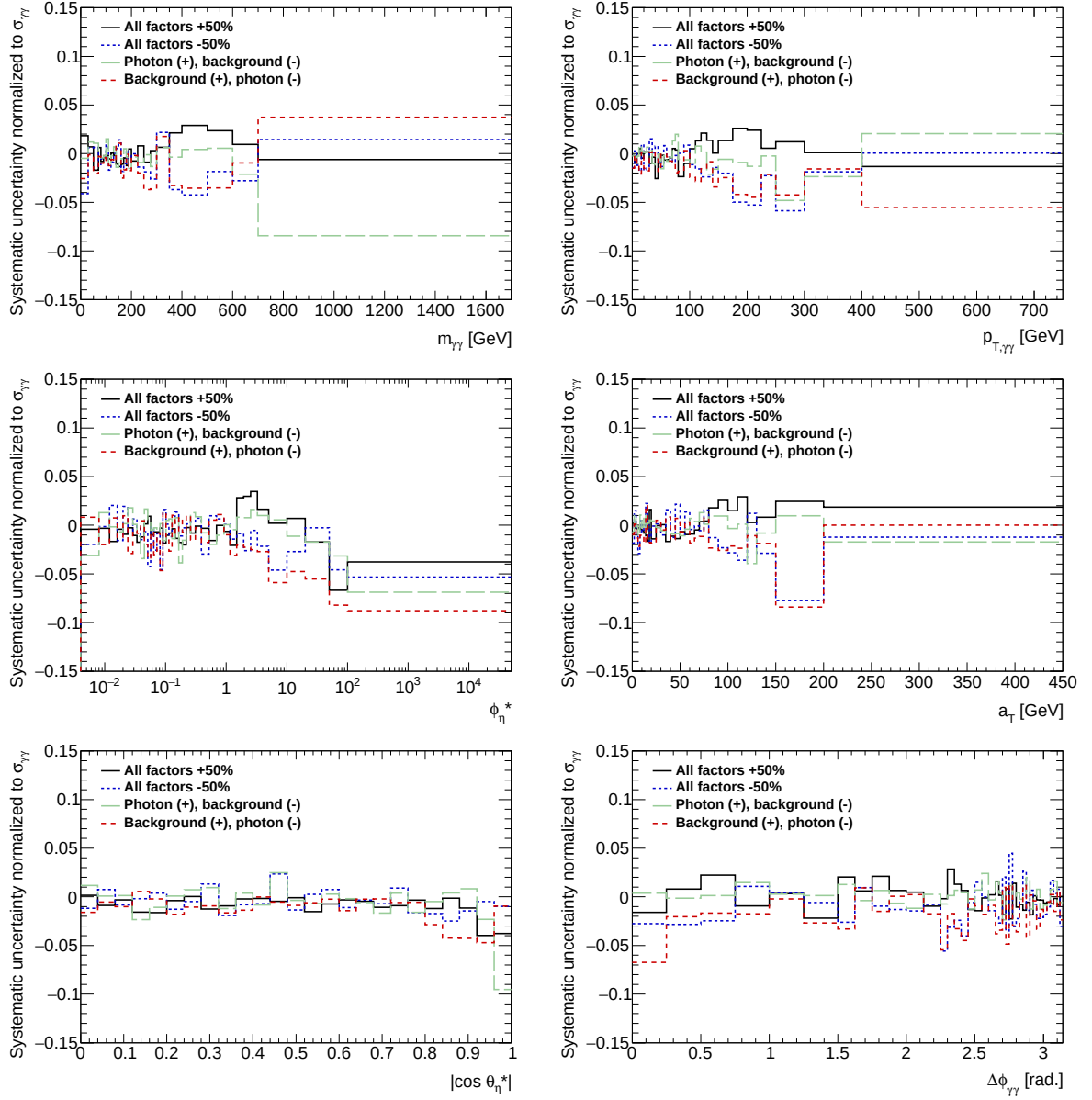


FIG. 166 – Variations des sections efficaces différentielles issues du choix arbitraire des bandes-passantes des KDEs dans le *template fit* 2D (voir section 15.6.4). La soustraction du bruit de fond est effectuée avec la méthode du *template fit* 2D (voir section 15. Les variations maximales sont considérées pour les incertitudes systématiques finales.

Dans le cas inclusif, les différences entre la soustraction inclusive et les soustractions différentielles resumées (voir section 15.6.4) ont également un certain effet sur l'incertitude, tandis que les distributions sont plus affectées par le choix de bande-passante des KDEs mais également par la fusion des pdfs de photons montrées précédemment dans la figure 163. Dans tous les cas, la modélisation des électrons joue un rôle sous-dominant.

Ces effets restent néanmoins en général de second ordre par rapport aux incertitudes traitées dans la section 16.6.2 et aux incertitudes statistiques dans les queues de distribution (voir section 16.6.6). Dans la section 16.6.4, les incertitudes spécifiques aux matrices de réponse sont détaillées.

16.6.4 Incertitudes de reconstruction

Dans cette section, les effets systématiques ayant un impact significatif uniquement sur les matrices de réponses sont détaillés. Étant donné que ceux-ci n'ont pas été discutés dans la section 15, chacun d'entre eux est introduit ci-dessous dans un court paragraphe.

Efficacité de la sélection sur l'isolation des traces

La sélection sur l'isolation des traces n'a pas d'effet significatif sur la soustraction du bruit de fond. En revanche, la connaissance limitée de son efficacité mène à un certain effet sur les matrices de réponse. De façon similaire à la section 16.6.2, les corrélations catégorie par catégorie sont supposées à 100% pendant la déconvolution afin de simplifier le traitement des incertitudes. Étant donné les migrations très faibles (voir section 16.2), aucun effet significatif sur la déconvolution n'est attendu de corrélations éventuelles entre les catégories.

L'incertitude sur l'efficacité de la sélection sur l'isolation des traces n'a pas été quantifiée précisément par le groupe de performance. Par conséquent, celle-ci a été estimée dans le cadre de l'analyse $\gamma\gamma + X$ dans la section 16.3.3. Une incertitude sur cette incertitude a également été extraite afin de permettre la mise en place d'une procédure de lissage limitant les fluctuations pour les décompositions selon $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc. Ainsi, la déconvolution est exécutée à nouveau avec les matrices de réponse déformées de $\pm 1\sigma$, mais également avec des matrices de réponse déformées de $\pm 1\sigma \pm 1\Sigma$, où Σ est "l'incertitude sur l'incertitude" pour les différentes observables. Une procédure de lissage similaire à celle décrite en détail dans la section 15.6.1 est finalement appliquée après déconvolution. Les résultats sont résumés dans la figure 167.

Dans le cas inclusif, l'incertitude liée à la sélection sur l'isolation des traces est de $\pm 0.28\%$ par rapport à la section efficace fiducielle totale, affectant donc très peu la précision finale de la mesure (voir section 16.3.3). Elle est également minimale dans la plupart des catégories, avec des valeurs comprises entre 1 et 2%. Elle atteint néanmoins 5% pour $\phi_\eta^* > 2$ et à bas $\Delta\phi_{\gamma\gamma}$, qui sont des régions particulièrement dominées par les photons de fragmentation. La précision de la mesure dans ces queues de distribution est néanmoins largement dominée par les incertitudes statistiques, comme nous le verrons dans la section 16.6.6.

Dans le paragraphe suivant, les effets reliés à la calibration du calorimètre électromagnétique sont discutés.

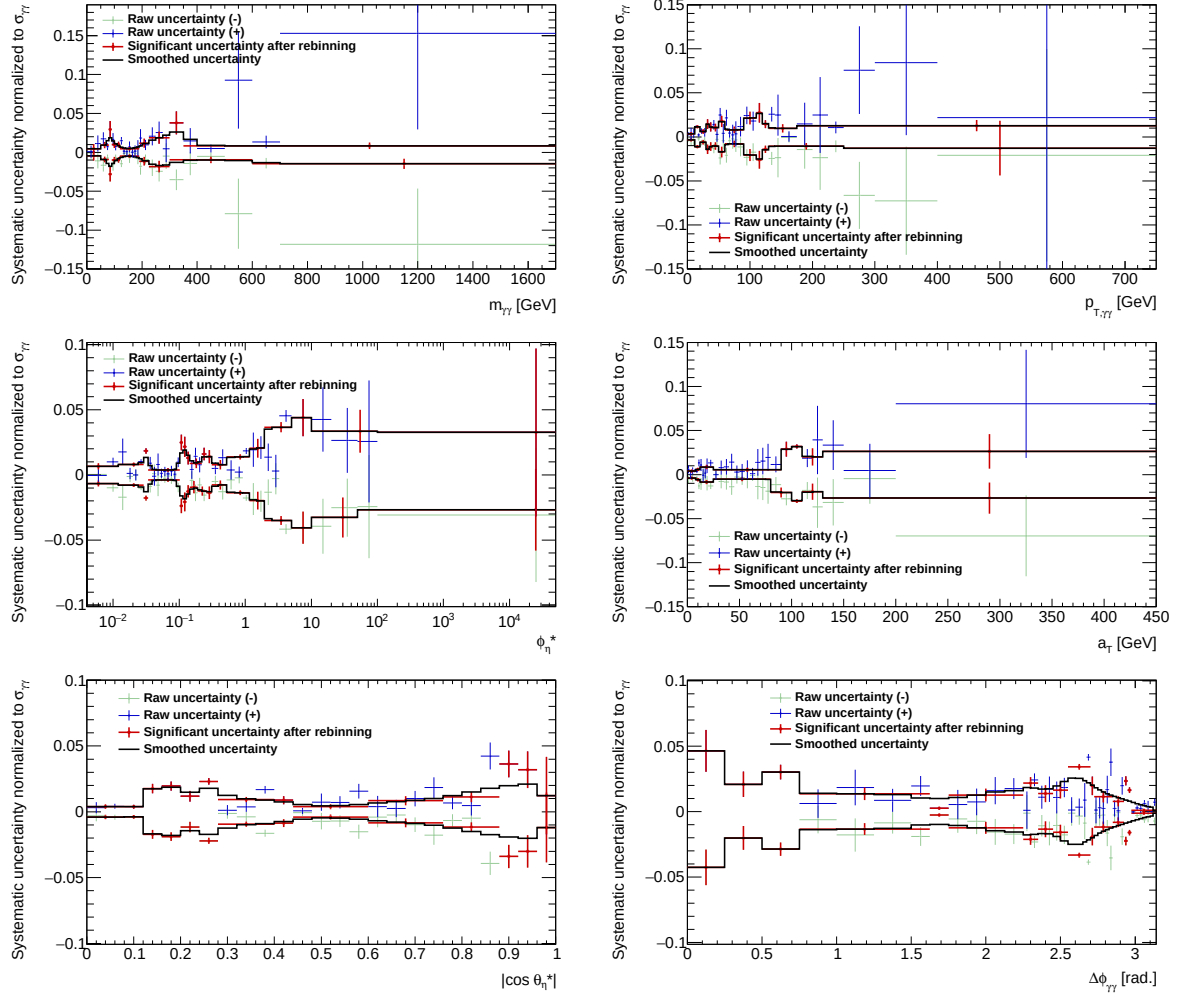


FIG. 167 – Incertitudes sur les sections efficaces différentielles issues de la sélection sur l'isolation des traces au cours des différentes étapes de la procédure de lissage. Les incertitudes brutes ainsi que leur barres d'erreur ont été calculées dans la section 16.3.3 et sont tout d'abord propagées dans la procédure de déconvolution (*raw uncertainties*). Une procédure similaire à celle décrite dans la section 15.6.1 est ensuite appliquée afin d'obtenir les incertitudes significatives (*significant uncertainties*), puis les incertitudes finales (*smoothed uncertainties*).

Calibration des photons

La chaîne de calibration des photons pour les données 2012 d'ATLAS a été détaillée dans la section 8. L'incertitude sous-jacente peut être décomposée en deux facteurs distincts, le premier étant relié à la détermination de l'échelle absolue en énergie du calorimètre électromagnétique et le second à sa résolution.

Le modèle de calibration final comporte respectivement 65 et 7 paramètres de nuisance indépendants pour l'échelle et la résolution. Ainsi, pour chacun de ces paramètres une nouvelle déconvolution est effectuée avec des matrices de réponse déformées à $\pm 1\sigma$ pour l'ensemble des catégories. Les variations avec le résultat nominal sont finalement ajoutées en quadrature aux incertitudes finales. Comme précédemment, les variations sont effectuées de façon cohérente entre toutes les catégories pour chacune des observables de mesure afin de simplifier le traitement des incertitudes. Dans le cas inclusif, l'effet résultant sur la section efficace fiducielle totale est sous-dominant et respectivement de $^{+0.65}_{-0.54}\%$ et $\pm 0.01\%$ pour l'échelle et la résolution.

Pour l'échelle en énergie, l'incertitude est dominée par l'effet du gain de la seconde couche du calorimètre à haut E_T^γ et par les constantes d'intercalibration entre le pré-échantillonneur et la première couche et entre la première et la seconde couche à plus bas E_T^γ , comme attendu des résultats de la section 8. La résolution est affectée légèrement par chacun des 7 paramètres de nuisance, avec une prépondérance du terme constant à haut E_T^γ .

Les résultats concernant les incertitudes issues de la sélection sur l'isolation des traces et de la calibration sont résumés dans le tableau 34 pour la section efficace fiducielle totale et dans la figure 168 pour les sections efficaces différentielles. Les incertitudes associées à ces effets sont globalement faibles, de l'ordre de quelques pourcents au maximum ($< 4\%$), et ont donc un impact très réduit sur la précision finale de la mesure. L'échelle absolue en énergie et la résolution du calorimètre sont moins bien déterminées à haut E_T^γ à cause du manque d'échantillons de photons purs dans les données, engendrant des extrapolations et donc une perte de précision (voir section 8). La largeur importante des catégories dans les queues de distribution, motivée d'abord pour des raisons statistiques, permet de limiter les effets de la calibration.

Paramètre de nuisance	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (pb)	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (%)
Échelle en énergie des photons	+0.107/-0.090	+0.65/-0.54
Résolution en énergie des photons	+0.002/-0.002	+0.01/-0.01
Description MC de l'isolation des traces	± 0.046	± 0.28
Total	+0.116/-0.101	+0.71/-0.61

TAB. 34 – Incertitudes systématiques finales sur la section efficace fiducielle totale reliées à la connaissance limitée de l'efficacité de la sélection sur l'isolation des traces et à la calibration du calorimètre, séparées en terme d'échelle en énergie absolue et de résolution. Ces paramètres de nuisance ont un impact uniquement sur les matrices de réponse et n'affectent pas de manière significative la précision de la mesure.

Dans le paragraphe suivante, l'effet des déformations géométriques et du mauvais alignement du calorimètre sont évalués.

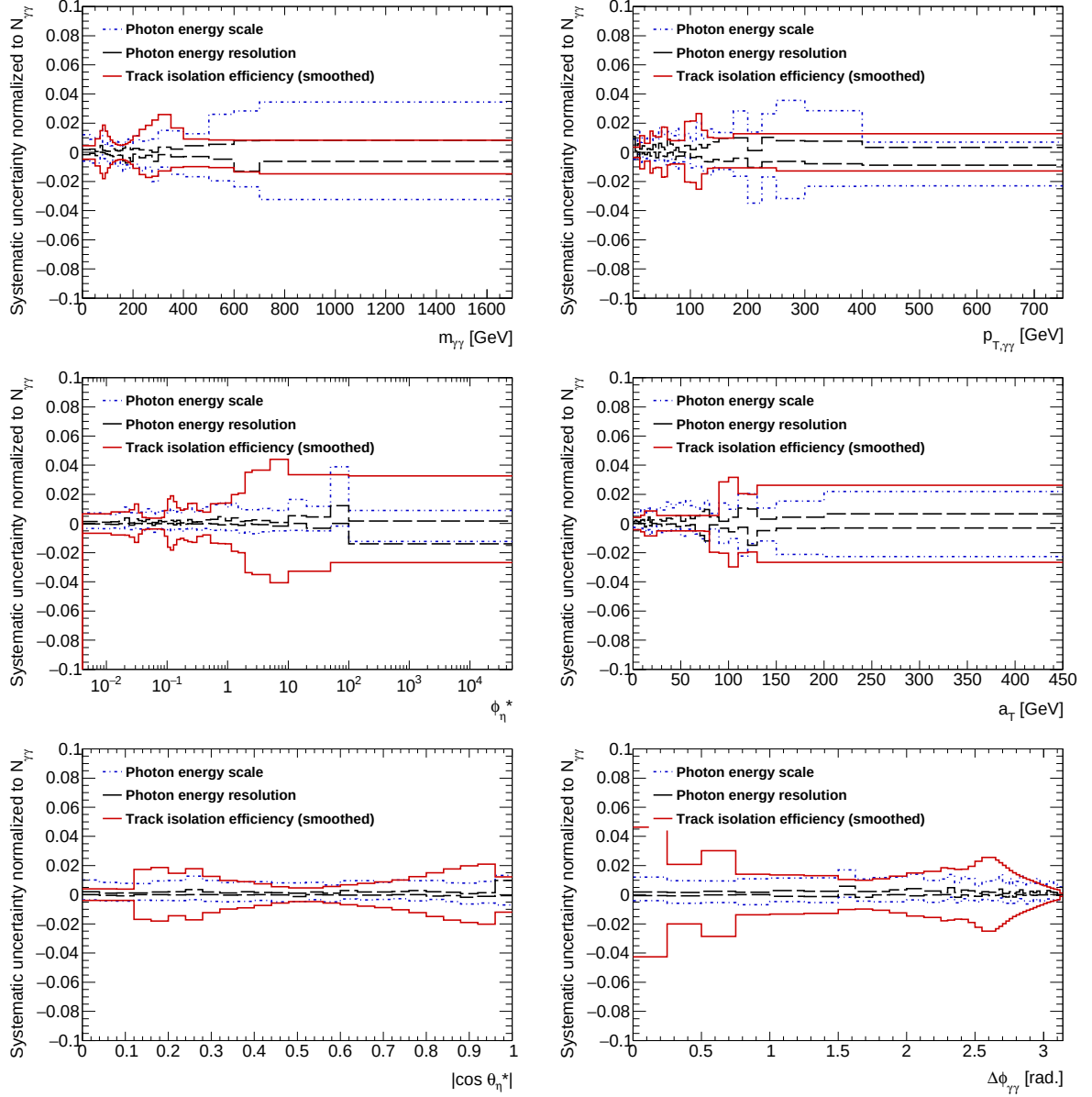


FIG. 168 – Incertitudes systématiques finales sur les différentes sections efficaces différentielles reliées à la connaissance limitée de l'efficacité de la sélection sur l'isolation des traces et à la calibration, séparées en terme d'échelle en énergie absolue et de résolution. Ces paramètres de nuisance ont un impact significatif uniquement sur les matrices de réponse et n'affectent pas de manière significative la précision de la mesure.

Déformations géométriques et mauvais alignement du calorimètre

Étant donné la dépendance des observables $|\cos \theta_\eta^*|$ et ϕ_η^* de l'analyse aux effets angulaires et en particulier à la pseudo-rapidité des photons (voir section 12.3), plusieurs effets systématiques liés à la géométrie du calorimètre électromagnétique ont été évalués. Ceux-ci sont négligés *de facto* dans la plupart des analyses d'ATLAS.

D'après les dernières études d'alignement du calorimètre d'ATLAS [253], les effets systématiques les plus significatifs susceptibles de jouer sur la résolution de $|\eta^\gamma|$ sont respectivement pour la partie tonneau et les bouchons une possible dilatation d'environ 3 mm tous les 3 m et une translation de l'ordre de ± 5 mm. Par conséquent, la procédure de déconvolution est exécutée à nouveau avec des matrices de réponses déformées correspondant à ces biais. L'alignement des deux bouchons n'étant pas corrélé, une translation de ± 5 mm est appliquée à l'un des bouchons dans la simulation et l'effet résultant est ajouté deux fois en quadrature à l'incertitude finale.

Les résultats sont résumés dans la tableau 35 pour la section efficace fiducielle totale et dans la figure 169 pour les sections efficaces différentielles. Ceux-ci sont toujours inférieurs à 1% et la plupart du temps de l'ordre de 0.3% ou moins. Ils n'ont par conséquent aucun impact sur les mesures finales. La dilatation du calorimètre engendre des variations positives de sections efficaces car elle a tendance à faire sortir les photons de l'acceptance en $|\eta^\gamma|$ (augmentation de $|\eta^\gamma|$) et donc à augmenter les corrections d'acceptance.

Paramètre de nuisance	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (pb)	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (%)
Dilatation du tonneau (calo.)	+0.030/-0.000	+0.18/-0.00
Mauvais alignement des bouchons (calo.)	+0.012/-0.013	+0.07/-0.08
Total	+0.032/-0.013	+0.19/-0.08

TAB. 35 – Incertitudes systématiques finales sur la section efficace fiducielle totale reliées aux déformations géométriques et aux incertitudes d'alignement du calorimètre. Ces paramètres de nuisance ont un impact uniquement sur les matrices de réponse et n'affectent pas de manière significative la précision de la mesure.

Dans le paragraphe suivant, l'effet systématique dû à la connaissance limitée de l'efficacité d'identification du vertex di-photon est évalué.

Efficacité d'identification du vertex di-photon

Pour chaque événement passant la présélection (voir section 14.5), le vertex d'émission de la paire de candidats photons est reconstruit à partir d'un algorithme basé sur un réseau de neurones décrit dans la section 14.4. La position en z de ce vertex est ensuite employée afin de corriger les pseudo-rapidités et impulsions transverses des candidats. Ces corrections sont propagées pour le calcul des différentes observables de mesure.

Cependant, dans certains cas, un mauvais vertex est identifié et le z_{PV} reconstruit diffère significativement du z_{PV} original. Par exemple, avec une tolérance de 0.3 mm (3 mm) sur z_{PV} par rapport à l'original, l'efficacité d'identification du vertex di-photon est de 82.6% (85.5%) selon l'échantillon de signal simulé SHERPA. D'après les études du $H \rightarrow \gamma\gamma$ discutées dans la section 14.4 [189, 241] basées sur des mesures dans les données autour du pic du boson Z^0 après repondération en $E_{\text{T},1}$, $E_{\text{T},2}$, l'efficacité devrait être

supérieure à 75%, en bon accord avec la simulation. Cela correspond bien à ce qui est obtenu avec l'échantillon de signal simulé SHERPA pour $\gamma\gamma + X$.

Néanmoins, afin d'évaluer l'effet qu'aurait une différence d'efficacité entre données et MC, la démarche suivante est adoptée : les poids des événements simulés présents dans l'échantillon **TITI** tels que $|z_{\text{PV}}^{\text{reco}} - z_{\text{PV}}^{\text{truth}}| > 0.3$ mm sont augmentés/abaissés de 5%, chiffre motivé par les valeurs d'efficacité de 75% et $\simeq 80\%$ mentionnées plus haut. Finalement, un facteur de normalisation est appliqué afin de préserver le nombre total d'événements dans la région **TITI** complète. Cela est justifié par le fait que la région **TITI** complète n'emploie aucune observable de mesure dans sa définition, or celles-ci représentent les seules variables montrant une certaine sensibilité aux corrections associées à z_{PV} . Par ailleurs la coupure fiducielle sur $\eta_1^\gamma, \eta_2^\gamma$ utilise la pseudo-rapacité calculée à partir de la seconde couche du calorimètre et donc non-corrigée de z_{PV} , pour des raisons géométriques. L'effet est donc par définition nul dans le cas inclusif. Pour les sections efficaces différentielles, les différentes matrices de réponse déformées sont produites et la procédure de déconvolution est répétée. L'incertitude résultante est montrée dans la figure 169.

Celle-ci est toujours inférieure à 1%, sauf à bas $p_{\text{T},\gamma\gamma}$ où elle atteint 1.4%. Elle passe par zéro une fois par observable à cause du facteur de renormalisation. L'effet est quasi nul pour $|\cos \theta_\eta^*|$ car cette distribution est très peu affectée par l'impulsion transverse des photons individuels ($E_{\text{T},1}$, $E_{\text{T},2}$) et du système di-photon ($p_{\text{T},\gamma\gamma}$), qui sont les paramètres principaux déterminant l'efficacité de sélection du vertex (voir section 14.4). En revanche, $\Delta\phi_{\gamma\gamma}$ y est sensible car l'algorithme d'identification est moins performant lorsque les particules formant le *recoil* sont très dispersées et de basse énergie ($\Delta\phi_{\gamma\gamma} \simeq \pi$ soit $p_{\text{T},\gamma\gamma} \simeq 0$).

Bien que n'ayant aucun effet sensible sur la précision de la mesure, il est intéressant de remarquer que cette incertitude est dans la plupart des catégories supérieure aux incertitudes de résolution du calorimètre en E^γ, η^γ (voir paragraphes précédents). Cela est directement relié au fait que les photons ne laissent souvent aucune trace dans le détecteur interne. Dans la section 16.6.5, la propagation des différents effets systématiques factorisables de la méthode de déconvolution est discutée brièvement.

16.6.5 Incertitudes factorisées

L'efficacité reliée au système de déclenchement et la luminosité intégrée collectée à $\sqrt{s} = 8$ TeV sont des quantités mesurées à partir d'échantillons indépendants dans les données et constantes pour l'ensemble des catégories employées dans le cadre de l'analyse $\gamma\gamma + X$. Par conséquent, celles-ci sont factorisées dans les différentes méthodes de déconvolution (voir sections 16.2 et 16.5). Elles sont également 100% corrélées entre toutes les catégories de la mesure ($m_{\gamma\gamma}$, $p_{\text{T},\gamma\gamma}$, **TITI** complète, etc).

L'efficacité du système de déclenchement a été étudiée dans la section 16.3.5 et le chiffre $99.44^{+0.14}_{-0.17}\%$ obtenu est adopté dans l'analyse. Par ailleurs, la luminosité intégrée collectée en 2012 par la chaîne `g35_loose_g25_loose` s'élève à $20.24 \pm 0.38 \text{ fb}^{-1}$ (mesure finale) [134]. L'effet du système de déclenchement sur la précision de la mesure est quasi-nul ($^{+0.17}_{-0.13}\%$). En revanche, la luminosité a un impact non-négligeable sur la section efficace fiducielle totale ($\simeq 1.9\%$). Un bilan des différentes incertitudes sera donné dans la section 16.6.7.

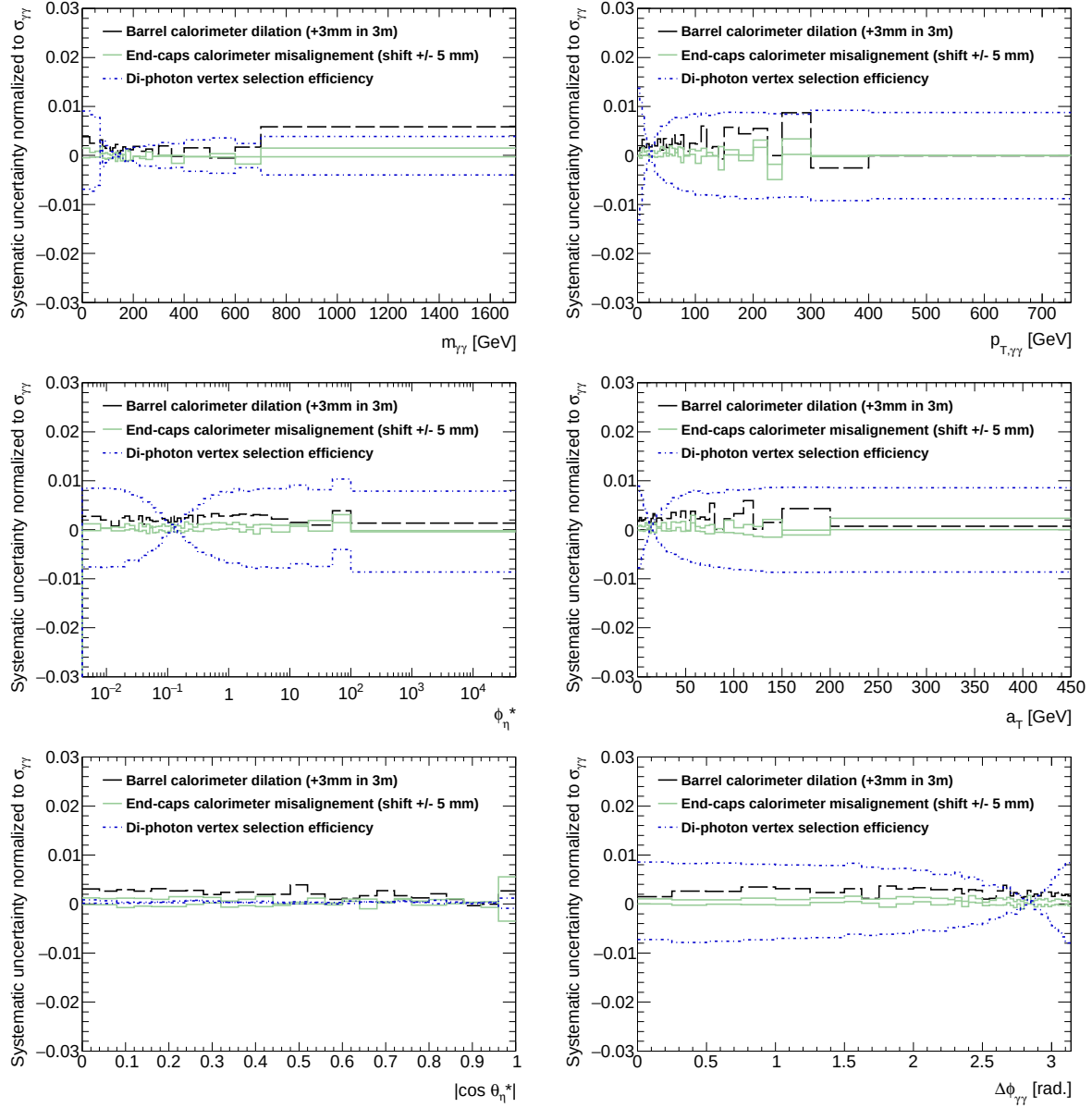


FIG. 169 – Incertitudes systématiques finales sur les différentes sections efficaces différentielles liées aux déformations géométriques, aux incertitudes d'alignement du calorimètre, ainsi qu'à l'efficacité d'identification du vertex di-photon. Ces paramètres de nuisance ont un impact significatif uniquement sur les matrices de réponse et n'affectent pas de manière significative la précision de la mesure.

Dans la section 16.6.6, la propagation des différents effets statistiques dans les résultats finals est abordée.

16.6.6 Incertitudes statistiques

Plusieurs effets statistiques indépendants affectent la précision de la mesure, à savoir la statistique limitée dans les données, incluant région de signal **TITI** et régions de contrôle de jets ($\tilde{\mathbf{T}}$), et la statistique limitée dans l'échantillon de signal simulée. La statistique des données a un effet uniquement sur les distributions observées tandis que celle du MC a un impact à la fois sur la soustraction du bruit de fond et sur les matrices de réponses. L'effet de la statistique MC sur la soustraction du fond est négligeable dans le cas inclusif et non nul uniquement dans les queues de distribution pour les différentes observables de mesure. (voir section 15.6).

Ainsi, dans le cas inclusif, l'incertitude statistique des données est propagée en répétant simplement l'opération de déconvolution avec un nombre d'événements observés après soustraction du bruit de fond augmenté et diminué des différents effets (voir section 15.6) tandis que l'effet du MC est estimé en faisant varier le facteur de correction C_{total} . Les résultats sont présentés dans le tableau 36. L'incertitude statistique est complètement dominée par la statistique limitée des données dans la région de signal.

Paramètre de nuisance	Variation de $\sigma_{tot}^{fid.}$ (pb)	Variation de $\sigma_{tot}^{fid.}$ (%)
Statistique (région de signal, données)	± 0.072	± 0.44
Statistique (régions de contrôle, données)	± 0.018	± 0.11
Statistique (MC)	± 0.003	± 0.02
Total	± 0.074	± 0.45

TAB. 36 – Incertitudes statistiques finales sur la section efficace fiducielle totale.

Pour la décomposition selon les différents spectres de mesure, la propagation est en principe plus compliquée car les itérations corrélient les distributions observées et les matrices de réponse au fur et à mesure de la procédure. Cependant, le code RooUnfold intègre une procédure propageant les effets statistiques en tenant compte des corrélations créées par les différentes itérations [251], ce qui permet de s'affranchir de l'utilisation de répliques *bootstrap* ou de *toy MC* tout en reproduisant correctement l'augmentation de l'incertitude avec le nombre d'itérations (voir section 16.5.2).

Pour utiliser cette implémentation, il faut néanmoins que les effets statistiques affectant les distributions observées et les matrices de réponse soient indépendants au début de la procédure, ce qui n'est a priori pas le cas pour la composante MC. Afin de simplifier la propagation des incertitudes statistiques pour les mesures de sections efficaces différentielles finales, les corrélations entre les effets de l'échantillon simulé sur les distributions observées et sur les matrices de réponse sont négligées et la procédure implémentée dans RooUnfold est utilisée. Cette simplification est justifiée par le fait que l'incertitude statistique globale sur la soustraction du bruit de fond est largement dominée par la statistique limitée dans la région de signal dans la très grande majorité des catégories considérées pour la mesure, qui est une composante non-corrélée avec les matrices de réponse. Cela est illustré dans la figure 170.

Dans la section 16.6.7, un bilan est effectué sur les incertitudes finales de la mesure.

En particulier, les distributions des incertitudes statistiques et systématiques finales sont comparées.

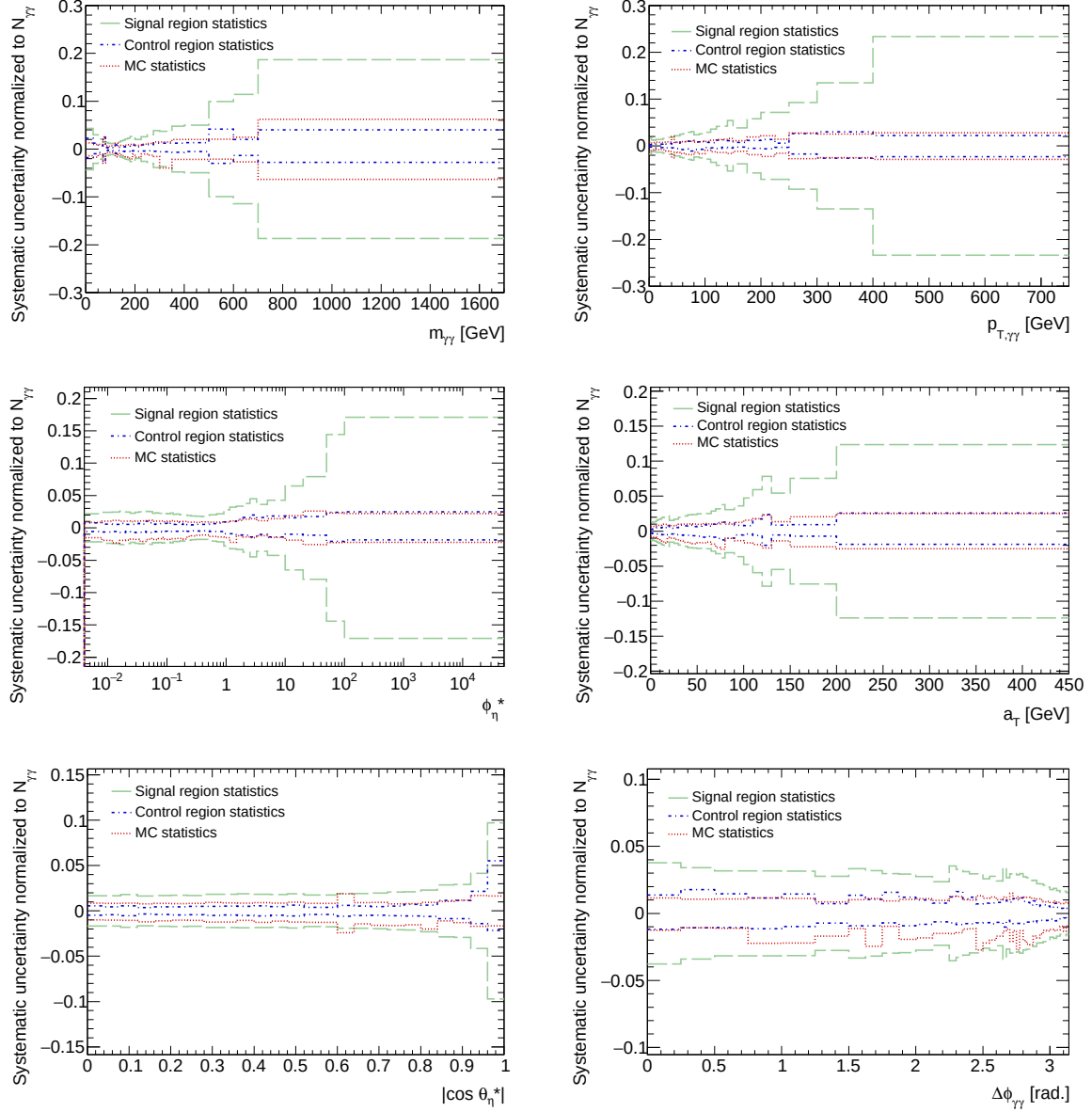


FIG. 170 – Comparaison des différents effets statistiques ayant un effet sur la soustraction du bruit de fond avec la méthode du *template fit* 2D. Les résultats sont exprimés en terme de variation d'événements $\gamma\gamma$ attendus et tirés de la section 15.6.

16.6.7 Bilan

La mesure finale de la section efficace fiducielle totale $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV effectuée avec le détecteur ATLAS est donnée ci-dessous :

$$\sigma_{\text{tot}}^{\text{fid.}} = 16.51 \pm 0.07(\text{stat}) \begin{smallmatrix} +0.75 \\ -0.74 \end{smallmatrix}(\text{syst}) \begin{smallmatrix} +0.32 \\ -0.31 \end{smallmatrix}(\text{lumi}) \begin{smallmatrix} +0.03 \\ -0.02 \end{smallmatrix}(\text{trigger}) \text{ pb} \quad (88)$$

Comme attendu, la précision expérimentale est limitée par l'incertitude systématique ($\pm 4.5\%$), suivie de la luminosité ($\pm 1.9\%$). Les incertitudes statistiques ($\pm 0.4\%$) et reliées à l'efficacité du système de déclenchement ($\pm 0.1\%$) sont en revanche négligeables. Les incertitudes dominantes dans le cas inclusif sont classées par ordre décroissant dans le tableau 37.

Incertitudes principales	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (pb)	Variation de $\sigma_{\text{tot}}^{\text{fid.}}$ (%)
Description MC de l'identification des photons	+0.465/-0.450	+2.81/-2.73
Dépendance vis-à-vis du modèle MC (SHERPA/PYTHIA)	+0.393/-0.336	+2.38/-2.04
Description MC de l'isolation (corr. de translation dans le calo.)	+0.324/-0.327	+1.97/-1.98
Luminosité intégrée (incertitude finale)	+0.320/-0.308	+1.94/-1.87
Définition des régions de contrôle	+0.252/-0.281	+1.52/-1.70
Total	+0.801/0.772	+4.85/-4.68

TAB. 37 – Incertitudes dominantes sur la section efficace fiducielle totale.

L'effet venant de l'identification des photons pourrait probablement être facilement réduit du côté de la soustraction du bruit de fond en effectuant des mesures du rapport $\tilde{\mathbf{T}}/\mathbf{T}$ dans les données $Z^0 \rightarrow l^+l^-\gamma$. Malheureusement, les mesures réalisées avec les données par le groupe de performance concernent pour le moment uniquement la sélection TIGHT (voir section 15.6.1). Des discussions avec le groupe de performance sont en cours à ce sujet. Des pistes pour réduire l'erreur expérimentale associée à la différence SHERPA/PYTHIA ont été données dans la section 15, tandis que quelques réflexions personnelles sur le traitement des photons dans ATLAS qui pourraient permettre une amélioration de l'isolation calorimétrique ont été évoquées dans la section 11 (soustraction de *clusters* topologiques au lieu de *clusters* de taille fixe, etc).

L'incertitude totale citée ($\simeq \pm 5.0\%$) est fortement réduite comparée à la mesure précédente d'ATLAS ($^{+7.3}_{-9.5}\%$) [203], grâce notamment à une région de signal mieux enrichie en photons immédiats (isolation des traces, sélection en $E_{\text{T},1}$, $E_{\text{T},2}$) mais également grâce à une meilleure description de l'isolation calorimétrique (corrections de fuites supplémentaires) et à une réduction de l'incertitude reliée à l'identification des photons par le groupe de performance (*ID scale factors*), même si des mesures complémentaires du rapport $\mathbf{T}/\tilde{\mathbf{T}}$ dans les données auraient bénéficiées à l'analyse.

Les incertitudes finales pour les différentes observables de mesure sont représentées dans la figure 171. Les conclusions changent peu par rapport à la section 15.6.5 car

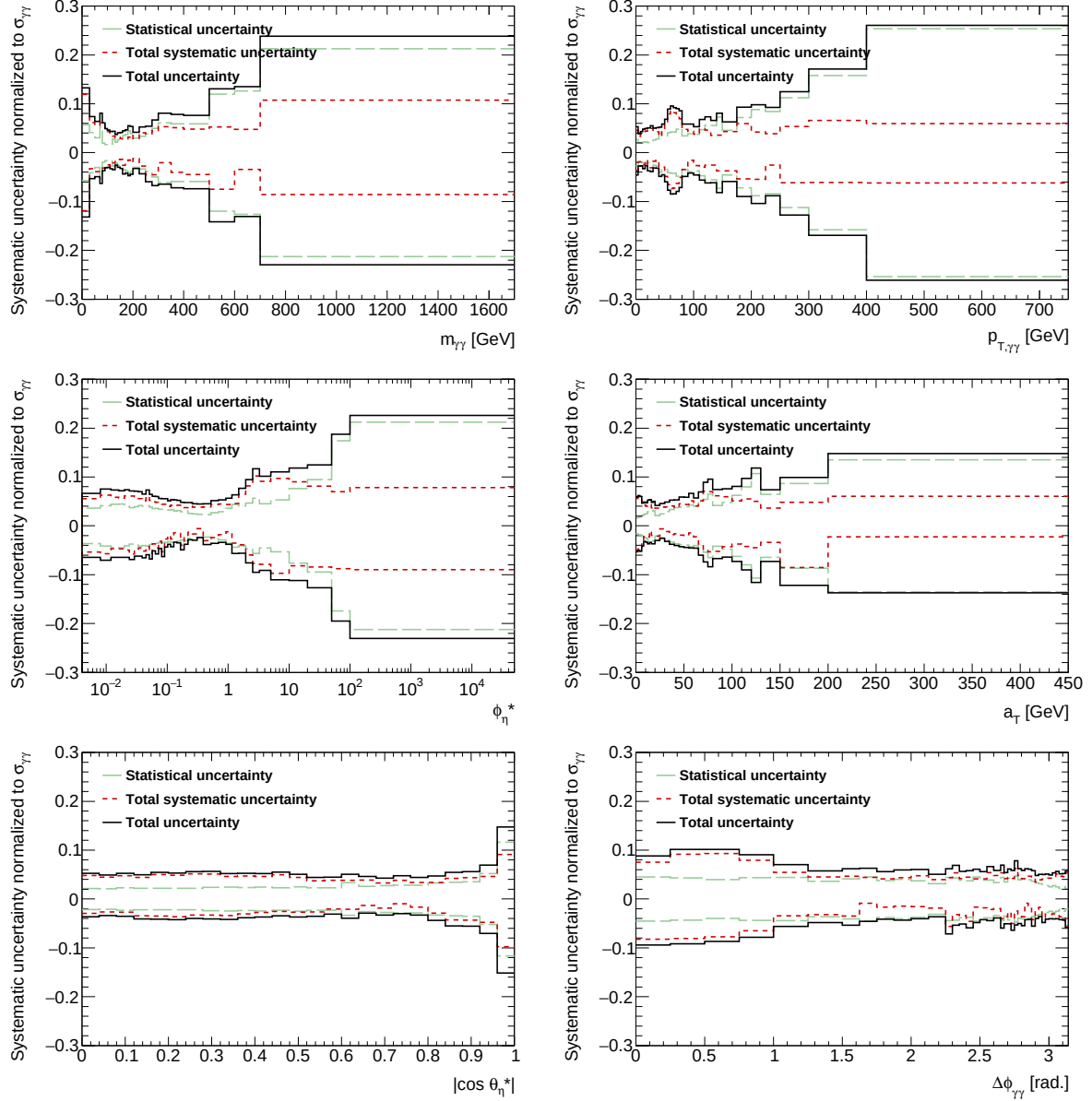


FIG. 171 – Incertitudes systématiques finales sur les différentes sections efficaces différentielles, décomposées en terme d'effets statistiques (trait pointillé vert, voir section 16.6.6) et systématiques (trait pointillé rouge). Les effets statistiques incluent ici statistique limitée des régions de signal et des régions de contrôle (données) ainsi que celle du MC. Ils sont non-corrélés entre les différentes catégories (*bin-by-bin uncorrelated*). L'incertitude totale (trait plein noir) correspond à la somme en quadrature des incertitudes statistiques et systématiques.

les incertitudes dominantes sont souvent issues de la soustraction du bruit de fond. Les distributions des incertitudes statistiques (trait pointillé vert) incluent désormais la statistique limitée des régions de signal et des régions de contrôle (données) ainsi que celle du MC, tandis que les incertitudes systématiques incluent les incertitudes de reconstruction et les incertitudes factorisables (voir sections 16.6.4 et 16.6.5). Cette classification est motivée par le fait que tous les effets statistique sont non corrélés entre les différentes catégories (*bin-by-bin uncorrelated*).

La mesure est dominée par les effets systématiques, sauf dans les queues de distribution où les incertitudes statistiques ont en général un poids plus important. À nouveau, les incertitudes sont du même ordre de grandeur ou meilleures comparées à la mesure précédente d'ATLAS [203], avec par ailleurs une portée étendue en $m_{\gamma\gamma}$ et $p_{T,\gamma\gamma}$ et une segmentation doublée pour $\Delta\phi_{\gamma\gamma}$ et $|\cos \theta_{\eta}^*|$.

Les régions incluant une fraction importante de photons de fragmentation et/ou d'états finals avec de nombreux partons additionnels ($m_{\gamma\gamma} < 70$ GeV, épaule de Guillet, $\phi_{\eta}^* > 1$, $\Delta\phi_{\gamma\gamma} \ll \pi$) ont généralement une incertitude plus élevée, attendu de la moins bonne discrimination entre photons de fragmentation et jets hadroniques dans le calorimètre par rapport aux photons directs mais également de la fraction de photon de fragmentation présents dans l'échantillon et de leur profil d'isolation plus incertains.

D'autre part, toutes les mesures réalisées avec la méthode de soustraction de bruit de fond alternative (méthode *4x4 matrix*, voir section 15.7) sont compatibles avec les résultats obtenus avec le *template fit* 2D, assurant d'autant plus de la fiabilité des résultats. Au contraire, des incertitudes supplémentaires avaient dû être considérées dans la mesure précédente d'ATLAS à $\sqrt{s} = 7$ TeV [203] à cause de différences pas bien comprises entre les différentes méthodes de soustraction, illustrant les difficultés posées par l'estimation précise du bruit de fond pour cette analyse. Le test des deux méthodes de soustraction avec le jeu de pseudo-données (voir section 15.4) a été déterminant afin d'arriver à des chiffres compatibles.

Dans la section 17, les mesures sont comparées avec plusieurs calculs théoriques. Une conclusion générale sur l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV et ses résultats est ensuite donnée en fin de partie.

17 Comparaison des mesures aux prévisions théoriques et conclusion

17.1 Prédictions théoriques

17.1.1 Considérations générales

Dans les sections 17.1.2 à 17.1.5, quelques détails sont donnés sur les différentes prédictions théoriques considérées pour la production $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV, incluant lorsque cela est possible le traitement des incertitudes théoriques associées. Les prédictions ont été sélectionnées de façon à constituer un éventail représentatif de l'état de l'art concernant les calculs $pp \rightarrow \gamma\gamma + X$ pour le LHC.

En règle générale, les mêmes coupures fiducielles que celles utilisées dans l'analyse sont appliquées pour générer les prédictions (voir section 12.2). Cependant, les générateurs partoniques (DIPHOX, RESBOS, 2 γ RES, 2 γ NNLO et MCFM) n'ont pas accès à l'énergie d'isolation transverse au niveau particule. De plus, les prédictions NNLO (2 γ RES, 2 γ NNLO et MCFM) incluent uniquement la composante directe et donc utilisent obligatoirement un critère d'isolation de Frixione pour assurer la convergence du calcul (voir section 3.5.3).

Par conséquent, la démarche suivante est adoptée :

- ◊ DIPHOX [111] est un calcul NLO partonique à ordre fixe en α_s incluant toutes les composantes : directe, 1-fragmentation, 2-fragmentations (voir section 3.5). Par conséquent, l'énergie d'isolation transverse des photons au niveau partonique (et non particule) est accessible. La coupure fiducielle d'isolation est donc effectuée sur cette variable au lieu de l'énergie d'isolation transverse au niveau particule. Le volume fiduciel obtenu est ensuite corrigé des effets de physique non-perturbative entrant en jeu lors du passage parton $\rightarrow$ particule (hadronisation, événement sous-jacent) en utilisant les échantillons de signal simulés SHERPA 1.4.0 et PYTHIA 8 (voir section 13.2).

Note : les corrections non-perturbatives sont actuellement en cours de finalisation. Elles ne sont pas incluses dans cette première version du manuscrit. L'effet attendu est de l'ordre de quelques pourcents.

- ◊ RESBOS [98] est un calcul NLO à ordre fixe en α_s avec resommation au NNLL pour $p_{T,\gamma\gamma} \simeq 0$. Pour $p_{T,\gamma\gamma} > 0$, la section efficace resommée est combinée avec les ordres supérieurs NLO. Uniquement la composante 1-fragmentation est considérée. Celle-ci est obtenue à l'aide d'une paramétrisation approchée des fonctions de fragmentation au NLO de manière à régulariser la production directe [101]. l'énergie d'isolation transverse des photons au niveau partonique étant accessible, les corrections non-perturbatives calculées pour DIPHOX sont également employées pour corriger le volume fiduciel de RESBOS.

Note : voir note ci-dessus.

- ◊ Les prédictions NNLO se limitent au calcul de la composante directe et donc utilisent nécessairement un critère d'isolation de Frixione pour en assurer la conver-

gence ($2\gamma\text{RES}$, $2\gamma\text{NNLO}$, MCFM). Dans ce cas, les recommandations mentionnées dans [121, 254] sont appliquées pour le choix des paramètres, soit :

$$\forall r < R_0, \quad E_T^{\text{had}}(r) < E_T^{\text{iso,max}}(r) = E_T^{\text{iso,max}} \left(\frac{1 - \cos r}{1 - \cos R_0} \right)^n, \quad (89)$$

avec $E_T^{\text{iso,max}} = 11 \text{ GeV}$, $R_0 = 0.4$ et $n = 1$. Cependant, la bonne correspondance entre ce critère et le critère appliqué dans l'analyse basée sur un cône de taille fixe ($E_T^{\text{had}} < 11 \text{ GeV}$) est avérée uniquement pour les petites valeurs de coupure ($E_T^{\text{had}} < 5 \text{ GeV}$ ou moins). Pour les valeurs de coupure supérieures, des écarts de l'ordre de 10% peuvent être obtenus, le critère de Frixione étant plus restrictif que celui basé sur un cône de taille fixe. On s'attend à ce que ces écarts dépendent significativement de la cinématique. Les corrections NNLO étant cependant de l'ordre de 40 à 50% et donc bien supérieures aux différences dues au critère d'isolation, elles devraient permettre dans tous les cas une nette amélioration de l'accord entre données et théorie.

Rappel : dans l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8 \text{ TeV}$, il a été décidé de considérer une valeur de coupure élevée sur l'isolation au niveau particule (11 GeV) afin de ne pas faire dépendre la mesure des corrections de déconvolution basées sur l'échantillon de signal simulé SHERPA (voir section 16.4), qui inclut a priori une description naïve de l'énergie d'isolation au niveau particule.

Note : des études sont actuellement en cours afin de déterminer le critère de Frixione correspondant le mieux au volume fiduciel de la mesure (valeur du paramètre n). Pour cela, l'idée est de comparer les prédictions NLO de DIPHOX (cône fixe) et $2\gamma\text{NNLO}$ (cône de Frixione). Pour le moment, le critère mentionné ci-dessus est utilisé sans corrections additionnelles ($n = 1$).

- ◇ Pour les générateurs incluant un algorithme de cascade de partons (SHERPA 1.4.0, PYTHIA 8, HERWIG 7 et SHERPA 2.3 α), les coupures fiducielles exactes de l'analyse sont appliquées, *ie.* la sélection sur l'énergie d'isolation transverse est effectuée au niveau particule dans un cône de taille fixe pour chacun des photons ($E_{T,1,2}^{\text{part.iso.}} < 11 \text{ GeV}$, voir section 12.2). Cependant, les générateurs utilisant des éléments de matrice au NLO (HERWIG 7 et SHERPA 2.3 α) nécessitent une coupure sur l'isolation de type Frixione pour assurer leur convergence. Celle-ci est alors suffisamment peu contraignante de manière à ne pas affecter le volume fiduciel final de la mesure.

Les différents calculs considérés sont abordés dans les sections 17.1.2 à 17.1.6. Dans la section 17.2, les comparaisons entre données et théorie seront commentées. Finalement, la partie IV s'achèvera par une conclusion générale sur l'analyse et ses résultats dans la section 17.3.

17.1.2 SHERPA 1.4.0 (*ME+PS merged at LO*) et PYTHIA 8 (*LO+PS*)

Dans la section 17.2, la mesure commencera par être comparée aux échantillons de signal simulés produits centralement par la collaboration ATLAS et employés dans l'analyse, soit l'échantillon $\gamma\gamma$ généré avec SHERPA 1.4.0 (*ME+PS merged at LO*) [83] et celui correspondant généré avec PYTHIA 8 (*LO+PS*) [82]. La description complète

de ces échantillons est disponible dans la section 13.2.2.

L'avantage principal de ces prédictions est qu'elles ont été générées avec une simulation complète du détecteur ATLAS, ce qui permet une comparaison avec les données au niveau reconstruit directement et ainsi une réduction des incertitudes expérimentales. La plupart des incertitudes de déconvolution devraient alors théoriquement être comptées au niveau des prédictions (incertitudes de simulation), sauf l'incertitude liée au choix de la *prior* pour la déconvolution, contribuant à l'incertitude intitulée "*Model dependence*" dans la section 16.6, qui elle, disparaît bel et bien (sous-dominante, voir section 16.6.2).

À cause de leurs éléments de matrice limités ainsi que des différents retours d'expérience des analyses $\gamma\gamma + X$ précédentes au LHC [202, 203, 224], on ne s'attend pas à ce que les prédictions de SHERPA 1.4.0 et PYTHIA 8 reproduisent l'échelle des sections efficaces mesurées. En particulier, aucune correction virtuelle n'est prise en compte par les éléments de matrice LO de ces générateurs alors que celles-ci contribuent significativement à l'augmentation de la section efficace [111]. Par ailleurs, des prédictions théoriquement bien meilleures (NLO, NNLO, etc) sont disponibles et seront discutées dans les sections 17.1.3 à 17.1.6. Par conséquent, les incertitudes théoriques liées aux prédictions de SHERPA 1.4.0 et PYTHIA 8 ne sont pas calculées pour les comparaisons finales.

En revanche, la forme des distributions en $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc, de SHERPA 1.4.0 et PYTHIA 8 est importante pour l'étape de déconvolution car celle-ci a un impact sur les migrations entre les différentes catégories de la mesure et donc sur les corrections de résolution finales (voir section 16). Les formes des distributions des observables de mesure sont en théorie plus sensibles à la modélisation de la composante de fragmentation (soit l'algorithme de cascades de partons ici) et aux corrections réelles en α_s , correspondant à l'émission de partons supplémentaires dans l'état final, plutôt qu'aux corrections virtuelles. SHERPA 1.4.0 inclut par exemple les éléments de matrice $\gamma\gamma + 1, 2$ jets au LO, correspondant à des corrections réelles en α_s . Par conséquent, les prédictions seront repondérées à la section efficace fiducielle totale mesurée pour les comparaisons afin de comparer les formes uniquement. Les sections efficaces totales fiducielles prédites par SHERPA 1.4.0 et PYTHIA 8 étant respectivement de 12.53 pb et 11.81 pb avec une incertitude statistique négligeable, les facteurs d'échelles sont d'environ 1.3 et 1.4, respectivement.

Dans la section 17.1.3, les prédictions issues de générateurs MC avec cascades de partons de dernière génération sont décrits.

17.1.3 HERWIG 7 et SHERPA 2.3 α (*ME+PS merged at NLO*)

Les générateurs MC de dernière génération incluent désormais des algorithmes de fusion et d'appariement entre éléments de matrice et algorithmes de cascade de partons permettant l'inclusion d'éléments de matrice NLO dans les prédictions, c'est-à-dire tenant compte des diagrammes correspondant aux premières corrections d'ordre supérieur en α_s à la fois réelles et virtuelles. Ce véritable tour de force permet de disposer de calculs robustes à la fois dans les régions de l'espace des phases sensibles aux instabilités infrarouges grâce aux cascades de partons ($p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* \simeq 0$, $\Delta\phi_{\gamma\gamma} \simeq \pi$) et dans les régions dominées par les corrections d'ordre supérieur en α_s grâce aux corrections NLO

(haut $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, etc). Ces calculs ont déjà permis de décrire avec succès les productions de $W^\pm$, Z et $t\bar{t}$ + jet(s) au LHC [106–108]. Ils représentent l'état de l'art et bon nombre de processus mesurés au LHC ne bénéficient pas encore de telles prédictions.

Deux générateurs MC permettent à l'heure actuelle de générer de telles prédictions pour la production $\gamma\gamma + X$ au LHC : le tout nouveau HERWIG 7 [85] et une version particulière de SHERPA en cours de développement dont les résultats n'ont pas encore été publiés, correspondant à la *release r25839*⁹⁵ et appelée SHERPA 2.3 α dans la suite. Les deux implémentations sont décrites dans les paragraphes suivants.

HERWIG 7

Les prédictions de HERWIG 7 concernant la production $\gamma\gamma + X$ n'ont encore jamais été confrontées aux données. Les auteurs ont accepté de les produire pour l'analyse $\gamma\gamma + X$ d'ATLAS à $\sqrt{s} = 8$ TeV et devraient nous les faire parvenir courant 2016. Plus d'informations sur HERWIG 7 sont disponibles dans les références [85].

SHERPA 2.3 α

Les prédictions de SHERPA 2.3 α ont été produites par le Dr. Frank Siegert, qui est à la fois membre de la collaboration ATLAS et l'un des auteurs principaux du générateur SHERPA. Une introduction de cette toute nouvelle implémentation pour le processus $\gamma\gamma + X$ est disponible dans la référence [255]. Celle-ci est basée sur la technique générale de fusion au NLO implémentée dans SHERPA 2.2 : les éléments de matrice correspondant à la composante directe $\gamma\gamma$ et $\gamma\gamma + 1$ parton additionnel sont générés au NLO tandis que ceux correspondant à $\gamma\gamma + 2$ partons et $\gamma\gamma + 3$ partons sont générés au LO à l'aide de Comix [256] et OpenLoops [257]. Ils sont ensuite combinés à la cascade de partons de SHERPA [258] à l'aide de l'algorithme de ME+PS@NLO [119] en un échantillon unique. Les photons issus de la cascade de partons permettent alors une simulation approchée des photons de fragmentation et une resommation approximative de la série perturbative à $p_{T,\gamma\gamma} \simeq 0$ (voir section 3.4).

Les paramètres par défaut du générateur SHERPA sont utilisés pour la modélisation du processus d'hadronisation et de l'événement sous-jacent. Les PDFs sont basées sur le jeu NNPDF3.0 au NNLO [88]. La valeur limite de α_{QED} à basse énergie est employée (limite de Thomson), ce qui est classique pour la production de photons.⁹⁶ En revanche, les échelles de factorisation et de renormalisation sont dynamiques du fait de l'algorithme interne de combinaison de SHERPA [258].

Une séparation fixe entre partons et photons au niveau de la génération (*cutoff*) n'étant plus suffisante pour la convergence de la composante directe au NLO (voir section 3.5.2), la coupure au niveau générateur $\Delta R_{\gamma q(g)} > 0.3$ appliquée pour la production de l'échantillon SHERPA 1.4.0 (*ME+PS merged at LO*, voir section 17.1.2) n'est plus pertinente pour SHERPA 2.3 α . Ainsi, une coupure minimale basée sur un cône de Frixione

⁹⁵<https://sherpa.hepforge.org>

⁹⁶**Note** : la première production montrée dans la section 17.2 a été produite par erreur à $\alpha_{\text{em}}(M_Z)$. Par conséquent, un facteur d'échelle global de 0.88344 est employé.

est appliquée au niveau générateur à la place de celle-ci, correspondant à :

$$\forall r < R_0, \quad E_T^{\text{had}}(r) < E_T^{\text{iso,max}}(r) = \epsilon_\gamma E_T^\gamma \left(\frac{1 - \cos r}{1 - \cos R_0} \right)^n, \quad (90)$$

avec $\epsilon_\gamma = 0.025$, $R_0 = 0.1$ et $n = 2$. Par ailleurs, les deux photons doivent avoir une distance minimale $\Delta R_{\gamma\gamma} > 0.2$ et des impulsions transverse individuelles $E_T^\gamma > 25$ GeV dès la génération. La sélection fiducielle de l'analyse $\gamma\gamma + X$ est ensuite appliquée au niveau particule via une procédure basée sur le logiciel RIVET [259]. Il a été vérifié que la sélection d'isolation au niveau générateur est suffisamment peu contraignante pour ne pas affecter les prédictions finales. Elle est de plus moins contraignante que celle appliquée pour SHERPA 1.4.0. La section efficace fiducielle totale est de 16.17 ± 0.02 (*stat.*) pb.

La génération des prédictions $\gamma\gamma + X$ avec SHERPA 2.3 α est cependant encore peu optimisée et par conséquent requiert un temps CPU très important, les différents événements étant générés avec des poids très faibles ($\simeq 100$ kiloheures CPU par production). L'estimation des incertitudes théoriques est donc pour le moment techniquement difficile. Celles-ci ne seront donc vraisemblablement pas incluses dans la publication finale, sauf si des progrès majeurs d'optimisation sont réalisés d'ici là.⁹⁷

Dans les section 17.1.4 à 17.1.5, les différents calculs partoniques à ordre fixe considérés pour les comparaisons avec les données sont décrits.

17.1.4 DIPHOX (ordre fixe en α_s au NLO)

Les prédictions du programme MC DIPHOX [111] ont été générées à partir de la même version que celle utilisée dans l'analyse $\gamma\gamma + X$ d'ATLAS à $\sqrt{s} = 7$ TeV (v1.3.3).⁹⁸ DIPHOX implémente un calcul NLO à ordre fixe au niveau partonique (donc sans algorithme de cascade de partons et d'hadronisation) pour la section efficace $pp \rightarrow \gamma\gamma + X$, ce qui correspond à $O(\alpha_s)$ pour la composante directe. Il constitue le seul générateur traitant dans leur intégralité les composantes directe, 1-fragmentation et 2-fragmentations au NLO (voir section 3.5). Les composantes de fragmentation sont générées à partir des diagrammes de production de quarks et de gluons dans la limite colinéaire à l'aide des fonctions de fragmentation NLO BFG II [112]. Ceux-ci sont indispensables à la convergence du calcul.

Bien qu'étant déjà un $O(\alpha_s^2)$ car comportant une boucle de quark, le diagramme $gg \rightarrow \gamma\gamma$ au LO pour la composante directe est néanmoins inclus dans le calcul DIPHOX à cause de la luminosité gluon-gluon très importante au LHC, reliée à l'énergie dans le centre de masse élevée de la machine et aux fonctions de structure du proton (voir section 3.3). En effet, ce diagramme représente à lui seul environ 15% de la section efficace totale. En revanche, les premières corrections de ce diagramme, correspondant à un ordre N³LO en α_s , ne sont pas incluses. Celles-ci ont été évaluées dans [228] et ont

⁹⁷**Note :** d'après les dernières conclusions de F. Siegert, les prédictions $\gamma\gamma + X$ *merged at NLO* sont finalement calculables à partir de la version SHERPA 2.2 classique. Des études sont toujours en cours à ce sujet. Si confirmé, cela permettrait d'atteindre une production très optimisée.

⁹⁸Je tiens à remercier le Prof. Juan Terrón Cuadrado et le Dr. Marco Delmastro pour leur aide très appréciée concernant la génération des prédictions de DIPHOX.

un effet modéré sur la section efficace ($\simeq 10\%$, mais plus élevée à bas $m_{\gamma\gamma}$).

Les prédictions nominales sont générées avec le jeu de PDFs CT10 NLO [236], bien que l'on envisage de les produire également avec un des derniers jeux de PDFs disponibles pour le LHC, comportant en particulier une meilleure détermination du flux de gluons à grand x grâce notamment aux dernières mesures reliées au quarks top et à la production de jets au LHC (NNPDF3.0 [88] ou CT14 [87], voir section 3.3). Les échelles de factorisation (μ_f), renormalisation (μ_R) et fragmentation (μ_F) sont prises égales à la masse des di-photons, $m_{\gamma\gamma}$. La section efficace fiducielle totale prédite par DIPHOX est de 10.80 ± 0.04 (*stat.*) pb.

Dans la figure 172, les prédictions de DIPHOX sont représentées pour les différentes observables de mesure, décomposées en termes de composante directe, 1-fragmentation et 2-fragmentations. Il est rappelé que celles-ci n'ont pas grand sens physique prises indépendamment car chacune d'elle dépend significativement de l'échelle de fragmentation arbitraire μ_F . Uniquement leur somme permet une annulation partielle de cette dépendance. Cependant, la décomposition de la figure 172 est intéressante afin d'évaluer les régions de l'espace des phases où les diagrammes de fragmentation contribuent de façon significative à la section efficace DIPHOX.

Comme attendu, les régions correspondant à $m_{\gamma\gamma} < 100$ GeV, $\Delta\phi_{\gamma\gamma} < \pi/2$, $\phi_\eta^* > 1$ et à l'épaule de Guillet ($p_{T,\gamma\gamma}$, a_T) sont très affectées par la fragmentation et particulièrement par la composante 1-fragmentation, la composante 2-fragmentations étant toujours négligeable sauf éventuellement à bas $p_{T,\gamma\gamma}/a_T/\phi_\eta^*$ où elle atteint quelques pourcents des prédictions totales. D'autre part, l'épaule de Guillet est bien issue d'une augmentation à la fois de la fragmentation et de la composante directe (voir section 3.5.4).

Lors du calcul des prédictions de DIPHOX, des sections efficaces négatives sans sens physique sont obtenues dans la première (dernière) catégorie en a_T et ϕ_η^* ($\Delta\phi_{\gamma\gamma}$). Par conséquent, ces catégories ont été exclues de la figure 172 et seront exclues des figures de comparaison. Cet effet est attribué aux instabilités infrarouges des prédictions directes dans les régions a_T , $\phi_\eta^* \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$ associées à la catégorisation très fine adoptée dans la mesure, possible grâce aux très bonnes résolutions des nouvelles variables a_T , ϕ_η^* et des variables angulaires ($\Delta\phi_{\gamma\gamma}$). Les diagrammes de fragmentation inclus dans le calcul ne sont alors apparemment pas suffisants pour régulariser la production totale et une section efficace non physique est obtenue.

D'autre part, il est rappelé qu'au LO $p_{T,\gamma\gamma} = a_T = \phi_\eta^* = 0$ et $\Delta\phi_{\gamma\gamma} = \pi$ exactement au niveau parton. Ainsi, hors de ces régions, les prédictions NLO correspondent finalement au premier terme de production non-nul (soit LO) et par conséquent un bon accord des données avec DIPHOX n'est absolument pas garanti. En revanche, $m_{\gamma\gamma}$ et $\cos \theta_\eta^*$ devraient être moins affectés par ce phénomène, les termes LO et NLO étant non nuls dans l'ensemble de l'espace des phases pour ces observables. Un meilleur accord entre les données et les prédictions de DIPHOX devrait donc être observé pour $m_{\gamma\gamma}$ et $\cos \theta_\eta^*$ comparé aux autres observables de mesure.

Les incertitudes théoriques du calcul DIPHOX ont également été estimées afin de pouvoir mieux quantifier l'accord entre les prédictions pQCD NLO à ordre fixe et les données. Les sources d'incertitudes suivantes ont été considérées :

◊ **Contributions des ordres supérieurs manquants** : L'incertitude théorique

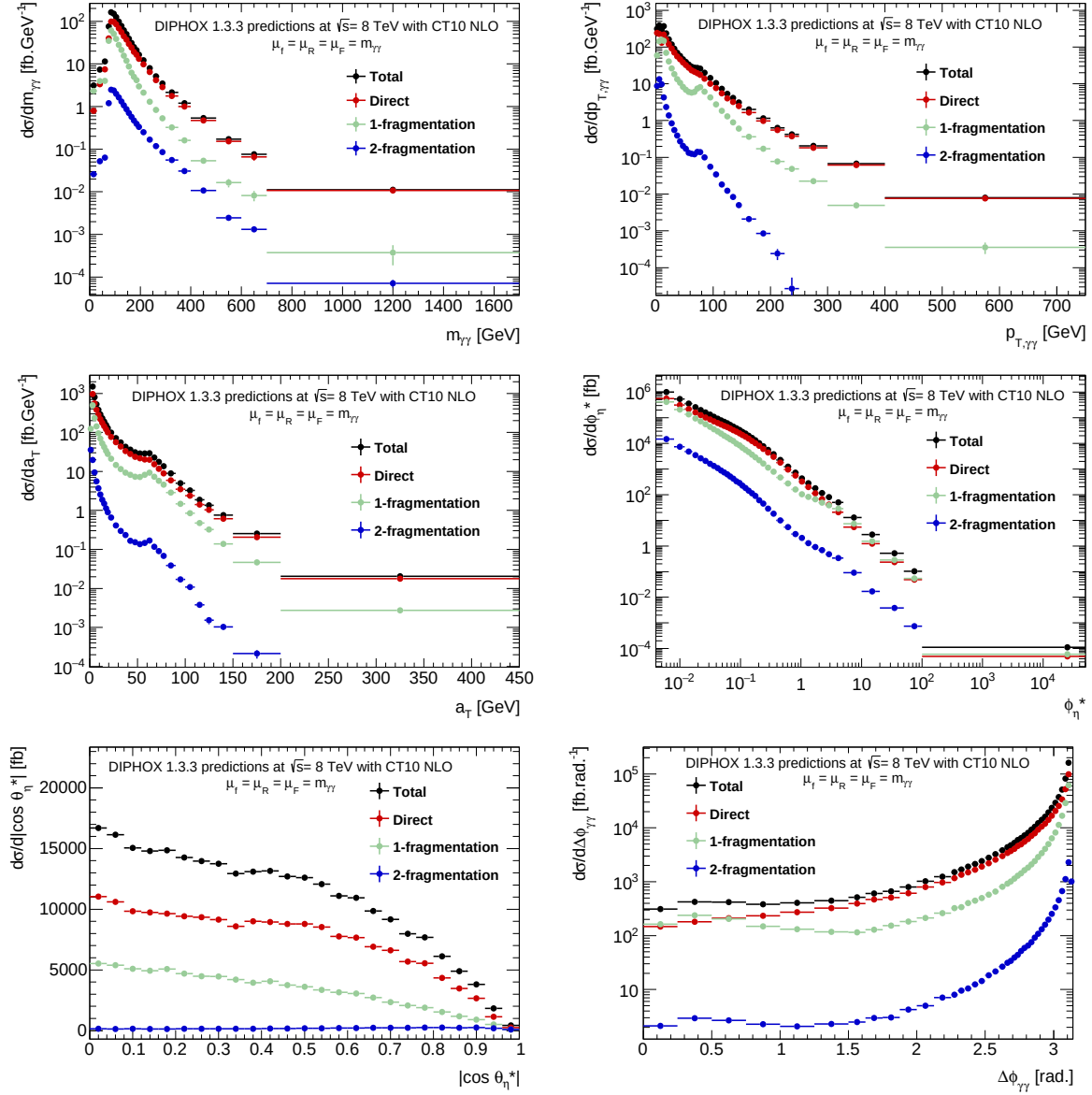


FIG. 172 – Prédications à ordre fixe en α_s issues du générateur DIPHOX à $\sqrt{s} = 8$ TeV basées sur les PDFs CT10 NLO. Les prédictions totales sont décomposées en termes de composantes directe, 1-fragmentation et 2-fragmentations. Les échelles de fragmentation, factorisation et renormalisation sont fixées à $m_{\gamma\gamma}$.

dominante au NLO est issue des corrections d'ordre supérieur en α_s manquants dans le calcul DIPHOX (NNLO, N³LO, etc), souvent appelée incertitude d'échelle. Celle-ci est évaluée de manière approximative en variant arbitrairement les différentes échelles μ_f , μ_R et μ_F de la manière suivante :

- **Variations cohérentes** : les prédictions sont régénérées avec μ_f , μ_R et μ_F fixées toutes les trois à $m_{\gamma\gamma}/2$, puis à $2 m_{\gamma\gamma}$.
- **Variations anticonhérentes** : les prédictions sont régénérées avec les échelles reliées aux divergences infrarouges μ_f et μ_F fixées à $m_{\gamma\gamma}/2$ et l'échelle reliée aux divergences ultraviolettes μ_R à $2 m_{\gamma\gamma}$, et inversement.

Les différences entre les prédictions nominales et les deux types de variations sont ensuite évaluées. L'incertitude finale est alors prise comme l'enveloppe asymétrique maximale des variations. Les différentes variations ainsi que l'enveloppe finale sont représentées pour les différentes observables de mesure dans la figure 173. Comme attendu, l'effet est très important, en particulier pour $p_{T,\gamma\gamma}$, a_T , ϕ_η^* et $\Delta\phi_{\gamma\gamma}$, où les variations s'étalent de 20 à 40%. Pour $\cos \theta_\eta^*$, l'effet est stable à environ 15% tandis que la région à basse masse ($m_{\gamma\gamma} < 100$ GeV) est beaucoup plus affectée par les variations d'échelle que la région à plus haute masse, où des écarts similaires à $\cos \theta_\eta^*$ de l'ordre de 15% sont observés. Dans la grande majorité des cas, la variation anticonhérente est celle menant aux plus grandes variations, sauf à grand $p_{T,\gamma\gamma}$ et a_T .

- ◇ **Structure du proton** : les incertitudes issues de la connaissance limitée des fonctions de structure du proton (PDFs) sont estimées en variant chacun des 20 paramètres de nuisances indépendants implémentés dans le jeu de PDFs CT10 NLO via l'interface LHAPDF [260]. Les incertitudes équivalentes à un intervalle de confiance de 90% ($\simeq 1.6 \sigma$ dans la limite gaussienne) renvoyées par le logiciel sont représentées dans la figure 174. Les limites à 68% ($\simeq 1 \sigma$ dans la limite gaussienne) sont obtenues en appliquant un facteur d'échelle de 1/1.645 (modèle gaussien) et également représentées. L'effet des PDFs sur les prédictions DIPHOX est de l'ordre de 2%, constant dans l'ensemble de l'espace des phases. Il est donc totalement négligeable par rapport aux incertitudes d'échelle.
- ◇ **Constante de couplage de l'interaction forte** : le traitement de α_s dans les différentes productions suit les recommandations disponibles dans la référence [124]. Ainsi, la valeur centrale considérée est $\alpha_s = 0.118$ et l'incertitude correspondante est évaluée en utilisant les ajustements supplémentaires fournis par CT10 NLO avec $\alpha_s = 0.118 \pm 0.02$, *ie.* 0.116 et 0.120. Les résultats sont montrés dans la figure 175. Comme pour les incertitudes reliées aux fonctions de structure du proton, les incertitudes équivalentes à un intervalle de confiance de 90% ($\simeq 1.6 \sigma$ dans la limite gaussienne) et 68% ($\simeq 1 \sigma$ dans la limite gaussienne) sont représentées. L'effet de α_s sur les prédictions DIPHOX est de l'ordre de 1 à 2%, constant dans l'ensemble de l'espace des phases, sauf pour la masse ($m_{\gamma\gamma}$) où il est décroissant. Il reste dans tous les cas totalement négligeable par rapport aux incertitudes d'échelle.
- ◇ **Effets non-perturbatifs sur l'espace fiduciel** : [Note : Les corrections non-

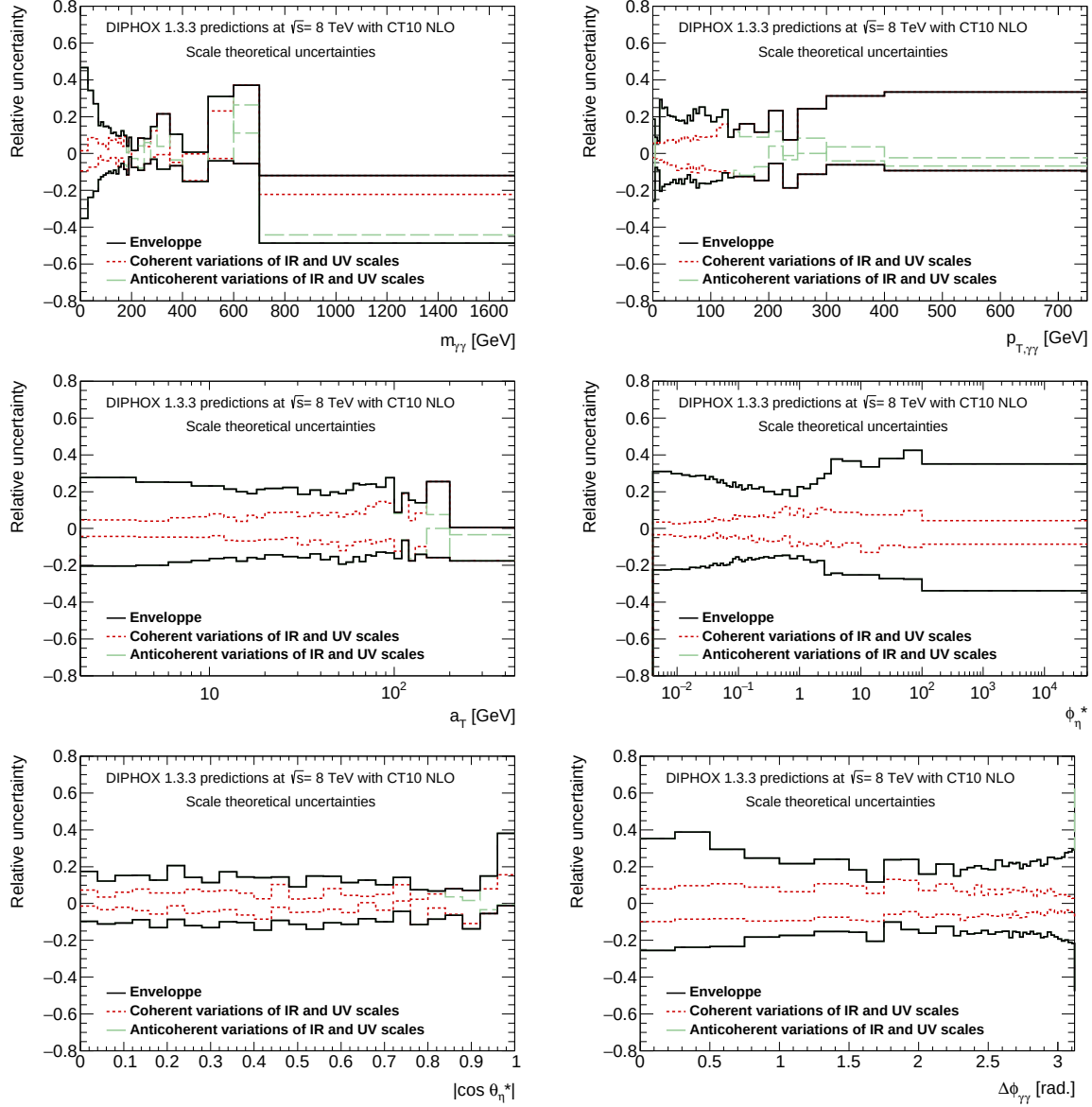


FIG. 173 – Incertitudes théoriques des prédictions NLO à ordre fixe du générateur DIPHOX liées à la connaissance limitée des corrections d'ordre supérieur en α_s . Les incertitudes sont estimées en faisant varier l'échelle de factorisation et de fragmentation d'une part et l'échelle de renormalisation d'autre part à $m_{\gamma\gamma}/2$ et $2m_{\gamma\gamma}$. Ces variations sont effectuées de manière cohérente (trait pointillé rouge) ou anticohérente (trait pointillé vert). L'incertitude finale est prise comme la déviation maximale des résultats avec les prédictions nominales. **Note** : des productions complémentaires sont actuellement en cours afin d'éliminer les fluctuations statistiques restantes (ex : $m_{\gamma\gamma}$).

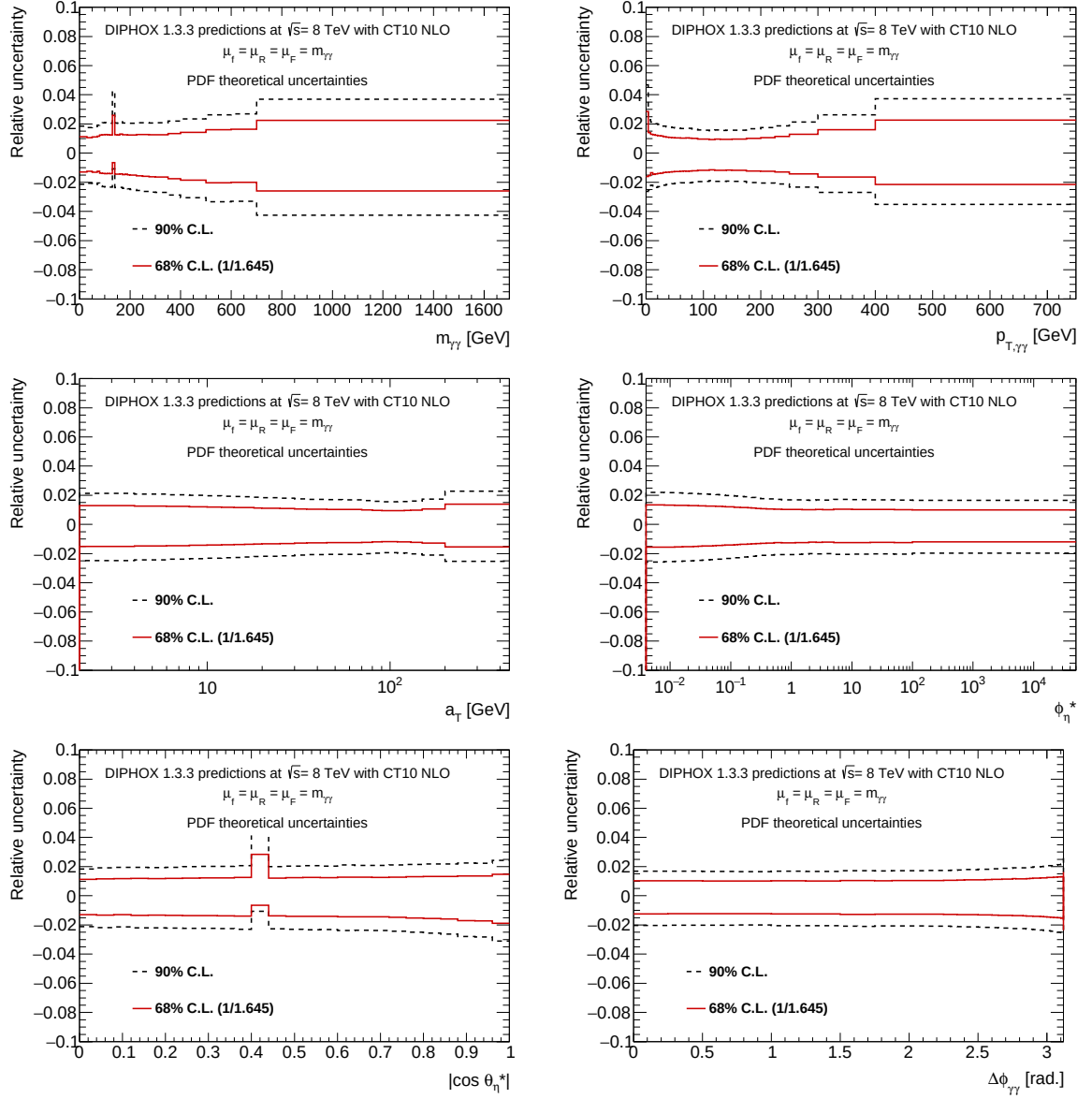


FIG. 174 – Incertitudes théoriques des prédictions NLO à ordre fixe du générateur DIPHON reliées à la connaissance limitée des fonctions de structure du proton (PDFs). Les incertitudes sont estimées à l'aide des paramètres de nuisance implémentés dans le jeu de PDFs CT10 NLO. **Note :** des productions complémentaires sont actuellement en cours afin d'éliminer les fluctuations statistiques restantes.

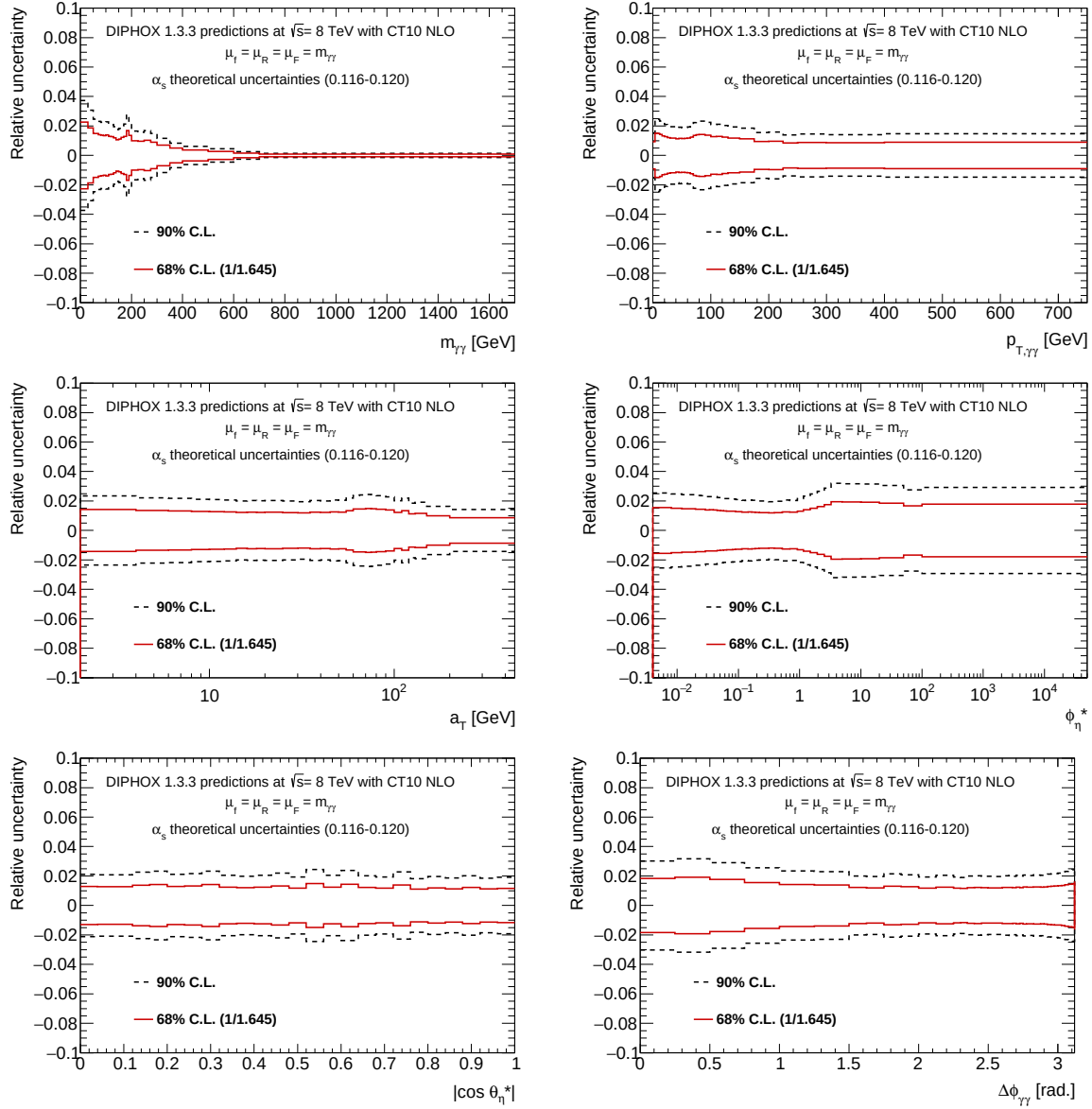


FIG. 175 – Incertitudes théoriques des prédictions NLO à ordre fixe du générateur DIPHOX liées à la connaissance limitée du couplage de l'interaction forte (α_s). Les incertitudes sont estimées à l'aide des ajustements supplémentaires considérant $\alpha_s = 0.118 \pm 0.02$ fournis dans le jeu de PDFs CT10 NLO. **Note :** des productions complémentaires sont actuellement en cours afin d'éliminer les fluctuations statistiques restantes.

perturbatives sont actuellement en cours de finalisation. Elles ne sont pas incluses dans cette première version du manuscrit. L'effet attendu est de l'ordre de quelques pourcents.]

La section efficace fiducielle totale prédite par DIPHOX s'élève à 10.80 ± 0.04 (*stat.*) $^{+1.50}_{-1.18}$ (*syst.*) pb.

Dans la section 17.1.5, les calculs NNLO partoniques à ordre fixe sont décrits. Ceux-ci sont disponibles uniquement pour la composante directe et donc utilisent nécessairement un critère d'isolation basé sur un cône de Frixione pour assurer la convergence des prédictions.

17.1.5 2γ NNLO et MCFM (ordre fixe en α_s au NNLO)

2γ NNLO

Les premières prédictions $\gamma\gamma + X$ à avoir été réalisées au NNLO ont été publiées en 2012 [117]. Celles-ci sont implémentées dans un code partonique appelé 2γ NNLO, qui peut désormais être obtenu sur demande en contactant les auteurs. Les diagrammes de production de $\gamma\gamma + 0$ parton (1 parton, 2 partons) sont inclus jusqu'aux contributions mettant en jeu 2 boucles (1 boucle, 0 boucle), incluant donc toutes les contributions directes à $O(\alpha_s^2)$. Les divergences infrarouges survenant au cours des différents stades intermédiaires du calcul NNLO sont traitées avec le formalisme de soustraction q_T décrit dans la référence [261].

Les composantes de fragmentation n'étant pas considérées dans le calcul, un critère d'isolation de Frixione correspondant à celui décrit dans la section 17.1 est utilisé pour la convergence des prédictions. Celles-ci sont générées à l'aide des PDFs CT10 NNLO [236], bien que l'on envisage de les produire également avec un des derniers jeux de PDFs disponibles pour le LHC, comportant en particulier une meilleure détermination du flux de gluons à grand x (NNPDF3.0 [88] ou CT14 [87]). α_s est fixée à 0.118 tandis que les échelles de factorisation et de renormalisation sont prises égales à $m_{\gamma\gamma}$.⁹⁹

Seuls les résultats d'une production préliminaire présentant une statistique limitée seront présentés dans ce manuscrit, les prédictions finales étant toujours en cours de production. La section efficace fiducielle totale est de 14.42 pb. Les incertitudes statistiques et théoriques sont pour le moment indisponibles, de même que la distribution correspondant à l'observable a_T . Les prédictions et incertitudes complètes seront incluses dans la publication finale.

MCFM

Très récemment, une implémentation au NNLO du processus $\gamma\gamma + X$ a été réalisée dans le cadre du programme MCFM [118]. Celle-ci incluant uniquement la production directe, un critère d'isolation de Frixione est également nécessaire pour assurer la convergence du calcul. Le code n'ayant pas encore été rendu public, les auteurs ont été

⁹⁹Remerciements particuliers au Dr. Leandro Cieri ainsi qu'au Prof. Juan Terrón Cuadrado pour leur aide concernant l'exécution du code 2γ NNLO et la génération des prédictions.

contactés et ont accepté de nous fournir leurs prédictions $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV. Celles-ci devraient nous parvenir courant 2016.

Dans la section 17.1.6, les calculs NLO et NNLO partoniques à ordre fixe incluant une resommation à tous les ordres en α_s au NNLL (voir section 3.4) dans la région $p_{T,\gamma\gamma} \simeq 0$ sont décrits.

17.1.6 ResBos et 2 γ RES (ordre fixe en α_s au NLO/NNLO avec resommation au NNLL)

ResBos

RESBOS est un générateur MC partonique au NLO créé tout d'abord pour fournir des prédictions fiables pour le processus Drell-Yan $q\bar{q} \rightarrow l^+l^-$ dans la limite où les deux leptons sont dos-à-dos dans le plan transverse [98]. Pour ce faire, il met en œuvre un procédé de resommation QCD à tous les ordres en α_s au NNLL (*next-to-next-to-leading logarithm*) pour $p_{T,\gamma\gamma} \simeq 0$ et combine la section efficace resommée avec les ordres supérieurs NLO pour $p_{T,\gamma\gamma} > 0$.

L'implémentation de RESBOS pour la section efficace $pp \rightarrow \gamma\gamma + X$ [101, 262–265] inclut les diagrammes de production directe NLO à la fois pour les modes de production $q\bar{q}$, qg (*born + fragmentation*) et gg , gq (*box + fragmentation*), contrairement à DIPHOX qui inclut uniquement le diagramme à l'ordre le plus bas pour $gg \rightarrow \gamma\gamma$ (voir section 17.1.4). Cependant, DIPHOX calcule explicitement les deux composantes de fragmentation (1-fragmentation, 2-fragmentations) permettant la stabilisation du mode de production direct qg/gq tandis que RESBOS simule approximativement la composante 1-fragmentation uniquement, à l'aide d'une paramétrisation permettant de reproduire les taux de fragmentation attendus au NLO. En revanche, DIPHOX ne procède à aucune resommation.

Les prédictions RESBOS nous ont été fournies par les auteurs.¹⁰⁰ Elles ont été générées avec les PDFs CT10 NLO et les échelles de factorisation et de renormalisation ont été fixées à la masse invariante des di-photons, $m_{\gamma\gamma}$. Pour des raisons techniques, le calcul est effectué pour $m_{\gamma\gamma} > 20$ GeV uniquement. L'effet de cette coupure additionnelle sur les prédictions est négligé, la section efficace mesurée dans pour $m_{\gamma\gamma} \in [0, 20 \text{ GeV}]$ étant inférieure à 0.8% de $\sigma_{\text{tot}}^{\text{fid.}}$. La sélection fiducielle de l'analyse est ensuite appliquée au niveau partonique. La section efficace fiducielle totale prédite par RESBOS est de 12.26 ± 0.05 (*stat.*) pb, significativement plus élevée que celles de DIPHOX (10.80 pb) probablement grâce aux corrections NLO de $gg \rightarrow \gamma\gamma$. Les auteurs ne prévoient pas d'estimer les incertitudes théoriques reliées aux prédictions RESBOS et nous ont conseillé d'utiliser celles de DIPHOX comme ordre de grandeur.

Les composantes correspondant à $q\bar{q}$, $qg \rightarrow \gamma\gamma$ et gg , $gq \rightarrow \gamma\gamma$ sont représentées dans la figure 176. Contrairement à DIPHOX, des prédictions valides sont produites pour l'ensemble des catégories utilisées pour la mesure, incluant la première (dernière) catégorie en a_T , ϕ_η^* ($\Delta\phi_{\gamma\gamma}$), probablement grâce au procédé de resommation. La contribution $gg \rightarrow \gamma\gamma$ est généralement de 5 à 10 fois plus faible que $q\bar{q}$, $qg \rightarrow \gamma\gamma$, ce qui correspond bien à l'ordre de grandeur de la contribution LO $gg \rightarrow \gamma\gamma$ observée pour DIPHOX

¹⁰⁰Remerciements particuliers au Prof. Pavel Nadolsky.

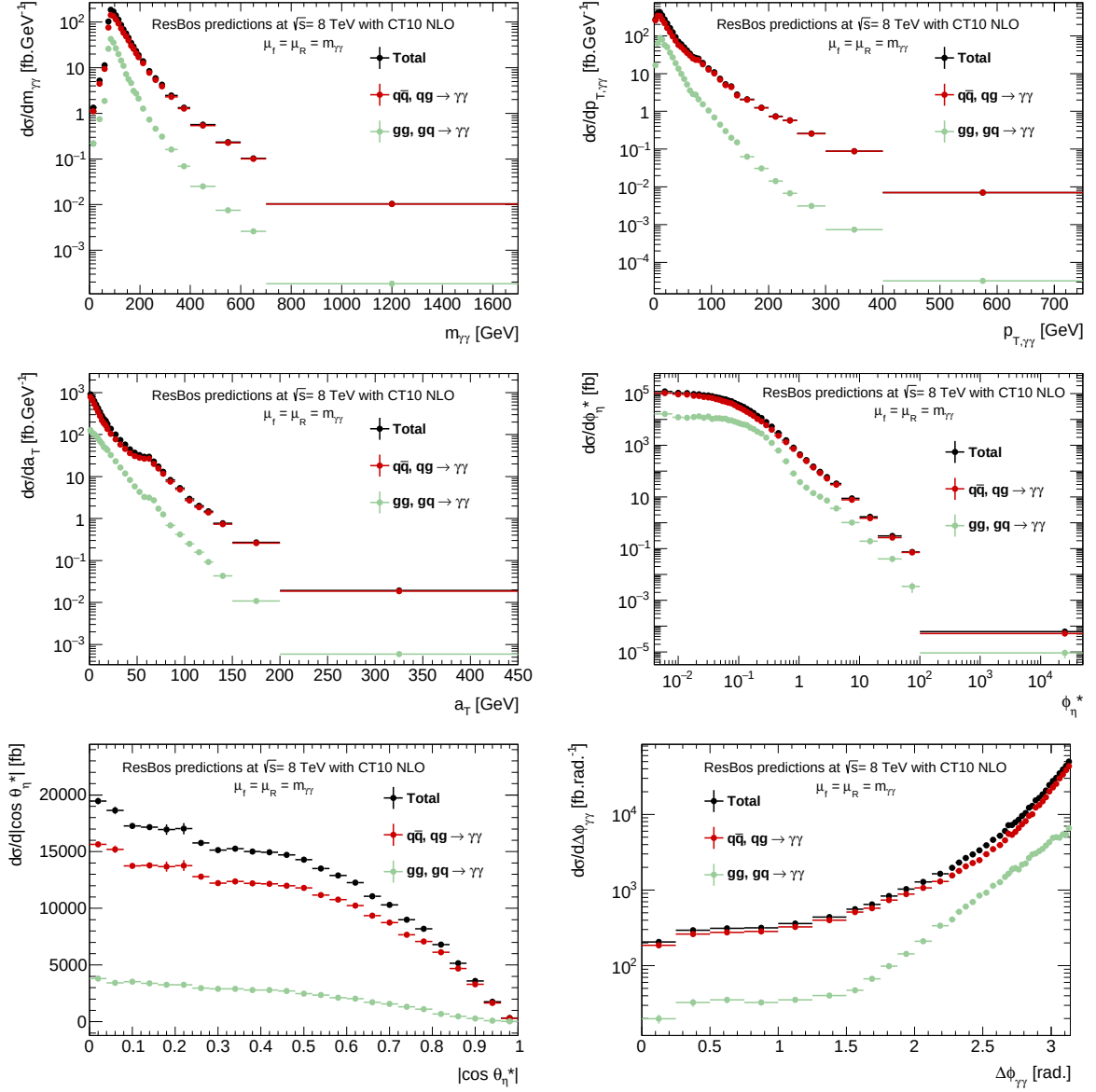


FIG. 176 – Prédications à $\sqrt{s} = 8$ TeV du calcul RESBos (NLO+NNLL) produites à l'aide des PDFs CT10 NLO. Les échelles de factorisation et renormalisation sont prises égales à $m_{\gamma\gamma}$. Les prédictions totales (points noirs) sont décomposées selon les modes de production $q\bar{q}, qg$ (points rouges, *born + fragmentation*) et gg, gq (points verts, *box + fragmentation*).

($\simeq 15\%$ de la section efficace fiducielle totale). Par ailleurs, la production semble bien mieux se stabiliser par rapport aux calcul DIPHOX à petit $p_{T,\gamma\gamma}$, a_T et haut $\Delta\phi_{\gamma\gamma}$ mais surtout à petit ϕ_η^* , attendu du procédé de resommation au NNLL pour $p_{T,\gamma\gamma} \simeq 0$ (voir section 3.4).

2 γ RES

Les résultats d'un nouveau calcul NNLL concurrent de RESBOS nommé 2 γ RES [102] ont été publiés en 2015 par les mêmes auteurs que 2 γ NNLO [117] (voir section 17.1.5). Comme RESBOS, celui-ci procède à une resommation à tous les ordres au NNLL pour $p_{T,\gamma\gamma} \simeq 0$, basée toutefois sur un formalisme différent décrit dans les références [266–269]. Cependant, pour $p_{T,\gamma\gamma} > 0$, la section efficace resommée est combinée avec les termes perturbatifs correspondant à la composante directe à $O(\alpha_s^2)$, *ie.* NNLO en α_s pour $p_{T,\gamma\gamma} = 0$ (espace des phases à deux corps) et NLO pour $p_{T,\gamma\gamma} \neq 0$ (espace des phases à trois corps). 2 γ RES inclut donc en général un ordre en α_s de plus comparé à RESBOS.

Par ailleurs, le calcul est mené de manière à ce que la section efficace fiducielle totale de 2 γ RES soit égale à celle équivalente prédite sans resommation (*ie.* 2 γ NNLO). Ainsi, la resommation affecte uniquement la forme des distributions différentielles, en particulier celles sensibles aux divergences infrarouges ($p_{T,\gamma\gamma}$, a_T , ϕ_η^* , $\Delta\phi_{\gamma\gamma}$).

2 γ RES incluant uniquement la composante directe (NNLO), un critère d'isolation de Frixione est appliqué afin d'assurer la convergence du calcul. D'après la publication associée [102], la resommation au NNLL permet de décrire les régions de l'espace des phases où le calcul à ordre fixe NNLO échouait à $\sqrt{s} = 7$ TeV (comparaison faite avec 2 γ NNLO). Il est donc particulièrement prometteur pour la mesure $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV.

Le code n'ayant pas encore été rendu public, les auteurs ont été contactés et ont accepté de nous fournir les prédictions de 2 γ RES à 8 TeV courant 2016. La question de l'évaluation des incertitudes théoriques n'a pas encore été abordée avec les auteurs.

Dans la section 17.2, les mesures issues de l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV sont comparées avec les différentes prédictions théoriques introduites tout au long de la section 17.1.

17.2 Comparaisons aux mesures expérimentales

17.2.1 SHERPA 1.4.0 (*ME+PS merged at LO*) et PYTHIA 8 (*LO+PS*)

Il est rappelé que la section efficace fiducielle totale mesurée est :

$$\sigma_{\text{tot}}^{\text{fid.}} = 16.51 \pm 0.07(\text{stat}) \pm 0.75(\text{syst}) \pm 0.32(\text{lumi}) \pm 0.03(\text{trigger}) \text{ pb} \quad (91)$$

En comparaison, SHERPA 1.4.0 et PYTHIA 8 prédisent pour le même volume fiduciel respectivement 12.53 pb et 11.81 pb (incertitude statistique négligeable), ce qui correspond à environ 76% et 72% de la section efficace totale mesurée seulement. La section efficace de SHERPA 1.4.0 est plus élevée grâce aux diagrammes $\gamma\gamma + 1,2$ parton(s), absents dans PYTHIA8.

Étant donné que ces échantillons MC sont utilisés pour déconvoluer la mesure et que l'échelle de leurs prédictions est fortement sous-évaluée à cause entre autres de l'absence des corrections virtuelles en α_s (voir section 17.2.1), les sections efficaces de SHERPA 1.4.0 et PYTHIA 8 sont comparées aux mesures différentielles après repondération à $\sigma_{\text{tot}}^{\text{fid.}}$. Ainsi, uniquement les formes des distributions sont comparées. Les comparaisons au niveau reconstruit et après déconvolution sont représentées pour les différentes observables de mesure dans les figures 177 à 182. Le facteur d'échelle varie légèrement entre les différentes variables, attendu de l'effet systématique sur la soustraction du fond discuté dans la section 15.6.4.

Globalement, SHERPA 1.4.0 et PYTHIA 8 décrivent de façon correcte la forme des différentes distributions. Cependant, PYTHIA 8 a tendance à sous-estimer de façon très importante les données dans les régions de l'espace des phases dominées par la fragmentation et/ou les ordres supérieurs en α_s , soit $m_{\gamma\gamma} < 80$ GeV, $p_{T,\gamma\gamma}, a_T > 50$ GeV, $\phi_\eta^* > 0.5$, $|\cos \theta_\eta^*| > 0.8$ et $\Delta\phi_{\gamma\gamma} < 2$ radians, mais également l'épaule de Guillet qui n'est absolument pas reproduite par les prédictions car issue des diagrammes $\gamma\gamma +$ partons (voir section 3.5.4). Pour les régions sensibles aux divergences infrarouges, *ie.* $p_{T,\gamma\gamma}, a_T, \phi_\eta^* \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$, le facteur d'échelle biaise légèrement la comparaison et les prédictions absolues de PYTHIA 8 sont seulement légèrement sous-estimées (et non sur-estimées).

En général, SHERPA 1.4.0 reproduit mieux les formes des distributions dans les régions de l'espace des phases dominées par la fragmentation et/ou les ordres supérieurs que PYTHIA 8 grâce aux diagrammes $\gamma\gamma + 1,2$ parton(s), qui correspondent à des corrections réelles en α_s . En particulier, l'épaule de Guillet est plutôt correctement reproduite par SHERPA 1.4.0. Cependant, les régions $p_{T,\gamma\gamma}, a_T, \phi_\eta^* \gg 0$ et $\Delta\phi_{\gamma\gamma} \simeq 0$ sont sur-estimées (ou décrites à peu près correctement si on n'applique pas de facteur d'échelle), tandis que les régions $m_{\gamma\gamma} \gg 0$ et $|\cos \theta_\eta^*| > 0.8$ sont sous-estimées quoiqu'il advienne. Ce problème est connu des auteurs de SHERPA qui pour le régler n'ont pas eu d'autre choix que de faire évoluer l'implémentation à “*ME+PS merged at NLO*” (SHERPA 2.3 α), comme nous le verrons dans la section 17.2.2.

Les variations de SHERPA 1.4.0 s'étalant en général dans une gamme plus restreinte par rapport à PYTHIA 8, l'usage de l'échantillon de signal simulé SHERPA 1.4.0 pour l'analyse nominale (soustraction du fond, déconvolution, voir sections 15 et 16) est définitivement validé. Par ailleurs, il est intéressant de remarquer qu'une fois les facteurs d'échelle appliqués, la mesure est presque systématiquement encadrée par les prédictions SHERPA 1.4.0 et PYTHIA 8 (à l'exception de $|\cos \theta_\eta^*| > 0.8$), ce qui motive l'utilisation de PYTHIA 8 pour l'évaluation des incertitudes systématiques dans l'analyse. Finalement, les comparaisons au niveau reconstruit ou après déconvolution diffèrent peu, attendu des matrices de réponses quasi-diagonales.

Les figures 177 à 182 ont également été confrontées aux comparaisons équivalentes effectuées lors de la mesure à $\sqrt{s} = 7$ TeV d'ATLAS [203] afin de vérifier la cohérence des deux résultats. La description de PYTHIA 8 est stable malgré le changement de PDFs (MRST2007 [270] $\rightarrow$ CTEQ6L1 [237], tout deux LO), cependant SHERPA 1.4.0 utilise désormais CT10 NLO [236] au lieu de CTEQ6L1 (LO), ce qui a nécessité un nouveau réglage des paramètres de la cascade de partons. Par conséquent, sa description des données a légèrement évolué. Une légère amélioration est observée pour $m_{\gamma\gamma} \gg 0$, au détriment de $p_{T,\gamma\gamma} \gg 0$. D'autre part, la région $\Delta\phi_{\gamma\gamma} \simeq 0$ est désormais surestimée

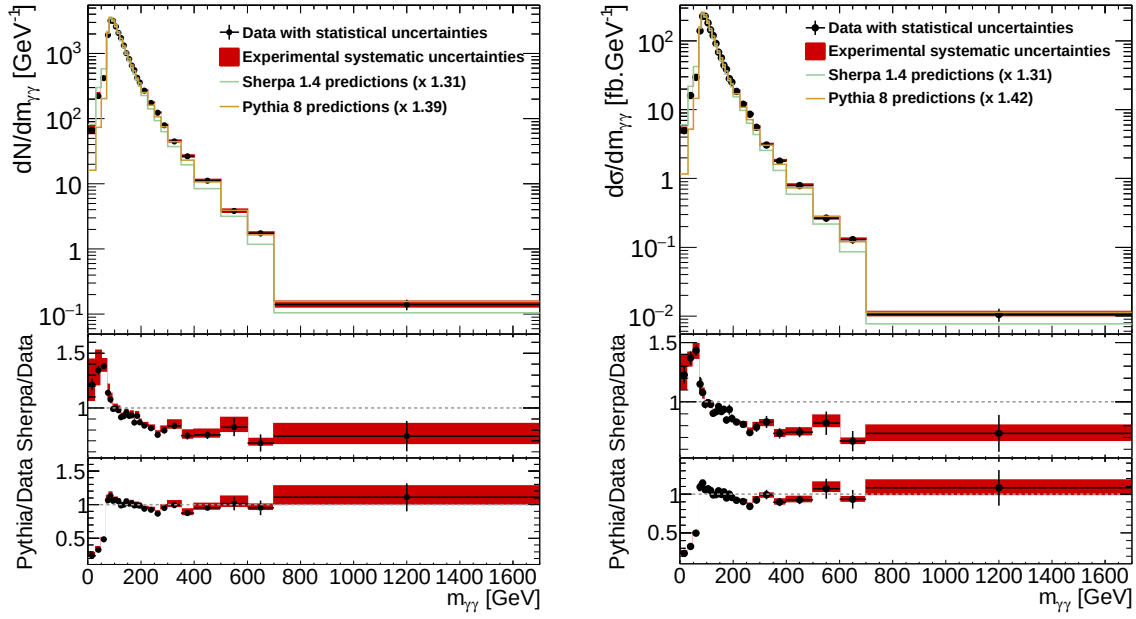


FIG. 177 – Comparaisons entre la section efficace différentielle $pp \rightarrow \gamma\gamma + X$ mesurée à $\sqrt{s} = 8$ TeV et les prédictions correspondantes des générateurs SHERPA 1.4.0 et PYTHIA 8 avant (gauche) et après (droite) déconvolution des effets de détecteur en fonction de $m_{\gamma\gamma}$. Les distributions MC sont repondérées à la section efficace totale mesurée par un facteur global mentionné dans la légende. Les incertitudes théoriques n'ont pas été estimées pour ces prédictions.

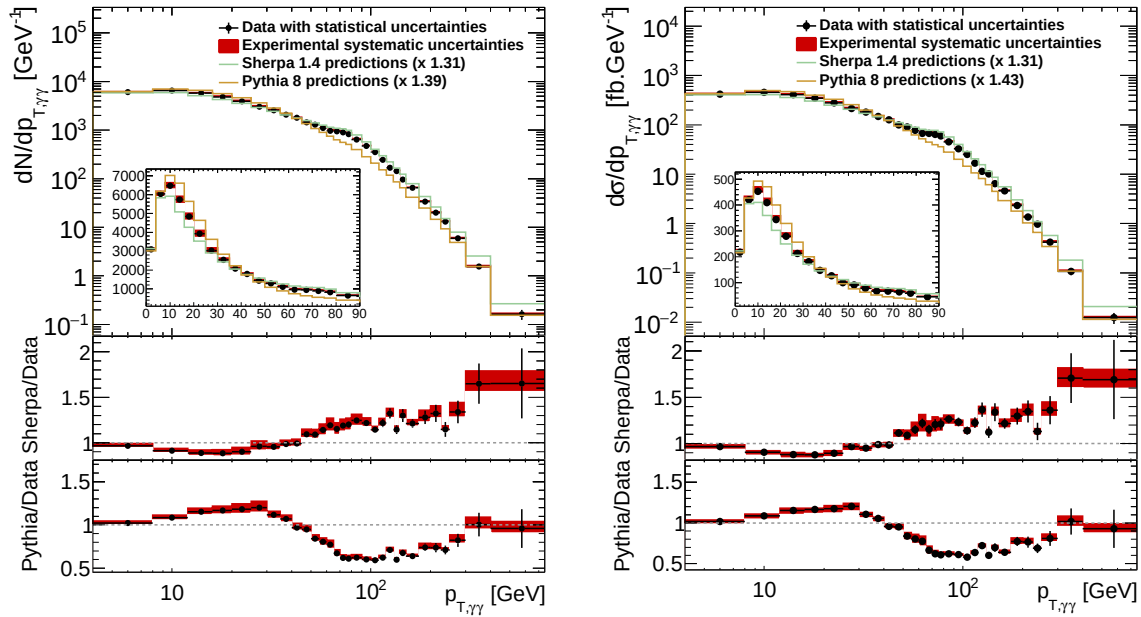
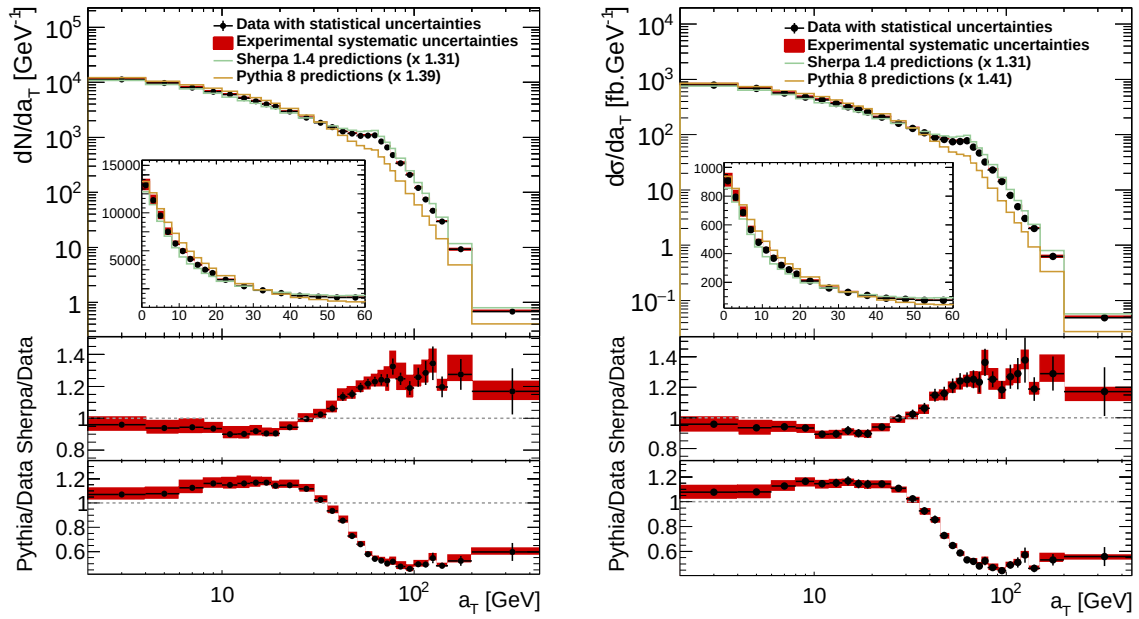
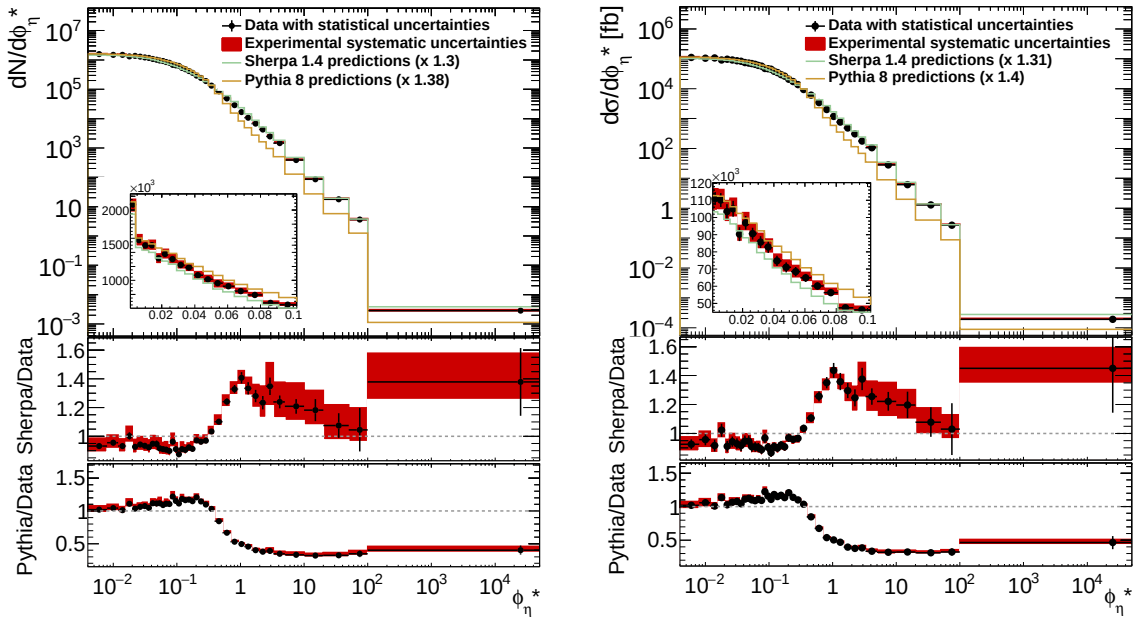
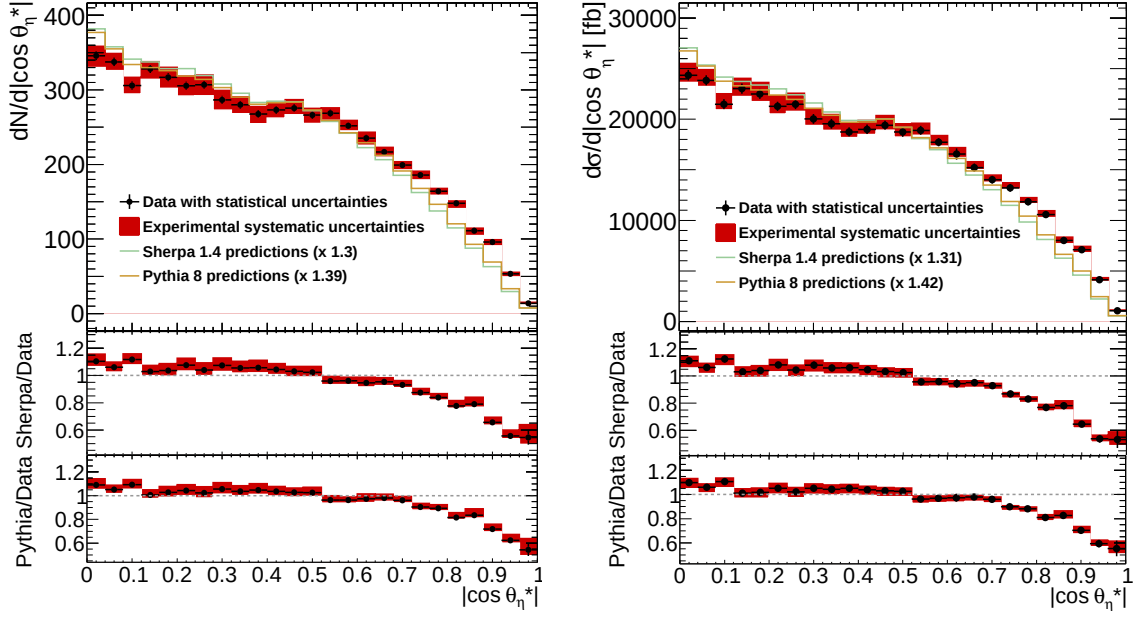
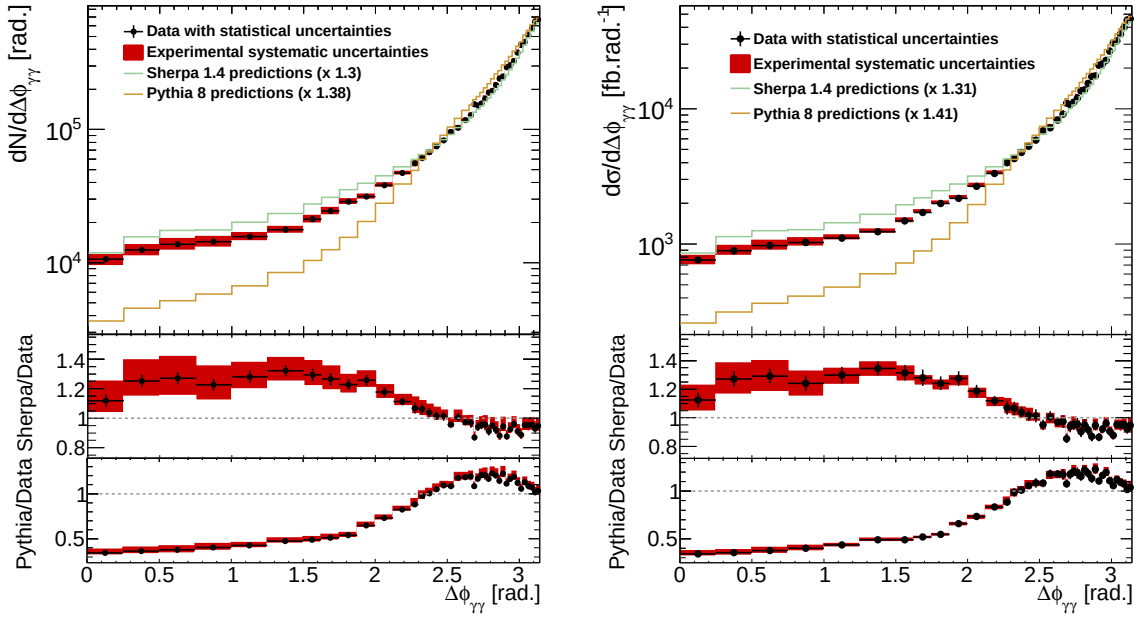


FIG. 178 – Figure similaire à la figure 177 pour $p_{T,\gamma\gamma}$.

FIG. 179 – Figure similaire à la figure 177 pour a_T .FIG. 180 – Figure similaire à la figure 177 pour ϕ_η^* .

FIG. 181 – Figure similaire à la figure 177 pour $|\cos \theta_\eta^*|$.FIG. 182 – Figure similaire à la figure 177 pour $\Delta\phi_{\gamma\gamma}$.

après application du facteur d'échelle tandis qu'elle avait tendance à être sous-estimée à $\sqrt{s} = 7$ TeV. Ce changement est probablement relié à l'évolution de $p_{T,\gamma\gamma}$ du fait des corrélations entre les régions $p_{T,\gamma\gamma} \gg 0$ et $\Delta\phi_{\gamma\gamma} \simeq 0$. a_T et ϕ_η^* n'ont pas été mesurés à $\sqrt{s} = 7$ TeV. Finalement, les facteurs d'échelle ont légèrement augmenté, passant de 1.2 ($\sqrt{s} = 7$ TeV) à 1.3 et 1.4 ($\sqrt{s} = 8$ TeV) pour SHERPA 1.4.0 et PYTHIA 8, respectivement. Cela est probablement simplement dû à l'augmentation des coupures fiducielles en $E_{T,1}$, $E_{T,2}$ (voir section 12.2), qui sont passées de 25 à 40 GeV et de 22 à 30 GeV entre l'analyse à $\sqrt{s} = 7$ et 8 TeV pour le premier et le deuxième photon, respectivement.

Ainsi, il semble très délicat d'arriver à une description correcte de la production $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV uniquement à l'aide d'éléments de matrice LO, que ce soit au niveau de l'échelle ($\sigma_{\text{tot}}^{\text{fid.}}$) ou de la forme des distributions différentielles. L'état de l'art pour les calculs basés sur des cascades de partons est désormais *ME+PS merged at NLO*, *ie.* incluant les corrections réelles et virtuelles au NLO. Cela devrait permettre d'améliorer significativement la description des données. Les résultats de comparaison au niveau particule avec les deux générateurs permettant de produire de telles prédictions pour $pp \rightarrow \gamma\gamma + X$ à l'heure actuelle (HERWIG 7 et SHERPA 2.3 α) sont présentés dans la section 17.2.2.

17.2.2 HERWIG 7 et SHERPA 2.3 α (*ME+PS merged at NLO*)

Note : les prédictions de HERWIG 7 ne sont pas encore disponibles et devraient nous être transmises courant 2016.

La section efficace fiducielle totale prédite par la version en développement de SHERPA appelée SHERPA 2.3 α (voir section 17.1.3) s'élève à 16.17 ± 0.02 (*stat.*) pb, en bien meilleur accord avec la mesure expérimentale que SHERPA 1.4.0 et désormais compatible dans l'incertitude expérimentale ($16.51^{+0.82}_{-0.81}$ pb). SHERPA 2.3 α reproduit donc avec succès l'échelle de la mesure, probablement grâce aux corrections virtuelles et réelles en α_s supplémentaires par rapport à la version précédente. La section efficace se décompose en fonction des quatre processus durs implémentés dans le générateur de la façon suivante :

- ◇ $pp \rightarrow \gamma\gamma$ (NLO) : 4.27 pb
- ◇ $pp \rightarrow \gamma\gamma + 1 \text{ parton}$ (NLO) : 5.84 pb
- ◇ $pp \rightarrow \gamma\gamma + 2 \text{ partons}$ (LO) : 2.87 pb
- ◇ $pp \rightarrow \gamma\gamma + 3 \text{ partons}$ (LO) : 3.18 pb

Du fait de l'algorithme de fusion employé pour combiner les éléments de matrice et la cascade de partons, le processus dur généré ne correspond pas forcément à la production expérimentale $\gamma\gamma + 0, 1, 2$ et 3 jets, aucun algorithme de reconstruction de jet n'étant considéré ici et la cascade de partons pouvant également conduire à la reconstruction de jets additionnels. La décomposition présentée ci-dessus est donc montrée uniquement pour illustrer le fonctionnement de l'algorithme d'appariement au NLO de SHERPA.

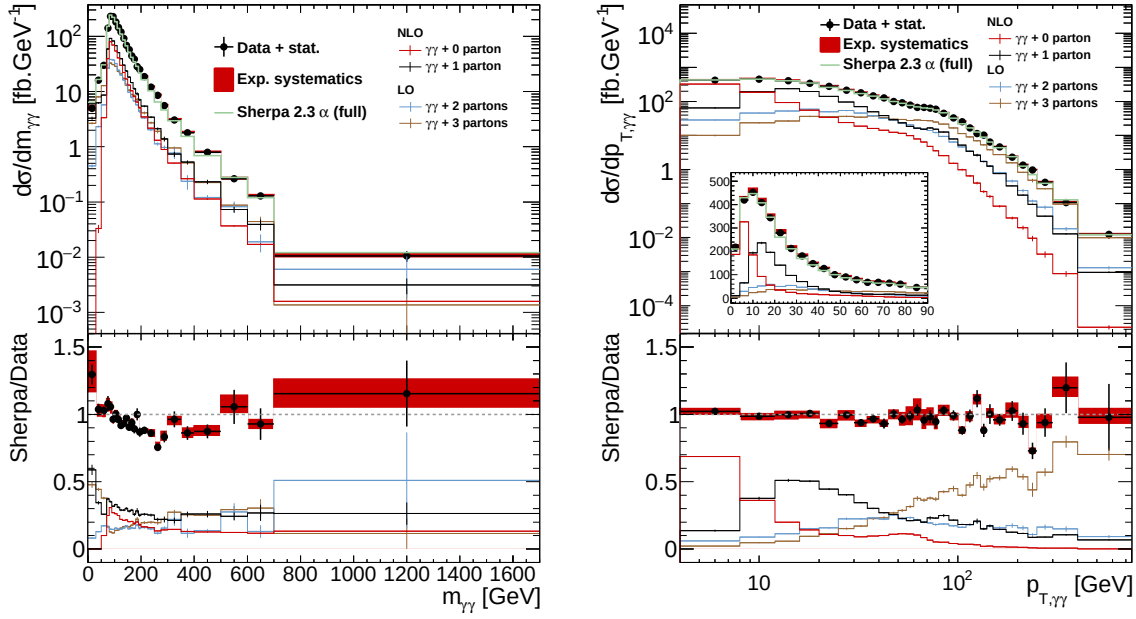
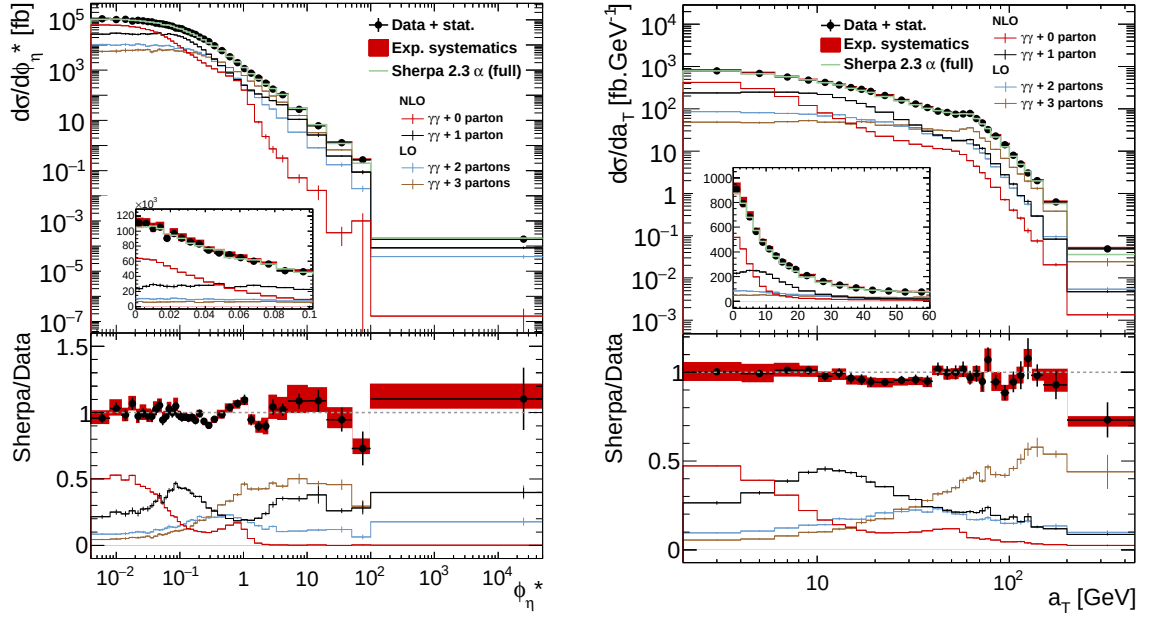
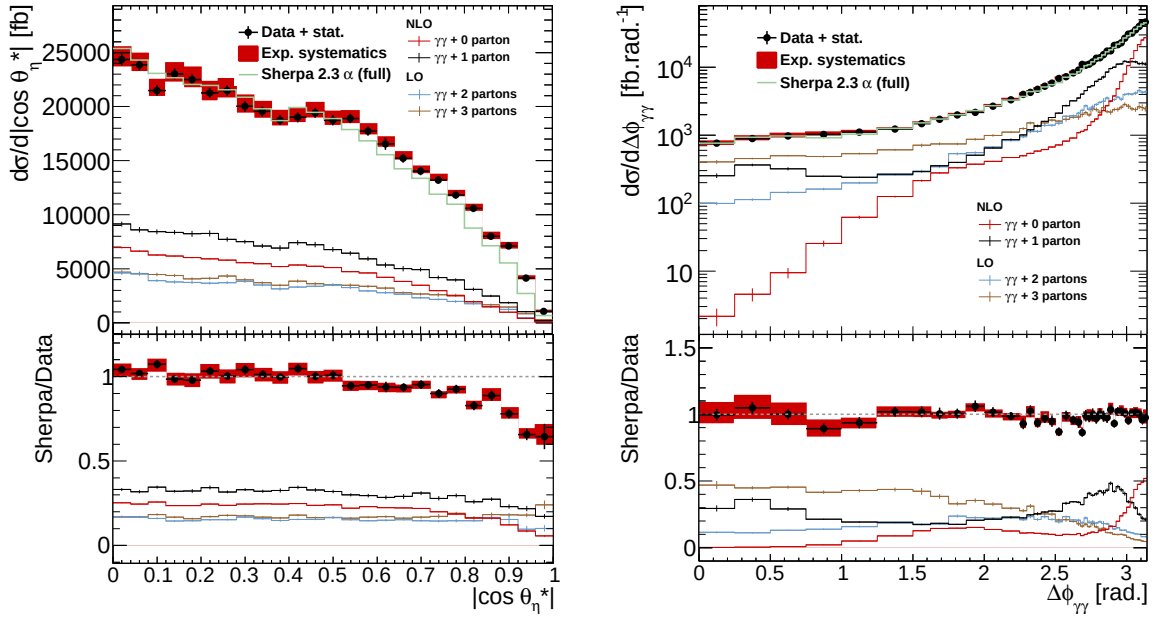


FIG. 183 – Comparaisons entre la section efficace différentielle $pp \rightarrow \gamma\gamma + X$ mesurée à $\sqrt{s} = 8$ TeV en fonction de $m_{\gamma\gamma}$ (gauche) et $p_{T,\gamma\gamma}$ (droite) et les prédictions correspondantes du générateur SHERPA 2.3 α , après déconvolution des effets de détecteur. Les prédictions MC sont décomposées en fonction du processus dur généré : $\gamma\gamma + 0, 1$ (NLO) ou 2, 3 partons (LO). Les incertitudes théoriques n'ont pas été estimées pour ces prédictions.

Les résultats de comparaison (sans facteur d'échelle) pour les différentes observables de mesure sont présentées dans les figures 183 à 185. À nouveau, les prédictions finales sont décomposées en fonction du processus dur généré : $\gamma\gamma + 0, 1$ (NLO) et 2, 3 partons (LO). On constate tout de suite la très nette amélioration apportée par SHERPA 2.3 α par rapport à SHERPA 1.4.0. En particulier, les variables $p_{T,\gamma\gamma}$, a_T , ϕ_η^* et $\Delta\phi_{\gamma\gamma}$ sont très bien reproduites et déjà compatibles dans la plupart des cas avec la mesure dans l'incertitude expérimentale sans même considérer les incertitudes théoriques, que ce soit dans les régions sensibles à l'émission de gluons mous dans l'état initial ($p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$) ou dans celles dominées par les ordres supérieurs en α_s ($p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* \gg 0$ et $\Delta\phi_{\gamma\gamma} \ll \pi$). Il n'est pas exclu que les incertitudes théoriques puissent être estimées pour la publication finale.

Il est également intéressant d'observer les dépendances des contributions des différents processus durs ($\gamma\gamma + 0, 1, 2, 3$ partons), même si le passage parton $\rightarrow$ jet n'est pas trivial à cause de la cascade de partons. Un examen détaillé des différentes distributions montre quelques rares endroits où la composante $pp \rightarrow \gamma\gamma + 3$ partons a un comportement légèrement suspect en dehors des fluctuations statistiques, comme $\phi_\eta^* \simeq 1$ (correspondant à $a_T = m_{\gamma\gamma}$ dans la limite $\Delta\phi_{\gamma\gamma} = \pi$) et $100 \text{ GeV} < a_T < 200 \text{ GeV}$. Des études complémentaires par les auteurs de SHERPA seront nécessaires afin d'en déterminer l'origine exacte, par exemple lors de l'estimation des incertitudes théoriques.

$m_{\gamma\gamma}$ et $|\cos \theta_\eta^*|$ présentent un accord légèrement moins bon quoique toujours très satisfaisant par rapport aux autres variables. Un léger déficit est observé dans la région $150 \text{ GeV} < m_{\gamma\gamma} < 400 \text{ GeV}$, la statistique ne permettant pas de conclure à plus haute valeur. Un déficit est également observé pour $|\cos \theta_\eta^*| > 0.7$. Cette région est dans tous

FIG. 184 – Figure similaire à la figure 183 pour ϕ_η^* et a_T .FIG. 185 – Figure similaire à la figure 183 pour $|\cos \theta_\eta^*|$ et $\Delta\phi_{\gamma\gamma}$.

les cas particulièrement difficile à décrire car correspond à des di-photons émis à grand $\Delta\eta_{\gamma\gamma}$ et donc habituellement accompagnée de radiations QCD importantes.

L'optimisation de la production SHERPA 2.3 α (actuellement en cours) permettrait d'améliorer la statistique des prédictions et ainsi très probablement de préciser les observations rapportées dans cette section. On s'attend donc à une meilleure compréhension de l'accord entre SHERPA 2.3 α et les données dans les prochains mois. Par ailleurs, des prédictions similaires sont en cours de production pour le processus $\gamma\gamma + \text{jet(s)}$, qui fait l'objet d'une mesure dédiée avec les données à $\sqrt{s} = 8$ TeV du détecteur ATLAS.

Dans la section 17.2.3, les mesures expérimentales sont comparées avec les calculs NLO à ordre fixe, soit DIPHOX (NLO) et RESBos (NLO+NNLL).

17.2.3 DIPHOX et ResBos (ordre fixe NLO en α_s)

DIPHOX et RESBos ont été décrits respectivement dans les sections 17.1.4 et 17.1.6. Ils incluent tous les deux les éléments de matrice $pp \rightarrow \gamma\gamma + X$ au NLO, bien que $gg \rightarrow \gamma\gamma$ soit inclus à LO dans DIPHOX et NLO dans RESBos, DIPHOX étant basé sur la logique d'un ordre fixe en $O(\alpha_s)$. DIPHOX inclut un traitement plus complet de la fragmentation tandis que RESBos procède à une resommation à tous les ordres pour $p_{T,\gamma\gamma} \simeq 0$. Ils génèrent tous les deux des prédictions au niveau partonique.

Les sections efficaces fiducielles totales prédites par DIPHOX et RESBos sont respectivement 10.80 ± 0.04 (*stat.*) $^{+1.50}_{-1.18}$ (*syst.*) pb et 12.26 ± 0.05 (*stat.*) pb. Aucune incertitude théorique n'a été fournie pour le moment pour RESBos par les auteurs, cependant elle devrait être du même ordre de grandeur que celle de DIPHOX. Les sections efficaces NLO représentent donc environ 65 et 74% de la valeur mesurée seulement, ce qui représente un écart de près de 4σ en ce qui concerne DIPHOX, montrant que les variations d'échelle arbitraires effectuées dans la section 17.2.3 pour estimer les ordres supérieurs manquant sous-estiment probablement l'effet réel. RESBos prédit une section efficace fiducielle totale 14% plus élevée que DIPHOX grâce aux corrections NLO du processus $gg \rightarrow \gamma\gamma$.

Les résultats de comparaison (sans facteur d'échelle) pour les différentes observables de mesure sont présentés dans les figures 186 à 188. RESBos et DIPHOX donnent des résultats similaires à haut $m_{\gamma\gamma}$, $p_{T,\gamma\gamma}$, a_T , ϕ_η^* , $|\cos \theta_\eta^*|$ et $\Delta\phi_{\gamma\gamma} < \pi$, largement déficitaires par rapport à la mesure à cause des corrections d'ordre supérieur manquantes dans le calcul (NNLO, N³LO, etc). Les variables $m_{\gamma\gamma}$ et $|\cos \theta_\eta^*|$ sont cependant légèrement mieux décrites en général grâce aux contributions LO et NLO non nulles pour l'ensemble de l'espace des phases. Il est également intéressant de remarquer que DIPHOX fait tout de même mieux que RESBos à grand ϕ_η^* , $|\cos \theta_\eta^*|$ et petit $\Delta\phi_{\gamma\gamma}$, $m_{\gamma\gamma}$ probablement grâce à son meilleur traitement des composantes de fragmentation, importantes dans ces régions (voir section 17.1.4). L'épaule de Guillet est également mieux reproduite par DIPHOX, probablement pour les mêmes raisons.

Cependant, le procédé de resommation de tous les ordres en α_s au NNLL employé par RESBos pour $p_{T,\gamma\gamma} \simeq 0$ permet une bien meilleure stabilisation des résultats dans les régions sensibles aux divergences infrarouges telles que $p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$, où des valeurs de sections efficaces en accord ou quasi-accord avec les mesures expérimentales sont observées. Au contraire, DIPHOX ne parvient pas à compenser efficacement

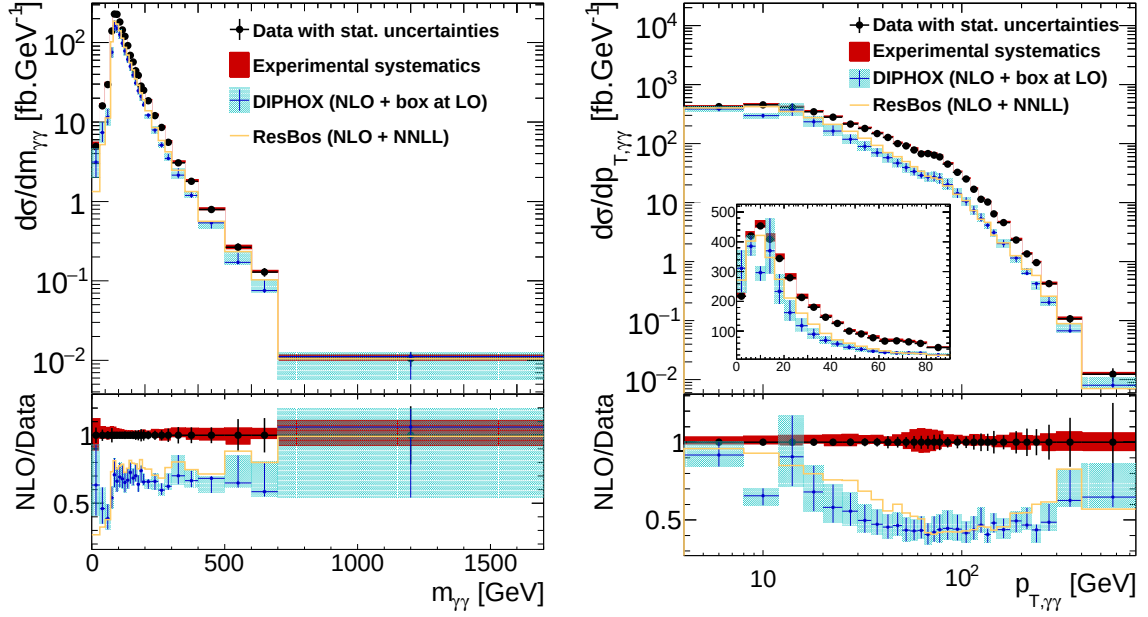


FIG. 186 – Comparaison entre la section efficace différentielle $pp \rightarrow \gamma\gamma + X$ mesurée à $\sqrt{s} = 8$ TeV en fonction de $m_{\gamma\gamma}$ (gauche) et $p_{T,\gamma\gamma}$ (droite) et les prédictions correspondantes de DIPHOX (NLO à ordre fixe) et RESBOS (NLO+NNLL), après déconvolution des effets de détecteur.

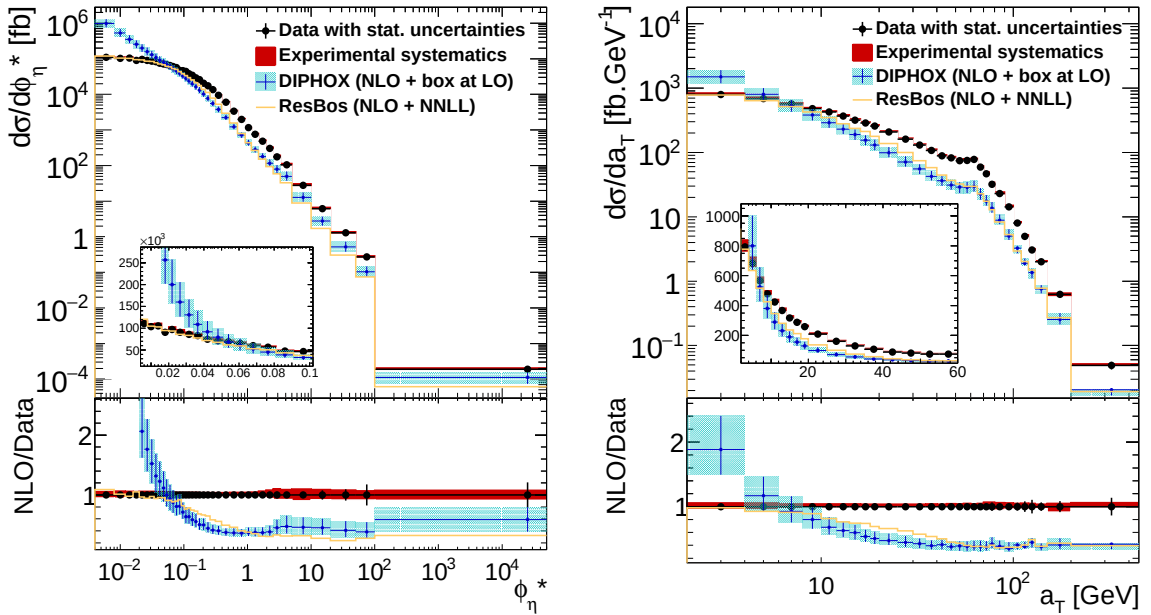


FIG. 187 – Figure similaire à la figure 186 pour a_T et ϕ_{η}^* .

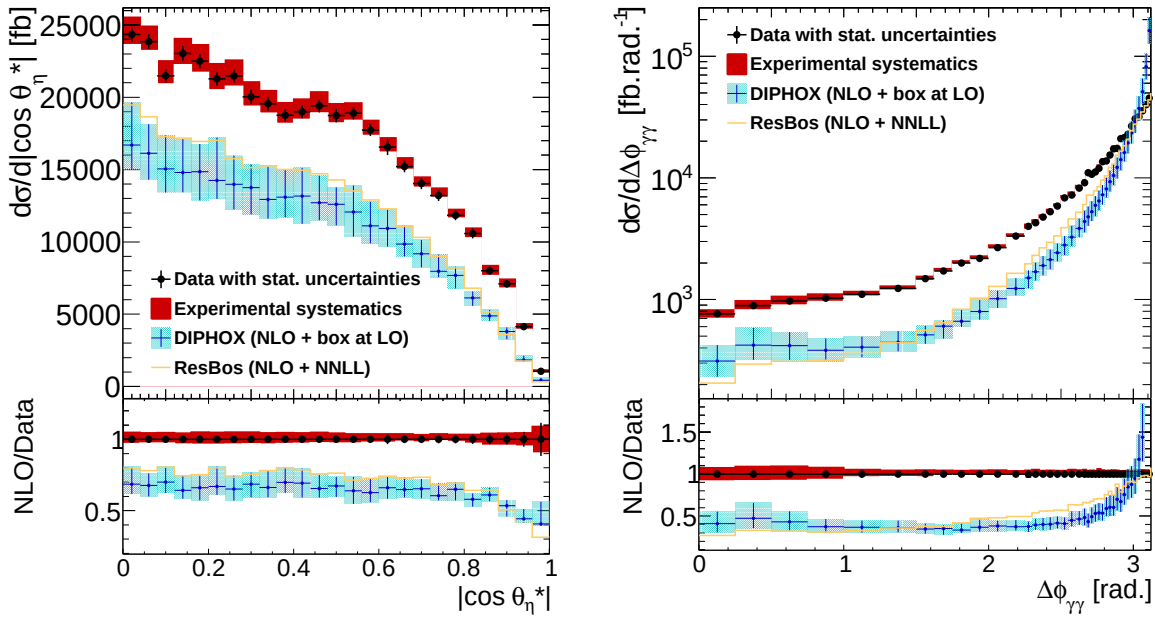


FIG. 188 – Figure similaire à la figure 186 pour $\cos \theta_\eta^*$ et $\Delta\phi_{\gamma\gamma}$.

l'augmentation de la section efficace due à l'émission colinéaire photon/parton des diagrammes initiés par qg (voir section 3.5.2) et mène ainsi à une description très pauvre de ces régions.

En général, l'incertitude théorique de DIPHOX ne couvre pas la différence observée avec les données, indiquant que les variations arbitraires des différentes échelles adoptées dans la section 17.1.4, pourtant classiques au LHC, ne permettent pas d'évaluer l'incertitude théorique de manière fiable au NLO. De façon assez suprenante, celles-ci semblent permettre de couvrir uniquement les corrections NLO de $gg \rightarrow \gamma\gamma$, qui sont de l'ordre de grandeur des différences entre DIPHOX et RESBOS. Cet effet est particulièrement visible pour $|\cos \theta_\eta^*|$. Ainsi, des écarts à plus de 5σ sont observés entre NLO et données pour toutes les observables.

Dans le cadre de l'analyse $\gamma\gamma + X$ d'ATLAS réalisée à $\sqrt{s} = 7$ TeV [203], les prédictions de DIPHOX avaient également été générées afin d'être comparées avec la mesure. Cependant, les prédictions de RESBOS n'avaient pu être générées que pour une partie réduite de l'espace des phases. Comme attendu, l'accord NLO/données est très similaire à celui obtenu à $\sqrt{s} = 8$ TeV. Par ailleurs, les corrections NLO de $gg \rightarrow \gamma\gamma$ ont été ajoutées aux prédictions de DIPHOX à $\sqrt{s} = 7$ TeV via le générateur GAMMA2MC [228]. Une augmentation de la section efficace fiducielle totale de 18% avait alors été observée, en très bon accord avec les 14% de différence observée entre RESBOS et DIPHOX à $\sqrt{s} = 8$ TeV. L'augmentation est légèrement plus faible à $\sqrt{s} = 8$ TeV, probablement simplement parce que RESBOS sous-estime les composantes de fragmentation, comme cela a déjà été mentionné plus haut.

Dans la section 17.2.4, les données sont comparées avec les calculs NNLO à ordre fixe, soit $2\gamma\text{NNLO}$, MCFM (NNLO) et $2\gamma\text{RES}$ (NNLO+NNLL).

17.2.4 2γ NNLO, MCFM et 2γ RES (ordre fixe NNLO en α_s)

Note : les prédictions de MCFM ne sont pas encore disponibles et devraient nous être transmises courant 2016.

Note : les prédictions de 2γ RES ne sont pas encore disponibles et devraient nous être transmises courant 2016.

2γ NNLO [117] constitue le premier calcul à ordre fixe au NNLO, *ie.* $O(\alpha_s^2)$, à avoir été réalisé pour la section efficace $pp \rightarrow \gamma\gamma + X$. Il a été décrit dans la section 17.2.4. Il se limite au calcul des composantes directes de sorte qu'un critère d'isolation de type Frixione est nécessaire pour la convergence des prédictions. La section efficace fiducielle totale prédite par 2γ NNLO est de 14.42 pb, ce qui correspond à environ 87% de la section efficace mesurée et donc constitue une nette amélioration par rapport aux prédictions NLO ($\simeq 70\%$). Les différentes sections efficaces différentielles sont comparées aux mesures dans les figures 189 à 191. Une production préliminaire avec une statistique limitée est employée. Les incertitudes théoriques et statistiques sont pour le moment indisponibles, de même que la distribution correspondant à l'observable a_T . Les prédictions finales incluant l'ensemble de ces résultats sont actuellement en cours de production.

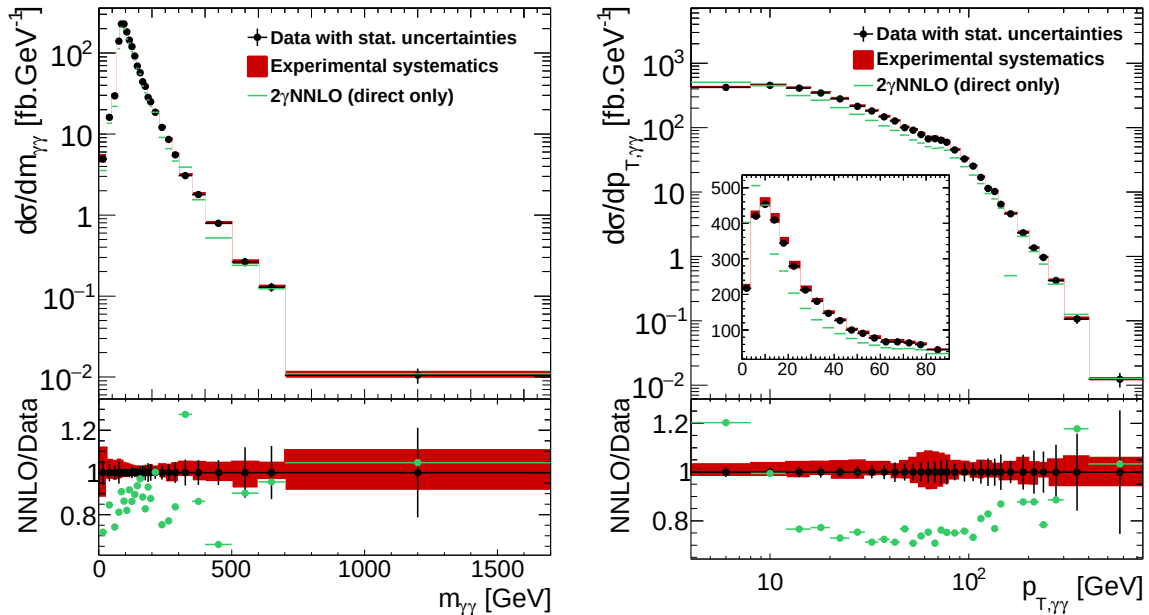


FIG. 189 – Comparaison entre la section efficace différentielle $pp \rightarrow \gamma\gamma + X$ mesurée à $\sqrt{s} = 8$ TeV en fonction de $m_{\gamma\gamma}$ (gauche) et $p_{T,\gamma\gamma}$ (droite) et les prédictions correspondantes de 2γ NNLO, correspondant à des prédictions à ordre fixe au NNLO pour la composante directe avec critère d'isolation de Frixione. **Note :** une production préliminaire avec une statistique limitée est employée. Les prédictions finales incluant l'ensemble de ces résultats sont actuellement en cours de production.

L'accord des prédictions au NNLO en fonction des différentes observables de mesure est en général bien meilleur par rapport aux prédictions NLO, bien qu'il soit difficile à quantifier plus précisément pour le moment à cause des incertitudes manquantes. La distribution en masse semble en bon accord avec la mesure expérimentale pour $m_{\gamma\gamma} > 200$ GeV, tandis que $|\cos \theta_\eta^*|$ est sous estimé d'environ 10 à 20% sur l'ensemble de l'espace des phases, sauf pour $|\cos \theta_\eta^*| > 0.8$ où le désaccord est plus important,

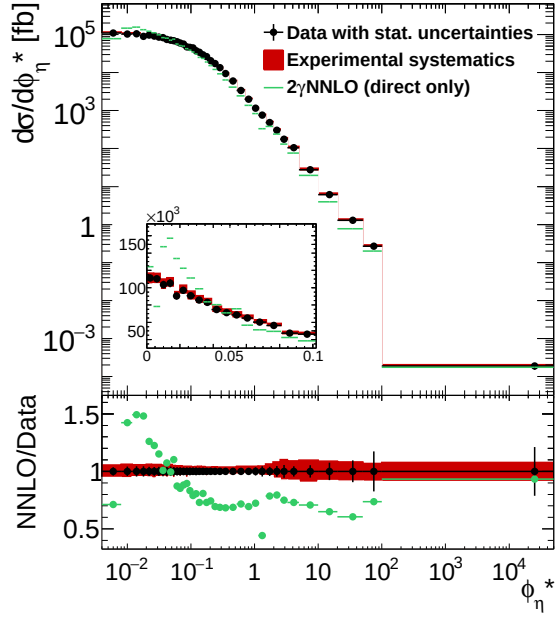


FIG. 190 – Figure similaire à la figure 189 pour ϕ_η^* . **Note :** une production préliminaire avec une statistique limitée est employée. Les prédictions finales incluant l'ensemble de ces résultats sont actuellement en cours de production.

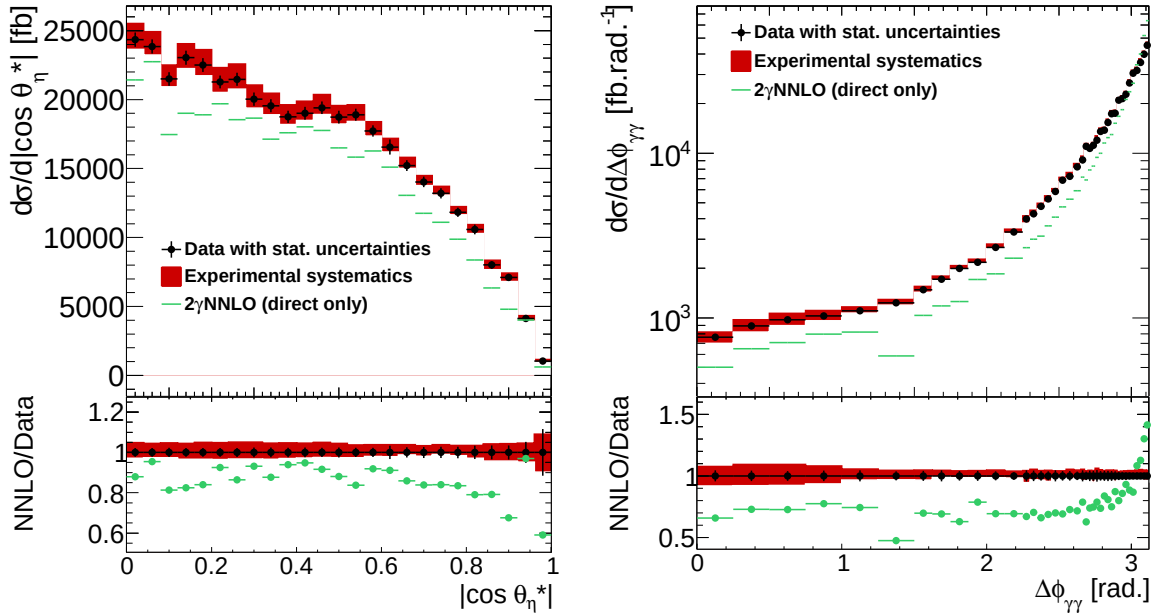


FIG. 191 – Figure similaire à la figure 189 pour $\cos \theta_\eta^*$ et $\Delta\phi_{\gamma\gamma}$. **Note :** une production préliminaire avec une statistique limitée est employée. Les prédictions finales incluant l'ensemble de ces résultats sont actuellement en cours de production.

comme attendu. Tout comme la région $|\cos \theta_\eta^*| > 0.8$, la région à basse masse souffre probablement de la fragmentation et des ordres supérieurs manquants.

Bien que le désaccord baisse significativement comparé aux prédictions NLO, les variables sensibles aux instabilités infrarouges semblent toujours surestimées pour $p_{T,\gamma\gamma}$, $\phi_\eta^* \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$, de l'ordre de 50% au maximum. Il est intéressant de remarquer que le reste de ces distributions est sous-estimé pour ces variables, en particulier la partie intermédiaire. Cela ouvre de bonnes perspectives d'amélioration pour 2 γ RES, qui implémente une resommation au NNLL et permet d'après [102] de rééquilibrer efficacement les désaccords de ce type observés avec 2 γ NNLO. En revanche, les queues de distribution seront probablement peu affectées par la resommation et donc toujours sous-estimées dans 2 γ RES à cause des ordres supérieurs manquants. Le comportement de ϕ_η^* au voisinage de 0 est également quelque peu surprenant et sera examiné plus en détail avec la production finale.

Dans la section 17.3, une conclusion générale est donnée sur la mesure $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV, présentée tout au long de la partie IV.

17.3 Conclusion

La mesure de la section efficace $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV a été réalisée avec les données du détecteur ATLAS accumulées pendant l'année 2012, correspondant à une luminosité intégrée de 20.24 fb⁻¹ (voir figure 192). L'analyse complète des données a pu être effectuée pendant les années de thèse. Elle est en ce moment passée en revue par un comité éditorial interne à la collaboration ATLAS, démarré l'hiver dernier. Celui-ci nous a récemment demandé de commencer à travailler sur la rédaction de l'article associé à la mesure, qui devrait donc être prêt dans les prochains mois. Plusieurs prédictions théoriques devraient également être finalisées dans les prochains mois et incluses dans la publication finale.

La légère augmentation de l'énergie dans le centre de masse ainsi que la quantité bien plus importante de données à $\sqrt{s} = 8$ TeV par rapport à la mesure précédente [203] (20.24 fb⁻¹ vs 4.9 fb⁻¹) a permis non seulement d'augmenter la portée et la résolution de la mesure, mais également de pouvoir étudier plus en détail un certain nombre d'effets systématiques (voir sections 15.6 et 16.6). En particulier, les différentes dépendances de l'isolation calorimétrique des photons (E_T^γ , $|\eta^\gamma|$, $< \mu >$, etc.) ont pu être identifiées, parfois corrigées (voir section 10.4) et prises en compte pour la soustraction du bruit de fond. Une coupure minimale sur l'isolation des traces a également été mise en place afin d'enrichir davantage la région de signal.

Par ailleurs, des distributions d'isolation significativement différentes des photons ont été observées pour les électrons mal identifiés, à la fois dans le détecteur interne et dans le calorimètre électromagnétique, ce qui impose par exemple la soustraction des bruits de fond électron et jet de façon simultanée dans le cas d'un *template fit* 2D de l'isolation. Finalement, des corrélations faibles mais néanmoins importantes entre les énergies d'isolation transverse des deux photons de la paire ont été mises en évidence, bien que leur origine n'ait pas pu être clairement identifiée (événements sous-jacents, corrections de fuites et d'empilement).

L'emploi de pseudo-données afin de valider la procédure de soustraction du bruit de

fond a été déterminante pour l'identification de ces différents effets (voir section 15.4), j'espère donc qu'elle sera reproduite pour les analyses futures à $\sqrt{s} = 13$ TeV. Elle a déjà été reproduite pour les mesures $\gamma + c$ et $\gamma + b$ à $\sqrt{s} = 8$ TeV ainsi que dans le cadre des mesures de pureté effectuées à $\sqrt{s} = 13$ TeV pour les recherches de résonances di-photons autour de 750 GeV [245, 246].

Ces nouvelles études et améliorations de l'analyse ont logiquement débouché sur une baisse globale des incertitudes systématiques par rapport à la mesure précédente [203], parfois d'un facteur 2. En particulier, les deux méthodes de soustraction du bruit de fond (*template fit* 2D, *4x4 matrix*) produisent désormais des résultats compatibles dans les incertitudes après prise en compte des corrélations (voir section 15.7), ce qui permet d'éviter l'ajout en quadrature d'une incertitude systématique additionnelle venant du choix de la méthode de soustraction.

Des pistes d'améliorations pour les analyses futures ont également été suggérées afin de réduire à nouveau les incertitudes systématiques, en particulier dans la section 15.6 : étude et optimisation de la description de l'empilement dans la simulation à partir de variables calorimétriques, mesure du rapport NON-TIGHT/TIGHT dans les désintégrations $Z^0 \rightarrow l^+l^-\gamma$. Des études MC annexes pourraient également se révéler intéressantes afin d'en apprendre plus sur le signal QCD $\gamma\gamma + X$: étude des corrélations d'isolation $\gamma\gamma$, amélioration de la résolution entre l'énergie d'isolation au niveau particule et au niveau reconstruit permettant éventuellement la création de plusieurs catégories pour les mesures futures.

Les distributions d'événements observés après soustraction du bruit de fond ont finalement été corrigées des effets de sélection et de détecteur via une méthode de déconvolution itérative (voir section 16). Cette étape n'est pas la plus problématique grâce à l'excellente résolution du calorimètre d'ATLAS et au travail impressionnant réalisé par le groupe de performance e/γ . Néanmoins, il est important de bien comprendre la forme des différentes corrections et d'étudier les corrélations entre les effets d'acceptance et d'efficacité afin de s'assurer que les chiffres fournis par le groupe de performance pour les photons s'appliquent au signal $\gamma\gamma + X$ (voir section 16.3.4). La propagation des incertitudes dans la procédure de déconvolution est également délicate, certains effets systématiques agissant à la fois sur la soustraction du bruit de fond et sur les matrices de réponse (voir section 16.6.2).

Finalement, la mesure a été comparée avec différentes prédictions représentatives de l'état de l'art pour les énergies du LHC (voir section 17). Certaines sont toujours en cours de production. Bien qu'il puisse paraître inquiétant aux premiers abords que l'analyse principale utilise des échantillons simulés ayant une description très naïve du signal (LO+PS, voir section 17.2.1), les très faibles migrations ainsi que la forme acceptable des distributions de l'échantillon SHERPA 1.4.0 une fois $\sigma_{\text{tot.}}^{\text{fid.}}$ corrigée permettent de garder les incertitudes systématiques sous contrôle. Les calculs $\gamma\gamma + X$ avec cascade de partons de dernière génération apparus récemment dans le domaine (ME+PS *merged at NLO*, voir section 17.1.3) ont également pu être comparés aux données au niveau particule pour la première fois. Ils permettent d'améliorer grandement l'accord données/MC, avec en particulier une bonne échelle pour la section efficace fiducielle totale et une description correcte à la fois des distributions en $m_{\gamma\gamma}$ et en $p_{T,\gamma\gamma}$. Ils apparaissent ainsi comme très prometteurs pour les analyses futures. Des tests sont actuellement en cours au sein de la collaboration ATLAS afin de produire l'échantillon $\gamma\gamma + X$ avec SHERPA 2.2

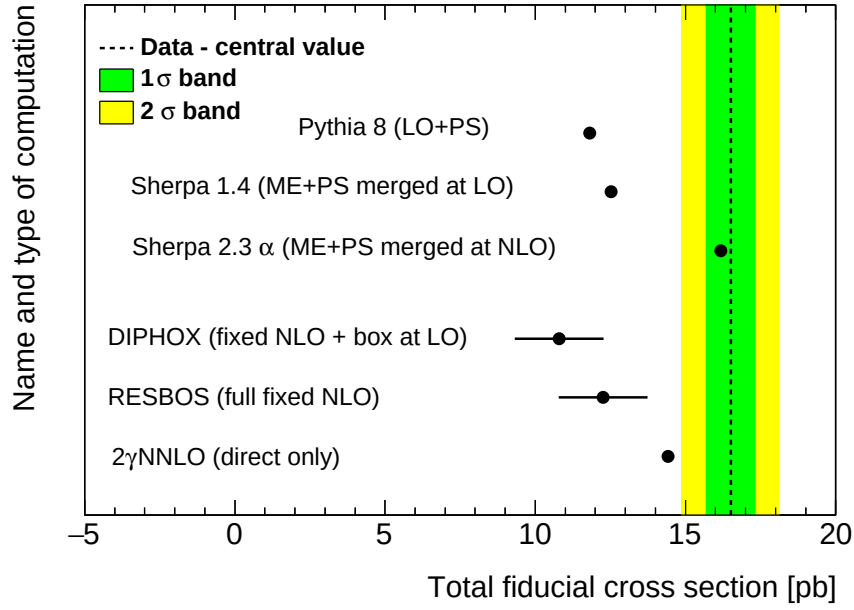


FIG. 192 – Résultats principaux de la production $pp \rightarrow \gamma\gamma + X$ en terme de section efficace fiducielle totale. La mesure est représentée par une ligne pointillée noire. Deux bandes de couleur vertes et jaunes représentent les contours expérimentaux à 1 et 2 σ . Les différentes prédictions considérées dans la section 17 sont également représentées : PYTHIA8 (*LO+PS*), SHERPA 1.4.0 (*ME+PS merged at LO*), SHERPA 2.3 α (*ME+PS merged at NLO*), DIPHOX (*fixed NLO* avec $gg \rightarrow \gamma\gamma$ au LO), RESBOS (*fixed NLO* incluant les premières corrections pour $gg \rightarrow \gamma\gamma$) et 2 γ NNLO (*fixed NNLO* incluant uniquement la production directe). Les incertitudes théoriques sont pour le moment uniquement disponibles au NLO et représentées par des barres horizontales de couleur noire.

(*ME+PS merged at NLO*) et une simulation complète du détecteur à $\sqrt{s} = 13$ TeV. Celui-ci pourrait alors être employé comme échantillon de signal simulé principal dans l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 13$ TeV, mais également comme échantillon de bruit de fond principal pour la recherche de résonances dans le canal di-photon, dont les résultats sont particulièrement attendus cet été après l'observation d'un excès modéré autour de $m_{\gamma\gamma} = 750$ GeV par ATLAS et CMS [8, 9].

Des prédictions partoniques à ordre fixe en α_s , plus rigoureuses du point de vue de la QCD, ont également été comparées avec la mesure. Les prédictions complètes incluant composante directe et composantes de fragmentation ne sont disponibles qu'au NLO, *ie.* $O(\alpha_s)$, via le générateur DIPHOX. Celles-ci présentent des incertitudes très importantes dues aux ordres supérieurs manquants et sous-estiment l'échelle de la mesure de plus de 30%. Par ailleurs, les régions $p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* \gg 0$ et $\Delta\phi_{\gamma\gamma} \ll \pi$ impliquent nécessairement l'émission d'au moins un parton supplémentaire et ne sont par conséquent décrites strictement qu'au LO par DIPHOX, menant à une description pauvre des données. D'autre part, les régions très sensibles aux infrarouges ($p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$) sont instables à cause du développement fixe et limité en α_s , en particulier pour les nouvelles observables de mesure (a_T , ϕ_η^*) qui présentent une excellente résolution à basse valeur et donc une description précise de cette région. Dans ce contexte, le procédé de resommation au NNLL mis en œuvre par RESBOS permet une nette amé-

loration des prédictions et l'obtention de résultats compatibles avec les données dans ces régions. L'ajout des premières corrections du processus $gg \rightarrow \gamma\gamma$ dans RESBOS permet également d'améliorer légèrement l'échelle des prédictions ($\simeq +15\%$ par rapport à DIPHOX). La fragmentation est en revanche moins bien reproduite que dans DIPHOX à cause d'approximations supplémentaires.

Au NNLO, les théoriciens ont pour le moment mis de côté l'idée de produire les trois composantes (directe, 1-fragmentation, 2-fragmentations) et se limitent à des prédictions pour la composante directe uniquement. Un critère d'isolation différent de celui appliqué dans les analyses expérimentales est alors nécessaire pour la convergence du calcul, basé sur un cône à rayon variable dit de Frixione (voir section 3.5.3). Ainsi, des différences de l'ordre de quelques pourcents sont attendues par rapport au volume fiduciel de la mesure.¹⁰¹ Malgré ces différences, les corrections NNLO représentent 30% à 40% de la section efficace fiducielle totale au NLO et par conséquent améliorent considérablement les prédictions. Néanmoins, des écarts avec les données de 10 à 20% en moyenne sont toujours visibles, signes probables des ordres N³LO et de la fragmentation manquante. Les régions sensibles aux infrarouges montrent également certaines difficultés à se stabiliser (écarts NNLO/données jusqu'à 50%), inévitables sans l'implémentation d'un procédé de resommation au regard de l'excellente résolution expérimentale atteinte à $\sqrt{s} = 8$ TeV. Comme évoqué dans [102], une resommation au NNLL pourrait ainsi permettre non seulement d'améliorer considérablement les prédictions NNLO pour $p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$ mais également dans la région intermédiaire $p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* > 0$ et $\Delta\phi_{\gamma\gamma} < \pi$. Dans ce contexte, les résultats de 2 γ RES (NNLO+NNLL) pourraient se montrer particulièrement intéressants. Ils devraient nous parvenir courant 2016.

La mesure est donc généralement en bon accord avec les prédictions du modèle standard dans les incertitudes considérées. Néanmoins, les prédictions montrent quelques difficultés à décrire précisément les données et notamment la distribution en masse, d'abord pour plusieurs raisons spécifiques au signal $\gamma\gamma + X$ (fragmentation, ordres supérieurs, cône de Frixione). En effet, même les calculs MC de dernière génération incluant pourtant les productions jusqu'à $\gamma\gamma + 3$ partons présentent une description limitée de $m_{\gamma\gamma}$ (voir section 17.2.2). Dans le cadre de cette analyse, il n'est par conséquent absolument pas exclu que de nouveaux phénomènes au-delà du modèle standard et faiblement couplés soient déjà présents dans les données. Ceux-ci sont recherchés spécifiquement dans le cadre d'analyses séparées où le bruit de fond QCD est modélisé à l'aide d'un ajustement de la distribution à basse masse. Ces recherches n'ont pour le moment apporté aucun résultat significatif bien qu'un excès modéré autour de $m_{\gamma\gamma} \simeq 750$ GeV ait été observé à $\sqrt{s} = 13$ TeV, à la fois par ATLAS et CMS [8, 198]. En revanche, les régions sensibles aux infrarouges ($p_{T,\gamma\gamma}$, a_T , $\phi_\eta^* \simeq 0$ et $\Delta\phi_{\gamma\gamma} \simeq \pi$) semblent être bien comprises du point de vue de la QCD grâce aux procédés de resommation à tous les ordres au NNLL, ce qui constitue un nouveau succès pour la pQCD. La mesure ainsi que les différentes prédictions théoriques pour la section efficace totale sont rappelées dans la figure 192.

Dans la partie V, le projet phénoménologique principal accompli pendant les années de thèse, visant à l'estimation du potentiel de découverte de nouveaux phénomènes via l'étude de couplages à quatre photons au LHC, est abordé en détail.

¹⁰¹**Note :** celles-ci sont actuellement en cours d'estimation.

Cinquième partie

Étude de la sensibilité des expériences ATLAS et CT-PPS aux couplages à quatre photons

Sommaire

18 Introduction	371
19 Calcul des termes de couplage à quatre photons	373
19.1 Utilisation d'un lagrangien effectif	373
19.1.1 Présentation	373
19.1.2 Application dans le cadre de nouvelles particules chargées . .	374
19.1.3 Application dans le cadre de nouvelles résonances neutres . .	376
19.1.4 Limites de la théorie effective et facteurs de forme	377
19.2 Calcul complet des amplitudes dans le cas générique de boucles de particules chargées	379
19.3 Calcul complet des amplitudes dans le cas générique de résonances de spin 0 et 2	380
20 Le canal di-photon exclusif au LHC avec détection des protons intacts	381
20.1 Les différents modes de production	381
20.1.1 Présentation dans le cadre du modèle standard	381
20.1.2 Le mode gluon-gluon	382
20.1.3 Le mode photon-photon	382
20.2 Mise en évidence expérimentale	385
20.2.1 Utilisation de détecteurs de protons à l'avant du faisceau . .	385
20.2.2 Stratégies adoptées à haute et basse énergie	386
20.2.3 État des lieux des différentes mesures	388
20.3 Les prédictions du modèle standard	389
20.3.1 Présentation	389
20.3.2 Discussion d'une mesure au LHC	391
21 Étude des sensibilités aux couplages anormaux à quatre photons à haute énergie	393
21.1 Scénarios considérés	393
21.2 Composition et simulation du bruit de fond	394
21.2.1 Les différents types de bruits de fond	394
21.2.2 Bruits de fond avec protons intacts à l'avant	395
21.2.3 Bruits de fond avec protons intacts d'empilement à l'avant . .	397
21.3 Prise en compte des effets de détecteurs	402
21.3.1 Taux de conversion des photons	402

21.3.2	Efficacité de présélection des photons et des électrons	403
21.3.3	Taux de mauvaise identification des électrons et des jets . . .	403
21.3.4	Résolution sur les observables mesurées	404
21.4	Sélection des pseudo-données	405
21.4.1	Distribution du signal et des bruits de fond sans sélection supplémentaire	405
21.4.2	Sélection des hautes masses et des productions dos-à-dos dans le détecteur central	406
21.4.3	Identification des systèmes centraux et à l'avant	407
21.4.4	Bilan	407
22	Sensibilités finales, comparaisons aux limites actuelles et perspectives futures	410
22.1	Estimation de la sensibilité des données	410
22.2	Résultats obtenus par la théorie effective	411
22.3	Limites génériques sur l'existence de nouvelles particules chargées . .	413
22.4	Limites sur l'existence de particules neutres de spin 0/2	414
22.5	Conclusion et perspectives	416

18 Introduction

Il existe un certain nombre d'arguments théoriques plaidant en faveur de l'existence de nouveaux phénomènes à l'échelle du TeV (voir section 2.3). Cependant, après la confirmation de l'existence du boson de Higgs prévu dans le cadre du modèle standard [2, 3] et l'analyse des données du *run 1* du LHC, un grand nombre de modèles favorisés jusqu'à maintenant ont été exclus ou réduits à un espace des paramètres limité et moins naturel (CMSSM, ...) [46]. Il est alors intéressant d'explorer le potentiel de découverte de nouvelle physique d'une part pour des processus encore peu étudiés à l'échelle du TeV, et d'autre part de façon plus robuste en s'affranchissant des dépendances aux différents scénarios. Les processus mettant en jeu des couplages à quatre photons (4γ), déjà introduits dans la section 2.4, présentent ces caractéristiques.

La partie V consiste en un compte-rendu détaillé des résultats obtenus sur la sensibilité des expériences ATLAS-AFP et CT-PPS (voir partie II) aux couplages anormaux à quatre photons à haute énergie. Ce sujet constitue le travail phénoménologique principal effectué pendant les années de thèse. Déjà introduite brièvement dans la section 2.4.3, l'étude se décompose comme suit.

Dans la section 19, les calculs théoriques des termes de couplage à quatre photons développés dans les références [7, 55–57] et utilisés pour l'étude sont présentés. Partant d'une approche effective valide uniquement pour des termes au-delà du modèle standard apparaissant à très haute énergie, le calcul complet des amplitudes est ensuite exposé dans le cas de contributions génériques correspondant à des boucles de particules chargées et à des résonances de spin 0 et 2. Contrairement à la théorie effective, ce dernier est valide pour une très large gamme d'énergie et peut donc être utilisé à la fois pour les processus de nouvelle physique et pour ceux du modèle standard. Par ailleurs, quelques modèles spécifiques donnant lieu à des couplages anormaux à 4γ sont discutés.

Dans la section 20, le canal de production exclusif de paires de photons avec protons intacts au LHC est introduit. C'est l'état final qui a été retenu pour étudier les sensibilités des expériences ATLAS-AFP et CT-PPS aux couplages anormaux à quatre photons à haute énergie. Son étude nécessite l'installation de détecteurs de protons à l'avant, prévue au plus tard début 2017 selon la programmation actuelle (voir section 6.2.1),¹⁰² et présente une section efficace plus faible comparée aux canaux classiques sans proton intact mais comprend par ailleurs de nombreux avantages : sa contamination par le mode de production gluon-gluon (voir section 2.4.1) est quasi-nulle à haute énergie, il permet un excellent contrôle des autres bruits de fond (voir section 21), souvent limitant dans ce genre d'étude [271], et bénéficie de prédictions théoriques précises [272]. Des études basées sur des canaux similaires pour la recherche des couplages anormaux entre photons et bosons W ou Z sont déjà présentes dans la littérature [74, 75].

La section 20 se conclut par la présentation des sections efficaces de di-photons exclusifs attendues dans le cadre du modèle standard (section 20.3). La totalité des contributions est prise en compte avec interférences comprises, en particulier celles associées aux boucles de W^+W^- , souvent négligées dans la littérature. La mesure d'un tel état final dans les collisions de protons du LHC à $\sqrt{s} = 14$ TeV est alors discutée. Cette discussion fait écho à la proposition récente d'observer la diffusion $\gamma\gamma$ via ce même état

¹⁰²CT-PPS devrait être prêt à prendre des données dès l'été 2016.

final avec les données de collision plomb-plomb du LHC [59]. Par ailleurs, la collaboration CDF au TEVATRON et la collaboration CMS au LHC ont récemment publié des mesures de production exclusive de paires de photons dans des collisions protons, mais à des énergies de collision moins élevées et sans avoir pu mettre en évidence le mode de production photon-photon [273, 274].

Dans la section 21, des détails sont apportés sur l'étude principale, qui vise à estimer la sensibilité d'ATLAS-AFP et CT-PPS aux couplages anormaux à quatre photons à haute énergie. La composition et la simulation des différents bruits de fond, la prise en compte des effets de détecteurs et la sélection des données sont abordées. La sélection tire parti à la fois des mesures du détecteur central et des futurs détecteurs de protons à l'avant afin de rejeter très efficacement les différents bruits de fond. Le nombre d'événements observés au fur et à mesure des coupures est donné pour les différents processus (signal, bruits de fond).

Finalement, les résultats finals de sensibilité pour la découverte de couplages anormaux à quatre photons à haute énergie dans ATLAS-AFP et CT-PPS via les di-photons exclusifs sont dévoilés dans la section 22. Leur compétitivité avec les mesures classiques est démontrée en confrontant les limites obtenues aux résultats actuels et à venir au LHC pour un certain nombre de modèles de nouvelle physique de référence. D'autres canaux potentiellement sensibles sont ensuite discutés, préparant la voie pour des études futures. Une parenthèse est également faite sur le rôle que pourrait jouer les couplages anormaux à quatre photons en cas de confirmation du léger excès présent dans le spectre en masse des di-photons, rapporté à la fois par les expériences ATLAS-AFP et CMS dans les données à $\sqrt{s} = 13$ TeV de 2015 [8, 9].¹⁰³

¹⁰³J'ai participé à deux études effectuées par la collaboration ATLAS-AFP dans le cadre de cette analyse : la première traite d'un certain nombre de vérifications et de comparaisons faites avec les données à $\sqrt{s} = 8$ TeV [275] et la seconde consiste en différentes mesures de bruit de fond à haute masse réalisées à $\sqrt{s} = 13$ TeV [245].

19 Calcul des termes de couplage à quatre photons

19.1 Utilisation d'un lagrangien effectif

19.1.1 Présentation

Dans cette section, un calcul effectif tiré des références [7, 55–57] et mettant en jeu un terme de couplage à quatre photons correspondant à la figure 193 est introduit, correspondant à la diffusion $\gamma\gamma$ ou *light-by-light scattering* (LbyL, voir section 2.4). En effet, dans le cas d'une physique au-delà du modèle standard située à une échelle en énergie Λ très supérieure à l'énergie E accessible expérimentalement, il est possible de décrire les manifestations associées à basse énergie par le biais d'un lagrangien effectif, valide pour $\Lambda \gg E$ (voir section 2.4.2). Parmi l'ensemble des choix possibles, il a été démontré dans [55] que dans le cas de l'existence du boson de Higgs (observé expérimentalement depuis par ATLAS et CMS en 2012 [2, 3]), l'opérateur de dimension 8 présenté dans l'équation 92 est celui de dimensions minimales respectant les symétries du modèle standard tout en introduisant un terme de couplage à quatre photons :¹⁰⁴

$$\mathcal{L}_{4\gamma} = \zeta_1 F_{\mu\nu} F^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} + \zeta_2 F_{\mu\nu} F^{\nu\rho} F_{\rho\lambda} F^{\lambda\mu}. \quad (92)$$

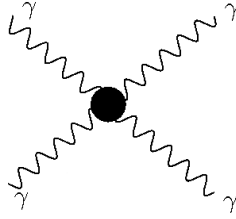


FIG. 193 – Diagramme correspondant à un couplage effectif à quatre photons.

$\mathcal{L}_{4\gamma}$ est composé de deux termes distincts formés par des combinaisons du tenseur électromagnétique $F_{\mu\nu}$ et paramétrés par les couplages anormaux ζ_1 , ζ_2 , homogènes à des TeV^{-4} et nuls dans le cas du modèle standard. Ces deux paramètres sont étudiés séparément dans le cadre de la théorie effective, bien qu'ils soient généralement tout deux prédits comme non-nuls dans un modèle complet de nouvelle physique.

Partant du lagrangien effectif de l'équation 92, il est possible d'écrire de façon simple la section efficace non-polarisée de la diffusion $\gamma\gamma$ à haute énergie associée en fonction des variables de Mandelstam s , t et des paramètres de couplages [57] :¹⁰⁵

$$\frac{d\sigma}{d\Omega} = \frac{1}{16\pi^2 s} (s^2 + t^2 + st)^2 [48(\zeta_1)^2 + 40\zeta_1\zeta_2 + 11(\zeta_2)^2], \quad (93)$$

où $\frac{d\sigma}{d\Omega}$ est la section efficace de production différentielle à intégrer sur un volume fiduciel donné, s et t sont les variables de Mandelstam et ζ_1 , ζ_2 sont les paramètres des

¹⁰⁴Les opérateurs de dimension 6 sont interdits par l'invariance de jauge pour les couplages quartiques neutres $\gamma\gamma\gamma\gamma$, $gg\gamma\gamma$ et $ggZZ$ [55].

¹⁰⁵Les variables de Mandelstam sont des invariants de Lorentz définis de la façon suivante pour une réaction de type $1+2 \rightarrow 3+4$: $s = (p_1 + p_2)^2$, $t = (p_1 - p_3)^2$, $u = (p_1 - p_4)^2$, où les p_i sont les vecteurs quadri-impulsions des particules correspondantes.

couplages anormaux introduits au paragraphe précédent. Cette formulation ne prend pas en compte les interférences potentielles des processus de nouvelle physique avec la production 4γ du modèle standard et n'est donc valide qu'à suffisamment haute énergie, lorsque la production prédite par le modèle standard est négligeable. Cette condition est cependant toujours vérifiée au LHC à haute énergie, comme cela sera démontré dans la section 21.

Il est alors possible de se limiter à l'étude des deux termes de couplage séparément, *ie.* $(\zeta_1 \neq 0, \zeta_2 = 0)$ et $(\zeta_1 = 0, \zeta_2 \neq 0)$, afin de reconstituer les sensibilités du plan 2D complet (ζ_1, ζ_2) . Pour cela on introduit le paramètre ζ , défini à partir de l'équation 93 par la formule suivante :

$$\zeta^2 \equiv 48(\zeta_1)^2 + 40\zeta_1\zeta_2 + 11(\zeta_2)^2. \quad (94)$$

ζ est de même dimension que les ζ_i et s'exprime en TeV^{-4} . Une sensibilité donnée en terme de ζ (ou $\zeta_1 \neq 0, \zeta_2 = 0$ par exemple) définit une ellipse dans le plan (ζ_1, ζ_2) . Cette propriété permet de simplifier grandement l'analyse en réduisant l'espace des paramètres à une dimension unique sans perte d'information et sera utilisée pour produire les résultats finals de l'étude à haute énergie (voir section 22).

Cependant, cette approche ne permet pas de distinguer expérimentalement le rapport des contributions entre les termes associés à ζ_1 et ζ_2 (voir équation 92), ni même leur signe. Aucune différence fondamentale dans les distributions angulaires des signaux issus des couplages ζ_1 et ζ_2 n'a été observée au cours de l'étude. En principe, des observables plus complexes pourraient néanmoins apporter ce genre d'information, par exemple en exploitant les différents modes de polarisation de l'état final.

Dans un contexte expérimental, la théorie effective permet de calculer simplement des limites sur le couplage de la nouvelle physique tout en s'affranchissant du détail des différents scénarios. Elle est également très utile afin de comparer avec un même formalisme les contributions au processus de LbyL attendues pour un certain nombre de modèles génériques allant au-delà du MS. Dans les sections 19.1.2 et 19.1.3, différents exemples d'application illustrant ces avantages sont abordés.

19.1.2 Application dans le cadre de nouvelles particules chargées

Dans le cas de l'existence de nouvelles particules chargées, des diagrammes supplémentaires prenant part au processus de LbyL sont attendus et schématisés dans la figure 194, gauche. La section efficace associée est alors entièrement définie par la charge électrique Q , la masse m , le spin S et la multiplicité N des nouveaux objets [56, 57]. La multiplicité correspond au nombre de degrés de liberté des nouvelles particules vis-à-vis de l'électromagnétisme. Elle est souvent supérieure à un dans les modèles complets : par exemple, la multiplicité attendue est de $N = 3$ si une nouvelle particule porte une charge de couleur. Il est alors commode d'introduire la charge effective Q_{eff} , définie par :

$$Q_{\text{eff}}^4 = \text{tr } Q^4 \simeq N \cdot Q^4, \quad (95)$$

où la trace est calculée sur l'ensemble des nouvelles particules dégénérées approximativement en masse. Par exemple, dans le modèle de Higgs composite minimal [33], la

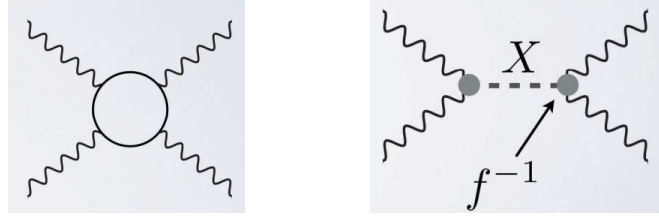


FIG. 194 – Contributions typiques de nouvelles particules chargées (gauche) et résonances neutres (droite) au processus de diffusion $\gamma\gamma$. Ces dernières font apparaître un nouveau couplage f^{-1} associé au vertex $\gamma\gamma X$.

théorie prédit un certain nombre de nouvelles particules colorées et chargées dont de nouveaux fermions dit *vector-like* (VL), approximativement dégénérés en masse. Leur charge effective correspond alors à $Q_{\text{eff}} = 2.22$ et 3.80 dans les deux formulations les plus simples, et peut augmenter considérablement selon la taille du groupe et le type de représentation choisi [57]. La charge effective n'est pas forcément entière à cause des différentes particules ayant des charges fractionnelles (ex : top $5/3$).

Il est alors possible de traduire les contributions génériques de nouvelles particules chargées en termes de couplages effectifs associés à $\mathcal{L}_{4\gamma}$ par la relation correspondante :

$$\zeta_i = \frac{\alpha_{\text{em}}^2 Q_{\text{eff}}^4}{m^4} c_{i,S}, \quad (96)$$

avec

$$c_{1,S} = \begin{cases} \frac{1}{288} & S = 0 \\ -\frac{1}{36} & S = \frac{1}{2} \\ -\frac{5}{32} & S = 1 \end{cases}, \quad c_{2,S} = \begin{cases} \frac{1}{360} & S = 0 \\ \frac{7}{90} & S = \frac{1}{2} \\ \frac{27}{40} & S = 1 \end{cases}, \quad (97)$$

où les ζ_i sont les paramètres de couplages anormaux, α_{em} est la constante de couplage électromagnétique, Q_{eff} est la charge effective, m la masse et S le spin. Ainsi, la présence de nouveaux quarks VL comportant une masse située autour de 500 GeV et une charge effective de 3 (4), pas totalement exclue par les limites actuelles du LHC, se traduirait par des couplages anormaux à quatre photons ζ_i de l'ordre de 0.01 (0.1) TeV^{-4} .¹⁰⁶ Les contributions attendues pour des particules scalaires ($S = 0$) se situent un ordre de grandeur en-dessous de celles des fermions et des vecteurs et ne sont pas considérées dans l'étude. Il est également intéressant de remarquer que dans le cas de fermions ($S = \frac{1}{2}$), $\mathcal{L}_{4\gamma}$ se réduit au lagrangien bien connu d'Euler-Heisenberg [49].

D'autre part, les coefficients de l'équation 97 suggèrent une augmentation rapide de la section efficace avec le spin des nouvelles particules. Cela offre donc a priori une excellente opportunité de tester simplement les extensions du modèle standard prédisant l'existence de résonances avec des spins exotiques, comme les modèles introduisant un nouveau secteur fort ou ceux impliquant des cordes de basses énergies [276]. Cependant, les états de spin élevés impliquent de façon implicite des interactions d'ordre

¹⁰⁶Comme nous le verrons dans la section 19.1.4, la théorie effective nécessite néanmoins d'être corrigée pour des masses de cette valeur.

supérieur nécessitant des développements théoriques importants afin de pouvoir être décrits correctement dans le cadre d'une théorie effective. Des études sont en cours mais ne devraient pas voir le jour avant plusieurs années.

Dans la section 19.1.3, le cas de contributions venant de nouvelles résonances neutres est discuté.

19.1.3 Application dans le cadre de nouvelles résonances neutres

Un autre cas intéressant provient de contributions génériques venant de nouvelles résonances neutres couplées aux photons, représentées schématiquement dans la figure 194, droite.¹⁰⁷ Celles-ci ne peuvent pas être de spin 1 par conservation de l'hélicité (théorème de Landau-Yang), par conséquent uniquement les spins 0^+ , 0^- (scalaires et pseudo-scalaires) et spin 2 sont étudiés, notés respectivement φ , $\tilde{\varphi}$ and $h^{\mu\nu}$. Leur description implique nécessairement l'introduction d'un terme de couplage f_S non-renormalisable et homogène à une masse au carré. Le lagrangien effectif associé est totalement contraint si l'on demande la préservation des symétries de Lorentz et de CP [7]. Il peut alors s'écrire :

$$\begin{aligned} \mathcal{L}_{\gamma\gamma} = & f_{0^+}^{-1} \varphi (F_{\mu\nu})^2 + f_{0^-}^{-1} \tilde{\varphi} F_{\mu\nu} F_{\rho\lambda} \epsilon^{\mu\nu\rho\lambda} \\ & + f_2^{-1} h^{\mu\nu} [-F_{\mu\rho} F_\nu{}^\rho + \eta_{\mu\nu} (F_{\rho\lambda})^2 / 4]. \end{aligned} \quad (98)$$

Il génère des interactions à quatre photons supplémentaires à l'ordre le plus bas de la théorie des perturbations correspondant à des couplages anormaux $\zeta_i = (f_S m)^{-2} d_{i,s}$, où :

$$d_{1,S} = \begin{cases} \frac{1}{2} & S = 0^+ \\ -4 & S = 0^- \\ -\frac{1}{8} & S = 2 \end{cases}, \quad d_{2,S} = \begin{cases} 0 & S = 0^+ \\ 8 & S = 0^- \\ \frac{1}{2} & S = 2 \end{cases}, \quad (99)$$

avec S le spin. Le signe de la parité est également mentionné en indice pour les spins 0 (scalaire ou pseudo-scalaire). Contrairement aux particules chargées, le couplage est alors décroissant vis-à-vis du spin.

De telles contributions sont prédites dans le cadre d'un grand nombre de modèles, notamment ceux postulant l'existence de dimensions supplémentaires de l'espace. Par exemple, des gravitons de Kaluza-Klein (spin 2) [277, 278] tels que ceux décrits par les modèles de Randall-Sudrum [279] contribuent à hauteur de :¹⁰⁸

$$\zeta_1 = -\frac{\kappa^2}{64\tilde{k}^4}, \quad \zeta_2 = \frac{\kappa^2}{16\tilde{k}^4}, \quad (100)$$

¹⁰⁷Dans le contexte de l'étude basée sur le canal $\gamma\gamma$ exclusif à haute énergie, on s'attend à obtenir des sensibilités compétitives uniquement pour les résonances larges, là où les recherches directes présentent une sensibilité limitée (voir section 22.2).

¹⁰⁸Cet ordre de grandeur est valide dans le cas où le modèle standard se situe sur la membrane infra-rouge du modèle.

où $\tilde{k}$ est relié à la courbure des dimensions supplémentaires et fixe la masse du premier graviton à $m_{\text{KK}} \simeq 3.8\tilde{k}$, tandis que κ représente l'intensité du couplage aux dimensions supplémentaires et peut être considéré comme un $O(1)$. Ainsi, dans un scénario réaliste où $\kappa = 2$ et le premier graviton a une masse de 5 TeV, on obtient des couplages anormaux de l'ordre de -0.02 et 0.1 TeV^{-4} pour ζ_1 et ζ_2 , respectivement.

On peut également citer le cas du dilaton fortement couplé : dans le cas de l'existence d'une nouvelle interaction fortement couplée respectant la symétrie conforme dans la limite à haute énergie [280–282], il est naturel que celle-ci soit brisée explicitement à plus basse énergie afin de différencier l'échelle électrofaible ($\simeq 200 \text{ GeV}$) de l'échelle de confinement de la QCD ($< 1 \text{ GeV}$). Cette brisure de symétrie (voir section 2.2) provoque l'apparition d'une particule scalaire 0^+ appelée dilaton, dont la masse se situe naturellement au voisinage de l'échelle d'énergie de la brisure. Afin de le différencier du dilaton léger souvent cité dans la littérature qui fait appel à des réglages fins (*fine-tuning*), on appellera le dilaton lourd de masse naturelle le *strongly-interacting heavy dilaton* (SIHD) dans la suite. Si le photon est lui-même un état composite de cette nouvelle interaction, il est fortement couplé au SIHD. Notons qu'il est possible dans certaines situations de faire correspondre le SIHD au radion¹⁰⁹ par le biais d'une correspondance AdS/CFT [57]. Dans tous les cas, on obtient des couplages anormaux à quatre photons de la forme suivante :

$$\zeta_1 = \frac{\pi^2}{2m_\varphi^4}, \quad \zeta_2 = 0. \quad (101)$$

Ainsi, un SIHD de 4 TeV donnerait lieu à des couplages anormaux de l'ordre de $\zeta_1 = 0.02 \text{ TeV}^{-4}$. Une des spécificités de ce modèle vient de la prédiction nulle pour ζ_2 . Malheureusement, cette caractéristique n'est pas exploitable facilement en utilisant le formalisme actuel (voir section 19.1.1).

Dans cette section, une théorie effective a été utilisée afin de générer des couplages anormaux à quatre photons. Toute théorie effective présente toutefois un domaine de validité bien déterminé qu'il convient de bien définir. Cela est fait dans la section 19.1.4.

19.1.4 Limites de la théorie effective et facteurs de forme

Les prédictions des théories effectives ne sont valides que dans un domaine limité de l'espace des phases. La première condition de validité, déjà énoncée dans la section 19.1.1, stipule que l'échelle en énergie de la nouvelle physique Λ doit être très supérieure à l'énergie E accessible expérimentalement, $\Lambda \gg E$. Elle représente la limite de découplage [7]. En effet, à partir d'une certaine échelle d'énergie $\Lambda \simeq E$, les effets de nouvelle physique ne peuvent plus être traités comme des petites perturbations du modèle standard et des corrections d'ordres supérieurs deviennent nécessaires à la régularisation des sections efficaces. Le calcul de ces corrections requiert alors un modèle plus complet, tout comme la description de systèmes comportant des vitesses proches de celle de la lumière requiert les équations complètes de la relativité lorsque les effets relativistes deviennent trop importants.

¹⁰⁹Le radion est une particule scalaire apparaissant dans le cadre de scénarios de dimensions supplémentaires de Randall-Sudrum.

En général, les acceptances expérimentales n'influent pas directement l'échelle en énergie des interactions observées. Le domaine de validité exact de limites tirées de comparaisons entre données et prédictions effectives n'est alors pas trivial à déterminer à cause de la limite de découplage. Comme nous le verrons dans la section 20, le canal proposé pour l'étude des couplages anormaux à quatre photons à haute énergie utilise les détecteurs de protons à l'avant. Ceux-ci comportent une acceptation en masse limitée dans le cas de deux détections en coïncidence ($m_{\gamma\gamma}^{\max} \leq 2 \text{ TeV}$, voir section 6.2.2), restreignant de facto la mesure à des processus ayant une échelle en énergie inférieure à $E_{\max} \simeq 2 \text{ TeV}$. Cette caractéristique est commode pour définir de façon conservatrice le domaine de validité de la théorie effective, en considérant $E_{\max}$ comme limite de découplage. En effet, tous les événements observés dans les détecteurs à l'avant ont automatiquement une échelle en énergie $E < E_{\max}$. On adoptera donc la limite de découplage suivante :

$$\Lambda > E_{\max} \quad \text{decoupling limit.} \quad (102)$$

La seconde condition de validité vient de la nature quantique de la théorie effective considérée ici. Toute théorie effective étant par définition non-renormalisable, on s'attend à ce que ses prédictions divergent à haute énergie et finissent par aller à l'encontre de la propriété d'unitarité, qui fixe le module de la fonction d'onde associée à tout système quantique à 1. L'unitarité permet de s'assurer que la théorie respecte bien le principe de causalité. Il est possible d'estimer la limite d'unitarité en utilisant une méthode appelée analyse des ondes partielles, détaillée dans [35]. On obtient alors les deux conditions suivantes [57] :

$$(4\zeta_1 + 3\zeta_2)s^2 < 4\pi, \quad (4\zeta_1 + \zeta_2)s^2 < \frac{12}{5}\pi, \quad (103)$$

à partir des amplitudes invariantes correspondant aux états de polarisations $(++++)$ et $(++--)$, respectivement. Comme la grande majorité des événements que l'on étudiera dans la section 21 sont produits avec un $\sqrt{s}$ inférieurs à environ 1 TeV à cause de l'acceptation limitée des détecteurs à l'avant (voir paragraphe précédent), on obtient comme ordre de grandeur pour la limite d'unitarité :

$$\zeta_i < (1 - 10) \text{ TeV}^{-4} \quad \text{unitarity limit.} \quad (104)$$

Comme nous le verrons dans la section 21, celle-ci n'est pas contraignante en pratique car les sensibilités de l'étude se situent un à deux ordres de grandeur plus bas et donc sont intéressantes uniquement pour des modèles ne violant pas l'unitarité.

Si l'une des deux conditions énoncées ci-dessus (découplage, unitarité) n'est pas vérifiée, il est courant de corriger très grossièrement la théorie effective en utilisant un facteur de forme [283], appliqué au niveau des amplitudes invariantes. La forme exacte de celui-ci est largement arbitraire et vise à imiter le processus de régularisation de la section efficace à l'approche du seuil de production de nouvelle physique. Ainsi, les facteurs de forme constituent souvent des fonctions de Green. Dans la suite de l'étude,

la forme suivante sera considérée :

$$1/(1 + (m_{\gamma\gamma}^2/\Lambda'^2)^2), \quad (105)$$

où $m_{\gamma\gamma}$ est la masse invariante du système di-photon et $\Lambda' = 1$ TeV. Dans le cadre des modèles discutés dans les sections précédentes (19.1.2 et 19.1.3), le facteur de forme est nécessaire uniquement pour des nouvelles particules de masses inférieures à 2 TeV (1 TeV) pour les résonances neutres (particules chargées). Au-delà, le seuil de production n'est jamais atteint grâce à l'acceptance limitée des détecteurs de protons à l'avant et la théorie effective reste valide. Les résultats de sensibilité des expériences ATLAS-AFP et CT-PPS sur les couplages anormaux ζ_i seront présentés dans le cas de la théorie effective avec et sans facteur de forme dans la section 22.

Afin d'éviter l'utilisation d'un facteur de forme *ad hoc* pour des modèles de nouvelle physique à plus basse énergie, il est possible de calculer les amplitudes invariantes complètes dans le cas des deux processus génériques introduits respectivement dans les sections 19.1.2 et 19.1.3. Cela permet de tenir compte de la totalité des dépendances intervenant dans les diagrammes associés et donc de s'affranchir des conditions de validité de la théorie effective, tout en retrouvant les mêmes résultats dans la limite à haute énergie ($\Lambda \gg E_{\max}$). Cela permet également de pouvoir générer les différentes contributions prévues par le modèle standard avec interférences incluses au LO (voir section 20). Des détails sur le calcul des amplitudes complètes sont donnés dans les sections 19.2 et 19.3.

19.2 Calcul complet des amplitudes dans le cas générique de boucles de particules chargées

Considérons tout d'abord le cas de nouvelles particules chargées. En toute généralité, les diagrammes qui leur sont associés mettant en jeu un couplage à quatre photons sont représentés dans la figure 195 à l'ordre le plus bas de la théorie des perturbations.

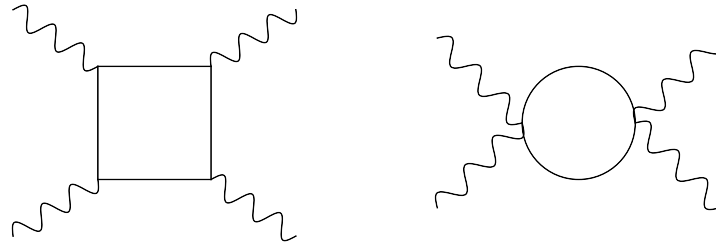


FIG. 195 – Diagrammes typiques correspondant aux contributions de nouvelles particules chargées aux couplages à quatre photons.

Ces diagrammes ont été calculés dans le cas de fermions dans [284] et dans le cas de bosons dans [52]. Par ailleurs, les corrections à un ordre supérieur en QED et QCD ont été évaluées dans [54] et ont un impact négligeable sur les sections efficaces. Seules les corrections électrofaibles n'ont pas été estimées, mais leur effet est supposé du même ordre de grandeur que les corrections QED/QCD et donc négligé.

Dans la suite, les notations et les conventions de normalisation de [284] sont adoptées. La section efficace non-polarisée de diffusion $\gamma\gamma$ correspondant aux processus génériques représentés dans la figure 195 peut être décomposée en fonction des différentes configurations d'hélicité :

$$\frac{d\sigma}{d\Omega} = \frac{\alpha_{\text{em}}^4 Q_{\text{eff}}^8}{2\pi^2 s} (|\mathcal{M}_{++++}|^2 + |\mathcal{M}_{++--}|^2 + |\mathcal{M}_{+-+-}|^2 + |\mathcal{M}_{+--+}|^2 + 4|\mathcal{M}_{+++-}|^2), \quad (106)$$

où les $\mathcal{M}_{ijkl}$ sont les amplitudes invariantes des différentes configurations d'hélicité (fonction des variables de Mandelstam s , t et u), α_{em} est la constante de couplage électromagnétique et Q_{eff} est la charge effective. Les variables de Mandelstam et Q_{eff} ont été définies précédemment (voir section 19.1.2). Finalement, seules les configurations $\mathcal{M}_{++++}$, $\mathcal{M}_{++--}$ et $\mathcal{M}_{+++-}$ ont besoin d'être calculées grâce aux relations suivantes, conséquences des propriétés de parité et de symétrie de Bose des photons [52] :

$$\begin{aligned} \mathcal{M}_{+-+-}(s, t, u) &= \mathcal{M}_{++++}(u, t, s), \\ \mathcal{M}_{+--+}(s, t, u) &= \mathcal{M}_{++++}(t, s, u) \end{aligned} \quad (107)$$

Le résumé complet des différentes amplitudes à une boucle (LO) utilisées dans les simulations des couplages à quatre photons durant les années de thèse est disponible dans les appendices de la référence [57]. Les cas limites (basse énergie $s, |t|, |u| \ll 1$, haute énergie $s, |t|, |u| \gg 1$, limite avant $|t| \ll s$, limite arrière $|u| \ll s$) et le lien avec la théorie effective y sont également explicitement mentionnés. Les contributions dues à tous les fermions chargés (y compris celles faisant intervenir le quark top) sont incluses dans la simulation, ainsi que les boucles de bosons $W^\pm$ et l'ensemble des termes d'interférences associées. Les amplitudes à deux boucles (NLO), non-utilisées pour l'étude, sont disponibles dans l'appendice de la référence [54].

19.3 Calcul complet des amplitudes dans le cas générique de résonances de spin 0 et 2

Le calcul complet des amplitudes a également été réalisé dans le cas de nouvelles résonances neutres pour des particules scalaires, pseudo-scalaires ou de spin 2. Les résultats sont actuellement en cours d'analyse et devraient mener prochainement à une publication. Ils ne seront pas discutés dans le manuscrit.

Dans la section 20, le canal de production exclusive de paires de photons avec protons intacts au LHC est présenté. C'est l'état final qui a été retenu afin d'étudier les sensibilités des expériences ATLAS-AFP et CT-PPS aux couplages anormaux à quatre photons. Les sections efficaces attendues par le modèle standard dans les collisions proton-proton du LHC à $\sqrt{s} = 14$ TeV sont également discutées.

20 Le canal di-photon exclusif au LHC avec détection des protons intacts

20.1 Les différents modes de production

20.1.1 Présentation dans le cadre du modèle standard

Le canal di-photon exclusif au LHC correspond aux interactions proton-proton menant à la production d'une paire de photons uniquement, les protons restant intacts dans l'état final : $pp \rightarrow p + \gamma\gamma + p$. Brièvement introduite dans la section 2.4.3, une discussion plus complète de ce canal est proposée ici.

Les deux diagrammes principaux de production exclusive de paires de photons prévus dans le cadre du modèle standard sont représentés dans la figure 196. Le processus situé à gauche fait intervenir un terme de couplage à quatre photons par diffusion $\gamma\gamma$. C'est celui qui nous intéresse et sera par conséquent appelé "signal" dans cette section.

À droite, le bruit de fond irréductible initié par une interaction gluon-gluon est représenté. Il est qualifié d'irréductible car présente un état final identique au signal, ce qui permet de le différencier de celui-ci uniquement par des moyens statistiques. Comme nous le verrons, le mode de production gluon-gluon est dominant au LHC à basse énergie, mais décroît rapidement à plus haute énergie.

Le mode photon-gluon est interdit par conservation de la conjugaison de charge C tandis que l'extraction de quarks à partir du proton mène nécessairement à la destruction de ce dernier. L'analyse des autres bruits de fond, dits réductibles, sera faite dans la section 21. Quelques explications sur la description théorique du mode gluon-gluon et photon-photon sont apportées respectivement dans les sections 20.1.2 et 20.1.3.

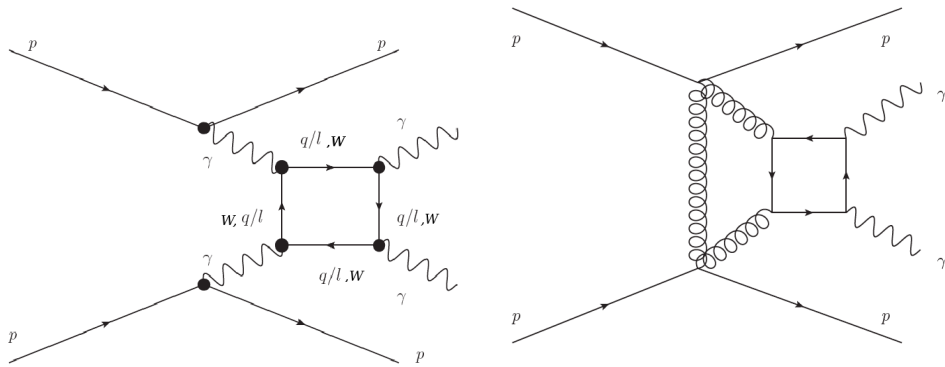


FIG. 196 – Diagrammes de production exclusive de paires de photons au LHC prédits par le modèle standard. La diffusion $\gamma\gamma$ intervient dans le diagramme de gauche, tandis qu'à droite le processus QCD exclusif "KMR" mettant en jeu un échange non-coloré de gluons est représenté. Tous deux sont donnés à l'ordre le plus bas du développement perturbatif. Les protons rentrant en collision sortent intacts de l'interaction.

20.1.2 Le mode gluon-gluon

Le mode de production gluon-gluon est décrit théoriquement par exemple dans la référence [285] et sera appelé “KMR” dans la suite, du nom des trois auteurs de l’article. Le calcul KMR est celui examiné le plus attentivement au LHC et le seul considéré dans l’étude, bien que d’autres modèles existent pour décrire la production QCD exclusive [286, 287]. Dans le calcul KMR, la section efficace est calculée à l’ordre le plus bas de la QCD perturbative¹¹⁰ en considérant l’émission d’un gluon supplémentaire entre les deux protons. Les émissions perturbatives de partons additionnels par les gluons sont supprimées par un facteur de type Sudakov, ce qui a pour effet de rendre l’interaction globale non-colorée et de garder les protons intacts. La paire de photons est finalement produite à partir de deux gluons extraits des protons par l’intermédiaire d’une boucle de quarks. Les autres modèles adoptent un point de vue plus proche de la théorie de Regge [289] en considérant ce processus comme le résultat d’un échange double de pomérons dans la limite élastique, la production exclusive étant cependant fortement amplifiée par des effets perturbatifs [286, 287].

Une modélisation minutieuse de l’événement sous-jacent (non-perturbatif, voir section 3.2) est indispensable au calcul complet de la section efficace KMR, car celui-ci retire à l’interaction son caractère exclusif dans la très grande majorité des cas par l’émission de radiations supplémentaires par les partons spectateurs ($\simeq 97\%$). En général, les effets non-perturbatifs sont encapsulés dans un facteur appelé *survival probability* et noté $\hat{S}_{gg}^2$. Ce facteur représente le taux d’événements exclusifs restant après prise en compte des effets de QCD non-perturbative par rapport au calcul de pQCD classique. Pour les processus gluon-gluon au LHC, $\hat{S}_{gg}^2$ est de l’ordre de 0.03 [290]. Cependant, il est attendu qu’il dépende du processus dur considéré (ici $gg \rightarrow \gamma\gamma$) ainsi que de la cinématique (s, t, u). Des efforts importants ont été accomplis très récemment afin d’améliorer le traitement théorique de $\hat{S}_{gg}^2$, avec la sortie du nouveau générateur SUPERCHIC2 qui l’évalue désormais de façon différentielle en fonction de l’impulsion transverse des protons sortants [291]. Les premiers résultats montrent des déformations importantes des distributions cinématiques mais une section efficace totale proche comparée à un $\hat{S}_{gg}^2$ fixe.

SUPERCHIC2 n’étant pas encore sorti au moment de l’étude réalisée pendant la thèse, la valeur fixe $\hat{S}_{gg}^2 = 0.03$ recommandée dans [292] a été adoptée dans les simulations (voir section 20.3). Le processus KMR constituant un bruit de fond négligeable dans l’étude phénoménologique à haute énergie (voir section 21), $\hat{S}_{gg}^2$ a un impact négligeable et n’est pas en position d’en modifier les conclusions de façon significative.

Dans la section suivante, le mode photon-photon est discuté plus en détail.

20.1.3 Le mode photon-photon

La diffusion $\gamma\gamma$ est prédite dans le cadre du modèle standard uniquement par l’intermédiaire de boucles pour les énergies du LHC (voir figure 196, gauche, et la section 2.4.1). Toutes les particules élémentaires chargées peuvent alors a priori circuler dans la boucle et contribuent à la section efficace totale : quarks, leptons et bosons $W^\pm$.

¹¹⁰Des facteurs correctifs sont néanmoins disponibles pour quelques processus, comme la production QCD exclusive du Higgs [288].

Chaque contribution est calculable indépendamment en utilisant les amplitudes invariantes présentées dans la section 19.2 et les termes d'interférence peuvent être obtenus lors du passage au module au carré. Cependant, le processus complet présenté dans la figure 196, gauche, correspond à un état initial constitué de deux protons et non de deux photons. Il est donc nécessaire d'étudier plus en détails l'émission de photons par les protons.

L'approximation des photons équivalents (EPA)

Comme tout objet chargé électriquement, les protons rentrant en collision à l'intérieur du LHC déforment de façon importante le champ électromagnétique dans leur environnement immédiat, ce qui a pour conséquence la création d'une multitude de photons pendant des instants très courts, dits virtuels. Ces productions sont autorisées en mécanique quantique grâce au principe d'incertitude d'Heisenberg. Il est possible de montrer que dans le cas limite où les protons sont à très haute énergie et les photons émis de façon colinéaire, les flux de photons sortants deviennent quasi-réels : c'est l'approximation des photons équivalents de Weizsäcker-Williams ou EPA (*equivalent photon approximation*) [293, 294]. Le photon est alors émis de façon cohérente par le proton, qui sort a priori intact de l'interaction. Au contraire, le flux tend très rapidement vers 0 et est beaucoup plus incertain pour des émissions non-colinéaires, les corrections non-perturbatives dues aux radiations supplémentaires émises par les partons spectateurs devenant alors beaucoup plus importantes (voir la sous-section suivante).

Dans le cadre de l'EPA, il est possible d'écrire la probabilité pour qu'un proton émette un photon portant une fraction longitudinale z de son impulsion de la façon suivante :

$$f_\gamma(z) = \frac{\alpha_{\text{em}}}{2\pi} \log \frac{\hat{s}}{m_p^2} \left[\frac{1 + (1 - z)^2}{z} \right], \quad (108)$$

où α_{em} est la constante de couplage électromagnétique, m_p est la masse du proton et $\hat{s}$ est une échelle d'énergie arbitraire correspondant à l'impulsion maximum transférée au photon. Dans les faits, les contributions pour $\hat{s} > 2 \text{ GeV}^2$ sont très petites et seront négligées dans la suite. En intégrant le produit des flux de photons venant des deux protons rentrant en collision tout en demandant la masse invariante des deux photons constante, on obtient la luminosité photon-photon élastique effective dans des collisions proton-proton en fonction de l'énergie de collision initiale et de la masse produite [74] :

$$\frac{dL_{\text{el.}}^{\gamma\gamma}(W)}{dW} = \int_{z_1, z_2, \sqrt{z_1 z_2 s} = W} f_\gamma(z_1) f_\gamma(z_2) dz_1 dz_2, \quad (109)$$

où W et $\sqrt{z_1 z_2 s}$ sont les masses invariantes du système di-photon et $\sqrt{s}$ est l'énergie de collision proton-proton dans le centre de masse. Le concept de luminosité effective est proche de celui de luminosité introduit dans la partie II et homogène à l'inverse d'une masse. Il permet d'écrire de façon simple la formule de la section efficace de production

exclusive de paires de photons par diffusion $\gamma\gamma$ au LHC :

$$\frac{d\sigma}{d\Omega} = \int \frac{d\sigma_{\gamma\gamma \rightarrow \gamma\gamma}}{d\Omega} \frac{dL_{\text{el}}^{\gamma\gamma}}{dW} dW. \quad (110)$$

La luminosité photon-photon élastique effective est représentée dans la figure 197 pour des collisions proton-proton à $\sqrt{s} = 14$ TeV, avec et sans coupure d'acceptance pour les détecteurs de protons à l'avant (voir section 6.2). La coupure d'acceptance appliquée dans la figure est moins restrictive que celle prévue actuellement pour le LHC (0.0015 au lieu de 0.015) et donnée à titre illustratif uniquement.¹¹¹ Dans les deux cas, la distribution décroît relativement faiblement avec la masse produite. De plus, les deux distributions restent du même ordre de grandeur à haute masse ($400 \text{ GeV} < W < 1 \text{ TeV}$), point sur lequel nous reviendrons dans la section 20.2.1.

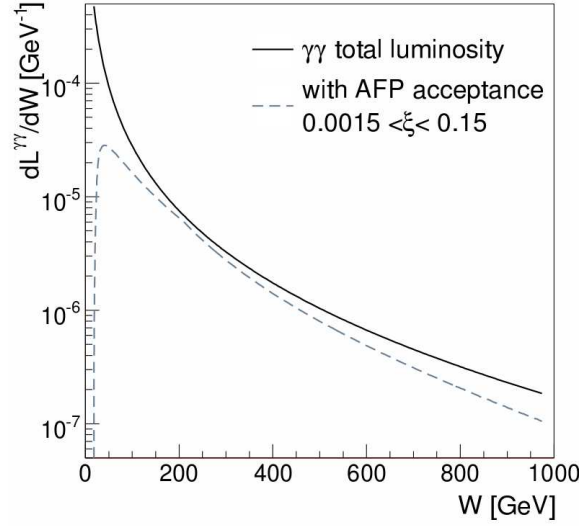


FIG. 197 – Luminosité photon-photon élastique effective (voir équation 109) pour des collisions proton-proton à $\sqrt{s} = 14$ TeV en fonction de la masse invariante produite. Celle-ci est représentée en prenant en compte (trait pointillé) ou non (trait plein) de l'acceptance des détecteurs de protons à l'avant (voir section 6.2). La coupure d'acceptance appliquée dans la figure est moins restrictive que celle prévue actuellement pour le LHC (0.0015 au lieu de 0.015) et donnée à titre illustratif uniquement [74].

Corrections non-perturbatives

Des corrections non-perturbatives similaires à celles présentées dans la section 20.1.2 pour le mode de production gluon-gluon doivent néanmoins être considérées afin de tenir compte d'effets dus à la taille du proton. Ces corrections incluent un veto sur les radiations supplémentaires pouvant être émises par les partons spectateurs, à l'image du mode de production gluon-gluon. Le formalisme utilisé pour le calcul de $\hat{S}^2$ est en grande partie commun entre le cas gluon-gluon et photon-photon, bien que $\hat{S}_{\gamma\gamma}^2$ soit mieux connu

¹¹¹Des détecteurs supplémentaires à 420 m du point d'interaction principal avaient été envisagés dans le projet de départ d'ATLAS-AFP, ce qui aurait permis d'étendre l'acceptance à $\xi > 0.0015$. Ceux-ci ne sont désormais plus d'actualité pour des raisons techniques liées à la machine du LHC.

en général. La *survival probability* attendue dans le cas photon-photon est de l'ordre de $\hat{S}_{\gamma\gamma}^2 \simeq 0.9$, soit bien plus élevée que pour le processus QCD ($\hat{S}_{gg}^2 \simeq 0.03$). Au premier ordre, cela rend les calculs photon-photon moins dépendants des effets non-perturbatifs par rapport aux processus gluon-gluon et donc plus robustes. Des études récentes ont amélioré l'estimation de $\hat{S}_{\gamma\gamma}^2$ [291,295]. Elles estiment désormais $\hat{S}_{\gamma\gamma}^2$ comprise entre 0.9 et 0.6 environ, avec une dépendance plus ou moins forte en fonction de la masse invariante et de la pseudo-rapacité des deux photons ($\hat{S}_{\gamma\gamma}^2$ plus basse à haut $m_{\gamma\gamma}$ et $|y_{\gamma\gamma}|$).

Pour tous les processus utilisant l'EPA considérés au cours de la thèse, une valeur fixe $\hat{S}_{\gamma\gamma}^2 = 0.9$ est adoptée. Un facteur 2 sur cette valeur ne change pas significativement les conclusions de l'étude, comme nous le verrons dans la section 22. Néanmoins, il sera important de mesurer au préalable la production d'un processus photon-photon bien connu au LHC à $\sqrt{s} = 14$ TeV afin de réduire l'incertitude sur $\hat{S}_{\gamma\gamma}^2$.

Ce point est discuté entre autres dans la section 20.2, qui aborde plus généralement les aspects expérimentaux concernant la détection de paires de photons exclusives avec protons intacts.

20.2 Mise en évidence expérimentale

20.2.1 Utilisation de détecteurs de protons à l'avant du faisceau

D'un façon générale, tous les processus gardant les protons intacts ne dévient ceux-ci que très légèrement de leur trajectoire initiale (EPA/photon-photon, KMR/gluon-gluon, etc ...). Ils constituent des processus dits à bas t , en référence à la variable de Mandelstam, et ne perdent qu'une fraction limitée de leur impulsion ($\xi < 0.2$). Ils peuvent ainsi être étudiés dans le cadre de la physique à l'avant (voir section 6.1).

Cela se comprend facilement intuitivement, néanmoins il est possible de le démontrer rapidement dans le cas photon-photon. En effet, comme nous l'avons vu dans la section 20.1.3, l'émission cohérente d'un photon par un proton est fortement amplifiée dans le cas quasi-colinéaire, cadre de l'EPA. L'impulsion transverse du proton après interaction est alors de l'ordre de quelques centaines de MeV seulement. Ainsi, les protons intacts ne sont que très légèrement déviés par rapport à leur trajectoire initiale et restent dans le tube faisceau après l'interaction. Le processus KMR présente les mêmes caractéristiques sur ce point à cause des effets non-perturbatifs, dominants à plus haut t et ξ dans les deux cas de figure.

Le système d'aimants très précis du LHC situé le long du tube faisceau peut alors être utilisé comme spectromètre et les protons intacts légèrement déviés, après avoir été éloignés du faisceau principal par les champs magnétiques des aimants, peuvent être détectés dans des stations dédiées situées à environ 200 mètres du point d'interaction principal et à quelques millimètres du faisceau sortant (voir section 6.2). Il est courant de définir l'acceptance des détecteurs de protons à l'avant en terme de fraction d'impulsion perdue par le proton, noté ξ . L'acceptance typique d'un détecteur à l'avant d'ATLAS-AFP et CT-PPS lors de prises de données nominales au LHC est de $0.015 < \xi < 0.15$ et non contraignante pour les impulsions transverses des protons considérées (voir section 6.2.2). Dans le cas d'une détection de deux protons en coïncidence, cela équivaut à une masse invariante créée dans le détecteur central $\sqrt{\xi_1 \xi_2 s}$ comprise entre 200 GeV

et 2 TeV environ (voir figure 43, section 6.2.2), d'où la nécessité de produire une masse centrale importante afin de tomber dans l'acceptance des stations à l'avant.¹¹² La faible déformation de la luminosité photon-photon élastique effective par l'acceptance des stations à l'avant pour le domaine $400 \text{ GeV} < W < 1 \text{ TeV}$ observée dans la figure 197 est donc cohérente avec ce qui est attendu.

Les événements exclusifs avec production d'un système massif dans le détecteur central ($s \gg t$, soit $W \gg 1 \text{ GeV}$) sont appelés centraux exclusifs (CEP, pour *central exclusive production*). Les processus considérés dans la section 20.1.1 en sont des exemples typiques. Une revue récente concernant les productions CEP est disponible dans le chapitre 5 de la référence [132].¹¹³

Les études expérimentales traitant de la production exclusive de paires de photons et des processus CEP en général peuvent être menées de plusieurs façons différentes, par exemple avec ou sans détection de protons intacts à l'avant. Dans ce dernier cas, le signal inclut alors également des processus avec protons dissociés. Dans tous les cas, ces études doivent être adaptées aux sections efficaces et à la distribution en masse attendue pour le signal. Dans la section 20.2.2, les différentes stratégies possibles sont abordées plus en détail.

20.2.2 Stratégies adoptées à haute et basse énergie

Recherche de nouvelle physique et mesures à haute masse ($m_{\gamma\gamma} > 200 \text{ GeV}$)

Dans le cadre d'une étude à haute énergie comme celle présentée dans la section 21 ($m_{\gamma\gamma} > 200 \text{ GeV}$), le signal attendu est constitué par la détection en coïncidence d'au moins deux protons intacts dans les stations à l'avant et d'au moins une paire de photons de haute énergie dans le détecteur central. En effet, le fort taux d'empilement attendu dans les collisions nominales du LHC à $\sqrt{s} = 14 \text{ TeV}$ ($\langle \mu \rangle \simeq 50$) et les faibles sections efficaces rendent délicate la mise en œuvre de critères d'exclusivité, qui consisteraient à considérer seulement les événements ayant une paire de photons reconstruite dans le détecteur central et rien d'autre. Dans ce contexte, la détection des protons intacts permet de garantir la sélection d'un échantillon pur en événements exclusifs (voir section 21). Par ailleurs, le déclenchement peut se baser uniquement sur la présence d'une paire de photons dans le détecteur central sans saturer l'électronique d'acquisition étant donné les faibles taux attendus pour $m_{\gamma\gamma} > 200 \text{ GeV}$. Les premières mesures suivant cette stratégie, adaptée entre autres à la recherche de nouveaux phénomènes, devraient débuter lors de la phase-2 des projets ATLAS-AFP et CT-PPS (voir section 6.2).

Mesures à plus basse masse ($m_{\gamma\gamma} < 200 \text{ GeV}$)

À plus basse masse et dans le cas de collisions nominales au LHC, les stations à l'avant n'ont pas l'acceptance requise pour pouvoir détecter les protons intacts (voir

¹¹²Il est rappelé que les acceptances citées ici sont valables uniquement pour des optiques configurées pour produire des collisions nominales au LHC. Elles peuvent être largement modifiées lors de *runs* dédiés (voir section 20.2.2).

¹¹³J'ai assumé avec Valery Khoze et Lucian Harland-Lang la responsabilité de la constitution et de l'édition finale du chapitre 5 de ce document, comme cela a déjà été mentionné dans la section 6.1.

section 6.2.2). Cependant, l'augmentation des sections efficaces de signal a permis de mesurer certains processus à $\sqrt{s} = 7$ TeV grâce à des critères d'exclusivité sans détection des protons intacts, comme par exemple la section efficace de production exclusive EPA de di-lepton, mesurée par ATLAS et CMS [274, 296, 297] (voir section 20.2.3). Néanmoins, ces mesures souffrent d'erreurs systématiques importantes, notamment liées au contrôle du bruit de fond et aux taux de protons dissociés, et sont très délicates à mettre en œuvre. Par ailleurs, leur faisabilité avec un taux d'empilement plus élevé (données à $\sqrt{s} = 8$ et 13/14 TeV) reste à établir.

Des approches alternatives reposant sur des prises de données dédiées avec un taux d'empilement beaucoup plus bas ($\langle \mu \rangle \simeq 1$ ou $\langle \mu \rangle \simeq 0.01$) peuvent alors être adoptées, à l'exemple de celle qui sera suggérée dans la section 20.3 pour réaliser la mesure de la production KMR standard des di-photons exclusifs :

- ◇ **Si le taux d'empilement est assez bas ($\langle \mu \rangle \simeq 0.01$)**, il est possible de se passer totalement des stations à l'avant et de reposer en intégralité sur des critères d'exclusivité dans le détecteur central pour la sélection du signal, à l'image de certaines mesures réalisées par LHCb et ALICE pendant le *run 1* (voir section 20.2.3). Ces critères peuvent également être intégrés au niveau du système de déclenchement. Par exemple, considérer uniquement les événements présentant des grands intervalles en rapidité vides de toute radiation supérieure à un certain seuil est un critère présentant un bon taux de rejet et une très bonne efficacité à $\langle \mu \rangle \simeq 0.01$. Les événements avec protons dissociés doivent dans ce cas être pris en compte pour décrire les données.
- ◇ **Le changement de configuration des optiques du LHC lors des *runs* dédiés modifie néanmoins les cartes d'acceptance des détecteurs à l'avant** et permet la détection des plus petites masses dans les stations verticales, *ie.* placées selon y ($0 < m_{\gamma\gamma} < 2$ TeV). Les stations à l'avant peuvent alors être utilisées dans les analyses et permettre des mesures dans des environnements légèrement plus dense ($\langle \mu \rangle \simeq 1$), tandis que l'efficacité des critères d'exclusivité décroît en général très fortement avec l'empilement. Les détecteurs à l'avant peuvent également être intégrés dans le système de déclenchement si les taux sont suffisamment bas.

Dans les deux cas de figure, le bruit de fond attendu est généralement assez faible, particulièrement lors de l'utilisation des détecteurs à l'avant. Le facteur limitant principal est alors la faible luminosité intégrée, qui est prévue de se situer respectivement aux alentours de 0.01 et 0.1 fb⁻¹ pour $\langle \mu \rangle \simeq 0.01$ et $\langle \mu \rangle \simeq 1$, à comparer à 300 fb⁻¹ dans le cas des *runs* nominaux (voir section 4.4). En pratique, les *runs* dédiés sont donc exploitables uniquement pour l'étude de processus ayant des sections efficaces significatives à basse masse. Ils apparaissent ainsi comme très complémentaires à l'exploitation des *runs* nominaux (haute masse et accès aux plus faibles sections efficaces). Des détails supplémentaires sur l'acceptance des détecteurs à l'avant pour les différentes configurations des optiques du LHC et sur les *runs* dédiés prévus sont disponibles dans le chapitre 1 de la référence [132].

Dans la section 20.2.3, les mesures de di-photons exclusifs déjà effectuées dans les collisionneurs hadroniques sont rapportées. De façon plus générale, un bref état des lieux

est dressé concernant les différentes mesures de processus CEP, incluant EPA et KMR.

20.2.3 État des lieux des différentes mesures

Mode de production gluon-gluon (KMR)

L'observation de 43 paires de photons exclusives à basse masse ($m_{\gamma\gamma} < 25$ GeV) a été rapportée récemment par la collaboration CDF au TEVATRON dans des collisions proton-antiproton à $\sqrt{s} = 1.96$ TeV, en bon accord avec les prédictions du calcul KMR qui dominent totalement l'espace fiduciel de mesure [273]. Elle représente aujourd'hui l'unique observation directe de ce canal. Une tentative de mesure similaire à $\sqrt{s} = 7$ TeV a été rapportée par la collaboration CMS au LHC en 2012 [274], mais n'a pas mené à une observation significative d'événements pour des raisons statistiques, toujours en bon accord avec les prédictions KMR. LHCb et ALICE envisagent de réaliser une mesure de ce processus pendant la seconde prise de données du LHC, tandis que sa faisabilité pour ATLAS et CMS à $\sqrt{s} = 14$ TeV est étudiée dans la section 20.3.

La mise en évidence de productions KMR a également été faite dans le canal di-jet et $\pi^+\pi^-$ par CDF [298, 299]. D'autres canaux KMR avec des états finals de mésons ont été mesurés par la collaboration ALICE et LHCb au LHC, leur revue complète étant disponible dans le chapitre 5 de la référence [132]. Des mesures similaires basées sur la combinaison des données de CMS et TOTEM prises lors du *run 1* sont attendues et actuellement en cours.

Mode de production photon-photon (EPA)

Le mode de production photon-photon n'a encore jamais été mis en évidence pour les di-photon exclusifs, que ce soit à haute ou à plus basse énergie (voir section 2.4). Il a cependant pu être mesuré dans les états finals di-leptons ($\gamma\gamma \rightarrow ee, \mu\mu$) par les collaborations ATLAS et CMS à $\sqrt{s} = 7$ TeV pendant la première prise de données du LHC [274, 296, 297], car ils constituent des signaux propres pour le détecteur et comportent une section efficace beaucoup plus élevée que $\gamma\gamma \rightarrow \gamma\gamma$. Les premières contraintes sur les facteurs de suppression $S_{\gamma\gamma}^2$ au LHC ont ainsi pu être établies. Des mesures similaires sont prévues à $\sqrt{s} = 13$ TeV afin de contraindre davantage $S_{\gamma\gamma}^2$ dans les données du *run 2*. La mesure de ce mode de production dans les collisions proton-proton du LHC à $\sqrt{s} = 14$ TeV pour les di-photon exclusifs est discutée dans la section 20.3.

Par ailleurs, comme cela a été discuté dans la section 2.4 ainsi que tout au long de cette partie, la recherche de tels événements à haute énergie, possible pendant le *run 2* grâce aux détecteurs d'ATLAS-AFP et de CT-PPS, présente un excellent potentiel pour la découverte de nouvelle physique. L'étude des événements exclusifs $\gamma\gamma \rightarrow WW$ par CMS à $\sqrt{s} = 7$ TeV [271] offre dans ce contexte des perspectives encourageantes. En effet, bien que la section efficace fiducielle attendue soit trop faible pour effectuer une mesure de la production $\gamma\gamma \rightarrow WW$ standard, cette étude a permis d'établir les meilleurs limites actuelles sur les couplages anormaux quartiques $WW\gamma\gamma$.

Dans la section 20.3, les prédictions du modèle standard concernant la production exclusive de paires de photons au LHC sont présentées dans le cadre de collisions proton-

proton à $\sqrt{s} = 14$ TeV. La faisabilité de la mesure d'un tel état final au LHC ainsi que sa capacité à en différencier les modes de production sont ensuite discutées.

20.3 Les prédictions du modèle standard

20.3.1 Présentation

La théorie effective et les amplitudes complètes présentées dans la section 19 pour le calcul des termes de couplages à quatre photons ont été implémentées dans le générateur de Monte Carlo FPMC [300] pendant les années de thèse. Celui-ci utilise l'EPA introduit dans la section 20.1.3 pour la production élastique des deux photons initiaux à partir de protons et considère un facteur de suppression non-perturbatif constant $S_{\gamma\gamma}^2 = 0.9$, ce qui est raisonnablement optimiste comparé aux dernières estimations disponibles [291, 295]. Il est alors possible pour la première fois de simuler l'intégralité des événements di-photons exclusifs initiés par photon-photon au LHC avec le même outil. Pour la génération du mode gluon-gluon (voir section 20.1.2), le code EXHUME MC [292] basé sur les calculs KMR [285] est employé. Celui-ci utilise un facteur de suppression non-perturbatif constant $S_{gg}^2 = 0.03$ à 14 TeV. Il est désormais possible avec le générateur SUPERCHIC2, récemment rendu public, de calculer $S_{\gamma\gamma}^2$ et S_{gg}^2 de façon différentielle événement par événement [291].

Les sections efficaces intégrées prédites par le modèle standard pour les différentes composantes du canal di-photon exclusif au LHC à $\sqrt{s} = 14$ TeV sont présentées dans la figure 198 en fonction de la coupure sur la masse du système di-photon final. Des valeurs numériques sont également rapportées dans le tableau 38 pour une comparaison plus directe. Elles sont en bon accord avec des estimations similaires ayant été publiées depuis ou en parallèle [59, 291]. Les deux photons présentant une topologie “dos-à-dos” typique des productions exclusives avec deux particules centrales dans l'état final, une coupure C sur la masse invariante implique de fait une coupure $C/2$ sur les énergies transverses individuelles des deux photons.¹¹⁴ Dans le cadre d'une véritable analyse de données, des coupures additionnelles devraient être appliquées, comme l'acceptance géométrique exacte du calorimètre $|\eta| < 2.37$ (voir section 5.4). Ces détails sont négligés ici, où seul l'ordre de grandeur des sections efficaces intégrées est considéré.

Selon la figure 198, le processus KMR/gluon-gluon domine les basses masses, confirmant les propos déjà avancés au cours des sections précédentes. À partir d'une coupure correspondant à $m_{\gamma\gamma} > 90$ GeV environ, le processus EPA/photon-photon devient majoritaire. La production QED est quant à elle dominée par les boucles de fermions jusqu'à une coupure correspondant à $m_{\gamma\gamma} > 150$ GeV environ (soit juste inférieure à $2 m_W$), pour ensuite laisser place aux boucles de $W^\pm$, qui dominent les plus hautes masses.

Ainsi, la section efficace prédite par le modèle standard à haute masse ($m_{\gamma\gamma} > 200$ GeV) est totalement dominée par la contribution QED s'effectuant via des boucles de $W^\pm$. Cela est attendu à cause de la masse importante de ces bosons ($m_W \simeq 80$ GeV). Elle est de l'ordre de $\sigma = 0.01$ fb, ce qui est très faible par rapport à la

¹¹⁴Cette affirmation est strictement vraie uniquement dans le cas où l'impulsion transverse des protons est totalement nulle dans l'état final ($p_T^{pp} = p_T^{\gamma\gamma} = 0$), en considérant la relation suivante : $m_{\gamma\gamma}^2 = 2E_{T,1}^\gamma E_{T,2}^\gamma [\cosh(\eta_1^\gamma - \eta_2^\gamma) - \cos(\phi_1 - \phi_2)]$. Nous nous contenterons de cette simplification dans cette section, où seul l'ordre de grandeur des différentes composantes importe.

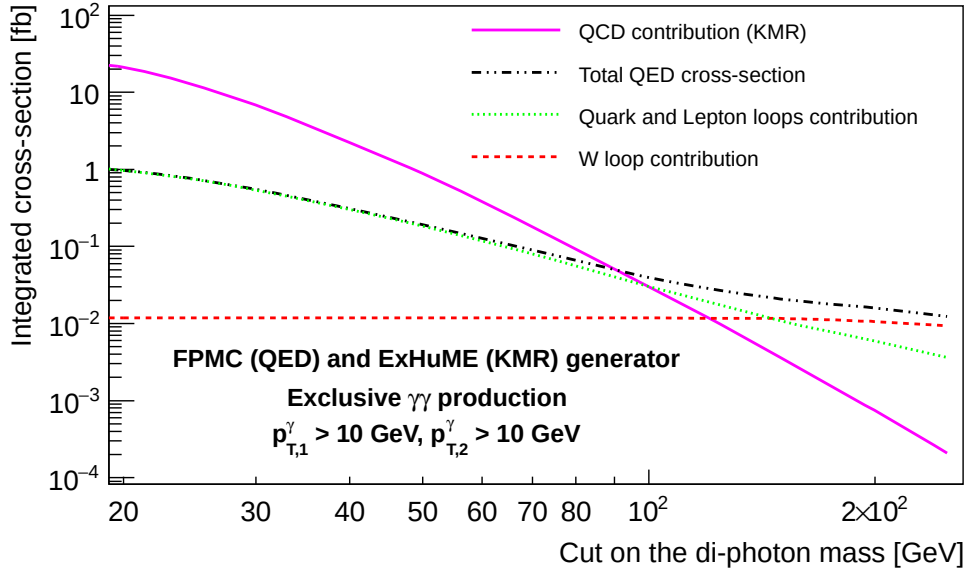


FIG. 198 – Sections efficaces intégrées des deux composantes du modèle standard participant à la production exclusive de paires de photons au LHC à $\sqrt{s} = 14$ TeV : QCD exclusive (KMR/gluon-gluon, trait plein rose) et QED exclusive (EPA/photon-photon, trait pointillé noir). La composante QED est elle-même décomposée en contributions impliquant des boucles de quarks et de leptons (trait pointillé vert) et des boucles de bosons $W^\pm$ (trait pointillé rouge). Différentes coupures sur la masse invariante de la paire de photons présente dans l'état final sont considérées. Dans tous les cas, les photons dans l'état final ont une impulsion transverse supérieure à 10 GeV.

Processus : Coupures	Exclusif QCD (KMR)	Exclusif QED (boucles de fermions)	Exclusif QED (boucles de $W^\pm$)
$m_{\gamma\gamma} > 10$ GeV, ($E_{T1,2} > 5$ GeV)	372.1 fb	5.5 fb	0.012 fb
$m_{\gamma\gamma} > 20$ GeV, ($E_{T1,2} > 10$ GeV)	20.4 fb	1.0 fb	0.012 fb
$m_{\gamma\gamma} > 50$ GeV, ($E_{T1,2} > 25$ GeV)	0.87 fb	0.18 fb	0.012 fb
$m_{\gamma\gamma} > 100$ GeV, ($E_{T1,2} > 50$ GeV)	0.030 fb	0.03 fb	0.012 fb
$m_{\gamma\gamma} > 200$ GeV, ($E_{T1,2} > 100$ GeV)	$7.4 \cdot 10^{-4}$ fb	$5.0 \cdot 10^{-3}$ fb	0.010 fb
$m_{\gamma\gamma} > 500$ GeV, ($E_{T1,2} > 250$ GeV)	$3.2 \cdot 10^{-6}$ fb	$3.0 \cdot 10^{-4}$ fb	0.004 fb

TAB. 38 – Sections efficaces intégrées des différents processus de production exclusive de paires de photons prévues dans le cadre du modèle standard au LHC à $\sqrt{s} = 14$ TeV en fonction de la coupure opérée sur leur masse invariante, $m_{\gamma\gamma}$.

statistique prévue lors des prises de données nominales du LHC à 14 TeV ($\mathcal{L} \simeq 300 \text{ fb}^{-1}$), le nombre d'événements attendu étant alors $\sigma \times \mathcal{L} \simeq 3$. Cela confirme la forte suppression de la production du modèle standard à haute masse, déjà annoncée dans la section 2.4.2. La production MS constituant un bruit de fond irréductible dans le cadre de la recherche de nouvelle physique à haute énergie, sa forte suppression augmente significativement la sensibilité du canal di-photon exclusif à de nouveaux phénomènes. Comme nous le verrons, l'analyse à haute énergie présentée dans la section 21 utilise une sélection qui permettra même d'exclure l'intégralité des bruits de fond (réductibles + irréductibles), accordant ainsi à toute observation une significativité statistique très haute concernant l'observation de nouveaux phénomènes.

20.3.2 Discussion d'une mesure au LHC

Mesurer la production de di-photons exclusifs telle que prédite par le modèle standard nécessiterait donc de descendre au maximum le seuil en masse invariante de l'analyse afin d'augmenter la section efficace de production. Cependant, lors des collisions nominales au LHC, les contraintes en terme d'empilement et de débit de données en sortie limitent les seuils de déclenchement des expériences à des niveaux élevés. Par exemple, le seuil minimum du *trigger* di-photon d'ATLAS pour les collisions nominales du *run 2* se situe à 35 et 25 GeV pour l'impulsion transverse du premier et deuxième photons respectivement, ce qui donne dans le meilleur des cas une section efficace exclusive de 1 fb environ ($m_{\gamma\gamma} > 50 \text{ GeV}$). Ce seuil en masse étant en-dessous de l'acceptance des détecteurs à l'avant dans le cas de collisions nominales ($m_{\gamma\gamma} > 200 \text{ GeV}$), les rares événements exclusifs seront donc totalement noyés dans le bruit de fond, constitué principalement de la production QCD non-exclusive (sans protons intacts) de paires de photons (voir partie IV). La mise en place de critères d'exclusivité (voir section 20.2.2) paraît également très délicate à cause du niveau d'empilement ($\langle \mu \rangle \simeq 50$) et du fait que les photons ne laissent le plus souvent aucune trace dans le détecteur interne. Par conséquent, on s'attend à ce que ni la composante gluon-gluon/KMR ni la composante photon-photon/EPA ne puissent être observées lors des prises de données nominales à $\sqrt{s} = 14 \text{ TeV}$ au LHC.

Néanmoins, une mesure de la composante gluon-gluon/KMR ne semble pas impossible lors de prises de données dédiées à basse luminosité ($\langle \mu \rangle \simeq 1$) avec reconfiguration des optiques, où il est prévu de collecter un échantillon de 0.1 fb^{-1} [132]. En effet, sous réserve de pouvoir implémenter un menu de déclenchement di-photon particulier avec des seuils abaissés à au moins $E_T > 5 \text{ GeV}$, environ une trentaine d'événements di-photon exclusifs KMR sont attendus (voir tableau 38). Par ailleurs, grâce à la reconfiguration des optiques du LHC par rapport au cas nominal, les détecteurs à l'avant verticaux présentent l'acceptance nécessaire à la détection des protons intacts de plus basse masse (voir section 20.2.2). Ils peuvent alors être intégrés au système de déclenchement afin d'aider à l'abaissement des seuils et au rejet du bruit de fond. Étant donné le faible taux d'empilement, il est également possible d'imaginer des critères supplémentaires sur la topologies des événements centraux (dos-à-dos) ou sur leur exclusivité (pas de radiation dans des intervalles importants en rapidité) pour abaisser davantage les seuils.

Le mode de production photon-photon paraît cependant définitivement hors de portée d'une mesure dans le cas de collisions proton-proton, à cause de sa section efficace

particulièrement faible. Néanmoins, comme récemment suggéré dans la littérature, il pourrait être observé à partir des collisions Plomb-Plomb au LHC [59]. En effet, celles-ci bénéficient d'une luminosité élastique effective photon-photon (voir section 20.1.3) bien plus importante que dans le cas des protons grâce au numéro atomique élevé du Plomb ($Z = 82$), qui intervient à la puissance 4 dans le calcul du flux ($Z^4 = 5 \cdot 10^7$). Cela permet de compenser largement la faible luminosité intégrée collectée lors des collisions d'ions lourds ($\mathcal{L} \simeq \text{nb}^{-1}$).

Dans cette section, le canal considéré pour l'étude phénoménologique à haute énergie a été introduit et des premiers ordres de grandeur sur les sections efficaces associées à ses différents modes de production prédits dans le cadre du modèle standard ont été rapportés. Dans les deux sections suivantes, l'étude à haute énergie et les résultats associés sont présentés. Dans la section 21, la composition et la simulation des différents bruits de fond (réductibles) à haute énergie, la modélisation des effets de détecteurs et la sélection effectuée dans les données simulées sont discutés. Finalement, les résultats finals de sensibilité à des signaux génériques de physiques au-delà du modèle standard sont dévoilés dans la section 22.

21 Étude des sensibilités aux couplages anormaux à quatre photons à haute énergie

21.1 Scénarios considérés

Dans le cadre de l'étude des couplages anormaux à quatre photons à haute énergie au LHC, on suppose disponible l'ensemble de la statistique prévue d'être collectée dans les collisions proton-proton nominales de la phase-1 à l'énergie nominale de $\sqrt{s} = 14$ TeV (2015-2023). Cela correspond à environ 300 fb^{-1} de données (voir section 4.4), qui est pris comme chiffre de référence. Par ailleurs, on considère une moyenne de 50 interactions additionnelles à chaque croisement de faisceaux en plus du processus de haute énergie, ce qui se situe dans la fourchette haute de l'empilement attendu.

Dans le cadre de la théorie effective, une extrapolation des sensibilités pour le LHC de haute luminosité (HL-LHC) est également effectuée. 3000 fb^{-1} de données proton-proton à $\sqrt{s} = 14$ TeV devraient être collectées entre 2024 et 2034 avec cette machine (voir section 4.4). Le taux d'empilement prévu lors ces collisions est bien plus haut et la valeur conservatrice $\langle \mu \rangle = 200$ est adoptée pour l'extrapolation. Il est cependant aujourd'hui difficile de savoir si les stations à l'avant seront en mesure d'être opérées à de tels niveaux d'empilement.

Par ailleurs, deux scénarios concernant les critères de sélection sur le statut de conversion des photons (voir partie III) sont considérés : au moins l'un des deux photons convertis, ou aucune sélection sur le statut de conversion. La présence d'au moins un photon converti permet une identification plus précise du vertex d'interaction ($\Delta z < 1 \text{ mm}$), grâce à l'utilisation des traces associées dans le trajectographe du détecteur central.¹¹⁵ Cela peut s'avérer utile pour vérifier la compatibilité du vertex central avec les mesures de temps de vol issues des détecteurs de protons à l'avant et ainsi rejeter le bruit de fond venant des protons d'empilement (voir section 6.3.2). Néanmoins, l'utilisation des temps de vol des protons ne semble pas nécessaire pour l'étude au regard des prévisions de bruits de fond (voir section 21.4), d'excellentes sensibilités aux modèles de nouvelle physique étant obtenues sans critères particuliers (voir section 22). Le bruit de fond relié à la machine (*beam-induced*) n'est cependant pas pris en compte dans l'étude car largement inconnu pour les prochains *runs* du LHC. Bien qu'a priori fortement réduit par la sélection principale (voir section 21.4), les mesures de temps de vol permettraient néanmoins un taux de rejet supplémentaire en cas de besoin. Par conséquent, les deux cas de figure sont considérés dans le cadre de la théorie effective. En revanche, seul le scénario sans sélection particulière sur le statut de conversion des photons sera cité pour les sensibilités obtenues à partir des amplitudes complètes.

Le canal étudié est la production exclusive de paires de photons avec protons intacts à l'avant. Les caractéristiques du signal ont déjà été énoncées dans la section 20.2.2 : détection en coïncidence d'au moins deux protons intacts dans les stations à l'avant et d'au moins une paire de photons de haute énergie dans le détecteur central. L'étude étant axée sur la recherche de nouvelle physique, les productions prévues par le modèle

¹¹⁵Cela est particulièrement vrai pour CMS, dont le calorimètre électromagnétique ne présente pas de segmentation longitudinale et ne peut donc pas implémenter la méthode du *pointing* [190] utilisée par ATLAS ($\Delta z \simeq 20 \text{ mm}$).

standard à partir du mode photon-photon et gluon-gluon sont considérées ici comme des bruits de fond irréductibles. Des détails sur la méthode adoptée pour estimer la sensibilité des données à un signal de nouvelle physique seront donnés dans la section 22.1.

Dans la section 21.2, l'analyse des différents bruits de fond attendus est effectuée et des détails sur leur simulation sont fournis. La section 21.3 traite de la modélisation des effets de détecteurs, étape indispensable afin d'obtenir des estimations réalistes sur les sensibilités. Ceux-ci incluent les taux de conversion des photons, les effets de résolution sur les observables, les probabilités de mauvaise identification de jets et d'électrons en photons ainsi que les efficacités de présélection des photons véritables (reconstruction, identification, isolation). Finalement, la section 21.4 détaille la sélection finale adoptée dans les pseudo-données afin de séparer le signal du bruit de fond.

21.2 Composition et simulation du bruit de fond

21.2.1 Les différents types de bruits de fond

Le bruit de fond représente l'ensemble des processus pouvant imiter les caractéristiques du signal. Il est possible de le décomposer en trois types principaux :

- ◇ **Les bruits de fond irréductibles** : constitués des deux processus de production exclusive de paires de photons prévus dans le cadre du modèle standard (voir section 20.3), ils comportent exactement les mêmes caractéristiques que le signal. Par conséquent, ils ne peuvent en être séparés que par des moyens statistiques.
- ◇ **Les bruits de fond réductibles avec protons intacts à l'avant** : constitués de processus gardant les deux protons intacts mais présentant un état final central différent. Ils peuvent imiter les caractéristiques du signal lors d'une mauvaise identification de l'état final central par le détecteur (jets ou électrons identifiés comme photons) et/ou lorsqu'ils présentent un état final central non exclusif ($\gamma\gamma + X$, ...).
- ◇ **Les bruits de fond réductibles avec un ou plusieurs protons intacts d'empilement à l'avant (dominants)** : constitués de processus inclusifs classiques avec destruction des protons à l'issue de l'interaction, ils peuvent être identifiés comme événements de signal lors de la détection en coïncidence de deux protons intacts issus d'interactions d'empilement indépendantes. Il existe également des situations hybrides où seul l'un des deux protons de l'interaction principale reste intact et où le second vient d'une interaction d'empilement.

Les deux bruits de fond irréductibles sont composés des processus KMR/gluon-gluon et EPA/photon-photon prédits par le modèle standard. Leur description ainsi que leur simulation ont déjà été traitées en détail dans la section 20. Bien que l'on s'attende à ce qu'ils soient totalement négligeables à haute énergie, il sont intégrés à la simulation finale de bruit de fond à des fins illustratives. On rappelle ici que le mode de production photon-gluon est interdit dans le modèle standard par conservation de la conjugaison de charge C .

Dans les sections 21.2.2 et 21.2.3, les deux catégories de bruits de fond réductibles sont analysées en détail et la simulation finale de bruit de fond utilisée dans l'étude est explicitée et justifiée.

21.2.2 Bruits de fond avec protons intacts à l'avant

Productions exclusives avec mauvaise identification du système central

Les premiers bruits de fond réductibles avec protons intacts à l'avant venant à l'esprit sont probablement ceux formés de productions exclusives présentant des états finals différents des di-photons, mais pouvant être reconstruits comme tels dans le détecteur. Cependant, on devine déjà que ceux-ci vont très vite tendre vers zéro à haute masse, comme nous l'avons déjà constaté dans le cas des bruits de fond irréductibles dans la section 20.3.

Par conséquent, seule la production exclusive de di-électron est considérée dans l'étude en dehors du bruit de fond irréductible, à des fins illustratives. Son diagramme principal, issu du mode de production photon-photon, est représenté dans la figure 199 et simulé avec le générateur FPMC en utilisant la même EPA et le même facteur de suppression que dans le cas des di-photons exclusifs ($S_{\gamma\gamma}^2 = 0.9$). Les productions exclusives KMR di-jet/photon+jet et EPA di-jet avec mauvaise identification des jets en photons sont supposées inférieures aux bruits de fond irréductibles et négligées.

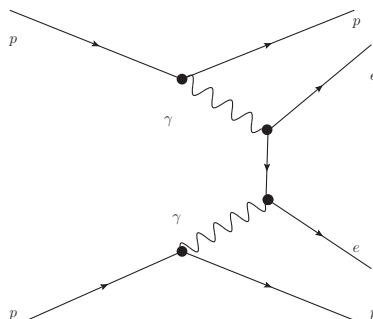


FIG. 199 – Diagrammes de production exclusive de paires d'électrons à partir de deux photons, à l'ordre le plus bas de la théorie des perturbations.

Productions diffractives inclusives

Par ailleurs, un certain nombre de processus diffractifs traditionnellement étudiés en physique à l'avant (voir section 6.1), comme l'échange double de pomérons (DPE pour *double pomeron exchange*), sont attendus. Ceux-ci ont été mesurés par exemple au TEVATRON [301] et mettent en jeu l'échange d'un objet hadronique non coloré appelé poméron, décrit d'abord dans la cadre de la théorie de Regge [289]. Son caractère non coloré permet de garder les deux protons intacts dans le cas d'un échange double, au prix d'un facteur de réduction non perturbatif S_{PP}^2 similaire à ceux déjà vus dans la section 20. Dans toutes les simulations incluant un échange double de pomérons, $S_{\text{PP}}^2 = S_{gg}^2 = 0.03$ est considéré. Toutes deux souffrent d'incertitudes significatives déjà discutées dans la section 20.1.2.

Les diagrammes de production DPE de paires de photons et de jets hadroniques sont représentés dans la figure 200. Ils constituent des bruits de fond réductibles dans le sens où les pomérons se désintègrent au cours de l'interaction en émettant des radiations supplémentaires à basse énergie, et donc brisent le caractère exclusif de l'événement. La production de di-jets est considérée dans le cas où les deux jets sont reconstruits comme deux photons dans le détecteur central.

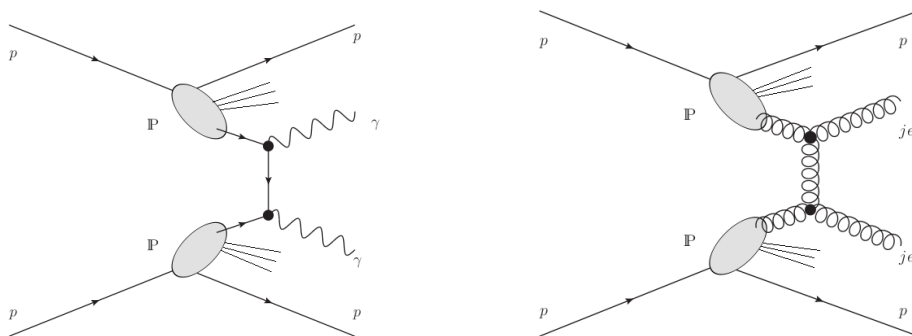


FIG. 200 – Diagrammes de production de paires de photon (gauche) et de paires de jets hadroniques issues de deux gluons (droite) via un échange double de pomérons à l'ordre le plus bas de la théorie des perturbations.

Il est possible de décrire le poméron par des techniques de calcul perturbatif lorsque l'énergie mise en jeu lors de l'interaction devient assez haute ($\gg 1$ GeV), comme dans le cas qui nous intéresse. Celui-ci est alors souvent traité comme un état composite de quarks et de gluons et des fonctions de structure associées dites “diffractives” sont définies (dPDF pour *diffractive proton density function*) en supposant le théorème de factorisation valide (voir section 3.3). Les lecteurs intéressés peuvent consulter [302] pour une revue des mesures des dPDFs et [303] pour une description plus complète du modèle théorique. Même si la structure du poméron est encore très incertaine [163], on sait qu'il est composé très majoritairement de gluons, ce qui donne une section efficace de production de paires de jets de gluons non-négligeable au LHC. Cette section efficace devrait être mesurée très prochainement à $\sqrt{s} = 13$ TeV par ATLAS-AFP et CMS-TOTEM (voir [132], chapitre 4).

La production DPE de photon+jet constitue également un bruit de fond réductible, mais sera négligée ici. Malgré le fait que seul un jet ait besoin d'être reconstruit comme photon pour imiter le signal, l'émission d'un photon et d'un jet à partir de deux pomérons nécessite l'extraction d'au moins un quark au lieu de deux gluons, or les quarks sont beaucoup moins abondants que les gluons dans le poméron d'après les fonctions de structure diffractives [163, 302]. Ainsi, un ordre de grandeur similaire entre la section efficace DPE photon+jet avec un jet reconstruit comme photon et la section efficace DPE di-jet avec les deux jets reconstruits comme deux photons ne serait pas surprenant. Cependant, comme nous le verrons dans la section 21.4, l'intégralité de ce bruit de fond sera largement rejetée dès lors que l'on se place à assez haute énergie, ce qui rend inutile un développement plus approfondi de sa simulation.

Un processus proche de DPE initié par photon-poméron est également attendu mais sera négligé dans la suite. En effet, la production de di-photon par ce mode est interdite afin d'assurer la conservation de la conjugaison de charge C , tout comme les di-photons

exclusifs ne peuvent pas être produits par photon-gluon.¹¹⁶ Il est cependant possible de considérer la production de photon+jet initiée par photon-poméron, qui pourrait représenter une production comparable au DPE di-photon. Néanmoins, compte tenu de la taille négligeable de ces bruits de fond dès lors que l'on se place à assez haute énergie, le processus photon+jet via photon-poméron sera négligé dans la simulation finale.

Finalement, seuls les deux diagrammes de la figure 200 sont inclus dans la simulation de bruit de fond et générés avec FPMC. Celui-ci utilise un flux de pomérons factorisé issu de la théorie de Regge combiné avec les PDFs diffractives mesurées à HERA [300]. Les reliquats des pomérons dans l'état final sont modélisés grâce à l'algorithme de cascade de parton de HERWIG 6.4 [304]. Comme déjà mentionné tout au long de ce paragraphe, les sections efficaces associées sont négligeables dès lors que l'on se place à assez haute énergie. Ainsi, leur simulation est réalisée uniquement afin de démontrer le faible impact des processus de diffraction inclusive sur la sensibilité finale de la mesure aux couplages anormaux à 4γ .

21.2.3 Bruits de fond avec protons intacts d'empilement à l'avant

Productions diffractives et empilement au LHC

En réalité, le bruit de fond dominant est constitué de productions inclusives non-diffractives ayant une grande section efficace à haute énergie et menant à la destruction des deux protons dans l'état final ($\gamma\gamma$, Drell-Yan e^+e^- , di-jet, γ +jet, etc), associé à la détection à l'avant de protons intacts émis en coïncidence de façon fortuite par des interactions d'empilement. En effet, celles-ci surviennent en grand nombre à chaque croisement de faisceaux ($\langle \mu \rangle = 50$ est considéré) et sont composées en grande partie de processus diffractifs inélastiques à bas t donnant lieu à un nombre significatif de protons intacts diffractés à petit angle dans l'état final.

Afin d'illustrer ce propos, les sections efficaces proton-proton inélastiques non-diffractives (ND, les deux protons sont détruits), simple-diffractives (SD, un proton est dissocié et le second intact), double-diffractives (DD, les deux protons sont dissociés) et diffractives centrales (CD, les deux protons sont intacts) sont rapportées dans le tableau 39 à $\sqrt{s} = 14$ TeV. Les chiffres sont extraits des prédictions des générateurs MC PYTHIA6 [305], PYTHIA8 [82] et PHOJET [306]. Seul PHOJET inclut les productions diffractives centrales, qui ont cependant une section efficace limitée. Comme attendu, les résultats des deux versions de PYTHIA sont assez proches. Ils indiquent qu'environ 25% des interactions d'empilement viennent de processus diffractifs, et donc en constituent une part conséquente. PHOJET prédit une section efficace ND 10% supérieure et DD presque deux fois plus faible que PYTHIA, avec un taux d'interaction diffractive restant tout de même autour de 20%. [180, 307]

Étant donné l'incertitude sur ces chiffres, il sera important de réaliser la mesure de ces productions à 13 ou 14 TeV, comme prévoit de le faire ATLAS-ALFA et CMS-TOTEM. Différentes mesures SD et DD ont déjà été réalisées par ALICE à $\sqrt{s} = 0.9, 2.76$ et 7 TeV [308] et par TOTEM à 7 TeV [309]. Les résultats se situent en général entre les chiffres prédits par PYTHIA et PHOJET.

¹¹⁶Cette propriété s'applique pour les processus DPE car les pomérons portent les nombres quantiques du vide.

Section efficace [mb] à $\sqrt{s} = 14$ TeV			
Processus	PYTHIA6	PYTHIA8	PHOJET
ND	54.7	57.5	67.9
SD	14.3	12.9	11.0
DD	10.2	8.9	4.1
CD	—	—	1.4

TAB. 39 – Sections efficaces inélastiques totales au LHC à $\sqrt{s} = 14$ TeV selon les générateurs PYTHIA6, PYTHIA8 et PHOJET, décomposées en processus non-diffractifs (ND), diffractifs (SD) et double-diffractifs (DD). PHOJET inclut également une catégorie pour la diffraction centrale (CD) [180, 307].

Les protons intacts et les états dissociés très collimatés émis par les interactions d’empilement mènent à des détections de protons à l’avant non-corrélées avec les objets à haute masse potentiellement émis dans le détecteur central, créant un bruit de fond important pour l’étude des couplages anormaux. Par conséquent, il est nécessaire d’inclure dans la simulation du bruit de fond une simulation fidèle de l’empilement afin d’obtenir des estimations réalistes de sensibilité. Le modèle adopté pour l’étude est explicité dans le paragraphe suivant.

Modèle adopté pour la simulation des interactions d’empilement

Afin de simuler la production de protons intacts à l’avant issus des interactions d’empilement, la démarche suivante est adoptée :

- ◊ 10 000 000 d’événements de biais minimaux sont générés sans empilement à $\sqrt{s} = 14$ TeV avec PYTHIA8. Ces événements représentent un échantillon réaliste de collisions proton-proton non biaisées par les systèmes de déclenchement des expériences. Ils permettent d’obtenir un échantillon représentatif des collisions d’empilement avec les proportions attendues pour les différents processus (élastiques, ND, SD, DD).
- ◊ PYTHIA8 inclut un algorithme de cascade de partons et d’hadronisation permettant de reproduire de façon réaliste toutes les particules présentes dans l’état final. Pour chaque événement, les protons intacts ainsi obtenus sont alors transportés jusqu’aux détecteurs à l’avant ($\simeq 200$ mètres du point d’interaction) grâce au code FPTRACKER [310], qui simule le trajet des particules à travers les différentes optiques du LHC, positionnées et paramétrées dans leur configuration nominale pour l’étude. Les résultats de FPTRACKER ont été recoupés au préalable avec ceux du logiciel MAD-X [311], qui est le code utilisé pour les opérations du LHC.
- ◊ On suppose ensuite des détecteurs à l’avant situés à 206 mètres du point d’interaction principal et à 2.5 mm du faisceau (environ 15 fois sa largeur), ce qui est conforme aux prévisions des projets ATLAS-AFP et CT-PPS [171]. La taille des détecteurs est très peu limitante pour la détection des protons les plus éloignés de la ligne de faisceau principale (grand ξ), qui sont écartés dans tous les cas en amont par les optiques du LHC. De tels détecteurs présentent l’acceptance

attendue en terme d'impulsion manquante pour les protons ($0.015 < \xi < 0.15$) déjà citée entre autres dans la section 6.2.2.

- ◇ Il est alors possible d'extraire la probabilité d'avoir une détection dans l'une des deux stations (gauche ou droite) ou une double détection (à gauche et à droite simultanément) pour une unique interaction de biais minimal (sans empilement pour le moment). Ces probabilités sont représentées dans la figure 201. Les distributions de probabilité de ξ associées, notées respectivement $f(\xi)$ et $f(\xi_1, \xi_2)$, sont également extraites et présentées dans la figure 202.
- ◇ La probabilité pour qu'un événement de biais minimal mène à une détection simple se situe entre 1 et 2%, ce qui est non négligeable si on considère le nombre d'interactions simultanées attendu ($\langle \mu \rangle = 50$). En revanche, la probabilité d'une double détection est de l'ordre de 0.01% et négligée. Cela permet de se restreindre à l'étude de la distribution de ξ à une dimension présentée dans la figure 202, gauche. La fonction densité de probabilité $f(-\log \xi) \simeq 0.05$ étant quasi-plate dans l'intervalle correspondant à l'acceptance des détecteurs à l'avant, on obtient par changement de variable $f(\xi) \simeq \frac{1}{\xi}$.
- ◇ Ainsi, pour chaque événement de bruit de fond généré, la démarche suivante est adoptée : un tirage poissonien de paramètre $\mu = 50$ est effectué afin de déterminer le nombre d'interactions d'empilement dans l'événement considéré. Pour chacune d'entre elles, un tirage uniforme indépendant est réalisé afin de déterminer si elles mènent à une détection à l'avant à gauche ou à droite du point d'interaction (*left/right tag*), ou pas du tout (*no tag*). Finalement, un ξ est tiré aléatoirement pour chaque interaction d'empilement menant à une détection de façon à suivre une densité de probabilité en $f(\xi) = \frac{1}{\xi}$ (voir point précédent).
- ◇ Chaque événement de bruit de fond présente alors une simulation réaliste de l'empilement, avec un nombre variable de détections sur le détecteur à gauche et à droite du point d'interaction principal, chacune associée à un ξ particulier.

L'effet de l'empilement sur la fréquence des détections à l'avant est illustré dans la figure 203, où la probabilité de détecter N protons sur l'un des côtés du point d'interaction principal est représentée pour $\langle \mu \rangle = 50, 100$ et 200 . Aucune détection particulière n'est demandée sur le détecteur opposé. À $\langle \mu \rangle = 50$ ($100, 200$), la probabilité d'avoir au moins une détection due au *pileup* est de 55% (80%, 95%) et d'en avoir exactement une de 35% (30%, 10%). Les détections multiples ($N > 1$) ne surviennent que dans 10% des cas pour $\langle \mu \rangle = 50$, mais arrivent dans plus de 50% (85%) des cas pour $\langle \mu \rangle = 100$ (200).

Cela ne devrait pas poser de problèmes majeurs pour la reconstruction des traces grâce à la très bonne segmentation des trajectographes à l'avant, issus de la même technologie que ceux utilisés pour le détecteur central (voir section 6.2). Cependant, le fort taux de remplissage à l'avant impose une segmentation très fine pour les détecteurs et l'électronique de temps de vol, ce qui nécessite l'emploi de technologies particulières comme des détecteurs ultra-rapides à silicium et la carte de lecture SAMPIC [172] (voir partie VI). L'estimation du taux de remplissage basé uniquement sur l'empilement doit par ailleurs être considéré comme une borne inférieure au chiffre réel, le bruit de fond

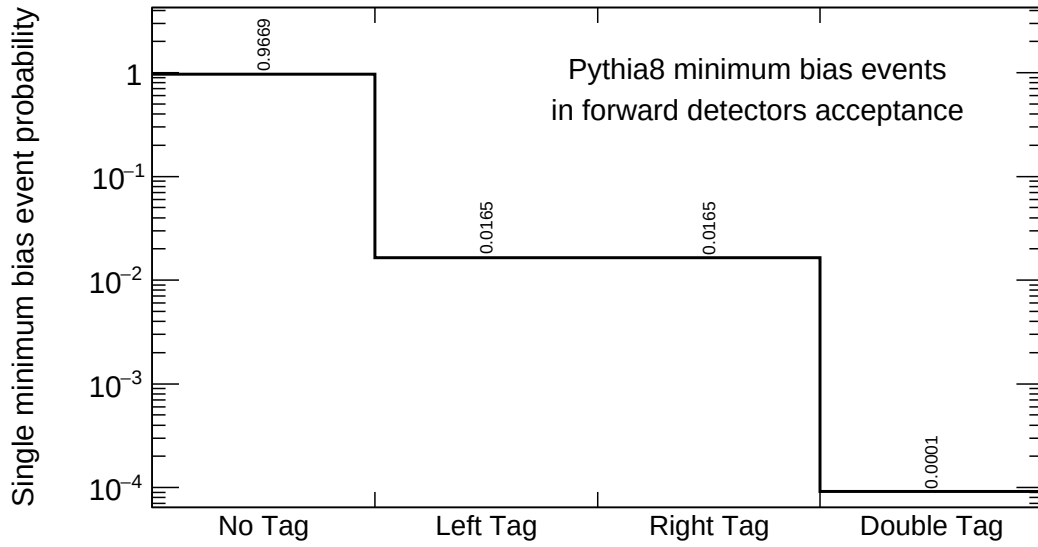


FIG. 201 – Probabilité de non-détection (No Tag), détection simple (Left/Right Tag) et détection double (Double Tag) dans les détecteurs à l'avant pour des événements de biais minimaux générés avec PYTHIA8 à $\sqrt{s} = 14$ TeV. Aucun empilement n'est considéré pour le moment [180, 307].

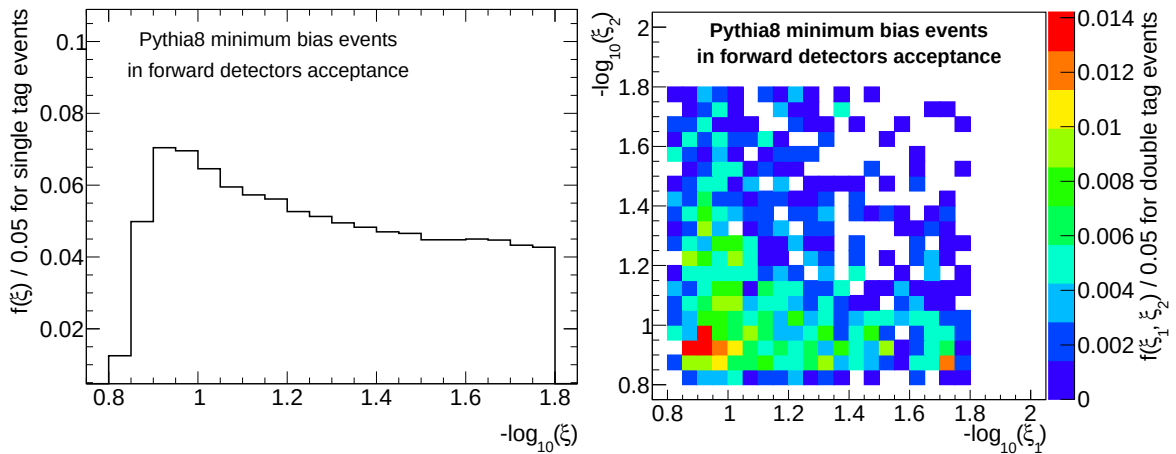


FIG. 202 – Densité de probabilité pour ξ dans le cas d'une détection simple (gauche) ou double (droite) dans les détecteurs à l'avant pour des événements de biais minimaux générés avec PYTHIA8 à $\sqrt{s} = 14$ TeV. Aucun empilement n'est considéré pour le moment [180, 307].

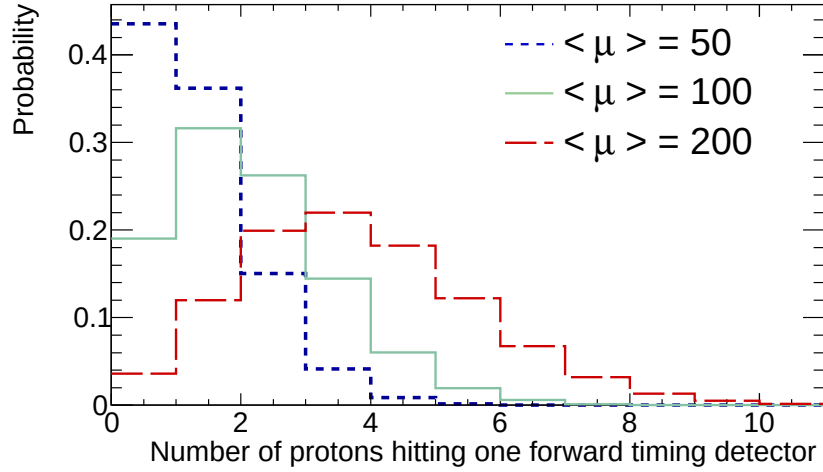


FIG. 203 – Probabilité de détection de N protons d’un côté du point d’interaction en fonction du nombre moyen d’interactions d’empilement considéré : $\langle \mu \rangle = 50$ (bleu), 100 (vert) ou 200 (rouge). Aucune détection particulière n’est demandée sur le détecteur opposé [172, 180, 307].

issu du faisceau (*beam-induced*), en grande partie inconnu, affectant très probablement le taux de remplissage à bas ξ (région la plus proche du faisceau). Celui-ci devrait être néanmoins très réduit après application de la sélection finale qui sera détaillée dans la section 21.4.

Il est intéressant de revenir sur la forme de la distribution $f(-\log \xi)$ présentée dans la figure 202, gauche. Celle-ci a été supposée constante pour l’étude, permettant d’obtenir la relation simple $f(\xi) = \frac{1}{\xi}$. Néanmoins, un léger pic est observé autour de $-\log \xi \simeq 1$ (soit $\xi \simeq 0.1$) et une décroissance rapide pour $\xi > 0.11$. Cette forme est également observée dans les simulations de l’expérience ATLAS à 13 TeV (voir [136], figure 12a) et attribuée à la section efficace SD.¹¹⁷ Il est donc attendu que l’approximation mentionnée plus haut mène à une légère sous-estimation (sur-estimation) du fond dû à l’empilement pour $\xi \simeq 0.1$ ($\xi > 0.11$). Néanmoins, comme nous le verrons dans la section 21.4), l’étude est très robuste vis-à-vis de l’empilement et donc ne peut être que très faiblement pénalisée par ces effets.

Bruit de fond avec protons intacts d’empilement à l’avant inclus dans la simulation finale

La plupart des états finals de collisions comportant des photons, des électrons ou des jets peuvent donc imiter le signal di-photon exclusif au LHC à cause des protons d’empilement et/ou de mauvaises identifications dans le détecteur central. Afin de limiter l’étude à un cadre raisonnable, seules les productions inclusives de di-photon, di-jet et Drell-Yan (e^+e^-) sont considérées et générées avec HERWIG 6.4. L’état final complet est modélisé grâce à l’algorithme de cascade de partons du générateur, tandis que les effets d’empilement sont simulés en utilisant la stratégie présentée au paragraphe précédent.

En réalité, la production correspondant à l’état final γ +jet avec un jet mal identifié

¹¹⁷Plus particulièrement, cet effet est issu de la courbe de flux du poméron.

devrait constituer un bruit de fond compris entre ceux issus des di-photons et des di-jets (voir partie IV). Par ailleurs, HERWIG 6.4 intégrant uniquement les éléments de matrice LO des différentes productions, une sous-estimation des sections efficaces de l'ordre de 20 à 40% est attendue. L'incertitude sur le bruit de fond d'empilement est donc assez grande et la robustesse de la sélection présentée dans la section 21 sera testée par de variations multiples afin de s'assurer de la fiabilité des résultats finals de sensibilité. Ces variations affecteront finalement très peu l'analyse grâce à la très bonne efficacité et le taux de rejet excellent de la sélection.

Dans la section 21.3, la modélisation des effets de détecteur adoptée dans l'étude, incluant les taux de conversion des photons, les efficacités de présélection (reconstruction, identification, isolation), les fréquences de mauvaises identifications et les effets de résolution sur les observables, est présentée.

21.3 Prise en compte des effets de détecteurs

21.3.1 Taux de conversion des photons

Comme mentionné dans la partie III, certains photons se convertissent en paires électron-positron lors de leur passage dans le détecteur par interaction avec le matériel environnant. Au niveau de l'analyse, cela permet d'avoir des traces associées aux dépôts calorimétriques des photons et donc de pouvoir identifier le vertex d'émission de la paire de photons avec une précision supérieure au millimètre [190]. Sans traces associées aux dépôts, celui-ci ne peut être situé qu'avec une précision de l'ordre de 10 à 20 mm, en utilisant par exemple la segmentation longitudinale du calorimètre (ATLAS) ou un algorithme d'apprentissage artificiel basé sur toutes les traces reconstruites de l'événement, issues par exemple de partons supplémentaires émis au cours de l'interaction (CMS).

La détermination précise de la position du vertex d'interaction par le détecteur central permet de rejeter jusqu'à 40 fois plus de protons d'empilement grâce à la comparaison avec les mesures de temps de vol à l'avant, qui devraient présenter des résolutions de l'ordre de 10 à 15 ps et donc permettre l'identification d'un vertex d'émission avec une précision d'environ 2 mm [171]. Ainsi, deux scénarios concernant le statut de conversion des photons sont considérés dans l'étude : au moins un photon converti, et aucune sélection particulière sur le statut de conversion. La réduction supplémentaire du bruit de fond apportée par les mesures de temps de vol n'est jamais appliquée car celui-ci est déjà réduit à zéro à l'issue de la sélection, comme nous le verrons dans la section 21.4. Néanmoins, il est intéressant d'estimer la sensibilité dans les deux scénarios afin de garder une marge confortable pour l'étude.

Il est donc nécessaire de simuler le taux de conversion des photons. Un taux réaliste de 15% de conversion dans la partie tonneau ($|\eta| < 1.4$) et 30% dans les bouchons ($1.4 < |\eta| < 2.37$) est adopté [43], confirmé par les mesures récentes réalisées à 13 TeV [312]. Les conversions sont attendues plus nombreuses dans les bouchons à cause de la plus grande quantité de matière à traverser (voir section 5.3).

21.3.2 Efficacité de présélection des photons et des électrons

Le traitement des photons et des électrons dans ATLAS a été abordé dans la partie III. Les performances attendues à $\sqrt{s} = 14$ TeV pour $\langle \mu \rangle = 80$, publiées dans le cadre du rapport rendu en 2013 au comité européen pour les accélérateurs du futur (ECFA) dans le cadre de prévisions pour le HL-LHC [313], sont utilisées pour l'étude. Suivant les conclusions du rapport, l'efficacité de présélection des photons, incluant la reconstruction et les critères d'identification et d'isolation, est estimée grâce à la formule suivante :

$$\epsilon_{\text{presel}}^{\text{EM}}(p_T) = 0.76 - 1.98 \exp \frac{-p_T [\text{GeV}]}{16.1} . \quad (111)$$

La dépendance (relativement faible) des efficacités vis-à-vis de la pseudo-rapidité est donc négligée et la formule de l'équation 111 est appliquée pour l'ensemble des photons tombant dans l'acceptance ($|\eta| < 2.37$). La même formule est extrapolée de façon inchangée pour les électrons, qui ont une chaîne de traitement proche de celle des photons en termes de reconstruction, identification et isolation.

À très haute impulsion transverse ($p_T > 100$ GeV) $\epsilon_{\text{presel}}^{\text{EM}}$ atteint 76%, ce qui paraît conservateur au regard des résultats à 8 TeV et $\langle \mu \rangle \simeq 20$ présentés pour des paires de photons ($\epsilon_{\text{presel}}^{\text{EM}} \times \epsilon_{\text{presel}}^{\text{EM}}$) dans la section 16, partie IV. Les conclusions rendues à l'ECFA se basent sur $\langle \mu \rangle = 80$. Elles sont ensuite extrapolées de façon inchangée à $\langle \mu \rangle = 140$ (HL-LHC) en considérant les performances stables via une ré-optimisation des coupures. Ainsi, l'étude présentée ici, considérant $\langle \mu \rangle = 50$ comme valeur nominale (LHC phase-1), supposera $\epsilon_{\text{presel}}^{\text{EM}}$ inchangé lors de l'extrapolation des résultats au HL-LHC (3000 fb^{-1} à $\langle \mu \rangle = 200$).

L'efficacité de présélection pour les jets sera comprise dans le taux de mauvaise identification en photons introduit dans la section 21.3.3.

21.3.3 Taux de mauvaise identification des électrons et des jets

Électrons

Les événements comportant N électrons et/ou photons issus de l'interaction dure dans l'état final passent tout d'abord l'étape de présélection en étant pondérés par $(\epsilon_{\text{presel}}^{\text{EM}})^N$ (voir section 21.3.2). Un poids supplémentaire fixe de 0.01^N est alors multiplié au poids final dans le cas des électrons afin de tenir compte du faible taux de mauvaise identification des électrons en photons par le détecteur. Ce chiffre correspond à l'ordre de grandeur obtenu dans le cadre de la mesure de la masse du boson de Higgs dans ATLAS [208]. Depuis, il a été ré-évalué à la hausse et se situe plutôt autour de 0.05 dans ATLAS, stable pour les différents taux d'empilement disponibles ($0 < \mu < 33$) [234]. L'effet d'une telle hausse ne change pas les résultats de l'étude, comme nous le verrons dans la section 22.

Par ailleurs, les électrons mal-identifiés comme photons sont considérés à 100% comme des photons convertis alors qu'ils peuvent également être reconstruits comme des photons non-convertis lorsque la trace associée est perdue dans le trajectographe.

Cependant, aucun chiffre n'a été rendu public à ce sujet, d'où la simplification considérée ici. Cela mène probablement à une légère sur-estimation du bruit de fond lié aux électrons en cas de sélection sur le statut de conversion, compensant partiellement la sous-estimation du taux de mauvaise identification. La dépendance faible et mal connue du nombre de mauvaises identifications des électrons vis-à-vis de l'impulsion transverse et de la pseudorapidité est négligée dans tous les cas.

Jets hadroniques

Les jets hadroniques sont d'abord reconstruits au niveau particule en utilisant l'algorithme anti- k_T [314] avec un paramètre $R = 0.4$, couramment utilisé dans les analyses d'ATLAS [78]. Cette étape permet d'extraire leur quadri-vecteur impulsion.

La proportion de jets passant les critères d'identification des photons est ensuite estimée grâce à la formule suivante, tirée du rapport de l'ECFA mentionné dans la section 21.3.2 :

$$\epsilon_{\text{fake}}^{\text{jet}}(p_T) = 0.0093 \exp \frac{-p_T[\text{GeV}]}{27.5} \quad (112)$$

où p_T représente l'impulsion transverse du jet renvoyé par anti- k_T . Cependant, dans le cadre de l'étude des couplages à 4γ $\epsilon_{\text{fake}}^{\text{jet}}(p_T = 200 \text{ GeV})$ est utilisé de façon conservatrice pour tous les jets ayant des impulsions transverses supérieures à 200 GeV. En effet, le rapport de l'ECFA se base sur des événements $H \rightarrow \gamma\gamma$ qui n'atteignent probablement pas des impulsions transverses aussi élevées et pourrait donc sous-estimer le taux de mauvaise identification à haut p_T .

Contrairement aux électrons, les jets mal-identifiés sont considérés comme des photons convertis ou non convertis selon les taux exposés dans la section 21.3.1. En effet, ceux-ci sont composés en grande partie de mésons légers ($\pi^0 \rightarrow \gamma\gamma, \dots$) se désintégrant en photons avant leur entrée dans le détecteur interne, et donc sont soumis a priori aux mêmes effets de conversion que les photons immédiats en première approximation.

21.3.4 Résolution sur les observables mesurées

Finalement, des altérations en énergie, pseudo-rapidité, angle d'azimut et ξ sont appliquées de manière aléatoire afin de reproduire les effets dus à la résolution limitée des détecteurs centraux et à l'avant.

Pour tous les objets centraux considérés dans l'étude (photons, jets et électrons reconstruits comme photons), un tirage gaussien centré autour de la valeur réelle de l'observable au niveau particule et de largeur 0.001 (valeur absolue) est appliqué aux pseudo-rapidités et angles d'azimut.¹¹⁸ En effet, ceux-ci sont très bien déterminés grâce à la très fine granularité du calorimètre électromagnétique (voir section 5.4) et ne souffrent que d'effets secondaires reliés à l'empilement, à l'alignement et à des déformations légères diverses (poids des cellules, ...).

Pour les jets, une fois le poids relié à la probabilité de mauvaise identification appliqué (voir section 21.3.3), les impulsions transverses sont recalculées en effectuant un

¹¹⁸Les angles d'azimut ϕ sont exprimés en radians dans l'intervalle $[0, \pi]$.

tirage gaussien centré en $0.75 p_T^{\text{jet}}$ avec une largeur correspondant à $0.13 p_T^{\text{jet}}$. Ces chiffres sont à nouveau extraits des résultats contenus dans le rapport rendu à l'ECFA [313] et permettent de tenir compte à la fois du biais sur l'échelle en énergie dû à l'application de la calibration des photons à des jets et des effets de résolution, venant principalement des variations topologiques des jets.

Comme nous le verrons dans la section 21.4, l'étude des couplages à 4γ considère uniquement les photons et les électrons ayant un $p_T > 100$ GeV. Ainsi, la résolution du calorimètre est dominée par le terme constant (voir section 5.4 et 8), qui est fixé de façon conservatrice à 1%. Une déformation gaussienne centrée autour de la valeur réelle de l'énergie E et présentant une largeur de $0.01.E$ est alors appliquée et propagée dans le calcul des différentes observables (E_T , $m_{\gamma\gamma}$, ...).

La résolution considérée sur la mesure de l'impulsion manquante des protons à l'avant (ξ) est de 2%, ce qui est conservateur vis-à-vis des estimations faites dans le cadre des projets ATLAS-AFP et CT-PPS (voir section 6.2.2). Ces déformations sont également propagées dans le calcul des différentes observables à l'avant (m_{pp}^{miss} , y_{pp}).

La simulation du bruit de fond et des effets de détecteurs ont été détaillés et justifiés pas à pas au cours des sections précédentes. Dans la section 21.4, la sélection opérée dans les pseudo-données finale, constitué du signal et des différents bruits de fond, est discutée en détail.

21.4 Sélection des pseudo-données

21.4.1 Distribution du signal et des bruits de fond sans sélection supplémentaire

Dans la partie 21.1, le signal a été défini par la détection en coïncidence d'au moins deux protons intacts dans les stations à l'avant ($0.015 < \xi < 0.15$) et d'au moins une paire de photons de haute énergie dans le détecteur central. Dans la suite, on considèrera $E_{T,1(2)}^\gamma > 100$ (200) GeV. Ce critère sur l'impulsion transverse des photons devrait être largement suffisant pour maintenir dans tous les cas les taux de déclenchement des expériences acceptables sans facteur de *prescale* (voir section 5.7), le premier *trigger* di-photon LOOSE officiel d'ATLAS à 13 TeV sans facteur de *prescale* (voir section 9) ayant comme seuil $E_{T,1(2)}^\gamma > 50$ GeV pour le *run 2*.

Les distributions des événements de signal et de bruit de fond après l'application de cette sélection minimale sont montrées dans la figure 204 en fonction de la masse invariante de la paire de photons ($m_{\gamma\gamma}$) dans le cas du scénario principal : $\mathcal{L} = 300 \text{ fb}^{-1}$, $\langle \mu \rangle = 50$ et aucune sélection sur le statut de conversion. Deux signaux différents sont générés par l'intermédiaire de la théorie effective présentée dans la section 19.1 ($\zeta_1 = 0.1$, 0.01 GeV^{-4} et $\zeta_2 = 0$).

Les distributions des signaux et des bruits de fond exclusifs ($pp \rightarrow p + \gamma\gamma + p$ et $p + e^+e^- + p$) ont une forme suivant l'acceptance en masse des détecteurs à l'avant (voir section 6.2.2), tandis que les bruits de fond nécessitant la détection de protons intacts d'empilement ($pp \rightarrow \gamma\gamma$, e^+e^- , di-jet) en sont totalement décorrélés. Les événements DPE présentent également une corrélation vis-à-vis de l'acceptance des détecteurs à l'avant mais le pic est décalé à plus basse masse comparé aux protons à cause de l'énergie

portée par les produits de désintégration des pomérons, non inclus dans $m_{\gamma\gamma}$.

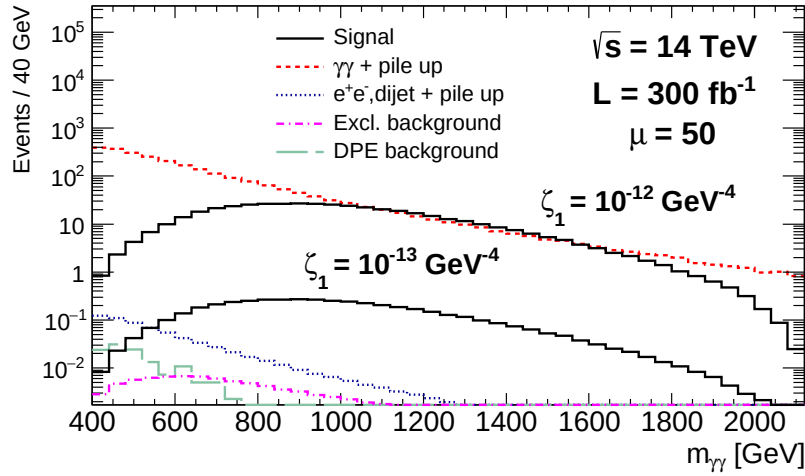


FIG. 204 – Nombre d'événements attendus pour 300 fb^{-1} de données à $\sqrt{s} = 14 \text{ TeV}$ pour deux signaux générés avec la théorie effective ($\zeta_1 \neq 0, \zeta_2 = 0$) et les différents types de bruit de fond, en fonction de la masse invariante de la paire de photons reconstruite. Seuls les événements avec $E_{T,1(2)}^\gamma > 100$ (200) GeV et au moins un proton intact de chaque côté du point d'interaction principal dans l'acceptance des détecteurs à l'avant ($0.015 < \xi < 0,15$) sont considérés. Excl.=Exclusive et DPE=*double pomeron exchange*.

Comme attendu, le bruit de fond lié à l'empilement domine, et en particulier la production inclusive de di-photons correspondant au processus mesuré à $\sqrt{s} = 8 \text{ TeV}$ dans la partie IV. Une sélection supplémentaire apparaît donc essentielle afin d'obtenir un signal sur bruit acceptable pour des valeurs de couplages anormaux intéressants ($\zeta < 0.1 \text{ TeV}^{-4}$, voir section 19).

Dans les sections 21.4.2 et 21.4.3, les coupures supplémentaires choisies pour la sélection finale sont décrites.

21.4.2 Sélection des hautes masses et des productions dos-à-dos dans le détecteur central

Les événements de signal étant émis à très haute masse (voir figure 204), seules les masses supérieures à 600 GeV ($m_{\gamma\gamma}$ calculé à partir du détecteur central) seront considérées dans la suite. Ce critère supplémentaire permet de supprimer complètement le bruit de fond exclusif et DPE tout en présentant une efficacité supérieure à 95% pour le signal.

À l'issue de la sélection des hautes masses, uniquement le bruit de fond relié à l'empilement contribue toujours de façon non-négligeable à la contamination des données. Une première façon simple d'en éliminer une partie consiste à sélectionner les paires de photons présentant une topologie dos-à-dos dans le plan transverse, typique des processus avec exactement deux particules centrales de haute énergie dans l'état final. Dans les faits, il est demandé à ce que les deux photons aient une impulsion similaire ($E_{T,2}^\gamma/E_{T,1}^\gamma > 0.95$) et des angles d'azimut opposés ($\Delta\phi_{\gamma\gamma} > \pi - 0.01$). Cela permet de supprimer une partie non-négligeable du bruit de fond tout en gardant une efficacité

très proche de 100% pour le signal.¹¹⁹

Cependant, ces coupures ne permettent pas de se débarrasser entièrement du bruit de fond restant. En effet, la plupart des productions inclusives au LO, comme par exemple la production non-exclusive de di-photons ($\gamma\gamma + \text{pileup}$), est elle-même constituée d'exactly deux particules de haute énergie dans l'état final ($q\bar{q} \rightarrow \gamma\gamma$) et donc passe en général de tels critères. Ainsi, il est absolument nécessaire d'exploiter les mesures à l'avant très précises de ξ ($\simeq 2\%$) afin de décontaminer entièrement l'échantillon du bruit de fond, comme cela est montré dans la section 21.4.3.

21.4.3 Identification des systèmes centraux et à l'avant

La détection des protons à l'avant donne un avantage déterminant pour rejeter le bruit de fond non-exclusif par rapport à une analyse classique utilisant uniquement le détecteur central. En effet, dans le cas d'un événement de signal, l'intégralité des particules émises dans l'état final sont détectées, soit une paire de protons et une paire de photons. Par conservation de Lorentz, la masse invariante et la rapidité de la paire de photons, notées respectivement $m_{\gamma\gamma}$ et $y_{\gamma\gamma}$, correspondent nécessairement aux mesures effectuées à l'avant avec la paire de protons, $m_{pp}^{\text{miss}} = \sqrt{\xi_1 \xi_2 s}$ et $y_{pp} = 0.5 \ln(\xi_1/\xi_2)$ (voir section 6.2.2), moyennant les effets de résolution. En revanche, ces variables n'ont aucune raison de correspondre pour le bruit de fond restant, où les protons d'empilement sont totalement décorrélés de l'événement central de haute énergie.

Ces caractéristiques sont illustrées dans la figure 205, où les distributions des événements de signal et de bruit de fond sont représentés en fonction du rapport entre les masses (gauche) et de la différence entre les rapidités (droite) des systèmes centraux et à l'avant, après sélection sur les hautes masses et les topologies dos-à-dos (voir section 21.4.2). La largeur non-nulle observée pour les distributions de signal provient des effets de résolution introduits dans la section 21.3.4. L'excellente résolution du calorimètre électromagnétique central associée aux très bonnes mesures de ξ dans les stations à l'avant permettent une discrimination impressionnante entre signal et bruit de fond. Seuls les événements présentant $\sqrt{\xi_1 \xi_2 s} = m_{\gamma\gamma} \pm 3\%$ et $|y_{\gamma\gamma} - y_{pp}| < 0.03$ seront considérés pour la sélection finale.

Un bilan résumant les différentes étapes de la sélection est donné dans la section 21.4.4.

21.4.4 Bilan

Les nombres d'événements attendus après chaque coupure pour les différents bruits de fond et un exemple de signal sont rapportés dans le tableau 40. Le signal est généré en utilisant le calcul des amplitudes complètes pour un boson exotique de charge effective $Q_{\text{eff}} = 4$ et de masse $m = 340$ GeV, ce qui correspond à un signal au-dessus des limites de sensibilité de l'étude mais au-delà des conditions de validité de la théorie effective (voir section 19.1.4). Les signaux correspondants générés à partir de la théorie effective avec et sans facteur de forme sont également représentés (voir équation 96). La théorie

¹¹⁹Comme nous les verrons, ces deux coupures ne sont pas essentielles étant donné le taux de rejet très élevé des coupures utilisant les mesures à l'avant (voir section 21.4.3).

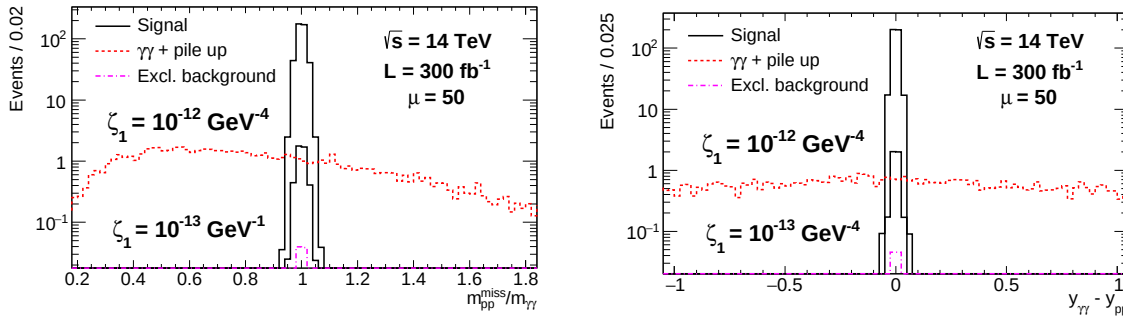


FIG. 205 – Nombre d'événements attendus pour 300 fb^{-1} de données à $\sqrt{s} = 14 \text{ TeV}$ pour deux signaux générés avec la théorie effective ($\zeta_1 \neq 0, \zeta_2 = 0$) et les différents types de bruits de fond, en fonction du rapport entre les masses (gauche) et de la différence entre les rapidités (droite) des systèmes centraux et à l'avant. La sélection sur les hautes masses et les topologies dos-à-dos décrites dans la section 21.4.2) est appliquée en plus de la sélection initiale.

effective n'étant pas valide pour ce domaine en masse, on s'attend à ce que les résultats avec et sans facteur de forme encadrent le calcul complet, ce qui est bien le cas. Les chiffres affichés dans le tableau 40 incluent le critère sur le statut de conversion, *ie.* au moins un des deux photons convertis. Ils sont décrits dans le paragraphe suivant.

L'efficacité totale de sélection du signal vis-à-vis des coupures supplémentaires implémentées pour rejeter le bruit de fond est très élevée, de l'ordre de 90%.¹²⁰ Le taux de rejet du bruit de fond dû à l'empilement n'est pas directement calculable à partir du tableau 40 car le bruit de fond attendu est très inférieur à 1 après sélection. Le chiffre exact pour $\gamma\gamma + \text{pileup}$ est 0.003 événements, ce qui donne un facteur de rejet d'environ 10^6 . Ces très bonnes performances sont atteintes tout d'abord grâce à l'excellente résolution sur les mesures de ξ (1-2%), qui est absolument essentielle à l'obtention d'un rapport signal sur bruit acceptable. Ainsi, il sera très important de mesurer ces performances dès le début de la prise de données d'ATLAS-AFP et CT-PPS. En effet, sans l'application des coupures décrites dans la section 21.4.3, le taux de rejet perd environ trois ordres de grandeur et le nombre d'événements attendus venant du bruit de fond $\gamma\gamma + \text{pileup}$ devient alors comparable aux événements de signal.

La robustesse de ces résultats est illustrée par plusieurs aspects :

- ◇ Le taux de rejet à $\langle \mu \rangle = 50$ (10^6) assure une suppression totale de celui-ci, y compris dans le cas d'une sous-estimation des sections efficaces inclusives de plusieurs ordres de grandeurs.
- ◇ L'exercice a été reproduit avec $\langle \mu \rangle = 200$ et mène à la détection de moins de 2 (5) événements de bruit de fond d'empilement sans ré-optimisation des coupures dans le cas d'une sélection (ou non) sur le statut de conversion. Cela donne une marge très confortable sur le modèle d'empilement adopté (voir section 21.2.3).

¹²⁰L'efficacité de présélection a été détaillée dans la section 21.3 et n'est pas incluse dans ces chiffres, qui sont donc valides après présélection et sélection sur $\xi_{1,2}$ ($> 60\%$, dépend de la masse considérée) et $E_{T1,2}^\gamma$ ($\simeq 0.76^2 \simeq 0.6$, voir équation 21.3.2).

Coupure / Processus	Signal complet	Signal EFT avec (sans) f.f	Excl.	DPE	DY, di-jet + <i>pileup</i>	$\gamma\gamma$ + <i>pileup</i>
$[0.015 < \xi_{1,2} < 0.15,$ $E_{T1,(2)}^\gamma > 200, (100) \text{ GeV}]$	65	18 (187)	0.13	0.2	1.6	2 968
$m_{\gamma\gamma} > 600 \text{ GeV}$	64	17 (186)	0.10	0.03	0.2	1 023
$[p_{T2}/p_{T1} > 0.95,$ $ \Delta\phi > \pi - 0.01]$	64	17 (186)	0.10	0.003	0.02	80.2
$\sqrt{\xi_1 \xi_2 s} = m_{\gamma\gamma} \pm 3\%$	61	16 (175)	0.09	$< 10^{-6}$	$< 10^{-6}$	2.8
$ y_{\gamma\gamma} - y_{pp} < 0.03$	60	12 (169)	0.09	$< 10^{-6}$	$< 10^{-6}$	0.003

TAB. 40 – Nombres d'événements de signal et de bruit de fond attendus après chaque étape de la sélection dans le cas de 300 fb^{-1} de données au LHC à $\sqrt{s} = 14 \text{ TeV}$. Le signal donné en exemple correspond à un boson exotique de charge effective $Q_{\text{eff}} = 4$ et de masse $m = 340 \text{ GeV}$. Les prédictions correspondantes obtenues à partir de la théorie effective avec et sans facteurs de formes sont également fournies. Au moins un photon converti est demandé lors de la sélection.

- ◇ Les mesures de temps de vol à l'avant n'ont pas été nécessaires afin de supprimer totalement le bruit de fond d'empilement. Celles-ci permettent de gagner un facteur de rejet supplémentaire proche de 40 à $\langle \mu \rangle = 50$ dans le cas d'une résolution sur la mesure de temps de 10 ps (voir section 6.2).
- ◇ Le nombre d'événements de signal peut être multiplié par un facteur 3 à 4 si aucune sélection n'est faite sur le statut de conversion des photons, ce qui permet d'augmenter du même facteur environ le rapport signal sur bruit vis-à-vis des bruits de fond issus de jets mal identifiés, ces derniers bénéficiant d'un taux de rejet très élevé. Cela engendre néanmoins une légère dégradation de la résolution sur la masse centrale ($m_{\gamma\gamma}$), inférieure à 1% dans le cas d'ATLAS grâce à l'utilisation du pointing [190] (voir section 21.1). L'effet pour CMS devrait être plus important à cause de l'absence de segmentation longitudinale pour le calorimètre.
- ◇ Le bruit de fond exclusif est celui présentant le taux de rejet le plus faible car ni la sélection topologique, ni les mesures de ξ ne permettent de le discriminer des événements de signal. Cependant, il est extrêmement faible aux énergies considérées (voir section 20) et bénéficie de prédictions théoriques stables et en particulier indépendantes de la détermination des PDFs à haute énergie (KMR, EPA).

Dans la section 22, la méthode statistique pour estimer la sensibilité des données aux signaux de nouvelle physique est décrite. Les résultats finals de sensibilité sont ensuite présentés et commentés dans les différents cas (théorie effective, bosons et fermions chargés, spin 0 et spin 2 neutres). Finalement, une conclusion générale est donnée, incluant une discussion des autres canaux potentiellement sensibles aux couplages anormaux à quatre photons et du rôle que pourrait jouer le canal di-photon exclusif dans le cas d'une confirmation du léger excès rapporté par ATLAS et CMS dans les premières données à $\sqrt{s} = 13 \text{ TeV}$ prises en 2015.

22 Sensibilités finales, comparaisons aux limites actuelles et perspectives futures

22.1 Estimation de la sensibilité des données

Comme cela a été démontré dans la section 21.4, la sélection effectuée dans les pseudo-données permet de supprimer l'intégralité du bruit de fond dans le cas de 300 fb^{-1} de donnée de collisions proton-proton à $\sqrt{s} = 14 \text{ TeV}$ au LHC, avec une très bonne marge d'erreur sur la simulation et ce quelque soit la sélection sur le statut de conversion des photons. Par exemple, dans le cadre d'au moins un photon converti on attend environ 0.09 événements issus du bruit de fond exclusif et 0.003 événement $\gamma\gamma + \text{pileup}$. Les autres processus se situent au moins deux ordres de grandeur en dessous.

On se trouve alors dans une situation où la moindre observation présente une significativité statistique très haute. En effet, le nombre d'événements N passant une sélection en physique des particules est en général décrit comme une variable aléatoire suivant une distribution de Poisson, l'échantillon sélectionné étant très réduit par rapport à celui de départ (loi des événements rares). L'incertitude statistique sur le nombre d'événements attendus après une sélection comme celle décrite dans la section 21.4 est alors de $\sqrt{N}$ pour N assez grand (voir paragraphe suivant). Il est courant d'évaluer la sensibilité d'un canal vis-à-vis de nouveaux phénomènes en comparant le nombre total d'événements observés à l'incertitude sur le nombre total d'événements de bruit de fond attendus (hypothèse 100% modèle standard) :

$$\sigma = \frac{N_{\text{tot}}}{\sqrt{N_{\text{tot}}^{\text{bkg}}}}, \quad (113)$$

où N_{tot} représente le nombre total d'événements observés, se réduisant dans notre cas aux événements de signal, et $N_{\text{tot}}^{\text{bkg}}$ représente le nombre total d'événements de bruit de fond attendus, soit environ 0.093 dans le cas décrit plus haut. σ est appelée significativité et peut être utilisée pour calculer une valeur p , définie comme étant la probabilité pour qu'une mesure similaire effectuée sur un échantillon de données indépendant mène à un désaccord identique ou moins bon. En physique des particules, on considère souvent le critère 5σ comme étant le critère minimum pour revendiquer une découverte, correspondant à une valeur p de $3 \cdot 10^{-7}$ si $N_{\text{tot}}^{\text{bkg}}$ est assez grand (limite gaussienne).

Cependant, deux facteurs empêchent de calculer directement les sensibilités avec l'équation 113 :

- ◇ La loi de Poisson est une loi discrète qui n'a pas d'approximation continue simple dans le cas $N_{\text{tot}}^{\text{bkg}} < 1$ en statistiques fréquentistes.
- ◇ La simulation du bruit de fond comporte un certain nombre d'incertitudes systématiques dont les sources ont été discutées dans la section 21.2. Elles devraient être sommées en quadrature au dénominateur de la formule 113 afin d'estimer correctement les sensibilités mais sont difficiles à évaluer précisément. Ces incertitudes supplémentaires devraient cependant constituer un très petit effet.

Par conséquent, la démarche conservatrice suivante est adoptée : lorsque $N_{\text{tot}}^{\text{bkg}}$ est très inférieur à 1, on le surestime à 1, ce qui revient à remplacer le terme $\sqrt{N_{\text{tot}}^{\text{bkg}}}$ par $\sqrt{N_{\text{tot}}^{\text{bkg}} + 1}$ dans l'équation 113. Ainsi la formule de l'équation 113 est utilisable directement et se réduit à $\sigma = N_{\text{tot}}$, ce qui constitue un critère très simple et transparent d'estimation de la sensibilité des données (5 événements observés = 5σ).

Dans les sections 22.2 à 22.4, les résultats finals de sensibilité évalués avec le critère introduit dans cette section sont présentés pour des signaux générés en utilisant le modèle basé sur la théorie effective (sections 22.2 et 22.4) et les calculs complets des amplitudes (section 22.3).

22.2 Résultats obtenus par la théorie effective

Les sensibilités finales du canal di-photon exclusif aux couplages anormaux à quatre photons définis dans le cadre du modèle effectif présenté dans la section 19.1 sont résumées dans le tableau 41. Différents scénarios sont considérés : $300 \text{ fb}^{-1}/<\mu> = 50$ ou $3000 \text{ fb}^{-1}/<\mu> = 200$ (LHC ou HL-LHC), sélection sur le statut de conversion ou non, significativité de 5σ (découverte) ou 1.96σ (intervalle de confiance de 95% dans le cadre de l'approximation gaussienne, sera pris comme 2σ dans la suite).

$\mathcal{L}$	300 fb^{-1}	300 fb^{-1}	300 fb^{-1}	3000 fb^{-1}
$<\mu>$	50	50	50	200
couplage (TeV^{-4})	$\geq 1 \gamma \text{ conv.}$ 5σ	$\geq 1 \gamma \text{ conv.}$ 95% CL	all γ 95% CL	all γ 95% CL
$\zeta_1 \text{ f.f.}$	0.15	0.075	0.04	0.035
$\zeta_1 \text{ no f.f.}$	0.035	0.025	0.01	0.01
$\zeta_2 \text{ f.f.}$	0.25	0.15	0.085	0.07
$\zeta_2 \text{ no f.f.}$	0.075	0.045	0.025	0.025

TAB. 41 – Sensibilités finales du LHC à $\sqrt{s} = 14 \text{ TeV}$ aux couplages anormaux à quatre photons paramétrés dans le cadre de la théorie effective en (ζ_1, ζ_2) pour différentes hypothèses en terme de statistiques et de conditions d'empilement. Les sensibilités sont estimées à partir du canal di-photon exclusif avec détection des protons intacts. Les chiffres sont donnés pour une significativité de 5σ ou 1.96σ , cette dernière étant équivalent à un intervalle de confiance de 95% dans le cadre de l'approximation gaussienne (95% CL). Une sélection sur le statut de conversion des photons peut être effectuée ou non.

Les couplages anormaux à quatre photons peuvent être observés jusqu'à 0.01 TeV^{-4} environ, ce qui permet de tester directement un grand nombre de modèles de nouvelle physique apparaissant à l'échelle de plusieurs TeV, en particulier ceux incluant des dimensions supplémentaires de l'espace (voir section 19.1.3 et 22.4). Une extrapolation naïve de l'étude montre que les sensibilités ne s'améliorent pas significativement au HL-LHC à cause de l'apparition d'événements de bruit de fond venant de l'empilement en quantité non négligeable (de 20 à 50 selon les cas). Cependant, une ré-optimisation des coupures pourraient améliorer la situation en augmentant le taux de rejet associé. Par ailleurs, il n'est pas encore établi que les détecteurs de protons à l'avant pourront supporter de telles conditions de *pileup* ($<\mu> = 200$), ni que le réglage des optiques du

HL-LHC permettront une bonne acceptance à l'avant. Les résultats de l'étude plaident donc en faveur de la réalisation d'une telle mesure pendant les *runs 2* et *3* du LHC.

Aucune limite directe sur ces couplages n'existant à l'heure actuelle (voir section 2.4),¹²¹ cette mesure constituerait un pas important permettant d'observer sous un angle différent la dynamique de la brisure de symétrie électrofaible (voir section 2.2), dans la prolongation des mesures de précision électrofaibles débutées au LEP (voir section 2.4.2).

Il est difficile de comparer directement les résultats obtenus ici avec les projections de sensibilité des recherches directes faites au LHC sans protons intacts dans la mesure où ces dernières sont la plupart du temps très dépendantes des modèles de référence utilisés. Néanmoins, il est intéressant de mentionner les deux points généraux suivants :

- ◊ Dans le cas de recherches de nouveaux phénomènes dans un spectre inclusif, celles-ci sont souvent basées sur des signaux de nouvelle physique correspondant à des résonances fines car ils présentent localement le meilleur rapport signal sur bruit [198, 315]. Des recherches de déviations fines sur des intervalles plus larges peuvent également être effectuées mais perdent significativement en sensibilité, à l'image des limites actuelles sur les plus petites valeurs de κ dans les modèles de dimensions supplémentaires (voir section 19.1.3 et [198, 315]). Au contraire, on s'attend à ce que l'étude du canal di-photon exclusif permette d'obtenir des limites du même ordre de grandeur pour une très large variété de signaux étant donné que le bruit de fond attendu est nul dans l'ensemble de la région de signal considérée. Le bruit de fond nul permet également de s'affranchir du *look-elsewhere effect* [316], qui réduit fortement la sensibilité des recherches standards basées sur la reconnaissance de signaux particuliers (différence entre significativité locale et globale).
- ◊ Le canal exclusif $pp \rightarrow p + W^+W^- + p$ a été étudié sans la détection des protons intacts par CMS à $\sqrt{s} = 7$ TeV et permet déjà d'obtenir la meilleure limite actuelle sur les couplages quartiques anormaux $\gamma\gamma WW$ [271] (voir section 2.4.2, figure 11). Cependant, la réalisation d'une telle mesure à plus haut taux d'empilement sans détection des protons intacts est très difficile à cause de la baisse en efficacité des coupures d'exclusivité (voir section 20.2.2), ce qui rend la projection des sensibilités correspondantes à $\sqrt{s} = 8$ TeV et 13 TeV incertaine.¹²² D'autre part, la mesure à $\sqrt{s} = 7$ TeV est limitée par les incertitudes sur le bruit de fond, qui serait sous un bien meilleur contrôle si l'on détectait les protons intacts à l'avant grâce aux mesures de ξ des stations à l'avant (voir section 21.4.3). Cela permettrait également d'exclure de l'échantillon les états finaux avec protons dissociés (voir section 20.2). Il est raisonnable de penser qu'il en serait de même pour l'état final di-photon et les couplages anormaux à 4γ . La situation pourrait même être meilleure, les di-photons bénéficiant d'une meilleure résolution en masse que

¹²¹Des limites indirectes peuvent être calculées sur la base de certaines mesures de précision QED comme la mesure $g - 2$ (voir section 2.4) [63]. Comme nous le verrons dans la section 22.3, celles-ci ne sont en général pas compétitives avec la mesure proposée.

¹²²La mesure de ce processus à $\sqrt{s} = 13$ TeV avec détection des protons intacts est une des mesures phares prévues par ATLAS-AFP et CT-PPS. Les sensibilités attendues vis-à-vis des couplages anormaux $\gamma\gamma WW$ sont détaillées par exemple dans [75].

les paires de $W^\pm$, permettant une meilleure identification entre mesures centrales ($m_{\gamma\gamma}$, $y_{\gamma\gamma}$) et à l'avant (m_{pp}^{miss} , y_{pp}).

Dans la section 22.3, les limites directes obtenues sur l'existence de nouvelles particules chargées sont présentées.

22.3 Limites génériques sur l'existence de nouvelles particules chargées

Les plans d'exclusion en masse-charge effective obtenus sur l'existence de nouvelles particules chargées dans le cas de vecteurs et de fermions sont présentés respectivement dans la figure 206, gauche et droite. Les résultats correspondant à des excès de 5σ (région jaune), 3σ (région verte) et 2σ (région rouge, environ équivalente à un intervalle de confiance de 95% dans le cadre de l'approximation gaussienne) sont rapportés. Les signaux nominaux sont basés sur le calcul complet des amplitudes présenté dans la section 19.2. Comme attendu, les sensibilités deviennent semblables à celle obtenues à partir de la théorie effective à haute énergie, visible sur la figure 206 par un trait pointillé noir dans le cas d'un excès à 5σ . Celle-ci est en principe valide au moins pour les masses supérieures à 1 TeV dans le cadre de nouvelles particules chargées (voir section 19.1.4). En réalité, elle semble valide également pour les masses comprises entre 700 GeV et 1 TeV d'après la figure 206. Cela confirme le caractère conservateur des conditions de validité adoptées dans la section 19.1.4. Par ailleurs, les sensibilités à 5σ en termes de charge effective Q_{eff} pour plusieurs valeurs de masse sont résumées dans la tableau 42. Aucune sélection sur le statut de conversion des photons n'a été demandé pour produire ces résultats.

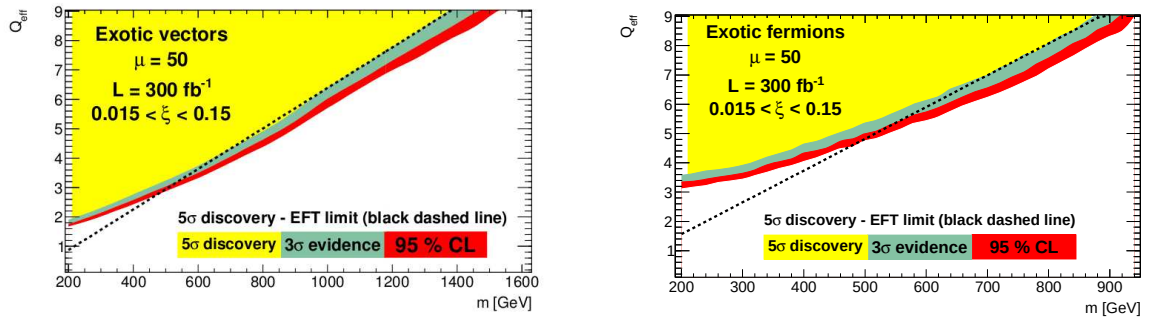


FIG. 206 – Sensibilités finales au LHC à $\sqrt{s} = 14$ TeV correspondant à l'observation d'un excès à 5, 3 et 2 σ (ce dernier équivalent environ à un intervalle de confiance de 95% dans l'approximation gaussienne) dans le cas de l'existence de nouveaux bosons vecteurs (gauche) et fermions (droite) chargés. La luminosité et l'empilement considérés sont respectivement de 300 fb^{-1} et $\langle \mu \rangle = 50$. Les résultats sont obtenus à partir du calcul complet des amplitudes et recoupés à haute énergie avec la théorie effective sans facteur de forme.

Les vecteurs (fermions) avec une charge effective de $Q_{\text{eff}} = 5$ peuvent être découverts pour des masses inférieures à 810 (460) GeV, confirmant l'augmentation des sensibilités avec le spin (voir section 19.1.2). La sensibilité à 5σ dans la limite à haute énergie correspond à un rapport $Q_{\text{eff}}/m > 0.0065$ (0.01) GeV^{-1} pour les vecteurs (fermions).

Mass (GeV)	300	600	900	1200
Q_{eff} (vector)	2.3	3.7	5.6	7.8
Q_{eff} (fermion)	3.9	6.2	9.1	-

TAB. 42 – Limites à 5σ concernant la charge effective de nouveaux bosons vecteurs et fermions pour différentes masses. Le scénario considéré est similaire à celui de la figure 206.

Bien que la plupart des limites actuelles sur les modèles de Higgs composite soient typiquement légèrement meilleures [33], ces résultats présentent l'avantage d'être totalement indépendants des différents modèles et donc fournissent des contraintes complémentaires intéressantes.

Les sensibilités obtenues ont été comparées aux limites indirectes pouvant être extraites à partir de certaines mesures de précision actuelles comme celle du moment magnétique du muon. En effet, de nouvelles particules chargées aurait un certain effet sur celui-ci, comme le montre les diagrammes de la figure 207. Compte tenu de l'incertitude expérimentale actuelle sur la mesure $g - 2$ [63], la sensibilité de celle-ci vis-à-vis de l'existence de nouvelles particules chargées a été estimée à $m/Q_{\text{eff}} \simeq 5$ GeV [57]. Malgré la précision impressionnante de la mesure $g - 2$, ces résultats se situent entre un et deux ordres de grandeur en dessous des résultats résumés dans la figure 206 et décrits dans le paragraphe précédent.

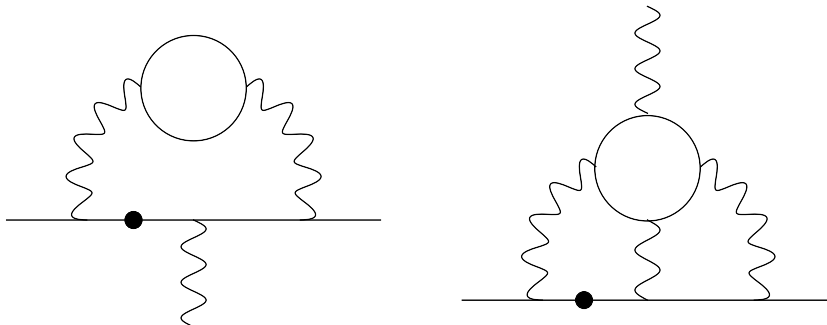


FIG. 207 – Contributions à deux (gauche) et trois boucles (droite) au moment magnétique gyromagnétique du muon. Le point représente l'insertion de la masse du muon et le cercle une particule chargée exotique.

Dans la section 22.4, les limites spécifiques à l'existence de nouvelles particules neutres de spin 0 et 2 sont détaillées. Les résultats utilisant les amplitudes complètes étant en cours de production, uniquement les limites obtenus dans le cadre de la théorie effective sont présentés.

22.4 Limites sur l'existence de particules neutres de spin 0/2

Dans le cas de contributions génériques aux couplages à quatre photons issues d'une nouvelle particule neutre X de spin 0 ou 2, il est nécessaire d'introduire un paramètre supplémentaire correspondant au couplage non-renormalisable $\gamma\gamma X$ noté f_S (voir section 19.1.3). Ainsi, dans le cadre de la relation simplifiée décrite dans la section 19.1.3,

la sensibilité en termes de (ζ_1, ζ_2) peut être retranscrite en termes de (f_S, m) , où m est la masse de la nouvelle particule neutre. Les limites à 5 et 2 σ (cette dernière correspondant environ à un intervalle de confiance de 95% dans l'approximation gaussienne) correspondent alors à des fonctions puissances, représentées dans la figure 208.

Les valeurs de masses et de couplages apparaissant dans la figure 208 démontrent la sensibilité du canal di-photon exclusif à des modèles de nouvelle physique situés à l'échelle de plusieurs TeV, y compris pour les grands f_S (ie. petit κ , voir section 19.1.3 et 22.2). Les sensibilités sont alors décroissantes avec le spin, comme déjà mentionné dans 19.1.3.

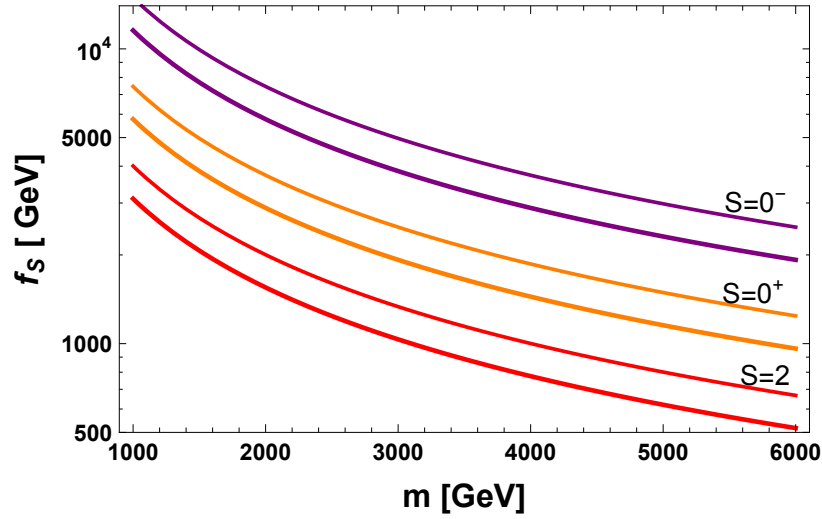


FIG. 208 – Sensibilités à 5 σ (lignes épaisses) et 2 σ (lignes fines) dans le cas d'un signal générique issu d'une nouvelle particule neutre générée à l'aide de la théorie effective et transposée en termes de (f_S, m) en utilisant les résultats de la section 19.1.3. Le cas des spins 0^- (violet), 0^+ (orange) et 2 (rouge) est représenté. Aucun facteur de forme n'est appliqué dans la théorie effective et aucune sélection n'est effectuée sur le statut de conversion des photons. 300 fb $^{-1}$ de données à $\sqrt{s} = 14$ TeV et $\langle \mu \rangle = 50$ sont considérés.

Finalement, le plan d'exclusion en (ζ_1, ζ_2) , obtenu en utilisant le paramètre ζ et l'équation 93 de la section 19.1.1 est représenté dans la figure 209. Les régions jaunes, grises et rouges correspondent respectivement à la mise en évidence expérimentale d'un excès à 5, 3 et 2 σ (cette dernière correspondant environ à un intervalle de confiance de 95% dans l'approximation gaussienne), tandis que la région blanche délimite la zone encore inaccessible. Les limites sont calculées pour 300 fb $^{-1}$ de données à $\sqrt{s} = 14$ TeV et $\langle \mu \rangle = 50$ avec aucune sélection sur le statut de conversion des photons.

L'emplacement de certains modèles particuliers de référence pris en exemple dans les sections 19.1.2 et 19.1.3 est également mentionné dans le plan (ζ_1, ζ_2) de la figure 209 en utilisant la théorie effective sans facteur de forme :

- ◊ les signaux correspondant à une nouvelle particule chargée de spin 1/2 et 1, charge effective $Q_{\text{eff}} = 3$ et masse $m = 1$ TeV sont représentés respectivement par les points verts et bleus.

- ◇ les signaux correspondant à un graviton de Kaluza-Klein¹²³ de masse et couplage $m_{\text{KK}} = 3 \text{ TeV}$, $\kappa = 2$ et à un dilaton SIHD de masse $m_\varphi = 3 \text{ TeV}$ fortement couplé à un photon composite¹²⁴ sont représentés respectivement en rouge et orange. Le signal SIHD présente la particularité d'avoir ζ_2 rigoureusement égal à 0.

Les limites de sensibilité à 5 et 2σ ($\simeq 95\%$ CL) sur la masse d'un graviton de Kaluza-Klein et d'un SIHD des scénarios évoqués ci-dessus sont respectivement $m_{\text{KK}} = 5240 \text{ GeV}$ (6230 GeV) et $m_\varphi = 3960 \text{ GeV}$ (4710 GeV). Celles-ci sont largement compétitives par rapport aux recherches directes menées dans les expériences ATLAS et CMS [198, 315, 317], en particulier dans le cas de résonances larges (voir section 22.2).

Dans la section 22.5, une conclusion générale sur l'étude de la sensibilité des couplages anormaux à quatre photons au LHC est donnée. D'autres canaux potentiellement sensibles sont ensuite discutés, préparant la voie pour des études futures. Une parenthèse est également faite sur le rôle que pourrait jouer les couplages anormaux à 4γ en cas de confirmation du léger excès présent dans le spectre en masse des di-photons à 13 TeV , rapporté à la fois par les expériences ATLAS et CMS dans les données de 2015 [8, 9].

22.5 Conclusion et perspectives

Dans cette partie, les sensibilités des expériences ATLAS-AFP et CT-PPS aux couplages à quatre photons ont été évaluées [56, 57]. Ce travail constitue le projet phénoménologique principal effectué pendant les années de thèse. La très grande suppression des contributions du modèle standard et le contrôle impressionnant du bruit de fond et de l'empilement grâce aux détections des protons intacts à l'avant permet d'obtenir d'excellentes sensibilités pour la découverte de couplages anormaux à quatre photons au LHC ($> 0.01 \text{ TeV}^{-4}$). En revanche, les contributions à quatre photons prévues par le modèle standard à basse énergie restent hors de portée d'une observation directe dans le cadre de collisions proton-proton.

Aucune limite directe n'existe sur les couplages anormaux à 4γ à l'heure actuelle. En plus de constituer un sujet d'étude particulièrement intéressant dans le cadre de la théorie des champs car totalement issue de la mécanique quantique, la détection éventuelle de couplages anormaux à haute énergie bouleverserait profondément notre compréhension actuelle de la brisure de symétrie électrofaible. Leur étude constitue un prolongement des mesures de précision électrofaibles débutées au LEP [318] et permet de placer des limites génériques sur l'existence de nouvelles particules chargées (vecteurs, fermions) et de résonances neutres (scalaire, pseudo-scalaire, spin 2).

Les résultats obtenus ont été démontrés comme très compétitifs avec les limites actuelles, y compris celles issues de certaines recherches directes, en particulier dans le cas de l'existence de dimensions supplémentaires de l'espace faiblement couplées (signal de type continuum). L'existence d'une grande variété de nouveaux phénomènes peut être testée jusqu'à une échelle en énergie correspondant à plusieurs TeV . Une théorie effective ainsi que le calcul complet des amplitudes dans différents cas particuliers ont été

¹²³Cet ordre de grandeur est valide dans le cas où le modèle standard se situe sur la membrane infra-rouge du modèle complet, voir section 19.1.3.

¹²⁴voir section 19.1.3.

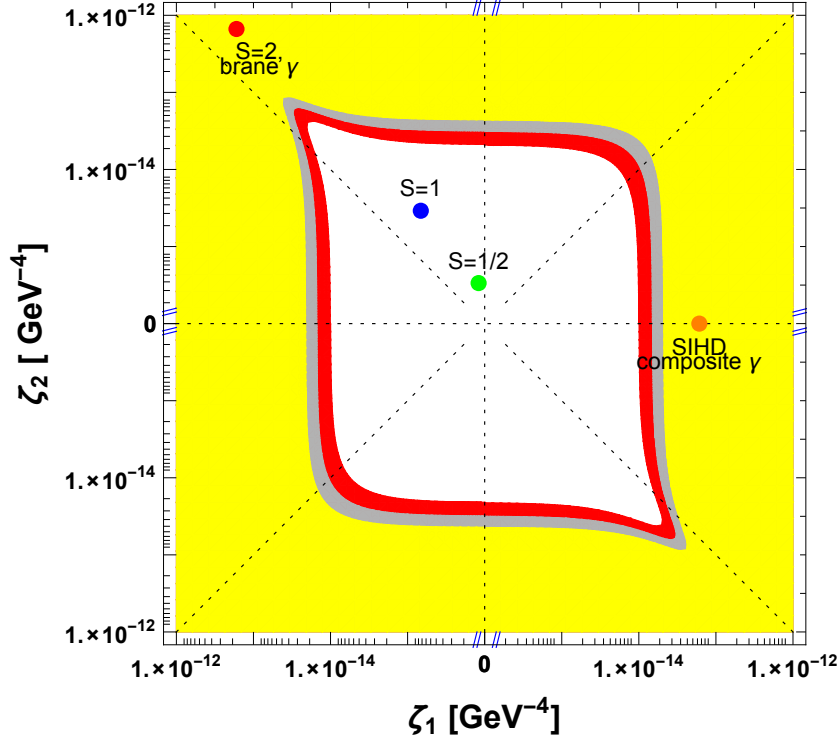


FIG. 209 – Sensibilités expérimentales et emplacements de différents modèles de nouvelle physique dans le plan des couplages anormaux (ζ_1, ζ_2) . Les axes suivent une échelle logarithmique pour $|\zeta_i| \in [10^{-12}, 10^{-16}]$. Les régions jaunes, grises et rouges correspondent respectivement à une mise en évidence expérimentale à 5, 3 et 2 σ (cette dernière correspondant environ à un intervalle de confiance de 95% dans l'approximation gaussienne). La région blanche reste inaccessible. Les limites sont calculées pour 300 fb $^{-1}$ de données à $\sqrt{s} = 14$ TeV et $\langle \mu \rangle = 50$ avec aucune sélection sur le statut de conversion des photons. Certains signaux de références sont calculés à partir de la théorie effective sans facteur de forme et placés dans le plan pour illustration (voir texte).

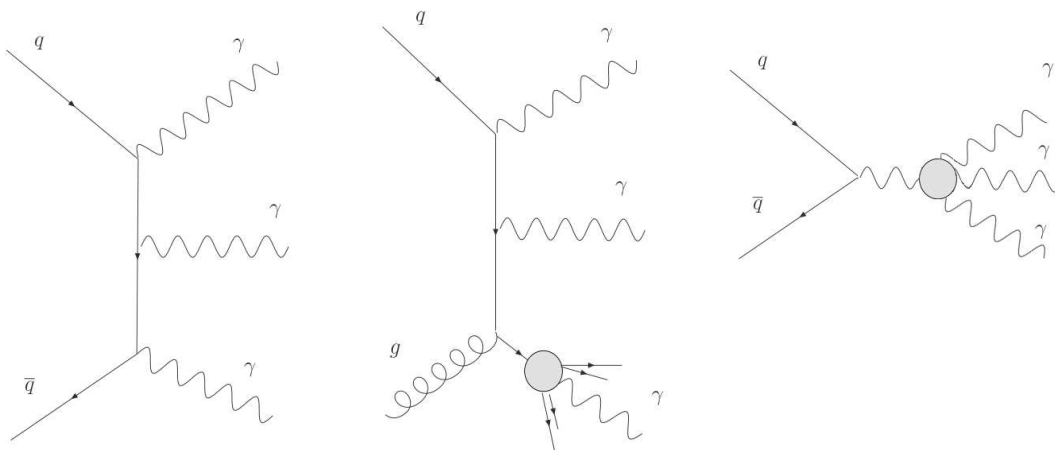


FIG. 210 – Diagramme de production de trois photons à l'ordre le plus bas de la théorie des perturbations au LHC (gauche, production 100% directe, centre, 1-fragmentation) et diagramme supplémentaire éventuel dû à l'existence de couplages anormaux (droite).

exploités afin d'obtenir des prédictions valides pour l'ensemble de l'espace des paramètres au LHC à $\sqrt{s} = 14$ TeV.

Pour ce faire, le canal di-photon exclusif avec protons intacts a été étudié. Celui-ci présente l'avantage de bénéficier de prédictions fiables pour le signal et les bruits de fond irréductibles grâce à l'utilisation de l'approximation des photons équivalents (EPA) et du calcul KMR, ce qui permet notamment de s'affranchir de la fonction de structure du photon dans le proton à haute énergie (PDFs), qui est mal déterminée [319]. Son caractère exclusif permet également un bien meilleur contrôle du bruit de fond, souvent limitant dans ce genre de mesure. L'étude du processus similaire correspondant à la production exclusive de paires de bosons W permet déjà de placer les contraintes les plus compétitives sur les couplages anormaux $\gamma\gamma WW$, démontrant l'excellent potentiel de ce type de canal [271].

Il pourrait également être intéressant d'évaluer la sensibilité aux couplages à 4γ du canal correspondant à la production inclusive de trois photons (sans protons intacts), voir figure 210. La présence du troisième photon dans l'état final devrait diminuer le bruit de fond, ce qui pourrait permettre d'obtenir des limites intéressantes. D'autre part, les collisions Plomb-Plomb fournissent un flux élastique photon-photon très important grâce à leur numéro atomique élevé, ce qui mène à un terrain privilégié pour l'observation de la diffusion $\gamma\gamma$ du modèle standard à basse énergie, encore jamais observée [59]. Malgré la faible luminosité intégrée attendue et la suppression du flux photon-photon à haute masse à cause du paramètre d'impact plus important dans les collisions de Plomb, il n'est pas totalement exclu d'obtenir des résultats intéressants également pour les plus hautes énergies.

Par ailleurs, dans le cas de l'existence de nouvelles particules chargées, les couplages à quatre photons semblent montrer une sensibilité croissante avec le spin des nouvelles particules (voir section 19.1.2), ce qui offre a priori une excellente opportunité pour tester les extensions du modèle standard prédisant l'existence de particules avec des spins exotiques, comme les modèles introduisant un nouveau secteur fort ou ceux impliquant

des cordes de basses énergies [276]. Cependant, les états de spin élevé génèrent de façon implicite des interactions d'ordre supérieur qui nécessitent des développements théoriques importants afin de pouvoir être décrits correctement dans le cadre d'une théorie effective. Des études sont en cours mais ne devraient pas être achevées avant plusieurs années.

Finalement, le canal di-photon exclusif pourrait jouer un rôle particulier en cas de confirmation du léger excès dans le spectre di-photon, rapporté à la fois par les collaborations ATLAS et CMS dans les premières données à $\sqrt{s} = 13$ TeV prises en 2015 [8, 9].¹²⁵ En effet, la légère tension entre les données à 13 et 8 TeV pourrait être le signe de modes de production électrofaible dominant (photon-photon, fusions de bosons vecteurs WW , ZZ , $Z\gamma$, ...) plutôt que parton-parton, alors que les analyses d'ATLAS et CMS mènent leurs études uniquement dans le cadre de nouvelles productions à partir de quarks et de gluons [320–325].

Dans ce cas, l'excès correspondrait alors en pratique majoritairement à des couplages anormaux à quatre photons et rentrerait dans le cadre de l'étude à haute énergie présentée dans cette partie. Par ailleurs, les couplages de la nouvelle résonance aux autres bosons électrofaibles pourraient être étudiés via les différents canaux exclusifs ($\gamma\gamma$, ZZ , WW , $Z\gamma$) bien plus précisément que dans le cadre des recherches classiques grâce à la détection des protons intacts [74, 75, 320, 321]. En effet, comme nous l'avons vu dans la section 20, l'unique mode de production exclusif à haute énergie est EPA/photon-photon, ce qui permet de mesurer les différents couplages ($\gamma\gamma\gamma\gamma$, $\gamma\gamma ZZ$, $\gamma\gamma WW$, $\gamma\gamma Z$) indépendamment, avec une très bonne précision et sans hypothèse supplémentaire sur la nouvelle physique sous-jacente.

Dans la partie VI, les résultats de tests de la carte de lecture SAMPIC pour des mesures de temps à la picoseconde sont détaillés. Ce travail a constitué ma tâche technique de qualification afin de devenir signataire officiel des articles de la collaboration ATLAS.

¹²⁵ J'ai participé à deux études effectuées par la collaboration ATLAS dans le cadre de cette analyse : la première traite d'un certain nombre de vérifications et de comparaisons faites avec les données à $\sqrt{s} = 8$ TeV [275] et la seconde consiste en différentes mesures de bruit de fond à haute masse réalisées à $\sqrt{s} = 13$ TeV [245].

Sixième partie

Test de la carte de lecture SAMPIC pour des mesures de temps à la picoseconde

Sommaire

23	Introduction	422
23.1	Considérations générales	422
23.2	La puce SAMPIC	423
23.2.1	TDCs	423
23.2.2	<i>Waveform</i> TDC	424
23.2.3	Architecture de la puce SAMPIC	425
23.3	Intégration et pilotage	426
23.4	Le traitement des données	430
23.4.1	Définition du format de données	430
23.4.2	Reconstruction du temps d'arrivée	430
24	Résultats des mesures expérimentales	434
24.1	Tests avec des signaux issus de générateurs	434
24.2	Tests avec des signaux issus de détecteurs ultra-rapides de silicium	438
24.2.1	Description générale	438
24.2.2	Premier montage, incluant deux amplificateurs à large bande	439
24.2.3	Second montage, incluant un amplificateur de charge rapide	442
25	Conclusion et perspectives	446

23 Introduction

23.1 Considérations générales

Au LHC, il est possible d'étudier des événements très particuliers qui, bien que mettant en jeu de très hautes énergies, laissent les protons intacts après collision (voir section 6). Ces événements sont appelés “diffractifs” lorsqu'ils sont initiés par des gluons. Ils peuvent également être initiés par des photons (voir partie V). Leur observation est motivée par une meilleure compréhension de la QCD (voir section 3) [163, 326] et la recherche de nouvelle physique au-delà du modèle standard via des couplages anormaux entre photons, bosons $W^\pm$ et bosons Z^0 (voir section 2.4 et partie V) [56, 57, 74, 75, 320, 321]. La recherche de nouvelle physique via de tels processus présenterait un intérêt particulier si l'excès modéré observé à la fois par les collaborations CMS et ATLAS dans le canal di-photon autour de 750 GeV est confirmé (voir section 22.5) [320–325].

Les protons intacts diffractés à petit angle peuvent être mesurés par des détecteurs dédiés intégrés dans des sections amovibles des chambres à vide (secondaire) du LHC en forme de pots romains (*roman pots*). Celles-ci sont placées à quelques millimètres du faisceau et à plus de 200 m de chaque côté du point d'interaction principal des détecteurs ATLAS et CMS. L'objectif étant de mesurer des événements rares, une opération de ces détecteurs à haute luminosité est prévue, c'est-à-dire avec en moyenne 40 à 70 interactions à chaque croisement de faisceaux au point d'interaction principal. Ces chiffres correspondent aux prévisions d'empilement pour la seconde prise de données du LHC pendant la période 2015-2018 (voir section 4.4). Les projets respectifs d'ATLAS et CMS-TOTEM pour l'installation de tels détecteurs sont AFP (*ATLAS forward proton detector*) [171] et CT-PPS (*CMS-TOTEM precision proton spectrometer*) [73]. Ceux-ci ont déjà été mentionnés dans la section 6.2 et au cours de la partie V.

La mesure des temps d'arrivée des protons est alors très utile afin de déterminer si ceux-ci proviennent de l'interaction principale de haute énergie ou de vertex secondaires (empilement). En effet, dans le cas de deux protons intacts détectés en coïncidence et voyageant à une vitesse proche de celle de la lumière, une résolution temporelle de 10 ps RMS permet de séparer les différents vertex jusqu'à un espacement de 2 mm environ, uniquement à l'aide des protons. Pour un empilement $\langle \mu \rangle = 40$, cette mesure additionnelle mène à une réduction du bruit de fond par un facteur proche de 40 lorsqu'elle est combinée avec les mesures du détecteur central [180].

La détection précise du temps de vol a également de nombreuses autres applications potentielles, par exemple en imagerie médicale. En effet, une résolution temporelle de 10 ps RMS sur les mesures de temps permet d'améliorer considérablement la reconstruction des tomographies par émission de positrons (TEP) grâce à l'augmentation du taux de rejet des artefacts de mesure (voir figure 211). Les données collectées seraient alors beaucoup moins importantes et si un signal sur bruit suffisant est atteint, une reconstruction directement en ligne pourrait même être envisageable, sans la nécessité d'un traitement ultérieur comme c'est le cas actuellement [327].

La puce SAMPIC (*sampler for picosecond time*) [178] est un circuit intégré à application spécifique (ASIC pour *application-specific integrated circuit*) qui a été conçu très récemment (première version sortie en 2013) pour lire des détecteurs de temps de vol et effectuer des mesures avec une résolution temporelle de 5 ps RMS. La mise en œuvre

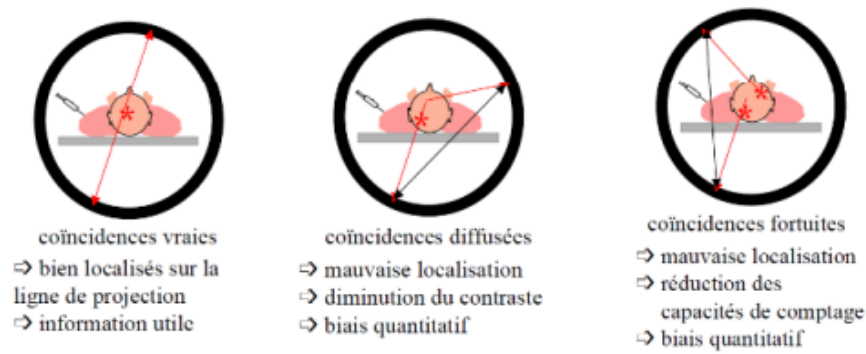


FIG. 211 – Les différents types de coïncidences dans le cadre d’une tomographie par émission de positrons (TEP). Une résolution en temps optimale permet de limiter les coïncidences diffusées et fortuites. [328]

des premiers tests de cette puce a constitué ma tâche technique dans la collaboration ATLAS et a été menée pendant les années de thèse.

Dans un premier temps, des mesures de résolution temporelle ont été réalisées avec des signaux gaussiens issus de générateur électronique. Finalement, des mesures similaires ont pu être réalisées avec des signaux venant de détecteurs au silicium atteints par des impulsions laser infrarouges.¹²⁶ L’analyse a nécessité le développement d’un logiciel particulier implémenté pendant les années de thèse et utilisé aujourd’hui par plusieurs groupes manipulant les données de SAMPIC.¹²⁷ Par ailleurs, de nombreuses améliorations du système d’acquisition et en particulier du logiciel de contrôle ont pu être apportées grâce aux différents retours d’expérience obtenus pendant les tests. Un article concernant les résultats principaux obtenus dans le cadre de ma tâche technique a été soumis au printemps 2016 à la revue *Nuclear Instruments and Methods in Physics Research* [329].

Dans la section 23.2, un bref descriptif de la puce SAMPIC est donné, complétant les informations disponibles dans la section 6.3. L’accent est mis sur les différences fondamentales de SAMPIC par rapport à des lecteurs de temps plus classiques. Dans la section 23.3, l’intégration de la puce et le logiciel de pilotage sont introduits, incluant notamment les différentes étapes de calibration. Finalement, le logiciel d’analyse implémenté pendant les années de thèse pour les mesures de résolution est détaillé dans la section 23.4. Les résultats des tests seront rapportés dans la section 24 et la partie VI s’achèvera sur une conclusion générale dans la section 25.

23.2 La puce SAMPIC

23.2.1 TDCs

La grande majorité des systèmes électroniques de précision pour la mesure de temps repose actuellement sur des TDCs (pour *time to digital converters*). Ceux-ci ont déjà

¹²⁶ Je tiens à remercier particulièrement le Prof. Nicolo Cartiglia pour son aide concernant les détecteurs au silicium.

¹²⁷ <https://svnweb.cern.ch/cern/wsvn/atlas-msaimper/SAMPICFastAnalysis>

été présentés rapidement dans la section 6.3.1. Les TDCs reposent sur un compteur qui fournit une mesure de temps grossière associée à une ligne à retard prédéfinie pour la mesure fine (*delay locked loops* ou DLL). La résolution finale de la mesure est alors souvent limitée par le pas de la DLL. Les puces les plus performantes arrivent néanmoins à une résolution d'environ 20 ps RMS. De nouveaux développements ayant lieu au CERN visent à descendre à 5 ps RMS [176].

Par ailleurs, les signaux en sortie d'un détecteur de particules sont analogiques alors qu'un TDC a besoin d'entrées numériques. Il est donc nécessaire d'opérer une conversion numérique avant la mesure du temps. Cette conversion est typiquement effectuée par un discriminateur rapide. Cependant, l'étalement en amplitude du signal (qui n'est pas une distribution delta de Dirac) induit souvent un effet dit de *time walk*, illustré dans la figure 212, même si celui-ci est en principe corrigé a posteriori des acquisitions. De plus, le discriminateur introduit un gigue (*jitter*) supplémentaire à ajouter à celui du TDC, sa précision affectant directement la mesure.

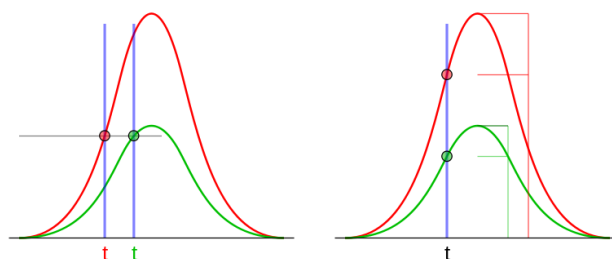


FIG. 212 – Illustration de l'effet de *time walk* (gauche) : dans le cas d'une discrimination basée sur un seuil de déclenchement fixe, le temps mesuré variera légèrement en fonction de l'amplitude du signal, engendrant une dégradation de la résolution de la mesure. La puce SAMPIC effectue un échantillonnage rapide du signal permettant d'implémenter des algorithmes plus complexes, basés par exemple sur des seuils normalisés à l'amplitude du signal (droite). Des détails supplémentaires sur ce point seront donnés dans la section 23.4 [330].

23.2.2 *Waveform* TDC

Afin de pallier à ces limitations, une nouvelle approche est adoptée dans la puce SAMPIC, basée sur le principe novateur de *waveform TDC*. Dans ce type de système, une mémoire analogique est ajoutée en parallèle de la ligne à retard. Un échantillonnage analogique rapide du signal d'entrée est alors effectué de manière continue. Un discriminateur est présent pour déclencher l'acquisition mais n'est plus critique pour la mesure de temps. En effet, plusieurs échantillons de part et d'autre de la position du signal de déclenchement sont enregistrés (64 dans SAMPIC) et exploitables par la suite. Ceux-ci permettent par ailleurs d'améliorer considérablement la résolution grâce à la possibilité d'un traitement hors-ligne plus poussé (interpolations, traitement de signal, ...).

Dans un *waveform TDC* comme SAMPIC, la mesure finale de temps est donnée par l'association de trois systèmes de mesure en parallèle :

- ◇ **Compteur grossier** : à la suite du déclenchement d'une acquisition, l'état d'un compteur basé sur un code Gray sur 12 bits connecté à l'horloge de la DLL est enregistré et écrit dans le fichier de sortie. Il présente un pas d'environ 6 ns.
- ◇ **Pas des lignes à retard** : la position du trigger parmi les 64 pas de la DLL est enregistrée. Elle peut alors être utilisée pour définir une région d'intérêt. Chaque pas est espacé d'environ 150 ps.
- ◇ **Traitement hors-ligne** : le contenu de la mémoire analogique est enregistré et ordonné en fonction de la position du *trigger*. Elle permet via une interpolation ou d'autres algorithmes de traitement du signal d'atteindre une résolution de l'ordre de quelques picosecondes RMS sur la mesure finale en temps.

23.2.3 Architecture de la puce SAMPIC

SAMPIC est un ASIC mettant en œuvre le principe de *waveform TDC* pour 16 canaux indépendants. Une représentation simplifiée d'un canal de SAMPIC est montrée dans la figure 213. Chaque canal présente une résistance de 50 Ω et est couplé en courant alternatif (*AC coupled*). Le signal d'entrée est envoyé à la fois dans :

- ◇ **64 cellules analogiques d'échantillonnage** : l'échantillonnage rapide du signal est basé sur 64 cellules et contrôlé par une DLL commune à toute la puce. Dans le cas du déclenchement d'une acquisition, les 64 échantillons sont figés et numérisés en parallèle via un convertisseur analogique/numérique (ADC, pour *analog-to-digital converter*) de Wilkinson, à nouveau propre à chaque canal. La conversion peut être configurée par l'utilisateur pour s'exécuter sur 8, 9, 10 ou 11 bits.
- ◇ **Discriminateur** : l'échantillonnage de chaque canal est commandé soit par le discriminateur propre au canal, soit par un signal de déclenchement externe. Un retard variable peut être ajouté à la sortie du discriminateur afin de garder plus d'échantillons après le passage du seuil. Ce dernier est ajustable indépendamment pour chaque canal.

À l'issue du déclenchement de l'acquisition dans un canal, l'état du compteur Gray, la position du *trigger* parmi les éléments de la DLL et le contenu des 64 cellules analogiques d'échantillonnage sont figées. Ces informations sont ensuite combinées pendant la lecture des données, après numérisation du contenu des cellules analogiques.

Le fait d'utiliser un ADC basé sur une conversion de Wilkinson permet une très bonne linéarité tout en gardant une consommation électrique faible. Cependant, il induit un temps mort conséquent par canal à cause de la durée de conversion. Celle-ci varie en fonction de la précision voulue sur la conversion (nombre de bits) et va de moins de 200 ns pour une conversion sur 8 bits à environ 1.3 μ s pour 11 bits. En pratique, la conversion peut être ajustée en fonction du rapport signal sur bruit afin de trouver le meilleur compromis entre temps mort et précision.

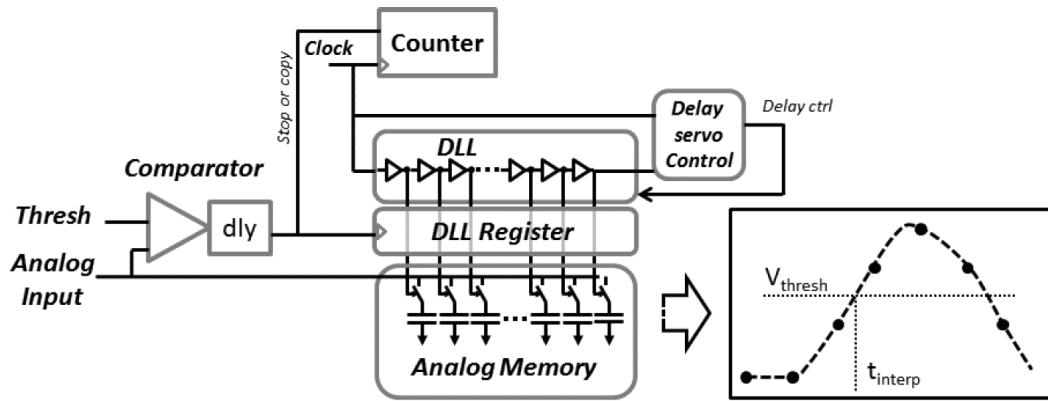


FIG. 213 – Représentation schématique d’un canal de la puce SAMPIC : un échantillonnage du signal sur 64 cellules analogiques associé à une ligne à retard est effectué en permanence. L’acquisition est commandée par un discriminateur ou un signal externe. En cas de déclenchement, deux mesures de temps sont enregistrées : le temps grossier (compteur Gray) et la position du *trigger* dans la DLL. Par ailleurs, les mesures analogiques du signal sont figées et converties en données numériques, utilisables par la suite pour atteindre une résolution de l’ordre de quelques picosecondes RMS via des algorithmes de traitement du signal.

Le temps mort est déjà réduit d’environ un facteur 2 dans la dernière version de la puce SAMPIC,¹²⁸ grâce à l’implémentation d’une nouvelle fonctionnalité appelée méthode “ping-pong” permettant d’utiliser deux canaux de SAMPIC pour enregistrer deux événements consécutifs arrivant sur un même canal de détecteur. Cette méthode a déjà été mise en œuvre avec succès pour une paire de puces ARS dans le cadre de l’expérience ANTARES [331]. D’autres techniques pour améliorer le temps mort sont en cours d’étude pour les prochaines versions. Un résumé général des caractéristiques de SAMPIC est présenté dans le tableau 43.

Un des avantages majeurs de SAMPIC vient également de son prix (environ 10 euros par canal), très raisonnable comparé aux technologies concurrentes, où le coût par canal est généralement de l’ordre de quelques milliers d’euros pour une précision similaire (modules commercialisés). Cela rend en principe l’utilisation de SAMPIC possible pour des expériences de très grande taille comme les détecteurs installés autour du LHC.

Dans la section 23.3, quelques détails sur l’intégration de SAMPIC et son logiciel de pilotage sont donnés. En particulier, les différentes calibrations recommandées au préalable de l’utilisation sont précisées.

23.3 Intégration et pilotage

La puce SAMPIC est livrée dans un boîtier incluant une ou deux carte filles, intégrant chacune une puce. Les cartes filles sont elles-mêmes reliées à une carte mère permettant le contrôle général du système ainsi que la lecture des données en sortie. Un logiciel contrôlant les paramètres d’acquisition et permettant la visualisation et la

¹²⁸sortie en décembre 2015.

	Performances mesurées	Commentaires
Technologie	AMS CMOS 0.18 μm	
Nombre de canaux	16	
Consommation électrique	180 mW (à partir des 1.8 V fournis)	
Profondeur du circuit à capacités commutées	64	
Bruit du discriminateur	2 mV	
Vitesse d'échantillonnage	1.6 - 8.4 GS/s (= milliard d'échantillons par seconde)	jusqu'à 10.2 GS/s sur les huit premiers canaux
Bande passante d'entrée à 3dB	1.6 GHz	
précision de l'ADC	8 à 11 bits	compromis entre temps mort et resolution
Bruit	< 1.3 mV RMS	
Gamme d'entrée dynamique	1 V (0.1 V à 1.1 V)	
Temps de conversion	200 ns sur 8 bits - 1.6 μs sur 11 bits	
Temps d'exécution du système de lecture	37.5 ns + 6.25 ns par échantillon à 160 MHz	Une lecture jusqu'à 200 Mhz est possible
Précision sur la détermination du temps avant correction	15 ps	
Précision sur la détermination du temps après correction	< 5 ps	voir section 23.3

TAB. 43 – Résumé des caractéristiques principales de la puce SAMPIC.

sauvegarde des données enregistrées par la puce est également disponible. Le logiciel de contrôle peut également procéder à quelques mesures en ligne basiques (charge, amplitude, différence de temps entre deux canaux, ...) afin de vérifier rapidement la bonne qualité de l'acquisition en cours.

Une photographie du système complet est montrée dans la figure 214. Il nécessite uniquement une alimentation électrique de 5 V et un ordinateur pour fonctionner. L'ordinateur peut être connecté au boîtier par USB 2.0, Ethernet Gigabit ou fibre optique. Par ailleurs, il est possible de synchroniser le dispositif avec une horloge externe, qui peut alors être utilisée pour le déclenchement ou au contraire pour mettre un veto sur le déclenchement. Le boîtier présente 16 ou 32 canaux d'entrée en fonction du nombre de cartes filles installées. Les entrées se font à l'aide de connecteurs MCX, qui ont l'avantage d'être très compacts.

Le logiciel de contrôle associé au dispositif est basé sur les bibliothèques LABWINDOWS¹²⁹ et fonctionne sur les plateformes Windows. Une version multi-plateformes est également en cours de développement. Une prise d'écran de l'interface graphique du logiciel est montrée dans la figure 215. Les principaux paramètres configurables via le logiciel sont :

◇ **Paramètres d'échantillonnage** : la fréquence d'échantillonnage, allant de 1.6 à 10.2 milliard d'échantillons par seconde (GS/s), la ligne de base, allant de 0.1 à 1.1 V et la gamme de conversion pour l'ADC (8, 9, 10 ou 11 bits).

◇ **Paramètres de déclenchement** : mode de déclenchement (externe ou basé

¹²⁹Documentation disponible en ligne : <http://www.ni.com/pdf/manuals/320682d.pdf>



FIG. 214 – Photographie du boîtier complet intégrant la puce SAMPIC, qui est le composant rectangulaire noir situé sur la face supérieure bien visible sur la photographie. Le boîtier inclut une ou deux cartes filles (ici une seule, correspondant à la carte placée sur le dessus) et une carte mère (carte du dessous) permettant le contrôle général du système ainsi que la lecture des données en sortie.

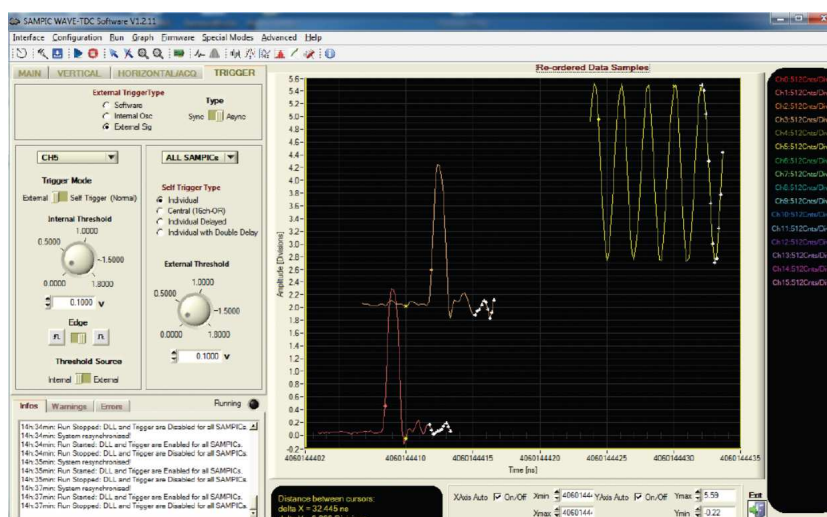


FIG. 215 – Prise d'écran dans l'interface graphique du logiciel de contrôle de SAMPIC. Celui-ci permet entre autres la configuration de la puce (panneau de gauche) et la visualisation de données échantillonnées à l'écran (panneau de droite).

sur le discriminateur du canal), seuils de déclenchement et polarité du signal (positif/négatif).

- ◇ **Paramètres d'acquisition** : format de données (ASCII ou binaire), taille maximum des fichiers d'acquisition, longueur de l'acquisition, etc.

Par ailleurs, en plus de la calibration des pieds-taux, deux types de corrections supplémentaires implémentées dans le logiciel de contrôle permettent à la puce d'atteindre ses performances optimales [178]. La première concerne les petites non-linéarités déformant les fonctions de transfert des ADCs. Celles-ci sont mesurées de façon statique dans chaque cellule en convertissant 20 différences de potentiel connues à l'avance réparties dans la plage de mesure générées directement à partir de la carte fille. Les écarts de linéarité sont ensuite ajustés par un modèle polynomial d'ordre 1 ou 2, utilisé par la suite lors de la prise de données en conditions réelles pour corriger les mesures.

La seconde correction est reliée aux petites non-uniformités présentes dans la ligne à retard contrôlant le remplissage des cellules analogiques. Ces composants étant utilisés pour l'acquisition du signal, les non-uniformités de la DLL affectent directement le temps entre les différents échantillons et donc la gigue de la puce. La procédure de calibration s'appliquant à la première version de SAMPIC (version utilisée pour les tests de la section 24) nécessite un générateur externe afin d'envoyer un sinus à haute fréquence très précis dans chaque canal. L'échantillonnage de celui-ci permet de corriger les petites inhomogénéités de différence de temps entre les cellules. Dans la dernière version de SAMPIC sortie tout récemment (voir section 23.2), le générateur de sinus est directement intégré dans la puce de sorte qu'aucun signal externe n'est nécessaire.

La calibration a besoin d'être répétée pour chaque fréquence d'échantillonnage et chaque configuration d'ADC. Le logiciel de contrôle enregistre alors les différentes corrections en fonction des paramètres de SAMPIC et corrige automatiquement les données lors des prochaines acquisitions. Il est important de souligner qu'avant calibration la puce permet déjà d'atteindre d'excellentes résolutions, de l'ordre de 15 ps RMS à 6.4 GS/s [178].

Finalement, le logiciel peut également être utilisé afin de visualiser des données pendant l'acquisition ou enregistrées au préalable. Il permet aussi de réaliser quelques mesures préliminaires simples comme une estimation du bruit dans un ou plusieurs canaux (*RMS noise*) ou une évaluation rapide de la différence de temps entre deux canaux. Ces fonctionnalités se sont avérées très utiles pour valider la calibration et préparer les tests détaillés dans la section 24.

Dans la section 23.4, le logiciel développé pendant les années de thèse pour l'analyse des données de SAMPIC est introduit. Celui-ci a entre autres servi à produire les résultats exposés dans la section 24. Il est utilisé aujourd'hui par plusieurs groupes manipulant les données de SAMPIC.

23.4 Le traitement des données

23.4.1 Définition du format de données

Les données échantillonnées par SAMPIC sont traitées par un réseau de portes programmables (FPGA, pour *field-programmable gate array*) situé sur la carte fille. Elles peuvent être enregistrées au format binaire ou ASCII selon les paramètres choisis via le logiciel de contrôle. Bien que des mesures préliminaires de différences de temps puissent déjà être effectuées avec ce logiciel, une analyse plus poussée nécessite un traitement hors-ligne. En effet, cela autorise une plus grande flexibilité permettant par exemple de comparer facilement différents algorithmes de reconstruction du temps et de tester plusieurs jeux de coupures de qualité pour éliminer les artefacts de mesure.

Le logiciel d'analyse hors-ligne est un code C++ basé sur l'environnement ROOT [160]. Tout d'abord, les différents signaux sont extraits à partir des données enregistrées à l'aide du logiciel de contrôle et stockés dans une structure de base de données (TTree). Comme c'est traditionnellement le cas en physique des hautes énergies, les données commencent par être organisées par "événement". Si le déclenchement est basé sur un signal externe, deux signaux ou plus sont considérés comme faisant partie du même événement lorsque les temps d'arrivée du signal externe et des signaux enregistrés sont inférieurs à un certain temps $\Delta t_{\max}$ après correction du délai de latence. Dans le cas contraire, tous les signaux ayant des différences de temps de départ inférieures à $\Delta t_{\max}$ sont considérés comme faisant partie du même événement. On parle alors de signaux en coïncidence.

$\Delta t_{\max}$ doit être fixé par l'utilisateur. Une petite valeur (ex : 20 ns) est généralement utilisée lorsqu'un taux de déclenchement élevé est attendu, tandis qu'une valeur plus grande est souvent préférable dans le cas contraire afin d'assurer une certaine tolérance, par exemple lors de l'utilisation de câbles de longueurs différentes pour les canaux d'entrée.

La ligne de base, *ie.* la différence de potentiel moyenne dans le canal lorsqu'un événement se produit, est calculée pour chaque signal à partir des n_b premiers points enregistrés, où n_b doit être spécifié par l'utilisateur. Elle est alors inscrite dans la base de données en même temps que le signal en question. D'autres informations importantes comme la fréquence d'échantillonnage, le numéro du canal ou le nombre de coups d'ADC sont également sauvegardées.

Cette étape intermédiaire permet de séparer la définition du format de données de l'analyse en tant que telle, ce qui rend le code plus polyvalent et utilisable facilement avec d'autres appareils (*ie.* oscilloscope ou autres *waveform TDC*).

23.4.2 Reconstruction du temps d'arrivée

Deux algorithmes principaux peuvent alors être exécutés afin de reconstruire les temps d'arrivée des différents signaux :

- ◇ **Discriminateur à fraction constante** (CFD, pour *constant fraction discriminator*) [332].

◇ **Corrélation-croisée (CC)** [333].

L'approche la plus courante pour reconstruire le temps d'arrivée d'un signal échantillonné se fait par le biais d'un algorithme de type CFD [332]. Le temps d'arrivée est alors défini comme l'instant où le signal passe un seuil correspondant à une certaine fraction de son amplitude maximale (voir section 23.2, figure 212, droite). La fraction est un paramètre libre choisi par l'utilisateur. Cette définition permet de s'affranchir en grande partie de l'effet de *timewalk* introduit dans la section 23.2.

Le code d'analyse détermine l'amplitude maximale du signal à l'aide d'une interpolation parabolique à partir des trois points les plus proches du maximum et la valeur exacte de passage de seuil par une interpolation linéaire (2 points utilisés) ou une spline (3 points utilisés). L'incertitude théorique d'un tel algorithme peut être évaluée facilement grâce à la formule suivante [332] :

$$\sigma_{\text{CFD}}^2 = \left(\sigma_{\text{noise}} \cdot \frac{\tau}{A} \right)^2 + \sigma_{\text{jitter}}^2, \quad (114)$$

où σ_{CFD} est l'incertitude totale, σ_{noise} correspond au bruit RMS du canal pendant l'acquisition, $\frac{\tau}{A}$ est le temps de montée du signal divisé par l'amplitude maximale (calculé au niveau du seuil choisi) et σ_{jitter} représente l'incertitude sur le temps de chaque point de donnée, intrinsèque à la chaîne d'acquisition.

Cependant, l'équation 114 fait l'hypothèse d'une montée linéaire du signal ($\frac{\tau}{A}$), ce qui n'est souvent pas entièrement vrai en particulier au début et à la fin du front montant, où un phénomène d'amortissement est souvent observé. Ainsi, le CFD est généralement exécuté par défaut avec une fraction de 0.5 afin de placer le seuil au milieu du front montant, le plus souvent linéaire. Le passage du seuil peut être évalué avec une spline dans le code, ce qui permet une détermination plus robuste lorsque des non-linéarités sont présentes entre les points encadrant le seuil. L'équation 114 sera utilisée dans l'analyse des tests afin d'estimer l'écart entre les signaux considérés et le cas linéaire idéal (voir section 24).

Une version améliorée du CFD a également été implémentée dans le code. Celle-ci consiste à moyenner le résultat renvoyé par plusieurs CFD de fractions différentes. Cela permet bien souvent d'améliorer la résolution grâce à l'utilisation de points supplémentaires disponibles le long du front montant. Dans tous les cas, le temps d'exécution du CFD est de l'ordre de 1 ms par événement (processeur 2.3 GHz avec 4Go de mémoire vive).

La deuxième méthode implémentée dans le code d'analyse est basée sur un algorithme de corrélation-croisée (CC) [333]. Pour chaque signal, une fonction de corrélation est calculée entre celui-ci et un modèle extrait en moyennant un jeu de données de haute statistique similaire mais indépendant. La fonction de corrélation est alors maximisée de façon à obtenir une superposition optimale entre le signal et le modèle. Ce dernier peut finalement être utilisé afin de reconstruire le temps du signal initial de façon plus robuste.

Cette technique s'apparente à une procédure d'ajustement et requiert la construction d'un modèle général applicable à tous les signaux d'un canal donné. Cela peut s'avérer délicat lorsque ceux-ci présentent des fluctuations statistiques importantes, comme cela

est attendu par exemple dans le contexte de la détection de particules à partir de détecteurs fins. En revanche si les fluctuations sont raisonnables, la CC permet en principe une meilleure exploitation des différents points de mesure.

Les étapes principales de l'algorithme de corrélation-croisée sont résumées ci-dessous :

- ◊ Un modèle de signal est généré indépendamment pour chaque canal à partir de données similaires mais indépendantes de la mesure contenant une statistique élevée. Pour cela, tous les signaux sont normalisés par leur amplitude maximale et synchronisés ensemble avec un CFD. Un modèle est alors construit canal par canal en moyennant la forme des différents signaux.
- ◊ Au cours de l'analyse, une synchronisation grossière entre le signal considéré et le modèle est tout d'abord effectuée à l'aide d'un CFD.
- ◊ Le modèle est ensuite translaté par rapport au signal jusqu'à ce que la fonction de corrélation atteigne sa valeur maximum. Le temps d'exécution est limité grâce à la définition d'une fenêtre réduite de translation centrée sur la synchronisation grossière décrite ci-dessus.
- ◊ La translation optimale entre signal et modèle est sauvegardée dans la base de données.

Comme déjà mentionné ci-dessus, la CC permet donc en principe une meilleure exploitation des différents points de mesure par rapport au CFD avec fraction unique, ce dernier n'en utilisant que quelques-uns, *ie.* trois points autour du maximum (parabole) plus deux ou quatre autour du seuil (interpolation linéaire ou spline). Par ailleurs, la CC dépend moins de la ligne de base, la soustraction de celle-ci étant nécessaire uniquement lors de la construction du modèle. La valeur de la fonction de corrélation au maximum fournit également un estimateur de qualité individuel pour chaque signal intéressant, même si elle n'a pas été exploitée pour le moment dans les tests présentés dans la section 24.

Néanmoins, on s'attend à ce que les performances de la CC dépendent de la reproductibilité du signal, ce qui défavorise certaines applications. Ce point faible peut être minimisé en considérant uniquement un sous-intervalle de points échantillonnés généralement plus stables, comme par exemple ceux correspondant au front montant du signal. Par ailleurs, dans le cas de l'exécution de l'algorithme sur un processeur unique, l'algorithme de CC présente un temps d'exécution significativement plus important que le CFD, de l'ordre d'un facteur 100.¹³⁰

Le logiciel d'analyse permet donc la définition d'un format de données pouvant être commun à plusieurs dispositifs (*waveform* TDCs, oscilloscopes, ...) et la reconstruction de temps de vol à l'aide de différents algorithmes. En particulier, il peut mesurer la différence de temps entre deux signaux arrivant dans des canaux distincts de SAMPIC.

¹³⁰Ces performances ont été obtenues après réduction de la fenêtre d'exécution de l'algorithme de corrélation-croisée grâce à un premier CFD (voir texte).

Pour les tests de la section suivante, la différence de temps entre les deux signaux de départ est fixe (gigue négligeable) et la mesure de temps est répétée un grand nombre de fois à la suite dans des conditions similaires. L'étalement de la mesure, correspondant alors à la RMS de l'histogramme des Δt , constitue dans ce cas une mesure de la résolution temporelle de la chaîne d'acquisition.

Dans la section 24, les résultats de différentes mesures de résolution effectuées avec SAMPIC sont présentés.

24 Résultats des mesures expérimentales

Plusieurs tests de la puce SAMPIC ont été entrepris afin de mesurer ses performances et de comparer les résultats des différents algorithmes implémentés dans le code d'analyse (voir section 23.4). Ces résultats constituent les premiers tests de SAMPIC et ont donc été produits à partir de la première version fonctionnelle de la puce.

Dans la section 24.1, les tests préliminaires réalisés avec des signaux directement issus d'un générateur électronique sont décrits et un compte-rendu des résultats est présenté. Dans la section 24.2, une chaîne complète d'acquisition plus proche des conditions réelles envisagées, incluant des détecteurs ultra-rapides au silicium [334,335]¹³¹ et une chaîne de pré-amplification, est testée à l'aide d'un faisceau laser infrarouge.

Finalement, une conclusion générale sur les performances de SAMPIC est donnée dans la section 25. Les développements à venir pour les prochaines versions de la puce sont également discutés.

24.1 Tests avec des signaux issus de générateurs

Les premiers tests à avoir été réalisés utilisent des signaux d'entrée issus d'un générateur LeCroy.¹³² Ceux-ci permettent de caractériser la puce et les performances des différents algorithmes implémentés dans le code d'analyse dans des conditions idéales.

Le générateur est configuré pour produire des signaux gaussiens montant en moins d'une nanoseconde avec une largeur fixe de 0.90 ns à une fréquence de 1 kHz. Leur amplitude peut être réduite grâce à un atténuateur à large bande (Wavetek, 3 GHz). Ils sont ensuite dédoublés grâce à un simple raccordement en T et finalement envoyés dans deux canaux de SAMPIC distincts via des câbles de longueurs variables, afin de moduler la différence entre leur temps d'arrivée respectifs. La gigue entre les deux signaux à l'entrée de la puce est supposée inférieure à 1 ps.

La fréquence d'échantillonnage de SAMPIC est fixée à 6.4 GS/s et les conversions des ADCs sont effectuées sur 11 bits pour l'intégralité des tests présentés dans cette section. Le déclenchement de l'acquisition est assuré séparément dans chaque canal par leur discriminateur respectif avec un seuil dépendant de l'amplitude des signaux. Pendant la première phase de test, le retard entre les deux signaux est fixé à environ 5 ns tandis que l'amplitude du signal est réduite progressivement à l'aide de l'atténuateur (test en amplitude). Pendant la seconde phase de test, l'amplitude (avant dédoublement) est maintenue à 1.2 V tandis que la différence de temps entre les deux signaux est variée de quelques picosecondes à quelques centaines de nanosecondes à l'aide de câbles de tailles variables (test de retard).

Des exemples de signaux échantillonnés sur l'un des canaux de SAMPIC avec une réduction d'amplitude de 27 dB (*ie.* $\simeq 33$ mV à l'entrée de la puce) sont montrés dans la figure 216. Quelques reflexions après le pic principal sont visibles car aucune adaptation d'impédance n'est effectuée au niveau du raccordement en T. Par ailleurs,

¹³¹ Je tiens à remercier particulièrement le Prof. Nicolo Cartiglia pour son aide concernant les détecteurs au silicium.

¹³² modèle : LeCroy 9211 - 250 MHz.

une atténuation allant jusqu'à 30% de l'amplitude maximale des signaux est observée lors de l'utilisation de câbles longs (test de retard, $\Delta t > 100$ ns), ce qui est attribué à la bande passante limitée des câbles ainsi qu'à leur effet de peau.

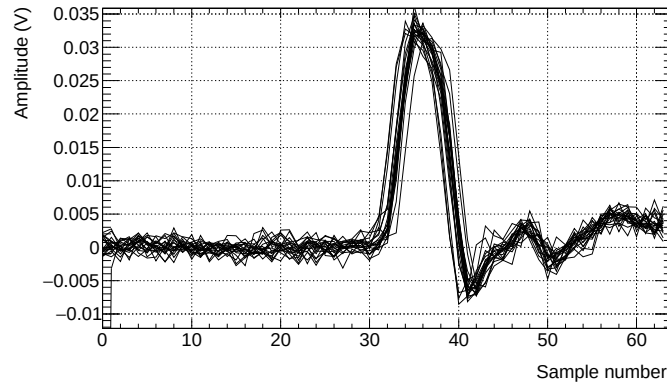


FIG. 216 – Signaux produits par un générateur gaussien et échantillonnés avec la puce SAMPIC à 6.4 GS/s après atténuation. L'amplitude mesurée dans la puce est d'environ 33 mV. Les échantillons sont interpolés linéairement pour la figure. Le recouvrement des différentes courbes illustre la très bonne reproductibilité du signal. Le premier point est parfois perdu (sample 0), ce qui est une caractéristique connue de la puce reliée au temps d'arrivée du signal de déclenchement.

Chaque jeu de données est constitué de 10 000 événements en coïncidence (voir section 23.4). Ceux-ci sont traités par le logiciel d'analyse a posteriori de l'acquisition afin de reconstruire les temps d'arrivée des différents signaux avec les trois algorithmes disponibles (CFD, CFD amélioré et CC). Le CFD utilise une fraction de 50%, tandis que le CFD amélioré moyenne les résultats de plusieurs CFD effectués avec des seuils de 30 à 70% par pas de 5%. Un exemple de synchronisation entre le signal et le modèle par corrélation croisée est présenté dans la figure 217.

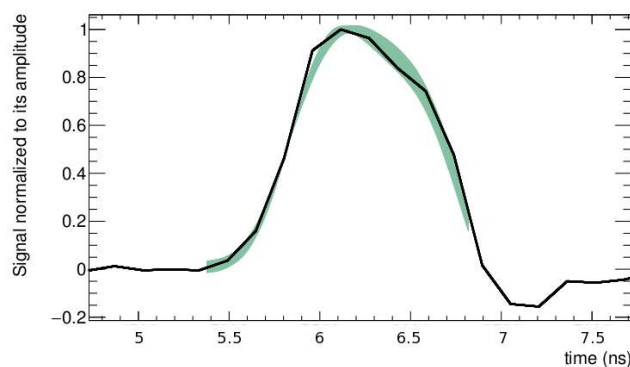


FIG. 217 – Exemple de synchronisation par corrélation-croisée entre le modèle (bande verte) et un signal (ligne noire). La bande du modèle correspond à sa dispersion statistique (RMS). Celle-ci atteint 0 au point utilisé pour la synchronisation temporelle des différents signaux utilisés lors de sa construction (voir section 23.4). Le signal montré ici correspond à l'un des échantillons de la figure 216.

Les histogrammes contenant les 10 000 mesures de différence de temps entre les deux canaux estimées à partir des trois algorithmes sont représentés dans la figure 218 pour

les jeux de données correspondant à des signaux d'environ 33 mV (haut) et 660 mV (bas) espacés d'environ 5 ns (test en amplitude). Les RMS de ces histogrammes pour l'intégralité des amplitudes testées sont montrées dans la figure 219. Celles-ci sont de l'ordre de 4 à 6 picosecondes pour les hautes amplitudes (> 200 mV), correspondant aux données avec un signal sur bruit élevé (SNR, pour *signal over noise ratio*). Cependant, la RMS augmente de façon importante pour les amplitudes inférieures à 100 mV à cause du SNR faible. Dans cette région, l'algorithme de corrélation-croisée permet une nette amélioration des résultats jusqu'à un facteur proche de 2. En revanche, le CFD amélioré ne permet pas d'améliorer significativement la résolution. Cela vient probablement de la montée très rapide des signaux, qui limite le nombre de points échantillonné le long du front montant à 1 ou 2 et donc ne permet pas au CFD amélioré d'exploiter plus de points qu'un CFD classique.

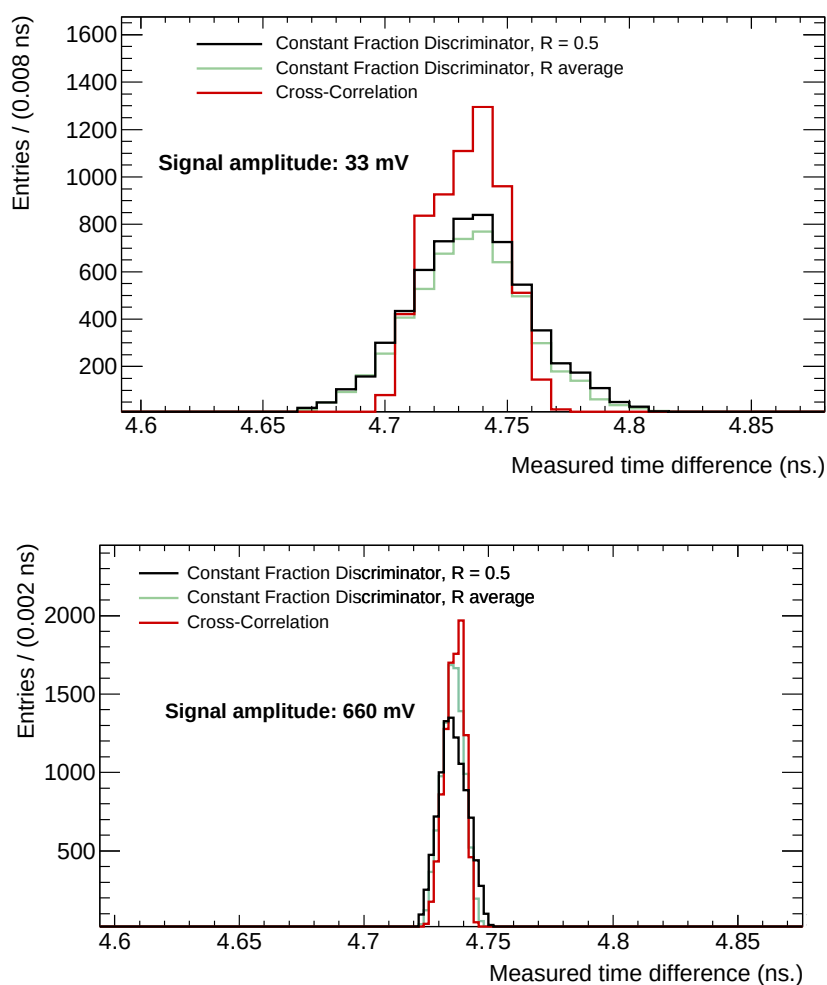


FIG. 218 – Distribution des différences de temps d'arrivée entre deux signaux de 33 mV (haut) et 660 mV (bas) dans SAMPIC. Les deux signaux sont issus d'un générateur gaussien et espacés d'environ 5 ns. Les distributions obtenues avec le CFD (ligne noire), le CFD amélioré (ligne verte) et la CC (ligne rouge) sont représentées.

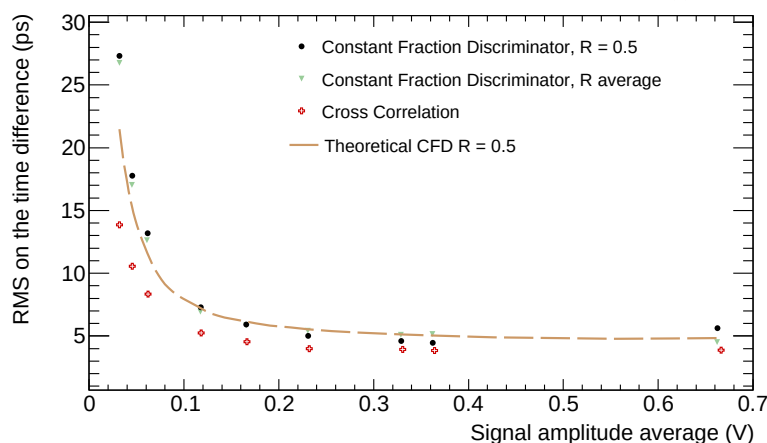


FIG. 219 – RMS de la différence de temps entre deux signaux mesurée à partir de 10 000 événements échantillonnés par la puce SAMPIC à une fréquence de 6.4 GS/s en fonction de l'amplitude moyenne des signaux. Les deux signaux sont issus d'un générateur gaussien et espacés d'environ 5 ns.

Les résultats du CFD classique avec une fraction à 0.5 sont en bon accord avec les prédictions théoriques pour ces signaux idéaux très rapides (voir équation 114, section 23.4), à l'exception de la toute dernière mesure ($\simeq 33$ mV) où le signal sur bruit très bas engendre parfois des non-linéarités importantes sur le front montant.

La courbe théorique est ici calculée pour une différence de temps, *ie* $\Delta t = t_1 - t_2$ avec l'équation 114 s'appliquant pour chacun des deux termes t_1 et t_2 . Le bruit RMS nécessaire au calcul des prédictions théoriques (voir équation 114) est extrait dans chaque canal à partir des points employés pour la détermination de la ligne de base. La valeur observée est d'environ 1.2 mV pour l'ensemble des points de mesure, conformément à ce qui est attendu en l'absence de source de bruit extérieur (voir tableau 43). σ_{jitter} est fixée de façon à ce que les mesures correspondent aux prédictions à SNR élevé et prise égale dans les deux canaux ($\sigma_{\text{jitter}}^1 = \sigma_{\text{jitter}}^2$). Si les deux canaux sont non corrélés (σ_{jitter}^1 et σ_{jitter}^2 indépendants), $\sigma_{\text{jitter}}^{1,2} \simeq 4 \text{ ps} / \sqrt{2} \simeq 3 \text{ ps}$, validant l'objectif de conception de SAMPIC (5 ps, voir section 23.2). Une corrélation légère n'est cependant pas exclue car la même puce est utilisée pour l'acquisition des deux signaux. En tout généralité la gigue pour un canal est alors donnée par $4 \text{ ps} / \sqrt{2 \cdot (1 - \rho)}$, où ρ est le facteur de corrélation entre deux canaux. Il serait nécessaire de répéter ces mesures avec deux SAMPIC différents afin de s'assurer de la non-corrélation complète entre les canaux d'une même puce.¹³³

Finalement, les résultats du test de retard sont dévoilés dans la figure 220. La résolution de SAMPIC montre une dépendance quasi-nulle vis-à-vis de la différence en temps mesurée, la légère pente visible sur la figure 220 étant attribuée à l'atténuation des signaux lors de l'utilisation de câbles longs mentionnée en début de section.

Ainsi, dans ces conditions idéales la puce SAMPIC fait preuve d'une résolution intrinsèque excellente, de l'ordre de 4 ps RMS sur le Δt entre deux signaux de 400 mV. De plus, l'algorithme de corrélation-croisée permet une nette amélioration des performances à bas signal sur bruit (< 200 mV) par rapport à un CFD classique ou amélioré, avec une

¹³³Cette remarque est également valide sur σ_{noise}^1 et σ_{noise}^2 , qui sont supposés indépendants ici. Cependant, le bruit joue un rôle négligeable à SNR élevé.

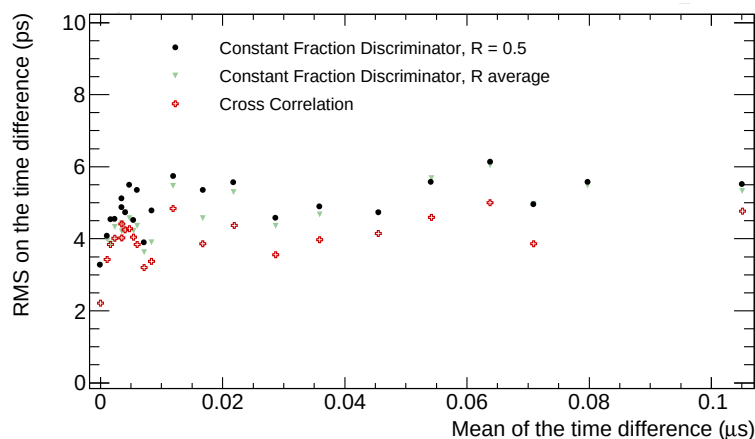


FIG. 220 – RMS de la différence de temps entre deux signaux mesurée à partir de 10 000 événements échantillonnés par la puce SAMPIC à une fréquence de 6.4 GS/s. Les résultats sont représentés en fonction de la différence de temps entre les deux signaux. L'amplitude des signaux est fixe et d'environ 660 mV (SNR élevé).

résolution améliorée jusqu'à un facteur proche de 2. Dans la section 24.2, les résultats de tests effectués à partir de signaux de détecteurs ultra-rapides au silicium sont présentés.

24.2 Tests avec des signaux issus de détecteurs ultra-rapides de silicium

24.2.1 Description générale

Afin de tester SAMPIC dans des conditions plus proches des conditions réelles envisagées dans le cadre d'une application en physique des hautes énergies ou en imagerie médicale (voir section 23.1), des mesures complémentaires ont été effectuées à partir de signaux générés par de véritables détecteurs de particules, toujours dans des conditions idéales cependant. Les détecteurs considérés représentent un type particulier de détecteurs à bas gain basés sur des diodes à avalanche (LGAD, pour *low gain avalanche diode*) [336] appelés détecteurs ultra-rapides au silicium (UFSD, pour *ultra-fast silicon detectors*) [334, 335].

Les LGADs sont des nouveaux détecteurs au silicium qui combinent l'avantage d'un gain interne, à l'image des photodiodes à avalanche (APD, pour *avalanche photodiodes*), et les propriétés classiques bien connues des détecteurs au silicium. L'idée principale est qu'un gain faible permet de limiter au maximum le bruit de comptage (*dark counts*)¹³⁴ et les fuites de courant souvent observées dans les APDs, tout en produisant un signal amplifié par exemple très utile pour des applications de temps de vol. Les UFSDs représentent un type de LGADs optimisé pour les mesures de temps de vol.

¹³⁴le bruit de comptage représente la réponse d'un détecteur en l'absence de signal.

24.2.2 Premier montage, incluant deux amplificateurs à large bande

Dans le premier test proposé, deux USFDs sont soumis à un faisceau laser infrarouge dédoublé au préalable. Les signaux sont ensuite amplifiés avant d'être envoyés dans deux canaux de SAMPIC distincts. Le montage complet est résumé dans les figures 221 et 222 et détaillé ci-dessous :

- ◇ Un faisceau laser picoseconde PiLas de longueur d'onde 1060 nm¹³⁵ est dédoublé à l'aide d'une fibre optique et envoyé à travers de deux USFDs. Le laser présente une bande passante de largeur 2 nm et peut être varié en intensité. La gigue entre les deux signaux laser est supposée inférieure à 1 ps.
- ◇ Des USFDs de gain 10 [335, 337] polarisés grâce à une différence de potentiel de 800 V sont utilisés.
- ◇ Les signaux en sortie des détecteurs sont amplifiés à 40 dB à l'aide de deux amplificateurs à large bande (BDA, pour *broadband amplifier*) de marque CIVIDEC.¹³⁶
- ◇ Les signaux sont finalement envoyés dans deux canaux distincts de SAMPIC. L'acquisition est déclenchée par les discriminateurs des canaux. La fréquence d'échantillonnage est fixée à 6.4 GS/s et les conversions des ADCs sont effectuées sur 11 bits.
- ◇ Les données échantillonnées par SAMPIC sont enregistrées sur un ordinateur à l'aide du logiciel de contrôle (voir section 23.3) et traitées avec le logiciel d'analyse (voir section 23.4).

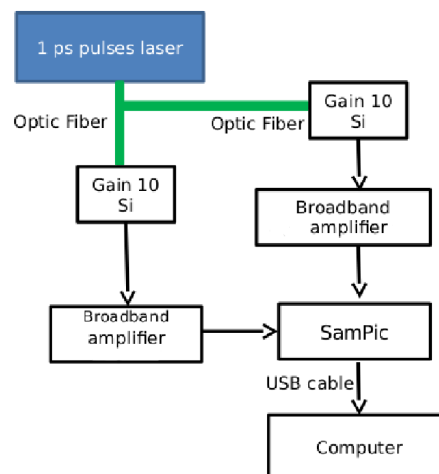


FIG. 221 – Schéma général du premier montage utilisé pour les tests de SAMPIC avec des détecteurs ultra-rapides au silicium. Deux USFDs associés à des amplificateurs à large bande sont utilisés [328].

¹³⁵modèle PiL106X.

¹³⁶modèle C2.

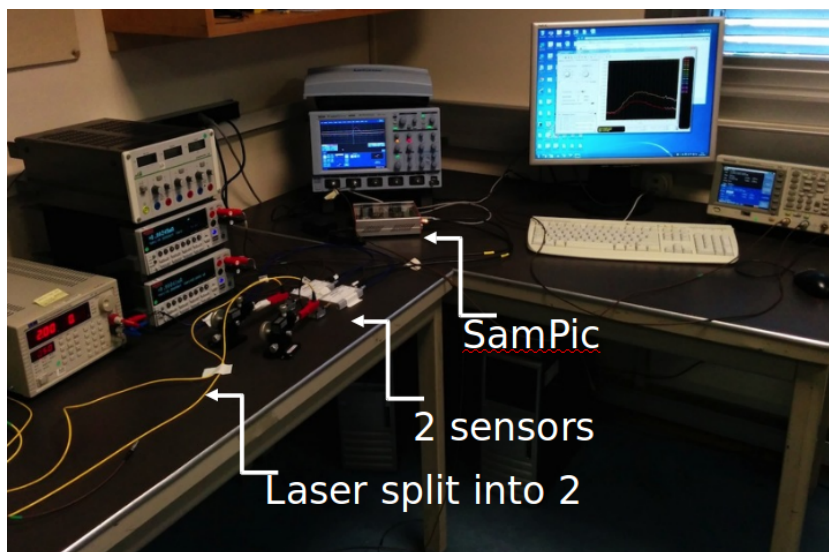


FIG. 222 – Photographie du banc de test monté au CERN (juin 2014). Le montage inclut un faisceau laser picoseconde dédoublé à l'aide d'une fibre optique, 2 UFSDs connectés à des BDAs et un boîtier contenant SAMPIC.

La première partie des tests est effectuée avec des BDAs, qui amplifient le courant lu en entrée à 40 dB (canaux de 50 Ω). L'intensité du laser est variée dans un intervalle correspondant au dépôt d'énergie moyen de 2 à 4 MIPs (*minimum ionizing particles*), permettant ainsi d'imiter grossièrement le passage de particules relativistes étudiées en physique des hautes énergies (LHC, ...). Malheureusement, le système d'amplification à disposition n'est pas suffisant pour produire un signal conforme à SAMPIC correspondant à l'énergie d'1 MIP. Le deuxième montage utilise un autre type d'amplificateur qui permettra d'effectuer des mesures correspondant au dépôt d'énergie moyen d'1 MIP (voir section 24.2.3).

Des exemples de signaux échantillonnés par SAMPIC produits par un faisceau laser équivalent à 2 MIPs sont présentés dans la figure 223. La polarité est inversée comparée aux tests précédents (voir section 24.1) à cause des porteurs de charge négatifs des UFSDs. La fenêtre d'acquisition à 6.4 GS/s est centrée sur le front montant du signal. D'autre part, un exemple de synchronisation entre un signal et son modèle par corrélation-croisée est montré dans la figure 224. Le signal est bien plus lent comparé aux tests précédents, avec un temps de montée de quelques nanosecondes (typiquement 3-4 ns). La queue de distribution fluctue également considérablement et est par conséquent exclue du modèle.

Le test en amplitude présenté dans la section 24.1 est alors répété avec le nouveau montage. Les distributions des différences de temps entre les deux signaux mesurées à partir de 10 000 événements échantillonnés par SAMPIC sont montrées dans la figure 225 pour un signal équivalent à 2 MIPs, tandis que la distribution de la RMS pour les différentes amplitudes testées est présentée dans la figure 226. Les résolutions mesurées sont toujours inférieures à 110 ps RMS. Le CFD amélioré permet cette fois une nette amélioration des résultats, jusqu'à 20 ps environ ($\simeq 15\%$), permettant ainsi de rapprocher la courbe expérimentale du CFD théorique valide pour un signal montant linéairement. En effet, les signaux étant bien plus lents que lors des tests précédents (voir section 24.1), de nombreux points d'échantillonnage figurent cette fois sur le pre-

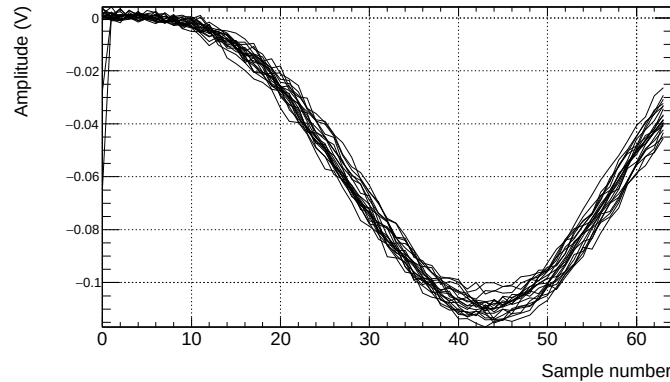


FIG. 223 – Signaux générés par des UFSDs amplifiés par un BDA CIVIDEC C2 à 40 dB et échantillonnées par la puce SAMPIC à 6.4 GS/s. Le signal monte en environ 3.5 ns (mesurée entre 10 et 90% de l'amplitude) et présente une amplitude équivalente à 2 MIPs, représentant une différence de potentiel à l'entrée de SAMPIC d'environ 110 mV. Les signaux sont interpolés linéairement pour la figure. Le premier point est parfois perdu (sample 0), ce qui est une caractéristique connue de la puce reliée au temps d'arrivée du signal de déclenchement.

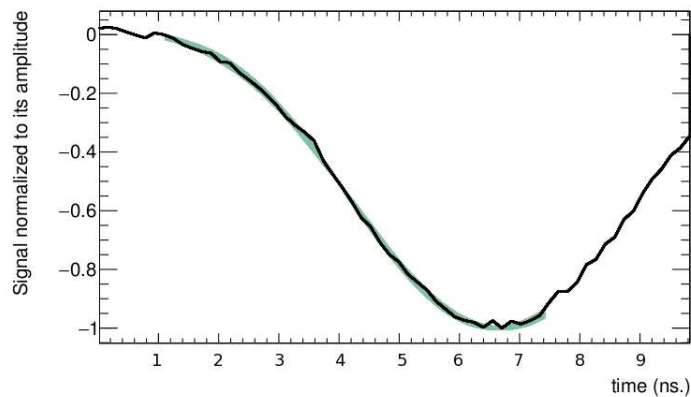


FIG. 224 – Exemple de synchronisation par corrélation-croisée entre le modèle (bande verte) et un signal (ligne noire). La bande du modèle correspond à sa dispersion statistique (RMS). Celle-ci atteint 0 au point utilisé pour la synchronisation temporelle lors de la construction du modèle (voir section 23.4). Le signal montré ici correspond à l'un des échantillons de la figure 223.

mier front (> 10). Cela permet au CFD amélioré d'augmenter la résolution en exploitant l'information contenue dans plus de points par rapport au CFD classique.

Comme dans la partie précédente, la courbe théorique est obtenue à partir de l'équation 114 à l'aide du bruit RMS mesuré dans chaque canal pour chaque jeu de données. Celui-ci est légèrement plus élevé que lors du test précédent ($1.4 \text{ mV} > 1.2 \text{ mV}$), probablement à cause de l'existence d'une composante de bruit supplémentaire venant de la chaîne UFSD + BDA.

D'autre part, le σ_{jitter} sur le Δt est bien plus élevé que précédemment et de l'ordre de 30 ps. Cependant, il représente la gigue de toute la chaîne d'acquisition, qui ne se limite plus à SAMPIC ici mais inclut l'ensemble UFSD + BCA + SAMPIC. La gigue totale du système d'acquisition est donc totalement dominée par l'UFSD + BCA, σ_{SAMPIC} ayant été estimé à environ 4 ps lors des tests précédents (voir section 24.1). Puisque deux UFSDs et deux BCAs distincts sont utilisés pour les deux canaux et que la gigue de SAMPIC est négligeable, on peut ici déduire que $\sigma_{\text{jitter}}^1 \simeq \sigma_{\text{jitter}}^2 = \frac{30 \text{ ps}}{\sqrt{2}} \simeq 21 \text{ ps}$. Ainsi, la précision de la mesure à haut SNR est limitée par la partie détecteur + amplificateur de la chaîne d'acquisition, qui est de l'ordre de 20 ps par canal.

Finalement, la corrélation-croisée donne des résultats comparables au CFD amélioré. Cela est probablement dû à l'exclusion des queues des signaux pour le calcul de la fonction de corrélation à cause de leur trop grandes fluctuations statistiques, ce qui limite le nombre d'échantillons utilisés par la CC à un chiffre proche de celui du CFD amélioré. Le plateau en résolution, correspondant à un domaine où le SNR n'est plus un facteur limitant, semble être atteint pour des amplitudes supérieures à 300 mV environ. Cette valeur est du même ordre de grandeur quoique légèrement supérieure à celle observée lors des tests précédents ($\simeq 200 \text{ mV}$, voir section 24.1).

24.2.3 Second montage, incluant un amplificateur de charge rapide

Une seconde phase de test a été réalisée à l'INFN de Turin en octobre 2014 [328] afin de mesurer la résolution d'un canal unique de SAMPIC avec UFSD et non plus la résolution sur une différence de temps entre deux canaux. Les résultats obtenus ici ne sont donc pas directement comparables aux résultats exposés dans les sections 24.1 et 24.2. Pour cela, un canal de SAMPIC est directement connecté au signal de déclenchement fournit avec le système du laser tandis qu'un autre canal échantillonne la réponse d'un détecteur UFSD amplifié par un amplificateur de charge rapide (CSA, pour *charge sensitive amplifier*) CIVIDEC de gain 5.4 mV/fC .¹³⁷ Cette configuration permet alors de produire des signaux conformes à SAMPIC équivalant 1 à 4 MIPs. Un schéma récapitulatif du montage est fourni dans la figure 227.

Le signal étant beaucoup plus lent que pour les tests précédents à cause de la bande passante plus limitée du CSA (temps de montée d'environ 6 ns), la fréquence d'échantillonnage de SAMPIC est abaissée à 3.2 GS/s afin de garder une fenêtre d'acquisition acceptable. Des exemples de signaux correspondant à 1 MIP ($\simeq 120 \text{ mV}$) et échantillonnés à 3.2 GS/s sont représentés dans la figure 228. La polarité du signal est à nouveau positive car le CSA inverse la polarisation des signaux pendant l'amplification.

Finalement, le test en amplitude est répété avec ce nouveau montage, les résultats

¹³⁷modèle C6.

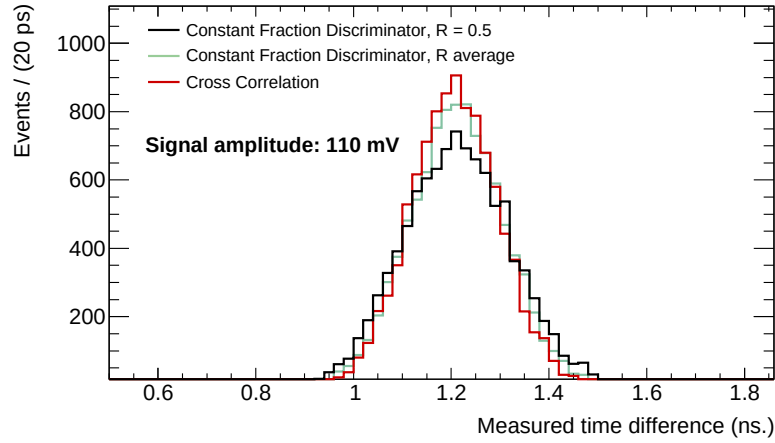


FIG. 225 – Distribution de la différence de temps d'arrivée entre deux signaux de 100 mV ($\simeq 2$ MIPs) dans SAMPIC. Les deux signaux sont issus d'UFSDs amplifiés par des BDAs CIVIDEC C2 à 40 dB et échantillonnés à 6.4 GS/s.

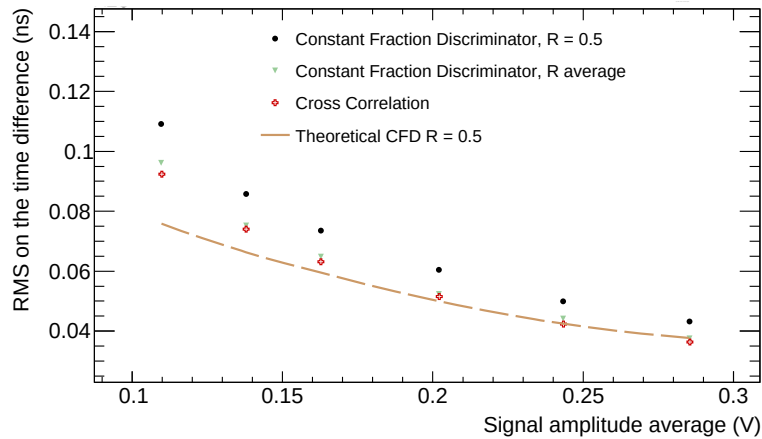


FIG. 226 – RMS de la différence de temps entre deux signaux mesurées à partir de 10 000 événements échantillonnés par la puce SAMPIC à une fréquence de 6.4 GS/s en fonction de l'amplitude moyenne des signaux ($\simeq 2$ à 4 MIPs). Les deux signaux sont issus d'UFSDs amplifiés par des BDAs CIVIDEC C2 à 40 dB.

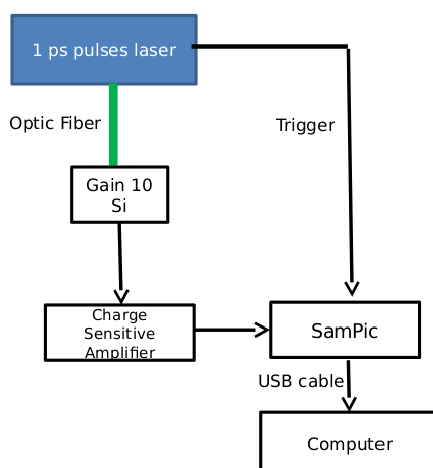


FIG. 227 – Schéma du second montage utilisé pour les tests de SAMPIC avec des détecteurs ultra-rapides au silicium. Un amplificateur de charge rapide est utilisé pour l'amplification du signal UFSD tandis qu'un autre canal est directement relié au signal de déclenchement du laser [328].

étant présentés dans la figure 229. Pour un dépôt d'énergie moyen correspondant à 1 MIP ($\simeq 120$ mV avec le CSA), la chaîne de mesure présente une résolution temporelle de 85 ps avec la corrélation croisée alors que le signal monte en 6 ns environ. La CC permet ainsi une amélioration des résultats de l'ordre de 20 ps ($\simeq 15\%$) comparé à un CFD classique et 10 ps comparé au CFD amélioré. À haut SNR, le plateau de résolution se situe à 40 ps environ. Il est atteint pour des amplitudes supérieures à 300 mV, de façon similaire au test précédent. Dans ce cas, les différents algorithmes donnent des résultats comparables.

Le bruit RMS et ses fluctuations en fonction du point de fonctionnement testé sont du même ordre de grandeur que lors des tests précédents et se situent autour de 1.3 mV. Les légères fluctuations de bruit conduisent ici à de légères fluctuations sur les résultats, à la fois pour la courbe théorique et les courbes expérimentales (voir figure 229). En effet, les signaux ayant des temps de montée beaucoup plus lents, la résolution de la mesure est naturellement plus sensible aux fluctuations du bruit (voir équation 114). À nouveau le CFD amélioré permet un rapprochement significatif des résultats du CFD avec le cas idéal, tandis que la CC fait même mieux à bas SNR. Le σ_{jitter} , représentant alors la gigue d'un canal unique UFSD + CSA + SAMPIC, est ici de l'ordre de 40 ps.

Dans la section 25, une conclusion générale est donnée sur l'ensemble des tests réalisés avec SAMPIC. Les limitations actuelles de la puce et ses améliorations futures sont également discutées.

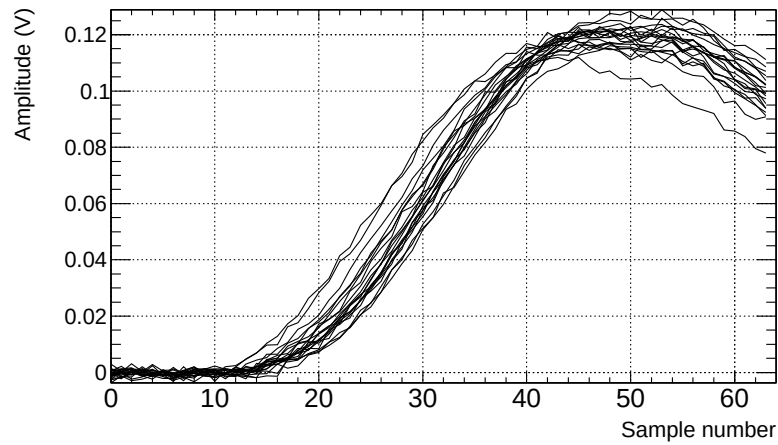


FIG. 228 – Signaux générés par un UFSD amplifié par un CSA CIVIDEC C6 à 5.4 mV/fC et échantillonnées par la puce SAMPIC à 3.2 GS/s. Le signal monte en environ 6 ns (mesurée entre 10 et 90% de l'amplitude) et présente une amplitude équivalente à 1 MIP, ce qui représente une différence de potentiel à l'entrée de SAMPIC d'environ 120 mV. Les signaux sont interpolés linéairement pour la figure. Le premier point est parfois perdu, ce qui est une caractéristique connue de la puce liée au temps d'arrivée du signal de déclenchement.

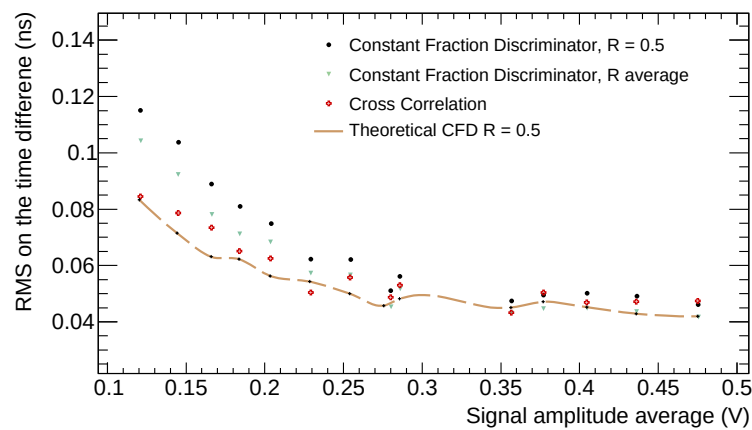


FIG. 229 – Résolution RMS du temps d'arrivée d'un signal unique mesurée à partir de 10 000 événements échantillonnés par la puce SAMPIC à une fréquence de 3.2 GS/s en fonction de l'amplitude moyenne des signaux, équivalent 1 à 4 MIPs. Le signal est issu d'un UFSD amplifié par un CSA CIVIDEC C6 de gain 5.4 mV/fC.

25 Conclusion et perspectives

Dans cette partie, les premières mesures de temps réalisées avec la puce SAMPIC ont été présentées. La puce est intégrée dans un boîtier fonctionnel prêt à l'emploi. Elle se base sur un échantillonnage très rapide de signaux de détecteurs (*waveform* TDC) afin d'extraire le temps de passage de particules avec une résolution de l'ordre de la picoseconde.

Un logiciel de contrôle (en ligne) et un code d'analyse (hors-ligne) dédiés sont utilisés pour l'acquisition et l'exploitation des données, ce dernier ayant été développé au cours des années de thèse. Le logiciel d'analyse implémente plusieurs algorithmes de traitement de signal (CFD classique, CFD amélioré, CC), ce qui permet de comparer leur performance sur des jeux de données identiques. Des études similaires ont déjà été réalisées avec d'autres *waveform* TDCs, mais uniquement à haut signal sur bruit [338].

Les premiers tests sont menés à l'aide de signaux idéaux produits par un générateur afin de caractériser la puce seule. Des mesures similaires sont ensuite réalisées avec des signaux de détecteurs ultra-rapides au silicium (UFSD) produits à partir d'un faisceau laser infrarouge.

Dans tous les cas, la résolution temporelle des mesures est largement indépendante du temps mesuré. En revanche, le signal sur bruit (SNR) influe considérablement sur la précision des mesures, attendu de l'équation 114 (voir section 23.4). Les résultats principaux des tests effectués au cours des années de thèse et décrits précédemment dans la section 24 sont résumés dans le tableau 44.

À haut SNR, tous les algorithmes démontrent des performances similaires et les mesures effectuées avec SAMPIC atteignent une résolution de l'ordre de 4 (40) ps pour des signaux issus de générateurs (UFSDs). Les différences obtenues dans les deux cas de figure montrent que la précision sur le temps est limitée par les UFSDs et les amplificateurs et non par SAMPIC, qui tient ses objectifs de conception correspondant à une gigue inférieure à 5 ps RMS. Par ailleurs, aucune dégradation significative des performances n'a été constatée entre l'échantillonnage à 6.4 et à 3.2 GS/s avec les UFSDs si on tient compte du fait que le signal dans le second cas est bien plus lent.

Le temps d'exécution d'un CFD (CC) est de l'ordre de 1 ms (100 ms) par événement sur un processeur 2.3 GHz avec 4 Go de mémoire. Celui-ci pourrait être réduit d'un facteur 100 en étant implémenté sur un réseau de porte programmable (FPGA) [339]. Par ailleurs, l'algorithme de corrélation croisée bénéficierait grandement d'un calcul parallélisé grâce à la structure de la fonction de corrélation, ce qui permettrait probablement de réduire à nouveau le temps d'exécution de manière significative.

La corrélation-croisée et le CFD amélioré permettent généralement une amélioration des performances à bas SNR. Le CFD amélioré est particulièrement intéressant dans le cas de signaux lents car il permet alors d'exploiter l'ensemble des points échantillonnés le long du front montant. La CC permet en principe une exploitation optimale des différents points de mesure sous la réserve d'une fluctuation limitée des signaux d'entrée. Ces résultats sont très encourageants et motivent le développement d'une algorithmique propre à la reconstruction du temps de vol en parallèle des *waveform* TDC.

Finalement, le plateau en résolution de SAMPIC est atteint pour des signaux d'en-

Fréquence d'échantillonnage	Type de signal	Mesure réalisée	Temps de montée (ns)	Résolution CFD (haut/bas SNR) (ps)	Résolution CC (haut/bas SNR) (ps)	Plateau de résolution (mV)
6.4 GS/s	Générateur	différence de temps	$\simeq 0.3$	4/27	4/14	> 200
6.4 GS/s	UFSD + BDA ($\simeq 2$ à 4 MIPs)	différence de temps	$\simeq 3$	40/95	40/93	> 300
3.2 GS/s	UFSD + CSA ($\simeq 1$ à 4 MIPs)	canal unique	$\simeq 6$	40/105	45/85	> 300

TAB. 44 – Résumé des différents tests réalisés avec la puce SAMPIC décrits en détail dans la section 24.

trée ayant des amplitudes supérieures à 300 mV environ pour les trois séries de tests réalisés. Ce résultat pourrait constituer une première référence dans le cadre de la conception d'une chaîne de détection complète intégrant SAMPIC.

Cependant, quelques développements supplémentaires sont nécessaires afin de pouvoir utiliser SAMPIC dans un grand détecteur de physique des hautes énergies comme ceux du LHC ou en imagerie médicale. En effet, les tests rapportés ici présentent quelques différences notables avec ces deux applications potentielles, résumées ci-dessous :

- ◇ Les signaux attendus dans le cadre d'applications en physique des hautes énergies ou en imagerie médicale sont produits respectivement à partir de protons hautement relativistes (ATLAS-AFP, CT-PPS) et de photons gamma (TEP) et non à partir de photons infrarouges comme ceux considérés dans la section 24.2. Lors des tests avec les UFSDs, le faisceau laser est réglé afin de reproduire l'énergie moyenne déposée par une ou plusieurs MIPs, cependant les fluctuations autour de la moyenne devraient être différentes dans le cas des protons chargés, affectant la variabilité des signaux. Par ailleurs, les photons gamma ne peuvent pas être détectés efficacement par des UFSDs ce qui implique l'utilisation d'autres types de détecteurs pour les TEPs, par exemple basés sur des cristaux scintillateurs [340] ou des cônes de Tcherenkov [341]. Des tests faisceaux dédiés incluant les bonnes particules et une chaîne de détection adaptée sont donc absolument nécessaires avant toute application spécifique.
- ◇ Afin de pouvoir utiliser SAMPIC dans des détecteurs de haute luminosité comme ceux du LHC, des efforts importants doivent également être fournis afin de réduire le temps mort par canal. En effet, celui-ci est d'au moins 200 ns pour la version de la puce considérée pour les tests. La nouvelle version de la puce, sortie en décembre 2015, implémente une méthode "ping-pong" (voir section 23.2) permettant déjà de réduire le temps mort d'un facteur 2. Des efforts complémentaires seront cependant nécessaires, l'objectif étant d'arriver à un temps mort par canal de l'ordre de quelques dizaines de nanosecondes. D'autre part, un *waveform* TDC produit un débit de données bien plus important qu'un TDC classique, ce qui nécessite des développements particuliers concernant le système de transmission et de traitement des données.

La collaboration TOTEM a réalisé en 2014 et 2015 plusieurs tests faisceaux intégrant SAMPIC et des détecteurs de temps de vol en diamant [342, 343]. Des résultats ont été rendus publics en juin 2015 [342], annonçant des résolutions de l'ordre de 100 ps. Ce chiffre peut être facilement réduit d'un facteur $\sqrt{n}$ dans le cas de détecteurs présentant une épaisseur fine tels que les détecteurs en diamant ou les UFSDs en empilant n couches de détection. Ainsi, il est réaliste d'imaginer 4 couches de détecteurs de temps de vol empilées, ce qui permettrait en principe d'obtenir une résolution améliorée d'un facteur 2, soit 50 ps. Une amélioration également du côté des détecteurs est prévue dans les prochaines années, ce qui pourrait éventuellement permettre de descendre encore légèrement ce chiffre (40 ps?).

Une chaîne de détection similaire a été mise en place par TOTEM au LHC et a pu prendre des données pendant les *runs* spéciaux effectués fin 2015, ce qui montre la détermination de la collaboration TOTEM à utiliser SAMPIC. Cela est très encourageant pour les utilisations futures de *Waveform* TDC au LHC. Les résultats des *runs* spéciaux de TOTEM n'ont pas encore été dévoilés mais sont attendus avec impatience par l'équipe de développement de SAMPIC.

La partie VII clôt ce manuscrit par une conclusion générale concernant les différents projets réalisés au cours de la thèse. Quatre appendices ainsi que la bibliographie complète sont finalement incluses.

“J’ouvris les yeux par-dessus l’épaule d’Oroshi, par-delà le plateau tranché net – un coup de hache – et je ne pouvais pas me résoudre à l’accepter. Du fond de mon être, quelque chose continuait à avancer sur la lande, à s’arc-bouter contre la coulée matinale des brouillards ascendants pour y tracer vent debout, choon ou pas, avec la grêle rapeuse aux pommettes, les yeux à la recherche guettant la prochaine colline, le corps en manque d’espace déjà, qui se projetait au loin sur ce relief en bouloches blanches, parmi cette terre neigeuse de nuages sans consistance de sol, où le talon ne pourrait jamais résonner et les crampons mordre, cette mer crémeuse pour rêveur abrité qui signait donc le bout du chemin ?”

Sov Strochnis, scribe
La horde du contrevent, Alain Demasio (2004).

Septième partie

Conclusion générale, appendices et références

Sommaire

26 Conclusion générale	452
A Corrections supplémentaires de l'isolation calorimétrique basées sur les données	455
B Estimations du taux de mauvaise identification des jets en photons dans les données	464
C Étude des différences entre les effets d'empilement dans les données et dans la simulation d'un point de vue calorimétrique	465
D Résumé des différentes mesures $pp \rightarrow \gamma\gamma + X$ réalisées à $\sqrt{s} = 8$ TeV	469
Références	476

26 Conclusion générale

Au cours de mes années de thèse, je me suis intéressé à la production de photons à très haute énergie dans les collisions proton-proton du LHC avec le détecteur ATLAS (voir parties I et II). Les photons sont des objets délicats à étudier pour plusieurs raisons. Tout d’abord, leur masse nulle leur permet d’être émis à des énergies et à des angles infiniment petits, ce qui pose de sérieux problèmes de divergences infrarouges lors d’un développement perturbatif (théorie) et mène à une certaine ambiguïté entre partons et photons au niveau expérimental dans les collisionneurs hadroniques, dans la limite d’une émission colinéaire (photons de fragmentation). Par ailleurs, les photons étant des particules neutres, ils ne laissent en général pas de traces dans le détecteur interne et ne peuvent ainsi être reconstruits qu’à partir de dépôts calorimétriques. Ils pourraient cependant constituer le premier portail vers une physique au-delà du modèle standard, comme l’a récemment montré l’excès modéré dans le spectre en masse des paires de photons autour de 750 GeV, observé à la fois par les collaborations ATLAS et CMS.

Dans un premier temps, je me suis intéressé à l’énergie d’isolation transverse des photons, qui constitue une mesure de leur environnement hadronique immédiat (voir partie III). L’énergie d’isolation est une variable clé pour discriminer les photons dont la production est gouvernée par la pQCD des photons secondaires issus de la désintégration de hadrons et de taus (jets). Au LHC, cette variable est affectée non seulement par le bruit de l’électronique du calorimètre et par les dépôts dans le cône d’isolation du candidat photon lui-même, mais également par le taux d’empilement très élevé du LHC et les événements sous-jacents. J’ai calculé des corrections de fuite supplémentaires pour l’isolation afin de tenir compte des différences de développement latéral entre les gerbes électromagnétiques simulées et réelles [201]. Celles-ci ont un effet particulièrement important pour les photons de haute impulsion transverse, messagers potentiels d’une physique au-delà du modèle standard, et donc ont un impact direct sur les recherches actuellement en cours au LHC. L’emploi de *clusters* de taille fixe pour les photons et la présence de différences de forme significatives entre les gerbes réelles et simulées entraînent cependant certaines limitations dans les corrections actuelles, discutées à la fin de la partie III.

J’ai ensuite réalisé la mesure de la section efficace de production de paires de photons isolés à partir de l’ensemble des données de collisions proton-proton collectées en 2012 avec une énergie dans le centre de masse de $\sqrt{s} = 8$ TeV par la collaboration ATLAS (voir partie IV) [211]. Le couplage important entre photons et partons permet d’effectuer un test complet de la QCD via la comparaison de ces résultats avec plusieurs calculs théoriques, allant des corrections perturbatives en α_s (LO, NLO, NNLO) aux techniques de resommation perturbatives des logarithmes à tous les ordres (NNLL), en passant par la mise en évidence de photons issus de la fragmentation de partons (composante directe, 1-fragmentation et 2-fragmentations) et de la structure interne du proton (PDFs). De façon plus pragmatique, cette mesure permet également d’évaluer la précision des calculs approchés basés sur la combinaison d’éléments de matrice et de cascades de partons (ME+PS *merged at LO, NLO*) décrivant l’intégralité de l’état final, et donc pouvant être associés à une simulation complète du détecteur. D’un point de vue expérimental, en plus de fournir des résultats de physique importants avec une portée en énergie augmentée et des incertitudes réduites, cette mesure a permis de mieux comprendre les variations de l’énergie d’isolation des photons en fonction de la région du détecteur

(E_T^γ , $|\eta^\gamma|$) et d'ouvrir des perspectives d'amélioration concernant la description de l'empilement dans la simulation d'ATLAS et les mesures d'identification effectuées par le groupe de performance. Des pistes ont également été suggérées afin d'étudier plus en détail le rôle de l'événement sous-jacent dans les événements $\gamma\gamma + X$ et les taux de production en fonction de l'énergie d'isolation transverse. La publication associée à la mesure devrait être soumise à une revue internationale dans les prochains mois.

Les couplages à quatre photons au LHC sont ensuite étudiés en profondeur dans la partie V, où un point de vue phénoménologique est adopté [56, 57]. En effet, il est souvent oublié que le LHC est également un collisionneur photon-photon à haute énergie très efficace grâce à la charge non nulle des protons et à sa luminosité importante. Les collisions photon-photon à haute énergie du LHC pourront par ailleurs bientôt être étudiées précisément grâce à l'installation imminente de détecteurs de protons intacts à l'avant pouvant être opérés lors des prises de données nominales (ATLAS-AFP, CT-PPS). Je montre qu'il sera alors possible de contraindre fortement de nombreux modèles de nouvelle physique grâce à l'étude à haute énergie du processus $pp \rightarrow p + \gamma\gamma + p$ par échange de photons, qui fait intervenir un terme de couplage à quatre photons via une diffraction $\gamma\gamma$, mieux connue sous le nom de *light-by-light scattering* (LbyL). La LbyL, qui fut beaucoup étudiée à basse énergie lors de la construction de la QED car purement issue de fluctuations quantiques, n'a pour le moment jamais pu être observée directement. Très peu de contraintes existent donc sur ses taux de production en général et quasiment aucune aux énergies du LHC. En particulier, la LbyL est sensible à haute énergie à l'existence de particules chargées massives telles que celles apparaissant dans les modèles de Higgs composite ou encore à l'existence de dimensions supplémentaires de l'espace. Une étude complète des différents bruits de fond attendus est réalisée et montre que les futurs détecteurs de protons à l'avant permettront son élimination totale, y compris pour les taux d'empilement très importants attendus au LHC ($\langle \mu \rangle \simeq 50$). Une étude de sensibilité générique vis-à-vis des différents modèles de nouvelle physique est menée à $\sqrt{s} = 14$ TeV à partir de signaux paramétrés à l'aide d'un modèle effectif, mais également à partir du calcul complet des amplitudes au LO dans le cas de nouvelles particules chargées (vecteur, fermion). Les amplitudes génériques complètes pour les résonances de spin 0 ou 2 ont également été implémentées et devraient mener à une publication prochaine. Les sensibilités sont démontrées comme étant compétitives avec les recherches directes actuelles et à venir au LHC et même parfois supérieures, en particulier pour les modèles de dimensions supplémentaires. Par ailleurs, les sensibilités obtenues sont largement transposables à n'importe quel modèle de nouvelle physique affectant la diffusion $\gamma\gamma$. La LbyL présente également un très bon potentiel de découverte dans le cas de l'existence de particules chargées avec des spins élevés, qui nécessitent néanmoins le développement d'un formalisme adapté afin de pouvoir être étudiés plus en détail.

Finalement, les résultats principaux obtenus lors des tests effectués avec la puce SAMPIC sont rapportés dans la partie VI [329]. SAMPIC se base sur le concept novateur du *Waveform TDC* pour effectuer des mesures de temps avec une résolution de l'ordre de quelques picosecondes. De telles résolutions sur la mesure de temps sont très utiles dans de nombreux domaines, en particulier en physique des hautes énergies afin de pouvoir par exemple distinguer le vertex primaire des vertex d'empilement à l'aide de mesures à l'avant du faisceau, et en imagerie médicale, par exemple pour les tomographies par émission de positrons pour rejeter les coïncidences diffusées et

fortuites. J'ai participé aux premiers tests de la puce à l'aide de signaux de générateurs et de signaux de détecteurs ultra-rapides au silicium. Les données ont ensuite pu être analysées grâce à un code dédié réalisé pendant les années de thèse. Des résultats très encourageants ont été obtenus, montrant que la puce SAMPIC est capable de réaliser des mesures de temps avec une résolution RMS de 4 (40) ps dans le cas des signaux de générateurs (de détecteurs). Différents algorithmes de reconstruction de temps ont également pu être comparés, ouvrant des perspectives intéressantes pour les mesures effectuées par des *Waveform TDC* à bas signal sur bruit. SAMPIC prend désormais des mesures au LHC dans le cadre de l'expérience TOTEM et devrait être utilisé pour les futurs détecteurs de protons intacts à l'avant de CMS-TOTEM (CT-PPS).

Dans l'attente des nouvelles données du LHC et des possibles surprises qu'elles apporteront, j'espère que ce manuscrit, résumant le travail accompli pendant mes années de thèse, aura été utile à ses lecteurs et donnera envie à de futurs étudiants de poursuivre les mesures de précision basées sur les photons, au LHC et ailleurs.

A Corrections supplémentaires de l'isolation calorimétrique basées sur les données

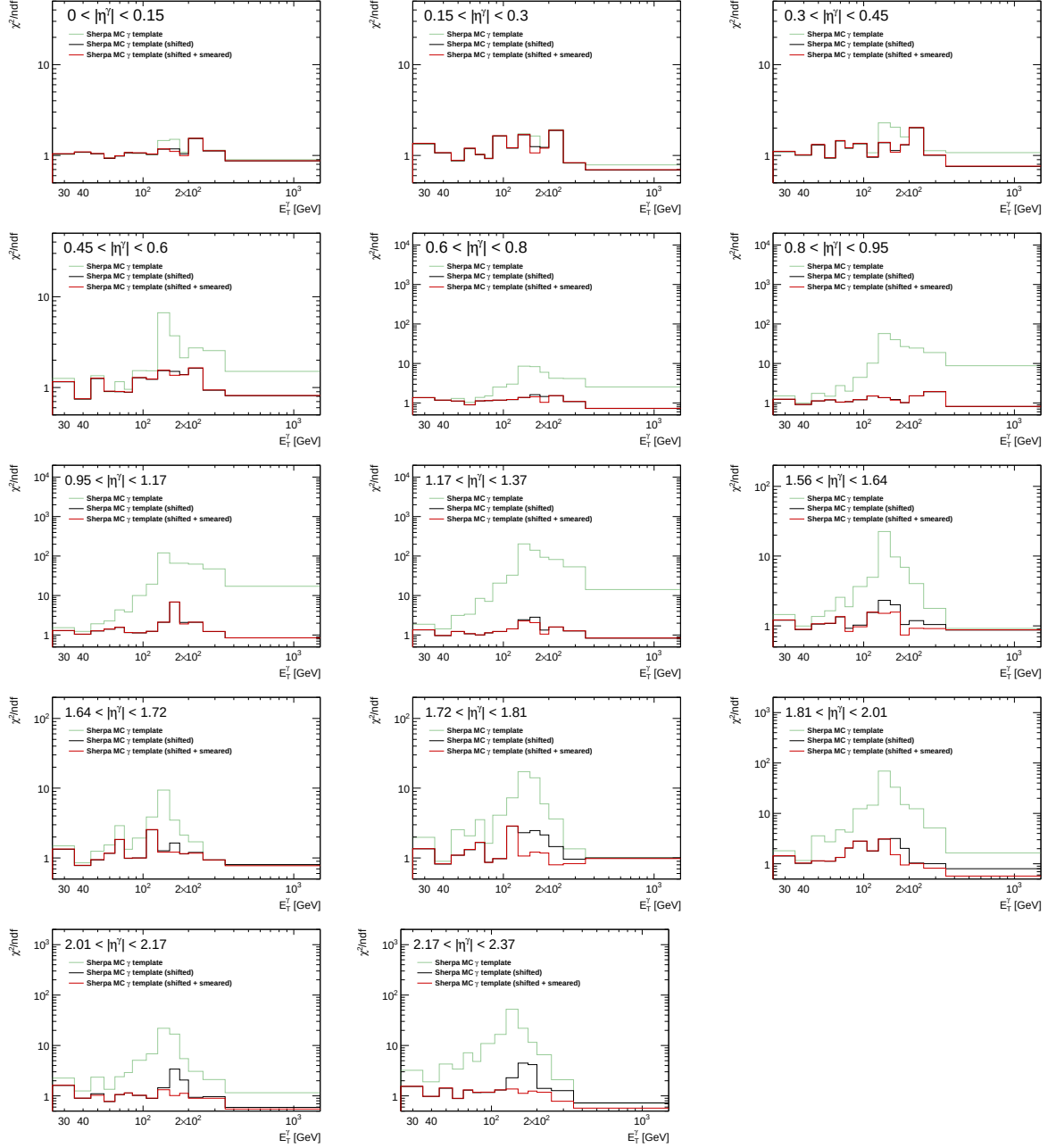


FIG. 230 – χ^2 par degré de liberté des différents ajustements aux données en fonction de l'impulsion transverse du candidat photon dans les différentes catégories en $|\eta^\gamma|$ utilisées dans l'étude. Les résultats basés sur les trois pdfs d'isolation décrites dans la section 10.4 sont représentés : SHERPA MC γ template (vert), SHERPA MC γ template (shifted) (noir) et SHERPA MC γ template (shifted + smeared) (rouge).

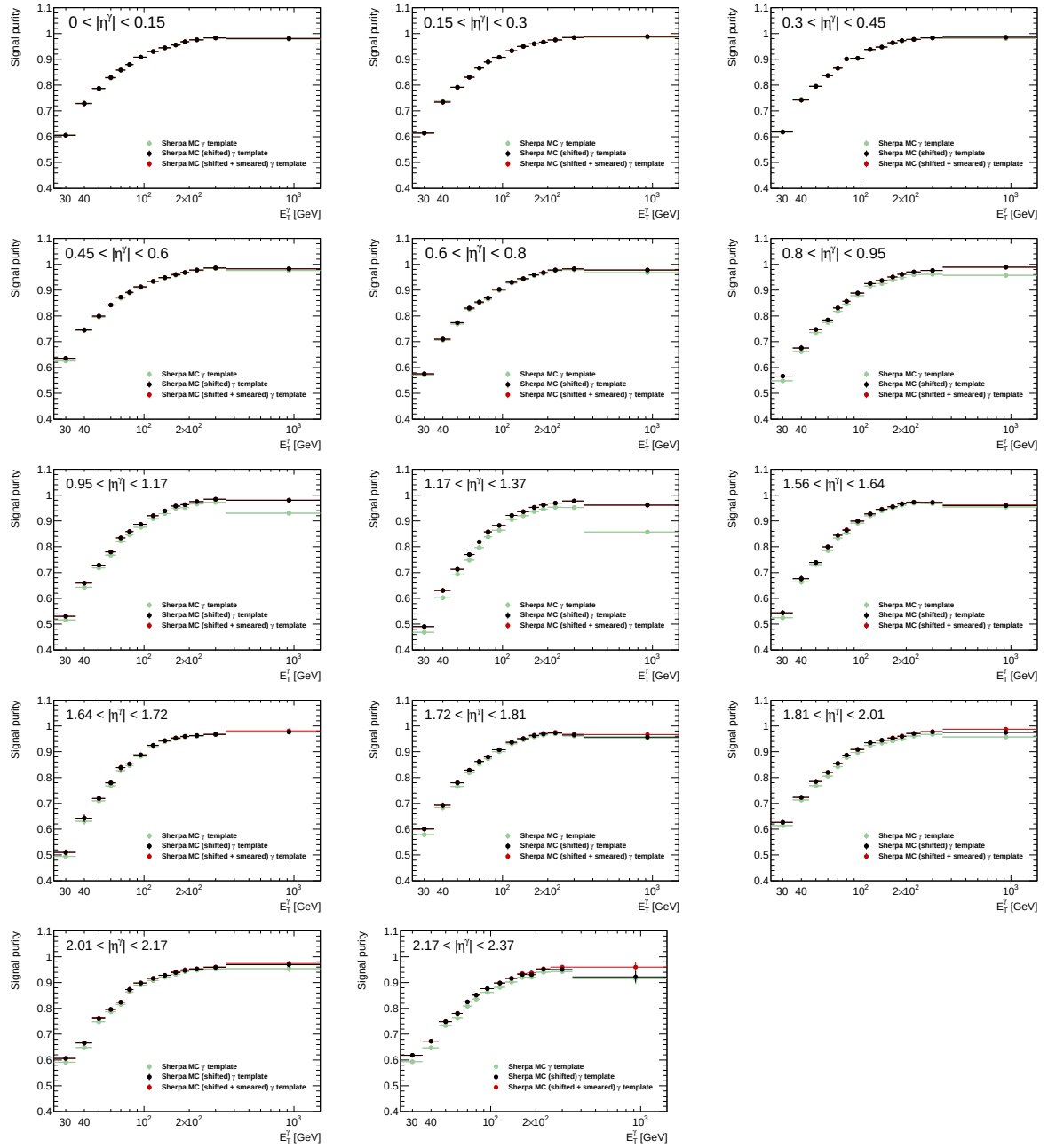


FIG. 231 – Pureté en photon immédiat dans la gamme d'isolation calorimétrique $[-2 \text{ GeV}, 4 \text{ GeV}]$ en fonction de l'impulsion transverse du candidat photon dans les différentes catégories en $|\eta^\gamma|$ utilisées dans l'étude. Les résultats basés sur les trois pdfs d'isolation décrites dans la section 10.4 sont représentés : SHERPA MC (vert), SHERPA MC translaté (noir) et SHERPA MC translaté et corrigé et largeur (rouge). Les barres d'incertitudes sont statistiques uniquement. Les résultats ne sont pas corrigés de la fuite de signal dans la région de contrôle.

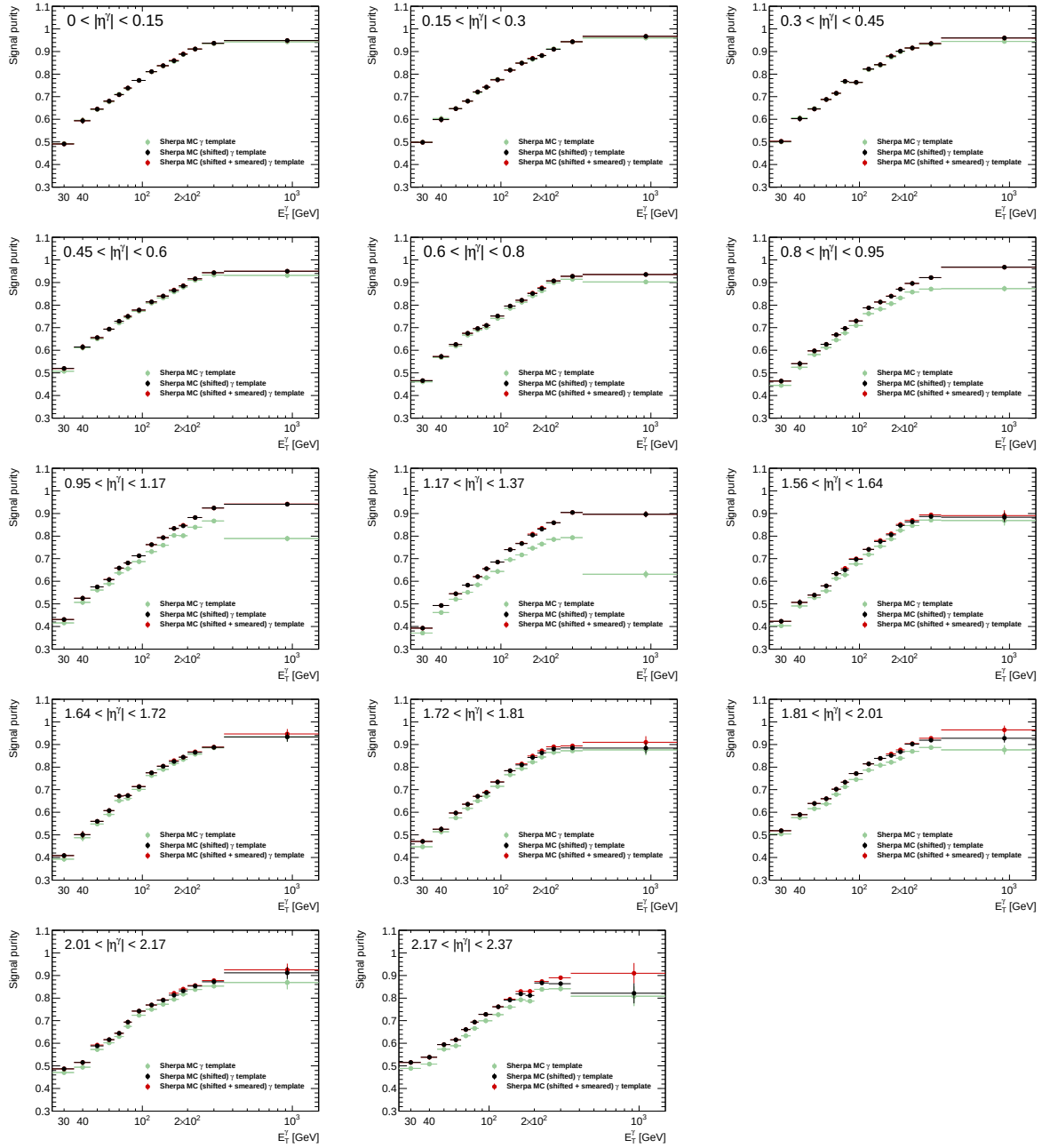


FIG. 232 – Figure similaire à la figure 231 pour une pureté définie dans la gamme d'isolation calorimétrique étendue à $[-2 \text{ GeV}, 10 \text{ GeV}]$.

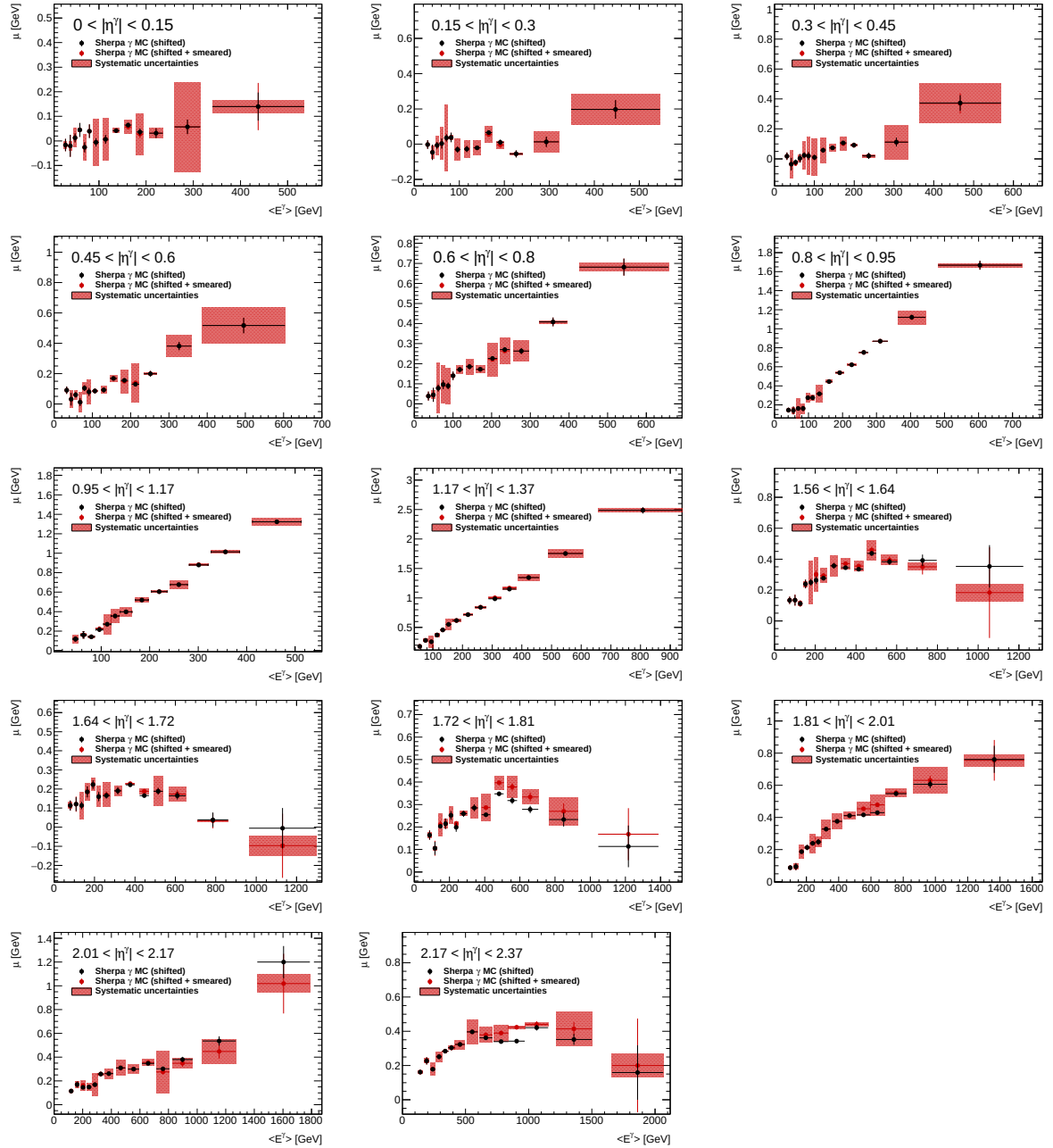


FIG. 233 – Valeurs du paramètre μ (voir section 10.4) obtenues par un ajustement du modèle aux données en fonction de l'énergie moyenne de chaque catégorie en E_T^γ dans les différentes catégories en $|\eta^\gamma|$ utilisées dans l'étude. Les résultats basés sur les deux pdfs d'isolation corrigées décrites dans la section 10.4 sont représentés : SHERPA MC translaté (noir) et SHERPA MC translaté et corrigé et largeur (rouge). Les barres d'incertitudes sont statistiques uniquement tandis que les incertitudes systématiques sont représentées pour le cas SHERPA MC translaté et corrigé et largeur par des rectangles rouges.

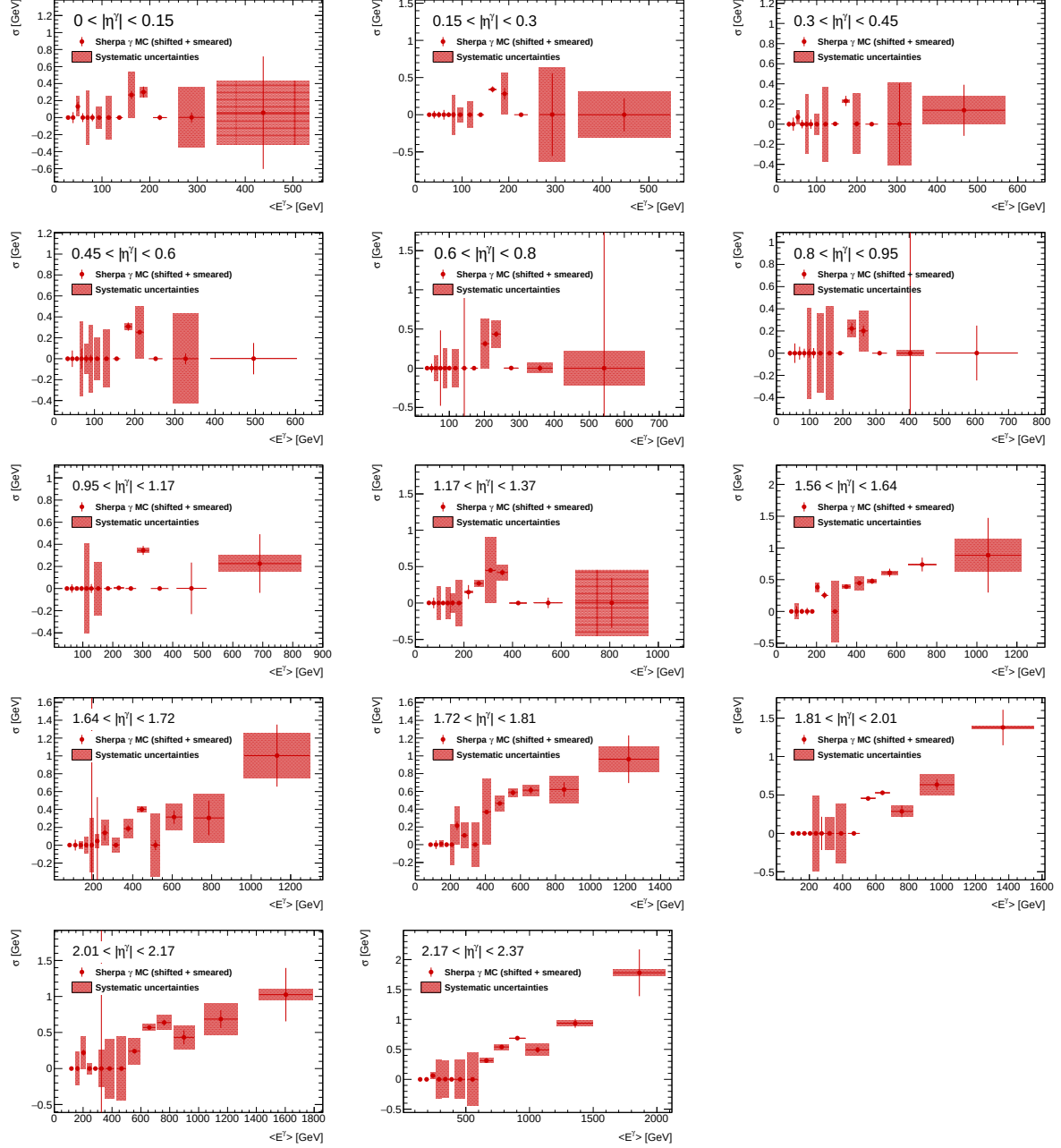


FIG. 234 – Valeurs du paramètre σ (voir section 10.4) obtenues par un ajustement du modèle aux données en fonction de l'énergie moyenne de chaque catégorie en E_T^γ dans les différentes catégories en $|\eta^\gamma|$ utilisées dans l'étude. Les résultats basés sur la pdf d'isolation corrigée en largeur sont représentés. Les barres d'incertitudes sont statistiques uniquement tandis que les incertitudes systématiques sont représentées par des rectangles rouges.

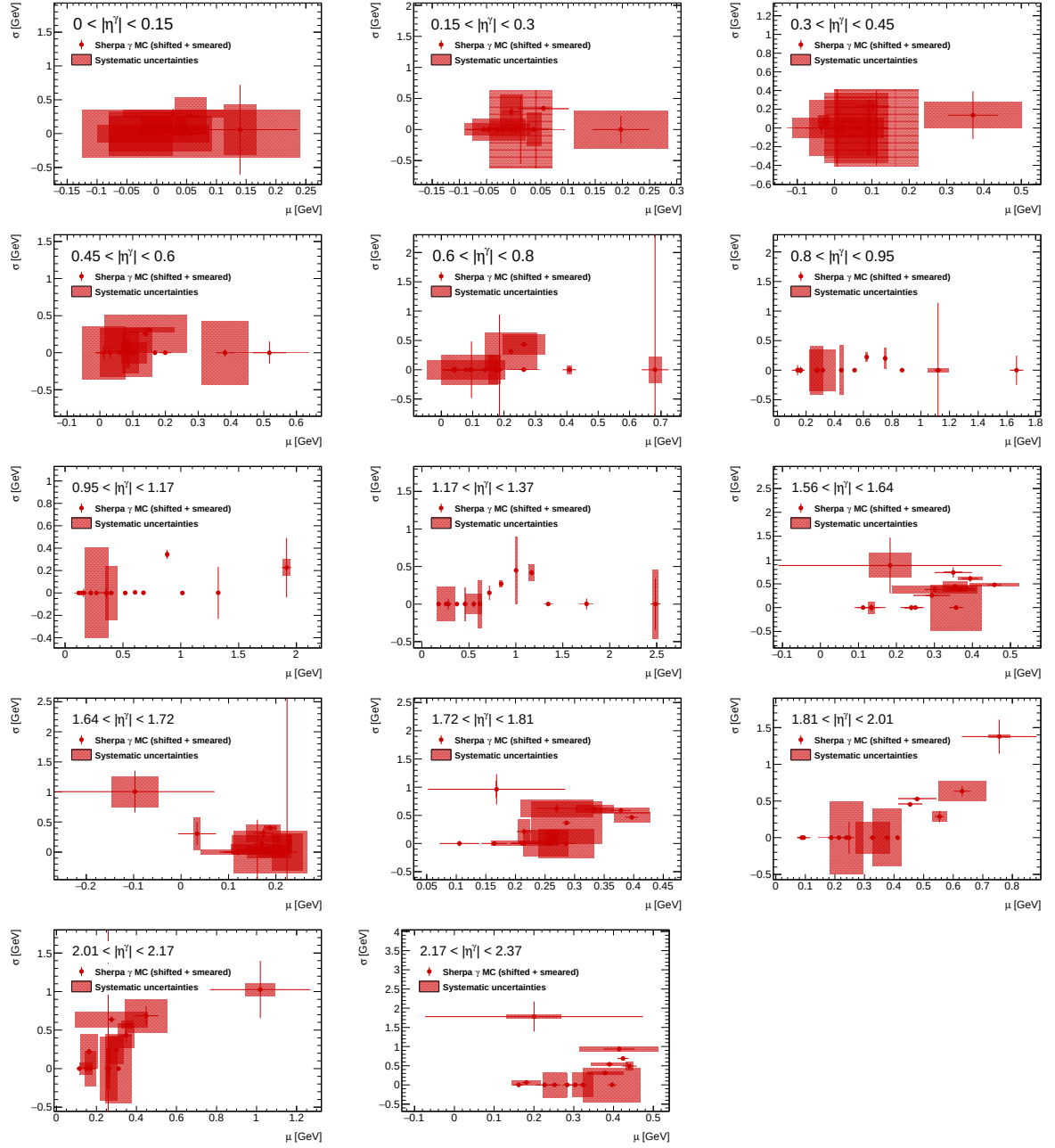


FIG. 235 – Paramètre σ donné en fonction du paramètre μ dans les différentes catégories en $|\eta^\gamma|$ utilisées dans l'étude. Les résultats basés sur la pdf d'isolation corrigée en largeur sont représentés. Les barres d'incertitudes sont statistiques uniquement tandis que les incertitudes systématiques sont représentées par des rectangles rouges.

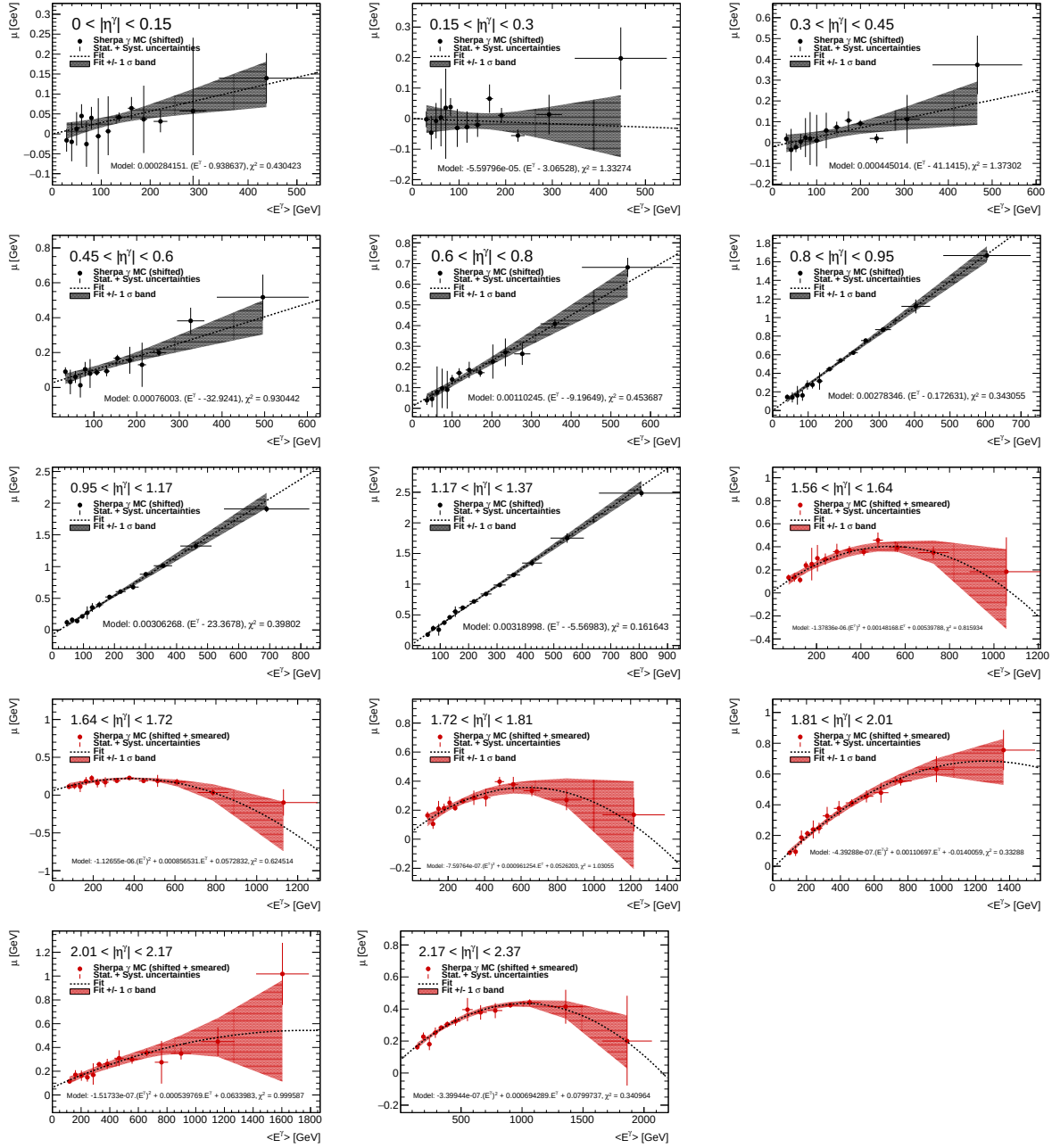


FIG. 236 – Valeurs du paramètre μ (voir section 10.4) obtenues par un ajustement du modèle aux données en fonction de l'énergie moyenne de chaque catégorie en E_T^γ dans les différentes catégories en $|\eta^\gamma|$ utilisées dans l'étude, représentées avec l'interpolation utilisée pour le calcul des corrections finales. Celle-ci est une fonction linéaire dans le tonneau et une parabole dans les bouchons. Les résultats utilisant la pdf d'isolation translatée sont représentés pour le tonneau tandis que ceux utilisant la pdf d'isolation translatée et corrigée en largeur sont représentés dans les bouchons. Les barres d'incertitudes représentent l'incertitude totale, obtenue en sommant en quadrature les incertitudes statistiques et systématiques. Le χ^2 par degré de liberté de l'ajustement final est également mentionné.

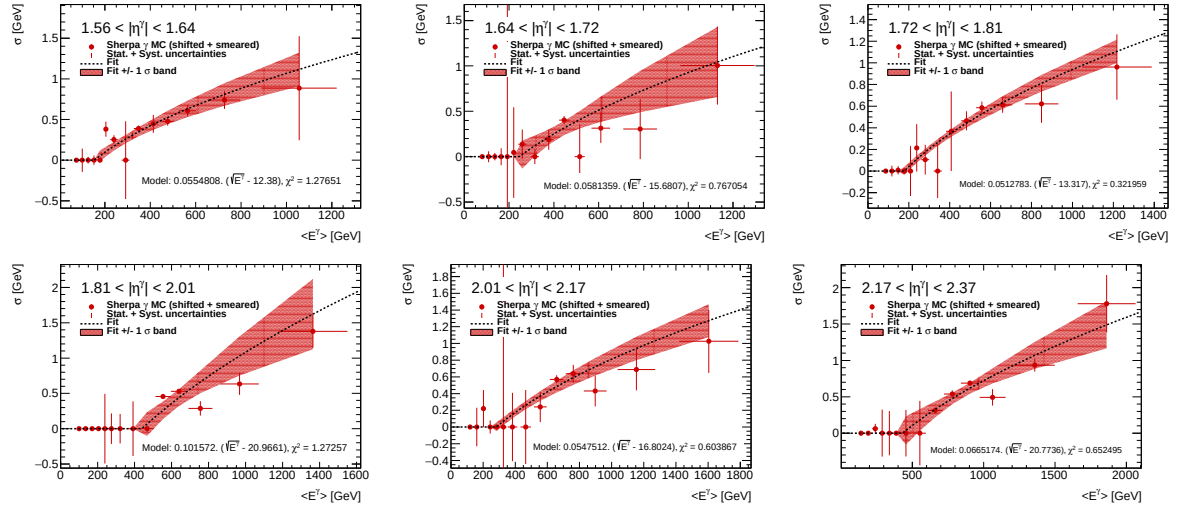


FIG. 237 – Valeurs du paramètre σ (voir section 10.4) obtenues par un ajustement du modèle aux données en fonction de l'énergie moyenne de chaque catégorie en E_T^γ dans les différentes catégories en $|\eta^\gamma|$ des bouchons utilisées dans l'étude, représentées avec l'interpolation utilisée pour le calcul des corrections finales. σ est fixé à 0 dans le tonneau. Les résultats utilisant la pdf d'isolation translatée et corrigée en largeur sont représentés. Les barres d'incertitudes représentent l'incertitude totale, obtenue en sommant en quadrature les incertitudes statistiques et systématiques. Le χ^2 par degré de liberté de l'ajustement final est également mentionné.

B Estimations du taux de mauvaise identification des jets en photons dans les données

Il est possible de déterminer grossièrement la probabilité pour un jet mal identifié en photon LOOSE' de passer la sélection TIGHT à partir des données, en se plaçant dans un région suffisamment non-isolée pour que la contamination en photons immédiats soit négligeable.

Ainsi, des régions non-isolées ($\tilde{\mathbf{I}}$) sont définies par le critère d'isolation suivant : $9 < E_{\mathbf{T}}^{\text{calo.iso.}} < 24$ GeV et $6 < p_{\mathbf{T}}^{\text{track.iso.}} < 22$ GeV. Celui-ci a été réglé afin que la contamination en photons immédiats soit négligeable. Si, de plus, on considère l'identification non-corrélée avec l'isolation, on obtient :

$$\lambda = \frac{n_{\text{fake}}^{\tilde{\mathbf{T}}}}{n_{\text{fake}}^{\mathbf{T}}} = \frac{n_{\text{fake}}^{\tilde{\mathbf{T}}\mathbf{I}}}{n_{\text{fake}}^{\mathbf{T}\mathbf{I}}} = \frac{n_{\text{fake}}^{\tilde{\mathbf{T}}\tilde{\mathbf{I}}}}{n_{\text{fake}}^{\mathbf{T}\tilde{\mathbf{I}}}} \quad (115)$$

On peut alors définir l'efficacité pour un jet mal identifié en photon LOOSE' de passer TIGHT comme :

$$\epsilon_{T|L'}^{\text{fake}} = \frac{n_{\text{fake}}^{\mathbf{T}\mathbf{I}}}{(n_{\text{fake}}^{\mathbf{T}\mathbf{I}} + n_{\text{fake}}^{\tilde{\mathbf{T}}\mathbf{I}})} = \frac{1}{(1 + \lambda)} \quad (116)$$

L'exercice peut être réalisé à la fois pour le premier et le second photon. Le statut de l'autre candidat est alors fixé de façon à sélectionner des jets mal identifiés émis dans le cadre d'un processus γ +jet ($\mathbf{T}\mathbf{I}\tilde{\mathbf{T}}/\mathbf{T}\tilde{\mathbf{I}}\tilde{\mathbf{I}}$), jet+ γ ($\mathbf{T}\tilde{\mathbf{I}}\mathbf{I}/\tilde{\mathbf{T}}\tilde{\mathbf{I}}\mathbf{I}$) ou di-jet ($\tilde{\mathbf{I}}\tilde{\mathbf{I}}$), suivant une méthode de type *tag and probe*.

Les chiffres donnés dans la figure 98, section 15.1 et la méthode de la *4x4 matrix* discutée dans la section 15.7.3 s'appuient sur ces principes.

C Étude des différences entre les effets d'empilement dans les données et dans la simulation d'un point de vue calorimétrique

La simulation de l'empilement dans ATLAS est effectuée à partir d'événements de biais minimaux générés avec PYTHIA8 (*minimum bias events*, voir section 21.2.3) [82, 235]. Les distributions exactes de $\langle \mu \rangle$ dans les données complètes étant souvent encore partiellement inconnues lors du démarrage de la production MC, une distribution proche et généralement légèrement plus large que celle attendue est adoptée. Une repondération des événements MC est alors implémentée a posteriori de la production afin d'identifier leur nombre moyen d'interactions par croisement de faisceaux à celui des données. Cette identification est faite période d'acquisition par période d'acquisition (*run number*).

Cependant, il a été observé dans ATLAS qu'une repondération de $\langle \mu \rangle_{MC}$ à $\langle \mu \rangle_{data}$ ne permettait pas de reproduire de façon satisfaisante la distribution du nombre de vertex reconstruits N_{PV} des données. Cela est attribué principalement à la différence entre la couverture d'ATLAS mesurée dans les données et celle renvoyée par PYTHIA8 pour la production inélastique totale proton-proton ($d\sigma_{inel.}/d\eta$).

Ainsi, il est nécessaire de définir un facteur d'échelle de telle sorte que $\langle \mu \rangle_{data} = SF_{\mu} \cdot \langle \mu \rangle_{MC}$, avec SF_{μ} dépendant de la couverture en η de l'analyse considérée. Ce facteur d'échelle a été mesuré à bas $\langle \mu \rangle$ à l'aide du scintillateur à l'avant (MBTS, $|\eta| < 3.9$) et à bas et haut $\langle \mu \rangle$ avec le détecteur interne ($|\eta| < 2.37$, voir section 5.3). L'étude utilisant le détecteur interne est validée sur la bonne description de la distribution du nombre de vertex reconstruits N_{PV} après application du SF. La valeur obtenue avec les données 2012 est de 1.09 ± 0.04 , ce qui signifie qu'un événement d'empilement MC a la même activité que 1.09 événement d'empilement dans les données en moyenne. Après repondération, les distributions de $\sum p_T^{track.}$ et $\sum (p_T^{track.})^2$ présentent un accord satisfaisant avec les données. Le facteur d'échelle est plus proche de 1.13 pour l'acceptance du MBTS.

Cependant, le détecteur interne est transparent pour les particules neutres et présente une fenêtre réduite en η . Ainsi, l'activité pourrait également être mesurée à partir du calorimètre, qui lui est sensible à l'ensemble des particules (sauf les muons et les neutrinos) et couvre une gamme en pseudo-rapacité plus large ($|\eta| < 3.2$). Cette comparaison pourrait se faire par exemple à l'aide de la médiane et de la RMS de la densité d'énergie ambiante (voir section 10.3.2).

Ces distributions¹³⁸ sont comparées entre données et MC après l'application du facteur d'échelle de 1.09 pour les événements de la région **TITI** (voir section 14) dans la figure 238. Les distributions de l'observable "médiane/RMS" sont également comparées. On observe en moyenne (médiane) une densité d'énergie ambiante plus importante et plus large (RMS) dans le MC repondéré que dans les données. En revanche, les distributions médiane/RMS sont en très bon accord. Ainsi, la "forme" des distributions de densité d'énergie ambiante paraît en bon accord, mais pas leur échelle absolue.

¹³⁸La densité d'énergie ambiante correspond à une distribution et non à une valeur unique pour un événement donné. On parle donc de la distribution de sa médiane et de sa RMS. Plus de détails sont disponibles dans la section 10.3.2.

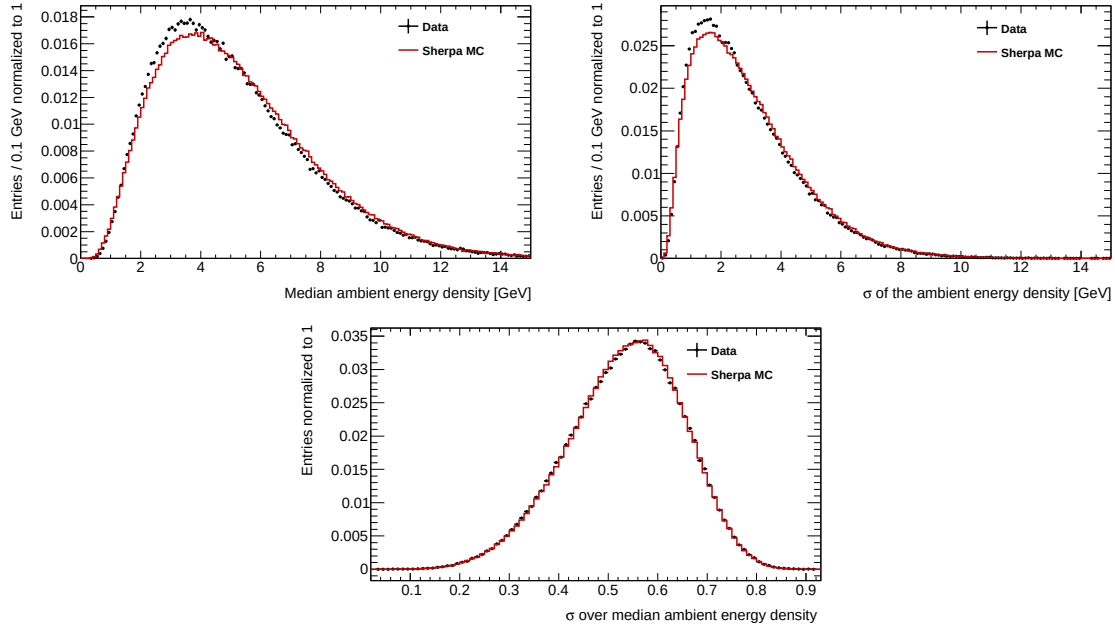


FIG. 238 – Comparaison des distributions de valeur médiane (gauche) et de RMS (droite) de la densité d'énergie ambiante entre les données et les échantillons simulés après repondération nominale ($SF = 1.09$). La distribution médiane/RMS est également représentée (bas) et trouvée en bon accord entre données et simulation.

Dans le cadre du calcul de l'isolation calorimétrique, la médiane est soustraite de $E_{T}^{\text{calo.iso.}}$ de sorte que seule la RMS joue sur les distributions finales. On observe ici une RMS absolue (en GeV) plus importante dans le MC que dans les données tandis que le *template fit* 2D discuté dans la section 15 semble privilégier une augmentation de $\langle \mu \rangle_{MC}$ dans la simulation (baisse du facteur d'échelle) et donc une augmentation des fluctuations de la densité d'énergie ambiante autour de la moyenne (voir figure 239 et la section 15.6.1). Ces deux observations semblent contradictoires à ce stade.

Cependant, la figure 238 n'est pas parfaite car la densité d'énergie ambiante employée dans le calcul de l'isolation calorimétrique est évaluée dans deux catégories indépendantes en pseudo-rapidity, $|\eta| < 1.5$ et $1.5 < |\eta| < 3.0$, en fonction de la pseudo-rapidity du candidat photon. Or, la figure 238 a été obtenue à partir du premier candidat photon de la région **TITI** définie dans la cadre de l'analyse $\gamma\gamma + X$ à $\sqrt{s} = 8$ TeV, qui considère la région $|\eta_1^\gamma| < 2.37$ avec $1.37 < |\eta_1^\gamma| < 1.56$ exclu et donc mélange les deux catégories citées précédemment. Ainsi, une différence au niveau des bouchons et un bon accord au niveau du tonneau ou vice-versa pourrait expliquer cette contradiction apparente.

Dans un premier temps, la figure 239 a été reproduite pour différentes catégories en η_1^γ , présentée dans la figure 240. Selon l'ajustement des distributions d'isolation, la valeur de 1.09 est adaptée uniquement pour $|\eta_1^\gamma| < 0.6$. Pour les autres catégories, une valeur inférieure à 1.05 améliore la qualité de l'ajustement. Ainsi, il semble que le calorimètre indique qu'un simple facteur d'échelle sur μ ne suffit pas pour décrire correctement l'empilement pour le calorimètre. Pour conclure, il serait nécessaire de produire de nouvelles figures correspondant exactement aux deux catégories en pseudo-rapidity utilisées pour la correction de l'énergie d'isolation, *ie.* $|\eta| < 1.5$ et $1.5 < |\eta| < 3.0$. Il pourrait être intéressant également de répéter l'étude effectuée avec le détecteur interne dans des

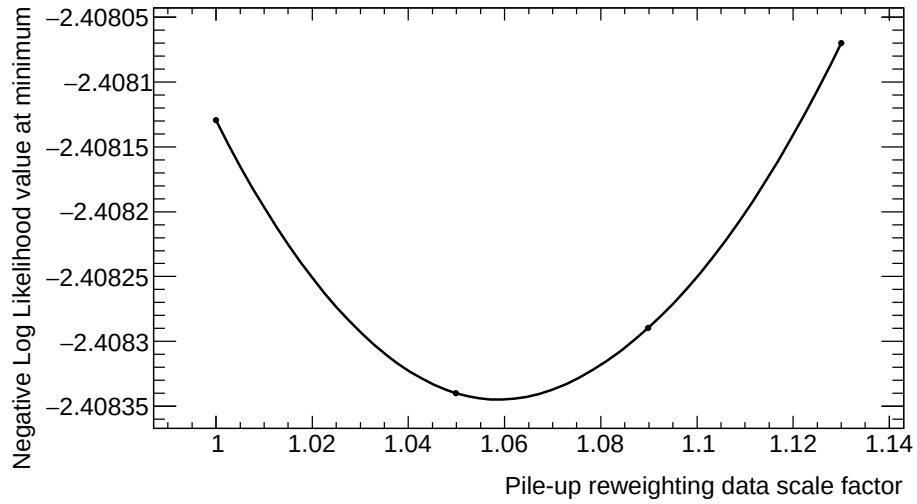


FIG. 239 – Profil de l’estimateur de maximum de vraisemblance pour la région **TITI** complète après ajustement en fonction du facteur d’échelle appliqué lors de la repondération en empilement des échantillons simulés. La valeur recommandée par le groupe de performance du détecteur interne est 1.09 tandis que la valeur privilégiée par l’ajustement se situe autour de 1.06.

catégories en pseudo-rapidité.

Une particularité supplémentaire du calorimètre par rapport au détecteur interne réside dans sa sensibilité plus importante vis-à-vis de l’empilement hors-temps. Il serait intéressant d’estimer l’accord données/simulation sur ce point également, par exemple en regardant l’accord sans aucun empilement hors-temps. Cela est réalisable en pratique à partir de données prises avec très peu d’empilement ou par exemple en se limitant aux croisements correspondant aux premiers de chaque train de paquets dans les données nominales.

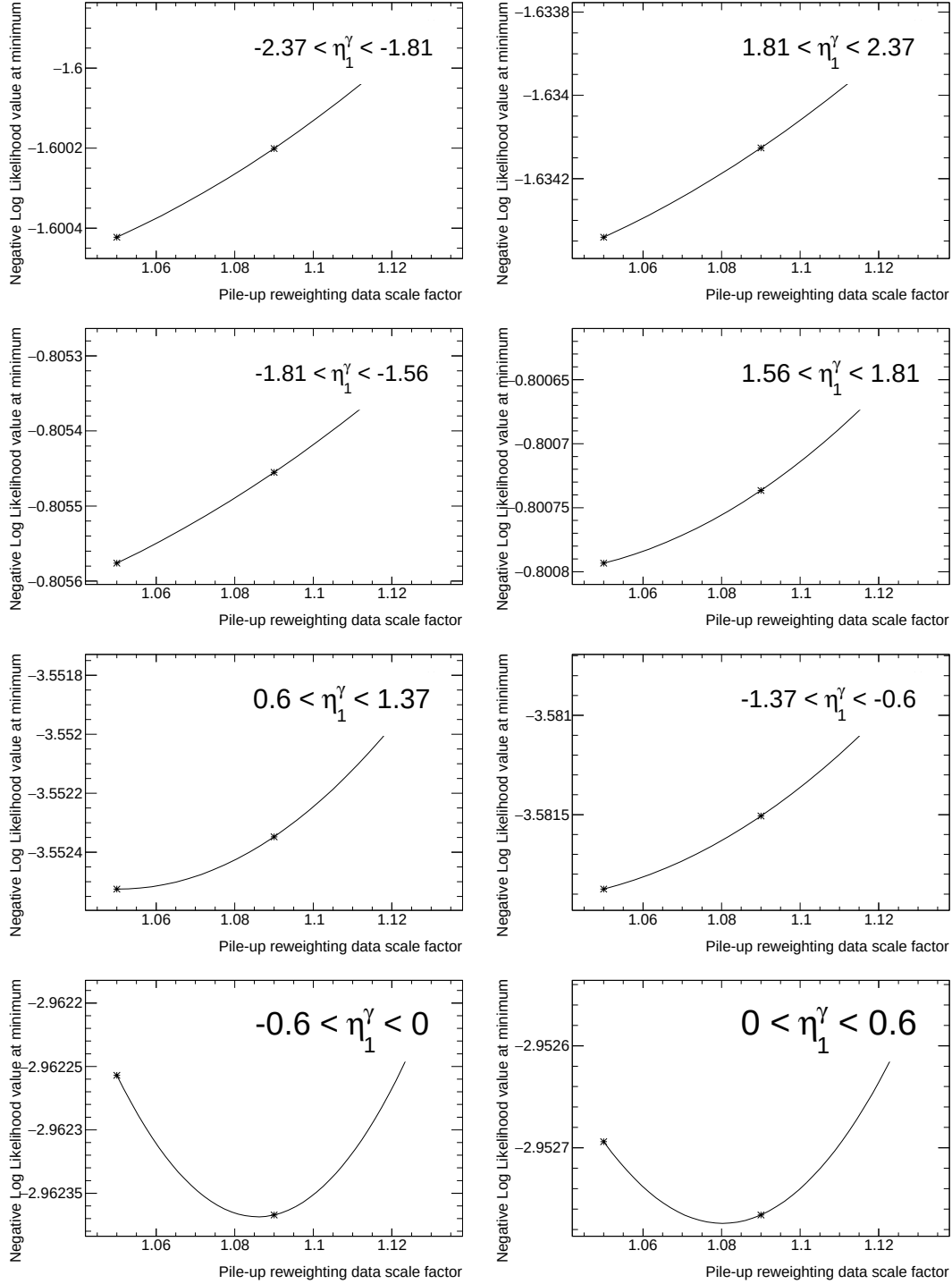


FIG. 240 – Figure similaire à la figure 239 pour la région **TITI** restreinte à un domaine en η_1^γ particulier.

D Résumé des différentes mesures $pp \rightarrow \gamma\gamma + X$ réalisées à $\sqrt{s} = 8$ TeV

Section efficace $pp \rightarrow \gamma\gamma + X$ fiducielle totale (voir section 12.2) à $\sqrt{s} = 8$ TeV :

$$\sigma_{\text{tot}}^{\text{fid.}} = 16.51 \pm 0.07(\text{stat}) \begin{smallmatrix} +0.75 \\ -0.74 \end{smallmatrix}(\text{syst}) \begin{smallmatrix} +0.32 \\ -0.31 \end{smallmatrix}(\text{lumi}) \begin{smallmatrix} +0.03 \\ -0.02 \end{smallmatrix}(\text{trigger}) \text{ pb} \quad (117)$$

Les différentes mesures de sections efficaces différentielles sont résumées dans les tableaux 45 à 50 pour les différentes observables de mesure (voir section 12.3).

$m_{\gamma\gamma}$ [GeV]	$d\sigma/dm_{\gamma\gamma}$ [fb/GeV]	statistical uncertainty [fb/GeV]	systematic uncertainty [fb/GeV]
0-30	4.94	± 0.28	+0.59 -0.59
30-50	16.02	± 0.66	+0.99 -0.53
50-70	29.69	± 0.89	+1.70 -1.15
70-80	139.86	± 7.40	+8.81 -4.82
80-90	228.99	± 4.46	+13.30 -7.14
90-100	227.76	± 3.81	+10.63 -5.90
100-110	182.25	± 2.96	+7.98 -4.52
110-120	144.40	± 2.52	+5.77 -4.06
120-130	119.75	± 3.31	+4.12 -2.53
130-140	91.87	± 1.95	+2.85 -1.31
140-150	68.89	± 1.83	+2.18 -0.98
150-160	56.95	± 1.66	+1.68 -0.77
160-170	44.26	± 1.55	+1.24 -1.15
170-180	38.67	± 1.44	+1.12 -0.72
180-190	28.39	± 1.26	+0.92 -0.43
190-200	25.05	± 0.98	+0.93 -0.51
200-225	18.51	± 0.56	+0.55 -0.23
225-250	12.05	± 0.40	+0.50 -0.33
250-275	8.56	± 0.33	+0.33 -0.39
275-300	5.55	± 0.28	+0.25 -0.25
300-350	3.07	± 0.19	+0.16 -0.06
350-400	1.79	± 0.11	+0.09 -0.07
400-500	0.79	± 0.05	+0.04 -0.04
500-600	0.26	± 0.03	+0.01 -0.02
600-700	0.129	± 0.016	+0.006 -0.004
700-1700	0.011	± 0.002	+0.001 -0.001

TAB. 45 – Section efficace différentielle $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV selon la masse invariante de la paire de photons $m_{\gamma\gamma}$, avec incertitudes statistiques et systématiques finales.

$p_{T,\gamma\gamma}$ [GeV]	$d\sigma/dp_{T,\gamma\gamma}$ [fb/GeV]	statistical uncertainty [fb/GeV]	systematic uncertainty [fb/GeV]
0-4	217.64	± 5.86	+9.96 -7.99
4-8	419.87	± 8.38	+13.99 -7.88
8-12	453.22	± 8.74	+17.65 -8.50
12-16	408.89	± 8.79	+17.64 -10.19
16-20	343.89	± 7.55	+15.26 -5.03
20-25	279.51	± 5.73	+12.90 -8.67
25-30	212.81	± 5.09	+10.54 -6.52
30-35	180.60	± 4.68	+7.73 -3.17
35-40	147.04	± 4.13	+6.30 -2.22
40-45	126.58	± 5.03	+4.23 -3.57
45-50	100.01	± 3.48	+4.68 -2.92
50-55	90.81	± 3.10	+5.39 -4.40
55-60	77.77	± 3.15	+6.16 -5.09
60-65	66.92	± 2.81	+5.75 -4.92
65-70	67.26	± 2.85	+5.49 -4.88
70-75	63.87	± 3.08	+4.87 -4.02
75-80	59.23	± 2.98	+4.11 -2.99
80-90	45.01	± 1.54	+2.14 -1.55
90-100	32.61	± 1.27	+1.43 -0.51
100-110	25.12	± 0.93	+0.95 -0.66
110-120	16.80	± 0.79	+0.82 -0.50
120-130	11.38	± 0.64	+0.53 -0.29
130-140	10.16	± 0.57	+0.36 -0.26
140-150	6.44	± 0.47	+0.23 -0.24
150-175	4.58	± 0.21	+0.20 -0.17
175-200	2.32	± 0.17	+0.14 -0.12
200-225	1.35	± 0.12	+0.06 -0.07
225-250	0.97	± 0.08	+0.04 -0.02
250-300	0.42	± 0.05	+0.02 -0.03
300-400	0.107	± 0.017	+0.007 -0.007
400-750	0.012	± 0.003	+0.001 -0.001

TAB. 46 – Section efficace différentielle $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV selon l'impulsion transverse de la paire de photons $p_{T,\gamma\gamma}$, avec incertitudes statistiques et systématiques finales.

a_T [GeV]	$d\sigma/da_T$ [fb/GeV]	statistical uncertainty [fb/GeV]	systematic uncertainty [fb/GeV]
0-2	906.11	± 14.53	+50.97 -45.80
2-4	792.05	± 14.00	+47.13 -37.29
4-6	684.55	± 12.61	+39.10 -33.24
6-8	565.99	± 11.81	+23.55 -20.11
8-10	478.06	± 11.14	+20.51 -10.54
10-12	424.23	± 10.66	+16.97 -9.17
12-14	366.87	± 10.59	+16.07 -8.51
14-16	317.80	± 10.12	+13.67 -3.45
16-18	287.55	± 9.02	+10.53 -5.63
18-20	257.73	± 7.86	+10.26 -5.24
20-25	207.93	± 4.40	+7.51 -4.35
25-30	160.43	± 3.94	+6.14 -1.01
30-35	129.96	± 3.84	+5.01 -1.50
35-40	107.87	± 3.59	+4.42 -1.93
40-45	88.68	± 3.28	+3.36 -1.41
45-50	82.23	± 3.13	+3.72 -1.55
50-55	74.73	± 2.95	+3.25 -1.49
55-60	75.21	± 3.21	+3.83 -1.09
60-65	77.53	± 3.11	+3.17 -1.74
65-70	59.15	± 2.78	+3.02 -2.19
70-75	46.61	± 2.53	+3.29 -2.34
75-80	32.12	± 2.09	+2.25 -1.71
80-90	22.93	± 0.95	+1.41 -1.20
90-100	14.26	± 0.68	+0.85 -0.61
100-110	7.99	± 0.50	+0.41 -0.30
110-120	4.98	± 0.40	+0.28 -0.21
120-130	3.06	± 0.33	+0.15 -0.14
130-150	2.00	± 0.13	+0.07 -0.07
150-200	0.62	± 0.05	+0.03 -0.05
200-450	0.049	± 0.007	+0.003 -0.001

TAB. 47 – Section efficace différentielle $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV selon la variable a_T (voir section 12.3), avec incertitudes statistiques et systématiques finales.

ϕ_η^*	$d\sigma/d\phi_\eta^*$ [fb]	statistical uncertainty [fb]	systematic uncertainty [fb]
0.001-0.004	111102.27	± 3768.44	+5636.74 -7037.19
0.004-0.008	110218.34	± 3985.46	+6116.84 -5888.93
0.008-0.012	103463.49	± 4283.98	+6506.61 -5899.25
0.012-0.016	105145.33	± 4658.36	+6153.43 -4926.98
0.016-0.020	90300.05	± 4018.00	+5120.70 -4241.04
0.020-0.024	96938.14	± 3968.35	+5964.97 -4801.25
0.024-0.029	90655.48	± 3391.90	+5526.62 -5233.89
0.029-0.034	85711.40	± 3369.77	+5053.30 -4081.09
0.034-0.039	82864.14	± 3472.86	+4609.17 -3764.23
0.039-0.045	74837.95	± 3017.17	+3830.66 -3871.00
0.045-0.051	71022.32	± 2919.40	+4309.91 -2996.70
0.051-0.057	68546.84	± 2623.32	+3571.76 -3845.18
0.057-0.064	64925.52	± 2304.57	+3175.30 -2375.80
0.064-0.072	60120.47	± 2161.99	+2894.14 -2706.54
0.072-0.081	56251.75	± 1897.10	+2524.02 -1762.02
0.081-0.091	47519.12	± 1609.39	+2103.14 -2526.39
0.091-0.102	46307.09	± 1562.05	+2055.61 -1239.23
0.102-0.114	44291.55	± 1435.35	+1755.86 -856.65
0.114-0.128	37831.23	± 1252.37	+1598.06 -1022.99
0.128-0.145	34238.67	± 1043.98	+1484.02 -475.30
0.145-0.165	29328.99	± 912.89	+1233.28 -449.58
0.165-0.189	26115.54	± 752.86	+1011.38 -1035.83
0.189-0.219	20656.49	± 624.02	+766.97 -461.38
0.219-0.258	17301.54	± 440.53	+698.83 -224.82
0.258-0.312	13504.85	± 328.10	+513.69 -225.78
0.312-0.391	9380.58	± 218.25	+356.63 -53.50
0.391-0.524	5996.92	± 137.41	+232.14 -182.23
0.524-0.695	3377.08	± 90.93	+148.32 -80.25
0.695-0.918	1991.11	± 57.47	+89.28 -35.95
0.918-1.153	1160.49	± 41.73	+50.43 -14.11
1.153-1.496	755.47	± 33.13	+34.66 -26.86
1.496-1.947	484.13	± 19.78	+31.55 -18.74
1.947-2.522	304.62	± 14.28	+24.95 -18.14
2.522-3.277	176.05	± 9.93	+17.99 -13.51
3.277-5	104.94	± 4.74	+9.55 -8.29
5-10	27.85	± 1.48	+2.70 -2.70
10-20	6.13	± 0.47	+0.55 -0.50
20-50	1.29	± 0.12	+0.10 -0.11
50-100	0.27	± 0.05	+0.02 -0.02
100-50000	0.00019	± 0.00004	+0.00001 -0.00002

TAB. 48 – Section efficace différentielle $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV selon la variable ϕ_η^* (voir section 12.3), avec incertitudes statistiques et systématiques finales.

$ \cos \theta_\eta^* $	$d\sigma/d \cos \theta_\eta^* $ [fb]	statistical uncertainty [fb]	systematic uncertainty [fb]
0.00-0.04	24351.60	± 519.25	+1174.39 -714.67
0.04-0.08	23851.03	± 505.58	+1063.22 -625.65
0.08-0.12	21491.38	± 494.67	+1020.82 -573.67
0.12-0.16	23035.46	± 508.13	+1018.16 -787.62
0.16-0.20	22505.46	± 496.15	+1123.82 -781.19
0.20-0.24	21279.69	± 463.64	+1034.56 -767.54
0.24-0.28	21470.22	± 474.65	+1105.02 -707.62
0.28-0.32	20028.14	± 483.87	+1027.55 -653.96
0.32-0.36	19550.79	± 475.14	+905.08 -660.65
0.36-0.40	18750.35	± 441.89	+880.14 -565.52
0.40-0.44	19003.60	± 461.15	+878.21 -523.90
0.44-0.48	19407.39	± 448.92	+964.74 -495.04
0.48-0.52	18729.58	± 465.78	+828.98 -503.25
0.52-0.56	18904.74	± 435.71	+708.86 -500.30
0.56-0.60	17721.84	± 409.92	+674.58 -365.05
0.60-0.64	16557.16	± 548.98	+633.73 -350.12
0.64-0.68	15216.20	± 392.97	+586.25 -200.29
0.68-0.72	14029.35	± 381.52	+458.88 -259.11
0.72-0.76	13208.76	± 372.08	+506.15 -128.66
0.76-0.80	11834.79	± 337.82	+402.78 -194.20
0.80-0.84	10582.77	± 346.01	+359.11 -304.74
0.84-0.88	8013.27	± 276.63	+338.68 -342.06
0.88-0.92	7108.08	± 249.87	+310.40 -307.85
0.92-0.96	4136.79	± 215.12	+190.44 -197.27
0.96-1.00	1050.63	± 122.06	+95.49 -102.34

TAB. 49 – Section efficace différentielle $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV selon la variable $|\cos \theta_\eta^*|$ (voir section 12.3), avec incertitudes statistiques et systématiques finales.

$\Delta\phi_{\gamma\gamma}$ [rad.]	$d\sigma/d\Delta\phi_{\gamma\gamma}$ [fb/rad.]	statistical uncertainty [fb/rad.]	systematic uncertainty [fb/rad.]
0.00-0.25	763.59	± 34.58	+57.59 -62.95
0.25-0.50	893.29	± 38.68	+82.03 -72.18
0.50-0.75	973.25	± 38.60	+90.79 -75.08
0.75-1.00	1028.68	± 44.87	+81.63 -66.71
1.00-1.25	1105.95	± 49.05	+60.26 -38.24
1.25-1.50	1236.09	± 44.60	+55.64 -39.49
1.50-1.62	1483.37	± 60.81	+69.32 -51.50
1.62-1.75	1715.70	± 77.15	+75.83 -14.96
1.75-1.88	1995.79	± 75.54	+82.63 -32.56
1.88-2.00	2177.11	± 87.94	+94.84 -32.03
2.00-2.12	2675.28	± 99.50	+128.03 -41.43
2.12-2.25	3322.44	± 103.69	+131.57 -61.53
2.25-2.30	3987.03	± 176.92	+173.48 -223.33
2.30-2.35	4279.53	± 185.64	+229.23 -136.44
2.35-2.40	4757.55	± 182.68	+198.85 -159.93
2.40-2.45	5276.16	± 196.49	+234.73 -240.04
2.45-2.50	5854.14	± 263.16	+240.73 -97.86
2.50-2.55	6847.26	± 261.73	+281.19 -123.33
2.55-2.60	7235.35	± 278.98	+389.91 -156.75
2.60-2.65	8247.08	± 278.61	+395.40 -159.16
2.65-2.67	9102.39	± 402.65	+486.08 -365.04
2.67-2.70	10994.95	± 384.40	+456.54 -400.80
2.70-2.73	10690.38	± 459.02	+471.56 -285.88
2.73-2.75	11160.06	± 446.51	+473.54 -557.31
2.75-2.77	11993.18	± 564.83	+747.78 -295.24
2.77-2.80	13563.60	± 532.20	+654.63 -427.22
2.80-2.83	13786.42	± 561.65	+576.50 -499.67
2.83-2.85	15376.19	± 639.15	+707.22 -131.95
2.85-2.88	17339.17	± 684.70	+808.94 -243.59
2.88-2.90	17449.93	± 635.85	+784.24 -733.97
2.90-2.92	20970.78	± 698.24	+728.03 -447.55
2.92-2.95	21457.45	± 676.46	+847.27 -361.54
2.95-2.98	22812.54	± 666.39	+927.80 -874.17
2.98-3.00	26697.34	± 737.41	+1141.94 -824.97
3.00-3.02	30606.13	± 761.57	+1304.38 -962.03
3.02-3.05	31763.37	± 873.85	+1466.48 -959.73
3.05-3.08	35584.05	± 879.70	+1811.56 -1127.51
3.08-3.10	39830.77	± 856.97	+1863.28 -1438.26
3.10-3.12	45182.60	± 1127.24	+2062.43 -1604.76
3.12-3.14	46257.71	± 941.40	+2464.35 -2520.34

TAB. 50 – Section efficace différentielle $pp \rightarrow \gamma\gamma + X$ à $\sqrt{s} = 8$ TeV selon la différence d'angles d'azimut $\Delta\phi_{\gamma\gamma}$, avec incertitudes statistiques et systématiques finales.

Références

- [1] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, LHC Design Report. CERN, Geneva, 2004.
- [2] **ATLAS** Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1–29, [[arXiv :1207.7214](#)].
- [3] **CMS** Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61, [[arXiv :1207.7235](#)].
- [4] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (Aug, 1964) 321–323.
- [5] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (Oct, 1964) 508–509.
- [6] F. Wilczek, *Problem of Strong p and t Invariance in the Presence of Instantons*, Phys. Rev. Lett. **40** (1978) 279–282.
- [7] S. Fichtel and G. von Gersdorff, *Anomalous gauge couplings from composite Higgs and warped extra dimensions*, JHEP **03** (2014) 102, [[arXiv :1311.6815](#)].
- [8] **ATLAS** Collaboration, *Search for resonances decaying to photon pairs in 3.2 fb⁻¹ of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, .
- [9] **CMS** Collaboration, *Search for new physics in high mass diphoton events in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Tech. Rep. CMS-PAS-EXO-15-004, CERN, Geneva, 2015.
- [10] M. E. Peskin and D. V. Schroeder, An introduction to quantum field theory. Advanced book program. Westview Press Reading (Mass.), Boulder (Colo.), 1995. Autre tirage : 1997.
- [11] F. Halzen and A. Martin, Quarks and leptons : an introductory course in modern particle physics. Wiley, 1984.
- [12] R. P. Feynman, *Space-time approach to quantum electrodynamics*, Phys. Rev. **76** (Sep, 1949) 769–789.
- [13] R. P. Feynman, *Mathematical formulation of the quantum theory of electromagnetic interaction*, Phys. Rev. **80** (Nov, 1950) 440–457.
- [14] J. Schwinger, *On quantum-electrodynamics and the magnetic moment of the electron*, Phys. Rev. **73** (Feb, 1948) 416–417.
- [15] S. Tomonaga, *On a relativistically invariant formulation of the quantum theory of wave fields*, Prog. Theor. Phys. **1** (1946) 27–42.
- [16] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961), no. 4 579 – 588.
- [17] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (Nov, 1967) 1264–1266.
- [18] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367–377.
- [19] R. Ellis, W. Stirling, and B. Webber, QCD and Collider Physics. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 1996.

- [20] LEP design report. CERN, Geneva, 1984. Copies shelved as reports in LEP, PS and SPS libraries.
- [21] TeVI Group, *Design Report Tevatron 1 project*, .
- [22] “Universität Zürich, Physik-Institut website.”
<http://www.physik.uzh.ch/groups/serra/StandardModel.html>.
- [23] E. Noether, *Invariante Variationprobleme*. *Nach. v.d. Ges. d. Wiss zu goettingen*, Mathphys. Klasse (1918) 235–257.
- [24] “Hotmath.” http://hotmath.com/hotmath_help/topics/transformation-of-graphs-using-matrices-reflection.html.
- [25] M. Gell-Mann, *Symmetries of baryons and mesons*, Phys. Rev. **125** (Feb, 1962) 1067–1084.
- [26] G. C. Branco, R. González Felipe, and F. R. Joaquim, *Leptonic cp violation*, Rev. Mod. Phys. **84** (Apr, 2012) 515–565.
- [27] “Standard model page, wikipedia.” https://en.wikipedia.org/wiki/Standard_Model#Construction_of_the_Standard_Model_Lagrangian.
- [28] T. Behnke et al., *The International Linear Collider Technical Design Report - Volume 1 : Executive Summary*, arXiv :1306.6327.
- [29] O. Dominguez and F. Zimmermann, *Beam Parameters and Luminosity Time Evolution for an 80-km VHE-LHC*, Tech. Rep. EuCARD-CON-2013-010, 2013.
- [30] W. K. H. Panofsky, *Evolution of particle accelerators and colliders*, SLAC Beam Line **27N1** (1999) 36–44.
- [31] M. Besancon, “Series of lectures at the new trends in high energy physics and QCD summerschool, 21-25th Oct. 2014.”
<https://indico.cern.ch/event/346738/contribution/1>.
- [32] E. Ponton, “Series of lectures at the new trends in high energy physics and QCD summerschool, 21-25th Oct. 2014.”
<https://indico.cern.ch/event/346738/contribution/4>.
- [33] K. Agashe, R. Contino, and A. Pomarol, *The minimal composite higgs model*, Nuclear Physics B **719** (2005), no. 1–2 165 – 187.
- [34] K. Agashe, A. Delgado, M. J. May, and R. Sundrum, *$RS1$, custodial isospin and precision tests*, JHEP **08** (2003) 050, [[hep-ph/0308036](#)].
- [35] **Particle Data Group** Collaboration, K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.
- [36] **L3** Collaboration, P. Achard et al., *Measurement of the mass and the width of the W boson at LEP*, Eur. Phys. J. **C45** (2006) 569–587, [[hep-ex/0511049](#)].
- [37] **SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3 Collaboration**, S. Schael et al., *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257–454, [[hep-ex/0509008](#)].
- [38] M. Kobayashi, “Nobel lectures on CP Violation and Flavour Mixing.”
http://www.nobelprize.org/nobel_prizes/physics/laureates/2008/kobayashi-lecture.html.
- [39] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (Oct, 1970) 1285–1292.

- [40] <http://quantstocktrader.blogspot.fr/2012/07/cern-experiments-observe-particle.html>.
- [41] H. Yukawa, *On the interaction of elementary particles. i*, Progress of Theoretical Physics Supplement **1** (1955) 1–10, [<http://ptps.oxfordjournals.org/content/1/1.1.full.pdf+html>].
- [42] M. Baak, J. Cúth, J. Haller, A. Hoecker, R. Kogler, K. Mönig, M. Schott, and J. Stelzer, *The global electroweak fit at NNLO and prospects for the LHC and ILC*, The European Physical Journal C **74** (2014), no. 9.
- [43] **ATLAS** Collaboration, *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at 7 and 8 TeV in the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2015-007, CERN, Geneva, Mar, 2015.
- [44] **CMS** Collaboration, V. Khachatryan et al., *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, Eur. Phys. J. **C75** (2015), no. 5 212, [[arXiv :1412.8662](#)].
- [45] M. Pierini, “Lhc run 1 anomalies at high mass and summary of run2 observations. 51st rencontres de moriond ew 2016.” <https://indico.in2p3.fr/event/12279/session/12/contribution/164/material/slides/0.pdf>.
- [46] O. Buchmueller et al., *The CMSSM and NUHM1 in Light of 7 TeV LHC, $B_s \rightarrow \mu^+ \mu^-$ and XENON100 Data*, Eur. Phys. J. **C72** (2012) 2243, [[arXiv :1207.7315](#)].
- [47] H. Euler and B. Kockel, *Ueber die Streuung von Licht an Licht nach der Diracschen Theorie*, Naturwiss. **23** (1935) 246.
- [48] A. Akhiezer Physik. Z. Sowjetunion **11** (1937) 263.
- [49] W. Heisenberg and H. Euler, *Folgerungen aus der diracschen theorie des positrons*, Zeitschrift für Physik **98** (1936), no. 11-12 714–732.
- [50] R. Karplus and M. Neuman, *The scattering of light by light*, Phys. Rev. **83** (1951) 776–784.
- [51] J. S. Schwinger, *On gauge invariance and vacuum polarization*, Phys. Rev. **82** (1951) 664–679.
- [52] G. Jikia and A. Tkabladze, *Photon-photon scattering at the photon linear collider*, Phys. Lett. **B323** (1994) 453–458, [[hep-ph/9312228](#)].
- [53] G. J. Gounaris, P. I. Porfyriadis, and F. M. Renard, *Light by light scattering at high-energy : A Tool to reveal new particles*, Phys. Lett. **B452** (1999) 76–82, [[hep-ph/9812378](#)]. [Erratum : Phys. Lett.B513,431(2001)].
- [54] Z. Bern, A. De Freitas, L. J. Dixon, A. Ghinculov, and H. L. Wong, *QCD and QED corrections to light by light scattering*, JHEP **11** (2001) 031, [[hep-ph/0109079](#)].
- [55] R. S. Gupta, *Probing Quartic Neutral Gauge Boson Couplings using diffractive photon fusion at the LHC*, Phys. Rev. **D85** (2012) 014006, [[arXiv :1111.3354](#)].
- [56] S. Fichet, G. von Gersdorff, O. Kepka, B. Lenzi, C. Royon, and M. Saimpert, *Probing new physics in diphoton production with proton tagging at the Large Hadron Collider*, Phys. Rev. **D89** (2014) 114004, [[arXiv :1312.5153](#)].

- [57] S. Fichet, G. von Gersdorff, B. Lenzi, C. Royon, and M. Saimpert, *Light-by-light scattering with intact protons at the LHC : from Standard Model to New Physics*, JHEP **02** (2015) 165, [[arXiv :1411.6629](#)].
- [58] S. Fichet, G. von Gersdorff, C. Royon, and S. Matthias, *Light-by-light scattering via exotic neutral resonances*, to be submitted (2016).
- [59] D. d’Enterria and G. G. da Silveira, *Observing light-by-light scattering at the large hadron collider*, Phys. Rev. Lett. **111** (Aug, 2013) 080405.
- [60] N. Watanabe, Y. Kurihara, K. Sasaki, and T. Uematsu, *Higgs Production in Two-Photon Process and Transition Form Factor*, Phys. Lett. **B728** (2014) 202–205, [[arXiv :1311.1601](#)].
- [61] D. Bernard et al., *Search for Stimulated Photon-Photon Scattering in Vacuum*, Eur. Phys. J. **D10** (2000) 141, [[arXiv :1007.0104](#)].
- [62] R. S. Van Dyck, P. B. Schwinberg, and H. G. Dehmelt, *New High Precision Comparison of electron and Positron G Factors*, Phys. Rev. Lett. **59** (1987) 26–29.
- [63] **Muon g-2** Collaboration, G. W. Bennett et al., *Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL*, Phys. Rev. **D73** (2006) 072003, [[hep-ex/0602035](#)].
- [64] D. Eriksson, G. Brodin, M. Marklund, and L. Stenflo, *A Possibility to measure elastic photon-photon scattering in vacuum*, Phys. Rev. **A70** (2004) 013808, [[physics/0411054](#)].
- [65] F. Della Valle et al., *Measurements of vacuum magnetic birefringence using permanent dipole magnets : the PVLAS experiment*, New J. Phys. **15** (2013) 053026, [[arXiv :1301.4918](#)].
- [66] G. Jarlskog, L. Jönsson, S. Prünster, H. D. Schulz, H. J. Willutzki, and G. G. Winter, *Measurement of delbrück scattering and observation of photon splitting at high energies*, Phys. Rev. D **8** (Dec, 1973) 3813–3823.
- [67] M. Schumacher, *Delbrück scattering*, Radiation Physics and Chemistry **56** (1999), no. 1–2 101 – 111.
- [68] S. J. Brodsky and P. M. Zerwas, *High-energy photon-photon collisions*, Nucl. Instrum. Meth. **A355** (1995) 19–41, [[hep-ph/9407362](#)].
- [69] **CMS** Collaboration, “Limits on anomalous triple and quartic gauge couplings.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSMPaTGC>.
- [70] K. Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld, *Low-energy effects of new interactions in the electroweak boson sector*, Phys. Rev. **D48** (1993) 2182–2203.
- [71] A. Aspect, J. Dalibard, and G. Roger, *Experimental test of Bell’s inequalities using time varying analyzers*, Phys. Rev. Lett. **49** (1982) 1804–1807.
- [72] **ATLAS** Collaboration, *The ATLAS experiment at the CERN large hadron collider*, Journal of Instrumentation **3** (2008), no. 08 S08003.
- [73] M. Albrow et al., *CMS-TOTEM Precision Proton Spectrometer*, Tech. Rep. CERN-LHCC-2014-021. TOTEM-TDR-003. CMS-TDR-13, CERN, Geneva, Sep, 2014.

- [74] O. Kepka and C. Royon, *Anomalous $WW\gamma$ coupling in photon-induced processes using forward detectors at the LHC*, Phys. Rev. **D78** (2008) 073005, [[arXiv :0808.0322](#)].
- [75] E. Chapon, C. Royon, and O. Kepka, *Anomalous quartic $W W$ gamma gamma, $Z Z$ gamma gamma, and trilinear WW gamma couplings in two-photon processes at high luminosity at the LHC*, Phys. Rev. **D81** (2010) 074003, [[arXiv :0912.5161](#)].
- [76] <https://plus.maths.org/content/strong-free-part-ii>.
- [77] <http://conferences.fnal.gov/lp2003/forthepublic/qcd/>.
- [78] **ATLAS** Collaboration, *Measurement of multi-jet cross-section ratios and determination of the strong coupling constant in proton-proton collisions at $\sqrt{s}=7$ TeV with the ATLAS detector.*, Tech. Rep. ATLAS-CONF-2013-041, CERN, Geneva, Apr, 2013.
- [79] J. M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, Int. J. Theor. Phys. **38** (1999) 1113–1133, [[hep-th/9711200](#)]. [Adv. Theor. Math. Phys.2,231(1998)].
- [80] **LHCb** Collaboration, R. Aaij et al., *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) 072001, [[arXiv :1507.0341](#)].
- [81] ATLAS collaboration, “ATLAS Stand-Alone Event Displays.” https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayStandAlone#High_pT_leading_jet.
- [82] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, [[arXiv :1410.3012](#)].
- [83] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, [[arXiv :0811.4622](#)].
- [84] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639–707, [[arXiv :0803.0883](#)].
- [85] J. Bellm et al., *Herwig 7.0 / Herwig++ 3.0 Release Note*, [arXiv :1512.0117](#).
- [86] G. Soyez, “BND Summer school (6 hours), September 7-17 2010, Oostende, Belgium.” www.lpthe.jussieu.fr/~soyez/lectures/2010-09-BND.pdf.
- [87] **CT14** Collaboration, S. Dulat et al., *The CT14 Global Analysis of Quantum Chromodynamics*, [arXiv :1506.0744](#).
- [88] **NNPDF** Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, [[arXiv :1410.8849](#)].
- [89] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era : MMHT 2014 PDFs*, Eur. Phys. J. **C75** (2015), no. 5 204, [[arXiv :1412.3989](#)].
- [90] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, *The Pomeron singularity in Nonabelian Gauge Theories*, Sov. Phys. JETP **45** (1977) 199–204. [Zh. Eksp. Teor. Fiz.72,377(1977)].
- [91] I. I. Balitsky and L. N. Lipatov, *The Pomeron singularity in Quantum Chromodynamics*, Sov. J. Nucl. Phys. **28** (1978) 822–829. [Yad. Fiz.28,1597(1978)].

- [92] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1–91, [[hep-ph/0409313](#)].
- [93] C. Marquet, “Talk given at the III Workshop on QCD and Diffraction at the LHC.” <https://indico.cern.ch/event/273342/session/5/contribution/42/attachments/492101/680218/marquet.pdf>.
- [94] T. Kinoshita, *Mass singularities of feynman amplitudes*, Journal of Mathematical Physics **3** (1962), no. 4 650–677.
- [95] T. D. Lee and M. Nauenberg, *Degenerate systems and mass singularities*, Phys. Rev. **133** (Mar, 1964) B1549–B1562.
- [96] F. Bloch and A. Nordsieck, *Note on the radiation field of the electron*, Phys. Rev. **52** (Jul, 1937) 54–59.
- [97] J. C. Collins, D. E. Soper, and G. F. Sterman, *Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production*, Nucl. Phys. **B250** (1985) 199.
- [98] G. A. Ladinsky and C. P. Yuan, *The Nonperturbative regime in QCD resummation for gauge boson production at hadron colliders*, Phys. Rev. **D50** (1994) 4239, [[hep-ph/9311341](#)].
- [99] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, *Sudakov resummation of multiparton QCD cross sections*, Physics Letters B **575** (2003), no. 3–4 268 – 278.
- [100] Y. Hatta, C. Marquet, C. Royon, G. Soyez, T. Ueda, and D. Werder, *A QCD description of the ATLAS jet veto measurement*, Phys. Rev. **D87** (2013), no. 5 054016, [[arXiv :1301.1910](#)].
- [101] C. Balazs, E. L. Berger, P. M. Nadolsky, and C. P. Yuan, *Calculation of prompt diphoton production cross-sections at Tevatron and LHC energies*, Phys. Rev. **D76** (2007) 013009, [[arXiv :0704.0001](#)].
- [102] L. Cieri, F. Coradeschi, and D. de Florian, *Diphoton production at hadron colliders : transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy*, JHEP **06** (2015) 185, [[arXiv :1505.0316](#)].
- [103] S. D. Drell and T.-M. Yan, *Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies*, Phys. Rev. Lett. **25** (1970) 316–320. [Erratum : Phys. Rev. Lett. 25,902(1970)].
- [104] S. Frixione, F. Stoeckli, P. Torrielli, B. R. Webber, and C. D. White, *The MCanLO 4.0 Event Generator*, [arXiv :1010.0819](#).
- [105] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs : the POWHEG BOX*, JHEP **06** (2010) 043, [[arXiv :1002.2581](#)].
- [106] **ATLAS** Collaboration, G. Aad et al., *Measurements of the W production cross sections in association with jets with the ATLAS detector*, Eur. Phys. J. **C75** (2015), no. 2 82, [[arXiv :1409.8639](#)].
- [107] **ATLAS** Collaboration, G. Aad et al., *Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **07** (2013) 032, [[arXiv :1304.7098](#)].
- [108] **ATLAS** Collaboration, G. Aad et al., *Measurement of the $t\bar{t}$ production cross-section as a function of jet multiplicity and jet transverse momentum in 7*

- TeV proton-proton collisions with the ATLAS detector*, *JHEP* **01** (2015) 020, [arXiv :1407.0891].
- [109] R. Ichou and D. d’Enterria, *Sensitivity of isolated photon production at TeV hadron colliders to the gluon distribution in the proton*, *Phys. Rev.* **D82** (2010) 014015, [arXiv :1005.4529].
- [110] ATLAS collaboration, “ATLAS Standard model summary plots.” https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/index.html#ATLAS_b_SMSummary_FiducialXsect.
- [111] T. Binoth, J. P. Guillet, E. Pilon, and M. Werlen, *A Full next-to-leading order study of direct photon pair production in hadronic collisions*, *Eur. Phys. J.* **C16** (2000) 311–330, [hep-ph/9911340].
- [112] L. Bourhis, M. Fontannaz, and J. P. Guillet, *Quarks and gluon fragmentation functions into photons*, *Eur. Phys. J.* **C2** (1998) 529–537, [hep-ph/9704447].
- [113] S. Catani, M. Fontannaz, and E. Pilon, *Factorization and soft gluon divergences in isolated photon cross-sections*, *Phys. Rev.* **D58** (1998) 094025, [hep-ph/9803475].
- [114] S. Catani, M. Fontannaz, J. P. Guillet, and E. Pilon, *Cross-section of isolated prompt photons in hadron hadron collisions*, *JHEP* **05** (2002) 028, [hep-ph/0204023].
- [115] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, *HERWIG 6 : An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *JHEP* **01** (2001) 010, [hep-ph/0011363].
- [116] S. Frixione, *Isolated photons in perturbative QCD*, *Phys. Lett.* **B429** (1998) 369–374, [hep-ph/9801442].
- [117] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini, *Diphoton production at hadron colliders : a fully-differential QCD calculation at NNLO*, *Phys. Rev. Lett.* **108** (2012) 072001, [arXiv :1110.2375].
- [118] J. M. Campbell, R. K. Ellis, Y. Li, and C. Williams, *Predictions for diphoton production at the LHC through NNLO in QCD*, arXiv :1603.0266.
- [119] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *QCD matrix elements + parton showers : The NLO case*, *JHEP* **04** (2013) 027, [arXiv :1207.5030].
- [120] L. Cieri, “Presentation given at the workshop on photon physics at the LHC 2015.” <https://indico.in2p3.fr/event/11495/contribution/22/material/slides/0.pdf>.
- [121] J. R. Andersen et al., *Les Houches 2013 : Physics at TeV Colliders : Standard Model Working Group Report*, arXiv :1405.1067.
- [122] **SM MC Working Group, SM and NLO MULTILEG Working Group** Collaboration, J. Alcaraz Maestre et al., *The SM and NLO Multileg and SM MC Working Groups : Summary Report*, in *Proceedings, 7th Les Houches Workshop on Physics at TeV Colliders*, pp. 1–220, 2012. arXiv :1203.6803.
- [123] T. Binoth, J. P. Guillet, E. Pilon, and M. Werlen, *Beyond leading order effects in photon pair production at the Tevatron*, *Phys. Rev.* **D63** (2001) 114016, [hep-ph/0012191].

- [124] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, [arXiv :1007.2241].
- [125] **CMS** Collaboration, S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.
- [126] **LHCb** Collaboration, A. A. Alves, Jr. et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [127] **ALICE** Collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [128] **TOTEM** Collaboration, G. Anelli et al., *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.
- [129] **LHCf** Collaboration, O. Adriani et al., *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.
- [130] J. L. Pinfold, *The MoEDAL Experiment at the LHC – a New Light on the Terascale Frontier*, J. Phys. Conf. Ser. **631** (2015), no. 1 012014.
- [131] “LHC beauty experiment public website.” <http://lhcb-public.web.cern.ch/lhcb-public/en/detector/Detector-en.html>.
- [132] *LHC Forward Physics*, Tech. Rep. CERN-PH-LPCC-2015-001, CERN, Geneva, 2015.
- [133] “CERN public website.” http://cern60-downloads.web.cern.ch/cern60-downloads/downloads/EXHIBITION%20CONTENT/3_LHC_Accelerators/3_2_Posters/3_2_1_Overview/CERN_Accelerators/3_PDF-text/CERN_Accelerators.pdf.
- [134] **ATLAS** Collaboration, *Luminosity Determination in pp Collisions at $\sqrt{s} = 8$ TeV using the ATLAS Detector at the LHC*, Tech. Rep. ATL-COM-DAPR-2015-014, CERN, Geneva, Aug, 2015. Journal draft for EdBoard review.
- [135] **ATLAS** Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73** (2013), no. 8 2518, [arXiv :1302.4393].
- [136] **ATLAS** Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, Tech. Rep. ATLAS-CONF-2015-038, CERN, Geneva, Aug, 2015.
- [137] “ATLAS LuminosityPublicResults.” https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#Integrated_luminosity_summary_pl.
- [138] “ATLAS LuminosityPublicResults.” https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#Integrated_luminosity_summary_p://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.
- [139] “Longer term LHC schedule - status in June 2015.” https://lhcb-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC%20schedule%20beyond%20LS1%20MTP%202015_Freddy_June2015.pdf.

- [140] **ATLAS** Collaboration, G. Aad et al., *Studies of the performance of the ATLAS detector using cosmic-ray muons*, Eur. Phys. J. **C71** (2011) 1593, [arXiv :1011.6665].
- [141] http://www.interactions.org/cms/?pid=2100\&image_no=CE0155.
- [142] “ATLAS public pictures : inner detector.”
<http://www.atlas.ch/photos/inner-detector-combined.html>.
- [143] L. Barranco Navarro et al., *Study of the mechanical stability of the ATLAS Insertable B-Layer*, Tech. Rep. ATL-COM-INDET-2015-022, CERN, Geneva, May, 2015.
- [144] S. McMahon, *The ATLAS Semi Conductor Tracker : The SCT operations and performance in the 3 years covering 2010-2012 operations, LS1 activities, lessons learned*, Tech. Rep. ATL-COM-INDET-2013-070, CERN, Geneva, Sep, 2013.
- [145] A. Romaniouk et al., *TRT performance in Run 1*, Tech. Rep. ATL-COM-INDET-2015-041, CERN, Geneva, Jun, 2015.
- [146] **ATLAS** Collaboration, G. Aad et al., *The ATLAS Inner Detector commissioning and calibration*, Eur. Phys. J. **C70** (2010) 787–821, [arXiv :1004.5293].
- [147] **ATLAS** Collaboration, G. Aad et al., *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. **C74** (2014), no. 10 3071, [arXiv :1407.5063].
- [148] “ATLAS masterclass website.”
<http://atlas.physicsmasterclasses.org/en/atlas.htm>.
- [149] **ATLAS** Collaboration, N. Nikiforou, *Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade*, in Proceedings, 3rd International Conference on Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA 2013), 2013. arXiv :1306.6756.
- [150] **ATLAS** Collaboration, G. Aad et al., *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. **C72** (2012) 1849, [arXiv :1110.1530].
- [151] “ATLAS public pictures : muon spectrometer.”
<http://www.atlas.ch/photos/muons-combined.html>.
- [152] **ATLAS** Collaboration, N. Soni, *ATLAS Forward Detectors and Physics*, in Lake Louise Winter Institute : Celebrating 25 years (LLWI 2010) Lake Louise, Alberta, Canada, February 15-20, 2010, 2010. arXiv :1006.5426.
- [153] O. Viazlo and A. L. Collaboration, *ATLAS LUCID detector upgrade for LHC Run 2*, Tech. Rep. ATL-FWD-PROC-2015-004, CERN, Geneva, Oct, 2015.
- [154] F. Kohn, Measurement of the charge asymmetry in top quark pair production in pp collision data at $\sqrt{s} = 7$ TeV using the ATLAS detector. PhD thesis, 2012. arXiv :1204.0952.
- [155] P. Fountas, *ATLAS Central Trigger Processor Input Module (CTPIN) Firmware Upgrade. CTPIN : Central Trigger Processor Input Module*, .
- [156] **ATLAS** Collaboration, G. Duckeck, D. Barberis, R. Hawkings, R. Jones, N. McCubbin, G. Poulard, D. Quarrie, T. Wenaus, and E. Obreshkov, *ATLAS computing : Technical design report*, .

- [157] **GEANT4** Collaboration, S. Agostinelli et al., *GEANT4 : A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.
- [158] D. Costanzo et al., *Validation of the GEANT4-Based Full Simulation Program for the ATLAS Detector*, .
- [159] **ATLAS** Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, [arXiv :1005.4568].
- [160] “ROOT Reference Guide.” <https://root.cern.ch/root/html/>.
- [161] J. Catmore, *A New Petabyte-scale Data Derivation Framework for ATLAS*, .
- [162] M. Saimpert, *Search of new states of matter at the ATLAS experiment at the LHC*, Master’s thesis, MINES ParisTech, projet final d’étude, 2013.
- [163] C. Marquet, C. Royon, M. Saimpert, and D. Werder, *Probing the Pomeron structure using dijets and γ +jet events at the LHC*, Phys. Rev. **D88** (2013), no. 7 074029, [arXiv :1306.4901].
- [164] E. Delagnes, D. Breton, H. Grabas, J. Maalmi, P. Rusquart, and M. Saimpert, *The SAMPIC Waveform and Time to Digital Converter*, in *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (2014 NSS/MIC)*, and *21st Symposium on Room-Temperature Semiconductor X-Ray and Gamma-Ray Detectors*, (Seattle, United States), Nov., 2014. See Electronique.
- [165] E. Delagnes, D. Breton, H. Grabas, J. Maalmi, and P. Rusquart, *Reaching a few picosecond timing precision with the 16-channel digitizer and timestamper SAMPIC ASIC*, Nucl. Instrum. Meth. **A787** (2015) 245–249.
- [166] H. Grabas, *Development of a picosecond time-of-flight system in the ATLAS experiment*. PhD thesis, U. Paris-Sud 11, Dept. Phys., Orsay, 2013.
- [167] **ATLAS** Collaboration, G. Aad et al., *Measurement of the total cross section from elastic scattering in pp collisions at $\sqrt{s}=7$ TeV with the ATLAS detector*, Nucl. Phys. **B889** (2014) 486–548, [arXiv :1408.5778].
- [168] P. Jenni, M. Nordberg, M. Nessi, and K. Jon-And, *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity*. Technical Design Report. CERN, Geneva, 2008.
- [169] **TOTEM** Collaboration, G. Antchev et al., *Luminosity-independent measurements of total, elastic and inelastic cross-sections at $\sqrt{s} = 7$ TeV*, Tech. Rep. TOTEM-2012-004. CERN-PH-EP-2012-353. 2, CERN, Geneva, Aug, 2015.
- [170] **TOTEM** Collaboration, G. Antchev et al., *Luminosity-Independent Measurement of the Proton-Proton Total Cross Section at $\sqrt{s} = 8$ TeV*, Phys. Rev. Lett. **111** (Nov, 2012) 012001. 8 p.
- [171] L. Adamczyk et al., *Technical Design Report for the ATLAS Forward Proton Detector*, Tech. Rep. CERN-LHCC-2015-009. ATLAS-TDR-024, CERN, Geneva, May, 2015.
- [172] C. Royon and N. Cartiglia, *The AFP and CT-PPS projects*, arXiv :1503.4632.
- [173] G. Chiodini, *Diamond Particle Detectors for High Energy Physics*, in *Proceedings, 4th IEEE International Workshop on Advances in Sensors and Interfaces (IWASI)*, 2011.

- [174] N. Cartiglia et al., *Performance of Ultra-Fast Silicon Detectors*, JINST **9** (2014) C02001, [[arXiv :1312.1080](#)].
- [175] *Letter of Intent for the Phase-I Upgrade of the ATLAS Experiment*, Tech. Rep. CERN-LHCC-2011-012. LHCC-I-020, CERN, Geneva, Nov, 2011.
- [176] L. Perktold and J. Christiansen, *A multichannel time-to-digital converter asic with better than 3 ps rms time resolution*, Journal of Instrumentation **9** (2014), no. 01 C01060.
- [177] E. Delagnes, “Precise Pulse Timing based on Ultra-Fast Waveform Digitizers.” http://irfu.cea.fr/en/Phoceia/file.php?class=cours&file=/eric.delagnes/ED_Waveform_digitizers_Valencia.pdf.
- [178] E. Delagnes, D. Breton, H. Grabas, J. Maalmi, and P. Rusquart, *Reaching a few picosecond timing precision with the 16-channel digitizer and timestamper SAMPIC ASIC*, Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment **787** (2015) 245 – 249.
- [179] R. Staszewski, *Study of Diffraction with the ATLAS detector at the LHC*. PhD thesis, U. Paris-Sud 11, Dept. Phys., Orsay, 2012. presented 24 Sep 2012.
- [180] C. Royon, M. Saimpert, K. O., and R. Zlebcik, *Timing detectors for proton tagging at the LHC*, Acta Physica Polonica B, Proceedings Supplement **Vol. 7, Number 4** (2014) 735.
- [181] *Expected photon performance in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2011-007, CERN, Geneva, Apr, 2011.
- [182] T. Koffas and K. Tackmann, *ATLAS Tracker Material using Photon Conversions*, Tech. Rep. ATL-COM-PHYS-2011-120, CERN, Geneva, Feb, 2011.
- [183] J.-F. Marchand, *Stability of photon conversion reconstruction with pile-up*, Tech. Rep. ATL-COM-PHYS-2012-823, CERN, Geneva, Jun, 2012.
- [184] **ATLAS** Collaboration, G. Aad et al., *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur. Phys. J. **C74** (2014), no. 7 2941, [[arXiv :1404.2240](#)].
- [185] **ATLAS** Collaboration, G. Aad et al., *Measurements of the photon identification efficiencies using 20.3 fb^{-1} of pp collisions collected by ATLAS at $\sqrt{s} = 8 \text{ TeV}$ in 2012*, submitted to EPJC [[arXiv :1606.0181](#)].
- [186] **ATLAS** Collaboration, G. Aad et al., *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector*, [arXiv :1605.0349](#).
- [187] M. Delmastro, K. Liu, G. Marchiori, E. Petit, M. Pitt, C. Hengler, E. Soldatov, and K. Tackmann, *Measurements of the photon identification efficiencies using 20.3 fb^{-1} of pp collisions collected by ATLAS at $\sqrt{s} = 8 \text{ TeV}$ in 2012*, Tech. Rep. ATL-COM-PHYS-2014-542, CERN, Geneva, May, 2014.
- [188] **ATLAS** Collaboration, G. Aad et al., *Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8 \text{ TeV}$ with ATLAS*, JHEP **09** (2014) 112, [[arXiv :1407.4222](#)].
- [189] **ATLAS** Collaboration, G. Aad et al., *Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8*

- TeV with the ATLAS detector*, Phys. Rev. **D90** (2014), no. 11 112015, [arXiv :1408.7084].
- [190] M. Escalier, L. Fayard, and J.-F. Marchand, *Reconstruction of the z vertex and direction of the photon*, Tech. Rep. ATL-PHYS-INT-2010-013, CERN, Geneva, Jan, 2010.
- [191] A. Morley, P. Ward, F. Djama, and A. Alonso, *Inner Detector performance in 25 ns and 50 ns runs*, Tech. Rep. ATL-COM-PHYS-2013-1348, CERN, Geneva, Sep, 2013.
- [192] **ATLAS** Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, arXiv :0901.0512.
- [193] **ATLAS** Collaboration, G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, arXiv :1603.0293.
- [194] J.-B. De Vivie de Regie, S. Laplace, and D. Olivito, *Topocluster based calorimeter isolation dependence on pile-up*, Tech. Rep. ATL-COM-PHYS-2012-362, CERN, Geneva, Mar, 2012.
- [195] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys.Lett. **B659** (2008) 119–126, [arXiv :0707.1378].
- [196] M. Cacciari, G. P. Salam, and S. Sapeta, *On the characterisation of the underlying event*, JHEP **1004** (2010) 065, [arXiv :0912.4926].
- [197] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. **D48** (1993) 3160–3166, [hep-ph/9305266].
- [198] **ATLAS** Collaboration, G. Aad et al., *Search for high-mass diphoton resonances in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Phys. Rev. **D92** (2015), no. 3 032004, [arXiv :1504.0551].
- [199] X. Anduaga, F. Arduh, T. Berry, J. Brown, Q. Buat, D. Hayden, F. Malek, J. Stark, and P. Waller, *Search for Extra Dimensions in the Diphoton Final State using 20.3 fb^{-1} of pp Collisions recorded by the ATLAS detector at the LHC*, Tech. Rep. ATL-COM-PHYS-2013-740, CERN, Geneva, Jun, 2013.
- [200] M. Stoebe, S. Chekanov, J. Cantero, B. Vachon, M. Stockton, C. Glasman, J. Terron, O. Rifki, S. Cheatham, and D. Layden, *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector using 20.2 fb^{-1}* , Tech. Rep. ATL-COM-PHYS-2014-744, CERN, Geneva, Jun, 2014. Supporting note for the 2012 inclusive cross section measurement.
- [201] B. Lenzi, C. Royon, and M. Saimpert, *Data-driven corrections for the calorimeter isolation energy of photons computed from topological clusters*, Tech. Rep. ATL-COM-PHYS-2015-258, CERN, Geneva, Apr, 2015.
- [202] **ATLAS** Collaboration, G. Aad et al., *Measurement of the isolated di-photon cross-section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys.Rev. **D85** (2012) 012003, [arXiv :1107.0581].
- [203] **ATLAS** Collaboration, G. Aad et al., *Measurement of isolated-photon pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **01** (2013) 086, [arXiv :1211.1913].
- [204] **ATLAS** Collaboration, G. Aad et al., *Evidence of $W\gamma\gamma$ Production in pp Collisions at $\sqrt{s}=8$ TeV and Limits on Anomalous Quartic Gauge Couplings*

- with the ATLAS Detector, *Phys. Rev. Lett.* **115** (2015), no. 3 031802, [arXiv :1503.0324].
- [205] “Roofit Reference Guide.” <https://root.cern.ch/roofit>.
- [206] “Minuit User’s Guide.” <http://seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/doc/doc.html>.
- [207] K. S. Cranmer, *Kernel estimation in high-energy physics*, *Comput.Phys.Commun.* **136** (2001) 198–207, [hep-ex/0011057].
- [208] ATLAS Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4l$ channels with the ATLAS detector using 25fb^{-1} of pp collision data*, *Phys.Rev.* **D90** (2014) 052004, [arXiv :1406.3827].
- [209] CMS Collaboration, V. Khachatryan et al., *Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8\text{ TeV}$* , *JINST* **10** (2015), no. 08 P08010, [arXiv :1502.0270].
- [210] A. R. Cueto Gomez, C. Glasman, J. Terron, M. Saimpert, and S. Chekanov, *Study of inclusive isolated-photon production in pp collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2015-603, CERN, Geneva, Jul, 2015.
- [211] B. Lenzi, X. Li, C. Royon, M. Saimpert, J. Terron, T. R. Wyatt, F. Siegert, and M. Delmastro, *Measurement of the isolated di-photon cross section in 20.24 fb^{-1} of pp collisions at $\sqrt{s} = 8\text{ TeV}$ with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2015-1166, CERN, Geneva, Sep, 2015. Supporting documentation of the ATLAS isolated di-photon measurement at 8 TeV.
- [212] F. Siegert, “Matching and merging QCD matrix elements and parton showers at NLO accuracy, SLAC Theory Seminar, 11 Jan 2013.” <http://fsiegert.web.cern.ch/fsiegert/talks/2013-01-10-SLAC.pdf>.
- [213] J. Cantero, *Measurement of the isolated tri-photon cross section in pp collisions at $\sqrt{s} = 8\text{ TeV}$ with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2016-165, CERN, Geneva, Feb, 2016.
- [214] G. Salam, *Thoughts on qcd for run 2*, 2014. Slides of a talk given at the ATLAS Standard Model Workshop at LAPP, Annecy-le-Vieux, on the 2nd-5th February 2015.
- [215] ATLAS Collaboration, G. Aad et al., *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, *Phys. Lett.* **B726** (2013) 120–144, [arXiv :1307.1432].
- [216] ATLAS Collaboration, *Study of the spin of the new boson with up to 25 fb^{-1} of ATLAS data*, Tech. Rep. ATLAS-CONF-2013-040, CERN, Geneva, Apr, 2013.
- [217] J. C. Collins and D. E. Soper, *Angular Distribution of Dileptons in High-Energy Hadron Collisions*, *Phys. Rev.* **D16** (1977) 2219.
- [218] A. Banfi, S. Redford, M. Vesterinen, P. Waller, and T. Wyatt, *Optimisation of variables for studying dilepton transverse momentum distributions at hadron colliders*, *Eur.Phys.J.* **C71** (2011) 1600, [arXiv :1009.1580].
- [219] T. W. M. Vesterinen, *A novel technique for studying Z boson transverse momentum distribution at hadron colliders*, *Instrum. Meth.* **A602** (2009) 432–437, [arXiv :0807.4956].

- [220] **ATLAS** Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*, Phys. Rev. **D90** (2014), no. 5 052004, [arXiv :1406.3827].
- [221] **ATLAS** Collaboration, G. Aad et al., *Measurement of angular correlations in Drell-Yan lepton pairs to probe Z/γ boson transverse momentum at $\sqrt{s}=7\text{ TeV}$ with the ATLAS detector*, Phys. Lett. **B720** (2013) 32–51, [arXiv :1211.6899].
- [222] **D0** Collaboration, V. M. Abazov et al., *Measurement of the differential cross sections for isolated direct photon pair production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\text{ TeV}$* , Phys. Lett. **B725** (2013) 6–14, [arXiv :1301.4536].
- [223] **CDF** Collaboration, T. Aaltonen et al., *Measurement of the cross section for prompt isolated diphoton production using the full CDF Run II data sample*, Phys. Rev. Lett. **110** (2013), no. 10 101801, [arXiv :1212.4204].
- [224] **CMS** Collaboration, S. Chatrchyan et al., *Measurement of differential cross sections for the production of a pair of isolated photons in pp collisions at $\sqrt{s} = 7\text{ TeV}$* , Eur.Phys.J. **C74** (2014), no. 11 3129, [arXiv :1405.7225].
- [225] **CMS** Collaboration, *Production of pairs of isolated photons in association with jets in pp collisions at $\sqrt{s} = 7\text{ TeV}$* , Tech. Rep. CMS-PAS-SMP-14-021, CERN, Geneva, 2015.
- [226] C. Balazs and C. P. Yuan, *Soft gluon effects on lepton pairs at hadron colliders*, Phys. Rev. **D56** (1997) 5558–5583, [hep-ph/9704258].
- [227] J. M. Campbell and R. K. Ellis, *An Update on vector boson pair production at hadron colliders*, Phys. Rev. **D60** (1999) 113006, [hep-ph/9905386].
- [228] Z. Bern, L. J. Dixon, and C. Schmidt, *Isolating a light Higgs boson from the diphoton background at the CERN LHC*, Phys. Rev. **D66** (2002) 074018, [hep-ph/0206194].
- [229] **ATLAS** Collaboration, G. Aad et al., *Dynamics of isolated-photon plus jet production in pp collisions at $\sqrt{s} = 7\text{ TeV}$ with the ATLAS detector*, Nucl. Phys. **B875** (2013) 483–535, [arXiv :1307.6795].
- [230] Z. Barnovska, M. F. Bessner, M. Delmastro, E. A. Lamperti, E. Petit, K. Tackmann, J. Terron, and M. Fanti, *Supporting Note : Measurement of the diphoton in association with jets cross sections with the ATLAS detector at $\sqrt{s} = 8\text{ TeV}$* , Tech. Rep. ATL-COM-PHYS-2015-721, CERN, Geneva, Jul, 2015.
- [231] G. D’Agostini, *A multidimensional unfolding method based on bayes’ theorem*, Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment **362** (1995), no. 2–3 487 – 498.
- [232] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, [arXiv :1405.0301].
- [233] T. Gehrmann, N. Greiner, and G. Heinrich, *Photon isolation effects at NLO in $\gamma\gamma + \text{jet}$ final states in hadronic collisions*, JHEP **06** (2013) 058, [arXiv :1303.0824]. [Erratum : JHEP06,076(2014)].

- [234] A. Bocci, L. Shi, and Z. Liang, *Measurements of the $e \rightarrow \gamma$ fake rate*, Tech. Rep. ATL-COM-PHYS-2014-026, CERN, Geneva, Jan, 2014.
- [235] **ATLAS** Collaboration, *Summary of ATLAS Pythia 8 tunes*, Tech. Rep. ATL-PHYS-PUB-2012-003, CERN, Geneva, Aug, 2012.
- [236] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, [[arXiv :1007.2241](#)].
- [237] J. P. et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, <http://arxiv.org/abs/hep-ph/0201195v3>.
- [238] S. Hoeche, S. Schumann, and F. Siegert, *Hard photon production and matrix-element parton-shower merging*, Phys. Rev. **D81** (2010) 034026, [[arXiv :0912.3501](#)].
- [239] H. HSG1 Working Group, *Background Studies for the Search of Higgs Boson Decaying to Two Photons with 4.9 fb⁻¹ of 7TeV data with the ATLAS Experiment*, Tech. Rep. ATL-COM-PHYS-2012-515, CERN, Geneva, May, 2012.
- [240] **ATLAS** Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, Tech. Rep. ATLAS-CONF-2012-042, CERN, Geneva, Mar, 2012.
- [241] B. Lenzi and D. Delgove, *Selection of the diphoton production vertex using multivariate techniques for $H \rightarrow \gamma\gamma$ and other analyses*, Tech. Rep. ATL-COM-PHYS-2015-1321, CERN, Geneva, Oct, 2015.
- [242] *Characterisation and mitigation of beam-induced backgrounds observed in the atlas detector during the 2011 proton-proton run*, Journal of Instrumentation **8** (2013), no. 07 P07004.
- [243] **ATLAS** Collaboration, G. Aad et al., *Measurement of the production cross section of an isolated photon associated with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. **D85** (2012) 092014, [[arXiv :1203.3161](#)].
- [244] G. Cowan, Statistical Data Analysis. Oxford science publications. Clarendon Press, 1998.
- [245] L. Carminati, M. Delmastro, M. Fanti, R. Lafaye, B. Lenzi, S. M. Mazza, G. Marchiori, C. Meyer, M. Saimpert, and A. Vallier, *Purity measurements - Supporting documentation for the "Search for resonances decaying to photon pairs in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector"*, Tech. Rep. ATL-COM-PHYS-2016-194, CERN, Geneva, Feb, 2016.
- [246] **ATLAS** Collaboration, *Search for new phenomena in diphoton events with the ATLAS detector at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATLAS-COM-CONF-2016-016, CERN, Geneva, Mar, 2016.
- [247] L. Demortier and L. Lyons, *Everything you always wanted to know about pulls*, Tech. Rep. CDF/ANAL/PUBLIC/5776, CDF, February, 2002.
- [248] D. Damazio, T. Kono, F. Monticelli, and G. Pasztor, *Performance of the ATLAS Electron and Photon Triggers in p-p Collisions at $\sqrt{s} = 8$ TeV in 2012*, Tech. Rep. ATL-COM-DAQ-2013-121, CERN, Geneva, Oct, 2013.
- [249] B. Malaescu, *An Iterative, Dynamically Stabilized(IDS) Method of Data Unfolding*, 2011. [arXiv :1106.3107](#).

- [250] “Roounfold : Root unfolding framework.”
- [251] http://hepunix.rl.ac.uk/~adye/software/unfold/bayes_errors.pdf.
- [252] S. Schmitt, *TUnfold : an algorithm for correcting migration effects in high energy physics*, JINST **7** (2012) T10003, [[arXiv :1205.6201](#)].
- [253] Y. Nakahama and J. de Vivie, *Calorimeter - Tracker Inter-Alignment*, Tech. Rep. ATL-COM-LARG-2010-051, CERN, Geneva, Nov, 2010.
- [254] L. Cieri, *Progress in parton-level sm xsections calculations and comparison with data*, 2015. Slides of a talk given at the Workshop on Photon Physics at the LHC at LPNHE, Paris, on the 18-19th of February 2015.
- [255] F. Siegert, *Parton shower simulations of sm processes with photons and comparison with lhc data*, 2015. Slides of a talk given at the Workshop on Photon Physics at the LHC at LPNHE, Paris, on the 18-19th of February 2015.
- [256] Gleisberg, Tanju and Höche, Stefan, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, [[arXiv :0808.3674](#)].
- [257] Cascioli, Fabio and Maierhöfer, Philipp and Pozzorini, Stefano, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, [[arXiv :1111.5206](#)].
- [258] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, [[arXiv :0709.1027](#)].
- [259] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz, and F. Siegert, *Rivet user manual*, Comput. Phys. Commun. **184** (2013) 2803–2819, [[arXiv :1003.0694](#)].
- [260] M. R. Whalley, D. Bourilkov, and R. C. Group, *The Les Houches accord PDFs (LHAPDF) and LHAGLUE*, in HERA and the LHC : A Workshop on the implications of HERA for LHC physics. Proceedings, Part B, 2005. [hep-ph/0508110](#).
- [261] S. Catani and M. Grazzini, *An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC*, Phys. Rev. Lett. **98** (2007) 222002, [[hep-ph/0703012](#)].
- [262] C. Balazs, E. L. Berger, S. Mrenna, and C. P. Yuan, *Photon pair production with soft gluon resummation in hadronic interactions*, Phys. Rev. **D57** (1998) 6934–6947, [[hep-ph/9712471](#)].
- [263] C. Balazs, P. M. Nadolsky, C. Schmidt, and C. P. Yuan, *Diphoton background to Higgs boson production at the LHC with soft gluon effects*, Phys. Lett. **B489** (2000) 157–162, [[hep-ph/9905551](#)].
- [264] C. Balazs, E. L. Berger, P. M. Nadolsky, and C. P. Yuan, *All-orders resummation for diphoton production at hadron colliders*, Phys. Lett. **B637** (2006) 235–240, [[hep-ph/0603037](#)].
- [265] P. M. Nadolsky, C. Balazs, E. L. Berger, and C. P. Yuan, *Gluon-gluon contributions to the production of continuum diphoton pairs at hadron colliders*, Phys. Rev. **D76** (2007) 013008, [[hep-ph/0702003](#)].
- [266] S. Catani, D. de Florian, and M. Grazzini, *Universality of nonleading logarithmic contributions in transverse momentum distributions*, Nucl. Phys. **B596** (2001) 299–312, [[hep-ph/0008184](#)].

- [267] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, *Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC*, Nucl. Phys. **B737** (2006) 73–120, [[hep-ph/0508068](#)].
- [268] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, *Higgs boson production at the LHC : Transverse-momentum resummation and rapidity dependence*, Nucl. Phys. **B791** (2008) 1–19, [[arXiv :0705.3887](#)].
- [269] S. Catani and M. Grazzini, *QCD transverse-momentum resummation in gluon fusion processes*, Nucl. Phys. **B845** (2011) 297–323, [[arXiv :1011.3918](#)].
- [270] A. D. Martin, W. J. Stirling, and R. S. Thorne, *MRST partons generated in a fixed-flavor scheme*, Phys. Lett. **B636** (2006) 259–264, [[hep-ph/0603143](#)].
- [271] **CMS** Collaboration, S. Chatrchyan et al., *Study of exclusive two-photon production of W^+W^- in pp collisions at $\sqrt{s} = 7$ TeV and constraints on anomalous quartic gauge couplings*, JHEP **07** (2013) 116, [[arXiv :1305.5596](#)].
- [272] G. G. da Silveira and V. P. Goncalves, *Constraining the photon flux in two-photon processes at the LHC*, Phys. Rev. **D92** (2015), no. 1 014013, [[arXiv :1506.0135](#)].
- [273] **CDF** Collaboration, T. Aaltonen et al., *Observation of Exclusive Gamma Gamma Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **108** (2012) 081801, [[arXiv :1112.0858](#)].
- [274] **CMS** Collaboration, S. Chatrchyan et al., *Search for exclusive or semi-exclusive photon pair production and observation of exclusive and semi-exclusive electron pair production in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **11** (2012) 080, [[arXiv :1209.1666](#)].
- [275] L. Carminati, I. Wingerter-Seez, L. Barak, J. B. Beacham, N. Berger, M. A. Chelstowska, R. Peterson-Hall, P. G. Hamnett, R. Lafaye, B. Lenzi, F. Malek, S. Manzonni, S. M. Mazza, T. Meideck, C. Meyer, E. Petit, M. Saimpert, K. Sekhon, J. Stark, L. Roos, J. Terron, B. Trocme, G. Unal, A. Vallier, and Y. C. Yap, *Cross-checks on 2015 diphoton data samples.*, Tech. Rep. ATL-COM-PHYS-2015-1432, CERN, Geneva, Nov, 2015. The author list will be added.
- [276] L. A. Anchordoqui, I. Antoniadis, D.-C. Dai, W.-Z. Feng, H. Goldberg, X. Huang, D. Lust, D. Stojkovic, and T. R. Taylor, *String Resonances at Hadron Colliders*, Phys. Rev. **D90** (2014), no. 6 066013, [[arXiv :1407.8120](#)].
- [277] T. Kaluza, *On the Problem of Unity in Physics*, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) **1921** (1921) 966–972.
- [278] O. Klein, *Quantum Theory and Five-Dimensional Theory of Relativity. (In German and English)*, Z. Phys. **37** (1926) 895–906. [Surveys High Energ. Phys.5,241(1986)].
- [279] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, Phys. Rev. Lett. **83** (1999) 3370–3373, [[hep-ph/9905221](#)].
- [280] Z. Chacko and R. K. Mishra, *Effective Theory of a Light Dilaton*, Phys. Rev. **D87** (2013), no. 11 115006, [[arXiv :1209.3022](#)].
- [281] Z. Chacko, R. K. Mishra, and D. Stolarski, *Dynamics of a Stabilized Radion and Duality*, JHEP **09** (2013) 121, [[arXiv :1304.1795](#)].

- [282] J. Serra, *A higgs-like dilaton : viability and implications*, EPJ Web Conf. **60** (2013) 17005, [[arXiv :1312.0259](#)].
- [283] O. J. P. Eboli, M. C. Gonzalez-Garcia, S. M. Lietti, and S. F. Novaes, *Anomalous quartic gauge boson couplings at hadron colliders*, Phys. Rev. **D63** (2001) 075008, [[hep-ph/0009262](#)].
- [284] V. Costantini, B. De Tollis, and G. Pistoni, *Nonlinear effects in quantum electrodynamics*, Nuovo Cim. **A2** (1971) 733–787.
- [285] V. A. Khoze, A. D. Martin, M. G. Ryskin, and W. J. Stirling, *Diffractional gamma-gamma production at hadron colliders*, Eur. Phys. J. **C38** (2005) 475–482, [[hep-ph/0409037](#)].
- [286] J. R. Cudell, A. Dechambre, O. F. Hernandez, and I. P. Ivanov, *Central exclusive production of dijets at hadronic colliders*, Eur. Phys. J. **C61** (2009) 369–390, [[arXiv :0807.0600](#)].
- [287] R. Peschanski, M. Rangel, and C. Royon, *Hybrid Pomeron Model of exclusive central diffractive production*, Acta Phys. Polon. **B40** (2009) 2323–2344, [[arXiv :0808.1691](#)].
- [288] V. A. Khoze, A. D. Martin, and M. G. Ryskin, *Prospects for new physics observations in diffractive processes at the LHC and Tevatron*, Eur. Phys. J. **C23** (2002) 311–327, [[hep-ph/0111078](#)].
- [289] J. R. Forshaw and D. A. Ross, Quantum Chromodynamics and the Pomeron. Cambridge University Press, 1997. Cambridge Books Online.
- [290] V. A. Khoze, A. D. Martin, and M. G. Ryskin, *Physics with tagged forward protons at the LHC*, Eur. Phys. J. **C24** (2002) 581–587, [[hep-ph/0203122](#)].
- [291] L. A. Harland-Lang, V. A. Khoze, and M. G. Ryskin, *Exclusive physics at the LHC with SuperChic 2*, Eur. Phys. J. **C76** (2016), no. 1 9, [[arXiv :1508.0271](#)].
- [292] J. Monk and A. Pilkington, *ExHuME : A Monte Carlo event generator for exclusive diffraction*, Comput. Phys. Commun. **175** (2006) 232–239, [[hep-ph/0502077](#)].
- [293] H. Terazawa, *Two photon processes for particle production at high-energies*, Rev. Mod. Phys. **45** (1973) 615–662.
- [294] V. M. Budnev, I. F. Ginzburg, G. V. Meledin, and V. G. Serbo, *The Two photon particle production mechanism. Physical problems. Applications. Equivalent photon approximation*, Phys. Rept. **15** (1975) 181–281.
- [295] M. Dyndal and L. Schoeffel, *The role of finite-size effects on the spectrum of equivalent photons in proton–proton collisions at the LHC*, Phys. Lett. **B741** (2015) 66–70, [[arXiv :1410.2983](#)].
- [296] CMS Collaboration, S. Chatrchyan et al., *Exclusive photon-photon production of muon pairs in proton-proton collisions at $\sqrt{s} = 7$ TeV*, JHEP **01** (2012) 052, [[arXiv :1111.5536](#)].
- [297] ATLAS Collaboration, G. Aad et al., *Measurement of exclusive $\gamma\gamma \rightarrow \ell^+\ell^-$ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Lett. **B749** (2015) 242–261, [[arXiv :1506.0709](#)].
- [298] CDF Collaboration, T. Aaltonen et al., *Observation of Exclusive Dijet Production at the Fermilab Tevatron $p\bar{p}$ Collider*, Phys. Rev. **D77** (2008) 052004, [[arXiv :0712.0604](#)].

- [299] **CDF** Collaboration, M. Albrow, A. Swiech, and M. Zurek, *Exclusive Central $\pi^+\pi^-$ production in CDF*, in Proceedings, 15th conference on Elastic and Diffractive scattering (EDS Blois 2013), 2013. [arXiv :1310.3839](#).
- [300] M. Boonekamp, A. Dechambre, V. Juranek, O. Kepka, M. Rangel, C. Royon, and R. Staszewski, *FPMC : A Generator for forward physics*, [arXiv :1102.2531](#).
- [301] **CDF** Collaboration, T. Affolder et al., *Dijet production by double pomeron exchange at the Fermilab Tevatron*, Phys. Rev. Lett. **85** (2000) 4215–4220.
- [302] P. R. Newman and F.-P. Schilling, *HERA diffractive structure function data and parton distributions*, in HERA and the LHC : A Workshop on the implications of HERA for LHC physics : Proceedings Part A, 2005. [hep-ex/0511032](#).
- [303] G. Ingelman and P. E. Schlein, *Jet Structure in High Mass Diffractive Scattering*, Phys. Lett. **B152** (1985) 256.
- [304] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, *HERWIG 6.4 release note*, [hep-ph/0201201](#).
- [305] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, [[hep-ph/0603175](#)].
- [306] “The PHOJET Generator.”
<https://opal.web.cern.ch/opal/group/twophoton/tpg/gen/phojet.html>.
- [307] O. Kepka, C. Royon, M. Saimpert, and R. Zlebcik, *Multiple proton-proton interactions in Atlas Forward Detector pp at $\sqrt{s} = 14$ TeV*, Tech. Rep. ATL-COM-PHYS-2013-847, CERN, Geneva, Jun, 2013.
- [308] **ALICE** Collaboration, B. Abelev et al., *Measurement of inelastic, single- and double-diffraction cross sections in proton–proton collisions at the LHC with ALICE*, Eur. Phys. J. **C73** (2013), no. 6 2456, [[arXiv :1208.4968](#)].
- [309] **TOTEM** Collaboration, *Double diffractive cross-section measurement in the forward region at the lhc*, Phys. Rev. Lett. **111** (Dec, 2013) 262001.
- [310] “FPTracker svn repository.”
<svnweb.cern.ch/trac/atlasoff/browser/ForwardDetectors/FPTracker>.
- [311] W. Herr and F. Schmidt, *A MAD-X primer*, in Intermediate accelerator physics. Proceedings, CERN Accelerator School, Zeuthen, Germany, September 15-26, 2003, pp. 505–528, 2004.
- [312] K. Grevtsov, I. Wingerter-Seez, M. Delmastro, G. Unal, and K. Tackmann, *Photon reconstruction using 18.6 pb^{-1} of pp collisions collected by ATLAS at $\sqrt{s} = 13$ TeV in 2015*, Tech. Rep. ATL-COM-PHYS-2015-725, CERN, Geneva, Jul, 2015.
- [313] *Performance assumptions based on full simulation for an upgraded ATLAS detector at a High-Luminosity LHC*, Tech. Rep. ATL-PHYS-PUB-2013-009, CERN, Geneva, Sep, 2013.
- [314] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$ jet clustering algorithm*, JHEP **04** (2008) 063, [[arXiv :0802.1189](#)].
- [315] **CMS** Collaboration, *Search for High-Mass Diphoton Resonances in pp Collisions at $\sqrt{s}=8$ TeV with the CMS Detector*, .

- [316] L. Lyons, *Open statistical issues in particle physics*, Ann. Appl. Stat. **2** (09, 2008) 887–915.
- [317] **CMS** Collaboration, S. Chatrchyan et al., *Search for heavy narrow dilepton resonances in pp collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV*, Phys. Lett. **B720** (2013) 63–82, [[arXiv :1212.6175](#)].
- [318] The ALEPH, DELPHI, L3, OPAL Collaborations, the LEP Electroweak Working Group, *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP*, Phys. Rept. **532** (2013) 119, [[arXiv :1302.3415](#)].
- [319] **NNPDF** Collaboration, R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, and J. Rojo, *Parton distributions with QED corrections*, Nucl. Phys. **B877** (2013) 290–320, [[arXiv :1308.0598](#)].
- [320] S. Fichtel, G. von Gersdorff, and C. Royon, *Scattering Light by Light at 750 GeV at the LHC*, [arXiv :1512.0575](#).
- [321] S. Fichtel, G. von Gersdorff, and C. Royon, *Measuring the diphoton coupling of a 750 GeV resonance*, [arXiv :1601.0171](#).
- [322] L. A. Harland-Lang, V. A. Khoze, and M. G. Ryskin, *The production of a diphoton resonance via photon-photon fusion*, [arXiv :1601.0718](#).
- [323] C. Csaki, J. Hubisz, and J. Terning, *The Minimal Model of a Diphoton Resonance : Production without Gluon Couplings*, Phys. Rev. **D93** (2016) 035002, [[arXiv :1512.0577](#)]. [Phys. Rev.D93,035002(2016)].
- [324] H. Ito, T. Moroi, and Y. Takaesu, *Studying 750 GeV Di-photon Resonance at Photon-Photon Collider*, [arXiv :1601.0114](#).
- [325] C. Csaki, J. Hubisz, S. Lombardo, and J. Terning, *Gluon vs. Photon Production of a 750 GeV Diphoton Resonance*, [arXiv :1601.0063](#).
- [326] A. Chuinard, C. Royon, and R. Staszewski, *Testing Pomeron flavour symmetry with diffractive W charge asymmetry*, [arXiv :1510.0421](#).
- [327] M. Dosanjh, M. Cirilli, S. Myers, and S. Navin, *Medical applications at cern and the enlight network*, Frontiers in Oncology **6** (2016), no. 9.
- [328] V. De Cacqueray, *La mesure de temps de vol avec SAMPIC*, Master’s thesis, MINES ParisTech, projet final d’étude, 2015.
- [329] D. Breton, V. De Cacqueray, E. Delagnes, H. Grabas, J. Maalmi, N. Minafra, C. Royon, and M. Saimpert, *Measurements of timing resolution of ultra-fast silicon detectors with the SAMPIC WTDC*, [arXiv :1604.0238](#).
- [330] “Constant fraction discriminator, from wikipedia, the free encyclopedia.” https://en.wikipedia.org/wiki/Constant_fraction_discriminator.
- [331] F. Druillolle, D. Lachartre, F. Feinstein, E. Delagnes, H. Lafoux, and C. Hadamache, *The analogue ring sampler : an asic front-end electronics of the antares neutrino telescope*, in Nuclear Science Symposium Conference Record, 2001 IEEE, vol. 1, pp. 138–142 vol.1, Nov, 2001.
- [332] H. Grabas, *Development of a picosecond time-of-flight system in the ATLAS experiment*. PhD thesis, U. Paris-Sud 11, Dept. Phys., Orsay, 2013.

- [333] J. Lewis, *Fast normalized cross-correlation*, Vision interface **10** (1995), no. 1 120–123.
- [334] H.-W. Sadrozinski, S. Ely, V. Fadeyev, Z. Galloway, J. Ngo, C. Parker, B. Petersen, A. Seiden, A. Zatserklyaniy, N. Cartiglia, et al., *Ultra-fast silicon detectors*, Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment **730** (2013) 226–231.
- [335] N. Cartiglia, R. Arcidiacono, M. Baselga, R. Bellan, M. Boscardin, F. Cenna, G. Dalla Betta, P. Fernández-Martínez, M. Ferrero, D. Flores, et al., *Design optimization of ultra-fast silicon detectors*, Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment **796** (2015) 141–148.
- [336] G. Pellegrini, P. Fernández-Martínez, M. Baselga, C. Fleta, D. Flores, V. Greco, S. Hidalgo, I. Mandić, G. Kramberger, D. Quirion, et al., *Technology developments and first measurements of low gain avalanche detectors (lgad) for high energy physics applications*, Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment **765** (2014) 12–16.
- [337] N. Cartiglia, F. Cenna, A. Picerno, F. Ravera, V. Fadeyev, H. Sadrozinski, A. Seiden, P. Freeman, Z. Galloway, G. Pellegrini, P. Fernández-Martínez, M. BASELGA, V. Greco, S. Hidalgo, D. Quirion, C. Royon, E. Delagnes, M. Saimpert, D. Breton, J. Maalmi, and N. Minafra, *Measurement of the Time Resolution of Ultra-Fast Silicon Detectors*, in 2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (2014 NSS/MIC), and 21st Symposium on Room-Temperature Semiconductor X-Ray and Gamma-Ray Detectors, (Seattle, United States), Nov., 2014. See Electronique.
- [338] D. Breton, E. Delagnes, J. Maalmi, K. Nishimura, L. Ruckman, G. Varner, and J. Va'vra, *High resolution photon timing with mcp-pmts : A comparison of a commercial constant fraction discriminator (cfd) with the asic-based waveform digitizers {TARGET} and wavecatcher*, Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment **629** (2011), no. 1 123 – 132.
- [339] D. Breton, E. Delagnes, J. Maalmi, and P. Rusquart, *The wavecatcher family of sca-based 12-bit 3.2-gs/s fast digitizers*, in Real Time Conference (RT), 2014 19th IEEE-NPSS, pp. 1–8, May, 2014.
- [340] E. Auffray, B. Frisch, F. Geraci, A. Ghezzi, S. Gundacker, H. Hillemanns, P. Jarron, T. Meyer, M. Paganoni, K. Pauwels, M. Pizzichemi, and P. Lecoq, *A comprehensive amp ; systematic study of coincidence time resolution and light yield using scintillators of different size and wrapping*, IEEE Transactions on Nuclear Science **60** (Oct, 2013) 3163–3171.
- [341] D. Yvon, J. P. Renault, G. Tauzin, P. Verrecchia, C. Flouzat, S. Sharyy, E. Ramos, J. P. Bard, Y. Bulbul, J. P. Mols, P. Starzynski, D. Desforge, A. Marcel, J. M. Reymond, S. Jan, C. Comtat, and R. Trebossen, *Calipso : An novel detector concept for pet imaging*, IEEE Transactions on Nuclear Science **61** (Feb, 2014) 60–66.

- [342] **TOTEM** Collaboration, “Presentation at the 122nd LHCC meeting.”
https://indico.cern.ch/event/389857/session/1/contribution/5/attachments/779712/1069147/TOTEM_STATUS_REPORT_June2015.pdf.
- [343] **TOTEM** Collaboration, “Presentation at the 125th LHCC meeting.”
<https://indico.cern.ch/event/493464/session/1/contribution/19/attachments/1236914/1816688/Totem-LHCC-3-2016.pdf>.