UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Physique des Particules**

Arrêté ministérial : 7 Aout 2006

Présentée par

Zuzana Barnovská

Thèse dirigée par **Stéphane Jezequel**
et codirigée par **Marco Delmastro**

préparée au sein **Laboratoire d'Annecy-Le-Vieux de Physique des Particules**
et de **l'Ecole Doctorale de Physique de Grenoble**

Diphoton measurements with the ATLAS detector at the LHC: search for new resonances and study of diphoton production in association with jets

Thèse soutenue publiquement le **29 Septembre 2015**,
devant le jury composé de :

M. Marco Delmastro
LAPP - Annecy-Le-Vieux, Directeur de thèse
M. Mauro Donega
ETH - Zurich, Rapporteur
M. Jean-Pierre Lees
LAPP - Annecy-Le-Vieux, Examinateur
M. Eric Pilon
LAPTh - Annecy-Le-Vieux, Examinateur
Mme. Kerstin Tackmann
DESY - Hamburg, Rapporteur
M. Alessandro Tricoli
CERN, Examinateur

Acknowledgements

Annecy le Vieux, October 24th 2015

The work presented in this thesis was performed within the ATLAS group at LAPP (Laboratoire d'Annecy-Le-Vieux de Physique des Particules) in the past three years, starting in October 2012.

I would like to thank my supervisor, Marco Delmastro, for his guidance, patience and help throughout the PhD. I am very thankful for all that you have taught me in the last three years, working closely with me on all my thesis topics. I have learned a great deal, even though I could have learned more, I hope you know I tried my best.

I am also grateful to my supervisor Stéphane Jezequel, who has very kindly helped me with everything from the paperwork to the French obstacles and mostly many technical issues in my work as well as picking out cheese for my wedding.

I would like to thank Kerstin Tackmann and Mauro Donega for accepting to be my referees and reading the thesis, for all the comments on the manuscript that made it better in the end. I am also thankful to the jury members Alessandro Tricoli, Eric Pilon and Jean-Pierre Lees, who have also given me many detailed comments to improve the quality of the text.

During my first year I worked on my qualification task and I would like to thank the LAr community and especially Martin Aleksa, Hong Ma and Pavol Strizenec for the guidance and suggestions.

During the work on the resonance search, I got to work closely with Nicolas, Elisabeth, Jessica, Remi, Marco, Isabelle and Olivier, who have shown me how to do a full analysis working as a team.

For the measurement analysis, I would like to thank the whole team, Marco, Kerstin, Marcello, Elisabeth, Martin and Edoardo and also the inclusive diphoton analysis guys, Matthias, Bruno and Xingguo for the discussions on the topics we had in common. Thanks to Remi for finding the time to help me understand the code.

I would like to thank the whole ATLAS group at LAPP, for welcoming me and making me feel like I belonged there. I will never forget the teaching/therapy sessions in the car, when driving to CERN and back.

Another thank you goes to the administration at LAPP, namely Myriam, Chantal, Marie-Claude, Claudine, Brigitte, Natalie and Christine who have always been helpful with all the paperwork

necessary for my work at LAPP.

I would like to thank Kirill for being the best office mate and listening to my ****. I wish you get what you want in your career (because you already have the best wife) starting with a great PhD., just two years to go!

I would like to wish the newcomers to the ATLAS group, Angela and Sergii, all the best for their PhD. as well. Special mention goes to my fellow pizza lover Emmanuel, who has listened to most of my coding complaints and helped me solve many. Elena and Andre, I hope you enjoy your time in Annecy and make the most of your PostDocs.

I would like to thank Houry for her friendship, it means a lot to me. I hope we stay in touch and wish her all the best in the next chapter of her life.

I would like to thank my friends from home, a.k.a. The Club of Hanas, for making me laugh over our Facebook conversation, especially during the hard days.

And last but not least I would like to thank my family and my new husband Tibor. Mamka, tatko, babka, ďakujem za Vašu podporu počas celého môjho štúdia. Tibor, ďakujem že si šiel so mnou a vydržal si to so mnou aj keď to bolo miestami ťažké. Ľúbim Ťa a ďakujem za všetko.

Contents

Contents **iii**

Introduction **1**

1 Theoretical background and motivation **5**

1.1 The Standard Model 5

1.2 Limitations of the Standard Model 10

1.3 Extensions of the Standard Model 10

1.3.1 Supersymmetry 11

1.3.2 Supersymmetric models and the potential for another Higgs discovery in the diphoton channel 12

1.4 Diphoton pair production at the LHC 14

1.4.1 Perturbative QCD 15

1.4.2 Diphoton pair production processes at LHC 17

1.4.3 Reducible background of the prompt diphoton production 18

1.4.4 Theoretical predictions 19

1.5 Diphotons in association with jets at the LHC 20

1.5.1 Diphoton production in association with (at least) one jet 20

1.5.2 Diphoton production in association with (at least) two jets 21

1.5.3 Diphoton production in association with up to three jets 23

1.6 Previous measurements of diphoton production 23

1.6.1 Prompt diphoton and photons in association with jets cross section measurements 23

1.6.2 Previous searches for the SM Higgs and BSM physics in the diphoton channel 24

2 The LHC and the ATLAS detector **27**
2.1 The Large Hadron Collider . . . 27
2.1.1 LHC design and operation . . . 27
2.1.2 Experiments at the LHC . . . 30
2.1.3 LHC performance . . . 30
2.1.4 Future of the LHC . . . 32
2.2 The ATLAS detector . . . 33
2.2.1 Overview of ATLAS and its subsystems . . . 34
2.3 The Liquid Argon Calorimeter . . . 43
2.3.1 Signal detection and readout . . . 43
2.3.2 Optimal filtering method . . . 47

3 Photons in ATLAS **49**
3.1 Reconstruction of photons in ATLAS . . . 49
3.2 Photon calibration . . . 51
3.2.1 The *calibration hits* method . . . 51
3.3 Photon identification in ATLAS . . . 52
3.4 Photon isolation . . . 54
3.4.1 Calorimeter isolation . . . 55
3.4.2 Track isolation . . . 57
3.5 Photon cleaning . . . 57

4 Jets in ATLAS **59**
4.1 Jet reconstruction algorithms . . . 59
4.2 Jet reconstruction in ATLAS . . . 60
4.3 Jet quality selection in ATLAS . . . 61
4.4 Jet calibration in ATLAS . . . 62
4.5 Jet energy resolution . . . 66

5 Refined photon calibration **67**
5.1 Refined MVA-based electron and photon response calibration . . . 67
5.1.1 Training of MC-based e/γ calibration . . . 68
5.1.2 Longitudinal layer intercalibration . . . 68
5.1.3 MC-based e/γ energy calibration . . . 70
5.1.4 Uniformity corrections . . . 72
5.1.5 $Z \to ee$ resolution smearing and scale calibration . . . 72
5.1.6 $J/\psi \to ee$ electron energy scale validation . . . 74

5.1.7 $Z \to ll\gamma$ data-driven scale validation . . . 74
5.2 Summary and evaluation of the refined calibration . . . 75
5.3 Study of the PS and L1, L2 energy scale using photons . . . 76
5.4 Back layer energy scale . . . 80
5.5 ϕ intercalibration using photons . . . 85

6 Search for resonances decaying into two photons 91
6.1 Overview of the search for other resonances decaying into diphotons . . . 91
6.2 Event Selection . . . 93
6.2.1 Low-mass 60 $<m_{\gamma\gamma}<$ 120 GeV selection . . . 94
6.2.2 High-mass 100 $<m_{\gamma\gamma}<$ 800 GeV selection . . . 95
6.2.3 Summary of the selection criteria . . . 95
6.3 Signal modelling . . . 96
6.3.1 Definition of the double-sided Crystal Ball function . . . 97
6.3.2 Signal modelling at *high-mass* - inclusive analysis . . . 99
6.3.3 Signal modelling at *low-mass* - conversion categories . . . 99
6.3.4 Systematic uncertainties on the signal modelling . . . 101
6.4 Background modelling . . . 108
6.4.1 Jet background and photon purity . . . 108
6.4.2 Electron background . . . 109
6.4.3 Continuum background . . . 112
6.4.4 SM Higgs background . . . 115
6.5 Fiducial volume definition . . . 116
6.6 Systematic uncertainties . . . 117
6.6.1 Photon identification systematic uncertainties . . . 117
6.6.2 Photon isolation systematic uncertainties . . . 119
6.7 Fiducial limit extraction . . . 120
6.7.1 *Low-mass* analysis . . . 120
6.7.2 *High-mass* analysis . . . 122
6.7.3 Validity of the limits for narrow resonances . . . 124
6.8 Conclusions . . . 126

7 Measurement of the differential production cross sections of diphotons in association with jets 129
7.1 Overview of the analysis . . . 129
7.2 Event and object selection . . . 131
7.2.1 Monte Carlo simulation samples . . . 131

7.2.2 Data samples . . . 131
7.2.3 Photon candidate selection . . . 132
7.2.4 Jet selection . . . 134
7.2.5 Summary of the selection . . . 139
7.3 Jet background decomposition and diphoton yield extraction . . . 141
7.3.1 2x2D Sideband method . . . 141
7.3.2 Systematic uncertainties on the diphoton yield from the 2x2DSB method 151
7.3.3 Diphoton yields extracted from the 2x2DSB method . . . 152
7.4 Electron background . . . 167
7.4.1 Computation of the fake rates . . . 167
7.4.2 Estimation of the electron impurities as a function of an observable . . . 174
7.4.3 Note on the approach . . . 179
7.5 Contribution of the Higgs boson decaying into photon pairs . . . 182
7.6 Unfolding . . . 185
7.6.1 SVD unfolding . . . 185
7.6.2 Detector response matrices . . . 186
7.6.3 Regularization . . . 186
7.6.4 Unfolded distributions . . . 197
7.7 Systematic uncertainties . . . 203
7.7.1 Background subtraction methods . . . 203
7.7.2 Photon energy scale and resolution . . . 203
7.7.3 Photon *tight* ID scale factors . . . 205
7.7.4 Jet energy scale and resolution . . . 206
7.7.5 JVF signal efficiency uncertainty . . . 207
7.7.6 Jet definition and pile-up . . . 207
7.8 Theoretical predictions . . . 213
7.8.1 NLO calculations . . . 213
7.8.2 Sherpa and Pythia MC . . . 213
7.9 Comparison of measured cross sections to theoretical calculations . . . 213
7.10 Conclusion . . . 214

8 Optimization of the LAr signal reconstruction for the LHC Run 2 237
8.1 LAr signal reconstruction approches using the Optimal Filtering technique . . . 238
8.2 Simulation and reconstruction setup . . . 239
8.3 Mean energy and noise vs. pileup autocorrelation matrix . . . 240
8.3.1 EM calorimeters . . . 240
8.3.2 HEC calorimeters . . . 244

8.3.3 FCAL 244
8.3.4 E_{cell} distribution properties 245
8.3.5 Summary 245
8.4 Noise vs. number of LAr samples 253
8.4.1 Bunch-spacing = 50 ns, $\mu = 90$ 253
8.4.2 Bunch-spacing = 25 ns, $\mu = 60$ 259
8.4.3 Detector-specific choice of 4 signal samples 260
8.5 Conclusion 261

Conclusions and prospects **265**

Appendix **266**

A Photon isolation efficiencies **267**

B Photon *tight* ID efficiencies **275**

C Breakdown of the systematic uncertainties of the 2x2DSB method **283**

D Diphoton yields and their total uncertainties. **291**

E Uncertainty on the electron impurities due to $e \to \gamma$ fake rate variation **305**

F Uncertainty on the electron impurities due to $\gamma \to e$ fake rate variation **313**

G Diphoton yields before and after electron background subtraction **321**

Bibliography **329**

Résumé **349**

Introduction

On July 4th 2012, the ATLAS and CMS collaborations announced the discovery of a new particle, observed mainly in the diphoton and ZZ decaying to four leptons channels. After the first studies of its properties, it was presumed to be the last missing piece of the Standard Model (SM), which was predicted by Peter Higgs in 1964. The Standard Model (SM) itself was not yet fully developed at that time, but the mechanism proposed by professor Higgs and professors Brout and Englert at the same time independently, would later prove to be an essential ingredient for it. After collecting more data and perfecting their analyses, both collaborations revised their results and proceeded to call the particle a Higgs boson. All its properties so far point to the new particle being the Standard Model Higgs boson, but there are many models and several good reasons for it to be just the beginning of the next level of our understanding of the universe.

After the discovery in 2012, the analysis effort shifted towards the precise measurement of the new particle's properties. My personal contribution, in part, to the calibration of photon energy, directly translated into the precision measurement of the Higgs boson mass measured by ATLAS, which is of the order of 0.3% in the latest publication [1].

Focusing on physics in ATLAS and the diphoton final state in particular, one can then further choose a path towards either searching for other new resonances, or try to precisely measure the diphoton production. In this thesis both approaches are presented. A search for narrow resonances decaying into two photons mostly followed the same methods as the original Higgs search, but new challenges presented themselves along the way, mainly because of the wide mass range considered. On the other hand, measuring the cross section of the production of photon pairs in association with jets provides a test of perturbative Quantum Chromo Dynamics (QCD) and in addition can serve as a prerequisite of the search for Vector Boson Fusion (VBF) production of the SM Higgs and di-Higgs production, where one Higgs boson would decay into two photons and the other into a pair of a b quark and antiquark. Lastly, studying events where two photons and two jets were produced, the effective cross section can be measured, possibly constraining the gluon Parton Distribution Functions (PDFs).

This thesis is organized as follows. Chapter 1 gives a glimpse into the current understanding of the Standard Model and introduces some of the possible extensions of the Standard Model, which are necessary to explain phenomena beyond its scope. A motivation for studies of diphoton

events and events with diphotons and jets is discussed.

Chapter 2 introduces the experimental setup, the Large Hadron Collider and the ATLAS experiment with its subsystems.

The reconstruction, identification and the original calibration of photons in ATLAS is presented in Chapter 3. Chapter 4 gives a brief introduction into jet reconstruction and calibration in ATLAS.

In Chapter 5, the refined calibration of photons in ATLAS, developed and adopted for all precision measurements since 2012, is described. Also, a study performed on photons, towards a better understanding of the energy measurement in the LAr detector of ATLAS is detailed. This study was fundamental to the Higgs boson mass measurement.

Chapter 6 describes a search for narrow scalar resonances decaying into two photons, where my personal contribution was the signal modelling and computation of related systematic uncertainties.

A measurement of differential cross sections of diphotons in association with jets is presented in Chapter 7, comparing the measured values to theoretical predictions.

In Chapter 8, a study aiming to optimize the energy reconstruction in the Liquid Argon Calorimeter in view of the LHC Run2 is presented.

Chapter 8.5 lists future prospects for the studies and measurements presented in this thesis and serves as a final summary.

Personal contribution

The work presented in this thesis was performed within the ATLAS Collaboration. Here, I list my own personal contributions to the studies and analyses presented in the thesis, further reminders are written in the text as well.

I contributed to the photon calibration presented in Chapter 5 by performing the following studies:

- a study of the Liquid Argon (LAr) calorimeter energy scale using photons, described in Section 5.3;
- the study of the need for a relative calibration of the last longitudinal layer of the LAr calorimeter using photons, discussed in Section 5.4, the estimation of the impact on the energy of an EM object was performed within the $e/gamma$ group;
- the investigation of the possible need for a ϕ intercalibration using photons is evaluated in Section 5.5, using photons.

The analysis presented in Chapter 6 was performed within the ATLAS group at the LAPP laboratory and my personal contribution to the search for a narrow resonance decaying into two photons was the signal modelling and the computation of associated systematic uncertainties, presented in Section 6.3.

Chapter 7 describes the measurement of the differential cross section of diphotons in association with jets. My contributions are:

- Global event and object selection criteria were determined within the group, in this context, I studied the optimal transverse energy thresholds of the photons. I also investigated the mean isolation energy, and the isolation efficiencies as a function of jet multiplicity, as described in Section 7.2.3. I studied the Jet Vertex Fraction cut selection and the photon true vs. reconstruction isolation energy. I also implemented the full analysis selection code.

- In Section 7.3, I performed the initial studies on the definition of the *non-tight* control region in simulated Monte Carlo (MC) samples and performed the 2x2D Side-Band (2x2DSB) background decomposition and computed the associated systematic uncertainties. The data driven isolation efficiencies were extracted by another group member as a function of E_{T} and η. I performed the derivation of the efficiencies from MC as a function of each observable. I inherited the code for the 2x2DSB from the previous diphoton cross section measurement at $\sqrt{s} = 7$ TeV and was in close contact with the author. I implemented the necessary modifications related to the presence of jets, additional selection criteria and new observables that were not measured before. I also implemented the input of data-driven isolation efficiencies.

- Section 7.4, which describes the subtraction of electron background from the diphoton yields extracted in the previous Section, was entirely performed by me. I measured the fake rates in data and in MC and computed the true fake rate. Subsequently, I derived the bin-by-bin impurities as a function of the observables and subtracted the electron background. I also propagated the statistical and derived the systematic uncertainties.

- For the unfolding procedure, described in Section 7.6, I provided the data inputs - the yields after background subtractions, and the MC inputs - the detector response matrices. The unfolding procedure itself was performed within the group. I provided the response matrices for the systematic uncertainties.

- In Section 7.9, I made the comparison of the results to Sherpa and Pythia MC and discussed the results.

The final Chapter 8 describes a study of noise in the Liquid Argon calorimeter, which was my qualification task to become an ATLAS author. I performed all steps including simulation, reconstruction and the overall analysis.

1 Theoretical background and motivation

1.1 The Standard Model

The Standard Model of elementary particles and their interactions is an established theory, which aims to describe what the world is made of and how it all holds together on the elementary level. It is a gauge theory based on three symmetry groups: $SU(3)_C \otimes SU(2)_L \otimes U(1)_H$. The Standard Model describes elementary fermions, their anti-particles, and bosons acting as force carriers.

Figure 1.1 lists the three generations of quarks *up* and *down*, *charm* and *strange*, *top* and *bottom* (or *beauty*); three generations of leptons - *electron*, *muon*, *tau* and their respective *neutrinos*; and the gauge bosons, carriers of the three fundamental forces: the weak force - the Z and $W^{\pm}$ bosons, the electromagnetic force - *photon* γ and the strong force - the *gluons* g. Particles carrying forces (γ, g, Z, $W^{\pm}$) have integer spin and are called bosons, following the Bose-Einstein statistics. The quarks and leptons have half-integer spins, and are called fermions. They obey the principles of Fermi-Dirac statistics and follow the Pauli exclusion principle, according to which two particles cannot occupy the same quantum state.

The strong force is based on the $SU(3)_C$ group, where C stands for color. It is described by Quantum Chromodynamics[1](QCD). The octet of gluons, acting as mediators of the strong force, carry color charge. Quarks carry one color (or anticolor) and can exchange it by interacting with gluons. Gluons can be emitted or absorbed by quarks, changing their color. The color is to be understood just as another charge - the color charge. Quarks also carry electric charge of 1/3 or 2/3, giving integer charge when building hadrons, such as the proton. Due to the non-Abelian structure of the $SU(3)_C$ group [3], the strong interaction between two quarks gets stronger when the quarks are pulled apart. Quarks therefore cannot exist separately, they always form color neutral objects - color singlets. This is also called color confinement. Each of the three generations of quarks can be represented by a left-handed doublet and two right-handed singlets

$$\begin{pmatrix} u \\ d \end{pmatrix}, \; u_R, \; d_R. \tag{1.1}$$

The groups $SU(2)_L$ and $U(1)$ form the foundation for the combined electroweak (EW) theory, where the electromagnetic and the weak interactions are unified. This was done by Glashow [4],

[1] Chroma means color in Greek.

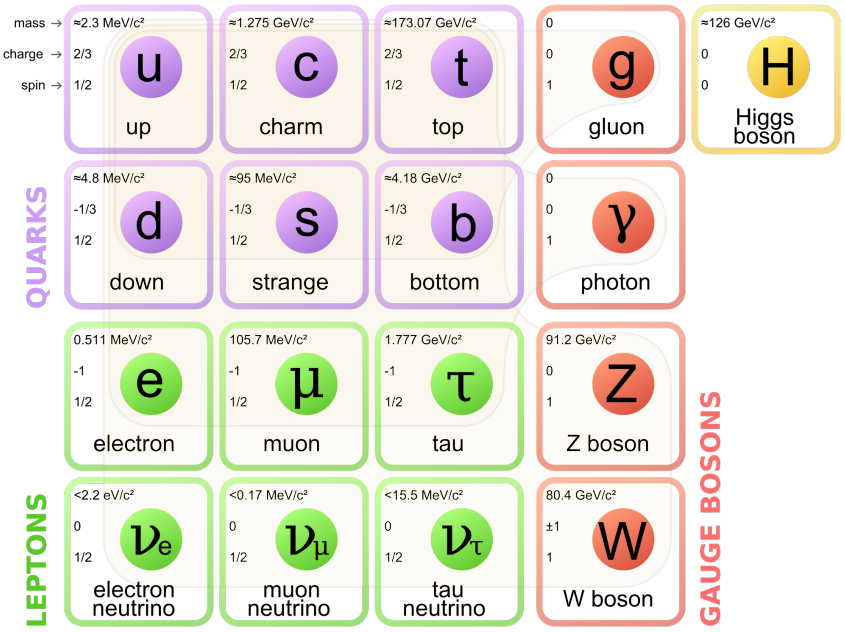


Figure 1.1: The Standard model of elementary particles [2].

Salam [5] and Weinberg [6] in the 1960s. The mediators of the electroweak interaction are the Z/γ and $W^{\pm}$ bosons. The charged leptons carry an electric charge of ± 1 and they, similarly to quarks, form a left-handed doublet and two right-handed singlets

$$\begin{pmatrix} \nu \\ e^- \end{pmatrix}, \; e_R^-, \; \nu_R, \tag{1.2}$$

while the right-handed neutrinos ν_R do not have to exist within the SM (they have not been experimentally observed), but are present in some of its extensions, explaining how neutrinos obtain mass [7].

The Quantum Electrodynamic (QED) theory, the QCD and the unified EW theory together give the description of the elementary particles and their forces. The SM Lagrangian can be summarized as

$$\mathcal{L}_{SM} = \mathcal{L}_{gauge} + \mathcal{L}_{fermions}, \tag{1.3}$$

where the $\mathcal{L}_{gauge}$ part includes the kinetic energies and self-interactions of the gauge bosons and the $\mathcal{L}_{fermions}$ part describes the quark and lepton kinetic energies and their interactions with the gauge bosons. No mass terms are present.

The Higgs field Φ and the Higgs boson H were introduced to break the EW symmetry. The mechanism of EW symmetry breaking, proposed by Robert Brout, Francois Englert [8], Peter Higgs [9, 10], and later Guralnik, Hagen and Kibble [11] now called the BEH-mechanism, allows the Z and $W^{\pm}$ bosons to obtain mass, while the photon remains massless.

The idea of spontaneous symmetry breaking comes from Ginzburg and Landau, who first applied it in their superconducting theory in 1950 [12]. Subsequently, in the 1960s, Nambu and Goldstone suggested the same mechanism for particle physics, producing massless Goldstone bosons [13, 14].

The BEH mechanism solves problem of particles obtaining mass by introducing a complex $SU(2)_L$ doublet to the SM Lagrangian term

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix} = \begin{pmatrix} \Phi^+ \\ \Phi^o \end{pmatrix}, \tag{1.4}$$

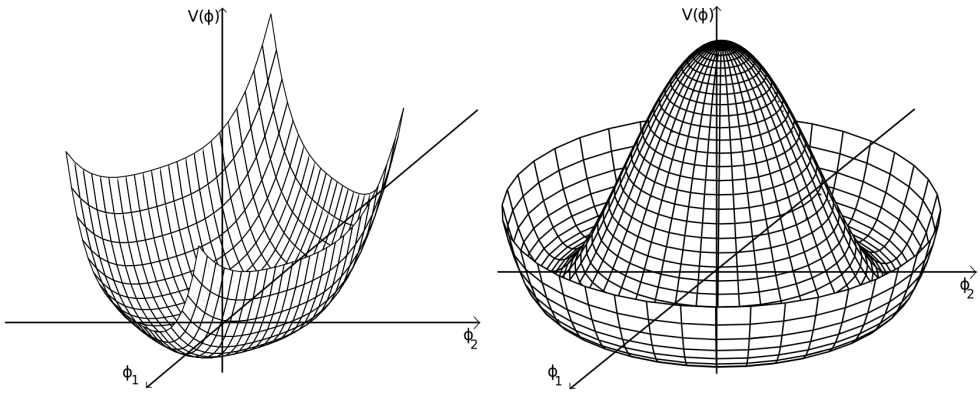
that couples to the massive gauge bosons. The Lagrangian

$$(D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \tag{1.5}$$

is added to the original SM Lagrangian (containing only the kinetic and interaction terms of fermions and gauge bosons), where the potential $V(\Phi)$ is defined as

$$V(\Phi) = \frac{1}{2}\mu^2 \Phi^\dagger \Phi + \frac{1}{4}\lambda(\Phi^\dagger \Phi)^2 \tag{1.6}$$

and has a shape shown in Figure 1.2(a) if $\mu^2 > 0$ and the shape of a Mexican hat, as in



(a) (b)

Figure 1.2: Higgs potential (a) before and (b) after symmetry breaking.

Figure 1.2(b) if $\mu^2 < 0$, with $\lambda > 0$ assumed. In the second case, the symmetry is broken and the minimum of the potential is located in the valley corresponding to $|\Phi| = \sqrt{-\frac{\mu^2}{2\lambda}}$. Defining $v = \sqrt{-\frac{2\mu^2}{\lambda}}$ as the vacuum expectation value, the ground potential can be written as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.7}$$

which explicitly breaks the electroweak symmetry [10]. Quantum excitations of the field can be expressed as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1.8}$$

where H is the physical Higgs boson. Using this scalar field within the EW Lagrangian, one obtains massive Z and $W^\pm$ and a massless photon, since the electromagnetic component is not broken.

Due to the local gauge invariance of the SM Lagrangian, necessary for the renormalizability of the SM theory, mass terms are not present in the SM Lagrangian, since they are not gauge invariant, and the left and right-handed fermions transform differently under $SU(2)_L$. The Higgs field interacts with fermions via the Yukawa couplings (term $\mathcal{L}_{Yukawa}$), and allows for fermionic mass terms to be present[2].

The mass of the Higgs boson itself can be expressed as

$$m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda v^2}, \tag{1.9}$$

and from the mass relations of the Z and $W^{\pm}$ bosons through the weak mixing angle θ_W

$$\cos\theta_W = \frac{m_W}{m_Z}, \tag{1.10}$$

which can be experimentally measured, one obtains the vacuum expectation value of $v = 246$ GeV. The SM does, however, not predict the H boson mass since the λ parameter, or the Higgs self-coupling strength, is a free parameter of the theory.

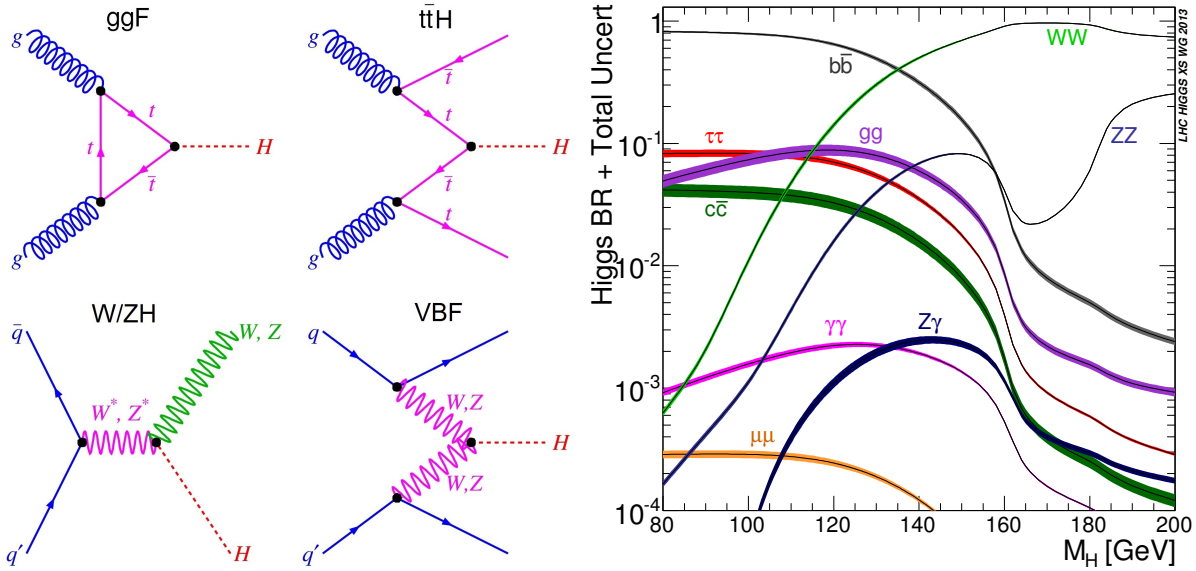
Bounds on the Higgs boson mass were put by the unitarity requirement ($m_H < 750$ GeV) [17]. Experimentally, the Higgs boson has been searched for at LEP, Tevatron and finally it has been discovered at the LHC in 2012 [18, 19], with the mass of $m_H = 125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst})$ GeV according to the latest combined measurements of the ATLAS and CMS experiments [20].

The SM Higgs boson can be produced by four different mechanisms, shown in Figure 1.3(a) and detected in many final states. The expected Higgs boson branching ratios for $\sqrt{s} = 8$ TeV collisions at the LHC are shown in Figure 1.3(b) for the relevant decay modes. The diphoton channel has a considerably lower branching fraction than the other channels, but photons produce a very clean signal in the detector and have a favorable signal to background ratio, with SM QCD production forming a falling exponential spectrum, so a resonant signal of the Higgs boson can be seen on top of it, thanks to the excellent resolution of the ATLAS EM calorimeter.

The $H \to ZZ \to 4l$ channel and the $H \to \gamma\gamma$ channel (decay via a *top* or W boson loop) were the two discovery channels in 2012 in both experiments. Later, results from other channels [22, 23, 24, 25, 26] as well as Higgs coupling measurements were presented [27]. The combined final LHC Run 1 Higgs mass measurement [20] is shown in Figure 1.4.

The knowledge of the Higgs boson mass and the very existence of such particle has implications on possible physics beyond the Standard Model (BSM), excluding some beyond the SM models or at least parts of their parameter phase space [28, 29, 30]. The combined EW fit [31, 32], performed after the Higgs boson discovery can also put constrains on physics beyond the SM. The mass of the H boson in relation to the *top* quark pole mass has several implications, e.g. that our universe is meta-stable [33], in case the SM is valid up to the Planck scale and there is no BSM physics. The question remains, whether the recently discovered particle is indeed the last missing SM particle - the SM Higgs boson or if it is just one of many new particles to be discovered at the LHC.

[2]Except the masses of neutrinos, which are not included in the SM. Theories forming extensions of the SM and including neutrino mass terms have been developed [7, 15, 16].



(a) (b)

Figure 1.3: Feynman diagrams of four SM Higgs boson production mechanisms (a) the gluon-gluon fusion (ggF), the top quark pair fusion (tt̄H), the Higgs-strahlung (ZH, WH) and the vector boson fusion (VBF); and SM Higgs branching ratios at $\sqrt{s} = 8$ TeV as a function of the Higgs mass for different decay modes (b) [21].

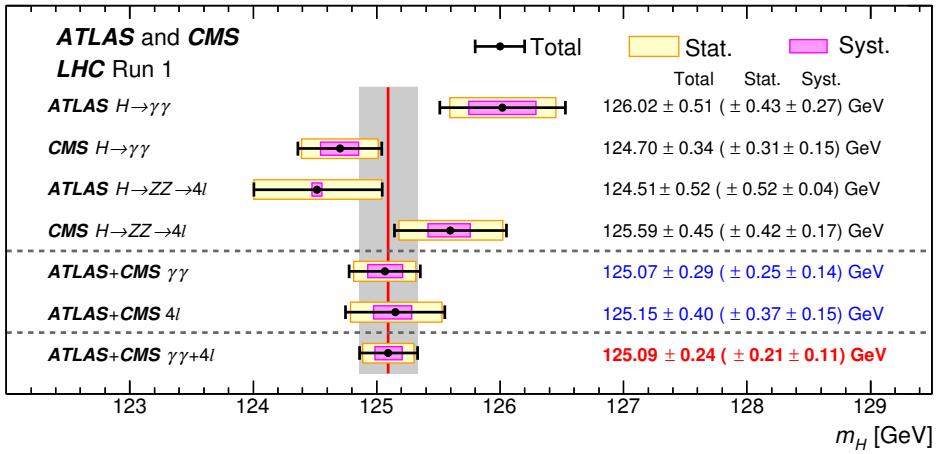


Figure 1.4: Summary of Higgs boson mass measurements from the individual analyses of ATLAS and CMS and from a combined analysis. The systematic (narrower, magenta-shaded bands), statistical (wider, yellow-shaded bands), and total (black error bars) uncertainties are indicated. The (red) vertical line and corresponding (gray) shaded column indicate the central value and the total uncertainty of the combined measurement, respectively [20].

1.2 Limitations of the Standard Model

As a theory, the Standard model does not explain all the phenomena observed in nature. The five main problems of the Standard Model are:

- Matter-antimatter asymmetry - the universe mostly consists of matter. CP-violation[3] observed in K^0 and B^0 meson systems agrees well with the predictions of the SM, but can only account for a small part of the asymmetry observed in the universe [35].
- Neutrino mass - neutrinos, within the scope of the SM, are massless, but recent experiments contradict this theory. The observation of neutrino oscillations [36] is proof that they indeed do have mass, albeit a very small one. Adding mass to the neutrinos is possible with some extensions to the SM and many experiments are trying to measure the masses, so far only being able to put upper bounds on them [37].
- Gravity - the SM does not explain gravity. At high energies, grand unification theories predict that all interactions should be unified [38], but when one tries to describe gravity as a quantum field theory, it becomes non-renormalizable. A superior theory is therefore needed to explain how gravity works [39, 40].
- Dark matter and dark energy - the universe, as we see it, is made of baryonic matter. In reality, this appears to be only about 4% of the universe's mass. Cosmological observations of galaxy clusters show that there must be something that is very massive but does not interact with SM particles, since we have not been able to observe it, that makes up $\approx 26\%$ of the universe. In addition, since the universe is expanding at an increasing rate, there must be something even more puzzling, for now called dark energy, that makes up $\approx 68\%$ of the universe [41].
- Naturalness - including the Higgs boson and the BEH mechanism, the problem of elementary particles obtaining mass is solved, but for the mass of the Higgs itself, the radiative loop corrections diverge quadratically with energy and the mass has to be adjusted by a correction of the order of 10^{17} in quadrature. This issue is also called the fine-tuning problem.

All these problems indicate that the SM has to be conquered by a new theory, which will be able to explain some or all of these phenomena. There are many theories and the experimental physicists at the LHC, but also experiments searching for Dark Matter, gravitational waves, neutrino experiments and cosmological studies are trying to find a hint of new physics.

1.3 Extensions of the Standard Model

Nowadays, the most studied extensions of the Standard Model among theoretical physicists are various Supersymmetry (SUSY) [42, 43, 44] extensions. Many other theories have been proposed, but so far, there has not been any hint of new physics coming from the LHC [45].

[3] CP-violation is a violation of the combination of charge (C) and parity (P) symmetry, first measured in 1964 by Christenson, Cronin, Fitch and Turlay in Ref [34].

1.3.1 Supersymmetry

Supersymmetry expands the Standard Model, by applying an additional spacetime symmetry, which links bosons and fermions by a SUSY transformation operator Q as

$$Q\,|\text{Boson}\rangle = |\text{Fermion}\rangle\,, \qquad Q\,|\text{Fermion}\rangle = |\text{Boson}\rangle\,. \tag{1.11}$$

For each particle of the SM, there is a supersymmetric partner of the opposite nature (bosons→fermions and vice versa). SUSY effectively doubles the amount of elementary particles as illustrated in Figure 1.5. It introduces spin 0 *squarks*, *sleptons* and *sneutrinos* and includes two complex

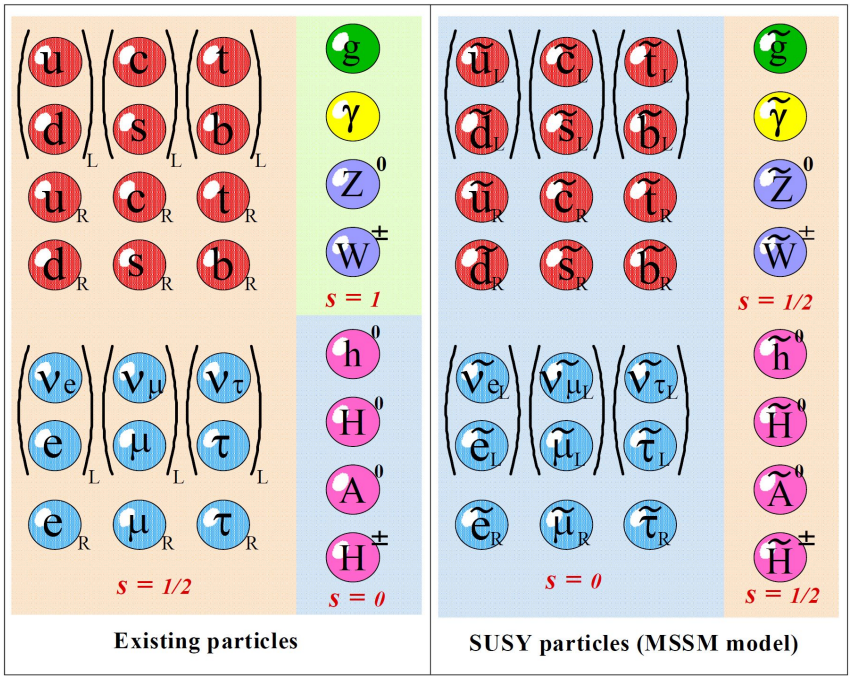


Figure 1.5: The Supersymmetric particles.

Higgs doublets (in some models), producing a total of five physical Higgs bosons h^0, H^0, A^0 and $H^\pm$ with their supersymmetric partners *Higgsinos*. The superpartners of the SM force particles and the *Higgsinos* would mix into four physical *neutralinos* and two *charginos* with spin 1/2.

SUSY particles are currently being searched for at the LHC. Experimentally, particles which are the easiest to search for, should be particles which are light in mass and have production cross sections within reach of the LHC (many models predict different masses, cross sections and branching fractions of SUSY particles). Producing heavy SUSY particles which desintegrate into a cascade can leave a light stable particle in the end. That would be the lightest stable supersymmetric particle - the *neutralino*. It could also be a possible candidate for a Dark Matter particle [46]. The superpartner of the *top* quark - the *stop* can also be light and within reach of the LHC [47, 48]. The extended Higgs sector and the graviton and its superpartner *gravitino* also have a strong discovery potential at the LHC [49, 50, 51]. In practice, the discoveries can be made in many different final states and simplified SUSY models with 100% branching fractions into a studied final state are being considered. Over the years, several simple approaches to

SUSY have been developed, where constraints are put on some of SUSY parameters, making them more prone for discovery at the LHC.

1.3.2 Supersymmetric models and the potential for another Higgs discovery in the diphoton channel

One of the basic SUSY extensions of the SM is the Minimal Supersymmetric Model (MSSM) [52]. Relating the particles with their superpartners if SUSY holds, would mean that the masses of the superpartners would be the same as their partners. Since this has not been observed experimentally, therefore, SUSY must be broken. While the mechanism of SUSY breaking is unknown, it should not, however, introduce any quadratic divergences to the scalar sector of the SM. The MSSM adds soft SUSY breaking terms with positive mass dimensions into the supersymmetric Lagrangian [53] with the superpotential

$$\mathcal{W}_{SUSY} = \mu H_d H_u + f_l Q H_d \bar{D} + f_u Q H_u \bar{U} + \lambda L L \bar{E} + \lambda' L Q \bar{D} + \lambda'' \bar{U} \bar{D} \bar{D} + \epsilon L H_u, \tag{1.12}$$

where H_d, H_u, L and Q are left-chiral superfields, doublets under $SU(2)_L$, and $\bar{U}$, $\bar{D}$, $\bar{E}$ are singlets under $SU(2)_L$ and f_l, f_u and f_d are 3x3 matrices representing Yukawa couplings. μ is the Higgs self-coupling. The last four terms of Eq. 1.12 are set to zero in MSSM, because they violate lepton and baryon number conservation. Instead of these terms, a new symmetry is introduced - the R-parity. The R-parity is a discrete symmetry, a new multiplicative quantum number defined as

$$P_R = (-1)^{3(B-L)+2s}, \tag{1.13}$$

where s is the spin, B is the baryon number and L is the lepton number. All SM particles have $R = 1$ and all new SUSY particles have $R = -1$. Consequences of the R-parity conservation are baryon and lepton number conservation (forbidding the last four terms of Eq. 1.12), SUSY particle pair production and one stable SUSY particle, which could be a Dark Matter candidate. In order for the cancellation of divergences of the mass terms of the SM particles to be effective, the lightest SUSY particles should not be heavier than 1 TeV [54].

Other than doubling the amount of particles, the MSSM predicts a larger Higgs sector, namely five Higgs bosons: h^0, H^0, A^0 and $H^{\pm}$. This is achieved by having two complex Higgs doublets instead of one, giving mass to the u-type and d-type quarks separately as

$$H_u : \begin{pmatrix} H_u^+ \\ H_u^0 \end{pmatrix}, \quad \begin{pmatrix} \tilde{H}_u^+ \\ \tilde{H}_u^0 \end{pmatrix} \tag{1.14}$$

and

$$H_d : \begin{pmatrix} H_d^0 \\ H_d^- \end{pmatrix}, \quad \begin{pmatrix} \tilde{H}_d^0 \\ \tilde{H}_d^- \end{pmatrix} \tag{1.15}$$

The MSSM predicts the lightest of the Higgs bosons h to be of mass $m_h < 135$ GeV, which so far, could be satisfied by the newly discovered Higgs boson. Further searches for the other Higgs bosons of higher (A) or even lower (h) masses and a good description of background processes for these searches are therefore of utmost importance.

The MSSM also features several problems, such as flavor invariance, CP invariance and the μ-problem. The μ-problem arises from the Higgsino mass parameter μ appearing in a $\mu H_u H_d$ term in the superpotential in Eq. 1.12. This μ parameter should have the same order of magnitude

as the electroweak scale, many orders of magnitude smaller than the natural cutoff scale - the Planck mass ($M_{Planck} = 10^{19}$ GeV). Again, this is a problem of naturalness. Many other SUSY models have been proposed to deal with the shortcomings of the MSSM. For example, the Next-to-Minimal Supersymmetric Model (NMSSM) solves the μ problem by adding a new singlet chiral superfield S [55, 56]. The NMSSM features seven physical Higgs bosons labeled as $H_1 \cdots H_7$ ordered by their mass, with the lightest $m_{H_1} \approx 1$ GeV.

In a specific NMSSM scenario presented in Ref. [57], the couplings of the light singlet-like scalar S to the b and τ are strongly suppressed. Due to that, the branching ratio of the S into two photons would be enhanced, and in agreement with the LEP limits [58], the light scalar could lie undiscovered in the diphoton channel at the LHC in the range of $60 < m_{\gamma\gamma} < 110$ GeV, with a possible signal strength up to two times larger than the Higgs boson at 126 GeV.

Two Higgs Doublet Models (2HDM) extend the SM with a second Higgs doublet, therefore obtaining five physical Higgs bosons. These models are often considered within SUSY, describing the Higgs sector. There are several types of 2HDM models [59, 60]:

- Type I fermiophobic model where charged fermions couple only to the second doublet;
- Type II MSSM-like model, with u and d type quarks couple to separate doublets;
- Type II Lepton-specific model, where charged leptons couple to the first doublet and quarks couple to the second doublet;
- Type II Flipped model, where u type quarks and charged leptons couple to the second doublet and d type quarks to the first doublet;
- Type III model, where both quarks and charged leptons couple to both doublets, making flavor changing neutral currents possible at tree level.

These models have six free parameters, namely the masses of the Higgs bosons (m_h, m_H, m_A and $m_{H^\pm}$), the ratio of the two Higgs fields vacuum expectation values conventionally written as

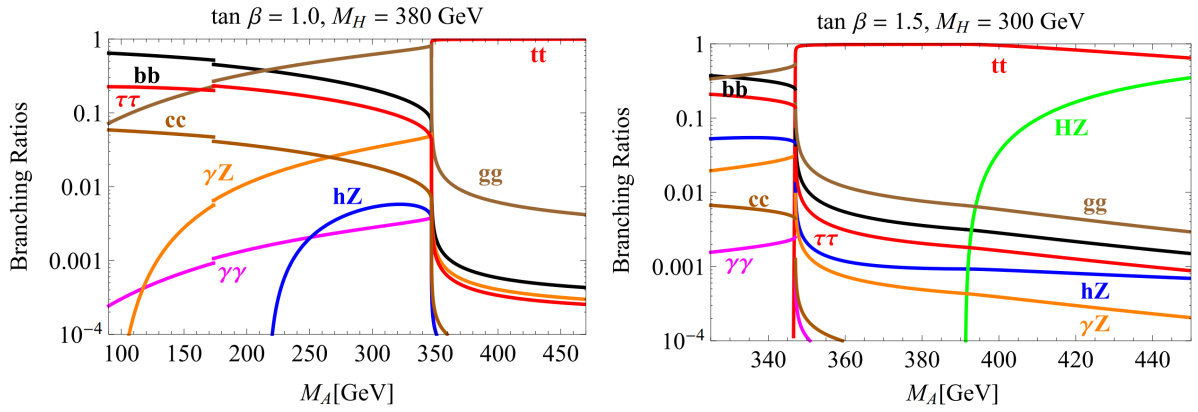
$$\tan\beta = \frac{v_u}{v_d}, \qquad 0 \leq \beta \leq \pi/2. \tag{1.16}$$

and a mixing angle of the two doublets α expressed as

$$\cos^2(\beta - \alpha) = \frac{m^2_{h^0}(m^2_Z - m^2_{h^0})}{m^2_{A^0}(m^2_{H^0} - m^2_{h^0})} \tag{1.17}$$

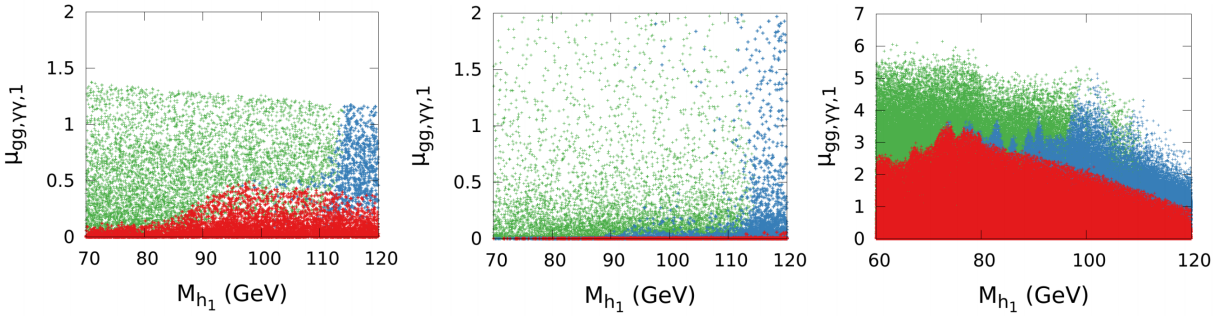
at tree level. Depending on the value of the α and β parameters, production of the pseudoscalar A boson can be enhanced. In case of the A being decoupled from the Z and $W^\pm$ bosons, the possibility of a discovery of a narrow low mass A (below the $t\bar{t}$ and HZ threshold) at the LHC in the diphoton channel emerges [61]. The branching ratios of such an A boson are shown in Figure 1.6. Photons present a much cleaner signal than other final states such as jets or neutrinos (missing-E_{T}) and can therefore serve as the discovery channel even with a much smaller branching fraction w.r.t. the other channels.

Another study of the discovery potential of a light scalar at the LHC was presented in Ref. [62], within the frame of the 2HDM and NMSSM models, but also parameterizing the light scalar in a model independent way, normalizing its couplings to those of a SM Higgs boson of the same mass. The results in Figure 1.7 show that there are regions of the phase space where a light scalar discovery at $m_{\gamma\gamma} < 110$ GeV is possible at the LHC.



(a) (b)

Figure 1.6: The pseudo-scalar CP odd Higgs boson branching ratios for (a) $\alpha = 0.78$ and $\tan\beta = 1.0$ and (b) $\alpha = 0.58$ and $\tan\beta = 1.5$ [61].



(a) (b) (c)

Figure 1.7: Signal strength normalized to the SM in the $gg \to h_1 \to \gamma\gamma$ channel ($m_{h_1} < m_H$, $H = h_2$). Points passing flavor and theoretical constraints (green), also passing the LEP constraints (blue) and points passing also LHC couplings constraint on h_2 (red) [62].

1.4 Diphoton pair production at the LHC

The spectrum of the invariant mass of two prompt photons produced at the LHC by QCD processes, has a falling exponential shape. Any possible resonance produced in addition to the SM diphoton production would be evident as a peak on top of the continuum background. Thus in a search for a new resonance one must properly know the background.

The SM prompt diphoton production can be summarized as direct photon production and production of fragmentation photons, where one or both of the two final state photons actually comes from the fragmentation of a quark or a gluon.

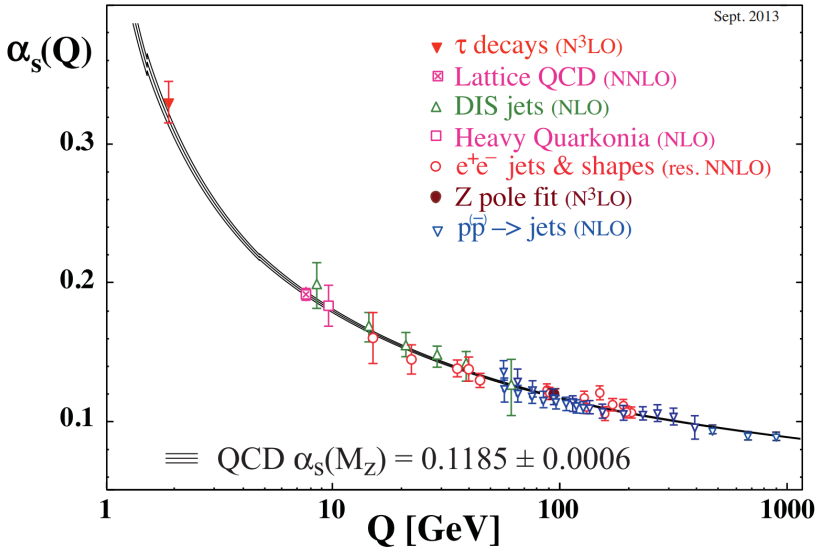


Figure 1.8: Running of the strong coupling constant as a function of the momentum transfer in the interaction [63].

1.4.1 Perturbative QCD

As mentioned in Section 1.1, QCD is the quantum field theory that describes strong interactions. Protons colliding at the LHC are composite particles. Therefore a scattering of two protons is always, in fact, a scattering between their constituents - quarks and gluons. Quarks inside a proton, which form the proton itself are called *valence quarks*. Gluons within the proton can split and produce so-called *sea quarks*, quickly annihilating back into a gluon. The behavior of the strong interaction, due to its confined nature, is not constant with energy and the strong coupling strength can be expressed as

$$g_s^2 = 4\pi\alpha_s, \tag{1.18}$$

where α_s is the strong coupling constant. The evolution of α_s with the momentum transferred in the interaction is illustrated in Figure 1.8. Processes occurring at lower energies can be described in a non-perturbative regime by soft QCD. These soft processes occur at LHC and form the so-called underlying event, present in pp collisions usually as background.

Partons carrying a higher momentum fraction of the proton produce hard interactions, which can be described by perturbative QCD. High energy cross sections can be calculated applying the factorization theorem [64]. Using a perturbation series, a physical quantity is expressed in different orders of α_s (leading order (LO), next-to-leading order (NLO), next-to-next-to-leading order (NNLO) and so on), which can be related to the QCD scale parameter Λ_{QCD} as

$$\alpha_s(Q^2) \propto \frac{1}{\ln\left(\frac{Q^2}{\Lambda^2_{QCD}}\right)}. \tag{1.19}$$

Within the factorization theorem, the proton is described as a collection of collinear partons, each carrying different fractions of the proton momentum. The probability of a parton of flavor

q carrying a certain fraction x of the proton momentum for a momentum transfer Q^2 can be expressed in terms of a Parton Distribution Function (PDF) $f_q(x, Q^2)$. The PDFs are obtained by fits of theoretical predictions to experimental data and provide a non-perturbative input to the perturbative calculation of cross sections. When factorizing the non-perturbative proton dynamics from the perturbative hard cross sections, one can vary the factorization scale and obtain a renormalization group equation (DGLAP equation [65]) for PDFs and derive their evolution as a function of the momentum transfer Q^2. The parameterizations of the PDFs are derived by PDF working groups, such as the CTEQ [66] and MSTW [67]. A recent calculation of the PDFs by the CTEQ group is shown in Figure 1.9.

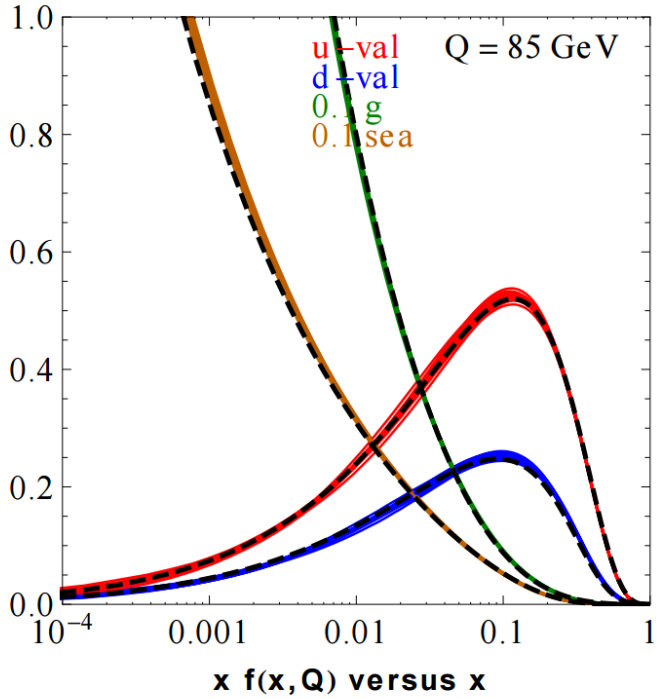


Figure 1.9: NNLO PDFs by the CT10 group for Q^2 value of 85 GeV^2 showing the Hessian error PDFs from the CT10 NNLO analysis with $xu_{valence} = x(u-\bar{u}), xd_{valence} = x(d-\bar{d}), 0.10xg$ amd $0.10xq_{sea}$ as function of x. The quark sea contribution is $q_{sea} = 2(\bar{d}+\bar{u}+\bar{s})$. The dashed curves are the central CT10 NLO fit [66].

The cross section of a hard process in a hadron-hadron collision can be written as

$$\sigma_{p_1p_2} = \int dx_1 f_{q_1/p_1}(x_1, \mu_F^2) \int dx_2 f_{q_2/p_2}(x_2, \mu_F^2) \times \hat{\sigma}(x_1, x_2, \mu_R^2), \tag{1.20}$$

including the proton PDFs and the partonic cross section $\hat{\sigma}$. μ_F is the factorization scale, separating long and short distance physics (describing high and low momentum transfer in an interaction) and μ_R is the running coupling renormalization scale. In case of a cross section of two protons into two photons, a fragmentation component $D(z, Q^2)$ needs to be considered, giving the probability that a parton will produce a fragmentation photon of a fraction of momentum z when hadronizing. It is visible in Figure 1.9 that the gluon PDF is very large and a gluon can

easily take 50% of the momentum of the proton. For this reason, processes involving gluons, such as the Higgs boson production via gluon-gluon fusion, can have high cross sections at the LHC, despite a modest value of the corresponding matrix-element cross section $\hat{\sigma}$.

1.4.2 Diphoton pair production processes at LHC

The production cross section of prompt diphoton events at the LHC can be mainly accounted for by the three processes, displayed in Figure 1.10:

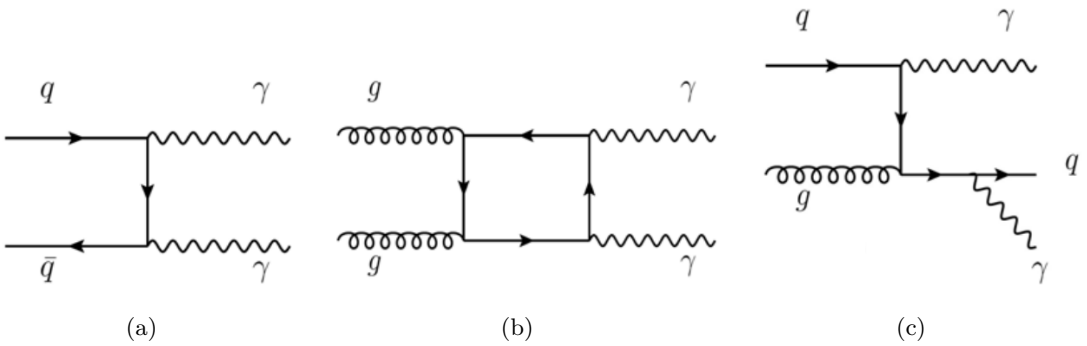


Figure 1.10: Feynman diagrams:(a) the Born diagram (LO), (b) the box diagram (NLO) and (c) the bremsstrahlung diagram (NLO).

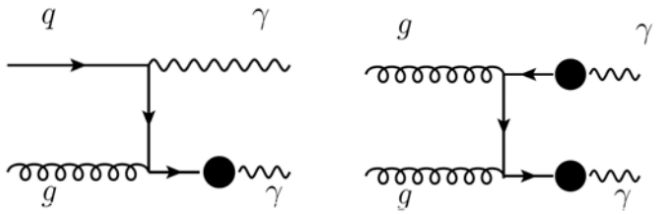
- the Born process - annihilation of two quarks $q\bar{q} \to \gamma\gamma$ of the order of α^2_{QED};
- the box diagram - $gg \to \gamma\gamma$ of the order of $\alpha^2_S\alpha^2_{QED}$. Even though this process is a next to leading order process, the gluon PDFs are enhanced at the LHC, therefore its cross section can be $\approx 30\%$ of the Born process cross section;
- the bremsstrahlung process - $qg \to q\gamma\gamma$ of the order of $\alpha_S\alpha^2_{QED}$ when a quark radiates a high energy photon. This process is also enhanced due to the large gluon PDFs.

Higher order processes can occur, where an extra parton can be emitted (radiative corrections) or emitted and absorbed back (virtual corrections). In case of the Born diagram, these increase its cross section by 30%. Virtual and radiative corrections to the box diagram have also been computed [68].

Experimentally, it is impossible to distinguish between a photon being emitted off a parton (Figure 1.10(b)) and a fragmentation photon (Figure 1.11(a)), that arose from the parton fragmentation. The distinction only holds at LO and becomes ambiguous at NLO.

A fragmentation photon can be produced in one or both partons of a two parton final state diagram and can even give a dominant contribution into the diphoton channel (for $80 < m_{\gamma\gamma} < 140$ GeV [69]). When calculating the fragmentation contribution, for example in the DIPHOX generator [69], the singularities occurring for very small angles between the parton and the photon are factorized for all orders in α_s according to the factorization property and absorbed into quark and gluon fragmentation functions of a photon ($D(z, Q^2)$). Corrections to the fragmentation contributions are also computed, since they make up a non-negligible component of the

total cross section. The two dominant fragmentation photon diagrams, due to the large gluon PDFs, are shown in Figure 1.11.



(a) (b)

Figure 1.11: Feynman diagrams with (a) one photon coming from fragmentation and (b) both photons coming from fragmentation.

1.4.3 Reducible background of the prompt diphoton production

In addition to the direct, bremsstrahlung and fragmentation diphoton production described in the previous section, events where one or both of the photons, considered to be from the hard process, are in fact coming from the decay of a neutral hadron in a jet, such as the π^0, η (later called jet background), compose a reducible background to the diphoton production. In experimental studies, jets can be misidentified as photons if their energy deposits mimic the properties of a photon in the detector. Special studies are devoted to separate jets from genuine prompt photons when measuring cross sections (see Section 7.3), since the cross sections of $\gamma - jet$ and $jet - jet$ production at the LHC are 10^3 and 10^7 times larger than the $\gamma\gamma$ cross section, respectively.

Figure 1.12 shows two processes that contribute at the order of $\alpha_{QED}\alpha_s$. The qg scattering contributes by 95% and the $q\bar{q}$ by 5%. The three main dijet contributions are shown in Figure 1.13. Initial or final state parton emission comprise higher order corrections.

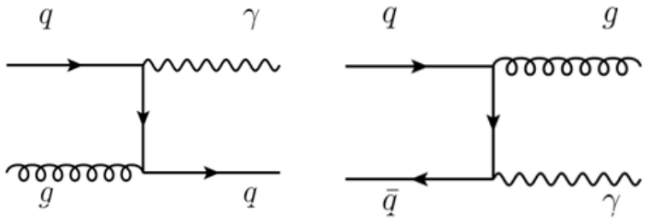


Figure 1.12: Feynman diagrams of the $\gamma - jet$ final state processes.

Another fraction of measured diphoton events at the LHC actually comes from electrons being misidentified as photons. The electron background consists of non-resonant (Drell-Yan) and resonant (Z) component and it is most prominent around the dielectron invariant mass peak at approximately 91 GeV, corresponding to the Z boson resonance, while the continuum

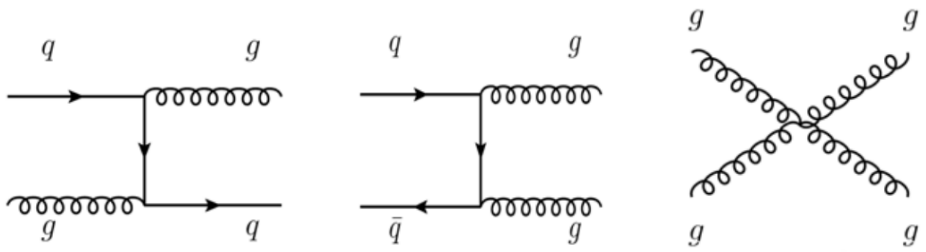


Figure 1.13: Feynman diagrams of the $jet - jet$ processes.

slowly fades away at high masses, representing $\approx 2 - 4\%$ of the $m_{\gamma\gamma}$ spectrum on average (see Section 7.4).

1.4.4 Theoretical predictions

In general, theoretical predictions can be calculated at fixed order, such as the LO, NLO, etc. and in addition, some generators can provide parton shower information [70], therefore providing the transition from parton level to the particle level.

Monte Carlo (MC) generators, such as Pythia [71] and Sherpa [72], provide a LO and NLO calculation, respectively, and use the parton shower to include initial and final state radiative corrections. This way, the possible double counting of the radiation from the initial or final state and that from the parton shower computation has to be taken into account [73, 74]. Following the parton shower, the obtained partons proceed by emitting gluons, that further split into quark pairs and so on. This process when multiple partons combine into the observable particles is called hadronization. These truth level particles can then be passed through the ATLAS detectors, simulating their passage and detection using Geant4 [75] and the resulting MC sample can be compared to data.

In the case of other generators, which can compute physics processes with more precision at the NLO or NNLO, but without the application of parton shower, one can compare various kinematic distributions to data. However, the data in such a comparison has to be corrected for detector effects, also known as the process of unfolding to the truth particle level, using a MC sample with parton shower, to be able to compare the production directly to the prediction, which only contains the partons produced without hadronization and interaction with the detector.

The diphoton production, from a theoretical point of view, is well known up to the NNLO level in QCD [76], for the Born and the bremsstrahlung processes, presented Section 1.4.2. The non-perturbative fragmentation component for a process involving two photons (and any number of jets), where one or both photons may come from fragmentation of a quark, is very difficult to compute due to the singularities arising at small angles.

Theoretical calculations deal with this problem by imposing a cut on the energy around the photon, called the isolation energy. In order to reduce fragmentation contributions, one imposes a veto on all photons where a parton is collinear to the photon, by applying a cone isolation criterion, requiring the hadronic activity around the photon to be smaller than a certain threshold (typically a few GeV). This, however, spoils the cancellation of Infrared (IR) singularities since

soft gluons cannot be emitted inside the cone. Another approach - the Frixione isolation [77], allows the cancellation of the fragmentation component and is IR safe at the same time. Inside a cone of radius r_γ around the photon the maximum hadronic energy must satisfy

$$E_{had,max}(r_\gamma) < \epsilon p_{\mathrm{T}}^{\gamma} \left(\frac{1-\cos r_\gamma}{1-\cos R} \right)^n , \tag{1.21}$$

where n, ϵ are usually set to one, n is the power with which the cone radius changes and ϵ is the fraction of the transverse momentum, also adjustable. Within a limit of $R \to 0$ the $E_{had,max}(r_\gamma)$ goes to zero. By applying progressively tighter cuts on smaller and smaller cones, the Frixione isolation prescription eliminates the collinear fragmentation component and leaves only the direct component. The Frixione isolation simplifies the theoretical calculations but it is not suitable for experimental measurements due to the finite segmentation of the detectors. A discrete version of the Frixione isolation was considered within the DIPHOX generator [78].

In experimental measurements, isolated photons are considered in order to best reject the jet background. A transverse isolation energy E_{T}^{iso} is defined in a finite cone around the photon and a cut on this variable is imposed. When comparing the experimental results to the theoretical predictions, the isolation requirement is usually not the same, since the theoretical calculations are preferentially computed using Frixione isolation. The comparison is therefore less meaningful, since in experiments, one cannot remove the fragmentation component, while the Frixione isolation guarantees its absence in theoretical calculations.

1.5 Diphotons in association with jets at the LHC

Production of two prompt photons with one or more additional jets can be achieved, for example, via the processes shown in Figure 1.10, with one or more gluons being radiated off of a quark leg(s) of the diagrams.

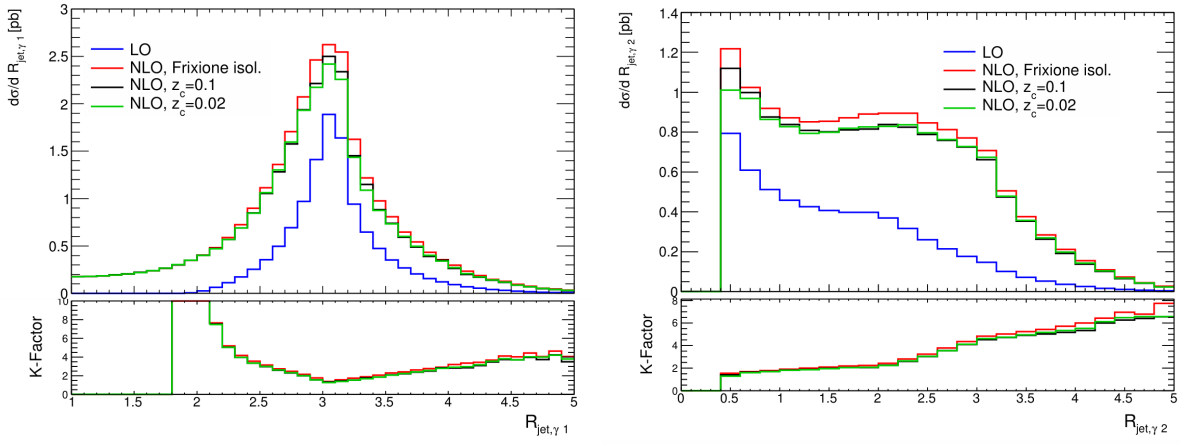
Measuring the cross sections of prompt diphotons in association with (at least) two jets can provide a good description of background to the $H \to \gamma\gamma$ produced via the VBF process. Searching for Higgs pair production in the $\gamma\gamma b\bar{b}$ channel and studying the Higgs self-coupling, will also need proper description of the background, even though these processes are not yet in reach of the LHC. In addition, when searching for the extended Higgs sector, the knowledge of background is also essential.

1.5.1 Diphoton production in association with (at least) one jet

The $pp \to \gamma\gamma + j$ process has been computed at NLO [79] especially for LHC conditions of $\sqrt{s} = 14$ TeV. These calculations comprise the both the LO ($q\bar{q} \to \gamma\gamma g$, $qg \to q\gamma\gamma$ and $\bar{q}g \to \bar{q}\gamma\gamma$) and NLO processes, where two gluons, two quarks or one quark and one gluon is produced together with the two photons in the final state.

A more recent publication [80], also at NLO using GOSAM [81] and MadGraph [82], explores the impact of the use of Frixione ($R = 0.4$, $n = 1$, $\epsilon = 0.5$) or fixed cone isolation criteria for the photons. As PDF input, the NLO parton distribution set from NNPDF2.3 [83] is used. The provided variables are presented for both the exclusive $\gamma\gamma + j$ and inclusive $\gamma\gamma + \geq 1j$. The difference between the LO and NLO predictions, as seen for example for the R separation (see

Eq. 2.6) of the leading photon and the highest-p_{T} jet for the inclusive case in Figure 1.14, is substantial and the spectrum should be measured experimentally at the LHC.



(a) (b)

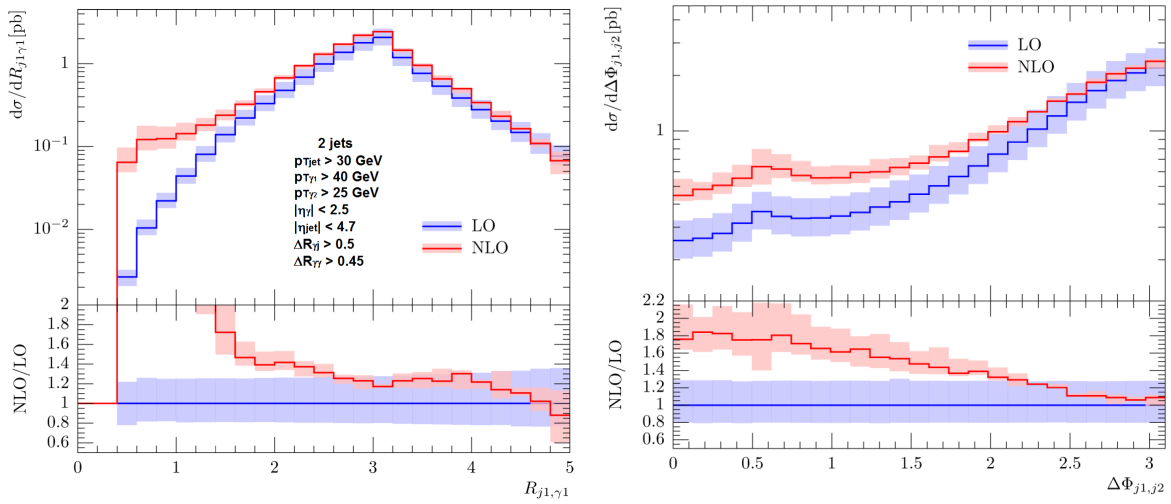
Figure 1.14: R-separation between the (a) leading photon and (b) subleading photon and the leading jet in the $\eta - \phi$ plane for $\gamma\gamma j$ inclusive cross section [80].

1.5.2 Diphoton production in association with (at least) two jets

The same group as in Ref. [79] also computed NLO corrections to the $\gamma\gamma + 2$ jets production [84], using the same tools as mentioned above, the CT10 set of parton PDFs [85], and Frixione isolation with parameters $R = 0.4$ as defined in Eq. 2.6, $n = 1$ and $\epsilon = 0.05$, in order to suppress the impact of fragmentation photons. Figure 1.15(a) shows the R-separation (defined in Eq. 2.6) of the leading photon and the leading jet being different for the LO and the NLO, same can be seen in Figure 1.15(b) for the azimuthal angle ϕ separation of the two leading jets, calling for an experimental measurement and comparison.

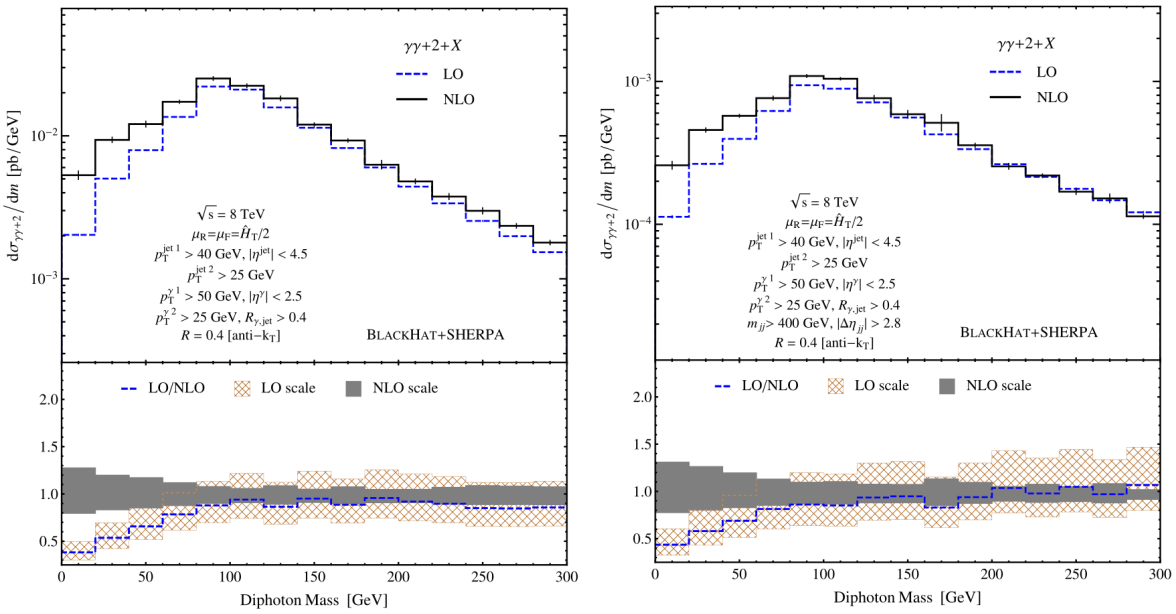
NLO calculations of diphoton production in association with two jets have been also performed by another group [86] using Sherpa [72] to simulate the partonic sub-processes at tree level and the Blackhat library [87] for virtual corrections. They also use Frixione isolation with the same parameters ($R = 0.4$, $n = 1$, $\epsilon = 0.5$) to remove the fragmentation component. In a previous paper, it was shown on prompt photon cross sections, when comparing them to measured CMS values [88], where a standard cone isolation was used with a cut of 5 GeV, that the difference between Frixione and cone isolations is below 1% of the cross sections [89].

The PDF input comes from the MSTW2008 LO and NLO PDFs [67]. The difference between the LO and NLO computation is evident in the diphoton invariant mass results, as seen in Figure 1.16(a). Ref. [86] was mostly aimed at the determination of the background to the Higgs boson, hence the m_{jj} and $|\Delta\eta_{jj}|$ requirement of the VBF production-specific cuts in Figure 1.16(b).



(a) (b)

Figure 1.15: (a) R-separation between the leading photon and the leading jet and (b) ϕ separation between the leading jet and the subleading jet for $\gamma\gamma jj$ cross section [84].



(a) (b)

Figure 1.16: Diphoton invariant mass distribution (a) without and (b) with VBF-specific kinematic cuts [86].

1.5.3 Diphoton production in association with up to three jets

In a recent paper [90], calculations of NLO corrections for processes with up to three jets were presented, using the NJET 2.0 code [91] and SHERPA, implementing up to four jets together with the production of two photons. A small effect of top quark loops in virtual amplitudes and in the case of $\gamma\gamma + 3j$ also a contribution from vector loops where photons couple directly to a virtual fermion loop are neglected, but should not amount to more than 0.5% of the total cross section in the $\gamma\gamma + 2j$ process, therefore it is negligible for the $\gamma\gamma + 3j$ process. Four different PDF sets are used in comparison. To suppress the fragmentation component, Frixione isolation is used with the parameters $R = 0.4$, $n = 1$ and $\epsilon = 0.05$. A ratio of the cross section of $pp \to \gamma\gamma + 3j$ over $pp \to \gamma\gamma + 2j$, as seen in Figure 1.17, shows a large difference between the LO and the NLO and is interesting to be measured at LHC due to the cancellation of many systematic uncertainties.

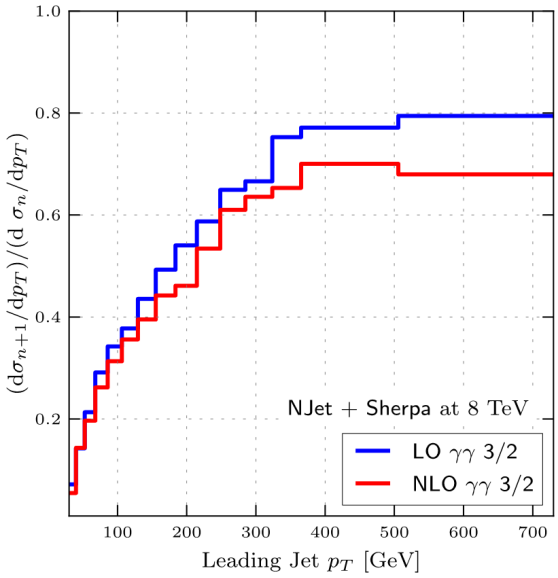


Figure 1.17: Ratio of $pp \to \gamma\gamma + 3j$ over $pp \to \gamma\gamma + 2j$ as a function of the leading jet p_{T} [90].

1.6 Previous measurements of diphoton production

1.6.1 Prompt diphoton and photons in association with jets cross section measurements

The diphoton final state was previously studied at Tevatron, in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV by the DØ experiment using 4.2 fb^{-1} [92] and 8.5 fb^{-1} of data [93]. The CDF experiment also measured the prompt isolated diphoton production cross section using 9.5 fb^{-1} of data [94].

At the LHC, the diphoton cross section was measured by the ATLAS Collaboration in pp collisions at $\sqrt{s} = 7$ TeV using 37 pb^{-1} of data [95] and the full set of 4.9 fb^{-1} of data [96]. The full dataset of 20.3 fb^{-1} collected at $\sqrt{s} = 8$ TeV will soon be published as well. The CMS

Collaboration measured the diphoton cross section using 36 pb^{-1} of $\sqrt{s} = 7$ TeV data [97] and with the full 5.0 fb^{-1} of pp collision dataset at $\sqrt{s} = 7$ TeV [98].

The single isolated photon production cross section in association with jets was measured at $\sqrt{s} = 7$ TeV using 37 pb^{-1} by ATLAS [99, 100]. In CMS, the photon plus jets production was measured using 2.1 fb^{-1} of $\sqrt{s} = 7$ TeV data [101]. At Tevatron, the DØ collaboration measured the differential cross section of photon plus jet events using 1 fb^{-1} [102] and 8.7 fb^{-1} of $\sqrt{s} = 1.96$ TeV data [103]. The CDF collaboration has measured the direct photon production in association with a heavy quark using 9.1 fb^{-1} of $\sqrt{s} = 1.96$ TeV data [104].

1.6.2 Previous searches for the SM Higgs and BSM physics in the diphoton channel

The search for the Higgs boson at LEP resulted into its final results [58], combined from all four collaborations (ALEPH, DELPHI, L3 and OPAL) set a limit on its mass at 114.4 GeV, using 2461 pb^{-1} of data collected in e^+e^- collisions at $\sqrt{s} = 189 - 209$ GeV. The ALEPH Collaboration observed a small excess at 114 GeV in the diphoton invariant mass spectrum, which was not significant enough for a discovery and set a limit for a fermiophobic Higgs boson at $m_H^{ferm.} > 109.7$ GeV [105]. A Higgs boson of the extended Higgs sector could, however, be still undetected even below 100 GeV. LEP did not really have sensitivity to a SM Higgs boson in the $\gamma\gamma$ channel, due to its low branching fraction which is why fermiophobic models were considered, where the diphoton channel could carry a large fraction of the branching ratio and be enhanced in this channel.

At Tevatron, the final search for a Higgs boson by the DØ Collaboration in the diphoton channel was performed with a 9.6 fb^{-1} dataset of $\sqrt{s} = 1.96$ TeV $p\bar{p}$ collision data and a 95% C.L. limit was put on the Higgs decaying into two photons branching fraction in the mass region of $100 < m_H < 150$ GeV [106], but no exclusion was made. A 95% C.L. exclusion was obtained for a fermiophobic Higgs for masses of $100 < m_H < 113$ GeV. The CDF Collaboration produced a similar limit and in addition excluded a fermiophobic Higgs decaying into diphotons below 114 GeV [107].

Leading up to the discovery of the Higgs boson in 2012, both ATLAS and CMS Collaborations released several iterations of the 95% C.L. exclusion limit in the diphoton channel [108, 109, 110, 111, 112, 113, 114, 115], but both experiments limited themselves to the last two “windows" in the invariant mass spectrum, which had not yet been excluded by Tevatron or LEP previously and only showed the region of $110 < m_H < 140 - 150$ GeV. The SM Higgs branching fraction into two photons (mediated by a loop of t or W) falls very fast above 150 GeV, so searching in this channel at $m_{\gamma\gamma} > 150$ GeV can only be useful when looking for BSM particles.

Shortly after the search for low and high mass diphoton resonances, described in this work in Chapter 6, was made public at the LHCP 2014 conference in June [116], the CMS Collaboration presented a similar analysis at ICHEP 2014 [117]. The focus was put on the high $m_{\gamma\gamma}$ spectrum and a spin 0 and spin 2 model was considered for a narrow resonance, and a spin 0 model for a wide resonance, as seen in Figure 1.18. No significant excess was observed.

Recently, a paper was published by ATLAS, looking for exotic resonances in the diphoton channel in the diphoton invariant mass range of $409 < m_{\gamma\gamma} < 3000$ GeV [118] using the full $\sqrt{s} = 8$ TeV dataset. The analysis was performed looking for a Randall-Sundrum (RS) graviton [51] or any similar high-mass diphoton resonance, and the signal was modelled with a width ranging from $\Gamma_G = 39$ MeV at $m_G = 300$ GeV up to $\Gamma_G = 31$ GeV at $m_G = 2250$ GeV according to

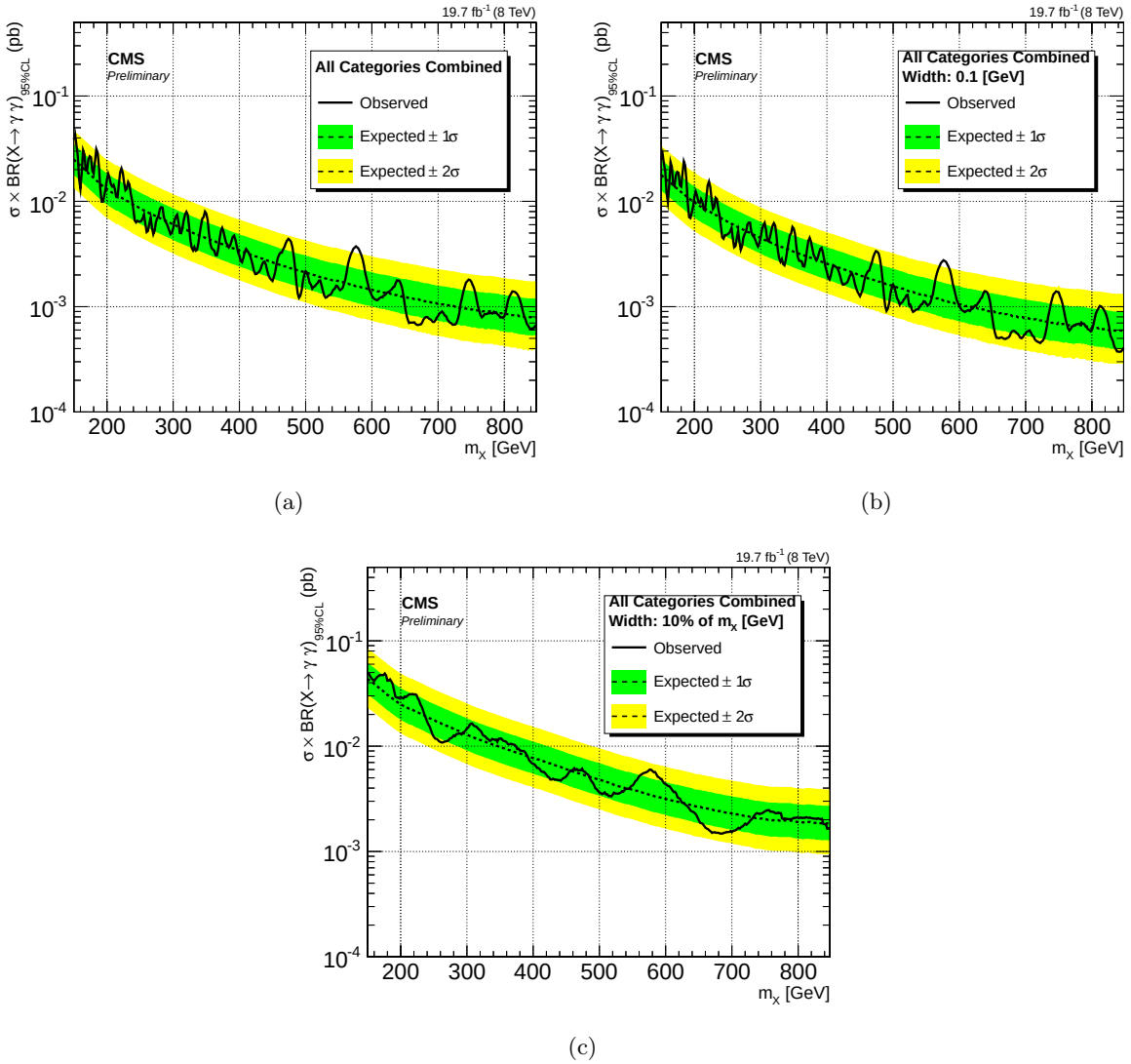


Figure 1.18: Exclusion limit at 95% C.L. on the cross section times the branching ratio of a new (a) spin 2 (b) narrow spin 0 and (c) wide spin 0 resonance decaying into two photons as a function of the resonance mass m_X [117].

the RS model. The data was not fit using sidebands (as was done in the analysis presented in Chapter 6), due to the lack of the high-mass events. PYTHIA and DIPHOX MC samples were normalized to a low-mass control region and used for the fit, which is why the analysis could be performed so high in $m_{\gamma\gamma}$. No significant excess was found and a limit was placed on the $\sigma \times BR(G^* \to \gamma\gamma)$ production.

So far, no new particle, except the Higgs of $m_H \approx 126$ GeV has been discovered.

2 The LHC and the ATLAS detector

2.1 The Large Hadron Collider

With its 27 km in circumference, the Large Hadron Collider (LHC) is the world's largest particle accelerator, capable of colliding protons and heavy ions at the TeV scale of the center of mass energy (c.m.). The concept of the LHC was first introduced in 1984, when the motivation for increasing the collision energy was supported by the recent discovery of the W and Z bosons. First, the Large Electron-Positron collider (LEP) was built, in order perform precision studies, only to be later replaced by the LHC, the new discovery machine. The construction of the LHC began in 2000, in the same underground tunnel that was used by LEP for the previous eleven years. The existing tunnel and other supporting facilities helped cutting the LHC construction time and costs.

2.1.1 LHC design and operation

The LEP collider was used to collide electrons and positrons at c.m. energy of up to $\sqrt{s} = 209$ GeV. Because electrons and positrons are elementary particles, colliding them results in a relatively clean final state, where all of the initial collision energy is effectively used in the hard process and can be precisely measured by detectors. On the downside, it is very difficult in practice to accelerate and keep electrons and positrons at a given energy, due to them being very light-weight, producing a large amount of bremsstrahlung radiation on a circular trajectory. In comparison, a hadron collider of similar dimensions and parameters can accelerate protons or heavy ions to higher energies but since those are composite particles, the c.m. energy obtained from an interaction of quarks or gluons will be smaller than the nominal beam energy.

The LHC was designed to collide two proton beams at energies up to $\sqrt{s} = 14$ TeV, meaning accelerating two beams of protons up to 7 TeV. The rise in c.m. energy between LEP and LHC was also enabled by technological progress. LHC was built using superconducting magnets and all its sections responsible for acceleration or bending of the beams have to be cooled down to 1.8K, using liquid Helium.

The assembly of LHC and its four experiments took eight years. First beams were present at the LHC on 10 September 2008. A faulty connection between two superconducting magnets resulted into an accident on 19 September 2008, and the machine had to be stopped for fourteen

months, when repairs and checks had to be done in order to prevent further damage [119]. On 20 November 2009, the LHC started operations again and proton-proton collisions were performed at a lower c.m. energy of $\sqrt{s} = 7$ TeV to guarantee the machine safety. The c.m. energy was increased to 8 TeV in 2012 and the LHC ran its proton-proton run until December 2012, following with a short proton-Pb run in the beginning of 2013. A technical stop planned for 2013 and 2014 allowed the experiments to upgrade their detectors and the LHC to prepare for higher c.m. energies [120], starting operations again in March 2015 and producing first pp collisions at $\sqrt{s} = 13$ TeV on 3 June, 2015.

2.1.1.1 LHC and its injection chain in nominal conditions

The LHC is the last part of a chain of acceleration devices which brings protons from very small energies to the TeV scale. The accelerator complex at CERN is shown in Figure 2.1. Starting

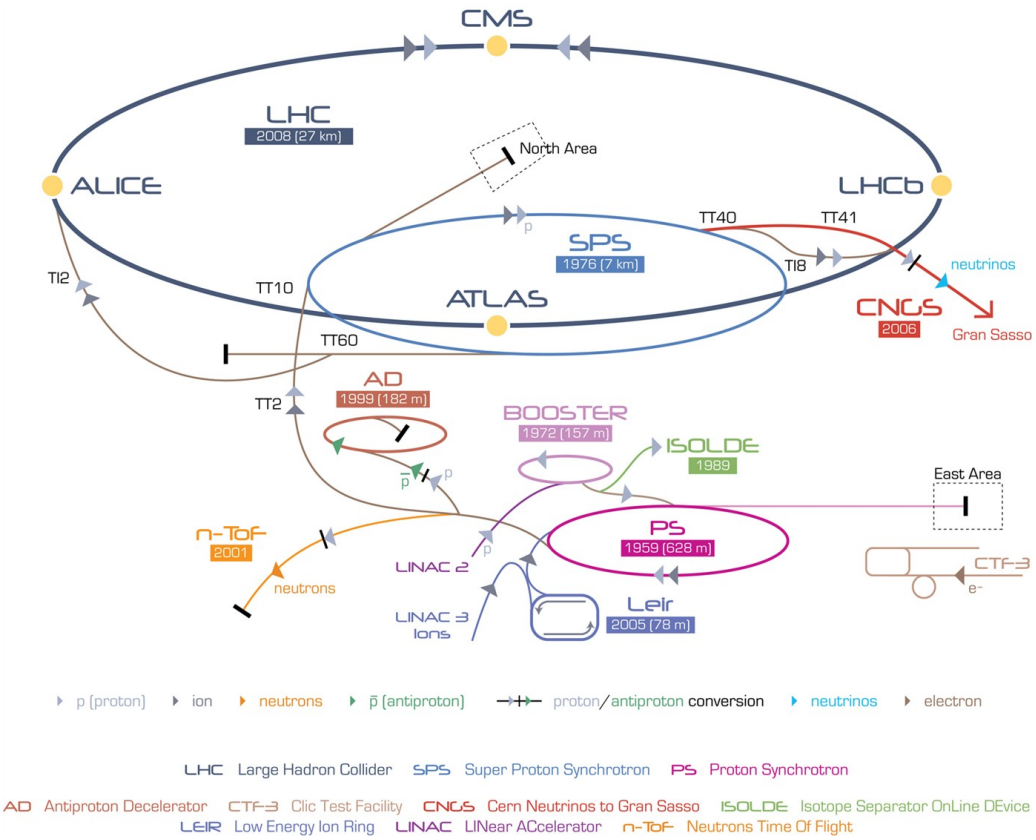


Figure 2.1: The accelerator complex at CERN [121]

with a bottle of hydrogen, the electrons are stripped off the protons by ionization and brought to an energy of 750 keV by the Radio Frequency Quadrupole (RFQ2) from a continuous stream into 6 bunches [122]. These are injected into the LINAC2, a 30 m long linear accelerator, which brings them to 50 MeV, transferring them into the Proton Synchrotron Booster (PSB). The PSB is a circular accelerator with a diameter of 50 m, with four layers of beam pipes, accelerating the protons up to 1.4 GeV in just around 1.2 s and injecting them into the Proton Synchrotron

(PS). This is another circular accelerator, 628.3 m in circumference, which brings protons from 1.4 GeV to 25 GeV and transforms the 6 bunches into 72 bunches, 4 ns long and 25 ns apart, leaving a hole of 320 ns so that the kicker magnets are able to turn on the magnetic field and inject the bunches into the Super Proton Synchrotron (SPS). The SPS, 7 km in circumference, is the last stage of circular acceleration, combining three bunch trains of the PS into 216 bunches, before the two proton beams are injected into the LHC at the energy of 450 GeV.

The LHC is designed to take in 13 fillings of the SPS, forming 2808 bunches, orbiting the LHC approximately 11000 times per second. It can increase the energy of the beams by 485 keV per turn. The beams are housed in vacuum tubes - beam pipes, in eight straight sections and eight bending sections. There are 1232 superconducting dipole magnets operating below 1.9 K in the bending sections of the LHC. These are responsible for bending the beams, unlike in previous accelerators, one dipole contains two beam pipes, the beams are therefore magnetically coupled. There are also 858 quadrupole magnets and higher level magnets to keep the beams as narrow and uniform as necessary. The accelerating section uses a 400 MHz superconducting cavity system to accelerate the beams from 450 GeV to up to 7 TeV. A schema of the LHC is shown in Figure 2.2.

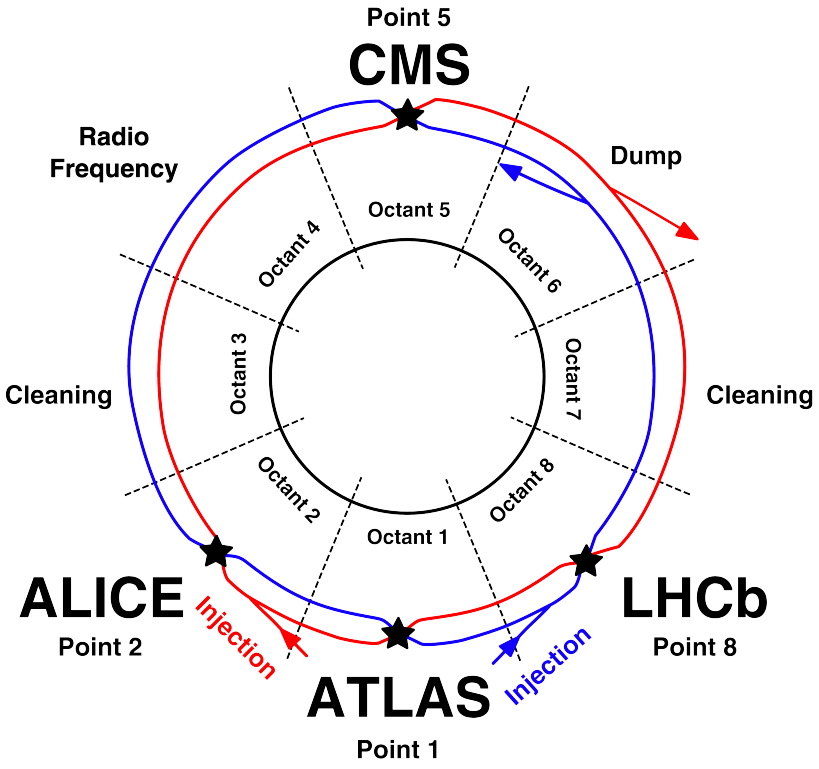


Figure 2.2: A schema of the Large Hadron Collider.

2.1.2 Experiments at the LHC

The LHC has four interaction points occupied by its four main experiments:

- ATLAS (A toroidal LHC ApparatuS) [123], at point 1, a large multi-purpose experiment built to study a wide range of physics phenomena in *pp* collisions;
- LHCb (Large Hadron Collider beauty) [124] an experiment at point 2, designed to study physics of the b quark and measure CP-violation;
- CMS (Compact Muon Solenoid) [125] the second multi-purpose experiment at point 5;
- ALICE (A Large Ion Collider Experiment) [126] at point 8, a heavy-ion oriented experiment, designed to probe the production of the quark-gluon plasma, a state of matter when quarks and gluons are deconfined.

Apart from the main four experiments, there are three smaller experiments dedicated to studying events very close to the beam - at very high pseudorapidity: TOTEM [127], measuring the total *pp* cross section with a luminosity-independent method and studying elastic and diffractive scattering at the LHC, positioned at point 5, close to the CMS detector; LHCf [128] measuring the cross sections and energies of neutral π^0 produced in ATLAS collisions in order to contribute to the understanding of ultra high-energy cosmic ray origin; and MoEDAL [129] searching for magnetic monopole and other highly ionizing particles in the LHCb cavern.

2.1.3 LHC performance

Previous hadron colliders, such as Tevatron or the SPS, used proton-antiproton collisions. Proton-proton collisions were chosen for the LHC, for two main reasons: higher luminosity and higher c.m. energy of the beams. Production of protons by hydrogen ionization is easier than production of antiprotons, and protons can be delivered by the injection chain at high quantities, therefore the LHC can operate at record luminosities - the nominal luminosity of the LHC is $10^{34}\text{cm}^{-2}\text{s}^{-1}$. Also, at LHC c.m. energies, gluon-gluon collisions have higher cross sections due to the gluon PDFs being enhanced (see Figure 1.9), so that new interesting processes, such as the production of the Higgs boson via gluon fusion, can be observed.

The performance of the LHC is characterized by the number of inelastic scattering events generated per second

$$N_{inel} = L \cdot \sigma_{inel}, \tag{2.1}$$

where σ_{inel} is the inelastic cross section and L is the luminosity, measured in $\text{cm}^{-2}\text{s}^{-1}$. The luminosity depends only on parameters of the beam and can be defined for a Gaussian distribution of the beam as

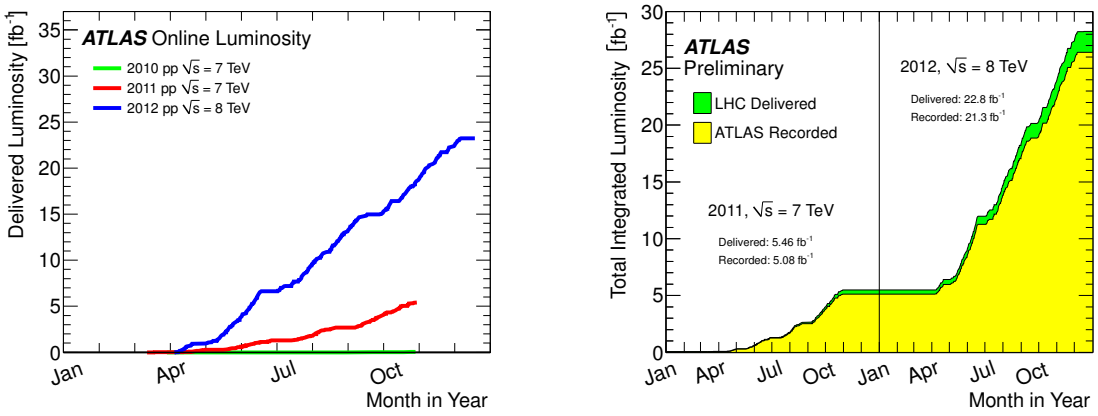
$$L = \frac{N_\text{b}^2 n_\text{b} f_\text{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F, \tag{2.2}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_rev the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalized transverse beam emittance, β^* the beta function[1] at the collision point and F the geometric luminosity reduction factor due to the crossing angle at the interaction point [122].

[1] $\beta^* = \gamma_r.\sigma^2/\epsilon_n$, where σ is the size of the beam in the transverse plane, γ_r is the relativistic gamma factor and ϵ_n is the beam emittance.

The measurement of luminosity is crucial for precision of cross section measurements performed using data from LHC collisions. The ATLAS experiment has two detectors for this purpose: the LUminosity measurement using Cerenkov Integrating Detector (LUCID) and the Beam Condition Monitor (BCM) based on diamond technology.

The luminosity in ATLAS is calibrated using Van der Meer scans [130], which are performed a few times per year. For data collected by the ATLAS detector at 8 TeV, the luminosity was measured with a precision of ±2.8%, derived using the same methodology as in 2011 [131]. The evolution of the luminosity collected by the ATLAS experiment over the whole data taking period of Run 1 (2010–2012) is shown in Figure 2.3(a). The total integrated luminosity recorded during the $\sqrt{s} = 8$ TeV data taking was 21.3 fb^{-1}, as seen in Figure 2.3(b), which came down to 20.3 fb^{-1} of data suitable for physics analysis after requiring a certain level of data quality such as all systems operational, no errors in data collection, transport or storage. The LHC was not



(a) (b)

Figure 2.3: (a) Cumulative luminosity delivered to the ATLAS detector during stable beams for pp collisions vs. time for 2010 data (green), 2011 (red) and 2012 (blue); (b) total integrated luminosity delivered by the LHC (green) and recorded by ATLAS (yellow) vs. time.

operating at its nominal parameters in 2010-2012 and a summary of beam conditions is listed in Table 2.1.

Parameter	2010	2011	2012	design value
Beam energy [TeV]	3.5	3.5	4	7
β^* in IP [m]	3.5	1.5/1.0	0.6	0.55
Bunch spacing [ns]	150	75/50	50	25
Max. number of bunches	368	1380	1380	2808
Max. bunch intensity [p per bunch]	1.2 x 10^{11}	1.45 x 10^{11}	1.7 x 10^{11}	1.15 x 10^{11}
Norm. emittance at start of fill [mm rad]	≈2.0	≈2.4	≈2.5	3.75
Peak luminosity [cm^{-2}s^{-1}]	2.1 x 10^{32}	3.7 x 10^{33}	7.7 x 10^{33}	1 x 10^{34}
Max. mean number of events per BC	4	17	37	19
Stored beam energy [MJ]	≈28	≈110	≈140	362

Table 2.1: LHC operating conditions in Run 1 compared to the desing values [132].

A disadvantage of the high number of closely-spaced bunches in the beam, and therefore high luminosity delivered by the LHC, is the occurrence of multiple collisions in the same bunch-crossing, the so-called *in time pile-up* and collisions taking place immediately before and after the considered collision, the so called *out of time pile-up*. The *in time pile-up*, the more problematic of the two, is expressed by the mean number of interactions per bunch crossing (BC) [133]

$$\mu = \frac{L_{bunch} \times \sigma_{inel}}{f_{rev}}, \tag{2.3}$$

where $L_{bunch} = L/n_b$ is the per bunch instantaneous luminosity . In Figure 2.4, the average number of interactions per bunch-crossing is shown, and it is evident that it increased by a factor of two between the two datasets. This was due to the increase of the number of protons per bunch and the decrease of the β^* at the interaction point (see Table 2.1).

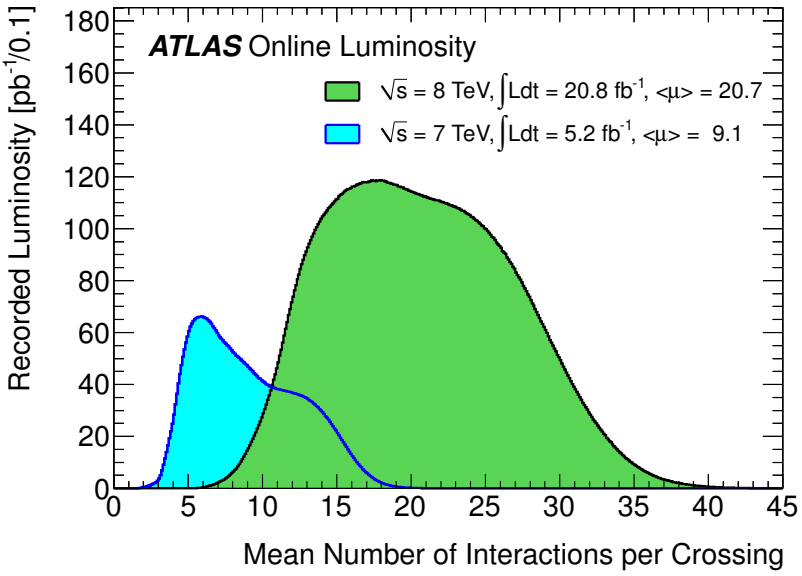


Figure 2.4: Luminosity-weighted distribution of the mean number of interactions per crossing for the full 2011 and 2012 *pp* datasets.

2.1.4 Future of the LHC

Currently, during the LHC Run 2, the experiments are collecting data at c.m. energy of 13 TeV. After the Run 2 of the LHC, which is expected to extend until 2018, another longer shut-down is forseen in order to allow repairs and upgrades of the machine, its injectors and the four detectors. In 2020, the LHC is expected to start its Run 3, running until the end of 2022 and altogether collecting 300 fb^{-1}. Later, some of the detector parts will have to be replaced due to radiation damage and the LHC will be upgraded into High Luminosity LHC (HL-LHC), expected to deliver 3000 fb^{-1}.

2.2 The ATLAS detector

The ATLAS detector is a multi-purpose detector, with full azimuthal coverage and a wide range of pseudorapidity coverage, depending on the subdetector. ATLAS is located at point 1 of the LHC tunnel, 100 m underground, where it was assembled from different parts produced all over the world. A schema of ATLAS is shown in Figure 2.5. With its cylindrical structure, ATLAS is

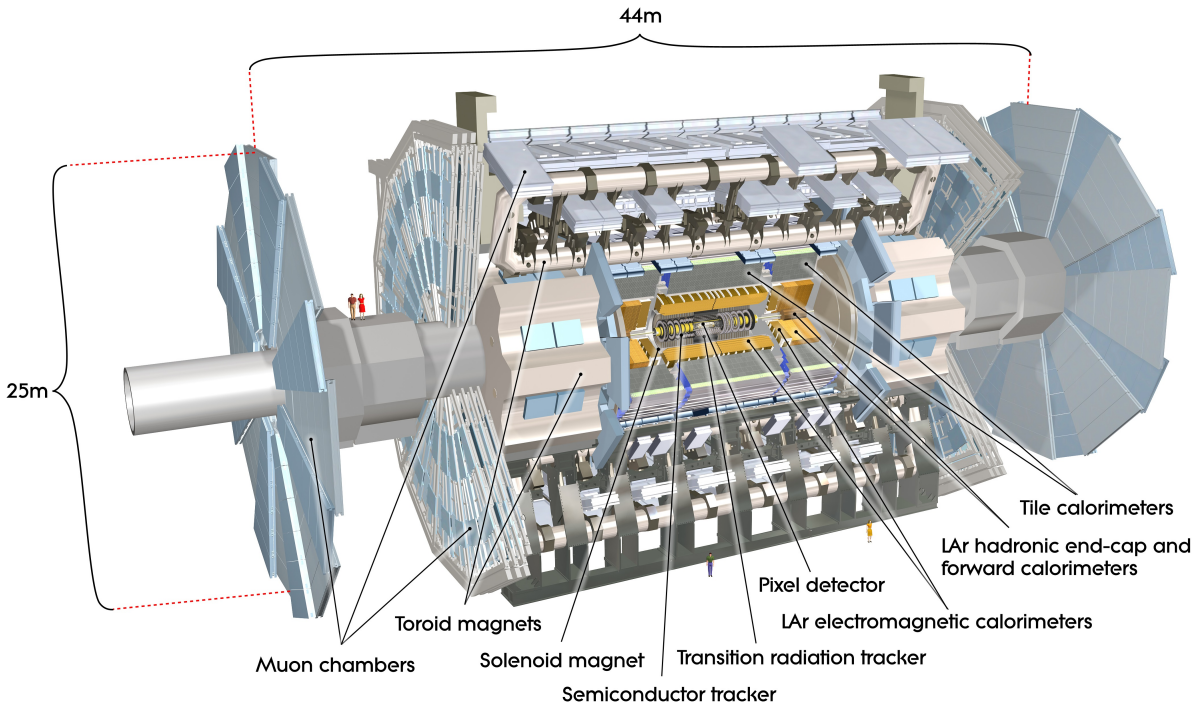


Figure 2.5: The ATLAS detector [134].

46 m long, has 25 m in diameter, and weights over 7000 t. The ATLAS detector was designed to withstand the extreme LHC running conditions of high luminosities and high energies. Emphasis was put on lowering the dead time of the detector and extending spatial acceptance, which means it can measure a wide range of final states produced in the collisions and focus on many areas of high energy physics. The ATLAS Collaboration consists of an unprecedented number of people: approximately 3000 scientists from 174 institutes in 38 countries, out of which roughly 1000 are graduate students. The first letter of intent for the ATLAS experiment was published in 1992 [135] and its construction started in 1997 with first collisions being observed in November 2009. The ATLAS experiment plans, with several upgrades, to take data until 2030.

2.2.1 Overview of ATLAS and its subsystems

2.2.1.1 ATLAS coordinate system

The ATLAS coordinate system starts at the nominal interaction point[2] and is a Cartesian coordinate system, with the z axis in the beam direction. The positive z axis is in the counterclockwise direction of the LHC tunnel. The ATLAS detector positive z axis side is labeled A and the negative is labeled C. The x axis points towards the center of the LHC ring and the y axis points upwards from the beam. A spherical coordinate system is also used, starting also at the nominal interaction point, with the azimuthal angle ϕ defined in the transverse plane to the beam and the polar angle θ defined between the z axis direction and the direction of the measured object. In practice, a Lorentz invariant quantity - the rapidity is used, defined as

$$y = \frac{1}{2}\ln\left[\frac{E + p_z}{E - p_z}\right], \tag{2.4}$$

where E and p_z are the energy and the z component of the momentum. In case of an ultra-relativistic object, pseudorapidity is used, defined as

$$\eta = -\ln\left|\tan\frac{\theta}{2}\right|, \tag{2.5}$$

therefore a pseudorapidity of zero stands for the upward direction and $|\eta| \to \infty$ denotes directions close to the beam. The pseudorapidity is a good approximation of rapidity for $m \ll p$, when $E \approx p$ ($E = p$ for $m = 0$). The $\eta - \phi$ coordinates are typically used to describe the separation of objects in the detector

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.6}$$

The ATLAS experiment itself consists of several subsystems, each specializing in measuring certain physics processes. A measurement of a pp collision event in the ATLAS experiment is a combination of measurements of its subsystems.

2.2.1.2 Magnet system

The ATLAS detector uses magnets to bend the trajectories of charged particles and to provide their charge and momentum measurement. Surrounding the inner detector system, as illustrated in Figure 2.5, is an Al-stabilized NbTi superconducting solenoid, 5.8 m long, 2.56 m in diameter, a 10 cm thick cylindrical magnet designed to produce a 2 T axial field [123]. The solenoid is positioned in front of the calorimeters, hence it was designed with minimal radiative thickness of ≈ 0.66 radiation lengths (X_0) at nominal incidence [136].

The second magnet system, designed to bend the trajectories of muons, is the toroid, which consists of a barrel [137] and two end-cap parts [138] and it is positioned after the calorimeters and before the muon detector. The barrel of the toroid is comprised of eight superconducting coils around the calorimeters and it is 25.3 m long and weighs 830 t. The magnetic field generated by the toroid has a field integral of 3-8 Tm. The end-caps are comprised of eight flat square coil units and eight keystone wedges in a cryostat.

[2] defined as the center of the Inner detector

2.2.1.3 Inner detector

The inner detector (ID) of the ATLAS experiment is a cyllindrical subsystem of detectors with the beam pipe as its axis. It measures 7 m in length and 1.15 m in diameter. The main function of the ID is to track the passage of charged particles through its layers. The trajectories of these particles are bent in the magnetic field generated by the solenoid. The information collected from the layers of the inner detector allows for particle momentum reconstruction and charge assessment. The ID was designed to reconstruct several primary (and secondary) vertices in a single bunch crossing, as seen in Figure 2.6.

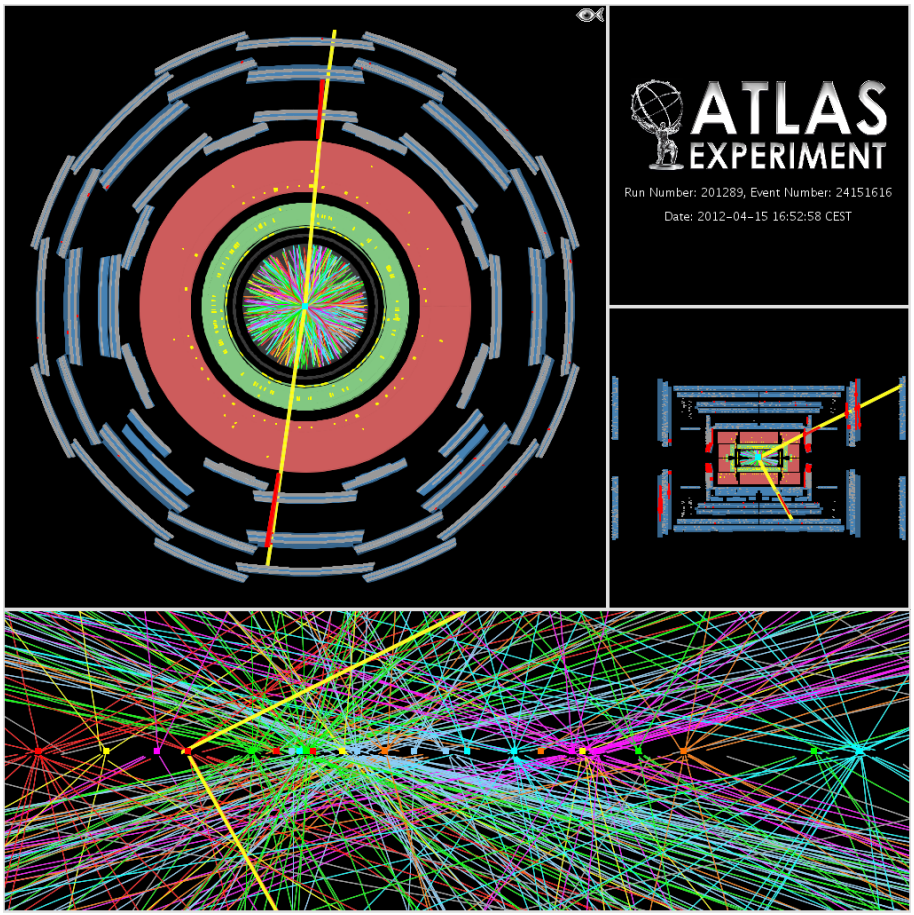


Figure 2.6: Event display of a $Z \to \mu\mu$ candidate with 25 reconstructed vertices recorded on April 15th 2012. The vertices shown are reconstructed using tracks with $p_{\mathrm{T}} > 0.4$ GeV and all tracks are required to have at least 3 Pixel and 6 SCT hits.

The inner detector covers $|\eta| < 2.5$ and it consists of three subsystems. The closest to the beam pipe is the pixel detector with its three layers and three end-cap disks perpendicular to the beam axis. It is built using silicon pixel layers and it has the highest granularity, with 1744 identical sensors. There are 47 232 pixels measuring 50x400μm^2 on each sensor with 46080 readout channels, which makes a total of $\approx$ 80 million channels. The sensors have oxygenated n-type wafers with readout pixels on the n$^+$-implanted side of the detector. This technology was

chosen due to its good charge collection efficiency and the high irradiation they can withstand in LHC collisions [139]. The precision of the track position, with typically three hits in the pixel, is of the order of 10 μm in the $R-\phi$ plane and 115 μm in the longitudinal direction z (or R in the end-caps) [140].

Forming the middle part of the inner detector, there are four layers of silicon micro strips (SCT). The SCT module contains 15912 sensors $285 \pm 15\mu$m thick, using a single-sided p-in-n technology with AC-coupled readout strips. A charged particle typically leaves four hits in the SCT, reaching a resolution of 17μm per layer in the direction transverse to the strips and around 580 μm in $z(R)$ for the barrel (end-caps) [140].

The outer part of the inner detector system is the transition radiation tracker (TRT), made of layers of gaseous straw tubes interweaved with a transition radiation material - polyethylene and polypropylene for the barrel and end-caps, respectively. The 4 mm in diameter TRT straw tubes are filled with a gas mixture of 70% Xe, 27% CO_2 and 3% O_2 at 5-10 mbar over-pressure. The cathodes are operated at -1530 V to provide a gain of 2.5×10^4 [123]. The anodes are 31 μm in diameter tungsten wires plated with $0.5-0.7$ μm of gold and they are kept at ground potential. The straws forming the TRT barrel are 144 cm long, positioned parallel to the beam pipe. In the end-caps, the 37 cm long straws are radially distributed in wheels. The total number of readout channels of the TRT is around 351 000. A track typically leaves thirty hits in the TRT, with an accuracy of 130 μm in the $R-\phi$ plane per straw [140].

The modules of the three subsystems are aligned to better than 1 μm [141]. The transverse momentum resolution of a track detected by the inner tracker with the typical number of hits in each layer (three, four and thirty for the pixel, SCT and TRT, respectively) can be expressed as

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} \cong 0.05\%\ p_\mathrm{T} \oplus 1\%. \tag{2.7}$$

The total material budget in the inner detector is shown in Figure 2.7 in terms of X/X_0 as a function of $|\eta|$. The amount of material determines the photon conversion rate in front of the calorimeter, therefore precise knowledge of the material distribution is crucial for measurements involving photons.

2.2.1.4 Calorimeters

The ATLAS Calorimeter system was designed to measure the energy of electrons, photons and hadrons and it covers up to $|\eta| < 4.9$ using several different technologies. An illustration of the ATLAS Calorimeter system is shown in Figure 2.8. It consists of the electromagnetic calorimeter (ECAL), used for electron and photon energy reconstruction and identification and the hadronic calorimeter (HCAL), designed to measure hadrons (protons, neutrons, pions, etc.). Both ATLAS Calorimeters are of the sampling type, meaning they consist of layers of active and passive (absorbing) material, where only a fraction of the energy is actually measured by the active part, the rest is computed according to the material budget of the absorber and other additional material related to the readout and other services, depending on the corresponding detector part. Different technologies are used for the different parts of the calorimeter. The electromagnetic barrel, the hadronic end-cap and the forward calorimeter uses Liquid Argon ionization, while the hadronic barrel and extended barrel use scintillation.

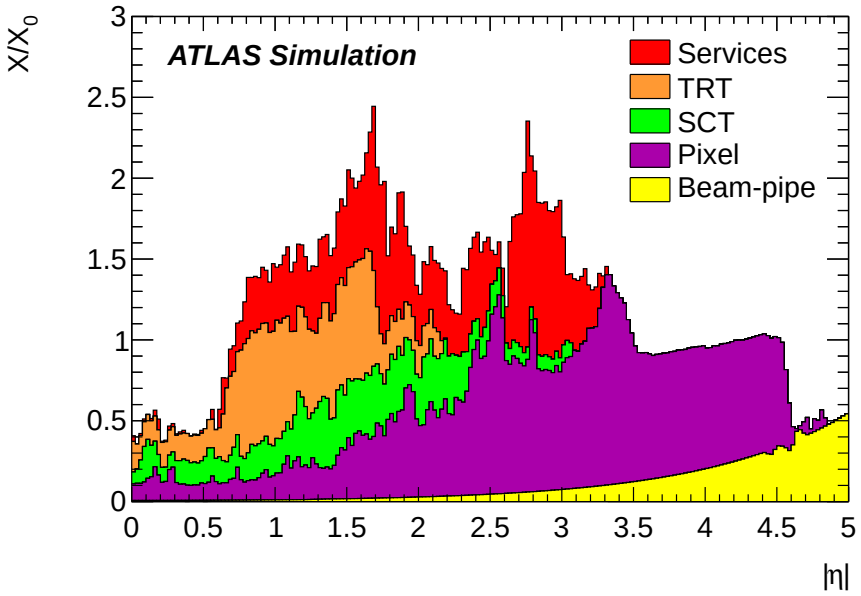


Figure 2.7: Amount of the material traversed by a particle in the inner detector, X/X_0, as a function of $|\eta|$, averaged over ϕ. The different colors show the contributions of each ID subdetector part [142].

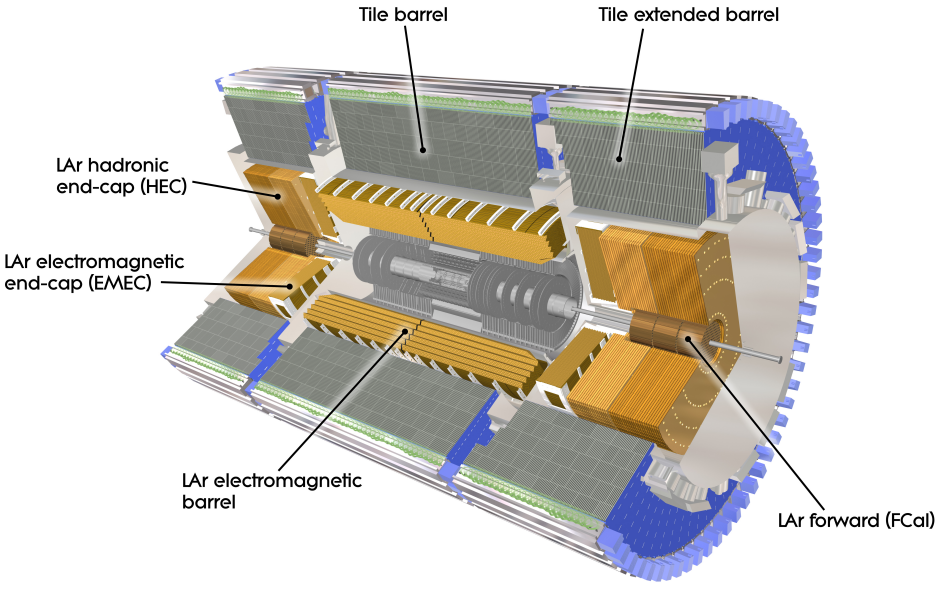


Figure 2.8: The ATLAS Calorimeter system composed of an electromagnetic calorimeter (Liquid Argon) and a hadronic calorimeter (scintillation tiles for the barrel part and Liquid Argon for the end-cap parts) [143].

2.2.1.4.1 Liquid Argon calorimeters The Liquid Argon calorimeters, using liquid Argon (at 87 K) as the active medium, are the electromagnetic barrel (EMB), electromagnetic end-caps (EMEC), hadronic endcaps (HEC) and forward calorimeters (FCAL) placed at high η inside the end-caps.

The EMB pseudorapidity coverage is up to $|\eta| < 1.475$ and the EMEC sections cover $1.375 < |\eta| < 3.2$ with a total of 22(24) X_0 in the barrel (endcap). The transition region between the two detector parts $1.37 < |\eta| < 1.52$[3] contains a lot of material corresponding to between 5 to 10 X_0 in front of the calorimeters, as seen in Figure 2.9, therefore the quality of energy reconstruction is reduced in this region. For photon precision measurements, this region is always excluded [142].

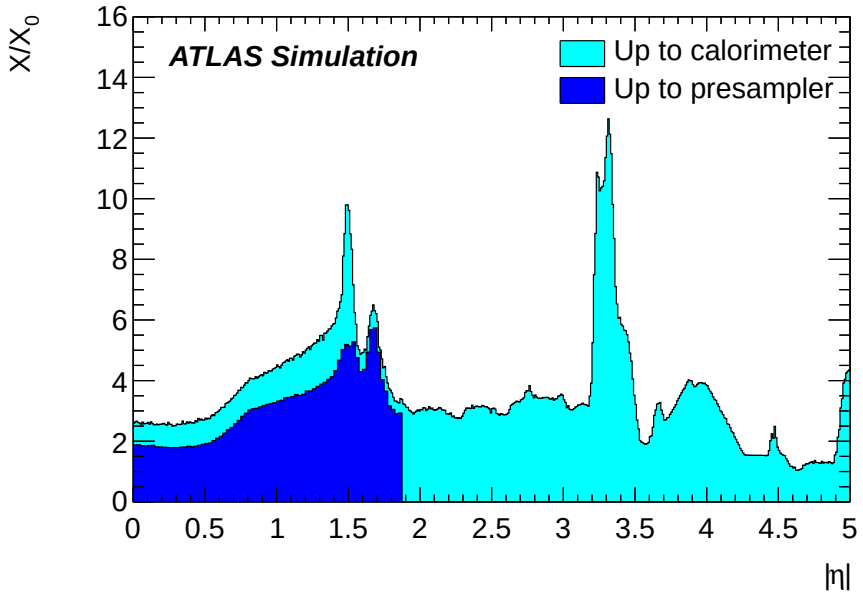


Figure 2.9: Amount of the material traversed by a particle, X/X_0, in the inner detector up to the presampler (light blue) and up to the calorimeter (dark blue) as a function of $|\eta|$, averaged over ϕ [142].

The EMB is split into two wheels in the middle of the detector, each of the wheels is 3.2 m long, has an outer radius of 4 m, an inner radius of 2.8 m and is enclosed in an aluminium cryostat which is vacuum insulated. The EMEC itself consists of two wheels, each in a separate cryostat and each 63.2 cm thick with a 207.7 cm outer radius. Each wheel consists of eight modules. Both the EMB and the EMEC are built with an accordion geometry, see Figure 2.10, that guarantees a continual coverage in ϕ, with readout services mounted on the inside and on the outside of the calorimeter cylinder in the barrel and on the outer base of the wheel of the end-cap.

The EMB and the EMEC calorimeters are built with lead absorbers interlaced with copper etched on kapton electrodes, separated by a 2.1 mm honeycomb structure, thereby defining the size of the gap. A high voltage of 2000 V is applied between the electrodes and the grounded

[3]This transition region definition, as $1.37 < |\eta| < 1.52$, was used for analyses performed on 2011 data at $\sqrt{s} = 7$ TeV. The definition is also used for all performance studies, but physics analyses exclude also the region of $1.52 \leq |\eta| < 1.56$ due to the additional material presence and therefore reduced performance.

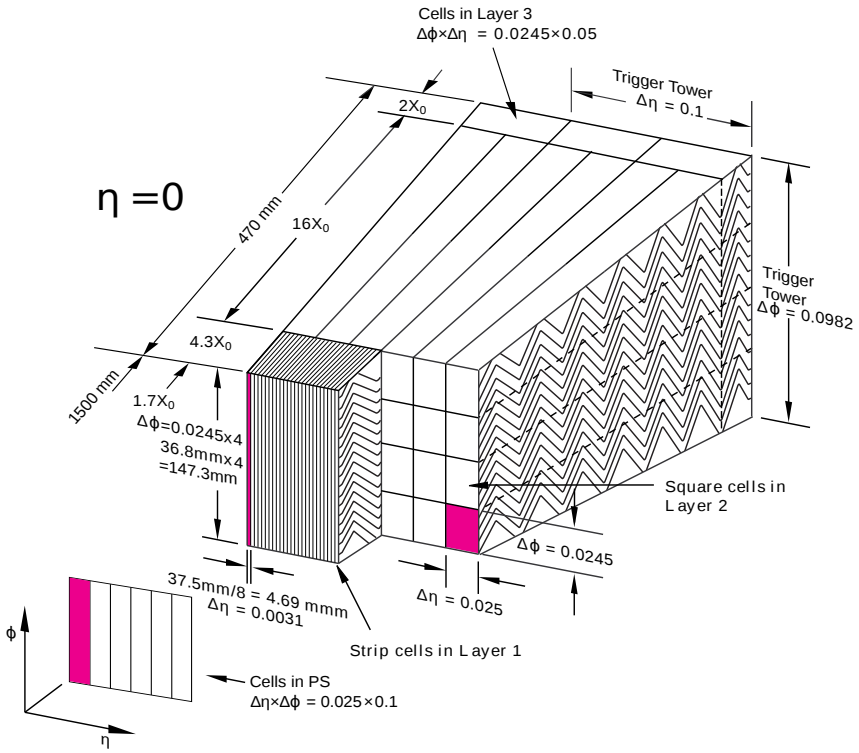


Figure 2.10: Sketch of an EMB module, longitudinal layers visible [144].

absorbers, which causes a drift of ionization charges induced by a passing particle, taking 450 ns for them to reach the electrodes. The high voltage is constant in the barrel but in the EMEC, the gap width varies with η and so does the high voltage (typically in steps of 0.2 in η) [145].

For $|\eta| < 2.5$, the EMB and EMEC are longitudinally segmented into three layers. This allows for a measurement of the longitudinal EM shower development. The first layer (L1), as seen in Figure 2.10, is finely segmented in η, in the barrel it is $0.025/8 \times 0.1$ in $\Delta\eta \times \Delta\phi$ for $|\eta| < 1.4$. Its thickness corresponds to about 4.4 X_0 at $\eta \approx 0$. Cells in the second layer (L2) are eight times wider in η and four times narrower in ϕ, namely 0.025×0.025 in $\Delta\eta \times \Delta\phi$ and have a thickness of approximately 17 X_0. The third longitudinal layer (L3) of the barrel calorimeter has a granularity of 0.050×0.025 in $\Delta\eta \times \Delta\phi$ for $|\eta| < 1.35$ and covers about two radiation lengths. The end-cap segmentation is more η-dependent. Most of the energy of an EM shower is deposited in the first two layers of the EM calorimeter. The first layer, thanks to its fine segmentation, can help distinguish single photon showers from two neighboring showers of photons coming from the decay of neutral hadrons in jets, such as a π^0 decaying into a photon pair.

In front of the first calorimeter layer, there is another layer, called the presampler (PS) reaching up to $|\eta| < 1.8$, which is used to correct the energy loss upstream of the calorimeter. It consists of an active LAr layer, 1.1 cm (0.5 cm) thick in front of the barrel (end-caps) and has a granularity of 0.025×0.01 in $\Delta\eta \times \Delta\phi$.

The hadronic end-cap wheels (HEC) of outer radius 2.03 m, covering $1.5 < |\eta| < 3.2$, are made of 32 modules in ϕ and two layers in depth, composed of copper and liquid argon flat plates.

The forward calorimeters (FCAL) [146] on both sides of the ATLAS system, are cylindrical sam-

pling calorimeters each consisting of three wheels.The first one, the electromagnetic calorimeter FCAL1, is the closest to the interaction point out of the three modules at 4.7 m. FCAL1 has a thickness of 27.6 X_0 measuring 444 mm in depth with an outer radius of 449 mm and it is made of tubes composed of copper electrodes with a liquid argon gap, wrapped with the PEEK[4] monofilament insulating fiber surrounded by copper absorber matrix build of separate plates stacked one behind the other. Altogether, there are 12 260 electrodes. The other two modules of the FCAL are optimized for high absorption lengths (each module has $X_0 \approx 90$), therefore tungsten was used as absorber and they are considered as hadronic calorimeters. Behind the two hadronic FCAL modules, a shielding plug made of copper alloy is positioned in order to reduce the backgrounds in the muon end-cap detector system.

The readout, collection and interpretation of the signal from the LAr calorimeters will be described in Section 2.3

2.2.1.4.2 Tile calorimeter is a sampling calorimeter measuring 5.8 m in length and 8.5 m in outer diameter. It uses steel as an absorber and a plastic scintillator as active medium. It was designed to measure the energy of hadrons and it is comprised of two barrel wheels (the barrel is split in the middle) and an extended barrel wheel at each side. The barrel part covers $|\eta| < 1$ while the extended barrels cover the region of $0.8 < |\eta| < 1.7$ [147]. The transition between the barrel and the extended barrel contains the Intermediate Tile Calorimeter (ITC), which contains more active material in order to maximize the overall calorimeter coverage. Each barrel has 64 modules of size $\Delta\phi \approx 0.1$ and is segmented longitudinally into three layers of tiles providing a segmentation of 0.1×0.1 in $\Delta\eta \times \Delta\phi$ in the first two layers and $\Delta\eta \times \Delta\phi = 0.1 \times 0.2$ in the third layer. Signal from the tile calorimeter is read-out using optical fibers lead along the smaller sides of the block scintillating tiles.

2.2.1.5 The muon system

The ATLAS muon spectrometer [148] was designed to trigger, identify and precisely measure the momentum of muons. It is located on the outer part of the ATLAS detector system, within the toroidal magnet system, which bends the trajectories of muons and helps with charge identification and momentum determination. The precision measurement of the muon tracks is done by the Monitored Drift Tube chambers (MDT), which cover the pseudorapidty range of $|\eta| < 2.7$ and consist of three to eight layers of drift tubes. The MDTs form three tracking layers both in the barrel part and the muon end-caps. In the forward region ($2.0 < |\eta| < 2.7$), the innermost tracking layer uses Cathode-Strip Chambers (CSC[5]), due to their higher rate capability and time resolution. To provide precision tracking, the muon system also contains fast trigger chambers. These are the barrel resistive plate chambers (RPC), providing coverage in $|\eta| < 1.05$ and end-cap thin gap chambers (TGC), stretching over $1.05 < |\eta| < 2.7$, positioned perpendicular to the beam-pipe.

2.2.1.6 Trigger and data acquisition system

The amount of collisions that the LHC delivers to ATLAS is much larger then what the detector with its electronics is capable of processing and storing. The trigger system, reducing the

[4] PolyEtherEtherKetone

[5] CSCs are multi-wire proportional chambers with cathode planes segmented into strips in orthogonal directions.

nominal event rate from 1 GHz down to 400 Hz, plays a key role in selecting events which are interesting from the physics point of view out of all the events being produced. At the nominal beam bunch-spacing of 25 ns, corresponding to a 40 MHz bunch-crossing frequency, with an average 25 collisions per crossing, the LHC delivers one billion events per second. For the Run 1 data-taking period, the delivered events were only half a billion per second, due to the 50 ns beam bunch-spacing. The trigger system consists of three levels of event selection.

The Level-1 (L1) trigger is a hardware-based selection system, which processes event information at the full beam crossing rate of 40 MHz, reducing it down to 100 kHz. It uses calorimeter information with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ and the muon spectrometer RPCs and TGCs and it identifies the so-called Regions of Interest (RoI), where basic calorimeter clustering or muon tracks, corresponding to a high transverse momentum object, were detected. The capacity time of the analogic pipeline to store data until the L1 decision is made, or latency, is 2.5 μs. After a L1 decision, there is a minimum dead time of five bunch crossings (125 ns), or 1.25% at an output rate of 100 kHz. The output information is sent to the next trigger level as a RoI in η and ϕ, where an interesting physics event was spotted.

The Level-2 (L2) is based on the information from the L1 trigger and provides an output rate of about 2 kHz. It is software-based and uses additional information from the inner detector on top of the calorimeter and muon information. Fast algorithms reconstruct leptons, photons and jets within the RoIs. The L2 initiates the readout of data from the entire detector.

Finally, the third trigger level, called the Event Filter (EF[6]), completes the selection leading to a final frequency of up to 400 Hz by using more sophisticated reconstruction algorithms, used also in offline physics analysis.

Chain point	Run 1	Nominal
Beam bunch spacing	50 ns	25 ns
Bunch crossing rate	20 MHz	40 MHz
Collision rate	0.5 GHz	1 GHz
L1 output rate	60 kHz	100 kHz
L2 output rate	2 kHz	up to 3.5 kHz
EF output rate	200 Hz	up to 400 Hz
Data output	450 MBs	up to 450 MBs

Table 2.2: Nominal and actual Run 1 event, trigger and detector output rates.

In Table 2.2, the nominal and Run 1 rates are listed for comparison. Triggers are typically organized in trigger chains, following a naming convention:

`<Trigger level>_<stream><momentum threshold>_<additional specifications>.`

For example, the trigger used for the Higgs boson searches in its decay into photon pairs was a diphoton trigger with loose shower shape criteria (for definition see Section 3.3) on the two photon candidates: `EF_g35_loose_g25_loose`. A schema of the trigger system is shown in Figure 2.11. The trigger selection will for each analysis presented in this thesis will be discussed later (Section 6.2.3, 7.2).

Due to the limited amount of data that can be processed and stored, some triggers are prescaled, meaning only a fraction of events triggering a certain configuration is accepted. For example,

[6]The L2 and the EF are together called the High Level Trigger (HLT)

in the 2012 data taking, the lowest unprescaled single photon trigger was `EF_g_120_loose`, requiring the photon to have $E_{\mathrm{T}}^{\gamma} > 120$ GeV. Triggers requiring a lower E_{T}^{γ} only accepted a pre-set fraction of such events. In an analysis, the number of events is normalized accordingly.

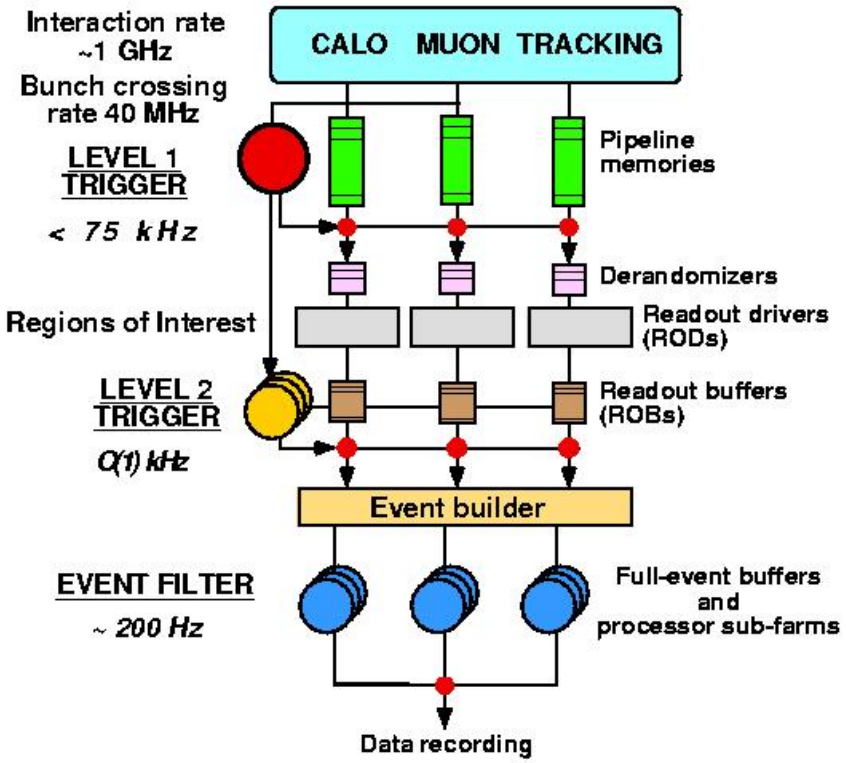


Figure 2.11: Trigger and data acquisition system in ATLAS [139].

The data acquisition is controlled by the Trigger and Data Acquisition system (TDAQ) and the Detector Control System (DCS). The TDAQ takes care of the detector configuration and the trigger system, aiming at the efficient transmission of event data from and to the trigger systems and to the acquisition of full events and their permanent storage. The DCS configures the ATLAS subdetectors in real time and identifies malfunctions, alerting the detector operation personnel - shift crew on any problems. A Run Control (RC) system controls the activation or deactivation of the data acquisition of each detector part. When the LHC does not have stable collisions, calibration runs can be taken, various detector parts have different calibrations that can be launched individually when needed.

The data in ATLAS is taken when LHC delivers stable beams and stable collisions. The typical unit of data is measured in luminosity blocks (later referred to as *lumi blocks*). In 2011 and 2012 data taking, a *lumi block* was 60 seconds long. *Lumi blocks* are then merged into runs with corresponding run numbers. A system of online and offline data quality checks then helps determine which runs and which *lumi blocks* within those runs are suitable for physics analysis. Events containing noise bursts in the LAr calorimeter cells, data integrity errors, HV trips, instabilities or HV ramping during a run are not considered in an analysis. Also events containing beam background halo (particles radiated off the beam), masked regions or cells of various detector parts and missing LAr readout are discarded for physics analyses.

A so called Good Run List (GRL) is produced, containing the run and *lumi block* numbers during which the detector operated without problems. The GRL can be evaluated for the whole ATLAS system or a specific GRL can be produced just for its subsystems. For example an analysis studying photons uses data collected by the calorimeters and the inner detector (to detect tracks and discriminate between electrons and photons that converted into e^+e^- pairs in the material before the calorimeter) does not need the muon system to be working if it does not use any information about muons. Physics analyses, however, usually use an *All good* GRL, where all ATLAS subsystems were fully operational. 95.5% of data collected at $\sqrt{s} = 8$ TeV passes the *All good* GRL [149].

2.3 The Liquid Argon Calorimeter

The Liquid Argon calorimeters altogether have 182 468 channels that need to be read out, providing a fine segmentation and excellent spatial resolution. This section gives the description of the signal detection, collection and read-out principle in the LAr calorimeters in more detail, because of its relevance for the physics analyses presented in this thesis.

2.3.1 Signal detection and readout

The liquid Argon calorimeters work on the principle of the detection of ionization charges produced by the charged particles in an EM shower induced by a particle depositing its energy in the calorimeter. Figure 2.12 shows a section of a module of the EMB LAr calorimeter, from which it is possible to illustrate the signal detection. When a charged particle of the EM shower

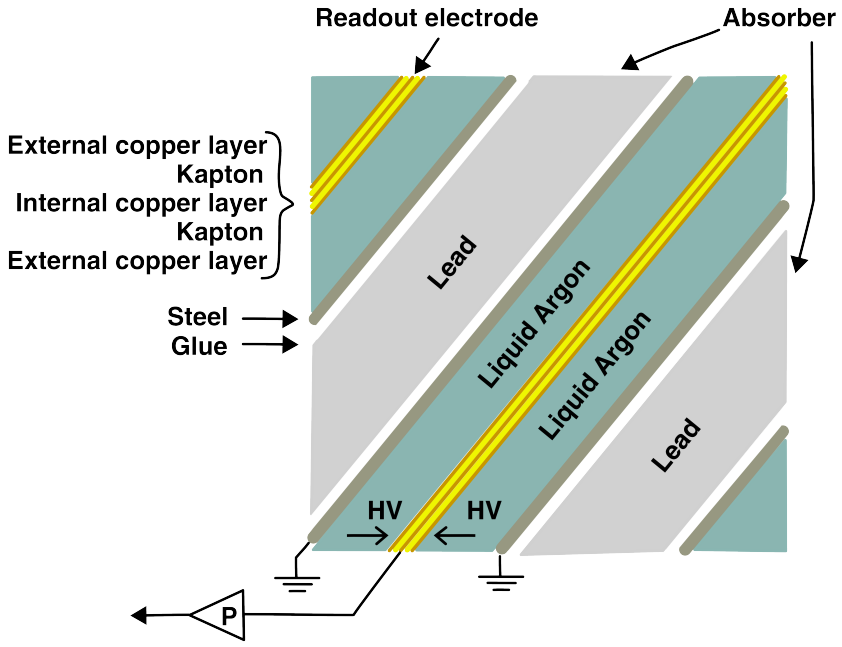


Figure 2.12: Schematic representation of a part of a LAr module in the barrel LAr calorimeter.

enters a calorimeter cell, it ionizes the LAr medium, generating charges (electrons and ions). Due to the high voltage applied between the electrodes and the grounding, the produced charges drift. The electrons (being lighter than the ions, electrons are used for charge collection) drift towards the cathode and accumulate, forming an electrical signal of a triangular shape. The drift time τ_d of the electrons, and therefore the duration of the pulse, is determined by the drift velocity in the medium and the width of the gap between the electrodes and the grounding, and it is ≈ 450 ns for the EMB calorimeter [150].

A scheme of the full of LAr read-out electronics chain is shown in Figure 2.13. The data acquisition chain can be divided into two sub-systems: the front-end part, situated directly on the detector cryostat, where the preamplification of signal is done; and the back-end part, situated in the neighboring ATLAS services cavern (USA 15). This part stores remote electronic responsible for signal processing, signal to energy conversion and calibration. It is accessible for intervention even during data-taking periods.

The signal processing can be summarized as follows [151]. The analog signal is routed from the electrodes to summing boards located on the front and back of the modules of the calorimeter for the first layer and for the second and third layer, respectively. The summing boards sum the signals in categories of azimuthal coverage, grouped into mother boards on the back of the cryostat. The signal is transmitted out of the cryostat via feedthroughs.

2.3.1.1 Front-end electronics

Entering the Frond End Boards (FEB), the signal is subject to several stages of analog processing, from preamplification to the digital conversion.

In the preamplifier, the raw signal is amplified and then split and further amplified by shaper chips, producing three linear gain scales with gain ratios of about ten (low ≈ 1, medium ≈ 10, and high gain ≈ 100). Each signal is shaped by a bipolar CR-$(\mathrm{RC})^2$ filter, changing the shape of the signal into a pulse, as illustrated in Figure 2.14(a). The bipolar shape, on average, minimizes the *out of time pile-up*, except for the beginning of the bunch train when the compensation has not had time to take place yet, since the LAr pulse integration takes ≈ 450 ns (ionization electron drift time in liquid Argon).

After passing the preamplifier, the analog signals are partially summed by the FEBs, before being sent to the Tower Builder Boards (TBB), where energies of four neighboring cells are summed and sent to the L1 trigger system for evaluation. The L1 trigger returns a decision in 2.8 μs.

The pulse is stored by switched capacitor array (SCA) analog pipeline chips, waiting for the trigger decision to arrive. Events accepted by the L1 trigger are read-out from the SCA at the LHC beam bunch crossing frequency of 40 MHz (every 25 ns) in five samples per channel, using an optimal gain scale, determined by the size of the given signal[7]. The samples are then digitized at a 5 MHz rate using a 12-bit Analog-to-Digital Converter (ADC). The digitized data are transmitted into the back-end electronics in the service cavern through optical fibers.

The time constant of the FEB shapers is chosen to minimize the overall noise level, which comes from the electronics noise and pile-up noise. The evolution of these two contributions is shown

[7]The gain is determined by the Gain Selector chips (GSEL) based on the value of the peak sample in medium gain compared to two reference thresholds, in order to keep the lowest non-saturated gain and use the minimum ADC→MeV conversion factor.

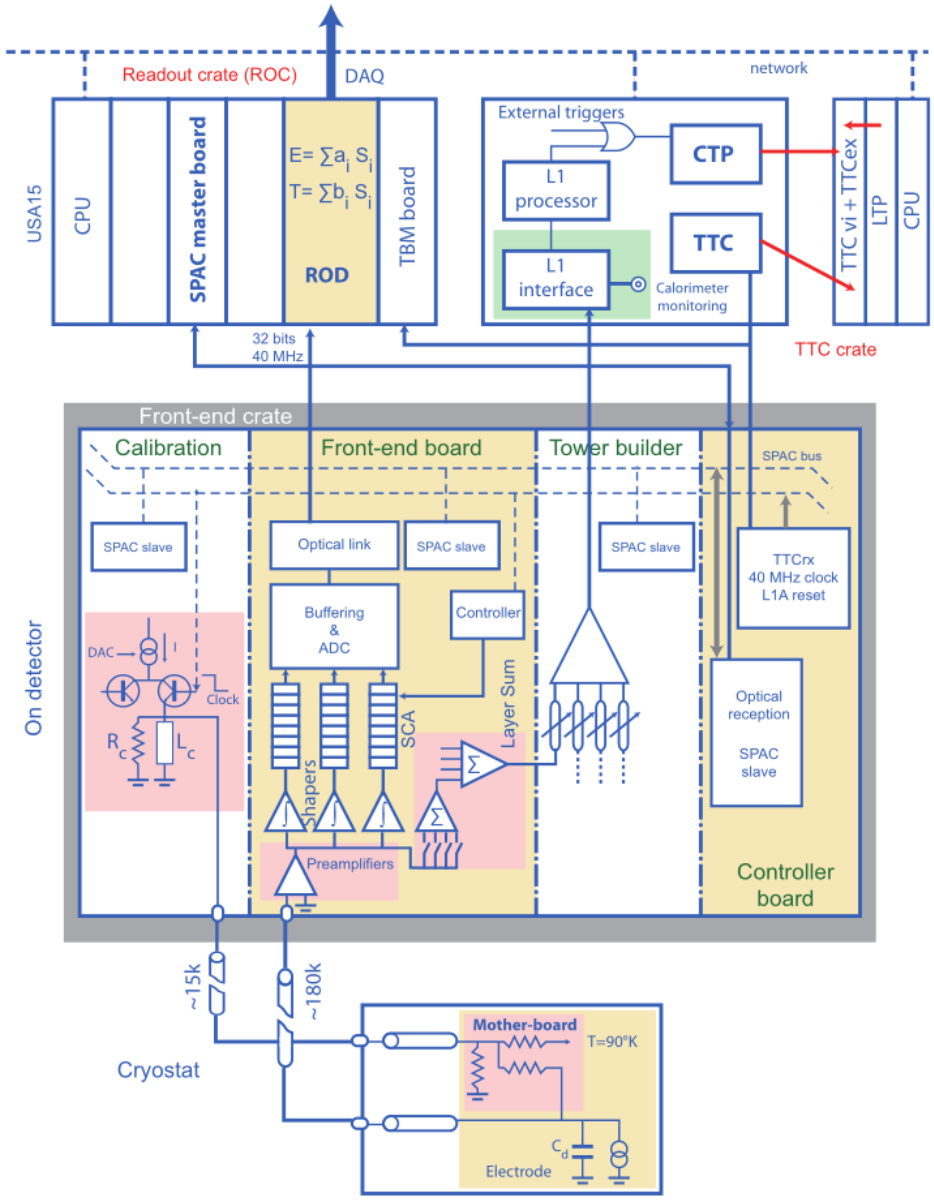


Figure 2.13: A diagram of the architecture of the electronics processing the signal from the LAr calorimeters [151].

as a function of the shaping time in Figure 2.14(b). The pile-up noise is luminosity dependent, and two configurations are shown. The fixed shaping time of 50 ns is not optimal for all possible configurations of pile-up and luminosity. Further optimization of the noise is performed by the back-end electronics, using the Optimal Filtering (OF) technique [153], which will be described in Section 2.3.2.

A FEB treats 128 channels from the detector. There are altogether 1524 FEBs mounted on the cryostat. Signal from eight FEBs is transmitted into a single Read Out Driver (ROD). There

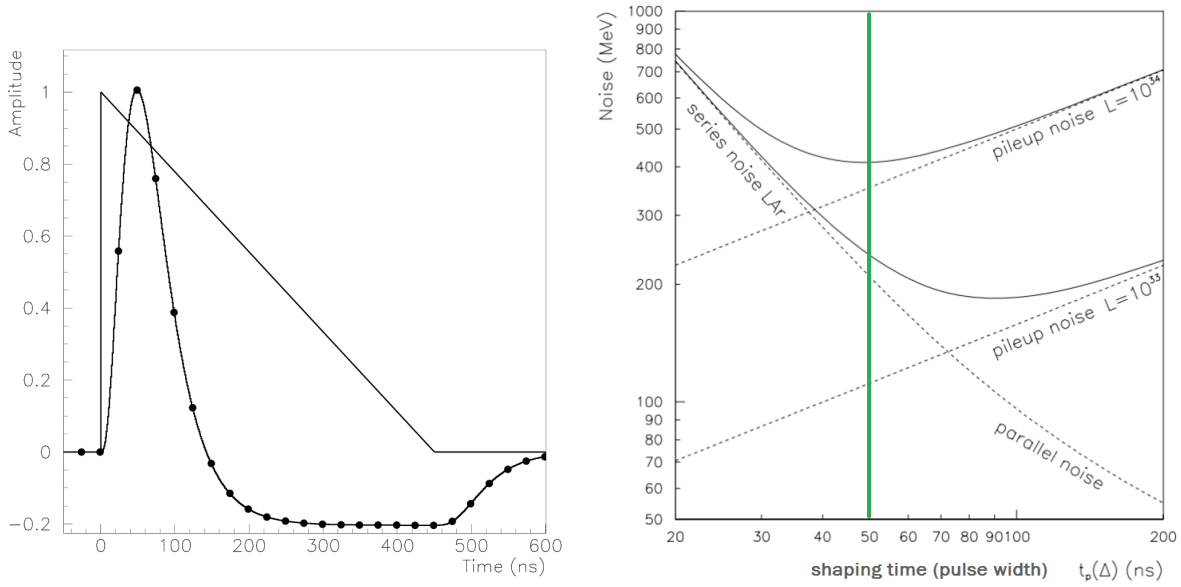


Figure 2.14: LAr pulse before and after shaping (left) and the LAr shaping time vs. the electronic and pile-up noise (right)[152].

are 192 ROD mother boards, combined into 16 ROD crates. The transfer between the FEBs and the RODs is facilitated by a 1.6 Gbit/s optical cable per FEB. This gives a fixed volume of data that the detector can give as output for further processing to the back-end electronics, therefore potentially limiting the acceptance of data by the detector in case of an increase of the trigger rate. For Run 2 of the LHC, the trigger rate is increased and the amount of LAr samples used to describe the pulse shape and the impact on the LAr calorimeter noise was studied. The results are presented in Chapter 8.

2.3.1.2 Back-end electronics

The back-end electronic system is composed of three subsystems: the RODs, Timing and Trigger Control (TTC) and the L1 trigger receiver. In physics data taking mode, the system reads the front-end electronics, receives the TTC signals, processes data and sends it further to the data acquisition system at the rate of the L1 trigger up to 90 kHz (for five samples). During Run 1 the trigger operated at a 75 kHz output rate.

A ROD mother board consists of three zones [152]. First, the eight optical receivers (ORx) receive the signal from the FEBs and transmit it to the four Field-Programmable Gateway Array (FPGA) receivers in 16 bits at 80 MHz frequency moving the signal further at the same frequency, to four Processing Units (PU). A processing unit contains two 720 MHz Digital Signal Processors (DSP) and it is responsible for online calculation of the energy E and timing τ of the LAr pulse using the Optimal Filtering (OF) technique [153]. The two DSP 16 bit output data are stored in two emission/reception First-In-First-Out's (FIFO), read out by an output controller (OC) at 80 MHz. The OCs give an output of 32-bit wide data at 40 MHz. The output data is then sent through serial links into a transition module to Read Out Buffers (ROBs) hosted on

PCs. The acquisition changes from the push-mode driven by the L1 trigger into a pull mode driven by the EF-farm, where event selection and building is performed and data is transmitted to the ATLAS data acquisition system and recorded [152].

2.3.2 Optimal filtering method

The Optimal Filtering method [153] enables the computation of the amplitude of the signal pulse and the time of the detection while minimizing the noise contribution and effects from the electronics noise. The ionization signal detected by the LAr calorimeters, amplified, shaped and sampled by the FEBs, enters the RODs and the five samples are used to compute the energy of a cell. Reading out multiple samples provides a better description of the signal shape, and 32 samples are normally used in calibration runs, precisely measuring the pulse shape, timing, amplitude and determining the noise. The amount of data that needs to be transferred and processed is, however, limited and a compromise of using five samples is reached.

The LAr signal pulse is characterized by its amplitude A, proportional to the deposited energy, and the time of arrival τ relative to the LHC clock. Using five samples, the amplitude and time is computed while minimizing the noise. Defining coefficients a and b (Optimal Filtering coefficients - OFCs) and linear sums $u = A$ (amplitude) and $v = A\tau$ (timing) of signal samples

$$u = \sum_{i}^{N_{samples}} a_i s_i \quad \text{and} \quad v = \sum_{i}^{N_{samples}} b_i s_i \tag{2.8}$$

The sample s at position i and at time t_i can be expressed as

$$s_i - p = A \cdot g(t_i - \tau) + n(t_i), \tag{2.9}$$

where g is the functional form of the ionization signal, p is the pedestal (from calibration runs) and n is the noise. The derivative of the signal functional form g' and the fact that τ is very small (a few ns) then gives

$$s_i - p = A \cdot g(t_i) - A\tau g'(t_i) + n(t_i) = Ag_i - A\tau g'_i + n_i \tag{2.10}$$

The assumption of a good knowledge of the pulse shape and its derivative and calibration run data [150], when a known current is injected and the detector response is measured, allow for determination of the amplitude A and time information $A\tau$ as

$$A = \langle u \rangle = \sum_{i}^{N_{samples}} \left(Aa_i g_i - A\tau a_i g'_i + \langle n_i \rangle \right) \tag{2.11}$$

and

$$A\tau = \langle v \rangle = \sum_{i}^{N_{samples}} \left(Ab_i g_i - A\tau b_i g'_i + \langle n_i \rangle \right) \tag{2.12}$$

The noise is zero on average and the OFCs have to satisfy the constraints

$$\sum_{i}^{N_{samples}} a_i g_i = 1 \qquad \sum_{i}^{N_{samples}} a_i g'_i = 0 \qquad \sum_{i}^{N_{samples}} b_i g_i = 0 \qquad \sum_{i}^{N_{samples}} b_i g'_i = -1. \tag{2.13}$$

The variances of u and v can be expressed as

$$\sigma_u = \sum_{ij} a_i a_j R_{ij} \qquad \sigma_v = \sum_{ij} b_i b_j R_{ij}, \tag{2.14}$$

where $R_{ij} = \langle n_i n_j \rangle$ is the total noise autocorrelation matrix, taking into account the electronics and pile-up noise.

The variances σ_u and σ_v are minimized, while satisfying the constraints of Eq. 2.13, using Lagrange multipliers. The functions to be minimized are

$$I_u = \sum_{ij} R_{ij} a_i a_j - \lambda \left(\sum_i a_i g_i - 1 \right) - \kappa \sum_i a_i g'_i, \tag{2.15}$$

and

$$I_v = \sum_{ij} R_{ij} b_i b_j - \mu \sum_i b_i g_i - \rho \left(\sum_i b_i g'_i + 1 \right). \tag{2.16}$$

By setting partial derivatives of I to zero, one obtains a set of linear equations that can be expressed in matrix form as

$$a = \lambda R^{-1} g + \kappa R^{-1} g' \qquad b = \mu R^{-1} g + \rho R^{-1} g'. \tag{2.17}$$

The Lagrange multipliers λ, κ, μ and ρ can be determined by applying the constraints from Eq. 2.13, therefore computing the OFCs. The amplitude and time can be obtained subsequently as

$$A = \sum_{i=0}^{N_{samples}-1} a_i (s_i - p) \qquad A\tau = \sum_{i=0}^{N_{samples}-1} b_i (s_i - p). \tag{2.18}$$

The energy of a LAr calorimeter cell can be computed from the amplitude A as

$$E = G_{ADC \to DAC} \times f_{DAC \to \mu A} \times f_{\mu A \to MeV} \times \left(\frac{M_{phys}}{M_{calib}} \right)^{-1} \sum_i a_i (s_i - p), \tag{2.19}$$

where

- $G_{ADC \to DAC}$ is a polynomial of first order $G(A) = R_0 + R_1 \times A$ and it describes the conversion of the analog signal into digital signal, using the optimal gain;
- $f_{DAC \to \mu A}$ converts the DAC to the injected current in the calibration;
- $f_{\mu A \to MeV}$ converts the ionization current on the calibration board into deposited energy in MeV, the scale being determined from test-beam studies;
- M_{phys}/M_{calib} evaluates the amplitude difference between the calibration pulse and the real physical pulse measured in the detector while data-taking [154].
- p is the ADC pedestal.

During Run 1 data-taking, when the beam bunch spacing was 50 ns instead of the nominal 25 ns, the standard OFC computation described above introduced a shift in the reconstructed energy at large $|\eta|$ values. This effect was interpreted as originating from the impact of the larger quantization of the pile-up contribution with 50 ns beam bunch spacing compared to 25 ns [155]. An extra constraint was added to the OFC computation in the regions of the calorimeter, where the shift was observed, for the 2012 data. This method will be described more in a context of calorimeter noise studies for Run 2 of the LHC in Chapter 8.

3 Photons in ATLAS

The electromagnetic calorimeter was designed so that the energy of the electrons and photons it is supposed to detect, would be, for the most part, deposited in its cells. Both of these particles, upon entering the calorimeter, produce an electromagnetic shower.

When electrons traverse the inner detector, they leave a track bent by the magnetic field of the solenoid. Photons do not interact in the tracker, but when passing through its material, there is a certain probability for them to convert into an electron-positron pair. Roughly 10-60% of photons convert, but this number strongly depends on their η position due to the different material budget of the inner detector, as shown in Figure 3.1. The conversion may happen anywhere before the calorimeter and a conversion vertex may or may not be reconstructed. The reconstruction combines data from the calorimeter layers (including the presampler) and the inner detector, which helps distinguishing between electrons, converted and unconverted photons.

The methods used for the reconstruction, identification and calibration of photons are summarized in the following.

3.1 Reconstruction of photons in ATLAS

When electrons and photons enter the EM calorimeter, they produce an EM shower and deposit their energy into multiple neighboring cells. The reconstruction of an EM object therefore begins with clustering of cells. The EM calorimeter is divided into towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. Inside a tower, the energy of all cells in all calorimeter layers is summed, giving the energy of the tower. Clusters are seeded by towers with transverse energy $E_{\mathrm{T}} > 2.5$ GeV, using a sliding-window algorithm [156], with a window size of 3×5 towers.

If a track detected by the ID with a $p_{\mathrm{T}} > 0.5$ GeV is associated to the cluster, the cluster is reconstructed as an electron. The track has to come from a vertex located in the beam interaction region and if extrapolated to the calorimeter middle layer, the track has to be close in ΔR to the cluster longitudinal barycenter.

If no track is matched to the cluster, it is reconstructed as an unconverted photon. In case a conversion vertex [157] with a radius smaller than 800 mm is associated to the cluster, it is reconstructed as a converted photon. Depending of whether one or two tracks were detected,

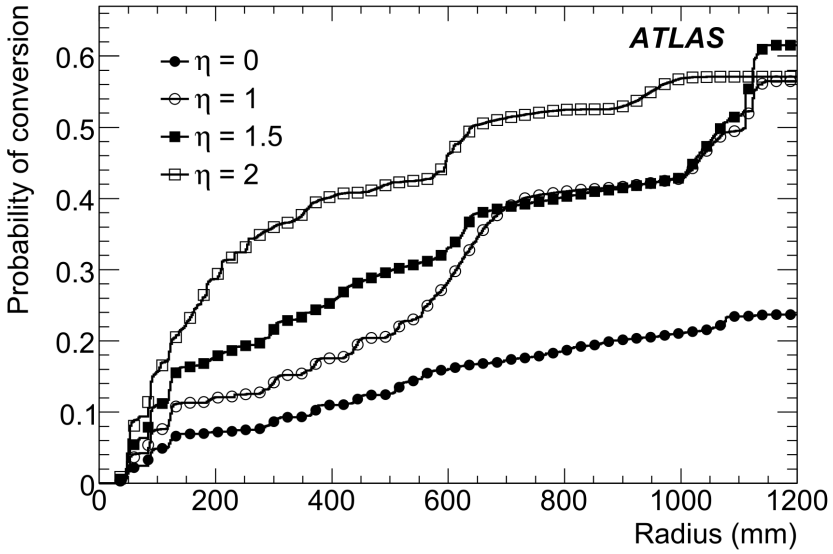


Figure 3.1: Probability for a photon to convert as a function of the radius for different values of $|\eta|$ in the calorimeter [140].

the converted photon is categorized as a one-track or a two-track conversion.

The size of the final cluster depends on the type of the particle and the detector part. The clusters of unconverted photon candidates in the barrel are built with a 3×5 cluster size, while for converted photons, a 3×7 cluster size is used, since they tend to have broader showers. Electrons in the barrel are reconstructed using a 3×7 cluster size. For the end-cap parts of the EM calorimeter, a cluster size of 5×5 is used for all electron and photon candidates. The efficiency of the electron reconstruction ranges from 97-99% as a function of η and p_{T} [158]. For photons with $p_{\mathrm{T}} > 25$ GeV, the expected reconstruction efficiency is around 98%, on average 99% for unconverted and 95% for converted photons [159].

Before moving on to the calibration of the EM cluster energy, a correction to η and ϕ positions of the photon candidate is applied. The position is measured independently for each calorimeter layer, then weighted according to the position of the cells in the cluster and corrected for systematic effects known from MC and test beam studies [160]. The finite granularity of the calorimeter cells causes a bias in the determination of the η position. This bias emerges in a so-called S-shape functional form. To correct for this fact, the calorimeter is divided in η and an unbinned fit using and empirical function is performed to simulated data. A small bias is introduced also in the ϕ direction due to the inter-module widening caused by gravity, and it is corrected for using L2 of the calorimeter, since it has the best granularity in ϕ. The position measurement after the η and ϕ correction is combined using L2 for ϕ and both L1 and L2 for η, weighting the finely segmented L1 three times more than L2 for the η position.

3.2 Photon calibration

After building the EM cluster and correcting its position, the energy of the photon candidate can be calibrated. The original electron and photon calibration used from 2009 to 2013 in ATLAS was the so-called *calibration hits* method[1]. This calibration is based on MC simulation studies and test beam studies [161] of energy deposits in the EM calorimeter layers and the evaluation of the material budget in front of the calorimeter.

3.2.1 The *calibration hits* method

The calibrated energy of a photon is parameterized as

$$E^{\gamma} = a\left(E_{tot}^{calo},\eta\right) + b\left(E_{tot}^{calo},\eta\right)\cdot E_{PS} + c\left(E_{tot}^{calo},\eta\right)\cdot (E_{PS})^2 + \frac{s_{cl}(X,\eta)}{f_{out}(X,\eta)}\left(\sum_{i=1}^{3} E_i\right)\times(1+f_{leak}(X,\eta))\times(F(\eta,\phi)), \tag{3.1}$$

where E^{γ} is the reconstructed photon energy; E_{tot}^{calo} is the total energy reconstructed in the calorimeter for the given candidate; the parameters a, b and c and their respective terms account for the energy loss in front of the calorimeter and are obtained from MC [161]. The c parameter is zero except for the $1.55 < |\eta| < 1.8$ region. The E_{PS} is the energy deposited in the presampler, defined as the energy deposited in the active LAr medium divided by the effective sampling fraction fixed to 0.05 in the barrel and 1/60 in the end-cap regions. The next term comprises the sum of the energy collected in the three layers of the accordion calorimeter corrected by $s_{cl}(X,\eta)$, a correction factor for the η and X variations in the accordion sampling fraction and f_{out} is a applied to account for lateral leakage of the shower outside the reconstructed cluster. The X is the longitudinal barycenter of the shower (shower depth) defined as

$$X = \frac{\sum_{i=0}^{3} E_i X_i}{\sum_{i=0}^{3} E_i}. \tag{3.2}$$

The $f_{leak}(X,\eta)$ coefficient accounts for longitudinal leakage correction meaning the energy deposited by the shower behind the EM calorimeter. $F(\eta,\phi)$ is the energy modulation correction due to the impact point inside a cell.

After the energy correction, the in-situ inter-calibration in η is performed, setting the energy scale. A data-driven method is applied, using electrons from $Z \to ee$ decays, which allow a comparison of the Z boson mass shape with the simulation. η-dependent coefficients $\alpha_i(\eta)$ are defined and the reconstructed energy is corrected w.r.t. to the true energy. This is described in detail in Section 5.1. In addition, validation is performed using events with $J/\psi \to ee$ decays. The relative energy resolution for 50 GeV photons is between 1.5 and 2% for $|\eta| < 0.7$ and slightly worse in the end-caps. Converted photons perform worse w.r.t the unconverted photons due to the energy loss by bremsstrahlung.

The energy resolution for photon candidates can be parametrized as

$$\frac{\sigma}{E} = \frac{A}{\sqrt{E}} \oplus \frac{B}{E} \oplus C, \tag{3.3}$$

[1] Also referred to as the *old calibration*.

where $A \approx 10\%$ is the sampling term known from test beam studies, which gets larger with η due to the presence of more material in front of the calorimeter, $B \approx 350 \times \cosh\eta$ MeV^2 is the noise term, also increasing with η, where it is dominated by pile-up and C is the constant term which has a design value of 0.7%, dominating at high energies. The measured values of the constant term range from $(1.2 \pm 0.1(\text{stat}) \pm 0.3(\text{syst}))$ in the EMB for $|\eta| < 1.37$ and increases up to 1.8% in the endcaps and 3% in the FCAL, as measured in 2010 data [161].

3.3 Photon identification in ATLAS

In order to separate prompt photons produced in hard collisions from background photons originating from neutral hadrons such as π^0 decays inside jets, a procedure of photon identification was developed. A set of shower shape variables is used to evaluate the photon quality and three cut-based selections: *loose*, *medium* and *tight* working points are established. The discriminating variables used for these selections are [140]:

Variables using the first EM calorimeter layer:

- $F_{side} = \frac{E_{7\times1} - E_{3\times1}}{E_{3\times1}}$, where $E_{x\times y}$ is the energy measured in the cells of size $x \times y$, describing the energy outside the core of the three central strips divided by the energy within the three central strips;
- $w_{s,3} = \sqrt{\frac{\sum E_i (i - i_{max})}{\sum E_i}}$ is the shower width for the three strips around the strip with the maximum energy deposit;
- w_{stot} computed the same way as $w_{s,3}$ but for a wider window of cells approximately 20×2 strip cells in $\eta \times \phi$;
- $\Delta E = \left[E^{S1}_{2^{nd}_{max}} - E^{S1}_{min} \right]$ evaluates the presence of a second energy deposit and is measured as the difference between the energy of the strip cell with the second largest energy deposit and the energy of the cell with the least energy between the largest and second largest energy deposit;
- $E_{ratio} = \frac{E^{S1}_{1^{st}_{max}} - E^{S1}_{2^{nd}_{max}}}{E^{S1}_{1^{st}_{max}} + E^{S1}_{2^{nd}_{max}}}$ is the front maxima (first and second largest energy deposit) relative ratio.

Variables using the second EM calorimeter layer:

- $R_\eta = \frac{E^{S2}_{3\times7}}{E^{S2}_{7\times7}}$ is the middle η ratio computed as a ratio of the energy in a 3×7 region of the second EM calorimeter cells to the energy in a 7×7 region;
- $R_\phi = \frac{E^{S2}_{3\times3}}{E^{S2}_{3\times7}}$ is the middle ϕ ratio, similar to R_η, but this variable has different behavior depending on the photon candidate conversion status, since converted photons - an e^+e^- pair bend their trajectories in the magnetic field of the solenoid and produce a wider energy deposit in ϕ;

[2] corresponding to a 3×7 EM cluster in $\eta \times \phi$ in the barrel region for a mean number of interactions per bunch-crossing of $\mu = 20$

- $w_{\eta 2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2}$ measuring the lateral width of the shower in the second layer of the EM calorimeter using cells within a window of 3×5;

and a variable using the hadronic calorimeter information:

- $R_{had} = \frac{E_{\mathrm{T}}^{had}}{E_{\mathrm{T}}}$ for a photon candidate is a ratio of the total transverse energy deposited in the hadronic calorimeter divided by the total transverse energy deposited in the EM calorimeter (in the region of $|\eta| < 0.8$ and $|\eta| > 1.37$ only the energy of the first layer of the EM calorimeter is considered for the numerator).

The *loose* selection uses only the variables R_{had}, R_η and $w_{\eta 2}$, their cuts being determined and optimized on MC simulation. It is used for triggering purposes and provides a prompt photon efficiency ranging from 97% for photons with $E_{\mathrm{T}} = 20$ GeV to above 99% for photons with $E_{\mathrm{T}} > 40$ GeV [162].

The *medium* selection is also used for triggering purposes, and uses the *loose* selection with additional requirements on the shower shape variables w_{stot} and E_{ratio}.

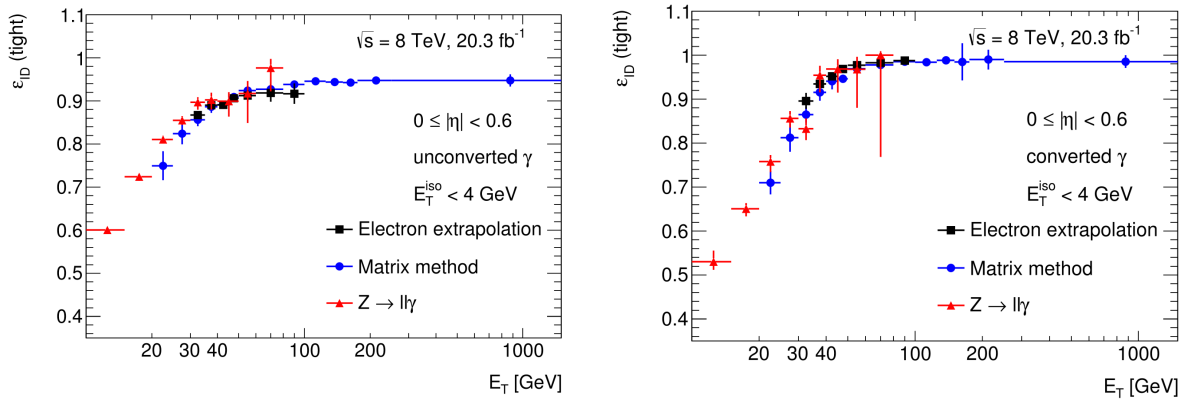
The *tight* selection considers all nine shower shape variables, adding the information of the first calorimeter layer to the variables used by the *loose* selection and tightening the *loose* selection criteria. The selection cuts were chosen to provide an expected efficiency of 85% for photons with $E_{\mathrm{T}} > 40$ GeV corresponding to an expected background rejection power of about 5000 [163]. The cuts do not depend on E_{T} of the photon candidates, but are optimized separately for seven η regions, due to the difference of the material budget in front of the calorimeter in different η positions, and independently for converted and unconverted photons.

Three data-driven methods have been used to compute the identification efficiencies for photons in ATLAS. The first method uses $Z \to ll\gamma$ radiative decays, a clean sample of prompt and isolated photons which have been produced during Run 1. As a function of E_{T}, the identification efficiency is estimated as a fraction of all selected probes passing the *tight* identification requirements. The second method, called electron extrapolation, uses $Z \to ee$ decays to deduce the distributions of photon discriminating variables based on the similarities between electron and photon EM showers. It provides precise identification efficiencies for photons of $30 \lesssim E_{\mathrm{T}}^{\gamma} \lesssim 100$. The third method, called the matrix method, bases its discrimination power on isolation requirements. Photons are required to be isolated from tracks in the inner detector and sample purity is extracted before and after applying the *tight* selection requirements, providing identification efficiencies up to several hundreds of GeV.

All three measurements use photon candidates detected within the fiducial region of the detector of $|\eta| < 1.37$ and $1.52 < |\eta| < 2.37$, with the result given in four $|\eta|$ regions of $|\eta| < 0.6$, $0.6 \leq |\eta| < 1.37$, $1.52 \leq |\eta| < 1.81$ and $1.81 \leq |\eta| < 2.37$. Photon candidates in the identification studies are also required to be isolated with $E_{\mathrm{T}}^{calo,iso} < 4$ GeV (for definition see Section 3.4).

The complete results of the three methods for $\sqrt{s} = 8$ TeV data are summarized in Ref. [164]. The comparison of the identification efficiencies measured for converted and unconverted photons in the most central region ($|\eta| < 0.6$) is shown in Figure 3.2.

An additional algorithm, the *ambiguity resolver*, described in Ref. [162], is used to resolve any residual ambiguities for photon candidates that have also been reconstructed as electrons. A different approach is taken depending on the conversion status of the photons. For converted



(a) (b)

Figure 3.2: Photon ID efficiencies obtained by the three methods for (a) unconverted and (b) converted photons in the $|\eta| < 0.6$ barrel region [164].

photon candidates that have also been reconstructed as electrons, the track associated to the electron object is evaluated against the track(s) origination from the conversion vertex candidate matched to the same cluster. In case the track matches a track coming from the conversion vertex, the candidate is retained as a converted photon. An exception is made for double-track conversion vertex candidates where the matched track has a hit in the b-layer, while the other track does not. The converted photon candidate is removed if the track does not coincide with any of the tracks assigned to the conversion vertex candidate, unless the track p_{T} is smaller than the converted photon candidate p_{T}.

For single track converted photon candidates, this procedure is done only if the candidate $p_{\mathrm{T}} > 2$ GeV and $E/p < 10$, where E is the cluster energy and p the track momentum.

Unconverted photon candidates are reconstructed from objects that have been recovered as electrons if the track $p_{\mathrm{T}} < 2$ GeV, or if $E/p > 10$ and no single-track converted photon was reconstructed from the same cluster. This procedure recovers roughly 86% of unconverted photons that have been initially categorized as electrons.

From MC studies, for photons of $E_{\mathrm{T}} > 25$ GeV 96% are reconstructed as photons only, and 4% are incorrectly reconstructed as electrons after applying the *ambiguity resolver* [159].

3.4 Photon isolation

Jets originating from non-perturbative QCD hadronization of quarks and gluons generally contain many hadrons. Isolation requirements imposed on photon candidates aim to distinguish photons coming from decays of neutral hadrons in jets. The principle is shown in Figure 3.3, where the neutral hadron produced through hadronization, has many tracks and also some energy deposits within a cone of ΔR around it. Prompt photons on the other hand should be well isolated. Isolation is defined by looking at the calorimetric deposits around the photon candidate or by studying the tracking information provided by the inner detector. In a physics analysis, one can require only the calorimeter isolation to be within a certain range (typically below 4-6

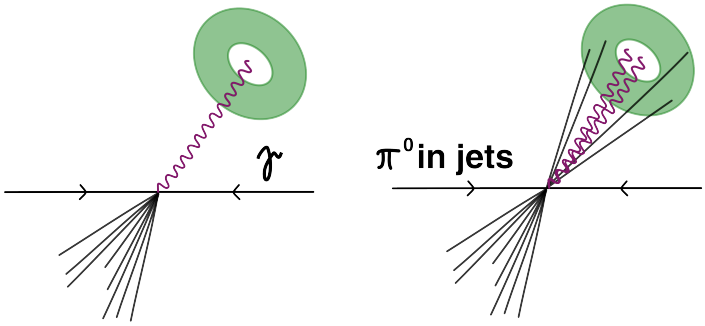


Figure 3.3: Illustration of the photon isolation for photons (left) and π^0 decaying into two photons inside a jet.

GeV), or combine the two isolations, based on what better suits the analysis in question. For a specific photon selection in an analysis, studies of the isolation efficiency are performed as a function of E_{T}^{γ}, η or other variables, e.g. variables one wants to use for cross section measurements, and a flat efficiency is typically required as a function of these variables.

3.4.1 Calorimeter isolation

The variable used to describe the calorimeter-based isolation is the transverse isolation energy $E_{\mathrm{T}}^{calo,iso}$. It is reconstructed by measuring the energy deposits in the cells surrounding the photon candidate. Two approaches have been used in ATLAS.

The first approach is the so-called `EtCone40` obtained by summing the energy of cells inside a cone of $\Delta R < 0.4$ around the photon candidate energy cluster and subtracting the energy of the candidate itself within the 5×7 cells cluster in which it was reconstructed. In addition, a leakage correction derived from MC as a function of E_{T} and η is applied to account for photon energy outside of the cluster.

Contributions from underlying event and pile-up noise are suppressed by estimating the ambient transverse energy density on an event by event basis [165, 166]. This energy density is defined as the median of the jet transverse energy distribution which is the ratio of the E_{T}^{jet} to the jet area. Its contribution to the transverse isolation energy is computed by multiplying the energy density to the total isolation cone area.

Underlying event contributions come from remnants of protons that were not involved in a hard collision. Contributions from pile-up, as previously described in Section 2.1.3, can originate from proton-proton collisions from the same beam bunch-crossing (in-time pile-up) or from collisions that occurred in bunch-crossings before or after (out of time pile-up). The time in which signal from one bunch crossing is integrated is larger than the beam bunch spacing, hence the out of time pile-up contribution.

The correction of the ambient energy density is based on jet reconstruction using free sized topological clusters. Topological clusters (topoclusters) start with a seed cell and iteratively add neighboring cells in all three dimensions if they pass a threshold, defined as a function of the expected noise [167].

The isolation energy, on the other hand, is computed by summing the energy of EM calorimeter cells. Therefore not all cells will necessarily belong to topoclusters, as illustrated in Figure 3.4. The residual dependence on pile-up is shown in Figure 3.5(a), where the isolation energy is shown as a function of the Bunch Crossing Identification (BCID).

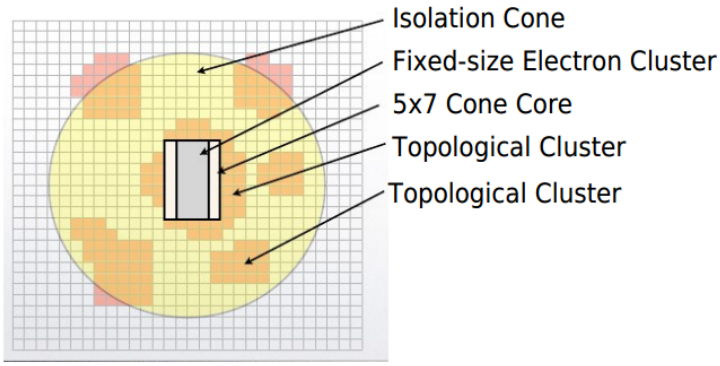
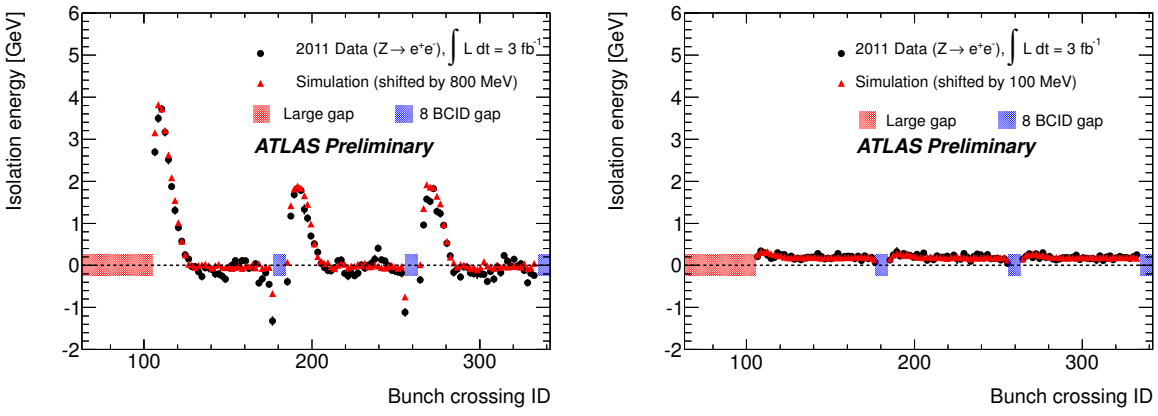


Figure 3.4: Illustration of the computation of isolation variables [168].



(a) (b)

Figure 3.5: Pile-up dependence of the (a) cell based calorimeter transverse isolation energy and the (b) topocluster based calorimeter transverse cell energy [168].

A second approach was derived, used mostly in 2012 data, where the transverse isolation energy is computed using topoclusters, called `topoEtCone40`, where the ambient energy density correction computation is coherent. Only the topoclusters with a barycenter within the cone of $\Delta R < 0.4$ are summed to give the $E_{\mathrm{T}}^{iso,topo}$. The change can be seen in Figure 3.5(b), where the pile-up dependence has been mitigated. The energy used to compute the `topoEtCone40` variable is calibrated to the EM scale, as described in Section 3.2. The `topoEtCone40` isolation variable is also corrected for the leakage and energy density as described above. This variable will be further referred to as $E_{\mathrm{T}}^{calo,iso}$ and is used in all 2012 analyses presented in this text. For 2011 studies, the first approach is used.

3.4.2 Track isolation

Using the information of the inner detector system, a track-based isolation variable can be computed [160].

The track isolation is defined as the sum of momenta of all tracks with $p_{\mathrm{T}} > 1$ GeV in a cone of $\Delta R < 0.2$ around the photon candidate and belonging to the primary vertex matched to the photon candidate[3]. The momentum threshold requirement for the tracks is applied to reduce most of the tracks coming from pile-up and underlying event. In addition, the tracks within $\Delta R < 0.1$ are required to have an impact parameter with respect to the beam $d_0 < 0.1$ mm [160]. The $p_{\mathrm{T}}^{track} < 15$ GeV requirement is imposed to remove tracks of very asymmetric conversions. Also, the tracks used must not be part of a reconstructed conversion vertex and must have a at least eight hits in the ID system with at least one of those hits in the b-layer (innermost layer of the ID). In case of a converted photon, the conversion tracks are excluded from the sum. The track isolation used in this text will be labeled $p_{\mathrm{T}}^{track,iso}$.

3.5 Photon cleaning

In order to reject events with EM topoclusters that have been reconstructed using fake energy deposits stemming from detector problems, a selection should be applied for all photon analyses. The sources of problems related (not only) to photons can be summarized as [169, 170]:

- LAr calorimeter noise bursts;
- data integrity errors;
- HV trips;
- instability in HV, ramping of the HV during the data-taking;
- masked calorimeter cells
- missing FEBs.

[3]The position of the track is taken at the vertex and the photon position is taken as the center of its reconstructed calorimeter cluster.

4 Jets in ATLAS

Quarks and gluons cannot be directly observed due to their colored nature, they are confined within color singlets - hadrons. They instead fragment into many particles, forming collinear flows of energy in the form of other non-color-confined particles (photons, electrons, pions and so on). These bursts of particles are called jets, they can be experimentally observed in a detector and if measured properly, they truly reflect the energy and angular properties of the original partons [171]. There are several jet-reconstruction algorithms, developed on various principles, trying to cluster all particles that came from one quark or gluon produced in a collision, combining the theoretical prediction of jets with their experimental observation.

4.1 Jet reconstruction algorithms

Jet reconstruction algorithms are tools designed to combine energy deposits measured by the calorimeters into jets, where the reconstructed jet should best reproduce the properties of the initial parton or gluon (such as energy, transverse momentum and direction). There are two main classes of jet reconstruction algorithms: cone algorithms and sequential recombination algorithms.

The first approach to jet reconstruction is the use of cone algorithms based on the fact that a jet consists of a large amount of hadrons in a small angular region. Therefore, the cone algorithm combines particles in the $\eta - \phi$ space with their closest neighbors within a cone of radius

$$R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{4.1}$$

Cone algorithms, such as the midpoint cone, start from a seed as a center of the cone and iterate until the cone is stable (as defined in Ref. [172] on page 3). This method, however, fails to identify all stable cones and leads to infrared collinear unsafety in perturbative calculations [173]. A seedless cone algorithm SISCone [172] was developed for infrared safety - low p_{T} particles are not falsely attached to jets, therefore not changing the momentum or direction of the reconstructed jet.

The second class of jet algorithms is called sequential recombination algorithms. The k_{T} algorithm attempts an approximate inversion of the QCD fragmentation process. It uses proximity in momentum and space by successively merging pairs of particles (or calorimeter clusters or

towers in four-momentum representation) in order of increasing relative transverse momentum squared defined as

$$d_{ij} = min\left(p_{T_i}^{2p} p_{T_j}^{2p}\right)\left(\frac{\Delta_{i,j}^2}{R^2}\right), \tag{4.2}$$

where the $\Delta_{i,j}$ is the distance between two particles i and j in the $y-\phi$ plane defined as

$$\Delta_{i,j}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2. \tag{4.3}$$

For each object i of energy reconstructed in the calorimeter, the k_{T} algorithm pairs it with its nearest neighbor, finds the minimal $\delta_{i,j}$ and if the minimal value $d_{min} = p_{T_i}^{2p}$, then the object becomes a jet, otherwise the two objects are merged and the process continues. The parameter p can be

$$p = \begin{cases} 1 & \mathrm{k_T\ algorithm} \\ 0 & \mathrm{Cambridge/Aachen\ algorithm} \\ -1 & \mathrm{anti-k_T\ algorithm} \end{cases}$$

The k_{T} family of jet reconstruction algorithms has collinear and infrared safety as its main advantages. The k_{T} algorithm, where $p = 1$, starts clustering from softer objects before harder objects are added to the jet. In case of $p = 0$, it is the Cambridge/Aachen algorithm, which only considers the geometrical distance of the two objects. The anti-k_{T} algorithm [174], with $p = -1$, starts from hard objects collecting adjacent soft objects, effectively reverting the effect of soft radiation. If a hard particle has no hard neighbors within cone of $2R$, all surrounding soft objects will be clustered into the jet, forming a conical shape. If two hard objects are within $R < \Delta R < 2R$, the objects in the overlapping region are associated to one of the hard objects depending on their $d_{i,j}$ leading to complex jet shapes. If they are within a radius R, they are merged into a single jet. Figure 4.1 shows the result of four different clustering algorithms mentioned.

4.2 Jet reconstruction in ATLAS

The anti-$\mathrm{k_T}$ algorithm is the most used in ATLAS and also jets used in this work were reconstructed by this algorithm. ATLAS typically uses two R parameters for reconstruction: $R = 0.4$ and $R = 0.6$.

The FastJet program [175] with four-momentum recombination scheme, is used to build jets from three-dimensional topological calorimeter cell clusters. The cells are calibrated at the electromagnetic scale (EM) determined from test beam measurements and Monte Carlo simulation. Topological clusters are formed starting from cells with high signal to noise ratio ($S/N > 4$), then iteratively adding surrounding cells with a signal to noise ratio above a certain threshold ($S/N > 2$) to the cluster. Only topoclusters with positive energy are considered as input for the jet-clustering algorithms. A second topocluster collection is built by taking the previous collection and calibrating the calorimeter cells in the clusters so that the response of the calorimeter to hadrons is correctly reconstructed. It uses the local cell signal weighting (LCW) method [176], aiming at improving the resolution compared to the EM scale.

Finally, jets are kept only if they reach $p_{\mathrm{T}} > 7$ GeV and are within $|\eta| < 4.5$. The resulting jets have a four-momentum equal to the sum of the four-vectors associated to the jets [176].

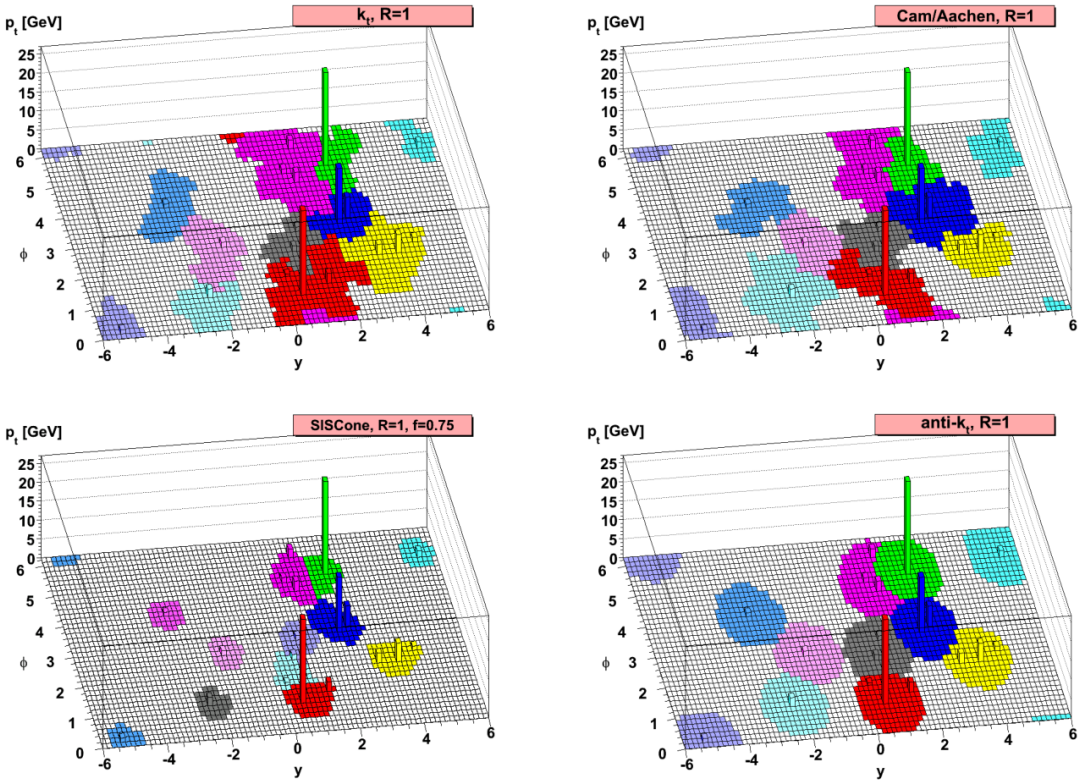


Figure 4.1: Illustration of the shapes of jet areas of different jet algorithms used on the same event [174].

Jets built from charged particle tracks (*track jets*) that originate from the PV and are matched to the calorimeter jets are used for validation, due to being insensitive to pile-up. These are only considered within the ID for $|\eta| < 2.5$. Tracks with $p_{\mathrm{T}}^{track} > 0.5$ GeV are selected, requiring at least one hit in the pixel and at least six hits in the SCT. Additional requirements on transverse and longitudinal impact parameters are applied. Detailed description of track reconstruction can be found in Ref. [177].

4.3 Jet quality selection in ATLAS

In order to study jets produced in hard scattering events in a physics analysis, the main backgrounds, meaning jets produced by processes other than hard scatter, must be identified and dealt with accordingly. Among the backgrounds that could be detected as jets are beam-gas events, where a proton from the beam collided with residual gas within the beam pipe; beam-halo events, e. g. events caused by interactions in the tertiary collimators in the beam-line far away from the ATLAS detector and cosmic ray muons or large calorimeter noise, such as noise bursts or rare coherent noise. These were all studied and a quality selection with an efficiency above 99% was provided [176].

The main source of background for hard scatter jets are jets coming from pile-up processes. *Track jets* are used to discriminate jets coming from the hard process from pile-up jets. The

discriminant used is the Jet Vertex Fraction (JVF) - fraction of a jet constituent transverse track momentum contributed by each vertex, defined as

$$\mathrm{JVF}(jet, vtx_j) = \frac{\sum_k p_{\mathrm{T}}(trk_k^{jet}, vtx_j)}{\sum_n \sum_l p_{\mathrm{T}}(trk_l^{jet}, vtx_n)}, \tag{4.4}$$

where vtx_j is the corresponding vertex, originally developed by the D‰ collaboration [178]. A large portion of jets have JVF$= -1$ due to not having any tracks. Pile-up jets exhibit a falling exponential shape of the JVF distribution starting at zero, while hard-scatter jets peak at one, as seen in Figure 4.2.

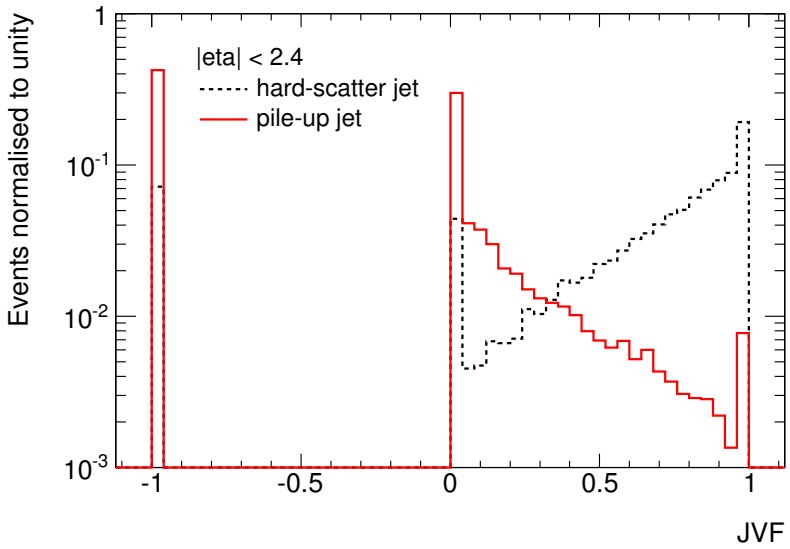


Figure 4.2: Distribution of the JVF variable defined in Eq. 4.4 for hard-scatter jets (black line) and pile-up jets (red line) for SHERPA MC with two photons and up to three jets in the final state [179].

A physics analysis should be insensitive to changes in jet multiplicity coming from pile-up. By requiring jets to pass a JVF cut at 0.5, the pile-up jet contribution is significantly reduced. The reduction depends on the selected p_{T}^{jet} cuts, and for a selection of jets with $p_{\mathrm{T}}^{jet} > 25$ GeV the residual pile-up jet contamination can be around 10%, as seen in Section 7.2.4. The jet multiplicity as a function of N_{PV}, after applying a JVF cut stays constant, as seen in Figure 4.3. The cut can be determined based on the individual analysis selection and is applied for jets for $|\eta| < 2.4$ and $p_{\mathrm{T}} < 50$ GeV to only reject pile-up jets, MC studies showed that the JVF cut applied on jets with $p_{\mathrm{T}} > 50$ GeV may remove also hard-scatter jets and create an inefficiency [180, 181]. For an application of this cut, see Section 7.2.4.

4.4 Jet calibration in ATLAS

Reconstructed jets need to be calibrated mainly to account for energy not measured by the calorimeter, since both calorimeters are of the sampling type. Energy losses due to inactive

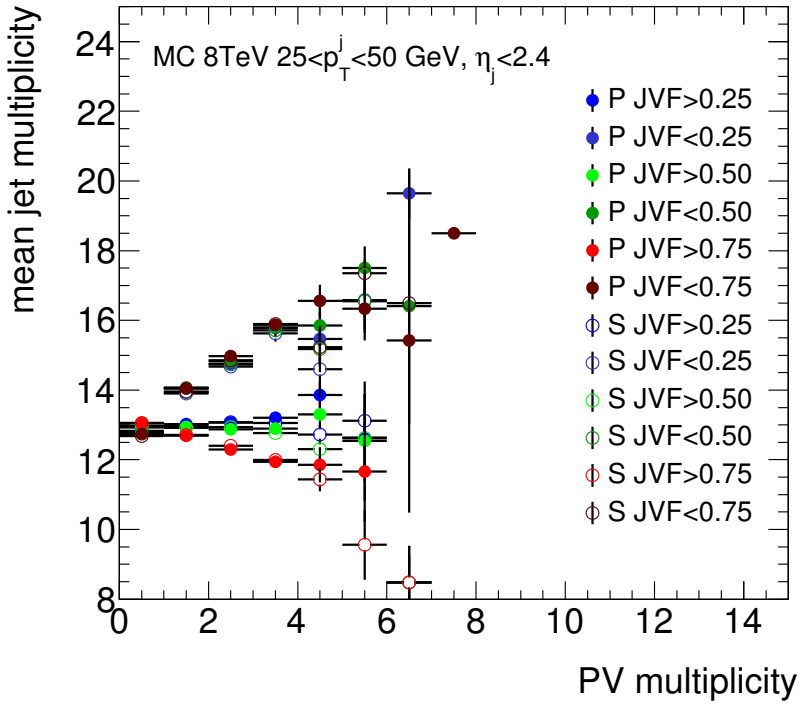


Figure 4.3: Primary vertex multiplicity as a function of the mean jet multiplicity for several configurations of the JVF selection and for PYTHIA (P, full points) and SHERPA (S, empty points) MC simulation events with two photons and up to two jets in the final state, passing the selection presented in Section 7.2.

regions of the detector, the so-called *dead material*, and leakage outside of the calorimeters, need to be accounted for as well. The energy leakage outside of the jet cone is considered too. Further, signal losses in the calorimeter clustering due to insufficient energy deposits to form a topocluster and out of cluster corrections need to be derived. Finally, the effect of pile-up needs to be evaluated and corrected for.

ATLAS has developed several jet calibration schemes [140] of different complexity and different sensitivity to systematic effects. The calibration scheme used for jet calibration in this work is the electromagnetic plus jet energy scale scheme (EM+JES). This jet calibration starts from the calorimeter energy at the EM energy scale, measuring the energy deposited in the calorimeter by the EM showers, and correcting it relative to the response of the calorimeter to the true energy corresponding to a jet of stable particles entering the detector in MC simulation. Jets in MC are reconstructed the same way as in data, using truth particles with a lifetime greater than 30 ps as input, excluding muons and neutrinos.

Jets are calibrated in four steps [176]:

- **Pileup correction** - energy coming from pile-up events is subtracted on average as a function of pseudorapidity η, bunch spacing τ_{bunch} (out of time pile-up) and number of primary vertices N_{PV}. Charged particle jets are compared with matched calorimeter jets, the correction being derived in situ. Low dependence on pile-up conditions is achieved by

considering only reconstructed tracks matched to the primary hard-scatter vertex when building charged-particle jets.

- **Origin correction** - direction of the jets is corrected so that they originate from the primary hard scatter vertex instead of the geometrical center of the detector, where they are first reconstructed at. This does not affect the energy of a jet.

- **Energy and direction correction** - applying calibration derived from MC simulation, to correct for jet energy and position. This correction is computed using all isolated calorimeter jets that have a matching isolated truth jet within $\Delta R = 0.3$. An isolated jet is defined as a jet, which within $\Delta R = 2.5 \times R$ has no other jet with $p_{\mathrm{T}}^{jet} > 7$ GeV. It has to be isolated with respect to the same jet type (calorimeter or truth jet). The jet energy scale calibration is parametrized as a function of calorimeter η_{calo} (not the origin-corrected one[1]) and energy is used instead of p_{T}, since the calorimeter measures energy. The EM-scale jet energy response is

$$R_{EM}^{jet} = \frac{E_{EM}^{jet}}{E_{truth}^{jet}}, \tag{4.5}$$

  where E_{jet}^{EM} is the jet energy at the EM scale and E_{truth}^{jet} is the truth jet energy. An average jet response $\left\langle R_{EM}^{jet} \right\rangle$ is defined as the peak position of a Gaussian fit to the distribution in Eq. 4.5. Also an average jet energy response $\left\langle E_{EM}^{truth} \right\rangle$ is derived and a response function is obtained for each E_{truth}^{jet} and η_{det} bin, parametrized as

$$F_{calib,k}(E_{EM}^{jet}) = \sum_{i=0}^{N_{max}} a_i \left(\ln E_{EM}^{jet}\right)^i, \tag{4.6}$$

  where a_i are free parameters and $N_{max} = 1 \ldots 6$ is chosen based on the goodness of fit. The final jet energy scale correction is

$$E_{EM+JES}^{jet} = \frac{E_{EM}^{jet}}{F_{calib}(E_{EM}^{jet})|_{\eta_{jet}}}. \tag{4.7}$$

  The average jet energy scale correction is shown in Figure 4.4 for $\sqrt{s} = 7$ TeV measurements.

- **In situ correction** - residual correction applied to data account for the accuracy of the MC simulation. Using a combination of $Z - jet$, $\gamma - jet$ and multi-jet balance techniques, central jet response is compared in data and MC. The central calibration is transferred to forward jets using η-intercalibration, balancing the p_{T} of a forward jet with a central jet. Generally, the difference between the calibration derived by using data and MC is around 2%.

For 2012 data, a global sequential correction, a multi-variate extension of the EM+JES calibration, is applied after the JES calibration due to non-trivial differences observed between the quark and gluon initiated jets. This correction significantly reduces the systematic uncertainties.

[1] Calorimeter η_{calo} shows a straightforward correspondence to a calorimeter region.

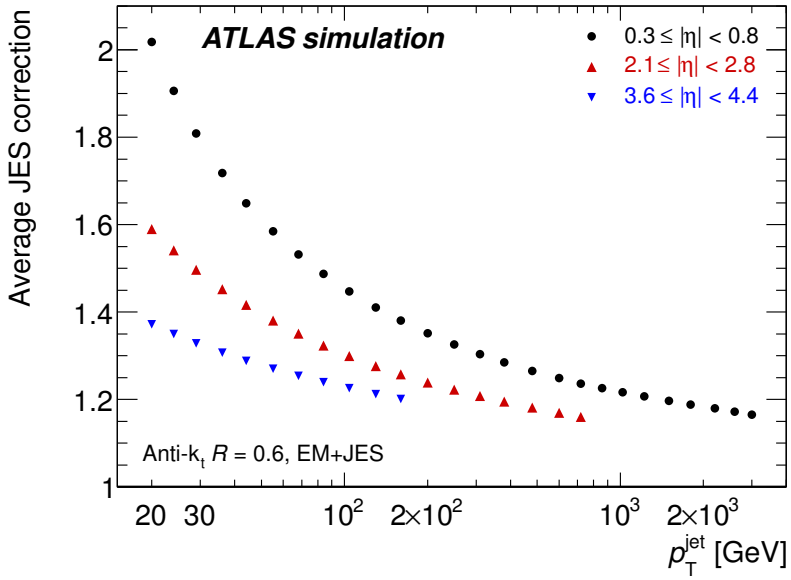


Figure 4.4: Average jet energy scale correction as a function of the calibrated jet transverse momentum for three η intervals obtained from MC simulation at $\sqrt{s} = 7$ TeV [176].

The corrections are applied as a function of the topology of energy deposits in the calorimeter, tracking information and muon spectrometer information. The sequential application of corrections leaves the mean value of the jet energy response unchanged.

The total jet energy scale uncertainty is a product of uncertainties coming from multiple sources. A total of 65 nuisance parameters are varied up and down within the uncertainty and applied separately for a specific analysis. The systematic and statistical uncertainties considered can be summarized as:

- 22 systematics and 34 statistical uncertainties from in situ methods;
- a single hadron response uncertainty only affecting high-p_T jets beyond the reach of in situ techniques;
- a systematic and a statistical uncertainty on the η-intercalibration;
- four sources of uncertainty due to pile-up calibration [182]
- two sources due to jet flavor

Simplified models have been developed for analyses not aiming at precision measurements, see more in Section 7.7.4. The calibration results in a correction at the level of 0.5% to the jet energy scale in data and an uncertainty of less than 1% for central anti-k$_T$ jets with $R = 0.4$ and $100 < p_T < 1500$ GeV, which is a significant improvement w.r.t. 2011 results shown in Figure 4.4.

4.5 Jet energy resolution

The jet energy resolution (JER) can be parameterized into three terms

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{N}{p_{\mathrm{T}}} \oplus \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus C, \tag{4.8}$$

where N is the noise term, S is the stochastic term due to the sampling nature of the calorimeter and C is a p_{T}-independent constant term. At low p_{T}, jets can be affected by the calorimeter electronics' noise.

Two methods are used to measure the noise term: Zero Bias[2] events using random cone p_{T} balance in opposite ϕ directions and distributions of soft jet momenta. The measured noise term ranges from ≈ 1.3 to ≈ 3.7, depending on the method and calibration scheme (EM or LCW[3]) and the two methods agree at the level of $\approx 15\%$ for $R = 0.4$ jets in data and a difference of up to 7.5% is observed in MC. The differences are taken as systematic uncertainties. The noise term is also derived for a no pile-up scenario in order to evaluate additional noise terms, other than the electronic and pile-up noise. The total noise term is then taken as a sum in quadrature of the noise term in a no pile-up sample and the one originating from pile-up (the two methods mentioned above).

The measurements of the JER using di-jet and multi-jet events [183] and vector bosons and jets [184] can be combined, keeping the noise term at the central value obtained from methods described above and propagating the uncertainties of the three measurements accordingly. The three methods are well in agreement. The final fitted values and uncertainties on the three JER terms for $R = 0.4$ EM+JES anti-$\mathrm{k_T}$ jets with N held constrained are $N = 3.33 \pm 0.63$, $S = 0.71 \pm 0.07$ and $C = 0.030 \pm 0.003$. The uncertainty on the jet energy resolution is less than 3% at 20 GeV and below 1% above 100 GeV. Previously this was as large as 5% at low η and up to 10% at high η [176].

Unfortunately, this improvement in resolution is not included in the calibration used for jets in the analysis presented in Chapter 7, due to technical difficulties, namely of some variables needed to estimate the new corrections not being present in the data samples available. The jet energy scale systematic uncertainties will therefore not be the most optimal that ATLAS is capable of.

[2]Zero bias events are events where no real particles are simulated passing the detector, only the so-called Geantinos. Such events are useful to study noise properties.

[3]LCW stands for Local Cluster Weighting, which is a different approach to jets, first calibrating clusters and then reconstructing jets. The EM scheme, used for jets in this work, first reconstructs jets using the anti-$\mathrm{k_T}$ algorithm and then calibrates them as a whole.

5 Refined photon calibration

The electron and photon calibration method described in Section 3.2 was used in ATLAS for most physics analyses using $\sqrt{s} = 7$ and 8 TeV data. The Higgs boson discovery analysis [18] and the search for additional resonances decaying into photon pairs described in Chapter 6 uses the *calibration hits* method. Following the Higgs boson discovery, a refined multi-variate-based calibration was derived between 2012 and 2013, mainly motivated by the possibility of improving the Higgs boson mass measurement. It was used for example for the final Higgs publications by ATLAS [1, 27] and also the diphoton in association with jets cross section measurement described in Chapter 7.

5.1 Refined MVA-based electron and photon response calibration

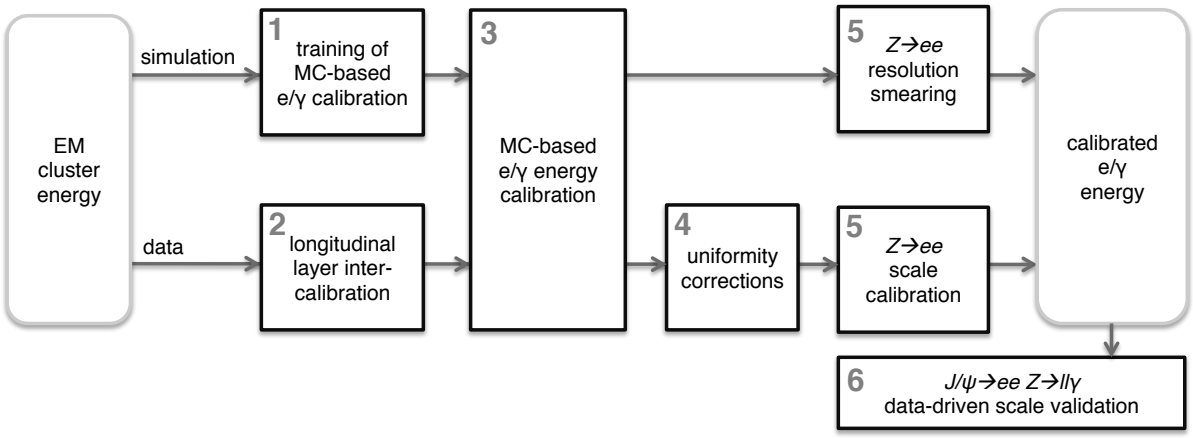


Figure 5.1: Schematic overview of the electron and photon energy calibration procedure in ATLAS [142].

The reconstructed energy of a photon candidate[1], described in Section 3.1, is to be corrected in

[1]The same goes for electrons, the calibration of the two objects is intertwined.

several steps, summarized in Figure 5.1 and briefly described in the following sections, according to Ref. [142].

5.1.1 Training of MC-based e/γ calibration

A multivariate algorithm (MVA) is used to calibrate the response of electrons and photons impinging the EM calorimeter using the longitudinal shower development and additional tracking information as inputs. The MVA aims to optimize the energy response and minimize the resolution. It is performed separately for converted and unconverted photons and the calibration constants are derived in η and E_{T} bins. All photons with $|\eta| < 2.37$ excluding the transition region between the barrel and the endcap calorimeter are calibrated with the MVA calibration. The photons that pass through the transition region are calibrated using the original EM calibration [160, 161] described in Section 3.2.1, but these photons are not used in physics analyses due to the large systematic uncertainties they carry.

Input variables for the method are the total energy measured in the calorimeter E_{calo}; the ratio of the energy measured by the presampler to the calorimeter energy E_{PS}/E_{calo}; the shower depth $X = \sum_i X_i E_i / \sum_i E_i$, where E_i is the cluster energy and X_i is the calorimeter thickness of layer i in radiation lengths; $\eta_{cluster}$ the cluster barycenter pseudorapidity in the ATLAS coordinate system and the η_{calo} the pseudorapidity in the calorimeter frame[2]. Photons are defined as converted if the conversion radius $R_{conv} < 800$ mm. If the vector sum of the conversion track momenta p_{T}^{conv} is larger than 3 GeV, the R_{conv} is also used as input for the MVA. If both tracks contain at least one hit in the Pixel or the SCT, $p_{\mathrm{T}}^{conv}/E_{calo}$ ratio and the fraction of the conversion momentum carried by the highest p_{T} track $p_{\mathrm{T}}^{max}/p_{\mathrm{T}}^{conv}$ is also included as input.

The MVA calibration is computed in 10×9 bins of $|\eta_{cluster}| \times E_{\mathrm{T}}^{calo}$ and additional 2×6 bins are used for the region where the two barrel wheels are combined. Further nonlinearities in the energy response are corrected by adjusting the peak position of the ratio of the output energy E_{MVA}/E_{true} to unity, corresponding to a correction ranging from +2–5% depending on η at $E_{\mathrm{T}} = 10$ GeV and decrease to zero at around 100 GeV. The most probable values (MPV) of E/E_{true} obtained from the MVA calibration are below 0.3% for $E_{\mathrm{T}}^{true} > 10$ GeV, better than 1% at lower transverse energies and reach up to 2% for converted photons in a few localized regions. An improvement of more than a factor of two is achieved at the high $|\eta|$ regions w.r.t the *calibration hits* method. The resolution improvement w.r.t. the *calibration hits* method ranges from 3–10% in the barrel and 10–15% in the end-caps for unconverted photons. For converted photons the improvement is up to 20%.

5.1.2 Longitudinal layer intercalibration

Muons from $Z \to \mu\mu$ decays are used to study the intercalibration of the response of L1 and L2 of the EM calorimeter. The determination of the PS energy scale is performed by studying energy distributions of electrons in data and MC. The L3 of the EM calorimeters is not intercalibrated due to its negligible contribution at the energies studied in Run 1 data (see also Section 5.4).

Muons serve as probes of the EM calorimeter due to the fact that they are almost insensitive to material in front of the calorimeter. The observed muon energy distribution for each calorimeter layer is described by a convolution of a Landau and a Gaussian distribution describing the

[2]The LAr calorimeter is shifted w.r.t. the center of the ATLAS detector system by a few mm.

energy deposit and the electronic noise, respectively. To extract the MPV of the deposited energy, two methods are used: an analytical fit and an estimation using a truncated mean. The ratio of the muon energy deposits in the two layers is studied for data and MC defined as $\left\langle E_{1/2} \right\rangle^{data} / \left\langle E_{1/2} \right\rangle^{MC}$, where $E_{1/2} = E_1/E_2$, and the obtained values are shown in Figure 5.2. The central value is given by the average and the difference between the two methods is taken as

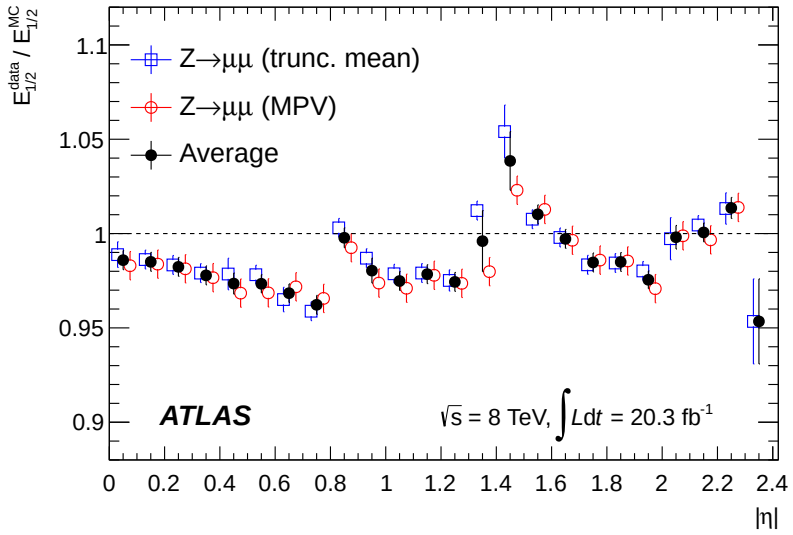


Figure 5.2: Ratio of the muon energy deposits in L1/L2 of the calorimeter for data/MC for the two methods used. The error bars represent the total (stat+syst) uncertainty specific to the $Z \to \mu\mu$ analysis [142].

a systematic uncertainty. Other systematic uncertainties considered include the muon path uncertainty, electric field changes at the layer transitions, difference between ionization current and deposited energy modelling and the uncertainty in the cross-talk between different calorimeter cells [185]. The statistical uncertainty is negligible. The bias observed from the ratio is up to 3% and varies greatly for different η regions. Due to the observed calibration bias, an η-dependent correction is applied to the energy measured in L2, since a direct comparison with data shows that the pattern observed in Figure 5.2 originates in the second layer of the EM calorimeter.

The presampler energy scale α_{PS} is derived from the ratio of PS energies measured using electrons from W and Z decays in data and MC simulation. The study of the expected correlation between $E_{1/2}$ and E_0 for electrons, done by varying the amount of material in various detector parts upstream of the calorimeters, aims to address the effects of passive material mismodelling, since this can affect the electron shower shapes description. Additional studies of the ratios of L1/L2 energy deposits between data and MC using unconverted photons from radiative Z decays and inclusive photons with PS energy deposits below 500 MeV, provide a deeper understanding of the upstream material budget. A study of the layer energy scale using photons, my personal contribution to the refined calibration, is described in Section 5.3. The PS energy scale is shown in Figure 5.3, and a visible improvement after the correction is observed. The measured PS scale α_{PS} is then applied on data as a correction factor.

After correcting the EM layers intercalibration, the distribution of $E_{1/2}$ can be used to determine

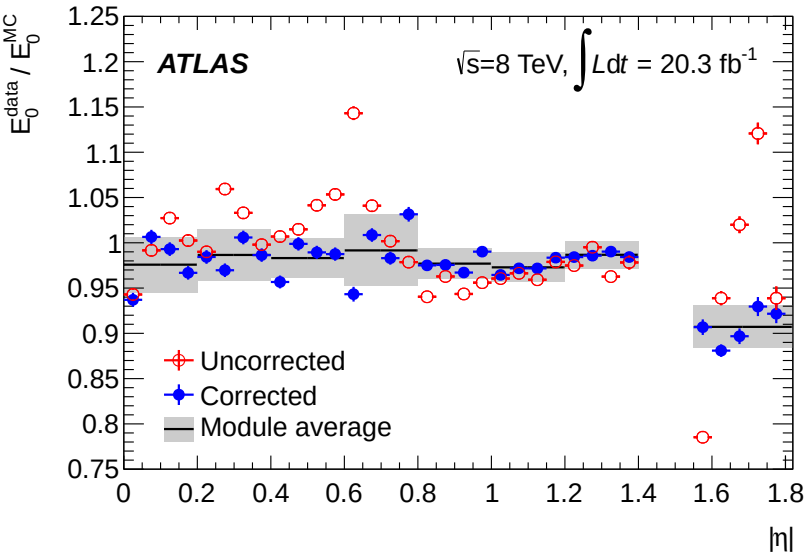


Figure 5.3: Ratio of the average PS energies, E_0^{data}/E_0^{MC} for electrons in data and simulation as a function of $|\eta|$ before and after corrections. The full line with a band represents the PS energy scale as a function of $|\eta|$, $\alpha_{PS}(\eta)$ and its uncertainty [142].

the passive material budget upstream of the calorimeter. Any additional material unaccounted for previously should be pointed out by higher values of the $E_{1/2}$ ratio in data. Additional studies using electrons and unconverted photons without any PS deposits were performed to examine the material description. Electrons can be sensitive to all material in front the calorimeter, while the unconverted photons are insensitive to the ID material upstream of the conversion radius[3]. Applying a PS veto on the unconverted photons will make them sensitive to the material between the PS and L1 of the calorimeter as illustrated in Figure 5.4. Simulated samples including variations of the material budget expressed in X_0 are used to compare with data and the observed difference defined as

$$\frac{\Delta X}{X_0} = \Delta E_{1/2}^{data} \left(\frac{\partial X/X_0}{\partial_{rel} E_{1/2}} \right) \tag{5.1}$$

is used to reevaluate the material determination. Following these studies, an improvement of the ATLAS material GEANT4 simulation was performed. The improved simulation and the measurement from data is compared in Figure 5.5 obtained from electrons for the L1 (red) and from both electrons and photons (blue).

5.1.3 MC-based e/γ energy calibration

The MC-based calibration using the new detector material description is applied to the cluster energies reconstructed from data and from MC simulations. Additional corrections, listed in subsequent steps are applied on top of it.

[3]Photons with $R_{conv} > 80$ cm are considered as unconverted in ATLAS, due to the fact that it is impossible to detect a conversion vertex outside the inner detector

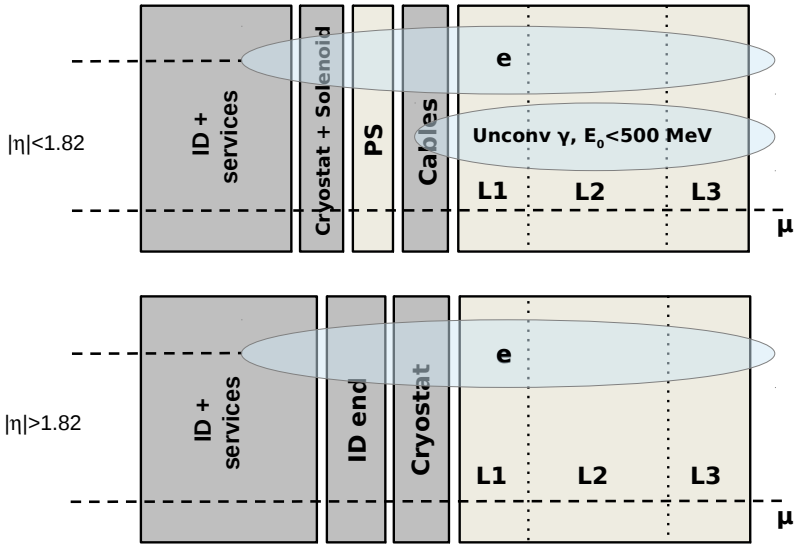


Figure 5.4: EM shower development illustration for electrons and unconverted photons with a PS veto in the $|\eta| < 1.82$ region (top) and electrons in the $|\eta| > 1.82$ region without the PS (bottom) [142].

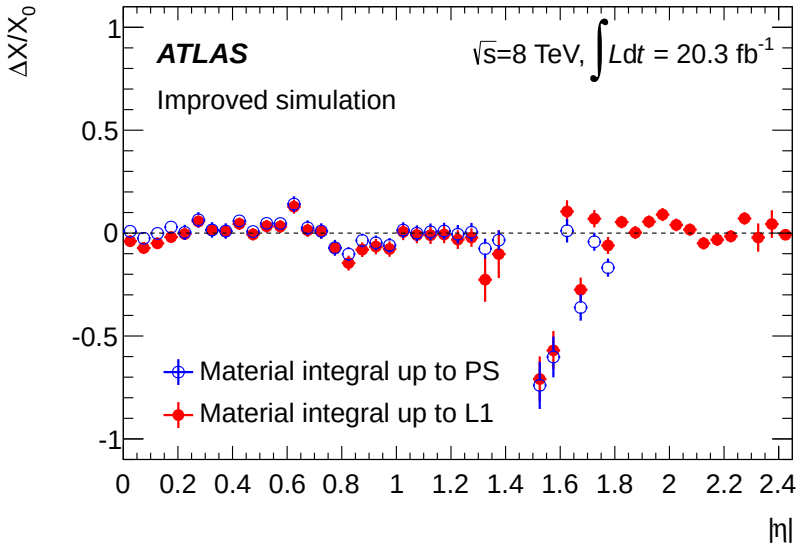


Figure 5.5: Difference between the material estimated from data and improved simulation as a function of $|\eta|$ [142].

5.1.4 Uniformity corrections

A set of studies has been performed to assure a good uniformity of the EM calorimeter energy reconstruction in η and ϕ and good stability of the calorimeter response w.r.t. the pile-up conditions and time [142].

- High voltage corrections - a few sectors of the EM calorimeter were set to a lower voltage than the nominal 2000 V due to short circuits occurring at the nominal voltage. The effect was studied on $Z \to ee$ events, looking at azimuthal profiles of the reconstructed di-electron invariant mass m_{ee}. A correction is derived and the response becomes uniform.
- Time dependence of the presampler response - a reduction of the high voltage was carried out in the PS due to the occurrence of sporadic electronic noise [161], from the nominal 2000 V down to 1600 V and later in 2012 down to 1200 V and some modules down to 800 V. First, a correction is derived at the cell level, using the expected HV dependence of the PS response, then it is verified by comparing PS response in $Z \to ee$ data and a residual η dependent variation is corrected by an empirical correction at the cluster level, reducing the bias down to 0.4% throughout the whole η range.
- Inter module widening - gravity induced inter module widening in the barrel region of 1(2)% for $\phi < (>)0$ has been observed by studying the energy response as a function of the azimuthal angle ϕ using $W \to e\nu$ events. A $\pi/8$ structure is adjusted by an empirical function, correcting the calorimeter response.
- An additional study of a potential asymmetry of photon detection as a function of ϕ, performed by me, is described in Section 5.5.
- Energy response difference between gains - the signal from the EM calorimeter, when treated by the electronics, can, depending on its amplitude, be processed by high, medium or low gain (see Section 2.3.1). For the $Z \to ee$ events, used to compute the absolute energy scale, both electrons have most of the cells of the second layer of the calorimeter treated with high gain. In di-photon analyses, especially in the $H \to \gamma\gamma$ analysis, more events contain a photon with at least one cell in the medium gain. $Z \to ee$ events were studied and a correction was derived by comparing the m_{ee} distribution in data and MC, for events with both electrons having all cells reconstructed with high gain to events with at least one cell using medium gain. The correction for low gain (very high energies) is assumed to be the same as the medium gain correction.

A residual azimuthal non-uniformity and stability of the energy response as a function of the mean number of interactions per bunch-crossing μ, time and number of reconstructed vertices was also studied and was found to be very well understood.

5.1.5 $Z \to ee$ resolution smearing and scale calibration

Studies performed on the non-uniformities in ϕ (as in Section 5.5) and in time showed that it is sufficient to derive the absolute scale only as a function of η. Defining the energy miscalibration as the difference in energy response in data and simulation, it can be parametrized as

$$E^{data} = E^{MC}(1 + \alpha_i), \tag{5.2}$$

where E^{data} (E^{MC}) is the electron energy in data (MC) and α_i is the difference w.r.t to the optimal calibration, in a given η bin i. The effect of the miscalibration on the invariant di-electron mass is

$$m_{ij}^{data} = m_{ij}^{MC}(1+\alpha_{ij}), \quad \text{where} \quad \alpha_{ij} = \frac{\alpha_i + \alpha_j}{2}. \tag{5.3}$$

m^{MC} is the invariant mass reconstructed in simulation and α_{ij} is the induced shift in η bins i and j of the two electrons. Second-order terms are neglected and the angle between the two electrons is assumed to be perfectly known. Assuming that the resolution curve for electrons is well modelled by the simulation up to a Gaussian constant term as

$$\left(\frac{\sigma_E}{E}\right)^{\text{data}} = \left(\frac{\sigma_E}{E}\right)^{\text{MC}} \oplus c \tag{5.4}$$

and the fact that the sampling term is well known from test beam studies [161] to be 10%, for each (η_i, η_j) category, the relative electron invariant mass must satisfy

$$\left(\frac{\sigma_m}{m}\right)_{ij}^{\text{data}} = \left(\frac{\sigma_m}{m}\right)_{ij}^{\text{MC}} \oplus c_{ij} = \frac{1}{2}\left[\left(\frac{\sigma_E}{E}\right)_i^{\text{MC}} \oplus c_i \oplus \left(\frac{\sigma_E}{E}\right)_j^{\text{MC}} \oplus c_j\right], \tag{5.5}$$

where the relative invariant mass resolution correction for (η_i, η_j) is $c_{ij} = (c_i \oplus c_j)/2$. The α and c parameters are determined by fitting separate templates of invariant masses including energy scale and resolution perturbations in narrow steps for (η_i, η_j) configurations and a two dimensional grid along (α_{ij}, c_{ij}) is constructed, with data categorized respectively. The optimal values of α and c parameters are obtained by χ^2 minimization and solving the system of Eq. 5.3-5.5.

Another method performing a likelihood fit to the energy scales using a parametrization instead of the templates is also used to obtain the corrections. A convolution of a Breit-Wigner and a Gaussian distribution is used and likelihood is

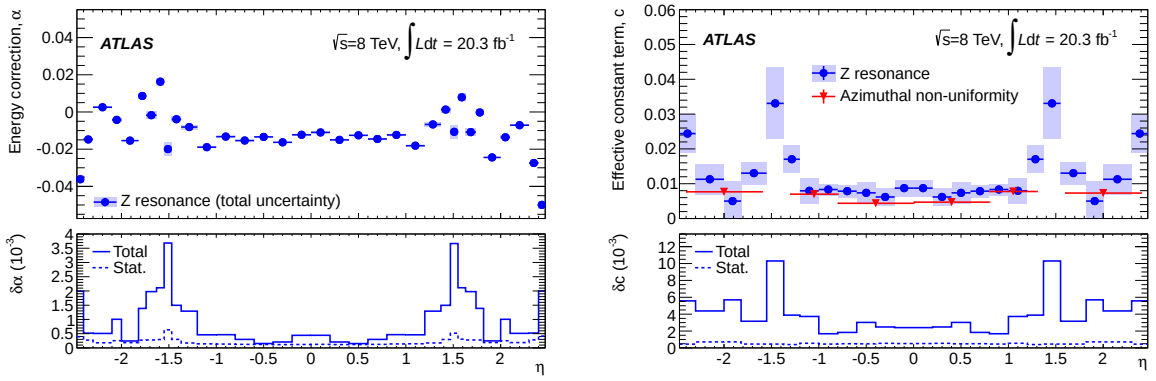
$$-\ln L_{\text{tot}} = \sum_{k=1}^{N_{\text{events}}} -\ln L_{ij}\left(\frac{m_k}{1+\frac{\alpha_i+\alpha_j}{2}}\right), \tag{5.6}$$

where N_{events} is the total number of selected events and $L_{ij}(M|\alpha_i, \alpha_j)$ is the probability density function quantifying the compatibility of an event with the expected Z boson line shape at the reconstruction level.

The two methods are applied to a $Z \to ee$ data sample and the η-dependent corrections defined w.r.t. the 2010 calibration scheme (described in Sec. 3.2.1) after uniformity and layer calibration corrections are within agreement to the level of 1.3×10^{-4} when the transition region between the barrel and the end-cap is not considered.

The template method result for the energy corrections α can be seen as a function of η in Figure 5.6(a). The effective constant term corrections c are shown in Figure 5.6(b), together with the contribution from azimuthal non-uniformity in ϕ, estimating its contribution to the effective constant term in four bins of η.

The uncertainties include energy scale and resolution distortion studies, selection criteria (tight ID, trigger, reconstruction efficiencies), bremsstrahlung momentum loss, pile-up, interaction point distribution, theoretical description of the Z line shape and the EW, top and multi-jet backgrounds. Further systematic uncertainties common for electrons and photons include the gain dependence of the energy response, the layer calibration uncertainty and the material modelling uncertainties. Photon-specific uncertainties include the conversion reconstruction inefficiency, fake conversions and lateral leakage mismodelling.



(a) (b)

Figure 5.6: Energy scale corrections α as a function of η (top left) with their stat. and total uncertainties (bottom left); effective constant term corrections c as a function of η (top right) with their stat. and total uncertainties (bottom right) and the contribution of the response uniformity [142].

5.1.6 $J/\psi \to ee$ electron energy scale validation

The electron energy scale is validated by looking at $J/\psi \to ee$ events in data. A sample of 185 000 events is selected, with the average electron energy of 11 GeV. The sample therefore provides a useful probe for low-E_{T} electrons, that cannot be studied with $Z \to ee$ decays. The m_{ee} distribution of this sample is modelled as an empirical function taking into account the Gaussian core of the J/ψ peak, a second order polynomial to extract the combinatorial background and the same shape as is used for J/ψ describes the $\psi(2S)$ resonance contribution. Events are categorized in bins of η of the two electrons ($\eta_i\eta_j$), and a PDF $L_{ij}(m_{ee})$ is built for each category. Electron energy scale factors α_i are extracted by using a likelihood function, as in Eq. 5.6 [142]. The resulting α_i factors are compared to the ones obtained from $Z \to ee$ events, mostly showing an agreement within the uncertainties. Agreement is also observed between data and MC.

The energy dependence of the calibration is also studied by categorizing electrons from $Z \to ee$ and $J/\psi \to ee$ in larger bins of η but also categorizing the $Z \to ee$ sample into five E_{T} bins. The procedure described in the first paragraph is repeated and the resulting scale factors all lie within the systematic uncertainty bands of the calibration, as shown in Figure 5.7(a) for central η.

5.1.7 $Z \to ll\gamma$ data-driven scale validation

The validation of the calibration and its corrections for photons, is done using $Z \to ll\gamma$ events with a large-angle separation requirement. The universality of the energy scale factors extracted in Section 5.1.5 is tested for unconverted, one-track and two-track converted photons separately [142].

The $m_{ee\gamma}$ distributions are compared for data and MC for both muons and electrons. Residual

miscalibrations between data and MC are parameterized as

$$E_i^{data} = (1 + \alpha_i) E_i^{MC} \tag{5.7}$$

and the three body invariant mass $m_{ll\gamma}(\alpha_i)$ is recomputed. The agreement between data and MC is quantified as

$$R(\alpha_i) = \frac{\left\langle m_{ll\gamma}(\alpha_i)^{data} \right\rangle / \left\langle m_{ll}^{data} \right\rangle}{\left\langle m_{ll\gamma}^{MC} \right\rangle / \left\langle m_{ll}^{MC} \right\rangle}, \tag{5.8}$$

where $\langle m \rangle$, is the mean value of the invariant mass in the radiative ($ll\gamma$) and non-radiative (ll) case. The α_i providing the best agreement between data and MC ($R(\alpha_i) = 1$) defines the photon energy scale. Systematic uncertainties considered for this method include the background contamination, fit range, electron and muon momentum scale and resolution. The combined photon energy scales compared to the calibration scales agree with the calibration systematic uncertainty bands [142], as shown for unconverted photons as a function of η and E_{T} in Figures 5.7(b) and 5.7(c).

Additional photon-specific uncertainties are assigned to the energy scale calibration due to conversion reconstruction inefficiency and fake conversions, by studying a selection of diphoton events typical for the Higgs search in the diphoton channel [186]. Systematic uncertainties on the $E_{1/2}$ scale and material description were propagated through the analysis and the impact on the energy scale of photons amounts to a few 10^{-4}.

Lateral leakage mismodelling is also considered, and the difference between the description of electrons and photons is evaluated by estimating the normalized difference

$$\Delta(\gamma - e) = \left(\frac{E_{7\times11} - E_{\mathrm{nom}}}{E_{\mathrm{nom}}} \right)^{data} - \left(\frac{E_{7\times11} - E_{\mathrm{nom}}}{E_{\mathrm{nom}}} \right)^{MC} \tag{5.9}$$

for three η intervals separately for converted and unconverted photons. The difference between this quantity for photons with and without the correction for the fraction of conversions described in the previous paragraph is estimated to range from $0.03 - 0.46\%$, η and conversion dependent. A systematic uncertainty is assigned on the photon energy scale as the largest difference between the converted and unconverted $\Delta(\gamma - e)$.

5.2 Summary and evaluation of the refined calibration

Comparing the electron pair invariant mass distribution after applying the calibration to the original and improved MC simulation, as seen in Figure 5.8, a small excess is visible in the low-mass region, it is nevertheless covered with systematic uncertainties of the calibration.

Overall, comparing the new calibration to the old *calibration hits* method used previously, the energy resolution is improved by 10% for unconverted and 20% for converted photons. For electrons, the improvement is between a few percent up to 30% in the region of $1.52 < |\eta| < 1.82$, where the material budget upstream of the calorimeter is the largest [142].

The photon energy scale uncertainty using the MVA calibration is typically 0.2% to 0.3% for $|\eta| < 1.37$ and $|\eta| > 1.82$. For the region of $1.52 < |\eta| < 1.82$, the uncertainty is 0.9% for converted and 0.4% for unconverted photons. The relative uncertainty on the energy resolution is better than 10% for $E_{\mathrm{T}} < 50$ GeV, and asymptotically rises to about 40% at high energies. The energy scale uncertainty model is expected to be valid up to $E_{\mathrm{T}} \approx 500$ GeV [142].

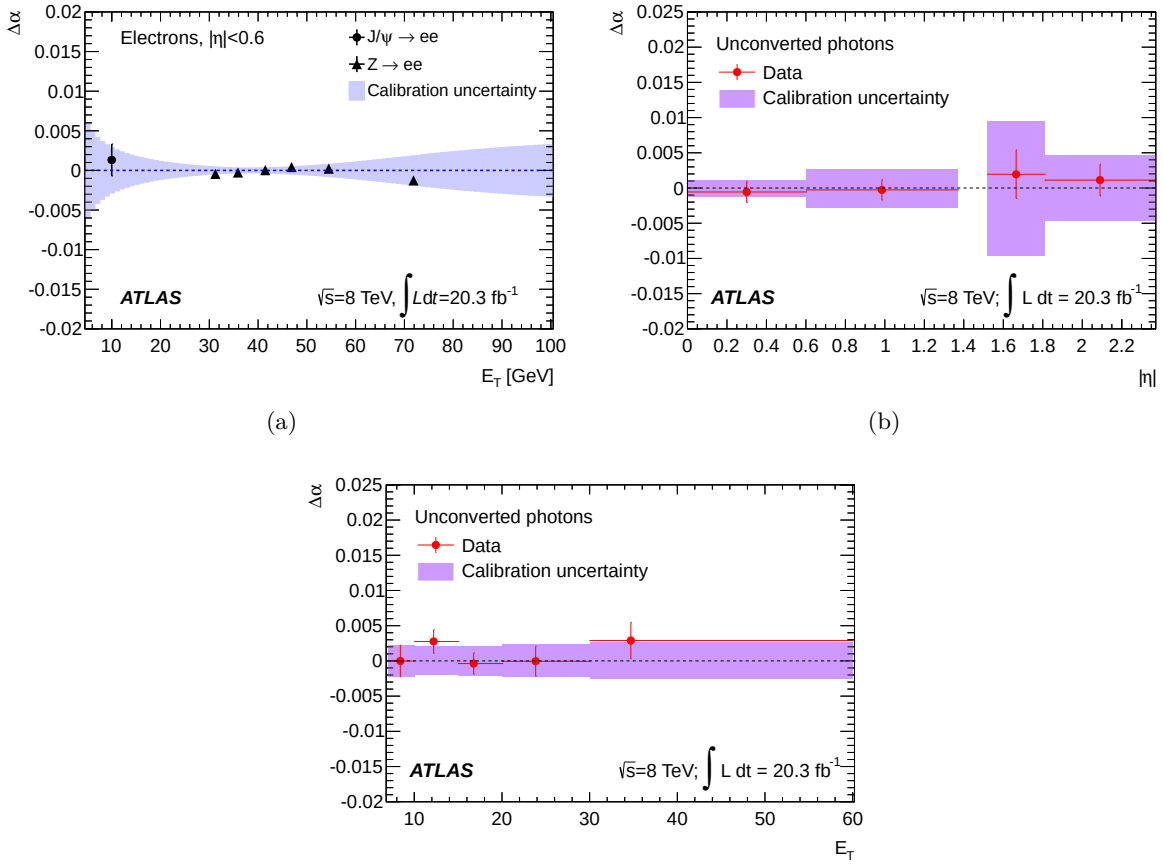


Figure 5.7: (a) Energy scale factors $\Delta\alpha$ obtained after the Z-based calibration from J/ψ and Z analyses [142], (b) combined photon energy scale factors $\Delta\alpha$ obtained after the calibration for unconverted photons as a function of η [142] and (c) E_{T} [142].

5.3 Study of the PS and L1, L2 energy scale using photons

The comparison of the energy deposited by electrons in the first and second layer of the LAr calorimeter (denoted E_1 and E_2, their ratio E_1/E_2 or simply E_{12}) in data and in MC can be sensitive to the amount of material upstream of the calorimeter. In case of a mismodelling of the material budget in the MC, differences can be observed. If there is more (less) material in reality w.r.t. the simulation, the shower can start developing earlier(later) and more (less) energy will be deposited in the first layer in data compared to the MC, the ratio of E_1/E_2 will be larger (smaller). Unconverted photons are insensitive to the material upstream of the calorimeter, and by applying a veto on energy deposits in the PS, one can avoid considering very late conversions with radius $R > 80$ cm.

Inclusive photons from pp collisions at $\sqrt{s} = 7$ TeV from 2011 and 8 TeV from 2012 are selected by requiring the events to pass a GRL selection, a trigger selection (`g120_loose` for 2012 and `g80_loose` for 2011) and to contain at least one primary vertex with at least three tracks.

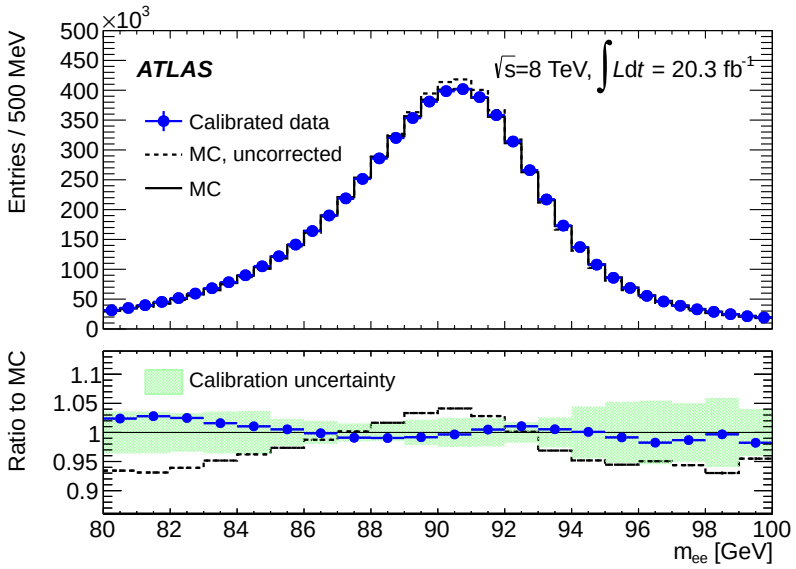


Figure 5.8: Electron pair invariant mass distribution after applying the calibration procedure compared to MC simulation without correction and with correction (top), ratio of the data and uncorrected MC distributions to the corrected MC distribution with the calibration uncertainty band [142].

Photon candidates in events passing these criteria must satisfy further requirements:

- $E_{\mathrm{T}}^{\gamma} > 125$ GeV - to avoid biases introduced by the threshold of the lowest unprescaled trigger;
- $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$;
- photons must satisfy the *tight* selection criteria;
- photons must be isolated with $E_{\mathrm{T}}^{calo,iso} < 4$ GeV;
- photons must be unconverted.

The purity of such a sample of photons should range between 90-99% and is E_{T} dependent [187] for 2012 data and 94-99% for 2011 data [188].

The mean value of the distribution of E_1/E_2 for photon candidates that were reconstructed as unconverted is compared for data and PYTHIA 8 [71] MC samples. The *double ratio* of $\langle E_1/E_2 \rangle$ for data/MC, defined as

$$\frac{\left\langle \frac{E_1}{E_2} \right\rangle^{data}}{\left\langle \frac{E_1}{E_2} \right\rangle^{MC}} \tag{5.10}$$

is shown in Figure 5.9, as a function of η. The difference of the behavior in the two end-caps (first and last five bins) can be quantified as below 2σ, where σ is the statistical uncertainty. By requiring in addition the PS layer energy to be $E_0 < 0.5$ GeV, only unconverted photon

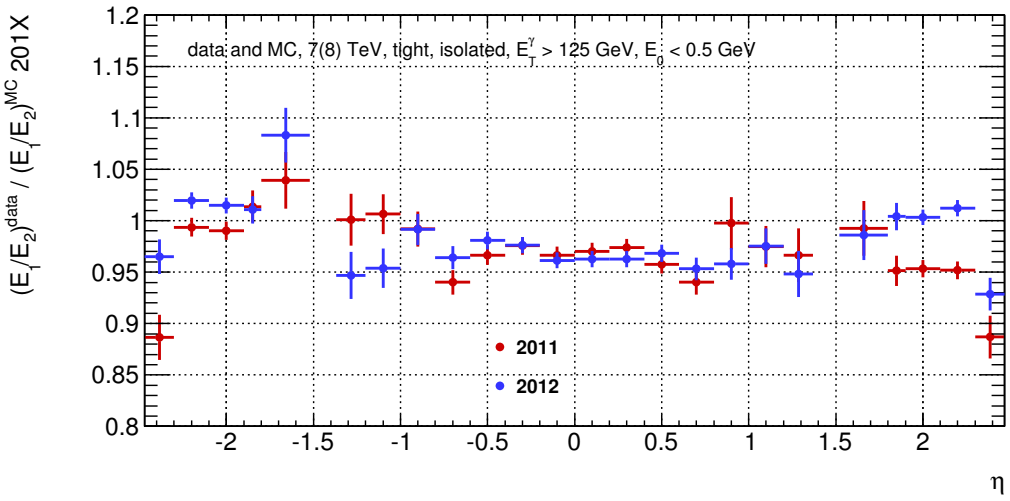
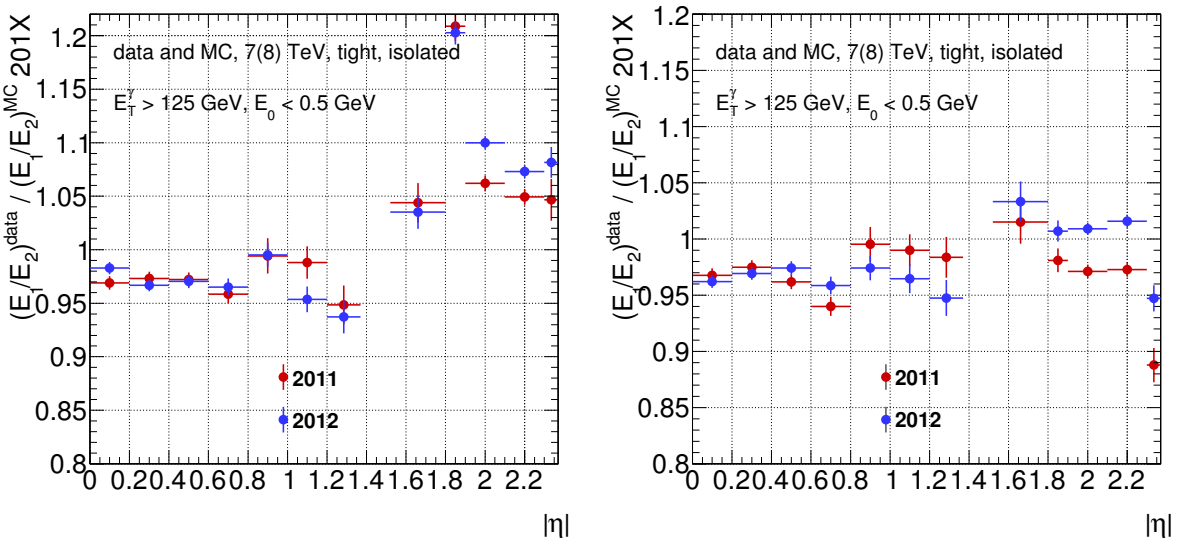


Figure 5.9: $\langle E_1/E_2 \rangle$ ratio for data/MC as a function of η when requiring in addition $E_0 < 0.5$ GeV for 2011 (red) and 2012 (blue) datasets.



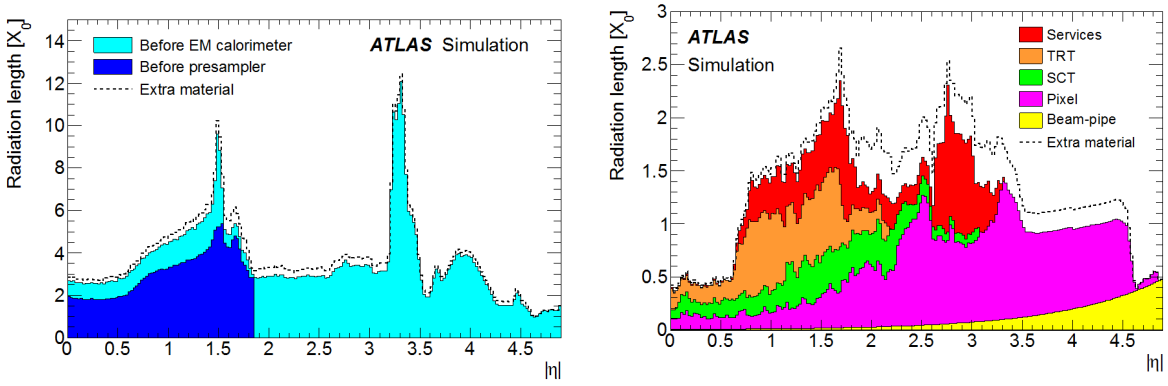
(a) (b)

Figure 5.10: $\langle E_1/E_2 \rangle$ ratio for data/MC as a function of $|\eta|$ when requiring $E_0 < 0.5$ GeV for the (a) old ATLAS geometry GEO-20 used in the *calibration hits* method and (b) for the new ATLAS geometry GEO-21 used in the MVA refined calibration, for 2011 (red) and 2012 (blue) datasets.

candidates for which the shower started very late are selected. These should be insensitive to the mismodelling of the material upstream of the calorimeter.

Figure 5.10 shows the same ratio as a function of $|\eta|$ and for the old (Figure 5.10(a)) and

new (Figure 5.10(b)) geometry used in ATLAS MC simulation, with a refined material budget upstream of the calorimeter as described in Section 5.1.2. In the central part of the calorimeter, the change is small, since the veto on the presampler energy is applied. On the other hand, a visible improvement is seen with the new geometry for the end-caps. Additional material has been added in the end-cap in MC simulation, as can be seen in the Figure 5.11, and the *double* ratio in Figure 5.10(b) is now closer to one with the new geometry.



(a) (b)

Figure 5.11: Amount of material traversed by a particle in X/X_0 for the (a) material up to the ID boundaries and (b) before the EM caloriemeters for the old (GEO-20, solid) and new (GEO-21, dashed) ATLAS geometry [142].

The evolution of the *double ratio* shows a similar trend as in Figure 5.2 measured using muons. The two datasets, collected in 2011 and 2012, show a good agreement, except for the end-cap region of $|\eta| > 1.8$, where a difference of 5% is observed. The same difference was seen studying electrons and it may be related to the change of the electronic calibration, more particularly the OFCs in 2012 w.r.t. 2011 data taking, as described in Chapter 8. A dedicated correction was derived to export the 2012 layer energy scale to the 2011 one.

To estimate the *double ratio* dependence on photon energy, the *double ratios* were evaluated in bins of $|\eta|$ as a function of E^{γ}, shown in Figure 5.12. The presampler energy veto was relaxed to $E_0 < 1$ GeV since the data was split into many bins in η and E_{T}, in order to gain statistical power for the study. The absence of a significant dependence of the *double ratio* on photon energy allows to deduce that energy scale corrections are only η-dependent. Figure 5.12 also shows the ratios for the two geometries and a visible improvement on the data and MC agreement for $|\eta| > 1.8$ when using the new geometry.

Looking at a *triple ratio* of the two layer energies, dividing the *double ratios* of the two datasets 2011/2012, it is sensitive to the calibration only and does not depend on material effects. This ratio is shown in Figure 5.13 for unconverted photon candidates with the presampler energy veto of $E_0 < 0.5$ GeV as a function of η. A clear drop is observed in high-η region (last four bins), which is consistent with what was observed for electrons and muons previously [189]. Comparing the *triple ratio* in bins of E_{T}, as seen in Figure 5.14 no significant E_{T} dependence was found.

The described study helped validate the corrections derived by using muons and electrons also

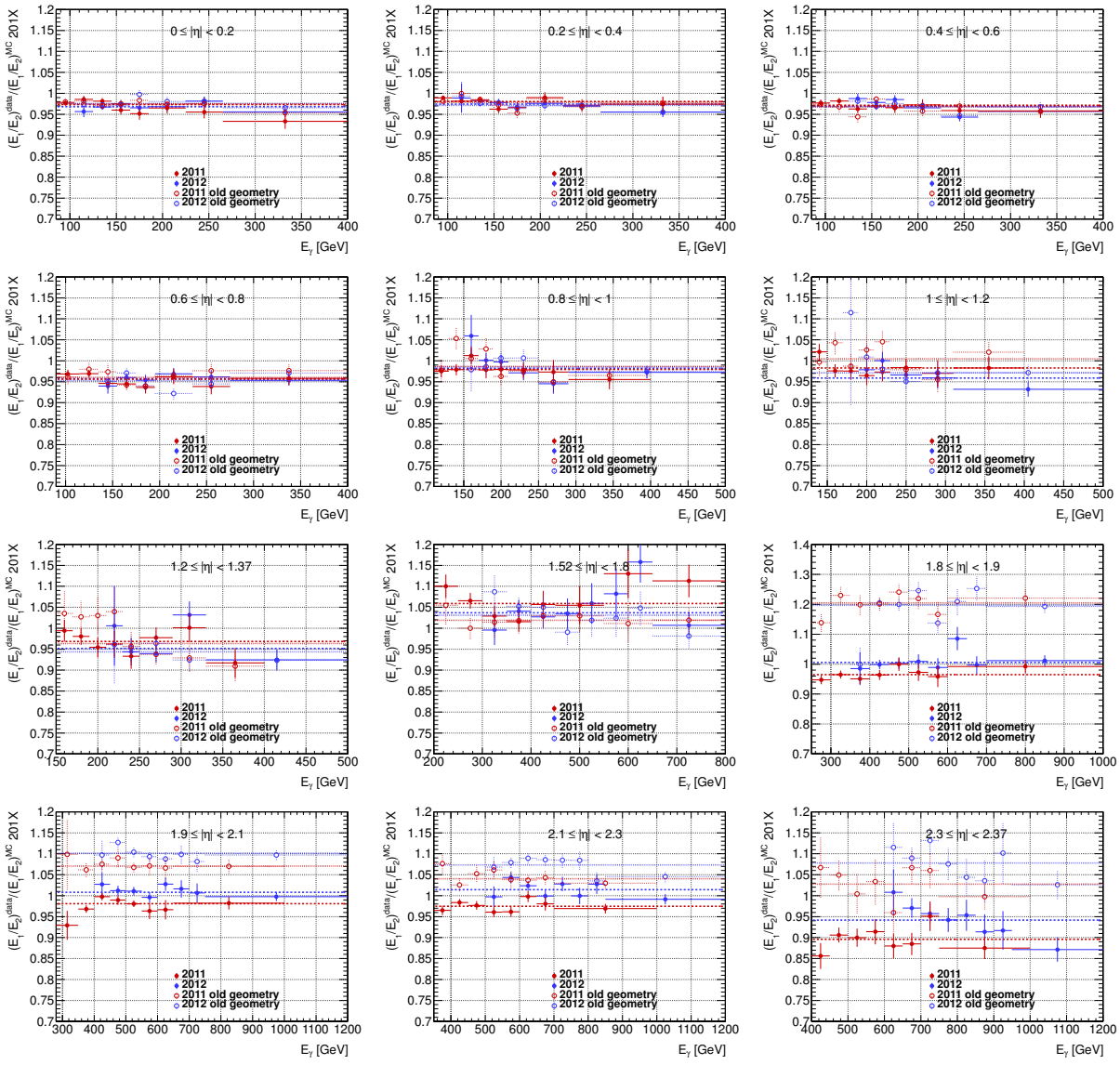


Figure 5.12: $< E_1/E_2 >^{data} / < E_{1/2} >^{MC}$ ratio as a function of photon energy with E_0 <1 GeV cut for the 2011 (red) and 2012 (blue) dataset for GEO-21 (new geometry, full points, dashed line) and GEO-20 (old geometry, open points, dotted line) for different η bins. Each $< E_1/E_2 >$ ratio is fitted with constant respectively.

for photons and validated the E_1/E_2 energy scale derived previously.

5.4 Back layer energy scale

At the energies of the $H \to \gamma\gamma$ decay and the $Z \to ee$, samples of which are used to derive the electron and photon calibration, the energy deposit in the back layer (L3) of the calorimeter is negligible. However, the shower depth increases with energy of the particles, which is why analyses using high energy electrons and photons may leave a no longer negligible energy deposit

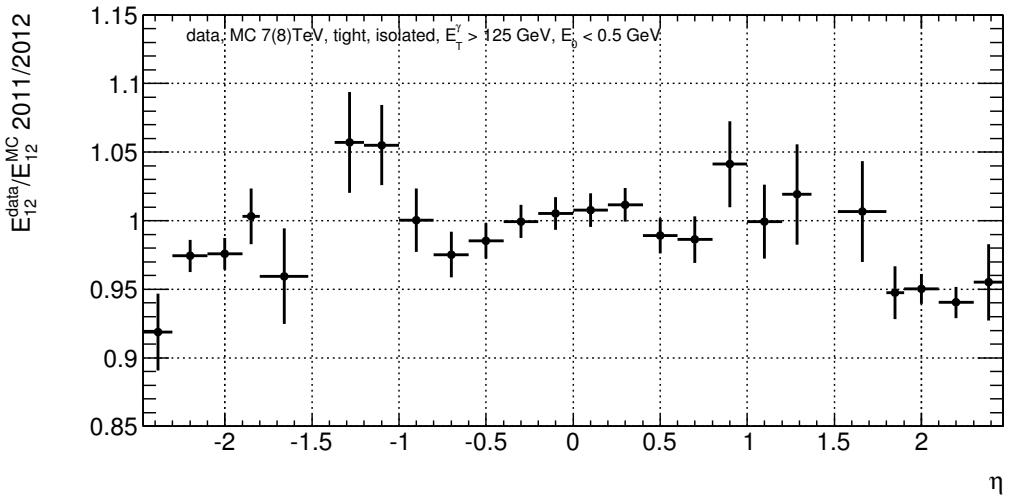


Figure 5.13: Ratio of the $\langle E_{1/2}^{data}\rangle/\langle E_{1/2}^{MC}\rangle$ values obtained from 2011 and 2012 data from unconverted photon candidates with $E_0 < 0.5$ GeV, for the full E_{T} range 125-650 GeV.

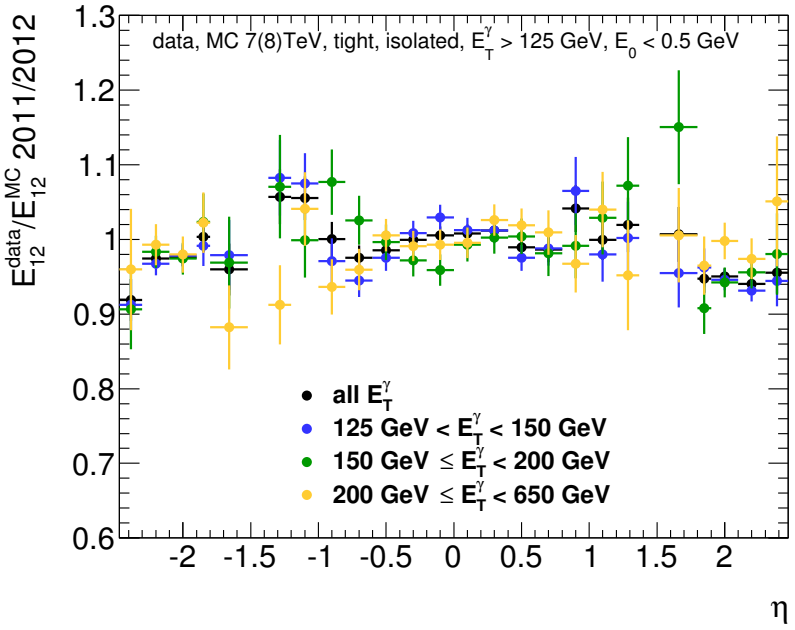


Figure 5.14: Ratio of the $\langle E_{1/2}^{data}\rangle/\langle E_{1/2}^{MC}\rangle$ values obtained from 2011 and 2012 data from unconverted photon candidates with $E_0 < 0.5$ GeV, in bins of E_{T}^{γ}.

in the third layer. As the relative calibration between the response in L1 and L2 might be imperfect, the same could occur for L3, e.g. for non-well modelled cross-talk effects. The consideration of the need for a L3 intercalibration, described in this section, is studied using electrons and a discrepancy between data and MC is observed, ranging from $\approx 10-25\%$ for $|\eta| < 1.8$ but a discrepancy quickly rises up to 350% at high η [189], as seen in Figure 5.15. This effect may be due to material mismodelling in the end-caps, but the passive material has been

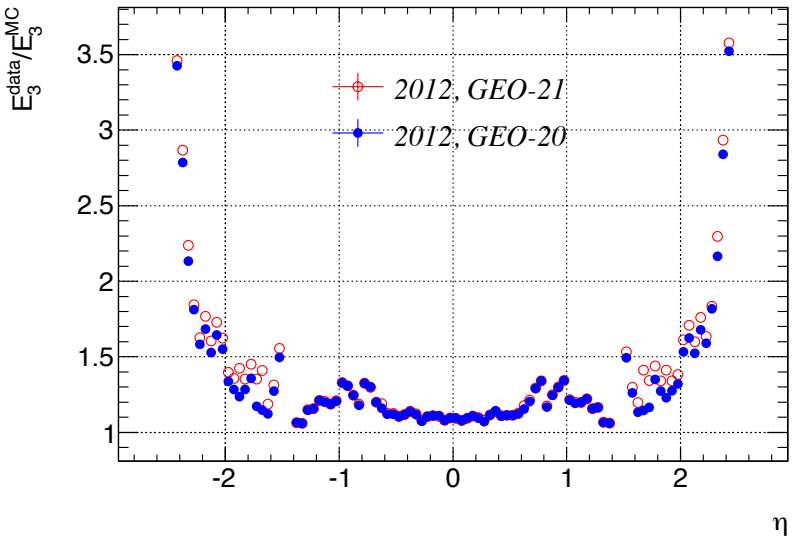


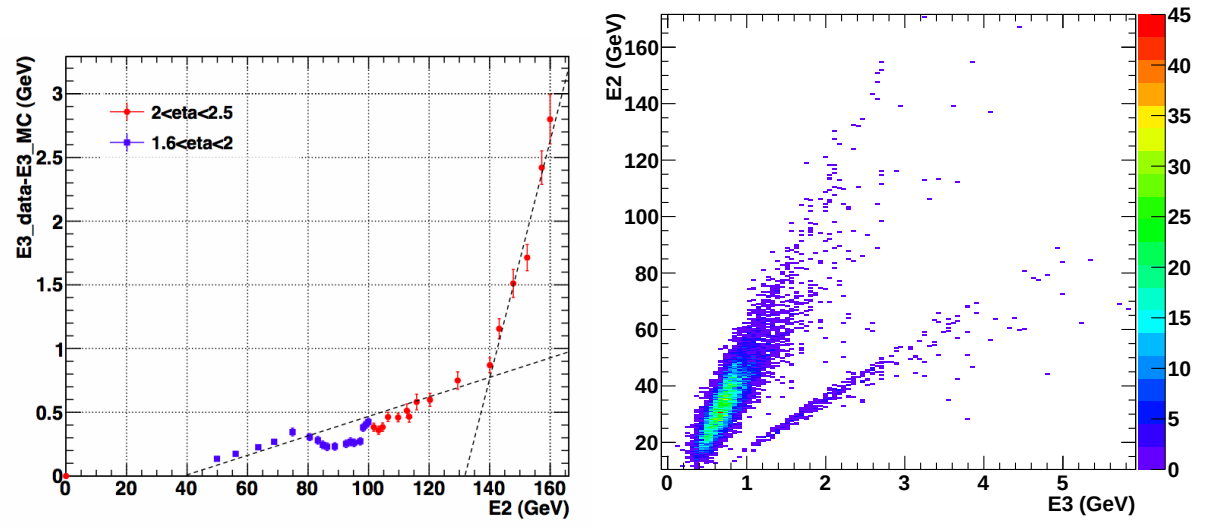
Figure 5.15: E_3^{data}/E_3^{MC} as a function of $|\eta|$ for GEO-20 (blue points) and GEO-21 (red points) derived on an electron sample, for the 2012 dataset [189].

evaluated previously in Ref. [190] using longitudinal shower shapes, and this was not observed. Also, the discrepancy is too large to stem only from this cause.

Other possible sources may be the difference in the electronic calibration or cross talk between the cells from different layers, studied previously in Ref. [191, 192, 193]. A study of L2 and L3 deposited energies on electrons showed two different regimes of correlation between the E_2 and E_3, one present in more events and the other in a smaller fraction of events, as seen in Figure 5.16 [189].

My personal contribution to this study is done looking at unconverted photon candidates. Applying the same selection as described in the previous section, the back layer energy profile is shown in Figure 5.17 for the 2011 and 2012 data and MC. A difference of up to 10-30% increasing with η is observed for 2011 data/MC ratio of the back layer energies, shown in Figure 5.18. For 2012, the data/MC difference is even higher, reaching above 40%, and up to 85 (95)% in the last bin at $2.1 < \eta < 2.37$ for negative (positive) η. These results, produced looking at unconverted photon candidates, show a discrepancy smaller than the study performed on electrons (up to $\approx 350\%$, see Figure 5.15), but still a significant difference is observed. The ratio of 2011/2012 data/MC for L3 energy deposits, shown in Figure 5.19, indicates a difference of up to 30% in the high-η regions of the end-caps between the 2011 and 2012 datasets, which is consistent with studies on electrons, described above. The L3 calibration will certainly need to be reevaluated for $\sqrt{s} = 13$ TeV data taking due to the increased c.m. energy for analyses using high energy photons and electrons.

Since the energy of an electron or a photon is derived from all calorimeter layers, this data/MC discrepancy also influences the final measured energy. A linear correlation is found between the relative bias on the total energy and the relative injected bias in E_3 [189], as seen in Figure 5.20. The relative bias propagated through to the total energy for a central η bin for unconverted photons is shown in Figure 5.20(b), where the bias is smaller than the uncertainty on the calibration up to a few hundred GeV(comparing to Figure 5.7(a)), considering a 10% bias, as



(a) (b)

Figure 5.16: (a) Correlation between the discrepancy of the back layer energy of data w.r.t. MC and the energy in the second layer of the calorimeter in data [189] and (b) correlation between the back and middle layer energy [189].

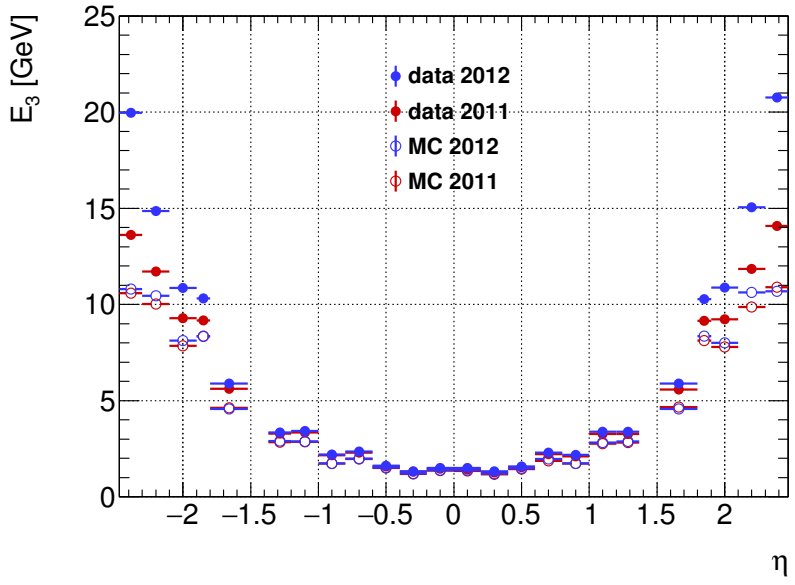


Figure 5.17: Back layer energy E_3 profile as a function of the photon pseudorapidity η for unconverted photons for data (filled circle) and MC (open circle) with the 2011 (red) and 2012 (blue) datasets.

observed in Figure 5.18 for the central η region. The bias on the total energy can be as high as 1% for end-cap regions. The extrapolation of the energy bias to ≈ 1 TeV energies is performed

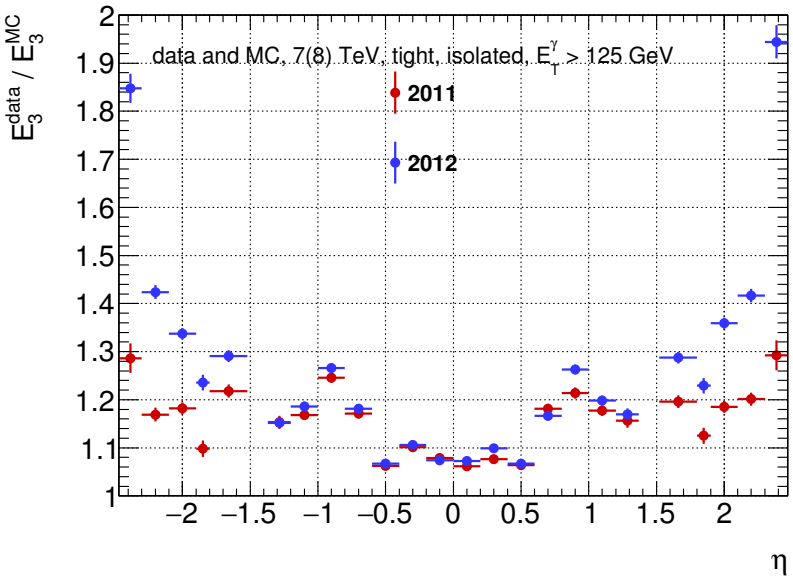


Figure 5.18: Data/MC ratio of the E_3 profile as a function of the photon pseudorapidity for unconverted photons for the 2011 (red) and 2012 (blue) datasets.

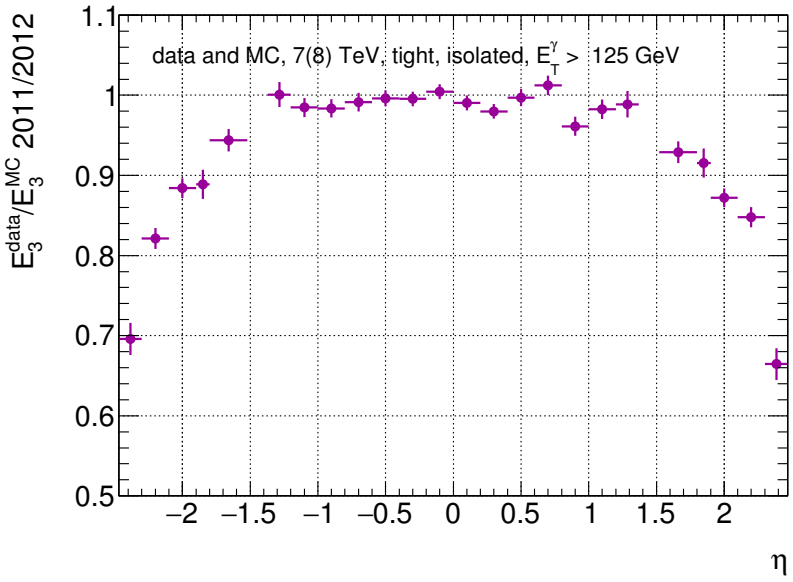
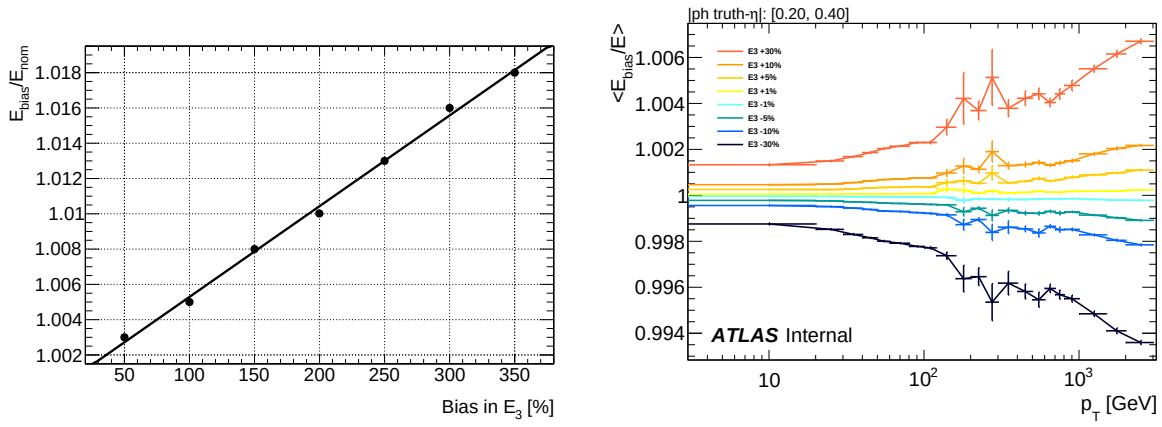


Figure 5.19: $E_3^{data/MC,2011}/E_3^{data/MC,2012}$ profile as a function of the photon pseudorapidity for unconverted photons.

for unconverted photons in the barrel, assuming a 30% bias in E_3 leads to a 0.8% bias on the total energy. For the end-cap, this rises up to 2%. For electrons and converted photons, assuming the same bias in E_3 induces a bias on the total energy of about 0.6% in the barrel and 1.5% in the end-cap [189].



(a) (b)

Figure 5.20: (a) Relative bias in the total energy obtained from a relative change in the E_3 energy for the 2012 dataset and the new geometry GEO-21 [189] and (b) resulting relative bias on the total energy due to a change in E_3 as a function of the unconverted photon energy for $0.2 < |\eta| \leq 0.4$ [189].

5.5 ϕ intercalibration using photons

The electron and photon calibration uses η-dependent factors to correct the energy between MC and data, as in Eq. 5.2. As described in Section 5.1.4, corrections are applied to data as a function of ϕ due to the asymmetries from HV changes and the intermodule widening. To investigate the possible need for a ϕ intercalibration, the following study was performed.

Events are selected from the full set of 2012 pp collisions at $\sqrt{s} = 8$ TeV, by requiring the GRL, no LAr errors (LAr calorimeter operated without any major problems) and at least one primary vertex with at least three reconstructed tracks. The events must also pass one of the photon triggers, namely the prescaled loose single photon triggers with thresholds of $E_{\mathrm{T}}^{\gamma} > 10$, 15, 20, 40, 60, 80, 100 GeV, the unprescaled single loose photon trigger with $E_{\mathrm{T}}^{\gamma} > 120$ GeV and the diphoton triggers, which are not prescaled, the `g35_loose_g25_loose` and `g30_medium_g20_medium`. After the selection, roughly 77 million events are kept. The study is done separately for converted and unconverted photons due to their differences in behavior inside the detector.

Photons in these events are required to pass object quality selection, have $E_{\mathrm{T}}^{\gamma} > 40$ GeV in order to have a photon sample of a reasonable purity. The purity rises with E_{T}^{γ} [165] and was measured to range from 70% to 99% at high E_{T} [187]. A cut on $|\eta| < 2.37$ is placed, excluding the barrel-end-cap transition region, pass the *tight* selection and be isolated with $E_{\mathrm{T}}^{calo,iso} < 4$ GeV. Due to the method used for the study, photons with $E_{\mathrm{T}}^{\gamma} \pm 5\sigma_{E_T}$ around the trigger thresholds are excluded, because they would create bin migrations. After the selection, around 3.5 million unconverted and 2 million converted photons are retained.

The photons are then split into 26 bins in η and 16 bins in ϕ. The η bins are chosen to be the same as the ones used in the η intercalibration, so that a simple correction can be applied to each η bin specifically. Figure 5.21 shows the energy spectrum of photons for the unconverted photon candidates for one bin of η separately for two bins in ϕ (black and green points). The method used for this study is the following: one chooses a bin in η and ϕ as an arbitrary reference, then,

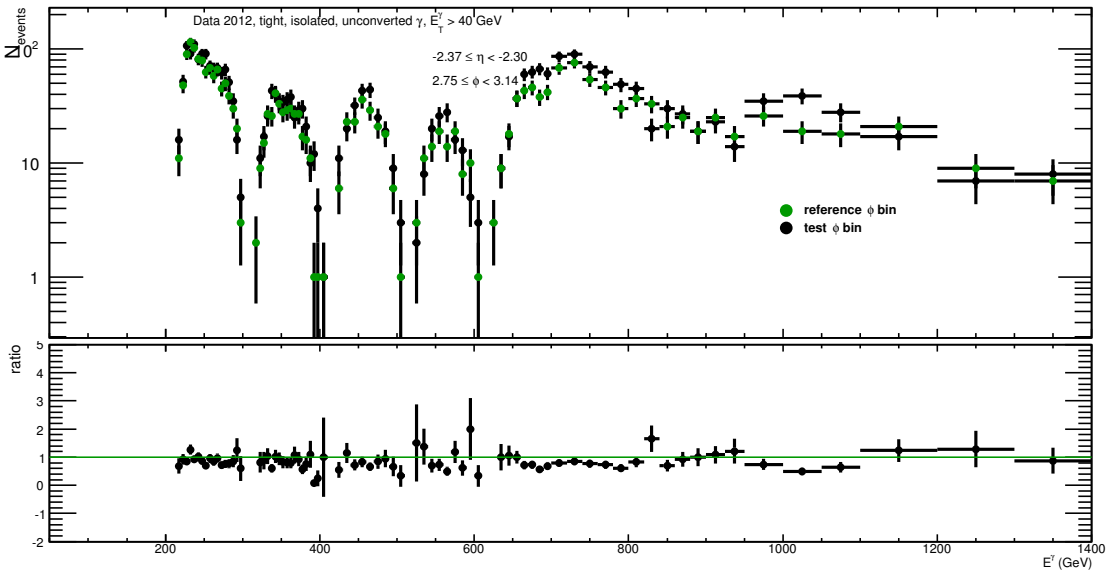
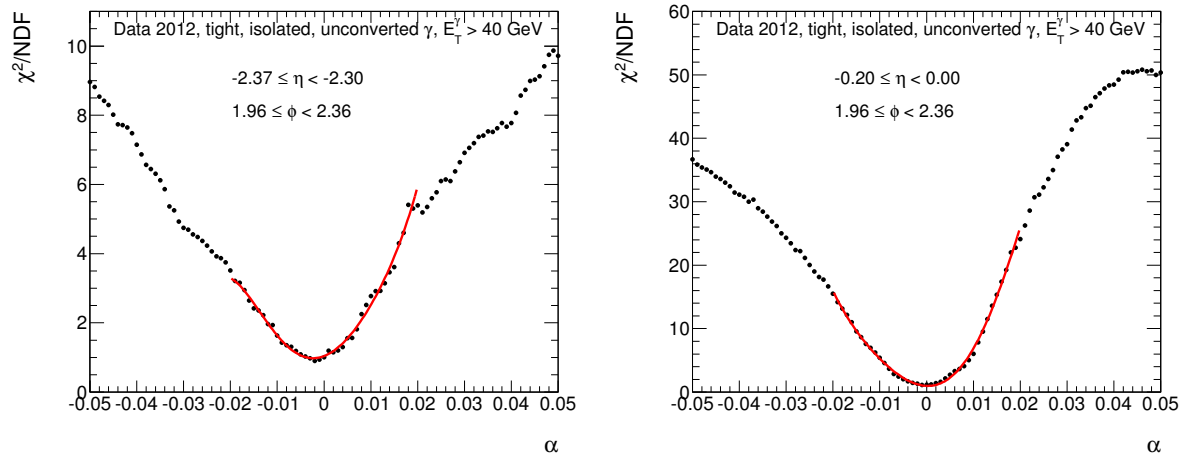


Figure 5.21: Top: Energy spectrum of unconverted photon candidates passing the selection, with $-2.37 \leq |\eta| < -2.30$ and the reference bin $-3.14 \leq \phi < -2.75$ (green points) and the test bin $2.75 \leq \phi < 3.14$ (black points). Bottom: Ratio of the two spectra (black points) and a constant fit (green line).

the ratio of the E^{γ} spectrum of the reference bin to the current test bin is derived and fitted with a straight line.



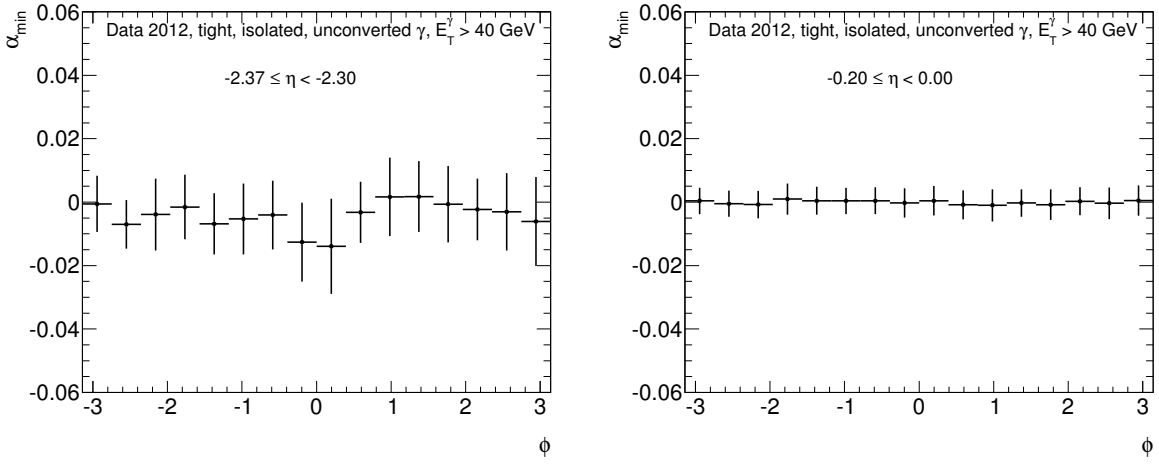
(a) (b)

Figure 5.22: χ^2/NDF of the fit of the photon energy ratio as a function of the α correction factor for unconverted photons of (a) $-2.37 \leq |\eta| < 2.30$ and $1.96 \leq \phi < -2.36$ and (b) $-0.20 \leq |\eta| < 0.00$ and $1.96 \leq \phi < -2.36$.

Any residual non-uniformity in ϕ should manifest itself as a slope in this ratio, an example of

which is shown in the bottom part of Figure 5.21, different from a straight line. By multiplying the energies of each photon by a correction factor $(1 + \alpha)$ in a range of $\pm 5\%$ of the photon energy and repeating the process, one can look at the χ^2/NDF[4] of each fit as a function of the α parameter, as shown in Figure 5.22 for two different bins of η. The χ^2 has a more stable shape with more statistics, which is clearly visible by comparing the two χ^2 distributions for the central η and peripheral case (more events are observed in the central bins). The χ^2 is then fitted using a polynomial of the third order, due to the asymmetric shape and a minimum is found. The value of the α parameter at the minimum - α_{min} is then the ϕ-dependent factor $\alpha(\phi)$ which could be needed to intercalibrate the energy of photons in a specific ϕ bin as

$$E^{\gamma'} = E^{\gamma}(1 + \alpha(\phi)). \tag{5.11}$$



(a) (b)

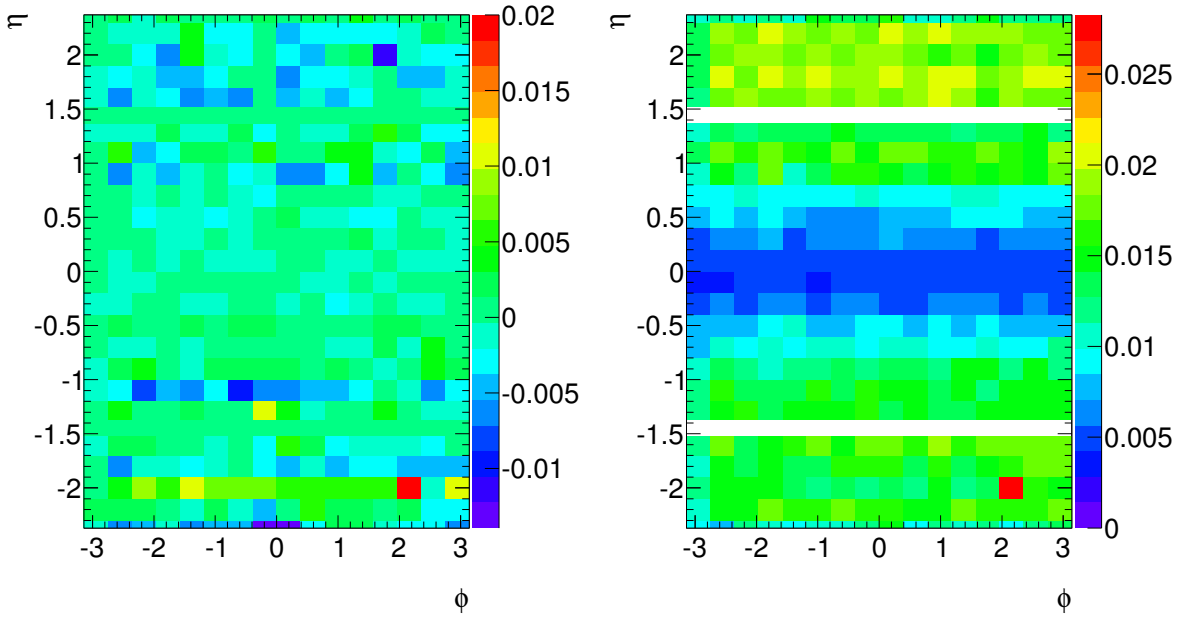
Figure 5.23: α_{min} as a function of ϕ for unconverted photons of (a) $-2.37 \leq |\eta| < 2.30$ and (b) $-0.20 \leq |\eta| < 0.00$.

Since the choice of the reference ϕ bin is arbitrary, a systematic uncertainty can be assigned to it by changing the reference bin. An example of the α factors is shown in Figure 5.23 for two representative η bins for unconverted photons. The error on the α factor is computed as the α interval where the χ^2 changes by one. An $\eta \times \phi$ map of the optimal $\alpha(\phi)$ correction factors, and their uncertainty is shown in Figure 5.24. In order to evaluate the origin of the spread of the correction as a function of ϕ and how much it is due to statistical fluctuations, one can subtract

$$RMS_{\alpha} - \sqrt{|RMS_{\alpha}^2 - \langle \sigma_{\alpha_{min}} \rangle^2|}, \tag{5.12}$$

where RMS is the spread of the α parameters in each bin and $\sigma_{\alpha_{min}}$ is the uncertainty on the α. The result is shown in Figure 5.25, indicating that the statistical power of this study is very limited. The fitted scale factors $\alpha(\phi)$, are compatible with zero and can be attributed to statistical fluctuations. There is therefore no need for additional intercalibration in ϕ. Figure 5.26 shows the same comparing the results for photon energies calibrated using the *calibration hits* method and the MVA calibration, with small differences between the two.

[4]The χ^2 is divided by the number of degrees of freedom (NDF) of the fit.



(a) (b)

Figure 5.24: The $\eta \times \phi$ map of the rescaling factors α_{min} (a) and the corresponding statistical uncertainty (b) for unconverted photons.

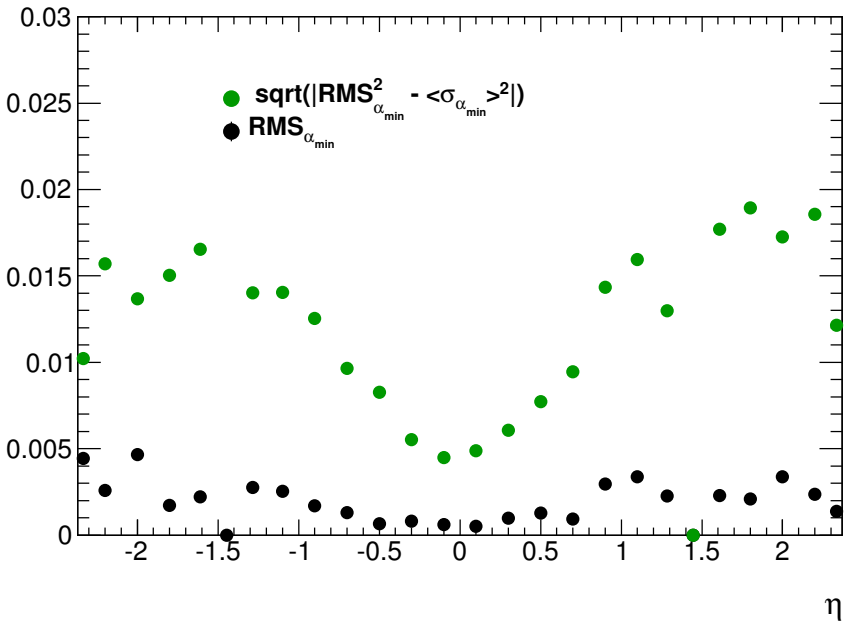


Figure 5.25: Estimation of the statistical power of the rescaling $\alpha_{min}(\phi)$ factors as a function of η.

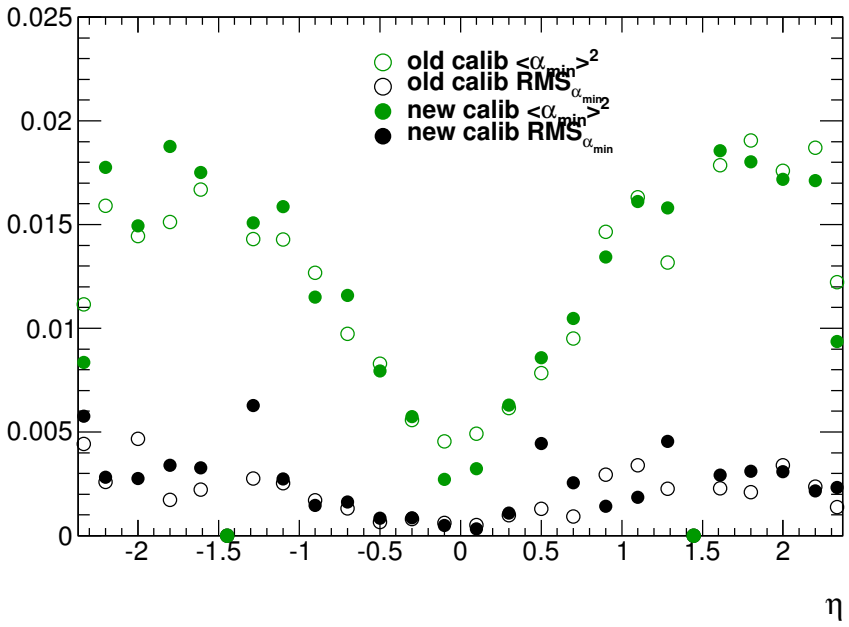


Figure 5.26: Estimation of the statistical power of the rescaling $\alpha_{min}(\phi)$ factors as a function of η for the old *calibration hits* method (open points) and the new MVA calibration (full points).

6 Search for resonances decaying into two photons

Moving on from the discovery of a Higgs boson of mass $m_H \approx 126$ GeV at the LHC, by the ATLAS and CMS collaborations [18, 19], and detailed measurement of its properties [27, 194, 195, 196, 197], the effort has been placed at searching for other resonances decaying into two photons. An extended Higgs sector including additional scalar resonances has been proposed by several models describing physics beyond the Standard Model [198, 199, 200, 201, 202, 203, 204] and as described in Section 1.3.2, searching for these in the diphoton channel is very well motivated.

This chapter describes an analysis that was performed within the group at LAPP, with my personal contribution being the signal modelling and evaluation of related systematic uncertainties. I also took part in the approval process of the analysis within the ATLAS Collaboration. The analysis was made public first as a conference note in June 2014 [205] and subsequently published in Physical Review Letters in August 2014 [206]. A more thorough description of the analysis is presented in the internal document [207].

6.1 Overview of the search for other resonances decaying into diphotons

The search for additional resonances decaying into two photons is based on techniques developed by ATLAS for the search of the SM Higgs boson with the $H \to \gamma\gamma$ decay [18]. The main difference is the extended diphoton invariant mass region of $65 - 600$ GeV compared to the previous searches for the SM Higgs boson, which were restricted to $110 - 160$ GeV in ATLAS, due to previous SM Higgs exclusions by LEP [58] and Tevatron [208] and the fact that the SM Higgs boson branching fraction into two photons falls fast above 160 GeV.

This analysis uses the same photon energy calibration as the ATLAS Higgs boson discovery publication [18], also referred to as the *calibration hits* method in Section 3.2.1. Because of the wide $m_{\gamma\gamma}$ range and the different backgrounds that have to be described while searching for a new resonance, the analysis is performed separately in two regions: a *low-mass* region, covering the masses of $60 - 120$ GeV with sensitivity to resonance masses between $65 - 110$ GeV and a

high-mass region, covering $100-800$ GeV, sensitive to resonance masses between $110-600$ GeV. The two regions overlap, in order to provide a continuous coverage of the mass of the new resonance m_{X}, since side-bands are needed on both sides of a tested mass point in order to be able to fit the data.

The search considered narrow resonances only, and is valid for resonances of a width negligible with respect to the diphoton invariant mass resolution in ATLAS. This will be further addressed in Section 6.7.3. Since there can be many models predicting a resonance decaying into two photons, the result aims to be as model-independent as possible, taking into account a wide range of final states when modelling the signal and unfolding the detector effects (see Section 6.5). Interference of the new resonance with the diphoton continuum is not considered.

The result of this analysis is presented in the form of a fiducial limit on the production cross section times branching fraction of the new resonance decaying into two photons. The analysis is model independent and interpretation of the result in term of any model is left for theorists.

The search for an additional resonance in the diphoton decay channel presented several challenges:

- **Diphoton continuum background:** SM photon pair production invariant mass spectrum in the studied range after the applied cuts (described in Section 6.4.3) has a falling exponential shape but finding a suitable functional form to describe it is not straightforward due to e.g. the treatment of *spurious signal* (defined in Section 6.4.3).
- **Trigger turn-on:** all events must pass a diphoton trigger, which is not 100% efficient, especially close to its E_{T}^{γ} thresholds. In the diphoton invariant mass distribution, the trigger turn-on shape distorts the diphoton continuum spectrum at the beginning of the studied range, at around $m_{\gamma\gamma} = 60$ GeV, as can be seen in Figure 6.1 for different photon E_{T} cuts on MC Simulation events. A function describing this shape needed to be chosen appropriately.
- **Drell-Yan (DY) background:** the resonant (Z boson peak) and non-resonant DY production of electrons misidentified as photons has a continuous falling shape with a peak at the Z mass at around $m_{\gamma\gamma} \approx 91$ GeV. A model describing this background needed to be derived, estimating the shape and the number of events, where electrons are misidentified as photons, especially in the *low-mass* region, where this contribution becomes prominent.
- **Higgs boson background:** the peak of the newly discovered Higgs particle of $m_H = 126$ GeV, with the full 2012 data, has to be added on top of the continuum background.
- **New resonance:** the new resonance peak has to be described by a shape evolving with $m_{\gamma\gamma}$ over a very wide $m_{\gamma\gamma}$ range, but stay model-independent at the same time.

Splitting the analysis into two $m_{\gamma\gamma}$ regions allows to better fit the background and introduce the optimal selection criteria for each mass range in order to improve the sensitivity to a new resonance. Due to the DY background being more prominent when one or even both photons are converted (appearing more like electrons, due to the presence of a track in the inner detector), the *low-mass* analysis is performed in categories of the number of conversions $N_{conv} = 0, 1, 2$. A blind analysis method is used throughout the analysis in order to stay unbiased, meaning all studies are done on MC samples and data is reviewed only after the whole analysis chain had been established, tested and finalized.

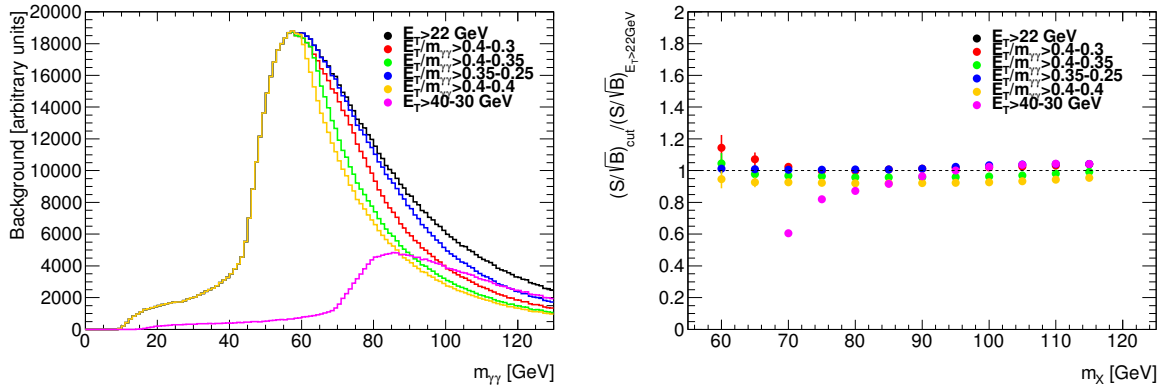


Figure 6.1: (a) The diphoton invariant mass spectrum in the low mass part after applying the trigger selection and various sets of E_{T} cuts applied on the two photon candidates [207]; and (b) Ratio between $S/\sqrt{B}$ with the tested cuts and $S/\sqrt{B}$ with the 22/22 GeV cuts [207].

6.2 Event Selection

The search for other resonances decaying into two photons is performed using the full 2012 dataset of pp collisions collected by ATLAS at $\sqrt{s} = 8$ TeV. When selecting events, the presence of two reconstructed photon candidates is required, both satisfying the tight identification criteria based on the EM shower shapes [162]. Both photon candidates have to be within the fiducial region of the calorimeter with $|\eta^{\gamma}| < 2.37$ excluding the transition region between the barrel and the end-cap calorimeters, located at $1.37 \leq |\eta| < 1.56$. The photon identification efficiency when applying the tight criteria is typically 85% at low E_{T}^{γ} and reaches a plateau of 95% at around $E_{\mathrm{T}} = 100$ GeV [164].

Furthermore, both photon candidates are required to be isolated, using the isolation criteria based on information from the calorimeter and inner detector. The track-based isolation (Section 3.4.2) transverse momentum (from now on called *track-based* isolation) is defined as the scalar sum of the transverse momenta of all tracks with $p_{\mathrm{T}} > 1$ GeV in a cone of size $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.2$ around each photon, and it is required to be smaller than 2.6 GeV. Only tracks originating from the diphoton vertex (as defined below) are used, while tracks associated to converted photon candidates are excluded from the sum.

The calorimeter isolation transverse energy (from now called *calorimeter-based* isolation) is defined as the transverse energy sum of the topological clusters with positive energy [167] reconstructed in the calorimeter around each photon in a cone of radius $\Delta R = 0.4$. The transverse energy sum excludes the energy deposits in the cluster region of the photon shower. In addition, corrections for the small energy leakage of the photon shower outside the cluster region are applied. The effects of the underlying event and additional minimum bias interactions (pile-up) occurring in the same or neighboring bunch crossings are corrected for, on an event-by-event basis [168, 209]. The value of the calorimeter isolation is required to be smaller than 6 GeV[1].

To improve the precision of the measurement of the diphoton invariant mass $m_{\gamma\gamma}$ and the computation of track-based quantities, such as the track isolation, the most precise location of the

[1]This will be reinvestigated in Section 6.2.2.2.

diphoton vertex is needed [210]. The diphoton vertex is determined using the *photon pointing* method, where the vertex position along the beam axis is obtained by combining trajectories of both photons (measured using the longitudinal segmentation of the calorimeter) with constraints from the average beam spot position and the tracker information. In case the photon has converted into an e^+e^- pair before reaching the calorimeter, and the conversion tracks have hits in the silicon detectors, the conversion vertex is used in the computation. For each vertex, the sum of the squared momentum $\sum p_{\mathrm{T}}^2$, the scalar sum of the momentum $\sum p_{\mathrm{T}}$ of all tracks associated to the reconstructed vertex, and the difference in the azimuthal angle $\Delta\phi$ between the 4-vector of the summed tracks momenta and the 4-vector of the diphoton system is defined. The best diphoton vertex is selected from all reconstructed vertices using a Neural Network (NN) alogorithm, combining the diphoton primary vertex z position estimated by the photon pointing, its uncertainty and track information [211].

The efficiency of the vertex selection was studied with $Z \to ee$ events in data and MC, by removing the electron tracks from the events and verifying the efficiency of finding the vertex associated to them. Corrections are applied in order to mimic the Higgs boson signal, by matching the dielectron and diphoton p_{T} spectra in the MC and the fraction of events with zero, one or two EM objects in the barrel region of the calorimeter. The efficiency of finding a primary vertex within 0.3 mm of the true one is higher than 75% [207].

The photon energy calibration used for this analysis is described in detail in Section 3.2.1 and is done separately for converted and unconverted photons.

The selection described above, together with E_{T}^{γ} cuts for the photons, was developed for the SM Higgs search in ATLAS and is therefore re-investigated for this analysis.

6.2.1 Low-mass $60 < m_{\gamma\gamma} < 120$ GeV selection

In order to be able to look at the lowest possible $m_{\gamma\gamma}$ events and considering the available ATLAS triggers, the photon transverse energy E_{T} selection is revisited. Instead of using the diphoton trigger requiring two photon candidates passing $E_{\mathrm{T}}^{\gamma_1} > 35$ GeV and $E_{\mathrm{T}}^{\gamma_1} > 25$ GeV and passing loose shower shape criteria (labeled as `EF_g35_loose_g25_loose`) used for the $H \to \gamma\gamma$ analysis [18], another unprescaled trigger is used, requiring two photon candidates, both passing $E_{\mathrm{T}}^{\gamma} > 20$ GeV and medium shower shape criteria (labeled as `EF_2g20vh_medium`), allowing the measurement to start using photons with $E_{\mathrm{T}}^{\gamma} > 22$ GeV, and therefore extending the $m_{\gamma\gamma}$ reach to 60 GeV.

At lower $m_{\gamma\gamma}$, the background contributions get larger when releasing the E_{T}^{γ} cuts. A study of the signal over background ratio using several different relative $E_{\mathrm{T}}^{\gamma}/m_{\gamma\gamma}$ cuts is performed, displayed in Figure 6.1, showing an up to 15% improvement of the expected significance $Z = S/\sqrt{B}$ using $E_{\mathrm{T}}^{\gamma_{1(2)}}/m_{\gamma\gamma} > 0.4(0.35)$ and $0.4(0.4)$, but also a distortion of the $m_{\gamma\gamma}$ spectrum, as seen in Figure 6.1. This would make it more challenging to find a function suitable for the fit of the *low-mass* $m_{\gamma\gamma}$ spectrum and so the $E_{\mathrm{T}}^{\gamma} > 22$ GeV cuts are chosen.

6.2.2 High-mass $100 < m_{\gamma\gamma} < 800$ GeV selection

6.2.2.1 Relative E_{T}^{γ} cuts

In previous SM Higgs searches in the diphoton channel, relative $E_{\mathrm{T}}^{\gamma}/m_{\gamma\gamma}$ cuts of 0.35 and 0.25 on the leading and subleading photon candidate are applied to maximize the signal to background ratio in the mass window of $110 < m_{\gamma\gamma} < 160$ GeV [212]. For this analysis, the search is performed up to 800 GeV and the study is performed again, scanning relative cut values between 0.1 and 0.5, using SM Higgs-like, produced via the gluon-fusion process (ggF) MC samples modelled with a narrow width of 4 MeV for X mass values ranging between 160 and 1000 GeV. Relative $E_{\mathrm{T}}^{\gamma}/m_{\gamma\gamma}$ cuts are chosen at 0.4 and 0.3 for the leading and subleading photons, respectively, giving an improvement in significance of up to 55% w.r.t absolute E_{T}^{γ} cuts of 40 and 30 GeV, which were used for the SM Higgs analyses [18].

6.2.2.2 Photon isolation

The efficiency of the track-based isolation is fairly stable with E_{T}^{γ}, but for the calorimeter-based isolation, a large dependence is observed, associated to the photon cluster leakage overcorrection (described in Section 3.2) and data to MC differences. To ensure a stable efficiency with E_{T}^{γ}, one can either introduce an E_{T}-dependent calorimeter-based isolation cut or correct the calorimeter isolation energy itself. In order to stay consistent with the SM Higgs analysis in ATLAS and to maintain a constant efficiency at higher E_{T}^{γ} values, a correction to the isolation energy is derived, only to be applied to photons with $E_{\mathrm{T}}^{\gamma} > 80$ GeV, which corresponds to $m_{\gamma\gamma}$ values above 160 GeV. The corrected isolation energy is defined as

$$E_{\mathrm{T}}^{calo,isocorr} = \begin{cases} E_{\mathrm{T}}^{calo,iso} - 0.7\%(E_{\mathrm{T}}^{\gamma} - 80 \text{ GeV}), & \text{for } E_{\mathrm{T}}^{\gamma} > 80 \text{ GeV} \\ E_{\mathrm{T}}^{calo,iso}, & \text{for } E_{\mathrm{T}}^{\gamma} \leq 80 \text{ GeV} \end{cases} \tag{6.1}$$

and a smaller than 6 GeV isolation requirement is applied. A cut on the track isolation is applied as $E_{\mathrm{T}}^{track,iso} < 2.6$ GeV, same as in the $H \to \gamma\gamma$ analysis [18].

6.2.3 Summary of the selection criteria

A summary of the selection cuts applied in the two analyses is shown in Table 6.1. The total signal selection efficiency as a function of the resonance mass m_{X} increases from 25% to 43% for the *low-mass* analysis and from 30% to 40% for the *high-mass* analysis. A dependence of the total efficiency on the SM Higgs production process is observed, mainly due to the difference in the kinematic properties of the photons, which will be further addressed in Section 6.5.

The diphoton triggers used for the *low-mass* and the *high-mass* part of the analysis are unprescaled. They require a photon candidate passing a corresponding E_{T} threshold and loose or medium shower shape criteria. The efficiency of the `EF_2g20vh_medium` trigger was measured to be $98.66\%^{+0.14}_{-0.16}(stat) \pm 0.10(syst)$ [213]. The efficiency of the `EF_g35_loose_g25_loose` trigger was measured as $99.61\%^{+0.06}_{-0.07}(stat) \pm 0.5(syst)$ [213].

After applying the standard data quality requirements (all systems operational, data collected and recorded properly) on the data sample of $\sqrt{s} = 8$ TeV data collected in 2012, it corresponds to a total integrated luminosity of 20.3 ± 0.57 fb^{-1}. With the selection criteria, described previ-

Selection	**Low-mass analysis**	**High-mass analysis**
trigger	`EF_2g20vh_medium`	`EF_g35_loose_g25_loose`
E_{T}^{γ}	$E_{\mathrm{T}}^{\gamma_{1,2}} > 22$ GeV	$E_{\mathrm{T}}^{\gamma_1}/m_{\gamma\gamma} > 0.4$ $E_{\mathrm{T}}^{\gamma_2}/m_{\gamma\gamma} > 0.3$
pseudorapidity	$\lvert\eta^{\gamma}\rvert < 2.37$, excluding $1.37 < \lvert\eta^{\gamma}\rvert < 1.56$	
calo-based isolation	$E_{\mathrm{T}}^{iso,calo} < 6$ GeV	same, $E_{\mathrm{T}}^{iso,calo}$ dependent for $E_{\mathrm{T}}^{\gamma} > 80$ GeV
track-based isolation	$E_{\mathrm{T}}^{iso,track} < 2.6$ GeV	$E_{\mathrm{T}}^{calo,track} < 2.6$ GeV
$m_{\gamma\gamma}$ region	$60 < m_{\gamma\gamma} < 120$ GeV	$m_{\gamma\gamma} > 100$ GeV
# of events	589 312	108 682

Table 6.1: Summary of the *low-mass* and *high-mass* analysis event selection.

ously, the *low-mass* region contains 589 312 collision events and the *high-mass* region contains 108 682 collision events.

6.3 Signal modelling

Even though the search for a narrow resonance has been limited to events with $m_{\gamma\gamma} < 600$ GeV, due to lack of events in the side-bands for a reliable final fit, the signal modelling is performed up to 1 TeV, because the MC samples used to model the resonance were available.

As a baseline, ggF Higgs-like MC samples are used to study the diphoton invariant mass shape following the detector response for each analysis selection. Since this search aimed to be as model independent as possible, different production modes leading to varying final state characteristics are explored. MC samples with events produced by the other SM Higgs production processes (namely VBF, ttH, ZH and WH) are studied, evaluating the systematic uncertainties, which is described in Section 6.3.4.The ggF MC samples were produced using POWHEG [214, 215] interfaced with PYTHIA8 [71] at several m_{X} points. Starting from $m_{\mathrm{X}} = 70$ GeV, in 5 GeV steps up to 150 GeV, the samples were produced using the SM Higgs width[2]. Above 150 GeV, the samples were produced using the SM Higgs width at 125 GeV, which is $\Gamma = 4.07$ MeV, and the steps between the mass points grow larger with m_{X}, as illustrated in Figure 6.2. The use of the 4 MeV width for these mass points will be further addressed as the Narrow Width Approximation (NWA). The validity of the final result for resonances of various natural widths is discussed in Section 6.7.3.

For the ggF samples, the Higgs boson production cross sections are computed up to next-to-next-to-leading order (NNLO) in QCD [216, 217, 218, 219, 220, 221]. Soft gluon resummation up to next-to-next-to-leading logarithm order is adopted [222]. The finite quark-mass effects are taken into account in POWHEG [223]. Next-to-leading order (NLO) EW corrections are applied [224, 225]. These results are compiled in [226, 227, 228] assuming factorization between QCD and EW corrections.

[2]These samples were already available due to the previous SM Higgs searches in the diphoton channel, and since their width is much smaller than the experimental resolution at those masses, the samples could also be used for this analysis.

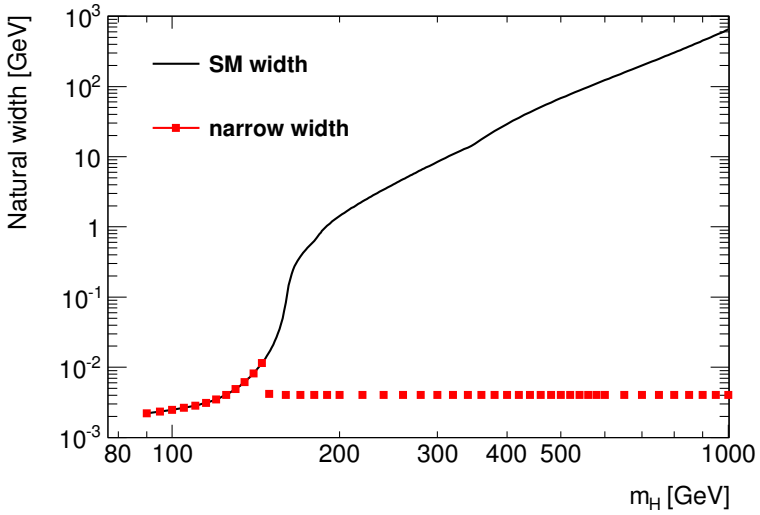


Figure 6.2: Width of the Higgs boson used in the SM Higgs samples (up to $m_H = 150$ GeV) and NWA Monte Carlo samples (for $m_H > 150$ GeV, as a function of the Higgs boson mass (red squares) compared to the natural width of the SM Higgs boson (black line) [207].

6.3.1 Definition of the double-sided Crystal Ball function

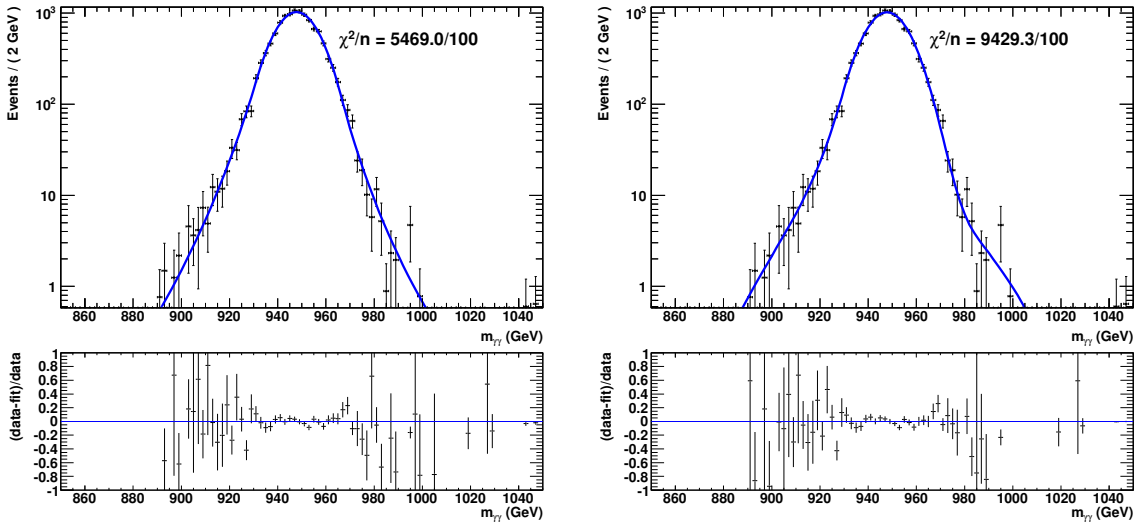
In the search for a SM Higgs boson decaying into two photons, the signal shape is modelled as the sum of a Crystal Ball function and a Gaussian distribution [110]. However, for large values of the resonance mass m_{X}, this function does not properly describe the shape of the high mass diphoton tail. A different and more suitable fit function is therefore used in this analysis - the double-sided Crystal Ball (DSCB), to properly take into account asymmetric and non-Gaussian low and high mass tails that can occur due to signal leakage outside of the LAr calorimeter and overcorrections at high E_{T}, respectively. A fit of the $m_{\gamma\gamma}$ distribution using this function is shown in Figure 6.3(a) compared to the Crystal ball plus Gaussian fit in Figure 6.3(b). The global fit χ^2 is better, and the residuals appear smaller in the high mass tail.

The signal modelling is performed separately for the categorized *low-mass* and the inclusive *high-mass* analyses. Figure 6.4 shows three DSCB function fits of low mass samples, to illustrate that the function properly describes the $m_{\gamma\gamma}$ distribution also at low masses.

The double-sided Crystal Ball function is defined as

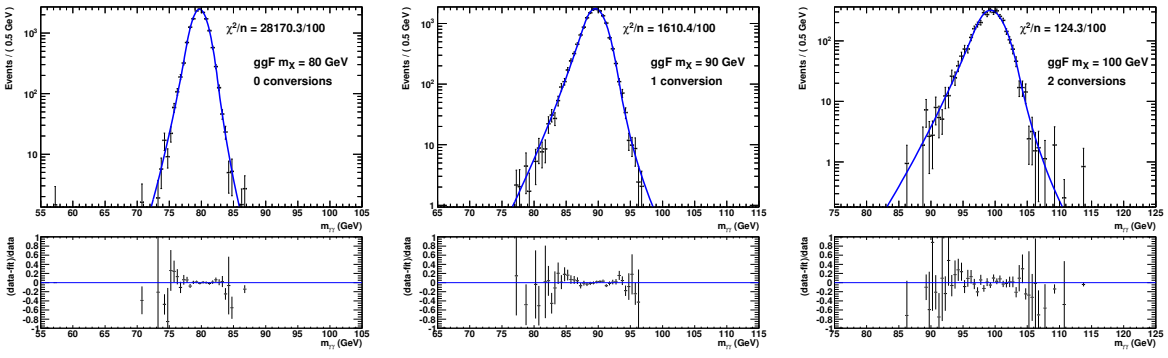
$$N \cdot \begin{cases} e^{-t^2/2} & \text{if } -\alpha_{Low} \geq t \geq \alpha_{High} \\ \frac{e^{-0.5\alpha_{Low}^2}}{\left[\frac{\alpha_{Low}}{n_{Low}}\left(\frac{n_{Low}}{\alpha_{Low}} - \alpha_{Low} - t\right)\right]^{n_{Low}}} & \text{if } t < -\alpha_{Low} \\ \frac{e^{-0.5\alpha_{High}^2}}{\left[\frac{\alpha_{High}}{n_{High}}\left(\frac{n_{High}}{\alpha_{High}} - \alpha_{High} + t\right)\right]^{n_{High}}} & \text{if } t > \alpha_{High}, \end{cases} \tag{6.2}$$

where $t = \Delta m_{\mathrm{X}}/\sigma_{CB}$, $\Delta m_{\mathrm{X}} = \mu_{CB} - m_{\mathrm{X}}$, N is a normalization parameter, μ_{CB} is the peak of the Gaussian distribution, σ_{CB} represents the width of the Gaussian part of the function, α_{Low} (α_{High}) is the point where the Gaussian becomes a power law on the low (high) mass side, n_{Low} (n_{High}) is the exponent of this power law. An illustrative drawing of the double-sided Crystal



(a) (b)

Figure 6.3: Fit of the $m_{\gamma\gamma}$ distribution for the ggF $m_X = 950$ GeV sample, using a (a) double-sided Crystal Ball function and (b) Crystal Ball plus Gaussian. The bottom plots show the normalized residuals.



(a) (b) (c)

Figure 6.4: Fit of the $m_{\gamma\gamma}$ distribution for ggF (a) $m_X = 80$ GeV sample for the 0 conversions category, (b) $m_X = 90$ GeV sample for the 1 conversion category and (c) $m_X = 100$ GeV sample for the 2 conversions category.

Ball function is provided in Figure 6.5. The Gaussian core connected to the two power laws in the tails give a continuous function as well as its derivative.

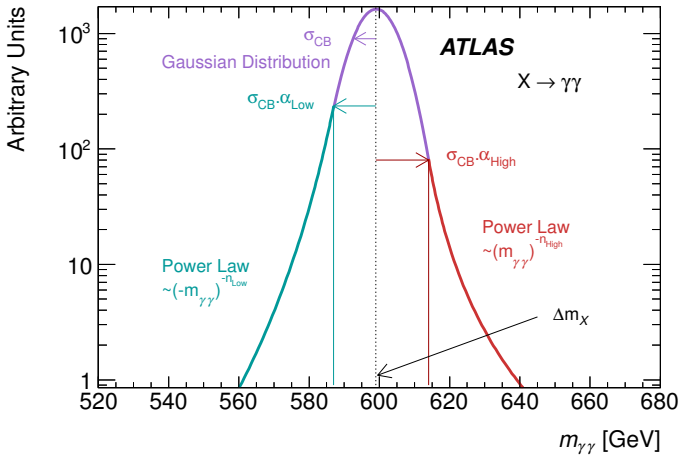


Figure 6.5: Illustration of the double-sided Crystal Ball function with its six parameters, for a signal mass $m_X = 600$ GeV.

6.3.2 Signal modelling at *high-mass* - inclusive analysis

In a first step, an unbinned fit of the $m_{\gamma\gamma}$ distribution of all the events passing the selection cuts in the ggF MC samples is performed using the double-sided Crystal Ball function, separately for each mass point m_X (single mass point fit). The evolutions of the DSCB parameters as a function of m_X are then fitted to extract parameterizations. As a second step, the parameterization functions of the DSCB parameters are used as input for a binned multiple mass point fit, where all the mass points are fitted simultaneously (simultaneous fit). The floating parameters of the fit are now the coefficients of the parameterization functions listed in Table 6.2. The n_{Low} and n_{High} parameters are also left free in the simultaneous fit, but since no particular trend is visible with mass in the single mass point fits, n_{Low} and n_{High} are parameterized with a constant.

Figure 6.6 shows the results of the simultaneous fit to the ggF samples, and the normalized residuals. Figure 6.7 shows the output parameterizations of the DSCB parameters from the simultaneous fit (pink line) compared to the single mass point fit results (blue points fitted with a blue line) where n_{Low} and n_{High} are fixed to the values obtained from the simultaneous fit. A very good agreement is found between both methods. The final parameterizations obtained from the simultaneous fit are given in Table 6.2, and are used for the *high-mass* analysis.

6.3.3 Signal modelling at *low-mass* - conversion categories

The simultaneous fit is repeated for the *low-mass* region, where the ggF samples are split into three subsamples according to the conversion status of the diphoton system: 0-converted, 1-converted and 2-converted. Figure 6.8 shows the evolution of the event fraction in each category as a function of m_X.

Figure 6.9 shows the output of the simultaneous fits for the three categories at *low-mass*. The need for a separate signal parameterization per category for the *low-mass* analysis comes from the calorimeter resolution being very different depending on the photon reconstruction status. This is true for all $m_{\gamma\gamma}$ values, but in the low-mass analysis, the DY background is categorized

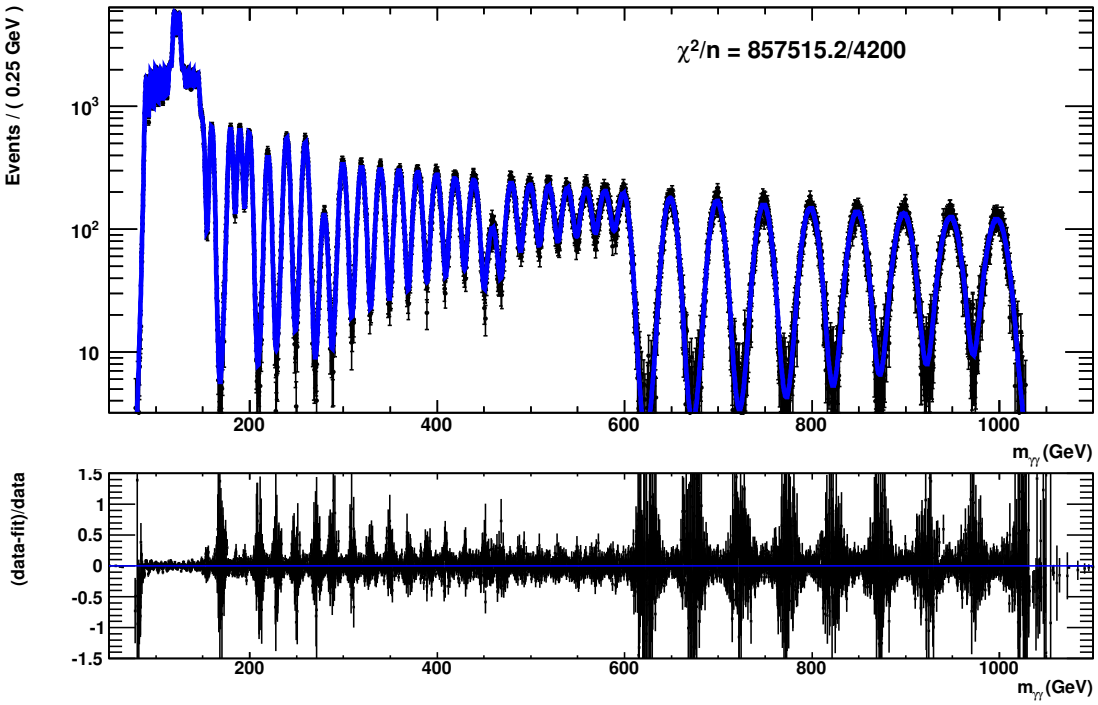


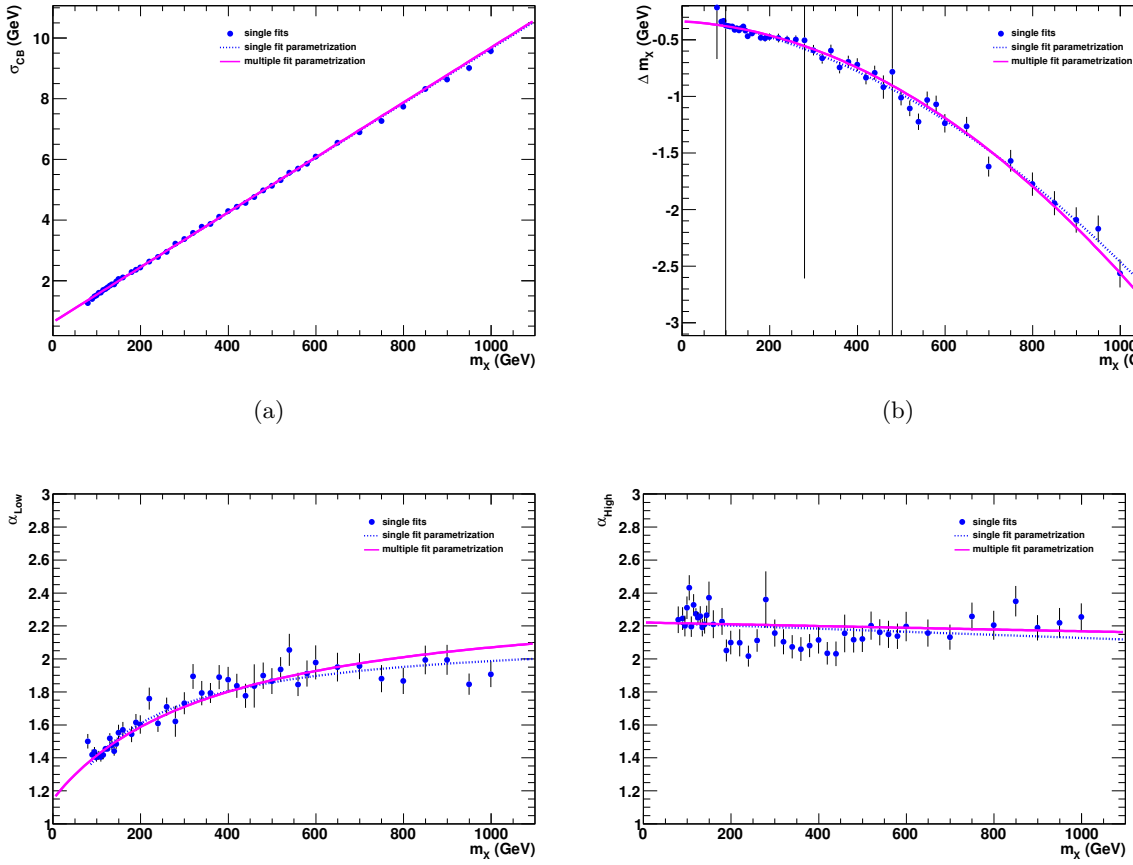
Figure 6.6: Simultaneous fit of the ggF MC samples, samples used: 90 - 1000 GeV.

Parameter	Parameterization	a	b	c
Δm_{X}	$a + bm_{nX} + cm_{nX}^2$	-0.379 ± 0.003	-0.062 ± 0.006	-0.020 ± 0.001
σ_{CB}	$a + bm_{nX}$	1.540 ± 0.003	0.905 ± 0.002	
α_{Low}	$a + b/(m_{\mathrm{X}} + c)$	2.420 ± 0.057	-483 ± 72	380 ± 48
n_{Low}	a	9.0 ± 0.2		
α_{High}	$a + bm_{nX}$	2.22 ± 0.02	-0.005 ± 0.003	
n_{High}	a	5.0 ± 0.1		

Table 6.2: Parameterizations of the DSCB parameters describing the signal shape, as a function of $m_{nX} = \frac{m_{\mathrm{X}}-100}{100}$. The result are extracted from the simultaneous fit. Errors are statistical only.

as well and a signal model is needed for all three categories separately (see Section 6.4.2).

The parameterizations obtained from the simultaneous fits for the three categories are given in Tables 6.3, 6.4 and 6.5, and will be used for the *low-mass* analysis. The n_{Low} and n_{High} parameters are constrained to constant parameterization in the simultaneous fit, and the single mass point fits are obtained using the resulting values of n_{Low} and n_{High} fixed. For the 2-conversion category, in the simultaneous fit, the n_{Low} is fixed to the same value that is obtained from the fit of the 1-conversion category, since the fit of the 2-conversion category did not converge to a reasonable value of n_{Low}, always reaching the upper limit given at input, which may be caused by a very low statistics of the MC sample in the 2 conversion category. The effect



(a) (b)

(c) (d)

Figure 6.7: Comparison of the simultaneous fit parameterization (pink line) to the output parameters of the single mass point fits (blue points) and to the parameterization of the single mass point fit parameters (blue line). The $m_{\mathrm{X}} = 80$ GeV blue point is not included for the single mass points fit. Few other mass points (namely $m_{\mathrm{X}} = 100, 280, 480$ GeV) have large errors on the Δm_{X} parameter, but these fits did converge and are therefore included in the DSCB parameters fit. The parameters are given as a function of m_{X} for the ggF samples: (a) σ_{CB}, (b) $\Delta m_{\mathrm{X}} = \mu_{CB} - m_{\mathrm{X}}$, (c) α_{Low}, (d) α_{High}

of the choice of this value $n_{Low} = 17$ is checked by fixing it to four different values ($n_{Low} = 13$, 15, 17 and 20) and comparing the fits on top of the MC sample, as seen in Figure 6.10. The variation of the n_{Low} parameter within the [13-20] range has no visible effect on the fitted signal shape.

6.3.4 Systematic uncertainties on the signal modelling

Due to the differences in the kinematic properties of the photons in the five production modes, the photon resolution can vary and hence the $m_{\gamma\gamma}$ distribution of the the new resonance X can change as well. The signal modelling is therefore repeated using VBF, $\mathrm{t\bar{t}X}$, ZX and WX

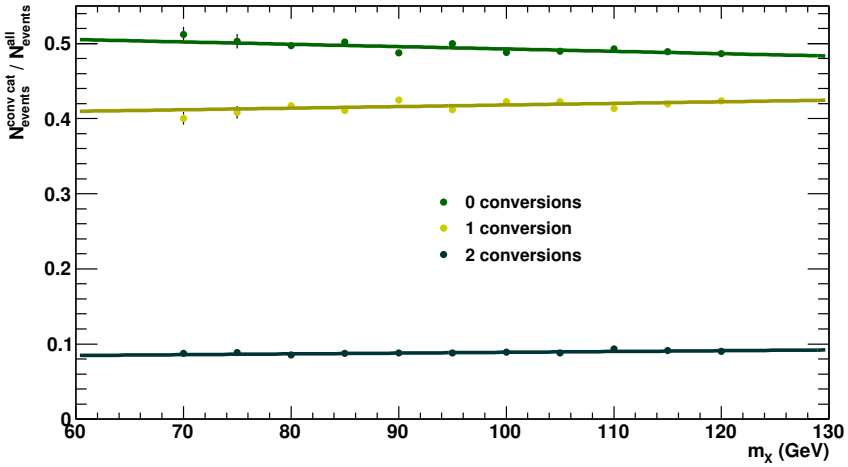


Figure 6.8: Fractions of 0-converted (green), 1-converted (yellow) and 2-converted (dark green) diphoton events passing all the selection cuts in the ggF samples as a function of m_{X}.

Parameter	Parameterization	a	b
Δm_{X}	$a + bm_{nX}$	-0.322 ± 0.004	-0.14 ± 0.03
σ_{CB}	$a + bm_{nX}$	1.418 ± 0.004	1.03 ± 0.02
α_{Low}	$a + bm_{nX}$	1.69 ± 0.02	-0.02 ± 0.07
n_{Low}	a	13 ± 1	
α_{High}	$a + bm_{nX}$	2.24 ± 0.04	0 ± 0.1
n_{High}	a	7.9 ± 0.8	

Table 6.3: Parameterizations of the double-sided Crystal Ball function describing the signal shape, result of the simultaneous fit at *low-mass*, for the 0-conversion category. $m_{nX} = \frac{m_{\mathrm{X}}-100}{100}$. Errors are statistical only.

Parameter	Parameterization	a	b
Δm_{X}	$a + bm_{nX}$	-0.468 ± 0.006	-0.12 ± 0.04
σ_{CB}	$a + bm_{nX}$	1.709 ± 0.006	1.07 ± 0.03
α_{Low}	$a + bm_{nX}$	1.23 ± 0.01	0.10 ± 0.05
n_{Low}	a	17 ± 1	
α_{High}	$a + bm_{nX}$	2.38 ± 0.05	-0.3 ± 0.2
n_{High}	a	5.2 ± 0.5	

Table 6.4: Parameterizations of the double-sided Crystal Ball function describing the signal shape, result of the simultaneous fit at *low-mass*, for the 1-conversion category. $m_{nX} = \frac{m_{\mathrm{X}}-100}{100}$. Errors are statistical only.

MC samples simulated with a narrow width. Figure 6.11 shows a significant production process dependence in the tail parameters of the DSCB, possibly due to an imperfect calibration for

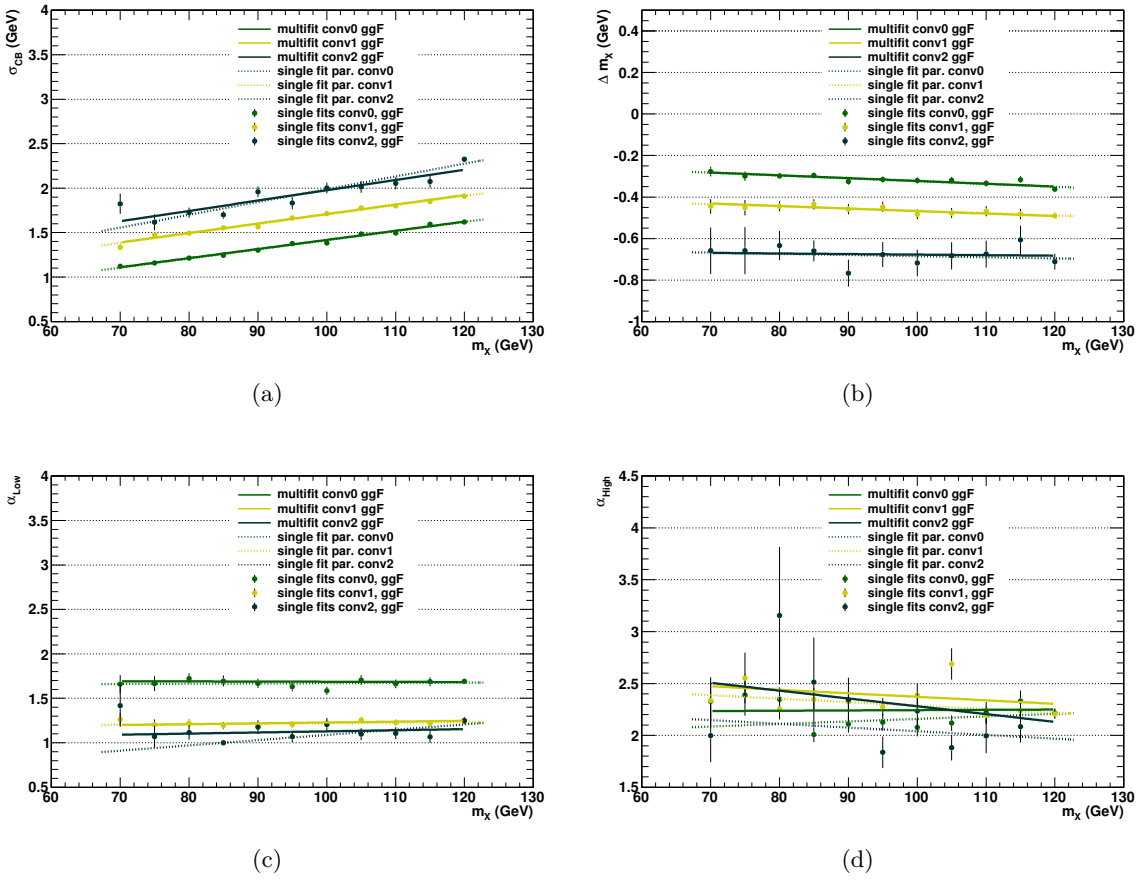


Figure 6.9: Parameterization obtained for the three *low-mass* categories (yellow and green lines). The dots show the DSCB parameters obtained from the single mass point fits for the three *low-mass* categories, the dashed line is a fit of the dots (the parameterization obtained from single mass point fits). The parameters are given as a function of m_{X} for the ggF samples: (a) σ_{CB}, (b) $\Delta m_{\mathrm{X}} = \mu_{CB} - m_{\mathrm{X}}$ (difference between the simulated and fitted peak position), (c) α_{Low}, (d) α_{High}.

photons with different E_{T}^{γ} spectra or the presence of jets in the event.

Aiming to be model independent in the search for a new resonance, this effect is considered as a possible source of a systematic uncertainty. It is computed as an event yield bias of the most-different result - the $\mathrm{t\bar{t}X}$ process w.r.t. the default ggF signal parameterization. Asimov datasets[3] are generated from the $\mathrm{t\bar{t}X}$ parameterization, together with and exponential background in steps of m_{X}. A DSCB plus exponential fit is performed using the signal parameters obtained from the ggF process MC samples. The resolution of the DSCB is allowed to vary within up to $\pm 40\%$ to account for the photon energy scale uncertainty [142]. The bias on the event yield is always below 1% and is therefore neglected.

Secondly, the effect of pile-up events on the signal shape is investigated by repeating the signal

[3]Asimov dataset replaces a set of data by one representative or median value [229].

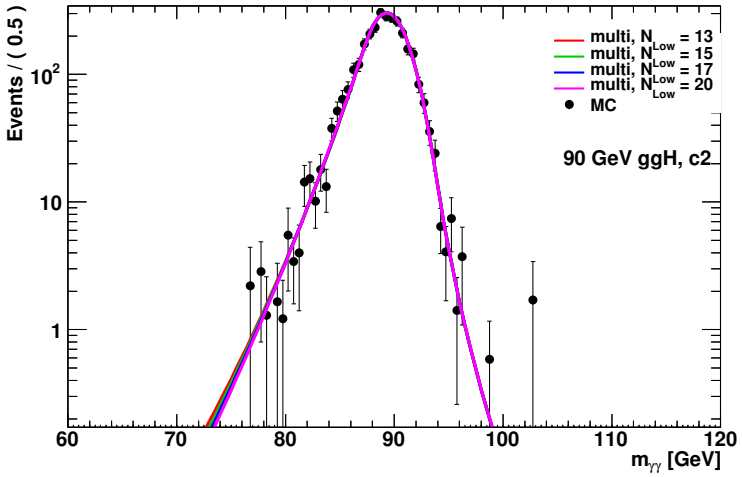


Figure 6.10: MC ggF sample at m_X= 90 GeV (black points) overlaid with a simultaneous fit result for the 2 conversion category obtained when fixing the n_{Low} parameter to 13 (red), 15 (green), 17 (blue) and 20 (pink).

Parameter	Parameterization	a	b
Δm_X	$a + bm_{nX}$	-0.68 ± 0.02	0.0 ± 0.1
σ_{CB}	$a + bm_{nX}$	1.98 ± 0.01	1.16 ± 0.09
α_{Low}	$a + bm_{nX}$	1.13 ± 0.02	0.1 ± 0.1
n_{Low}	a	17 ± 0	
α_{High}	$a + bm_{nX}$	2.28 ± 0.08	-0.8 ± 0.3
n_{High}	a	7 ± 2	

Table 6.5: Parameterizations of the double-sided Crystal Ball function describing the signal shape, result of the simultaneous fit at *low-mass*, for the 2-conversion category, $m_{nX} = \frac{m_X - 100}{100}$. Errors are statistical only.

parameterization for ggF MC samples separately for the average number of interactions per bunch crossing $\mu < 18$ and $\mu > 18$, being the low and high pile-up conditions, respectively. Figure 6.12 shows the parameterizations obtained. The procedure of the event yield bias determination is repeated for this case and it is found to be at the 0.1% level, therefore negligible.

The systematic effects due to the photon energy resolution are also studied. The noise term of the LAr calorimeter energy resolution, defined in Equation 3.3, is negligible for photons considered in this analysis. The uncertainty on the energy resolution is computed by shifting the sampling and constant terms of the energy resolution up and down within their uncertainties, applying them on the ggF signal MC samples and re-computing the signal parameterizations. The affected σ_{CB} parameter can be seen in Figure 6.13.

For the *low-mass* part of the analysis, in Figures 6.13(b)-6.13(d), the signal MC samples are split into the three conversion categories and the parameterization is recomputed. The bottom plots

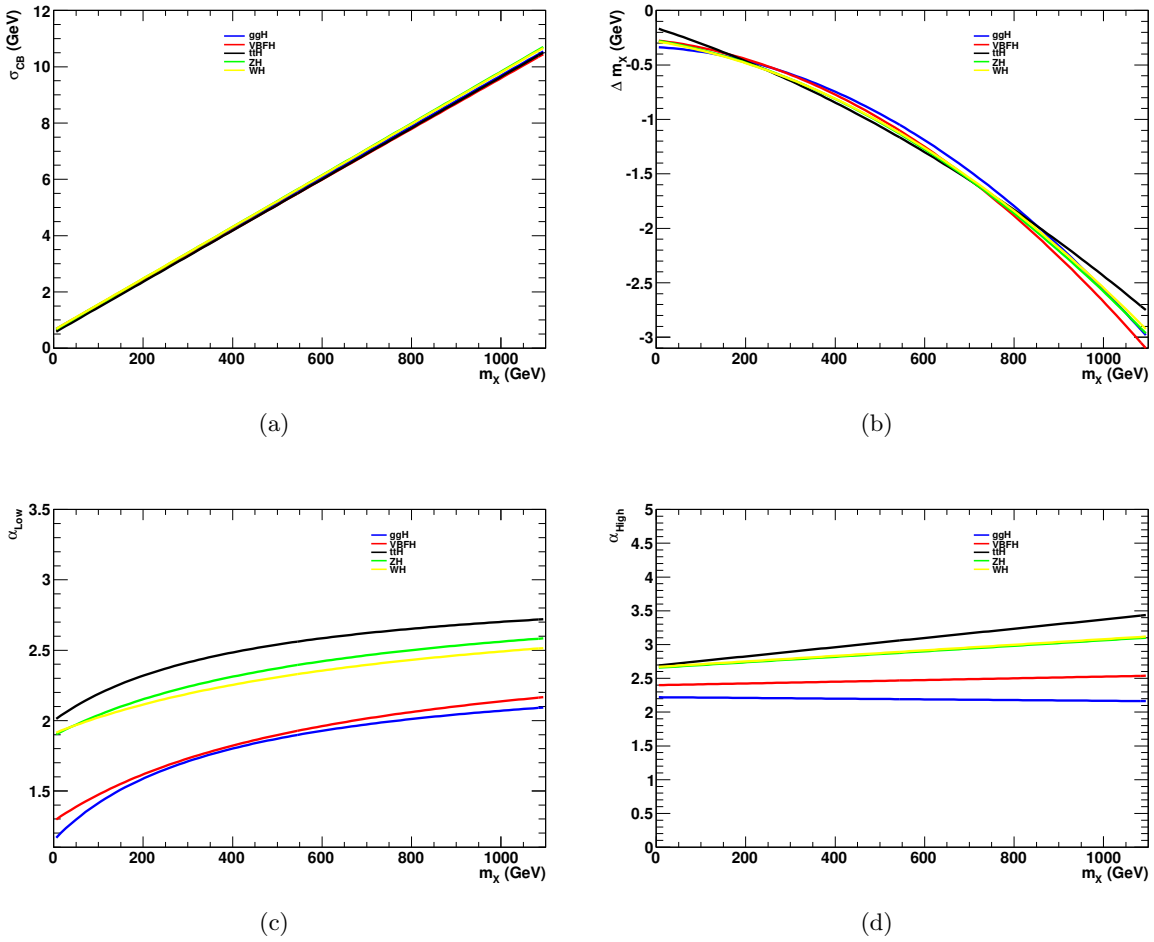


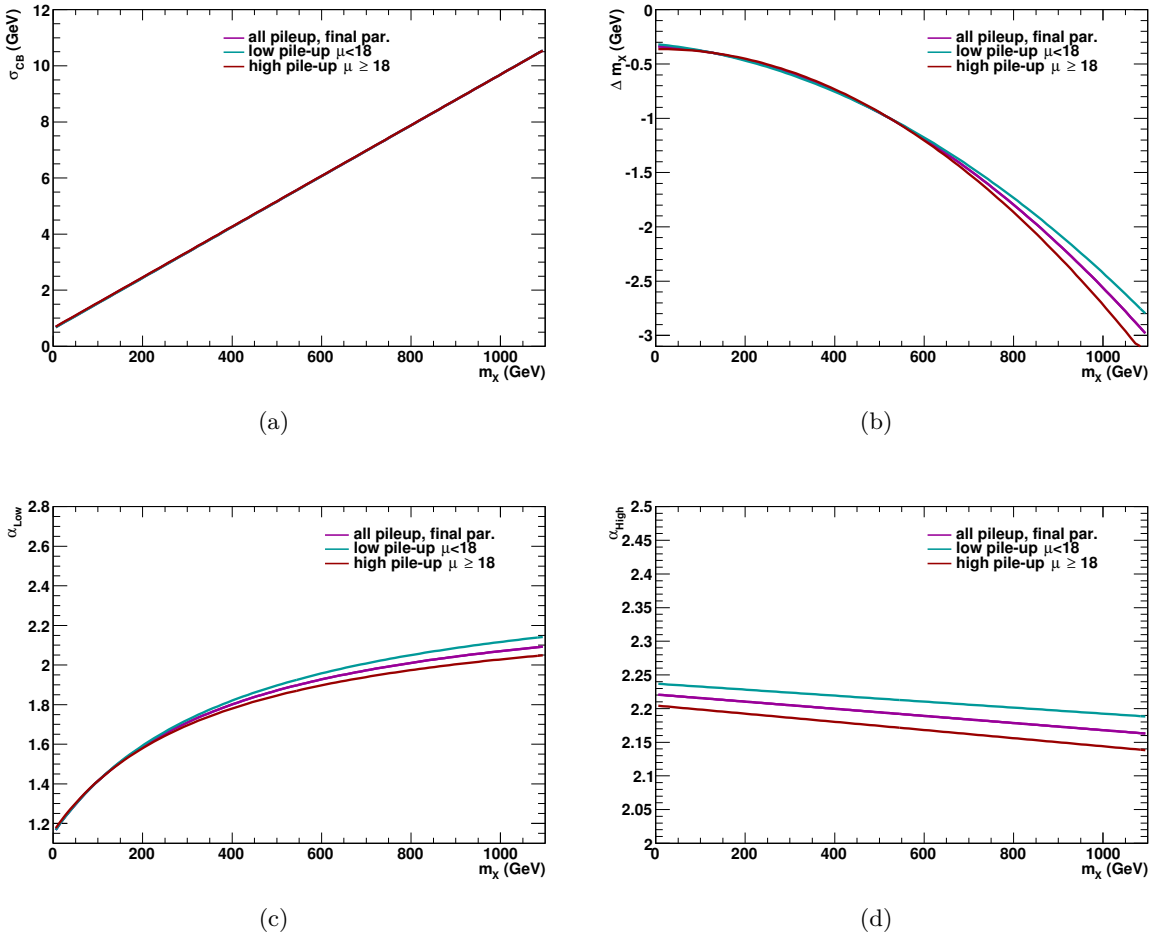
Figure 6.11: Simultaneous fits of the MC samples of the five main SM Higgs production processes modelled with NWA as a function of m_X.

show the relative difference w.r.t the default parameterization, spanning from 10% to 20% and heavily depending on the conversion category. Converted photons are more similar to electrons, therefore the uncertainty from extrapolating the energy resolution measured with electrons to photons is smaller. The *high-mass* part, displayed in Figure 6.13(a), shows a m_{X} dependence of the σ_{CB} of $10 - 40\%$.

The obtained systematic uncertainties on the photon energy resolution are parameterized as a function of m_{X} using a polynomial function and included as a component of the nuisance parameter on σ_{CB} in the final likelihood fit.

For the *low-mass* part of the analysis, an uncertainty on the migration between the conversion categories is also considered by comparing the number of events in each category using the SM $H \to \gamma\gamma$ MC sample at 125 GeV and the same sample produced with distorted ATLAS geometry[4]. A systematic uncertainty of 2.1%, 1.1% and 6% is assigned to the 0, 1 and 2-

[4]Distorted geometry is defined as a MC sample where the particle detection and reconstruction is performed with a different material budget w.r.t. the nominal configuration, such as a few % of X_0 added in between the pixel and SCT layers, increase of service material before the LAr calorimeter cryostat, ϕ-asymmetric material



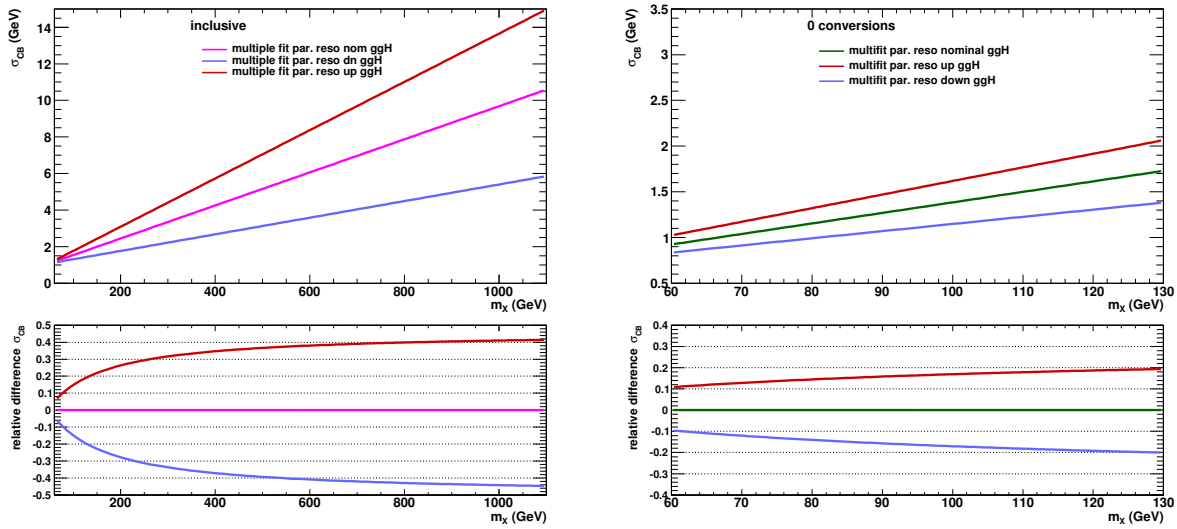
(a) (b)

(c) (d)

Figure 6.12: Simultaneous fit parameterizations of the inclusive ggF MC samples (pink), the low pile-up ggF MC samples with $\mu < 18$ (green) and the high pile-up ggF MC samples with $\mu \geq 18$ (red) as a function of m_X.

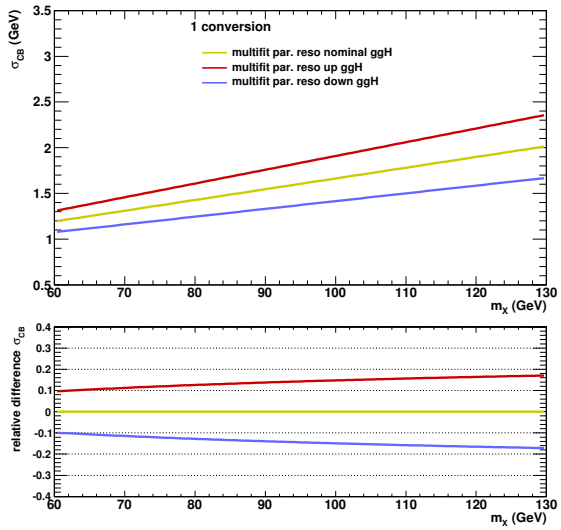
conversions categories, respectively.

added to the cryostat, material added behind the cryostat, density of the materials increased and so on, for explanation see for example [160].

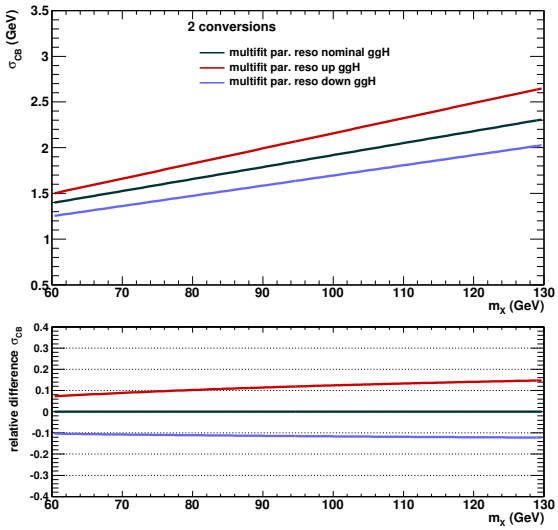


(a)

(b)



(c)



(d)

Figure 6.13: *High-mass* analysis (a). *Low-mass* analysis split in 0 conversion (b), 1 conversion (c) and 2 conversions (d) categories. Variation of the σ_{CB} parameter as a function of m_{X} for the nominal samples, and after the up/down shift of the calorimeter resolution parameters (top) and the relative difference between the shifted samples and the nominal sample, as a function of m_{X} (bottom).

6.4 Background modelling

6.4.1 Jet background and photon purity

The main reducible background for prompt diphoton production at the LHC are γ-jet, jet-γ and jet-jet events, where one or both jets are misidentified as photons. The composition of this reducible background can be estimated using several approaches, introduced in previous measurements [96]. For this analysis, the 2x2D Sideband method (2x2DSB) is used.

The 2x2DSB method uses the photon isolation and tight identification criteria to sort events into sixteen categories (two photons, each can be isolated or not and tight or not) and express the number of events in each category using photon tight identification and isolation efficiencies and jet fake rates. Simulated SHERPA [72] samples are used to obtain the photon isolation and ID efficiencies[5], with all other inputs being from data and extracted with a fit. Figure 6.14 shows the $m_{\gamma\gamma}$ spectrum decomposed into the $\gamma\gamma$, γ-jet+jet-γ and jet-jet components. The inclusive purity is measured to be $(77.5 \pm 2.5)\%$ [207]. The dominant systematic uncertainty comes from the definition of the control region based on the tight identification. The total fractions of $\gamma\gamma$, γ-jet and jet-jet events are 69.1%, 26.1% and 4.8%, respectively, for the *low-mass* range, while the purity increases from 65% to 75% with the invariant mass [207]. For the *high-mass* region, the fractions are 83.7%, 15.1% and 1.2% with a purity increasing from 80% to 96% as a function of $m_{\gamma\gamma}$ [207].

Since the $X \to \gamma\gamma$ search relies on a fit of the backgrounds and a potential signal, the jet background subtraction is not necessary, and is only shown here to illustrate the understanding of the background composition. However in the precision measurement of the diphoton + jet production cross section, presented in Chapter 7, the 2x2DSB method is used to precisely estimate the QCD diphoton yields, and it will be described in detail in Section 7.3.

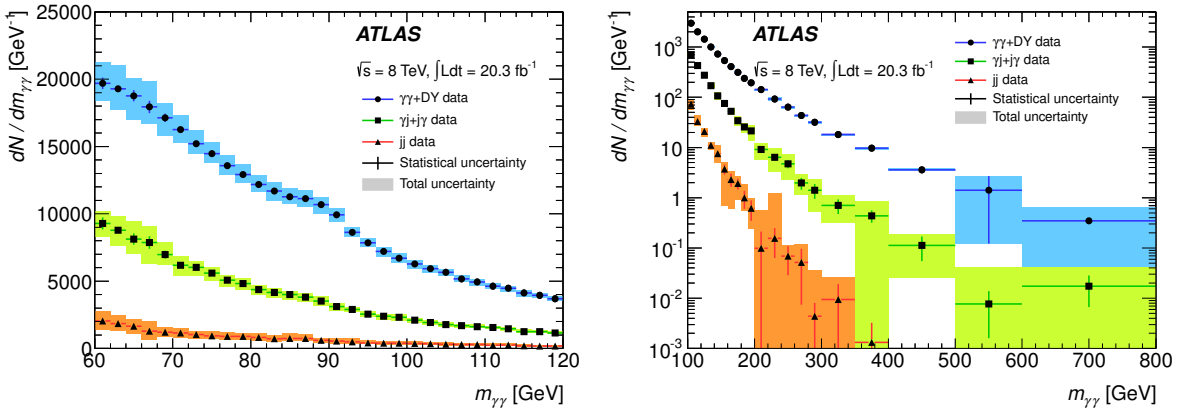


Figure 6.14: Decomposition of the reconstructed diphoton invariant mass spectrum into $\gamma\gamma$, γ-jet and jet-jet components, obtained using the 2x2DSB method for the (a) *low-mass* region and (b) for the *high-mass* region [206].

[5]Photon *tight* ID efficiencies are obtained from MC but the MC events are scaled by photon ID scale factors, which are data-driven and were derived by the e/gamma group [230].

6.4.2 Electron background

The second most prominent background source comes from electrons misidentified as photons. In the *low-mass* part of the analysis, electrons produced by the Drell-Yan process introduce a sizeable background component, with the resonant Z-shape being visible in the spectrum in Figure 6.14(a). In the *high-mass* analysis, the electron background is very small, taking a shape similar to the diphoton background and therefore does not need to be extracted, its description is neglected in this analysis, since it is absorbed in the diphoton continuum.

6.4.2.1 Photon conversion categories

The electrons misidentified as photons are expected to be mostly reconstructed as converted photons, since an electron track can be reconstructed for true electrons in the tracker and one or two tracks are typically reconstructed for converted photons. The electron background is therefore evaluated in conversion categories separately for unconverted-unconverted (UU) photons, converted-converted (CC) and converted-unconverted and unconverted-converted (CU+UC) together. Further categorization based on the number of conversion tracks detected was studied both in terms of statistical power and systematic uncertainties, which are dominant for this background, and it was determined to be unnecessary. The expected fractions of signal and DY events for each category, derived from the DY backgroud template computed in this section from MC and data-driven measurements, normalized to the luminosity in data, are shown in Table 6.6.

$\gamma\gamma$ category	UU	CU+UC	CC
N_{data}	272 184	253 804	63 224
N_{DY}	1 080 ± 260	3 400 ± 600	2 700 ± 250
f_{DY}	15.0%	47.3%	37.7%
f_X	48.7%	42.5%	8.8%

Table 6.6: Number of diphoton events in data (N_{data}), number of expected Drell–Yan events (N_{DY}), fractions of expected signal (f_X) and Drell–Yan (f_{DY}) in each conversion category for the *low-mass* analysis. The signal fraction is given for m_{X}= 90 GeV but the mass-dependence is negligible.

6.4.2.2 Electron reconstruction

The amount of DY background is estimated using dielectron events in data. Electrons are reconstructed using algorithms for converted and unconverted photon reconstruction. The electrons are required to pass the same E_{T} cuts as the photons and the `Tight++` identification criteria (a tighter selection based on shower shape variables and TRT hits information optimized for electrons) [231]. The isolation cut is not applied and events with electrons overlapping with a reconstructed photon within a cone of $\Delta R_{\gamma e} < 0.05$ are removed from the sample. This selection provides a Drell-Yan template, which needs to be normalized by electron to photon fake rates, to be computed separately for the leading and the subleading candidate.

As illustrated in Figure 6.15, electrons forming the Drell-Yan background mostly underwent large bremsstrahlung, and have therefore somewhat different kinematic properties than the electrons

used to produce the template. The dielectron peak position of electrons is shifted w.r.t. the fake photon peak. The kinematic properties of the electrons are studied extensively and a correction of E_{T} and ϕ is derived. Other effects, such as the η position of the electrons, have been found to be negligible. Since the shift in ϕ depends on the charge q of the electron (or positron) and on its transverse momentum, the correction is derived as a function of E_{T} and $q\phi E_{\mathrm{T}}$. MC samples of $Z \to ee$ events generated with POWHEG[214, 215] and interfaced with PYTHIA8[71] are used to derive these corrections.

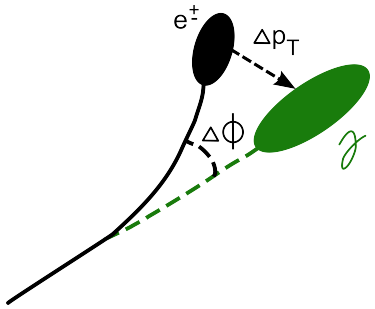


Figure 6.15: Illustration of the bremsstrahlung of electrons that can fake photons in the detector.

The difference between the reconstructed and true transverse energy $\Delta E_T = (E_T^{reco} - E_T^{truth})$ is fitted with a Gaussian distribution in E_T bins, separately for electrons and fake photons. The obtained mean μ and σ in each E_T bin are then used to obtain the shift μ and smearing σ parameters

$$\mu = \mu_{\gamma(e)} - \mu_e, \qquad \sigma = \sqrt{\sigma^2_{\gamma(e)} - \sigma^2_e}. \tag{6.3}$$

The E_T dependence of μ and σ is then parameterized as $aE_T/(1 + bE_T^2)$ so that the correction is zero at zero and infinity. Around E_T=40 GeV the dependence is mostly linear. Only a shift is applied for the unconverted photon category since the σ parameter is found to be compatible with zero. The electron candidate transverse momentum is then transformed to simulate fake photons with a Gaussian smearing

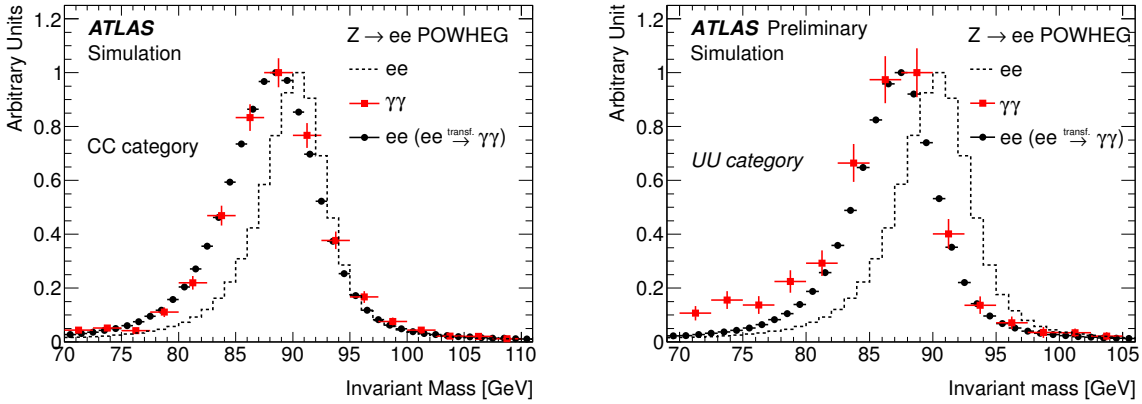
$$E_T \to \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(E_T-\mu)^2}{2\sigma^2}}. \tag{6.4}$$

For the azimuthal angle ϕ correction, the difference $(\phi^{reco} - \phi^{truth})qE_T^{reco}$ depends on η with a $\cosh(\eta)$ shape, hinting at an effect proportional to the length of the track in the longitudinal plane. The corresponding shift and smear parameters are therefore modelled with a $\cosh\eta$ dependence.

A similar procedure is followed to reproduce the fake unconverted photon kinematics. The obtained invariant mass distributions for converted and unconverted fake photons are shown in Fig 6.16.

6.4.2.3 $e \to \gamma$ fake rates

The normalization of the template is derived from $e \to \gamma$ fake rates, which are computed from data by using dielectron and electron-photon events, separately extracting fake rates for the



(a) (b)

Figure 6.16: Invariant mass distributions in the (a) CC category [206] and (b) UU category for simulated $Z \to ee$ events reconstructed as ee (dotted lines), reconstructed as $\gamma\gamma$ (red squares) and reconstructed as ee after the transformation of electrons to match the kinematic properties of the electrons misidentified as converted photons (black circles) [205], title of x-axis changed for consistency.

leading and subleading candidate for each conversion category. The fake rates, defined as

$$\rho_1 = \frac{N_{\gamma\gamma}}{N_{ee}} \qquad \rho_2 = \frac{N_{e\gamma}}{N_{\gamma e}}, \tag{6.5}$$

are measured in a window around the Z boson mass, its width being defined as four times the width of a Gaussian (as part of the double-sided Crystal Ball) fit of the ee mass distribution. The peak position and its width are different for each category. The non-resonant Drell-Yan background is modelled as a decreasing exponential and is subtracted from data before the fake rate extraction. The total number of Drell-Yan events contaminating the diphoton signal can be expressed as $\alpha\rho_1\rho_2 N_{ee}$, where α is a correction factor introduced to take into account the reconstruction efficiency differences between single photons and diphoton pairs. It is computed using MC $Z \to ee$ events as

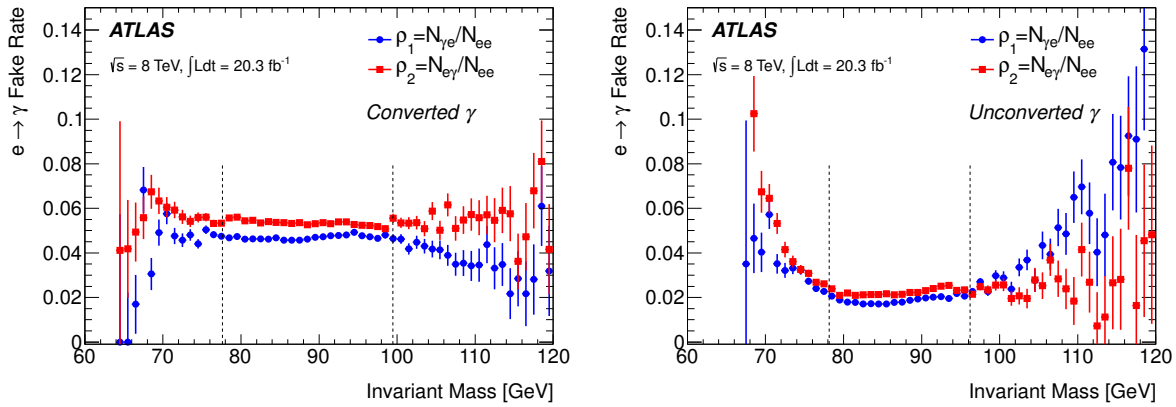
$$\alpha = \frac{1}{\rho_1^{\mathrm{MC}}\rho_2^{\mathrm{MC}}}\frac{N_{\gamma\gamma}^{\mathrm{MC}}}{N_{ee}^{\mathrm{MC}}} = \frac{N_{\gamma\gamma}^{\mathrm{MC}}N_{e\gamma}^{\mathrm{MC}}}{N_{ee}^{\mathrm{MC}}N_{\gamma e}^{\mathrm{MC}}}, \tag{6.6}$$

and is expected to be around one if the misidentification of the two candidates is uncorrelated. This assumption is not true in case of this analysis, since the track isolation requirement is applied, with the track isolation using the inner tracker information from both photon candidates to find the primary vertex. When building ee and γe pairs to derive the template and the fake rates, neither the dielectron vertex nor the electron-photon vertex were used, therefore α will not be one. The computed value is $\alpha \approx 0.5$.

Several systematic uncertainties are considered to take into account the difference of α from unity. The mass window of the DSCB fit is varied between two and four times the measured Z boson peak width and its position is shifted by $\pm 1\sigma$ of the Z boson peak position uncertainty, producing an uncertainty on the $\alpha\rho_1\rho_2$ product of $\pm 2\%$. Parameters of the electron kinematic

transformation are varied within their uncertainties, resulting in an uncertainty of ±1.5%. Forgoing the subtraction of the non-resonant component, a difference of +5% is observed, which is symmetrized to illustrate the fact that the knowledge of its description is poor. Finally, a SHERPA MC sample and a sample of PYTHIA with distorted ATLAS geometry is compared to the baseline PYTHIA sample, leading to an effect of −5%, also symmetrized.

The measured $e \to \gamma$ fake rates are shown for converted and unconverted photons in Figure 6.17. A very small mass-dependence is observed after the E_{T} shift and smearing is applied to the electrons to match the Z boson peak position in events where electrons are misidentified as a photon. Residual non-linearities are accounted for with the systematic uncertainties on the α factor mentioned above.



(a) (b)

Figure 6.17: $e \to \gamma$ fake rates for (a) converted and (b) unconverted photon candidates computed from reconstructed $Z \to ee$ events in data for the leading (blue circles) and subleading (red squares) photons as a function of the mass of ee, $e\gamma$ and γe pairs [206]. The dotted lines show the window used to determine the nominal value of the fake rates.

6.4.2.4 Final Drell-Yan templates

The Drell-Yan background templates extracted from dielectron data, transformed using the method described in Section 6.4.2.2 and normalized according to the measured $e \to \gamma$ fake rates are shown in Figure 6.18 together with a ratio plot, showing the differences between the conversion categories. The CU and UC templates are summed into the CU category and used as one in the final fit. The systematic uncertainties described previously amount to a relative 9-25% uncertainty that is category dependent, with the highest error found in the UU category, which is the least populated. Systematic uncertainties on the template shape are evaluated using other shapes by varying the shift and smearing transformation parameters within 1σ, producing a difference w.r.t. the default shape of 1.5 to 3% per bin, category dependent.

6.4.3 Continuum background

The choice of the fit function for the continuum background is driven by the *spurious signal* method, previously used in the SM Higgs boson searches [232]. Using high statistics background-

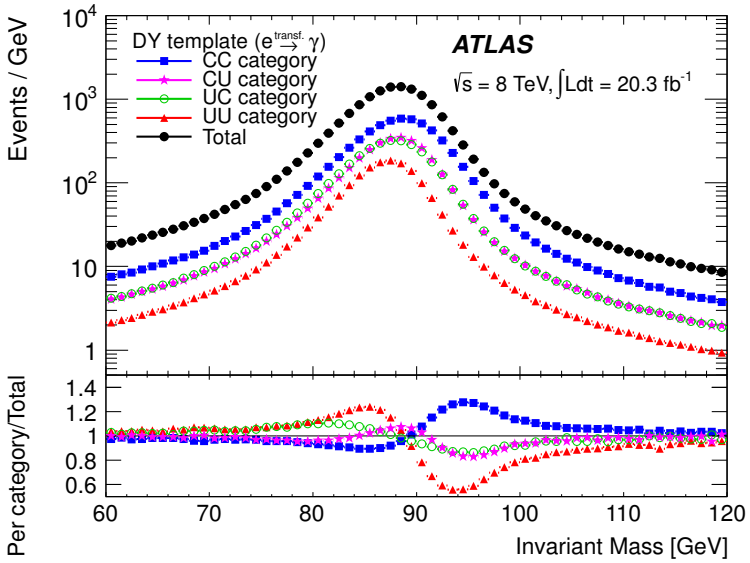


Figure 6.18: Drell-Yan invariant mass template shapes extracted from data (top) describing the expected diphoton background component and ratio of each category template to the total sum (bottom) [206].

only SHERPA MC samples, a fit of a signal and background function is performed. The obtained signal yield is called the spurious signal $N_{spurious}$. This procedure is repeated in steps of 1 GeVof $m_{\gamma\gamma}$ and the maximum spurious signal is required to be lower than 20% of the statistical uncertainty of the background. Different fitting functions are explored to find the best possible description of the data. If several fitting functions satisfy the 20% constrain on the spurious signal in the fit, the function with the smallest amount of degrees of freedom is chosen.

The size of the spurious signal is taken into account in the final fit, when computing the limit, as an additional component on top of the background and signal. It is described as a DSCB shape with a width equal to $N_{spurious}$ [233]. The size of the spurious signal is determined to be 1-67 events, depending on $m_{\gamma\gamma}$.

6.4.3.1 *Low-mass* continuum background

For the *low-mass* analysis, several functions are tested, separately for the three conversion categories. The background is modelled using two high statistics SHERPA $\gamma\gamma$ MC samples generated for a different range of the diphoton invariant mass, merged and reweighted according to the luminosities. Roughly 10% of the dataset went through ATLAS reconstruction, with the rest having the true photon quadrivectors corrected to mimic experimental effects. Photon energy is smeared and conversion status is randomized to mimic the data. The MC is normalized to match the statistics of data. Following the spurious signal method, a fit of signal and background is performed in the range of $60 < m_{\gamma\gamma} < 120$ GeV and the spurious signal is tested. A sum of a Landau and an exponential is chosen, with two degrees of freedom for the Landau part, one for the exponential and one for the fraction between the two. The Landau function should properly describe the trigger turn-on, where a change of concavity is expected (as illustrated in

Figure 6.1 on MC), while the exponential describes the rest of the diphoton continuum. The spurious signal computed with different functions for the *low-mass* part of the analysis is shown in Figure **??** for the three conversion categories.

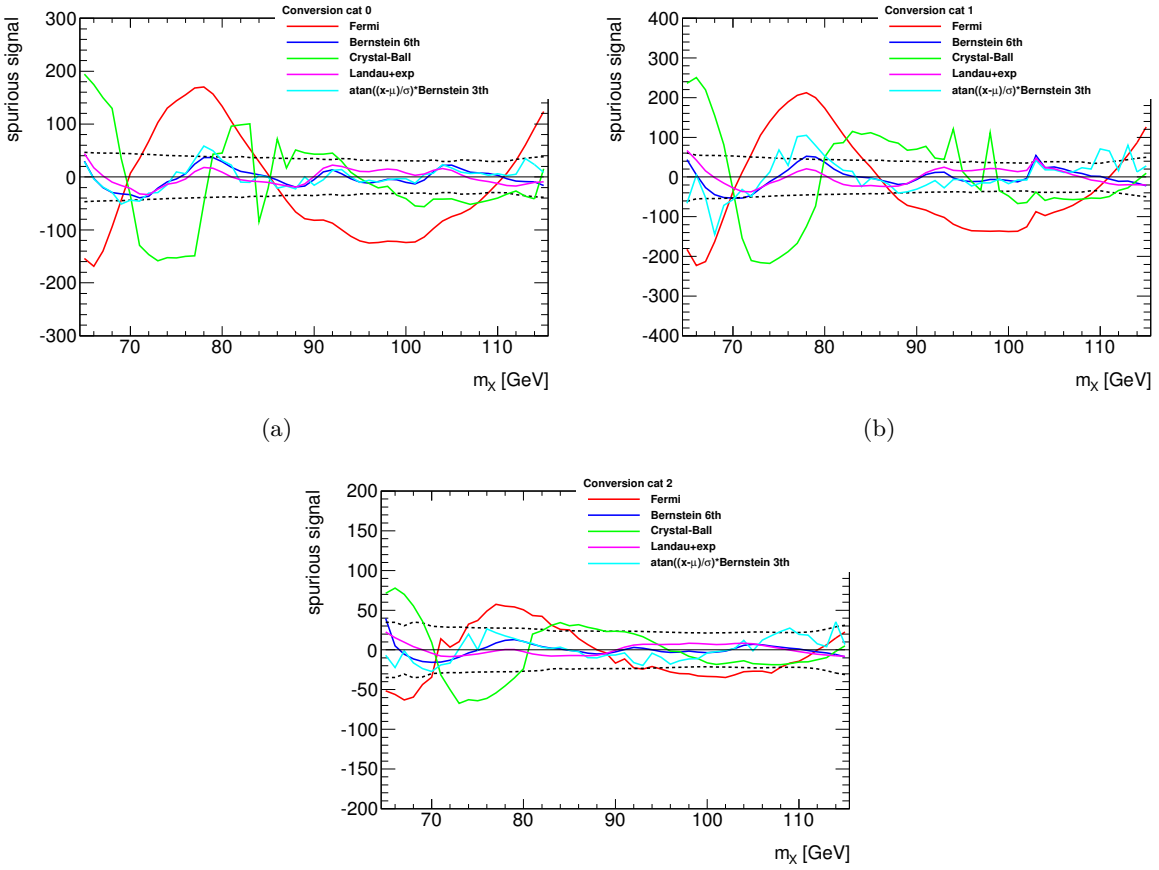


Figure 6.19: Size of spurious signal as a function of m_X for different functions for the (a) 0-conversions, (b) 1-conversion and (c) 2-conversions category. The dashed line corresponds to 20% of the expected background uncertainty [207].

6.4.3.2 *High-mass* continuum background

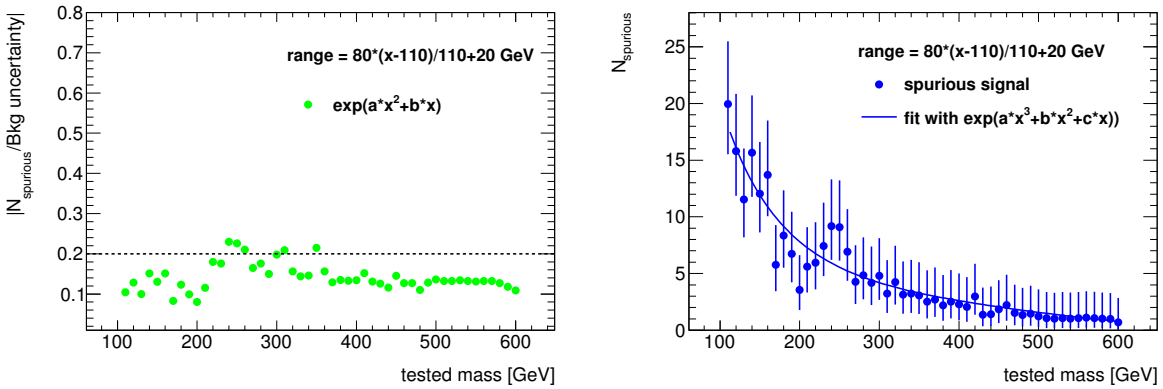
In the *high-mass* analysis, the choice of the function is more difficult. The shape of the spectrum is modelled on high statistics SHERPA MC samples including the $\gamma\gamma$, γ-jet and jet-jet processes, composed of three separate MC samples mixed according to the yield obtained from the 2x2DSB method described in Section 6.4.1. The full range of $100 < m_{\gamma\gamma} < 600$ GeV is tested, but it proved to be very difficult to find a function satisfying the spurious signal criteria with a fit in the full range.

A sliding window technique is used instead. The size of the sliding window is determined from the spurious signal requirement being smaller than 20% of the background uncertainty, and the condition that for a mass of 110 GeV, the window cannot be larger than 20 GeV due to the

relative E_{T} cuts creating a change of shape in $m_{\gamma\gamma}$ at 100 GeV. A simple exponential proved to be insufficient and an exponential of a second order polynomial is used instead. A possible improvement of this would be if asymmetrical windows were considered. The optimal range of the window is determined to vary with $m_{\gamma\gamma}$ as

$$80 \cdot (m_{\gamma\gamma} - 110)/110 + 20 \text{ GeV}. \tag{6.7}$$

Figure 6.20(a) shows the spurious signal computed using the exponential of a second order polynomial in the sliding window range. The spurious signal is computed in 10 GeV steps for the whole range, fitted by an exponential of a third order polynomial, shown in Figure 6.20(b) and used in the final limit computation.



(a) (b)

Figure 6.20: (a) Ratio between the size of the spurious signal and the statistical uncertainty on the background as a function of m_X. (b) Size of spurious signal as a function of m_X, fitted by an exponential of a third order polynomial [207].

6.4.4 SM Higgs background

The SM Higgs boson signal in the diphoton channel represents another background manifesting itself as a peak on top of the continuum at the mass of $m_{\gamma\gamma} \approx 126$ GeV, therefore affecting the *high-mass* part of the analysis.

To model the shape of the Higgs boson, a DSCB function, as described in Section 6.3.1, is used, in order to stay consistent throughout the analysis. The event yields for the five SM Higgs production modes are computed from the efficiencies derived using the event selection described in Section 6.2.2. The cross sections are taken from the CERN Yellow Report [21]. The value of $m_H = 125.9$ GeV is used for the mass, taken from PDG [63], as the most precise measured Higgs mass, at that time. The signal strength is taken as $\mu = 1$, even though the ATLAS experiment had, then, measured a μ value larger than one [186, 27]. This analysis still aims to be able to see another potential resonance close or overlapping in mass to the recently discovered Higgs boson.

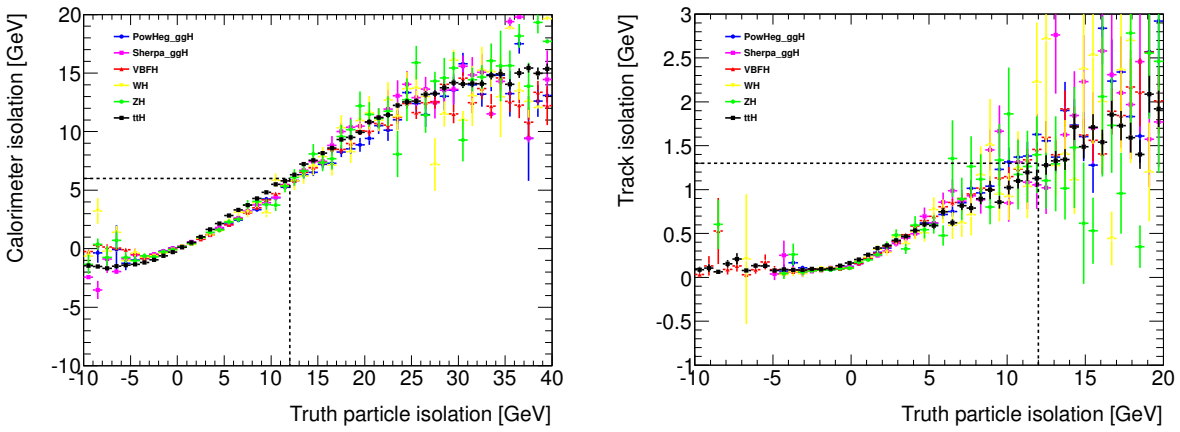
6.5 Fiducial volume definition

The search for a narrow resonance decaying into two photons is performed in a fiducial volume. To correct the number of fitted signal events in data for detector effects, a correction factor

$$C_X = \frac{N_{MC}^{reco}}{N_{MC}^{fid}} \tag{6.8}$$

is defined, where N_{MC}^{reco} is the number of simulated signal events passing all selection criteria and N_{MC}^{fid} is the number of simulated signal events generated within the fiducial volume.

To define the fiducial volume, the same E_{T} and η selections are applied on the two photons. Photons reconstructed in the presence of high-p_{T} jets can have a lower efficiency of the isolation cut, leading to a different C_X factor and model dependence. Applying a cut on the particle isolation, defined as the p_{T} sum of all stable particles, except neutrinos, found within a cone of $\Delta R = 0.4$ around the photon at the truth level, reduces this effect. The cut is chosen at 12 GeV, since this value corresponds to the calorimeter-based isolation energy cut applied at the reconstructed level of $E_{\mathrm{T}}^{iso,calo} < 6$ GeV, as illustrated in Figure 6.21(a). No additional cut is applied on the particle isolation due to the track-based isolation requirement at the reconstruction level, since the track-based isolation cut of $p_{\mathrm{T}}^{track,iso} < 2.6$ GeV applied at the reconstruction level corresponds to a much looser cut on particle isolation, as shown in Figure 6.21(b).



(a) (b)

Figure 6.21: Reconstructed calorimeter isolation (a) and reconstructed track isolation (b) as a function of the truth particle isolation, for all SM Higgs production modes and m_X=125 GeV [207].

The C_X factor is computed for all five SM Higgs production processes in order to be as model independent as possible. These five processes span a range of production topologies. The final state can be clean or populated with jets or bosons, with different angular and kinematic distributions. The final C_X factor is shown in Figure 6.22 as a fit to the blue points obtained from the ggF MC samples. The bottom part shows the ratio of the VBF, ttH, ZH and WH processes to the ggF process, from which a m_{X}-dependent systematic uncertainty of 3-15% is extracted. The uncertainty is 8-15% for the *low-mass* part of the analysis and a stable ±3% difference is observed for the *high-mass*. These differences come from the diverse kinematic

properties of each production process (higher E_{T} in ttX, more central photons in VBF). Small differences at *high-mass* can originate in the fraction of events with a photon in the excluded η region, or the small decrease of selection efficiency for high E_{T}^{γ} photons.

The dependence of the C_X factor on pile-up is evaluated using low and high pile-up events and an uncertainty of 1-3% is applied. Underlying event effects are also studied by turning off the multiple parton interaction (MPI) and a difference of 1% is found. Effects coming from the renormalization scale and PDF uncertainties are found to be negligible.

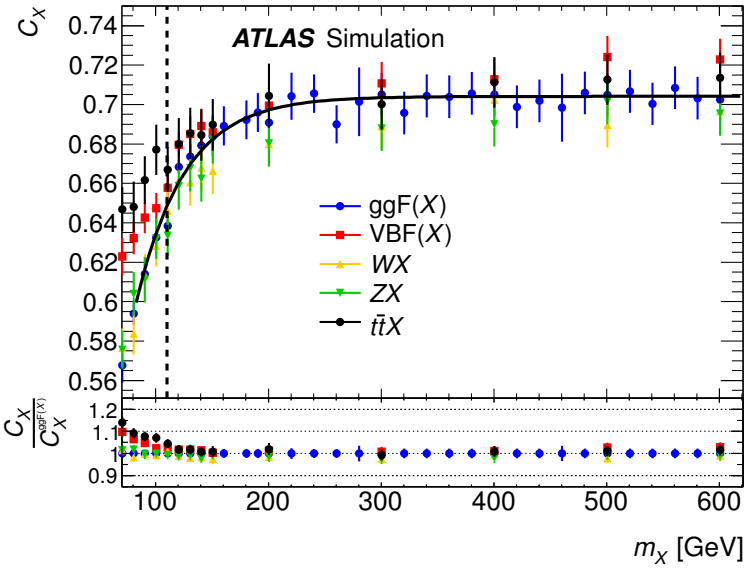


Figure 6.22: Efficiency correction factors C_X computed from fully simulated samples of the five main SM Higgs production modes modelled with NWA, as a function of m_X (top) and a ratio of the C_X factors to the ggF(X) C_X factor (bottom). The discontinuity at $m_X = 110$ GeV is due to the transition between the *low-* and *high-mass* analyses [206].

6.6 Systematic uncertainties

Systematic uncertainties not evaluated previously are described in this section. A summary of all considered systematic uncertainties in this analysis is shown in Table 6.7.

6.6.1 Photon identification systematic uncertainties

The computation of the uncertainty on photon identification is based on the uncertainties on data-driven measurements in Ref. [163]. The uncertainties depend on photon conversion status, E_{T} and η and range from 1.5-2.5%. As a function of $m_{\gamma\gamma}$, this uncertainty is parameterized using a 5th order polynomial, decreasing from 2.7% at low $m_{\gamma\gamma}$ down to 1.6% at 600 GeV, as shown in Figure 6.23.

Source	Uncertainty	Reference
Signal yield		
Luminosity	±2.8%	Ref. [131]
Trigger	±0.5%	Ref. [186]
Photon identification	±1.6 – 2.7%, mass-dependent	Section 6.6.1
Isolation efficiency	±1 – 6%, mass-dependent	Section 6.6.2
Photon energy scale	±0.25%	Ref. [186]
Signal modelling		
Photon energy resolution	±10 – 40% mass and category-dependent	Section 6.3.4
Pile-up	negligible	Section 6.3.4
Low-η material	±0.34%	Ref. [186]
High-η material	±0.39%	Ref. [186]
Barrel Presampler scale	±0.06%	Ref. [186]
Endcap Presampler scale	±0.10%	Ref. [186]
Other effects	±0.01%	Ref. [186]
C_X *factors*		
Production process	±3 – 15%, mass-dependent	Section 6.5
Pile-up	±1 – 3%, mass-dependent	Section 6.5
MPI	±1%	Section 6.5
non-resonant Background		
Spurious Signal	1 – 67 events, mass-dependent	Section 6.4.3
Higgs Background Theory		
ggF	−7.8 +7.2% (scale), −6.9 +7.5% (PDF)	Ref. [21]
VBF	−0.1 +0.3% (scale), −2.8 +2.6% (PDF)	Ref. [21]
WH	±1.0% (scale), ±2.3% (PDF)	Ref. [21]
ZH	±3.2% (scale), ±2.5% (PDF)	Ref. [21]
$t\bar{t}$H	−9.3 +3.8% (scale), ±8.1% (PDF)	Ref. [21]
Branching ratio	−4.87 +4.84%	Ref. [21]
DY Background modelling		
Peak position	±1.5 – 3.5%, category-dependent	Section 6.4.2.3
Normalisation	±9 – 25% category-dependent	Section 6.4.2.3
Template shape	±1.5 – 3% category-dependent	Section 6.4.2.4

Table 6.7: Summary of the main sources of systematic uncertainties for the measurement of the fiducial cross section.

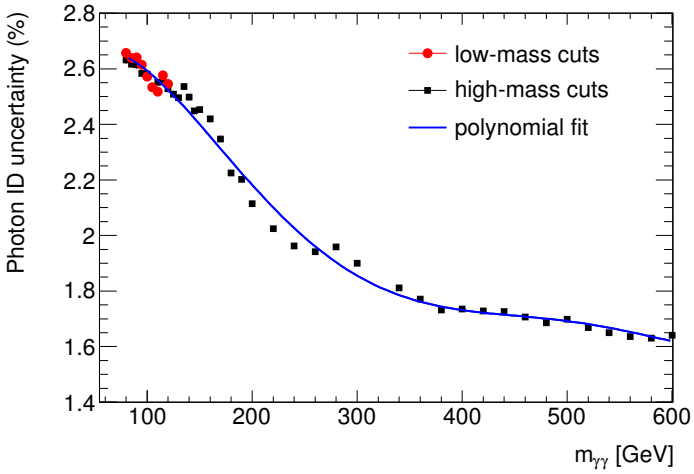


Figure 6.23: Photon ID uncertainty as a function of m_X [207].

6.6.2 Photon isolation systematic uncertainties

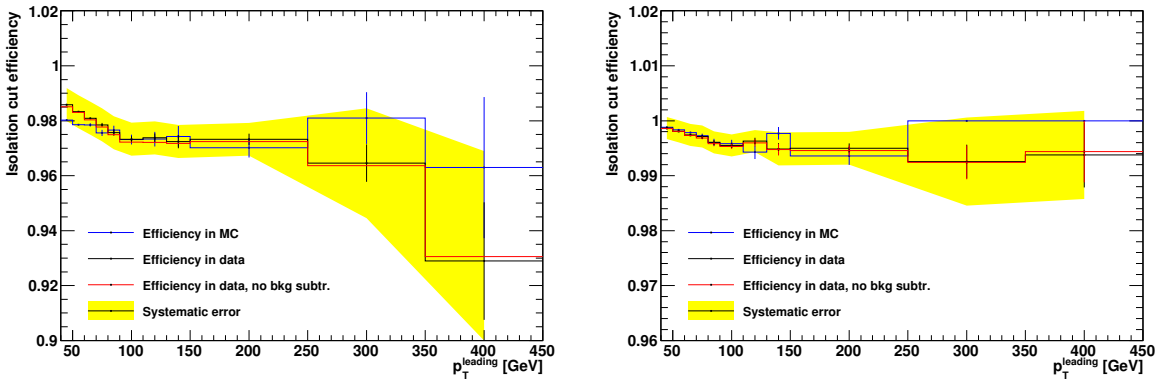
The systematic uncertainty on the photon isolation selection is computed from data and MC comparisons in $Z \to ee$ events. The isolation distributions of these electrons were previously found to be comparable to those of photons [234] and the effects of pile-up and underlying event are also similar [235].

The efficiency of the isolation selection is studied separately for the calorimeter-based and the track-based isolation for the leading electron in ten p_T^e bins between 40-350 GeV. The selected events had to have both electrons in the $|\eta| < 2.47$ region, with $p_T^{e1} > 40$ GeV and $p_T^{e2} > 30$ GeV, the `Tight++` identification criteria, $70 < m_{ee} < 110$ GeV and pass the same data quality requirements as the photon events selection.

The isolation efficiencies are obtained for data and MC by integrating isolation distributions up to the corresponding cut value, which is 6 GeV for the calorimeter-based isolation and 2.6 GeV for the track-based isolation. The efficiency of the calorimeter isolation as a function of p_T^{e1} has a falling trend for both data and MC, as shown in Figure 6.24(a). The agreement between data and MC is good up to 250 GeV where the MC statistics gets too small.

The observed disagreement is used as a systematic uncertainty on the isolation efficiency, ranging from 0.5% to 7% at high p_T^{e1}, which is a conservative estimate, since the differences are mostly based on the MC statistics being too low. The $m_{\gamma\gamma}$ dependence of the track-based isolation efficiency is negligible, as shown in Figure 6.24(b), and the difference between data and MC varies from 0.2% to 1% with p_T^{e1}. A background subtraction of the QCD contribution around the Z boson peak in data is performed, but isolation efficiencies obtained before the subtraction are very close to the efficiencies when no subtraction is performed, in the end, the values computed after the subtraction are used.

To obtain the total effect of these uncertainties on the $X \to \gamma\gamma$ signal, a set of pseudo-experiments is performed, generating twenty thousand pseudo-experiments for each event passing all selection cuts. The isolation energy is shifted by a random value from a Gaussian distribution with a width equal to the systematic uncertainty of the efficiencies. The RMS of the



(a) (b)

Figure 6.24: Efficiency of the isolation cuts as a function of electron E_T in $Z \to ee$ MC (blue line) and data with (black line) and without (red line) background subtraction. Plot (a) is for the 6 GeV calorimeter cut, and plot (b) is for the 2.6 GeV track isolation cut. The yellow band shows the systematic error, determined from the size of the data/MC discrepancies [207].

variations over all pseudo-experiments, combining in quadrature the effects from both photons and both isolation variables, is used to obtain the total effect as a function of m_X. The total error is dominated by the calorimeter-based isolation and exhibits a turn-on behavior from 1% at *low-mass* values up to 6% at *high-mass*. A Fermi function defined as

$$\sigma_{isol}(m_X) = \sigma_0 + \Delta\sigma \left(1 + e^{-\frac{m_X - m_0}{\Gamma}}\right) \tag{6.9}$$

is used to fit the shape, with the best-fit values of $\sigma_0 = (1.24 \pm 0.04)\%$, $\Delta\sigma = (3.8 \pm 0.1)\%$, $m_0 = 602 \pm 8$ GeV and $\Gamma = 81 \pm 7$ GeV.

6.7 Fiducial limit extraction

With all the ingredients ready for a final statistical assessment of the data, a maximum likelihood fit is performed, looking for a new narrow resonance in the selected diphoton events.

6.7.1 *Low-mass* analysis

The *low-mass* analysis covers the m_X range of $65 < m_X < 110$ GeV, while the fit is performed over the range of $60 < m_X < 120$ GeV in order for sidebands to be present. The data in the *low-mass* region include the diphoton continuum and the Drell-Yan background, while the SM Higgs background is considered to be negligible. The extended PDF used to describe the data can be written as

$$\mathcal{L} = \prod_{c=1}^{n_c} e^{-N_c^{total}} \prod_{i=1}^{n_c^{data}} \mathcal{L}_c\left(m_{\gamma\gamma}(i,c)\right) \left[\prod_{k=1}^{\dim \boldsymbol{\theta}} \exp\left(-\frac{1}{2}\left(\theta_i - \theta_i^{aux}\right)^2\right)\right] \tag{6.10}$$

where $n_c = 3$ is the number of categories, n_c^{data} is the number of data events, N_c^{total} is the sum of the fitted numbers of events in each component in a category c, and $m_{\gamma\gamma}(i,c)$ is the $m_{\gamma\gamma}$ value for event i of category c. The per-event term can be expressed as

$$\begin{aligned} & \mathcal{L}_c(m_{\gamma\gamma}; \sigma_{fid}, m_{\mathrm{X}}, N_{UU,c}, N_{UC,c}, N_{CU,c}, N_{CC,c}, N_{bkg,c}, \mathbf{c}_c, \boldsymbol{\theta}) \\ = \; & N_{X,c}(\sigma_{fid}, m_{\mathrm{X}}, \boldsymbol{\theta}_{N_X}, \boldsymbol{\theta}_{SS}) f_X(m_{\gamma\gamma}, m_X, \mathbf{x}_X(m_X), \theta_\sigma) \\ + \; & N_{UU,c}(\boldsymbol{\theta}_{N_{UU,c}}) f_{UU,c}(m_{\gamma\gamma}, \mathbf{x}_{UU,c}, \boldsymbol{\theta}_{UU,c}) \\ + \; & N_{UC,c}(\boldsymbol{\theta}_{N_{UC,c}}) f_{UC,c}(m_{\gamma\gamma}, \mathbf{x}_{UC,c}, \boldsymbol{\theta}_{UC,c}) \\ + \; & N_{CU,c}(\boldsymbol{\theta}_{N_{CU,c}}) f_{CU,c}(m_{\gamma\gamma}, \mathbf{x}_{CU,c}, \boldsymbol{\theta}_{CU,c}) \\ + \; & N_{CC,c}(\boldsymbol{\theta}_{N_{CC,c}}) f_{CC,c}(m_{\gamma\gamma}, \mathbf{x}_{CC,c}, \boldsymbol{\theta}_{CC,c}) \\ + \; & N_{bkg,c} f_{bkg,c}(m_{\gamma\gamma}, \mathbf{c}_c) \end{aligned} \tag{6.11}$$

where σ_{fid} is the fiducial production cross section of the new resonance of mass m_{X}; $N_{UU,c}$, $N_{UC,c}$, $N_{CU,c}$, and $N_{CC,c}$ are the number of Drell-Yan background events identified respectively as UU (0-conversion category), UC, CU (1-conversion category) and CC (2-conversion category). $N_{bkg,c}$ is the fitted number of background events and $\mathbf{c}_c$ collectively refers to the Bernstein polynomial parameters used to describe its shape. The nuisance parameters $\boldsymbol{\theta}$ include all uncertainties described in Section 6.6 relevant for the *low-mass* analysis.

The $N_{X,c}$ is the number of events for the new resonance, derived as the product of the total number of events and the category fractions defined in Section 6.3.3 and shown in Figure 6.8. f_X is the signal PDF, defined as a DSCB shape, with parameters $\boldsymbol{x}_X = \{\delta M, \sigma_{CB}, \alpha_{low}, \alpha_{high}, N_{low}, N_{high}\}$. The number of Drell-Yan events in each category is defined as

$$N_i = N_i^0 e^{\left(\sigma_{norm,i}^{stat} \theta_{norm,i}^{stat}\right)} e^{\left(\sigma_{norm,i}^{syst} \theta_{norm,DY}^{syst}\right)} \tag{6.12}$$

for each conversion category component ($i = UU, UC, CU, CC$), where N_i^0 is the normalization of the template, computed as described in Section 6.4.2.2 as $\alpha\rho_1\rho_2$, and $\sigma_{norm,i}^{stat}$ and $\sigma_{norm,i}^{syst}$ are the systematic uncertainties on the template normalization due to statistical and systematic effects. The PDFs of the Drell-Yan components are described as a sum of a DSCB, their nominal values being obtained from a fit of the Drell-Yan templates. Uncertainties are derived on the shape by varying the shift and smearing corrections in the fit, same is applied for the peak position uncertainty.

Finally, the continuum background is described by the background PDF $f_{bkg}(m_{\gamma\gamma,\mathbf{c}})$ in the shape of a Landau and an exponential, as described in Section 6.4.3.

Figure 6.25 shows the $m_{\gamma\gamma}$ distributions in the three conversion categories together with background-only fits, including the continuum background description and the Drell-Yan description with the parameters of the fit set to the best-fit values obtained from the maximum likelihood fit. The agreement of the data with the fits is very good and the Drell-Yan contribution is indeed dominant in the CC category.

The p-value p_0 is defined as the probability that the background can produce a fluctuation greater than or equal to the excess observed in data. The expected and observed p_0 for a background-only hypothesis as a function of the new resonance mass m_{X} is shown in Figure 6.26 for the *low-mass* range and the first part of the high-mass range up to 170 GeV. A $p_0 = 0.5$ corresponds to no signal observed, $p_0 < 0.5$ corresponds to positive and $p_0 > 0.5$ to negative signal. The look-elsewhere effect [236] is not taken into account. No significant excess is observed, therefore a limit is derived.

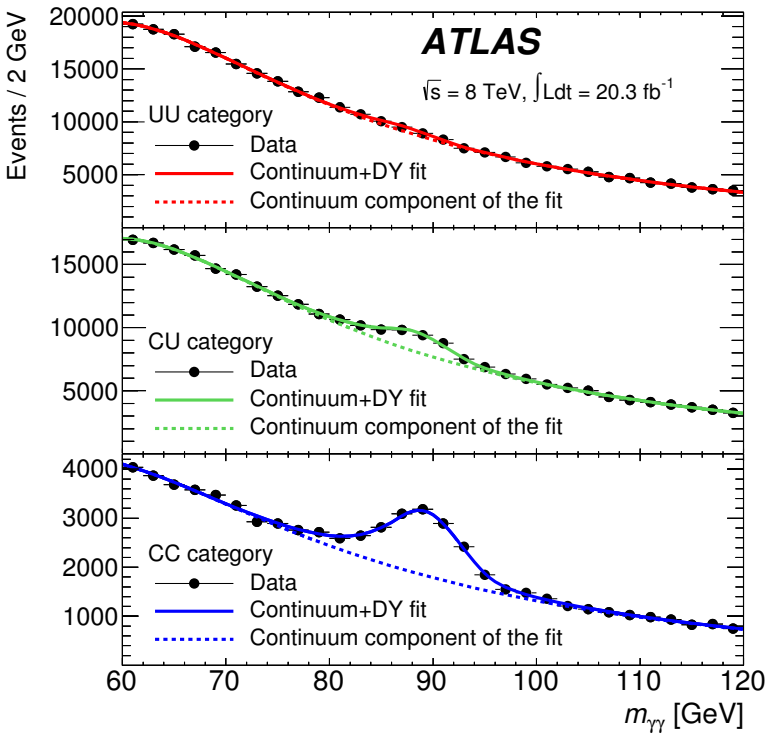


Figure 6.25: Background-only fits to the data (black dots) as functions of the diphoton invariant mass $m_{\gamma\gamma}$ for the three conversion categories in the low-mass range. The solid lines show the sum of the Drell-Yan and the continuum background components. The dashed lines show the continuum background component only [206].

Using the $\tilde{q}_\mu$ estimator, as in previous SM Higgs boson searches [237], a 95% limit on the fiducial cross section σ_{fid} is derived, using the CL_s technique [238], with p-values computed in the asymptotic approximation. A range of cross section hypotheses is checked, scanning the region close to 95% exclusion, determining the exact position of the crossing point of $\sigma_{fid}^{95\%}$ using a quadratic interpolation based on the cross section values with p-values closest to 5%. The expected and observed limits are shown in Figure 6.27, together with the uncertainty bands. The expected limit is higher around m_X= 90 GeV and 126 GeV due to the presence of the Z boson and the Higgs.

6.7.2 *High-mass* analysis

The *high-mass* analysis covers m_X values in the range of $110 < m_X < 600$ GeV. A sliding-window fit technique is used within the range $100 < m_{\gamma\gamma} < 800$ GeV, with a fit window defined in Section 6.4.3.2. An exponential of a second order polynomial is used to fit the continuum background and a DSCB is used for the Higgs boson background around $m_{\gamma\gamma} \approx 126$ GeV. The signal strength used as input to the fit is $\mu = 1$ and the starting point of its mass is $m_H = 125.9$ GeV. The electron background is neglected for this part of the analysis.

The data in the high-mass region are described using the same likelihood form as Eq. 6.10,

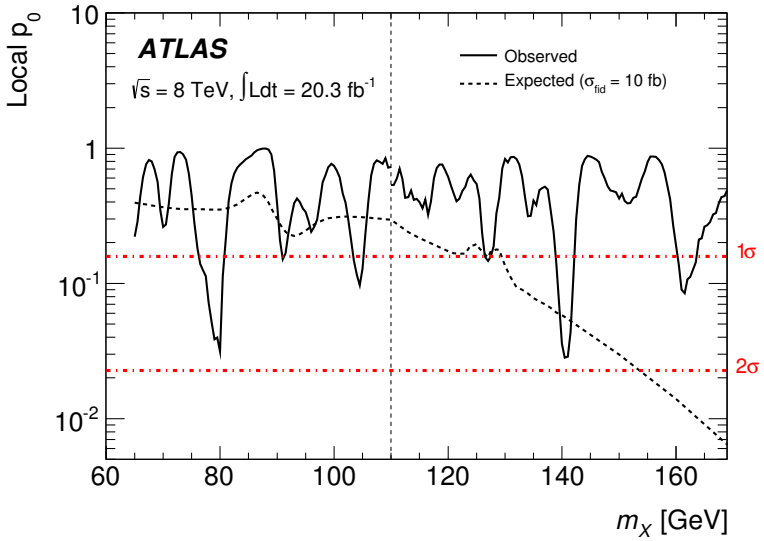


Figure 6.26: Observed and expected (for $\sigma_{fid} = 10$ fb) p-values for the background-only hypothesis p_0, for both the *low-mass* and *high-mass* analyses in the 65-170 GeV mass range, without taking the look-elsewhere effect into account. The dotted-dashed lines indicate the corresponding significance. The discontinuity in the limit at $m_X = 110$ GeV is due to the transition between the *low-* and *high-mass* analyses [206].

uncategorized, described as

$$\begin{aligned}\mathcal{L}(m_{\gamma\gamma};\sigma_{fid},m_X,\mu,m_H,N_{bkg},\xi,\boldsymbol{\theta}) &= N_X(\sigma_{fid},m_X,\boldsymbol{\theta}_{N_X},\theta_{SS})f_X(m_{\gamma\gamma},m_X,\mathbf{x}_X(m_X),\theta_\sigma)\\ &+ N_H(\mu,\boldsymbol{\theta}_{N_H})f_H(m_{\gamma\gamma},m_H,\mathbf{x}_H(m_H),\boldsymbol{\theta}_H)\\ &+ N_{bkg}f_{bkg}(m_{\gamma\gamma},a,b)\end{aligned} \tag{6.13}$$

where σ_{fid} is the fiducial production cross section of the new resonance of mass m_{X}, μ and m_{H} are the strength parameter and the mass of the Higgs boson resonance, N_{bkg} is the fitted number of background events and a and b are background shape parameters, $\boldsymbol{\theta}$ collectively designates the nuisance parameters used to describe the relevant systematic uncertainties. The N_X and N_H are the number of events of the new resonance and the Higgs boson, respectively. The N_{bkg} parameter is left free in the fit. The f_X PDF of the signal is described as a DSCB shape with its uncertainties. Same description is applied for the f_H PDF, with additional systematic uncertainties coming from theoretical predictions of the cross section and branching ratio of the SM Higgs boson. The f_{bkg} PDF is described as an exponential of a second order polynomial in a sliding window defined previously. The overall likelihood including the extended and constant terms is

$$\begin{aligned}&\mathcal{L}(\sigma_{fid},m_X,\mu,m_H,N_{bkg},\xi,\boldsymbol{\theta}) =\\ &e^{-\left(N_X+N_H+N_{bkg}\right)}\left[\prod_{i=1}^{n}\mathcal{L}(m_{\gamma\gamma};\sigma_{fid},m_X,\mu,m_H,N_{bkg},\xi,\boldsymbol{\theta})\right]\left[\prod_{k=1}^{\dim\boldsymbol{\theta}}\exp\left(-\tfrac{1}{2}\left(\theta_i-\theta_i^{aux}\right)^2\right)\right]\end{aligned} \tag{6.14}$$

The $m_{\gamma\gamma}$ distribution from data and three example points of its background fit are shown in Figure 6.28, demonstrating the width of the sliding window range for the fit at the mass points $m_X = 125$, 250 and 500 GeV.

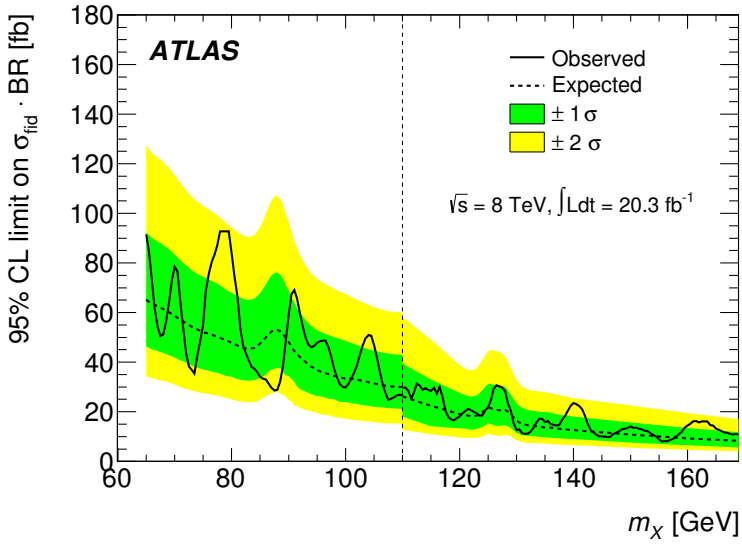


Figure 6.27: Expected and observed limits on the fiducial production cross section $\sigma_{fid} \cdot BR(X \to \gamma\gamma)$ of a narrow resonance as a function of the resonance mass m_X for both the *low-mass* and *high-mass* analyses in the 65-170 GeV mass range. The green and yellow bands show the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limit. The discontinuity in the limit at $m_X = 110$ GeV is due to the transition between the *low-* and *high-mass* analyses [206].

Figure 6.29 shows the background-only hypothesis p_0 value for the *high-mass* analysis range. The two minima at 201 GeV and 530 GeV correspond to observed p_0 values of 2.0% and 1.3% respectively. The local significance of a new resonance at $m_X = 201$ GeV is 2.06 σ. For the mass of $m_X = 530$ GeV, the local significance is 2.22 σ. Including the look-elsewhere effect, the obtained significances are too small to evaluate reliably ($< 0.1\sigma$). Since no significant excess is observed, a limit is derived, shown in Figure 6.30.

6.7.3 Validity of the limits for narrow resonances

The bias on the fitted signal event yield is studied as a function of the natural width of a new resonance X in order to determine the range of resonance widths, where the computed limits are still valid. A naive estimate of the validity would be that any resonance which has a natural width smaller than the detector resolution should be considered as narrow. This assumption is tested using Asimov datasets [229] of an exponential background component and a signal component with a varying width described by a Breit-Wigner shape for masses between 65 and 600 GeV. Each dataset is then fitted with the signal model DSCB described in Section 6.3 and an exponential

$$N_{sig}.DSCB(m_X; \sigma \cdot e^{\theta\epsilon}) + N_{bkg}.e^{\beta m_{\gamma\gamma}}, \tag{6.15}$$

where the number of signal and background events N_{sig} and N_{bkg} are left free in the fit. The slope β of the exponential is set to 0.002 and the DSCB width is allowed to vary within its uncertainty ϵ, which can be up to 40% at higher masses. One thousand signal events are injected for each mass point and the range of acceptable widths is defined so that the bias of the fit would be

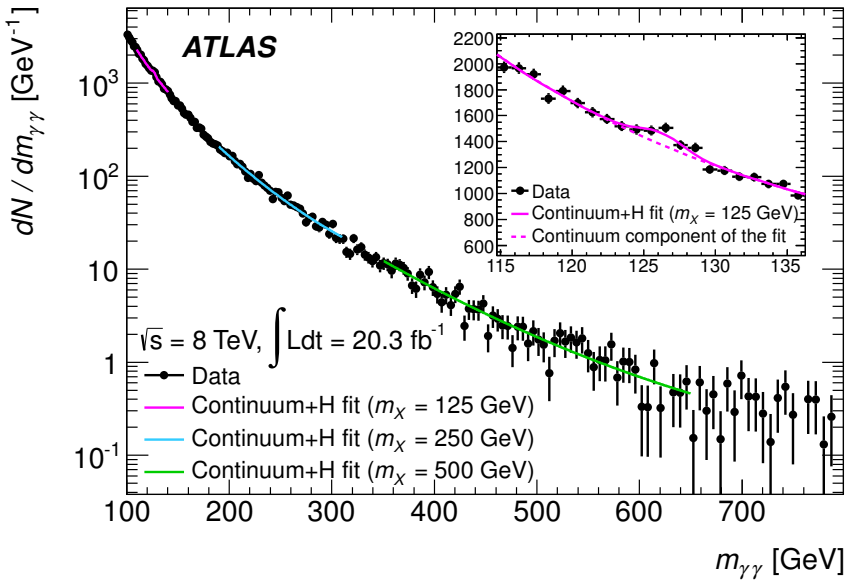


Figure 6.28: Background-only fits to the data (black dots) as a function of the diphoton invariant mass $m_{\gamma\gamma}$ for the inclusive *high-mass* analysis. The solid line shows the sum of the Higgs boson and the continuum background components. The dashed line shows the continuum background component only. Fits of H+continuum are shown for three different mass points [206].

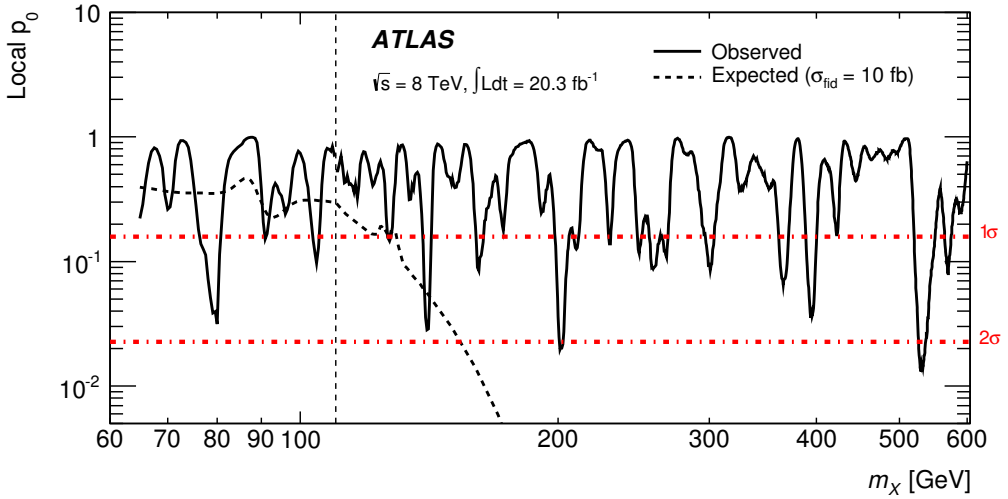


Figure 6.29: Observed and expected (for $\sigma_{fid} = 10$ fb) p-values for the background-only hypothesis p_0, for both the *low-* and *high-mass* analyses in log scale, without taking the look-elsewhere effect into account. The dotted-dashed lines indicate the corresponding significance. The discontinuity in the limit at m_X= 110 GeV is due to the transition between the *low-* and *high-mass* analyses [206].

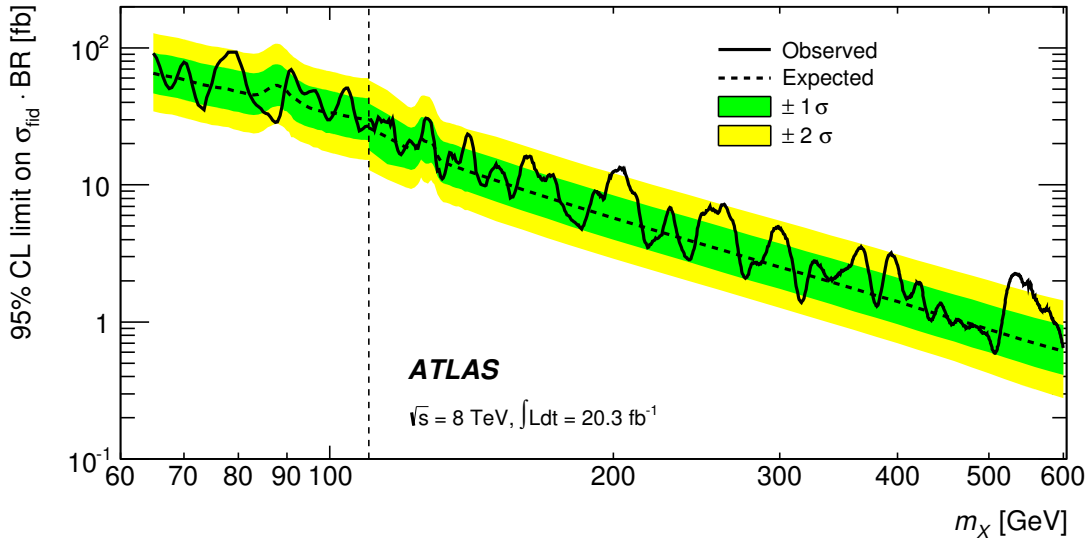


Figure 6.30: Expected (dashed line) and observed (solid line) limits on the fiducial production cross section times branching fraction $\sigma_{fid}.BR(X \to \gamma\gamma)$ of a narrow resonance as a function of the resonance mass m_X for both the *low-* and *high-mass* analyses in log scale. The green and yellow bands show the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limit. The discontinuity in the limit at m_X= 110 GeV is due to the transition between the low- and high-mass analyses [206].

kept below 10%. The obtained widths up to which the limit is considered to be valid, with a bias on the fitted signal yield below 10% are shown in Figure 6.31 as a function of m_X fitted with a line, representing the boundary of valid widths as

$$\Gamma_X < 0.09 \text{ GeV} + 0.01 \ m_X. \tag{6.16}$$

As expected, this corresponds to $\approx 1\% m_X$, which is roughly the experimental diphoton invariant mass resolution in ATLAS.

6.8 Conclusions

The observed and expected 95% C.L. limits, shown in Figures 6.27 and 6.30, are in good agreement, consistent with the absence of a signal. The limits on $\sigma_{fid}BR(X \to \gamma\gamma)$ for an additional scalar resonance range from 90 fb for $m_X = 65$ GeV to 1 fb for $m_X = 600$ GeV.

The word "scalar" is used in the final publication [206], because even though the analysis tries to be as model independent as possible, the MC samples used for signal modelling are generated using a specific model - a narrow but otherwise SM Higgs boson - a spin 0 particle. SM Higgs production processes are chosen to represent a wide spectrum of possible final states. A spin-2 model is not explicitly considered, but that does not mean the limit is of no use for theories including a spin-2 particle. In case of a discovery, the angular properties of the photons observed in the final state would have been studied more closely, in order to determine the spin of the particle.

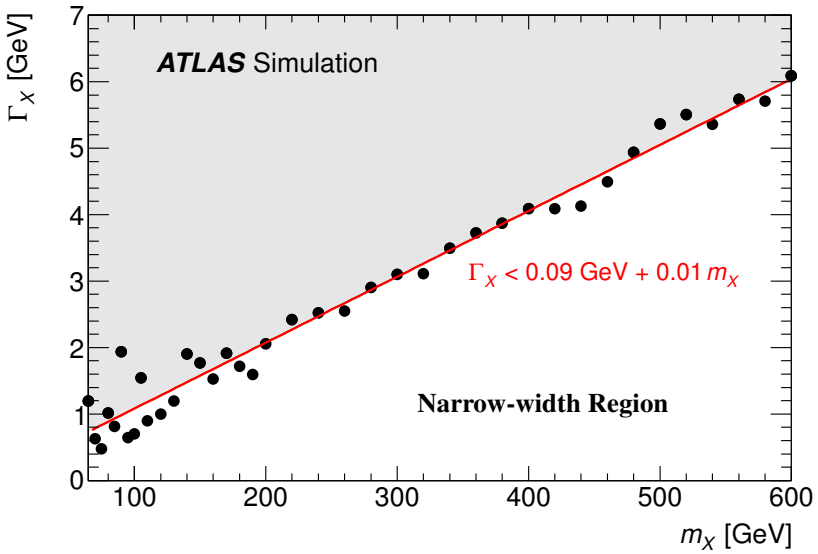


Figure 6.31: Maximal values of natural signal widths compatible with the fiducial limits, as a function of the resonance mass m_X(black dots). The red line is a fit to the black dots [206].

The published limits extend over a considerably wider diphoton invariant mass range than the previous searches performed by the ATLAS [239, 186] and CMS collaborations [240]. The limits produced in this work are the first limits independent of the event topology measured for the diphoton final state. The CMS analysis presented in Ref. [117] is restricted to $150 < m_X < 850$ GeV and did not explore the *low mass* region. The search for exotic resonances in the diphoton channel by ATLAS [118] covers the region $409 < m_X < 3000$ GeV.

A similar analysis is already in preparation within ATLAS for the data being collected during LHC Run 2 and will be performed as soon as the LHC delivers a few fb^{-1} of new data this year, in hope of a potential new discovery in the diphoton channel at $\sqrt{s} = 13$ TeV.

7 Measurement of the differential production cross sections of diphotons in association with jets

This chapter presents the measurement of the differential production cross section of diphotons in association with jets at $\sqrt{s} = 8$ TeV in pp collisions, for multiple observables related to the final state objects and in exclusive categories of the number of jets $N_{jet} = 0, 1, 2, \geq 3$.

The theoretical calculations mentioned in Sections 1.5.1-1.5.3 explore many kinematical properties of photons, the diphoton system and properties related to both photons and jets, such as the R-plane separation of photons and jets, $\Delta R_{\gamma j}$. This analysis aims to measure all properties mentioned in Ref. [79, 80, 84, 86, 90], in order to provide a complete comparison of experimental observables to theoretical predictions at the leading versus the next to leading order.

The angular properties of the dijet and the diphoton system are be studied, in particular by looking at the ΔS variable defined as the difference in azimuthal angle of the leading and subleading γj pair or the diphoton and dijet pair.

All observables considered in this analysis are listed in Table 7.1. The definition of the jet multiplicity N_{jet} is discussed in Section 7.2.4.

The final state containing two photons and jets, in terms of a SM cross section measurement, was not studied at $\sqrt{s} = 7$ TeV, due to the small amount of data collected at that c.m. energy. Previous measurements of the diphoton and photon in association with jets final states were listed in Section 1.6.1.

7.1 Overview of the analysis

To measure the cross section of diphotons in association with jets, a selection of photons and jets has to be first developed, based on studies of the efficiencies of individual selections. For photon candidates, the photon E_{T} selection, isolation definition and separation of the two photons in the R-plane are studied in Section 7.2.3. Selection of jets, mostly based on a requirement to reject a large fraction of pile-up jets, is discussed in Section 7.2.4.

Observable	$0-jet$	$1-jet$	$2-jet$	$\geq 3-jet$
$m_{\gamma\gamma}$	✓	✓	✓	✓
$p_{\mathrm{T}}^{\gamma\gamma}$	✓	✓	✓	✓
$\cos\theta^*_{\gamma\gamma}$	✓	✓	✓	✓
$\Delta\phi_{\gamma\gamma}$	✓	✓	✓	✓
$p_{\mathrm{T}}^{jet_1}$		✓	✓	✓
$p_{\mathrm{T}}^{jet_2}$			✓	✓
$p_{\mathrm{T}}^{jet_3}$				✓
m_{jj}			✓	
p_{T}^{jj}			✓	
$\Delta\phi_{jj}$			✓	
Δy_{jj}			✓	
$\Delta R_{\gamma_1 j_1}$		✓	✓	✓
$\Delta R_{\gamma_2 j_1}$		✓	✓	✓
$\Delta R_{\gamma_1 j_2}$			✓	✓
$\Delta R_{\gamma_2 j_2}$			✓	✓
$\Delta S_{\gamma j\gamma j} = \Delta\phi_{\gamma j\gamma j}$			✓	
$\Delta S_{\gamma\gamma jj} = \Delta\phi_{\gamma\gamma jj}$			✓	

Table 7.1: Variables considered for the diphoton in association with jets cross section measurement.

Prompt photons that originated from the hard process can be, to a certain level, separated from π^0 decays into two photons in jets by requiring them to be isolated and to pass the *tight* ID criteria. In ATLAS, several jet background decomposition methods have been used in the past and one of them is implemented for this analysis, as described in Section 7.3.

Electrons that have been misidentified as photons also represent a background, which has been previously estimated to amount to about 4% [96]. The electron to photon and photon to electron fake rates can only be properly estimated around the Z boson mass, where one can study the ee ($ee\gamma$) and $e\gamma$ (eee) distributions and estimate the probability of electron faking a photon (photon faking an electron). The peak is clearly visible in the diphoton spectrum, therefore the fake rate measurement can be validated in data in the Z boson mass peak region. A devoted measurement of the fake rates and the ee and $e\gamma + \gamma e$ yield as a function of all observables is presented in Section 7.4. For masses away from the Z peak, additional uncertainties have to be considered, as well as for other observables where this background can be spread more evenly.

The Higgs component of the $\gamma\gamma$ spectra for all observables is discussed in Section 7.5. Its yield is very small w.r.t. the QCD-induced diphoton yield, therefore it does not explicitly have to be subtracted, and interference between the two components should also be considered.

The unfolding procedure is described in Section 7.6. The Singular Value Decomposition (SVD) approach is adopted in two dimensions, which allows to properly take into account the migration between the N_{jet} categories as well as bin migrations as a function of a given observable.

Systematic uncertainties are evaluated in Section 7.7. The dominant systematic uncertainties come from the non-tight control region definition from the jet background decomposition method and from the photon and jet energy scale.

Theoretical predictions obtained from the group in Ref. [80] and the process of comparing them to experimental results are addressed in Section 7.8.

Finally, the results are presented in Section 7.9 and compared to MC predictions computed using diphoton samples of PYTHIA (up to two jets in the final state) and SHERPA (up to three jets in the final state).

7.2 Event and object selection

7.2.1 Monte Carlo simulation samples

In order to study true diphoton events and events with two photons and a certain number of jets in the final state, Monte Carlo simulation events are used. After the generation of the hard process, partons which carry color (quarks and gluons) undergo parton showering and hardonization. Stable particles are then passed through the ATLAS detector simulation using GEANT4 [75]. The simulated samples contain pile-up to give a realistic description of experimental conditions in the ATLAS detector when collecting data.

Two sets of MC samples are used in this analysis. PYTHIA diphoton samples with up to two jets are generated using PYTHIA8 [71] with LO CTEQ6L1 PDFs [241]. An E_{T} cut on the photons is applied at generation level at $E_{\mathrm{T}}^{\gamma} > 20$ GeV. The second sample is generated using SHERPA with CT10 PDFs and it is actually split into two samples covering the region of $55 < m_{\gamma\gamma} < 80$ GeV and $m_{\gamma\gamma} > 80$ GeV. The number of events, cross sections and event filter efficiencies are summarized in Table 7.2.

MC sample	events	cross section [nb]	EFE	luminosity [fb^{-1}]
Pythia	9 999 576	$1.38.10^{2}$	$6.48.10^{-4}$	111.8
Sherpa $55 < m_{\gamma\gamma} < 80$ GeV	14 887 973	$9.15.10^{-2}$	$3.40.10^{-1}$	478.3
Sherpa $80 < m_{\gamma\gamma} < \infty$ GeV	11 954 350	$6.16.10^{-2}$	$3.93.10^{-1}$	493.7

Table 7.2: Simulated MC samples used in the diphoton in association with jets cross section measurement analysis, EFE stands for Event Filter Efficiency.

7.2.2 Data samples

This analysis uses the full pp collision dataset, corresponding to the integrated luminosity of 20.3 ± 2.8 fb^{-1} [131], collected in 2012 at $\sqrt{s} = 8$ TeV by the ATLAS detector at the LHC. Events are required to pass a diphoton `2g20vh_medium` trigger with *medium* shower shape criteria and an E_{T} threshold of 20 GeV for both photon candidates. The efficiency of the `EF_2g20vh_medium` trigger was measured to be $98.66\%^{+0.14}_{-0.16}(stat) \pm 0.10(syst)$ [213]. The *All good* GRL is applied, as described in Section 2.2.1.6, removing runs and *lumi blocks*, where not all ATLAS subsystems were fully operational. Furthermore, events are required not to contain any noise bursts or data integrity errors in the LAr and Tile Calorimeters, as defined in Section 2.2.1.6. At least one primary vertex with at least three tracks must be present and there must be at least two reconstructed photon candidates. The number of events after passing each selection criterion are listed in Table 7.3.

Selection	Number of events
All	71 749 692
Trigger	27 882 247
GRL	26 623 904
No calorimeter errors	26 566 873
At least one PV	26 566 698
PV with 3 tracks	26 566 672
At least two reconstructed photons	19 767 185

Table 7.3: Number of events passing the event selection criteria.

7.2.3 Photon candidate selection

The reconstructed energy of the photon candidates is initially calibrated as described in Chapter 5. Subsequently, the transverse momentum of each candidate is computed from the calibrated cluster energy and the pseudorapidity measured in the second layer of the calorimeter η_{S2} as

$$E_{\mathrm{T}}^{\gamma,calib} = \frac{E^{\gamma,calib}}{\cosh(\eta_{S2}^{\gamma})}. \tag{7.1}$$

Each candidate has to satisfy $E_{\mathrm{T}}^{\gamma} > 20$ GeV in order to be further considered. The pseudorapidity acceptance is defined as

$$|\eta_{S2}| < 1.37 \quad \text{or} \quad 1.56 \leq |\eta_{S2}| < 2.37, \tag{7.2}$$

with the transition region of $1.37 \leq |\eta| < 1.56$ between the barrel and the end-cap calorimeters removed due to sub-optimal calibration performance. The pseudorapidity of the second layer of the calorimeter for each photon candidate cluster center is used. Photon candidates are also required to pass *loose* shower shape criteria, the ambiguity resolver (defined in Section 3.3) and basic photon cleaning, such as the HV spikes removal, as described in Section 3.5.

If preselected, photon candidates are ordered in E_{T}^{γ} and the two leading photons are retained. The diphoton vertex is determined using the *photon pointing* method [210], by combining the trajectories of both photons and looking for the closest vertex when extrapolating to the beam in the center of the detector, as explained in Section 6.2. After finding the vertex, the η^{γ} position of the photons is recomputed w.r.t. to the new vertex.

Using the newly computed η^{γ}, the E_{T}^{γ} of the photons is also recomputed, according to Eq. 7.1. The leading and subleading photon candidates are reordered in case the corrected $E_{\mathrm{T}}^{\gamma 2} > E_{\mathrm{T}}^{\gamma 1}$.

Both photon candidates are required to pass $E_{\mathrm{T}}^{\gamma} > 22$ GeV. This cut is chosen due to the trigger used, in order to avoid being too close to the trigger threshold, since the trigger efficiency at the threshold is lower than away from the threshold, but at the same time to be able to measure the largest possible phase space. The choice of E_{T}^{γ} cuts being the same for both photons, may create instabilities in the fixed order NLO theoretical calculations in the $0-jet$ category, in terms of infrared logarithmic divergencies when the E_{T}^{γ} is close to the E_{T}^{γ} threshold and one integrates over a third soft parton, as described in [242] and [243].

For this reason, most diphoton cross section measurements, if inclusive and dominated by the $0-jet$ category, apply an asymmetric cut on the E_{T}^{γ} of the two photons. This analysis however aims to measure the diphotons in association with jets and this problem would only arise in the $0-jet$ category. Since the focus of this analysis lies on the presence of jets in the event, the

symmetric cuts can be applied. The downside of this decision is that the theoretical predictions will have large uncertainties in the $0 - jet$ category. The choice of the diphoton trigger and subsequently the E_{T}^{γ} cuts was also motivated by increasing the statistical power of the analysis, due to very low statistics in the measured data when requiring the presence of two or three or more jets.

A separation in the R-plane, defined in Eq. 2.6, between the two photons is required to be

$$\Delta R_{\gamma\gamma} > 0.4, \tag{7.3}$$

due to the overlapping isolation cones of the two photons, where the energy deposits around one photon could pollute the isolation cone of the other and result in a lower photon isolation efficiency. A profile plot of the calorimeter based isolation energy $E_{\mathrm{T}}^{iso,topo}$ as a function of the separation $\Delta R_{\gamma\gamma}$ is shown in Figure 7.1 for the leading and subleading photon using SHERPA MC simulation. A visible drop in the mean isolation energy is seen after $\Delta R_{\gamma\gamma} > 0.4$ for both generators and both the leading and subleading candidate.

Further, the diphoton invariant mass is required to be $m_{\gamma\gamma} > 60$ GeV due to a cut of $m_{\gamma\gamma} > 55$ GeV applied at generation level in the SHERPA MC simulation sample used for this analysis. The cut in data has to be applied in order to be able to perform the unfolding procedure using this MC sample.

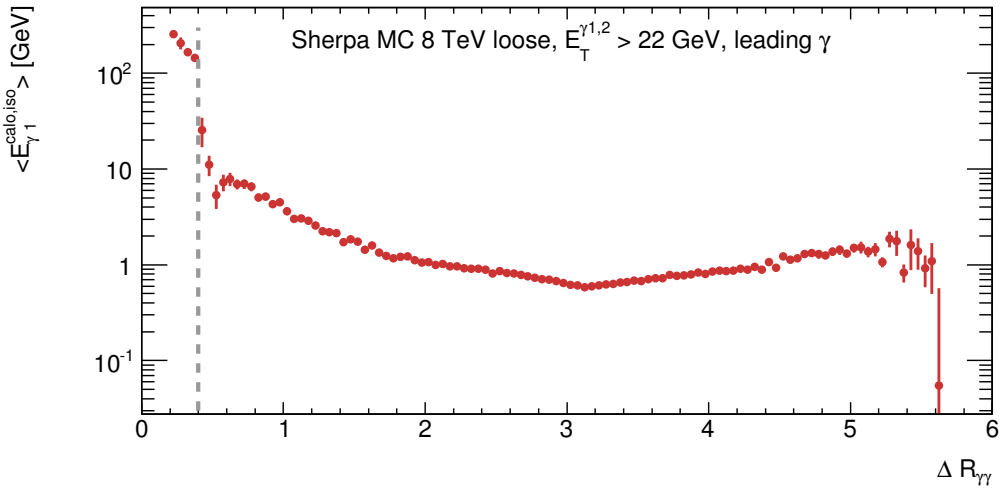
Photon candidates are also required to pass the *tight* shower shape criteria, defined in Section 3.3. This selection partially rejects the background coming from jets but also removes part of the signal.

Both photon candidates are required to pass a calorimeter based isolation requirement, namely to have $E_{\mathrm{T}}^{iso,calo} < 4$ GeV.

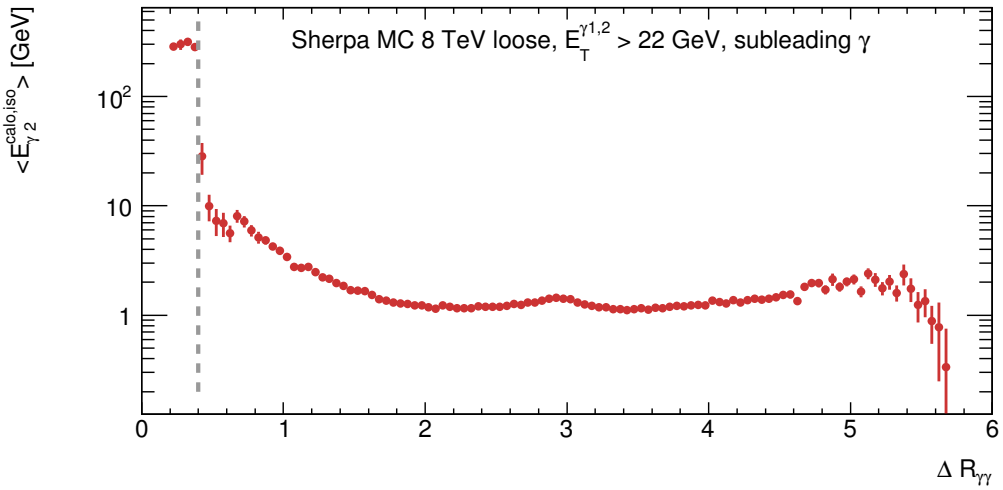
To determine the residual background coming from jets, after requiring the photons to be tight and isolated, the jet background decomposition methods can be used, as described in Section 7.3. The purity of the *tight* and isolated region was usually around 67-77% in previous measurements, depending on the selection criteria [96, 206]. The number of events retained after passing each photon selection is shown in Table 7.4.

Selection	Cut	Number of events
Event selection	Sec. 7.2.2	19 767 185
Preselection	above	10 584 833
$\Delta R_{\gamma\gamma}$	> 0.4	10 258 459
E_{T}^{γ}	> 22 GeV	7 695 170
$m_{\gamma\gamma}$	> 60 GeV	5 666 379
$\Delta R_{\gamma j}$	> 0.6	5 486 467
tight ID criteria	Sec. 3.3	1 590 358
$E_{\mathrm{T}}^{iso,calo}$	< 4 GeV	792 053

Table 7.4: Number of events passing each selection criteria (and the ones before them) related to photons.



(a)



(b)

Figure 7.1: (a) Mean calorimeter based isolation energy of the leading and (b) subleading photon candidate as a function of the photon separation $\Delta R_{\gamma\gamma}$, after passing the full selection except the isolation, *tight* ID and $\Delta R_{\gamma\gamma}$ requirements.

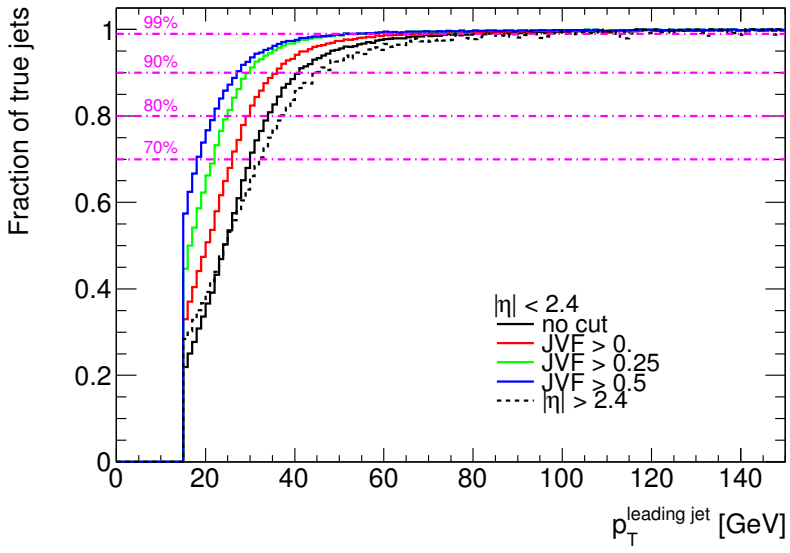
7.2.4 Jet selection

Jets used in this analysis have been reconstructed using the EM energy scale starting from topological clusters, and calibrated as described in Chapter 4. Jets are considered in the region of $|\eta| < 4.4$. Jets from regions of the calorimeter where some cells have been masked are removed.

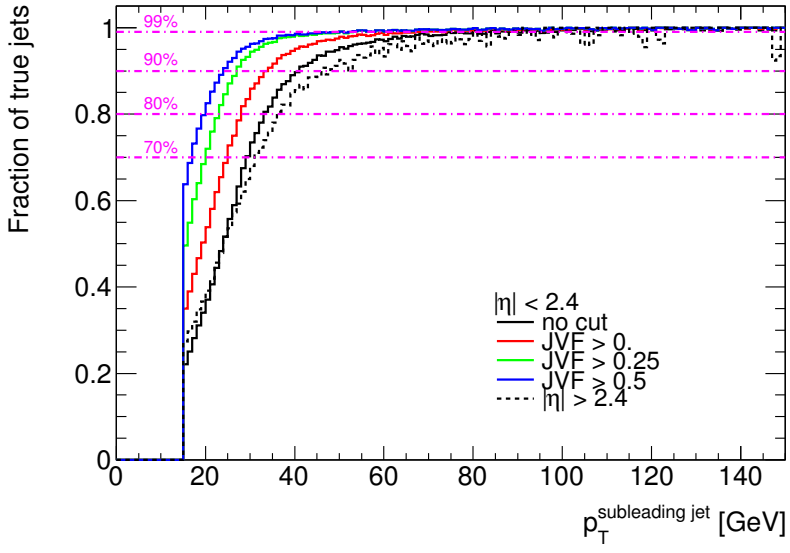
The jets are also required to be separated from the two selected photons by $\Delta R_{\gamma j} > 0.4$, in order to remove jets that are in fact photons that have been also reconstructed as jets. The p_{T}^{jet} selection of the jets is determined by studying SHERPA diphoton plus up to three jets MC events and looking at the fraction of true jets coming from the hard process of that event. The fraction

of true jets is required to be above 90% and the p_T^{jet} cut on the jets are determined according to Figure 7.2.

Figure 7.2 shows the fraction of true hard-scatter jets as a function of the leading and subleading jet candidate p_{T}. By requiring a purity of 90%, the p_{T}^{jet} threshold for the jets is determined.



(a)

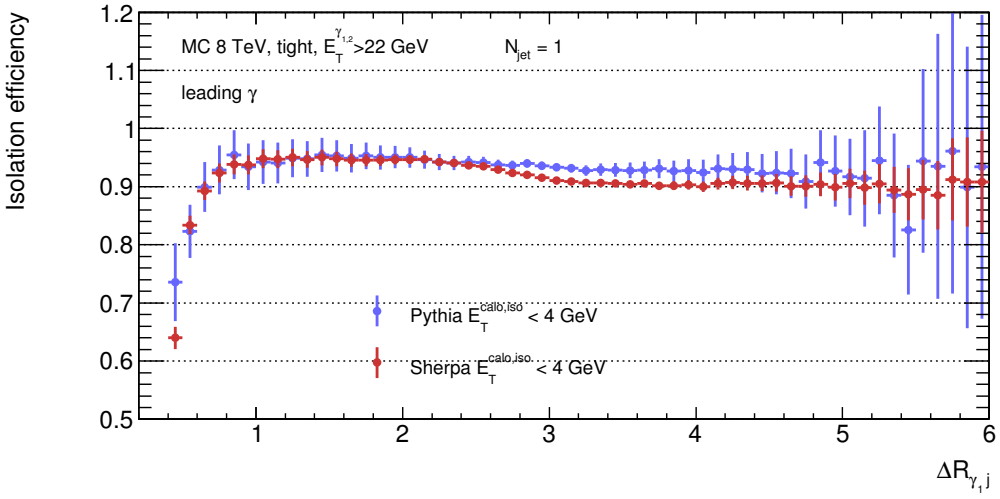


(b)

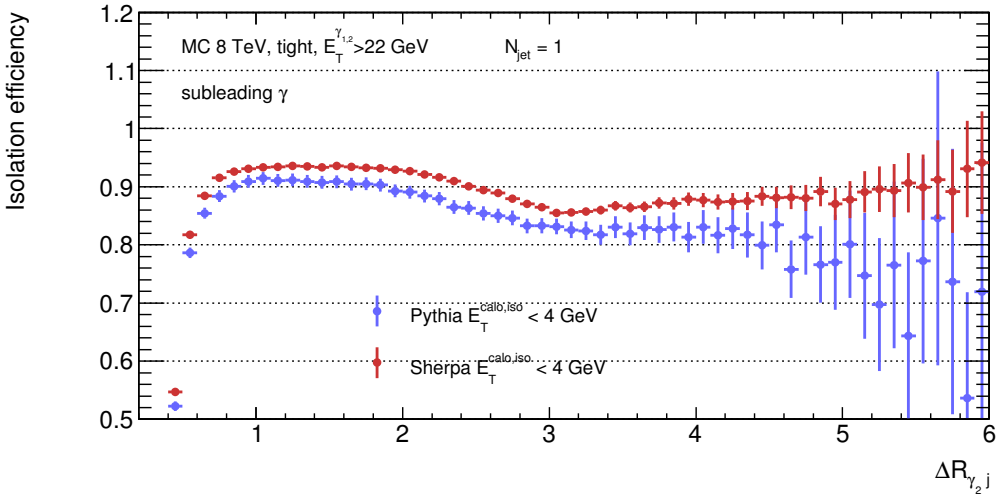
Figure 7.2: Fraction of hard-scatter jets for different η regions and different JVF selection criteria as a function of the minimal cut on the jet p_{T} for the (a) leading and the (b) subleading jet for events passing the diphoton selection [179].

The calibrated jets are required to pass $p_{\mathrm{T}}^{jet} > 25$ GeV in the region of $|\eta| < 2.4$. In this region

the JVF>0.5 cut (as defined in Section 4.3) is applied for jets with $p_{\mathrm{T}}^{jet} < 50$ GeV in order to reduce presence of pileup jets. Jets with $|\eta| > 2.4$, are required to pass $p_{\mathrm{T}}^{jet} > 50$ GeV, since there is no tracking information available for these jets, the JVF selection cannot be applied. To maintain the same requirement on the purity to be above 90%, a step in the p_{T}^{jet} is created. After passing these selections, a list of *good jets* is derived. If any of these jets have $\Delta R_{\gamma j} < 0.6$, the event is removed. The photon and jet separation was studied using MC simulation events. The study on MC uses both PYTHIA and SHERPA MC events.A selection of events, where both photons pass *tight* ID criteria is used to compute the isolation efficiency of the leading and subleading photon as a function of the separation of the photon and the closest jet. This is shown in Figures 7.3-7.5 for the $1-jet$, $2-jet$ and $\geq 3-jet$ categories, respectively.

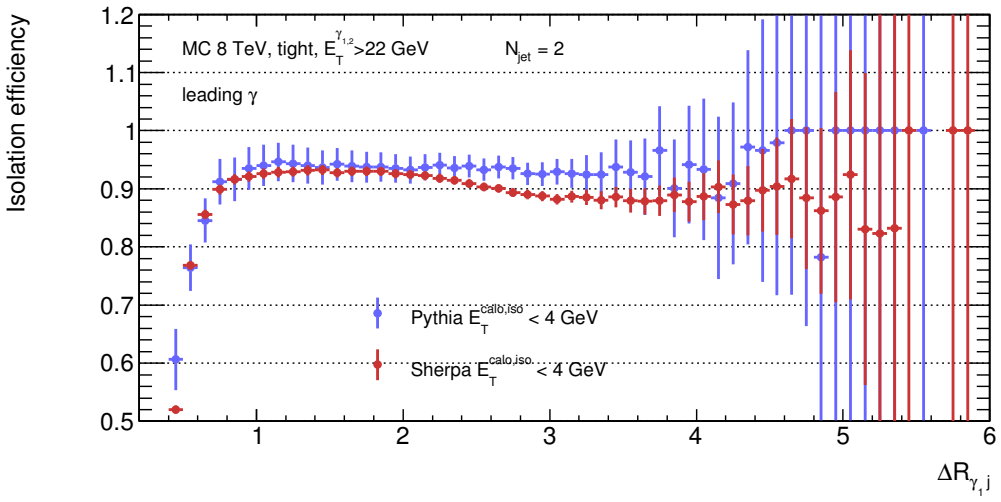


(a)

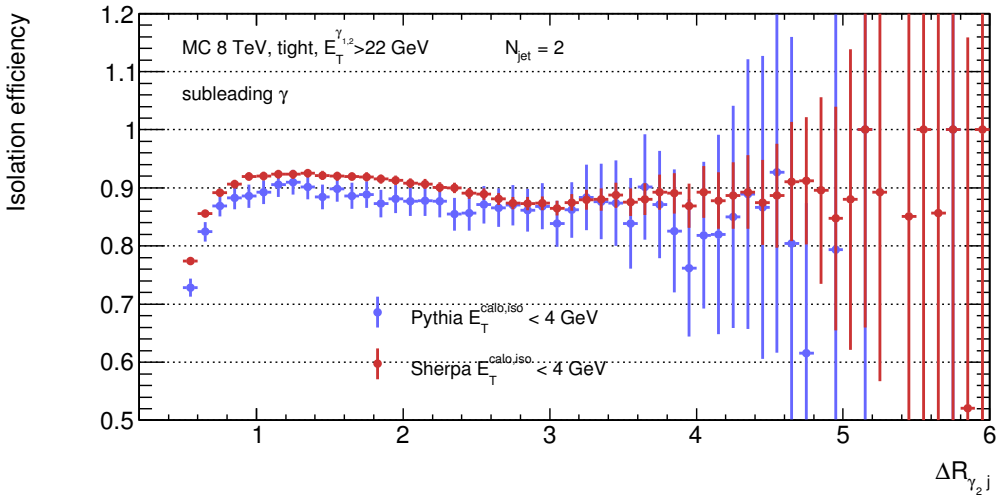


(b)

Figure 7.3: Isolation efficiency of the leading (a) and subleading (b) photon as a function of the R plane separation between the closest jet and the corresponding photon in the $1-jet$ category.



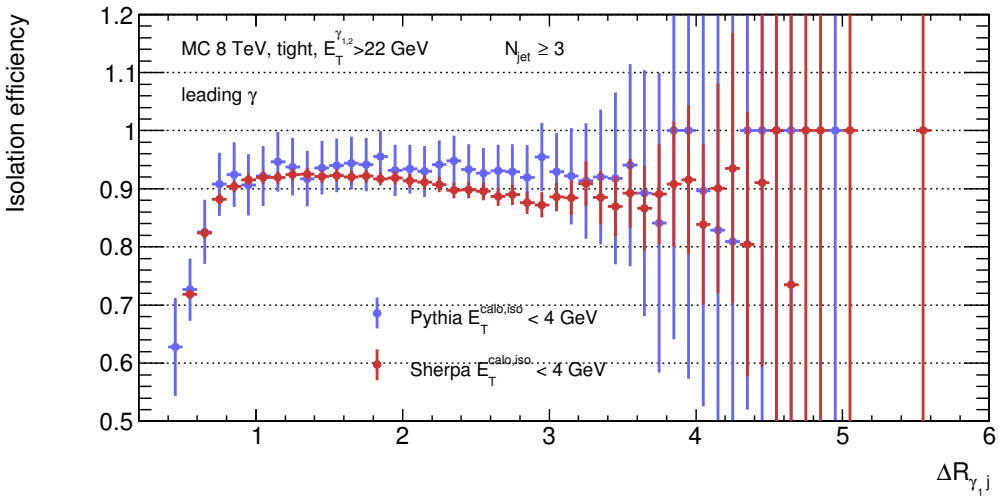
(a)



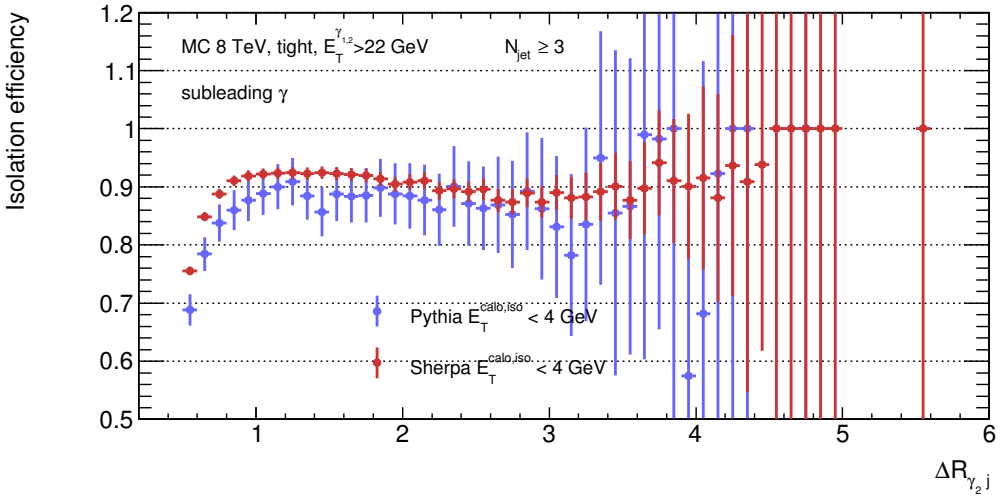
(b)

Figure 7.4: Isolation efficiency of the leading (a) and subleading (b) photon as a function of the R plane separation between the closest jet and the corresponding photon in the $2-jet$ category.

A significant drop of the isolation efficiency is observed for $\Delta R_{\gamma j} < 0.8$, which is even more pronounced for $\Delta R_{\gamma j} < 0.6$, where also a large difference between the two MC generators is observed. Therefore, one cannot properly rely on their description in this region. Events where a jet passing all the other criteria (*good jet*) has $\Delta R_{\gamma j} < 0.6$ are removed from the measurement. These events are removed from the $0-jet$ category as well, it is not just the jet that is removed, but the full event. The number of events passing the full selection in each N_{jet} category is listed in Table 7.5.



(a)



(b)

Figure 7.5: Isolation efficiency of the leading (a) and subleading (b) photon as a function of the R plane separation between the closest jet and the corresponding photon in the $\geq 3 - jet$ category.

Category	Number of events
$0 - jet$	562 335
$1 - jet$	163 958
$2 - jet$	49 033
$\geq 3 - jet$	12 628

Table 7.5: Number of events where both photon candidates are *tight* and isolated listed for the four N_{jet} categories.

Selection	Value
trigger	`EF_2g20vh_medium`
E_{T}^{γ}	$E_{\mathrm{T}}^{\gamma_{1,2}} > 22$ GeV
pseudorapidity η_{s2}	$\lvert\eta^{\gamma}\rvert < 2.37$, excluding $1.37 < \lvert\eta^{\gamma}\rvert < 1.56$
calo-based isolation	$E_{\mathrm{T}}^{iso,calo} < 4$ GeV
tight ID criteria	as defined in Section 3.3
diphoton invariant mass	$m_{\gamma\gamma} > 60$ GeV
photon separation	$\Delta R_{\gamma\gamma} > 0.4$
photon and jet separation	$\Delta R_{\gamma j} > 0.6$
Jet Vertex Fraction	JVF>0.5
p_{T}^{jet} for $\eta^{jet} < 2.4$	$p_{\mathrm{T}}^{jet} > 25$ GeV
p_{T}^{jet} for $2.4 < \eta^{jet} < 4.4$	$p_{\mathrm{T}}^{jet} > 50$ GeV

Table 7.6: Summary of the analysis event and object selection at reconstruction level. The JVF cut is only applied for $p_{\mathrm{T}}^{jet} < 50$ GeV and $\eta^{jet} < 2.4$.

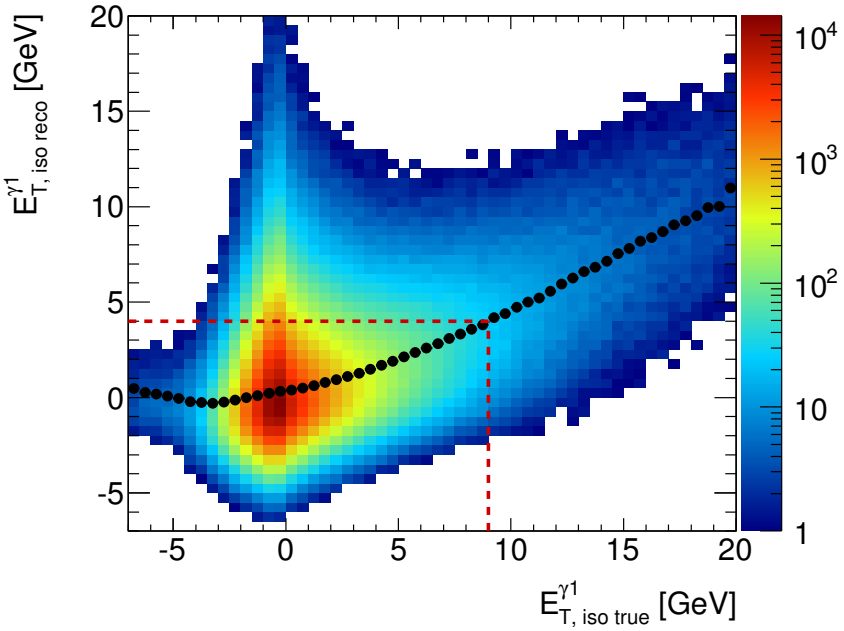
7.2.5 Summary of the selection

The summary of the event and object selection is shown in Table 7.6.

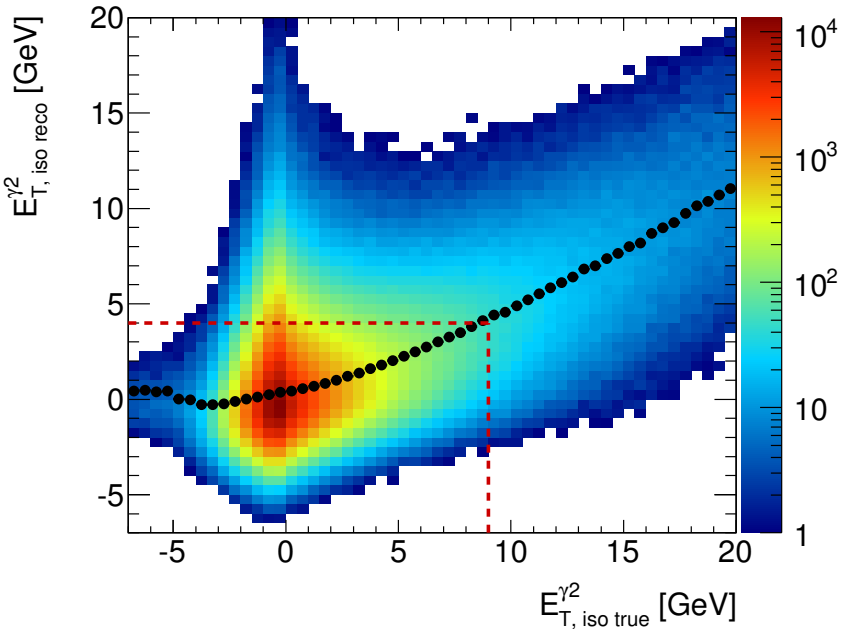
At the truth level, in MC, the same selection is applied as on the reconstruction level, except for the requirement of the trigger, photon isolation, *tight* ID and the JVF selection. The calorimeter-based photon isolation requirement has been investigated w.r.t. the truth particle isolation, defined as the sum of the energy of stable particles in a cone of $\Delta R < 0.4$ around the photon, as seen in Figure 7.6. This shows that the particle isolation which corresponds to the $E_{\mathrm{T}}^{calo,iso} < 4$ GeV criterion is 9 GeV, for the leading and the subleading photon, therefore this value is chosen for the particle isolation in MC. The selection criteria at truth level, defining the fiducial phase space that the measured yields are unfolded to, in Section 7.6, is summarized in Table 7.7.

Selection	Value
E_{T}^{γ}	$E_{\mathrm{T}}^{\gamma_{1,2}} > 22$ GeV
pseudorapidity η_{s2}	$\lvert\eta^{\gamma}\rvert < 2.37$, excluding $1.37 < \lvert\eta^{\gamma}\rvert < 1.56$
particle isolation	$E_{\mathrm{T}}^{iso,true} < 9$ GeV
diphoton invariant mass	$m_{\gamma\gamma} > 60$ GeV
photon separation	$\Delta R_{\gamma\gamma} > 0.4$
photon and jet separation	$\Delta R_{\gamma j} > 0.6$
p_{T}^{jet} for $\eta^{jet} < 2.4$	$p_{\mathrm{T}}^{jet} > 25$ GeV
p_{T}^{jet} for $2.4 < \eta^{jet} < 4.4$	$p_{\mathrm{T}}^{jet} > 50$ GeV

Table 7.7: Summary of the analysis object selection selection at truth level.



(a)



(b)

Figure 7.6: Truth particle isolation versus the calorimeter based isolation energy $E_{\mathrm{T}}^{iso,calo}$ at the reconstruction level in Sherpa MC for (a) the leading photon and (b) the subleading photon, passing the full selection except the isolation requirement at both the truth and reconstruction level. Mean reconstructed calorimeter isolation value in black points.

7.3 Jet background decomposition and diphoton yield extraction

Background mostly coming from π^0 decays or decays of other neutral hadrons into photons inside jets can be estimated by applying a background decomposition method, using *tight* identification and isolation criteria to distinguish photons from events where one or both photon candidates are in fact a jet. Events where isolated electrons have been misidentified as photons are treated together with the diphoton events as signal in this section. The background coming from isolated electrons will be dealt with and subtracted in the following Section 7.4. The diphoton yield obtained here also contains the contribution of the Higgs boson, which will be addressed in Section 7.5.

In ATLAS, three background decomposition methods have been used previously, the 2D Template fit method [96], which implements a likelihood fit of the two photon candidates' isolation energies; the 4x4 Matrix method [242], which relies on computing event-by-event weights; and the 2x2D Sideband method (2x2DSB), which counts events based on the *tight* ID and isolation criteria of the two photon candidates. For this analysis, the 2x2DSB method was used.

Previously, in the diphoton cross section measurement using the full dataset collected at $\sqrt{s} =$ 7 TeV, the 2x2DSB and 2D Template fit methods were used, producing comparable results [96]. The 4x4 Matrix method and the 2x2DSB method were applied in the diphoton cross section measurement using 37 pb^{-1} of $\sqrt{s} = 7$ TeV data [95].

7.3.1 2x2D Sideband method

Events where both photon candidates pass all selection criteria including the *tight* ID and the isolation requirement, are composed of $\gamma - \gamma$, $\gamma - jet$, $jet - \gamma$ and $jet - jet$ events, because the *tight* ID and the isolation selections do not reject 100% of the background. To distinguish the four components, the isolation and *tight* ID properties for events including the ones not passing *tight* ID or isolation criteria are evaluated.

The 2x2DSB method starts from a sample where both photons pass a modified definition of the *tight* ID criteria and no criteria on the isolation energy have been applied.

A set of so-called *loose′* definitions is obtained by removing the requirements of the *tight* criteria one-by-one. Removing the selection criteria on the $w_{s,3}$ and F_{side} variables (defined in Section 3.3), provides a *loose′*2 (**L'2**) selection. Further removing the cut on ΔE gives the *loose′*3 (**L'3**) selection. By omitting also the cut on E_{ratio}, the *loose′*4 selection is obtained. Finally, removing the cut on w_{stot} gives the *loose′*5 (**L'5**) selection. The *tight* selection is therefore a subset of the **L'2** which is a subset of **L'3** and so on. As baseline, the *loose′*4 selection is used and all events taken as input into the 2x2SB method have to pass the **L'4** selection. This choice is, to a certain level, arbitrary. The possible bias due to the choice of the **L'** selection is studied on Pythia Direct Photon sample by comparing the calorimeter-based isolation distributions of true *tight* jets and true *non-tight* jets (photons that have been truth-matched to jets). Ideally, the isolation distribution of true *tight* jets should match the *non-tight* true jets distribution. This way the background decomposition based on the *tight* and isolation properties of the candidates, is able to distinguish singnal from background. In Figure 7.7, the calorimeter-based isolation distribution of the true *tight* and true *non-tight* jets for various **L'** selections is investigated. The ratio w.r.t. the true *tight* jets shows a good agreement between the definitions up to about 8 GeV, but the sample statistics is very limited. Choosing the **L'4** as baseline, the other

selections are used to compute a systematic uncertainty, as described in Section 7.3.2. Table 7.8 lists a summary of the selection criteria for the definition of the **L'** selections and the number of events retained for each selection.

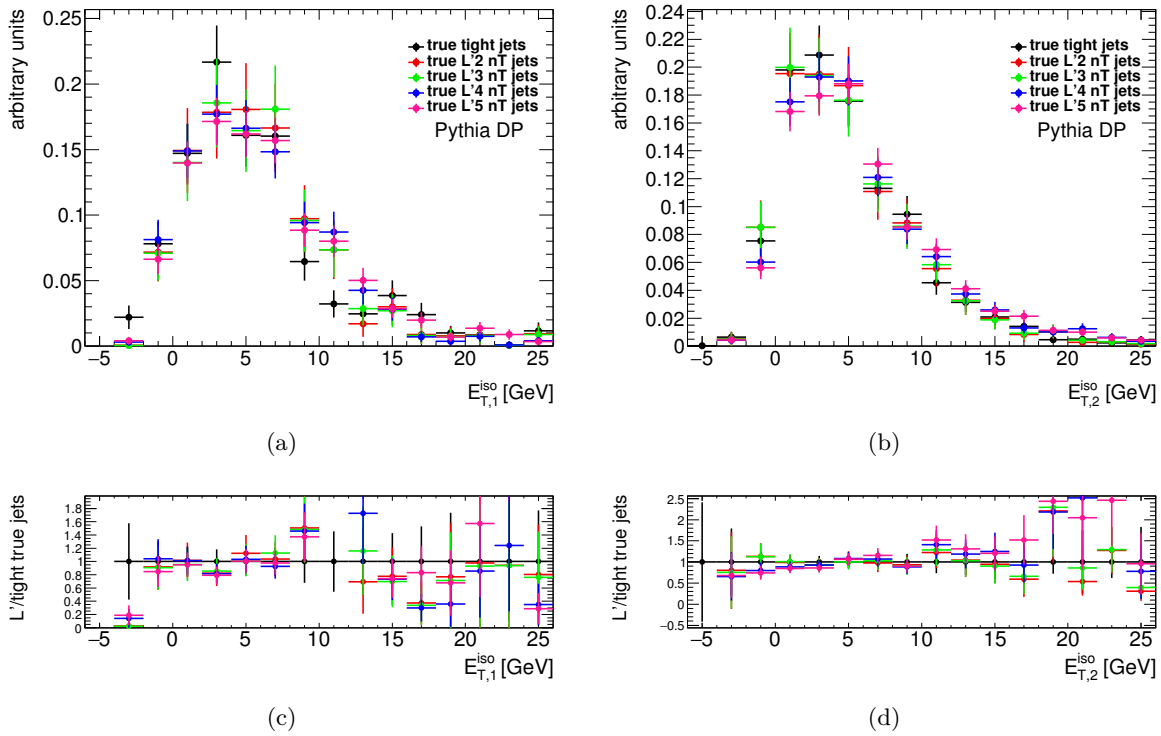


Figure 7.7: Calorimeter-based isolation energy of true *tight* jets (black) and for the four **L'** selections (colors) evaluated on PYTHIA DP MC samples for the (a) leading and (b) subleading jet. Ratio of **L'**to *tight* true jet isolation for the (c) leading and (d) subleading jet.

Sample	Abbrev.	Omitting cut on variable	Number of events
tight	**T**	-	1 590 358
*loose'*2	**L'2**	$w_{s,3}$ F_{side}	2 506 049
*loose'*3	**L'3**	$w_{s,3}$ F_{side} ΔE	2 571 527
*loose'*4	**L'4**(**L'**)	$w_{s,3}$ F_{side} ΔE E_{ratio}	2 856 548
*loose'*5	**L'5**	$w_{s,3}$ F_{side} ΔE E_{ratio} w_{stot}	3 077 436

Table 7.8: Overview of the **L'** selection definitions. No isolation requirement was applied, and the removal of events where a *good jet* has $\Delta R_{\gamma j} < 0.6$ was not required.

Events where both photon candidates pass all requirements listed in Section 7.2 except for the isolation criteria and the *tight* selection, but pass the **L'4** selection, constitute the starting sample for the method, also labeled as the **L'L'** selection. These events contain signal events $W_{\gamma\gamma}^{L'L'}$ and background events $W_{\gamma j}^{L'L'}$, $W_{j\gamma}^{L'L'}$ and $W_{jj}^{L'L'}$ as

$$N_{TOTAL}^{L'L'} = W_{\gamma\gamma}^{\mathrm{L'L'}} + W_{\gamma\mathrm{j}}^{\mathrm{L'L'}} + W_{\mathrm{j}\gamma}^{\mathrm{L'L'}} + W_{\mathrm{jj}}^{\mathrm{L'L'}}, \tag{7.4}$$

To extract the number of diphoton events, one considers a selection of events, where both photon candidates are well identified - passing the *tight* ID criteria and well isolated, passing the $E_{\mathrm{T}}^{iso,topo} < 4$ GeV (**TITI**). The **TITI** events have a different composition of signal and bacgkorund than the **L'L'** events

$$W_{\mathrm{tot}}^{\mathrm{TITI}} = W_{\gamma\gamma}^{\mathrm{TITI}} + W_{\mathrm{j}\gamma}^{\mathrm{TITI}} + W_{\gamma\mathrm{j}}^{\mathrm{TITI}} + W_{\mathrm{jj}}^{\mathrm{TITI}}. \tag{7.5}$$

The W_{ij}^{TITI}**TITI** events can be related to the $W_{ij}^{L'L'}$ **L'L'** events in terms of efficiencies, fake rates and correlation factors. The **L'L'** events can be sorted into sixteen categories based on the *tight* and isolation status of the two candidates, as illustrated in Figure 7.8. Events passing the **L'L'** selection but failing the *tight* (**T**) selection are labeled as $non-tight$ ($\tilde{\mathbf{T}}$). The category labeled as **AA** contains the **TITI** events where most of the signal is located.

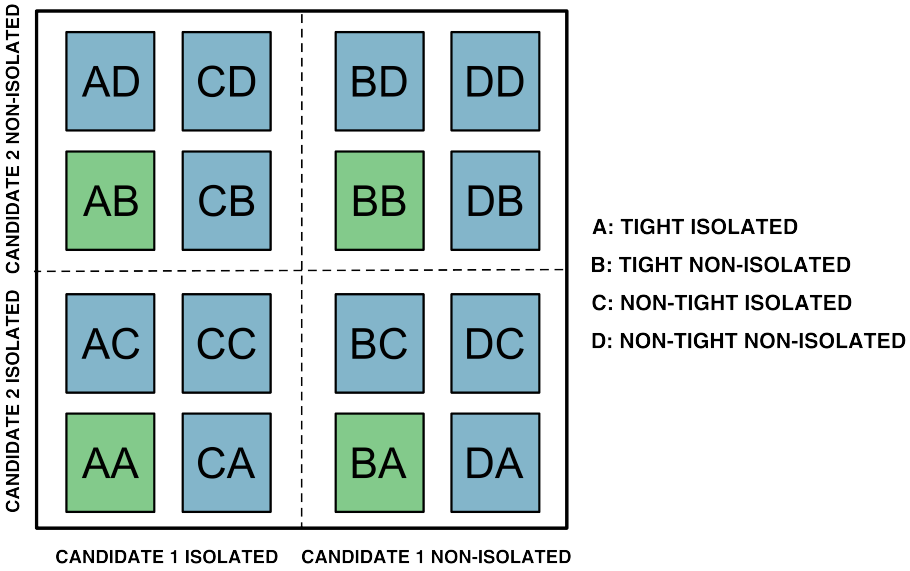


Figure 7.8: Schematic representation of the 2x2DSB method event categories for two photon candidates.

The number of events in each category can then be written as

$$\begin{aligned} N_{AA} &= W_{\gamma\gamma}^{L'L'}\varepsilon_{I1}\varepsilon_{T1}\varepsilon_{I2}\varepsilon_{T2} \\ &+ W_{\gamma j}^{L'L'}\varepsilon_{I1}\varepsilon_{T1}f_{I2}f_{T2}\xi_{j2} \\ &+ W_{j\gamma}^{L'L'}\varepsilon_{I2}\varepsilon_{T2}f_{I1}f_{T1}\xi_{j1} \\ &+ W_{jj}^{L'L'}f'_{I1}f'_{T1}f'_{I2}f'_{T2}\xi_{Ijj}\xi_{j1}\xi_{j2} \end{aligned} \tag{7.6}$$

$$\begin{aligned} N_{AB} &= W_{\gamma\gamma}^{L'L'}\varepsilon_{I1}\varepsilon_{T1}\left(1-\varepsilon_{I2}\right)\varepsilon_{T2} \\ &+ W_{\gamma j}^{L'L'}\varepsilon_{I1}\varepsilon_{T1}\left(1-f_{I2}\xi_{j2}\right)f_{T2} \\ &+ W_{j\gamma}^{L'L'}f_{I1}f_{T1}\left(1-\varepsilon_{I2}\right)\varepsilon_{T2}\xi_{j1} \\ &+ W_{jj}^{L'L'}f'_{I1}f'_{T1}\left(1-f'_{I2}\xi_{Ijj}\xi_{j2}\right)f'_{T2}\xi_{j1} \end{aligned} \tag{7.7}$$

$$
\begin{aligned}
N_{BA} &= W_{\gamma\gamma}^{L'L'}\left(1-\varepsilon_{I1}\right)\varepsilon_{T1}\varepsilon_{I2}\varepsilon_{T2} \\
&+ W_{\gamma j}^{L'L'}\left(1-\varepsilon_{I1}\right)\varepsilon_{T1}f_{I2}f_{T2}\xi_{j2} \\
&+ W_{j\gamma}^{L'L'}\left(1-f_{I1}\xi_{j1}\right)f_{T1}\varepsilon_{I2}\varepsilon_{T2} \\
&+ W_{jj}^{L'L'}\left(1-f'_{I1}\xi_{Ijj}\xi_{j1}\right)f'_{T1}f'_{I2}f'_{T2}\xi_{j2}
\end{aligned} \tag{7.8}
$$

$$
\begin{aligned}
N_{BB} &= W_{\gamma\gamma}^{L'L'}\left(1-\varepsilon_{I1}\right)\left(1-\varepsilon_{I2}\right)\varepsilon_{T1}\varepsilon_{T2} \\
&+ W_{\gamma j}^{L'L'}\left(1-\varepsilon_{I1}\right)\left(1f_{I2}\xi_{j2}\right)\varepsilon_{T1}f_{T2} \\
&+ W_{j\gamma}^{L'L'}\left(1-\varepsilon_{I2}\right)\left(1-f_{I1}\xi_{j1}\right)\varepsilon_{T2}f_{T1} \\
&+ W_{jj}^{L'L'}\left(1-f'_{I1}\xi_{j1}-f'_{I2}\xi_{j2}+f'_{I1}f'_{I2}\xi_{Ijj}\xi_{j1}\xi_{j2}\right)f'_{T1}f'_{T2}
\end{aligned} \tag{7.9}
$$

$$
\begin{aligned}
N_{AC} &= W_{\gamma\gamma}^{L'L'}\varepsilon_{I1}\varepsilon_{T1}\varepsilon_{I2}\left(1-\varepsilon_{T2}\right) \\
&+ W_{\gamma j}^{L'L'}\varepsilon_{I1}\varepsilon_{T1}f_{I2}\left(1-f_{T2}\xi_{j2}\right) \\
&+ W_{j\gamma}^{L'L'}f_{I1}f_{T1}\varepsilon_{I2}\left(1-\varepsilon_{T2}\right)\xi_{j1} \\
&+ W_{jj}^{L'L'}f'_{I1}f'_{T1}f'_{I2}\left(1-f'_{T2}\xi_{j2}\right)\xi_{Ijj}\xi_{j1}
\end{aligned} \tag{7.10}
$$

$$
\begin{aligned}
N_{CA} &= W_{\gamma\gamma}^{L'L'}\varepsilon_{I1}\left(1-\varepsilon_{T1}\right)\varepsilon_{I2}\varepsilon_{T2} \\
&+ W_{\gamma j}^{L'L'}\varepsilon_{I1}\left(1-\varepsilon_{T1}\right)f_{I2}f_{T2}\xi_{j2} \\
&+ W_{j\gamma}^{L'L'}f_{I1}\left(1-f_{T1}\xi_{j1}\right)\varepsilon_{I2}\varepsilon_{T2} \\
&+ W_{jj}^{L'L'}f'_{I1}\left(1-f'_{T1}\xi_{j1}\right)f'_{I2}f'_{T2}\xi_{Ijj}\xi_{j2}
\end{aligned} \tag{7.11}
$$

$$
\begin{aligned}
N_{AD} &= W_{\gamma\gamma}^{L'L'}\varepsilon_{I1}\varepsilon_{T1}\left(1-\varepsilon_{I2}\right)\left(1-\varepsilon_{T2}\right) \\
&+ W_{\gamma j}^{L'L'}\varepsilon_{I1}\varepsilon_{T1}\left(1-f_{I2}-f_{T2}+f_{I2}f_{T2}\xi_{j2}\right) \\
&+ W_{j\gamma}^{L'L'}f_{I1}f_{T1}\xi_{j1}\left(1-\varepsilon_{I2}\right)\left(1-\varepsilon_{T2}\right) \\
&+ W_{jj}^{L'L'}\xi_{j1}f'_{I1}f'_{T1}\left(1-\xi_{Ijj}f'_{I2}-f'_{T2}+f'_{I2}f'_{T2}\xi_{j2}\xi_{Ijj}\right)
\end{aligned} \tag{7.12}
$$

$$
\begin{aligned}
N_{DA} &= W_{\gamma\gamma}^{L'L'}\left(1-\varepsilon_{I1}\right)\left(1-\varepsilon_{T1}\right)\varepsilon_{I2}\varepsilon_{T2} \\
&+ W_{\gamma j}^{L'L'}\left(1-\varepsilon_{I1}\right)\left(1-\varepsilon_{T1}\right)f_{I2}f_{T2}\xi_{j2} \\
&+ W_{j\gamma}^{L'L'}\left(1-f_{I1}-f_{T1}+f_{I1}f_{T1}\xi_{j1}\right)\varepsilon_{I2}\varepsilon_{T2} \\
&+ W_{jj}^{L'L'}\left(1-f'_{I1}\xi_{Ijj}-f'_{T1}+f'_{I1}f'_{T1}\xi_{j1}\xi_{Ijj}\right)f'_{I2}f'_{T2}\xi_{j2}
\end{aligned} \tag{7.13}
$$

$$
\begin{aligned}
N_{BC} &= W_{\gamma\gamma}^{L'L'}\left(1-\varepsilon_{I1}\right)\varepsilon_{T1}\varepsilon_{I2}\left(1-\varepsilon_{T2}\right) \\
&+ W_{\gamma j}^{L'L'}\left(1-\varepsilon_{I1}\right)\varepsilon_{T1}f_{I2}\left(1-f_{T2}\xi_{j2}\right) \\
&+ W_{j\gamma}^{L'L'}\left(1-f_{I1}\xi_{j1}\right)f_{T1}\varepsilon_{I2}\left(1-\varepsilon_{T2}\right) \\
&+ W_{jj}^{L'L'}\left(1-f'_{I1}\xi_{Ijj}\xi_{j1}\right)f'_{T1}f'_{I2}\left(1-f'_{T2}\xi_{j2}\right)
\end{aligned} \tag{7.14}
$$

$$
\begin{aligned}
N_{CB} &= W_{\gamma\gamma}^{L'L'}\varepsilon_{I1}\left(1-\varepsilon_{T1}\right)\left(1-\varepsilon_{I2}\right)\varepsilon_{T2} \\
&+ W_{\gamma j}^{L'L'}\varepsilon_{I1}\left(1-f_{I2}\xi_{j2}\right)\left(1-\varepsilon_{T1}\right)f_{T2} \\
&+ W_{j\gamma}^{L'L'}f_{I1}\left(1-\varepsilon_{I2}\right)\left(1-f_{T1}\xi_{j1}\right)\varepsilon_{T2} \\
&+ W_{jj}^{L'L'}f'_{I1}\left(1-f'_{T1}\xi_{j1}\right)\left(1-f'_{I2}\xi_{Ijj}\xi_{j2}\right)f'_{T2}
\end{aligned} \tag{7.15}
$$

$$
\begin{aligned}
N_{CC} &= W_{\gamma\gamma}^{L'L'}\varepsilon_{I1}\left(1-\varepsilon_{T1}\right)\varepsilon_{I2}\left(1-\varepsilon_{T2}\right)\\
&+ W_{\gamma j}^{L'L'}\varepsilon_{I1}\left(1-\varepsilon_{T1}\right)f_{I2}\left(1-f_{T2}\xi_{j2}\right)\\
&+ W_{j\gamma}^{L'L'}f_{I1}\left(1-f_{T1}\xi_{j1}\right)\varepsilon_{I2}\left(1-\varepsilon_{T2}\right)\\
&+ W_{jj}^{L'L'}f'_{I1}f'_{I2}\left(1-f'_{T1}\xi_{j1}-f'_{T2}\xi_{j2}+f'_{T1}f'_{T2}\xi_{j1}\xi_{j2}\right)\xi_{Ijj}
\end{aligned}
\tag{7.16}
$$

$$
\begin{aligned}
N_{BD} &= W_{\gamma\gamma}^{L'L'}\left(1-\varepsilon_{I1}\right)\varepsilon_{T1}\left(1-\varepsilon_{I2}\right)\left(1-\varepsilon_{T2}\right)\\
&+ W_{\gamma j}^{L'L'}\left(\left(1-\varepsilon_{I1}\right)\varepsilon_{T1}\left(1-f_{I2}-f_{T2}+f_{I2}f_{T2}\xi_{j2}\right)\right.\\
&+ W_{\gamma g}^{L'L'}\left(1-f_{I1}\xi_{j1}\right)f_{T1}\left(1-\varepsilon_{T2}\right)\left(1-\varepsilon_{I2}\right)\\
&+ W_{jj}^{L'L'}f'_{T1}\left(1-f'_{I2}-f'_{T2}-f'_{I1}\xi_{j1}+f'_{I2}f'_{T2}\xi_{j2}+f'_{T2}f'_{I1}\xi_{j1}+f'_{I2}f'_{I1}\xi_{Ijj}\xi_{j1}-f'_{I1}f'_{I2}f'_{T2}\xi_{Ijj}\xi_{j1}\xi_{j2}\right)
\end{aligned}
\tag{7.17}
$$

$$
\begin{aligned}
N_{DB} &= W_{\gamma\gamma}^{L'L'}\left(1-\varepsilon_{I1}\right)\left(1-\varepsilon_{T1}\right)\left(1-\varepsilon_{I2}\right)\varepsilon_{T2}\\
&+ W_{\gamma j}^{L'L'}\left(1-\varepsilon_{I1}\right)\left(1-\varepsilon_{T1}\right)\left(1-f_{I2}\xi_{j2}\right)f_{T2}\\
&+ W_{j\gamma}^{L'L'}\left(1-f_{I1}-f_{T1}+f_{I1}f_{T1}\xi_{j1}\right)\left(1-\varepsilon_{I2}\right)\varepsilon_{T2}\\
&+ W_{jj}^{L'L'}\left(1-f'_{I1}-f'_{T1}-f'_{I2}\xi_{j2}+f'_{I1}f'_{T1}\xi_{j1}+f'_{T1}f'_{I2}\xi_{j2}+f'_{I1}f'_{I2}\xi_{Ijj}\xi_{j2}-f'_{I1}f'_{T1}f'_{I2}\xi_{Ijj}\xi_{j1}\xi_{j2}\right)f'_{T2}
\end{aligned}
\tag{7.18}
$$

$$
\begin{aligned}
N_{CD} &= W_{\gamma\gamma}^{L'L'}\varepsilon_{I1}\left(1-\varepsilon_{T1}\right)\left(1-\varepsilon_{I2}\right)\left(1-\varepsilon_{T2}\right)\\
&+ W_{\gamma j}^{L'L'}\varepsilon_{I1}\left(1-\varepsilon_{T1}\right)\left(1-f_{I2}-f_{T2}+f_{I2}f_{T2}\xi_{j2}\right)\\
&+ W_{j\gamma}^{L'L'}f_{I1}\left(1-f_{T1}\xi_{j1}\right)\left(1-\varepsilon_{I2}\right)\left(1-\varepsilon_{T2}\right)\\
&+ W_{jj}^{L'L'}f'_{I1}\left(\left(1-f'_{T1}\xi_{j1}\right)\left(1-f'_{I2}\xi_{Ijj}\right)-f'_{T2}\left(1-f'_{T1}\xi_{j1}\right)\left(1-f'_{I2}\xi_{j2}\xi_{Ijj}\right)\right)
\end{aligned}
\tag{7.19}
$$

$$
\begin{aligned}
N_{DC} &= W_{\gamma\gamma}^{L'L'}\varepsilon_{I2}\left(\left(1-\varepsilon_{T2}\right)\left(1-\varepsilon_{I1}\right)-\varepsilon_{T1}\left(1-\varepsilon_{T2}\right)\left(1-\varepsilon_{I1}\right)\right)\\
&+ W_{\gamma j}^{L'L'}f_{I2}\left(\left(1-\varepsilon_{I1}\right)\left(1-f_{T2}\xi_{j2}\right)-\varepsilon_{T1}\left(1-\varepsilon_{I1}\right)\left(1-f_{T2}\xi_{j2}\right)\right)\\
&+ W_{j\gamma}^{L'L'}\varepsilon_{I2}\left(\left(1-\varepsilon_{T2}\right)\left(1-f_{I1}\right)-f_{T1}\left(1-f_{I1}\xi_{j1}\right)\left(1-\varepsilon_{T2}\right)\right)\\
&+ W_{jj}^{L'L'}\left(\left(1-f'_{T2}\xi_{j2}\right)\left(1-f'_{I1}\xi_{Ijj}\right)-f'_{T1}\left(1-f'_{T2}\xi_{j2}\right)\left(1-f'_{I1}\xi_{j1}\xi_{Ijj}\right)\right)f'_{I2}
\end{aligned}
\tag{7.20}
$$

$$
\begin{aligned}
N_{DD} &= W_{\gamma\gamma}^{L'L'}\left(1-\varepsilon_{T1}-\varepsilon_{T2}+\varepsilon_{T1}\varepsilon_{T2}+\varepsilon_{I2}\left(1-\varepsilon_{T1}-\varepsilon_{T2}+\varepsilon_{T1}\varepsilon_{T2}\right)+\varepsilon_{I1}\left(-\left(1-\varepsilon_{I2}\right)\left(1-\varepsilon_{T1}\right)+\varepsilon_{T2}\left(1-\varepsilon_{I2}\right)\left(1-\varepsilon_{T1}\right)\right)\right)\\
&+ W_{\gamma j}^{L'L'}\left(1-\varepsilon_{T1}-f_{T2}+\varepsilon_{T1}f_{T2}+f_{I2}\left(1-\varepsilon_{T1}-f_{T2}\xi_{j2}+\varepsilon_{T1}f_{T2}\xi_{j2}\right)+\varepsilon_{I1}\left(\left(1-\varepsilon_{T1}\right)\left(1-f_{I2}\right)+f_{T2}\left(1-f_{I2}\xi_{j2}\right)\left(1-\varepsilon_{T1}\right)\right)\right)\\
&+ W_{j\gamma}^{L'L'}\left(1-\varepsilon_{T2}-f_{T1}+\varepsilon_{T2}f_{T1}+\varepsilon_{I2}\left(1-f_{T1}-\varepsilon_{T2}+\varepsilon_{T2}f_{T1}\right)+f_{I1}\left(\left(1-\varepsilon_{I2}\right)\left(1-f_{T1}\xi_{j1}\right)+\varepsilon_{T2}\left(1-\varepsilon_{I2}\right)\left(1-f_{T1}\xi_{j1}\right)\right)\right)\\
&+ W_{jj}^{L'L'}\left(1-f'_{T1}-f'_{T2}+f'_{T1}f'_{T2}+f'_{I2}\left(1-f'_{T1}-f'_{T2}\xi_{j2}+f'_{T1}f'_{T2}\xi_{j2}\right)\right)\\
&+ W_{jj}^{L'L'}\left(f'_{I1}\left(-\left(1-f'_{I2}\xi_{Ijj}\right)\left(1-f'_{T1}\xi_{j1}\right)+f'_{T2}\left(1-f'_{I2}\xi_{Ijj}\xi_{j2}\right)\left(1-f'_{T1}\xi_{j1}\right)\right)\right)
\end{aligned}
\tag{7.21}
$$

In these sixteen equations, there are nineteen unknowns, therefore the system is underconstrained. Out of the nineteen, six parameters are provided as input and are only varied within their uncertainties in the 2x2DSB fit:

- ε_{T1} and ε_{T2} are the *tight* identification efficiencies for the leading and subleading photons respectively. They are determined from MC simulation of the diphoton signal.
- ε_{I1} and ε_{I2} are the isolation efficiencies for the leading and subleading photons respectively, can be determined from MC or measured in data.

or are fixed:

- ξ_{j1} and ξ_{j2} are the correlation between the identification and the isolation fake rates for the leading and subleading jets. These parameters are fixed to 1.

The other 13 unknowns are outputs from the 2D sideband method, namely:

- $W_{\gamma\gamma}^{L'L'}$, $W_{\gamma j}^{L'L'}$, $W_{j\gamma}^{L'L'}$ and $W_{jj}^{L'L'}$ are the yields for the four categories of events after the LOOSE' preselection.
- f_{T1} and f_{T2} are the *tight* identification fake rates for the leading and subleading jets in γ-jet and jet-γ events.
- f'_{T1} and f'_{T2} are the *tight* identification fake rates for the leading and subleading jets in jet-jet events.
- f_{I1} and f_{I2} are the isolation fake rates for the leading and subleading jets in γ-jet and jet-γ events.
- f'_{I1} and f'_{I2} are the isolation fake rates for the leading and subleading jets in jet-jet events.
- ξ_{Ijj} is the correlation between the isolation of the jets in the jj component.

A fit method is applied using all of the above sixteen equations 7.6-7.21 to extract the thirteen unknowns: the yields ($W_{\gamma\gamma}^{\mathrm{L'L'}}$, $W_{\gamma\mathrm{j}}^{\mathrm{L'L'}}$, $W_{\mathrm{j}\gamma}^{\mathrm{L'L'}}$, $W_{\mathrm{jj}}^{\mathrm{L'L'}}$), fake rates ($f_{I1}$, f_{I2}, f'_{I1}, f'_{I2}, f_{T1}, f_{T2}, f'_{T1}, f'_{T2}) and the jj correlation ξ_{Ijj}. The **TITI** event yields for each component can be than extracted as

$$W_{\gamma\gamma}^{\mathrm{TITI}} = W_{\gamma\gamma}^{\mathrm{L'L'}}\epsilon_{I1}\epsilon_{T1}\epsilon_{I2}\epsilon_{T2} \tag{7.22}$$

$$W_{\gamma\mathrm{j}}^{\mathrm{TITI}} = W_{\gamma\mathrm{j}}^{\mathrm{L'L'}}\epsilon_{I1}\epsilon_{T1}f_{I2}f_{T2}\xi_{j2} \tag{7.23}$$

$$W_{\mathrm{j}\gamma}^{\mathrm{TITI}} = W_{\gamma\mathrm{j}}^{\mathrm{L'L'}}\epsilon_{I2}\epsilon_{T2}f_{I1}f_{T1}\xi_{j1} \tag{7.24}$$

$$W_{\mathrm{jj}}^{\mathrm{TITI}} = W_{\mathrm{jj}}^{\mathrm{L'L'}}f'_{I1}f'_{T1}f'_{I2}f'_{T2}\xi_{Ijj}\xi_{j1}\xi_{j2}. \tag{7.25}$$

7.3.1.1 Data driven isolation efficiencies

The isolation efficiencies of the two photon candidates are provided as input to the 2x2DSB method. These efficiencies can be either measured in data or taken from MC simulation as

$$\epsilon_{I1} = \frac{N_{XITI}}{N_{XITI} + N_{X\tilde{I}TI}} = \frac{N_{AA} + N_{CA}}{N_{AA} + N_{CA} + N_{BA} + N_{DA}} \tag{7.26}$$

$$\epsilon_{I2} = \frac{N_{TIXI}}{N_{TIXI} + N_{TIX\tilde{I}}} = \frac{N_{AA} + N_{AC}}{N_{AA} + N_{AC} + N_{AB} + N_{AC}}, \tag{7.27}$$

where X means any, I means isolated and $\tilde{I}$ means non-isolated. Isolation distributions of photons from MC have been observed to be shifted w.r.t. those from data in previous measurements [96] and this shift was also η-dependent. Therefore for this analysis, the data-driven isolation efficiencies are used in order to eliminate the dependence on MC simulation.

The data-driven efficiencies are computed as follows. To compute the isolation efficiency of one photon candidate, the other is required to be *tight* and isolated. The isolation distribution of the *tight* and *non-tight* candidates is built and the *non-tight* distribution is normalized to the number of events in the *tight* distribution in a region, where almost no signal is present, $10 < E_{\mathrm{T}}^{iso,topo} < 25$ GeV. Then, the normalized *non-tight* distribution is subtracted from the *tight* distribution and the fraction of events passing the $E_{\mathrm{T}}^{iso,topo} < 4$ GeV selection w.r.t. the total number of events, therefore the isolation efficiency, is computed.

This approach, however is not completely legitimate due to the different relative contribution of signal/background in the two distributions - the *tight* and *non-tight* distribution in the non-isolated region of $10 < E_{\mathrm{T}}^{iso,topo} < 25$ GeV and may result in oversubtraction of background from the *tight* distribution. In order to stay in a region where this ratio is more similar (above 99% of background), an anti-isolation region is determined using the track-based isolation variable $p_{\mathrm{T}}^{iso,track}$, defined in Section 3.4.2. The anti-isolation region is defined as $10 < p_{\mathrm{T}}^{iso,track} < 25$ GeV coincidentally corresponding to the $E_{\mathrm{T}}^{iso,topo}$ anti-isolation region. The boundaries for the track anti-isolation region were determined on MC requiring at the size of the signal contamination in the *non-tight* distribution to be the same as in the tight distribution.

The input efficiencies are extracted as described above and are η^{γ} and E_{T}^{γ}-dependent, shown in Figure 7.9 and 7.10. Due to statistical limitations, the efficiencies are only computed up to $E_{\mathrm{T}}^{\gamma} = 400$ GeV. For photons with higher E_{T}^{γ}, the efficiency of the last E_{T}^{γ} bin is applied. The input efficiencies for the 2x2DSB method have to be presented as a function of all the observables $\epsilon_{Ix}(obs.)$, such as $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$ and so on, where $x = 1, 2$. To obtain these, the SHERPA MC sample is used, where the $\epsilon_{Ix}(obs.)$ are derived by applying the data-driven isolation efficiencies event by event for photons according to their η^{γ} and E_{T}^{γ} and building the distributions for all observables. The final input efficiencies are obtained by dividing the observable distributions where efficiencies have been applied by the distributions with no efficiencies applied.

The data driven input efficiencies for the 2x2DSB method as a function of each observable (bin by bin) and for N_{jet} categories and the leading and subleading candidate separately, are shown in Figures 7.11 and 7.12 for the diphoton and dijet variables in the $2-jet$ category. Figures for the rest of the observables can be found in Appendix A. The isolation efficiencies typically range from 85 to 99% depending on the observable and jet multiplicity. Some of them, for example the isolation efficiency as a function of $m_{\gamma\gamma}$ or $p_{\mathrm{T}}^{\gamma\gamma}$, shown in Figure A.3(a) and A.3(b), show a siginificant difference in evolution as a function of the observable between the data driven and MC efficiencies. However, it has been tested, when performing the 2x2DSB fit, both sets of efficiencies (MC and data driven) give comparable results bin-by-bin, with maximum differences in the yields at the level of 2% in bins with sufficient statistics. The differences in integrated yields are found to be negligible. The integrated (independent on observables) efficiencies are listed in Table 7.9 for each N_{jet} category.

Isolation eff.	$0-jet$	$1-jet$	$2-jet$	$\geq 3-jet$
ϵ_{I1}	0.9710 ± 0.0006	0.9234 ± 0.0011	0.9118 ± 0.0017	0.8787 ± 0.0026
ϵ_{I2}	0.9297 ± 0.0006	0.9194 ± 0.0011	0.9256 ± 0.0018	0.9074 ± 0.0027

Table 7.9: Inclusive data driven input isolation efficiencies for the leading and subleading photons with their statistical uncertainty for exclusive N_{jet} categories.

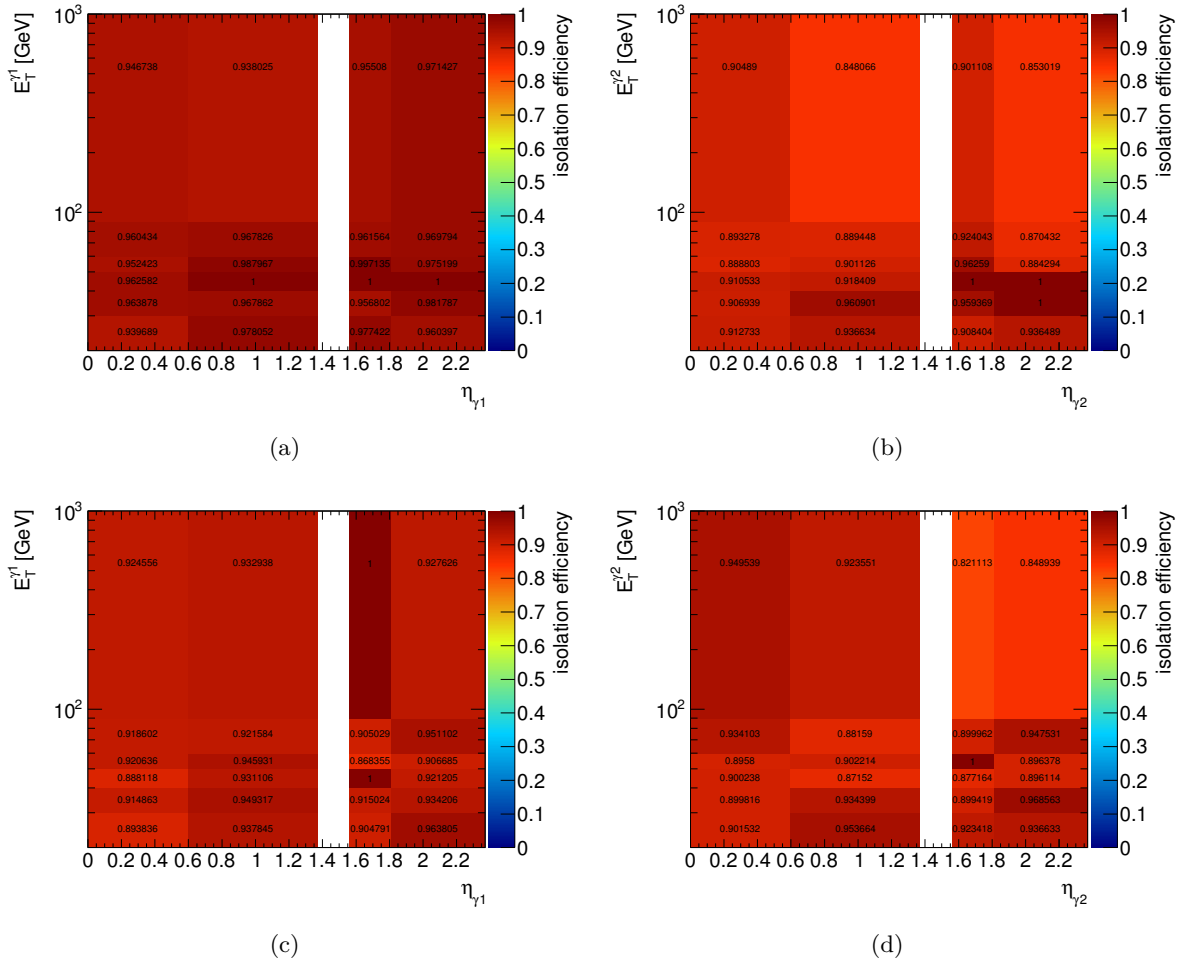


Figure 7.9: Data driven isolation efficiencies of the leading photon candidate (left) and subleading photon candidate (right) for the $0-jet$ (a) and (b) and $1-jet$ (c) and (d) category as a function of η and E_T of the photon candidate.

7.3.1.2 Tight ID efficiencies from MC

Efficiencies of the *tight* ID are extracted from MC simulation as

$$\epsilon_{T1} = \frac{N_{TXTI}}{N_{\tilde{T}XTI} + N_{TXTI}} = \frac{N_{AA} + N_{BA}}{N_{CA} + N_{DA} + N_{AA} + N_{BA}} \tag{7.28}$$

$$\epsilon_{T2} = \frac{N_{TITX}}{N_{TI\tilde{T}X} + N_{TITX}} = \frac{N_{AA} + N_{AB}}{N_{AC} + N_{AD} + N_{AA} + N_{AB}}, \tag{7.29}$$

where $\tilde{I}$ are non-isolated candidates and X means any. In fact, in events are weighted by scale factors in case the photons pass the *tight* selection. These scale factors were computed by matching the *tight* ID efficiencies in data and MC, therefore the *tight* ID efficiencies extracted here are data-driven. The correlation of the two photon candidates *tight* selection is treated according to [230]. The efficiencies extracted from SHERPA MC are shown in Figures 7.13-7.14

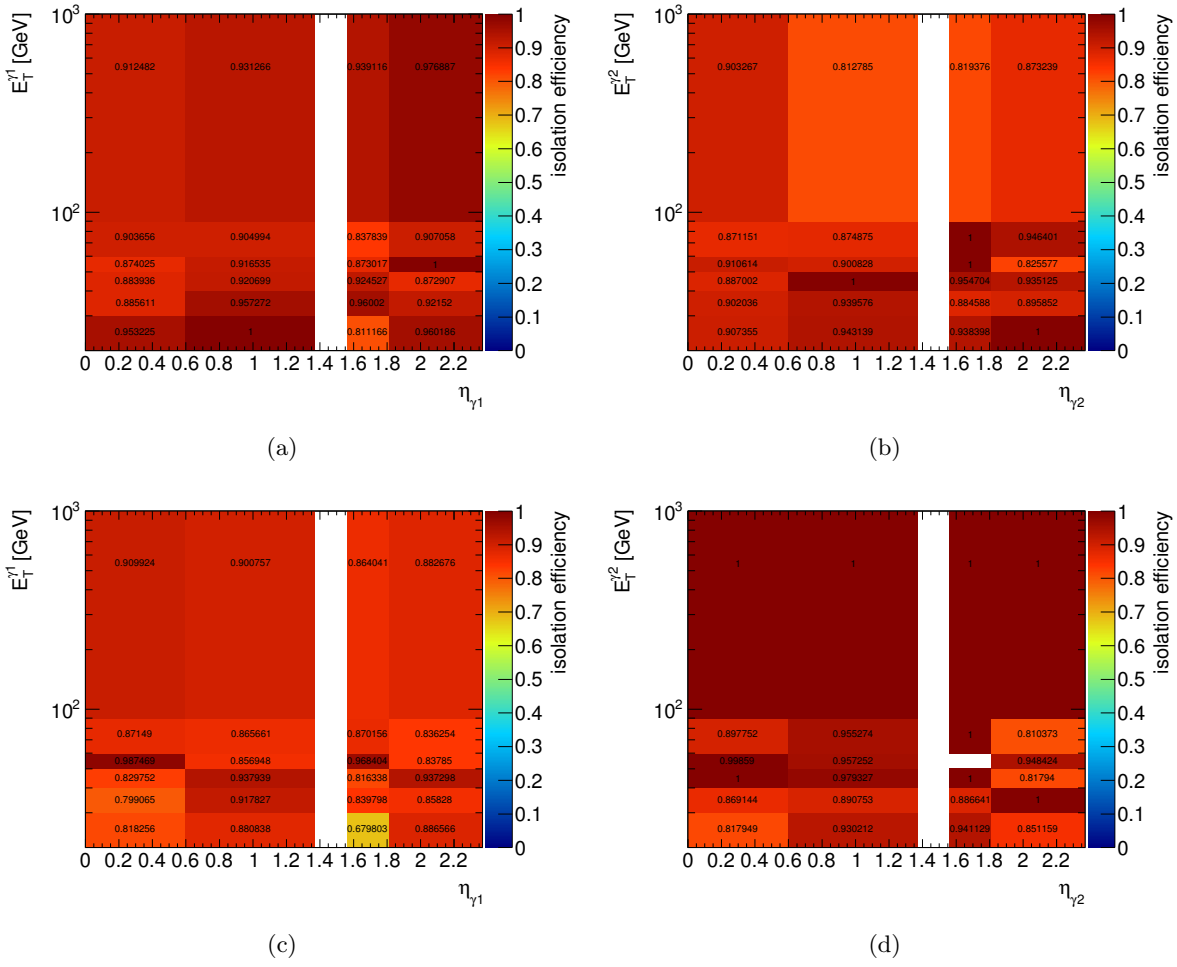


Figure 7.10: Data driven isolation efficiencies of the leading photon candidate (left) and subleading photon candidate (right) for the $2-jet$ (a) and (b) and $\geq 3-jet$ (c) and (d) category as a function of η and E_T of the photon candidate.

for the diphoton and dijet observables in the $2-jet$ category. The photon *tight* ID efficiencies as a function of the other observables are shown in Appendix B. The *tight* ID efficiencies for the inclusive sample (independent of the observables) are listed in Table 7.10.

Tight ID eff.	$0-jet$	$1-jet$	$2-jet$	$\geq 3-jet$
ϵ_{T1}	0.9602 ± 0.0001	0.9718 ± 0.0001	0.9736 ± 0.0002	0.9762 ± 0.0004
ϵ_{T2}	0.9442 ± 0.0001	0.9462 ± 0.0002	0.9492 ± 0.0003	0.9507 ± 0.0005

Table 7.10: Inclusive input *tight* ID Sherpa MC efficiencies for the leading and subleading photons with their statistical uncertainty (derived from MC).

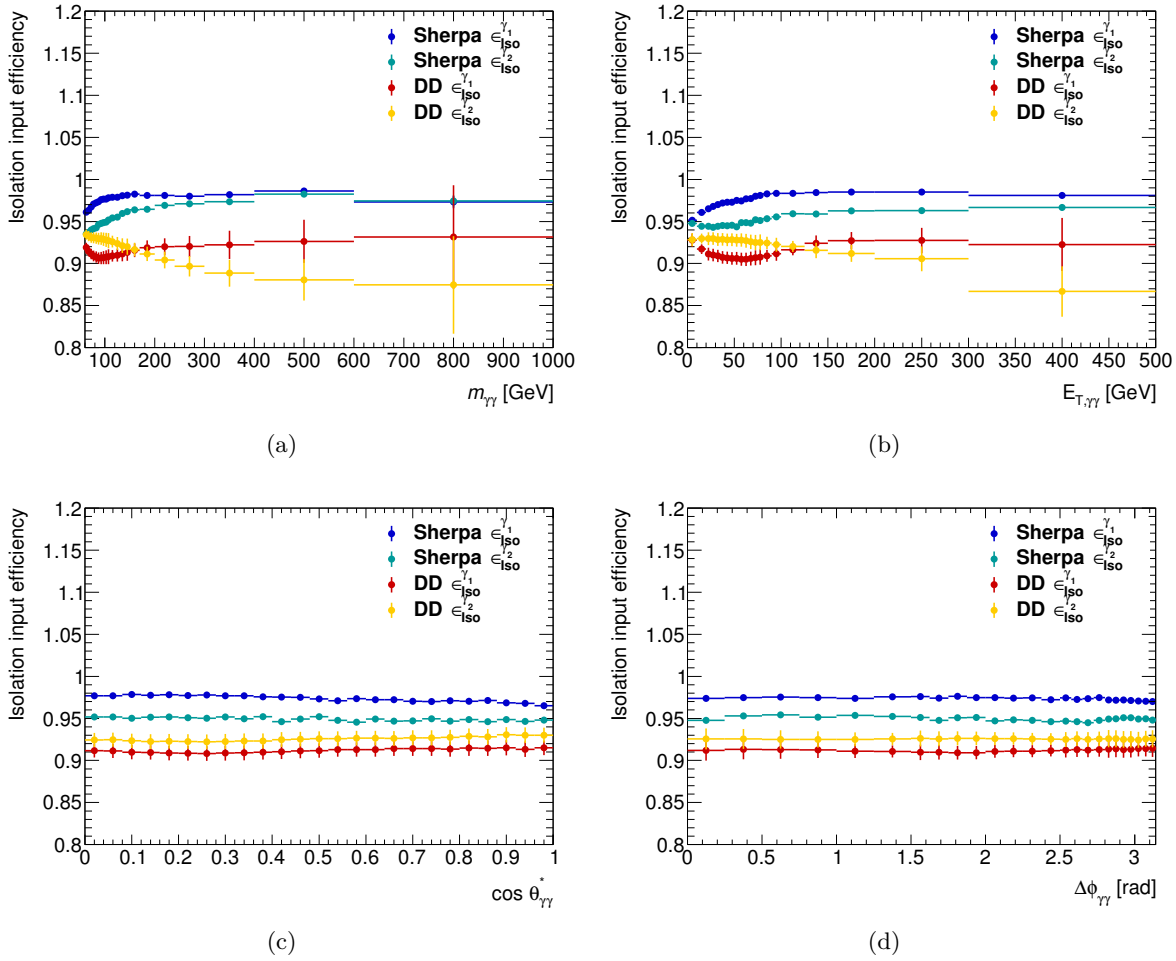


Figure 7.11: Data driven isolation efficiencies of the leading (red) and subleading (yellow) photon candidate, extracted as described in the text and SHERPA MC efficiencies for the leading (dark blue) and subleading (light blue) candidate as a function of the diphoton observables for the $2-jet$ category.

7.3.1.3 ID and isolation correlation parameters for jets

The correlation parameters between identification and isolation for jets cannot be measured from data directly since they would be computed using background events that have to fall into the signal region. The values of these parameters have to be derived from Monte Carlo simulation, which was further checked using a data-driven method in the jet-enhanced control region as detailed in [187]. The nominal di-photon and background yields are extracted assuming $\xi_{j1} = \xi_{j2} = 1$ and the corresponding systematic effect is estimated by varying those parameters in a range depending on the fake rates and the allowed range for the uncertainty on the correlations θ:

$$\theta_i = \frac{\xi_{ji}(1 - f_{Ii} - f_{Ti} + f_{Ii}f_{Ti}\xi_{ji})}{(1 - f_{Ii}\xi_{ji})(1 - f_{Ti}\xi_{ji})} \tag{7.30}$$

with $i = 1, 2$ and $\theta = 1 \pm 0.05$ as evaluated in [187]. With the average values obtained for the fake rates this translates into an allowed range for the ξ_j of about $\pm 2\%$.

7.3.2 Systematic uncertainties on the diphoton yield from the 2x2DSB method

The sources and derivation of systematic uncertainties related to the extraction of the *tight* and isolated diphoton yield (including the isolated electron background and the Higgs boson contribution) are described in this section.

7.3.2.1 Non-tight ID control region definition

One of the inputs to the 2x2DSB method are the data-driven *tight* ID efficiencies for the leading and subleading photon candidate. The *non-tight* control region is defined as events passing the **L'4** but failing the *tight* selection. In order to estimate the uncertainty on the choice of the definition, the other *loose'* definitions are used and the 2x2DSB method is repeated with the *non-tight* control region defined as **L'2**, **L'3** and **L'5**. The isolation efficiencies are also extracted again using the data-driven efficiencies computed the same way as described in Section 7.3.1.1 but using the *non-tight* control region definition with the other *loose'* selections instead of **L'4**. This uncertainty is the dominant systematic uncertainty inclusively amounting to roughly five percent and mostly also bin-by-bin as a function of the observables, which is shown in Figure C.1. This uncertainty is symmetrized, due to the lack of confidence in the **L'5** definition being stable enough to serve as an estimate of a systematic uncertainty.

7.3.2.2 Photon isolation shape from MC

The extraction of photon isolation efficiencies as a function of each observable from the data-driven η^γ and E_{T}^γ efficiencies is performed using Sherpa MC events, as described in Section 7.3.1.1. A systematic uncertainty due to the chosen MC sample is evaluated by repeating the extraction using a Pythia MC sample. This way, the dependence of the efficiencies on the shape of the observable distributions is taken into account. This uncertainty is very small, at the per mille level both inclusively and bin-by-bin for each observable.

7.3.2.3 Input efficiencies and fake rates binning

For the $\gamma - jet$, $jet - \gamma$ and $jet - jet$ components, the variation of the *tight* ID and isolation efficiencies as a function of the observables is taken into account by applying the efficiencies from the neighboring bins (left and right) to the currently computed bin. This procedure aims to account for bins with very low statistics, where the method is not able to perform a reliable fit. The uncertainty is symmetrized and is typically small (below 2%) when there are enough events, but gets larger and sometimes even dominant in bins with very few events.

7.3.2.4 Photon identification efficiencies

The uncertainty on the identification efficiencies for the *tight* ID photons is computed by applying photon ID scale factors computed according to [163, 159, 230] and varying their uncertainty

according to the correlation model between the two photon candidates. This uncertainty is always smaller than 0.5% bin-by-bin for each observable except for the $0-jet$ category, where the the last bins of $p_{\mathrm{T}}^{\gamma\gamma}$ and first bins of $\Delta\phi_{\gamma\gamma}$ lack statistics. Some of the p_T^{jet} observables exhibit the same behavior.

7.3.2.5 Data driven isolation efficiencies - normalization region definition

The uncertainty on the determination of the data-driven isolation efficiencies due to the definition of the normalization region, where the *non-tight* isolation distribution is normalized to the tight distribution and subtracted, is evaluated by moving the lower edge of the track-based isolation window from 10-25 GeV by 2 GeV, down to 8 GeV and up to 12 GeV. The 2x2DSB method is repeated and the uncertainty is extracted, inclusively it is comparable to the other small systematics, at the level of 0.2%. Bin-by-bin, it amounts to an uncertainty at the level of below 1%, except for bins with low statistics.

7.3.2.6 Systematic uncertainties - summary

All systematic uncertainties are shown in Figures 7.15 and 7.16 as a function of the diphoton and dijet observables in the $2-jet$ category. Additional figures for the other observables can be found in Appendix C.

7.3.3 Diphoton yields extracted from the 2x2DSB method

The nominal values of the diphoton yield are extracted using the **L'4** definition of the non-tight control region. All systematic uncertainties described in Section 7.3.2 are applied. The binning of each observable is determined with the requirement of the statistical uncertainty not being larger than the systematic uncertainties in the bin-by-bin case for each observable. The binning was also limited by the requirement of the $2-jet$ category having a sufficient number of events per bin to be able to perform the background decomposition with reasonable results. By choosing the same binning for the observables for the N_{jet} categories, the possibility of comparing the yields across the categories and computing cross section ratios is retained.

7.3.3.1 Integrated diphoton yields

The inclusive diphoton yields extracted by the 2x2DSB method are shown in Tables 7.11-7.14 for the N_{jet} categories. Systematic and statistical uncertainties are also listed in Tables 7.15-7.18 and amount to up to 5% of the diphoton signal yield in all jet multiplicity categories.

7.3.3.2 Differential diphoton yields

The diphoton yields are extracted as a function of the observables in exclusive N_{jet} categories. The decomposed spectra are shown in Figures 7.17-7.22 for the diphoton and dijet observables in the $2-jet$ category, together with the statistical, systematic and total (summed in quadrature) uncertainties for illustration. Additional figures for the other observables can be found in Appendix C.

$0-jet$	2x2DSB	stat	stat[%]	syst ±	syst ±[%]
Yield $\gamma\gamma$	338 341	1635.41	0.48	17 546.40	5.19
Yield γj	132 557	1198.56	0.90	6804.97	5.13
Yield $j\gamma$	43 386.9	856.77	1.97	2328.89	5.37
Yield jj	48 049.7	755.25	1.57	10 814.50	22.51
Purity	60.17%		0.32		3.12

Table 7.11: Integrated diphoton yields extracted from the 2x2DSB method for the $0-jet$ category with statistical and systematic uncertainties.

$1-jet$	2x2DSB	stat	stat[%]	syst ±	syst ±[%]
Yield $\gamma\gamma$	108 949	759.96	0.70	5526.62	5.07
Yield γj	28 643.5	477.93	1.67	1683.97	5.88
Yield $j\gamma$	17 165	372.37	2.17	1453.00	8.46
Yield jj	9 200.78	225.45	2.45	2558.68	27.81
Purity	66.45%		0.47		3.37

Table 7.12: Integrated diphoton yields extracted from the 2x2DSB method for the $1-jet$ category with statistical and systematic uncertainties.

$2-jet$	2x2DSB	stat	stat[%]	syst ±	syst ±[%]
Yield $\gamma\gamma$	34 908.60	374.35	1.07	1296.38	3.71
Yield γj	7 773.10	226.45	2.91	463.43	5.96
Yield $j\gamma$	4 225.33	171.27	4.05	431.12	10.20
Yield jj	2 125.40	99.38	4.68	676.11	31.81
Purity	71.19%		0.74		2.65

Table 7.13: Integrated diphoton yields extracted from the 2x2DSB method for the $2-jet$ category with statistical and systematic uncertainties.

$\geq 3-jet$	2x2DSB	stat	stat[%]	syst ±	syst ±[%]
Yield $\gamma\gamma$	12 601.4	202.13	1.60	427.33	3.39
Yield γj	2 283.83	114.68	5.02	126.95	5.56
Yield $j\gamma$	1 254.04	84.37	6.73	225.56	17.99
Yield jj	605.597	48.75	8.05	196.55	32.46
Purity	75.26%		1.11		2.56

Table 7.14: Integrated diphoton yields extracted from the 2x2DSB method for the $\geq 3-jet$ category with statistical and systematic uncertainties.

The systematic uncertainties are later symmetrized as the diphoton yields for each observable enter the unfolding procedure in Section 7.6. Table 7.19 lists the systematic uncertainties of the 2x2DSB method and their treatment when they are propagated through the unfolding, described in Section 7.6. The uncertainties treated as correlated between bins are propagated separately, uncertainties treated as uncorrelated between bins are first combined in quadrature and propagated through the unfolding procedure together. The photon ID scale factors, the isolation efficiency and the window definition systematic are negligible w.r.t. the **L'** and the

Systematic uncertainty	syst +	syst +[%]	syst −	syst −[%]
LoosePrime	17 531.40	5.18	17 531.40	5.18
Photon ID Scale Factors	696.27	0.21	696.27	0.21
DD Iso Efficiencies	173.31	0.05	173.31	0.05
DD Iso Window def.	121.03	0.04	121.03	0.04
Total	17 546.40	5.19	17546.4	5.18

Table 7.15: Breakdown of the systematic uncertainties of the 2x2DSB method for the integrated yields in the $0 - jet$ category.

Systematic uncertainty	syst +	syst +[%]	syst −	syst −[%]
LoosePrime	5525.72	5.07	5525.72	5.07
Photon ID Scale Factors	93.9737	0.09	50.3924	0.05
DD Iso Efficiencies	6.28086	0.01	6.28086	0.01
DD Iso Window def.	33.25	0.03	33.25	0.03
Total	5526.62	5.07	5526.62	5.07

Table 7.16: Breakdown of the systematic uncertainties of the 2x2DSB method for the integrated yields in the $1 - jet$ category.

Systematic uncertainty	syst +	syst +[%]	syst −	syst −[%]
LoosePrime	1295.35	3.71	1295.35	3.71
Photon ID Scale Factors	27.8843	0.08	9.52361	0.03
DD Iso Efficiencies	27.4944	0.08	27.4944	0.08
DD Iso Window def.	34.00	0.10	34.00	0.10
Total	1296.38	3.71	1296.38	3.71

Table 7.17: Breakdown of the systematic uncertainties of the 2x2DSB method for the integrated yields in the $2 - jet$ category.

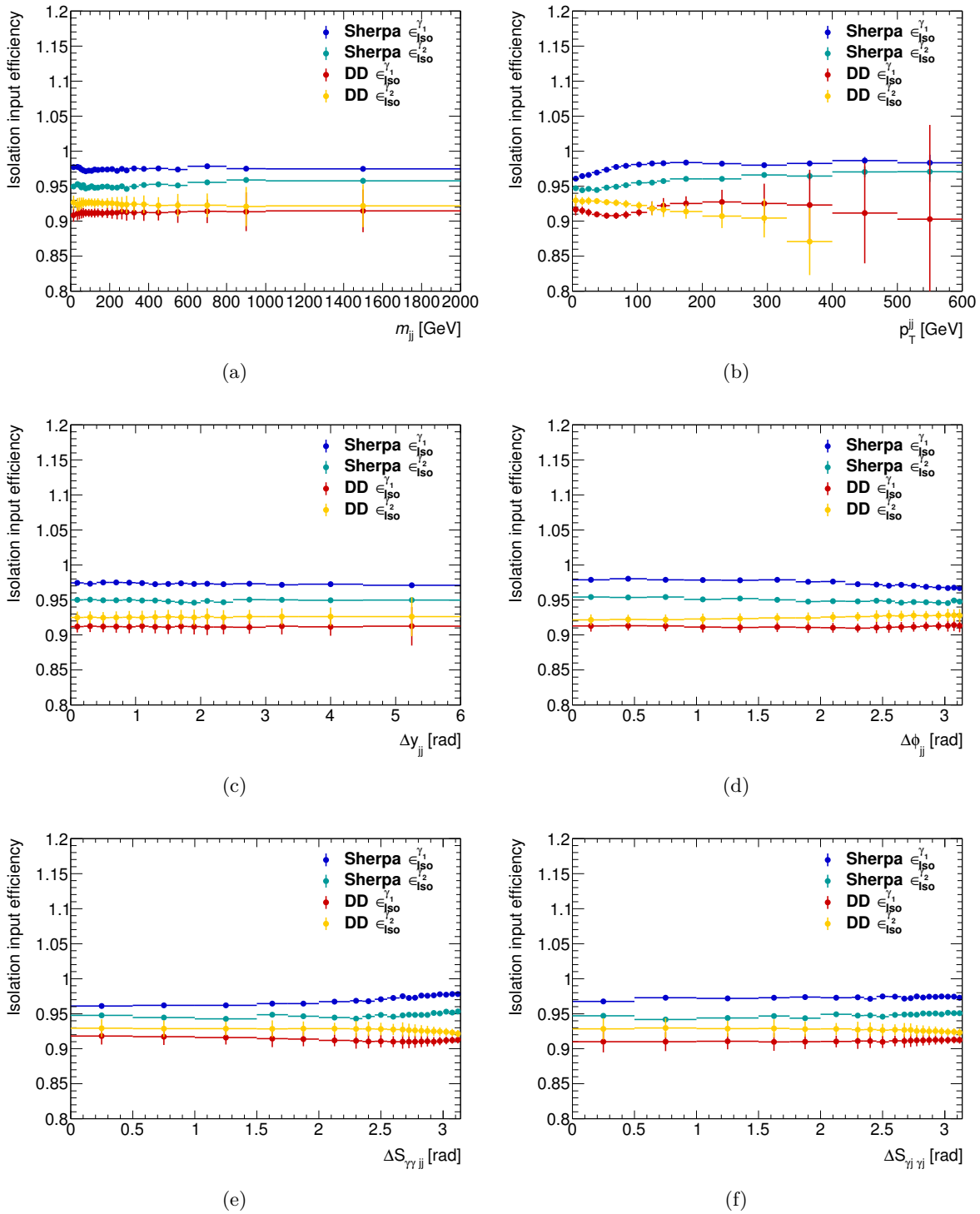
Systematic uncertainty	syst +	syst +[%]	syst −	syst −[%]
Loose Prime	426.762	3.39	426.762	3.39
Photon ID Scale Factors	3.29302	0.03	6.54467	0.05
DD Iso Efficiencies	1.29742	0.01	1.29742	0.01
DD Iso Window def.	20.89	0.17	20.89	0.17
Total	427.33	3.39	427.33	3.39

Table 7.18: Breakdown of the systematic uncertainties of the 2x2DSB method for the integrated yields in the $\geq 3 - jet$ category.

binning of efficiencies, therefore their impact on the final error on the measured cross section is be very small as well.

Systematic uncertainty	Treatment
Loose Prime	correlated between bins
Photon ID Scale Factors	uncorrelated between bins
DD Iso Efficiencies	uncorrelated between bins
DD Iso Window def.	uncorrelated between bins
Efficiencies/Fake-rates binning	uncorrelated between bins

Table 7.19: Treatment of the systematic uncertainties of the 2x2DSB method as they enter the unfolding procedure.



(a) (b) (c) (d) (e) (f)

Figure 7.12: Data driven isolation efficiencies of the leading (red) and subleading (yellow) photon candidate, extracted as described in the text and SHERPA MC efficiencies for the leading (dark blue) and subleading (light blue) candidate as a function of the dijet and the ΔS observables for the $2-jet$ category.

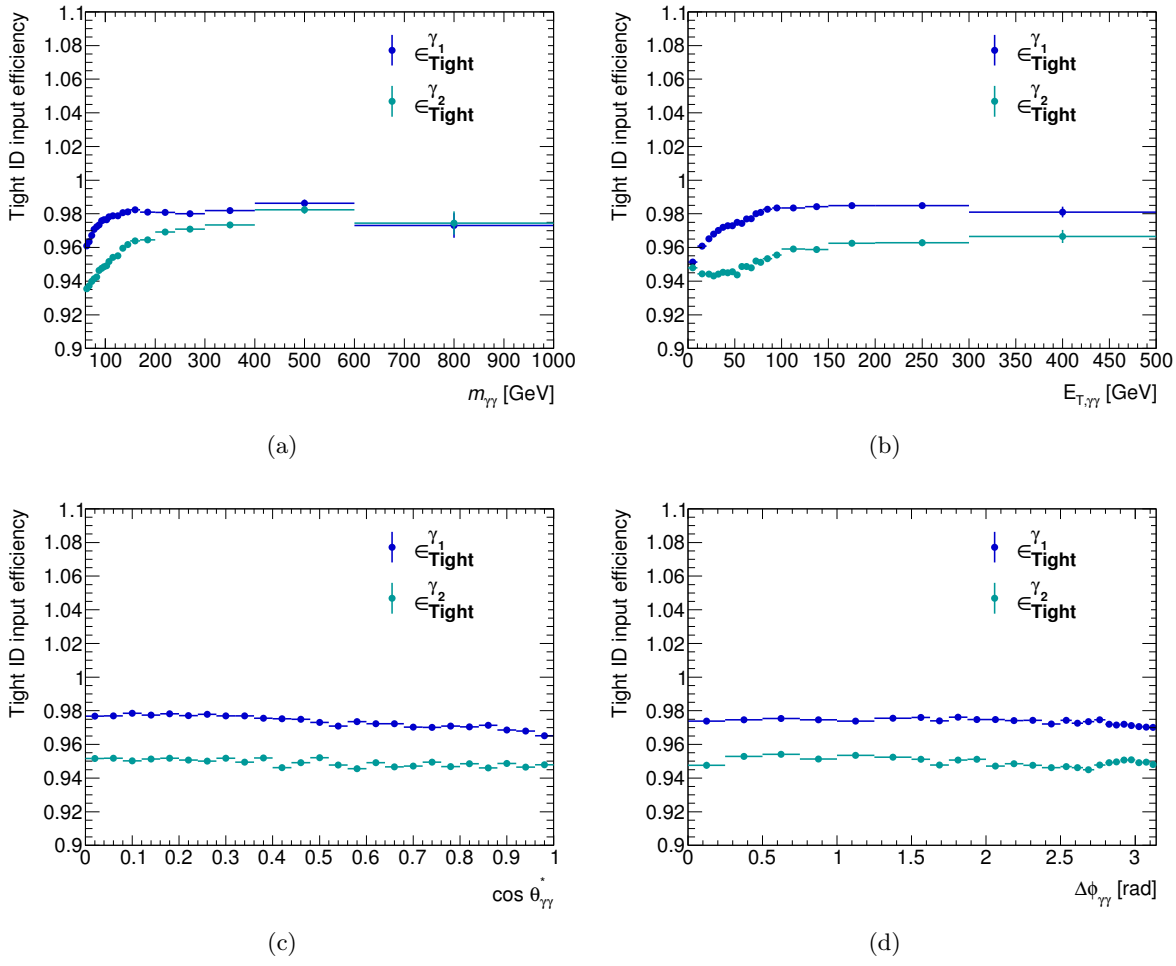
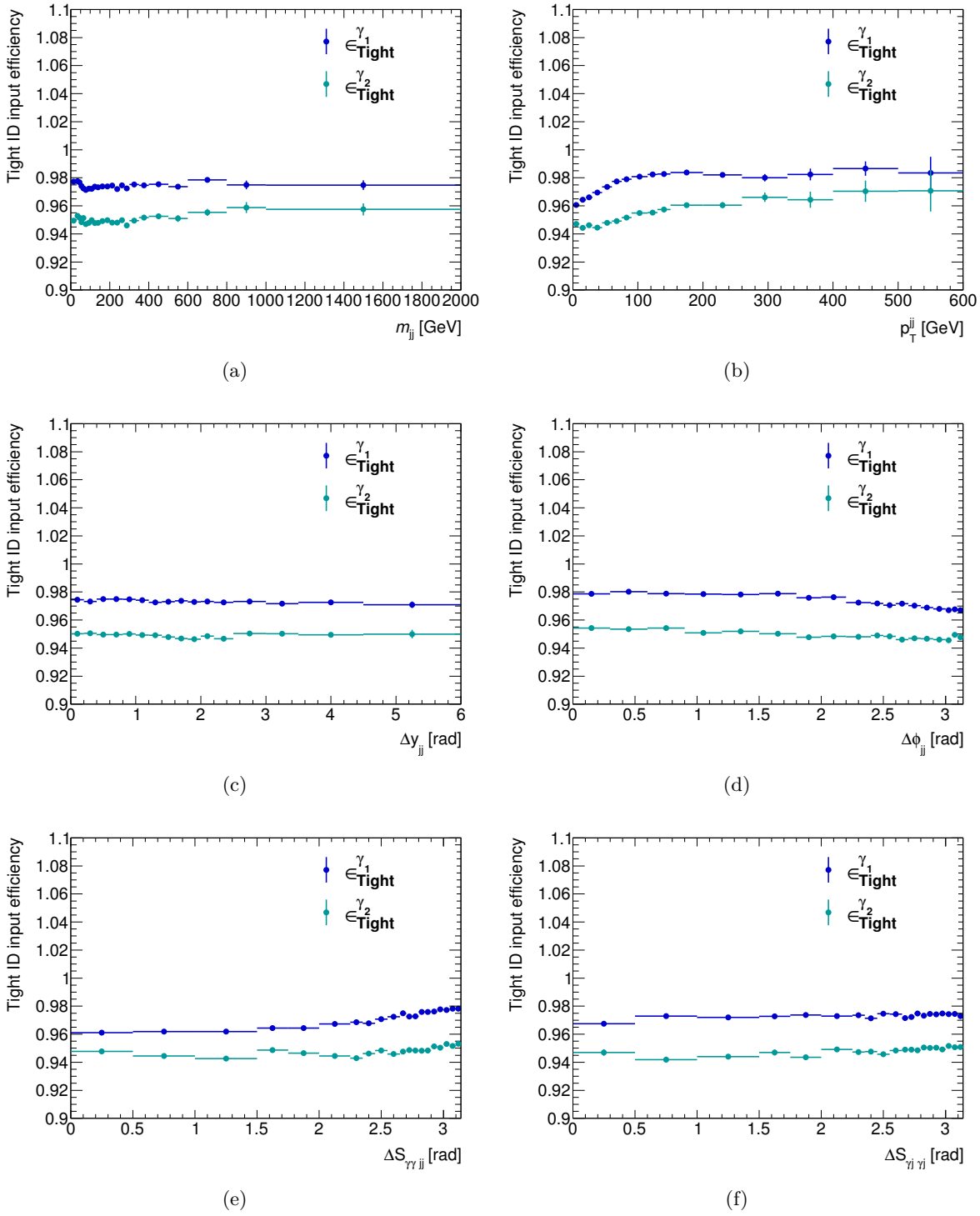
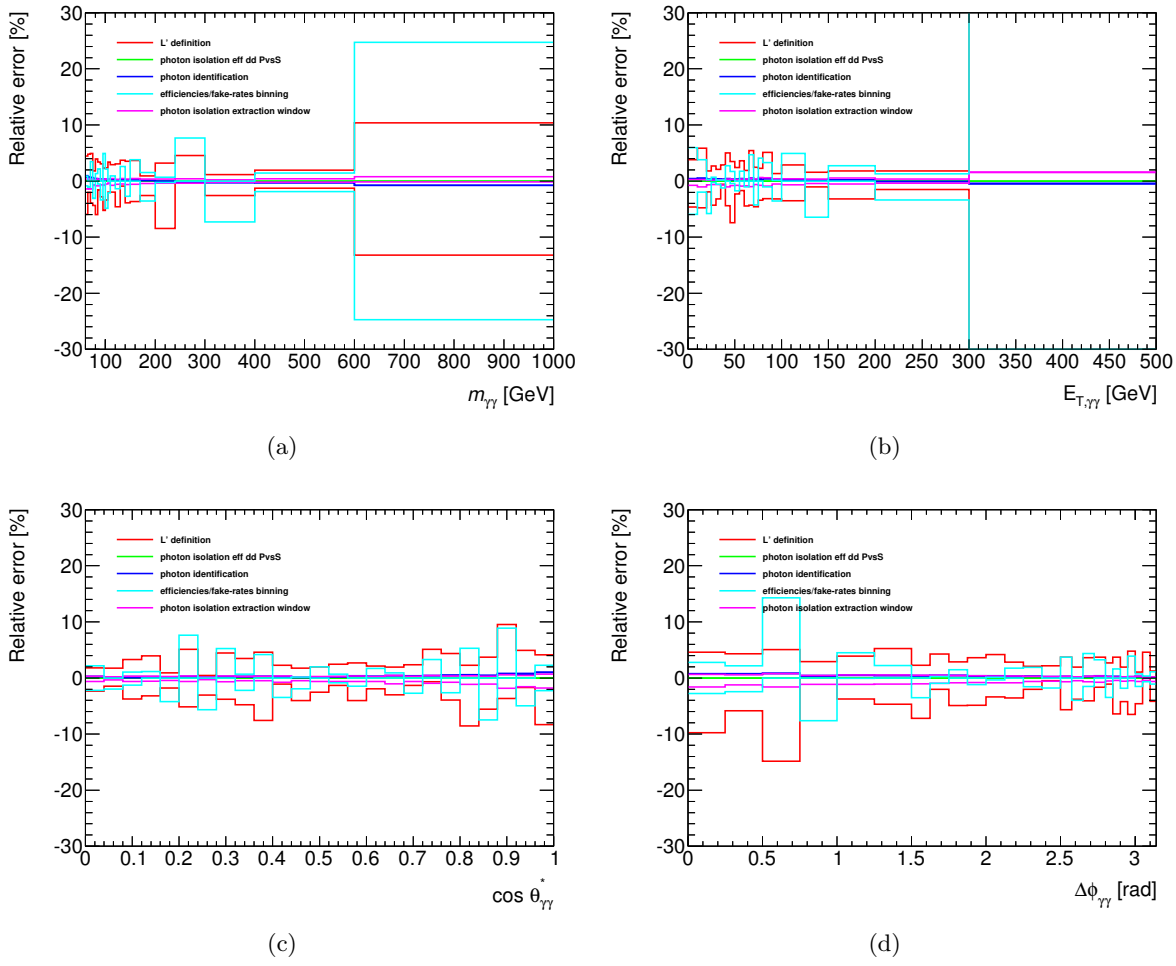


Figure 7.13: Input *tight* ID efficiency of the leading (dark blue) and subleading (light blue) photon as a function of the diphoton observables for the $2-jet$ category.



(a) (b) (c) (d) (e) (f)

Figure 7.14: Input *tight* ID efficiency of the leading (dark blue) and subleading (light blue) photon as a function of the dijet and the ΔS observables for the $2-jet$ category.



(a) (b)

(c) (d)

Figure 7.15: Breakdown of the systematic uncertainties on the diphoton yield extracted by the 2x2DSB method as a function of the diphoton observables for the $2-jet$ category.

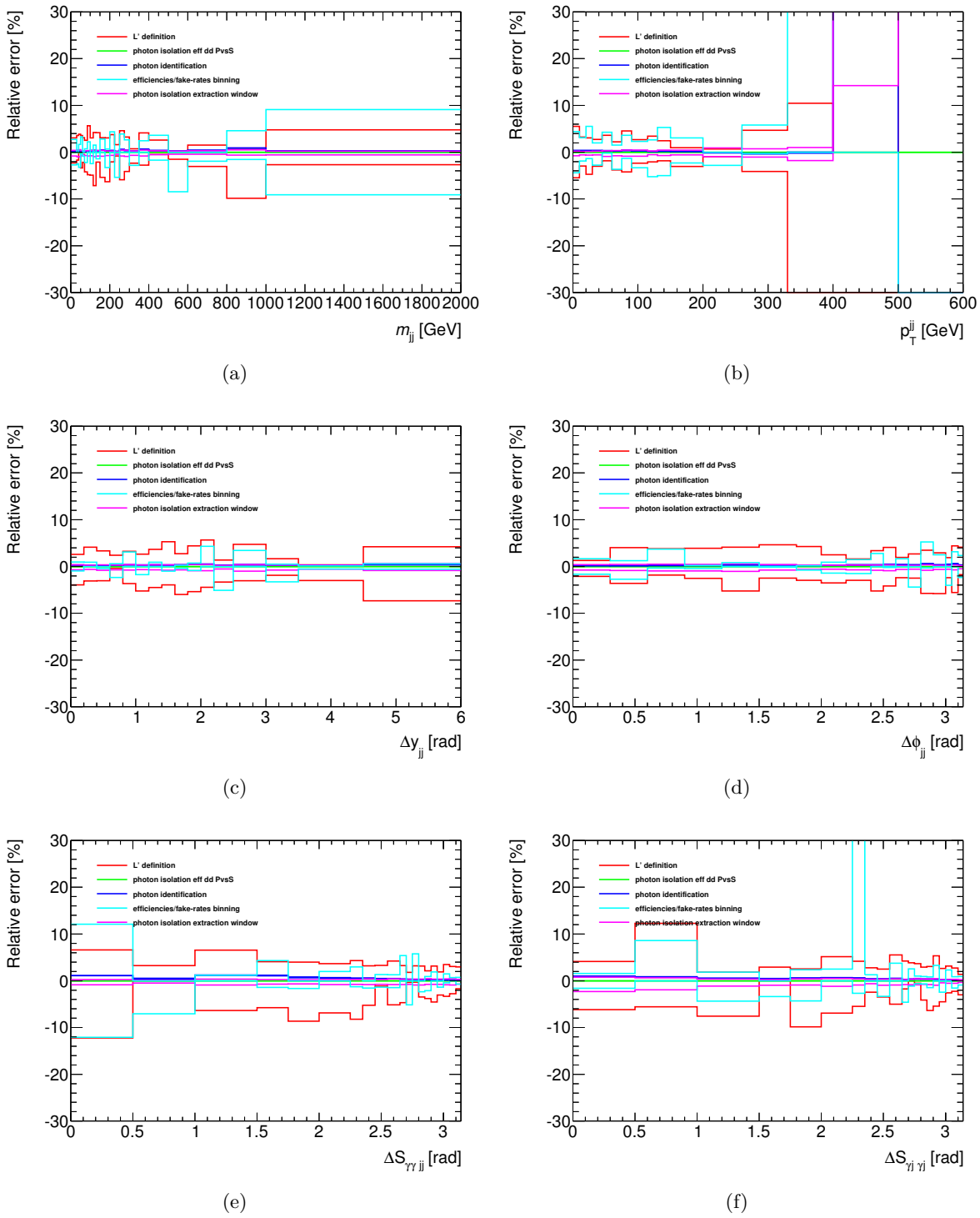


Figure 7.16: Breakdown of the systematic uncertainties on the diphoton yield extracted by the 2x2DSB method as a function of the dijet and the ΔS observables for the $2-jet$ category.

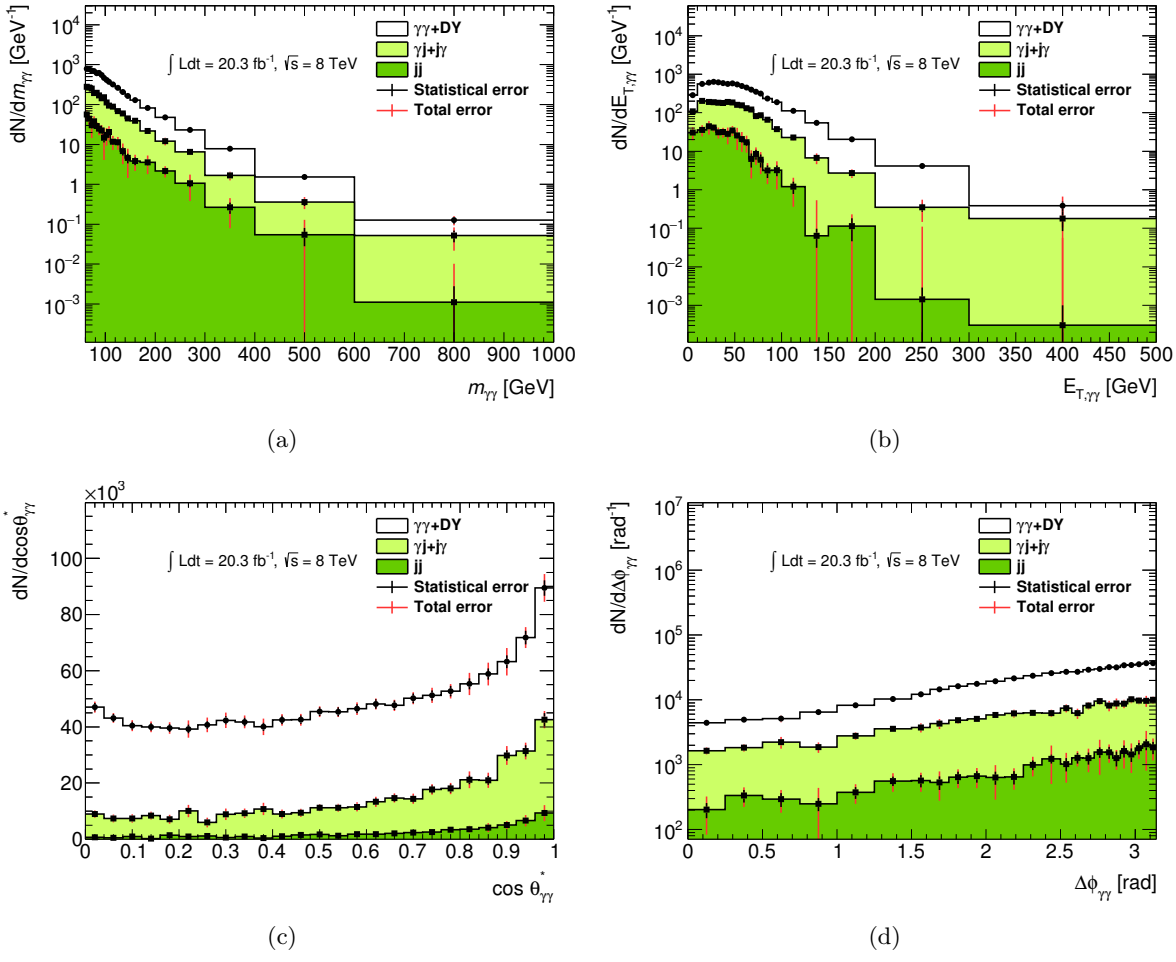


Figure 7.17: Diphoton yields as a function of the diphoton observables: $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$, $\cos\theta^{*}_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ for the $N_{jet} = 2$ category.

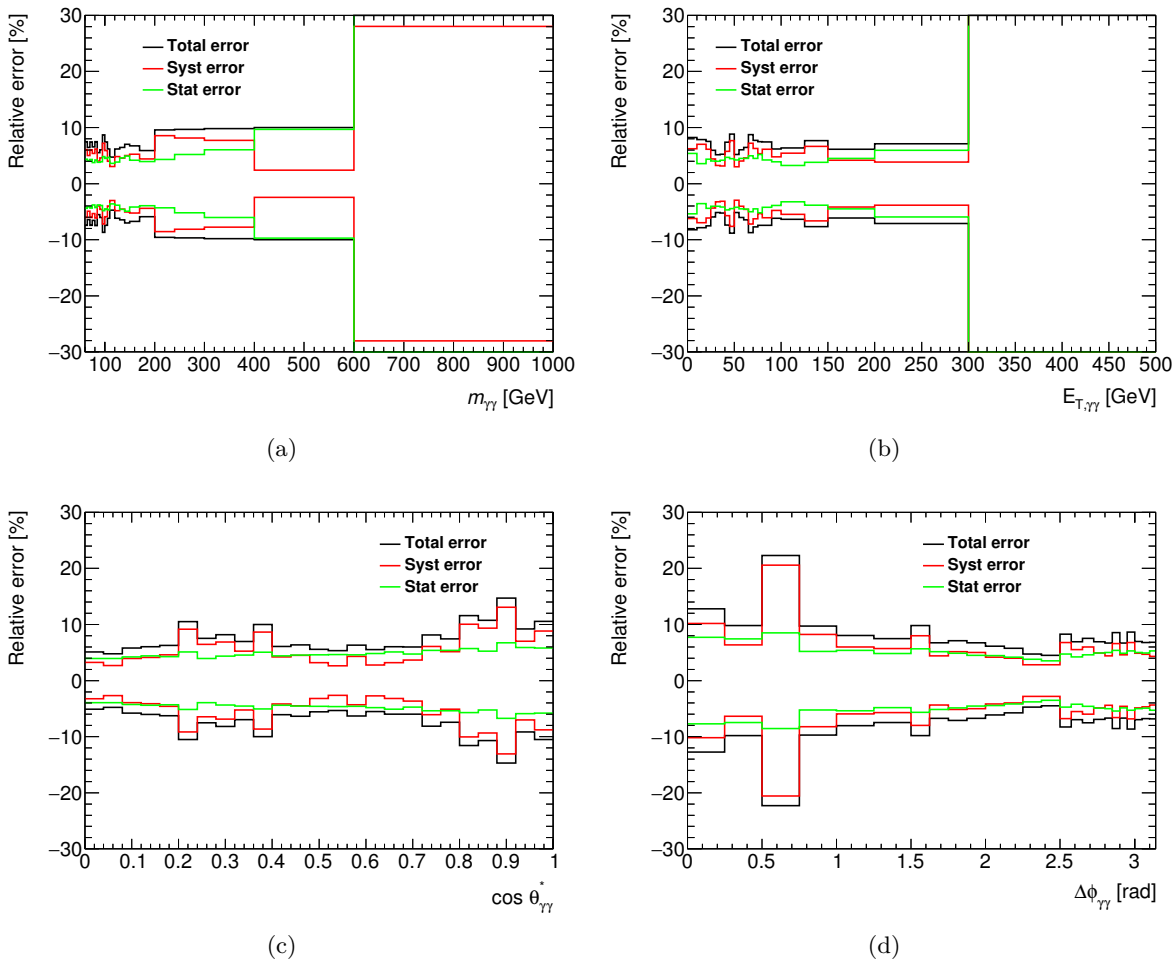


Figure 7.18: Statistical (green) and systematic (red) and total (black) uncertainties on the diphoton yields as a function of the diphoton observables: $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$, $\cos\theta^{*}_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ for the $N_{jet} = 2$ category.

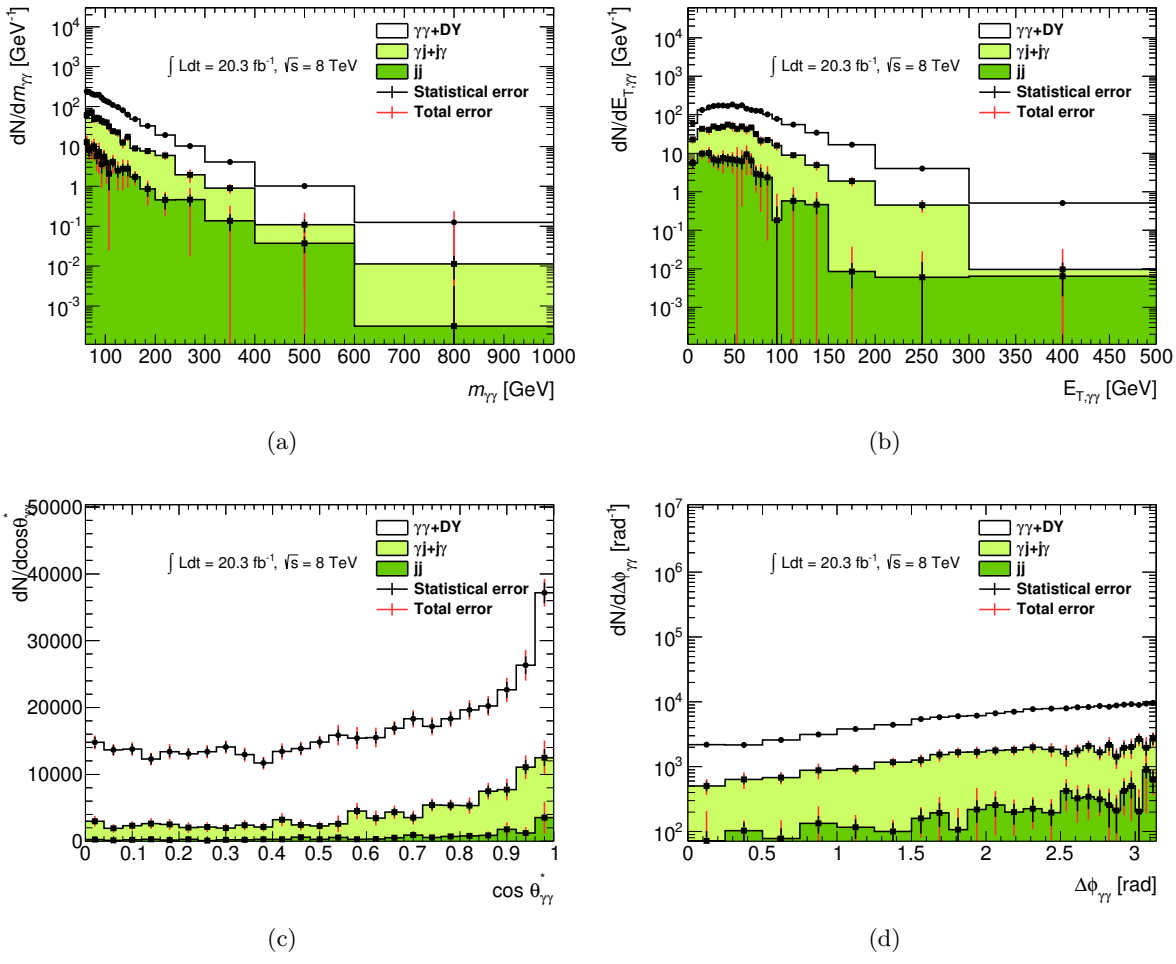


Figure 7.19: Diphoton yields as a function of the diphoton observables: $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$, $\cos\theta^{*}_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ for the $N_{jet} \geq 3$ category.

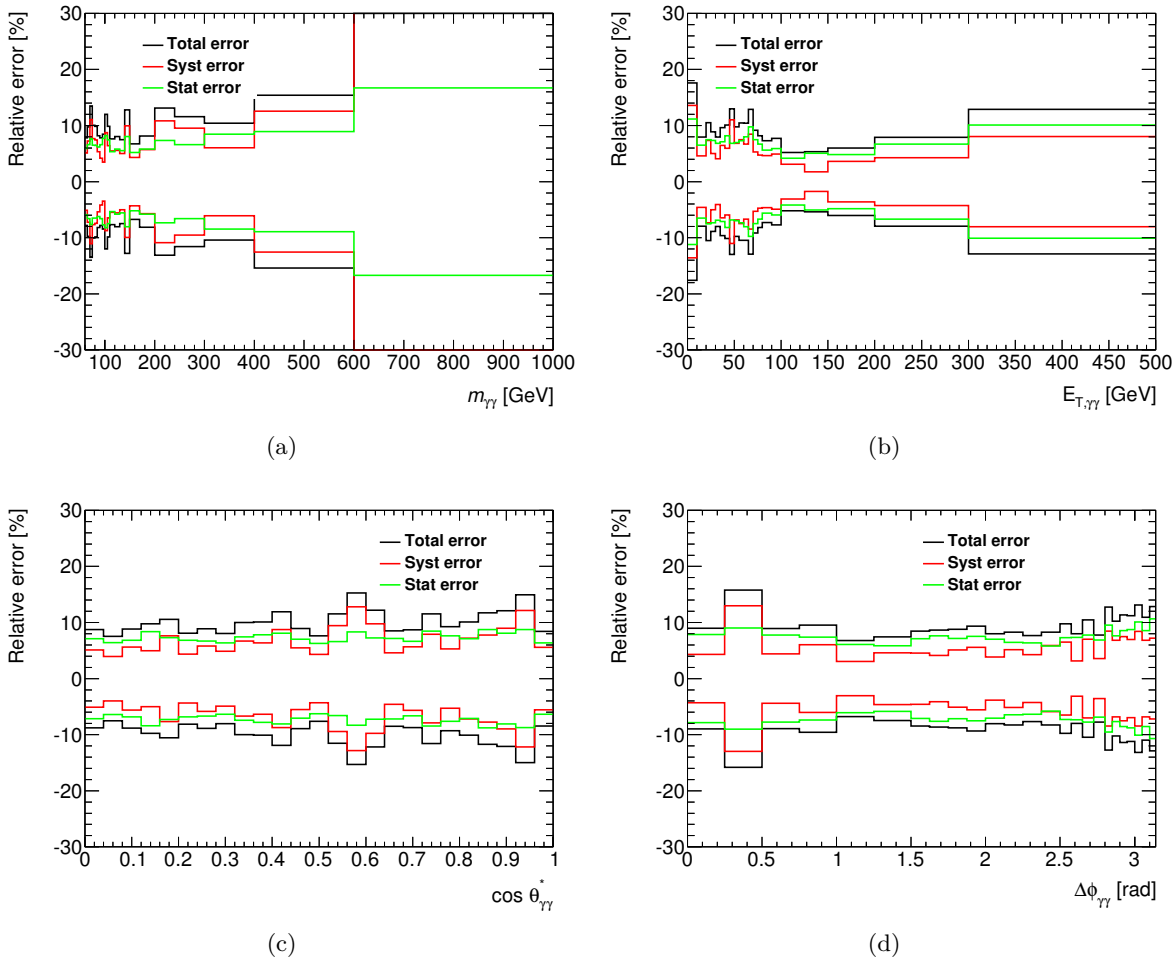


Figure 7.20: Statistical (green) and systematic (red) and total (black) uncertainties on the diphoton yields as a function of the diphoton observables: $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$, $\cos\theta^{*}_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ for the $N_{jet} \geq 3$ category.

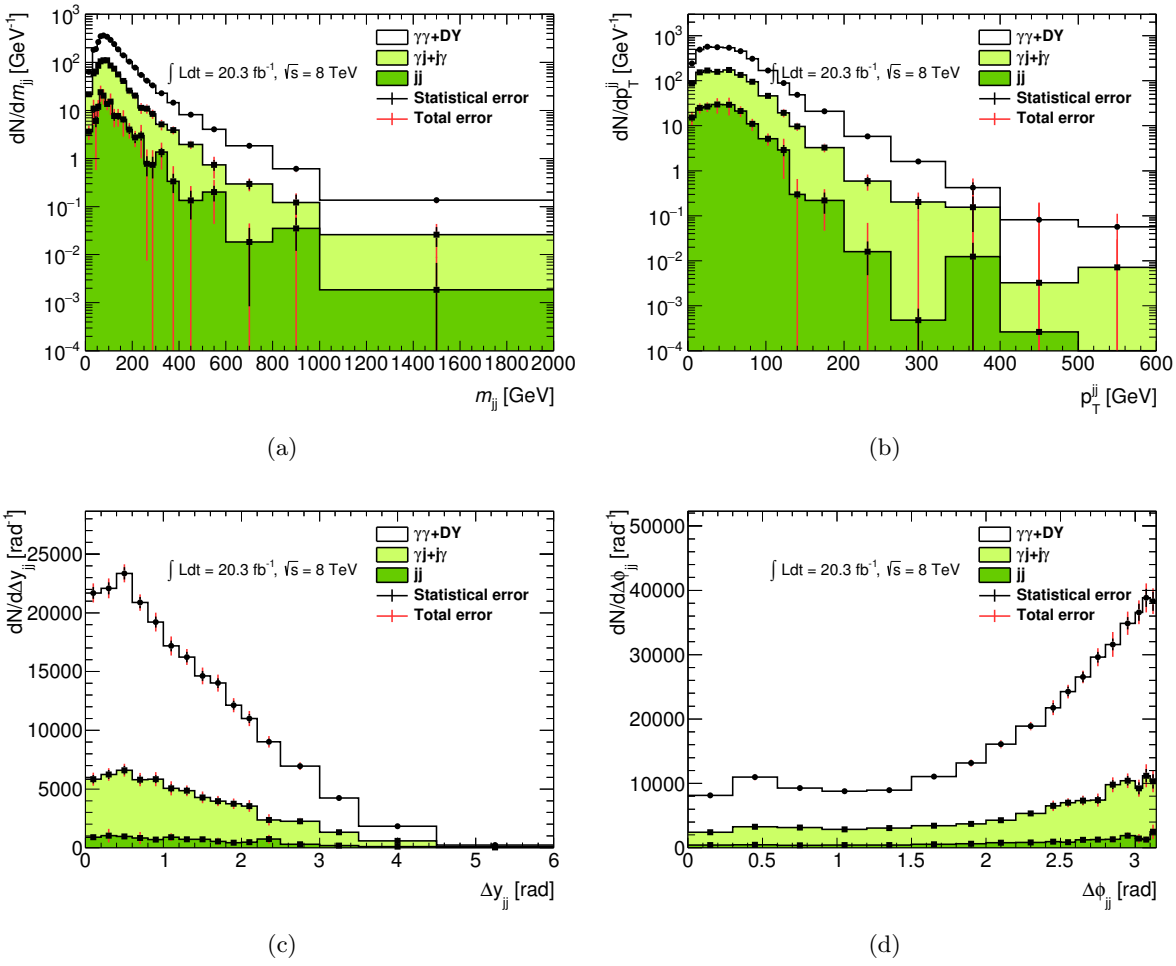


Figure 7.21: Diphoton yields as a function of the dijet observables: m_{jj}, p_{T}^{jj}, Δy_{jj} and $\Delta\phi_{jj}$ for the $N_{jet} = 2$ category.

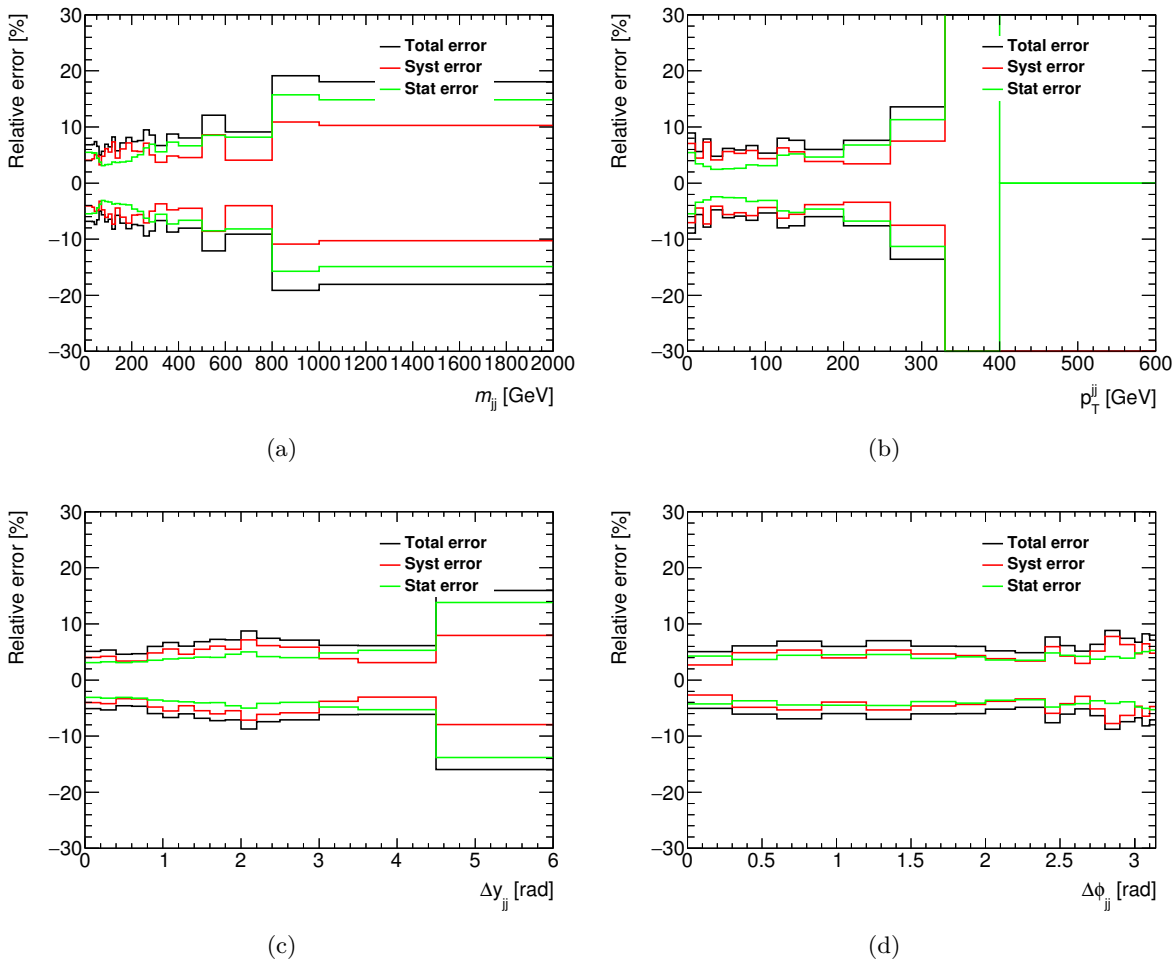


Figure 7.22: Statistical (green) and systematic (red) and total (black) uncertainties on the diphoton yields as a function of the dijet observables: m_{jj}, p_{T}^{jj}, Δy_{jj} and $\Delta\phi_{jj}$ for the $N_{jet} = 2$ category.

7.4 Electron background

The diphoton yields, extracted using the 2x2DSB method, contain a component associated to electrons and positrons wrongly reconstructed as photons. This occurs more frequently when an electron undergoes very large bremstrahlung or the inner detector fails to detect a B-layer hit of the track or a spurious conversion vertex is associated to the electron. The electron background is clearly visible in the low-mass region of the $m_{\gamma\gamma}$ spectrum, in correspondence to the $Z \to ee$ resonance, around $m_{\gamma\gamma} \approx 90$ GeV, as shown in Figure 7.23. These events clearly come from the Z boson decays into an electron-positron pair. Other processes contributing to the electron background are the Drell-Yan continuum or the $WW \to e\nu e\nu$ in the ee final state and the $\gamma W \to \gamma e\nu$ and $\gamma Z \to \gamma ee$ in the γe final state. Since it would be difficult to evaluate each contribution separately, a more generic approach is followed, which relies on computing the probabilities of electrons faking a photon and vice versa. A similar approach was used previously in the $\sqrt{s} = 7$ TeV measurement of the inclusive diphoton cross section [96]. In the following, both e^+ and e^- will be referred to as *electrons*.

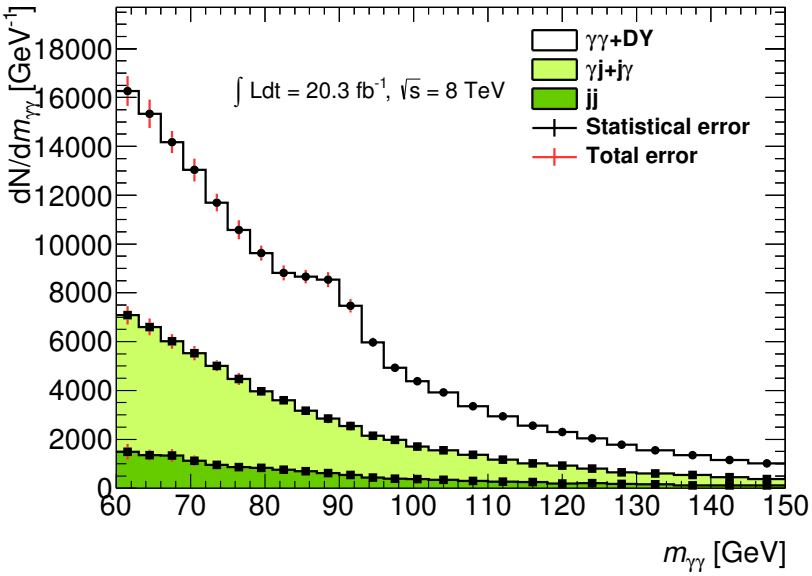


Figure 7.23: $m_{\gamma\gamma}$ spectrum in the $60 < m_{\gamma\gamma} < 150$ GeV region, the $\gamma\gamma$, γj and jj components obtained with the 2x2DSB method for the inclusive diphoton sample.

7.4.1 Computation of the fake rates

The $e \to \gamma$ fake rate can be extracted by looking at the number of ee and $e\gamma$ events observed with a pair invariant mass compatible with the Z boson mass peak (further referred to as the Z boson peak). For single electron and photon final state, N_x is the number of events containing an object classified as x (x being e or γ) by the reconstruction and passing the analysis selection criteria. n_x is the number of reconstructed events containing a true object x. The observed

numbers of N_γ and N_e can be related to n_γ and n_e as

$$\begin{pmatrix} N_\gamma \\ N_e \end{pmatrix} = \begin{pmatrix} \epsilon_\gamma & \rho_{e\to\gamma} \\ \rho_{\gamma\to e} & \epsilon_e \end{pmatrix} \times \begin{pmatrix} n_\gamma \\ n_e \end{pmatrix}, \tag{7.31}$$

where $\epsilon_{\gamma/e}$ is the probability that γ/e will be identified as such and will pass the analysis selection, $\rho_{e\to\gamma}$ is the probability that a true e will be identified as a γ and $\rho_{\gamma\to e}$ is the probability that a true γ will be identified as an e. Defining the number of signal events $N_x^{sig} = \epsilon_x n_x$ and expressing the fake rates as

$$f_{e\to\gamma} = \frac{\rho_{e\to\gamma}}{\epsilon_e} \qquad f_{\gamma\to e} = \frac{\rho_{\gamma\to e}}{\epsilon_\gamma} \tag{7.32}$$

then Eq. 7.31 can be rewritten as

$$\begin{pmatrix} N_\gamma \\ N_e \end{pmatrix} = \begin{pmatrix} 1 & f_{e\to\gamma} \\ f_{\gamma\to e} & 1 \end{pmatrix} \times \begin{pmatrix} N_\gamma^{sig} \\ N_e^{sig} \end{pmatrix}, \tag{7.33}$$

where $N_{\gamma/e}^{sig}$ is the number of true γ/e identified as such and passing the analysis selection. The physical meaning of $f_{e\to\gamma}$ and $f_{\gamma\to e}$ is described as the ratio between the number of objects wrongly classified and correctly classified

$$f_{e\to\gamma} = \frac{N_{e\to\gamma}}{N_{e\to e}} \tag{7.34}$$

$$f_{\gamma\to e} = \frac{N_{\gamma\to e}}{N_{\gamma\to\gamma}}. \tag{7.35}$$

When considering two-body final states and calling N_{xy} the number of final states classified as xy by the reconstruction and passing the analysis selection cuts, and N_{xy}^{sig} those for which the classification is correct, this matrix equation holds

$$\begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e} \\ N_{ee} \end{pmatrix} = \begin{pmatrix} 1 & f_{e\to\gamma} & f_{e\to\gamma}^2 \\ 2f_{\gamma\to e} & 1 + f_{e\to\gamma}f_{\gamma\to e} & 2f_{e\to\gamma} \\ f_{\gamma\to e}^2 & f_{\gamma\to e} & 1 \end{pmatrix} \times \begin{pmatrix} N_{\gamma\gamma}^{sig} \\ N_{\gamma e}^{sig} \\ N_{ee}^{sig} \end{pmatrix}, \tag{7.36}$$

where γe denotes the sum of $e\gamma$ and γe events. The Eq. 7.36 can be inverted to be able to compute the number of signal events as

$$\begin{pmatrix} N_{\gamma\gamma}^{sig} \\ N_{\gamma e}^{sig} \\ N_{ee}^{sig} \end{pmatrix} = \frac{1}{(1 - f_{e\to\gamma}f_{\gamma\to e})^2} \begin{pmatrix} 1 & -f_{e\to\gamma} & f_{e\to\gamma}^2 \\ -2f_{\gamma\to e} & 1 + f_{e\to\gamma}f_{\gamma\to e} & -2f_{e\to\gamma} \\ f_{\gamma\to e}^2 & -f_{\gamma\to e} & 1 \end{pmatrix} \times \begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e} \\ N_{ee} \end{pmatrix}. \tag{7.37}$$

The signal component $N_{\gamma\gamma}^{sig}$ of the $N_{\gamma\gamma}$ sample can be computed when the fake rates are known as

$$N_{\gamma\gamma}^{sig} = \frac{N_{\gamma\gamma} - \left[f_{e\to\gamma}N_{\gamma e} - f_{e\to\gamma}^2 N_{ee}\right]}{(1 - f_{e\to\gamma}f_{\gamma\to e})^2}. \tag{7.38}$$

The two fake rates $f_{e\to\gamma}$ and $f_{\gamma\to e}$ depend on several experimental effects, such as the material distribution in front of the LAr calorimeter and the inner detector tracking performance. The description of these effects in MC is imperfect [142]. Therefore, the measurement should be performed using data. A validation of this measurement can be done using MC simulation, as will be shown later.

7.4.1.1 Electron to photon fake rate

The fake rate $f_{e\to\gamma}$ can be estimated from the number of events with diphoton invariant mass compatible with that of objects coming from the decay of the Z boson. When looking at ee and $e\gamma$ events, where the γ is actually an electron that has been misidentified as a photon, in the invariant mass region close to the Z boson mass, Eq. 7.36 can be simplified as

$$N_{ee} = N_{ee}^{sig} \tag{7.39}$$

$$N_{\gamma e} = N_{ee}^{sig} \cdot 2 f_{e\to\gamma} \tag{7.40}$$

so that the electron to photon fake rate is expressed as

$$f_{e\to\gamma} = \frac{N_{\gamma e}}{2N_{ee}}. \tag{7.41}$$

Dielectron pairs are built from data, requiring the two electrons to pass the analysis selection detailed in Section 7.2 except the photon *tight* ID. Instead, they have to pass the electron *tight*++ ID requirement. Electrons are also required not to overlap with any reconstructed photon present in the same event. The $m_{e\gamma}$ and $m_{\gamma e}$ invariant mass distributions are also built, requiring the same criteria on the electrons as mentioned above. The photons in this case are reconstructed as photons but the calibration is applied as if they were electrons, in order to stay consistent between the four objects used in the fake rate computation. The ee invariant mass peak and the $\gamma e + e\gamma$ peak (further referred to as $e\gamma$) are shifted with respect to each other, as seen in Figure 7.24, and the ee peak also has a smaller width than the $e\gamma$ peak.

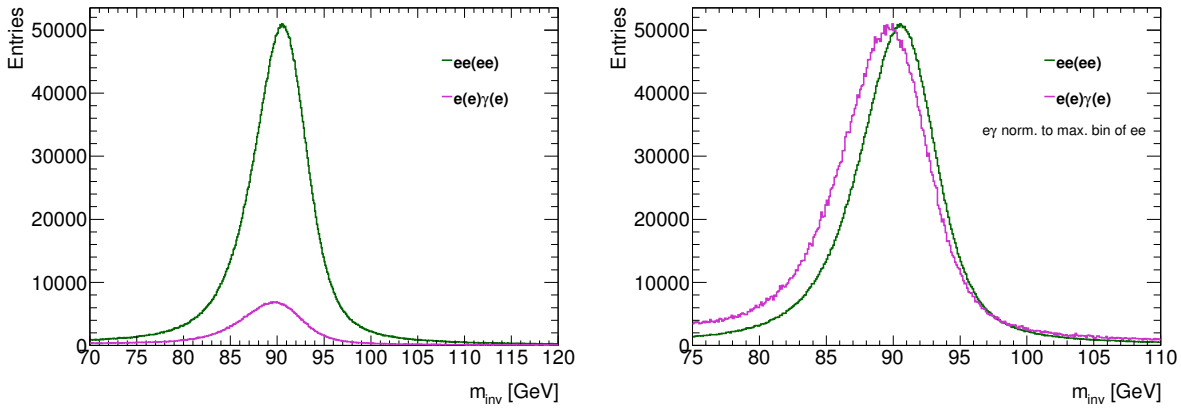
This is a consequence of two effects. First, electrons and photons are reconstructed differently, namely the cluster size varies for photons, as described in Section 3.1. The calibration is also different between the two objects (as discussed in Chapter 5) and calibrating one as the other results in miscalibration. The shift has to be accounted for by smearing and shifting the γ object in the $e\gamma$ distribution separately for $e\gamma$ and γe events, similarly to what was performed in Section 6.4.2.2, but it is left as is for the moment.

To extract the fake rate, the two peaks are fitted using a convolution of a double-sided Crystal ball function and a Breit-Wigner distribution with a fixed Z boson width of $\Gamma_Z = 2.495$ GeV [63], to describe the Z boson resonance and the detector resolution. The background is described by an exponential and the fit is performed in a window of $70 < m_{inv} < 180$ GeV in order to avoid the trigger turn-on shape. The fit result is shown in Figure 7.25. The background is subtracted.

To determine the precise properties of the Z boson peak, it is fitted by a simple Gaussian to determine its mean μ and width σ. This is performed for both the ee and $e\gamma$ events and the number of events $\pm 1.5\sigma$ around the Z peak is counted. Even without applying a shift to the photon E_{T}, by fitting the two peaks separately and using a different range to extract the number of events, the fake rate computation is consistent. The background computed from the fit as described above is subtracted and the fake rate is computed according to Eq. 7.41. The result is listed in Table 7.20.

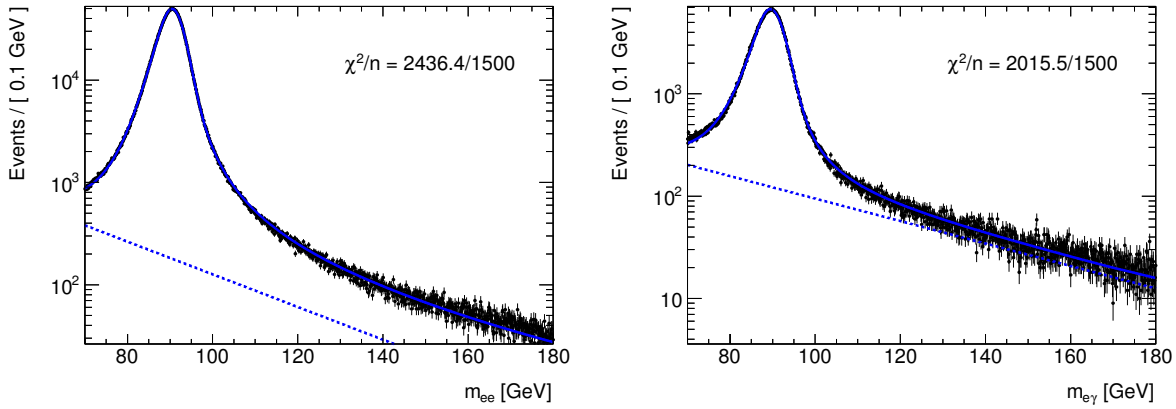
The systematic uncertainties on the electron to photon fake rate, that are considered, are the following:

- the fake rate extraction is repeated by counting events in a window of $\pm 2.0\sigma$ and $\pm 1\sigma$ around the Z mass;



(a) (b)

Figure 7.24: ee (green) and $e\gamma$ (pink) invariant mass (m_{inv}) distributions from data (a) as extracted and (b) $e\gamma$ normalized to the maximum bin of the ee distribution and zoomed. A shift of ≈ 1 GeV is observed.



(a) (b)

Figure 7.25: Fit of a DSCB function convoluted with a Breit-Wigner of a fixed width $\Gamma_Z = 2.495$ GeV and an exponential to the (a) ee and (b) $e\gamma$ invariant mass in the window of $70 < m_{inv} < 180$ GeV (full blue line - signal, dashed blue line - background) to the data (black points).

- the fit of the spectra using a DSCB convoluted with a BW and an exponential for background estimation is performed in different ranges, namely $65 < m_{inv} < 185$ GeV and $75 < m_{inv} < 175$ GeV;
- a different background subtraction method is applied. An exponential fit in a window of $120 < m_{inv} < 180$ GeV is performed to the ee and $e\gamma$ spectra and the background below the Z peak is subtracted. The low-mass side of the spectrum, for $m_{inv} < 60$ GeV is not fitted due to the shape of the trigger turn-on;

Sample	$N_{e\gamma}$	N_{ee}	extracted $f_{e\to\gamma}$
data	449 768	2 957 402	$0.0760 \pm 0.0001(\text{stat}) \,^{+0.0013}_{-0.0019}$ (syst)
$Z \to ee$ MC	412 898*	2 818 422*	$0.0732 \pm 0.0001(\text{stat}) \,^{+0.0014}_{-0.0014}$ (syst)
	$N^{true}_{e(\gamma)}$	$N^{true}_{e(e)}$	true $f_{e\to\gamma}$
Z$\to ee$ MC	751 580*	14 663 227*	0.0513 ± 0.0001 (stat)

Table 7.20: $e \to \gamma$ fake rates from data and Pythia MC. (*) These numbers are from three MC samples that have been reweighted according to the generated luminosities and normalized to the luminosity in data.

- the fake rate is also computed without any background subtraction at all.

All systematic uncertainties are listed in Table 7.21. The total uncertainty corresponds to the systematic uncertainty in Table 7.20.

Method	systematic down	systematic up
extraction window	0.00026	0.00138
fit range	0.00029	0.00011
subtraction method (sym)	0.00052	0.00052
no subtraction (sym)	0.00012	0.00012
Total	0.00134	0.00189

Table 7.21: $e \to \gamma$ fake rate systematic uncertainties breakdown for data. (sym) stands for symmetrized uncertainties.

The procedure of extracting the fake rate is also repeated on $Z \to ee$ Pythia MC events, where also a true fake rate is computed. The true fake rate $f^{true}_{e\to\gamma}$ is defined as the number of true electrons passing the object selection (η, p_{T}, *tight* and isolation), which have been reconstructed as photons, divided by the number of true electrons, passing the selection, which have been reconstructed as electrons. The trigger requirement and the requirement of having at least two reconstructed photons in the event are released for this computation.

The fake rate extracted from MC using the method described above and the true fake rate differ by 2.2%, as seen in Table 7.20. An additional uncertainty of 2.2% has to be considered on the measured $e \to \gamma$ fake rate in data, due to the potential bias of the method. The dependence of the fake rate on E_{T} is not considered and should be investigated more thoroughly. The fake rate measured in data differs from the one measured in MC as well, but this just proves that one cannot rely on the MC for this measurement completely. The fake rate measurement is not performed in N_{jet} categories as there is no reason to assume it to be dependent on N_{jet} and any dependence is assumed to be covered by the systematic uncertainties.

7.4.1.2 Photon to electron fake rate

Similarly, the photon to electron fake rate can be measured from three body decays of the $Z \to ee\gamma$. By counting the number of events identified as $ee\gamma$ and eee around the Z boson mass peak, the fake rate can be derived from

$$N_{ee\gamma} = N^{sig}_{ee\gamma} \tag{7.42}$$

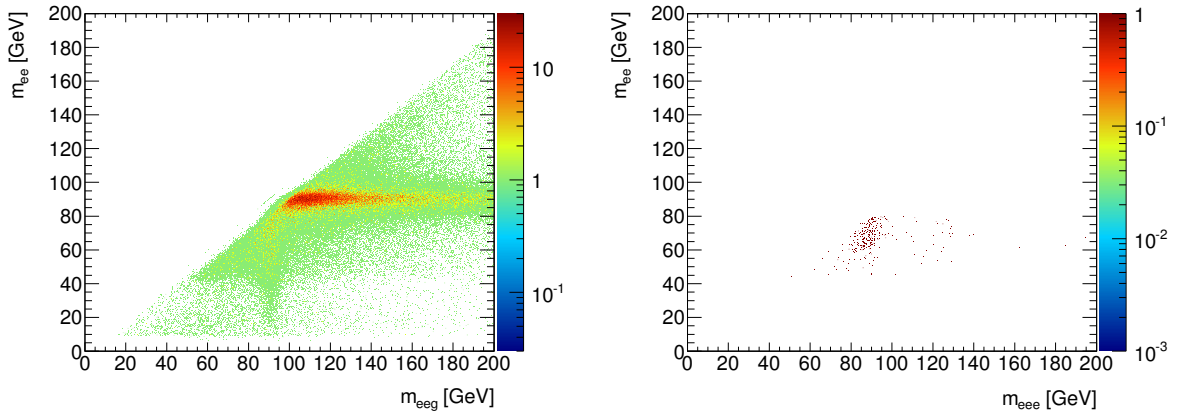
$$N_{eee} = N^{sig}_{ee\gamma} \cdot f_{\gamma \to e} \tag{7.43}$$

as

$$f_{\gamma \to e} = \frac{N_{eee}}{N_{ee\gamma}}. \tag{7.44}$$

The photon and electron selection in search for three-body decays of the Z boson is adjusted by lowering the E_{T} threshold of the three objects to 10 GeV. All objects are required to be *tight*, isolated and separated in the R-plane by $\Delta R < 0.4$. The invariant mass of the leading and subleading electrons is required to be $45 < m_{e_1 e_2} < 80$ GeV in order to reject two-body Z boson decays. Further, when building the $ee\gamma$ system, the third electron, if there is one present, cannot pass the selection. Same is required for the photons when building the eee system.

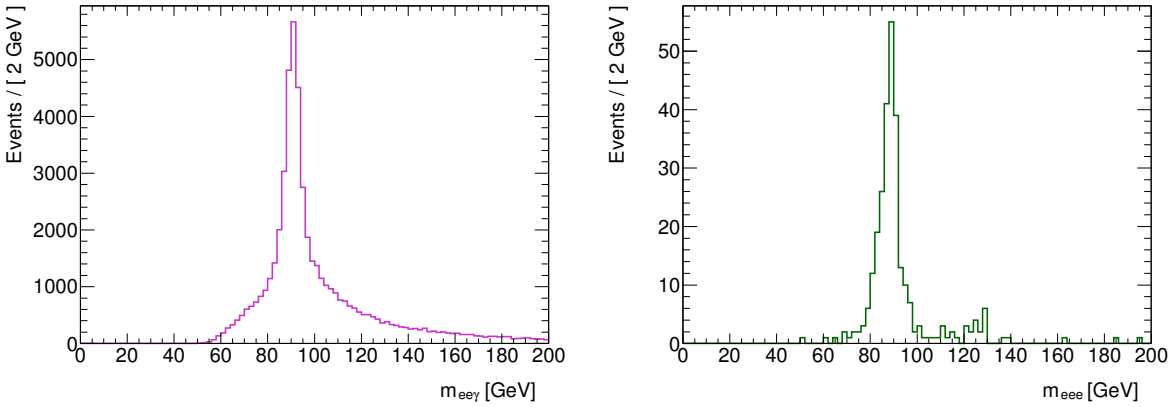
In Figure 7.26(a), the relation of the invariant mass of the $ee\gamma$ system and the ee system measured in data is shown. Here, the Z boson peak associated to two electrons and a random photon passing the selection is visible with masses of the $ee\gamma$ system above 100 GeV. By restricting the selection to the $m_{ee} < 80$ GeV region, these events are removed and an $ee\gamma$ invariant mass peak positioned at around 90 GeV is retained. The same selection is applied on three electrons, removing events where a photon would pass the selection as well, and the m_{eee} invariant mass distribution is derived. The relationship of the m_{eee} and m_{ee} distributions is shown in Figure 7.26(b) and a small amount of events at $m_{eee} \approx 90$ GeV is visible.



(a) (b)

Figure 7.26: The 2D plot of the invariant mass of the (a) $ee\gamma$ and (b) eee system vs. the ee system for data events passing the event selection and the three objects passing the selection described in the text.

After building the eee and $ee\gamma$ distributions, shown in Figure 7.27, the procedure described in Section 7.4.1.1 is used again to extract the number of events $N_{ee\gamma}$. The fit of the DSCB convoluted with a Breit-Wigner distribution and an exponential for the background is performed and the background is subtracted. The fit is shown in Figure 7.28. Since the fit does not properly describe the beginning of the peak, it was moved to start at 84 GeV. A systematic uncertainty is assigned on the fake rate due to the subtraction method by performing a simple exponential fit in the $120 < m_{ee\gamma} < 180$ GeV region and by not performing any subtraction at all. The extraction window is also shifted to be $\pm 2\sigma$ and $\pm 1\sigma$ instead of the nominal $\pm 1.5\sigma$ around the mean.



(a) (b)

Figure 7.27: The invariant mass of the (a) $ee\gamma$ and (b) eee system passing the selection as described in the text, extracted from data.

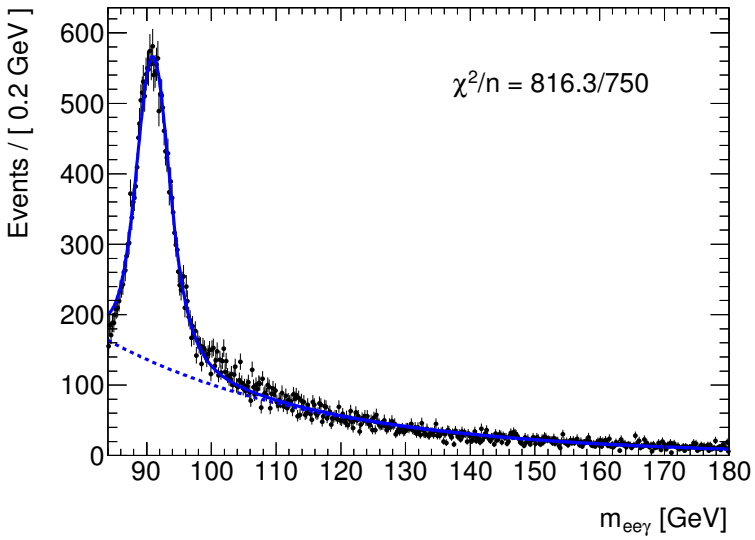


Figure 7.28: Fit of a DSCB function convoluted with a Breit-Wigner of a fixed width $\Gamma_Z = 2.495$ GeV and an exponential to the $ee\gamma$ invariant mass distribution in a window of $84 < m_{eee} < 180$ GeV (full blue line - signal, dashed blue line -background).

The m_{eee} distribution has very few events that could be considered as background, therefore a constant fit is performed in the same region and the background is subtracted from the peak. Omitting the subtraction is considered as a systematic uncertainty. The extraction window systematic is also considered. The fake rate is also re-computed when no veto is required on the third object of the other kind (e for $ee\gamma$ and photons for eee) and the systematic uncertainties on the fake rate are summarized in Table 7.22.

Method	systematic
extraction window	0.00021
fit range	0.00124
no subtraction	0.00303
no veto on third object	0.00034
Total	0.00330

Table 7.22: $\gamma \to e$ fake rate systematic uncertainties breakdown for data. All uncertainties are symmetrized to be more conservative.

The obtained $\gamma \to e$ fake rates computed according to Eq. 7.44 in data and MC are listed in Table 7.23. The two fake rates agree within their uncertainties, but nevertheless, to be conservative, when extracting the yield, a 100% uncertainty should be considered. The true fake rate is not computed for the moment.

Sample	N_{eee}	$N_{ee\gamma}$	extracted $f_{\gamma\to e}$
data	155	14 068	0.0110 ± 0.0009(stat) ±0.0033 (syst)
$Z \to ee$ MC	111*	14 447*	0.0077 ± 0.0007(stat) ±0.0042 (syst)

Table 7.23: $\gamma \to e$ fake rates measured on data and MC. (*) number of events for MC is computed from three MC samples that are combined and reweighted according to the generated luminosities and normalized to the luminosity in data.

7.4.2 Estimation of the electron impurities as a function of an observable

To estimate the contribution of the electron background as a function of any observable, the histograms of $\gamma\gamma$ events $h_{\gamma\gamma}$, $e\gamma$ events $h_{e\gamma}$ and ee events h_{ee} passing all selection criteria, as described in Section 7.2.5, are built. For each bin in a histogram, the impurity of the bin $I_e[bin]$ is computed, according to Eq. 7.38

$$\begin{aligned} I_e[bin] &= 1 - \frac{h^{sig}_{\gamma\gamma}[bin]}{h_{\gamma\gamma}[bin]} \\ &= \frac{f_{e\to\gamma}\frac{h_{\gamma e}[bin]}{h_{\gamma\gamma}[bin]} - (f_{e\to\gamma})^2\frac{h_{ee}[bin]}{h_{\gamma\gamma}[bin]}}{(1 - f_{e\to\gamma}f_{\gamma\to e})^2} - \frac{2f_{e\to\gamma}f_{\gamma\to e} - (f_{e\to\gamma}f_{\gamma\to e})^2}{(1 - f_{e\to\gamma}f_{\gamma\to e})^2}. \end{aligned} \tag{7.45}$$

Using the impurities for each bin, the electron background is subtracted from the diphoton yield as

$$h^{sig}_{\gamma\gamma}[bin] = h_{\gamma\gamma}[bin]\,(1 - I_e[bin])\,. \tag{7.46}$$

The statistical uncertainties of the ee and $e\gamma$ yields and the two fake rates are propagated through the subtraction and combined with the statistical uncertainty of the diphoton yield.

The systematic uncertainty on the $e \to \gamma$ fake rate is evaluated by re-computing the bin-by-bin impurities with the fake rate increased and decreased by its total systematic uncertainty, which is conservatively set at 50% of its size, due to the negligence of the possible E_{T} dependence and the fact that the fake rate measured in MC and the true fake rate do not agree within their uncertainties. The total systematic uncertainty considered, variated up and down is shown in

Figures 7.29 and 7.30 for the diphoton and dijet observables in the $2-jet$ category. The figures for the full list of observables can be found in Appendix E.

The uncertainty on the subtracted electron background coming from the $\gamma \to e$ fake rate is evaluated by recomputing the impurities bin by bin for the fake rate increased and decreased by 100% of its measured value. The contribution from the $f_{\gamma \to e}$ fake rate is very small due to the quadratic term in Eq. 7.38. The difference between the nominal $\gamma \to e$ fake rate and the varied fake rate up and down is shown in Figure 7.31 and 7.32 for the diphoton and dijet observables in the $2-jet$ category. Additional figures are shown in Appendix F.

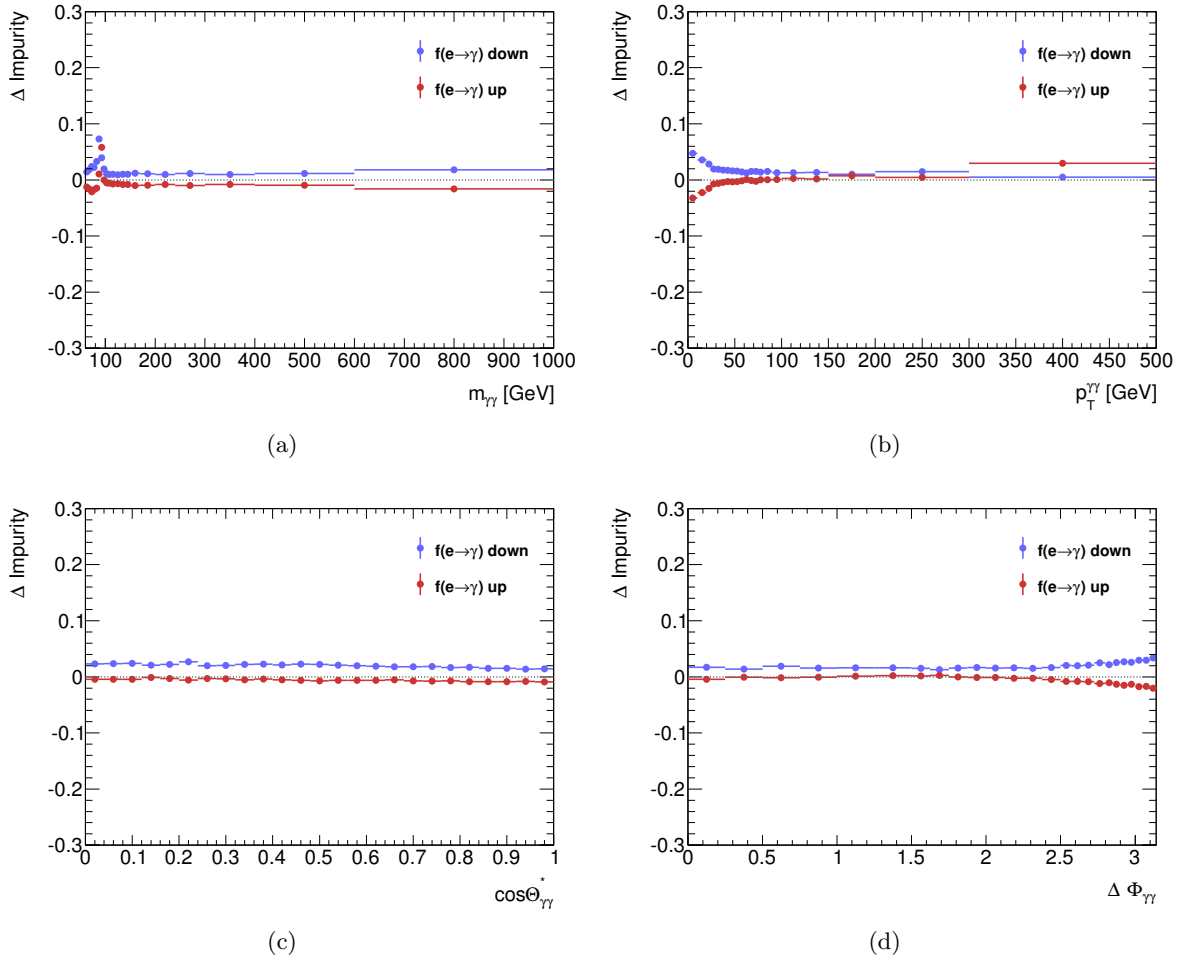


Figure 7.29: Relative change in the estimated impurities estimated as a 50% variation of the $e \to \gamma$ fake rate up and down, as a function of the diphoton observables for the $2-jet$ category.

Another uncertainty is assigned due to the calibration of the ee and $e\gamma$ yields. The electrons used to build these yields have a different E_{T} spectrum w.r.t. the photons in the diphoton yield. Consequently the Z peak of the ee and $e\gamma$ invariant mass distributions do not align with each other, nor with the Z boson peak visible in the diphoton yield. The electrons used to build the ee and $e\gamma$ yields are therefore reconstructed and calibrated as if they were photons to partially reduce the shift. The choice of what kind of photon (converted or unconverted) to use for this

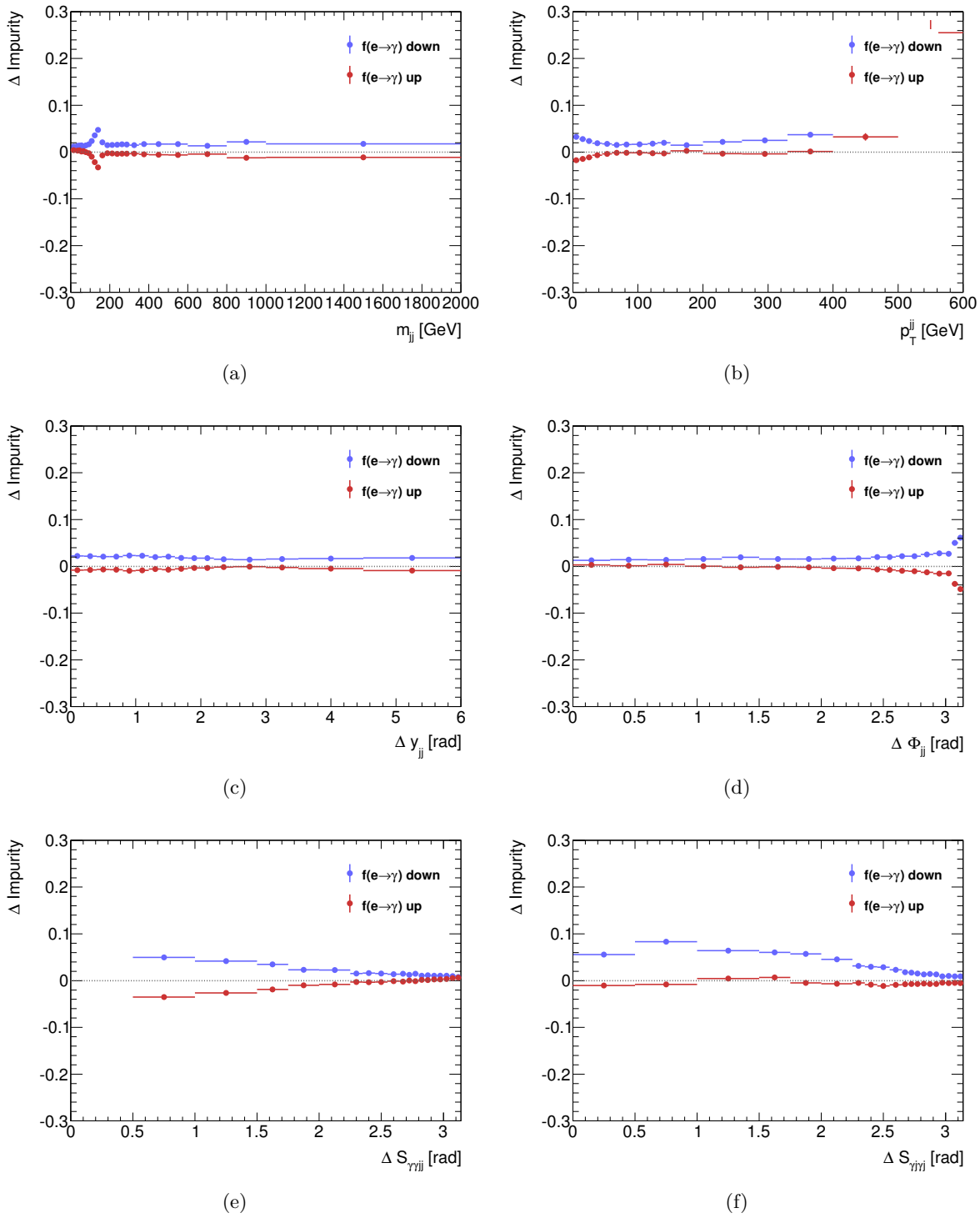


Figure 7.30: Relative change in the estimated impurities estimated as a 50% variation of the $e \to \gamma$ fake rate up and down, as a function of the dijet observables for the $2 - jet$ category.

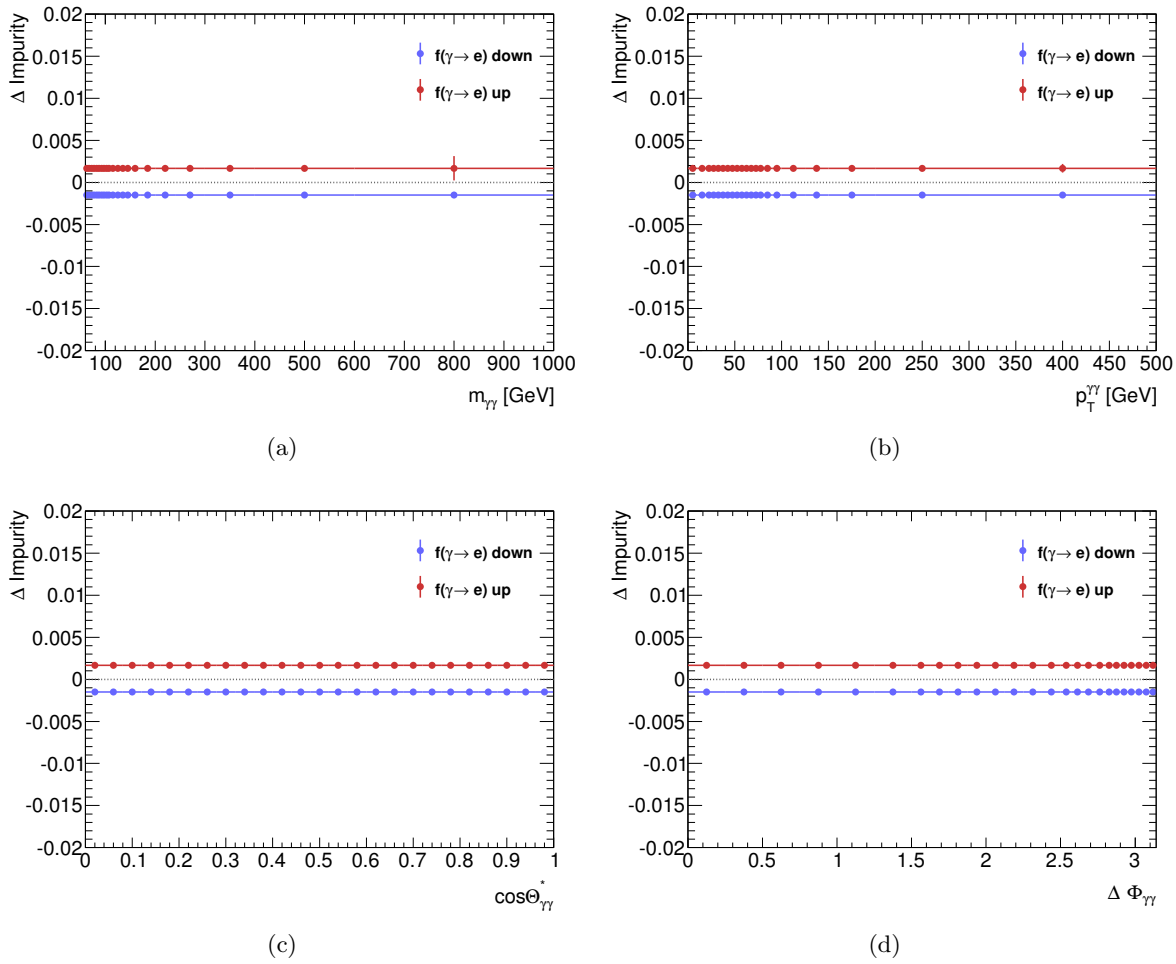


Figure 7.31: Relative change in the estimated impurities estimated as a 100% variation of the $\gamma \to e$ fake rate up and down, as a function of the diphoton observables for the $2-jet$ category.

should be studied in more detail. The conservative approach is taken in this work by deriving four sets of yields for the ee distributions, using combinations of two electrons reconstructed and calibrated as converted-converted (CC), unconverted-unconverted (UU), converted-unconverted (UC) and unconverted-converted (CU) photons. Two sets of yields are derived for the $e\gamma$ distributions as well. The average of the four (two) yields is taken as the central value and the maximum difference w.r.t. the other yields is taken as a systematic uncertainty, symmetrized. The background subtraction is illustrated in Figure 7.33 for the $0-jet$ category and the peaking background disappears with the subtraction. The difference between the subtracted yields using the average and other yields is very small.

All three sources of uncertainty described in this section are propagated through the unfolding procedure separately and the treatment of the uncertainties is again summarized in Section 7.7.1. The yields extracted from the 2x2DSB method and the yields with the electron background subtracted are shown in Figures 7.34-7.35 for the diphoton and dijet observables in the $2-jet$ category with their total (stat+syst) uncertainties. Additional figures containing the subtracted

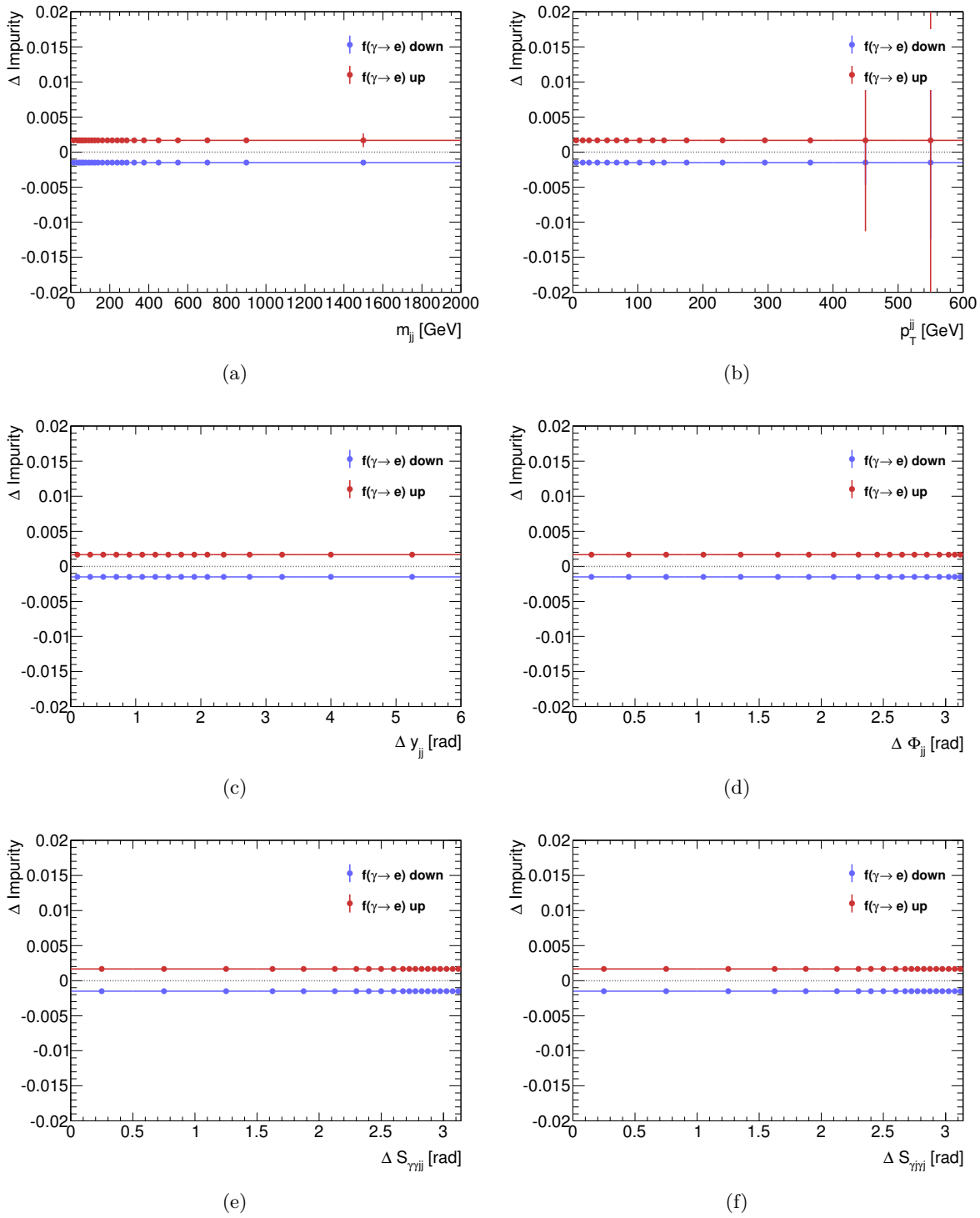


Figure 7.32: Relative change in the estimated impurities estimated as a 100% variation of the $\gamma \to e$ fake rate up and down, as a function of the dijet observables for the $2-jet$ category.

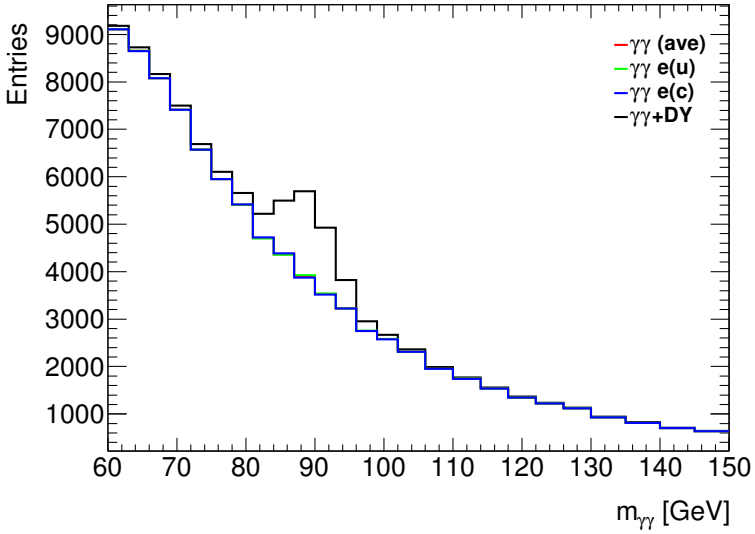


Figure 7.33: Subtraction of the electron background for the $m_{\gamma\gamma}$ observable in the $0 - jet$ category. Diphoton + DY yield as extracted by the 2x2DSB method (black), diphoton yield after the electron background subtraction using the average ee and $e\gamma$ yields computed from the different conversion categories (red), the subtracted yield using electrons reconstructed and calibrated as unconverted photons (green) and converted photons (blue).

yields as a function of the other variables can be found in Appendix G. The electron impurity is typically below 5% bin-by-bin, for the $m_{\gamma\gamma}$ observable for example, it amounts to 2% on average, except for the region of the Z boson mass, where it is roughly 30%.

7.4.3 Note on the approach

This section aims to address the differences between the treatment of the electron background in the analysis presented in Chapter 6 and here.

In the search for new resonances, the electron background was fitted together with the continuum and the potential signal of the new resonance X. The model used to describe the contribution was a template created by building the invariant mass distribution of two electrons. These electrons, passing the same selection as photons (except for *tight*, instead passing *tight*++), could not overlap with any reconstructed photon and only events which were not classified as diphoton events were used. The template was built for the three conversion categories separately, due to the shift observed between them, and the need to fit each category on its own in the final fit. The transformation of the electron E_{T} was done according to Eq. 6.3 and 6.4. In the fit, the normalization of the template was done by applying the measured fake rate of $e \to \gamma$, constrained within its uncertainties. The fake rate of $\gamma \to e$ was neglected.

It is true that the selection of events for the two analyses is very similar, the only difference being the isolation requirement. A very similar approach could be used for the cross section measurement as well, but in addition, this analysis needs to estimate the electron background as a function of all other observables and not just the invariant mass and consider the continuum

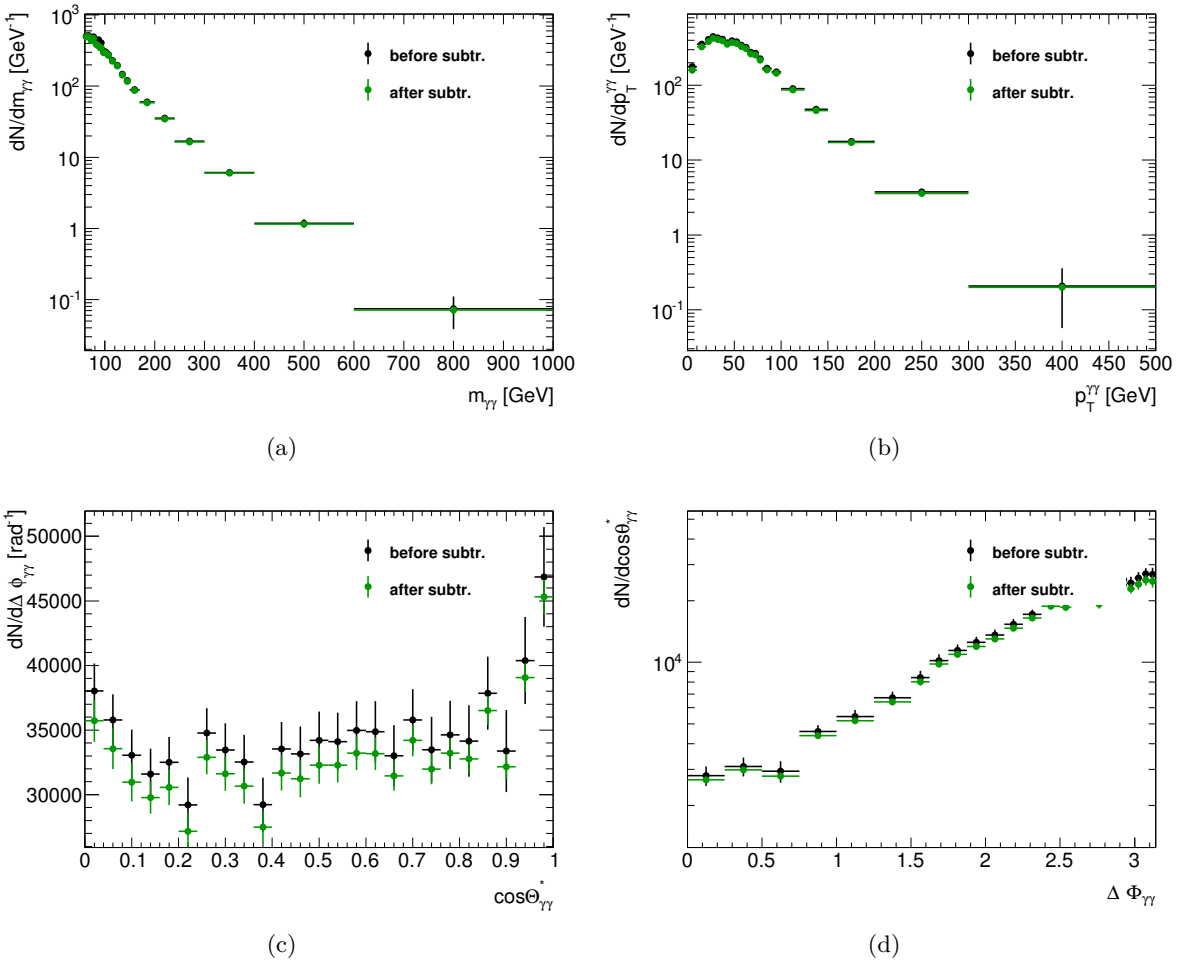


Figure 7.34: 2x2DSB diphoton yields (black) and yields after electron background subtraction (green), as a function of the diphoton observables for the $2-jet$ category.

background as well. This analysis could benefit from the separation of this background into categories, which would allow for a proper subtraction of the Z boson peak, since in Figure 7.33, the change of the slope around 90 GeVand further away from the original peak is visible. This indicates a not completely correct subtraction, which is covered by the systematic uncertainties, but could be estimated better.

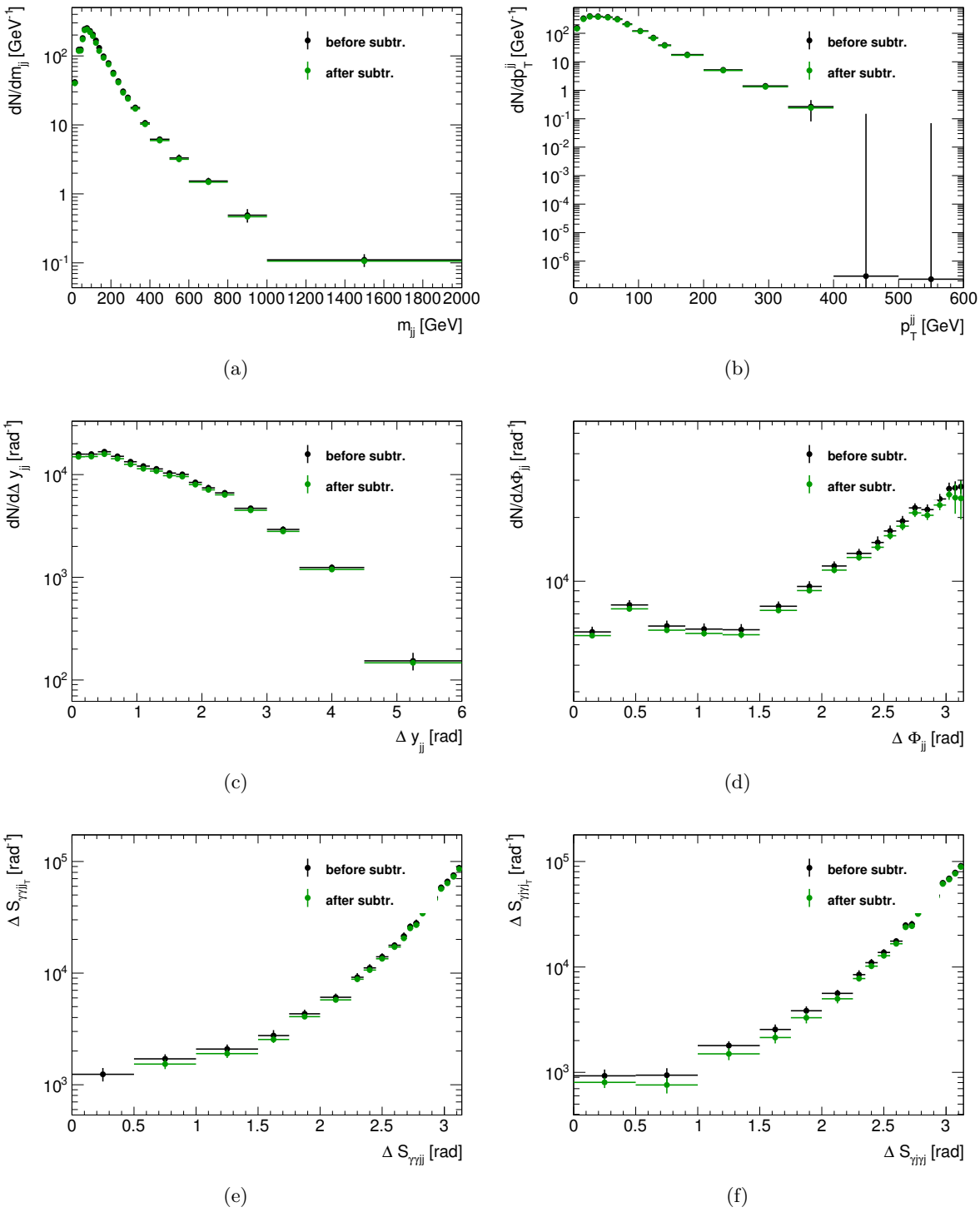


Figure 7.35: 2x2DSB diphoton yields (black) and yields after electron background subtraction (green), as a function of the dijet observables for the $2-jet$ category.

7.5 Contribution of the Higgs boson decaying into photon pairs

The recently discovered Higgs boson also contributes to the diphoton yield with its $H \to \gamma\gamma$ decay. The Higgs boson also interferes with the diphoton continuum [244].

The Higgs contribution is quantified by rerunning the full analysis selection on SM Higgs MC samples for the five different production processes, as described in Section 6.3, at the mass of $m_H = 125$ GeV. The obtained yields[1] are shown in Figures 7.36-7.38 and are always negligible in comparison to the QCD-induced diphoton yield. They are covered by systematic uncertainties assigned on the yield already, for example at the Higgs peak in the $m_{\gamma\gamma}$ observable. The Higgs contribution is therefore not subtracted in this analysis also due to the presence of interference between the diphoton continuum and the Higgs decays into diphotons. Considering the little knowledge there is about this interference [244], the subtraction would not be well defined.

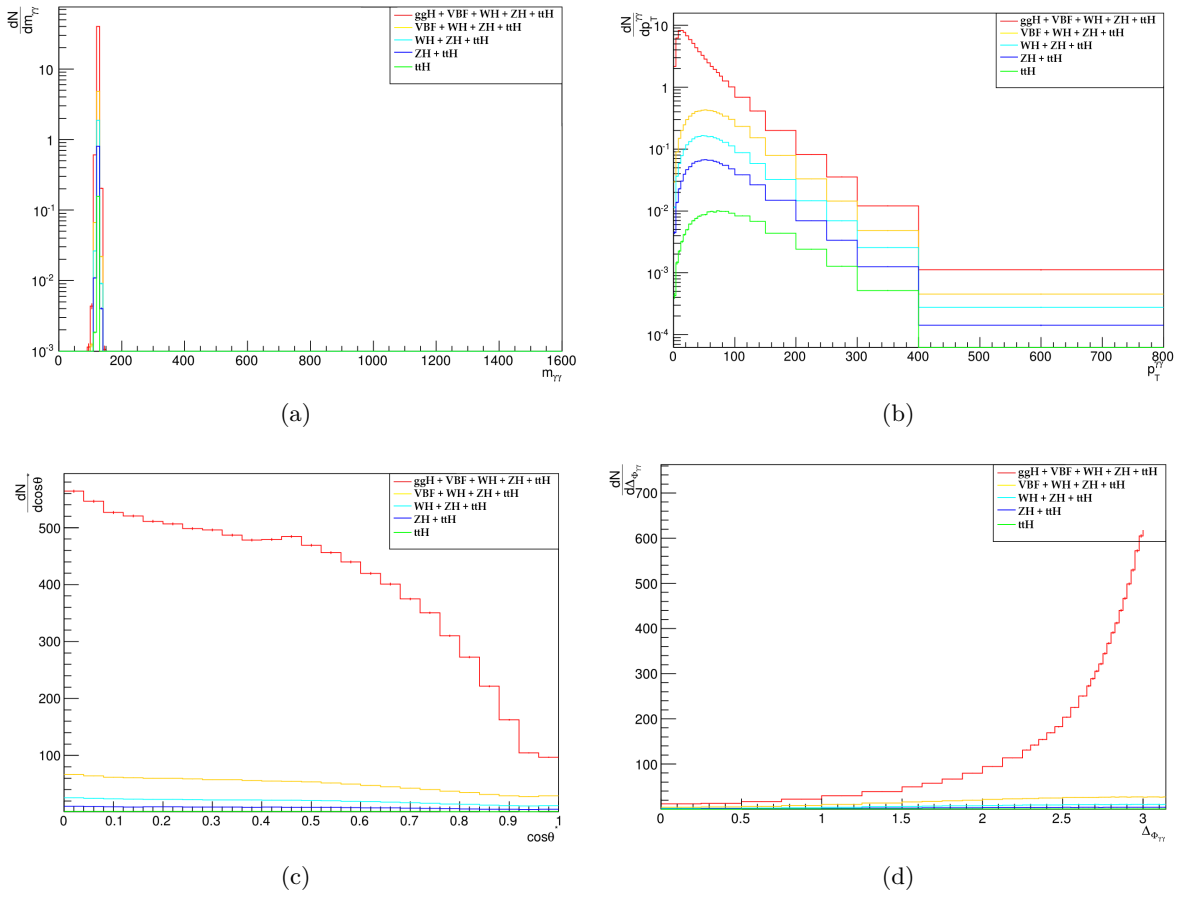


Figure 7.36: Higgs contribution for the five SM Higgs production processes as a function of the diphoton observables for the $0-jet$ category [179].

[1] These yields were computed for inclusive jet multiplicity categories. This will be updated.

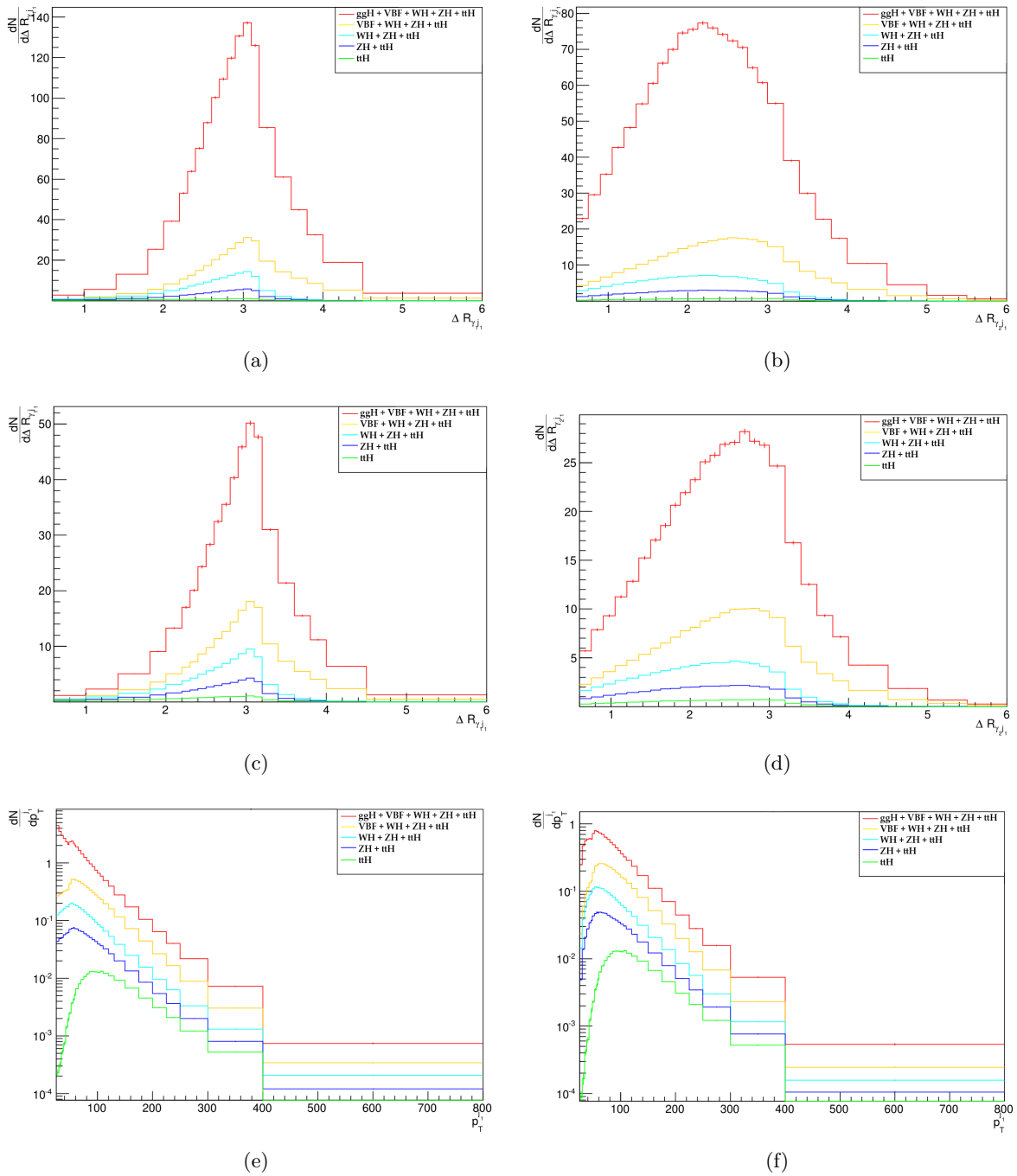


Figure 7.37: Higgs contribution for the five SM Higgs production processes as a function of the R-separation between the photons and the leading jet for the $1-jet$ (top) and $2-jet$ (middle) category and the p_{T} of the leading jet for the $1-jet$ (e) and $2-jet$ (f) category [179].

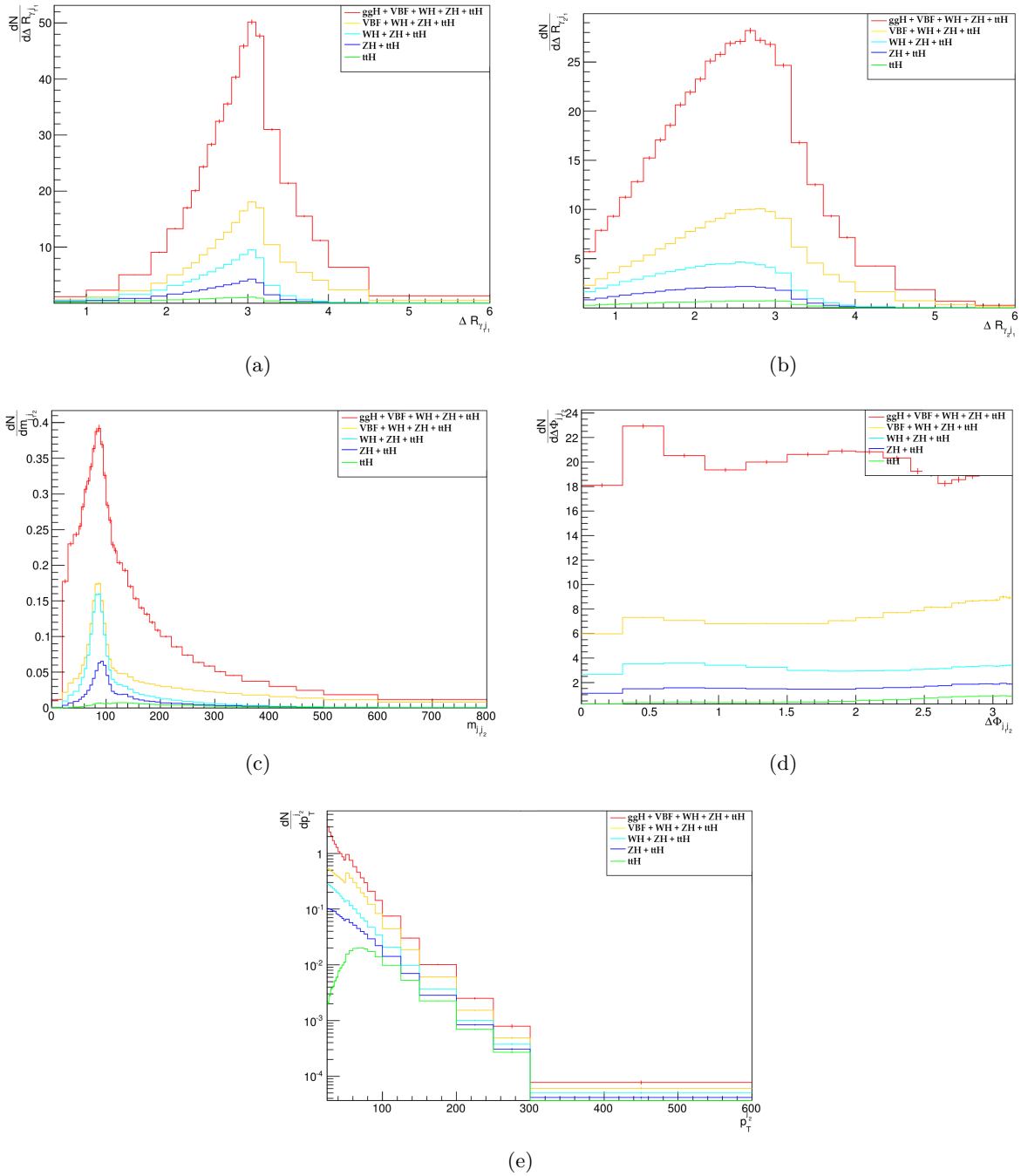


Figure 7.38: Higgs contribution for the five SM Higgs production processes as a function of the R-separation between the photons and the subleading jet, the dijet variables and the p_{T} of the subleading jet for the $2-jet$ category [179].

7.6 Unfolding

In order to be able to compare the measured data to theoretical predictions and also between different experiments, the signal reconstruction efficiency of the ATLAS detector, the efficiencies of the selection criteria, detector-induced distortions of the measured distributions and the presence of pile-up jets should be evaluated by the process of unfolding. For this, MC events are used to derive the mentioned efficiencies and bin migration effects by relating the reconstructed quantities to the true quantities.

Several methods are conventionally used to unfold measured distributions in ATLAS, such as the simple bin-by-bin unfolding and Bayesian iterative unfolding [245]. For this analysis, the SVD unfolding [246] is used, for reasons described in the following.

7.6.1 SVD unfolding

If the measured values are stored in vector $\boldsymbol{y}$ and the true values of an observable are stored in vector $\boldsymbol{x}$, the $\hat{\boldsymbol{A}}$ the detector response matrix, relates the true and measured quantities as

$$\hat{\boldsymbol{A}}\boldsymbol{x} = \boldsymbol{y}. \tag{7.47}$$

$\hat{\boldsymbol{A}}$ contains events passing both the true and the reconstruction selection and is obtained from MC. The Singular Value Decomposition (SVD) unfolding method [246], is based on the decomposition of the detector response matrix $\hat{\boldsymbol{A}}$ into a diagonal matrix $\boldsymbol{S}$ and two orthogonal matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ as

$$\hat{\boldsymbol{A}} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T, \qquad \hat{\boldsymbol{A}}^{-1} = \boldsymbol{V}\boldsymbol{S}^{-1}\boldsymbol{U}^T. \tag{7.48}$$

The diagonal elements of $\boldsymbol{S}$ are the singular values of $\hat{\boldsymbol{A}}$. Without any additional constrains, this can lead to largely fluctuating solutions. The SVD unfolding method, as implemented in the `TSVDUnfold` class in ROOT, uses a linear algorithm which is equivalent to minimizing the matrix equation [247]

$$\chi^2(\boldsymbol{x}) = (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y})^T \boldsymbol{V}_{yy}^{-1}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}) + \tau^2(\boldsymbol{x} - x_b)^T(\boldsymbol{L}^T\boldsymbol{L})(\boldsymbol{x} - x_b), \tag{7.49}$$

where $\boldsymbol{x}$ is a vector of unknowns, $\boldsymbol{y}$ is a vector of measured values, $\boldsymbol{V}_{yy}$ is the covariance matrix of $\boldsymbol{y}$. $\boldsymbol{A}$ is a matrix of probabilities and relates to $\hat{\boldsymbol{A}}$ as

$$\boldsymbol{A} = \frac{\hat{\boldsymbol{A}}}{\boldsymbol{x}} \tag{7.50}$$

The second part of the Eq. 7.49, the regularization term, ensures the stability of the result against large fluctuations. τ is the regularization strength parameter, which is typically chosen to be $\tau > N_{bins}/2$, and its value needs to be determined for each observable. The $\boldsymbol{L}$ is the regularization condition, and x_b the regularization bias. The regularization condition $\boldsymbol{L}$ is also called the curvature matrix, which can take into account the relationship between neighboring bins properly and dampen large oscillations in the unfolded spectra.

For this analysis, the diphoton yields differential in a certain quantity (e.g. $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$), are measured in jet categories[2]. In order to properly take into account the bin migration between

[2]Jet multiplicity categories for each observable measured are listed in Table 7.1.

the jet multiplicities at the same time as the bin migration due to detector resolution, the curvature matrix is modified to properly describe the relationship of neighboring bins when linearizing the 2D problem into 1D. Yields in the different N_{jet} categories are linearly combined into one histogram (not summed, just merged one after the other) and unfolded at the same time.

7.6.2 Detector response matrices

The 2D detector response matrices are filled for each observable, relating the reconstructed quantities to their true simulated values. For this, SHERPA MC events with two photons and up to three jets in the final state are used. True values are extracted by requiring the event to pass the selection at the truth level, while reconstructed values are derived after passing the analysis selection, as described in Section 7.2.5. The two highest-E_{T} photons are required to be true photons, but the E_{T} ordering is not matched with the true conditions. This way, the swap of the leading and subleading photons, which can occur in data at the reconstruction level is evaluated as a possible bin migration.

The reconstructed jets are not required to be matched to true jets. The jet selection is applied, together with the JVF selection which, as it was determined in Section 7.2.4, leaves up to about 10% of pile-up jets among the selected reconstructed jets. Since it is not possible to distinguish true and pile-up jets in data, the migration matrices have to account for the background of pile-up jets selected as one of the three jets from the hard scattering. An uncertainty will be associated both to the JVF cut and the inaccuracy of the MC prediction of the pile-up jets, as described in Section 7.7.5 and 7.7.6, respectively.

The response matrices are shown in Figures 7.39-7.55, and are mostly diagonal, meaning that the true and reconstructed quantities agree well. Some off-diagonal elements are also present, and are more pronounced in case of the angular properties of the photons and jets combined, where the possibility of a leading-subleading object misidentification arises. Migrations between N_{jet} categories are also sizeable, and are treated accordingly by the 2D unfolding. In the figures, the bins of the matrices are divided by the bin width of each bin (in one axis), to ensure the continuity of the depicted spectra, but for the unfolding procedure itself, the matrices are filled without dividing the bin content by the bin width. The measured diphoton yields are treated in the unfolding without the bin width division as well, and are normalized to the bin width only after the unfolding.

7.6.3 Regularization

To perform the unfolding itself, an optimal regularization parameter needs to be chosen. The regularization is optimized only as a function of the observables and not as a function of jet multiplicity. The regularization parameter is determined by varying the shape of the predicted reconstructed MC spectra using toy studies, therefore introducing different biases. Biases considered for this analysis are [179]:

- for each jet category, keep the first bin of an observable as is, and increase the content of the last bin by 100%, the rest of the bins follow a linear increase;
- same, but increasing the last bin by 75%, 50%, 25%, and decreasing by 25%, 50%;

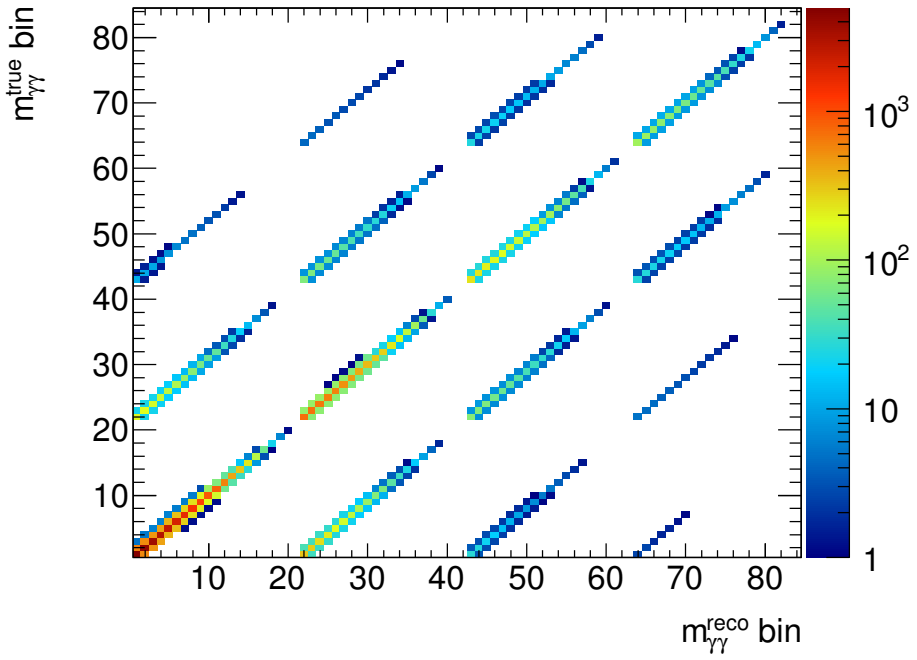


Figure 7.39: Response matrix associating the reconstructed and true simulated $m_{\gamma\gamma}$ for the four jet multiplicity categories. Bins start at 0.5 and the first 21 bins represent the $0-jet$ category, bins 22-43, the $1-jet$ category and so on.

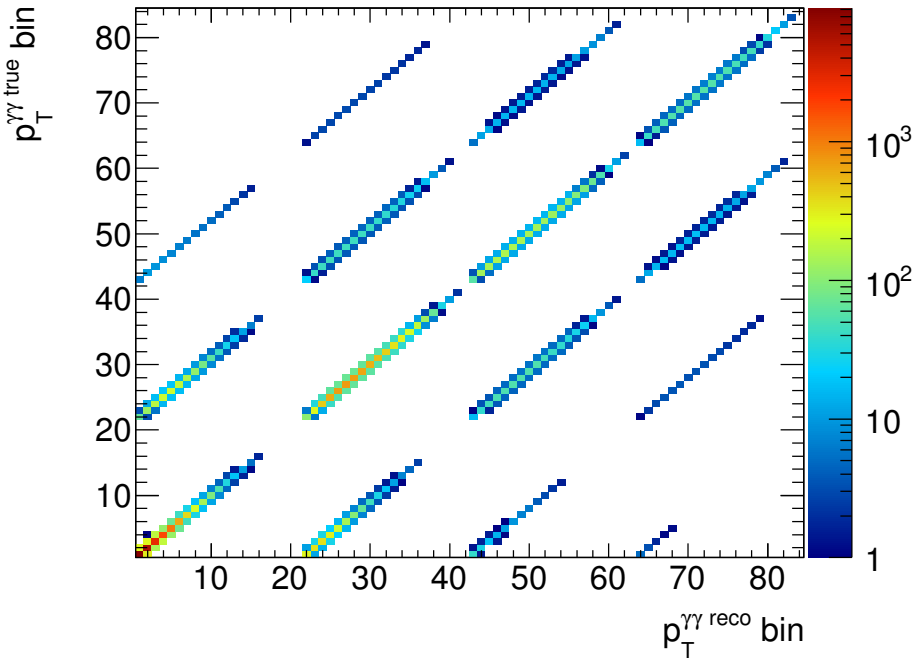


Figure 7.40: Response matrix associating the reconstructed and true simulated $p_{\mathrm{T}}^{\gamma\gamma}$ for the four jet multiplicity categories. Bins start at 0.5 and the first 21 bins represent the $0-jet$ category, bins 22-43, the $1-jet$ category and so on.

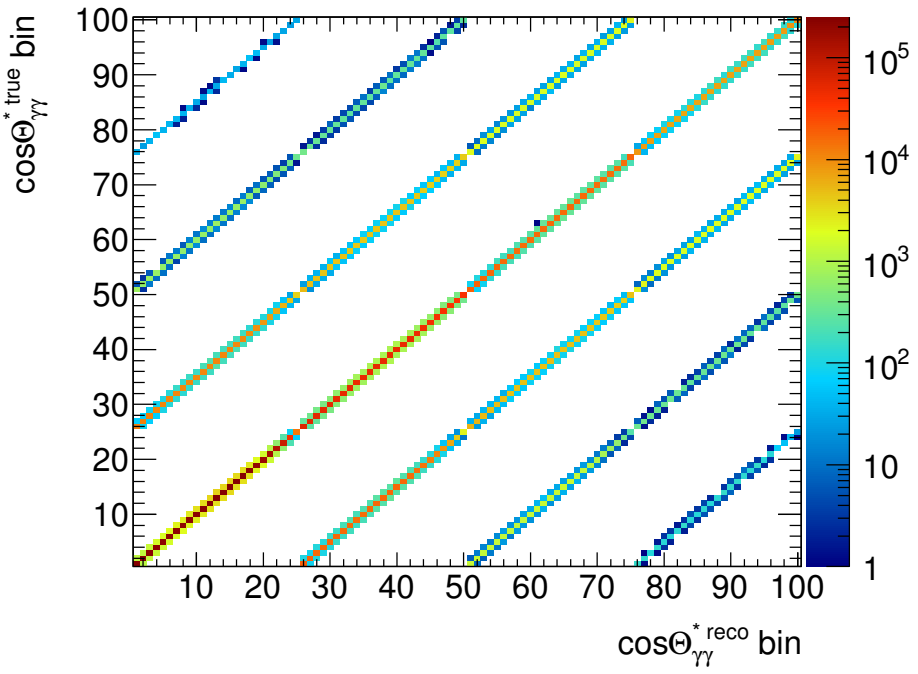


Figure 7.41: Response matrix associating the reconstructed and true simulated $\cos\theta^*_{\gamma\gamma}$ for the four jet multiplicity categories. Bins start at 0.5 and the first 25 bins represent the $0-jet$ category, bins 26-51, the $1-jet$ category and so on.

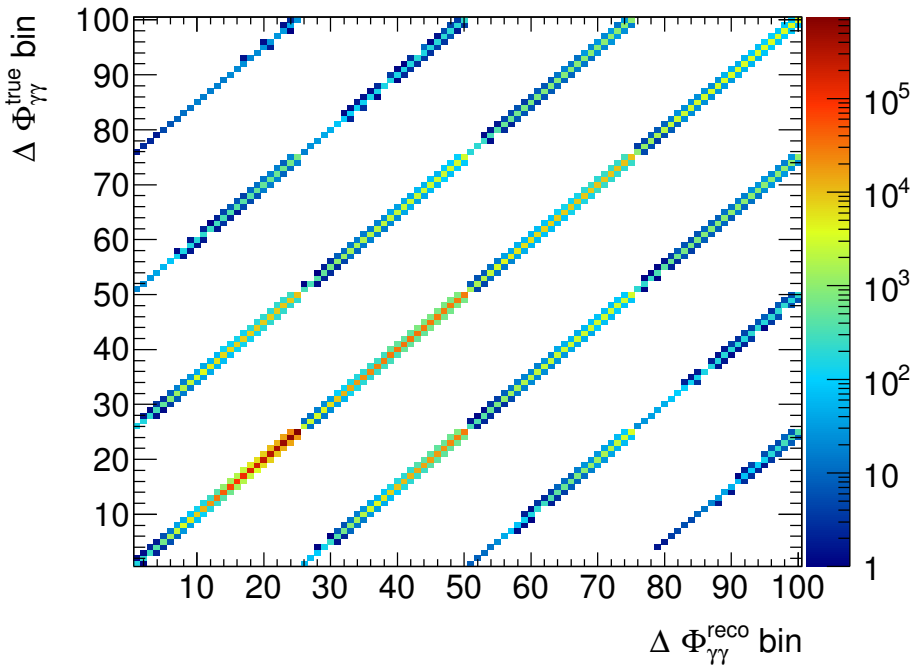


Figure 7.42: Response matrix associating the reconstructed and true simulated $\Delta\phi_{\gamma\gamma}$ for the four jet multiplicity categories. Bins start at 0.5 and the first 25 bins represent the $0-jet$ category, bins 26-51, the $1-jet$ category and so on.

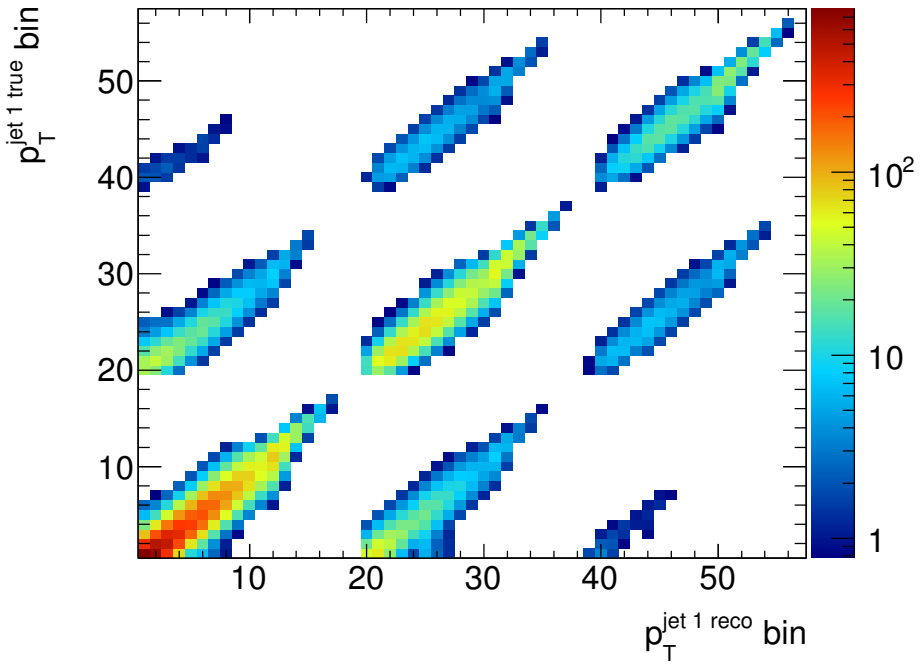


Figure 7.43: Response matrix associating the reconstructed and true simulated $p_{\mathrm{T}}^{jet_1}$ for three jet multiplicity categories, where this observable is defined ($1-jet$, $2-jet$ and $\geq 3-jet$).

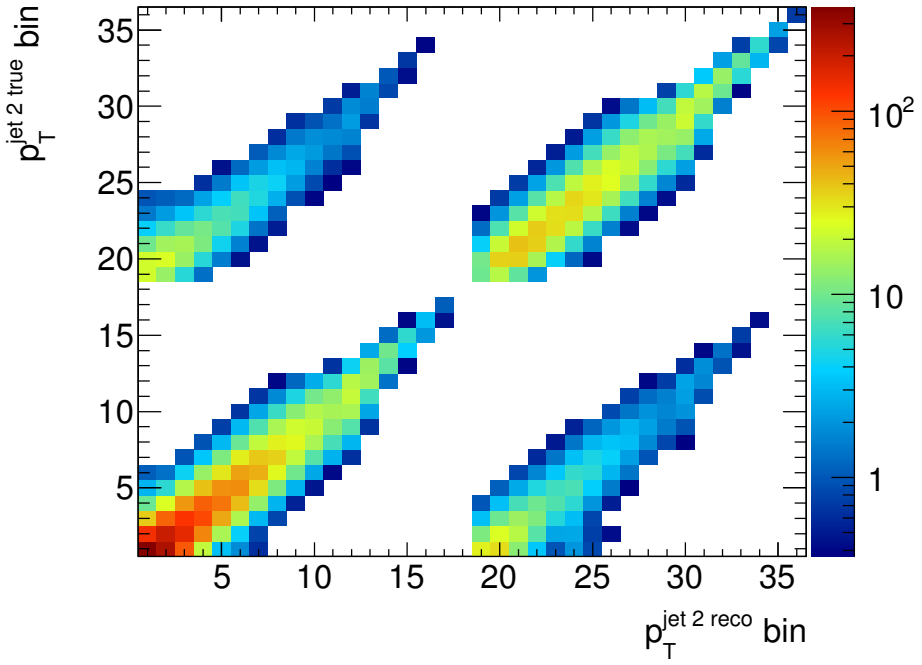


Figure 7.44: Response matrix associating the reconstructed and true simulated $p_{\mathrm{T}}^{jet_2}$ for two jet multiplicity categories, where this observable is defined ($2-jet$, $\geq 3-jet$).

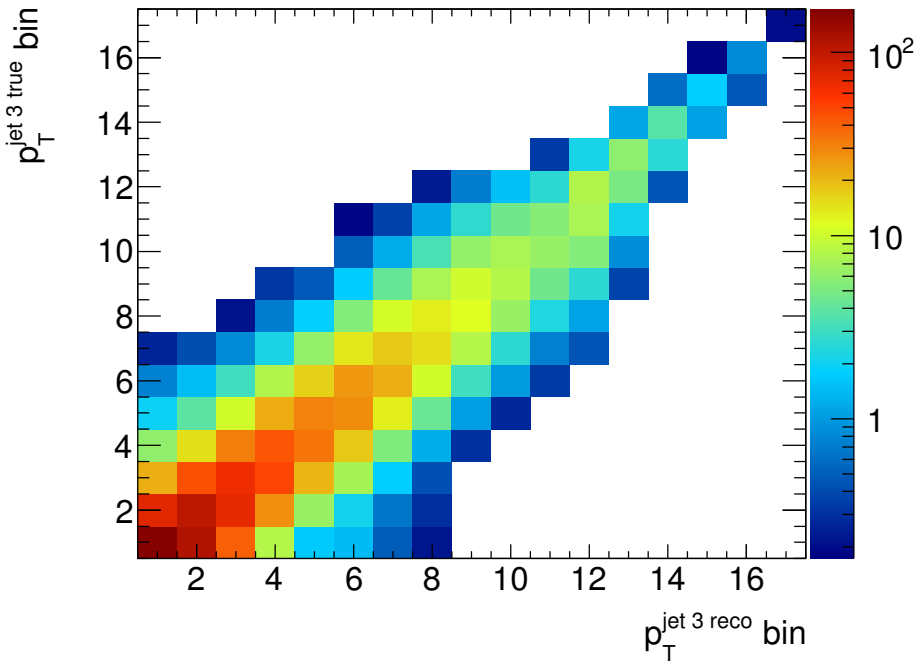


Figure 7.45: Response matrix associating the reconstructed and true simulated $p_{\mathrm{T}}^{jet_3}$ for the $\geq 3-jets$ jet multiplicity category.

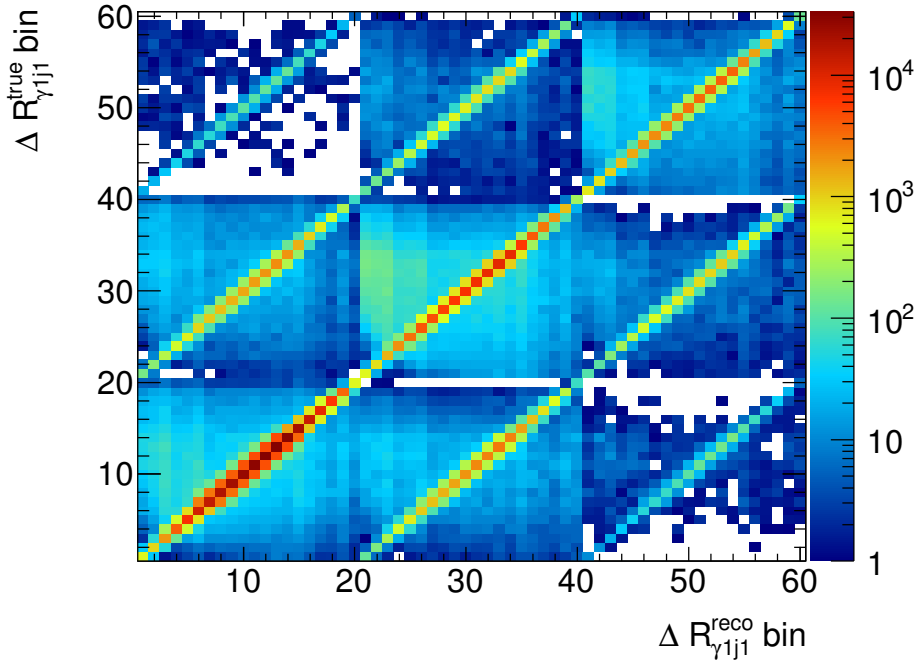


Figure 7.46: Response matrix associating the reconstructed and true simulated $\Delta R_{\gamma_1 j_1}$ for three jet multiplicity categories, where this observable is defined ($1-jet$, $2-jet$ and $\geq 3-jet$).

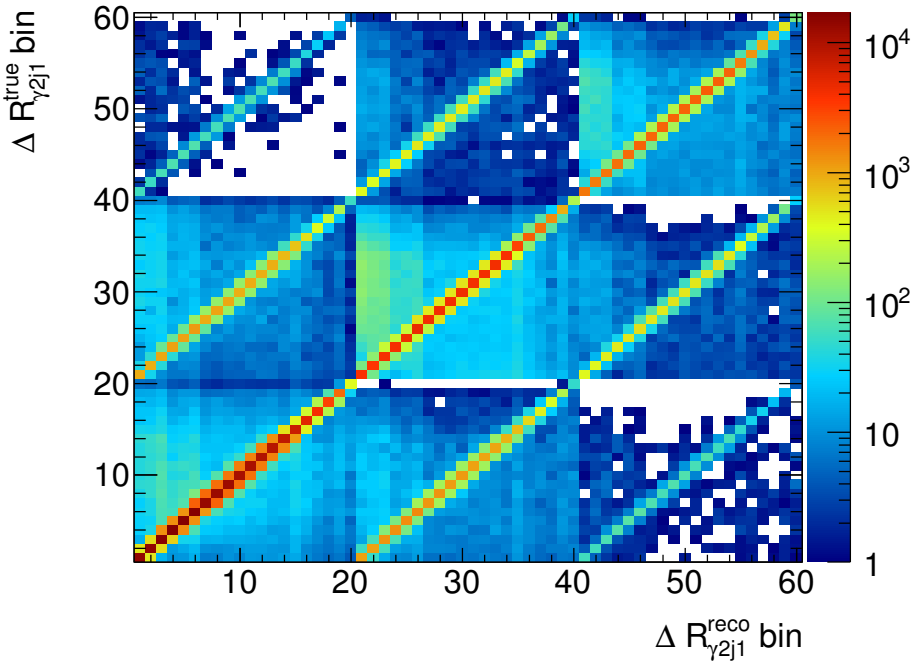


Figure 7.47: Response matrix associating the reconstructed and true simulated $\Delta R_{\gamma_2 j_1}$ for three jet multiplicity categories, where this observable is defined ($1-jet$, $2-jet$ and $\geq 3-jet$).

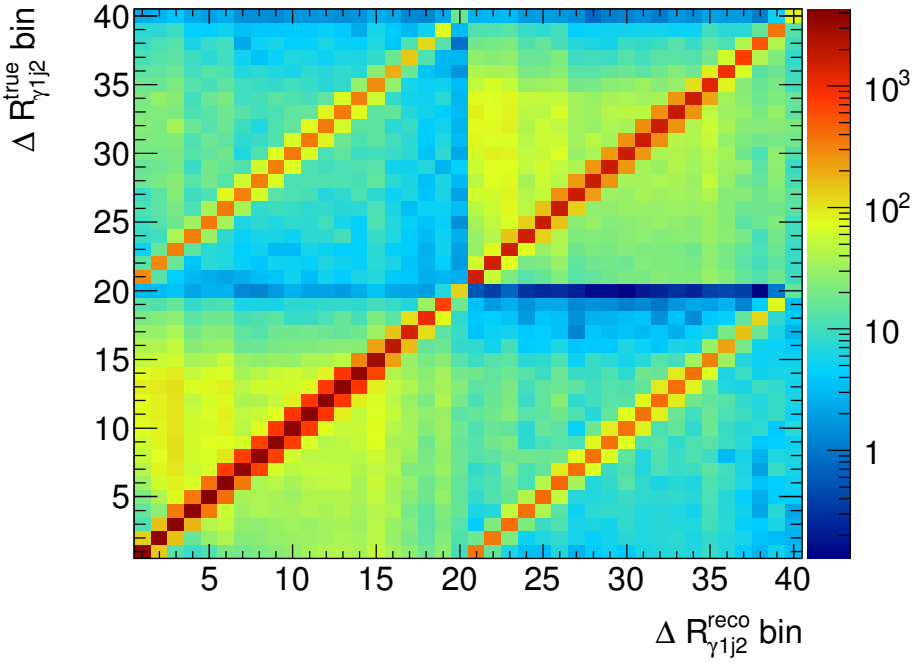


Figure 7.48: Response matrix associating the reconstructed and true simulated $\Delta R_{\gamma_1 j_2}$ for two jet multiplicity categories, where this observable is defined ($2-jet$ and $\geq 3-jet$).

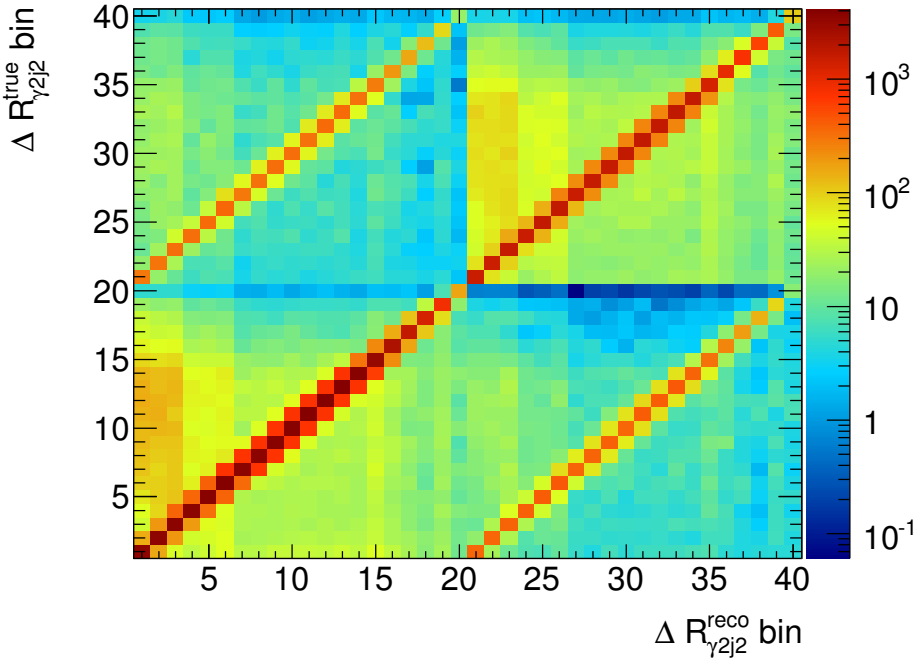


Figure 7.49: Response matrix associating the reconstructed and true simulated $\Delta R_{\gamma 2 j_2}$ for two jet multiplicity categories, where this observable is defined ($2-jet$ and $\geq 3-jet$).

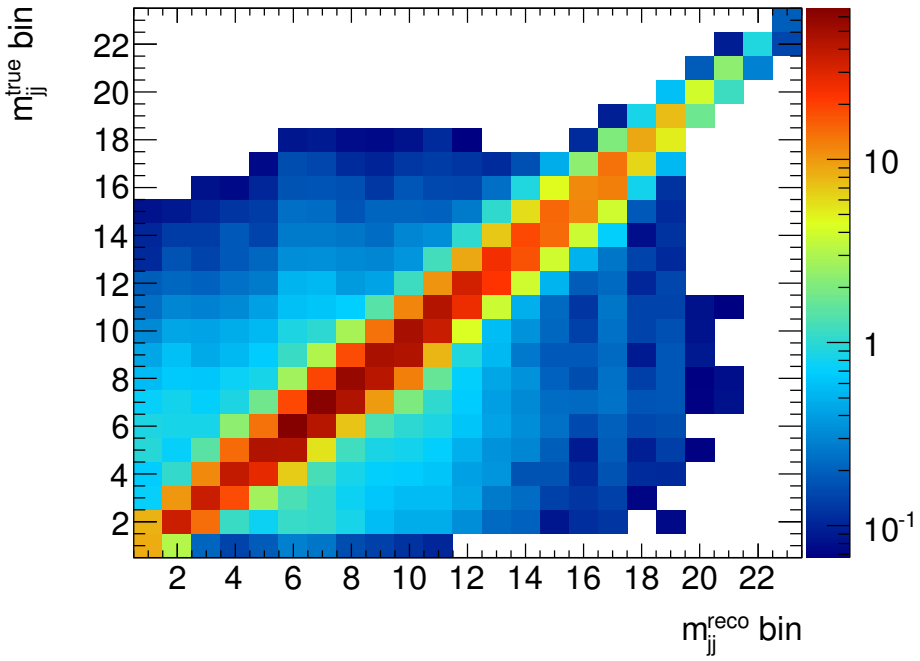


Figure 7.50: Response matrix associating the reconstructed and true simulated m_{jj} for the $2-jet$ category.

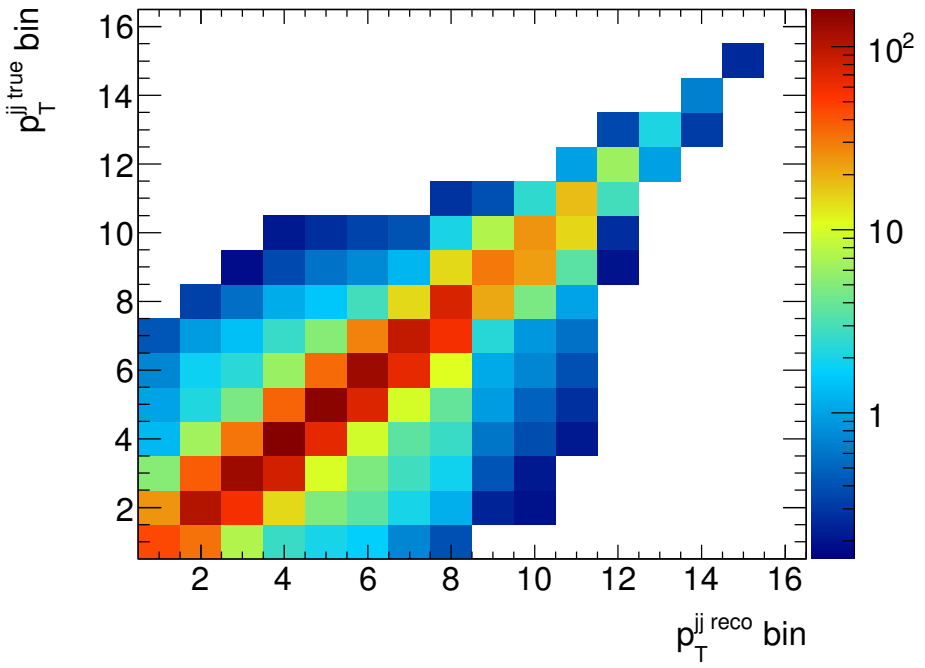


Figure 7.51: Response matrix associating the reconstructed and true simulated p_{T}^{jj} for the $2-jet$ category.

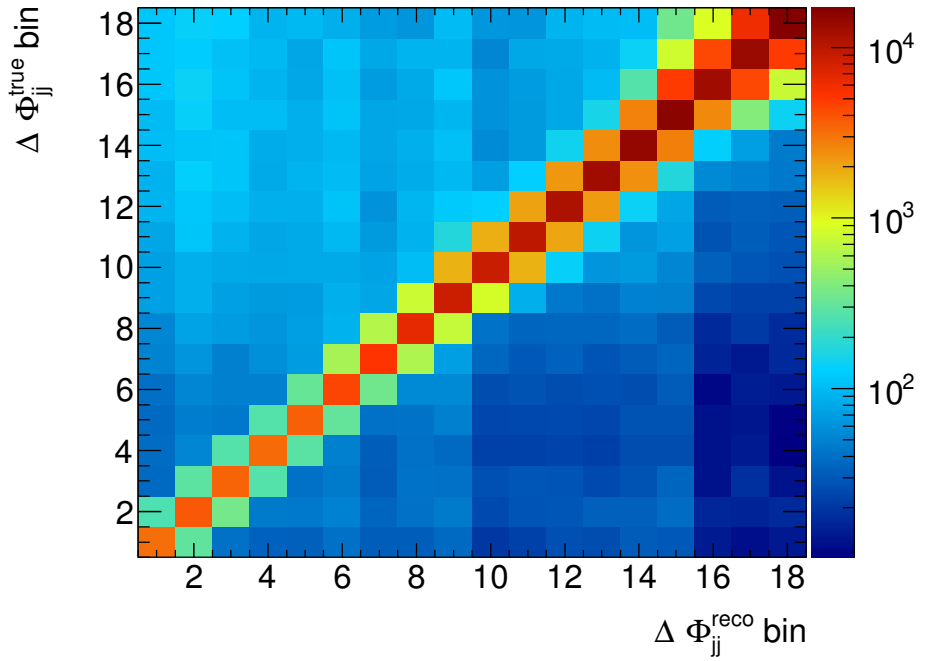


Figure 7.52: Response matrix associating the reconstructed and true simulated $\Delta\phi_{jj}$ for the $2-jet$ category.

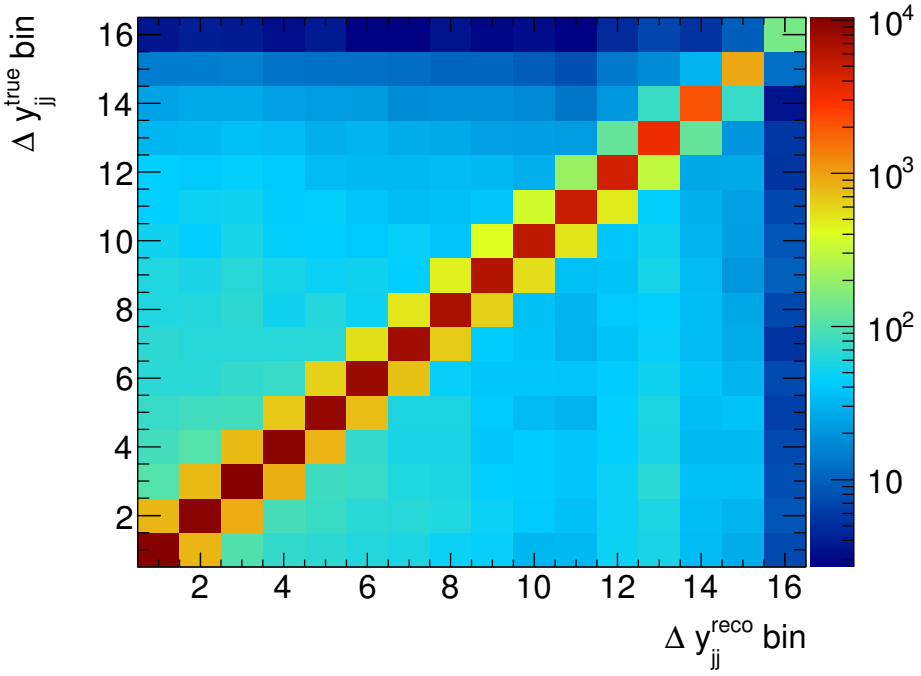


Figure 7.53: Response matrix associating the reconstructed and true simulated Δy_{jj} for the $2-jet$ category.

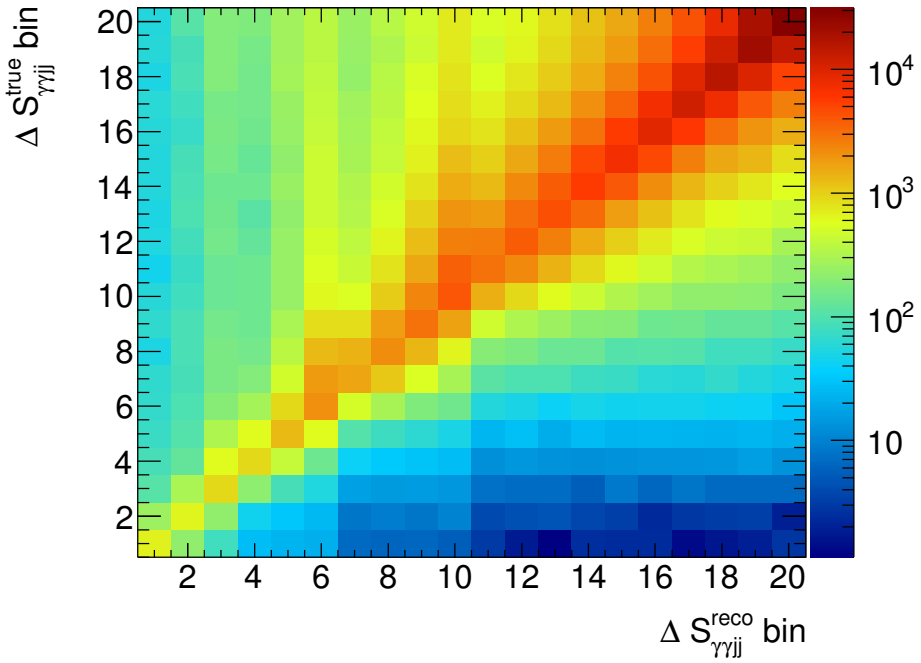


Figure 7.54: Response matrix associating the reconstructed and true simulated $\Delta S_{\gamma\gamma jj}$ for the $2-jet$ category.

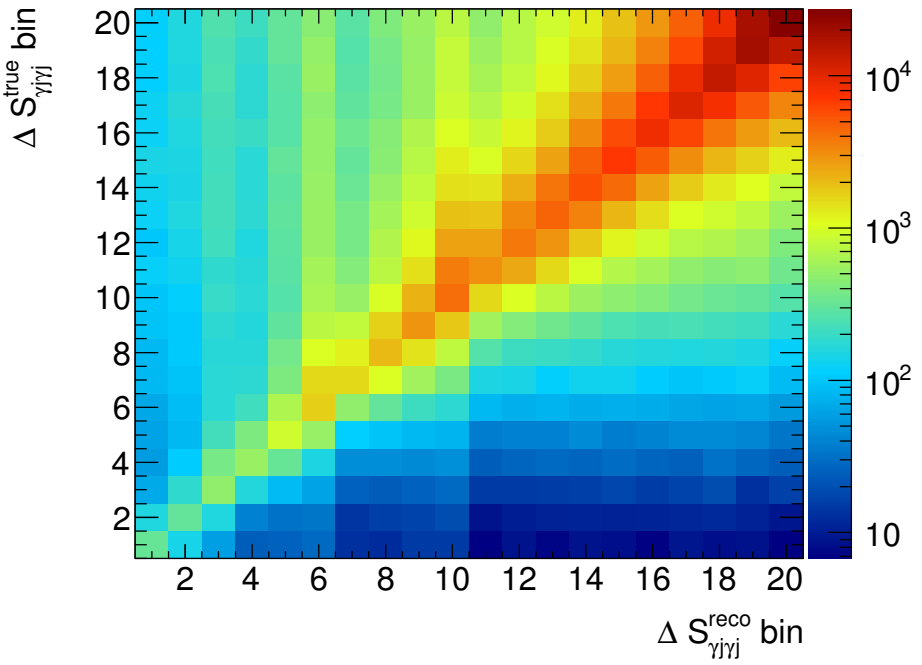


Figure 7.55: Response matrix associating the reconstructed and true simulated $\Delta S_{\gamma j \gamma j}$ for the $2-jet$ category.

- variation of the N_{jet} distribution, bins in the first jet category are unchanged, bins in the last jet category are increased by 50%;
- same, but reduced by 30%;
- a parabolic modification of the spectrum, where the central bin is increased by 50% and decreased by 30%.

Skewing the shape of the spectra in all possible directions allows to cover a wide range of potential biases associated to the imperfect description of the data spectra by the MC. The regularization parameter τ, or the regularization strength, can range from one to the number of bins in a given spectrum of an observable. In general, smaller values of the regularization parameter lead to a stronger regularization, therefore smaller statistical uncertainties but larger biases. The ideal τ is chosen as the value that has minimal bias for all biases considered, and the smallest spread between the biases. The maximum absolute bias of a given bin is taken as a systematic uncertainty on the unfolded yield. The absolute bias for $m_{\gamma\gamma}$ observable for the different modifications of the shape, is shown in Figure 7.56.

7.6.4 Unfolded distributions

Using the migration matrices and determining the appropriate regularization parameter, the unfolded distributions are obtained. The cross section can be defined in each bin as

$$\sigma = \frac{N^{true}}{L}, \tag{7.51}$$

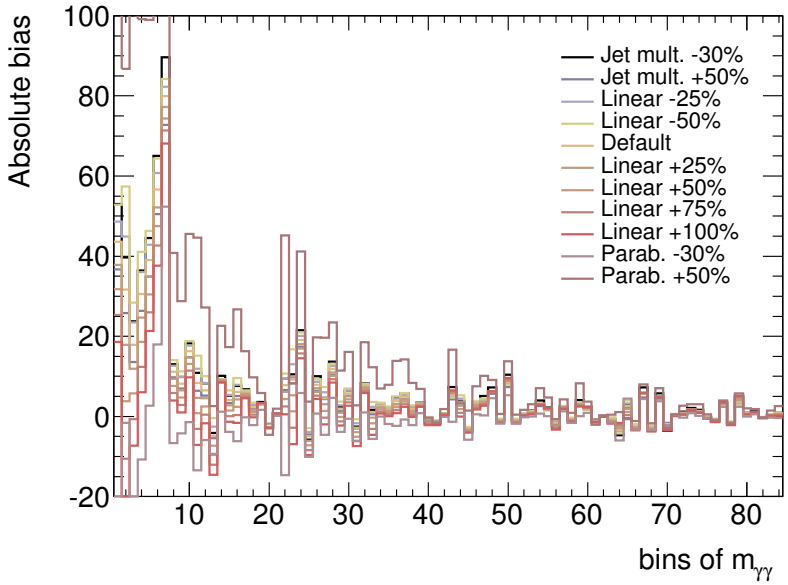


Figure 7.56: Illustration of the choice of the regularization parameter.

where N^{true} is the number of events after unfolding and L is the integrated luminosity. The exact central value for the luminosity used to obtain the cross sections is 20 276.9 pb^{-1} [248]. The integrated luminosity was measured by the luminosity group as 20.3 fb^{-1} $\pm$ 2.8% [131]. The luminosity uncertainty on the final cross sections is not propagated for the moment. Figures 7.57-7.60 show the unfolded spectra divided by the luminosity for each observable and for the available N_{jet} categories, uncertainties are statistical and total (statistical+systematic) uncertainties. The systematic uncertainties are missing part of the photon energy scale, the jet energy scale and the luminosity systematic.

The cross sections decrease with N_{jets} and the shape of most distributions changes significantly between the $0-jet$ category and categories with jets, especially for the diphoton observables, shown in Figure 7.57. The $p_{\mathrm{T}}{}^{\gamma\gamma}$ differential cross section in the $0-jet$ category falls very steeply towards zero, the distribution is practically depleted above $p_{\mathrm{T}}^{\gamma\gamma} > 80$ GeV. Jet p_{T} distributions become harder both with the number of jets and the jet p_{T} ordering, as seen in Figure 7.58. The $\gamma - j$ separation observables $\Delta R_{\gamma j}$, shown in Figure 7.59, illustrate the fact that the difference between the photon separation with the leading jet strongly depends on the photon p_{T} ordering, whereas this is no longer visible for the subleading jet and photon separation. The leading photon and jet separation peaks at π, whereas the subleading photon and any jet separation is more evenly spread. Some of the dijet observables, shown in Figure 7.60, exhibit features that are unexpected, such as the $\Delta\phi_{jj}$ in the first few bins, and should be further investigated.

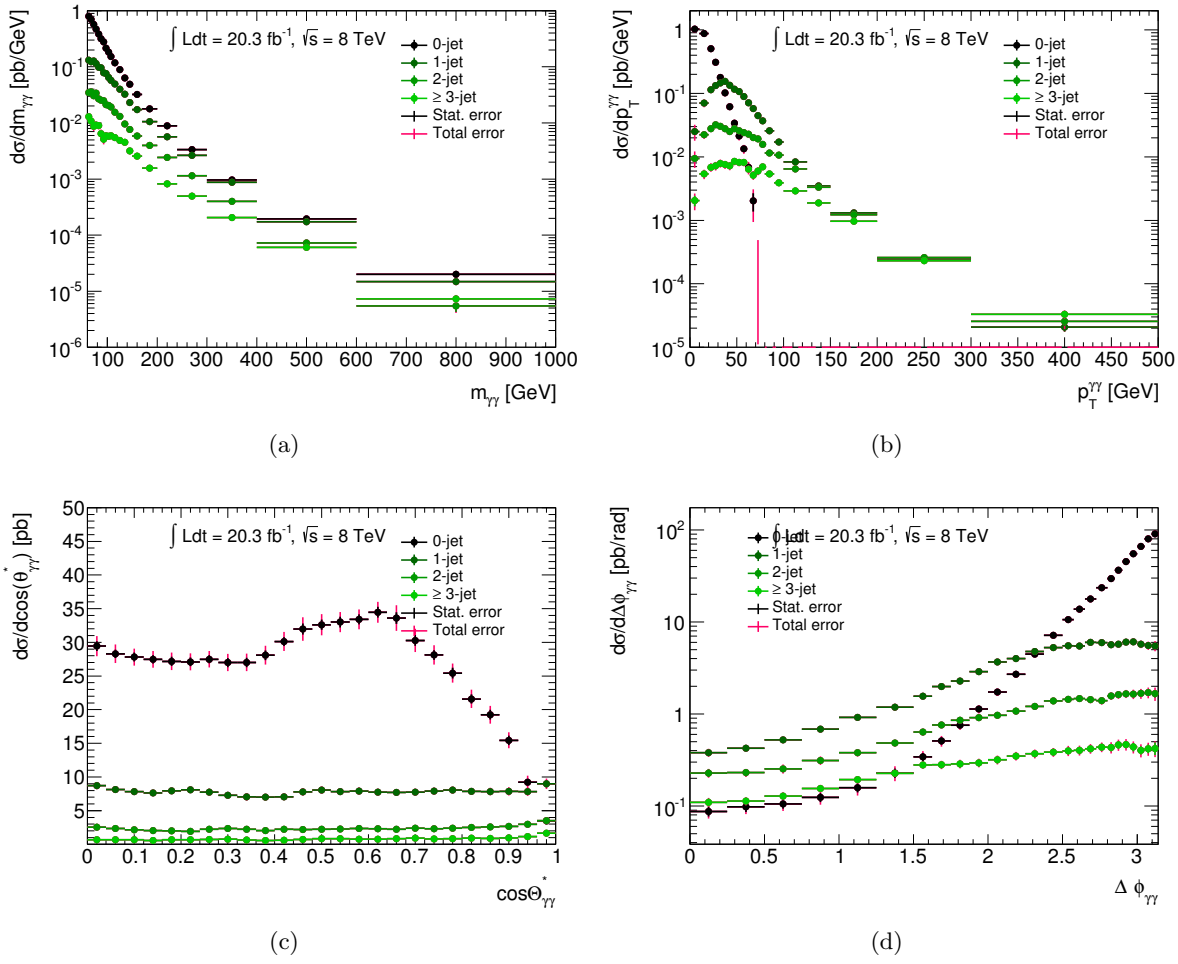


Figure 7.57: Unfolded spectra divided by the luminosity measured in data for the diphoton observables for the four jet multiplicity categories.

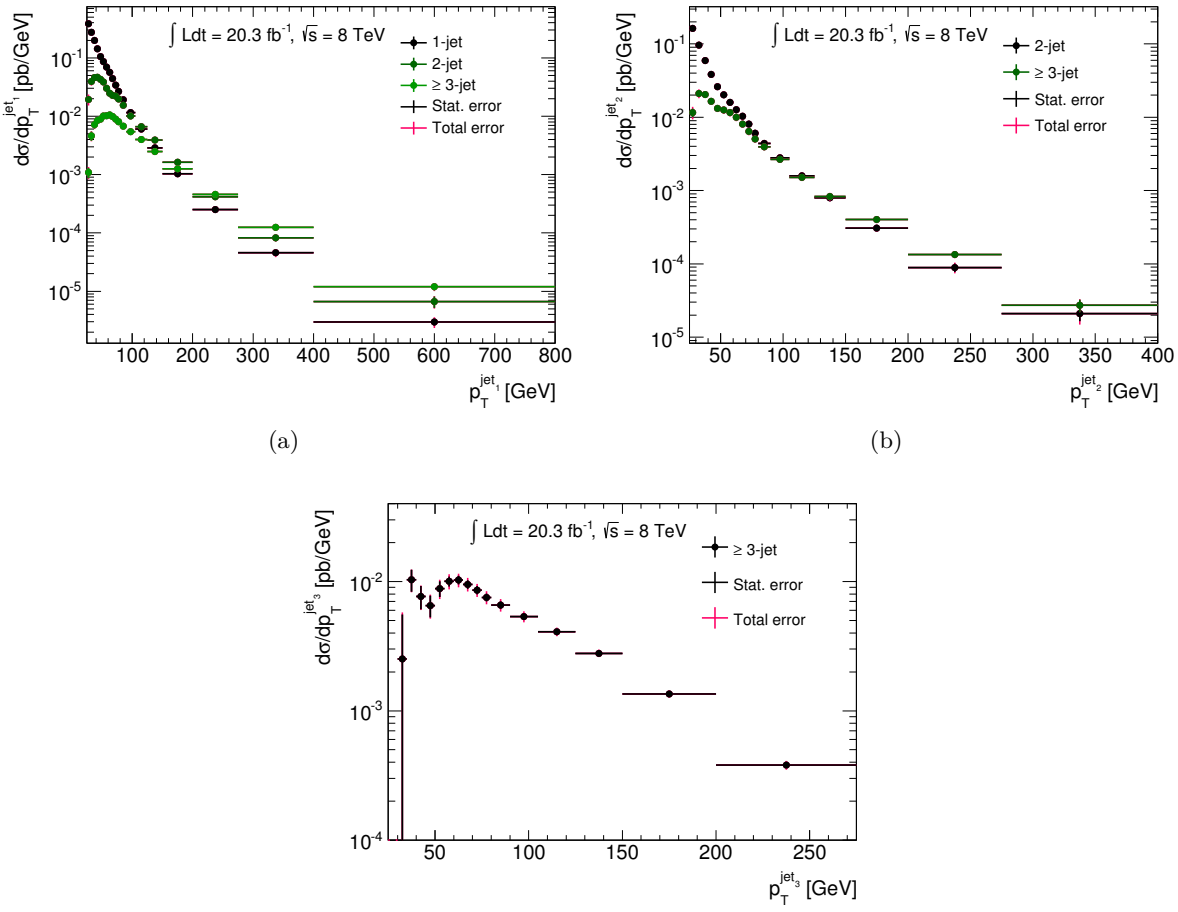


Figure 7.58: Unfolded spectra divided by the luminosity measured in data for the p_T^{jet} observables for the corresponding jet multiplicity categories.

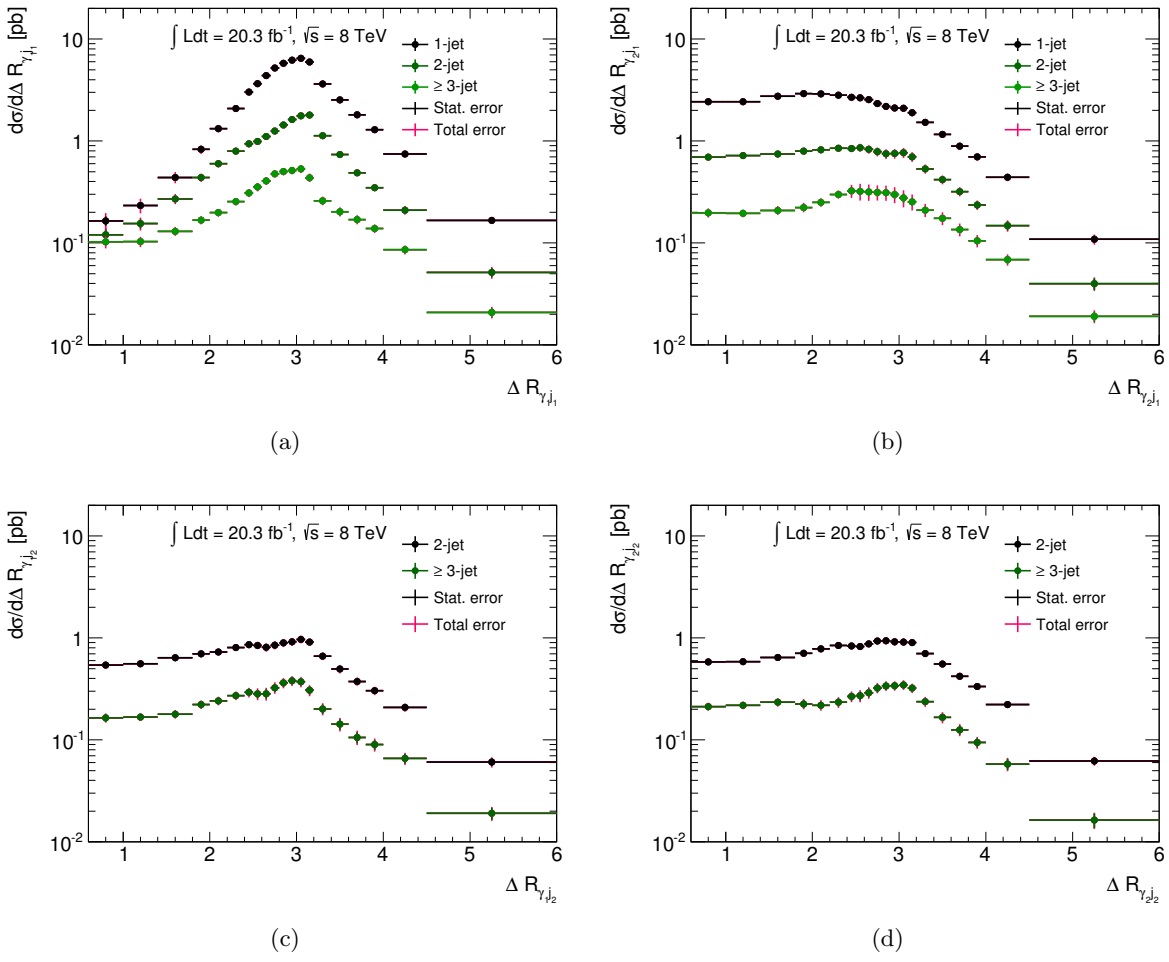
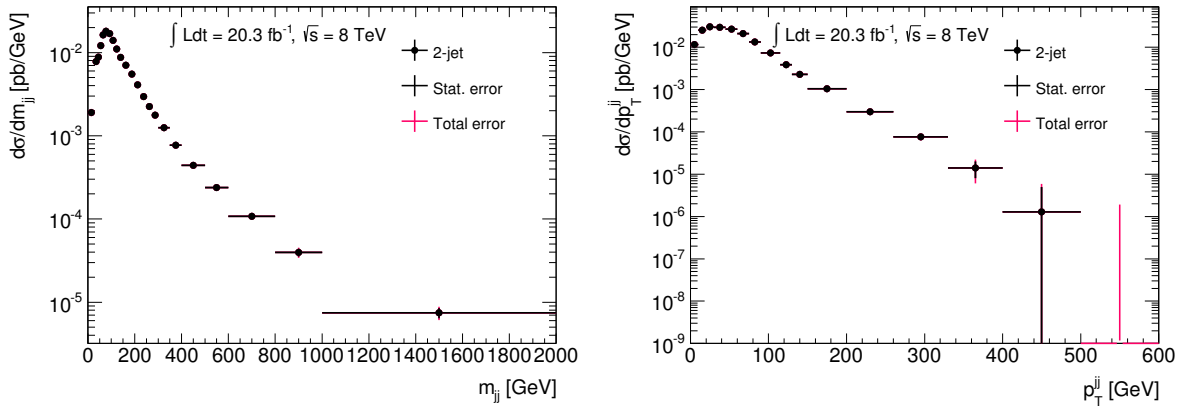
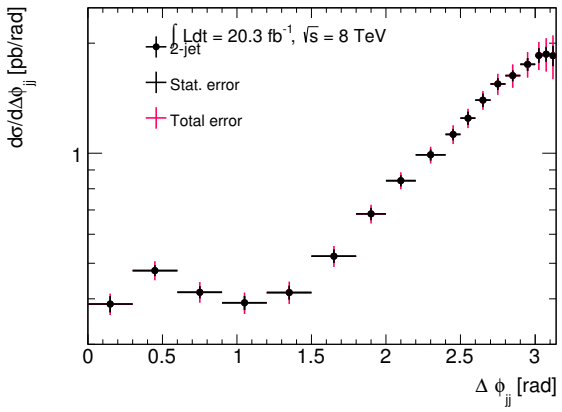


Figure 7.59: Unfolded spectra divided by the luminosity measured in data for the $\Delta R_{\gamma j}$ observables for the corresponding jet multiplicity categories.

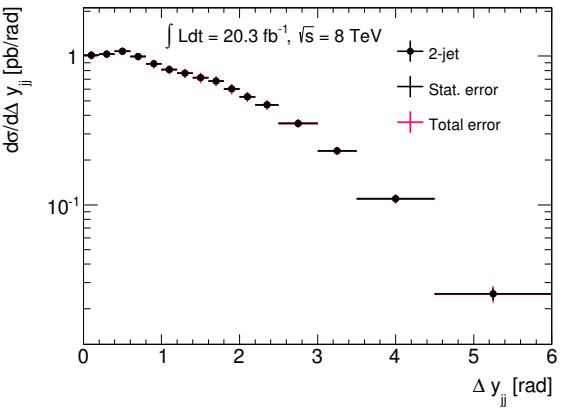


(a)

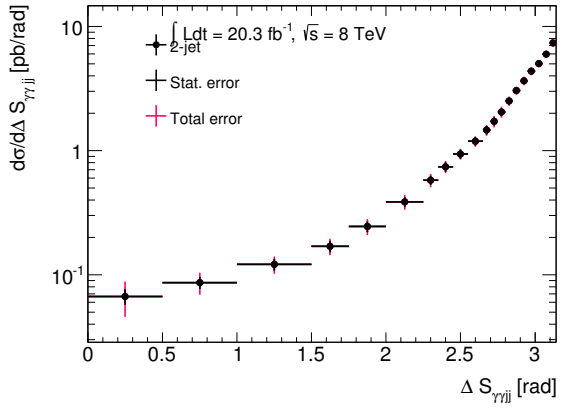
(b)



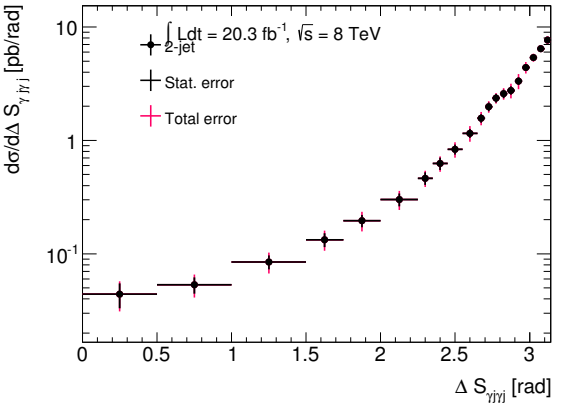
(c)



(d)



(e)



(f)

Figure 7.60: Unfolded spectra divided by the luminosity measured in data for the $\Delta R_{\gamma j}$ observables for the corresponding jet multiplicity categories.

7.7 Systematic uncertainties

The statistical uncertainties on the diphoton yield from the 2x2DSB method and the electron background subtraction method are combined in quadrature. This total statistical uncertainty is propagated through the unfolding procedure with help of toy studies. The systematic uncertainties contributing to the cross sections are described in the following sections. The final total systematic uncertainty up (down) is obtained by summing the positive (negative) variations in quadrature.

In addition to the systematic uncertainties described in the next sections, the already mentioned luminosity and absolute regularization bias systematic uncertainty are included in the total uncertainty. Statistical uncertainty of the MC is also included. The breakdown of the systematic uncertainties, the statistical and the total uncertainty bin-by-bin is shown in Figures 7.61-7.68.

In most cases, the dominant uncertainties come from the 2x2DSB method. The systematic uncertainty on the JVF selection is dominant for example for the $\cos\theta^*_{\gamma\gamma}$ observable in the $1-jet$ category, since the JVF selection may keep pile-up jets that would pass the analysis selection, making these events belong to the $1-jet$ category, where the shape of the $\cos\theta^*_{\gamma\gamma}$ distribution is very different w.r.t. the $0-jet$ category.

7.7.1 Background subtraction methods

The systematic uncertainties, described in Section 7.3.2, originating from the 2x2DSB method can be split into two categories. The systematic uncertainty associated to the *non-tight* ID control region definition (**L'**) must be considered as completely correlated between bins, due to the fact that the **L'** variations are defined in the same way across the events.

The other systematic uncertainties, namely the photon isolation shape, the binning of the input parameters, the photon *tight* ID scale factor variation and the determination of the data-driven isolation efficiencies (normalization window size), are much smaller than the **L'** uncertainties in almost all bins. The way they were defined, they are taken as uncorrelated between bins.

The systematic uncertainty on the diphoton yield originating in the $e \to \gamma$ and $\gamma \to e$ fake rate computation are propagated through the unfolding as completely correlated between bins, since their variation is defined consistently everywhere. The systematic uncertainties associated to the ee and $e\gamma$ yields are propagated through the unfolding separately and are treated as completely correlated between bins as well due to the fact that the change in calibration strategy variations (CC, UC, CU and UU combinations) are the same everywhere.

7.7.2 Photon energy scale and resolution

The migration matrices are re-derived for photon energy scale according to the ATLAS e/gamma group recommendations [249] and the correlations between the two photon energy scale uncertainties are treated according to the simplified model developed in Ref. [250]. This model includes 29 nuisance parameters (NP), summarized in Table 7.24. A full correlation model contains 56 nuisance parameters and decorrelates some of the material uncertainties and the E_1/E_2 scale uncertainty in 0.2 bins in η. The total systematic uncertainty due to photon energy scale is computed as a sum in quadrature of the positive and negative maxima of individual variations w.r.t. the nominal values. In this thesis, the photon energy scale systematics considered are

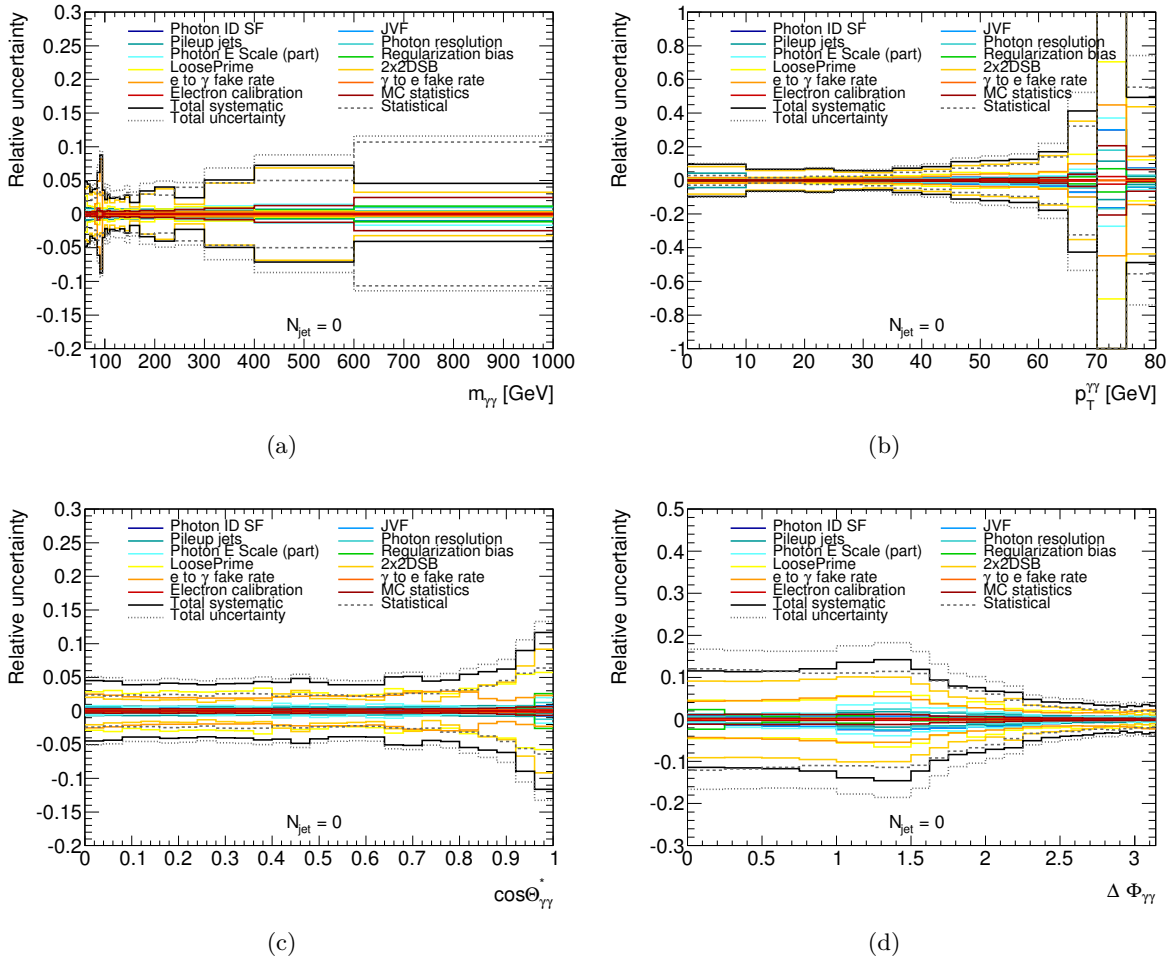


Figure 7.61: Breakdown of the relative systematic uncertainties on the measured yield, as a function of the diphoton observables for the $0-jet$ category.

only the fully correlated ones. The rest is in progress of being finalized within the group.

The systematic uncertainties on the photon energy resolution are also treated according to [249]. There are altogether seven variations performed both up and down:

- uncertainty associated to the measured Z boson peak width evaluated by smearing of the Z peak;
- uncertainty on the sampling term;
- uncertainty on the material budget in the ID, LAr calorimeter, the transition region between the barrel and end-cap and the cryogenic services;
- uncertainty associated to the modeling of pile-up.

The same sum in quadrature approach is taken for the final photon energy resolution uncertainties.

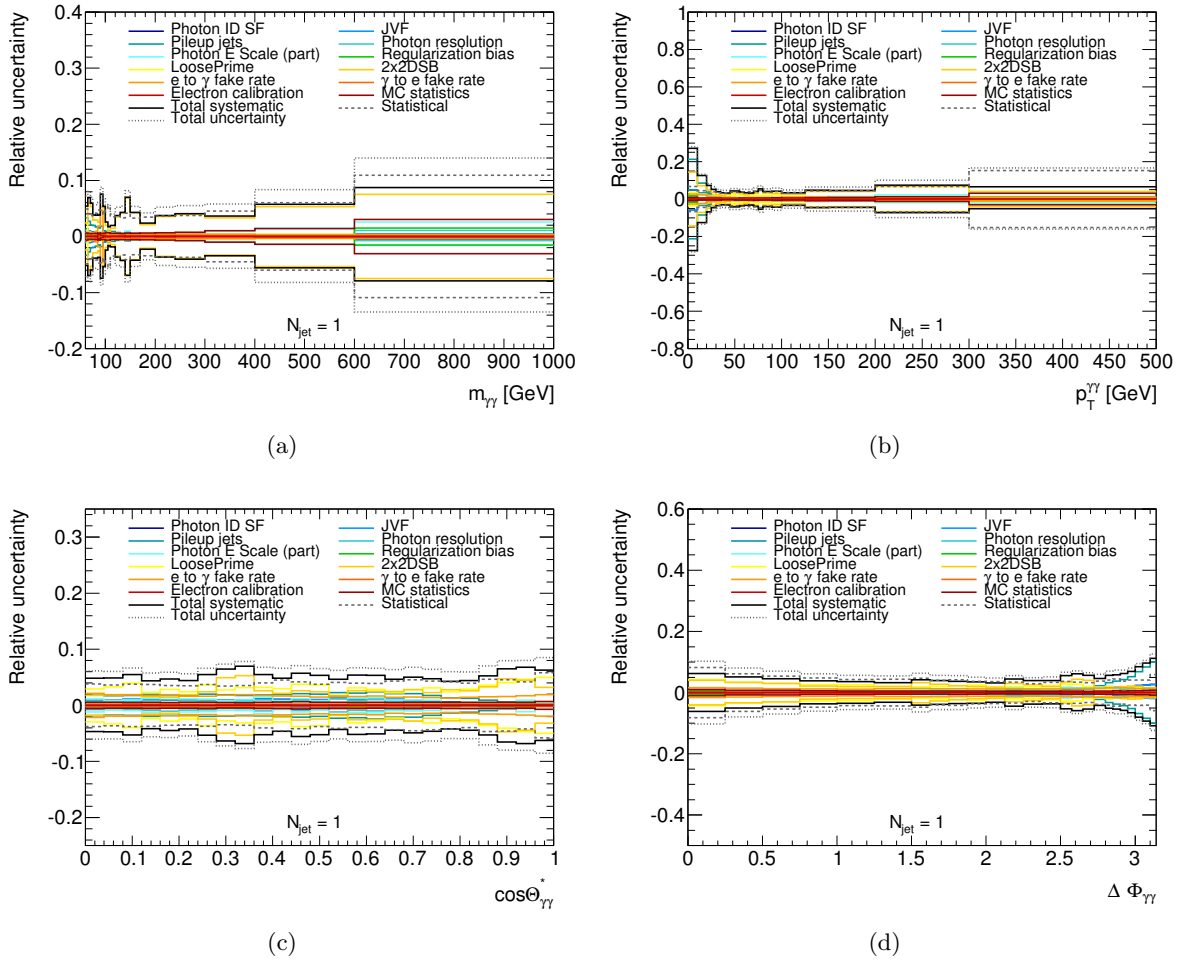


Figure 7.62: Breakdown of the relative systematic uncertainties on the measured yield, as a function of the diphoton observables for the $1-jet$ category.

7.7.3 Photon *tight* ID scale factors

When computing the *tight* ID of the photons, the shower shape variables in MC, as described in Section 3.3, are rescaled to match those in data. The photon ID efficiency scale factors are then applied to MC events to correct for the difference in the efficiency of the *tight* ID selection criterion as measured in data and as predicted by MC. Uncertainties on the scale factors are computed based on the combination of data-driven measurements using radiative Z tag-and-probe, electron extrapolation with $Z \to ee$ tag-and-probe and a matrix method [230].

When measuring two photons at the same time, correlations between the uncertainties have to be considered and a correlation model was developed by the e/gamma group [230]. In order to evaluate the impact of this uncertainty on the unfolding procedure, the response matrices are re-derived with the scale factors scaled up and down by their uncertainties, the unfolding procedure is then repeated and the largest positive and negative difference between the nominal and shifted result is taken as an uncertainty on the unfolded yields.

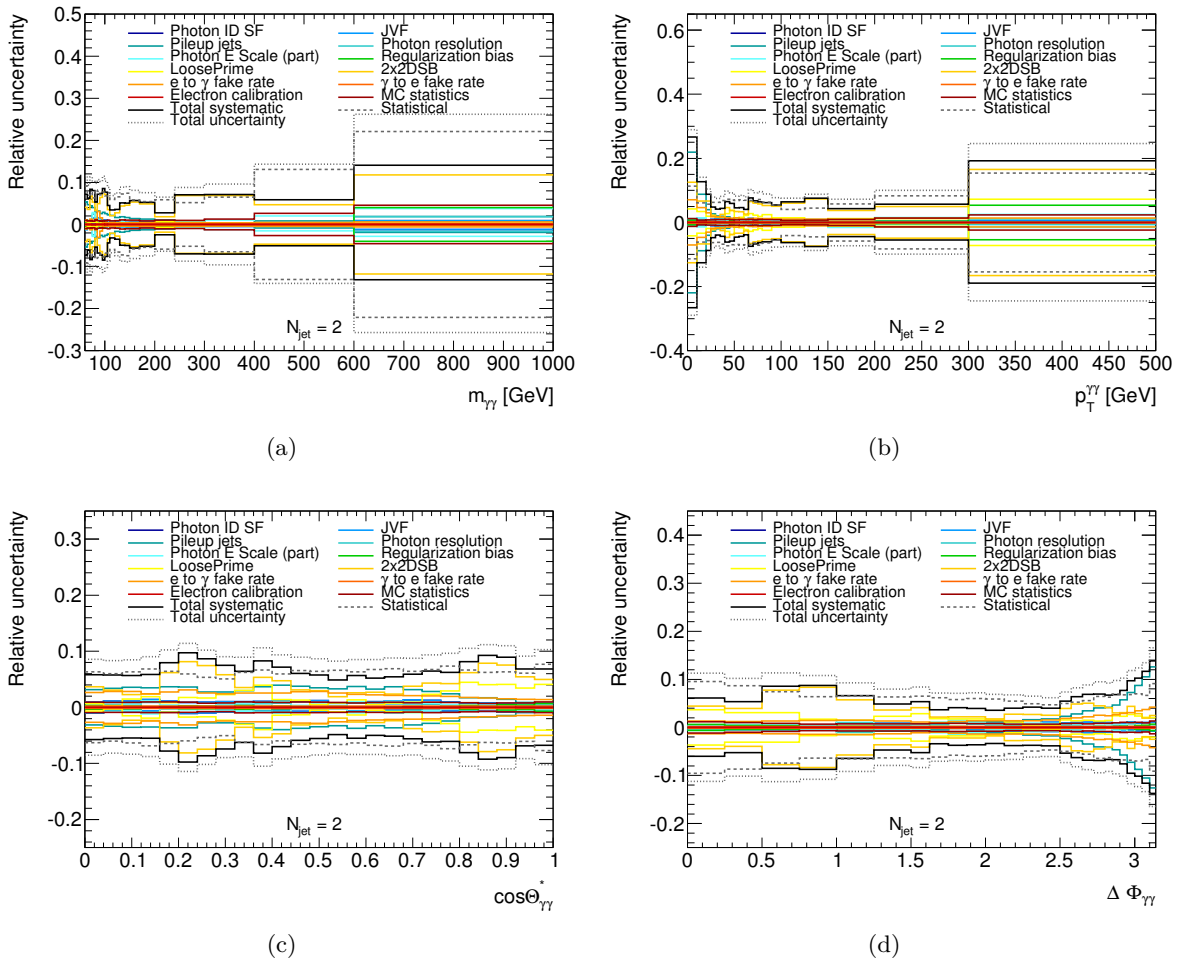


Figure 7.63: Breakdown of the relative systematic uncertainties on the measured yield, as a function of the diphoton observables for the $2-jet$ category.

7.7.4 Jet energy scale and resolution

To account for the uncertainties on the jet energy scale and resolution, the full model of 65 nuisance parameters was described in Section 4.4. A simplified model is used for this analysis, containing 14 nuisance parameters, namely [251]:

- six parameters from the reduction of the in-situ analyses nuisance parameters;
- two from η intercalibration, depending on the MC modelling and statistics;
- one from the behavior of high-p_{T} jets in propagation of single hadron uncertainties to the jet;
- one from the MC non-closure between the different MC samples used for calibration, with the main change being the material budget of the detectors;

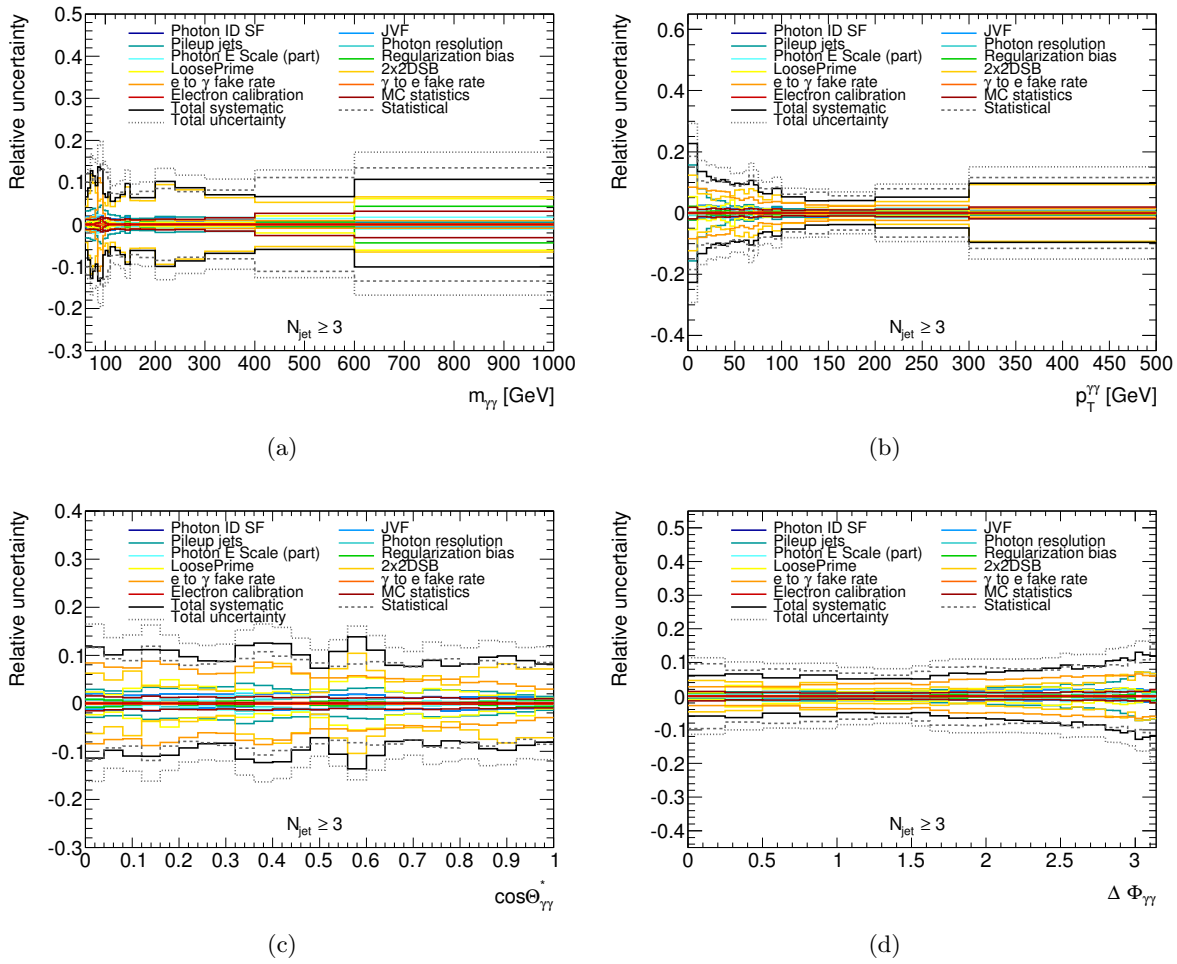


Figure 7.64: Breakdown of the relative systematic uncertainties on the measured yield, as a function of the diphoton observables for the $\geq 3-jet$ category.

- four parameters from pile-up, three of these are dependent on the number of primary vertices and μ (number of interactions per bunch crossing).

7.7.5 JVF signal efficiency uncertainty

The uncertainty on the MC description of the efficiency of the JVF selection cut at JVF>0.5 was derived by the jet-E_{T}^{miss} group within ATLAS and is applied according to [252], by re-deriving the response matrices again with the JVF uncertainty applied up and down, unfolding the distributions and computing the differences up and down w.r.t. the nominal values.

7.7.6 Jet definition and pile-up

The JVF selection applied on MC reconstructed jets, keeps about 10% of pile-up jets in the measured yield, as mentioned in Section 7.2.4, meaning out of the jets that are measured as

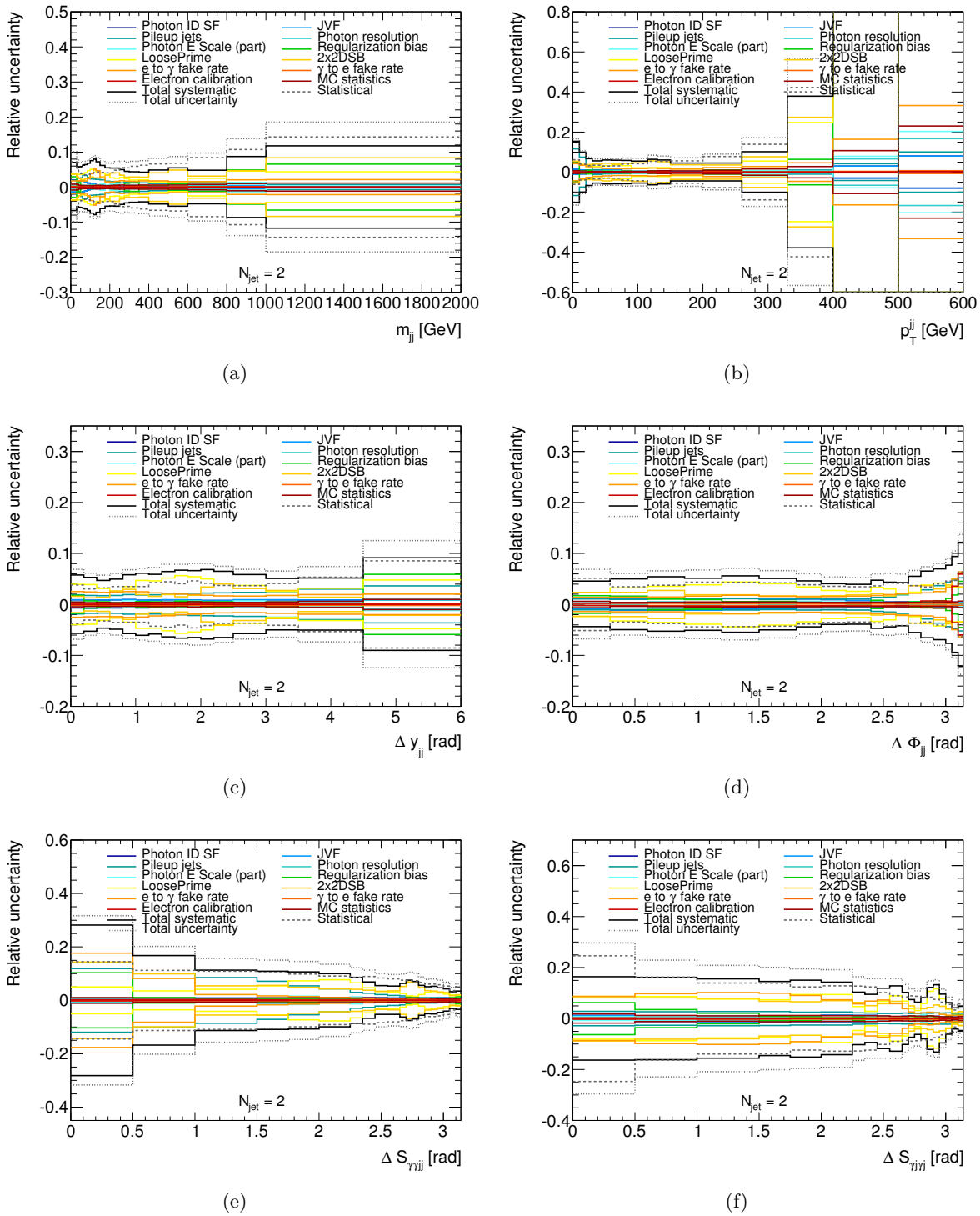


Figure 7.65: Breakdown of the relative systematic uncertainties on the measured yield, as a function of the dijet and the ΔS observables for the $2-jet$ category.

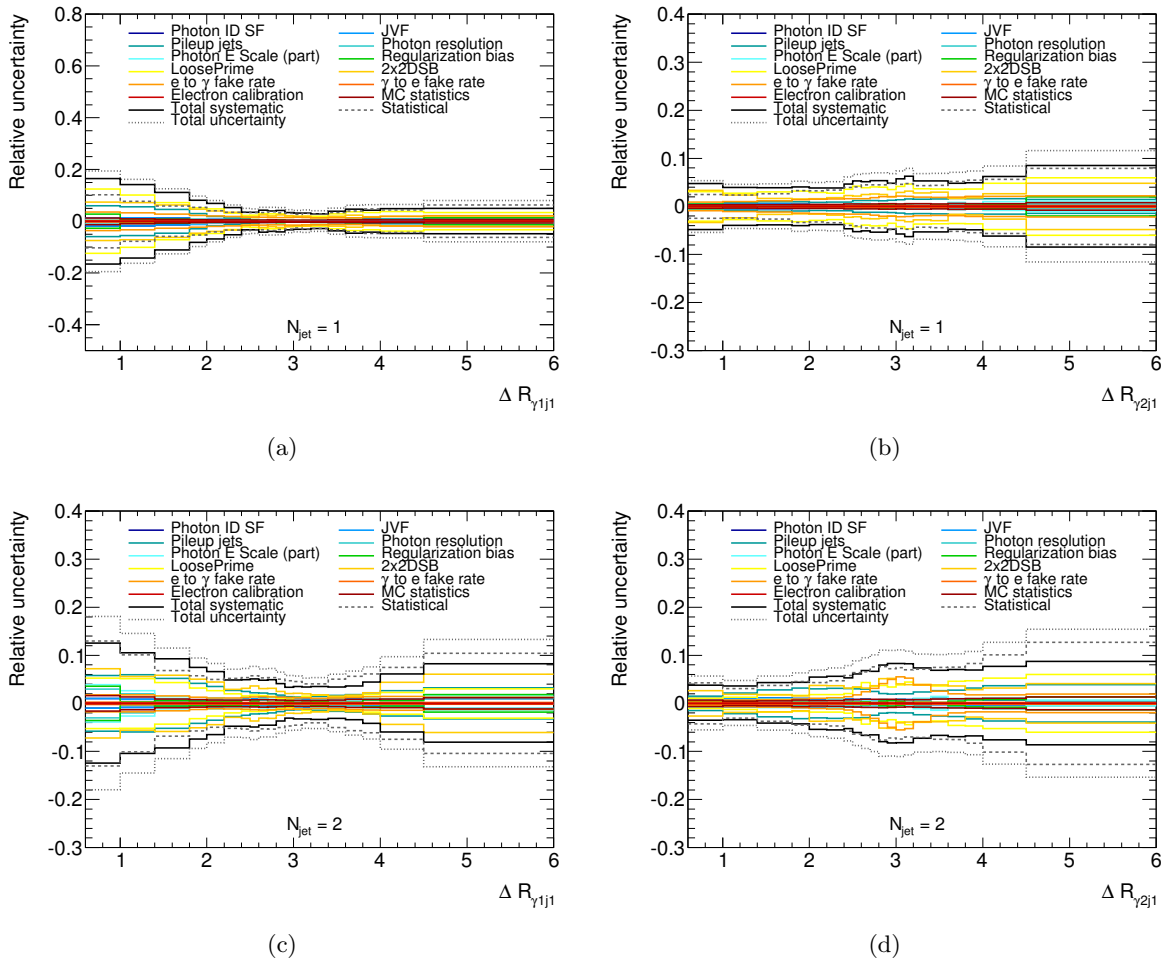


Figure 7.66: Breakdown of the relative systematic uncertainties on the measured yield, as a function of the R-separation between the leading photon and the leading jet (left) and the subleading photon and the leading jet (right) for the $1-jet$ (top) and $2-jet$ (bottom) category.

the leading, subleading or third jet, 10% could originate from pile-up instead of hard-scattering. An uncertainty on how well the MC describes the fraction of residual pile-up jets in data could be conservatively assigned by removing the pile-up jets completely from the MC reconstructed selection and repeating the unfolding procedure with new response matrices.

This, however may be too conservative, and an approach adopted previously in the Higgs cross section measurement [253, 248] will be taken instead. There, 35% of pile-up jets are randomly removed from the selection in MC at the reconstruction level and the unfolding is repeated. The 35% is determined by looking at the double ratio of the number of pile-up jets w.r.t. hard scatter jets not passing the JVF cut of 0.1 and 0.5 in MC. This uncertainty is estimated by filling the response matrices after the random pile-up jet removal, performing the unfolding procedure and comparing the differential spectra to the nominal values.

Source of systematic uncertainty	Simplified model	Number of NP
Method	fully correlated	1
LAr HG/MG miscalibration	fully correlated	1
L1 gain	fully correlated	1
E_1/E_2 scale	decorrelated barrel/end-cap	2
LAr E_1/E_2 calibration	decorrelated barrel/end-cap	2
PS scale	decorrelated barrel/end-cap	2
LAr E_1/E_2 modelling for e/ γ_{unconv}	decorrelated barrel/end-cap	2
Material in inner detector	decorrelated in four bins	4
Material in cryostat	decorrelated barrel/end-cap	2
Material in calorimeter	decorrelated barrel/end-cap	2
LAr E_1/E_2 modelling for γ_{unconv}	decorrelated barrel/end-cap	2
LAr E_1/E_2 modelling for e	fully correlated in end-cap	1
Lateral leakage γ_{unconv}	fully correlated	1
Lateral leakage γ_{conv}	fully correlated	1
Conversion inefficiency	fully correlated	1
Conversion fake rate	fully correlated	1
Conversion radius	fully correlated	1
Pedestal	fully correlated	1
GEANT4 simulation	fully correlated	1
Total number of nuisance parameters		29

Table 7.24: Simplified correlation model of the photon energy scale uncertainties for two photons.

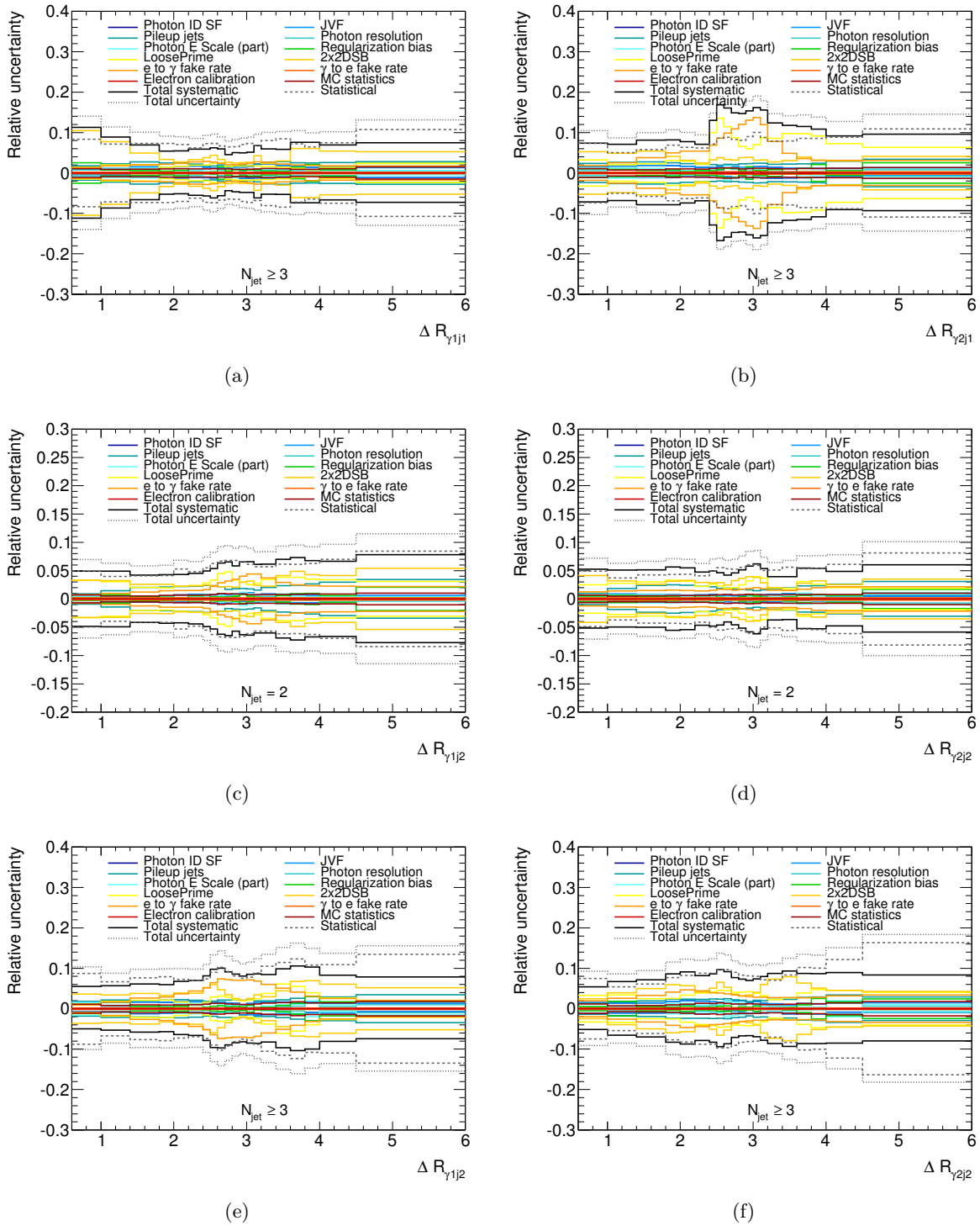


Figure 7.67: Breakdown of the relative systematic uncertainties on the measured yield, as a function of the R-separation for the leading jet in the $1-jet$ (top), $2-jet$ (middle) and $\geq 3-jet$ (bottom) category.

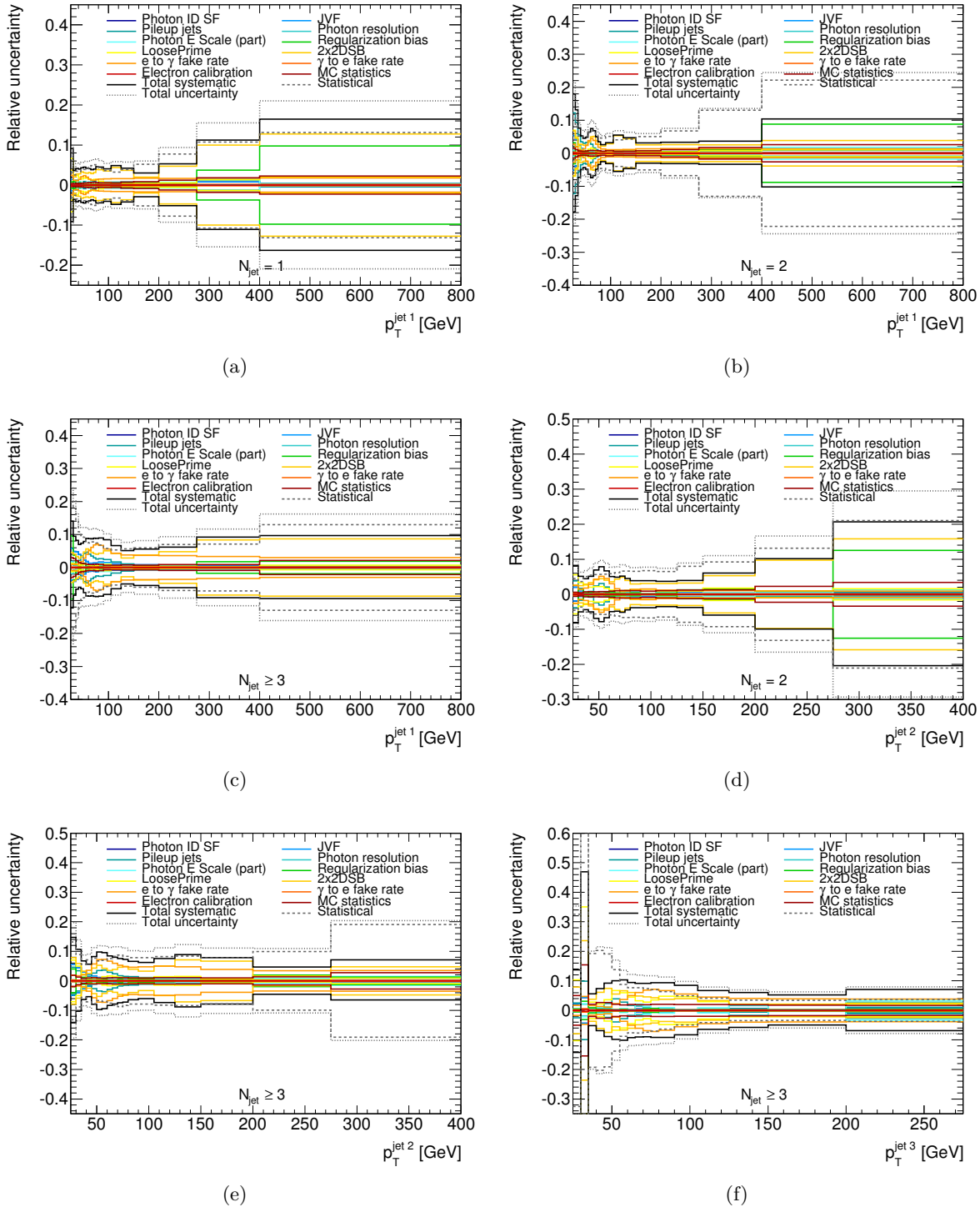


Figure 7.68: Breakdown of the relative systematic uncertainties on the measured yield, as a function of the leading jet $p_{\mathrm{T}}^{jet_1}$. (a) $1-jet$, (b),(d) $2-jet$, (c),(e),(f) $\geq 3-jet$ category.

7.8 Theoretical predictions

7.8.1 NLO calculations

Unfortunately, due to some issues with selection criteria applied at generation level in the generated NTUPLES from the group in Refs. [80, 84], the GoSam NLO predictions cannot be shown since they would not be meaningful. This work is still ongoing within the analysis group.

7.8.2 Sherpa and Pythia MC

Pythia and Sherpa are LO MC generators with an underlying event model and parton shower, therefore they provide a full list of particles in an event instead of just the partons coming from the hard process. Sherpa, in addition, includes higher order real-emission matrix elements.

The measured cross sections are compared to the Sherpa MC, which have been simulated for diphoton events in association with up to three jets in the final state, coming from the hard process. Pythia MC samples with up to two jets in the final state are also used for comparison. The result is shown in the next section. The samples used were already introduced in Table 7.2. The full analysis selection at truth level, as summarized in Table 7.7 has been applied to both Pythia and Sherpa and the cross sections differential in all measured observables are computed.

In general, Pythia and Sherpa are expected to describe well the low part of $p_{\mathrm{T}}^{\gamma\gamma}$ and high $\Delta\phi_{\gamma\gamma}$, where soft gluon emission is of importance. Overall, they are expected to underestimate the cross sections, therefore in the comparisons, they are scaled to match the integral of the measured distributions in data.

7.9 Comparison of measured cross sections to theoretical calculations

The unfolded cross sections are compared to Sherpa MC events, by rescaling them to the integral of the distribution measured in data. This way, one can compare the shape predicted by MC versus the unfolded shape in data. The comparisons are shown in Figures 7.69-7.79. Sherpa MC events are shown with statistical uncertainties only, whereas the data is shown with the statistical and total uncertainties. Ratio plots between the data and MC are also shown (bottom of each plot), where the errors are the total uncertainty on the measured cross section normalized to Sherpa. Statistical uncertainty of the MC is shown as a band around one.

The $m_{\gamma\gamma}$ observable shows the same behavior across the jet multiplicity categories, where at low $m_{\gamma\gamma}$, the agreement is relatively good, the difference rises up to 100% in the $1-jet$ category in high $m_{\gamma\gamma}$.

The $p_{\mathrm{T}}^{\gamma\gamma}$ predicted by Sherpa is overestimating the distribution in the high tail for all jet multiplicity categories. The largest discrepancy is found in the $0-jet$ category, where Sherpa has much harder spectrum, whereas in data, the cross section is depleted above 80 GeV. In the other jet multiplicity categories, the agreement is good, with high-$p_{\mathrm{T}}^{\gamma\gamma}$ bins showing a disagreement of up to 25%.

The agreement between data and MC for the $\cos\theta^{*}_{\gamma\gamma}$ is very good in all jet multiplicity categories.

$\Delta\phi_{\gamma\gamma}$ is described very well too, except for the low end of the distribution in the $0-jet$ category, where the difference rises up to 275% where the fact that SHERPA is missing the NLO and NNLO contributions is visible.

The m_{jj} and p_{T}^{jj} observables are not described well, the discrepancy ranges from a few percent to a MC overestimation by up to 100% in the high tails of the spectra. SHERPA predicts a harder spectrum for the dijets produced in association with diphotons. The diphoton p_{T} spectrum in the $2-jet$ category exhibits the same behavior. But comparing the diphoton and the dijet invariant mass shape between data and SHERPA, m_{jj} is overestimated, whereas $m_{\gamma\gamma}$ is underestimated in the high-mass tail.

The angular properties of the dijet system in the $2-jet$ category are described very well, with the ratio between data and MC stays close to one throughout the whole spectrum. The ΔS observables are systematically overestimated by the MC in the low-end of the spectrum, suggesting that the Multiple Parton Interaction (MPI) model in SHERPA overestimates the MPI rate.

The description of the $\gamma - j$ separation is exceptionally good. The p_{T}^{jet} is overestimated by SHERPA for each category and jet, except for the p_{T} of the third jet, which is completely different in data and MC, where SHERPA predicts a much softer spectrum than what is observed.

It is also interesting to compare the $p_{\mathrm{T}}^{\gamma\gamma}$ distribution to the $p_{\mathrm{T}}^{jet_1}$ distribution in the $1-jet$ category, since the leading (and only) jet in this case should balance the dijet system. SHERPA overestimates both distributions in the high tails.

The measured cross sections are also compared to Pythia MC simulation with up to two hard scatter jets in the final state, in Figures 7.80-7.90. One can therefore assume that the comparison with the $\geq 3-jet$ category is not meaningful, which is visible from the description of the p_{T} of the third jet, in Figure 7.90(c).

In general, PYTHIA MC predictions also have to be rescaled to the integral of data, where the overall cross section PYTHIA predicts is smaller than was seen in SHERPA and data. Overall, SHERPA is better at describing the shape of most observables.

7.10 Conclusion

The measurement of differential cross sections of diphotons in association with jets is presented, measured on data collected by the ATLAS experiment at $\sqrt{s} = 8$ TeV, corresponding to the total integrated luminosity of $20.3 \pm 2.8\%$ fb^{-1}.

First, the event and object selection is investigated, determining the fiducial phase space for the measurement. The analysis deals with subtraction of background coming from decays of neutral hadrons inside jets, which can mimic photons. The 2x2DSB method, a counting method, is used to compute diphoton yields, which still contain events with one or both photons actually being an electron. The electron background is subtracted by measuring the $e \to \gamma$ and $\gamma \to e$ fake rates and estimating the bin-by-bin impurities to be subtracted from the 2x2DSB yields. The contribution of the Higgs boson decaying into diphotons is evaluated but neglected w.r.t. to the size of the QCD-induced diphoton yields. SVD unfolding is used to unfold the measured distributions, based on SHERPA MC events generated as diphotons plus up to three hard-scatter jets in the final state. The obtained cross sections are compared to SHERPA MC predictions, normalized to the total cross sections. They are also compared to PYTHIA MC, but this sample

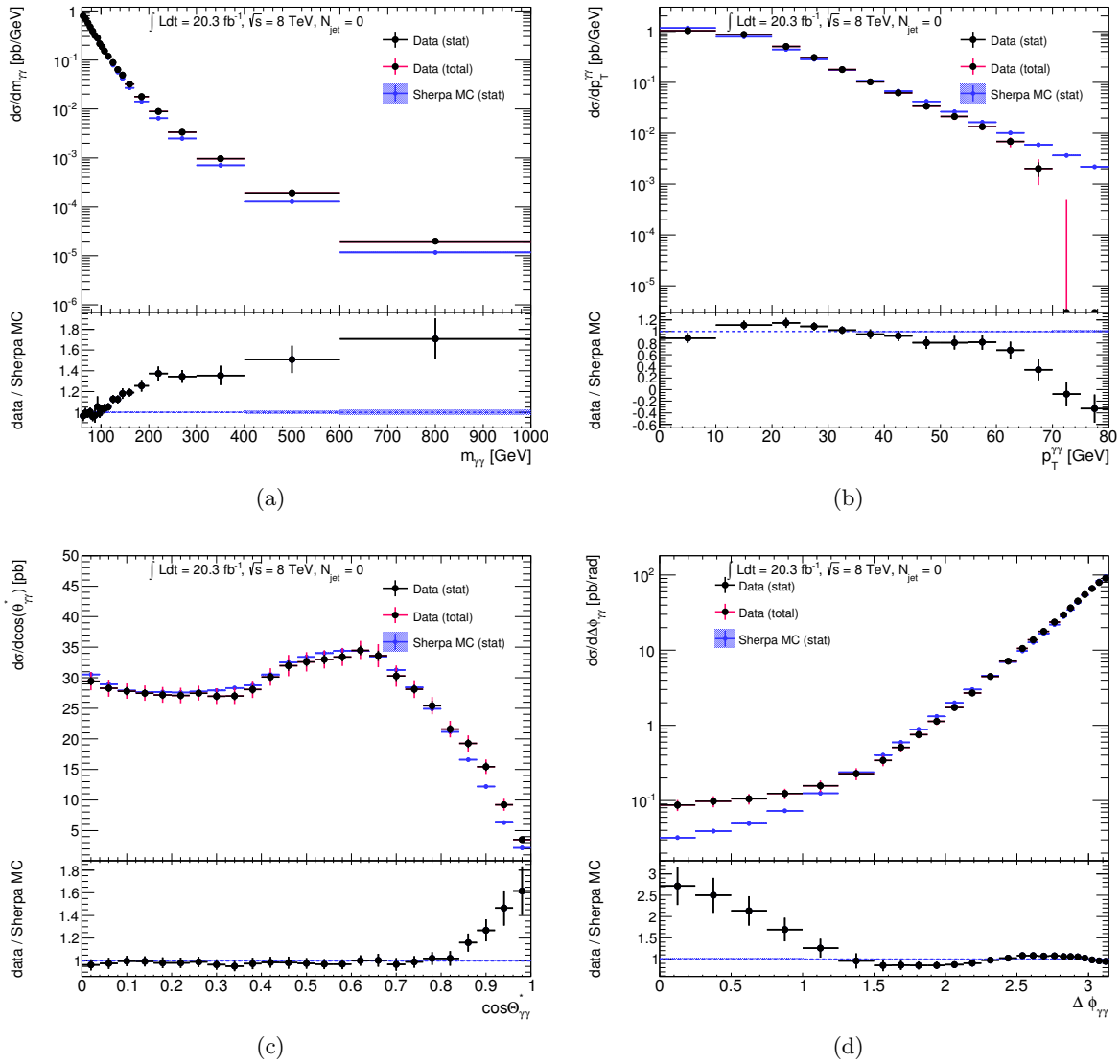


Figure 7.69: Comparison between the measured cross sections in data and SHERPA MC, as a function of the diphoton observables for the $0-jet$ category.

only has up to two jets from the hard process in the final state.

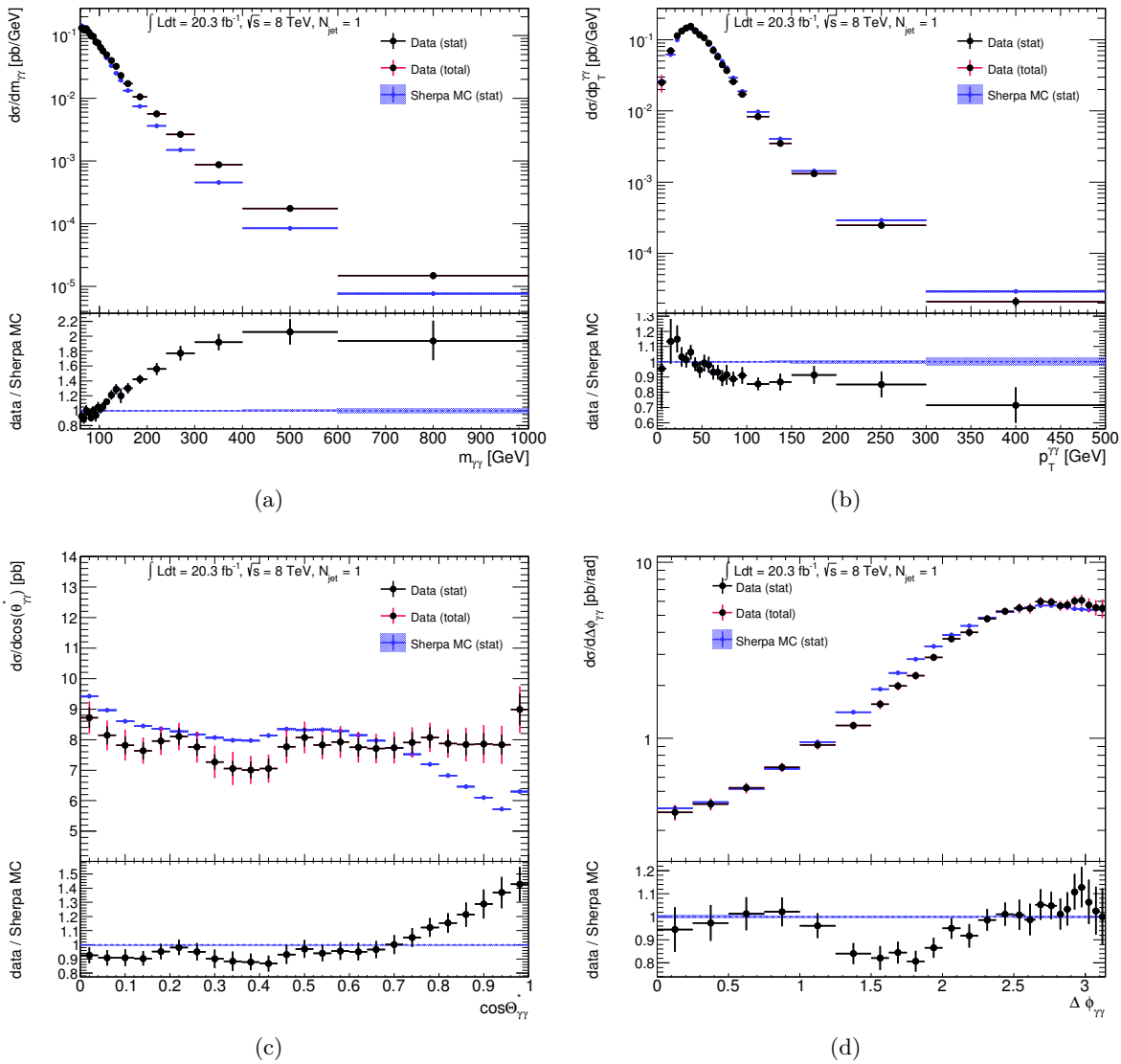


Figure 7.70: Comparison between the measured cross sections in data and SHERPA MC, as a function of the diphoton observables for the $1-jet$ category.

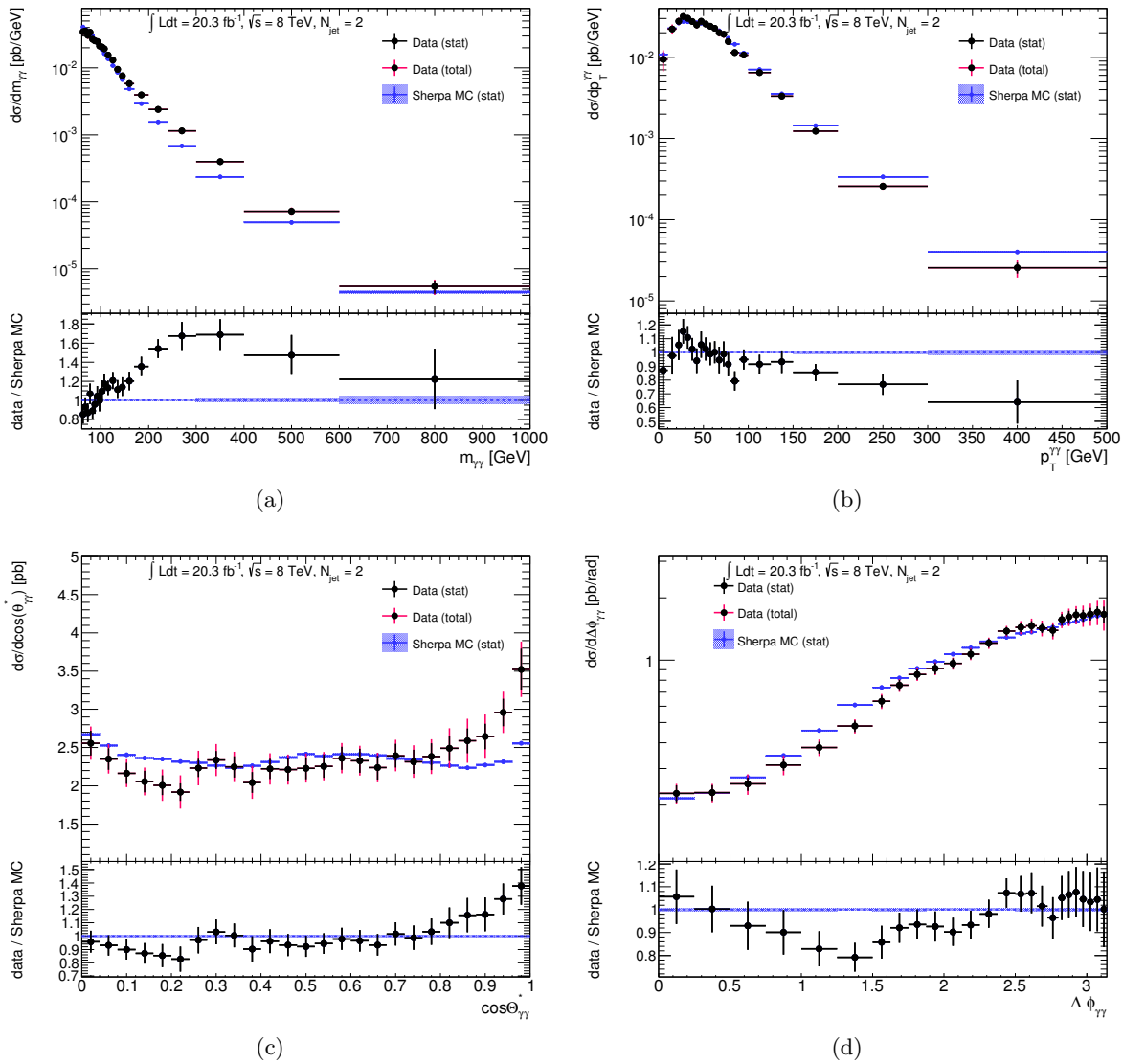


Figure 7.71: Comparison between the measured cross sections in data and SHERPA MC, as a function of the diphoton observables for the $2-jet$ category.

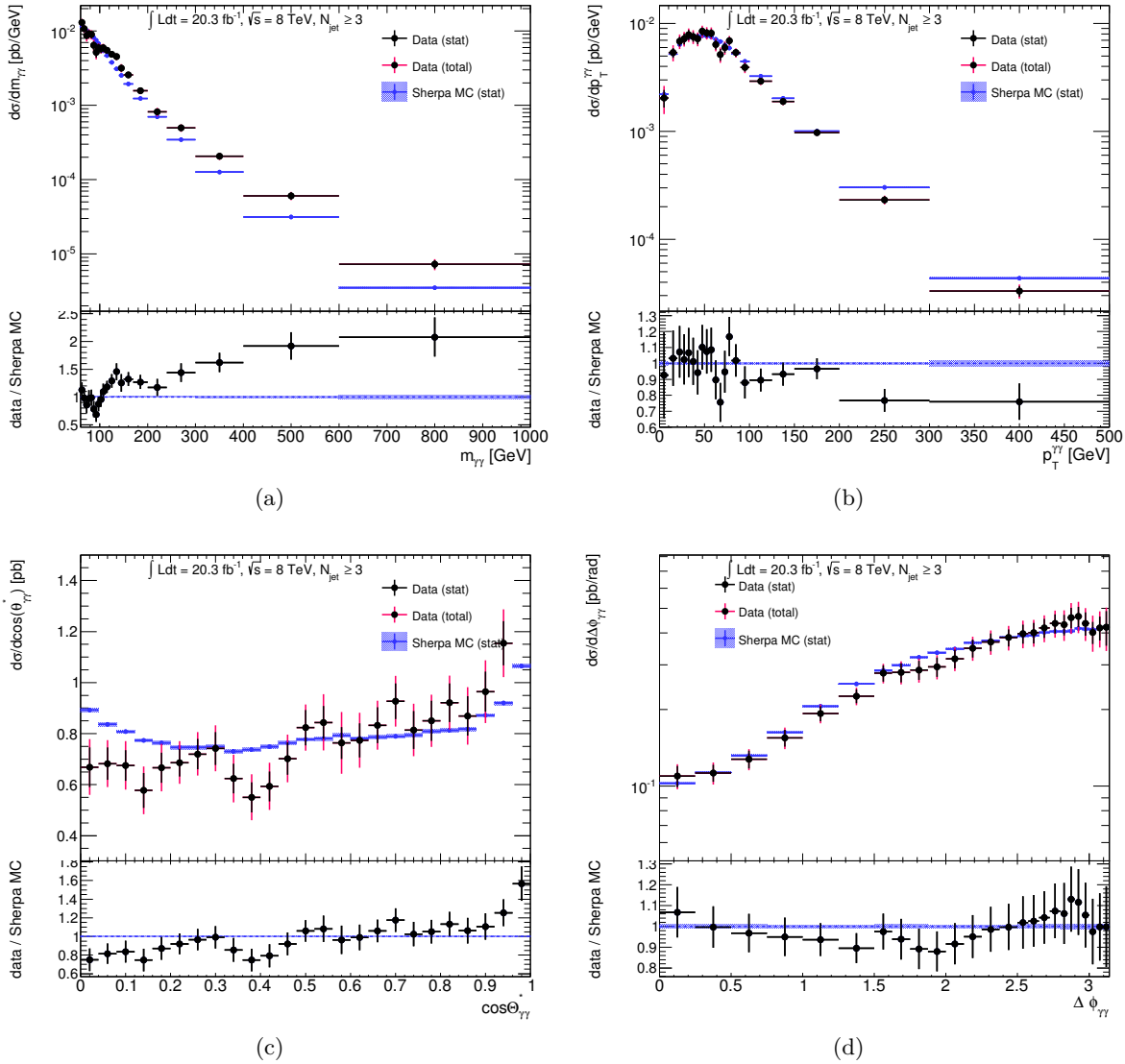


Figure 7.72: Comparison between the measured cross sections in data and SHERPA MC, as a function of the diphoton observables for the $\geq 3 - jet$ category.

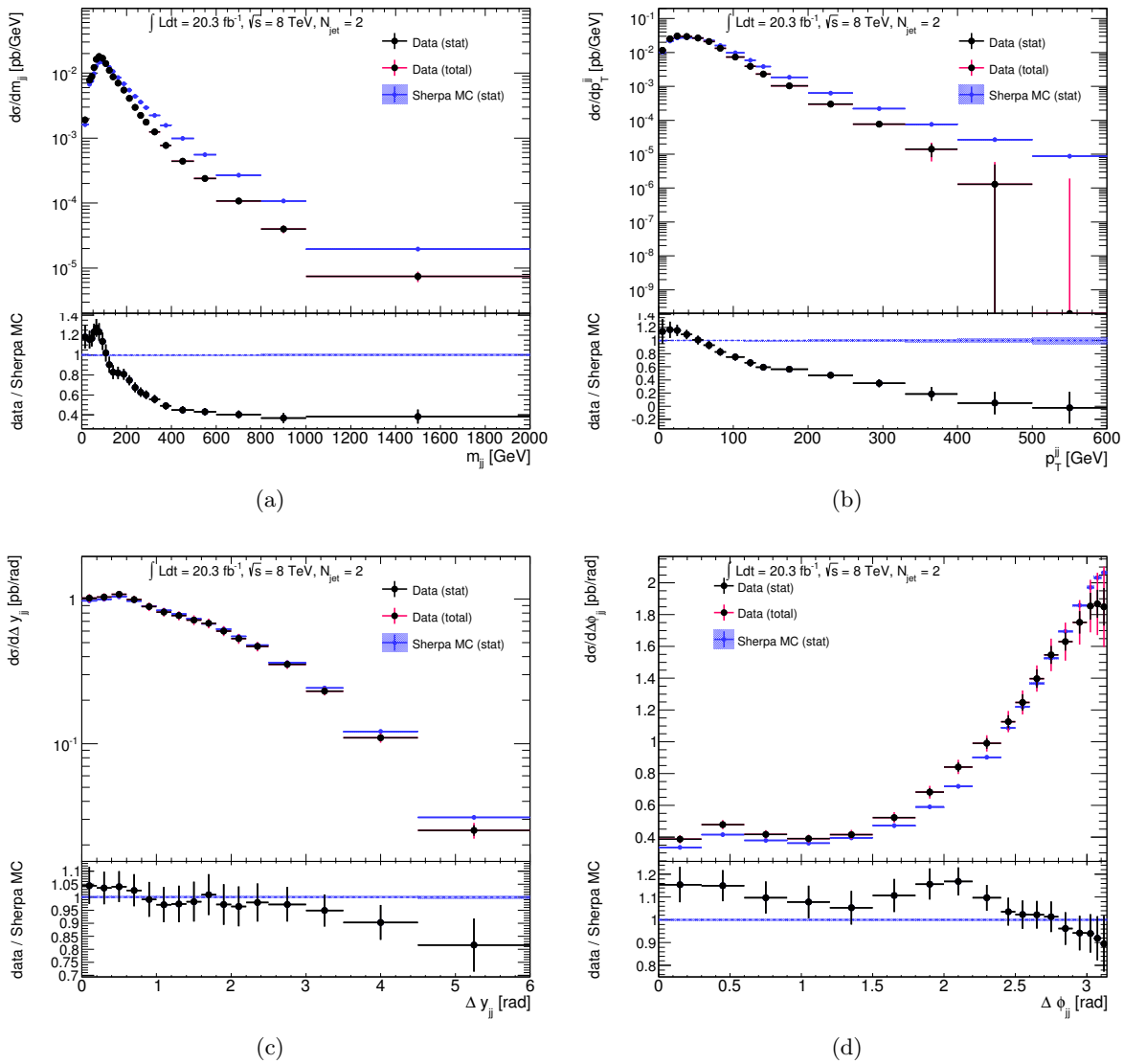
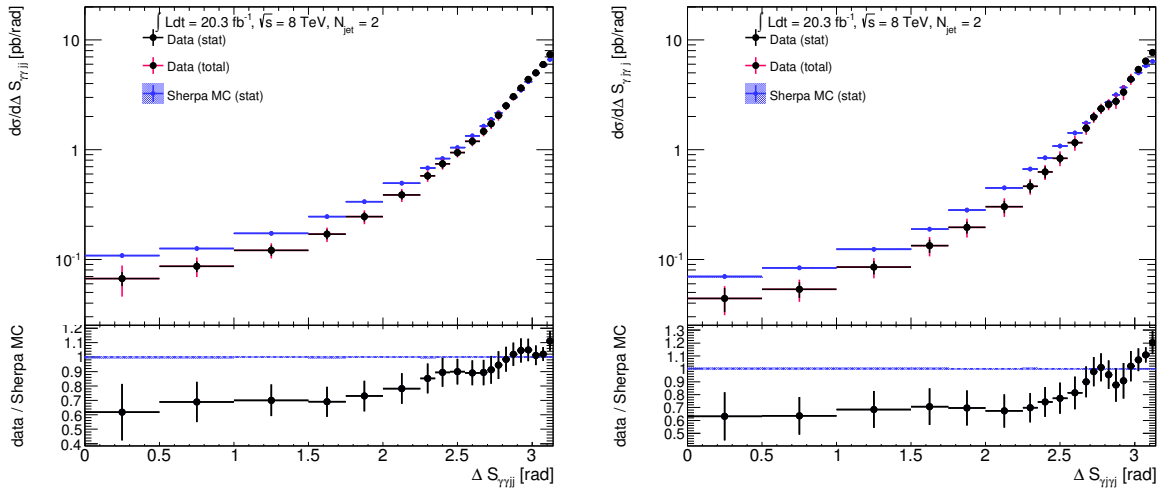


Figure 7.73: Comparison between the measured cross sections in data and SHERPA MC, as a function of the dijet observables for the $2-jet$ category.



(a) (b)

Figure 7.74: Comparison between the measured cross sections in data and SHERPA MC, as a function of the ΔS observables for the $2-jet$ category.

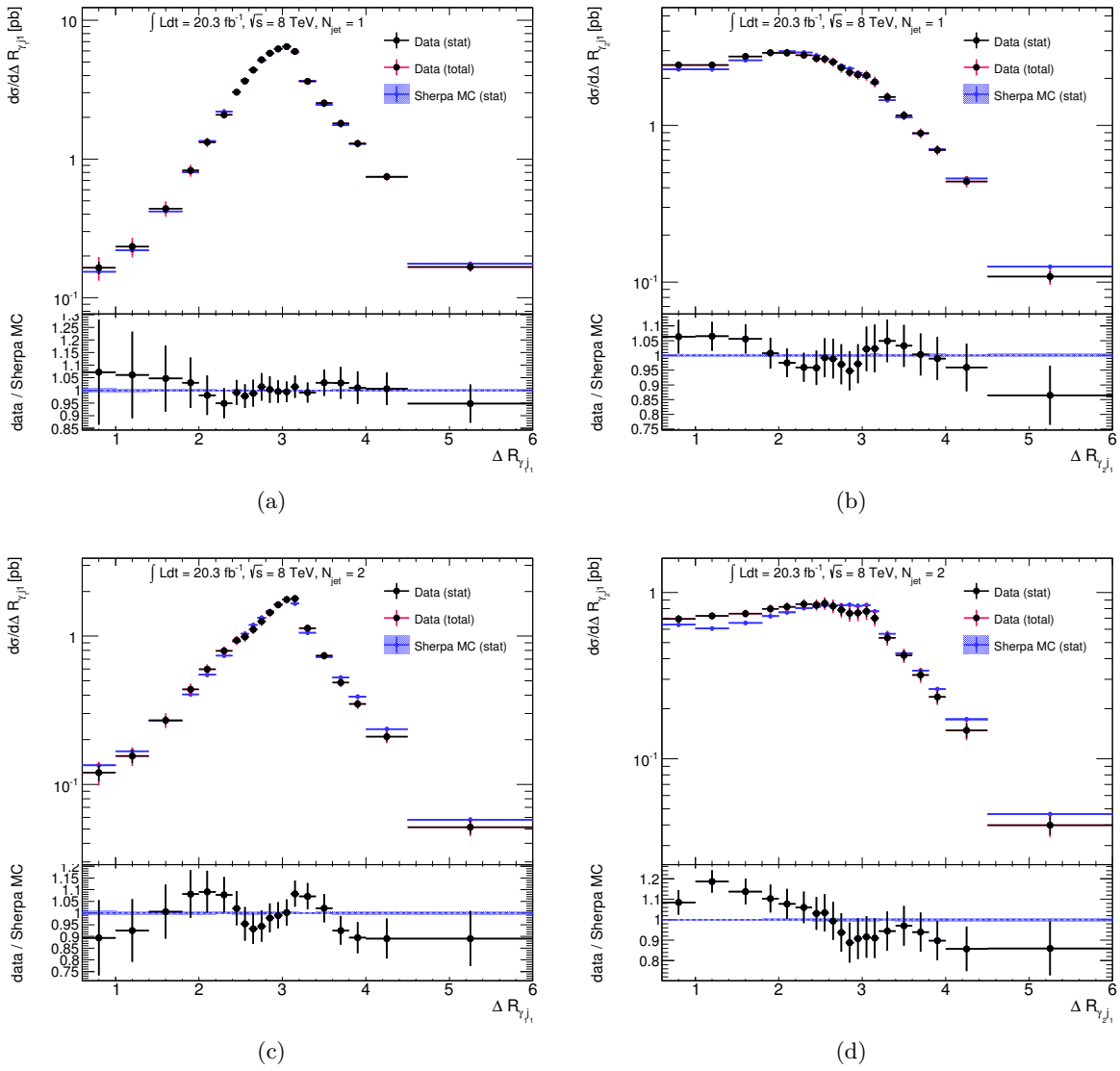
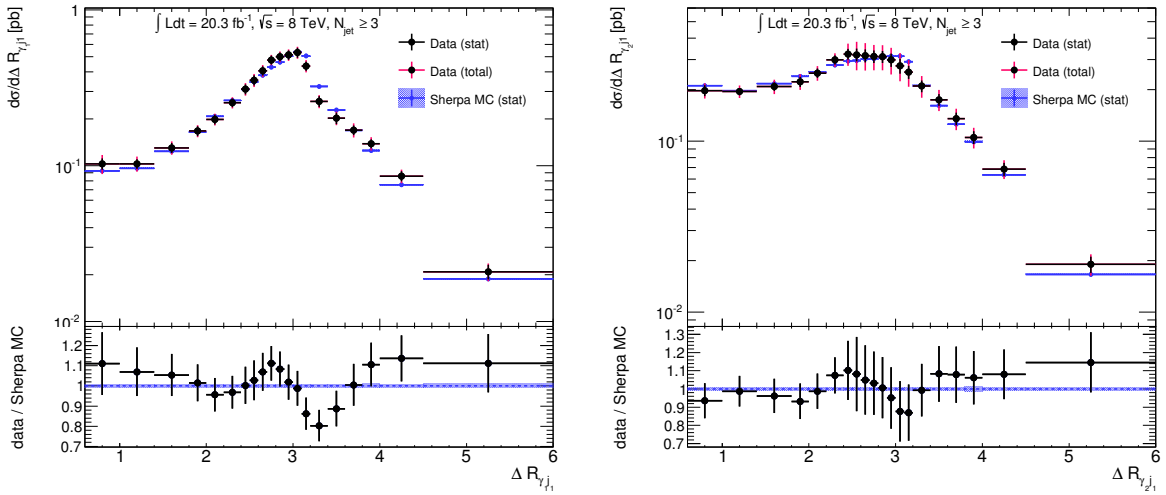


Figure 7.75: Comparison between the measured cross sections in data and SHERPA MC, as a function of the R-separation between the leading photon and the leading jet (left) and the subleading photon and the leading jet (right) for the $1-jet$ (top) and $2-jet$ (bottom) category.



(a) (b)

Figure 7.76: Comparison between the measured cross sections in data and SHERPA MC, as a function of the R-separation between the leading photon and the leading jet (left) and the subleading photon and the leading jet (right) for the $\geq 3 - jet$ (bottom) category.

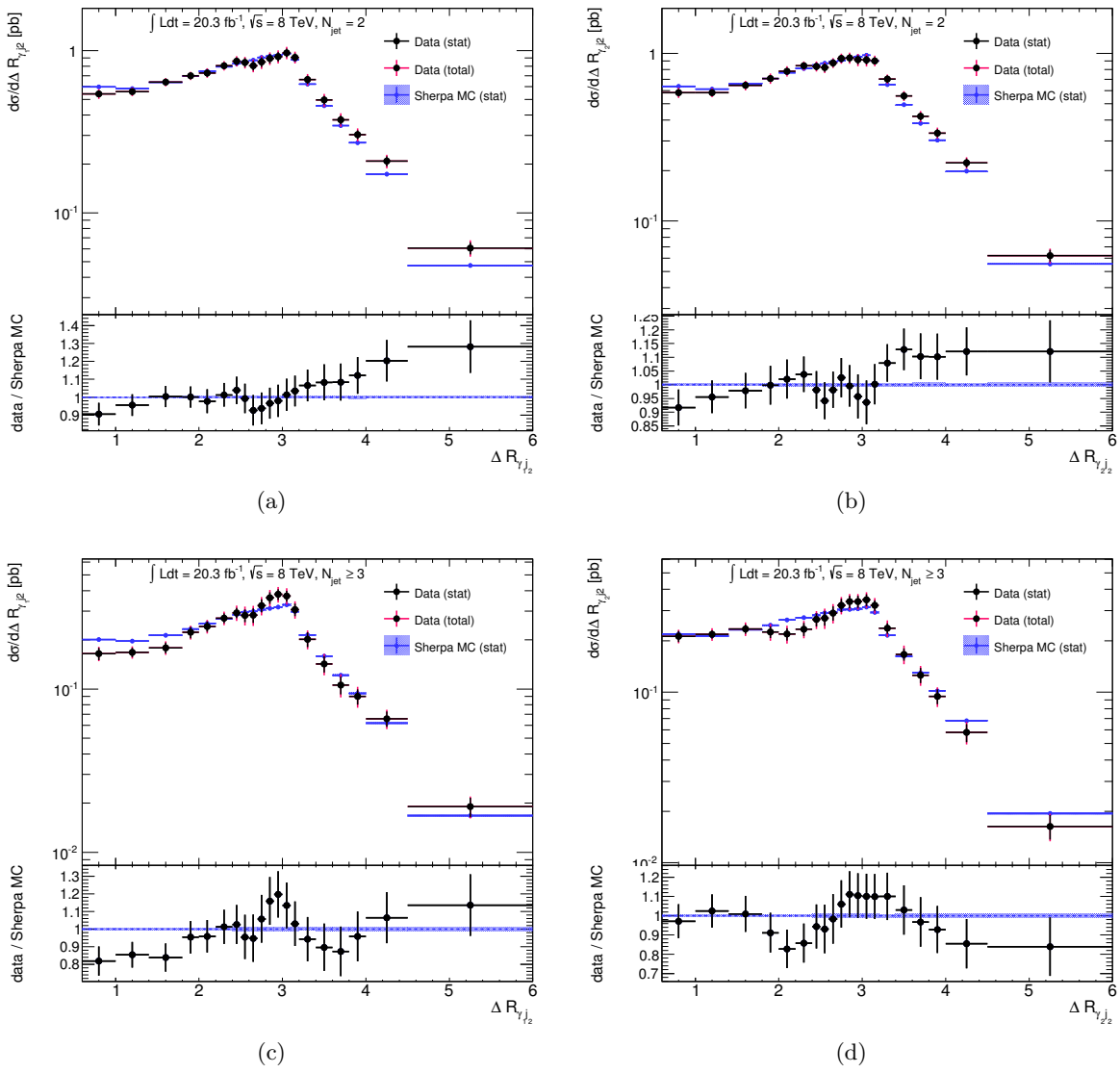


Figure 7.77: Comparison between the measured cross sections in data and SHERPA MC, as a function of the R-separation between the leading photon and the subleading jet (left) and the subleading photon and the subleading jet (right) for the $2-jet$ (top) and $\geq 3-jet$ (bottom) category.

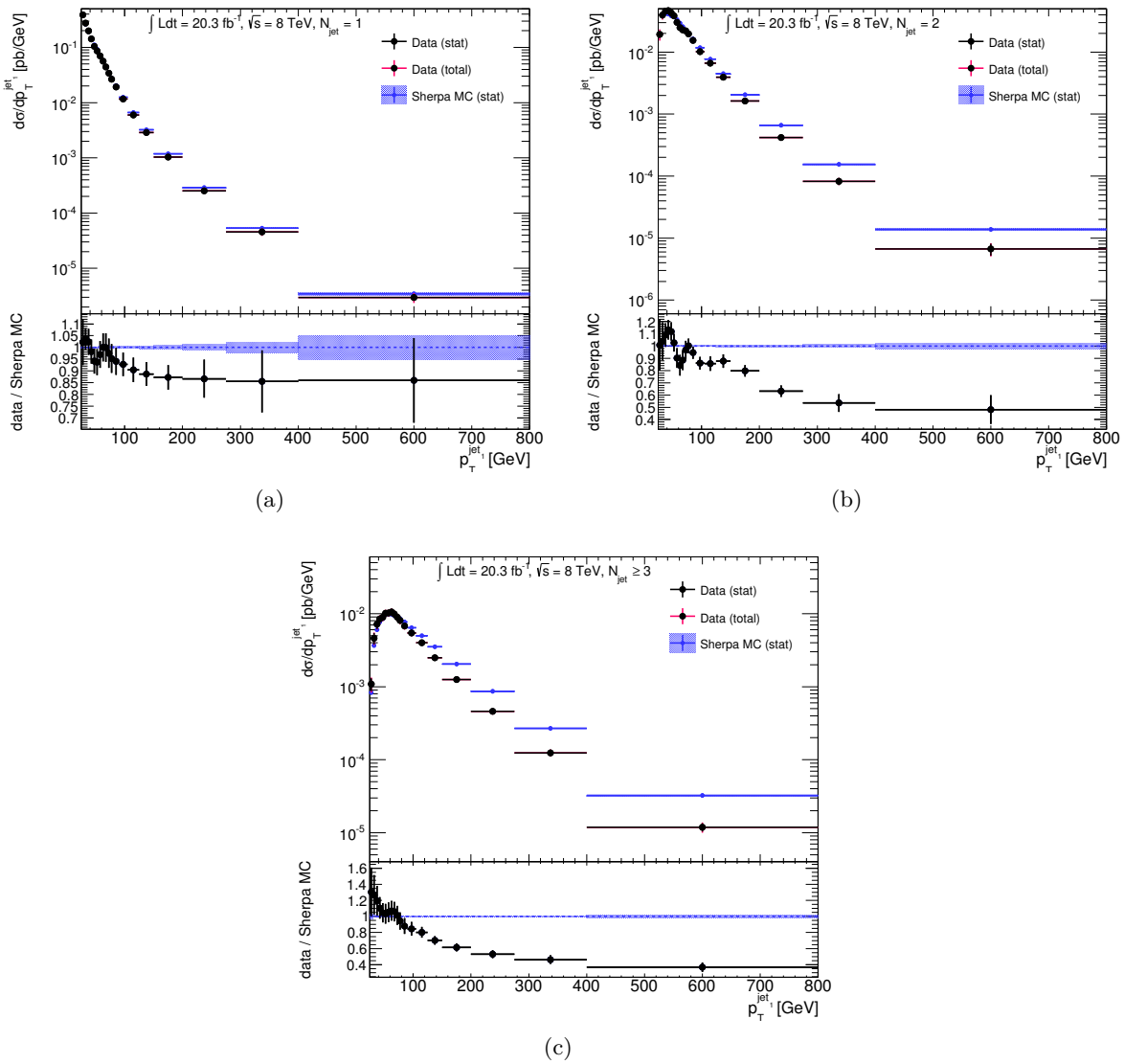
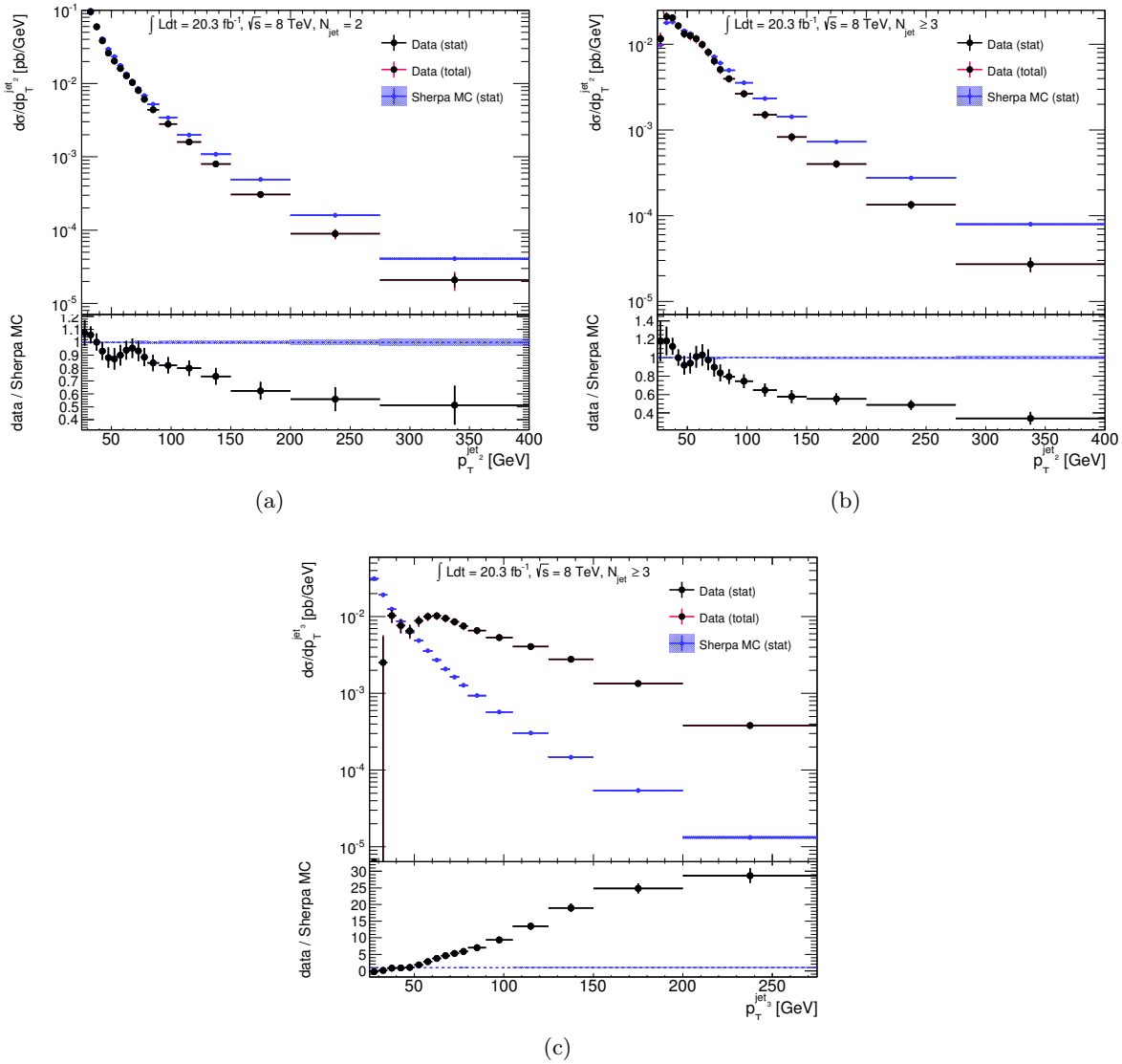


Figure 7.78: Comparison between the measured cross sections in data and SHERPA MC, as a function of the leading jet $p_{\mathrm{T}}^{jet_1}$. (a) $1-jet$, (b) $2-jet$, (c) $\geq 3-jet$ category.



(a) (b)

(c)

Figure 7.79: Comparison between the measured cross sections in data and SHERPA MC, as a function of the p_{T}^{jet} for the subleading jet in the (a) $2-jet$ and (b) $\geq 3-jet$ category and (c) the third jet in the $\geq 3-jet$ category.

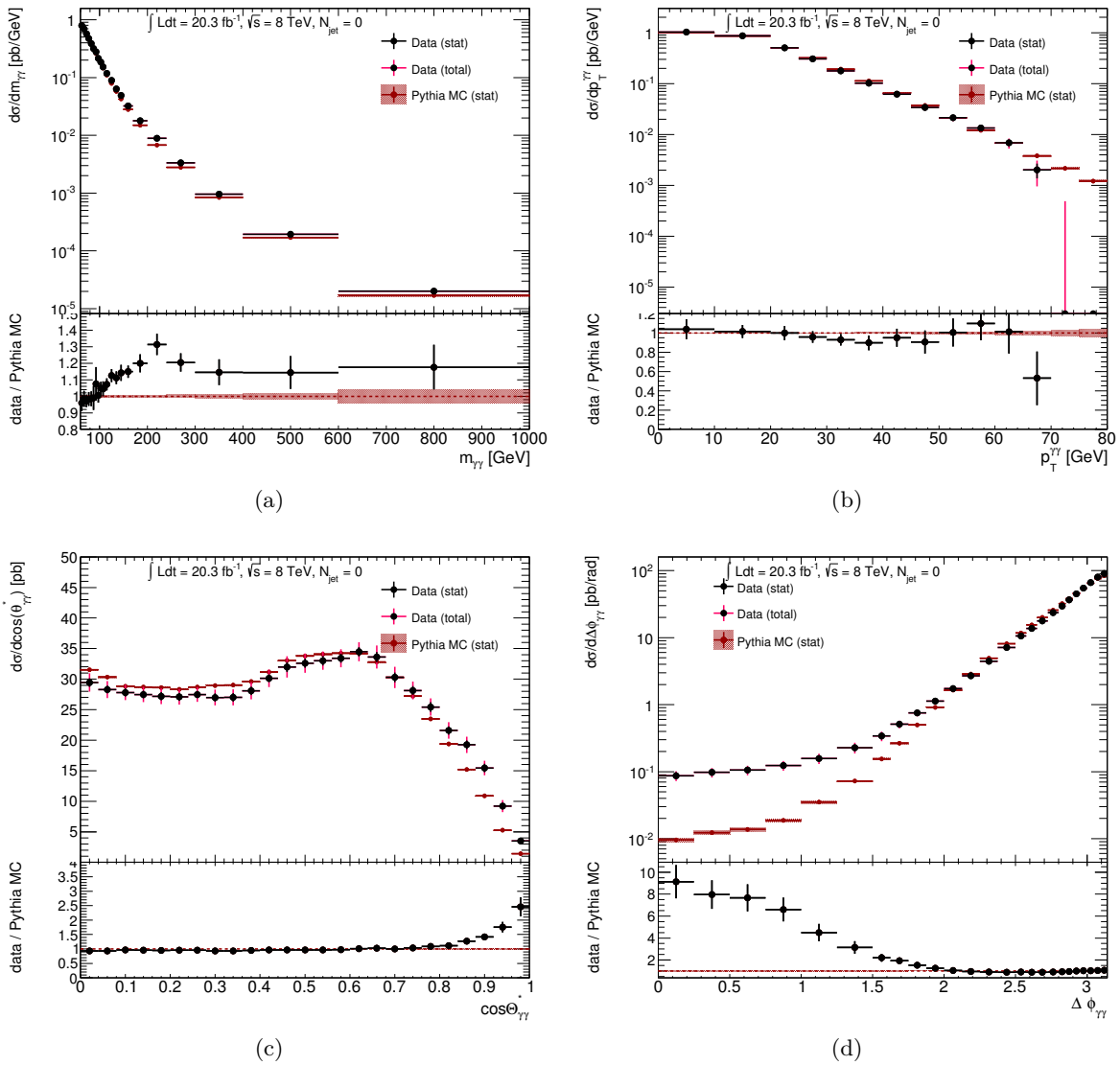


Figure 7.80: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the diphoton observables for the $0-jet$ category.

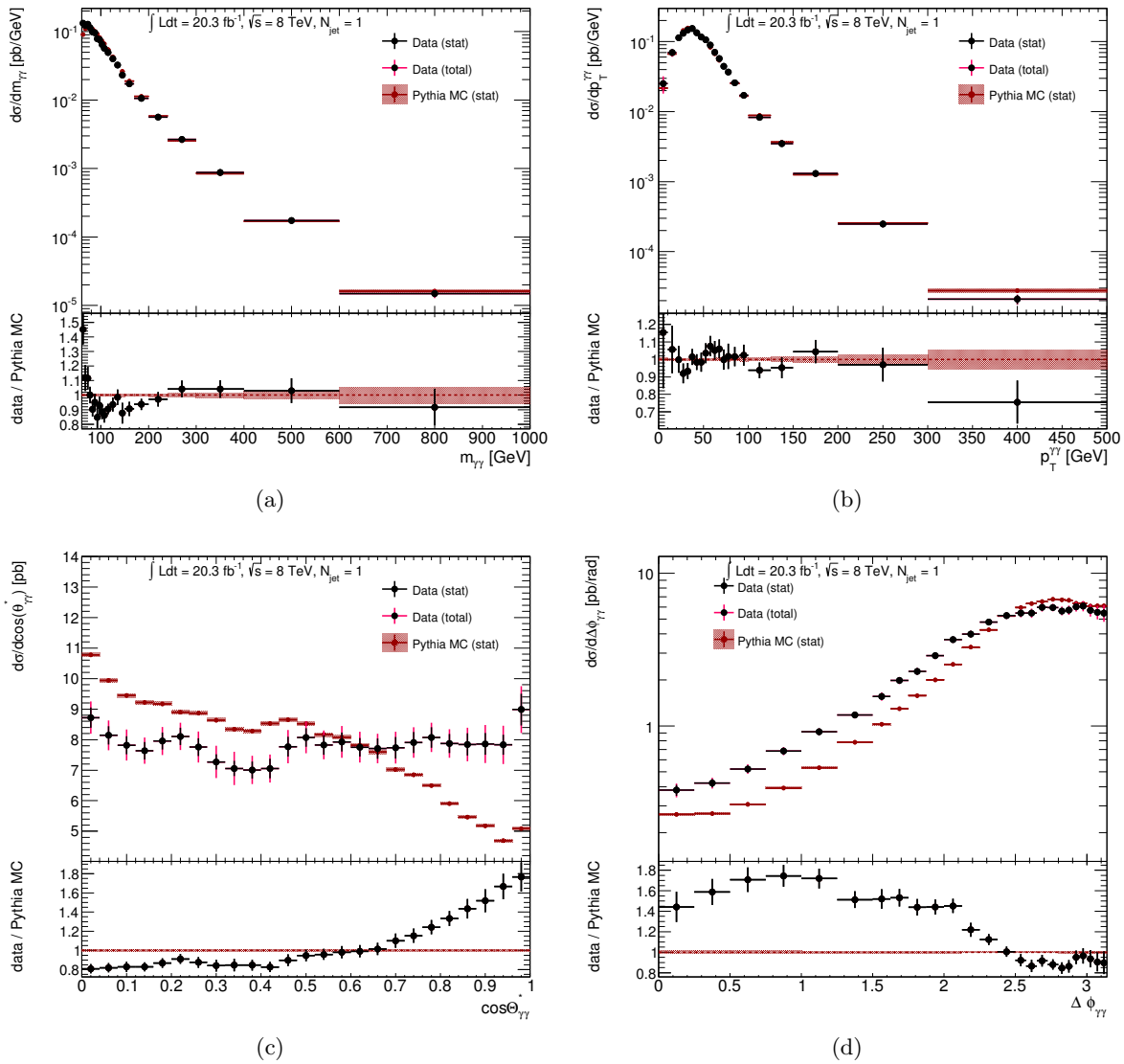


Figure 7.81: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the diphoton observables for the $1-jet$ category.

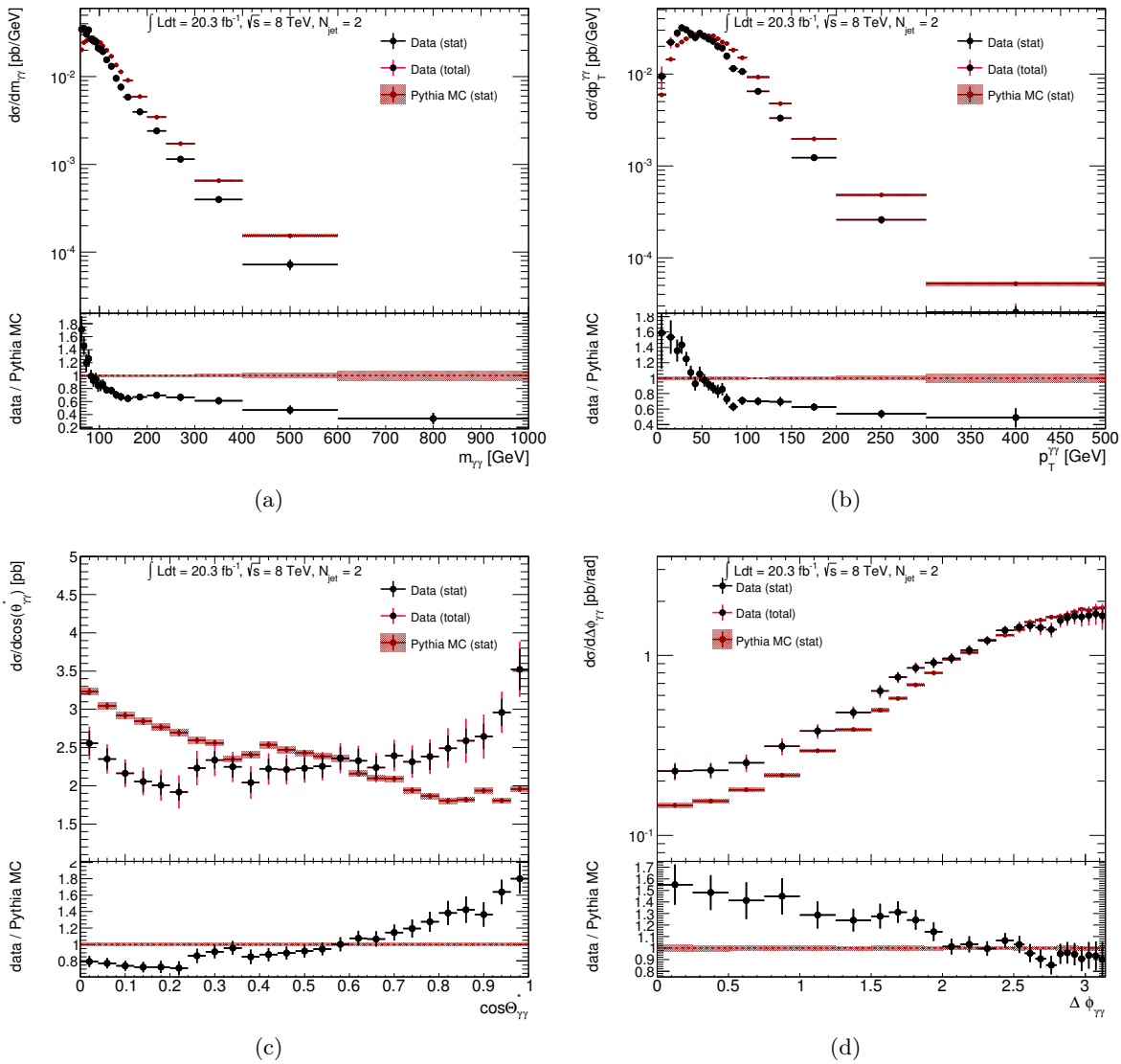


Figure 7.82: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the diphoton observables for the $2-jet$ category.

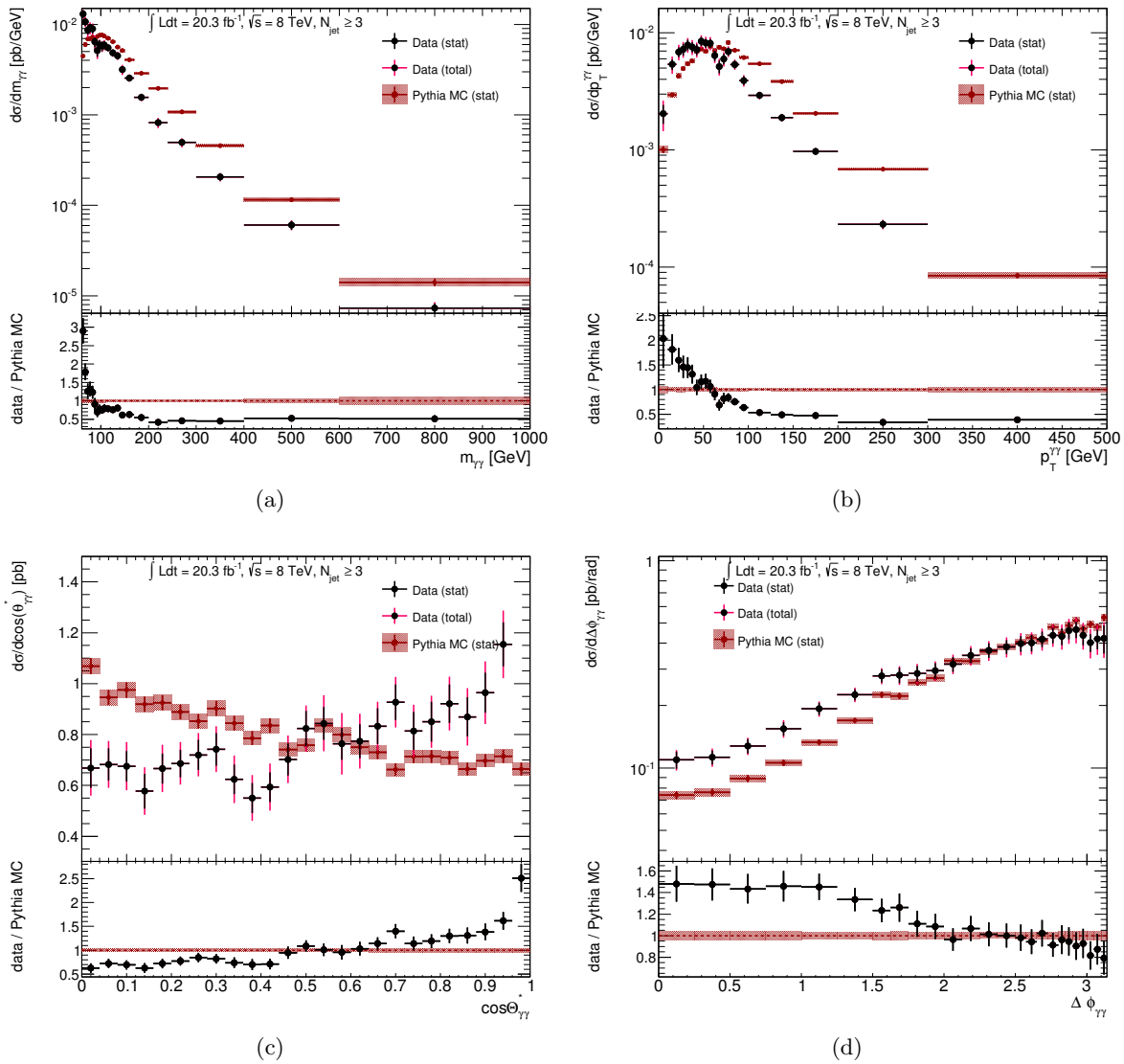


Figure 7.83: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the diphoton observables for the $\geq 3 - jet$ category.

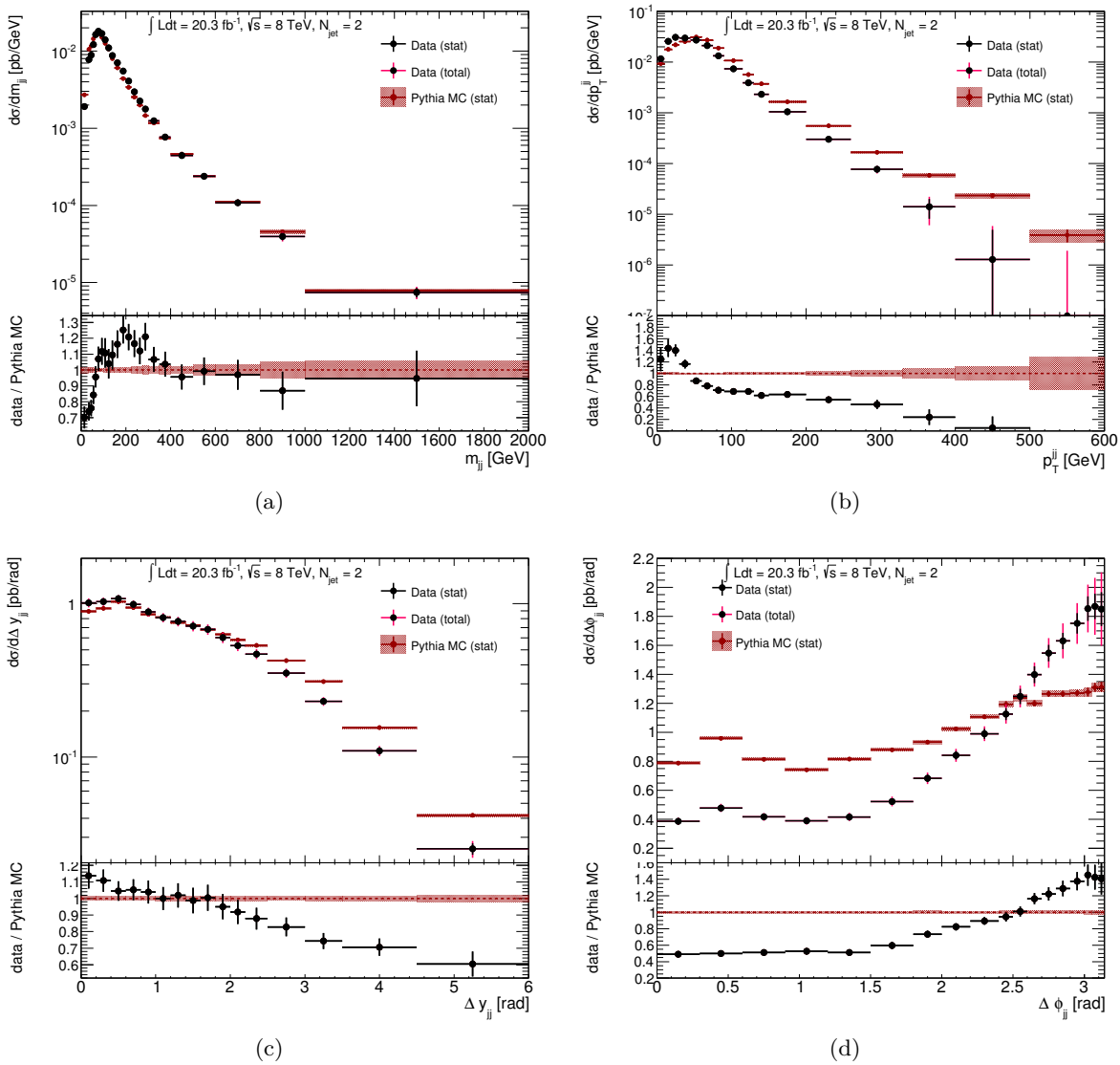
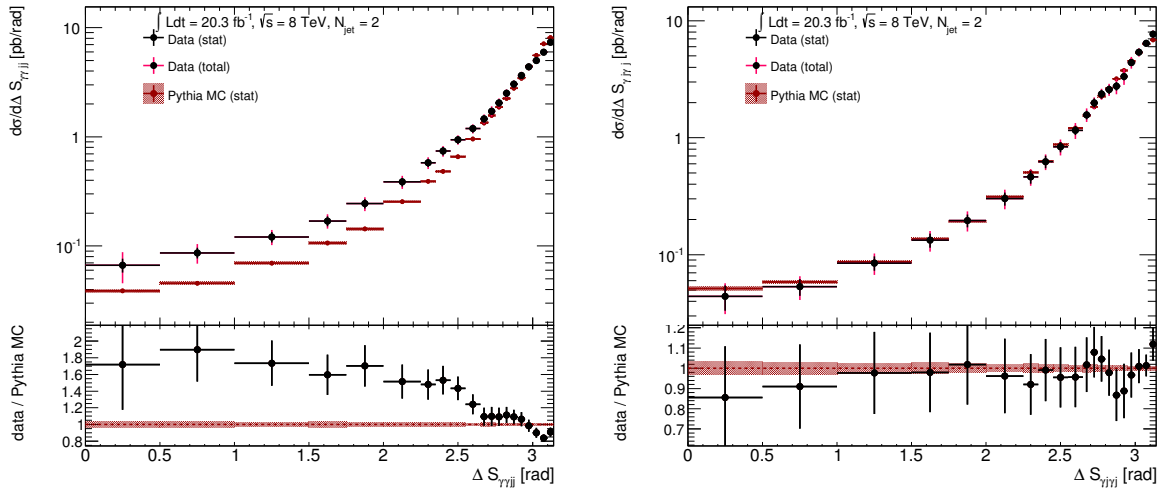


Figure 7.84: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the dijet observables for the $2-jet$ category.



(a) (b)

Figure 7.85: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the ΔS observables for the $2-jet$ category.

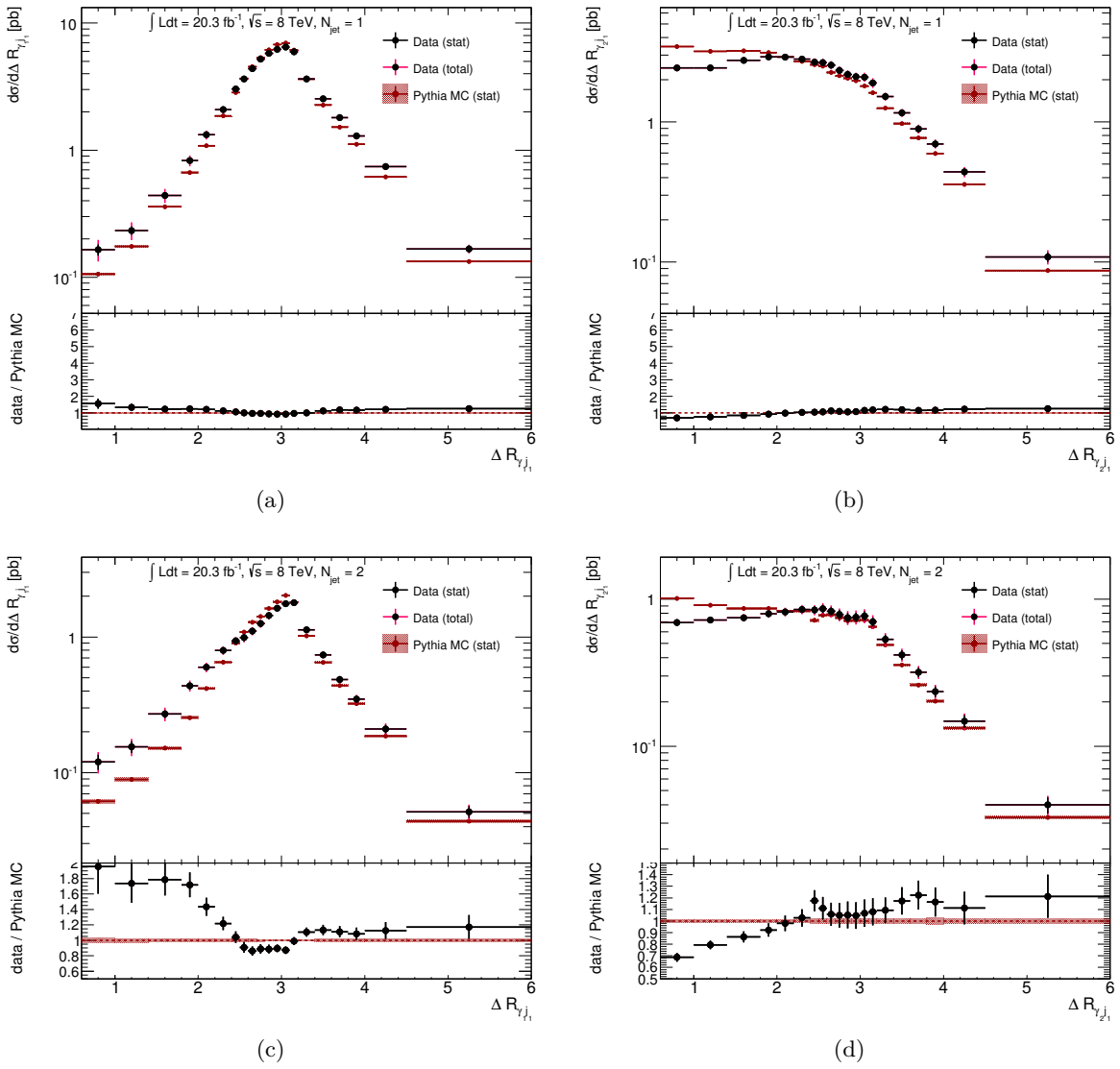
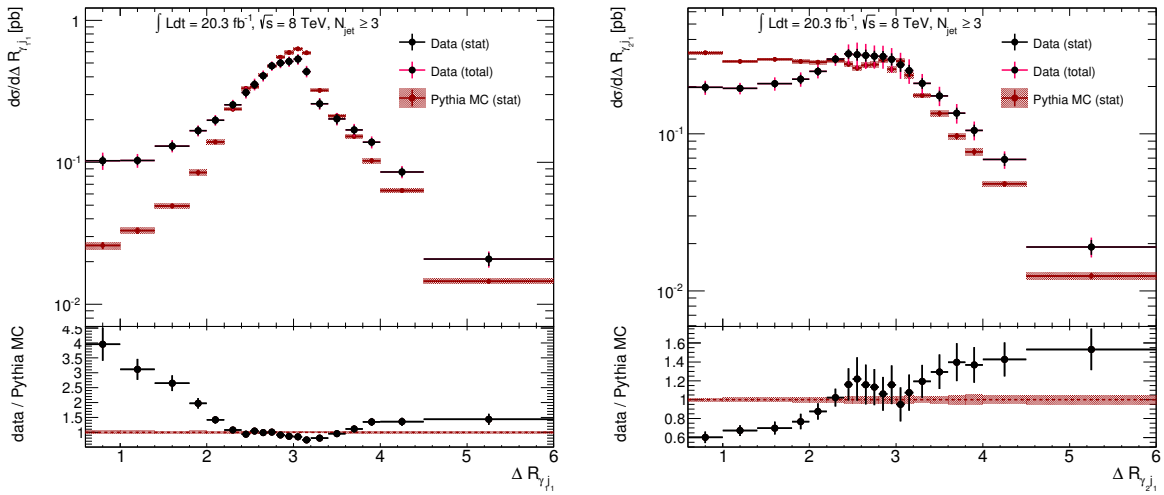


Figure 7.86: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the R-separation between the leading photon and the leading jet (left) and the subleading photon and the leading jet (right) for the $1-jet$ (top) and $2-jet$ (bottom) category.



(a) (b)

Figure 7.87: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the R-separation between the leading photon and the leading jet (left) and the subleading photon and the leading jet (right) for the $\geq 3-jet$ (bottom) category.

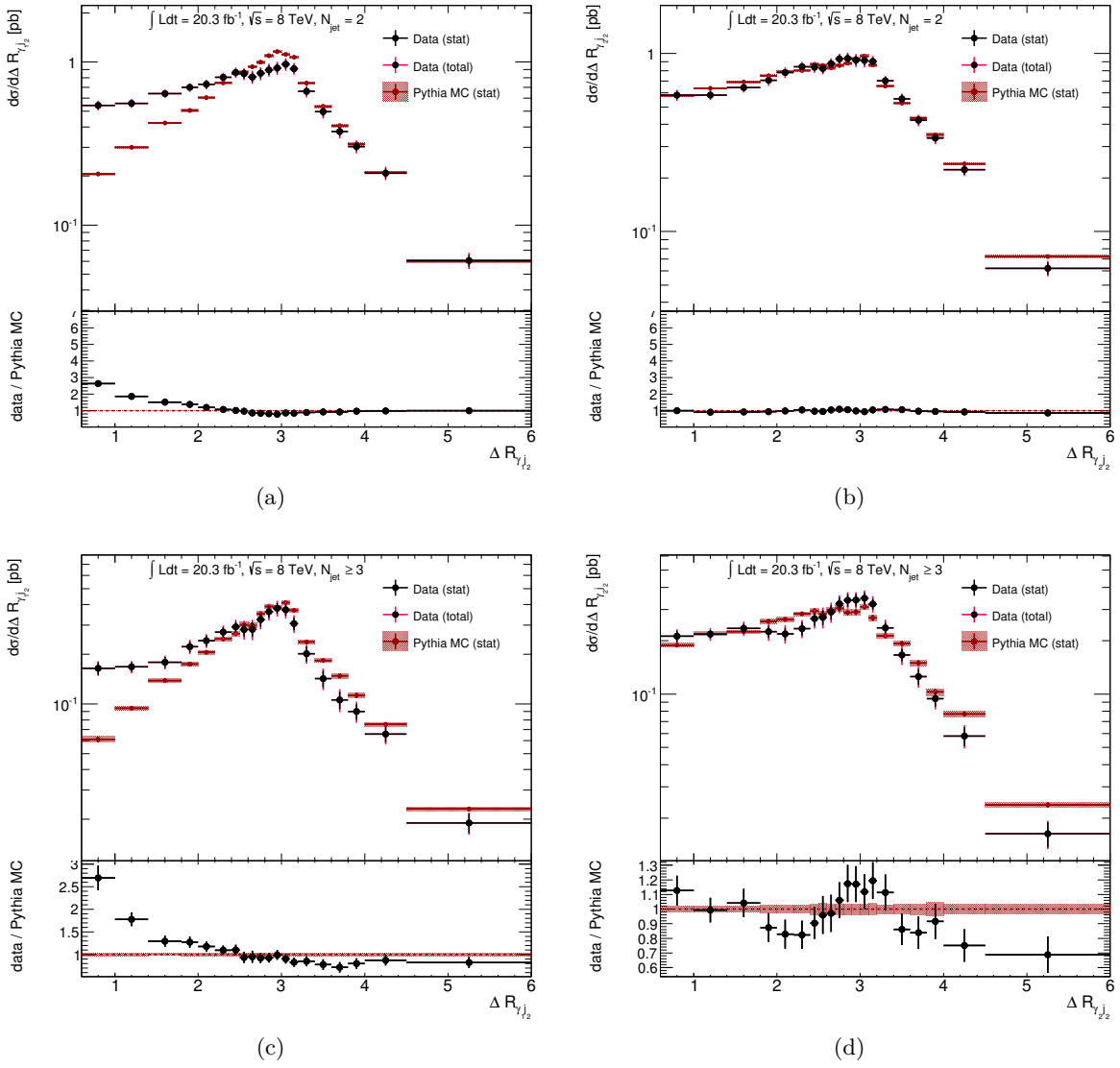
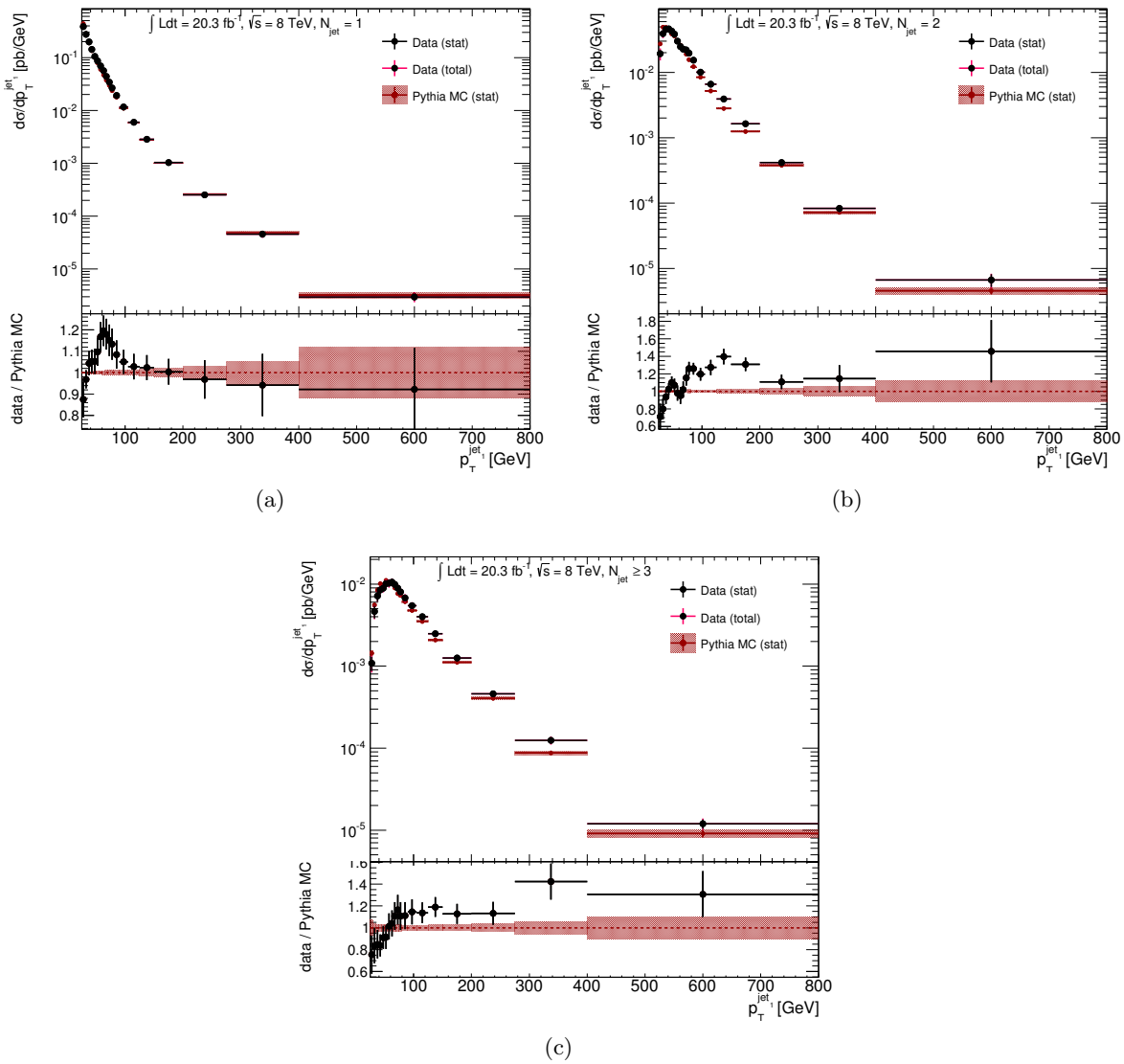


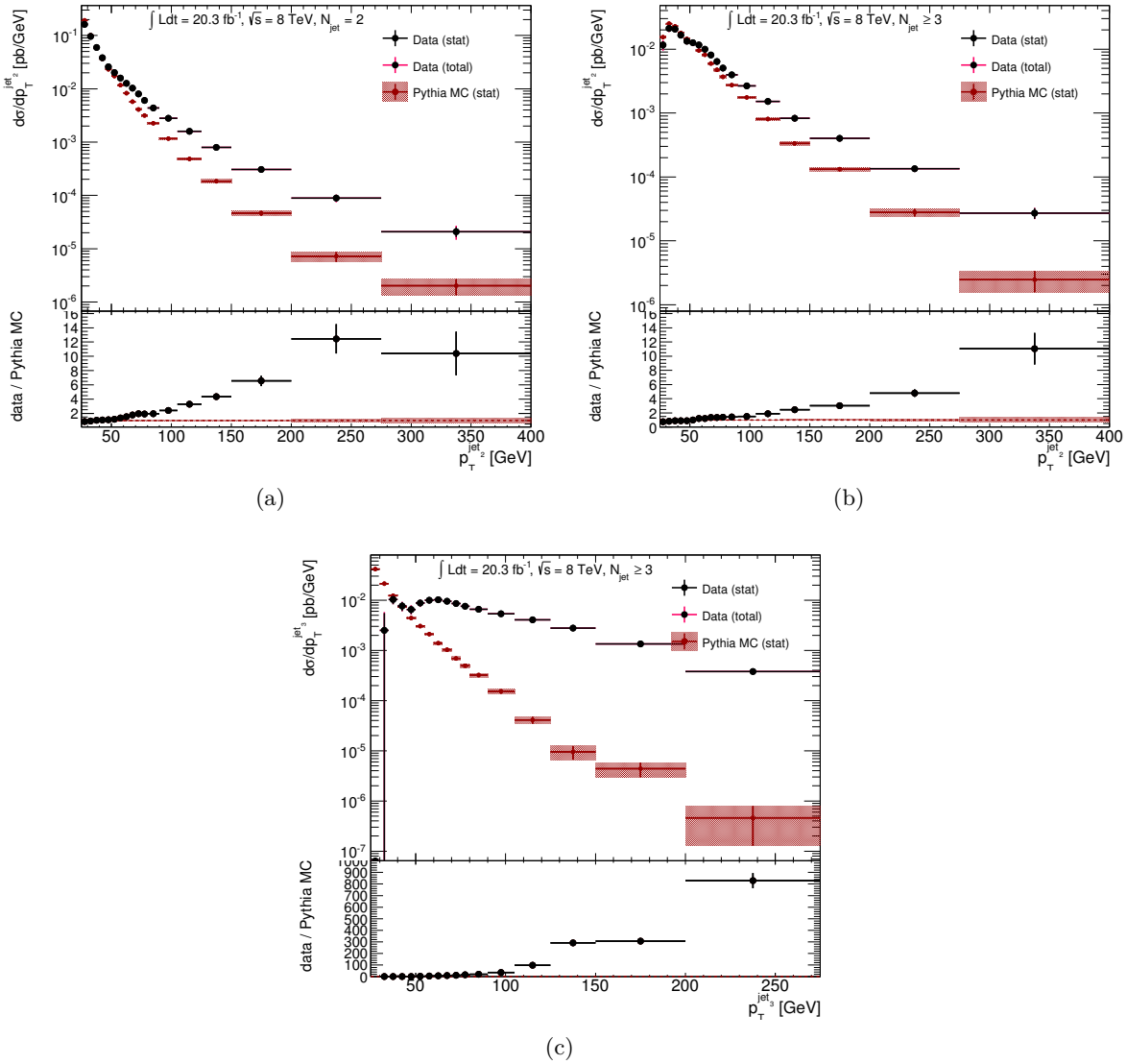
Figure 7.88: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the R-separation between the leading photon and the subleading jet (left) and the subleading photon and the subleading jet (right) for the $2-jet$ (top) and $\geq 3-jet$ (bottom) category.



(a) (b)

(c)

Figure 7.89: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the leading jet $p_{\mathrm{T}}^{jet_1}$. (a) $1-jet$, (b) $2-jet$, (c) $\geq 3-jet$ category.



(a) (b)

(c)

Figure 7.90: Comparison between the measured cross sections in data and PYTHIA MC, as a function of the p_{T}^{jet} for the subleading jet in the (a) $2-jet$ and (b) $\geq 3-jet$ category and (c) the third jet in the $\geq 3-jet$ category.

8

Optimization of the LAr signal reconstruction for the LHC Run 2

During the LHC proton-proton operations in 2011 and 2012, the average number of collisions μ increased from ~ 10 to ~ 20, as seen in Figure 2.4 on page 32. Because of the larger expected pileup in 2012, the reconstruction of the ATLAS Liquid Argon calorimeter signal amplitudes [254], based on the Optimal Filtering technique [153], was migrated from an optimization aiming to correct for the electronic noise only, to one targeting the combination of electronic and pileup noise, and the reduction of the bias introduced by the latter in the forward region [155].

The average number of collisions μ is expected to increase even more when the LHC restarts its operations at $\sqrt{s} = 13$ TeV in 2015, with peak values that can reach up to $\mu \approx 50-60$. The beam bunch spacing(BS) will also change. In Run 1, the proton beams consisted of 1380 bunches, distributed every 50 ns. For Run 2, the LHC plans to reach its nominal running conditions of colliding bunches every 25 ns, with a short period at the beginning of 2015 data-taking when the bunch spacing will still be 50 ns.

In order to cope with the increasing number of bunches and the change of collision frequency, the ATLAS detector trigger system at the L1 level will change its acceptance frequency from about 60 kHz to about 100 kHz. The LAr calorimetry system is currently capable to operate without becoming "busy"[1] up to around 90 kHz, when 5 LAr samples are transmitted from the frond-end to the back-end electronics to reconstruct the pulse amplitudes, due to a limited bandwidth of 1.6 Gbit/s between them. In order to keep the detector "busy" time as small as possible, the option of using a smaller number of LAr samples is investigated.

The use of a reduced number of samples for the LAr signal amplitude reconstruction would imply a worse noise reduction performance. For this reason, the noise behavior of the LAr calorimeters is studied using simulated Zero Bias (ZB) events, where only signals associated to electronic noise and energy deposits due to collision pileup are simulated. Samples for different μ (30, 60, 90) conditions, BS configurations (25 ns, 50 ns) and Optimal Filtering reconstruction (correct/wrong optimization for a given bunch spacing configuration) are used. The results are presented for the three main calorimeter parts separately: the electromagnetic calorimeter (EM), the hadronic end-caps (HEC) and the forward calorimeters (FCAL).

[1] During data-taking, ATLAS flags its subdetectors as "busy" when they are unable to accept, process and write out new events due to hardware limitations.

8.1 LAr signal reconstruction approches using the Optimal Filtering technique

The cell energy reconstruction of the LAr calorimeter is based on the Optimal Filtering Technique [153], briefly summarized in Section 2.3.2. For the 2010 and 2011 data taking periods, the LAr OFCs were optimized to minimize the electronics noise only (this optimization is further referred to as *Standard OFC* or STD). For the 2012 data-taking, when a larger pileup contribution was expected, the optimization of the OFCs for the electronic and pile-up noise was introduced [155] (here referred to as *Optimization 0* or OPT0).

Since the OFC signal reconstruction, in presence of very narrow pulses and a very large pileup contribution associated to a large bunch spacing, can lead to an artificial shift in the average reconstructed energy, an extra constraint was added to the traditional OFC computation to minimize this bias in high η regions of the LAr calorimeters (this optimization is labeled here as *Optimization 1* or OPT1) [155]. The observed shift is interpreted as coming from the impact of the imperfect average cancellation of the pile-up contribution under the mentioned conditions. The OFCs are optimized by minimizing

$$I = \sum_{ij} R_{ij} a_i a_j - \lambda \left(\sum_i a_i g_i - 1 \right) - \kappa \sum_i a_i g'_i - \gamma \sum_i a_i \delta_i, \tag{8.1}$$

as in Eq. 2.15 on page 48, but with an additional constrain, where $\boldsymbol{\delta}$ is the average shift for each sample - the extra constrain. The OFCs are obtained by minimizing I w.r.t. $\boldsymbol{a}$ as

$$a = R^{-1}(\lambda \boldsymbol{g} + \kappa \boldsymbol{g'} - \gamma \boldsymbol{\delta}). \tag{8.2}$$

Minimizing I w.r.t. λ κ and γ gives

$$\frac{\partial I}{\partial \lambda} = \sum_i a_i g_i - 1 = 0 \tag{8.3}$$

$$\frac{\partial I}{\partial \kappa} = \sum_i a_i g'_i = 0 \tag{8.4}$$

$$\frac{\partial I}{\partial \gamma} = \sum_i a_i \delta_i = 0. \tag{8.5}$$

By multiplying Eq. 8.2 by $\boldsymbol{g}^\dagger$, $\boldsymbol{g'}^\dagger$ and $\boldsymbol{\delta}^\dagger$ from the left and using the constrains of Eq. 8.3-8.5, one gets

$$\left(\boldsymbol{g}^\dagger R^{-1} \boldsymbol{g}\right) \lambda + \left(\boldsymbol{g}^\dagger R^{-1} \boldsymbol{g'}^\dagger\right) \kappa + \left(\boldsymbol{g}^\dagger R^{-1} \boldsymbol{\delta}\right) \gamma = 1, \tag{8.6}$$

$$\left(\boldsymbol{g'}^\dagger R^{-1} \boldsymbol{g}\right) \lambda + \left(\boldsymbol{g'}^\dagger R^{-1} \boldsymbol{g'}^\dagger\right) \kappa + \left(\boldsymbol{g'}^\dagger R^{-1} \boldsymbol{\delta}\right) \gamma = 0, \tag{8.7}$$

$$\left(\boldsymbol{\delta}^\dagger R^{-1} \boldsymbol{g}\right) \lambda + \left(\boldsymbol{\delta}^\dagger R^{-1} \boldsymbol{g'}^\dagger\right) \kappa + \left(\boldsymbol{\delta}^\dagger R^{-1} \boldsymbol{\delta}\right) \gamma = 0. \tag{8.8}$$

Solving for λ, κ, δ, together with Eq. 8.2, one obtains the OFC $\boldsymbol{a}$ [155].

The overview of the three optimizations and which regions of the calorimeter they were applied in, is shown in Table 8.1.

Method	Standard	Optimization 0	Optimization 1
Data	2010-2011	2012	2012
Region	all	EM ($\|\eta\| < 2.5$) FCAL2+3 ($3.2 < \|\eta\| < 4$)	EM ($2.5 < \|\eta\| < 3.2$) HEC, FCAL1 FCAL2+3 ($4 < \|\eta\| < 4.9$)
Noise	Electronics	Electronics + Pile-up	Electronics + Pile-up

Table 8.1: Definition of the different OFC optimization methods. The *Standard* method has been used to reconstruct 2010 and 2011 data, *Optimization 0* and *Optimization 1* were used for 2012 data in different detector regions.

8.2 Simulation and reconstruction setup

This study uses Monte Carlo samples simulating 1000 Zero Bias events at each μ value (30, 60, 90). No physics process is generated, and only minimum bias events, corresponding to the desired pileup level, are overlaid. The response of the ATLAS detector is then simulated, including the expected electronic noise levels. The minimum bias samples used for the pileup overlay are simulated at $\sqrt{s} = 7$ TeV[2] with Pythia8 [71], for low and high p_{T} minimum bias. Low p_{T} minimum bias events represent particles produced by soft interactions as part of pile-up, while the high p_{T} minimum bias events account for particles coming from harder, high p_{T} parton-parton scatterings that are still part of the pile-up and are well described by QCD. Each pileup configuration corresponds to a different mixture of the low p_{T} and high p_{T} minimum bias events, shown in Table 8.2. For each μ value, samples are produced using different bunch

μ	number of low p_{T} MBE	number of high p_{T} MBE
30	29.96	0.04
60	59.92	0.08
90	89.88	0.12

Table 8.2: Minimum bias events (MBE) low p_{T} and high p_{T} configuration for each pileup configuration - μ.

spacing configuration of 50 ns and 25 ns. The simulation included eighty bunches in a beam, in three bunch trains (bunches 73-77 were empty).

The simulated events are digitized and reconstructed using Athena version 17.2.0.2, applying the three OFC optimizations and changing the number of LAr samples used for reconstruction from five to four and three.

Samples generated for each μ point using the 25 ns and 50 ns bunch spacing configurations, are then reconstructed with OFC optimized either for the pileup autocorrelation matrix expected for the simulated bunch spacing, or for that computed for the alternative bunch spacing ("wrong" optimization). For each of the 4 bunch-spacing / pileup autocorrelation configurations, 3 alternative OFC optimizations were tried:

- **Standard**: OFC optimized only for the electronic noise, thus insensitive to the bunch

[2] $\sqrt{s} = 7$ TeV is chosen due to the availability of pile-up simulation samples at the time of the study.

spacing configuration;

- **Optimization 0**: OFC optimized for electronic noise and pile-up corresponding to the simulated μ, with pileup autocorrelation varying from the one used in the digitization steps, or the alternative "wrong" one;
- **Optimization 1**: OFC optimized for electronic noise and pile-up corresponding to the simulated μ, plus the additional constrain to minimize the energy bias, with pileup autocorrelation varying from that used in the digitization steps, or the alternative "wrong" one.

For the study of the noise performance as a function of the number of signal samples, only samples reconstructed with OFC optimized using the pileup autocorrelation corresponding to the simulated bunch-spacing are produced. In this case samples for all three μ values were produced, when only three samples (three central one out of the usual five) and four samples (first four and last four out of the usual five) were used both to compute the OFC and to perform the LAr energy reconstruction.

8.3 Mean energy and noise vs. pileup autocorrelation matrix

The results of the study of the impact of "correctly" and "wrongly" optimized OFCs when applied to a given bunch spacing configuration are presented in this section. When starting the LHC Run 2 in 2015, first a 50 ns BS run was scheduled, which delivered roughly 0.85 fb^{-1} of collision data to the experiments. After, the LHC started running on its nominal BS configuration of 25 ns. The "wrongly" optimized OFCs are studied here to investigate their implications on the noise behavior of the calorimeter. In case the noise is not significantly worse, the same OFCs may be used for both BS conditions.

Five signal samples are used to reconstruct the LAr cell energy. The noise reduction or increase is studied in three μ conditions (μ = 30, 60 and 90) for 25 ns and 50 ns bunch spacing. The notation for the bunch spacing conditions and the pileup autocorrelation used for the OFC optimization is shown in Table 8.3.

beam / optimization	25 ns	50 ns
25 ns	25_25	25_50
50 ns	50_25	50_50

Table 8.3: Notation of bunch spacing conditions for simulation and reconstruction.

8.3.1 EM calorimeters

The impact of the different OFC optimizations on the average cell energy $\langle E_{cell} \rangle$ and therefore the noise behavior in the EM part of the LAr calorimeters ($|\eta| < 2.5$) is shown in Figure 8.1 as a function of the Bunch Crossing IDentifier (BCID), integrated over η. Since both bunch spacing configurations foresee three bunch trains of 80 bunches, the properties are studied by superimposing the three trains. The increase of the average cell energy at the beginning and

at the end of the bunch-trains, associated to the imperfect cancellation of the response baseline despite the bipolar pulse shape, is reduced when using the two methods optimizing the OFCs for pile-up only or also including the constraint on the baseline shift (red and green points, respectively).

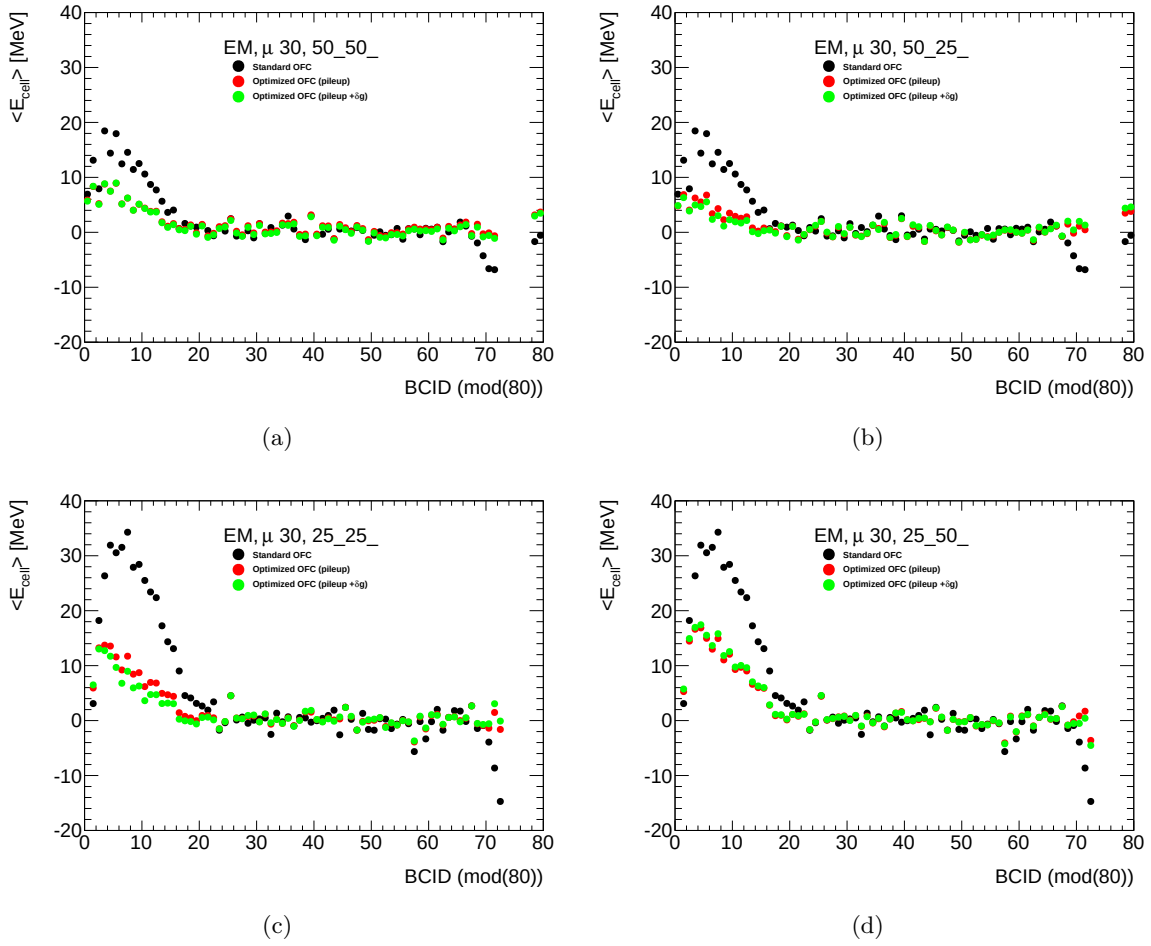


Figure 8.1: Average cell energy $\langle E_{cell} \rangle$ as reconstructed from ZB events at $\mu = 30$ in the EM part of the LAr calorimeters, for 50 ns (a)(b) and 25 ns (c)(d) BS configurations, reconstructed using different OFC optimizations and alternative pileup autocorrelation matrices for 25 ns BS (b)(d) and 50 ns BS (a)(c). Standard optimization (black points), Optimization 0 (red points) and Optimization 1 (green points).

Figure 8.2 shows the same quantities for $\mu = 60$. For this pileup level, even the OFCs using correct pileup autocorrelation matrix for 50 ns bunch spacing introduce a sizable bias in the average energy. Such effect is possibly associated to the peculiar properties of the 50 ns pileup autocorrelation used for the calculation of the OFCs. When the wrong pileup autocorrelation matrix is used to optimize the OFCs for data simulated at 50 ns bunch spacing, the bias is even larger. On the other hand, for the same level of pileup with 25 ns bunch spacing, the bias disappears, regardless of the pileup autocorrelation matrix used to optimize the OFCs. For even larger pileup level, $\mu = 90$, the effects observed at $\mu = 60$ are even more pronounced (Figure 8.3).

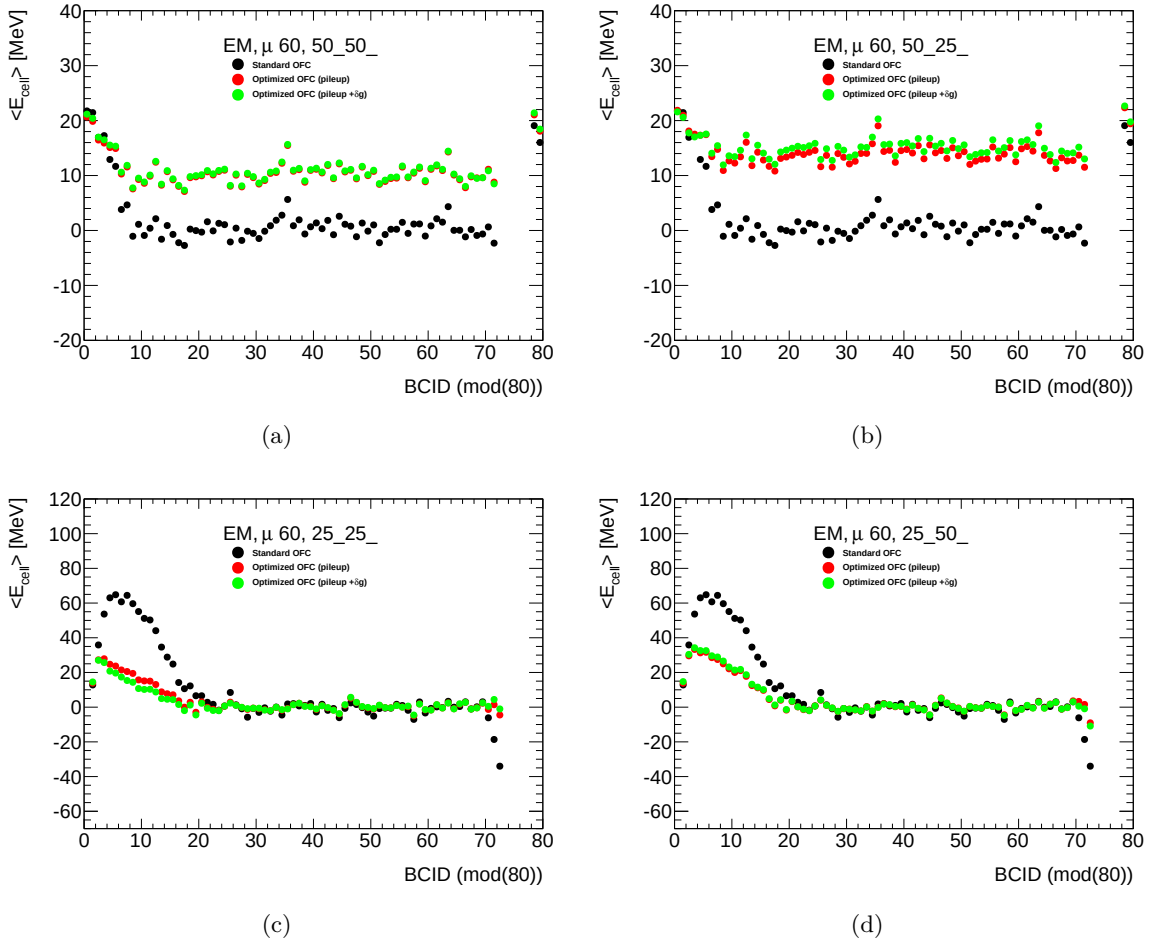


Figure 8.2: Average cell energy $\langle E_{cell} \rangle$ as reconstructed from ZB events at $\mu = 60$ in the EM part of the LAr calorimeters, for 50 ns (a)(b) and 25 ns (c)(d) BS configurations, reconstructed using different OFC optimizations and alternative pileup autocorrelation matrices for 25 ns BS (b)(d) and 50 ns BS (a)(c). Standard optimization (black points), Optimization 0 (red points) and Optimization 1 (green points).

The average cell energy shown in Figures 8.1, 8.2 and 8.3 is computed integrating over the relevant η range of the EM calorimeter ($|\eta| < 3.2$). In order to explore localized biases, the same quantity is studied as a function of η after having removed the first and last part of the bunch trains, where the imperfect cancellation of the baseline shift is expected. This is done by requiring

$$20 < \mathrm{mod}(BCID, 80) < 65. \tag{8.9}$$

The behavior of the average cell energy $\langle E_{cell} \rangle$ as a function of η is shown for the different bunch spacing and OFC optimization configurations, and all the μ values, in Figure 8.4. It is observed again that the reconstruction using OFCs optimized for the wrong bunch spacing provides a biased $\langle E_{cell} \rangle$ w.r.t. the one using correct bunch spacing, and the bias grows with increasing μ. The effect is particularly pronounced for $|\eta| > 2.5$.

A related effect is observed on the noise property, estimated with the RMS of the E_{cell} distribu-

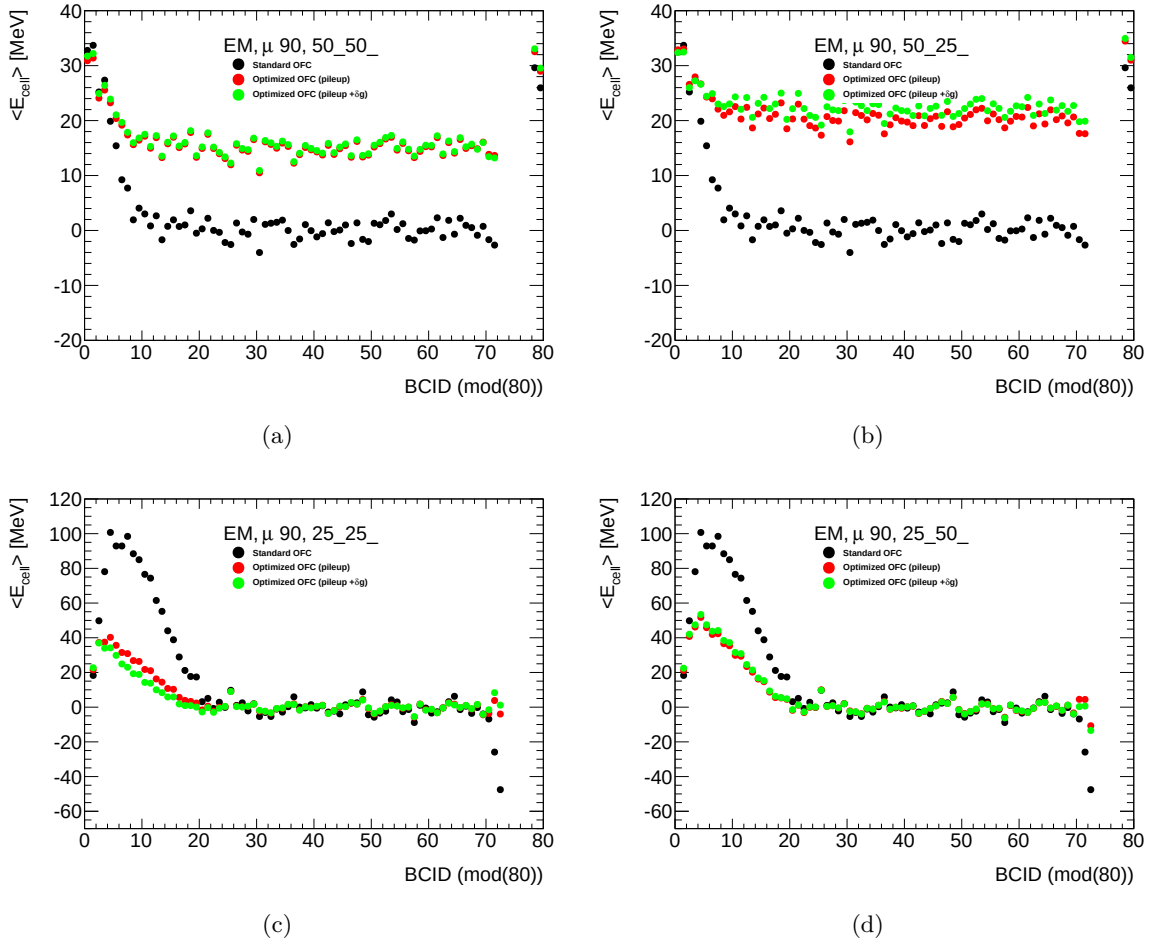


Figure 8.3: Average cell energy $\langle E_{cell} \rangle$ as reconstructed from ZB events at $\mu = 90$ in the EM part of the LAr calorimeters, for 25 ns and 50 ns BS configurations, as reconstructed using different OFC optimizations and alternative pileup autocorrelation matrices for 25 ns BS (b)(d) and 50 ns BS (a)(c) . Standard optimization (black points), Optimization 0 (red points) and Optimization 1 (green points).

tion, that are shown in Figure 8.5. As expected, the introduction of a specific optimization for the pileup noise reduces the noise level with respect to those obtained when OFCs optimized only for the electronic noise are used. On the other hand, for increasing pileup levels and especially in the $|\eta| > 2.5$ region, the noise performance worsens when the wrong pileup autocorrelation matrix is used for the OFC optimization.

In order to better evaluate the change in noise, ratios of the RMS corresponding to the different reconstruction configurations are computed, and shown in Figures 8.6 and 8.7 only for those configurations when the correct pileup autocorrelation matrix is used for the OFC optimization.

For the 50 ns bunch-spacing configuration (Figure 8.6), a better noise performance is observed when the OFC optimization is used (OPT0, OPT1) with respect to that including only the electronic noise (STD). For $\mu = 30$ at high η, the noise is smaller by up to 50%; the improvement at higher pile-up is however smaller, only up to 10% and in central η regions. When comparing

the two pileup OFC optimizations (pileup only, or the additional constrain for the baseline bias), a very similar noise behavior is observed.

For the 25 ns bunch-spacing configuration (Figure 8.7), the noise reduction obtained when applying OFCs optimized for pileup is more stable with μ, reaching up to ≈70% at high η all the way down to a few percent for the most central part. In this case, a slightly better performance is obtained when the OFCs are only optimized for pileup, since the addition of the extra constrain for the baseline shift slightly reduces the improvement.

8.3.2 HEC calorimeters

The same study, discussed in Section 8.3.1 is performed for the HEC section of the LAr calorimeter ($1.4 < |\eta| < 3.4$), and a similar behavior of the average cell energy and RMS is observed for the different reconstruction configurations. The cell energy is less biased when the pileup autocorrelation for a given bunch-spacing is used to optimize the OFCs, and a smaller noise is observed.

Figure 8.8 shows the RMS ratio for the 50 ns bunch spacing configuration. A noise decrease up to 20% is observed in high η regions at $\mu = 30$, when the OFC optimization is used (OPT0, OPT1) with respect to that including only the electronic noise (STD). This decrease reduces to only a few percent in the more central η regions. As in the EM case, a slightly better performance is obtained when the OFC are only optimized for pileup.

In higher pile-up conditions ($\mu = 60$ or 90), the noise increase is observed, by up to 10% in high η regions, when the OFCs optimized for pileup are used, with respect to those addressing only the electronic noise. This unexpected and counterintuitive behavior is believed to be related to the very peculiar properties of the noise autocorrelation matrix for 50 ns bunch-spacing, but it is not completely understood and will need further studies.

For the 25 ns bunch spacing conditions (Figure 8.9) the noise behavior is more stable with μ. A noise improvement up to 60% is observed at high η when OFCs optimized only for pileup are used. When OFCs including the extra constraint for the baseline shift are used, the noise decrease reduces to only 10-20%, and only in the high η regions. In the low η regions the noise is instead higher, suggesting that, while optimizing the OFCs for pileup is certainly necessary in the HEC, the use of the extra constrain for the baseline shift should be avoided in this region.

8.3.3 FCAL

In the FCAL part of the LAr calorimeters ($3 < |\eta| < 5$), using the pileup autocorrelation matrix coresponding to the correct bunch spacing is again the best choice.

Figure 8.10 shows the RMS ratio plots for the 50 ns bunch spacing configuration. A noise decrease is observed in high η regions at $\mu = 30$, when the OFC optimization (OPT0, OPT1) is used with respect to that including only the electronic noise (STD). As for the HEC case, in higher pile-up conditions ($\mu = 60$ or 90) the noise is observed to increase by up to 20%, when the OFCs optimized for pileup are used, with respect to those addressing only the electronic noise. This unexpected and counterintuitive behavior is believed to be related to the very peculiar properties of the noise autocorrelation matrix for 50 ns bunch-spacing, but it is not completely understood and will need further studies. Then introduction of the extra constraint for the baseline shift recovers the effect, showing an improvement slowly increasing with η, up

to $\approx 10\%$.

For the 25 ns bunch spacing conditions (Figure 8.11) the noise behavior is again more stable with μ. The noise improvement is visible for both OFC pileup optimizations, with that adding the extra constraint begin only about 1% worse.

8.3.4 E_{cell} distribution properties

The value of $\langle E_{cell} \rangle$ and RMS discussed in Sections 8.3.1, 8.3.2 and 8.3.3 are obtained by computing the arithmetic mean and RMS of the cell energy distribution in a given bin. This representation might lead to biases in case of very asymmetric distributions, or in the presence of large tails. For this reason, in order to validate the results discussed above, the E_{cell} distributions are locally studied in more detail.

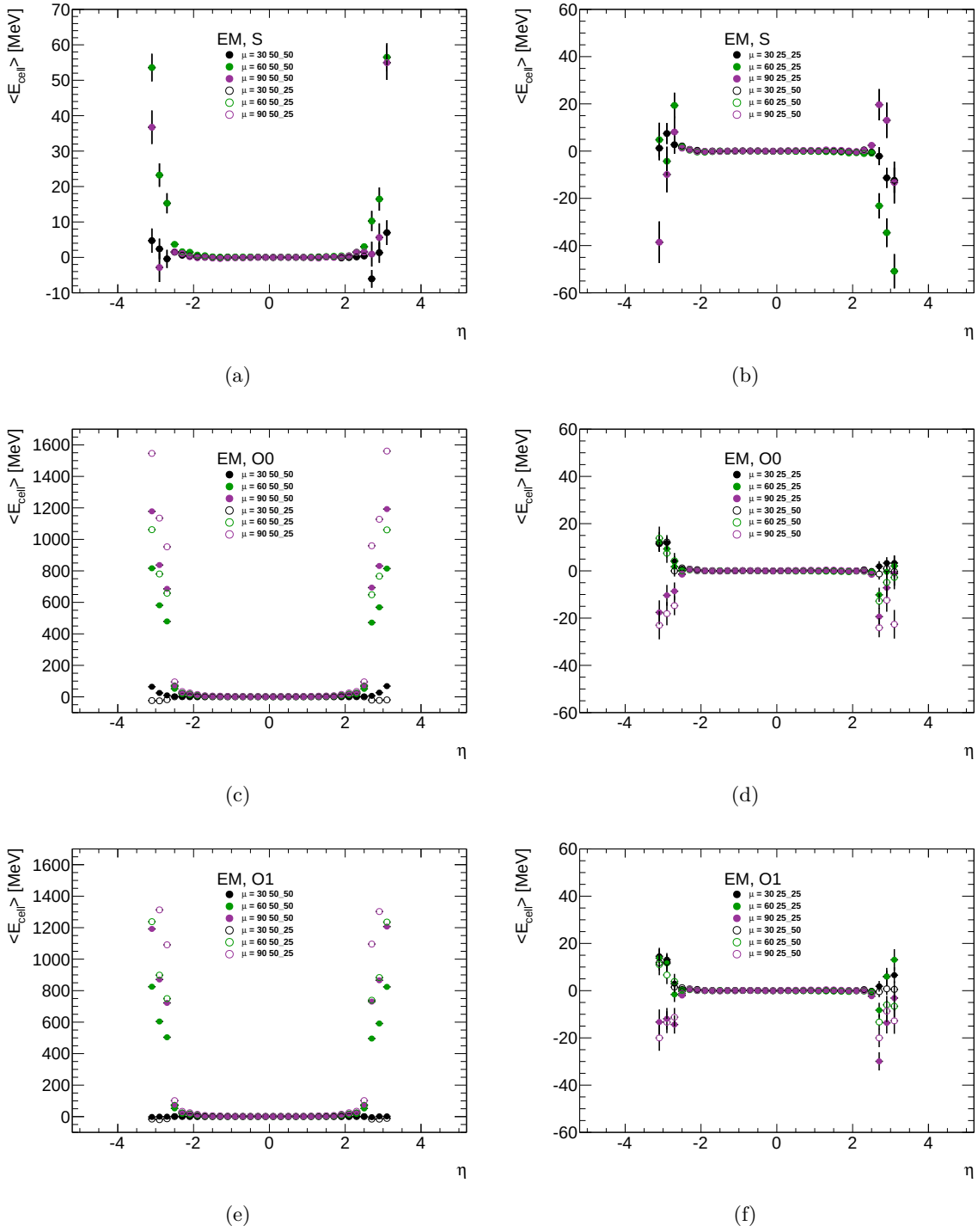
Figure 8.12 shows examples of the E_{cell} distributions for different bunch-spacing and OFC reconstruction configurations in the η regions where the largest biases are observed. All distributions show indeed very asymmetric tails. In order to evaluate the impact of such tails on the results discussed in the previous section, the distributions are therefore fit around their center, defined as the $\pm$ 1.5 RMS region centered around zero, with a Gaussian function. An example of such fit is shown in Figure 8.13. The mean and σ values obtained from the fits are then compared with the $\langle E_{cell} \rangle$and RMS obtained from the full samples. An example of such comparison is shown in Figure 8.14. The impact of the use of a Gaussian fit on the distribution cores is found to be small, and the results obtained using the simple mean and RMS remains representative of the effects. The validation was also performed for the HEC and FCAL parts of the calorimeter.

8.3.5 Summary

The results of this study show that, when optimizing the OFC for pileup, the use of the "wrong" autocorrelation matrix not according to the bunch spacing of the data to be reconstructed, generates a noise increase, and potentially a bias on the mean energy for the 50 ns bunch spacing. The use of the correct pileup autocorrelation matrix is therefore strongly advised, especially for high expected pileup levels.

In regards of which OFC pileup optimization is more suitable, the introduction of the additional constrain for the baseline shift gives similar if not slightly worse noise reduction than the optimization for pileup only, that remains the suggested optimization for the EM calorimeters. Similar conclusion holds for the HEC.

In the FCAL calorimeters, however, the situation is more complicated. At 50 ns bunch spacing, the optimization including the additional constrain for the baseline shift seems to recover some the bias, and it is therefore suggested. At 25 ns bunch spacing, both OFC pileup optimizations are very close in performance, with that having the extra constraint being only slightly better. In this configuration, one should carefully weigh the advantage of a minimal noise improvements versus a simpler optimization scheme, and the uniformity of approach among all LAr calorimeters.



(a) (b) (c) (d) (e) (f)

Figure 8.4: Average cell energy $\langle E_{cell} \rangle$ as a function of η, reconstructed from ZB events at $\mu =$ 30, 60 and 90 (black, green and purple, respectively) in the EM part of the LAr calorimeters, for 25 ns (a)(c)(e) and 50 ns (b)(d)(f) BS configurations, reconstructed using different OFC optimizations and alternative pileup autocorrelation matrices (same BS as simulation - full points, different - empty points) . Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

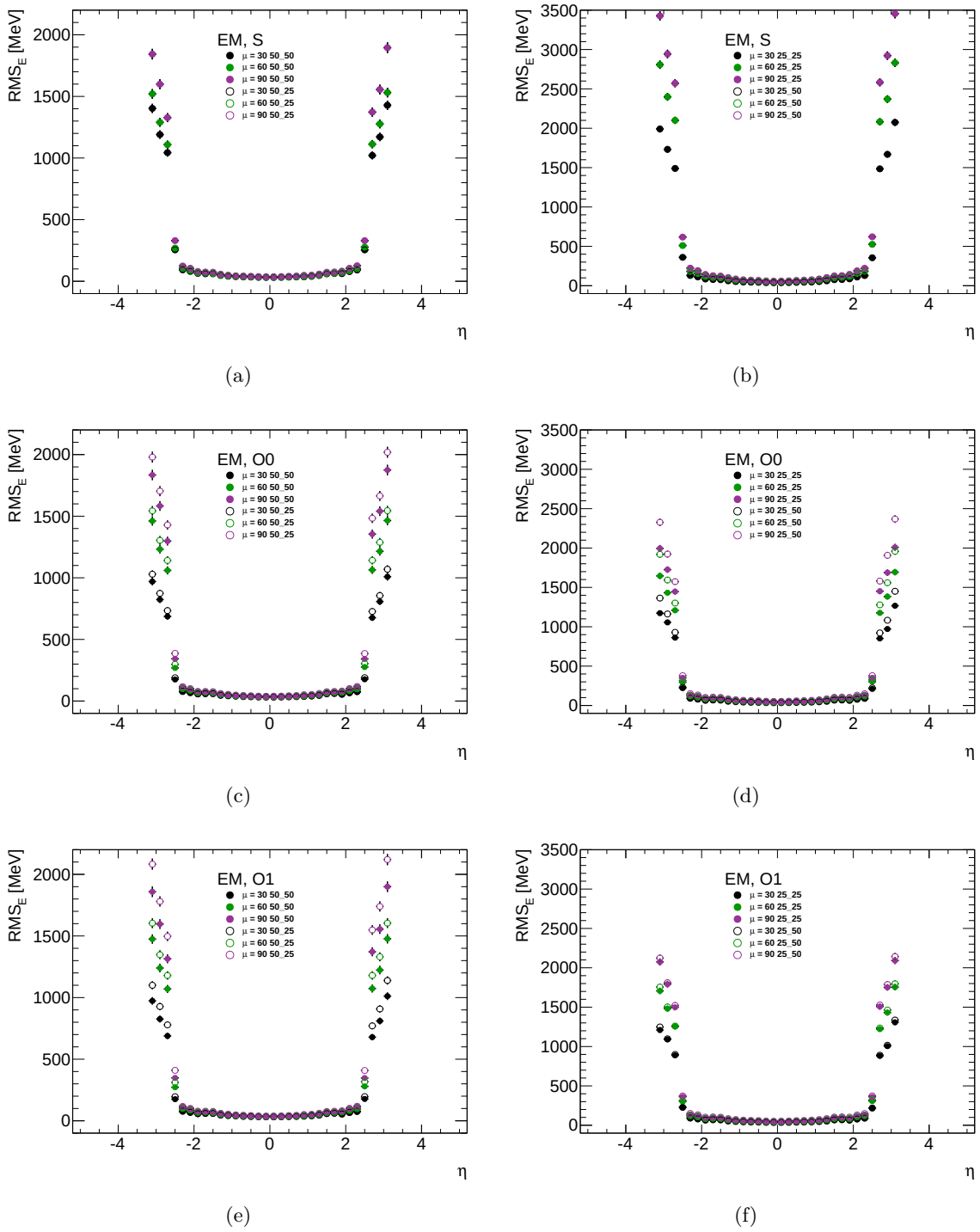


Figure 8.5: RMS of the cell energy as a function of η, reconstructed from ZB events at $\mu = 30$, 60 and 90 (black, green and purple, respectively) in the EM part of the LAr calorimeters, for 25 ns (a)(c)(e) and 50 ns (b)(d)(f) BS configurations, reconstructed using different OFC optimizations and alternative pileup autocorrelation matrices (same BS as simulation - full points, different - empty points). Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

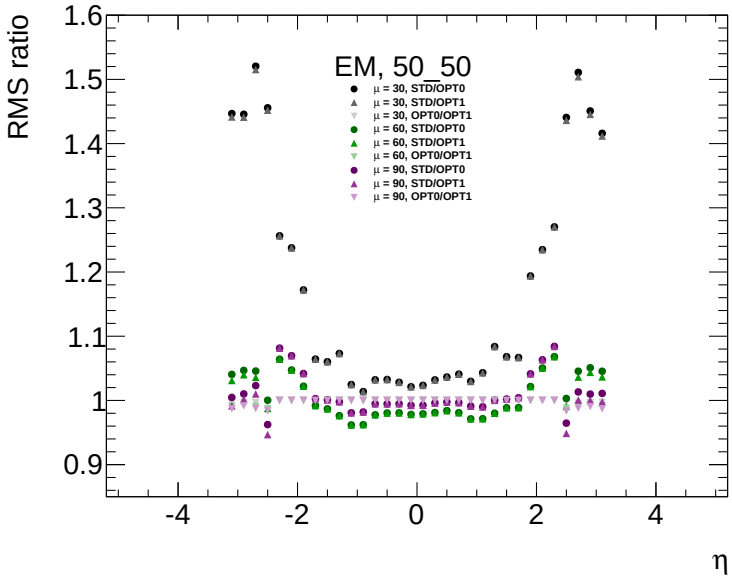


Figure 8.6: Ratio of the RMS of E_{cell} vs. η for the EM part of the calorimeter, reconstructed from ZB events at $\mu = 30$, 60 and 90 (shades of black and grey, green and purple, respectively) in the EM part of the LAr calorimeters for 50 ns BS configuration, reconstructed using different OFC optimizations. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

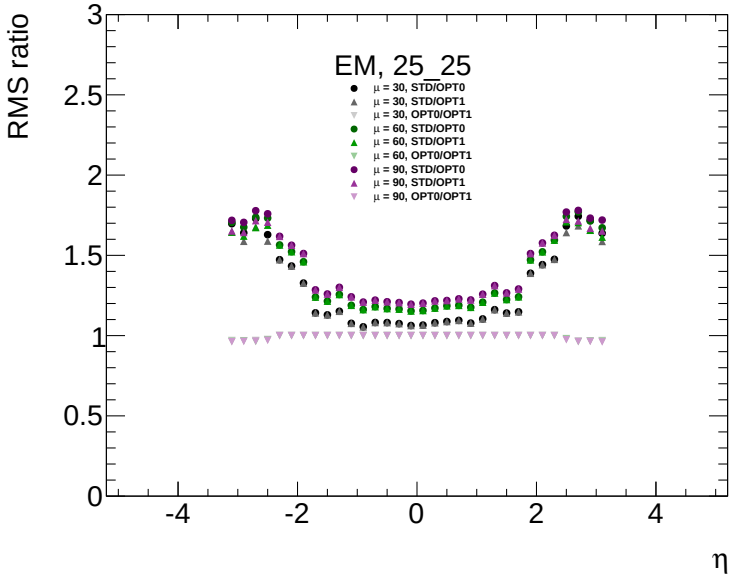


Figure 8.7: Ratio of the RMS of E_{cell} vs. η for the EM part of the calorimeter, reconstructed from ZB events at $\mu = 30$, 60 and 90 (shades of black and grey, green and purple, respectively) in the EM part of the LAr calorimeters for 25 ns bunch spacing configuration, reconstructed using different OFC optimizations. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

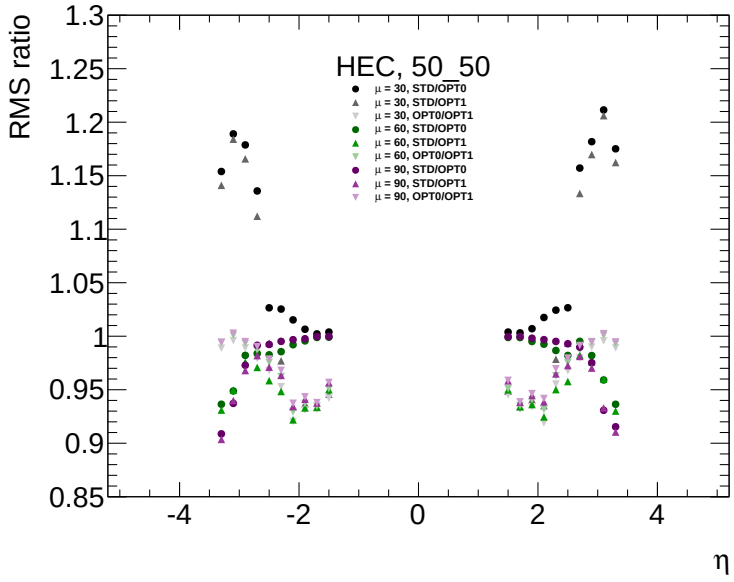


Figure 8.8: Ratio of the RMS of E_{cell} vs. η for the HEC part of the calorimeter, reconstructed from ZB events at $\mu = 30$, 60 and 90 (shades of black and grey, green and purple, respectively) in the HEC part of the LAr calorimeters for 50 ns bunch spacing configuration, as reconstructed using different OFC optimizations. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

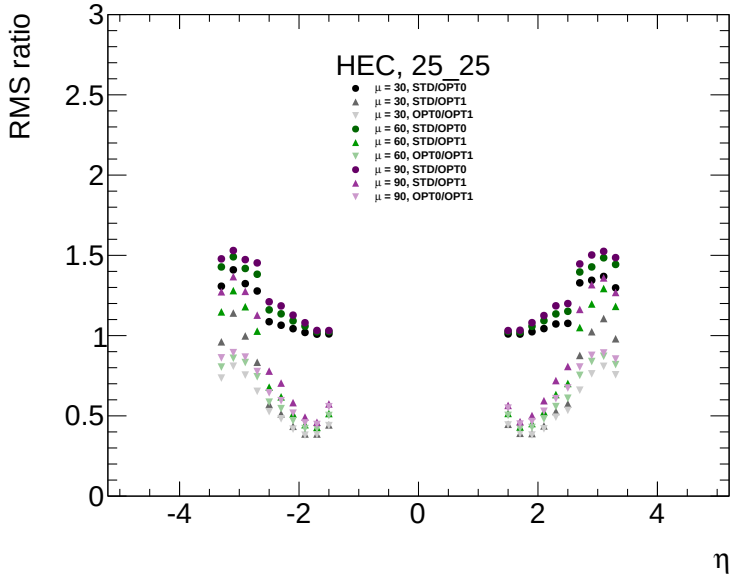


Figure 8.9: Ratio of the RMS of E_{cell} vs. η for the HEC part of the calorimeter, reconstructed from ZB events at $\mu = 30$, 60 and 90 (shades of black and grey, green and purple, respectively) in the HEC part of the LAr calorimeters for 25 ns bunch spacing configuration, reconstructed using different OFC optimizations. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

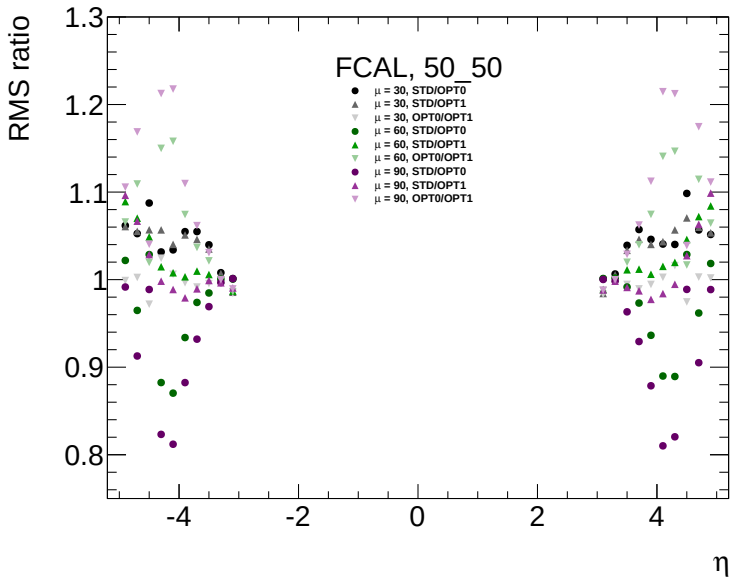


Figure 8.10: Ratio of the RMS of E_{cell} vs. η for the FCAL part of the calorimeter, reconstructed from ZB events at $\mu = 30$, 60 and 90 (shades of black and grey, green and purple) in the FCAL part of the LAr calorimeters for 50 ns bunch spacing configurations, as reconstructed using different OFC optimizations. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

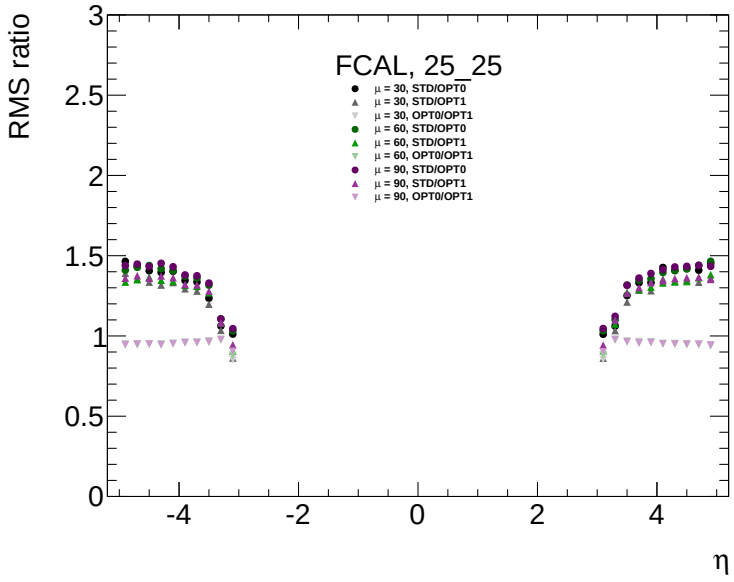


Figure 8.11: Ratio of the RMS of E_{cell} vs. η for the FCAL part of the calorimeter, as reconstructed from ZB events at $\mu = 30$, 60 and 90 (shades of black and grey, green and purple, respectively) in the FCAL part of the LAr calorimeters for 25 ns bunch spacing configurations, as reconstructed using different OFC optimizations. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

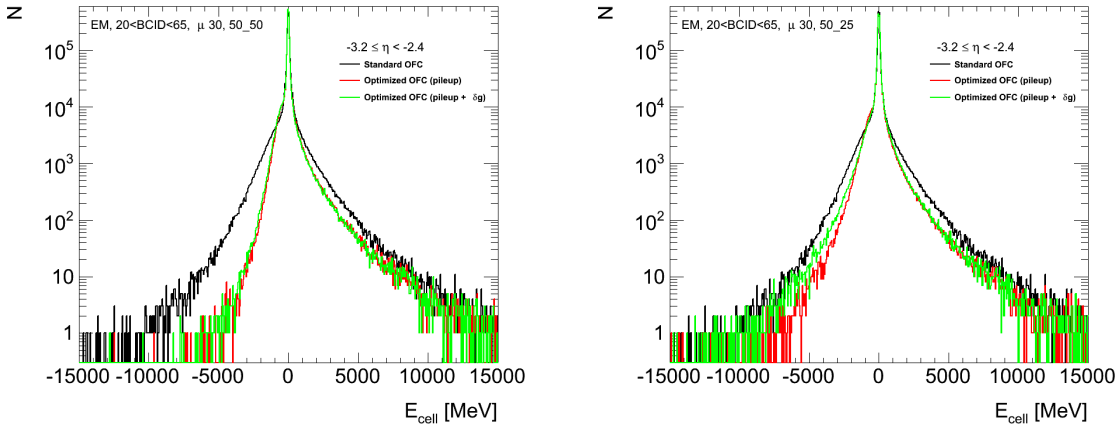


Figure 8.12: $\langle E_{cell} \rangle$ distribution for high η region $-3.2 \leq |\eta| < -2.4$ in the EM part of the calorimeter at $\mu = 30$ for different bunch spacing (a)(b) simulated with 50 ns BS and (c)(d) with 50 ns BS, reconstructed with OFCs optimized for 50 ns BS (a)(c) and 25 ns BS (b)(d) and three OFC reconstruction configurations (Standard optimization - black line, Optimization 0 - red line and Optimization 1 - green line).

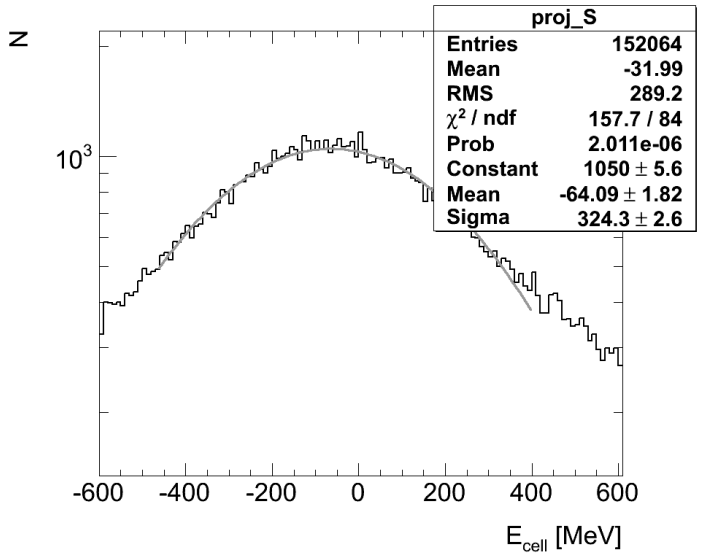
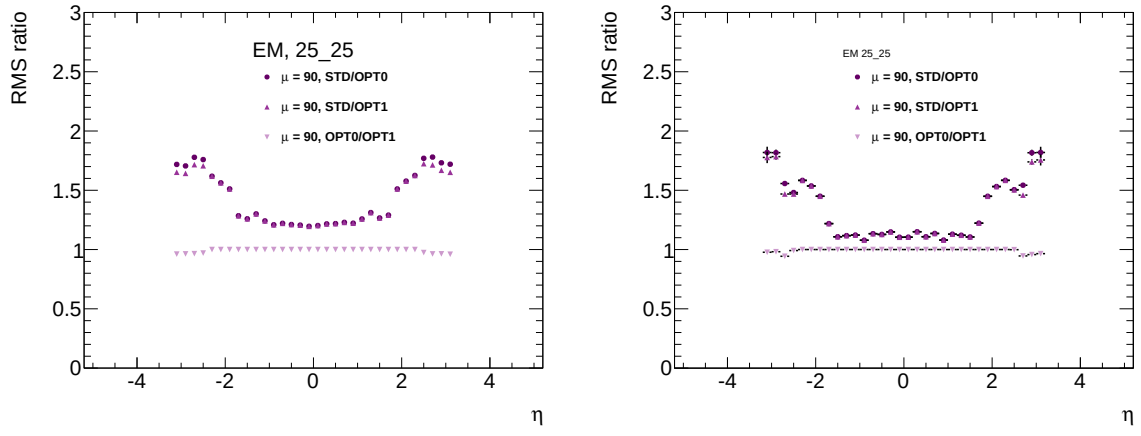


Figure 8.13: Cell energy E_{cell} distribution for the highest η bin for the EM part at $\mu = 30$, BS 25 ns simulated and reconstructed, using the Standard optimization method.



(a) (b)

Figure 8.14: Ratio of the RMS of E_{cell} as a function of η at $\mu = 90$ with bunch spacing 25 ns, reconstructed with OFC optimized using the correct pileup autocorrelation matrix. (a) values obtained from the simple mean and RMS of the distributions and (b) values obtained for the Gaussian fits of the distribution cores.

8.4 Noise vs. number of LAr samples

In this section, the evolution of the expected noise in the LAr cells is studied when the number of signal samples used for the OFC amplitude reconstruction is reduced from the five used in Run 1 operations to four and three. Following the results presented in Section 8.3, this study is only performed for the most convenient OFC reconstruction configurations (using the correct pileup autocorrelation matrix for a given bunch spacing configuration), and for a selection of representative LHC running conditions at $\sqrt{s} = 13$ TeV, namely:

- beam bunch spacing = 50 ns; $\mu = 90$;
- beam bunch spacing = 25 ns; $\mu = 60$;

The former running conditions are representative of what might occur in the beginning of the 2015 LHC operations, while the latter are more likely to represent the overall running conditions after the first period of consolidation[3].

In order to perform the studies, Zero Bias samples were simulated using the setup discussed in Section 8.2 for the two configurations listed above. These samples were then digitized and reconstructed using three LAr signal samples (the central three out for the usual five) and four LAr signal samples (the first four and the last four of the usual five). The reduced sample configurations are illustrated in Figure 8.15. The same noise studies discussed in Section 8.3 for the different pileup autocorrelation matrices are performed as a function of the signal sample configuration.

8.4.1 Bunch-spacing = 50 ns, $\mu = 90$

Figure 8.16 shows the cell energy distributions when a different number of LAr signal samples is used, for the three OFC optimizations in the EM part of the calorimeter. The noise levels, computed as the E_{cell} distribution RMS, are shown in Figure 8.17.

In order to evaluate the potential noise modification, a ratio of the RMS values is computed, by comparing the values obtained when using five samples to those computed when only three or four samples are selected. The comparison is done for the three different OFC optimizations, as shown in Figure 8.18. When using OFCs optimized only for electronic noise, a ≈10-20% noise increase when removing one or two samples is observed. When the OFC optimization for pileup is introduced, at low η the noise increase is up to ≈10-13%, while at high η it is only up by ≈5%, and even improves by up to ≈15%. In a situation in which the pileup noise reduction is suboptimal because of the imperfect cancellation of the out-of-time pulses, and the peculiar properties of noise autocorrelation, the reduction of number of samples helps to mitigate the pileup noise impact regardless of the optimization used.

The RMS ratios for the HEC are shown in Figure 8.19. The noise variation as a function of the number of samples is very similar to that observed in the EM calorimeters. When removing one or two samples for reconstruction and using the OFCs optimized only for electronic noise, the noise increase ranges from ≈7-20%. Similar increases are observed for OFCs optimized for

[3] After the study was performed, the predicted values of μ changed and LHC will most probably not produce pile-up with μ larger than 60. Therefore a study of BS= 50 ns with $\mu = 60$ and BS= 25 ns with $\mu = 45$ would have been more suitable.

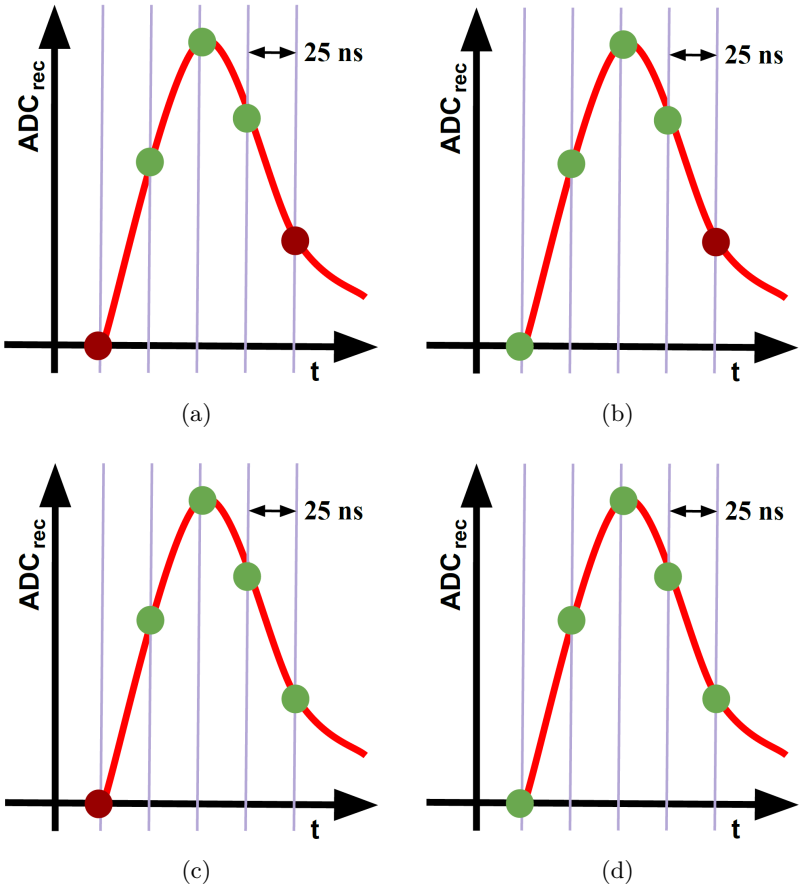


Figure 8.15: Illustration of different LAr signal samples configurations used in the noise study: (a) three samples (the central three out for the usual five); four samples (the first four (b) and the last four (c) of the usual five) and five samples (d).

pileup in the low η region. At high η, using the OFCs optimized only for pileup improves the noise levels up to $\approx$10%, regardless of the reduced number of samples.

FCAL RMS ratios are shown in Figure 8.20. When OFCs optimized only for electronic noise are used, the noise level changes by $\sim$ 5% when removing one or two samples. In this configuration, the total noise in the FCAL depends very little on the number of samples used: the effect is very likely associated to the narrowness of the FCAL pulses, and the larger OFC weight given to the central samples. When the optimization for pileup is introduced, removing one or two samples improves the noise levels by $\sim$ 37%.

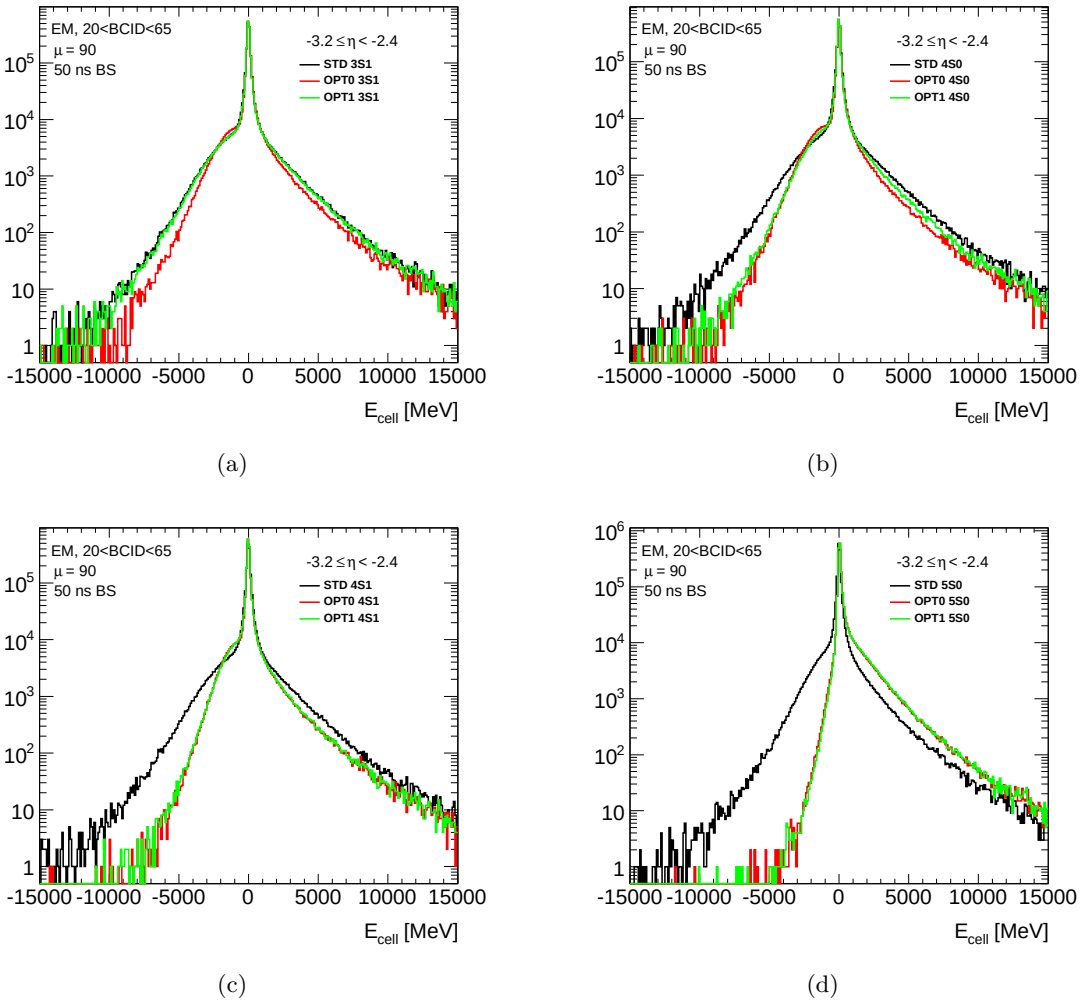


Figure 8.16: Cell energy E_{cell} distributions for a) three, b) first four, c) last four and d) five LAr samples used for the signal amplitude reconstruction, and for the three different OFC optimizations (Standard optimization - black line, Optimization 0 - red line, Optimization 1 - green line).

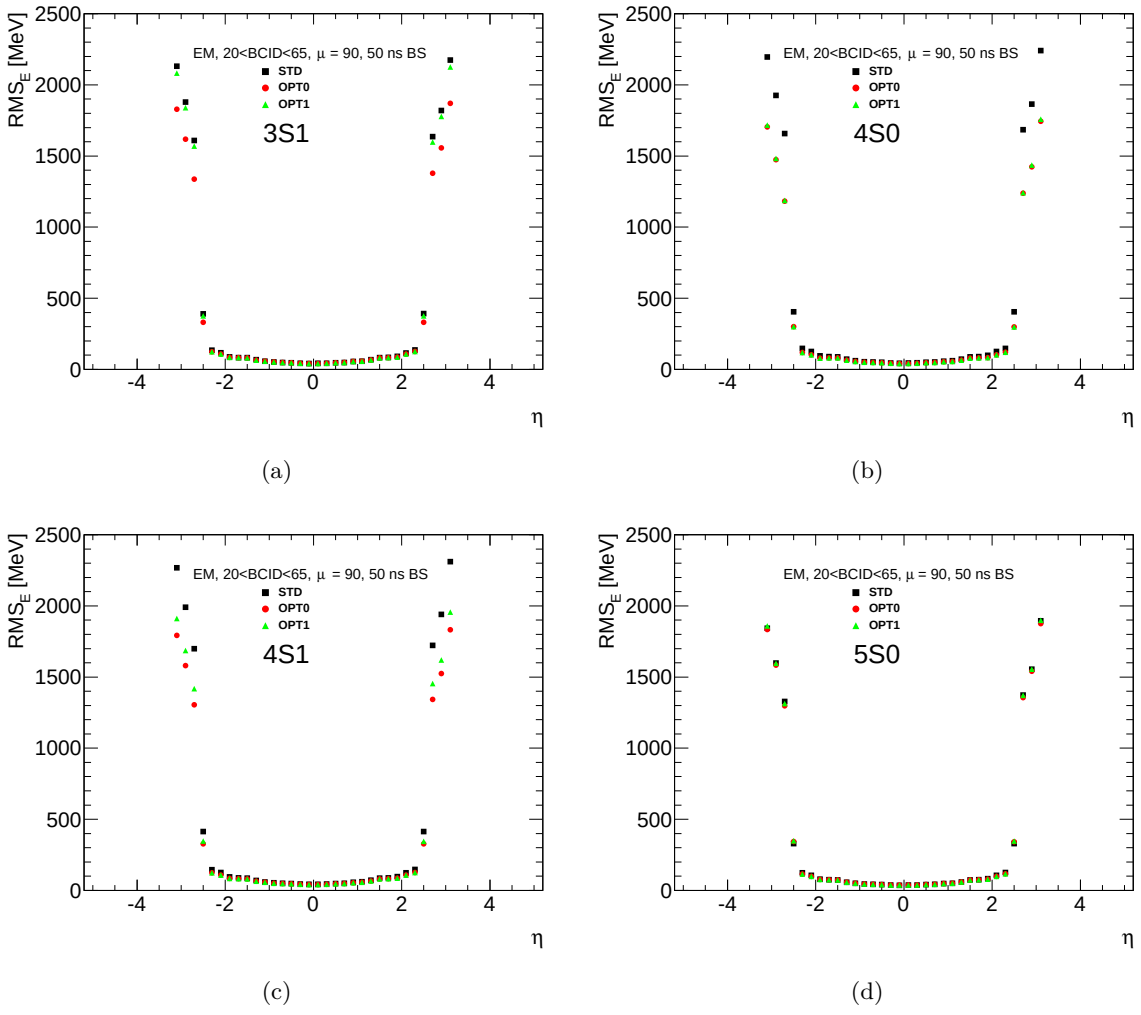


Figure 8.17: RMS of the E_{cell} vs. η when the central three samples (a), the first four samples (b), the last four samples (c) and the usual five samples used for the signal amplitude reconstruction, for the three different OFC optimizations (Standard optimization - black squares, Optimization 0 - red circles, Optimization 1 - green triangles.

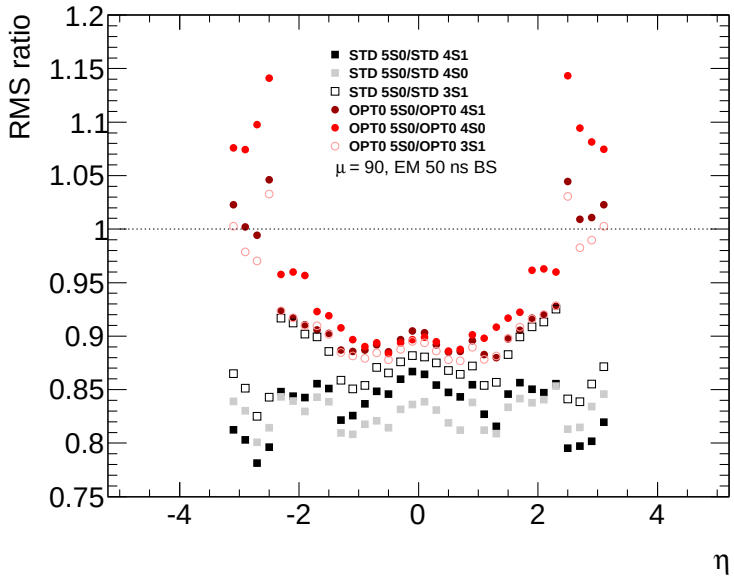


Figure 8.18: Ratio of the cell energy RMS as a function of η in the EM calorimeter, for different number of samples used in the signal amplitude reconstruction, and two different OFC optimizations (Standard optimization - shades of black and grey, Optimization 0 - shades of red). The values are obtained for a bunch spacing of 50 ns, and $\mu = 90$. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

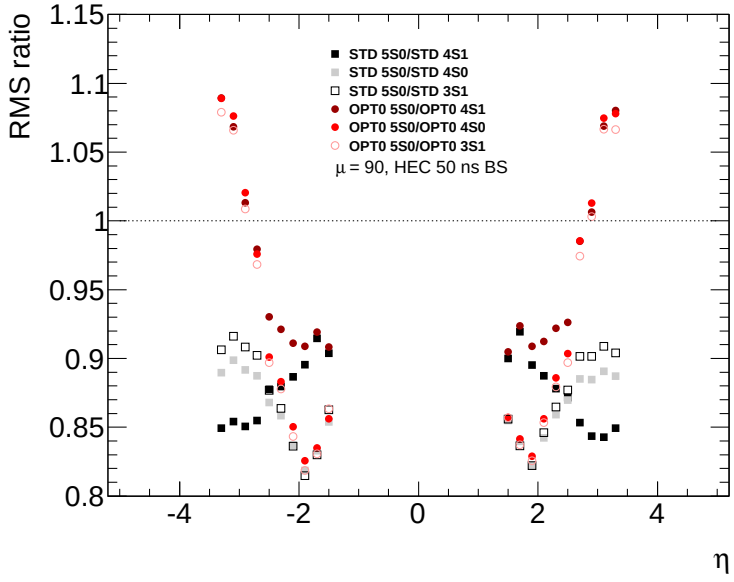


Figure 8.19: Ratio of the cell energy RMS as a function of η in the HEC calorimeter, for different number of samples used in the signal amplitude reconstruction, and two different OFC optimizations (Standard optimization - shades of black and grey, Optimization 0 - shades of red). The values are obtained for a bunch spacing of 50 ns, and $\mu = 90$. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

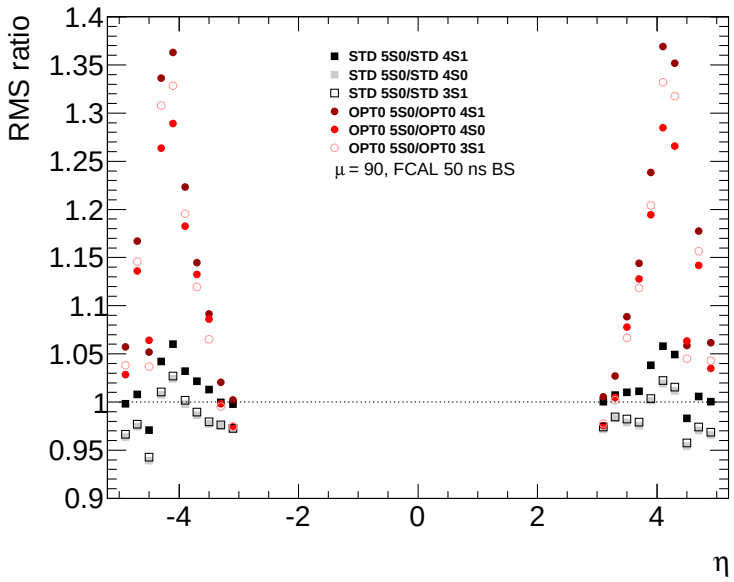


Figure 8.20: Ratio of the cell energy RMS as a function of η in the FCAL calorimeter, for different number of samples used in the signal amplitude reconstruction, and two different OFC optimizations (Standard optimization - shades of black and grey, Optimization 0 - shades of red). The values are obtained for a bunch spacing of 50 ns, and $\mu = 90$. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

8.4.2 Bunch-spacing = 25 ns, $\mu = 60$

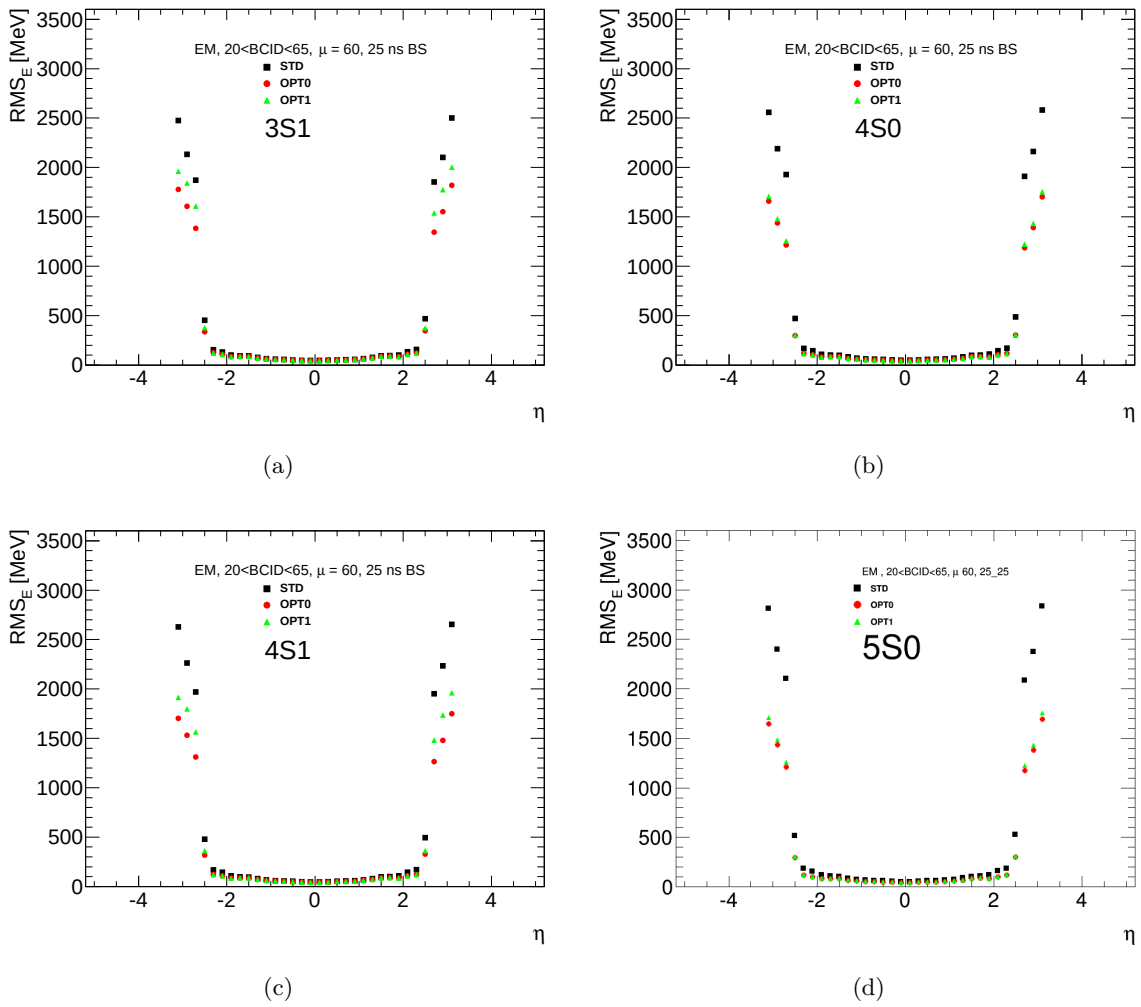


Figure 8.21: Cell energy distributions for a) three, b) first four, c) last four and d) five LAr samples used for the signal amplitude reconstruction, and for the three different OFC optimizations (Standard optimization - black squares, Optimization 0 - red circles, Optimization 1 - green triangles.

Figure 8.21 shows the noise levels, computed as the distribution RMS, for different number of signal samples and alternative OFC optimization, in the specific bunch spacing and pileup configuration discussed in this section.

The impact on the noise level of changing the number of signal samples used for the amplitude reconstruction for the EM part of the calorimeter is shown in Figure 8.22, where the E_{cell} RMS ratios are shown. When using OFCs optimized only for electronic noise, a ∼5-20% noise reduction is observed when removing one or two samples. When the OFC optimization for pileup is introduced, the noise levels are similar when the first four samples are used, while when the last four are selected the noise levels increased by up to ∼ 2-13%.

Results for the HEC are shown in Figure 8.23. When removing one or two samples for recon-

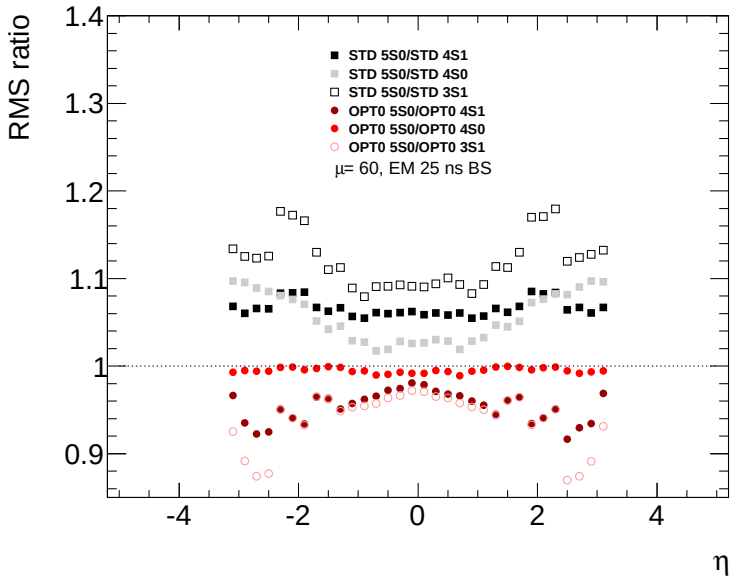


Figure 8.22: Ratio of the cell energy RMS as a function of η in the EM calorimeter, for different number of samples used in the signal amplitude reconstruction, and two different OFC optimizations (Standard optimization - shades of black and grey, Optimization 0 - shades of red). The values are obtained for a bunch spacing of 25 ns, and $\mu = 60$. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

struction and using the OFCs optimized only for the electronic noise, the noise increases by a few percent at low η but decreases by up to $\sim 15\%$ at high η. When OFC optimized for pileup are used and the first sample of the five is removed, the noise increases by up to $\sim 4\%$. For the first four and three sample configuration, the noise increases up to $\sim 13\%$.

Comparison of the noise levels in the FCAL calorimeter for five, four or three samples used for amplitude reconstruction is shown in Figure 8.24. When using the OFC optimized only for electronic noise, removing the first sample decreases the noise by 0.5%. The two OFC optimization for pileup increase the noise by up to $\sim 4\%$. In this case, removing the last sample increases the noise by up to $\sim 4\%$ at low η, but only around 1% at high η. When removing the first sample (last 4) and also the last sample (3 samples), the noise increases up to $\sim 12\%$ at high η.

8.4.3 Detector-specific choice of 4 signal samples

The evaluation of the choice of the number of samples was performed by making ratios of the RMS (noise) of the Zero Bias MC events reconstructed using the first four and the last four samples. The results are shown in Figure 8.25 for the EM part of the calorimeter. The HEC and FCAL part are shown in Figures 8.26 and 8.27. These ratios show that the optimal choice of samples used for LAr reconstruction depends on the calorimeter regions, bunch spacing and pile-up conditions. A summary of these observations is shown in Table 8.4.

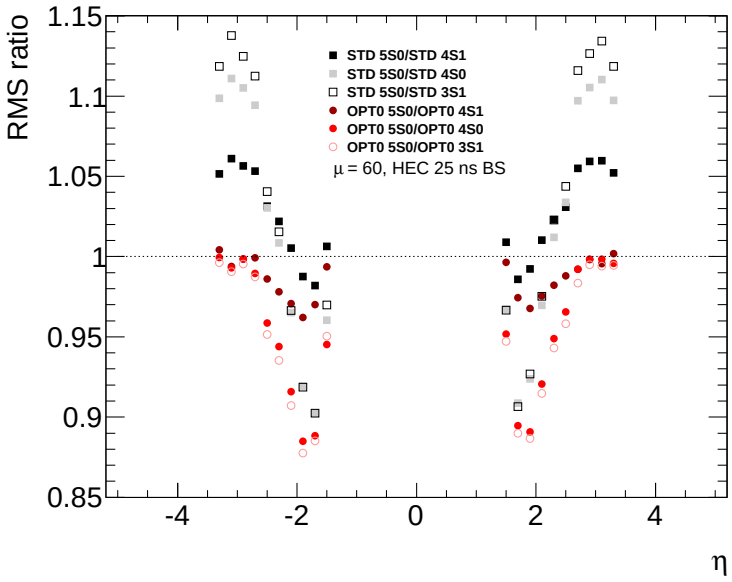


Figure 8.23: Ratio of the cell energy RMS as a function of η in the HEC calorimeter, for different number of samples used in the signal amplitude reconstruction, and two different OFC optimizations (Standard optimization - shades of black and grey, Optimization 0 - shades of red). The values are obtained for a bunch spacing of 25 ns, and $\mu = 60$. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.

OFC elec. noise only	**EM**	**HEC**	**FCAL**
$\mu = 90$, 50 ns	last 4	equal	last 4
$\mu = 60$, 25 ns	last 4	equal	last 4
OFC pile-up	**EM**	**HEC**	**FCAL**
$\mu = 90$, 50 ns	first 4	last 4	last 4
$\mu = 60$, 25 ns	first 4	last 4	first 4
OFC pile-up + baseline	**EM**	**HEC**	**FCAL**
$\mu = 90$, 50 ns	last 4	last 4	first 4
$\mu = 60$, 25 ns	equal	last 4	last 4

Table 8.4: Summary of recommended choices of samples used for LAr reconstruction for different bunch spacing and pile-up conditions.

8.5 Conclusion

Noise properties in the LAr calorimeter cells are studied for the different level of pileup, corresponding to the expected LHC running conditions in 2015. Alternative combinations of bunch spacing configurations (50 ns, 25 ns) and of average number of interactions per bunch crossing μ (30, 60, 90) are investigated, and the impact of different Optimal Filtering Coefficients optimization strategies is explored.

Regardless of the pileup level, it is always better to use the exact collision bunch spacing to

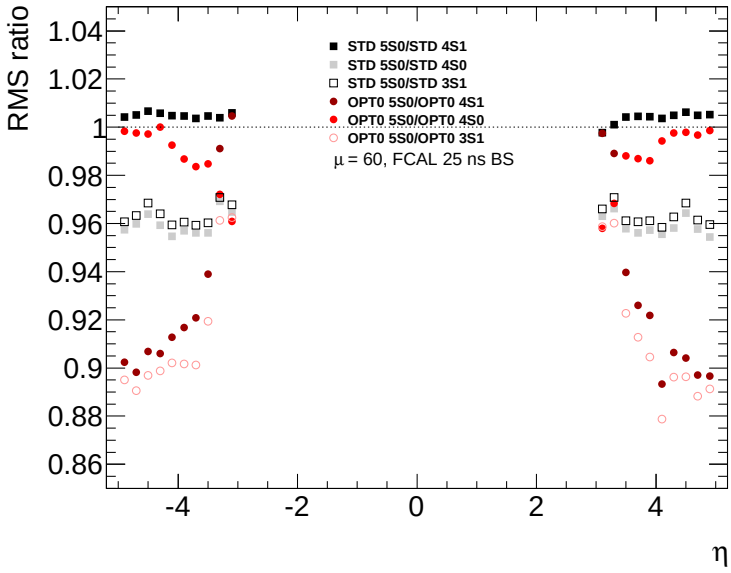
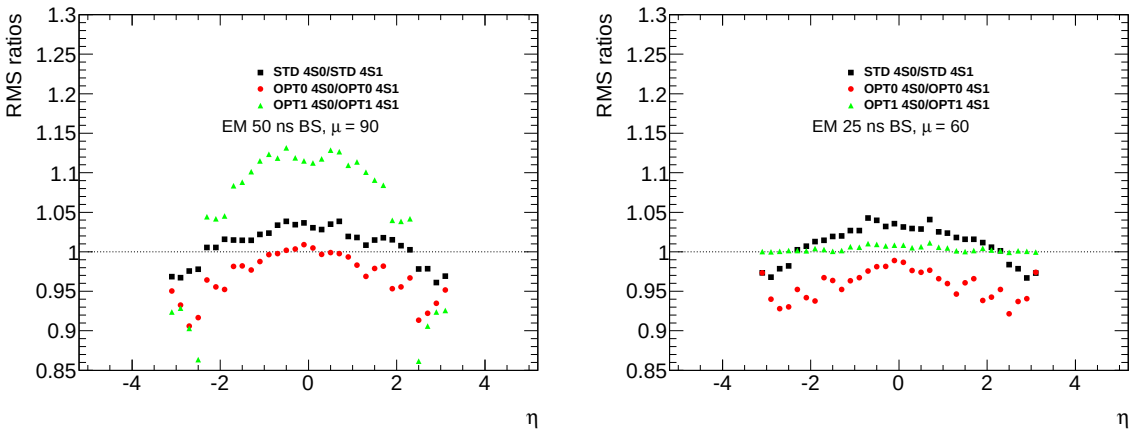


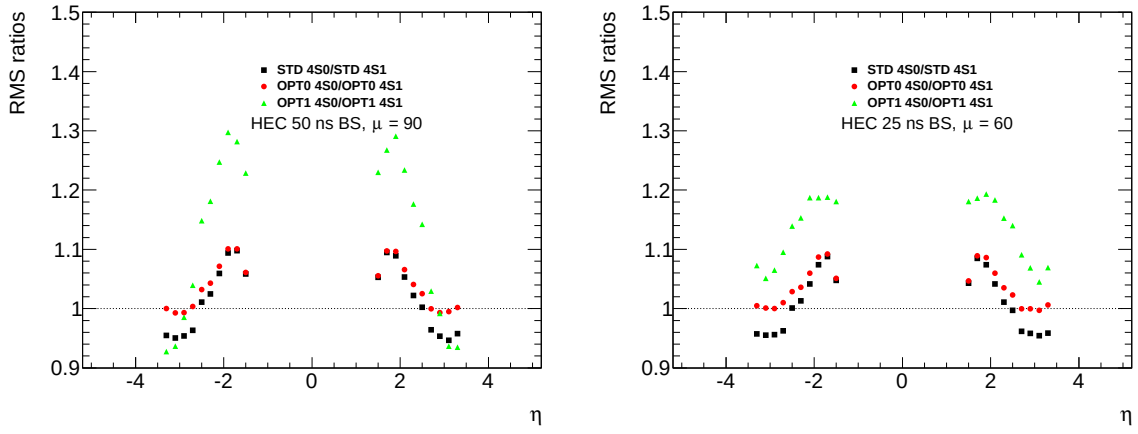
Figure 8.24: Ratio of the cell energy RMS as a function of η in the FCAL calorimeter, for different number of samples used in the signal amplitude reconstruction, and two different OFC optimizations (Standard optimization - shades of black and grey, Optimization 0 - shades of red). The values are obtained for a bunch spacing of 25 ns, and $\mu = 60$. Only events with $20 < \mathrm{mod}(BCID, 80) < 65$ are used.



(a) (b)

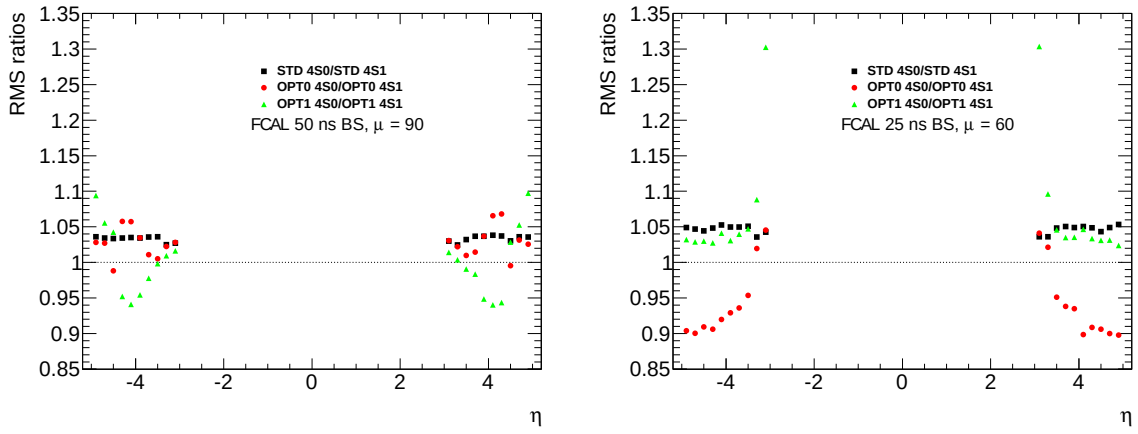
Figure 8.25: Ratios of the RMS of ZB MC events reconstructed using first four (4S0) and last four (4S1) samples for the EM part of the calorimeter for (a) 50 ns bunch spacing and $\mu = 90$ and (b) 25 ns bunch spacing and $\mu = 60$. Both shown for the three different OFC optimizations (Standard optimization - black squares, Optimization 0 - red circles, Optimization 1 - green triangles).

estimate the pileup autocorrelation matrix used to compute the OFC. When the wrong bunch spacing is used (e.g. the hypothesis of 50 ns bunch spacing is used to compute the OFC, that



(a) (b)

Figure 8.26: Ratios of the RMS of ZB MC events reconstructed using first four (4S0) and last four (4S1) samples for the HEC part of the calorimeter for (a) 50 ns bunch spacing and $\mu = 90$ and (b) 25 ns bunch spacing and $\mu = 60$. Both shown for three different OFC optimizations (Standard optimization - black squares, Optimization 0 - red circles, Optimization 1 - green triangles).



(a) (b)

Figure 8.27: Ratios of the RMS of ZB MC events reconstructed using first four (4S0) and last four (4S1) samples for the FCAL part of the calorimeter for (a) 50 ns bunch spacing and $\mu = 90$ and (b) 25 ns bunch spacing and $\mu = 60$. Both shown for three OFC different optimizations (Standard optimization - black squares, Optimization 0 - red circles, Optimization 1 - green triangles).

are then applied to data with 25 ns bunch spacing), the resulting noise level is always higher than that obtained with OFC using the correct pileup time autocorrelation. This is especially true if OFC corresponding to 50 ns bunch spacing are used to reconstruct data with 25 ns bunch spacing. Since the bulk of the LHC operations is supposed to use 25 ns bunch spacing, we urge the use of the correctly-optimized OFC.

Overall, the noise performance of the LAr calorimeters is always better when dealing with data at 25 ns bunch spacing, thanks to the better cancellation of the LAr bipolar pulses.

The impact of using a number of signal samples smaller than the five used during the LHC Run 1 operations (e.g. three or four) is also evaluated. When reducing the number of samples, a 10-15% increase of the noise RMS is observed as expected. When four samples are used, the change in performance between the first four (out of the original five) or the last four is studied. The choice happens to be strongly dependent on the calorimeter region, because of the different pulse width properties. The optimal position of these four samples (e.g. avoiding to have one sample at the signal peak, and having the two central ones shifted around the peak itself) was not studied and such fine-tuning would need a dedicated study.

The recommendations provided in this study were adopted by ATLAS for Run 2 data taking.

Conclusions and prospects

This thesis shows the use of photons in the ATLAS detector from the calibration studies, presented in Chapter 5, through the search for new resonances in the diphoton channel, presented in Chapter 6 and a measurement of the diphoton production cross section in association with jets, presented in Chapter 7. In addition, calorimeter studies with Run2 LHC running conditions in mind are presented in Chapter 8.

The e/gamma calibration effort in ATLAS is ongoing for the $\sqrt{s} = 13$ TeV, where the emphasis in the photon case will be put on further reducing the systematic uncertainties on the photon and jet energy scale and resolution. The precise measurement of the Higgs boson mass and properties will be important when enough data is collected as it can be 'rediscovered' with the Run2 data. Due to the increase of the c.m. energy, the back layer of the calorimeter will need to be intercalibrated with the other layers to avoid potential biases when studying high-p_{T} electrons and photons.

The search for additional resonances decaying into two photons, presented in Chapter 6, and the obtained limit on fiducial cross section times branching fraction of such a resonance, excludes parts of parameter space of some 2HDM and SUSY models, as described in Ref. [255, 256, 257]. But since there are many unknown parameters in each of those models, it is impossible to fully exclude any of them.

Since the branching fraction of an additional diphoton resonance may be increased by a factor of two or even more at $\sqrt{s} = 13$ TeV, the possibility of a discovery in this channel is still there, even with the first 10 fb^{-1} of data, which should be collected by summer of 2016.

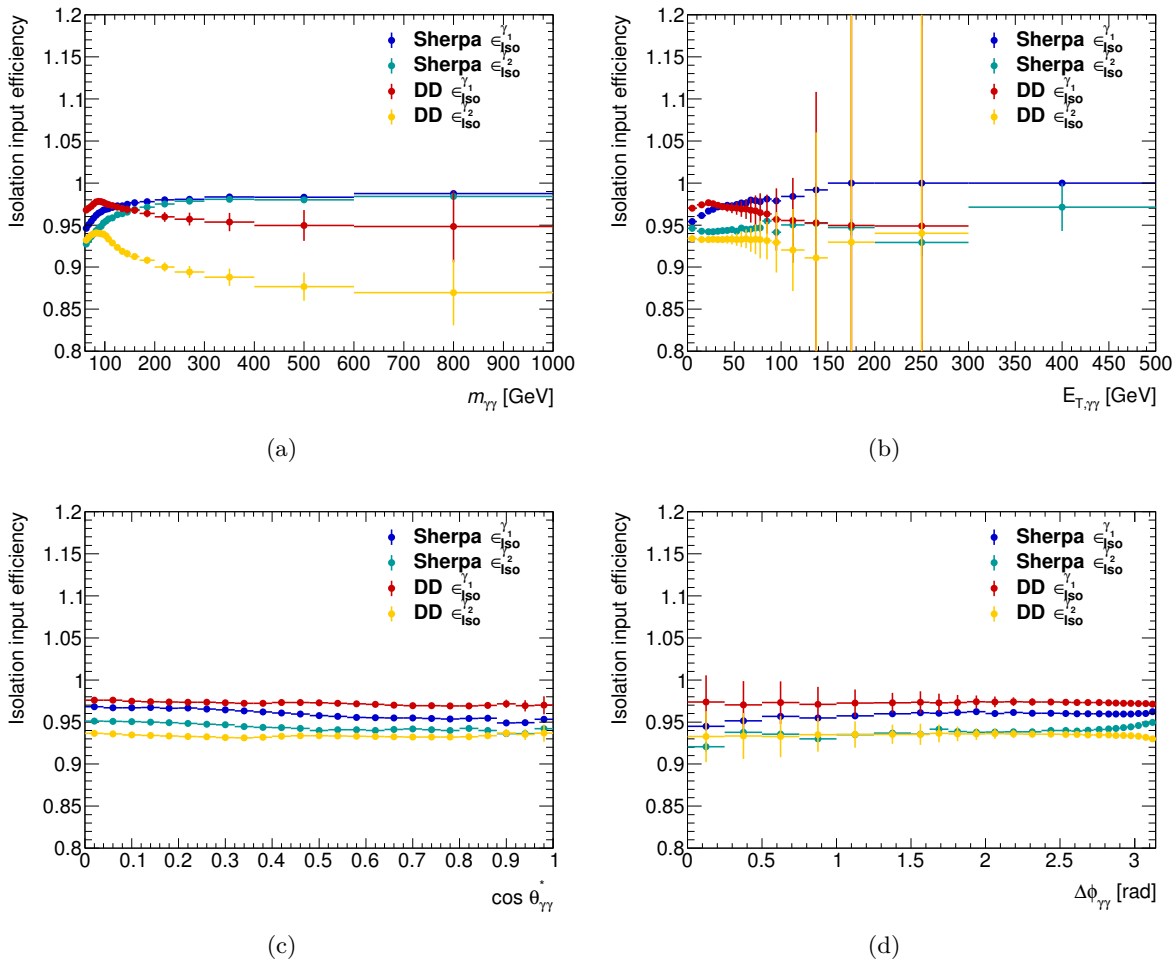
A search using the early 2015 data collected at $\sqrt{s} = 13$ TeV searching for high-mass ($m_X >$ 150 GeV) resonances in the diphoton channel is already implemented and in approval stages within the collaboration. The plan to repeat the low-mass analysis is underway as well, albeit slightly more complicated due to the presence of the Drell-Yan background.

Considering the SM measurement of the differential production cross sections of diphotons in association with jets, the aim is to finish the comparison with the NLO predictions and therefore give a comprehensive evaluation of SM NLO predictions of diphotons produced in association with jets. Any significant discrepancies in these comparisons, need to be investigated, mainly to improve the MC predictions for low-signal searches at the LHC.

A

Photon isolation efficiencies

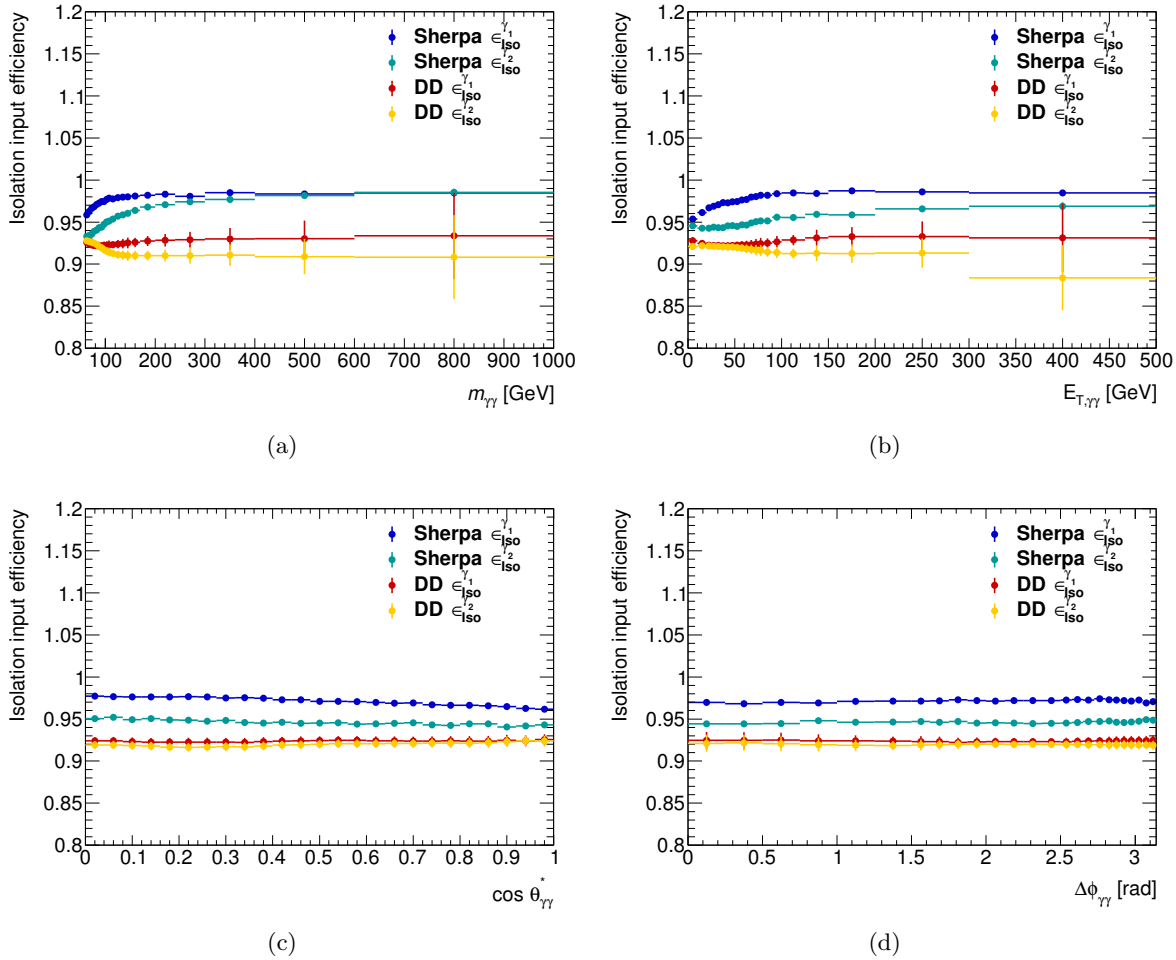
This Appendix contains plots of the photon isolation efficiencies from MC and the data driven isolation efficiencies, used as input for the 2x2DSB method for the leading and subleading photons separately, as described in Section 7.3.1.1, as a function of all observables except the ones already shown in the main text.



(a) (b)

(c) (d)

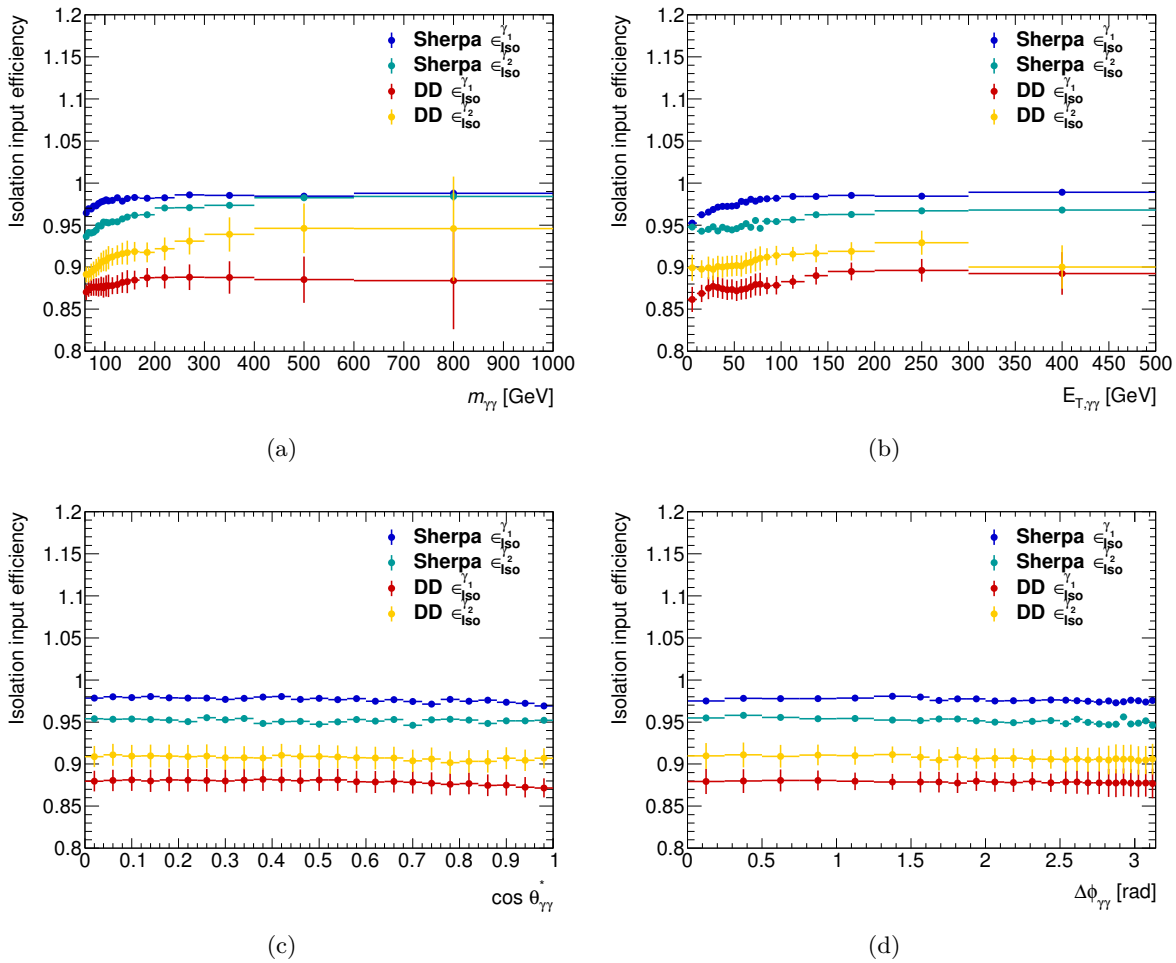
Figure A.1: Data driven isolation efficiencies of the leading (red) and subleading (yellow) photon candidate, extracted as described in the text and SHERPA MC efficiencies for the leading (dark blue) and subleading (light blue) candidate as a function of the diphoton observables for the $0-jet$ category.



(a) (b)

(c) (d)

Figure A.2: Data driven isolation efficiencies of the leading (red) and subleading (yellow) photon candidate, extracted as described in the text and SHERPA MC efficiencies for the leading (dark blue) and subleading (light blue) candidate as a function of the diphoton observables for the $1-jet$ category.



(a) (b)

(c) (d)

Figure A.3: Data driven isolation efficiencies of the leading (red) and subleading (yellow) photon candidate, extracted as described in the text and SHERPA MC efficiencies for the leading (dark blue) and subleading (light blue) candidate as a function of the diphoton observables for the $\geq 3-jet$ category.

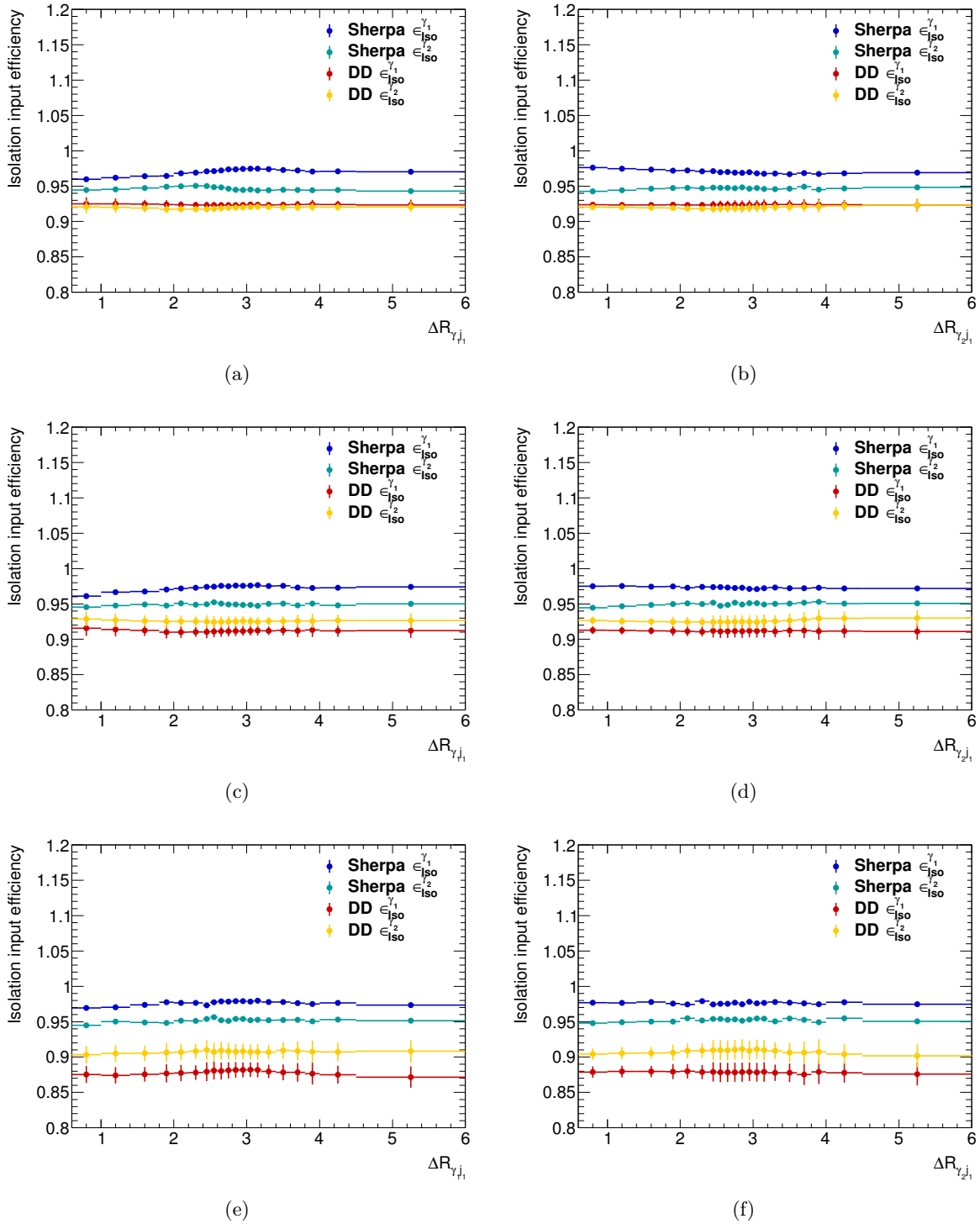
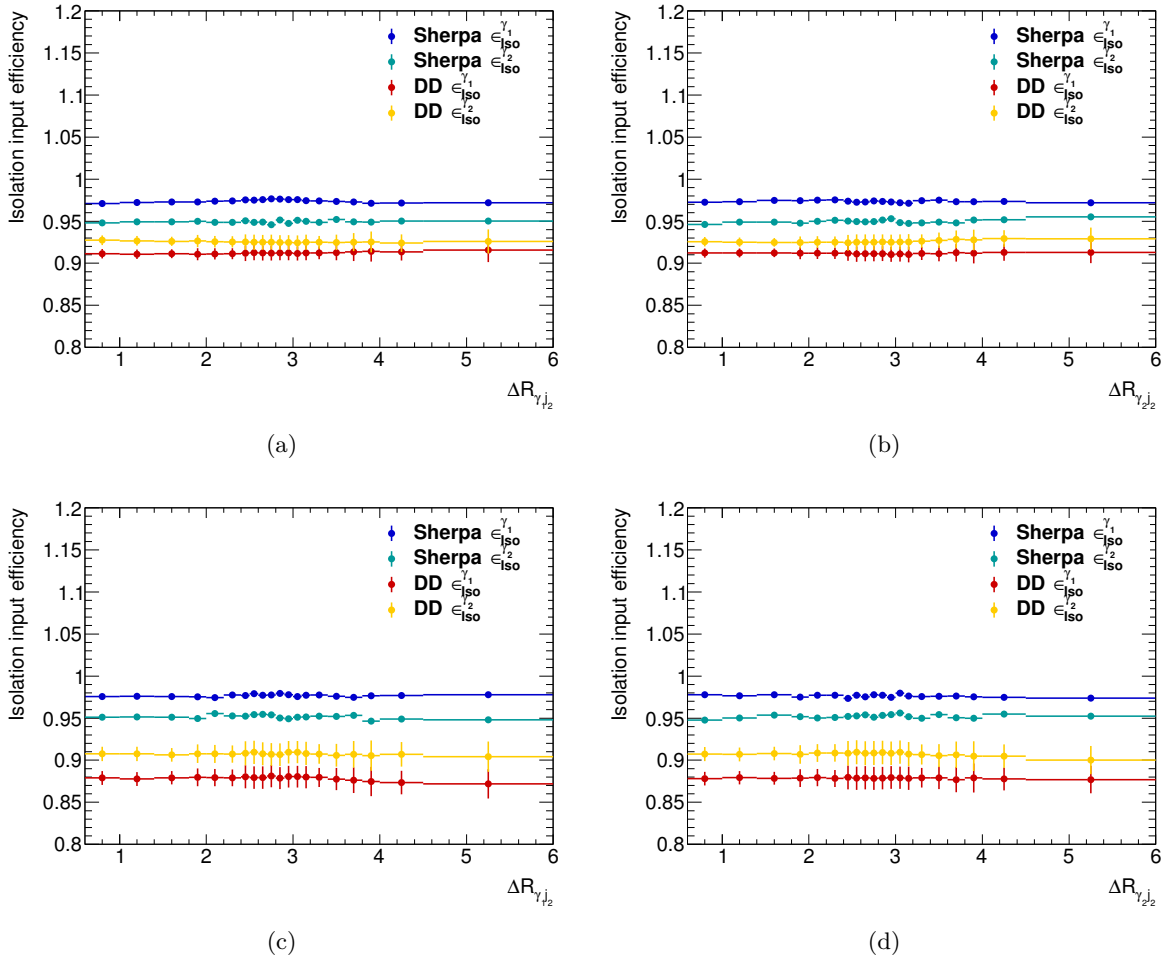


Figure A.4: Data driven isolation efficiencies of the leading (red) and subleading (yellow) photon candidate, extracted as described in the text and SHERPA MC efficiencies for the leading (dark blue) and subleading (light blue) candidate as a function of the R-separation between the leading photon and the leading jet (left) and the subleading photon and the leading jet (right) for the $1-jet$ (top), $2-jet$ (middle) and $\geq 3-jet$ (bottom) category.



(a) (b)

(c) (d)

Figure A.5: Data driven isolation efficiencies of the leading (red) and subleading (yellow) photon candidate, extracted as described in the text and SHERPA MC efficiencies for the leading (dark blue) and subleading (light blue) candidate as a function of the R-separation between the leading photon and the subleading jet (left) and the subleading photon and the subleading jet (right) for the $2-jet$ (top) and $\geq 3-jet$ (bottom) category.

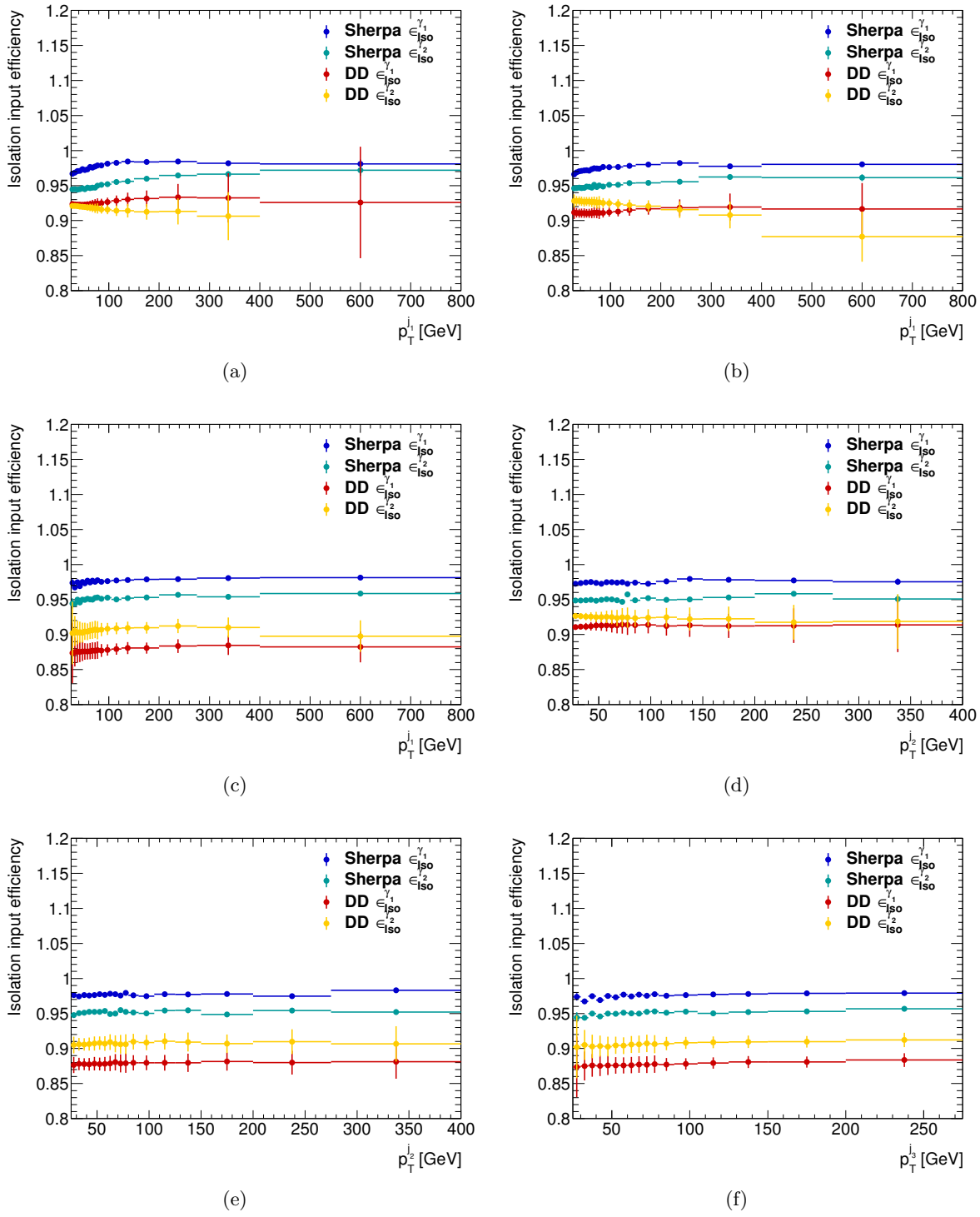
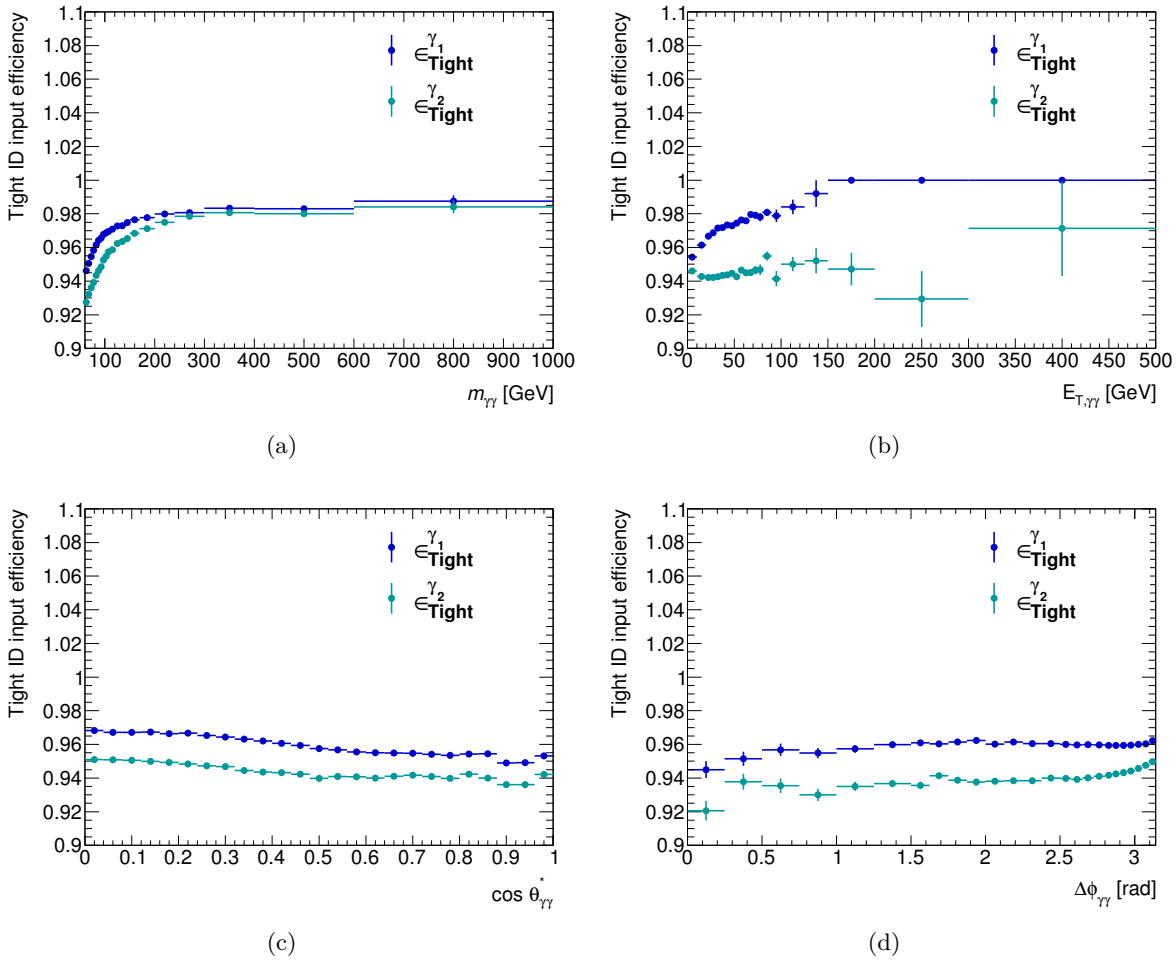


Figure A.6: Data driven isolation efficiencies of the leading (red) and subleading (yellow) photon candidate, extracted as described in the text and SHERPA MC efficiencies for the leading (dark blue) and subleading (light blue) candidate as a function of the p_{T}^{jet}. (a) $1-jet$, (b),(d) $2-jet$, (c),(e),(f) $\geq 3-jet$ category.

B

Photon *tight* ID efficiencies

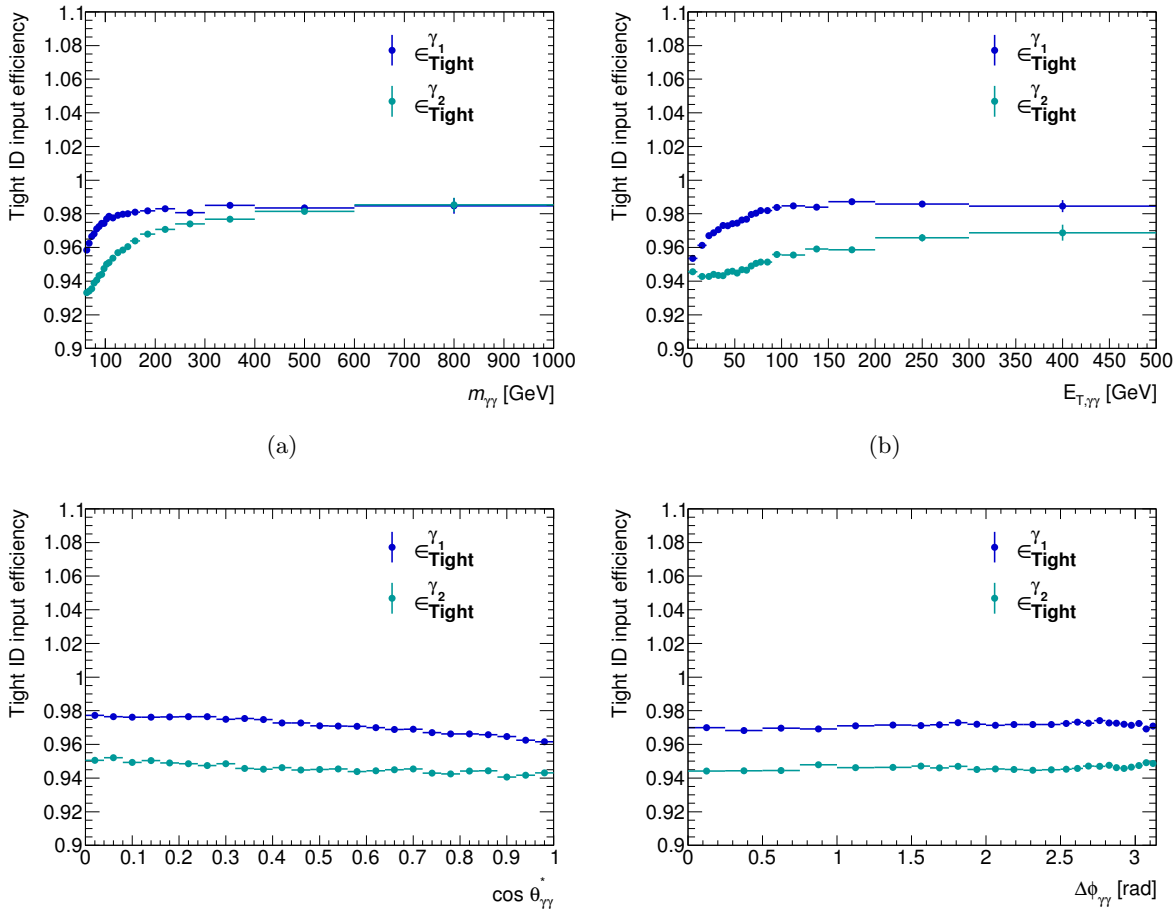
This Appendix contains plots of the photon *tight* ID efficiencies from Sherpa MC, used as input for the 2x2DSB method for the leading and subleading photons separately, as described in Section 7.3.1.2, as a function of all observables except the ones already shown in the main text.



(a) (b)

(c) (d)

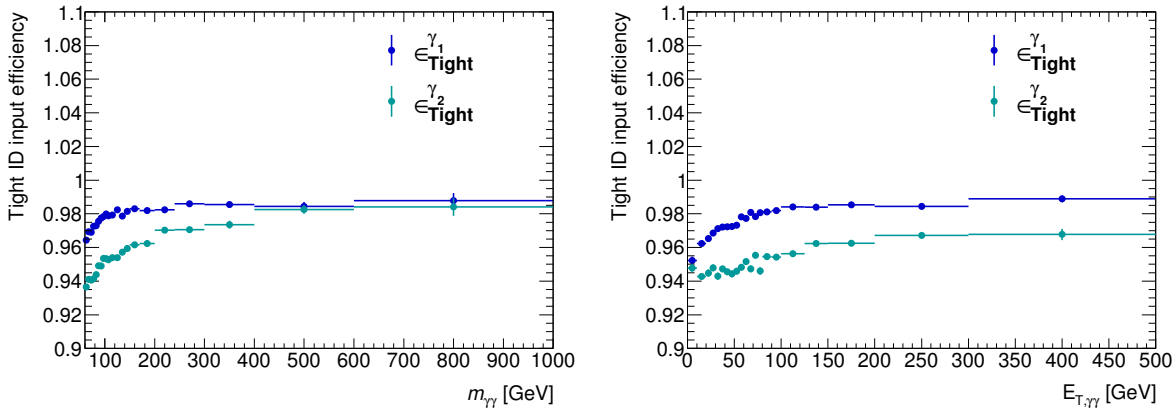
Figure B.1: Input *tight* ID efficiency of the leading (dark blue) and subleading (light blue) photon as a function of the diphoton observables for the $0-jet$ category.



(a) (b)

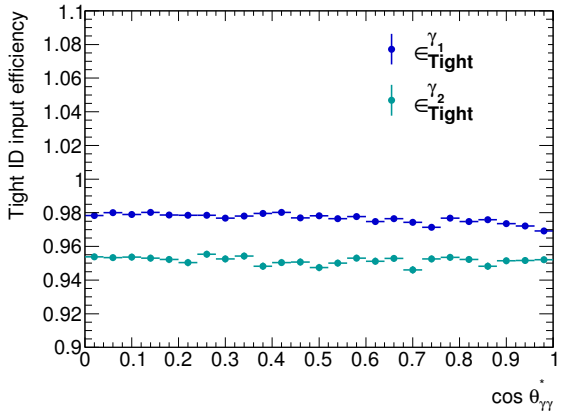
(c) (d)

Figure B.2: Input *tight* ID efficiency of the leading (dark blue) and subleading (light blue) photon as a function of the diphoton observables for the $1-$*jet* category.

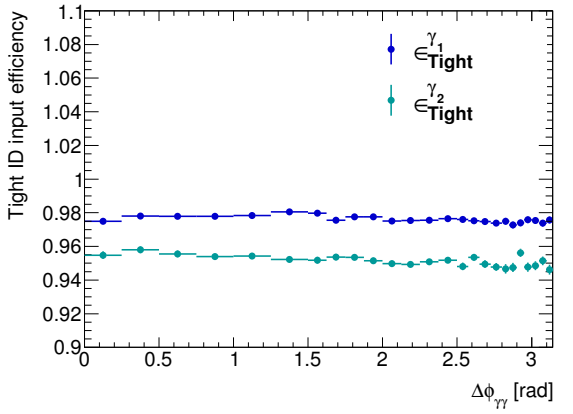


(a)

(b)



(c)



(d)

Figure B.3: Input *tight* ID efficiency of the leading (dark blue) and subleading (light blue) photon as a function of the diphoton observables for the $\geq 3-jet$ category.

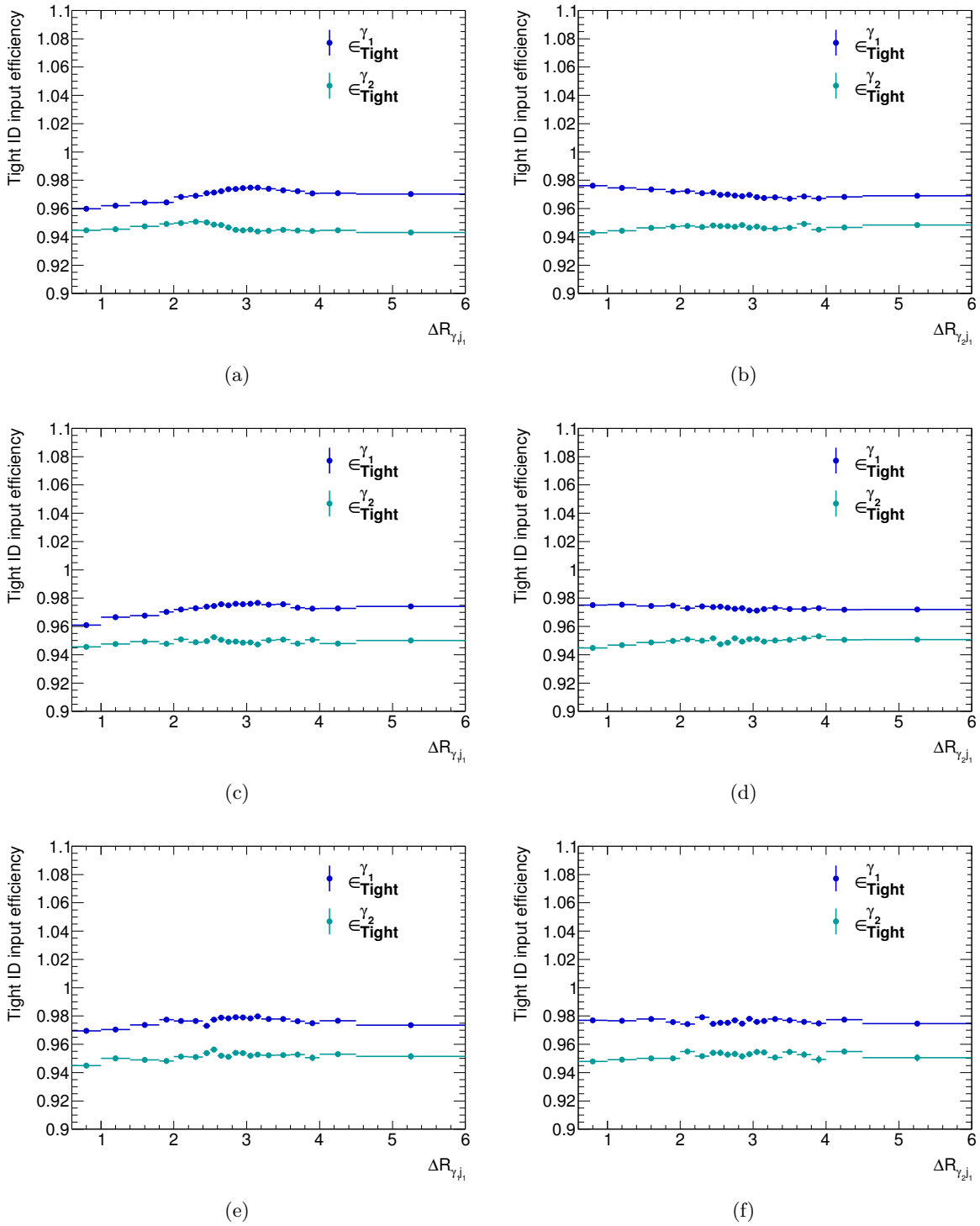
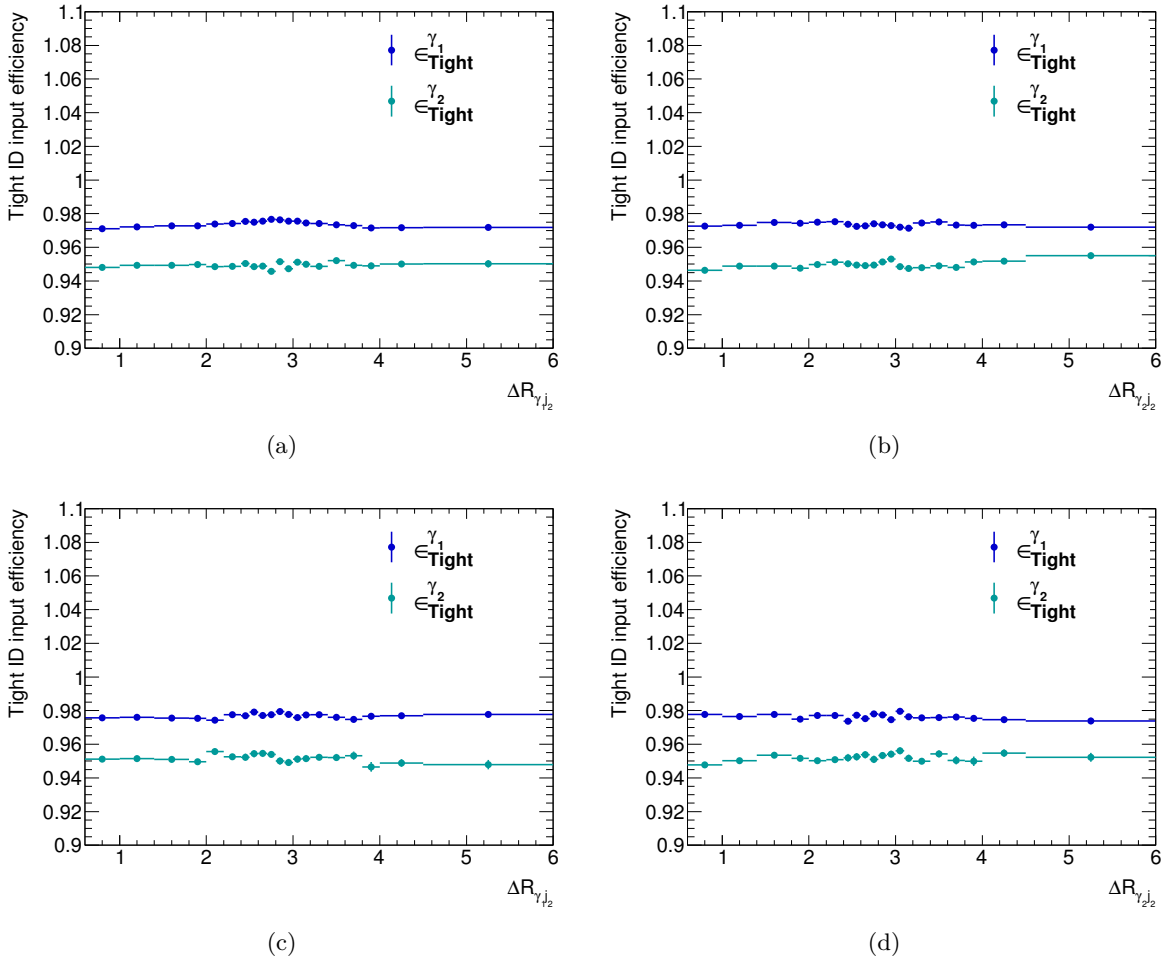


Figure B.4: Input *tight* ID efficiency of the leading (dark blue) and subleading (light blue) photon as a function of the R-separation between the leading photon and the leading jet (left) and the subleading photon and the leading jet (right) for the $1-jet$ (top), $2-jet$ (middle) and $\geq 3-jet$ (bottom) category.



(a) (b)

(c) (d)

Figure B.5: Input *tight* ID efficiency of the leading (dark blue) and subleading (light blue) photon as a function of the R-separation between the leading photon and the subleading jet (left) and the subleading photon and the subleading jet (right) for the $2-jet$ (top) and $\geq 3-jet$ (bottom) category.

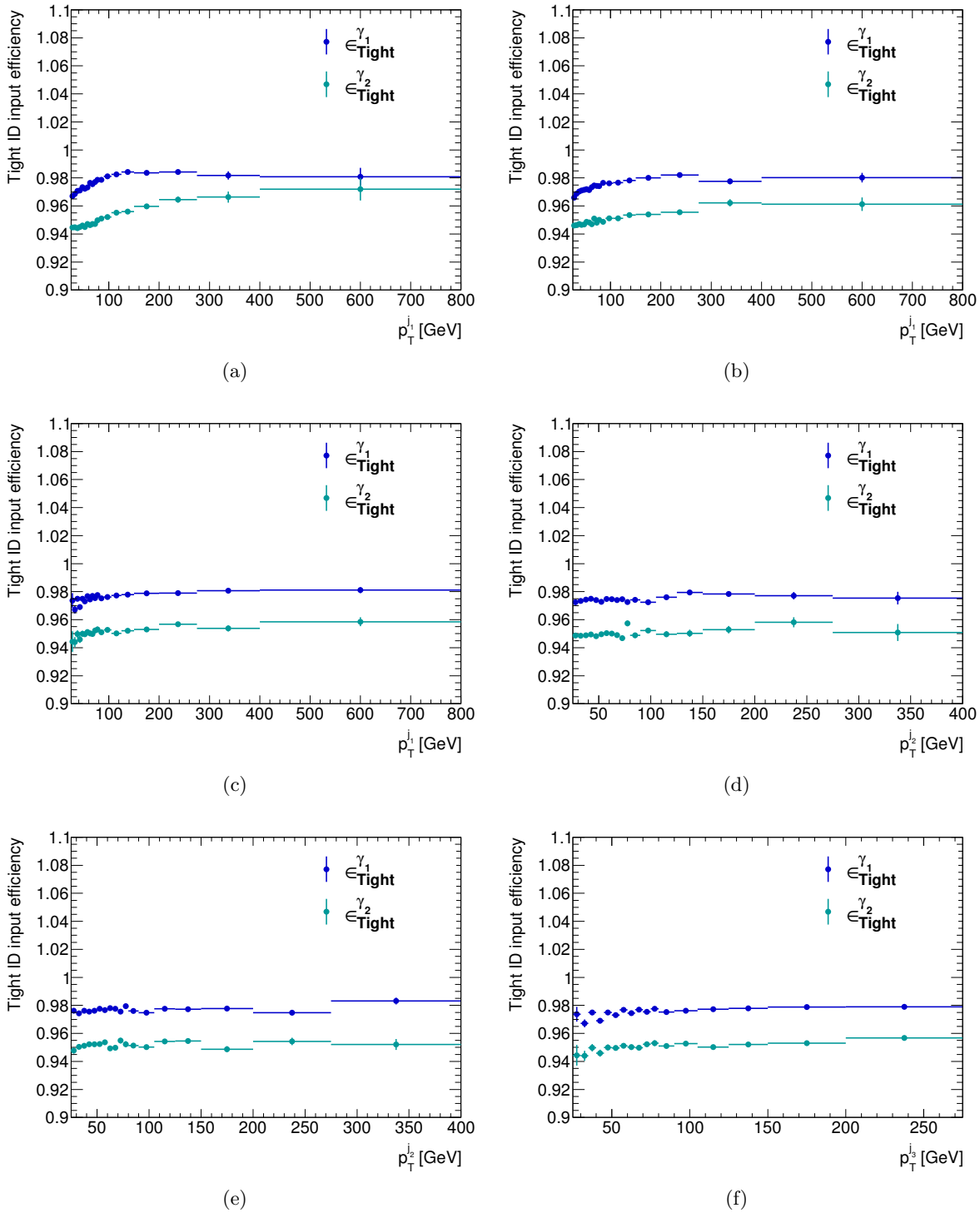


Figure B.6: Input *tight* ID efficiency of the leading (dark blue) and subleading (light blue) photon as a function of the p_{T}^{jet}. (a) $1-jet$, (b),(d) $2-jet$, (c),(e),(f) $\geq 3-jet$ category.

C

Breakdown of the systematic uncertainties of the 2x2DSB method

This Appendix contains plots of the breakdown of the systematic uncertainties on the 2x2DSB method as a function of all observables except the ones already shown in the main text in Section 7.3.2.

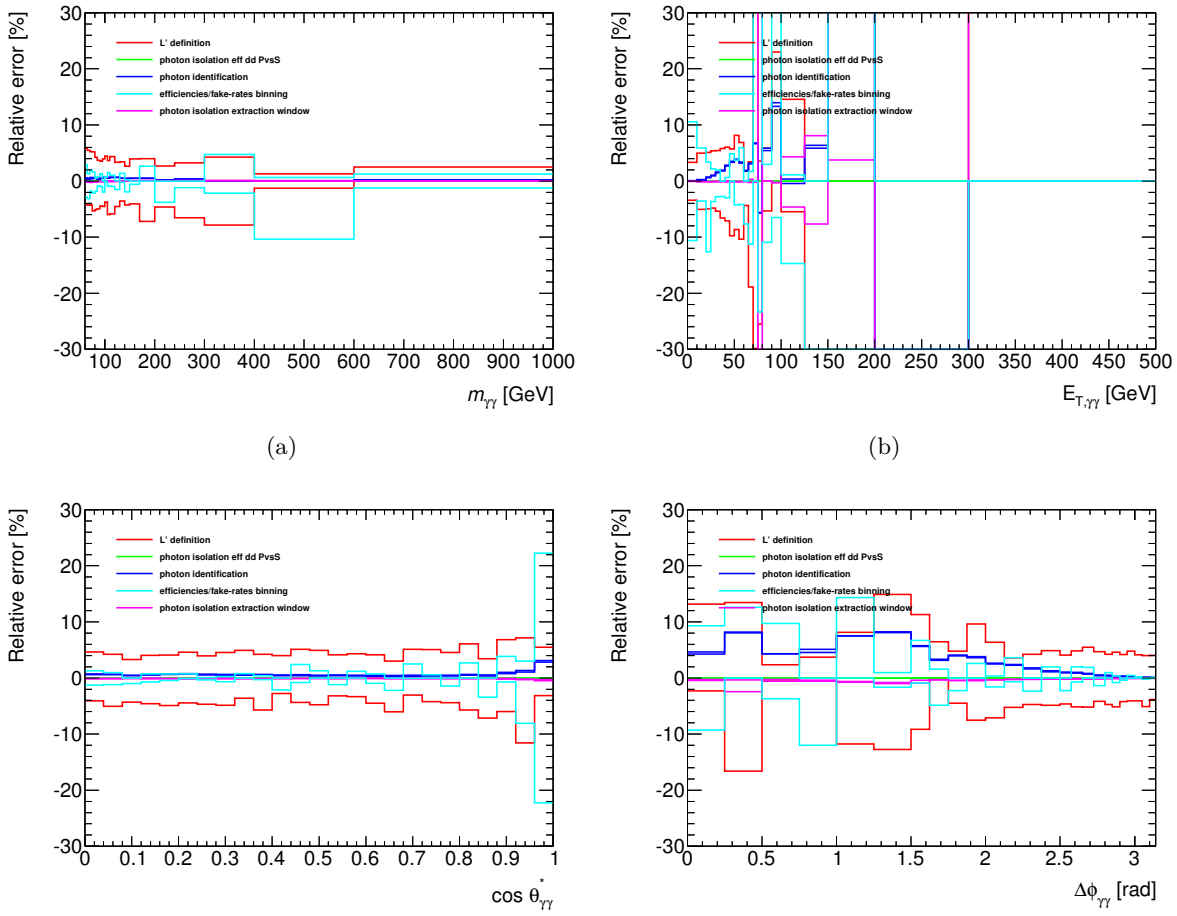


Figure C.1: Breakdown of the systematic uncertainties on the diphoton yield extracted by the 2x2DSB method as a function of the diphoton observables for the $0-jet$ category.

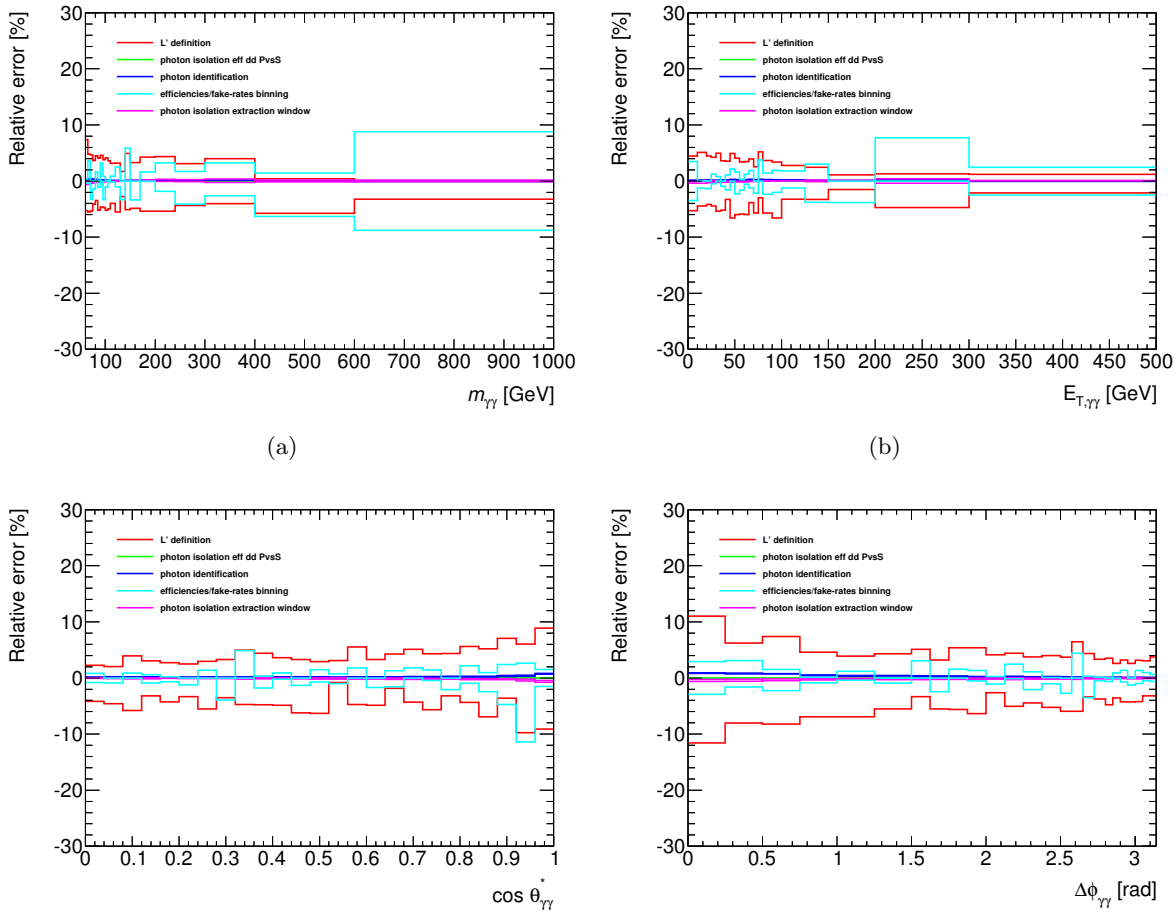


Figure C.2: Breakdown of the systematic uncertainties on the diphoton yield extracted by the 2x2DSB method as a function of the diphoton observables for the $1-jet$ category.

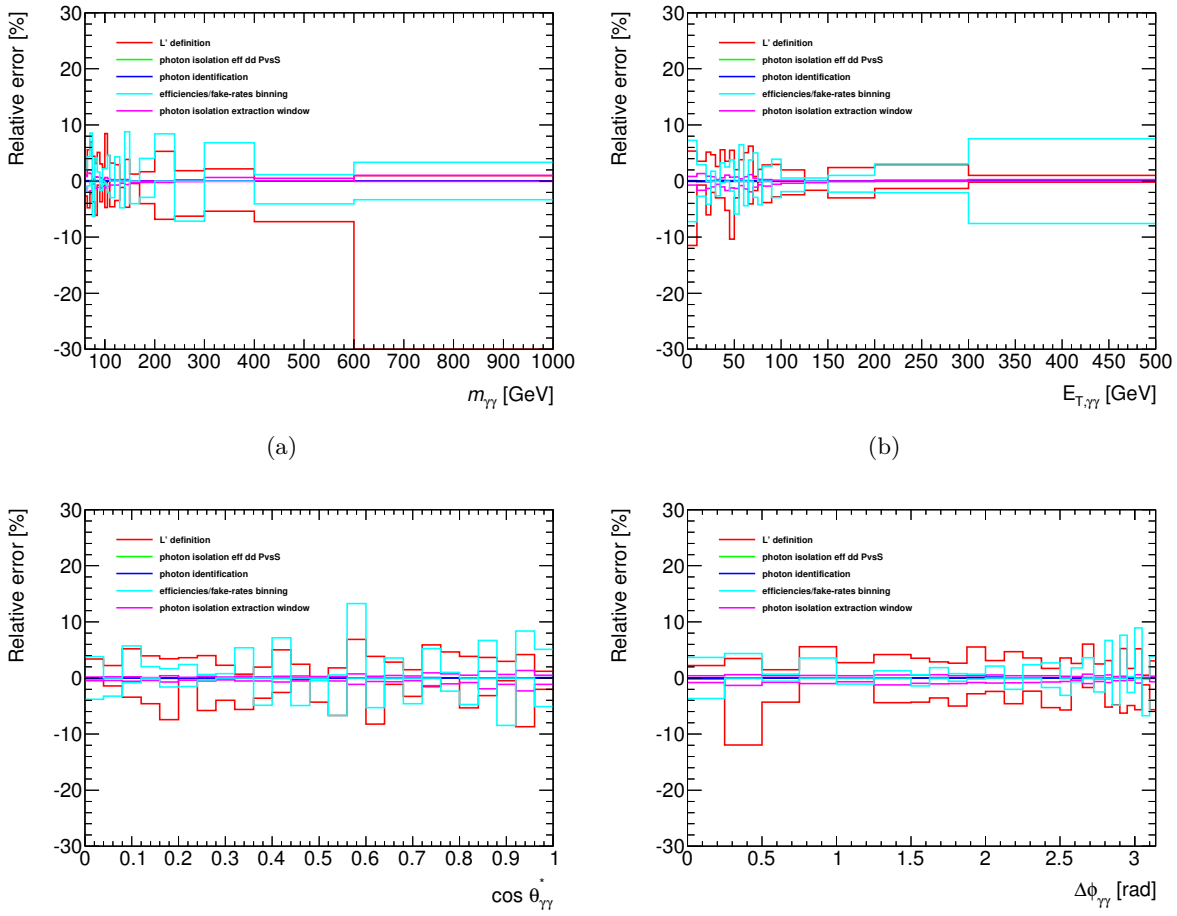


Figure C.3: Breakdown of the systematic uncertainties on the diphoton yield extracted by the 2x2DSB method as a function of the diphoton observables for the $\geq 3-jet$ category.

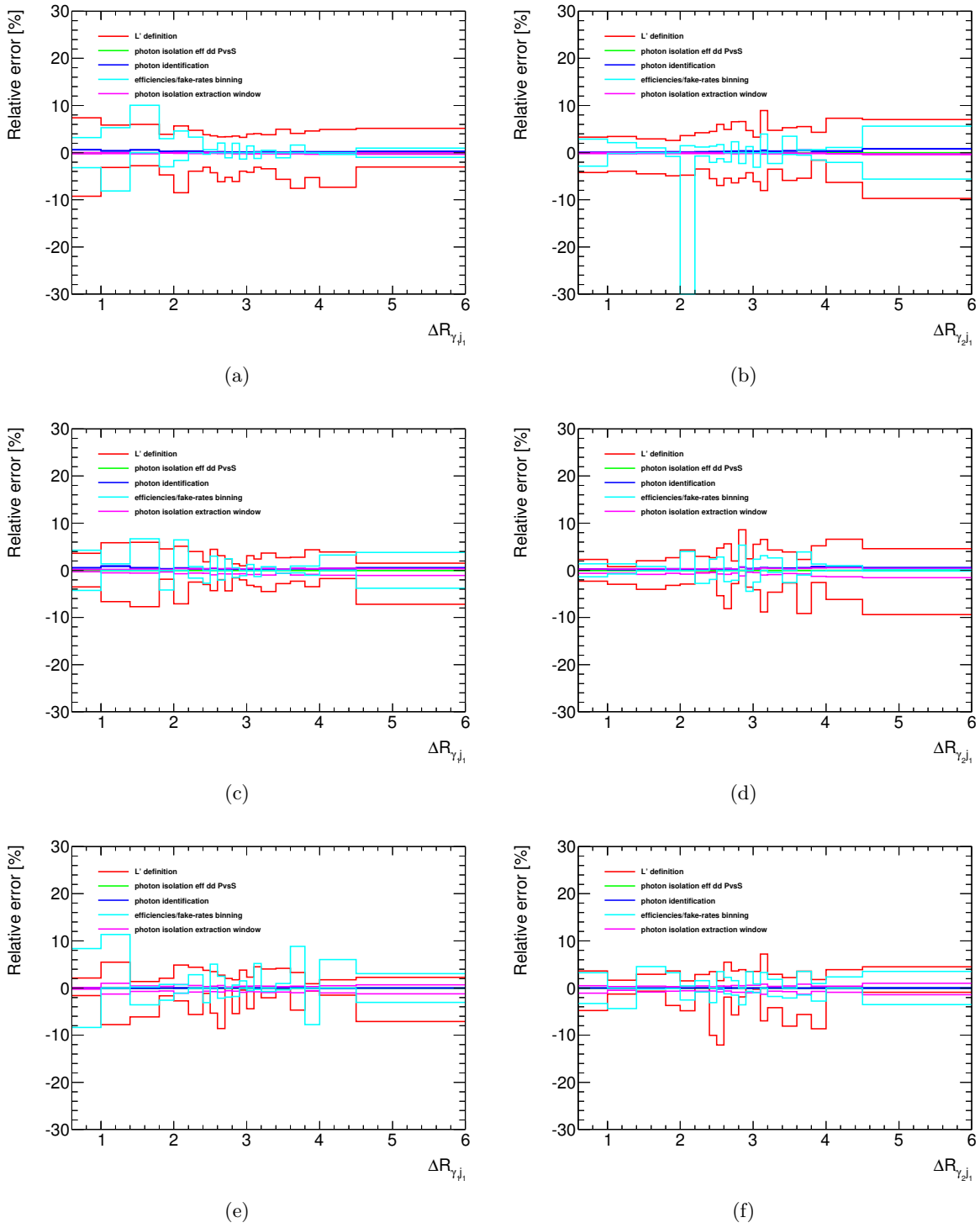


Figure C.4: Breakdown of the systematic uncertainties on the diphoton yield extracted by the 2x2DSB method as a function of the R-separation between the leading photon and the leading jet (left) and the subleading photon and the leading jet (right) for the $1-jet$ (top), $2-jet$ (middle) and $\geq 3-jet$ (bottom) category.

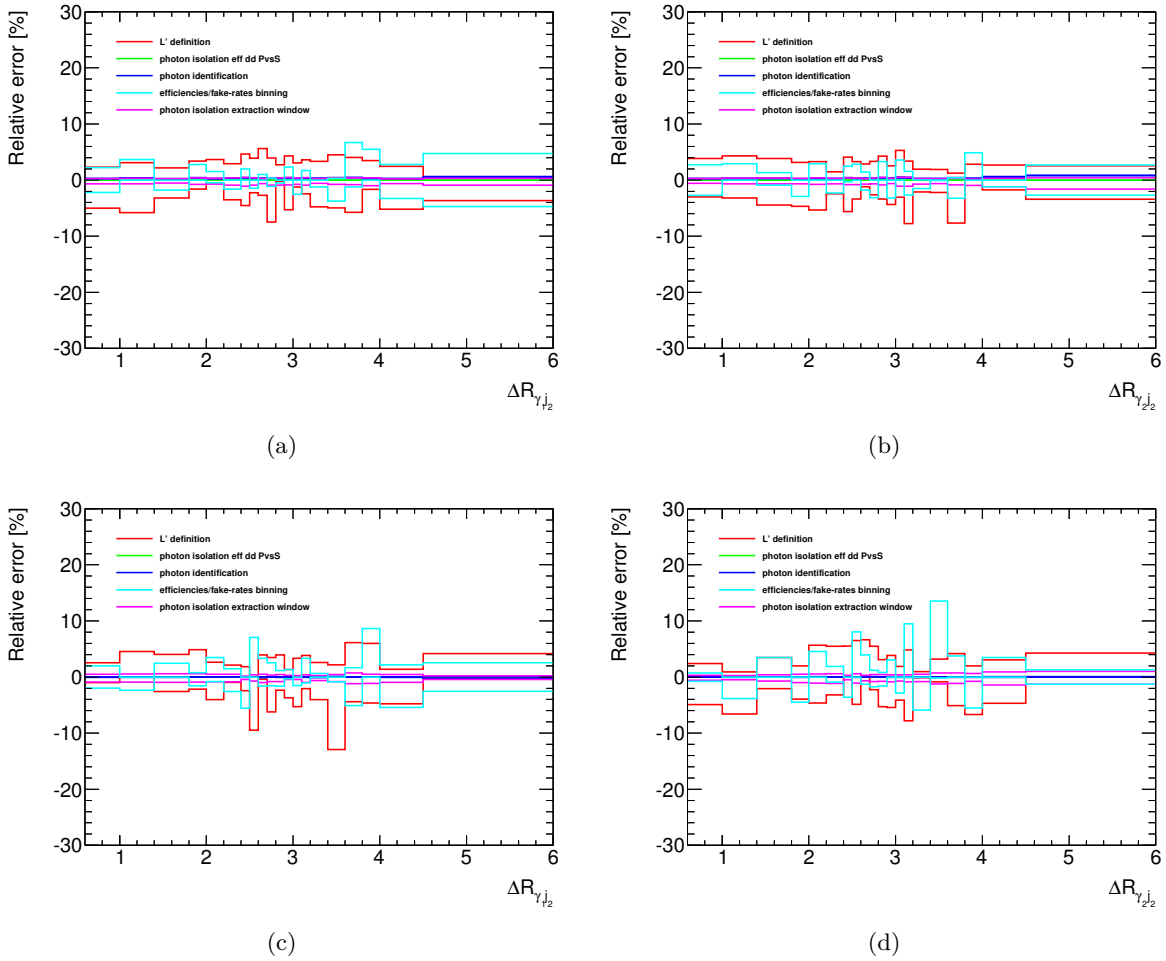


Figure C.5: Breakdown of the systematic uncertainties on the diphoton yield extracted by the 2x2DSB method as a function of the R-separation between the leading photon and the subleading jet (left) and the subleading photon and the subleading jet (right) for the $2-jet$ (top) and $\geq 3-jet$ (bottom) category.

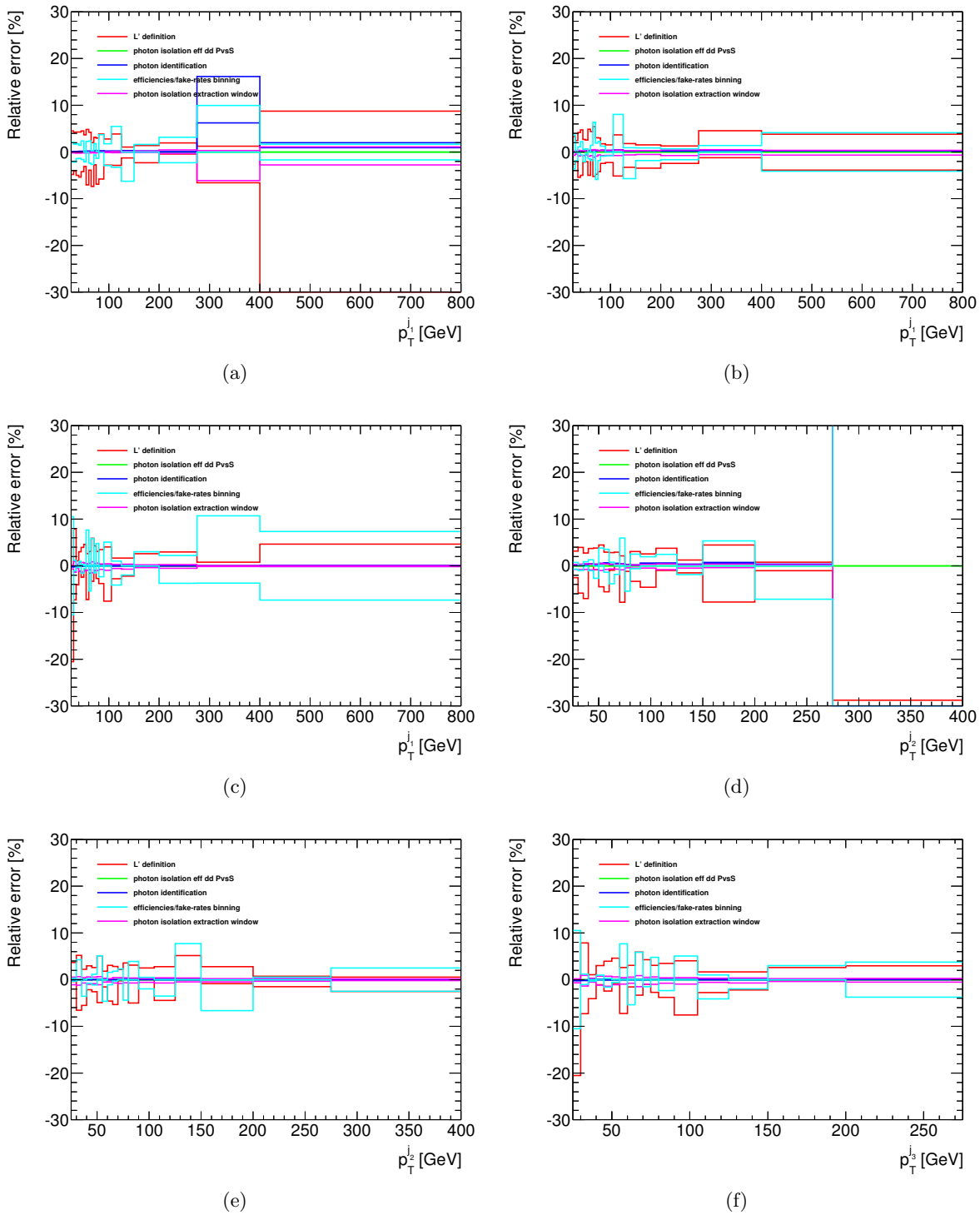


Figure C.6: Breakdown of the systematic uncertainties on the diphoton yield extracted by the 2x2DSB method as a function of the p_{T}^{jet}. (a) $1-jet$, (b),(d) $2-jet$, (c),(e),(f) $\geq 3-jet$ category.

D

Diphoton yields and their total uncertainties.

This Appendix contains plots of the decomposed diphoton spetra into the four components, as described in Eq. 7.4 and their statistical, systematic and total uncertainties, as a function of all observables except the ones already shown in the main text in Section 7.3.3.2.

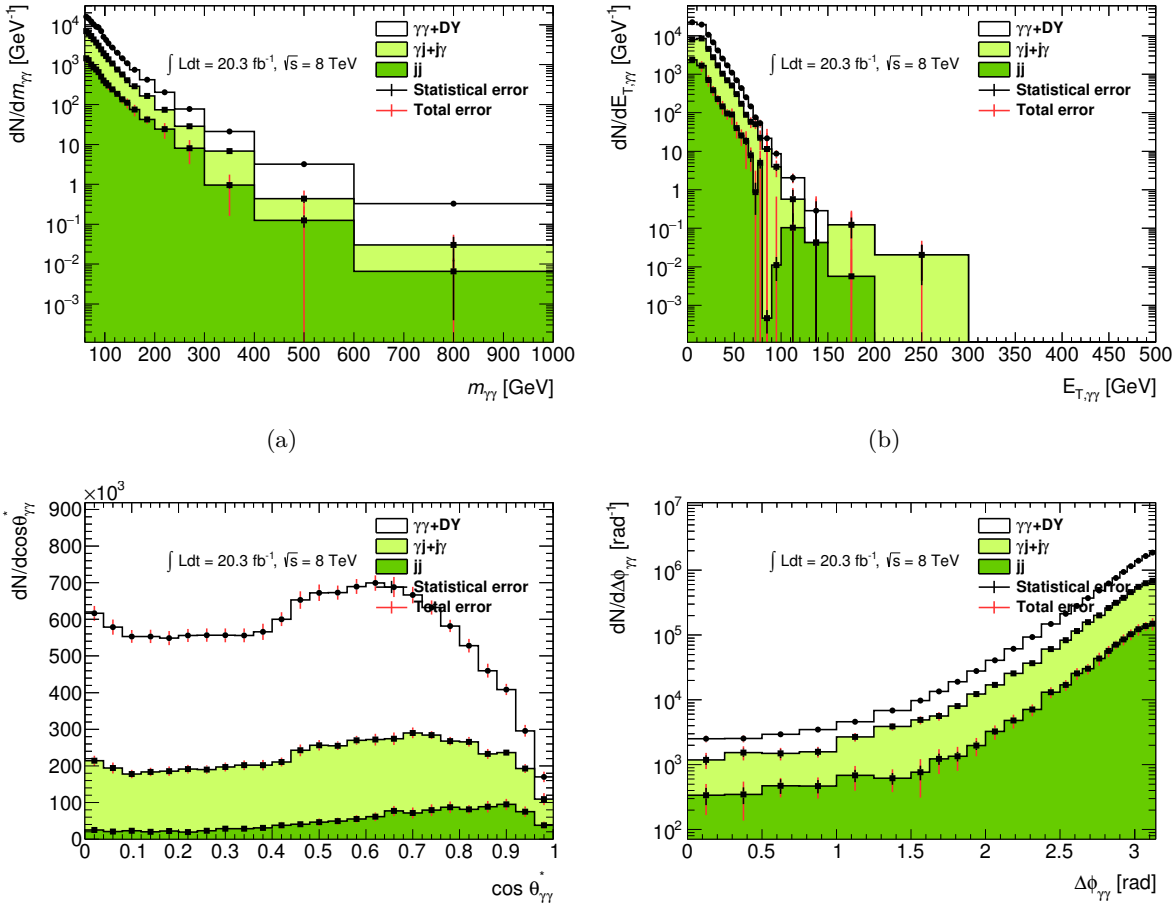


Figure D.1: Diphoton yields as a function of the diphoton observables: $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$, $\cos\theta^{*}_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ for the $N_{jet} = 0$ category.

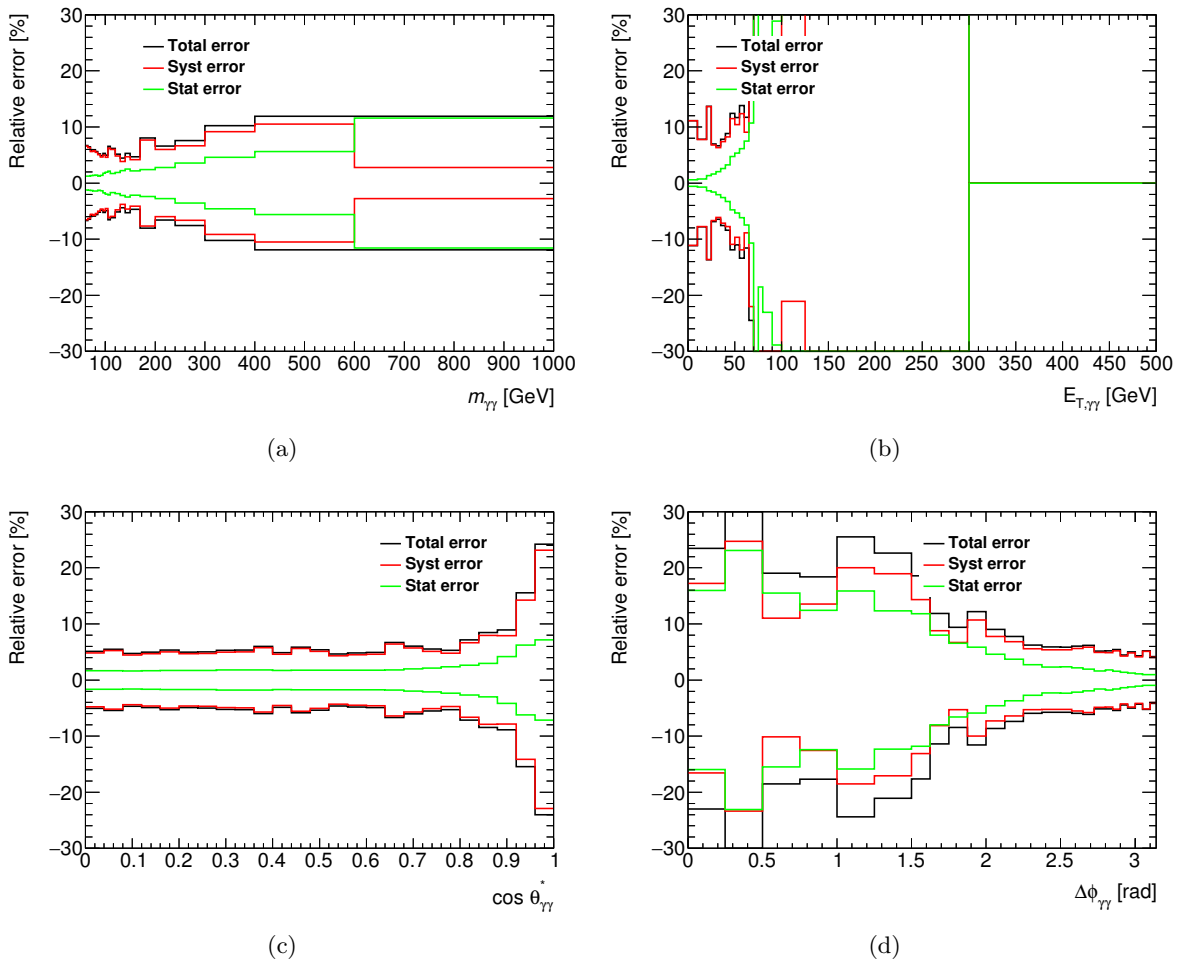


Figure D.2: Statistical (green) and systematic (red) and total (black) uncertainties on the diphoton yields as a function of the diphoton observables: $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$, $\cos\theta^{*}_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ for the $N_{jet} = 0$ category.

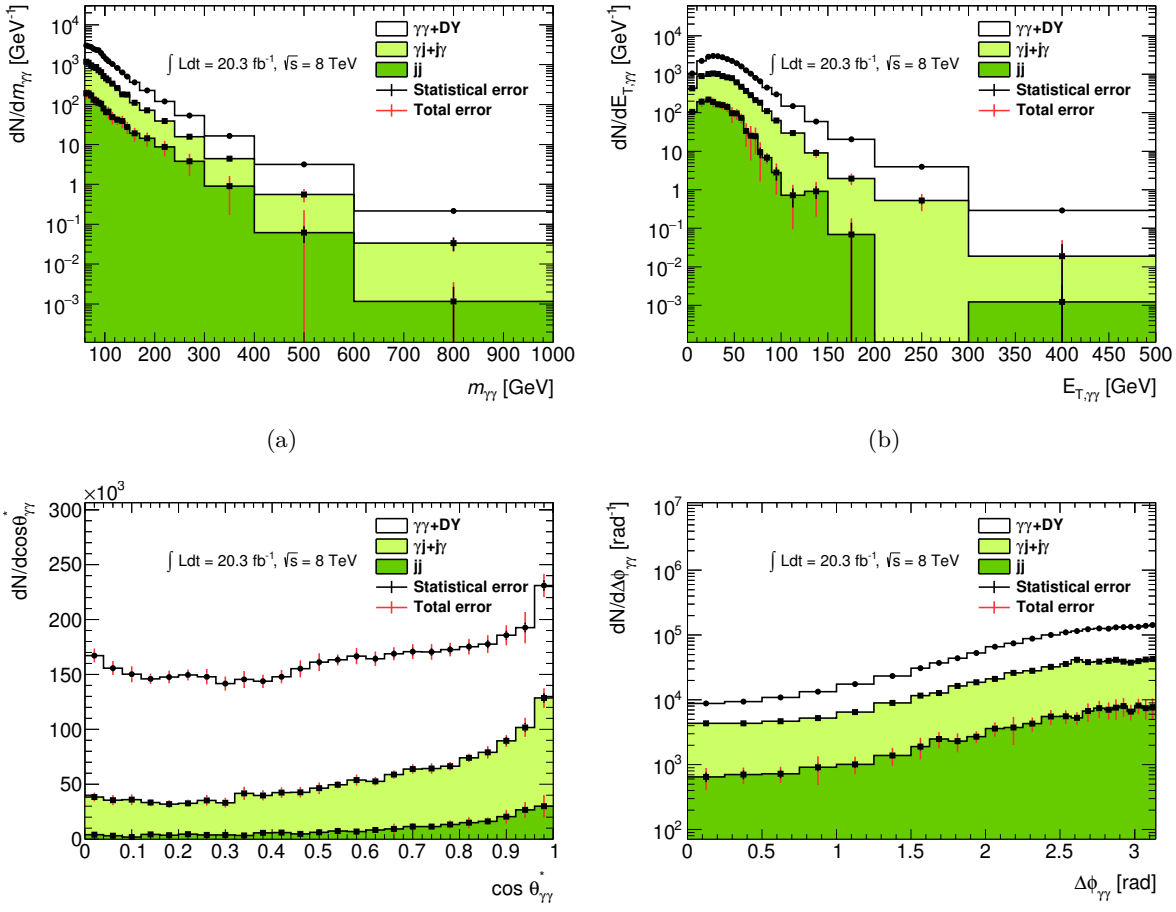


Figure D.3: Diphoton yields as a function of the diphoton observables: $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$, $\cos\theta^{*}_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ for the $N_{jet} = 1$ category.

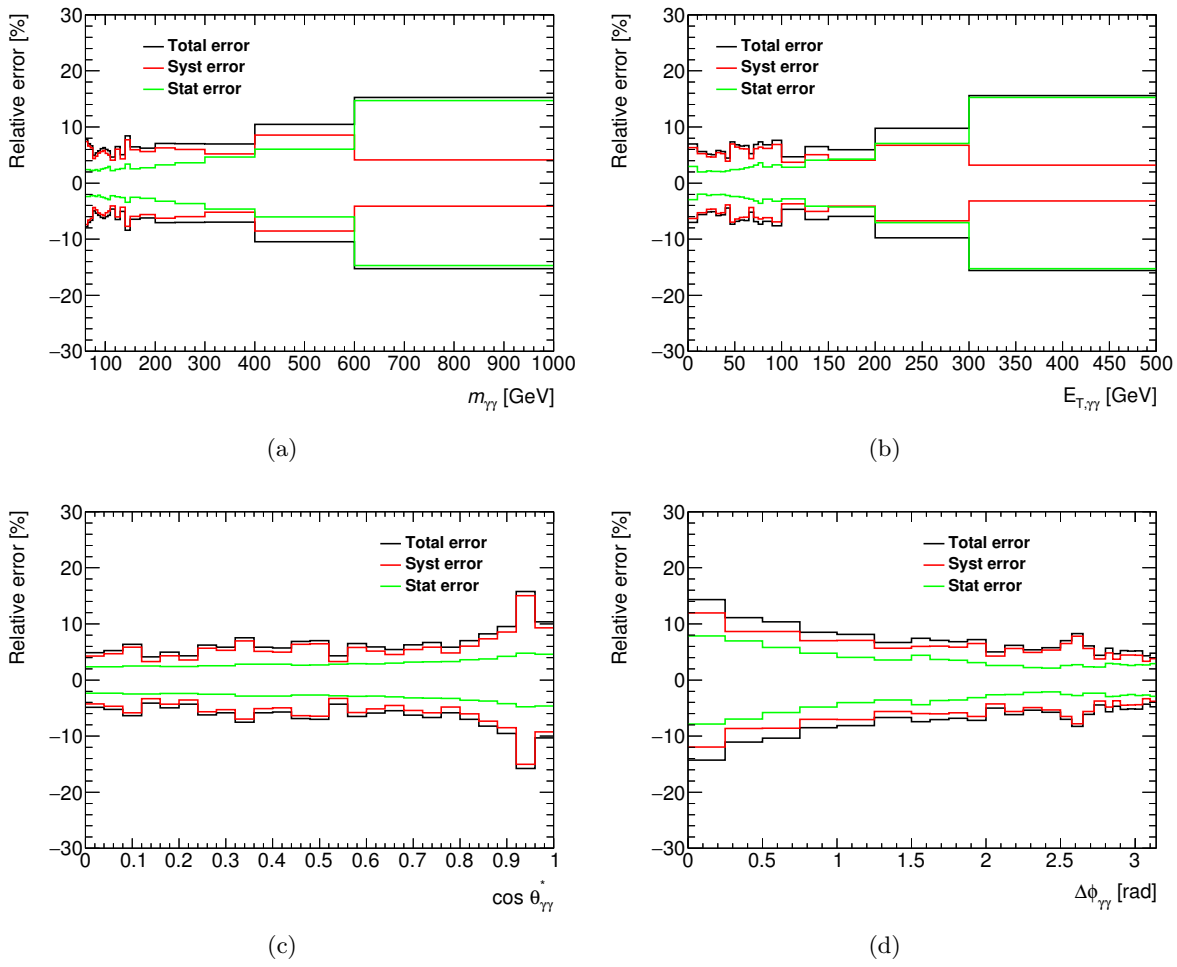


Figure D.4: Statistical (green) and systematic (red) and total (black) uncertainties on the diphoton yields as a function of the diphoton observables: $m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma\gamma}$, $\cos\theta^{*}_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ for the $N_{jet} = 1$ category.

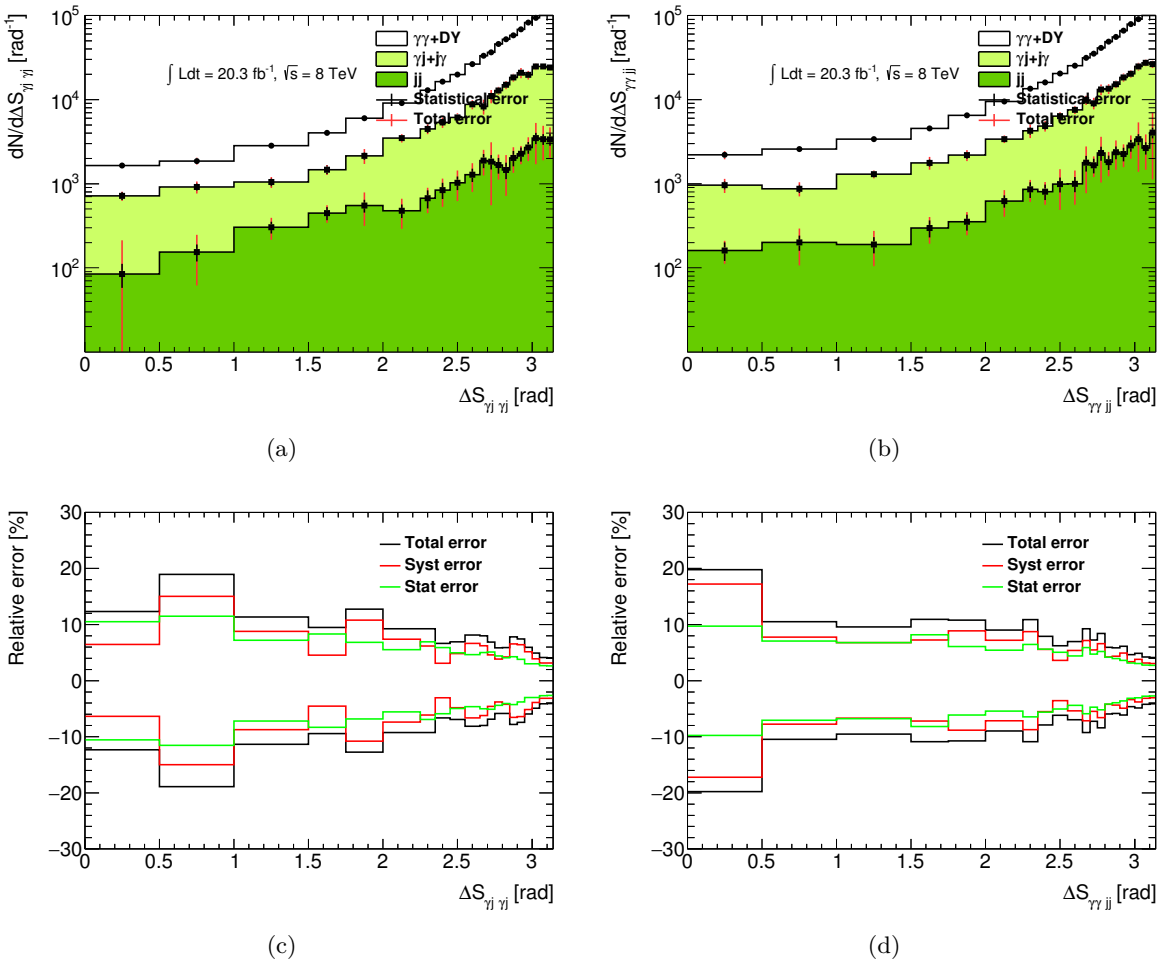
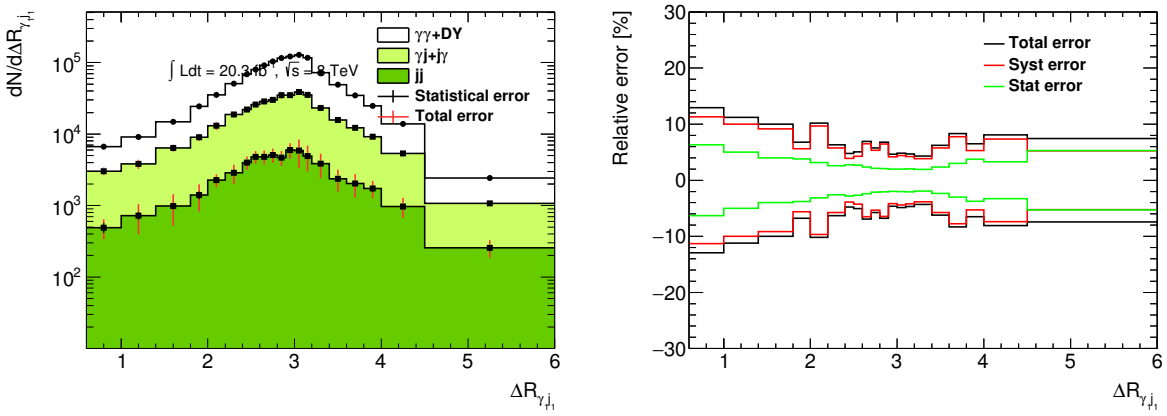
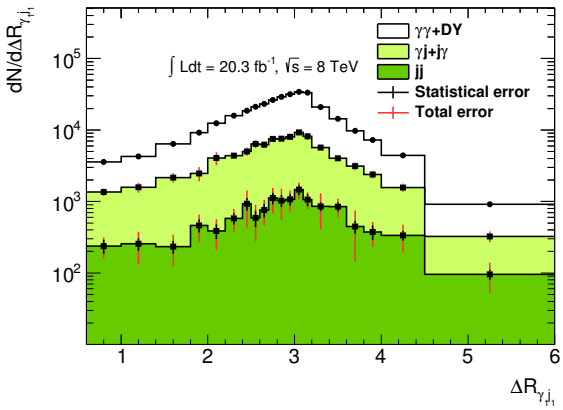


Figure D.5: Diphoton yields (top) and their uncertainties (bottom) as a function of the ΔS observables (a), (c) $\Delta S_{\gamma j \gamma j}$ and (b), (d) $\Delta S \gamma j \gamma j$ for the $N_{jet} = 2$ category.

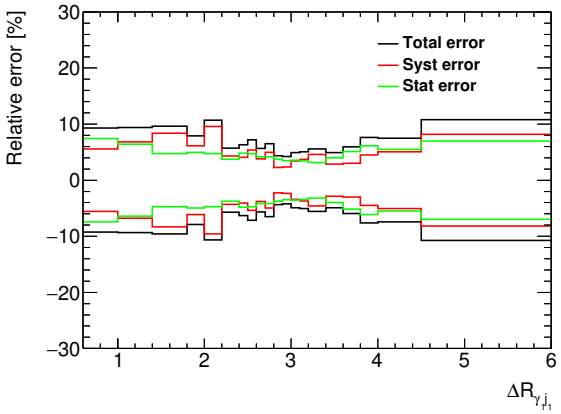


(a)

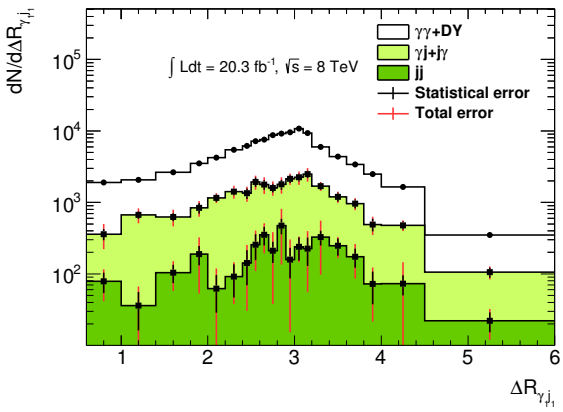
(b)



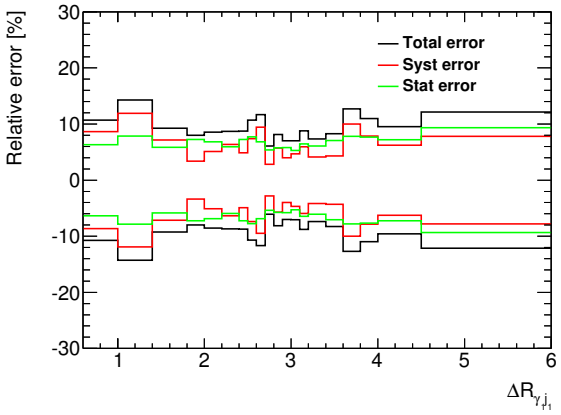
(c)



(d)



(e)



(f)

Figure D.6: Diphoton yield (left) as a function of the $\Delta R_{\gamma_1 j_1}$ variable and its uncertainties (right) for the $N_{jet} = 1$ (top), $N_{jet} = 2$ (middle) and $N_{jet} \geq 3$ (bottom) category.

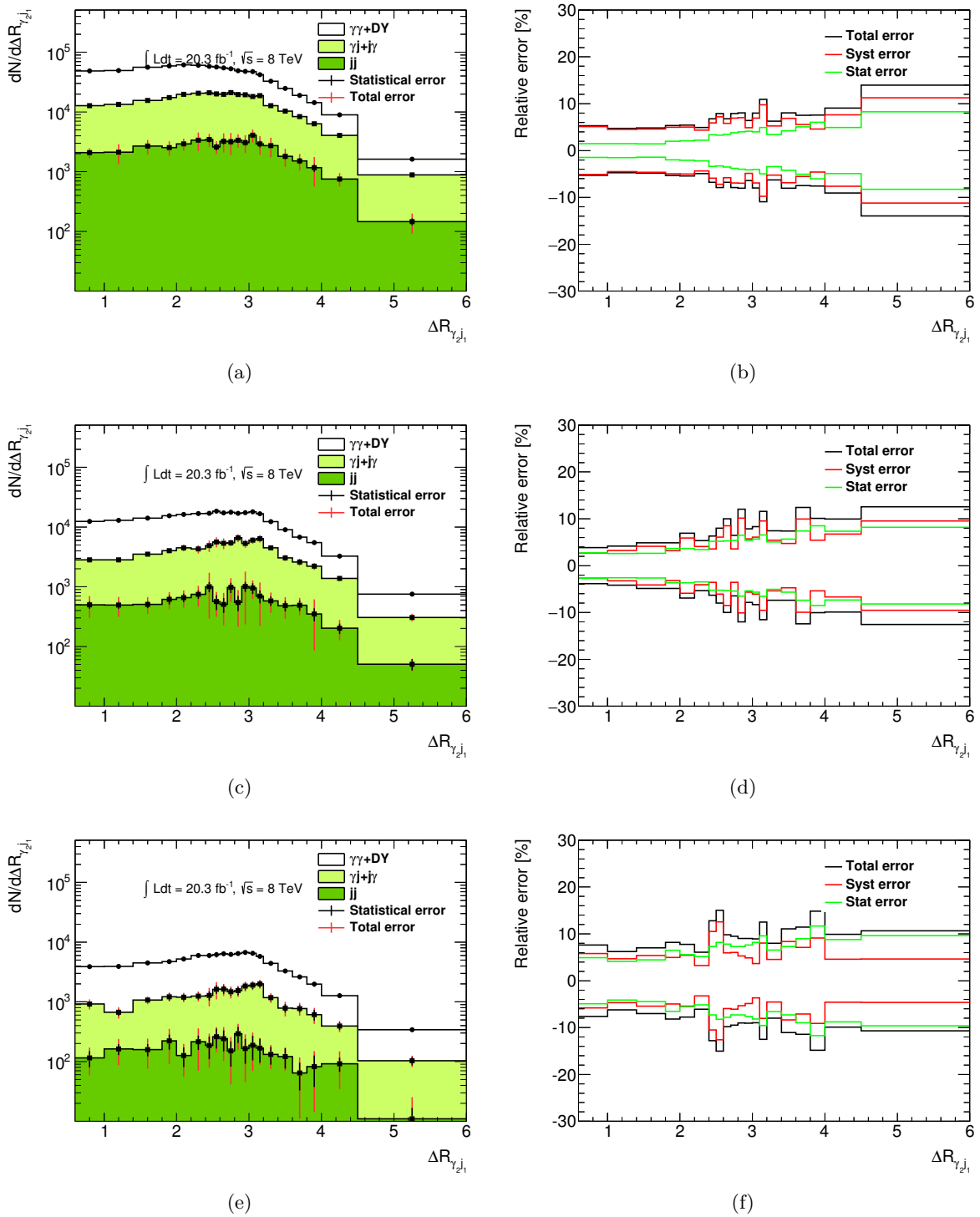


Figure D.7: Diphoton yield (left) as a function of the $\Delta R_{\gamma_2 j_1}$ variable and its uncertainties (right) for the $N_{jet} = 1$ (top), $N_{jet} = 2$ (middle) and $N_{jet} \geq 3$ (bottom) category.

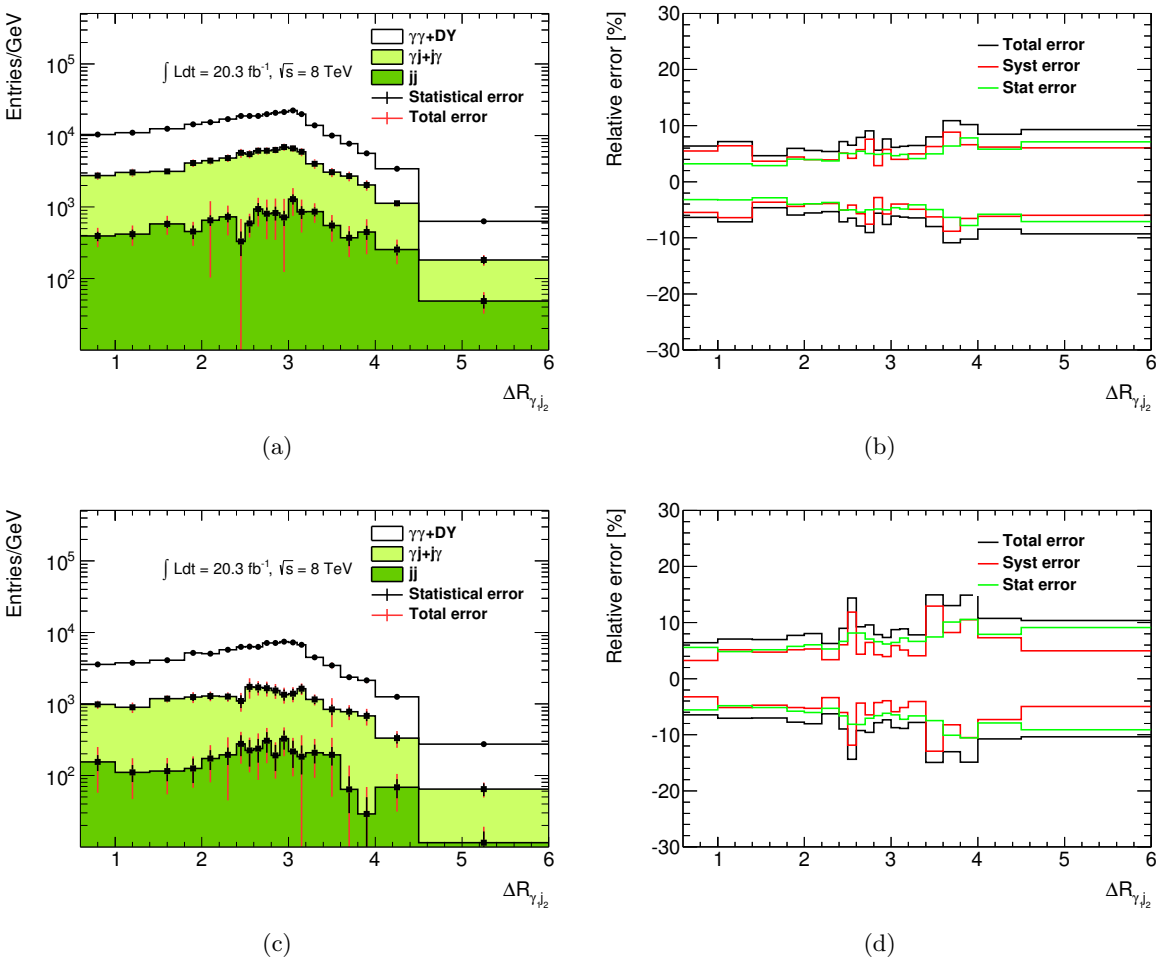


Figure D.8: Diphoton yield (left) as a function of the $\Delta R_{\gamma_1 j_2}$ variable and its uncertainties (right) for the $N_{jet} = 2$ (top) and $N_{jet} \geq 3$ (bottom) category.

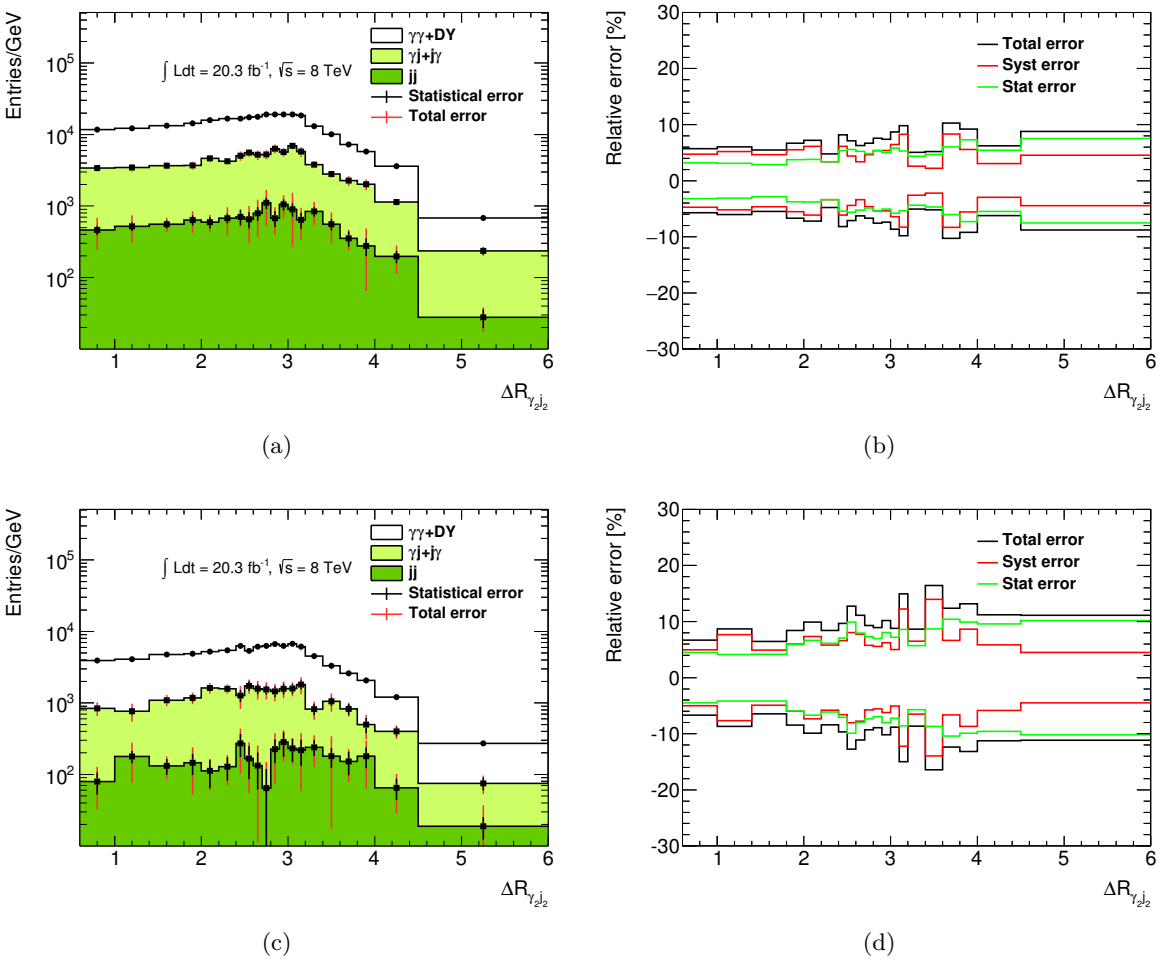


Figure D.9: Diphoton yield (left) as a function of the $\Delta R_{\gamma 2 j 2}$ variable and its uncertainties (right) for the $N_{jet} = 2$ (top) and $N_{jet} \geq 3$ (bottom) category.

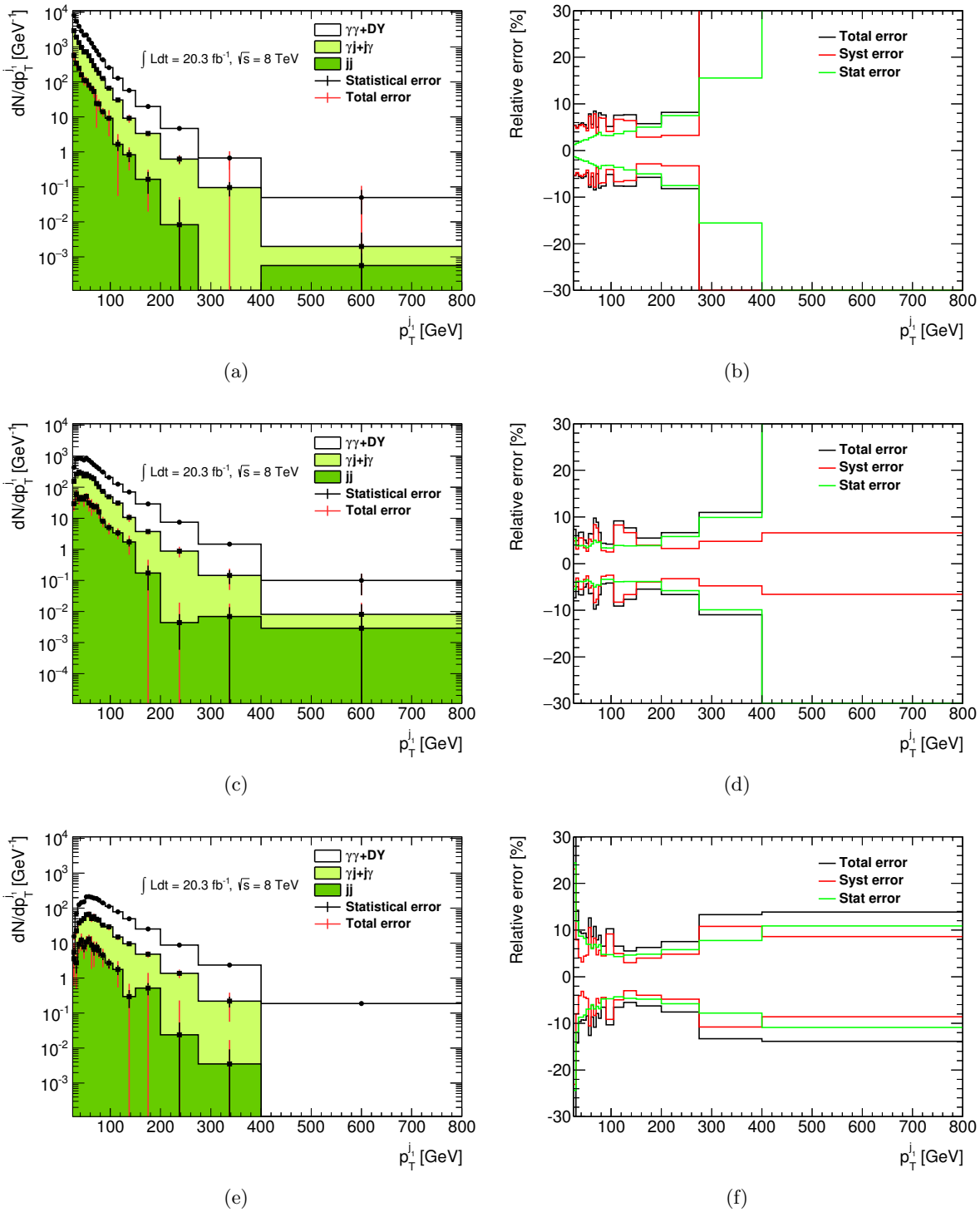
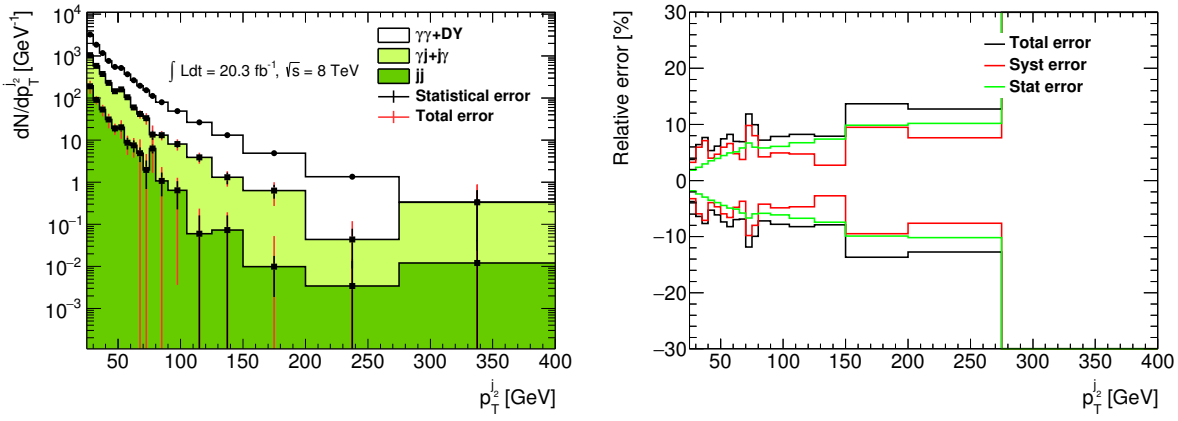
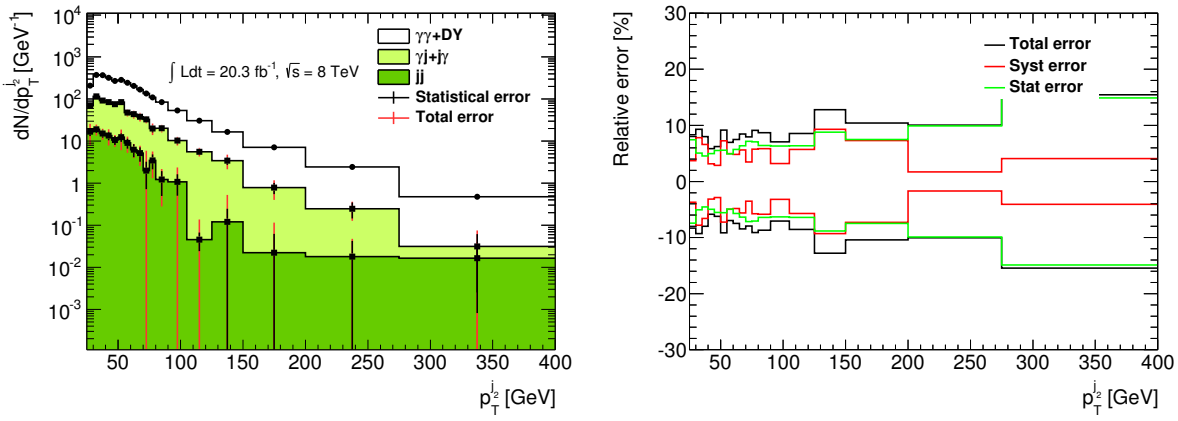


Figure D.10: Diphoton yield (left) as a function of the $p_{\mathrm{T}}^{jet_1}$ variable and its uncertainties (right) for the $N_{jet} = 1$ (top), $N_{jet} = 2$ (middle) and $N_{jet} \geq 3$ (bottom) category.

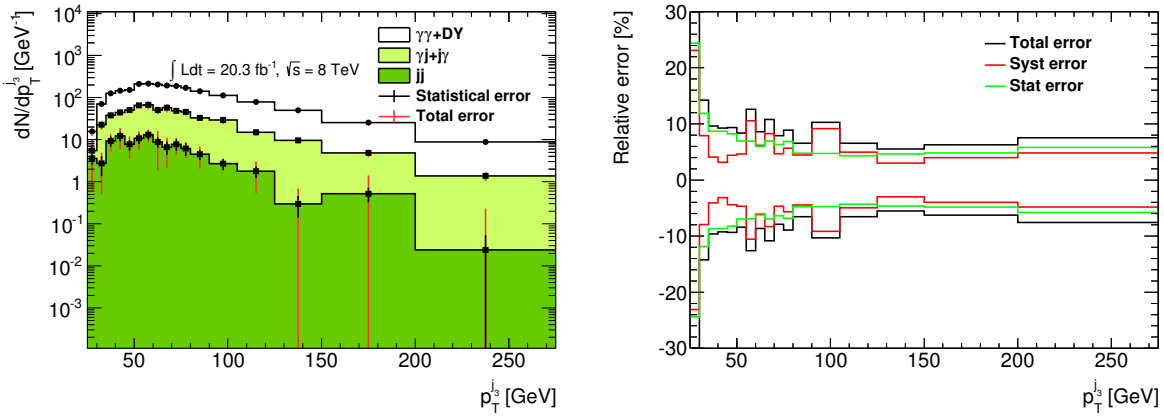


(a) (b)



(c) (d)

Figure D.11: Diphoton yield (left) as a function of the $p_{\mathrm{T}}^{jet_1}$ variable and its uncertainties (right) for the $N_{jet} = 1$ (top), $N_{jet} = 2$ (middle) and $N_{jet} \geq 3$ (bottom) category.



(a) (b)

Figure D.12: Diphoton yield (a) as a function of the $p_{\mathrm{T}}^{jet_3}$ variable and its uncertainties (b) for the $N_{jet} \geq 3$ category.

E

Uncertainty on the electron impurities due to $e \to \gamma$ fake rate variation

This Appendix contains plots of the uncertainties assigned on the diphoton yields due to the uncertainty on the measured $e \to \gamma$ fake rate as a function of all observables except the ones already shown in the main text in Section 7.4.2.

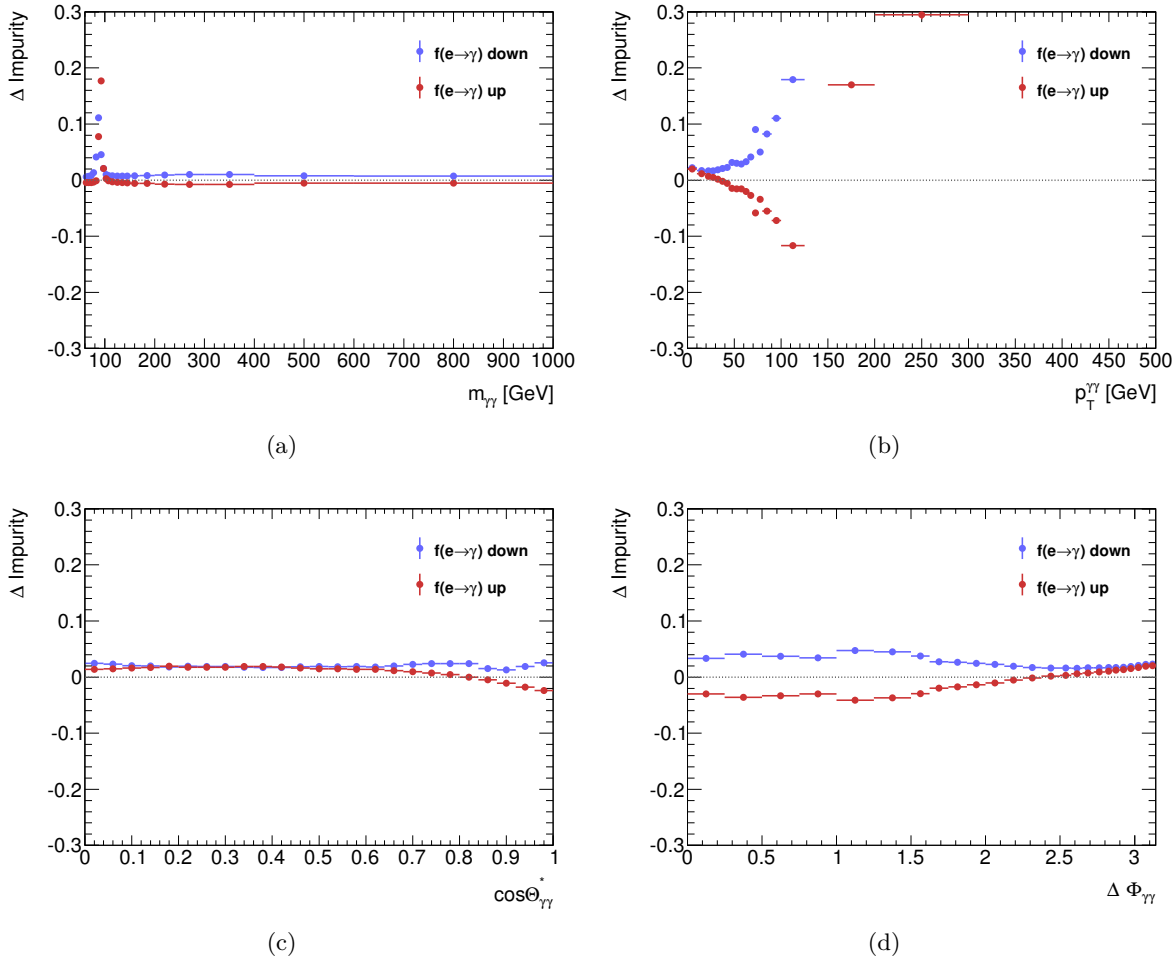
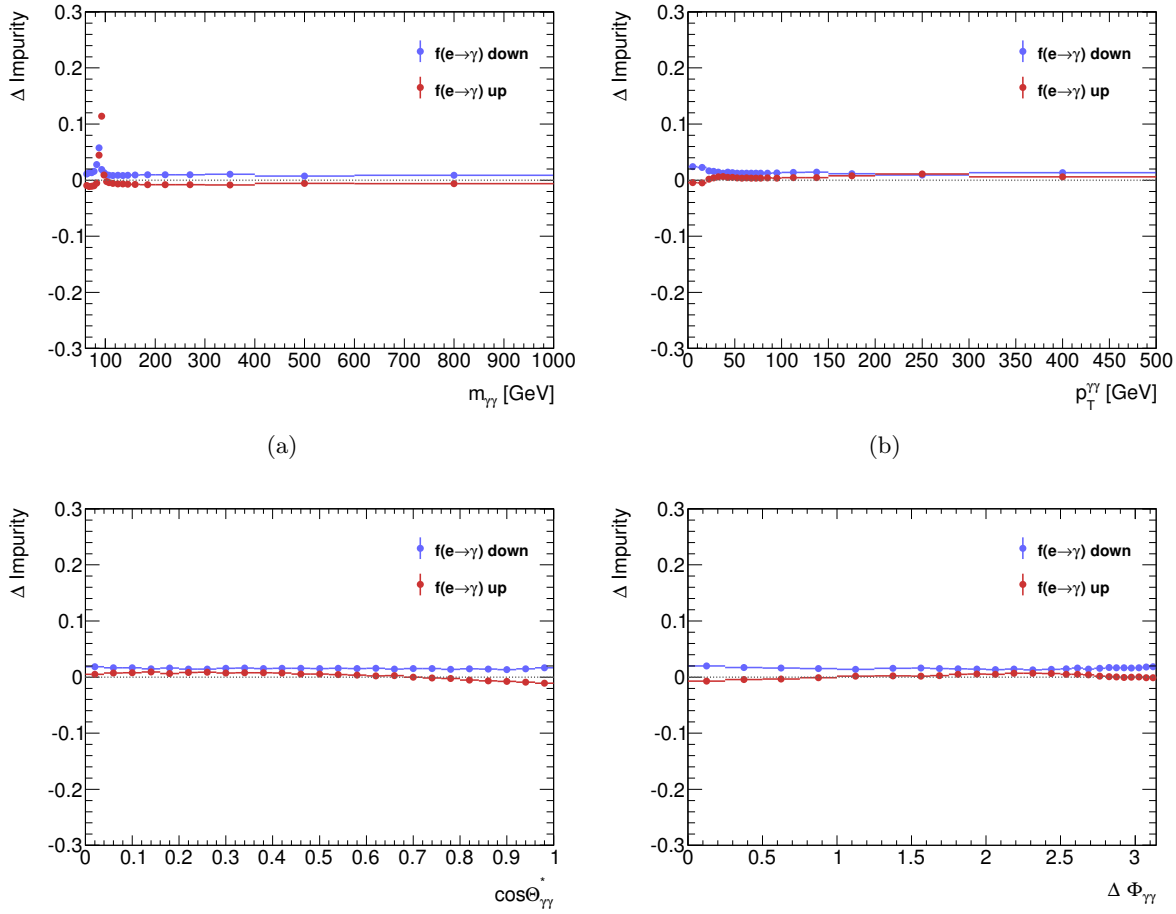
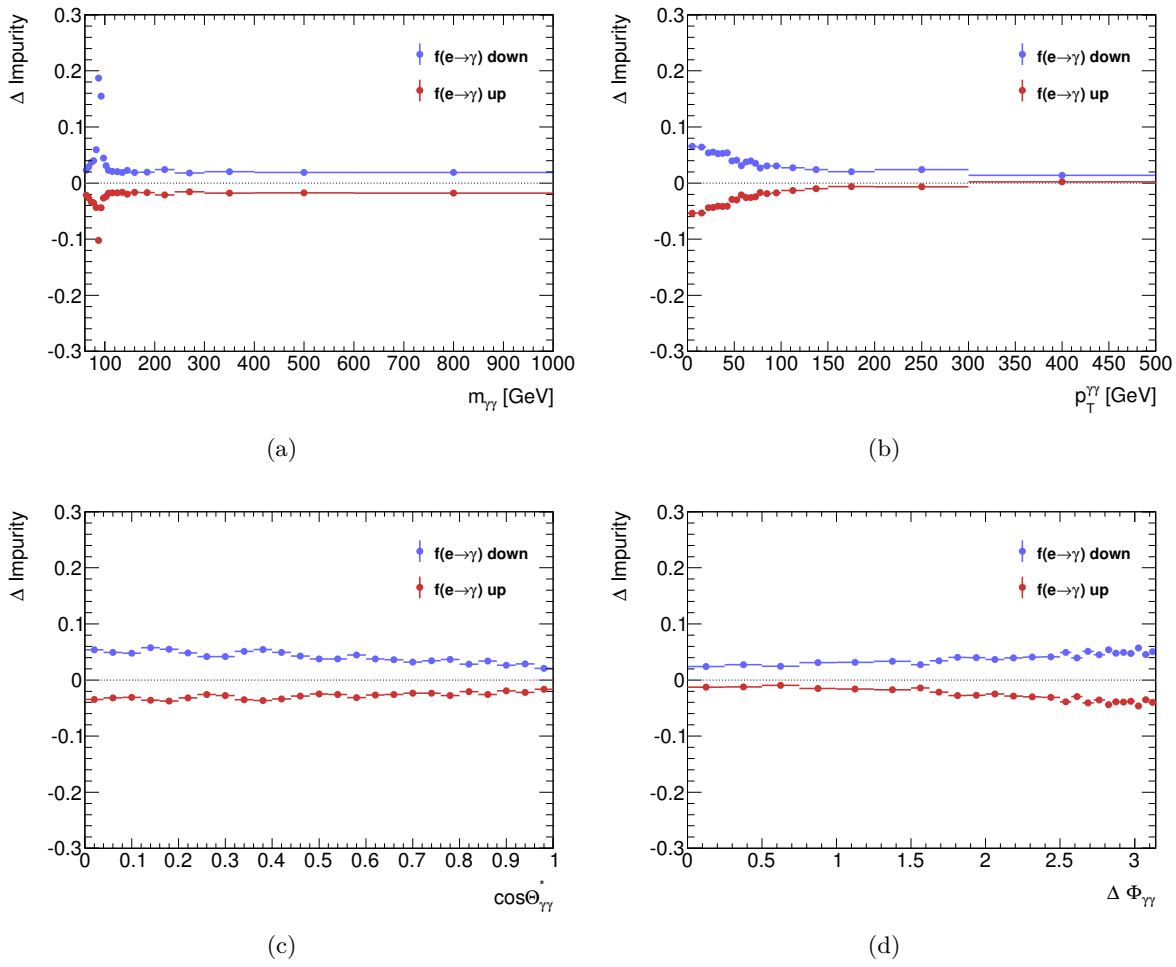


Figure E.1: Relative change in the estimated impurities estimated as a 50% variation of the $e \to \gamma$ fake rate up and down, as a function of the diphoton observables for the $0-jet$ category.



(a) (b) (c) (d)

Figure E.2: Relative change in the estimated impurities estimated as a 50% variation of the $e \to \gamma$ fake rate up and down, as a function of the diphoton observables for the $1-jet$ category.



(a) (b) (c) (d)

Figure E.3: Relative change in the estimated impurities estimated as a 50% variation of the $e \to \gamma$ fake rate up and down, as a function of the diphoton observables for the $\geq 3 - jet$ category.

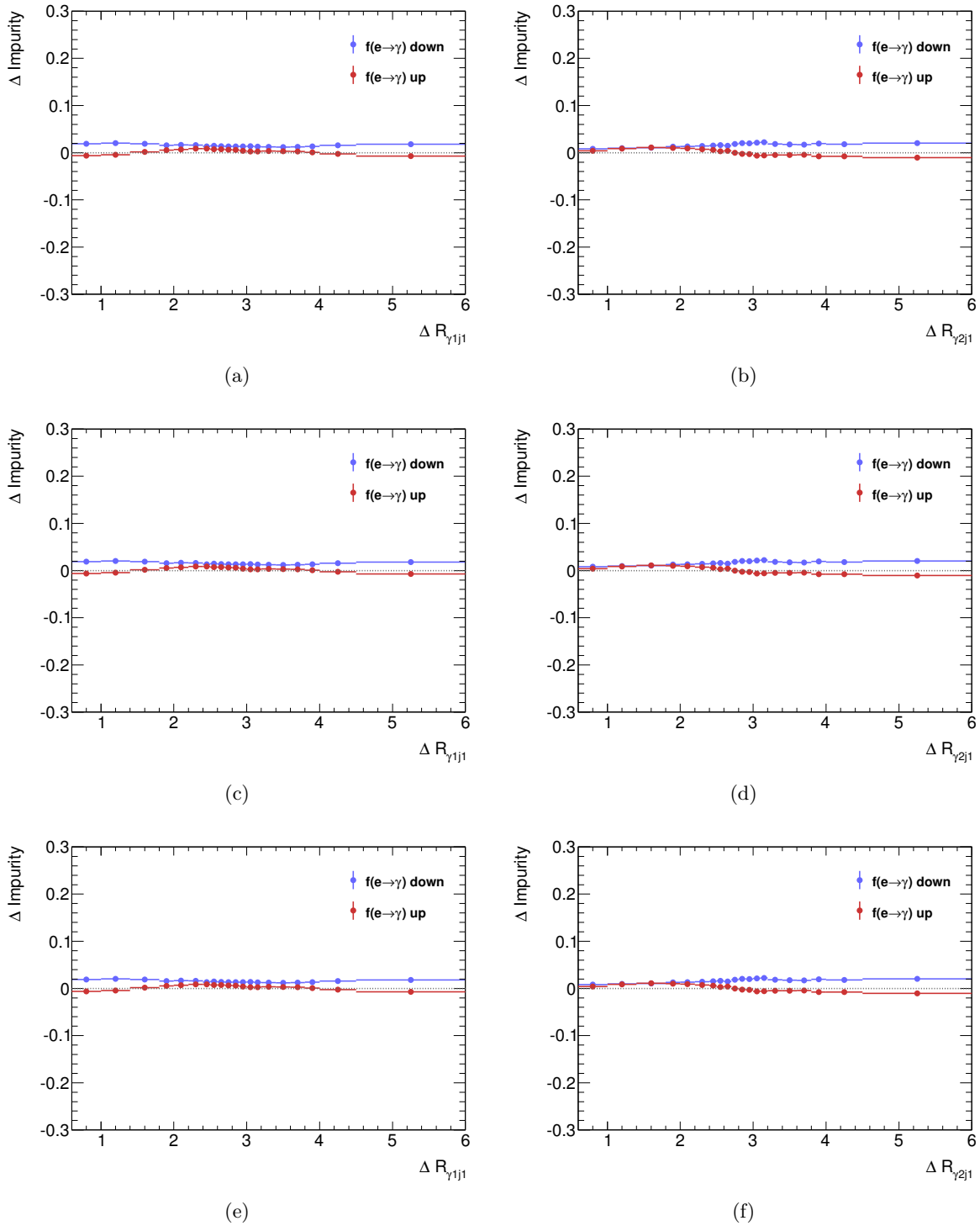
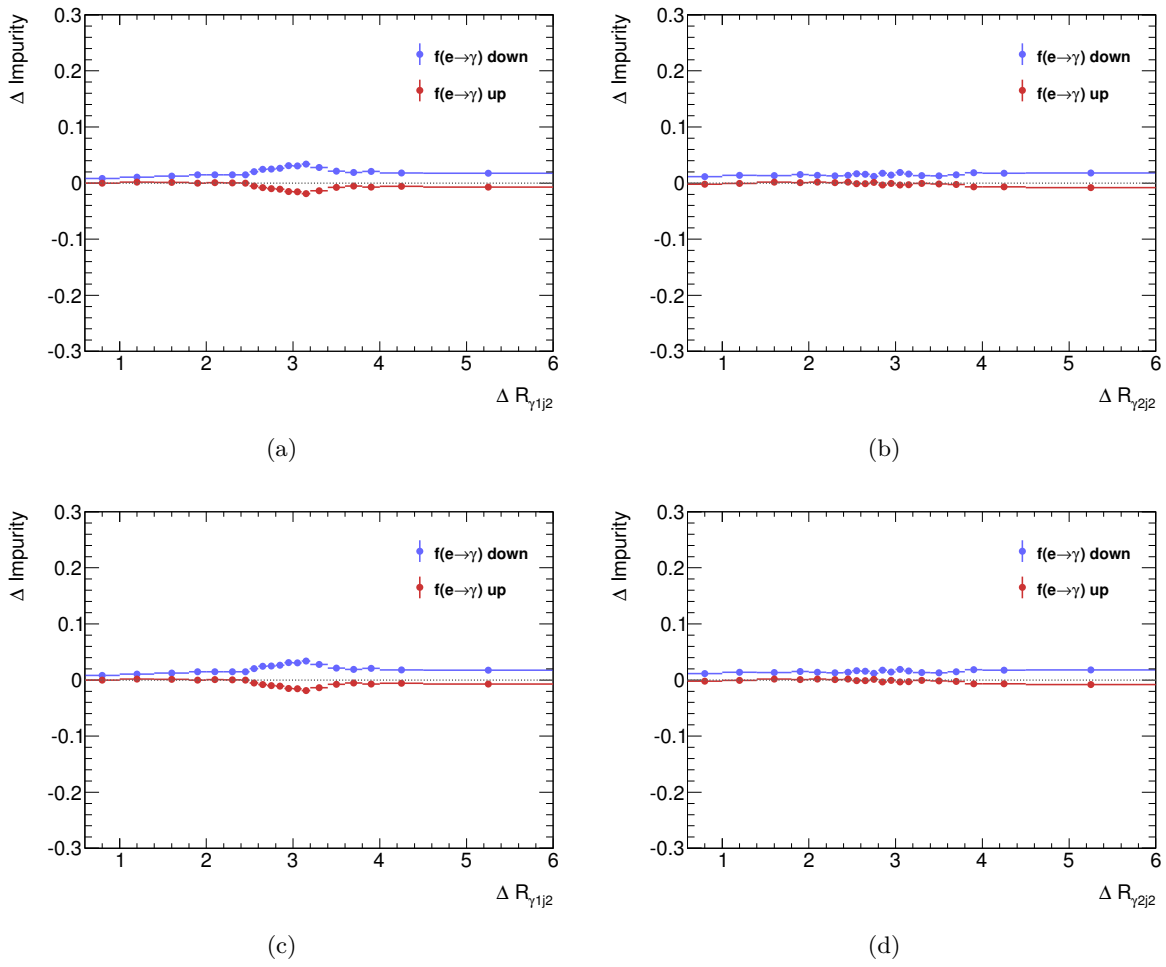


Figure E.4: Relative change in the estimated impurities estimated as a 50% variation of the $e \to \gamma$ fake rate up and down, as a function of the R-separation between the photons and the leading jet for the $1 - jet$ category.



(a) (b)

(c) (d)

Figure E.5: Relative change in the estimated impurities estimated as a 50% variation of the $e \to \gamma$ fake rate up and down, as a function of the R-separation between the photons and the subleading jet for the $2 - jet$ category.

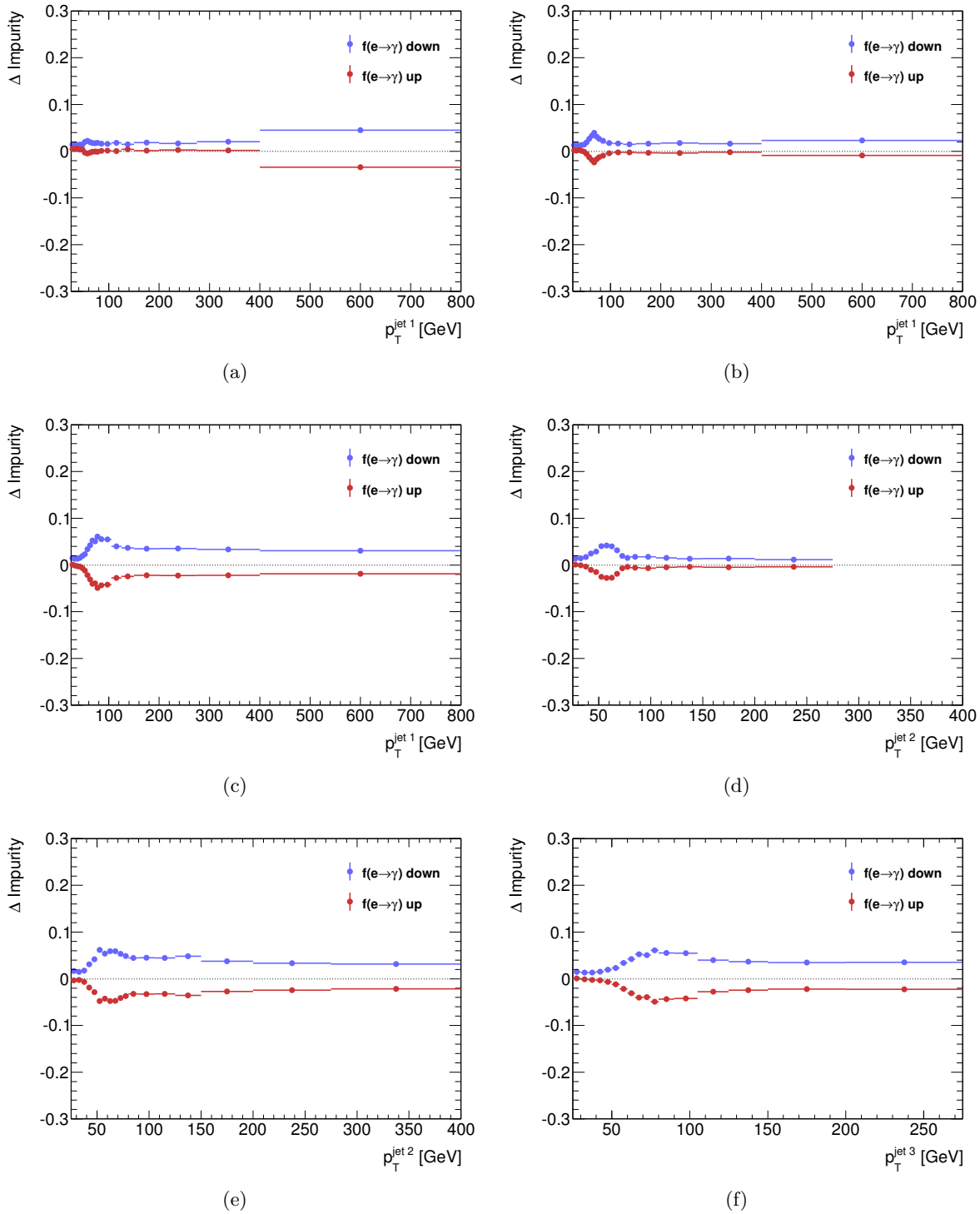
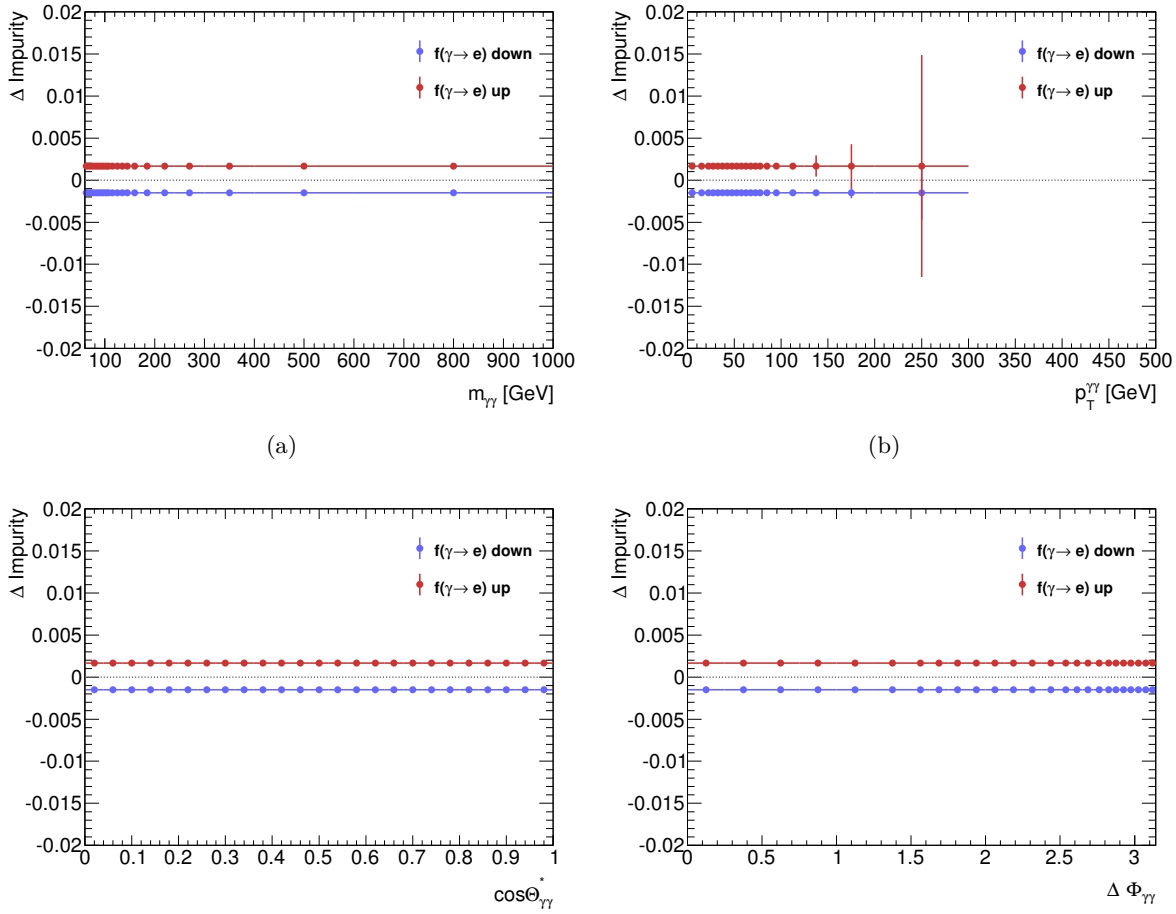


Figure E.6: Relative change in the estimated impurities estimated as a 50% variation of the $e \to \gamma$ fake rate up and down, as a function of the p_T^{jet}. (a) $1-jet$, (b),(d) $2-jet$, (c),(e),(f) $\geq 3-jet$ category.

F

Uncertainty on the electron impurities due to $\gamma \to e$ fake rate variation

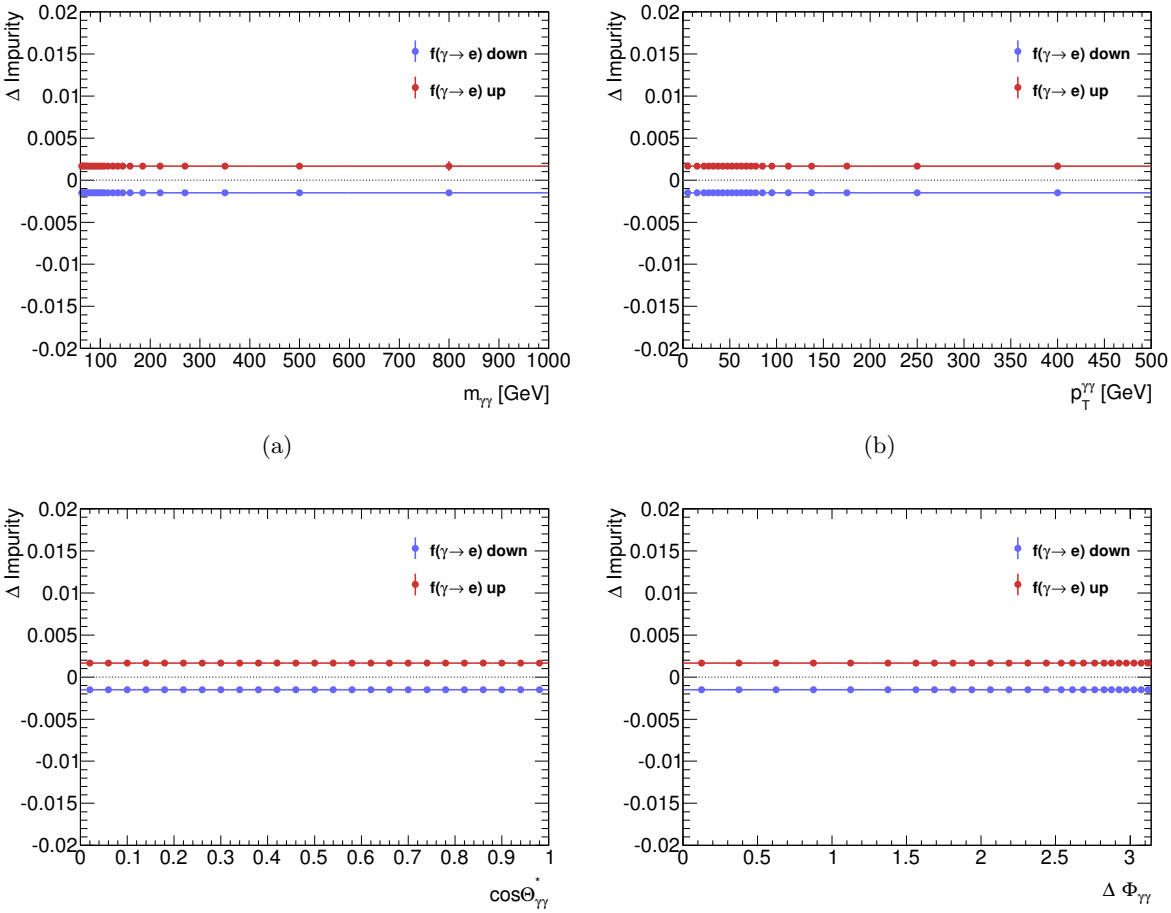
This Appendix contains plots of the uncertainties assigned on the diphoton yields due to the uncertainty on the measured $\gamma \to e$ fake rate as a function of all observables except the ones already shown in the main text in Section 7.4.2.



(a) (b)

(c) (d)

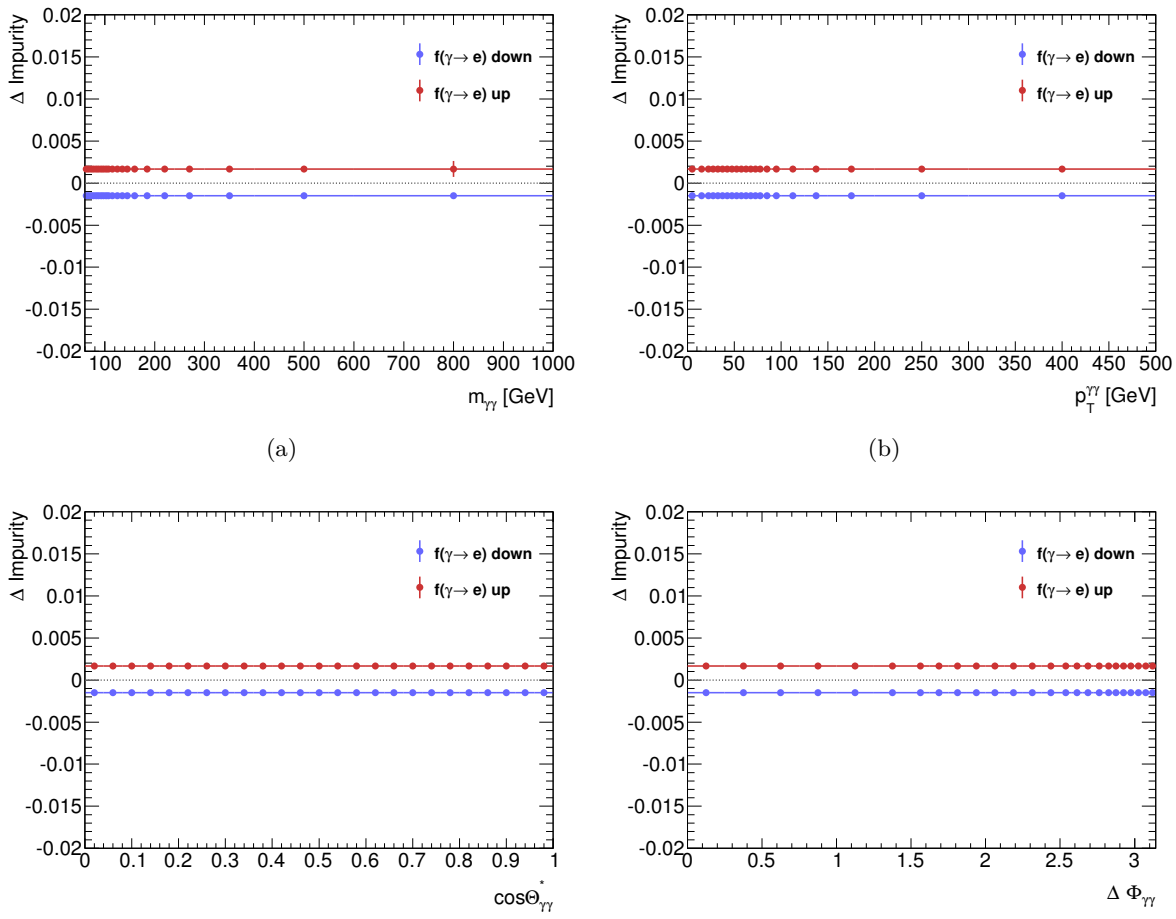
Figure F.1: Relative change in the estimated impurities estimated as a 100% variation of the $\gamma \to e$ fake rate up and down, as a function of the diphoton observables for the $0-jet$ category.



(a) (b)

(c) (d)

Figure F.2: Relative change in the estimated impurities estimated as a 100% variation of the $\gamma \to e$ fake rate up and down, as a function of the diphoton observables for the $1-jet$ category.



(a) (b)

(c) (d)

Figure F.3: Relative change in the estimated impurities estimated as a 100% variation of the $\gamma \to e$ fake rate up and down, as a function of the diphoton observables for the $\geq 3 - jet$ category.

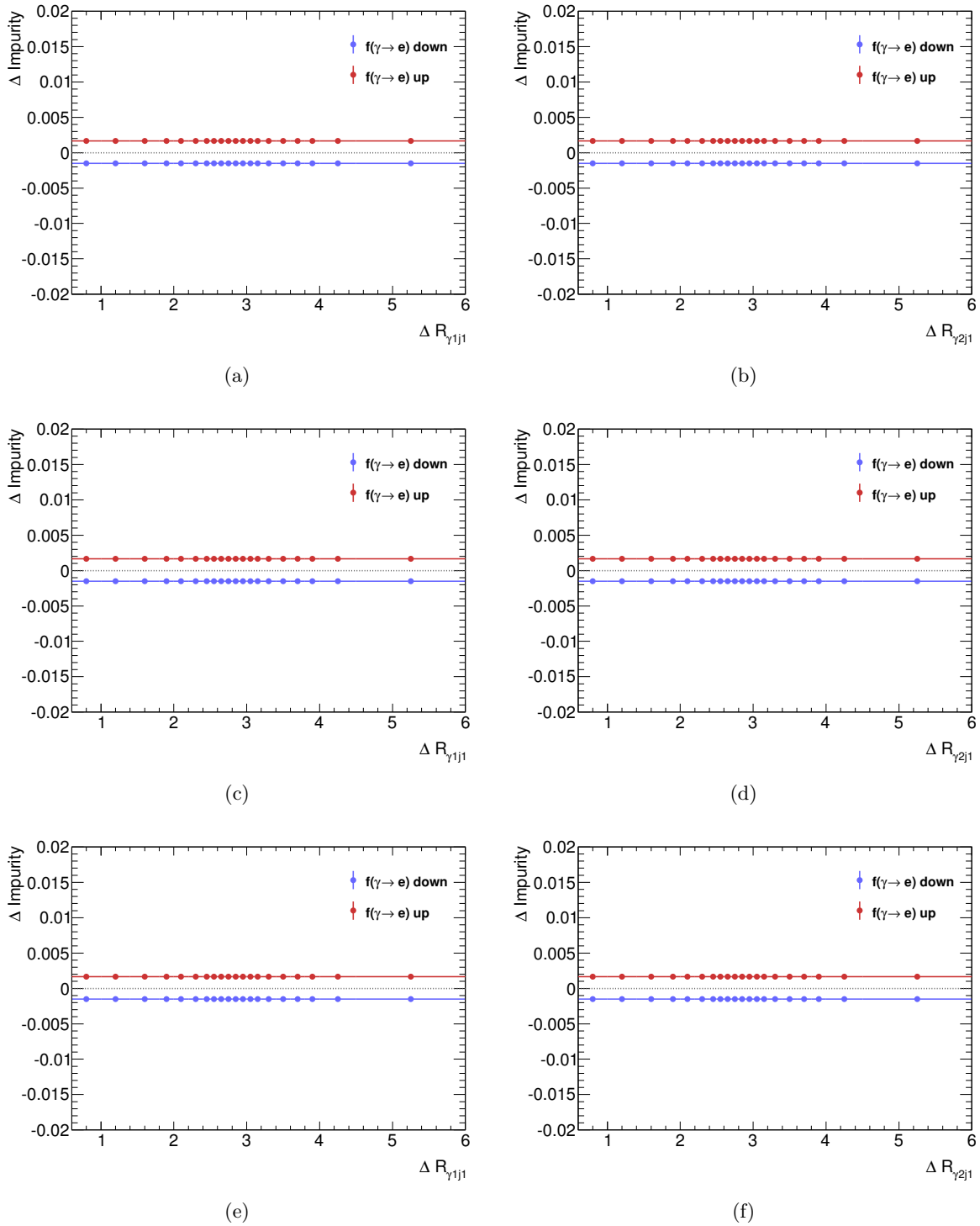
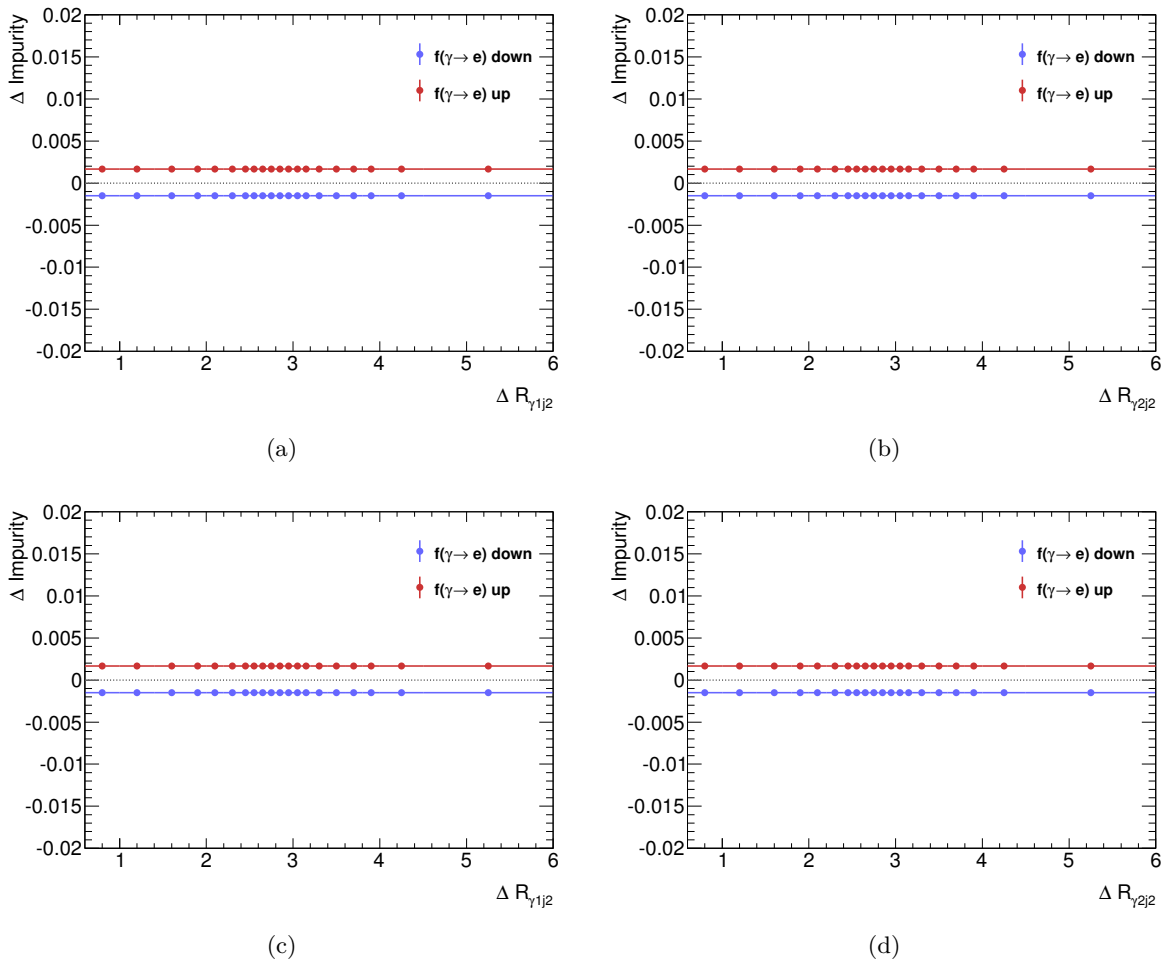


Figure F.4: Relative change in the estimated impurities estimated as a 100% variation of the $\gamma \rightarrow e$ fake rate up and down, as a function of the R-separation between the photons and the leading jet for the $1 - jet$ category.



(a) (b)

(c) (d)

Figure F.5: Relative change in the estimated impurities estimated as a 100% variation of the $\gamma \to e$ fake rate up and down, as a function of the R-separation between the photons and the subleading jet for the $2 - jet$ category.

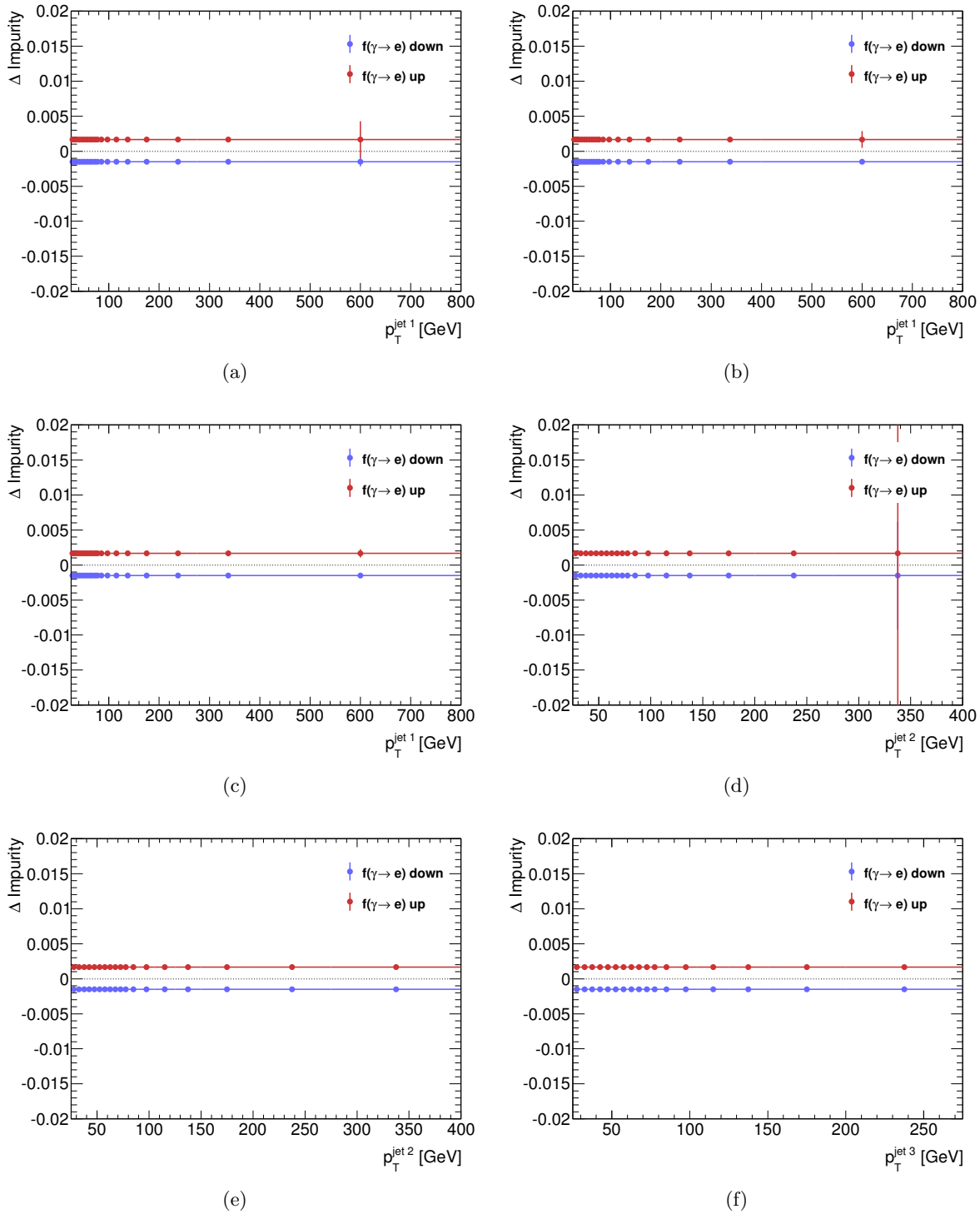


Figure F.6: Relative change in the estimated impurities estimated as a 100% variation of the $\gamma \to e$ fake rate up and down, as a function of the p_{T}^{jet}. (a) $1-jet$, (b),(d) $2-jet$, (c),(e),(f) $\geq 3-jet$ category.

G

Diphoton yields before and after electron background subtraction

This Appendix contains plots of the diphoton yields extracted from the 2x2DSB method and the diphoton yields after the electron background subtraction, described in Section 7.4.2, as a function of all observables except the ones already shown in the main text.

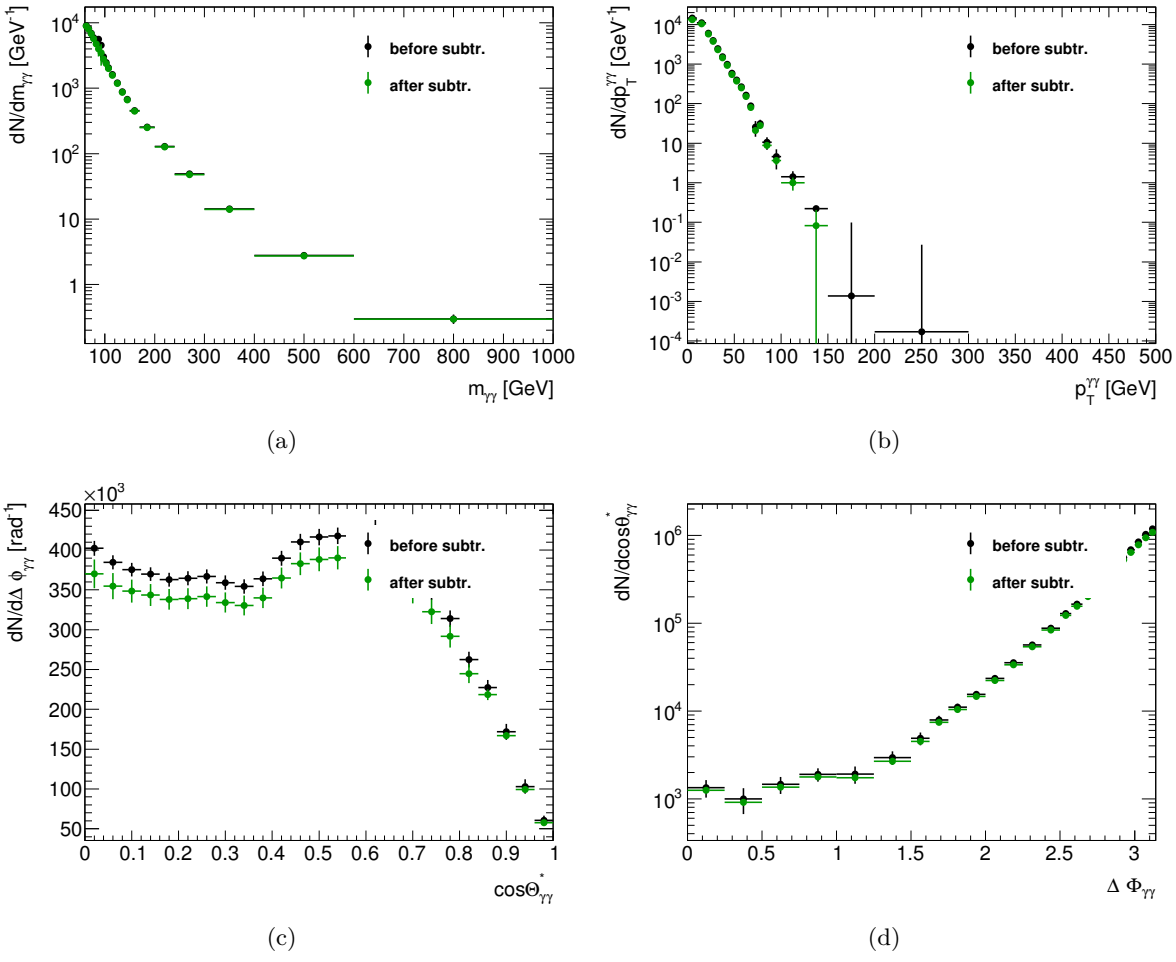


Figure G.1: 2x2DSB diphoton yields (black) and yields after electron background subtraction (green), as a function of the diphoton observables for the $0-jet$ category.

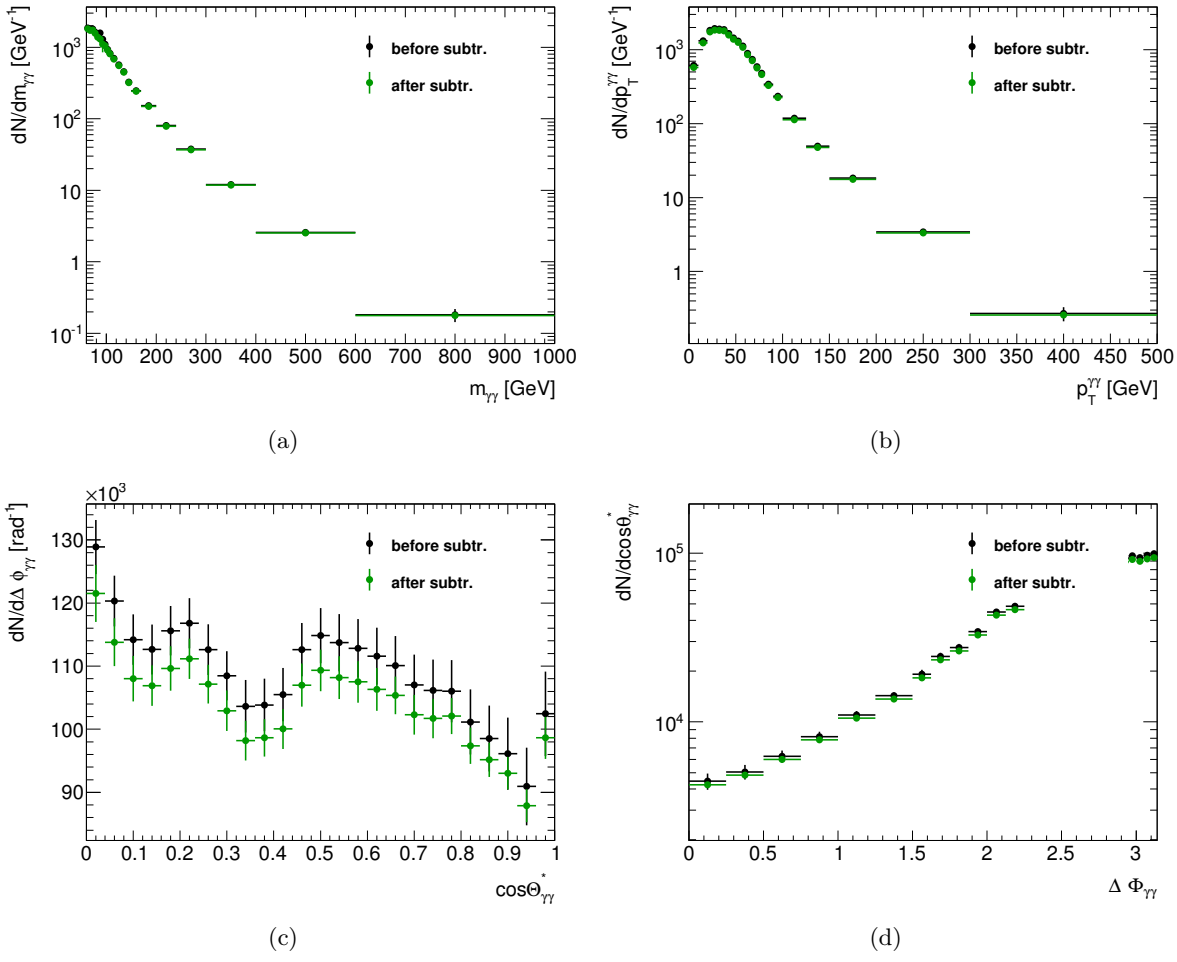


Figure G.2: 2x2DSB diphoton yields (black) and yields after electron background subtraction (green), as a function of the diphoton observables for the $1-jet$ category.

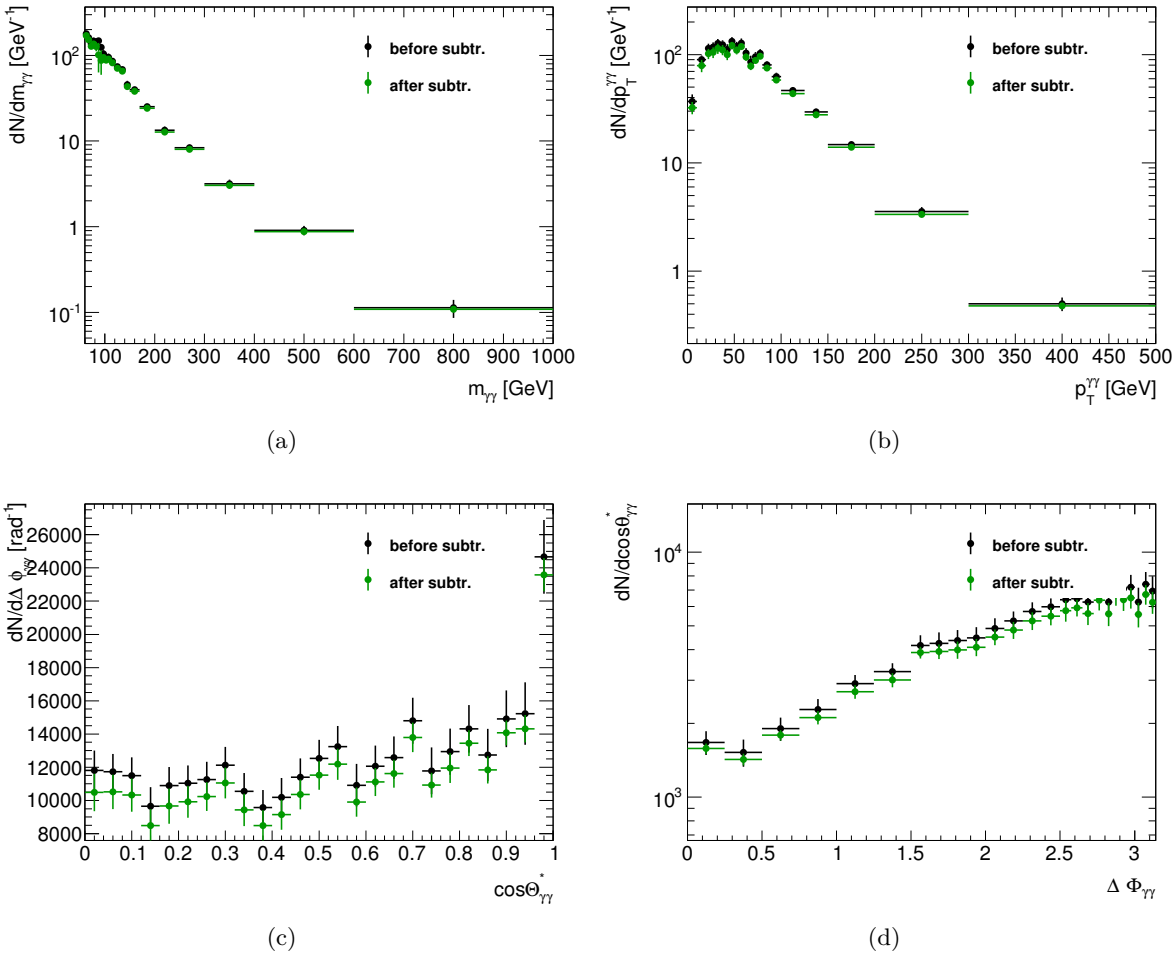


Figure G.3: 2x2DSB diphoton yields (black) and yields after electron background subtraction (green), as a function of the diphoton observables for the $\geq 3-jet$ category.

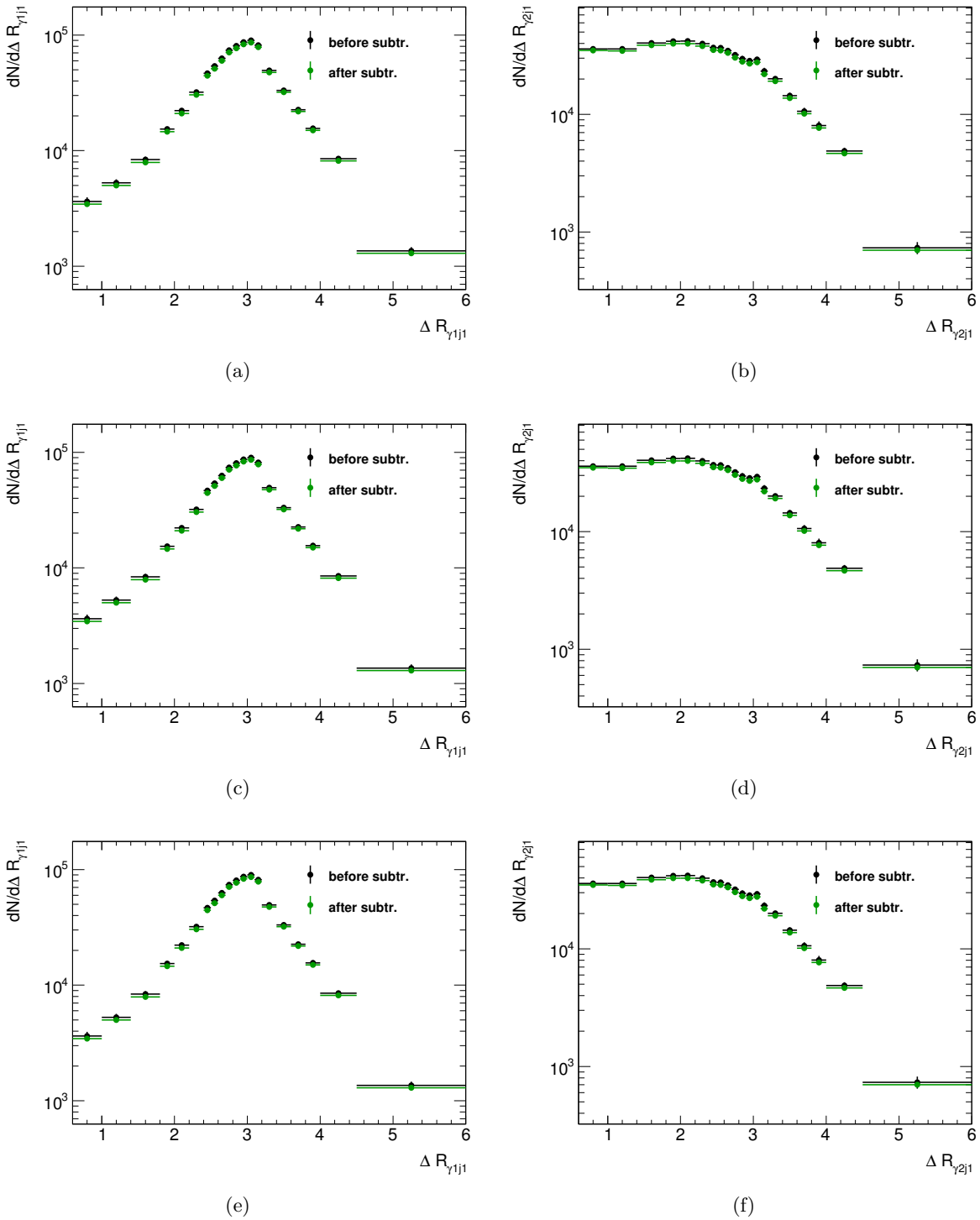
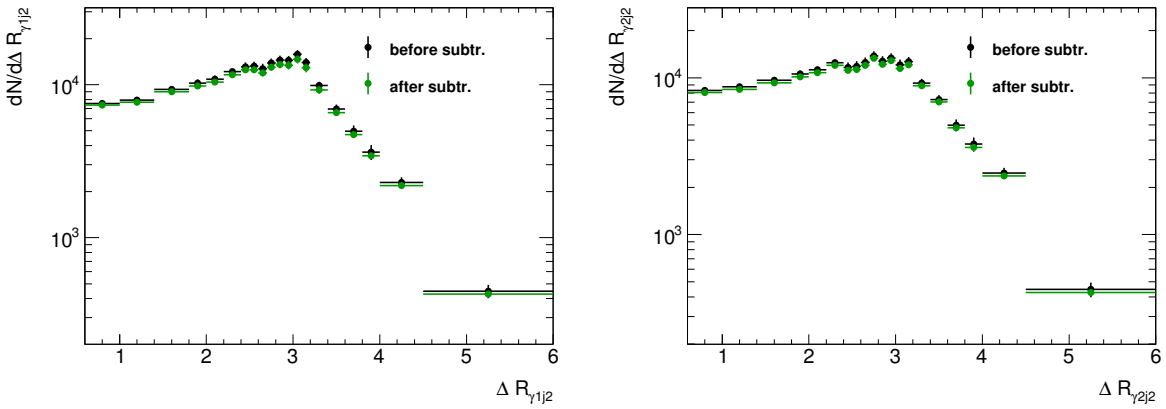
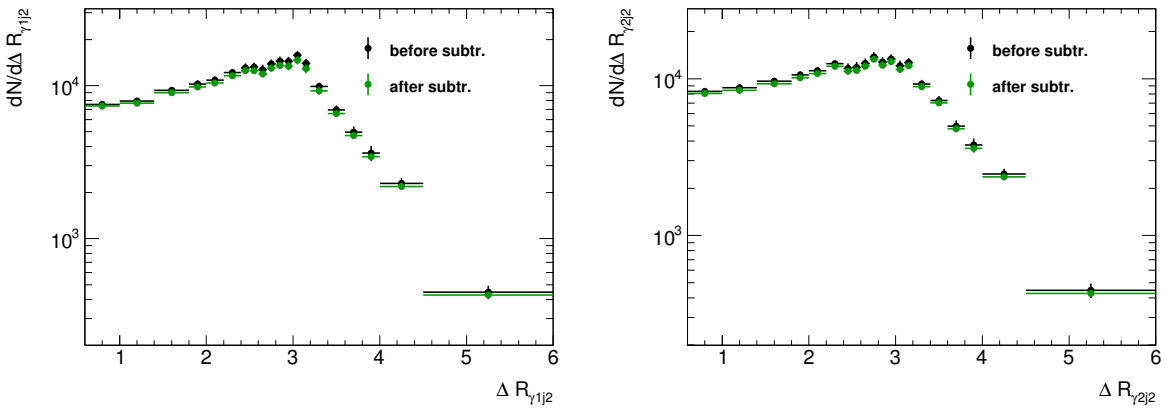


Figure G.4: 2x2DSB diphoton yields (black) and yields after electron background subtraction (green), as a function of the R-separation between the photons and the leading jet for the $1-jet$ category.



(a) (b)



(c) (d)

Figure G.5: 2x2DSB diphoton yields (black) and yields after electron background subtraction (green), as a function of the R-separation between the photons and the subleading jet for the $2-jet$ category.

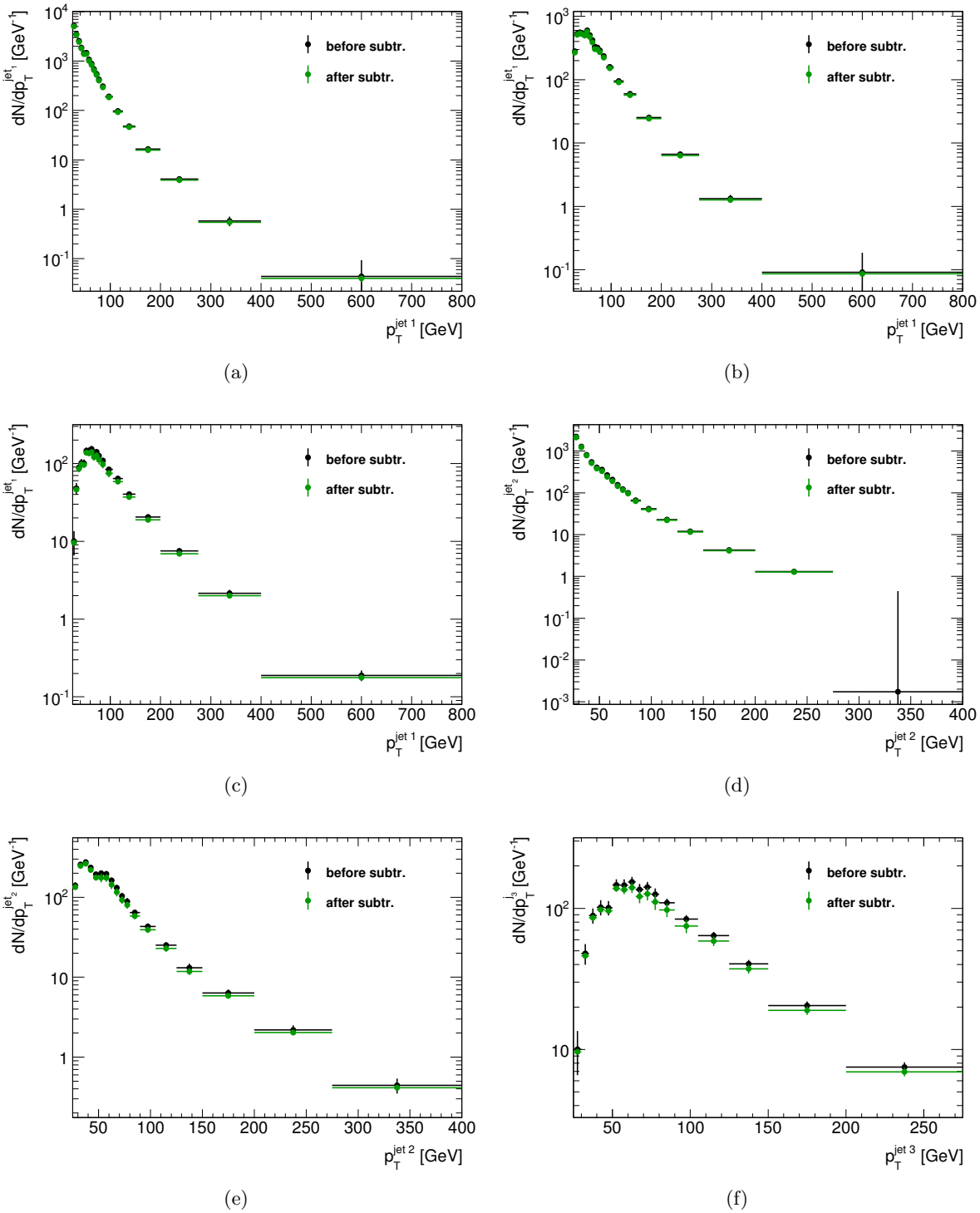


Figure G.6: 2x2DSB diphoton yields (black) and yields after electron background subtraction (green), as a function of the p_{T}^{jet}. (a) $1-jet$, (b),(d) $2-jet$, (c),(e),(f) $\geq 3-jet$ category.

Bibliography

[1] ATLAS Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*, Phys.Rev. **D90** (2014) no. 5, 052004, arXiv:1406.3827 [hep-ex].

[2] HTTPS://EN.WIKIPEDIA.ORG/WIKI/STANDARD_MODEL#/MEDIA/FILE:STANDARD_MODEL_OF_ELEMENTARY_PARTICLES.SVG.

[3] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*. Advanced book classics. Addison-Wesley Publishing Company, 1995. http://books.google.fr/books?id=i35LALN0GosC.

[4] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) no. 4, 579 – 588. http://www.sciencedirect.com/science/article/pii/0029558261904692.

[5] A. Salam and J. Ward, *Weak and electromagnetic interactions*, Il Nuovo Cimento Series 10 **11** (1959) no. 4, 568–577. http://dx.doi.org/10.1007/BF02726525.

[6] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (Nov, 1967) 1264–1266. http://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[7] P. Minkowski, *$\mu \to e\gamma$ at a rate of one out of 10^9 muon decays?*, Physics Letters B **67** (Apr, 1977) 421–428.

[8] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (Aug, 1964) 321–323. http://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[9] P. Higgs, *Broken symmetries, massless particles and gauge fields*, Physics Letters **12** (1964) no. 2, 132 – 133. http://www.sciencedirect.com/science/article/pii/0031916364911369.

[10] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (Oct, 1964) 508–509. http://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[11] G. S. Guralnik, C. R. Hagen, and T. W. Kibble, *Global Conservation Laws and Massless Particles*, Physical Review Letters **13** (nov, 1964) 585–587.

[12] V. Ginzburg and L. Landau Zh. Eksp. Teor. Fiz. **20** (1950) 1064.

[13] Y. Nambu, *Quasi-Particles and Gauge Invariance in the Theory of Superconductivity*, Phys. Rev. **117** (Feb, 1960) 648–663. http://link.aps.org/doi/10.1103/PhysRev.117.648.

[14] J. Goldstone, *Field theories with Superconductor solutions*, Il Nuovo Cimento **19** (1961) no. 1, 154–164. http://dx.doi.org/10.1007/BF02812722.

[15] M. Magg and C. Wetterich, *Neutrino mass problem and gauge hierarchy*, Physics Letters B **94** (1980) no. 1, 61 – 64. http://www.sciencedirect.com/science/article/pii/0370269380908254.

[16] J. Schechter and J. W. F. Valle, *Neutrino masses in SU(2)⊗U(1) theories*, Phys. Rev. D **22** (nov, 1980) 2227–2235.

[17] B. W. Lee, C. Quigg, and H. B. Thacker, *Weak interactions at very high energies: The role of the Higgs-boson mass*, Phys. Rev. D **16** (Sep, 1977) 1519–1531. http://link.aps.org/doi/10.1103/PhysRevD.16.1519.

[18] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv:1207.7214 [hep-ex].

[19] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30, arXiv:1207.7235 [hep-ex].

[20] ATLAS, CMS Collaboration, G. Aad et al., *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, arXiv:1503.07589 [hep-ex].

[21] LHC Higgs Cross Section Working Group Collaboration, S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347 [hep-ph].

[22] ATLAS Collaboration, G. Aad et al., *Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector*, arXiv:1412.2641 [hep-ex].

[23] ATLAS Collaboration, G. Aad et al., *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector*, JHEP **1501** (2015) 069, arXiv:1409.6212 [hep-ex].

[24] ATLAS Collaboration, G. Aad et al., *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, arXiv:1501.04943 [hep-ex].

[25] CMS Collaboration, S. Chatrchyan et al., *Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons*, JHEP **1405** (2014) 104, arXiv:1401.5041 [hep-ex].

[26] CMS Collaboration, S. Chatrchyan et al., *Evidence for the direct decay of the 125 GeV Higgs boson to fermions*, Nature Phys. **10** (2014) 557–560, arXiv:1401.6527 [hep-ex].

[27] *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2015-007, CERN, Geneva, Mar, 2015. https://cds.cern.ch/record/2002212.

[28] A. Arbey, M. Battaglia, A. Djouadi, F. Mahmoudi, and J. Quevillon, *Implications of a 125 GeV Higgs for supersymmetric models*, Physics Letters B **708** (2012) no. 1 - 2, 162 – 169. http://www.sciencedirect.com/science/article/pii/S0370269312000810.

[29] P. Draper, P. Meade, M. Reece, and D. Shih, *Implications of a 125 GeV Higgs boson for the MSSM and low-scale supersymmetry breaking*, Phys. Rev. D **85** (May, 2012) 095007. http://link.aps.org/doi/10.1103/PhysRevD.85.095007.

[30] P. Nath, *Supersymmetry after the Higgs*, arXiv:1501.01679 [hep-ph].

[31] M. Baak and R. Kogler, *The global electroweak Standard Model fit after the Higgs discovery*, arXiv:1306.0571 [hep-ph].

[32] 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, . http://dx.doi.org/10.1140/epjc/s10052-014-3046-5.

[33] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, JHEP **1208** (2012) 098, arXiv:1205.6497 [hep-ph].

[34] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (Jul, 1964) 138–140. http://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[35] W.-S. Hou, *Source of CP Violation for the Baryon Asymmetry of the Universe*, Chin.J.Phys. **47** (2009) 134, arXiv:0803.1234 [hep-ph].

[36] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, arXiv:hep-ex/9807003.

[37] J. Conrad, *Neutrino Experiments*, arXiv:0708.2446 [hep-ex].

[38] H. Georgi and S. L. Glashow, *Unity of All Elementary-Particle Forces*, Physical Review Letters **32** (feb, 1974) 438–441.

[39] C. Rovelli, *Zakopane lectures on loop gravity*, PoS **QGQGS2011** (2011) 003, arXiv:1102.3660 [gr-qc].

[40] Y. Nambu, *Quark Model and the Factorization of the Veneziano Model*, Proc. Intern. Conf. on Symmetries and Quark Models (1969) .

[41] Planck Collaboration, P. Ade et al., *Planck 2013 results. I. Overview of products and scientific results*, Astron.Astrophys. **571** (2014) A1, arXiv:1303.5062 [astro-ph.CO].

[42] J.-L. Gervais and B. Sakita, *Field theory interpretation of supergauges in dual models*, Nuclear Physics B **34** (nov, 1971) 632–639.

[43] Y. A. Golfand and E. P. Likhtman, *Extension of the Algebra of Poincare Group Generators and Violation of P Invariance*, JETP Letters-USSR **13** (1971) 323.

[44] D. V. Volkov and V. Akulov Pizma Zh. Eksp. Theor. Fiz. **16** (1972) 621.

[45] A. Collaboration, *Overall SUSY search summary*, ATLAS SUSY Public Plots (2015) , 1303.5062.

[46] A. Lahanas, *LSP as a Candidate for Dark Matter*, Lect.Notes Phys. **720** (2007) 35–68, arXiv:hep-ph/0607301 [hep-ph].

[47] C. Han, K.-i. Hikasa, L. Wu, J. M. Yang, and Y. Zhang, *Current experimental bounds on stop mass in natural SUSY*, JHEP **1310** (2013) 216, arXiv:1308.5307 [hep-ph].

[48] D. Chowdhury, R. M. Godbole, K. A. Mohan, and S. K. Vempati, *Charge and Color Breaking Constraints in MSSM after the Higgs Discovery at LHC*, JHEP **1402** (2014) 110, arXiv:1310.1932 [hep-ph].

[49] F. King, S. M. Mühlleitner, R. Nevzorov, and K. Walz, *Discovery prospects for NMSSM Higgs bosons at the high-energy Large Hadron Collider*, Phys. Rev. D **90** (Nov, 2014) 095014. http://link.aps.org/doi/10.1103/PhysRevD.90.095014.

[50] O. Eberhardt, U. Nierste, and M. Wiebusch, *Status of the two-Higgs-doublet model of type II*, JHEP **1307** (2013) 118, arXiv:1305.1649 [hep-ph].

[51] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, Phys.Rev.Lett. **83** (1999) 3370–3373, arXiv:hep-ph/9905221 [hep-ph].

[52] I. J. Aitchison, *Supersymmetry and the MSSM: An Elementary introduction*, arXiv:hep-ph/0505105 [hep-ph].

[53] P. G. Mercadante, *Search for supersymmetry at the LHC*, Brazilian Journal of Physics **34** (12, 2004) 1424 – 1433. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-97332004000700020&nrm=iso.

[54] J. A. Casas, J. M. Moreno, S. Robles, K. Rolbiecki, and B. Zaldivar, *What is a Natural SUSY scenario?*, arXiv:1407.6966 [hep-ph].

[55] P. Fayet, *Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino*, Nuclear Physics B **90** (1975) 104–124.

[56] J.-P. Derendinger and C. A. Savoy, *Quantum effects and SU(2)xU(1) breaking in supergravity gauge theories*, Nuclear Physics B **237** (may, 1984) 307–328.

[57] M. Badziak, M. Olechowski, and S. Pokorski, *125 GeV Higgs and enhanced diphoton signal of a light singlet-like scalar in NMSSM*, PoS **EPS-HEP2013** (2013) 257, arXiv:1310.4518 [hep-ph].

[58] LEP Working Group for Higgs boson searches, ALEPH, DELPHI, L3, OPAL Collaboration, R. Barate et al., *Search for the standard model Higgs boson at LEP*, Phys.Lett. **B565** (2003) 61–75, arXiv:hep-ex/0306033 [hep-ex].

[59] N. Craig and S. Thomas, *Exclusive signals of an extended Higgs sector*, Journal of High Energy Physics **11** (nov, 2012) 83, arXiv:1207.4835 [hep-ph].

[60] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rep. **516** (jul, 2012) 1–102, arXiv:1106.0034 [hep-ph].

[61] B. Grinstein and P. Uttayarat, *Carving Out Parameter Space in Type-II Two Higgs Doublets Model*, JHEP **1306** (2013) 094, arXiv:1304.0028 [hep-ph].

[62] G. Cacciapaglia, A. Deandrea, G. D. La Rochelle, and J.-B. Flament, *Searching for a lighter Higgs boson: Parametrization and sample tests*, Phys.Rev. **D91** (2015) no. 1, 015012, arXiv:1311.5132 [hep-ph].

[63] Particle Data Group Collaboration, K. Olive et al., *Review of Particle Physics*, Chin.Phys. **C38** (2014) 090001.

[64] S. Drell and T.-M. Yan, *Partons and their Applications at High-Energies*, Annals Phys. **66** (1971) 578.

[65] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126** (1977) 298.

[66] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, Phys.Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[67] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur.Phys.J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[68] Z. Bern, L. Dixon, and C. Schmidt, *Isolating a light Higgs boson from the diphoton background at the CERN LHC*, Phys. Rev. D **66** (Oct, 2002) 074018. http://link.aps.org/doi/10.1103/PhysRevD.66.074018.

[69] T. Binoth, J. 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, arXiv:hep-ph/9911340 [hep-ph].

[70] S. Catani, F. Krauss, R. Kuhn, and B. Webber, *QCD matrix elements + parton showers*, JHEP **0111** (2001) 063, arXiv:hep-ph/0109231 [hep-ph].

[71] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput.Phys.Commun. **178** (2008) 852, arXiv:0710.3820 [hep-ph].

[72] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007, arXiv:0811.4622 [hep-ph].

[73] Z. Nagy and D. E. Soper, *Matching parton showers to NLO computations*, JHEP **0510** (2005) 024, arXiv:hep-ph/0503053 [hep-ph].

[74] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].

[75] S. A. et. al., *Geant4 - a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) no. 3, 250 – 303. http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[76] 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 [hep-ph].

[77] S. Frixione, *Isolated photons in perturbative QCD*, Phys.Lett. **B429** (1998) 369–374, arXiv:hep-ph/9801442 [hep-ph].

[78] J. Guillet, *Photon Production at Hadronic Colliders*, Winter Workshop (2011) .

[79] V. Del Duca, F. Maltoni, Z. Nagy, and Z. Trocsanyi, *QCD radiative corrections to prompt diphoton production in association with a jet at hadron colliders*, JHEP **0304** (2003) 059, arXiv:hep-ph/0303012 [hep-ph].

[80] T. Gehrmann, N. Greiner, and G. Heinrich, *Photon isolation effects at NLO in $\gamma\gamma$ + jet final states in hadronic collisions*, JHEP **1306** (2013) 058, arXiv:1303.0824 [hep-ph].

[81] G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, et al., *Automated One-Loop Calculations with GoSam*, Eur.Phys.J. **C72** (2012) 1889, arXiv:1111.2034 [hep-ph].

[82] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, et al., *MadGraph/MadEvent v4: The New Web Generation*, JHEP **0709** (2007) 028, arXiv:0706.2334 [hep-ph].

[83] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, et al., *Parton distributions with LHC data*, Nucl.Phys. **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[84] T. Gehrmann, N. Greiner, and G. Heinrich, *Precise QCD predictions for the production of a photon pair in association with two jets*, Phys.Rev.Lett. **111** (2013) 222002, arXiv:1308.3660 [hep-ph].

[85] J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, et al., *CT10 next-to-next-to-leading order global analysis of QCD*, Phys.Rev. **D89** (2014) no. 3, 033009, arXiv:1302.6246 [hep-ph].

[86] Z. Bern, L. Dixon, F. Febres Cordero, S. Hoeche, H. Ita, et al., *Next-to-leading order diphoton+2-jet production at the LHC*, arXiv:1312.0592 [hep-ph].

[87] C. Berger, Z. Bern, L. Dixon, F. Febres Cordero, D. Forde, et al., *An Automated Implementation of On-Shell Methods for One-Loop Amplitudes*, Phys.Rev. **D78** (2008) 036003, arXiv:0803.4180 [hep-ph].

[88] CMS Collaboration, V. Khachatryan et al., *Measurement of the Isolated Prompt Photon Production Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV*, Phys.Rev.Lett. **106** (2011) 082001, arXiv:1012.0799 [hep-ex].

[89] Z. Bern, G. Diana, L. Dixon, F. Febres Cordero, S. Hoche, et al., *Driving Missing Data at Next-to-Leading Order*, Phys.Rev. **D84** (2011) 114002, arXiv:1106.1423 [hep-ph].

[90] S. Badger, A. Guffanti, and V. Yundin, *Next-to-leading order QCD corrections to di-photon production in association with up to three jets at the Large Hadron Collider*, Journal of High Energy Physics **2014** (2014) no. 3, . http://dx.doi.org/10.1007/JHEP03%282014%29122.

[91] *NJET 2.0*, https://bitbucket.org/njet/njet/downloads.

[92] D0 Collaboration, V. M. Abazov et al., *Measurement of direct photon pair production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Lett. **B690** (2010) 108–117, arXiv:1002.4917 [hep-ex].

[93] D0 Collaboration, V. Abazov et al., *Measurement of the differential cross sections for isolated direct photon pair production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys.Lett. **B725** (2013) 6–14, arXiv:1301.4536 [hep-ex].

[94] 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 [hep-ex].

[95] 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 [hep-ex].

[96] 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 **1301** (2013) 086, arXiv:1211.1913 [hep-ex].

[97] CMS Collaboration, S. Chatrchyan et al., *Measurement of the Production Cross Section for Pairs of Isolated Photons in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **01** (2012) 133, arXiv:1110.6461 [hep-ex].

[98] CMS Collaboration, S. Chatrchyan et al., *Measurement of differential cross sections for the production of a pair of isolated photons in pp collisions at $\sqrt{s} = 7$ TeV*, Eur.Phys.J. **C74** (2014) no. 11, 3129, arXiv:1405.7225 [hep-ex].

[99] 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 [hep-ex].

[100] ATLAS Collaboration, G. Aad et al., *Dynamics of isolated-photon plus jet production in pp collisions at $\sqrt{(s)} = 7$ TeV with the ATLAS detector*, Nucl.Phys. **B875** (2013) 483–535, arXiv:1307.6795 [hep-ex].

[101] CMS Collaboration, S. Chatrchyan et al., *Measurement of the triple-differential cross section for photon+jets production in proton-proton collisions at $\sqrt{s}$=7 TeV*, JHEP **06** (2014) 009, arXiv:1311.6141 [hep-ex].

[102] D0 Collaboration, V. M. Abazov et al., *Measurement of the differential cross-section for the production of an isolated photon with associated jet in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV*, Phys. Lett. **B666** (2008) 435–445, arXiv:0804.1107 [hep-ex].

[103] D0 Collaboration, V. M. Abazov et al., *Measurement of the differential cross section of photon plus jet production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys.Rev. **D88** (2013) 072008, arXiv:1308.2708 [hep-ex].

[104] CDF Collaboration, T. Aaltonen et al., *Measurement of the cross section for direct-photon production in association with a heavy quark in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV*, Phys.Rev.Lett. **111** (2013) no. 4, 042003, arXiv:1303.6136 [hep-ex].

[105] LEP Collaboration, A. Rosca, *Fermiophobic Higgs bosons at LEP*, Nucl.Phys.Proc.Suppl. **117** (2003) 743–745, arXiv:hep-ex/0212038 [hep-ex].

[106] D0 Collaboration, V. M. Abazov et al., *Search for a Higgs boson in diphoton final states with the D0 detector in 9.6 fb-1 of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys.Rev. **D88** (2013) no. 5, 052007, arXiv:1301.5358 [hep-ex].

[107] CDF Collaboration, T. Aaltonen et al., *Search for a Higgs boson in the diphoton final state using the full CDF data set from proton-antiproton collisions at $\sqrt{s} = 1.96$ TeV*, Phys.Lett. **B717** (2012) 173–181, arXiv:1207.6386 [hep-ex].

[108] ATLAS Collaboration, *Search for the Higgs boson in the diphoton final state with 38 pb^{-1} of data recorded by the ATLAS detector in proton-proton collisions at $\sqrt{s}$=7 TeV*, Tech. Rep. ATLAS-CONF-2011-025, CERN, Geneva, Mar, 2011. http://cds.cern.ch/record/1336758.

[109] ATLAS Collaboration, *Search for the Higgs Boson in the Diphoton Channel with the ATLAS Detector using 209 pb^{-1} of 7 TeV data taken in 2011*, Tech. Rep. ATLAS-CONF-2011-085, CERN, Geneva, Jun, 2011. http://cds.cern.ch/record/1356193.

[110] ATLAS Collaboration, G. Aad et al., *Search for the Standard Model Higgs boson in the two photon decay channel with the ATLAS detector at the LHC*, Phys.Lett. **B705** (2011) 452–470, arXiv:1108.5895 [hep-ex].

[111] ATLAS Collaboration, *Search for the Standard Model Higgs boson in the diphoton decay channel with 4.9 fb^{-1} of ATLAS data at $\sqrt{s} = 7$ TeV*, Tech. Rep. ATLAS-CONF-2011-161, CERN, Geneva, Dec, 2011. https://cds.cern.ch/record/1406356.

[112] CMS Collaboration, *Search for a Higgs boson decaying into two photons in the CMS detector*, Tech. Rep. CMS-PAS-HIG-11-010, CERN, Geneva, 2011. http://cds.cern.ch/record/1369553.

[113] CMS Collaboration, *Search for a Higgs boson decaying into two photons in the CMS detector*, Tech. Rep. CMS-PAS-HIG-11-030, CERN, Geneva, 2011. http://cds.cern.ch/record/1406346.

[114] CMS Collaboration, *Search for the fermiophobic model Higgs boson decaying into two photons*, Tech. Rep. CMS-PAS-HIG-12-002, CERN, Geneva, 2012. http://cds.cern.ch/record/1430018.

[115] CMS Collaboration, *Evidence for a new state decaying into two photons in the search for the standard model Higgs boson in pp collisions*, Tech. Rep. CMS-PAS-HIG-12-015, CERN, Geneva, 2012. http://cds.cern.ch/record/1460419.

[116] ATLAS Collaboration, N. Rompotis, *Beyond Standard Model Higgs boson physics with the ATLAS experiment at the LHC*, arXiv:1408.3521 [hep-ex].

[117] CMS Collaboration, *Search for an Higgs Like resonance in the diphoton mass spectra above 150 GeV with 8 TeV data*, . https://cds.cern.ch/record/1714076?ln=en.

[118] ATLAS Collaboration, G. Aad et al., *Search for high-mass diphoton resonances in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, arXiv:1504.05511 [hep-ex].

[119] *CERN releases analysis of LHC incident*, . http://press.web.cern.ch/press-releases/2008/10/cern-releases-analysis-lhc-incident.

[120] *LHC: Preparations for collisions at 13 TeV*, . http://home.web.cern.ch/about/updates/2015/04/lhc-preparations-collisions-13-tev.

[121] CERN. https://espace.cern.ch/acc-tec-sector/default.aspx.

[122] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, et al., *LHC Design Report. 1. The LHC Main Ring*, .

[123] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) no. 08, S08003. http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[124] T. L. Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008) no. 08, S08005. http://stacks.iop.org/1748-0221/3/i=08/a=S08005.

[125] T. C. Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) no. 08, S08004. http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[126] T. A. Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) no. 08, S08002. http://stacks.iop.org/1748-0221/3/i=08/a=S08002.

[127] T. T. Collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) no. 08, S08007. http://stacks.iop.org/1748-0221/3/i=08/a=S08007.

[128] LHCf Collaboration, O. e. a. Adriani, *LHCf experiment: Technical Design Report*. Technical Design Report LHCf. CERN, Geneva, 2006.

[129] MoEDAL Collaboration, J. e. a. Pinfold, *Technical Design Report of the MoEDAL Experiment*, Tech. Rep. CERN-LHCC-2009-006. MoEDAL-TDR-001, CERN, Geneva, Jun, 2009.

[130] S. van der Meer, *Calibration of the effective beam height in the ISR*, Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968.

[131] 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 [hep-ex].

[132] M. Lamont, *Status of the LHC*, Journal of Physics: Conference Series **455** (2013) no. 1, 012001. http://stacks.iop.org/1742-6596/455/i=1/a=012001.

[133] ATLAS Collaboration, G. Aad et al., *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC*, Eur.Phys.J. **C71** (2011) 1630, arXiv:1101.2185 [hep-ex].

[134] J. Pequenao, *Computer generated image of the whole ATLAS detector*, Mar, 2008.

[135] ATLAS Collaboration, *ATLAS: letter of intent for a general-purpose pp experiment at the large hadron collider at CERN*. Letter of Intent. CERN, Geneva, 1992.

[136] A. e. a. Yamamoto, *Progress in ATLAS central solenoid magnet*, IEEE T. Appl. Supercond. **10** (2000) 353.

[137] ATLAS Collaboration, J. P. Badiou, J. Beltramelli, J. M. Baze, and J. Belorgey, *ATLAS barrel toroid: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331065. Electronic version not available.

[138] ATLAS Collaboration, *ATLAS end-cap toroids: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331066. Electronic version not available.

[139] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1999. Electronic version not available.

[140] ATLAS Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, arXiv:0901.0512 [hep-ex].

[141] *Alignment of the ATLAS Inner Detector and its Performance in 2012*, Tech. Rep. ATLAS-CONF-2014-047, CERN, Geneva, Jul, 2014. https://cds.cern.ch/record/1741021.

[142] 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 [hep-ex].

[143] J. Pequenao, *Computer Generated image of the ATLAS calorimeter*, Mar, 2008.

[144] HTTPS://TWIKI.CERN.CH/TWIKI/BIN/VIEW/ATLASPUBLIC/LARCALOPUBLICRESULTSDETSTATUS.

[145] G. Aad, B. Abbott, J. Abdallah, A. A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, and et al., *Drift Time Measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using Cosmic Muons*, European Physical Journal C **70** (Dec., 2010) 755–785, arXiv:1002.4189 [physics.ins-det].

[146] A. Artamonov, *The ATLAS Forward Calorimeter*, Journal of Instrumentation **3** (2008) no. 02, P02010. http://stacks.iop.org/1748-0221/3/i=02/a=P02010.

[147] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.

[148] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. distribution.

[149] *ATLAS data quality information for Run 1*, . https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010.

[150] G. Aad, B. Abbott, J. Abdallah, A. A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, and et al., *Drift Time Measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using Cosmic Muons*, European Physical Journal C **70** (dec, 2010) 755–785, arXiv:1002.4189 [physics.ins-det].

[151] T. L. A. F. E. E. collaboration (N J Buchanan et al), *ATLAS liquid argon calorimeter front end electronics*, Journal of Instrumentation **3** (2008) no. 09, P09003. http://stacks.iop.org/1748-0221/3/i=09/a=P09003.

[152] T. L. A. B. E. E. collaboration (A Bazan et al), *ATLAS liquid argon calorimeter back end electronics*, Journal of Instrumentation **2** (2007) no. 06, P06002. http://stacks.iop.org/1748-0221/2/i=06/a=P06002.

[153] Cleland, W.E and Stern, E.G., *Signal processing considerations for liquid ionization calorimeters in a high rate environment*, Nuclear Instruments and Methods in Physics Research **A3382** (1994) 467–497.

[154] ATLAS Collaboration, H. A. et. al., *Performance of the electronic readout of the ATLAS liquid argon calorimeters*, Journal of Instrumentation **5** (2010) no. 09, P09003. http://stacks.iop.org/1748-0221/5/i=09/a=P09003.

[155] M. Delmastro, H. Ma, and G. Unal, *Liquid Argon Optimal Filtering Coefficients Optimization for 2012 data*, Tech. Rep. ATL-COM-LARG-2013-009, CERN, Geneva, Apr, 2013.

[156] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008.

[157] ATLAS Collaboration, E. B.-T. et al., *A framework for vertex reconstruction in the ATLAS experiment at LHC*, Journal of Physics: Conference Series **219** (2010) no. 3, 032019. http://stacks.iop.org/1742-6596/219/i=3/a=032019.

[158] ATLAS Collaboration, *Electron reconstruction efficiency*, . https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20120611/ElectronEfficiency2012/ATL-COM-PHYS-2011-783/ATL-COM-PHYS-2011-783.pdf.

[159] M. Delmastro, S. Gleyzer, C. Hengler, M. Jimenez, T. Koffas, M. Kuna, K. Liu, Y. Liu, G. Marchiori, E. Petit, M. Pitt, E. Soldatov, and K. Tackmann, *Photon identification efficiency measurements with the ATLAS detector using LHC Run 1 data*, Tech. Rep. ATL-COM-PHYS-2014-949, CERN, Geneva, Aug, 2014. https://cds.cern.ch/record/1747242.

[160] ATLAS Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, arXiv:0901.0512 [hep-ex].

[161] ATLAS Collaboration, G. Aad et al., *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur.Phys.J. **C72** (2012) 1909, arXiv:1110.3174 [hep-ex].

[162] *Expected photon performance in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2011-007, CERN, Geneva, Apr, 2011. http://cds.cern.ch/record/1345329.

[163] ATLAS Collaboration, *Measurements of the photon identification efficiency with the ATLAS detector using 4.9 fb^{-1} of pp collision data collected in 2011*, Tech. Rep. ATLAS-CONF-2012-123, CERN, Geneva, Aug, 2012. https://cds.cern.ch/record/1473426.

[164] 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 TeV in 2012*, Tech. Rep. ATL-COM-PHYS-2014-542, CERN, Geneva, May, 2014. https://cds.cern.ch/record/1704917.

[165] ATLAS Collaboration, G. Aad et al., *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys.Rev. **D83** (2011) 052005, arXiv:1012.4389 [hep-ex].

[166] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **0804** (2008) 005, arXiv:0802.1188 [hep-ph].

[167] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008. https://cds.cern.ch/record/1099735.

[168] ATLAS Collaboration, *Calorimeter isolation with topoclusters versus pile-up*, . https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20120404/\TopoIsolation/ATL-COM-PHYS-2012-362/index.html.

[169] HTTPS://TWIKI.CERN.CH/TWIKI/BIN/VIEW/ATLASPROTECTED/LARCLEANINGANDOBJECTQUALITY#LARERROR_EVENT_FLAG.

[170] Q. Buat, J. Leveque, N. Lorenzo Martinez, F. Polci, and B. Trocme, *Definition of the photon cleaning procedure for 2011 data*, Tech. Rep. ATL-COM-PHYS-2012-134, CERN, Geneva, Feb, 2012. https://cds.cern.ch/record/1422649.

[171] A. Ali and G. Kramer, *Jets and QCD: A Historical Review of the Discovery of the Quark and Gluon Jets and its Impact on QCD*, Eur.Phys.J. **H36** (2011) 245–326, arXiv:1012.2288 [hep-ph].

[172] G. P. Salam and G. Soyez, *A Practical Seedless Infrared-Safe Cone jet algorithm*, JHEP **0705** (2007) 086, arXiv:0704.0292 [hep-ph].

[173] G. P. Salam, *Towards Jetography*, Eur.Phys.J. **C67** (2010) 637–686, arXiv:0906.1833 [hep-ph].

[174] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ph].

[175] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur.Phys.J. **C72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[176] ATLAS Collaboration, G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, arXiv:1406.0076 [hep-ex].

[177] 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. https://cds.cern.ch/record/1435196.

[178] D0 Collaboration, V. Abazov et al., *Measurement of the $p\bar{p} \to t\bar{t}$ production cross section at $\sqrt{s} = 1.96$ TeV in the fully hadronic decay channel.*, Phys.Rev. **D76** (2007) 072007, arXiv:hep-ex/0612040 [hep-ex].

[179] Z. Barnovska, M. F. Bessner, M. Delmastro, M. Fanti, E. A. Lamperti, E. Petit, K. Tackmann, and J. Terron, *Supporting Note: Measurement of the diphoton in association with jets cross sections with the ATLAS detector at $\sqrt{s} = 8$ TeV*, Tech. Rep. ATL-COM-PHYS-2015-721, CERN, Geneva, Jul, 2015. https://cds.cern.ch/record/2034485.

[180] D. W. Miller, A. Schwartzman, and D. Su, *Jet-Vertex Association Algorithm*, Tech. Rep. ATL-COM-PHYS-2008-008, CERN, Geneva, Jan, 2008. https://cds.cern.ch/record/1082880.

[181] D. W. Miller, A. Schwartzman, and D. Su, *Pile-up jet energy scale corrections using the jet-vertex fraction method*, Tech. Rep. ATL-COM-PHYS-2009-180, CERN, Geneva, Apr, 2009. https://cds.cern.ch/record/1172636.

[182] *Pile-up subtraction and suppression for jets in ATLAS*, Tech. Rep. ATLAS-CONF-2013-083, CERN, Geneva, Aug, 2013. https://cds.cern.ch/record/1570994.

[183] *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at* $\sqrt{s} = 8\ TeV$, Tech. Rep. ATLAS-CONF-2015-017, CERN, Geneva, Apr, 2015. https://cds.cern.ch/record/2008678.

[184] A. Dattagupta, J. Horton, L. Kogan, S. Lammers, G. Lefebvre, B. Malaescu, A. Olariu, P. Starovoitov, and M. Vetterli, *Determination of the jet energy scale and resolution at ATLAS using* Z/γ*-jet events in data at* $\sqrt{s} = 8\ TeV$, Tech. Rep. ATL-COM-PHYS-2014-791, CERN, Geneva, Jul, 2014. https://cds.cern.ch/record/1741697.

[185] M. Aharrouche and et al., *Response uniformity of the {ATLAS} liquid argon electromagnetic calorimeter*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **582** (2007) no. 2, 429 – 455. http://www.sciencedirect.com/science/article/pii/S0168900207018591.

[186] ATLAS Collaboration, *Measurements of the properties of the Higgs-like boson in the two photon decay channel with the ATLAS detector using 25* fb^{-1} *of proton-proton collision data*, Tech. Rep. ATLAS-CONF-2013-012, CERN, Geneva, Mar, 2013. https://cds.cern.ch/record/1523698.

[187] 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. https://cds.cern.ch/record/1711874. Supporting note for the 2012 inclusive cross section measurement.

[188] ATLAS Collaboration, G. Aad et al., *Measurement of the inclusive isolated prompt photons cross section in pp collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector using 4.6* fb^{-1}, Phys. Rev. **D89** (2014) no. 5, 052004, arXiv:1311.1440 [hep-ex].

[189] M. Boonekamp, B. Lenzi, N. Lorenzo Martinez, P. Schwemling, F. Teischinger, G. Unal, and L. Yao, *Electromagnetic calorimeter layers energy scales determination*, Tech. Rep. ATL-COM-PHYS-2013-1423, CERN, Geneva, Oct, 2013. https://cds.cern.ch/record/1609068.

[190] M. Boonekamp, O. Fedin, D. Froidevaux, and G. Unal, *Passive material determination using electrons and photons*, Tech. Rep. ATL-COM-PHYS-2013-1644, CERN, Geneva, Dec, 2013. https://cds.cern.ch/record/1636887.

[191] J. Colas, C. de La Taille, R. Lafaye, N. Massol, and P. Pralavorio, *Crosstalk in the ATLAS electromagnetic calorimeter*, tech. rep., 1999.

[192] J. Labbe and R. Ishmukhametov, *Crosstalk Measurements in the Electromagnetic Calorimeter during ATLAS Final Installation*, Tech. Rep. ATL-LARG-INT-2009-004. ATL-COM-LARG-2008-012, CERN, Geneva, Dec, 2008. https://cds.cern.ch/record/1143373.

[193] F. Hubaut, P. Pralavorio, B. Dekhissi, J. E. Derkaoui, A. El-Kharrim, and F. Maaroufi, *Crosstalk in production modules of the Electromagnetic Endcap Calorimeter*, Tech. Rep. ATL-LARG-2003-012, CERN, Geneva, Oct, 2003. https://cds.cern.ch/record/685566.

[194] 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 [hep-ex].

[195] *Study of the spin and parity of the Higgs boson in HVV decays with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2015-008, CERN, Geneva, Mar, 2015. https://cds.cern.ch/record/2002414.

[196] CMS Collaboration, S. Chatrchyan et al., *Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs*, Phys.Rev.Lett. **110** (2013) no. 8, 081803, arXiv:1212.6639 [hep-ex].

[197] CMS Collaboration, V. Khachatryan et al., *Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs*, Phys.Lett. **B736** (2014) 64, arXiv:1405.3455 [hep-ex].

[198] A. Hill and J. van der Bij, *Strongly Interacting Singlet-Doublet Higgs Model*, Phys. Rev. D **36** (1987) 3463.

[199] M. Veltman and F. Yndurain, *Radiative Corrections to WW Scattering*, Nucl. Phys. B **325** (1989) 1.

[200] T. Binoth and J. van der Bij, *Influence of strongly coupled, hidden scalars on Higgs signals*, Z. Phys. C **75** (1997) 17, arXiv:hep-ph/9608245 [hep-ph].

[201] R. Schabinger and J. D. Wells, *A Minimal spontaneously broken hidden sector and its impact on Higgs boson physics at the large hadron collider*, Phys. Rev. D **72** (2005) 093007, arXiv:hep-ph/0509209 [hep-ph].

[202] B. Patt and F. Wilczek, *Higgs-field portal into hidden sectors*, arXiv:hep-ph/0605188 [hep-ph].

[203] G. M. Pruna and T. Robens, *The Higgs Singlet extension parameter space in the light of the LHC* , Phys. Rev. D **88** (2013) 115012, arXiv:1303.1150 [hep-ph].

[204] T. D. Lee, *A Theory of Spontaneous T Violation*, Phys. Rev. D **8** (1973) 1226.

[205] *Search for scalar diphoton resonances in the mass range 65-600 GeV with the ATLAS detector in pp collision data at $\sqrt{s}$ = 8 TeV*, Tech. Rep. ATLAS-CONF-2014-031, CERN, Geneva, Jun, 2014. https://cds.cern.ch/record/1706244.

[206] (ATLAS) Collaboration, *Search for Scalar Diphoton Resonances in the Mass Range 65 - 600 GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8$ TeV*, Phys. Rev. Lett. **113** (Oct, 2014) 171801. http://link.aps.org/doi/10.1103/PhysRevLett.113.171801.

[207] HSG1, *Supporting note: Limit on fiducial production cross-section of a narrow resonance decaying into two photons with the ATLAS detector using 20 fb^{-1} of proton-proton collision data*, Tech. Rep. ATL-COM-PHYS-2013-1324, CERN, Geneva, Sep, 2013. https://cds.cern.ch/record/1599990.

[208] TEVNPH (Tevatron New Phenomina and Higgs Working Group), CDF, D0 Collaboration, *Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data*, arXiv:1203.3774 [hep-ex].

[209] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Physics Letters B **659** (2008) 119.

[210] H. HSG1 Working Group, *Signal studies in H to gamma gamma search with 8TeV data*, Tech. Rep. ATL-COM-PHYS-2012-755, CERN, Geneva, Jun, 2012. https://cds.cern.ch/record/1453770.

[211] ATLAS Collaboration, *Measurement of the Higgs boson mass from the Hgg and HZZ4l channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. D **90** (Sep, 2014) 052004. http://link.aps.org/doi/10.1103/PhysRevD.90.052004.

[212] *Differential cross sections of the Higgs boson measured in the diphoton decay channel using 8 TeV pp collisions*, Tech. Rep. ATLAS-CONF-2013-072, CERN, Geneva, Jul, 2013. https://cds.cern.ch/record/1562925.

[213] 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. https://cds.cern.ch/record/1609629.

[214] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **0904** (2009) 002, arXiv:0812.0578 [hep-ph].

[215] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **1002** (2010) 037, arXiv:0911.5299 [hep-ph].

[216] A. Djouadi, M. Spira, and P. Zerwas, *Production of Higgs bosons in proton colliders: QCD corrections*, Phys. Lett. **B264** (1991) 440.

[217] S. Dawson, *Radiative corrections to Higgs boson production*, Nucl. Phys. **B359** (1991) 283.

[218] M. Spira, A. Djouadi, D. Graudenz, and P. Zerwas, *Higgs boson production at the LHC*, Nucl. Phys. **B453** (1995) 17, arXiv:hep-ph/9504378 [hep-ph].

[219] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, Phys. Rev. Lett. **88** (2002) 201801, arXiv:hep-ph/0201206 [hep-ph].

[220] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, Nucl. Phys. **B646** (2002) 220, arXiv:hep-ph/0207004 [hep-ph].

[221] V. Ravindran, J. Smith, and W. L. van Neerven, *NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions*, Nucl. Phys. **B665** (2003) 325, arXiv:hep-ph/0302135 [hep-ph].

[222] S. Catani, D. de Florian, M. Grazzini, and P. Nason, *Soft gluon resummation for Higgs boson production at hadron colliders*, JHEP **0307** (2003) 028, arXiv:hep-ph/0306211 [hep-ph].

[223] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*, JHEP **1202** (2012) 088, arXiv:1111.2854 [hep-ph].

[224] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini, *Two loop light fermion contribution to Higgs production and decays*, Phys. Lett. **B595** (2004) 432, arXiv:hep-ph/0404071 [hep-ph].

[225] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NLO electroweak corrections to Higgs boson production at hadron colliders*, Phys. Lett. **B670** (2008) 12, arXiv:0809.1301 [hep-ph].

[226] D. de Florian and M. Grazzini, *Higgs production at the LHC: updated cross sections at* $\sqrt{s} = 8$TeV, arXiv:1206.4133 [hep-ph].

[227] C. Anastasiou, S. Buehler, F. Herzog, and A. Lazopoulos, *Inclusive Higgs boson cross-section for the LHC at 8 TeV*, JHEP **1204** (2012) 004, arXiv:1202.3638 [hep-ph].

[228] J. Baglio and A. Djouadi, *Higgs production at the LHC*, JHEP **1103** (2011) 055, arXiv:1012.0530 [hep-ph].

[229] I. Asimov, *Franchise*. August, 1955.

[230] HTTPS://TWIKI.CERN.CH/TWIKI/BIN/VIEW/ATLASPROTECTED/PHOTONID2012.

[231] 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 [hep-ex].

[232] 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. https://cds.cern.ch/record/1447803.

[233] H. HSG1 Working Group, *Statistics studies for H to gamma gamma search in the full 2011 dataset*, Tech. Rep. ATL-COM-PHYS-2012-732, CERN, Geneva, Jun, 2012. https://cds.cern.ch/record/1453249.

[234] H. HSG1 Working Group, *Background Studies for the Search of Higgs Boson Decaying to Two Photons with 8TeV data*, Tech. Rep. ATL-COM-PHYS-2012-754, CERN, Geneva, Jun, 2012. https://cds.cern.ch/record/1453768.

[235] H. HSG1 Working Group, *Search for the Higgs boson in the diphoton decay channel with data collected at 7 TeV and 8 TeV*, Tech. Rep. ATL-COM-PHYS-2012-503, CERN, Geneva, May, 2012. https://cds.cern.ch/record/1447438.

[236] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, European Physical Journal C **70** (nov, 2010) 525–530, arXiv:1005.1891 [physics.data-an].

[237] Cowan, G. and Cranmer, K. and Gross, E. and Vitells, O., *Asymptotic formulae for likelihood-based tests of new physics*, Eur.Phys.J.C **71** (2011) 1554. http://arxiv.org/abs/1007.1727.

[238] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693–2704. [,11(2002)].

[239] ATLAS Collaboration, *Observation and study of the Higgs boson candidate in the two photon decay channel with the ATLAS detector at the LHC*, Tech. Rep. ATLAS-CONF-2012-168, CERN, Geneva, Dec, 2012. https://cds.cern.ch/record/1499625.

[240] CMS Collaboration, *Properties of the observed Higgs-like resonance using the diphoton channel*, Tech. Rep. CMS-PAS-HIG-13-016, CERN, Geneva, 2013. https://cds.cern.ch/record/1558930.

[241] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[242] H. e. a. Abreu, *Measurement of isolated di-photon cross section in pp collision at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Tech. Rep. ATL-PHYS-INT-2011-071, CERN, Geneva, Sep, 2011. https://cds.cern.ch/record/1381549.

[243] T. Binoth, J. 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, arXiv:hep-ph/9911340 [hep-ph].

[244] L. J. Dixon and Y. Li, *Bounding the Higgs Boson Width Through Interferometry*, Phys.Rev.Lett. **111** (2013) 111802, arXiv:1305.3854 [hep-ph].

[245] G. D'Agostini, *Improved iterative Bayesian unfolding*, ArXiv e-prints (Oct., 2010) , arXiv:1010.0632 [physics.data-an].

[246] A. Hocker and V. Kartvelishvili, *SVD approach to data unfolding*, Nucl. Instrum. Meth. **A372** (1996) 469–481, arXiv:hep-ph/9509307 [hep-ph].

[247] HTTP://WWW.DESY.DE/~SSCHMITT/TALKS/UNFOLDSTATSCHOOL2014.PDF.

[248] T. HSG1 Group, *Measurement of fiducial and differential cross sections of the Higgs boson in the diphoton decay channel using centre-of-mass energy of 8 TeV*, Tech. Rep. ATL-COM-PHYS-2013-1664, CERN, Geneva, Dec, 2013. https://cds.cern.ch/record/1639554.

[249] HTTPS://TWIKI.CERN.CH/TWIKI/BIN/VIEW/ATLASPROTECTED/EGAMMACALIBRATIONGEO21.

[250] N. Andari, M. Boonekamp, L. Carminati, N. Lorenzo Martinez, M. Rimoldi, R. Turra, and G. Unal, *Calibration systematic uncertainties : overview and correlations*, Tech. Rep. ATL-COM-PHYS-2013-1654, CERN, Geneva, Dec, 2013. https://cds.cern.ch/record/1637535.

[251] HTTPS://TWIKI.CERN.CH/TWIKI/BIN/VIEWAUTH/ATLASPROTECTED/JETUNCERTAINTIES2012.

[252] HTTPS://TWIKI.CERN.CH/TWIKI/BIN/VIEWAUTH/ATLASPROTECTED/JETETMISSRECOMMENDATIONS2012.

[253] ATLAS Collaboration, G. Aad et al., *Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at* $\sqrt{s} = 8$ *TeV with ATLAS*, JHEP **09** (2014) 112, arXiv:1407.4222 [hep-ex].

[254] The ATLAS collaboration, *Liquid Argon Calorimeter Technical Design Report*, Tech. Rep. CERN/LHCC/96-041, CERN, 1996.

[255] T. Robens and T. Stefaniak, *Status of the Higgs Singlet Extension of the Standard Model after LHC Run 1*, Eur. Phys. J. **C75** (2015) 104, arXiv:1501.02234 [hep-ph].

[256] J. Song and Y. W. Yoon, *Gigantic diphoton rate of heavy Higgs bosons in the aligned two Higgs doublet models with small* $\tan\beta$, Phys. Rev. **D91** (2015) no. 11, 113012, arXiv:1412.5610 [hep-ph].

[257] G. Cacciapaglia, A. Deandrea, G. D. La Rochelle, and J.-B. Flament, *Searching for a lighter Higgs boson: Parametrization and sample tests*, Phys. Rev. **D91** (2015) no. 1, 015012, arXiv:1311.5132 [hep-ph].

Résumé

Au lendemain de la découverte du boson de Higgs au LHC, l'état final diphoton offre encore beaucoup d'opportunités en termes de mesure de la section efficace de production diphoton du Modèle Standard et de recherche de résonances supplémentaires. Les deux approches sont présentées dans cette thèse. Comme la reconstruction de photons dans ATLAS repose sur un étalonnage précis de leur réponse dans le calorimètre EM, cette thèse adresse aussi l'étalonnage de la réponse des photons et ses possibles améliorations. Pour préparer le Run 2 du LHC démarrant en 2015, il est également intéressant de regarder la reconstruction du signal LAr avec pile up plus élevé et une plus grande fréquence de déclenchement obligeant à réduire le nombre d'échantillons de cinq à quatre.

Cette thèse est organisée comme suit. Le chapitre 1 présente la compréhension actuelle du Modèle Standard, et introduit certaines de ses extensions possibles. Les motivations pour la recherche d'événements diphoton et de faire des mesures avec les états finaux diphotons plus jets sont discutées. Le chapitre 2 présente le dispositif expérimental, le Grand collisionneur de hadrons et l'expérience ATLAS avec ses sous-systèmes. Dans le chapitre 3, la reconstruction et l'étalonnage des photons est décrite. Le chapitre 4 donne une brève introduction sur la reconstruction des jets et leur étalonnage dans ATLAS. Dans le chapitre 5, l'étalonnage plus précis de l'énergie des photons est présenté et une étude est effectuée sur les photons pour améliorer la mesure de l'énergie dans le détecteur LAr. Le chapitre 6 décrit la recherche de résonances scalaires étroites se désintégrant en deux photons, où ma contribution personnelle était la modélisation du signal et le calcul des incertitudes systématiques associées. Une mesure de sections efficaces différentielles de diphoton en association avec des jets est présentée dans le chapitre 7, avec comparaison des valeurs mesurées aux prédictions théoriques. Dans le chapitre 8, une étude visant à l'optimisation de la reconstruction de l'énergie dans le calorimètre à argon liquide pour le Run2 du LHC est présentée. Le chapitre 9 liste les perspectives des études et les mesures présentées dans cette thèse, et résume l'ensemble de la thèse.

1. Le Modèle Standard

Le Modèle Standard des particules élémentaires est une théorie de jauge sur la base de trois groupes de symétrie: SU (3) c x SU (2) L x U (1) H. Chaque groupe représente une force

portée par une partie des particules élémentaires. La force forte, portée par les quarks et les gluons est basée sur le groupe de couleur SU(3), qui a été bien décrit par la théorie de la chromodynamique quantique. Les groupes SU (2) L et U (1) combinés sont la base de la théorie électrofaible, unifiant les interactions électromagnétiques et faibles, portés par le photon et les bosons Z et W, qui interviennent dans les interactions entre fermions - trois générations de leptons et de neutrinos.

Un champ scalaire avec une valeur moyenne dans le vide non nulle, présenté par Peter Higgs, François Englert et Robert Brout en 1964, permet la brisure spontanée de la symétrie électrofaible et de prédire l'existence d'un boson scalaire supplémentaire. Les expériences ATLAS et CMS ont découvert une résonance d'une masse d'environ 126 GeV se désintégrant en deux photons, deux Z allant en quatre leptons et d'autres états finaux compatibles avec le boson de Higgs et ayant des propriétés compatibles avec celles du boson de Higgs du Modèle Standard. Le mécanisme de brisure de symétrie EW et l'obtention de masse par les bosons et les fermions peuvent être expliqués par le mécanisme de Higgs terminant le Modèle Standard (MS).

Il existe, cependant, plusieurs phénomènes observés dans la nature que le MS est incapable d'expliquer et nécessitant une extension au-delà du MS. Les principaux problèmes du MS comprennent l'asymétrie matière-antimatière, la masse des neutrinos, la gravité, la matière noire et l'énergie sombre et la naturalité. De nombreuses extensions du MS ont été proposées, la plus recherchée étant la supersymétrie, prédisant, entre autres choses, l'existence d'un secteur de Higgs étendu, avec deux Higgs neutres supplémentaires H et A, les deux pouvant se désintégrer en deux photons.

Le canal de diphoton a joué un rôle majeur dans la découverte du boson de Higgs (126). Plusieurs modèles prédisent l'existence d'un autre Higgs à découvrir dans le spectre de masse invariante diphoton aussi bien en dessus qu'en dessous de la masse du Higgs à 126 GeV. La recherche de résonances supplémentaires se désintégrant en deux photons est donc d'une importance capitale.

Il est également intéressant de mesurer la section efficace de production de diphoton au LHC, en raison de grandes différences entre les calculs théoriques de la section transversale à l'ordre dominant et à l'ordre suivant, en particulier quand ils sont associés à un ou plusieurs jets. Cet état final n'a pas encore été mesuré dans ATLAS.

2. LHC et le détecteur ATLAS

Le Grand collisionneur d'hadrons (LHC) est le plus grand accélérateur de particules du monde, capable d'accélérer des protons et de faire des collisions proton-proton jusqu'à 13 TeV. Pendant la Run1 du LHC, le détecteur ATLAS a recueilli 20.3 fb^{-1} de données sur les collisions à sqrt s = 8 TeV. Le détecteur ATLAS est une des deux expériences multi-usages du LHC, conçues pour la découverte du boson de Higgs et de la supersymétrie. Il se compose de plusieurs sous-systèmes. En partant du tube de faisceau arrive en premier le détecteur interne - trois systèmes distincts assurant la détection des particules chargées en provenance du point d'interaction. Un aimant supraconducteur à solénoïde facilite la reconstruction des particules chargées en déviant leurs trajectoires. Le calorimètre à argon liquide, composé de quatre couches dans la partie tonneau présente une excellente résolution spatiale et permet la détection des photons et des électrons et la mesure de leur énergie. Le calorimètre à tuiles mesure l'énergie des hadrons et des jets. La trajectoire des muons est déviée par les aimants toroïdaux et les muons sont détectés par le système du spectromètre à muons. En raison de la conservation de l'impulsion transverse dans une collision proton-proton, l'énergie transverse manquante peut être calculée, qui peut être crée

par les neutrinos ou d'autres particules encore inconnues.

Le détecteur ATLAS utilise un système de déclenchement pour collecter et un système d'acquisition de données pour traiter les événements qui peuvent avoir des états finaux intéressants. La qualité des données recueillies est évaluée ligne et hors ligne. Une liste de bon runs est produite qui est utilisée par toutes les analyses afin de supprimer les événements où des problèmes de détection sont apparus.

3. et 5. La détection et l'étalonnage des photons dans ATLAS

La détection de photons commence par la détection des dépôts d'énergie dans les cellules du calorimètre LAr. Les photons sont reconstruits à partir des tours du calorimètre - additionnant l'énergie de toutes les couches du calorimètre dans le tour de taille 0.025x0.025 en pseudorapidité et azimut. Pour reconstruire les amas électromagnétiques (EM), ils sont initiés par des tours d'énergie transversale dépassant 2.5 GeV et recherchés par un algorithme à fenêtre glissante avec une taille de fenêtre de 3x5. Si une trace est associée à l'amas, il est classé comme un électron. Si un vertex de conversion est reconstruit, le cluster est classé comme photon converti et dans le cas d'aucune trace n'est observée par le détecteur interne, le cluster est classé comme photon non converti. Les amas sont ensuite reconstruits en utilisant une zone de cellules du calorimètre de taille variable selon leur classification et la région du détecteur.

L'énergie des photons est calibrée en fonction de leur état de conversion et d'autres informations de détection tels que les dépôts d'énergie dans chaque couche du calorimètre LAr, en utilisant un étalonnage multivariée. La réponse relative des couches du calorimètre est généralement étalonnée avec les données et la simulation des muons et des électrons. Dans cette thèse, la réponse relative est également étudiée en utilisant des photons pour étendre la zone de validité de l'intercalibration à plus haute énergie, et d'évaluer une éventuelle dépendance de l'énergie résiduelle. Ensuite, le matériau passif devant le calorimètre utilisé par la simulation est optimisé à partir du profil longitudinal des gerbes dans les calorimètres. Les bosons Z se désintégrant en paires électron-positron sont utilisés pour aligner les échelles d'énergie absolue des données et de la simulation. Une étude sur la nécessité d'un étalonnage des photons en fonction de l'angle azimutal est aussi effectuée.

4. Jets dans ATLAS

En raison des propriétés de la QCD, les quarks et les gluons ne peuvent pas être observées directement, se fragmentant immédiatement en hadrons. A haute énergie, la production de quarks et de gluons donne lieu à un flux de hadrons colinéaires, appelé jet. Les jets d'ATLAS sont détectés par les calorimètres LAr et tuiles, tandis que les particules chargées peuvent être détectées par le système de détecteur interne. Les jets sont reconstruits en utilisant des clusters topologiques et des algorithmes de reconstruction de jet, à savoir l'algorithme anti-kt de rayon 0.4 dans le plan R. Les jets sont reconstruits en utilisant tous les dépôts d'énergie dans le calorimètre. Par conséquent, une grande partie des jets, surtout à bas p_{T}, viennent de l'empilement et non de partons produits lors de processus durs. L'empilement est rejeté en partie par l'application de l'algorithme 'Jet Vertex Fraction' (JVF) qui s'applique dans le cas où la fraction d'impulsion transverse du jet venant d'un vertex est inférieure à 0.5.

Les jets sont calibrés à partir de l'énergie du calorimètre à l'échelle de l'énergie EM, en mesurant les dépôts d'énergie des gerbes EM et en la corrigeant par rapport à la réponse du calorimètre

à la vraie énergie correspondant à un jet de particules stables dans la simulation.

6. Recherche d'une résonance étroite dans le canal diphoton

La recherche d'une résonance étroite présentée dans le chapitre 6 utilise toutes les données recueillies par ATLAS lors des collisions proton-proton à 8 TeV en 2012. Les événements sont tenus de passer le seuil de déclenchement diphoton avec un critère sur la forme de germe de gerbe et une énergie transversale d'au moins 20 GeV. Les événements doivent aussi satisfaire les critères de bonne qualité des données, tandis que les candidats de photons doivent passer une liste de critères incluant une énergie transversale supérieure à 22 GeV et une pseudorapidité dans l'acceptance des calorimètres LAr soit en dessous de 2.37 et hors de la région de transition tonneau-bouchons. Les photons sont également tenus de passer des critères d'identification strictes et avoir une isolation calorimétrique de 6 GeVet une isolation par rapport aux traces de 2.6 GeV. Un critère supplémentaire d'isolation calorimétrique variant avec l'énergie transversale du photon est ajouté pour les photons avec une énergie transverse supérieure à 80 GeV. L'analyse est effectuée dans deux régions de masse invariante des diphotons en raison des différentes caractéristiques du bruit de fond.

La partie de faible masse de l'analyse, considère les événements diphoton avec une masse invariante entre 55-120 GeV et est sensible aux résonances de masse entre 60-110 GeV. Le principal défi de cette partie de l'analyse est le bruit de fond des électrons isolés, qui forme un pic autour de la masse du boson Z à 91 GeV. Le taux d'électrons identifiés comme photon a été calculé en utilisant les événements de simulation et un modèle décrivant la forme du pic en est dérivé. Comme les électrons et les photons sont reconstruits et calibrés différemment dans ATLAS, la masse invariante diélectron utilisée pour calculer le modèle ne correspond pas à la position et la forme du pic de masse des diphotons et est décalée également en fonction des différentes catégories d'état de conversion de photons. Un décalage dans l'énergie transversale et un étalement en azimut est calculé à partir de la simulation et sont appliqués sur les données lors de l'extraction du modèle. Le bruit de fond continuum diphoton est modélisé par une convolution d'une Landau et d'une fonction exponentielle, décrivant le démarrage au seuil de déclenchement dans la zone de basse masse et la forme exponentielle décroissante du continuum diphoton loin du seuil de déclenchement. La partie haute de masse invariante de l'analyse couvre la région entre 100 et 700 GeV et est sensible aux nouvelles résonances pour des masses entre 110-600 GeV. Le pic dans le spectre de masse invariante diphoton, correspondant à la désintégration de Higgs en deux photons, est reproduit en utilisant une fonction 'double-sided Crystal Ball' avec une intensité de signal unitaire et une masse de 125.9 GeV, qui était la valeur la plus précise en vigueur au moment de l'analyse. Le fond de continuum diphoton est décrit par une exponentielle d'un polynôme du second degré dans une fenêtre glissante autour du point de masse testé.

La modélisation du signal d'une nouvelle résonance est faite pour plusieurs catégories de conversion utilisant des échantillons simulés du boson de Higgs entre 60-150 GeVet un échantillon avec une largeur étroite pour les points de masse au-dessus de 150 GeV, basée sur une largeur naturelle de 4 MeV. L'évolution de la forme du signal est modélisée par une fonction de la masse invariante diphoton afin d'obtenir une prédiction continue. La fonction 'double-sided Crytal Ball' est utilisée pour décrire la forme. Un ajustement simultané de tous les points de masse disponibles est effectué pour produire des résultats en accord avec le résultat d'ajustements individuels. La modélisation du signal dans la région de faible masse est effectuée séparément pour les différentes catégories de conversion.

Pour corriger le nombre d'événements de signaux calculés dans les données des effets de détection, un facteur de correction Cx est calculé comme rapport du nombre d'événements reconstruits dans la simulation sur le nombre d'événements de départ. Une coupure sur l'isolation des particules est appliquée à 12 GeV, ainsi que les sélections de photons mentionnés ci-dessus. Le facteur Cx est calculé, par défaut, pour la fusion gluon-gluon de production du Higgs et pour les autres modes de production du Higgs, la différence entre eux est considérée comme l'incertitude systématique.

D'autres incertitudes systématiques affectant la mesure sont le signal parasite - nombre d'événements identifiés comme signal lors de l'ajustement de la somme signal + fond provenant seulement de bruit de fond simulé; les incertitudes provenant de la modélisation du bruit de fond électron - principalement la normalisation et de la forme de la fonction de référence; l'incertitude sur la résolution du signal, allant de 10 à 40% pour les masses élevées et encore beaucoup d'autres.

Un ajustement du maximum de vraisemblance est effectué séparément dans les zones de faible et haute masse du spectre de de diphoton incluant les bruits de fond déjà cités et les incertitudes systématiques associées. Aucun excès significatif n'est trouvé sur toute la gamme. Une limite sur la section efficace fiducielle multipliée par le rapport de branchement pour une nouvelle résonance se désintégrant en deux photons est calculée, allant de 90 fb à 60 GeV à 1 fb à 600 GeV. En utilisant les données 2012 recueillies par l'expérience ATLAS, aucune résonance supplémentaire autre que le boson de Higgs n'est observée dans l'état final de diphoton.

7. Mesure de la section d'efficace diphotons en association avec des jets

Les sections efficaces différentielles de production de diphoton en association avec des jets sont mesurées, en fonction des propriétés cinématiques et angulaires des photons et des jets, à partir des données 2012 (20.3 fb^{-1}). Une sélection des événements similaire à celle de la recherche de résonances diphoton et présentée dans le chapitre précédent est appliquée. Les événements doivent passer le seuil de déclenchement des diphoton avec une énergie transverse supérieure à 20 GeV et des critères de forme de gerbes moyennes et passer les critères de qualité des données. Les photons sont choisis si leur ET est supérieure à 22 GeV et si ils sont dans l'acceptance en pseudorapidité des calorimètres LAr en excluant la zone de transition entre le tonneau et bouchons. Les paires de photons passant des critères d'identification sevères et ayant une énergie d'isolation calorimétrique inférieure à 4 GeV sont conservés. Les photons doivent également être séparés dans le plan R d'au moins 0.4. La masse invariante diphoton doit être supérieure à 60 GeV à cause d'une coupure à 55 GeV lors de la production des événements simulés utilisés pour le calcul de section efficace finale et l'extraction de l'efficacité de sélection sévere.

Les jets sont reconstruits en utilisant l'algorithme anti-kt avec R = 0.4 et calibrés selon l'échelle de l'énergie EM. Les jets de pseudorapidité inférieure à 2.4 et une impulsion transverse inférieure à 50 GeV doivent passer la coupure JVF (> 0.5). L'impulsion transverse pour les jets sélectionnés avec une pseudorapidité inférieure à 2.4 doit dépasser 25 GeV. Pour les jets de pseudorapidité entre 2.4 et 4.4, l'impulsion transverse doit dépasser 50 GeV car les traces chargées ne sont plus reconstruites et il est impossible de rejeter les jets de pile-up en appliquant la sélection JVF. Les jets sont choisis de telle sorte qu'ils soient distants de plus de 0.6 dans le plan R. Les catégories dans la multiplicité de jet sont définies après le passage de la sélection des jets. Les événements où un jet sélectionné est à une distance du photon de 0.4 à 0.6 dans le plan R sont éliminés en raison de la mauvaise description de l'efficacité de réjection de ces photons dans les données et dans la simulation.

Les événements diphoton sélectionnés contiennent encore une certaine fraction d'événements où un ou deux photons sont en fait des jets (principalement pi0 se désintégrant en deux photons à l'intérieur d'un jet) ou des électrons identifiés à tort comme des photons. Le Higgs, se désintégrant en paires de diphoton contribue également au taux de production de diphoton.

Le taux de production de diphoton, incluant le bruit de fond électron, est extrait en utilisant le procédé à bande latérale 2x2D, qui est une méthode de comptage basé sur la séparation des événements diphoton en seize catégories selon l'identification sévère (ID) et l'isolation des deux candidats de photons et en effectuant un ajustement pour extraire les contributions des diphoton, photon-jet, jet-photons et le jet-jet. Le procédé démarre avec un échantillon sans avoir appliqué le critère d'isolation et quatre des critères d'identifications sont relâchés. La méthode prend en entrée les efficacités d'identifications des photons à partir de la simulation et celui d'isolement à partir des données. La principale incertitude systématique vient de la définition de la région non-serrée et est évaluée en libérant deux à cinq des critères de sélection des ID serrés. Cette systématique varie pour chaque distribution différentielle d'une observable mais représente trois à cinq pour cent.

Le bruit de fond électron est soustrait en construisant les distributions contenant des diélectrons et de événements électron-photon venant des données et en appliquant un taux de confusion d'électron en photon ou de photon en électron. Le taux d'électrons confondus avec des photons est mesuré avec des diélectrons et des paires électron-photon autour du pic du Z and une incertitude est assignée à cause d'une possible dépendance en fonction de l'énergie transverse.

La contribution de Higgs est estimée en utilisant la simulation et en supposant que le taux de production est celle du Higgs.

Les distributions différentielles sont extraites à l'aide de la décomposition en valeurs singulières en deux dimensions, prenant en compte les migrations entre catégories de jets ou entre variable différentielles. Les sections efficaces sont comparées aux simulations des modèles Pythia et Sherpa.

8. OFC pour Run 2

Les impulsions du calorimètre LAr d'ATLAS ont été détectées et échantillonnées toutes les 25 ns pendant le Run1 en conservant cinq échantillons pour la reconstruction des amplitudes d'impulsion. Cette quantité de données a été transmise à l'électronique de back-end au taux de déclenchement de niveau 1 de 60 kHz sans problème. Le lien de transmission entre les électronique frontale et l'électronique de back-end du détecteur est cependant limité par le lien à 1.6 Gbit/s, et comme il est prévu que le taux de déclenchement augmente à 100 kHz pour la prise de données Run2, afin d'éviter les temps morts du détecteur, quand il serait incapable de traiter et stocker des données, la possibilité de garder seulement quatre échantillons est étudiée. L'impact de la réduction du nombre d'échantillons de cinq à quatre ou trois sur les performances de bruit du calorimètre dans des conditions de haute pileup est évalué. Une recommandation sur le choix des quatre échantillons parmi les cinq est également construite sur la base de la comparaison des niveaux de bruit attendus pour chacune des combinaisons d'échantillons et séparément pour les trois parties du calorimètre LAR - le tonneau, les bouchons et la partie avant du calorimètre, et a été adopté par ATLAS pour la prise de données à 25 ns.

Conclusions

Après une introduction théorique et une description du LHC et du détecteur ATLAS, dans la première partie de la thèse, plusieurs études consacrées à l'étalonnage de photons sont présentées.

La modélisation du signal d'une résonance étroite se désintégrant en diphoton est présentée dans le cadre de la recherche d'une nouvelle résonance. Après avoir effectué un ajustement du maximum de vraisemblance, aucune nouvelle résonance n'a été trouvée, et une limite sur la section efficace fiducielle multipliée par le rapport de branchement d'une nouvelle résonance se désintégrant en deux photons a été produite. Ces résultats ont été publiés comme publication ATLAS dans Physical Review Letters en 2014.

La mesure de la section efficace de production de diphoton isolés en association avec un, deux ou trois et plus de jets a été présentée en fonction de variables cinématiques et angulaires des diphotons et des jets. Cette mesure fera partie d'une publication ATLAS, qui est actuellement en préparation.

Enfin, les propriétés de bruit du calorimètre ATLAS LAr ont été étudiées avec un important pile-up et un nombre réduit d'échantillons LAr. Une recommandation a été donnée pour le choix des quatre échantillons pour minimiser le bruit.

Titre: Mesures dans l'état final diphoton au LHC avec le détecteur ATLAS: recherche des nouvelles résonances et étude de la production de paires de photons en association avec des jets.

Résumé: Cette thèse étudie les collisions pp à $\sqrt{s} = 8$ TeV avec des paires de photons dans l'état final, telles que recueillies par l'expérience ATLAS au LHC au cours de 2012. L'état final avec diphoton s'est avéré cruciale pour la découverte du boson de Higgs. La motivation pour la recherche de résonances supplémentaires se désintégrant en deux photons est très forte en raison l'existence de nombreux modèles prévoyant un secteur de Higgs étendu. Dans le même temps, des mesures détaillées des sections efficaces diphoton sont nécessaires pour établir la précision des prédictions théoriques actuellement disponibles pour ces processus.

La mesure des photons par ATLAS repose sur un excellent étalonnage de la réponse du calorimètre électromagnétique. Pour cette raison, les études liées à l'étalonnage de la réponse de photons dans le calorimètre Argon Liquide (LAr) sont initialement réalisées, y compris la mesure de l'étalonnage par rapport à l'échelle de l'énergie de la couche calorimètre utilisant des photons, et la nécessité éventuelle d'un inter-étalonnage de la réponse en énergie des photons en fonction de l'angle azimutal ϕ. Une étude de l'amplitude du bruit du calorimètre dans des conditions de pile-up élevés est également présentée, aboutissant à des recommandations pour les configurations d'échantillonnage d'impulsion LAr au Run2.

La recherche d'une résonance supplémentaire se désintégrant en deux photons est effectuée et aucun signal n'est observé: une limite est mise sur la section efficace fiducielle d'une résonance multipliée par son rapport de branchement de désintégration en diphoton. La mesure de la section efficace de production de diphoton isolés en association avec des jets est effectuée en fonction de la masse invariante de diphoton, l'impulsion transverse du système diphoton, la différence d'angles dans le plan azimutal de l'angle entre les deux photons et l'angle theta de coplanarité, en fonction de la multiplicité en jets et de différentes variables associés aux jets. Les résultats sont comparés aux dernières prévisions Sherpa et Pythia MC. Une telle mesure est la première du genre effectuée au LHC.

Mots-clés: diphoton, jets, section efficace, production, resonance, boson de Higgs, ATLAS

Title: Diphoton measurements with the ATLAS detector at LHC: search for new resonances and study of diphoton production in association with jets.

Abstract: This thesis studies pp collisions at $\sqrt{s} = 8$ TeV with pairs of photons in the final state, as collected by the ATLAS experiment at the LHC during 2012. The diphoton final state has proven to be crucial to the discovery of the Higgs boson. The motivation for searching for additional resonances decaying into two photons is very strong due to many models predicting an extended Higgs sector. At the same time, detailed measurements of diphoton cross sections are necessary to establish the quality of the theoretical predictions currently available for these processes.

The measurement of photons by ATLAS relies on an excellent calibration of the electromagnetic calorimeter response. For this reason, studies related to the calibration of the photon response in the ATLAS Liquid Argon (LAr) calorimeter are initially performed, including a measurement of the calorimeter layer energy scale relative calibration using photons, and the potential need of an inter-calibration of the photon energy response as a function of the azimuthal angle ϕ. A study of noise performance of the calorimeter in high pile-up conditions is also presented, resulting in recommendations for the LAr pulse sampling configurations for Run2.

A search for an additional resonance decaying into two photons is performed, leading to no significant excess being observed. A limit on the fiducial cross section times branching fraction of a new resonance decaying into two photons is therefore computed. The measurement of the production cross section of isolated prompt diphotons in association with jets is performed as a function of the diphoton invariant mass, the transverse momentum of the diphoton system, the azimuthal angle difference of the two photons and the coplanarity angle theta, for jet multiplicity categories and as a function of various jet variables. The results are compared to Sherpa and Pythia MC predictions. Such a measurement has not been previously published at the LHC.

Keywords: diphoton, jets, cross section, production, resonance, Higgs boson, ATLAS