Università degli Studi di Torino

Scuola di Dottorato in Scienze della Natura e Tecnologie Innovative

Indirizzo in Fisica ed Astrofisica

Investigation of the Higgs sector using $H \to ZZ^{(*)} \to 4\ell$ events with the CMS detector

Coordinatore
Prof. Mauro Gallio

Tutore
Prof. Ernesto Migliore

Candidato
Stefano Casasso

Ciclo XXVII

Date of dissertation
January 27^{th} 2015

Candidate
Stefano Casasso

Tutor
Prof. Ernesto Migliore
Università di Torino

Degree Board
Dr. Alessio Ghezzi
Università Milano-Bicocca
Prof. Lorenzo Magnea
Università di Torino
Dr. Bruno Mansoulié
CEA Saclay
Prof. Gabriella Sciolla
Brandeis University

Ph.D. Coordinator
Prof. Mauro Gallio
Università di Torino

Acknowledgements

Mi preme innanzitutto ringraziare Ernesto, in particolare per la pazienza che ha dimostrato nell' "insegnarmi il mestiere" e per l'attenzione che ha dedicato al mio lavoro in questi anni.

Grazie alla mia famiglia, che ha sempre sostenuto le mie scelte, che ha sempre creduto in me, che da sempre sopporta il mio egoismo e le mie lunghe assenze.

Grazie a tutto il gruppo di CMS Torino, che mi ha dato l'opportunita' di portare a termine questo lavoro. Ho passato con voi quattro anni magnifici.

Grazie a tutti gli scienziati con cui ho avuto l'onore di interagire in questi anni: sono sicuro che tutti, chi piu' chi meno, ognuno a modo suo, mi hanno insegnato qualcosa.

Grazie ai miei amici, per i bei momenti passati insieme, ma anche, soprattutto, per quelli brutti che insieme abbiamo esorcizzato.

Contents

Introduction **1**

I Background **3**

1 The Standard Model and beyond **4**

1.1 Introduction 5

1.2 The Standard Model 5

1.2.1 The particle content of the Standard Model 5

1.2.2 The electroweak theory 6

1.2.3 The Higgs mechanism 8

1.3 Beyond the Standard Model 10

1.3.1 Reasons to go beyond the SM 10

1.3.2 Electroweak singlet 13

2 The Compact Muon Solenoid experiment at the Large Hadron Collider **16**

2.1 The Large Hadron Collider 17

2.1.1 The LHC design 17

2.1.2 The LHC operations during Run 1 (2008-2013) 18

2.2 The physics of the Higgs boson at the LHC 19

2.2.1 Higgs production 21

2.2.2 Higgs decay channels 22

2.2.3 The discovery of a Higgs boson 24

2.3 The CMS detector 26

2.3.1 The coordinate system 27

2.3.2 The overall design 27

2.3.3 The magnet 28

2.3.4 The muon system 29

2.3.5 The electromagnetic calorimeter 30
2.3.6 The hadron calorimeter 32
2.3.7 The inner tracker system 34
2.3.8 The trigger and the data acquisition 37

II Research Work 40

3 Study of the muon momentum scale with the CMS detector 41
3.1 Introduction 42
3.2 The *MuScleFit* algorithm 43
3.3 Calibration of muon momentum scale and resolution 45
3.3.1 Signal and background model 45
3.3.2 Fit strategy 48
3.3.3 Datasets and selection 49
3.3.4 Extra-smearing 50
3.4 Results 53
3.5 Systematic uncertainty on the muon momentum scale and resolution 73
3.6 Impact of muon momentum corrections on physics analyses 89
3.6.1 Search for a SM Higgs boson in the $\mu^+\mu^-$ final state [33] . . 89
3.6.2 Measurement of the properties of a Higgs boson in the four-lepton final state [15] 89

4 Observation of a Higgs boson in the $H \to ZZ \to 4\ell$ channel and measurement of its properties 92
4.1 Introduction 93
4.2 Datasets 93
4.2.1 Experimental data 93
4.2.2 Simulated samples 93
4.3 Physics Objects 95
4.3.1 Leptons 95
4.3.2 Final state radiation 99
4.3.3 Jets 101
4.4 Event selection and categorization 102
4.4.1 Best candidate selection 102
4.4.2 Signal efficiency 102
4.4.3 Event categorization 103

4.5 Background estimation 107
4.5.1 Irreducible background 107
4.5.2 Reducible background 108
4.6 Kinematic discriminants 110
4.7 Yields and kinematic distributions 111
4.8 Combined fit model and inputs 114
4.9 Systematic Uncertainties 116
4.10 Results 117
4.10.1 Exclusion limits and local signal significance 117
4.10.2 Mass and width 118
4.10.3 Signal strength 119
4.10.4 Spin and Parity 120

5 Study of the interference between the Higgs signal and the ZZ continuum background with LO generators 123
5.1 Introduction 124
5.2 The interference in the ggH channel 124
5.3 The interference in the VBF channel 125
5.4 The extraction and parameterization of the interference 129

6 Search for a high mass Higgs boson 135
6.1 Introduction 135
6.2 BSM intepretation of the high mass data in the $H \to ZZ^{(*)} \to 4\ell$ channel 136
6.2.1 Signal lineshape 136
6.2.2 $qq' \to qq'ZZ$ background estimation 141
6.2.3 Matrix element discriminant in the 2-jet category 141
6.2.4 Effect of h(125) in the BSM high mass search 143
6.2.5 Results 145
6.3 The combination of the WW and ZZ channels 150

Summary 153

A Statistical procedure for the calculation of the upper limits 155

References 158

Introduction

The Large Hadron Collider (LHC) and its experiments were built to address some of the unsolved questions in particle physics, most of them not having a satisfactory answer in the framework of the Standard Model (SM) of fundamental interactions.
One of the main goals of the LHC, in particular, is to shade light on the mechanism of the electroweak symmetry breaking (EWSB), which is sensitive to new physics scenarios, often referred to as physics Beyond Standard Model (BSM), such as, for example, Supersymmetry (SUSY).
The EWSB is realized in the SM introducing a scalar field with non-null vacuum expectation value. The particle associated to such field, the *Higgs boson*, had not been observed in any experiment before the LHC.

On the 4th of July 2012 the CMS and ATLAS collaborations at the LHC announced the discovery of a new particle, with mass around 125 GeV, in the context of the searches for the SM Higgs boson. Among the channels which contributed to this exceptional discovery, the one in which the Higgs decays to two Z bosons, which subsequently decay to four leptons (electrons and muons), in short $H \to 4\ell$, has been one of the most important. Thanks to the low background rate and the good mass resolution, the four-lepton channel is infact the most sensitive for the observation of the SM Higgs boson in a large mass range. Moreover, the final state is fully reconstructed and the angular correlations of the decay products can be used to measure the spin-parity properties.
The properties of the newly discovered particle have been extensively investigated with the data collected by the experiments during the Run 1 of the LHC (2011-2012). So far, no significant deviation from the predictions of the SM has been observed. However, most of the measurements are still limited by the statistics: more data are needed to shade light on the nature of this new particle.

This thesis describes my work done within the Compact Muon Solenoid (CMS) collaboration during my three-year Ph.D. program at the Università di Torino (University of Turin). During this period, I have been working in the $H \to 4\ell$ analysis group, participating to the discovery of the new boson in July 2012, the early studies of its properties and the exclusion of a Higgs with non-SM couplings at high mass.
I studied the muon momentum scale and resolution and developed calibration constants which have been applied to the measurement of the Higgs properties in the four-lepton channel, as well as in several other analysis in CMS.
Throughout my Ph.D., I have also been involved in the group responsible for the validation of the constants of the alignment of the Tracker detector.
This thesis is organized as follows.

In chapter 1, the SM theory is briefly reviewed, with emphasis on the EWSB mechanism. A minimal extension of the SM scalar sector is also presented, which is referred to as the Electroweak Singlet (EWS) model.
In chapter 2, the LHC is briefly introduced, with some details on the operations during the Run 1. Some highlights on the Higgs physics at the LHC are also given and the observation of a new particle consistent with the SM Higgs boson is presented. Finally, the CMS detector is described in some detail.
In chapter 4, the so-called "legacy" $H \to 4\ell$ analysis, which describes the measurements of the properties of the new boson in the four lepton final state using the full dataset of the Run 1, is presented. This analysis has been published and represents the observation of the new boson in this channel alone, with a significance that exceeds the 6-sigma level.
One of the most important results in the Legacy analysis is the measurement of the mass of the new particle, with an accuracy of about 0.5%. In this measurement, the lepton momentum scale and resolution are crucial and respresent the largest systematic uncertainty. For this reason, I developed corrections to the muon momentum scale and resolution and I studied the systematic uncertainty associated to the muon scale. The method, the results and the validation of such results are discussed in chapter 3. These calibrations have been adopted in the Legacy $H \to 4\ell$ analysis, as well as in many other analyses in CMS.
In the context of the search for a high mass Higgs boson, one of the most important effects to be taken into account is the interference of the signal with the continuum, which affects the lineshape for masses above 400 GeV. I studied the interference using leading order (LO) generators and I developed a novel technique to implement the effect of the interference in the shape of the signal: these studies are detailed in chapter 5.
In chapter 6 the results of the search for a high mass Higgs state in the $H \to 4\ell$ channel in the context of the EWS model are presented and preliminary results are shown for the combination of several WW and ZZ channels.

Part I

Background

Chapter 1

The Standard Model and beyond

Abstract

The Standard Model, which describes the interactions between the fundamental particles, is one of the most succesful theory in the history of science. It has been tested extensively by the physicists all over the world and succeeded in describing the observations in many facilities and experiments during the last sixty years.

However, there are some theoretical prejudices in the community of the high energy physics about the fact that the Standard Model should break at a certain energy scale. Moreover, there are experimental evidences of new phenomena that do not find a satisfactory explanation in the Standard Model.

1.1 Introduction

To our knowledge, there exist four fundamental forces in nature: the gravitational, the electromagnetic, the weak and the strong force. Apart from the gravitation, the other three have been succesfully accomodated in the framework of a gauge invariant quantum field theory, known as Standard Model (SM) [1, 2, 3, 4, 5]. Since the unification of the electromegnetic and weak forces in the 1961 by Glashow, Weinberg and Salam, the SM has been verified experimentally by means of particle accelerators, nuclear reactors, cosmic rays and other apparatus. Even though no significant evidence of discrepancy with respect to the SM predictions have been reported in such experiments, still there are experimental and theoretical reasons to believe that there should be a more complete theory, able to describe phenomena at higher energy scales and being approximated by the SM at the electroweak scale ($v \simeq 246\,\mathrm{GeV}$).

In this chapter the theory of Standard Model is briefly reviewed in section 1.2, with some details given on the fundamental particles and their properties (1.2.1), the electroweak unification (1.2.2) and the Higgs mechanism (1.2.3).
In section 1.3.1 the main problems with the SM are outlined and a possible extension of the scalar sector, relevant for the analysis described in chapter 6, is described with some details in section 1.3.2.

1.2 The Standard Model

1.2.1 The particle content of the Standard Model

The matter is believed to be composed by twelve elementary particles with spin 1/2, the *fermions*. Fermions are classified according to their interaction properties: while *quarks* participate in the eletromagnetic, weak and color interactions, the *leptons* do not have color charge and interact only through electromagnetic and weak forces. Leptons are said to be "singlet" under the $SU(3)_C$ group, i.e. the Lie group describing the Quantum Chromo Dynamics (QCD).
Quarks cannot be observed as free particles due to the property of the QCD known as confinement. They can only be observed as quark-antiquark (*mesons*) or three-quark (*baryons*) bound states.
Fermions are also divided into three *generations* (or families) according to the *flavour*. There exists a strong hierarchy in the mass of the fermions among the different families, with the third family being much heavier than the other two, both in the lepton and in the quark sector. The meaning of this classification into generations will be more clear in section 1.2.2 when discussing the electroweak theory.
The properties of the fermions are summarized in table 1.1.

The fundamental forces are described in the SM as interactions mediated by *vector bosons*, i.e. spin 1 particles, which belong to the adjoint representation of the gauge group. In the SM the gauge bosons are the *photon*, mediator of the electromagnetic interaction, the *W and Z bosons*, responsible for the weak force and the *gluons* which are the carriers of the color interaction. The photon, the W and the Z

Table 1.1: The properties of the fermions.

1^{st} generation	2^{nd} generation	3^{rd} generation	Electric charge (in units of proton charge)	Interactons
		Leptons		
ν_e	ν_μ	ν_τ	0	weak
e	μ	τ	-1	E.M., weak
		Quarks		
u	c	t	+2/3	E.M., weak, strong
d	s	b	-1/3	

are the quanta of the boson fields associated with the $SU(2)_I \otimes U(1)_Y$ gauge group, which describes the electromagnetic and weak interactions in the unified electroweak theory. The gluons arise as excitations of the gauge fields associated to the $SU(3)_C$ of the QCD theory.
The properties of the fundamental interactions and the corresponding bosons are listed in table 1.2.

Table 1.2: The properties of the fundamental interactions and of their bosons.

Force	Electromagnetic	Weak	Strong
Quantum	photon (γ)	$W^{\pm}$, Z	gluons
Mass [GeV/c^2]	0	80.39, 91.188	0
Coupling constant	$\alpha_{\mathrm{e.m.}}(Q^2=0) \approx 1/137$	$G_F/(\hbar c)^3 \approx 1 \cdot 10^{-5}\,\mathrm{GeV}^{-2}$	$\alpha_{\mathrm{s}}(Q^2=m_Z) \approx 0.1$

The field content of the SM is completed with the scalar field (spin 0), whose quantum, the *Higgs boson* is not only the subject of the next sections in this chapter, but is the main character of the story told in this thesis.

1.2.2 The electroweak theory

At low energies, the weak force is described by the following effective lagrangian:

$$\mathcal{L}_{\mathrm{Fermi}} = -\frac{G_{\mathrm{F}}}{\sqrt{2}}\bar{\nu}_\mu \gamma^\alpha (1-\gamma_5)\mu \bar{e}\gamma_\alpha(1-\gamma_5)\nu_e \,, \tag{1.1}$$

which was firstly formulated by Fermi in 1933 and then completed to include the axial current, according to the experimental observartion of maximal parity violation in the weak interactions.
G_{F} is the Fermi constant reported in table 1.2.
The $1-\gamma_5$ negative helicity projector selects only left-handed fermions to participate in the interaction.
Even though it has been demonstrated to describe succesfully the weak interactions at low energies, the lagrangian in equation 1.1 is non-renormalizable and gives rise, at high energies, to non unitarities in the scattering amplitudes.

Both problems are overcome if the interaction is derived in the context of a gauge theory, i.e. a theory which is invariant under the local transformations induced by the generators of some Lie group. The choice of the group is driven by the observation of leptons behaving like left-handed doublets and right-handed singlets under the weak interactions, which suggests the use of the $SU(2)_I$ group:

$$L_L = \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L , \ \ell_R \ , \ Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}_L , \ u_R, \, d_R \tag{1.2}$$

The success of the Quantum Electrodynamics (QED), based on the abelian $U(1)_{\mathrm{e.m.}}$ group, in describing quantum-relativistic electromagnetic processes leads to the choice of the $SU(2)_I \otimes U(1)_Y$ as a gauge group to perform the unification of weak and electromagnetic interactions, with g and g' being the coupling constants of the two subgroups in the product. The generators of $SU(2)_I$ are $t^a = \frac{1}{2}\tau^a$ where τ^a are the Pauli matrices and the associated quantum number is the *weak isospin* I, while Y is the *weak hypercharge*. The electric charge, the hypercharge and the third component of the weak isospin (I_3) satisfy the following equation, known as the *Gell-Mann - Nishijima formula*:

$$Q = I_3 + \frac{1}{2}Y \tag{1.3}$$

The quantum numbers of the fermions of the SM are summarized in table 1.3.
The right-handed neutrino is, as far as we know, a "sterile" particle since it does not have any interaction in the SM.

Table 1.3: The quantum numbers of the SM fermions.

	I_3	Y	Q
ν_L	1/2	-1	0
ℓ_L	-1/2	-1	-1
ν_R	0	0	0
ℓ_R	0	-2	-1
u_L	1/2	1/3	2/3
d_L	-1/2	1/3	-1/3
u_R	0	4/3	2/3
d_R	0	-2/3	-1/3

The physical fields are given in terms of linear combinations of the $SU(2)_I$ gauge fields $W_\mu^{1,2,3}$ and the $U(1)_Y$ gauge field B_μ:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \\ A_\mu &= B_\mu \mathrm{cos}\theta_\mathrm{w} + W_\mu^3 \mathrm{sin}\theta_\mathrm{w} \\ Z_\mu &= -B_\mu \mathrm{sin}\theta_\mathrm{w} + W_\mu^3 \mathrm{cos}\theta_\mathrm{w} \end{aligned} \tag{1.4}$$

The rotation from the gauge fields to the physical fields is achieved with *Weinberg angle* θ_w.

The interaction Lagrangian is given as:

$$\begin{aligned}
\mathcal{L}_{\rm int} =& g\sin\theta_{\rm w} J_\mu^{\rm e.m.} A^\mu \\
&+ \frac{g}{\cos\theta_{\rm w}}(J_\mu^3 - \sin^2\theta_{\rm w} J_\mu^{\rm e.m.}) Z^\mu \\
&- \frac{g}{\sqrt{2}} \sum_i \left[\bar{u}_i \gamma^\mu \frac{1-\gamma^5}{2} M_{ij}^{\rm CKM} d_j + \bar{\nu}_i \gamma^\mu \frac{1-\gamma^5}{2} \ell_i \right] W_\mu^+ + {\rm h.c.}
\end{aligned} \tag{1.5}$$

In the first term in equation 1.5, the electromagnetic current is given by $J_\mu^{\rm e.m.} = \sum_f q_f \bar{f}\gamma_\mu f$. By imposing this coupling to be the same as the QED one, we obtain $g\sin\theta_{\rm w} = g'\cos\theta_{\rm w} = e$.
The second term is the weak neutral current where $J_\mu^3 = \sum_f I_{3,f} \bar{f}\gamma_\mu \frac{1-\gamma^5}{2} f$.
The third term is the charged weak current, in which the *Cabibbo-Kobayashi-Maskawa matrix* $M^{\rm CKM}$ is introduced, rotating the mass quark eigenstates into the interaction eigenstates.

1.2.3 The Higgs mechanism

In the Lagrangian of the SM no mass term can be explicitely introduced for the weak vector bosons, since it would violate the gauge invariance. For the fermions, a mass term of the type $m_f\bar{\psi}\psi$ is not invariant under *chirality* transformations, which is required by the GWS model. The *Higgs mechanism* is invoked to give masses to both weak bosons and fermions while keeping the photon massless, which implements the spontaneous symmetry breaking pattern:

$$SU(2)_I \otimes U(1)_Y \to U(1)_{\rm e.m.}$$

In its simplest realization, a complex scalar doublet field is introduced:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{1.6}$$

with $I = 1/2$ and $Y = 1$ quantum numbers.

The field in equation 1.6 enters the SM with the Lagrangian:

$$\mathcal{L}_{\rm EWSB} = (D_\mu\phi)^\dagger(D_\mu\phi) + V(\phi^\dagger\phi) \tag{1.7}$$

In equation 1.7 the covariant derivative $D_\mu = \partial_\mu - igt_a W_\mu^a - \frac{i}{2}g'YB_\mu$ is used to have gauge invariant interaction terms.
The potential is defined as:

$$V(\phi^\dagger\phi) = -\mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2,\ \mu^2 < 0,\ \lambda > 0 \tag{1.8}$$

The potential is minimized by a manifold of solutions given by:

$$\phi^\dagger\phi = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2} \tag{1.9}$$

The choice of the ground state or *vacuum* is arbitrary and it is referred to as the *spontaneous symmetry breaking*.
For instance, if one aligns the ground state along the ϕ^0 axis, the doublet can be written, in a generic gauge, as:

$$\phi = \frac{1}{\sqrt{2}} e^{\frac{i}{v}\phi_a t^a} \begin{pmatrix} 0 \\ \mathrm{h}+v \end{pmatrix}, \; a = 1,2,3 \tag{1.10}$$

In equation 1.10 $\phi_4 \equiv \phi_4(x) = h(x) + v$ and h is the *Higgs field*.

If one fixes the gauge with the transformation $\phi \to \phi' = e^{-\frac{i}{v}\phi_a t^a}\phi$ (*unitary gauge*) and rewrites the Lagrangian in equation 1.7 using the parameterization in 1.10 and the unitary gauge, then the masses of the weak bosons arise due to the non-zero vacuum expectation value (vev) of the scalar field v:

$$\begin{aligned} m_W &= \frac{1}{2} v g \\ m_Z &= \frac{1}{2} v \sqrt{g^2 + g'^2}, \end{aligned} \tag{1.11}$$

while the photon A_μ remains massless. This is connected to the fact that the ground state is still invariant under the $U(1)_{\mathrm{e.m.}}$ group and thus the QED is unbroken.

From equation 1.7 one gets also the interactions between the Higgs and the weak boson fields which are proportional to the mass of the boson squared.
It has to be noticed that in the electroweak symmetry breaking (EWSB) mechanism the three Goldston bosons $\phi_{1,2,3}$ have been "eaten" by the longitudinal polarization of the weak bosons thus preserving the total number of degrees of freedom and h is the only scalar field which survives.

The vev v can be extracted using the relation with the Fermi constant:

$$v = \left(\frac{1}{\sqrt{2}G_{\mathrm{F}}}\right)^{1/2} \simeq 246\,\mathrm{GeV} \tag{1.12}$$

The Higgs mechanism can also be used to generate the mass of the fermions through the interaction between the scalar and the fermion fields, called *Yukawa* interaction.
The lepton masses arise from the gauge invariant interaction:

$$\mathcal{L}_{Y,\ell} = g_{H\ell}\bar{L}_L\phi\ell_R + \mathrm{h.c.} \tag{1.13}$$

It can be noticed that from equation 1.10 only mass terms for the down component of the iso-doublets L_L (defined in equation 1.2) appear, which is consistent with the assumption of massless neutrinos.

The down quarks interact with the Higgs field with a term similar to the one in equation 1.13, while the up quarks must couple to the charge-conjugate of ϕ, defined as:

$$\phi^C = -i\tau_2\phi^* \tag{1.14}$$

Thus the Yukawa Lagrangian in the quark sector is:

$$\mathcal{L}_{Y,q} = g_{Hd}\bar{Q}_L\phi d_R + g_{Hu}\bar{Q}_L\phi^C u_R + \text{h.c.} \tag{1.15}$$

The masses of the fermions arising from equations 1.13 and 1.15 are of the form $m_f = \frac{g_{Hf}}{\sqrt{2}}v$ and the couplings of the fermions to the Higgs field are proportional to their mass.

1.3 Beyond the Standard Model

Despite its undisputed success, the SM is not believed to be the ultimate theory of the fundamental interactions. This consideration arises mainly from the observations of phenomena which are not explained within the SM theory, such as dark matter (DM) or matter-antimatter asymmetry. In addition, the SM has some features, which, even though are not problematic, in the sense that do not question the self-consistency of the theory, are believed to be unnatural. In the community of high energy physics there is a strong belief that a more complete theory should appear at some scale, to solve the internal issues of the SM and give a satisfactory answer to the open questions in particle physics.

In section 1.3.1 the most convincing points to go beyond the SM are briefly summarized and in section 1.3.2 a BSM model with an additional Higgs boson is described in some detail.

1.3.1 Reasons to go beyond the SM

Dark matter and dark energy

To our knowledge only around 5% of the universe is composed by ordinary baryonic matter. The remaining content is attributed to *dark matter* ($\approx 27\%$) and *dark energy* ($\approx 68\%$). The DM was first postulated in 1932 by Jan Oort to explain the orbital velocities of stars in the Milky Way, but the first to invoke it based upon robust evidences was Vera Rubin in the 1960s to describe the galaxy rotation curves.
The DM is believed to be made of weakly interacting massive particles (WIMPs) which interact with the ordinary matter only through the gravitational and weak force. This means that it does not emit neither absorb light, and thus is invisible to our telescopes. Other evidences for the presence of dark matter come from the measurements of the gravitational lensing of galaxy clusters.
The dark energy (DE) is an hypothetical form of energy which permeates the universe

and contributes to its expansion. It is believed to be evenly distributed in space in a way that does not have local gravitational effects. The presence of such large dark energy contribution is mainly inferred from the measurement of the anysotropy in the cosmic microwave background radiation (CMB).

DM particle candidates are hypothesized in many BSM theories. In Supersymmetry (SUSY), many models identify the *lightest supersymmetric particle* (LSP) with the cold dark matter. There are also non-supersymmetric models which attempt to describe dark matter features making use of *sterile neutrinos*. *Hidden valley* models instead predict the existence of a "hidden" sector, described by additional gauge groups, which interact with SM fermions only gravitationally and whose particle may be a good candidate for the DM.

Matter-antimatter asymmetry

The laws which describe the fundamental interactions predict particles and antiparticles to be always produced in such a way that the *baryon number* is conserved, which is defined as:

$$B = \frac{1}{3}(n_q - n_{\bar{q}}) \tag{1.16}$$

This means that at the time of Big Bang an equal amount of matter and antimatter should have been produced and, given that they annihilate in pairs, in the universe there should have remained nothing left but energy. The fact that we observe stars and galaxies made of matter is a hint that there had been some mechanism at work which caused a small unbalance in favour of matter.
Such processes, which proceed via CP-violation interactions, are predicted also in the SM, but they are too tiny to generate the needed amount of asymmetry. Other sources of CP-violation are considered in BSM theories.

The naturalness of the Higgs mass

In the SM the mass of the scalars receive loop corrections from diagrams as the ones shown in figure 1.1.

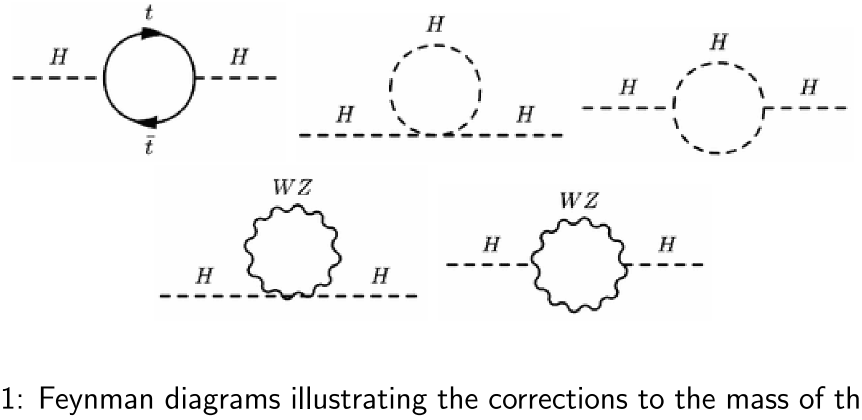


Figure 1.1: Feynman diagrams illustrating the corrections to the mass of the scalar.

If the SM is assumed to be valid up to a "cutoff" scale Λ, the physical mass of the scalar can be parameterized as:

$$m_h^2 = m_{h,0}^2 + \delta m_h^2$$
$$\delta m_h^2 = \frac{3\Lambda^2}{8\pi^2 v^2}(2m_W^2 + m_Z^2 + m_{h,0}^2 - 4m_t^2) \tag{1.17}$$

In the SM there is no symmetry that protects scalar masses from receiving large corrections. From equation 1.17, if the SM is assumed to be valid up to $\Lambda = 5\,\mathrm{TeV}$ then the physical Higgs mass m_h would result from an "unnatural" cancellation between two quantities, $m_{h,0}$ and δm_h^2, which are $\approx$ 100 times larger than m_h. This problem is referred to as the *naturalness* of the Higgs mass, the *hierarchy* problem and also *fine-tuning* problem.

Even though the fine-tuning of the Higgs mass parameter would not be itself an inconsistency of the SM, it is widely believed that there should be some new physics models providing a mechanism to avoid large corrections to the scalar masses. An elegant solution is found in Supersymmetry, where for each diagram in figure 1.1 there is a similar one involving the superpartner particle in the loop and which aquires opposite sign. The cancellation between the two is not perfect due to the fact that SUSY is broken, but anyway good enough to restore a natural behaviour.

The vacuum stability

The stability of the SM ground state depends on the potential in equation 1.8. In particular, going to high scales, the quartic term in the potential dominates and we get $V_{\text{eff}} \sim -\lambda(\mu)(\phi^\dagger\phi)^2$, where μ is the renormalization scale. The running quartic coupling $\lambda(\mu)$ receives decreasing contribution to its evolution from the large Yukawa top coupling, thus avoiding any Landau pole.
However, a full NNLO analysis of the scalar potential [6] shows that, assuming $m_h = 125\,\mathrm{GeV}$, the Higgs quartic coupling becomes negative at high energy, as it is shown in figure 1.2 (left). This means that the scalar potential is unbounded from below and the SM cannot be a consistent theory up to the Planck scale $M_{\text{Planck}} \approx 10^{19}\,\mathrm{GeV}$. Even if we assume some mechanism to restore the boundness of the potential at the Planck scale, we still may have negative λ in a region between $10^{10} - 10^{19}\,\mathrm{GeV}$, which would cause the potential to have a global minimum different from the ground state of the SM. In principle, the SM vacuum may decay into this global minimum state by means of quantum tunneling and this is what is called a *metastable configuration*. The SM phase diagram is shown in figure 1.2 (right): as one can see the measured values of the top and Higgs mass strongly disfavour the stability region. Even though the decay time may be far larger than the age of the universe, still it is true that the universe started in a configuration whose energy is far from the minimum, which is an intriguing observation by itself and leads to many speculations of new physics appearing at high scales.

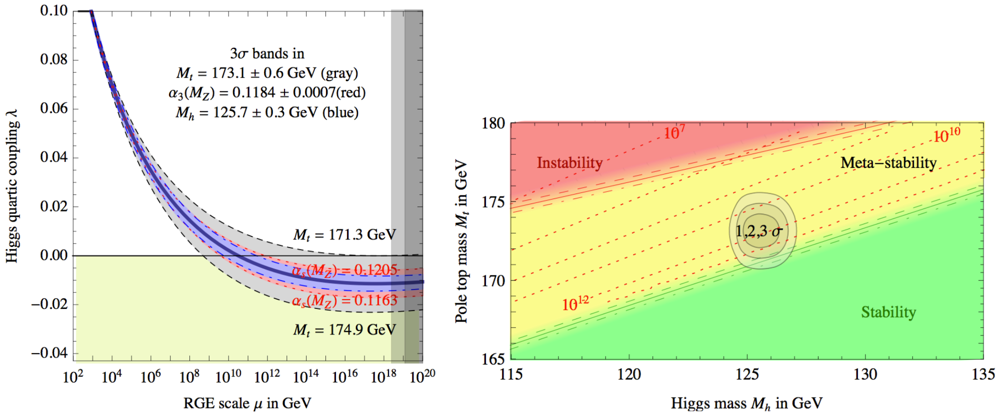


Figure 1.2: (left) Renormalization group (RG) evolution of the quartic coupling of the scalar potential λ. (right) Phase diagram of the SM showing the stability condition of the vacuum as a function of the Higgs and top masses [6].

1.3.2 Electroweak singlet

The theories that are generically labelled *electroweak singlet models* [7,8,9,10] include a variety of different scenarios in which the SM scalar sector is minimally extended with an additional state. These models are usually built upon the hypothesis that there is a "phantom" or "hidden" sector populated by SM singlet fields, with fermions, gauge bosons and scalars organized in gauge groups as in the SM. The simplest realization of the hidden sector is the abelian $U(1)_{\text{hid}}$.
The hidden sector can be connected to the SM via renormalizable interactions of two types:

- the kinetic mixing proportial to $B_{\mu\nu}C^{\mu\nu}$, where $\boldsymbol{B}, \boldsymbol{C}$ are the field strength tensors of $U(1)_Y$ and $U(1)_{\text{hid}}$ respectively;
- the mixing term between the SM scalar field Φ_{SM} and the hidden sector Higgs Φ_H, which is of the form $|\Phi_{\text{SM}}|^2|\Phi_H|^2$.

The first kind of models leads to the so-called Z' phenomenology. For the purpose of this work, in particular referring to the analysis described in chapter 6, we focus on the second type of models, which applies to a more generic gauge structure for the hidden sector.
The gauge theory of the hidden sector is at least partly broken by a non-null vacuum expectation value $< \Phi_H > \neq 0$: this requirement is needed to ensure the mixing with the SM Higgs, leading to an additional massive state which is the subject of the search discussed in chapter 6.

The Lagrangian of the extended scalar sector is:

$$\begin{aligned}\mathcal{L}_{\text{scalar}} =& |D_\mu \Phi_{\text{SM}}|^2 - \mu^2 \Phi_{\text{SM}}^\dagger \Phi_{\text{SM}} - \lambda(\Phi_{\text{SM}}^\dagger \Phi_{\text{SM}})^2 \\ &+ |\tilde{D}_\mu \Phi_H|^2 - \mu_P^2 |\Phi_H|^2 - \lambda_P |\Phi_H|^4 \\ &- \eta \Phi_{\text{SM}}^\dagger \Phi_{\text{SM}} |\Phi_H|^2 \end{aligned} \tag{1.18}$$

In equation 1.18 the first line is the SM scalar lagrangian (equation 1.7), the second line is the hidden sector symmetry breaking lagrangian (with $\tilde{D}_\mu$ containing the gauge fields of the hidden sector) and the third line is the interaction term which induces the mixing between the SM Higgs and the hidden sector Higgs.
The expansion of the two scalar fields around their vevs v, ξ can be written as:

$$\Phi_{\mathrm{SM}} = \frac{1}{\sqrt{2}} \begin{pmatrix} G^{\pm} \\ \phi_{\mathrm{SM}} + v + iG^0 \end{pmatrix}, \; \Phi_H = \frac{1}{\sqrt{2}}(\phi_H + \xi + iG') \tag{1.19}$$

where in equation 1.19 the fields G are the Goldstone bosons, which are removed from the calculation using the unitary gauge.
The two mass eigenstates h, H can be obtained by rotating the fields ϕ_{SM} and ϕ_H with the mixing angle ω:

$$\begin{cases} h & = c_\omega \, \phi_{\mathrm{SM}} - s_\omega \, \phi_H \\ H & = s_\omega \, \phi_{\mathrm{SM}} + c_\omega \, \phi_H \end{cases} \tag{1.20}$$

where $c_\omega \equiv \cos\omega$ and $s_\omega \equiv \sin\omega$. The mixing angle is given by:

$$\tan\omega = \frac{\eta v \xi}{(-\lambda v^2 + \lambda_P \xi^2) + \sqrt{(\lambda v^2 - \lambda_P \xi^2)^2 + \eta^2 v^2 \xi^2}} \tag{1.21}$$

The masses of the two Higgs states are:

$$m^2_{h,H} = (\lambda v^2 + \lambda_P \xi^2) \pm \sqrt{(\lambda v^2 - \lambda_P \xi^2)^2 + \eta^2 v^2 \xi^2} \tag{1.22}$$

If $m_H > 2m_h$ the decay mode $H \to hh$ is kinematically allowed. The partial width for this decay is:

$$\Gamma(H \to hh) = \frac{|\mu_T|^2}{8\pi m_H} \sqrt{1 - \frac{4m_h^2}{m_H^2}} \tag{1.23}$$

where in equation 1.23 μ_T is the effective coupling appearing in the trilinear mixing term $\mu_T h^2 H$:

$$\mu_T = -\frac{\eta}{2}(\xi c_\omega^3 + v s_\omega^3) + (\eta - 3\lambda) v c_\omega^2 s_\omega + (\eta - 3\lambda_P) \xi c_\omega s_\omega^2 \tag{1.24}$$

The mixing between the scalars results in a SM-like phenomenology also for the hidden sector Higgs, which has the same decay modes as the SM Higgs. The couplings with the SM particles will be suppressed with respect to the prediction of the SM, for both h and H, by a factor of $C \equiv c_\omega$ and $C' \equiv s_\omega$ respectively.
In the following, we will identify the quantum of the h field with the newly discovered Higgs-like particle with mass $m_h \simeq 125\,\mathrm{GeV}$, and we will assume $m_H > m_h$.
The signal strengths, defined as the ratio of the production cross section times branching fraction to the prediction of the SM, will scale as:

$$\begin{aligned} \mu_h &= C^2 \\ \mu_H &= C'^2(1 - \mathrm{BR}_{\mathrm{new}}) \end{aligned} \tag{1.25}$$

where the $\mathrm{BR}_{\mathrm{new}}$ parameter is introduced to account for all the undetected decay modes, including the $H \to hh$ channel.
The widths of the two resonances are also modified as:

$$\begin{aligned}
\Gamma_h &= C^2 \Gamma_{\mathrm{SM}}(m = m_h) \\
\Gamma_H &= \frac{C'^2}{1 - \mathrm{BR}_{\mathrm{new}}} \Gamma_{\mathrm{SM}}(m = m_H)
\end{aligned} \tag{1.26}$$

The phenomenology induced by the mixing between the two states thus provides one additional resonance at mass $m_H > 125\,\mathrm{GeV}$ with a narrower width and a smaller cross section. Clearly, this class of models can be constrained both from the measurement of the boson at 125 GeV and from the direct search for additional an resonance, which will be discussed in chapter 6.

Chapter 2

The Compact Muon Solenoid experiment at the Large Hadron Collider

Abstract

The Large Hadron Collider (LHC) is the world's biggest and highest energy particle accelerator ever built and is expected to shade light on the open questions in high energy physics. The Compact Muon Solenoid (CMS) experiment at the LHC is a multi-purpose detector designed to study a wide range of physics processes, like the production of a Higgs boson or supersymmetric particles.

Thanks to the incredible success of the LHC operations during the Run 1 and the effort of thousands of physicists around the world, the CMS and ATLAS collaborations were able to discover a Higgs boson.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest particle accelerator in the world colliding beams of protons and heavy ions at an unprecendented energy. It has been built by the European Organization for Nuclear Research (CERN) with the goal of investigating the open issues of the Standard Model. There are infact many hints that interactions at the TeV energy scale may help in understanding the mechanism of electroweak symmetry breaking or reveal the existence of new massive particles (Supersymmetry): a hadron collider with a target center-of-mass energy of 14 TeV is thus a wonderful opportunity to discover new physics. The LHC also smashes lead ions at an ultra-relativistic regime, allowing the study of the matter at extreme conditions of temperature and density, when the the Quark Gluon Plasma state is believed to occur.

This chapter is organized as follows. In section 2.1 the design of the LHC is outlined and a summary of the LHC operations during Run 1 is given. In section 2.2 the physics of the Higgs sector at the LHC is described with some detail and the discovery of a Higgs boson with mass around 125 GeV during the Run 1 is highlighted. In section 2.3 the CMS detector is introduced and a brief description of its main subsystems is given.

2.1.1 The LHC design

The LHC is a 27 km circular accelerator, which lies in the French-Swiss border in the Geneva area, at a depth varying between 50 and 175 m. The tunnel hosting the LHC was formerly built in 1983-1988 for the Large Electron Positron collider (LEP).
The LHC was designed to collide beams of protons at an energy of 7 TeV each, resulting in a total center-of-mass energy of 14 TeV. In addition to the physics program with protons, it collides lead ions at 2.76 TeV/NN.
The LHC ring is made of two beam pipes, which make the two beams circulate in opposite directions and make them collide in the four interaction points where the detectors are placed. A schematic layout of the LHC is in figure 2.1.

The beams are kept on their circular path by means of 1232 dipole magnets along the tunnel and are focused by 392 quadrupole magnets. The magnets of the LHC are superconducting and are kept at an operational temperature of 1.9 K using about 96 tonnes of liquid helium. The field in the magnet varies between 0.53-8.3 T during the acceleration of the protons from 450 GeV to 7 TeV.
Before reaching the LHC, the protons undergo a series of accelarating steps, which starts with the linear particle accelerator (LINAC2), which generates 50 MeV protons to be injected in the Proton Synchrotron Booster (PSB). In the PSB the protons reach the energy of 1.4 GeV and are transferred to the Proton Synchrotron (PS). They are then injected with an energy of 26 GeV inside the Super Proton Synchrotron (SpS) where their energy is increased up to 450 GeV. At this stage the proton are injected in the LHC ring, where bunches are accumulated and accelerated to the nominal energy over a period of about 20 minutes.

The luminosity is a crucial parameter for the LHC, since the physics program is mainly focused on the study of rare processes, like the production of the Higgs boson.

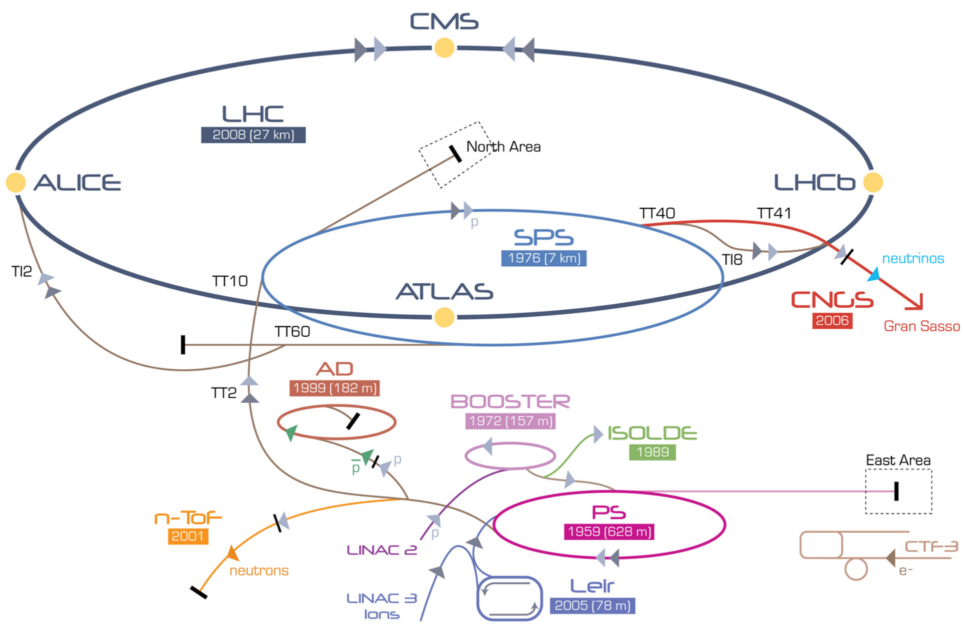


Figure 2.1: A schematic layout of the complex of accelartors at CERN.

The luminosity can be written in terms of the parameters of the LHC accelerator, assuming a gaussian beam shape:

$$\mathcal{L} = \frac{N_b^2 n_b f_{\mathrm{rev}} \gamma}{4\pi \epsilon_n \beta^*} F \tag{2.1}$$

In equation 2.1 the following quantities are defined:

- $N_b \simeq 10^{11}$ is the number of particles per bunch;
- $n_b = 2808$ is the number of bunches per beam;
- f_{rev} is the revolution frequency;
- $\gamma = 7461$ is the relativistic Lorentz factor at 7 TeV per beam;
- $\epsilon_n = 3.75\,\mu\mathrm{m\,rad}$ is the normalized transverse beam emittance;
- $\beta^* = 0.55\,\mathrm{m}$ is the optical Beta function at the collision point;
- $F = 0.836$ is the geometrical reduction factor due to the crossing angle at the interaction point.

The design peak luminosity of the LHC, referring to the ATLAS and CMS experiments, is $10^{34}\,\mathrm{cm^{-2}s^{-1}}$.

2.1.2 The LHC operations during Run 1 (2008-2013)

The *Run 1* of the LHC refers to the collisions recorded during the first phase of the physics program between 2008 and 2013. The original schedule of the accelerator

was delayed by about one year due to the accident on 19th September 2008, just 9 days after the first protons entered the main ring. The accident happened during the commissioning of the dipole magnet circuit in sector 3-4. A resistive zone developed in the electrical bus between a dipole and a quadrupole, causing the energy discharge of the magnet system. Shortly after, an electrical arc developed causing the damage of the helium enclosure and the release of the helium in the insulation vacuum of the cryostat. Quenches started developing in neighbouring subsector due to electrical noise in the quench detector. The quench relief valves started releasing the helium in the tunnel (as expected), but in subsectors 23-25 they were unable to contain the pressure rise below the nominal value and the pressure forces severely damaged the vacuum barriers and the mechanical supports of the magnets.
In total 6 tonnes of helium were lost and 5 quadrupoles plus 24 dipoles were damaged in the accident.

Most of 2009 was spent to repair the LHC machine and only in November 2009 the first collisions at 450 GeV energy in the center of mass were delivered. The energy was then increased until the November 30th, when the LHC became the world's highest energy particle accelerator, reaching 1.18 TeV per beam and beating the previous record by the Tevatron (0.98 TeV per beam).
On March 30th of 2010 two proton beams in the LHC collided at a center-of-mass energy of 7 TeV, marking the begin of the LHC physics program. In November of the same year, the LHC delivered the first lead ion collision run. On December, with the end of the heavy ion run, the LHC was shutdown and restarted on March 2011.
Another milestone was reached on the 21th of April 2011, when the LHC became the world's highest luminosity accelerator with a peak luminosity of $4.67 \cdot 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, beating once again the Tevatron's record of $4 \cdot 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.
In October 2011, the ATLAS and CMS experiments reached $5\,\mathrm{fb}^{-1}$ of collected data and reported the firt hints of a signal in the search for the SM Higgs boson.
With the real possibility of a discovery before the first 2-year long shutdown in 2013-2014, it was decided to increase the energy up to 4 TeV per beam, slightly changing the schedule of the LHC. On April 2012, the first 8 TeV energy collisions were delivered and on July the 4th the ATLAS and CMS collaborations announced the discovery of a new resonance with mass around 125 GeV. By the end of the year the ATLAS and CMS collaborations both collected around $20\,\mathrm{fb}^{-1}$ of data at this energy.

A graphical summary of the integrated luminosity and the peak luminosity achieved by the CMS experiment is shown in figure 2.2.

2.2 The physics of the Higgs boson at the LHC

One of the main points of the physics program of the ATLAS and CMS experiments at the LHC is the search for the Higgs boson. The phenomenology of the Higgs in proton-proton collisions is briefly reviewed in sections 2.2.1 and 2.2.2. In section 2.2.3 the results of the search for the SM Higgs boson are presented, culminated with the discovery of a new particle with mass around 125 GeV [11, 12, 13] and the first measurements of its properties.

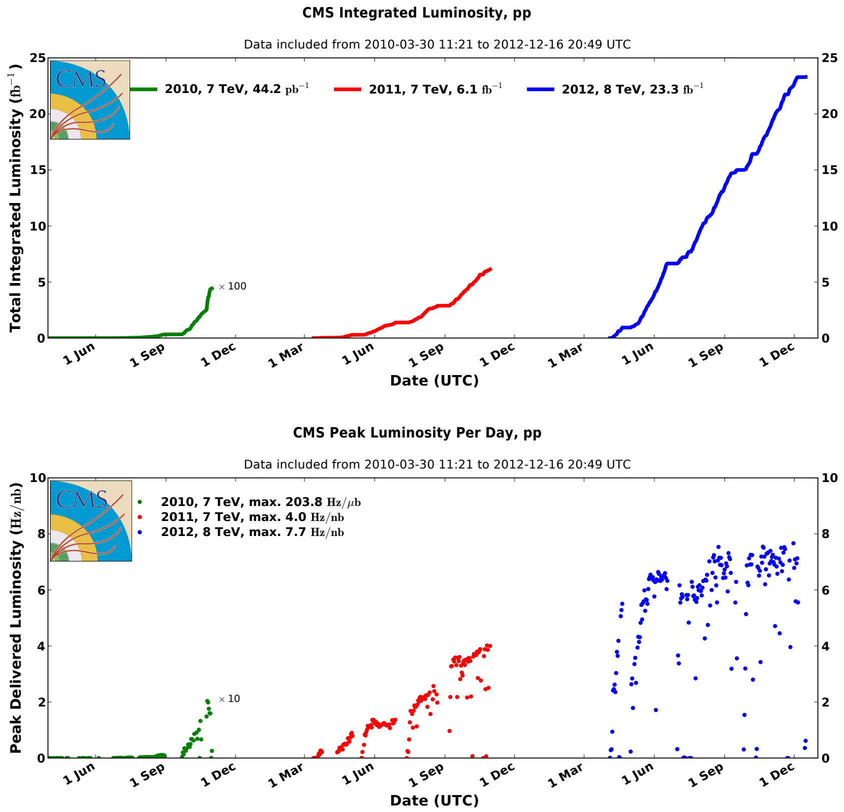


Figure 2.2: (top) The integrated luminosity collected by the CMS experiment during Run 1, as a function of time. (bottom) The peak luminosity recorded by the CMS experiment during Run 1, as a function of time.

2.2.1 Higgs production

There exist four main production mechanisms for the Higgs boson in the SM, which are sketched in figure 2.3:

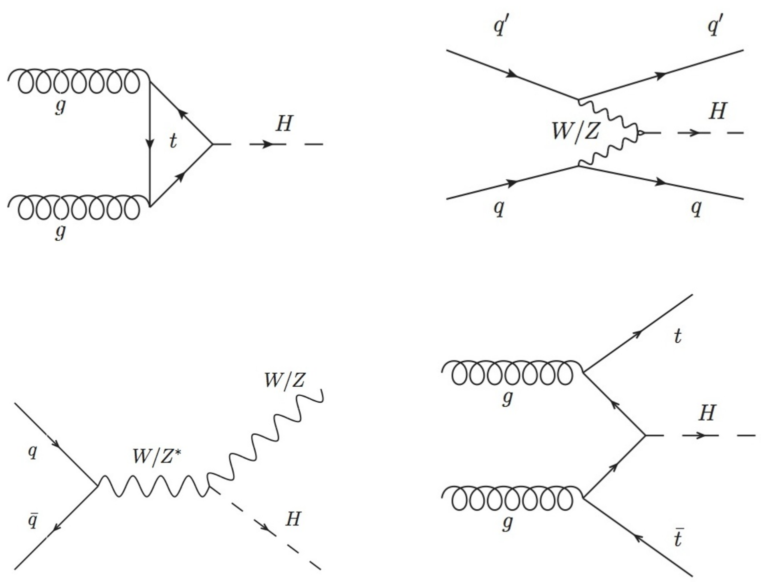


Figure 2.3: Feynman diagrams corresponding to the lowest order contributions to the gluon fusion (top left), the vector boson fusion (top right), the Higgs-strahlung (bottom left) and the $t\bar{t}$ associated production mechanism.

- The **gluon fusion** (ggH) is the dominant production mode at the LHC for every value of the mass of the Higgs (see figure 2.4). The main contribution to this process arises from the t quark loop, due to the Yukawa coupling being proportional to the mass of the fermion. This process is also particularly sensitive to new physics contributing to the loop amplitudes, like the presence of a fourth fermion family.
  The gluon-fusion cross section is known at NNLO+NNLL accuracy in QCD and NLO in electroweak. The radiative QCD corrections are of particular importance: the cross section increases by a factor of two going from LO to NLO, another 25% comes from the NNLO corrections and NNLL soft gluon resummation adds another 5%.

- The **vector boson fusion** (VBF) process accounts for $\approx 7\%$ of the total cross section for a Higgs of mass 125 GeV. Despite the small yield, this process provides a peculiar signature, with two jets with a large rapidity gap and large invariant mass.

- The **Higgsstrahlung** (associated production with a vector boson VH) is the process in which a vector boson (W or Z) radiates a Higgs. The VH contribution to the total cross section is less than 5%, but it is extremely important in some decay channels because the Higgs production can be "tagged" using the associated vector boson, thus reducing sensibly the background.

- The $t\bar{t}$ **associated production** (ttH) contributes to less than 1 % of the total signal yield at $\sqrt{s} = 8\,\mathrm{TeV}$. Nevertheless, the presence of a $t\bar{t}$-quark pair in the final state makes this process observable at the LHC, thanks to the use of techniques to identify the presence of jets arising from the fragmentation of b-quarks (b-tagging). Moreover, the cross section is expected to increase by a factor of approximately 4 going from 8 TeV to 13 TeV, thus increasing the possibility to observe this process.

The Higgs production cross section as a function of the Higgs mass is shown in figure 2.4.

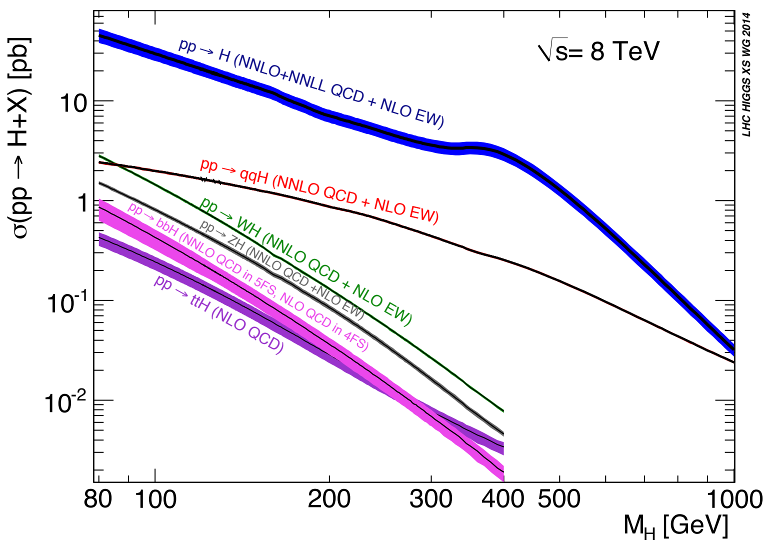


Figure 2.4: The cross section for the Higgs production in proton-proton collisions at a center-of-mass energy of 8 TeV as a function of the Higgs mass. The coloured bands represent the total theoretical uncertainty.

2.2.2 Higgs decay channels

The Higgs in the SM couples with a strength proportional to the mass and to the mass squared for fermions or gauge bosons respectively. For this reason the Higgs tends to decay into the heaviest particle which is kinematically available.
The branching ratio of the SM Higgs boson as a function of its mass is shown in figure 2.5.

The most important decay channels, from an experimental point of view, are not just those which provide the highest event yield, but rather those which have a high signal-to-background ratio. As an example, the decays into a pair of gluons or a pair of c-quarks, despite the relative large branching fraction for a low mass Higgs, are overwhelmed by the background from QCD processes, and thus have a little chance to be observed at the LHC.

In the following a brief review of the most sensitive decay channels is presented.

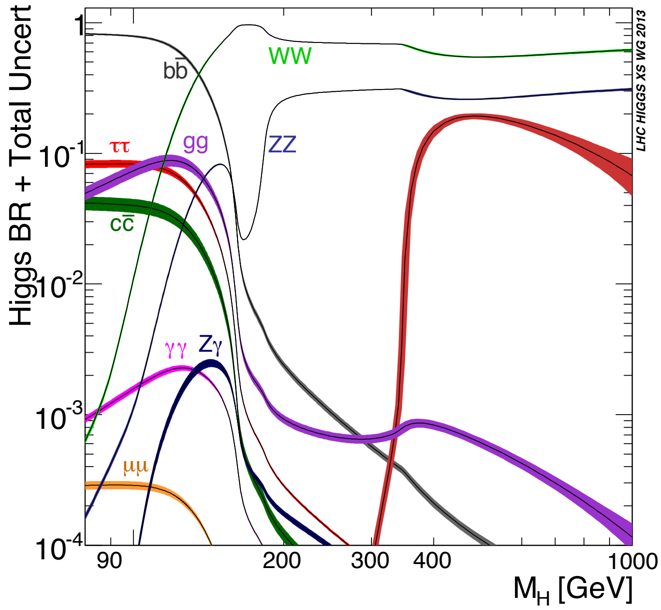


Figure 2.5: The SM Higgs boson branching ratio as a function of the Higgs mass. The coloured bands represent the total theoretical uncertainty.

$H \to ZZ \to 4\ell$

This channel looks for a narrow peak in the four-lepton invariant mass distribution appearing over a small continuum background. Small instrumental background may arise from the misidentification of jets as leptons. Due to different expected resolutions, the $4\mu, 4e, 2e2\mu$ final states are analyzed separately and then combined. The final state is completely reconstructed and measured, thus preserving the total kinematical information of the event. This feature makes it possible to study the mass and the spin of the Higgs boson with high precision.

$H \to \gamma\gamma$

In the $\gamma\gamma$ decay channel, a Higgs signal appears as a peak in the exponentially falling di-photon invariant mass spectrum. The large background comes from QCD diphoton production and misendification of jets which are reconstructed as photons. The signal yield is releatively small, compared to other channels, but the signature is clear (two high-E_{T} photons) and the mass resolution is very good ($\leq 2\,\mathrm{GeV}$): these two features makes this channel one of the most powerful to study a low mass Higgs boson.

$H \to WW \to \ell\nu\ell\nu$

The decay of the Higgs into a pair of W bosons has the largest branching ratio right above the WW threshold (160 GeV), as it is shown in figure 2.5. With the use of dedicated reconstruction algorithms for the transverse missing energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), this channel is sensitive also in the low mass regime, down to $m_H \simeq 120\,\mathrm{GeV}$.
The events are selected with two well isolated, high-p_{T} leptons and large missing transverse energy due to the presence of two neutrinos in the final state.

$H \to b\bar{b}$

Despite having the largest branching fraction for a low mass Higgs (see figure 2.5), the $b\bar{b}$ channel has a huge background coming from QCD di-jet production, whose cross section is several order of magnitudes larger than the Higgs signal. For this reason an inclusive search is not possibile with this final state and the events are selected requiring the additional presence of a W or Z boson, decaying leptonically, in order to target the VH associated production, which has a better signal to background ratio.

$H \to \tau\tau$

The branching ratio of a low mass Higgs into a τ-lepton pair is the second largest, after the $b\bar{b}$ mode. In this channel, the main challenge is the reconstruction of the τ leptons, which is difficult due to the presence of neutrinos in the leptonic decay and relatively soft pions in the hadronic decay. The channel is further divided into subchannels according to the different combinations of the τ decay modes. Since the resolution on the τ four-momentum is not as good as for muons, electron or photons, the Higgs signal would appear as a broad excess in the invariant mass of the τ pair.

2.2.3 The discovery of a Higgs boson

In December 2011, the ATLAS and CMS collaborations reported “intriguing hints” of the signal of a new particle in the search for the SM Higgs boson, using approximately $5\,\mathrm{fb}^{-1}$ of data at 7 TeV. The analysis was repeated in 2012 with the addition of the first $5\,\mathrm{fb}^{-1}$ at 8 TeV and indeed confirmed the hints seen in the 2011 dataset, leading to the announcement of the discovery in the famous public seminar given on July the 4th 2012 at CERN by the ATLAS spokeperson Fabiola Gianotti and the CMS spokeperson Joe Incandela.

The discovery of a new particle occurs when the p-value of the background-only hypothesis exceeds the level of a $5\,\sigma$ significance. This threshold corresponds to 1 chance over more than 3 millions to have an experimental outcome which is as less compatible with a fluctuation of the background as the one observed. In July 2012, both ATLAS and CMS reached the $5\,\sigma$ threshold with the combination of the most sensitive channels: $H \to ZZ^{(*)} \to 4\ell$, $H \to \gamma\gamma$, $H \to WW^{(*)} \to \ell\nu\ell\nu$. The significance of the excess seen by the two experiments is shown as a function of the hypothesized Higgs mass in figure 2.6.

After the discovery, the first measurements of the properties of the new resonances were performed.
The mass of the new boson, from the combination of the high resolution channels, is measured with the full dataset to be $m_H = 125.03\,^{+0.26}_{-0.27}\,\mathrm{stat.}\,^{+0.13}_{-0.15}\,\mathrm{syst.}\ \mathrm{GeV}$ in CMS. The mass measured by the ATLAS collaboration is in agreement with the CMS one within the current uncertainties.
In CMS, using the matrix element approach described in 4.10.4, the spin of the new particle was first tested against the pure pseudoscalar hypothesis [14] and later against other exotic models [15] in the $H \to ZZ^{(*)} \to 4\ell$ channel. The diphoton channel has also been shown to be sensitive to the spin properties of the new boson [16]. The

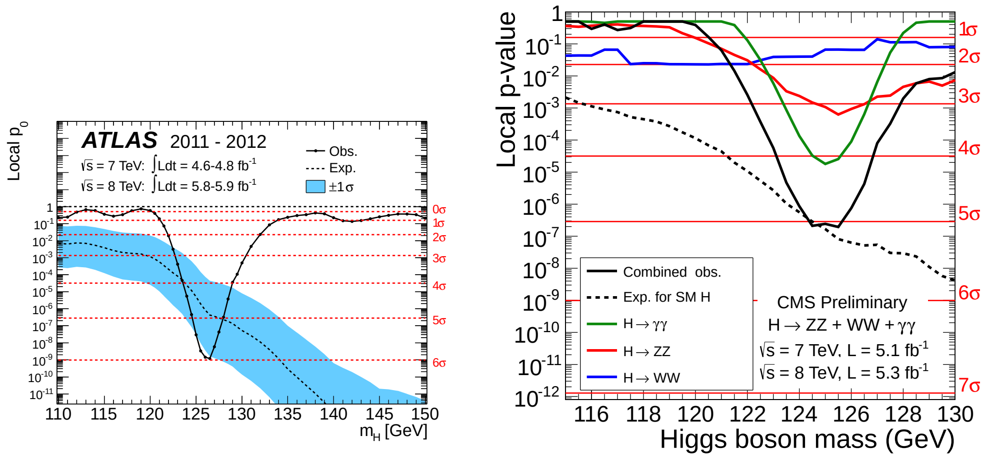


Figure 2.6: The local p-value as a function of the Higgs mass for the combination of the most sensitive channels in ATLAS (left) and CMS (right). The solid line is the observed p-value, the dashed one is the expected in the presence of a SM Higgs boson.

experimental data collected during Run 1 strongly favour the spin-parity quantum numbers (0^+) predicted by the SM.
The signal stregth μ, defined as the ratio of the measured rate to the prediction of the SM, has been measured in the combination of the main channels [17, 18]. The value measured in CMS is $\mu = 1.00 \pm 0.09\,(\mathrm{stat.})\,^{+0.08}_{-0.07}\,\mathrm{theo.} \pm 0.07\,\mathrm{syst.}$ and is found to be in good agreement with the measurement done in ATLAS, both being compatible with the value predicted in the SM ($\mu = 1$). Figure 2.7 (left) shows the two-dimensional scan of the $\mu_{\mathrm{VBF,VH}}$ and $\mu_{\mathrm{ggH,ttH}}$ parameters.
The couplings of the newly discovered Higgs-like particle were also measured with the full dataset of the Run 1. The approach used to test the compatibility of the couplings with the prediction from the SM makes use of the so-called k-framework [19]. Such method consists in defining effective coupling scale factors k_i for different particles or group of particles and rewriting the relevant cross sections in terms of those parameters. So far, no significant deviation is seen in the coupling structure of the new particle with respect to the SM expectations for a Higgs boson.

The width Γ_H of a SM Higgs boson of mass 125 GeV is very tiny, around 4 MeV, well below the reach of the experiments at the LHC for a *direct* measurement. However, it has been demonstrated that, by measuring the ratio of the off-shell to the on-shell cross section, one can obtain an *indirect* contraint on the total width [20], under the hypothesis that the couplings are the same off-shell and on-shell. This measurement has been recently performed in CMS [21] and ATLAS [22]. The CMS collaboration reports a 95% confidence level upper limit of $\Gamma_H < 22\,\mathrm{MeV}$, which corresponds to 5.4 times the width predicted by the SM. The ATLAS measurement yields limits in the same range of those from CMS.

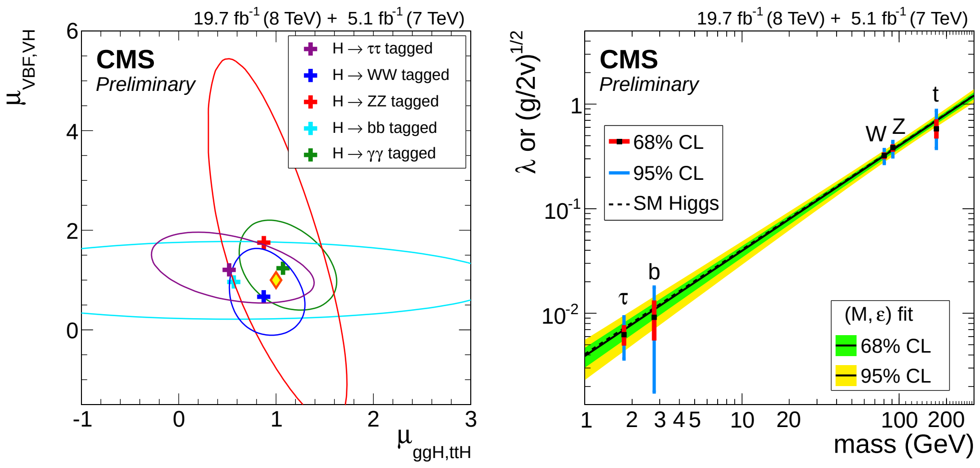


Figure 2.7: (left) The scan of the two signal strengths $\mu_{\mathrm{VBF,VH}}$ and $\mu_{\mathrm{ggH,ttH}}$ defined for the vector boson and fermion induced processes respectively. (right) The couplings of the SM Lagrangian as a function of the particle mass.

2.3 The CMS detector

The LHC is expected to deliver, at its design functioning, approximately 10^9 inelastic proton-proton interactions per second. This huge rate leads to an incredible variety of challenges for the detectors. First of all, the online event selection system (the *trigger*) must reduce the rate to about 100 events/s, in order to make it possible to record the interesting data. At the design peak luminosity, the experiments would record about 20 overlapping collisions in the same bunch crossing, referred to as pile-up (PU): a high-granularity detector with low occupancy is mandatory to properly reconstruct the events in such a crowded environment. The high flux of particle requires also the use of radiation-hard sensors and electronics.

In order to cope with these challenges and to fulfill the physics program of the LHC, the CMS detector must match some requirements on the reconstruction of the physics objects, which can be summarized as:

- good muon identification, charge distinction capability up to momenta of about TeV, good momentum resolution and dimuon mass resolution;
- good identification of electrons and photons and good diphoton, dielectron mass resolution;
- good dijet mass resolution;
- efficient b and τ tagging.

In the following a brief summary of the design of the CMS detector and of the technology used for each subsystem is presented. A detailed description can be found elsewhere [23].

2.3.1 The coordinate system

The CMS experiment uses a cylindrical coordinate system, whose origin is at the nominal collision point inside the detector. The y-axis points vertically upward, the x-axis points radially towards the centre of the LHC and the z-axis is along the beam direction. The azimuthal angle ϕ is measured in the $x-y$ plane from the x-axis, while the polar angle θ is measured from the z-axis. Instead of θ, the *pseudorapidity* η is often used, which is defined as:

$$\eta = -\ln\tan\frac{\theta}{2} \tag{2.2}$$

The transverse momentum p_{T} is defined as the magnitude of the projection of the momentum in the x, y plane. The transverse energy is defined as $E_{\mathrm{T}} = E\mathrm{sin}\theta$. The missing transverse energy, often referred to as MET and denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$, is the magnitude of the vectorial sum of the momentum of all the particles reconstructed in the event projected in the trasverse plane.
A variable which is often used to measure the distance between particle trajectories is ΔR, defined as:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \tag{2.3}$$

2.3.2 The overall design

The CMS detector is shown in figure 2.8.
The design of the CMS detector was mainly driven by the configuration of the magnetic field: a large bending power is mandatory to have a good momentum resolution for particle with high momentum.
For this reason a 13-m long, 6-m-diameter superconducting solenoid, provinding a magnetic field of 4 T has been placed at the heart of the experiment.
In the steel return yoke, 4 muon stations have been integrated, each made of several layers of aluminium drift tubes (DT) in the barrel and cathode strip chambers (CSC) and resistive plate chambers (RPC) in the endcap.
The inner tracker and the calorimetry systems are placed inside the solenoid. In the barrel, the tracking system consists of 10 layers of silicon strip detectors, complemented by 3 layers of silicon pixel detectors in the region closer to the interaction point.
The electromagnetic calorimeter (ECAL) is made of lead tungstate ($\mathrm{PbWO_4}$) crystals, whose scintillation light is detected using avalanche photodiodes (APDs) or vacuum phototriodes (VPTs) for the barrel and endcap respectively. A pre-shower detector is installed in front of the ECAL for the rejection of π^0.
The hadron calorimeter (HCAL) is a sampling calorimeter with brass as a passive material and scintillator as active material. The light is detected by hybrid photodiodes (HPDs). The outer HCAL (HO) further extends the hadron calorimeter in the barrel region to improve the containment of the tail of the hadronic showers. To improve the geometrical coverage of the HCAL, iron/quartz sampling calorimeter are placed in the forward region, where photomoltipliers (PMs) are used to detect the Cherenkov light emitted.

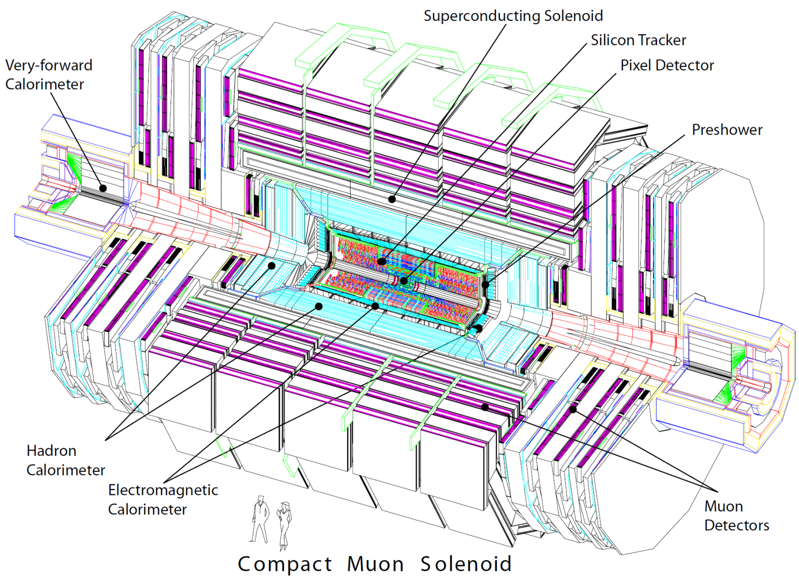


Figure 2.8: A section of the CMS detector, with the indication of each subdetector.

2.3.3 The magnet

The CMS experiment aims to achieve a good momentum resolution for particle momenta up to 1 TeV. The larger the bending power the better the momentum resolution, so in order to meet this goal a large solenoid with an high axial magnetic field is needed.
The superconducting CMS magnet is designed to provide a magnetic field of 4 T, but during the Run 1 of the LHC it has been decided to operate at 3.8 T. It features a free bore, with a diameter of 6 m and a length of 13 m, and a 10000 tonnes steel yoke, where the flux is returned, made of 5 3-layered wheels in the barrel and 3 disks in the endcaps. The yoke plays also the role of absorber in front of the muon stations. The axial direction of the field allows to start measuring the momentum at $r = 0$ (unlike a toroidal configuration), resulting in a more compact design of the whole spectrometer. In table 2.1 the main parameters of the CMS magnet are listed.

Table 2.1: The main parameters of the CMS magnet.

Magnetic field	4 T
Inner bore	5.9 m
Length	12.9 m
Number of turns	2168
Current	19.5 kA
Stored energy	2.7 GJ
Total weight	$\sim$ 12000 tons

The CMS experiment solenoid employs high purity aluminium conductors, with an overall cross section of the cables of 64 $\times 22\,\mathrm{mm}^2$. To keep the necessary low temperatures (about 4 K) an indirect termosyphon cooling is used together with epoxy resin impregnation.

2.3.4 The muon system

As the name of the experiment suggests, the system dedicated to the measurement of the muons is one of the most important features of the CMS detector and has heavily influenced its design. The muon detector has three important goals: muon momentum measurement, muon identification and muon triggering.
The measurement of muon momentum is done in cooperation with the inner tracker: for moderate-low transverse momenta the tracker measurement dominates, while for momenta above about 200 GeV the muon system starts becoming more and more important and its spacial resolution eventually dominates.
The muon identification is ensured by the absorbing material in front of the muon chambers, which exceeds 16 interaction lengths and makes the hadron punchthrough component negligible for the momentum range of interest for most of the physics analysis.
The muon triggering is achieved with the use of fast and segmented gas detector.
Due to the large number of detection planes needed, the muon chambers had to be inexpensive, but at the same time robust and reliable, in order to accomplish the tasks mentioned above. For this reason, gaseous detectors of three different technologies (DT, CSC and RPC) are employed.
The muon system, represented in figure 2.9, is divide in barrel ($|\eta| < 1.2$) and endcap ($1.0 < |\eta| < 2.4$) detector.

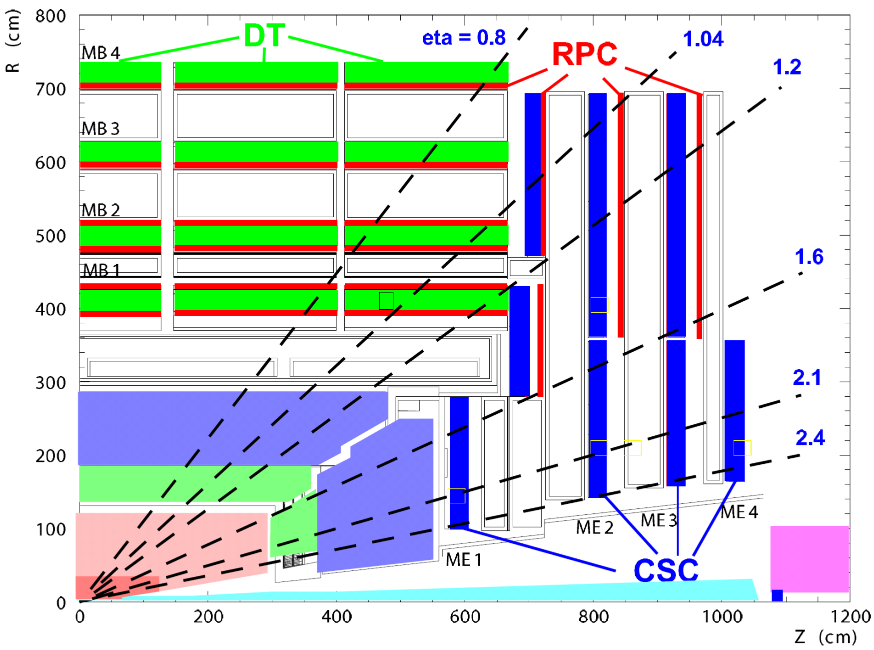


Figure 2.9: The layout of the muon system.

Barrel detector

The barrel detector consists of 250 chambers, placed in 4 concentric stations (MB1, MB2, MB3 and MB4 starting from the innermost one) at $r = 4.0$ m, 4.9 m, 5.9 m, 7.0 m respectively. Each station is divided in 5 wheels along the z-axis and 12 sectors in the azimuthal angle. MB1 and MB2 are made by "sandwiching" one DT chamber in between two RPCs, while in the two outermost stations, MB3 and MB4, each DT is coupled with one RPC.
In MB1, MB2 and MB3, each DT chamber is made of 12 layers of drift tubes, grouped in 3 Super Layers (SL) of 4 layers each: 2 SLs measure the $r - \phi$ coordinate while the other one measures the z coordinate. The first SL is separated from the other to have a larger arm for the measurement in the bending plane. In MB4, DT chambers consist of only 2 SLs measuring $r-\phi$ coordinate. A total of 480 RPCs are used in the barrel detector. Each RPC is a double-gap bakelite chamber operating in avalanche mode, with gap width of 2 mm. The strips, running along the beam direction, are segmented in order to match trigger requirements.

Endcap detector

Each endcap detector is composed of 4 stations labeled ME1 to ME4 starting from the station closest to the interaction point. They are mounted in disks perpendicular to the beam direction, each disk being divided into 2 or 3 concentric rings. There are 468 CSCs in total, each arranged in a trapezoidal shape and made of 6 gas gaps (7 layers) with planes of cathode strips in the radial directions and anode wires almost perpendicular to the strips. Most CSCs are overlapped in ϕ in order to avoid gaps in acceptance. There are 36 chambers for every ring, except the innermost ring of ME 2,3 and 4 which has 18 chambers. The ionization of a charge particle passing through the planes cause charge to form on the anode wire and image charge on the cathode strips, thus allowing to get (r, z, ϕ) hits in each layer. Three out of four chambers have a RPC plane in the endcaps (see figure 2.9).

2.3.5 The electromagnetic calorimeter

The measurement of electromagnetic objects, namely electrons and photons, is of fundamental importance for the physics program at the LHC. In particular, the channel in which the Higgs decays to a pair of photons was chosen as the main benchmark for the design of the electromagnetic calorimeter in CMS.
The need for a good resolution, in a demanding environment like the LHC, lead to the choice of an homogeneous calorimeter made of lead tungstate $\mathrm{PbWO_4}$ crystals. The $\mathrm{PbWO_4}$, infact, provides a fast response: about 80% of the light output is collected in 25 ns, which is the nominal bunch spacing at the LHC. The high density of the crystal results in a short radiation length ($X_0 = 0.9\,\mathrm{cm}$) and a small Moliere radius (2.2 cm), which allows for a fine granularity and compact layout. The light yield is, in turn, relatively low: at the operating temperature (18°C) about 4.5 photons per MeV of deposited energy are collected by the photodetectors. To exploit the internal reflection for an optimal light collection, the surfaces of the crystal are polished. The effect of the radiation damage, which produces a wavelength-dependent loss in the light transmission, is corrected for with the use of injected laser light.

Both the number of scintillating photons emitted and the amplification in the APDs decrease with increasing temperature: for this reason the temperature is kept stable within ± 0.05 °C thanks to a water flow cooling system which dissipates the heat due to the electronic readout.

The CMS ECAL, which is divided in barrel ECAL ($|\eta| < 1.479$) and endcap ECAL ($1.479 < |\eta| < 3.0$), is shown in a schematic view in figure 2.10.

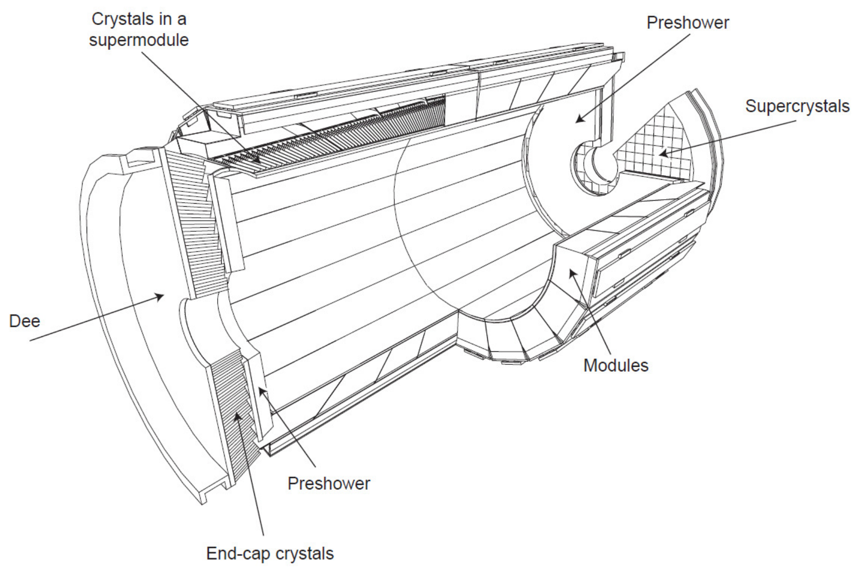


Figure 2.10: The layout of the electromagnetic calorimeter.

Barrel electromagnetic calorimeter

The ECAL barrel (EB) consists of 61200 crystals, with a 360-fold segmentation in ϕ and a granularity of 2×85 in η. They are arranged in a quasi-projective configuration with respect to the interaction point. They have a tapered shape, slightly varying with the pseudorapidity, with a cross-section of $22 \times 22\,\mathrm{mm}^2$ at the front face of the crystal and $26 \times 26\,\mathrm{mm}^2$ at the rear, corresponding to approximately 0.0174×0.0174 in $\eta - \phi$. The legnth is 230 mm, which is equivalent to 25.8 X_0.
The crystals are contained in an alveolar structure called submodule, with thin walls of aluminum (front) and glass fibre-epoxy resin (sides) separating each crystal by its neighbours by 0.35 mm. Submodules are organized into modules, each containing from 400-500 crystals depending on η. Four modules, separated by aluminium webs, form a supermodule, which contains 1700 crystals in total. There are 18 supermodules, each one covering 20°in ϕ.
In the barrel ECAL, the scintillating light is detected by APDs.

Endcap electromagnetic calorimeter

The ECAL endcap (EE) is made of identically shaped crystals grouped in 5×5 units called supercrystals (SC), by means of an alveolar structure similar to the one used in

the barrel. Each endcap is divided in two halves called "dees", each one containing 138 standard SCs plus 18 special SCs for the innermost and outermost circumferences, arranged in a grid in the $x-y$ plane. Crystals point towards the interaction point, with a small tilt of 2-8 degrees with respect to the exact projection.
The cross section of the crystals in the endcaps is $28.62 \times 28.62\,\mathrm{mm}^2$ in the front and $30 \times 30\,\mathrm{mm}^2$ in the rear, and their legnth is 220 mm (24.7 X_0).
The light output is readout by VPTs.

On-detector electronics

The ECAL read-out system is designed to read the signal of a trigger tower (5×5 crystals in $\eta \times \phi$) or a super-crystal in the EB and EE respectively.
The signal coming from the APDs and VPTs is amplified and shaped by a Multi-Gain Pre-Amplifier (MGPA). The MGPA is an ASIC developed in 0.25 μm technology and consists of three amplifiers, with nominal gains of 1, 6 and 12, a CR-RC shaper, three digital-to-analog converters and a test-pulse generator. The full scale signals correspond to about 1.5 TeV and 1.6-3.1 TeV for the APDs (EB) and VPTs (EE) respectively. The non-linearity of the output pulse is less than 1% in all the dynamic range. The noise for the highest gain is about 8000 and 4000 electrons for the APD configuration and the VPT configuration respectively.
The three analog signals are digitized in parallel by a 40-MHz, 12-bit ADC, developed in 0.25 μm technology. The highest non-saturated signal is selected, and the information is reported consisting of the 12-bit ADC count plus two bits coding the ADC number.
The digitized data are stored, during the L1 trigger latency, in 256-word-deep pipelines. Five of these pipelines and the logic to calculate (once every bunch crossing) the sum of the 5 channels corresponding to the trigger unit are implemented in a 0.25 μm ASIC: the energy is thus summed in strips of 5 crystals along ϕ.
The trigger tower energy sum is transmitted using Gigabit Optical Hybrids (GOH) to the Trigger Concentration Card (TCC) which completes the generation of "trigger primitives".
On receipt of a L1 trigger, the corresponding data are transmitted to the Data Concentrator Card (DCC), which is responsible for the zero suppression and the data reduction. The Clock and Control System (CCS) boards distributes the system clock, the trigger and configure the front end electronics.

2.3.6 The hadron calorimeter

The hadron calorimeter completes the calorimetry system of CMS with the measurement of the energy of hadron showers. The HCAL is required to have a good hermeticity for a precise determination of the missing transverse energy. The unbalance of the momentum in the transverse plane is infact a sign of neutral, stable particles which escaped the detection and is the typical signature of many important processes in the SM and in new physics, like SUSY or other dark matter models.
The CMS HCAL is a sampling calorimeter with brass (70% copper and 30% zinc) as a passive material and scintillator tiles as active material, coupled with wavelength shifting (WLS) fibres, which carry the light to the photodetectors. It is placed right

around the ECAL, and its design is strongly influenced by the presence of the solenoid, which restricts the amount of material budget due to the limited space in the coil. For this reason, in the barrel region an additional outer calorimeter (HO) is placed outside the solenoid.

A sketch of the hadron calorimeter is shown in figure 2.11.

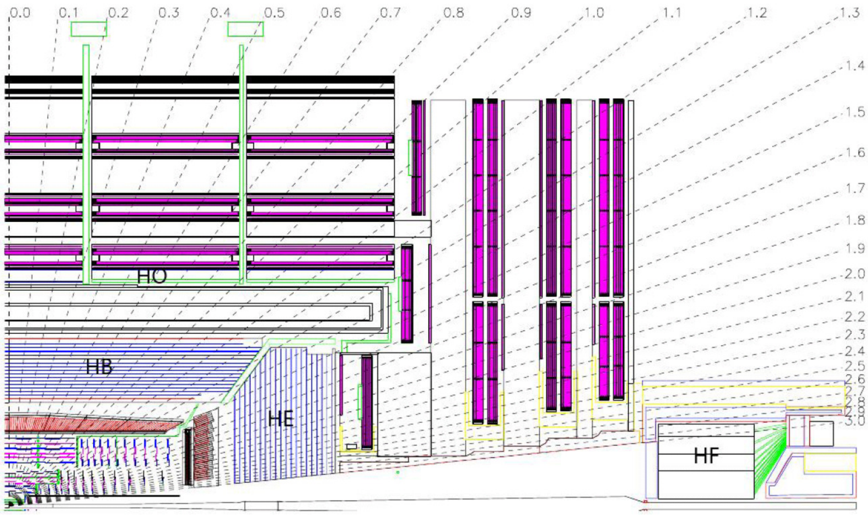


Figure 2.11: The layout of the hadron calorimeter.

It is divided four subsystems: the hadron barrel (HB), the hadron outer (HO), the hadron endcap (HE) and the hadron forward (HF) calorimeter.

Barrel hadron calorimeter

The HB inner radius is 177.7 cm, the outer one is 287.65 cm, and covers the region $|\eta| < 1.3$. It is made of two half barrels, each one divided in 18 20°-ϕ wedges. Each wedge is made of 17 layers of active scintillator (3.7 mm thick) interspersed with brass plates (about 50-60 mm depending on the radial position), while the innermost and the outermost layers are made of stainless steel, which ensures structural strength. The first active layer is placed directly behind the ECAL, in order to deal with low showering particles coming from the material between ECAL and HCAL. Each tile has a size of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ and is coupled with one WLS fibre. The overall HB consists then of 32 "towers" in η, readout by pixelated hybrid photodiodes (HPDs).

Outer hadron calorimeter

The HO is placed outside the HB, inside the barrel muon system, having the same pseudorapidity acceptance of the HB. It is divided in 5 sections, or "rings", along η, each one covering 2.5 m length in z. The ring 0 has 2 scintillator layers on both side of an iron absorber, while the other rings have a single layer. The scintillator tiles are 10 mm thick and match the ϕ segmentation of DT chambers. The tower geometry of the scintillator in η and ϕ is the same of the HB. The main motivation for the

HO is to measure the energy of penetrating hadron showers, in order to get the tails of the jets. Increasing the material budget up to about 11 interaction lengths in the mid-rapidity region improves also the resolution in $E_{\mathrm{T}}^{\mathrm{miss}}$.

Endcap hadron calorimeter

The HE detectors cover the region $1.3 < |\eta| < 3$ and each endcap is attached to the muon endcap yoke. The ϕ-design of HE matches the HB, even though for $|\eta| > 1.74$ the ϕ granularity is halved to host the bending radius of the WLS fibres. The $\eta - \phi$ size of the towers is the same as for the HB for $|\eta| < 1.6$ and $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ The thickness of the absorber plates is 79 mm, with 9 mm gaps to accomodate the scintillators. In total, the thickness of the HE is about 10 interaction lengths.

Forward hadron calorimeter

The forward hadron calorimeter HF is located at 11.2 m from the interaction point. The absorber layers are made of steel, while the active one is radiation hard quartz in fibres with varying length embedded to the steel, which provide fast Cherenkov light yield. The distance between 2 adjacent fibres is 5 mm and they are readout separately by phototubes in the rear of the detector. Each module of HF is made of 18 wedges, which are placed in a non-projective way, with fibres running along beam direction.

2.3.7 The inner tracker system

The inner tracker detector of CMS is designed to achieve a precise and efficient reconstruction of the charged particle trajectories coming out of the collisions. The measurement of the tracker is essential, in particular, to estimate the momentum of muons and electrons, identify tau leptons and reconstruct jets and secondary vertices. Tracking information is also used in the high-level trigger system of CMS.
The LHC presents several challeges for the operation of the tracking system, due to the huge flux of particles and the large number of PU interactions. In order to accomplish the tasks mentioned above, the tracker should have:

- **high granularity**, to ensure track reconstruction in a crowded environment;
- **fast response**, to assign the trajectories to the correct bunch crossing;
- **radiation hardness**, to survive about ten years of data taking;
- **low material budget**, to avoid gamma conversions and large contribution to the momentum resolution due to multiple scattering.

All these requirements lead to the choice to build the tracker using a silicon detector technology. The CMS tracker is so far the largest silicon detector ever built.

In order to keep a low occupancy ($\lesssim 1\%$) in the presence of a high charged particle flux and to reach the desired resolution on the impact parameter, pixelated tracker layers have to be used for radii below 10 cm. In the intermediate region

($20\,\mathrm{cm} < r < 55\,\mathrm{cm}$), the reduced rate density allows to employ micro-strip detectors. In the outer region ($55\,\mathrm{cm} < r < 110\,\mathrm{cm}$) the strip pitch is further increased together with the strip length, in order to reduce the number of read-out channels. In order to keep a signal-to-noise ratio of approximately 10:1, thicker silicon sensors are used in this region. The acceptance of the tracker extends up to $|\eta| < 2.5$.

The radiation damage causes an increase of the leakage current of the sensors, which turns to have an exponential dependence on the temperature. For this reason, the CMS tracker has operated at a temperature of about +4°C during Run 1. The temperature will be further decreased during the future high luminosity runs of the LHC, down to about -20°C. The cooling is ensured by cooling circuits which employ C_6F_{14} in the liquid phase.

The read-out chips of the CMS tracker makes use of a standard 250 nm CMOS technology, slightly modified to make the chip radiation-hard. This choice ensures a good radiation hardness of the read-out electronics: the lifetime of the strip tracker subsystem is then limited by the aging of the silicon sensors.

A schematic view of the CMS tracker is in figure 2.12. More details about the layout of the pixel detector and the the strip detector are given in the following sections.

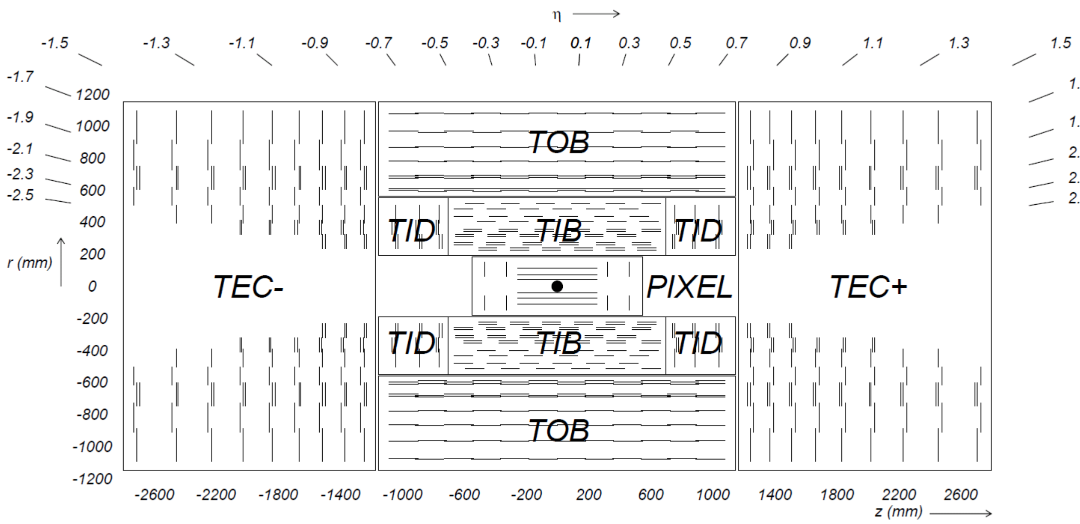


Figure 2.12: The layout of the inner tracker system.

The pixel tracker

In the barrel, three cylindrical layers of pixel detector modules (BPix) are placed surrounding the interaction point. The pixel detector is complemented with two disks for each endcap (FPix). The BPix layers are 53 cm-long and are placed at $r = 4.4, 7.3, 10.2\,\mathrm{cm}$. The FPix are located at radii between 6-15 cm and are placed at $z = \pm 34.5, \pm 46.5\,\mathrm{cm}$. In total the pixel detector consists of 66 million pixels, for

a total area of 1.06 m^2.
The layout of the pixel detectors is sketched in figure 2.13.

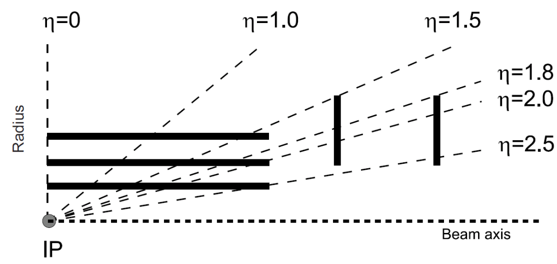


Figure 2.13: The layout of the pixel detector.

The pixel cell size is the same for both subsystems and is $100 \times 150\,\mu m^2$, which is a choice made to achieve a similar resolution in the $r - \phi$ and in the z directions. The pixel detector uses zero-supressed, anolog pulse height read-out: this feature allows to get a very good hit spatial resolution (15-20 $\mu\mathrm{m}$) due to the sharing of the charge between different pixels.
The mechanics and the cabling in the pixel detector has been designed to allow a yearly access.

The pixel read-out consists of three parts:

- a read-out data link from the modules to the front end driver (pxFED), which receives the data, digitizes it, formats it and sends it to the Data Acquisistion System (DAQ);
- a fast control link from the pixel front end controller (pFEC) to the modules. The pFEC sends the clock, fast control signals (trigger) and programs the front end devices;
- a slow control link from a standard FEC to the supply tube/service cylinder. The FEC configures the ASICs on the supply tube/service cylinder.

The FEC, pFEC and pxFED are VME modules located in the electronics room and connected to the front-end with optical fibres.
The read-out chip (ROC) is a full custom ASIC using a 250 nm CMOS technology and contains 52x80 pixels. It amplifies and buffers the charge signal, performs zero-suppression, verifies the Level 1 (L1) trigger, adjusts the voltage levels, currents and offsets and sends the hit information to a Token Bit Manager (TBM). The TBM controls a group of from 8 to 24 ROCs depending on the position in the detector.

The strip tracker

The strip tracker detector is divided in four different subsystems.
The tracker inner barrel (TIB), made of 4 layers, and the tracker inner disks (TID), made of 3 disks, extend up to a radius of 55 cm, providing up to 4 hit measurements in the $r - \phi$ plane. The tracker outer barrel (TOB) has an outer radius of 116 cm and consists of 6 layers, each one carrying another $r - \phi$ hit. The tracker endcaps (TEC)

cover the region between $124\,\mathrm{cm} < |z| < 282\,\mathrm{cm}$ and $22.5\,\mathrm{cm} < |r| < 113.5\,\mathrm{cm}$ and is composed of 9 disks, each made of up to 7 rings (depending on the $|z|$ position). The TEC detectors thus measure up to 9 ϕ hits of the trajectories.
The strips are parallel to the beam axis in the TIB and TOB and along the radial direction for the TID and TEC. Both TIB and TID use 320 μm-thick silicon micro-strip sensors. The strip pitch in the TIB is 80 (120) $\mu\mathrm{m}$ in the layers 1 and 2 (3 and 4), with a spatial resolution of about 23 (35) $\mu\mathrm{m}$. In the TID the pitch varies between 100-141 $\mu\mathrm{m}$.
The TOB is made of 500 $\mu\mathrm{m}$-thick sensors, with the strip pitch of 184 $\mu\mathrm{m}$ in the first 4 layers and 122 $\mu\mathrm{m}$ in the layers 5 and 6. The resulting single point resolution is about 53 $\mu\mathrm{m}$ and 35 $\mu\mathrm{m}$ respectively.
The four inner rings of the TEC have 320 $\mu\mathrm{m}$-thick sensors with strips of average pitch 97 $\mu\mathrm{m}$, while the outer 3 are 500 $\mu\mathrm{m}$-thick and have a 184 $\mu\mathrm{m}$ average pitch.
In the first two layers of the TIB and TOB, as well as in the two inner rings of the TID and rings 1,2 and 5 of the TECs, a second micro-strip module is mounted back-to-back, providing a measurement of the z coordinate in the barrel and the r coordinate in the disks. The resolution in this direction is about 230 $\mu\mathrm{m}$ in the TIB, and 530 $\mu\mathrm{m}$ in the TOB, while it varies with the pitch for the TID and TEC.
The silicon strip detector of CMS is expected to provide, in the region $|\eta| < 2.4$, about 9 hits per trajectory, with 4 of them being 2-dimensional measurements.
It consists of more than 9 million strips, covering an area of about 198 m^2.

The signals from the silicon strips detector are amplified, shaped and stored by a custom ASIC, the APV25. It is fabricated in a 250 nm CMOS process which is specifically designed to ensure radiation tolerance. The APV25 consists of a pre-amplifier, a shaper and a 192 element deep pipeline, which can store the data for a trigger latency of 4 μs.
After a positive L1 trigger decision, the analogue signals are trasmitted via optical fibres to the front end driver (FED), located in the service cavern. Here, an accurate pedestal subtraction takes place. All the other fast informations, such as the clock and the trigger are also transmitted by optical links.
This analogue read-out scheme ensures optimal spatial resolution, thanks to charge sharing, operational robustness and reduced material budget, since the analogue to digital conversion happens out of the tracker volume.

2.3.8 The trigger and the data acquisition

The LHC is expected to deliver proton-proton collisions with a bunch crossing of 25 ns, corresponding to a crossing frequency of 40 MHz. At the design peak luminosity ($10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$) each bunch crossing contains approximately 20 interactions. One of the main challenges of the LHC experiments was to design a trigger system able to reduce the rate of events to approximately 100 Hz, which is the maximum sustainable rate at which event information can be processed offline.
The CMS trigger achieves this goal with a combined system consisting of two steps: the level-1 (L1) trigger, which makes use of custom-designed electronics, and the high-level trigger (HLT), which is instead software-based.

The L1 trigger has an output rate of about 30 kHz, thus implementing a rejection factor of approximately 10^3. It uses only coarse informations from the calorimeters and the muon system to decide whether to discard or not the event, while the whole information is hold in pipelined memories in the front-end electronics. Since the latency of the L1 trigger is about 3.8 μs, such pipelines should be deep enough to store the data corresponding to approximately 150 bunch crossings.
The architecture of the L1 trigger is illustrated in figure 2.14.

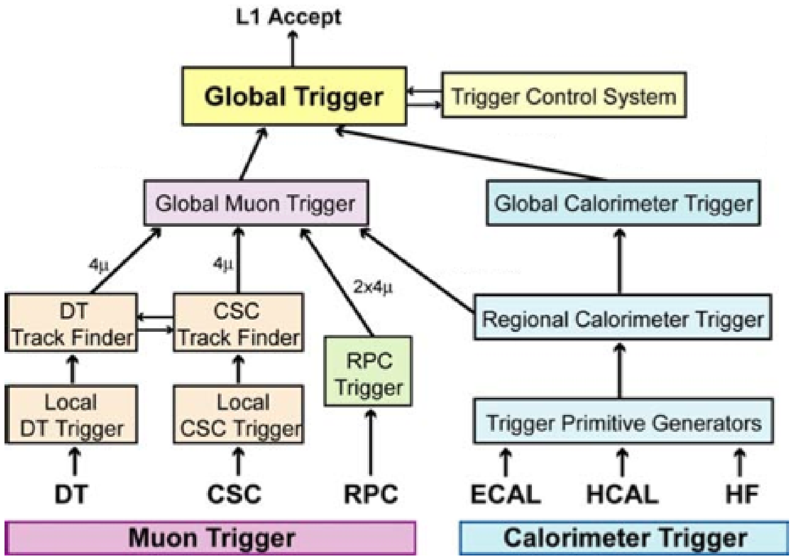


Figure 2.14: The architecture of the L1 trigger.

The L1 trigger is organized in a hierarchical structure, with local, regional and global components. The local triggers, or trigger primitives, are energy deposits in the calorimeters or track segments in the muon chambers. The regional triggers combined the information in each region of the detector to provide trigger objects sorted by momentum or quality. The global calorimeter and muon triggers determine the highest rank objects in the whole detector for each type and transfer the information to the global trigger: here the decision on whether to keep the event is taken based on the outcome of the algorithms. The decision is then communicated through the Timing, Trigger and Control (TTC) system.
The L1 trigger electronics is partly placed on the detectors themselves, partly in a control room which is 90 m distant from the experimental cavern.

The HLT trigger consists of a PC farm with about a thousands of commercial processors running a software similar to the one used in the offline analysis.
After the event has passed the L1 trigger selection, data are trasferred from the pipeline memories to the front-end read-out buffers. Here further signal processing, zero suppression and data compression are performed. The total size of an event at this point is 1.5 MB for p-p collisions. Each event, contained in several hundreds different buffers, is transferred to a processor, running a HLT software.
During Run 1 the nominal HLT rate was about 350 Hz: these events have been processed immediately and are reffered to as “streamed” data. An additional 300 Hz of bunch-crossings have been recorded and processed later, during the long shutdown in

2013-2014: these data are reffered to as “parked” data.

Part II

Research Work

Chapter 3

Study of the muon momentum scale with the CMS detector

Abstract

If high precision is needed in the measurement of final states with muons, then a precise calibration of the muon's momentum is mandatory, since many detector effects can spoil the measurement of charged tracks. A method to recover the correct scale of muons, based on a likelihood fit to the dimuon mass spectrum in the region of resonances, has been developed in CMS and used with the data of the Run 1 of the LHC.

In this chapter we review the method, the calibration strategy and the results obtained with 7 TeV and 8 TeV data of the LHC.

3.1 Introduction

Since the CMS detector is expected to provide precise measurements in many channels, a careful knowledge of the response of each subdetector is mandatory. In particular, on top of the "hardware" calibration at the level of the sensor and the electronics of the detector subsystem, a campaign of *in situ* calibration of the physics objects used in the analysis, like muons, electrons, photons and jets, has started as soon as the very first collisions data appeared in 2010.
Among the measurements that CMS is able to provide, the trajectory of charged particles measured in the silicon tracker is one of the most sensitive to the conditions of the detector, in particular to the assembly operations and the positioning of its sensors. The determination of the transverse momentum p_{T} is highly sensitive to the alignment of the tracker and the muon chambers, the amount of material along the trajectory and the knowledge of the magnetic field.
The momentum scale of tracks can be fruitfully studied thanks to the precise knowledge of the mass of resonances decaying to muons. In particular, for muon momenta from few GeVs up to ten hundreds of GeVs the use of resonances like the Z the J/Ψ or the Υs can be very helpful.

This chapter is organized as follows. In the first section 3.2 the algorithm used for the calibration of muons' momenta is presented. Then in section 3.3 the strategy used for the calibration with the 7 TeV and 8 TeV data is described and in the following section 3.4 the validation of the results of the calibration constants is presented. In section 3.5 the study of the systematic uncertainty on the momentum scale of the muons is illustrated. Finally in section 3.6 an overview is given of the impact of muon momentum scale corrections and systematics in the physics analyses in CMS.

3.2 The MuScleFit algorithm

The information about the muon momentum scale can be extracted from a likelihood fit to dimuon resonances, combining the parameters of the measured tracks in dimuon events and the knowledge of the mother particle. The *per-event* observable that provides such information is, of course, the mass of the muons' mother. Given that it is not a *per-track* variable, a probabilistic approach is needed to correlate a possible bias (offset from the true value) in the track momentum with a shift in the observed dimuon mass. Another complication is the finite mass resolution: the information on the scale can be inferred from the event only if an estimation of the uncertainty on the measured mass is provided. Anyway, the latter can be extrapolated from the single muon's resolution, which can be extracted itself from the fit, as it will be discussed later. The finite intrinsic width of the resonance in the case of the Z must be also taken into account.

Such a multi-parameter likelihood approach has been implemented in the *MuScleFit* algorithm, as a part of CMSSW, the official code used by the CMS collaboration for event reconstruction and physics analysis. A complete description of the MuScleFit algorithm can be found in this note [24]. Here we give an outline of the ingredients that enter the likelihood function, focusing on the ones which have been studied in more details during the calibration with Run 1 data.
The inputs needed to build the likelihood are essentially:

- The lineshape of the resonance (signal model)
- The background model
- The ansatz function for the momentum scale
- The ansatz function for the momentum resolution

The lineshape is provided as a two-dimensional histogram, which stores the probability density that tells how likely a reconstructed mass m is, given a mass resolution s. Such a probability, in general, reads as a convolution of a theoretical cross section $\sigma(m, m_0)$ (i.e. a *Breit-Wigner* distribution or the result of a more complicated computation) and a gaussian term:

$$P_{\mathrm{sig}}(m_{obs}, s, m_0) = \int dm \frac{\sigma(m, m_0)}{s} e^{-\dfrac{(m - m_{obs})^2}{2s^2}}$$

The observed reconstructed mass and its resolution are then written as a function of the kinematic variables of the muons which are taken, for simplicity, in the basis $(p_{\mathrm{T}}, \phi, \cot\theta)$.
The scale is parametrized as

$$p_{\mathrm{T}}^{\mathrm{corr}} = F(\vec{x}, \vec{\alpha}) p_{\mathrm{T}}$$

With $\vec{x}$ are denoted all the observables on which the scale may depend on, such as the track parameters. The ansatz function depends on the parameters $\vec{\alpha}$ which enters into the likelihood through the dependence of the reconstructed mass and are determined by the maximum likelihood fit.

$$m \equiv m(p_{\mathrm{T},1}^{\mathrm{corr}}, \phi_1, \cot\theta_1; p_{\mathrm{T},2}^{\mathrm{corr}}, \phi_2, \cot\theta_2)$$

Since the direction (θ, ϕ) is known with much better precision with respect to the p_{T}, the contribution of ϕ and $\cot\theta$ to the mass resolution are neglected and we can write:

$$s \simeq \sqrt{\left(\frac{\partial m}{\partial p_{\mathrm{T},1}}\right)^2 \sigma_{p_{\mathrm{T},1}}^2 + \left(\frac{\partial m}{\partial p_{\mathrm{T},2}}\right)^2 \sigma_{p_{\mathrm{T},2}}^2}$$

The resolution ansatz function introduces another set of parameters $\vec{\beta}$, to be eventually determined in the fit:

$$\sigma_{p_{\mathrm{T}}} = G(\vec{x}, \vec{\beta})$$

Finally, a probability density function $P_{\mathrm{bkg}}(m)$ is needed to describe the non-resonant background under the peak. The procedure to extract it, together with the fraction of the signal f_{sig}, is described in details in section 3.3.1.
The likelihood can be written, for a single resonance as:

$$ln\mathcal{L} = \sum_{i=0}^{N} ln\left[f_{\mathrm{sig}} P_{\mathrm{sig}}(m_i, s_i) + (1 - f_{\mathrm{sig}}) P_{\mathrm{bkg}}(m_i)\right]$$

In the equation above the sum is extended to all the N events in the dataset, and the the quantities m_i and s_i are understood to be computed with the measured characteristics of the event i and to be functions of $\vec{\alpha}, \vec{\beta}$. The extension to multiple resonances is straightforward.

In general, a maximization of $ln\mathcal{L}$ allows to determine scale, resolution and background model simultaneously. Unfortunately, in practice, the fit involves a large number of parameters, and, in order to have stable results a strategy has to be adopted to estimate the parameters in a suitable order in the fit.

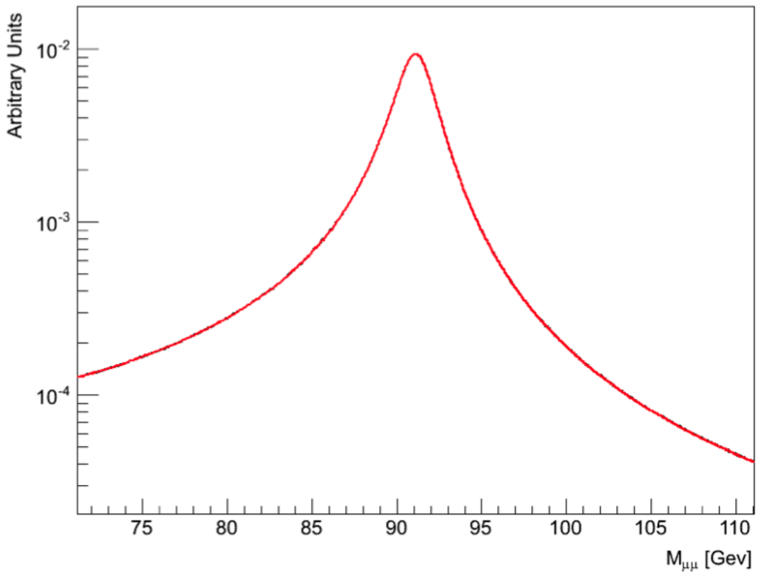
3.3 Calibration of muon momentum scale and resolution

Preliminary MC studies performed to validate the implementation of the MuScleFit algorithm are described in [24]. A preliminary calibration campaign was also carried on with the first $Z \to \mu\mu$ events in 2011 [25].
Here we describe in details the results of the extensive calibration that has been performed with the whole 7 and 8 TeV datasets collected by the CMS experiment during the Run 1 of the LHC and the corresponding Monte Carlo (MC) simulated samples.
Corrections function are derived using $Z \to \mu\mu$ decays, but J/ψ and Υ are also used for their validation.

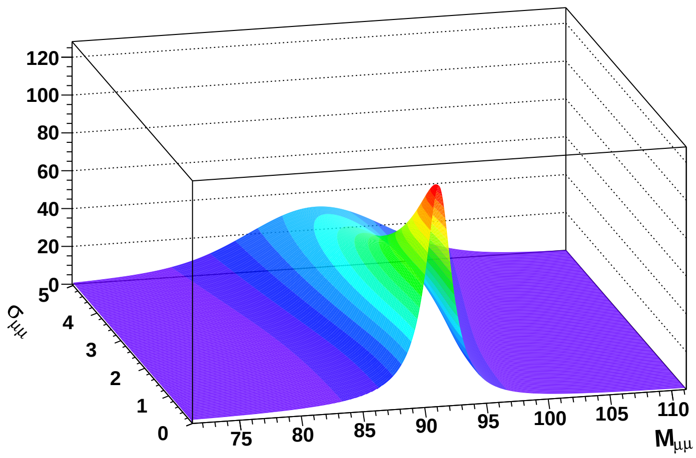
3.3.1 Signal and background model

For the signal model, i.e. the Z boson lineshape, a calculation at NNLO in $\alpha_{\rm s}$ [26] has been used. This approach has the advantage, with respect to the Breit-Wigner distribution, that describes correctly the radiative tail at low virtualities. While there is practically no difference between 7 and 8 TeV calculation in terms of shape, a mild dependence has been observed on the acceptance cuts applied on the muons, therefore they have been chosen to be close to the actual experimental selection cuts, i.e. $p_{\mathrm{T},\mu} > 18\,\mathrm{GeV},\ |\eta_\mu| < 2.4$. A numerical approach has been used then to convolute such a distribution with a gaussian to account for detector resolution.
In figure 3.1a the invariant mass spectrum corresponding to the theoretical calculation is shown, while on figure 3.1b the 2D distribution coming out from the convolution is presented.

An exponential function has been chosen to model the non resonant background under the Z peak, mainly due to $t\bar{t}$ fully leptonic events and $Z \to \tau\tau$ events surviving the $Z \to \mu\mu$ selection. The slope of the background function, as well as the fraction of background events are determined in a preliminary step, in order to spoil possible unphysical correlations with the parameters of the ansatz functions for the scale and the resolution.
The background model is determined in a grid of 4×4 pseudorapidity bins for each muon, with the bins defined as [-2.4,-0.8,0,0.8,2.4], in order to take into account topological difference between signal and background events. Examples of fits to extract background parameters are in figure 3.2. The signal model in those fits, shown in red dashed line using post-fit values of the parameters, is a Breit-Wigner function, convoluted with a double-sided Crystal Ball pdf to model the detector effects.



(a) NNLO calculation of the Z lineshape



(b) Convolution of the Z lineshape with the experimental resolution

Figure 3.1: Figure 3.1a shows the NNLO calculation of the Z lineshape, for 7 TeV (red) and 8 TeV (black), which are essentially indistinguable. Figure 3.1b shows the 2D probability distribution as a function of the mass ($x-\text{axis}$) and the resolution ($y-\text{axis}$) as a result of the convolution between the theoretical lineshape and the detector resolution.

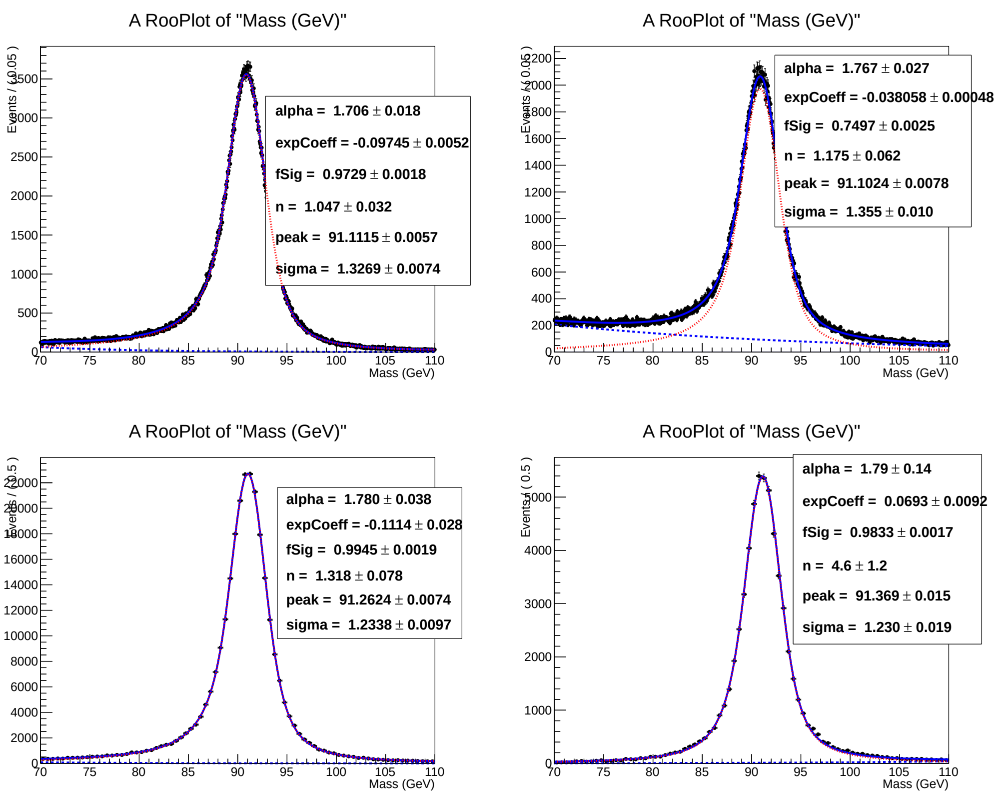


Figure 3.2: Examples of fits to extract background parameters, for a center-of-mass-energy of $\sqrt{s} = 8\,\mathrm{TeV}$ in a data sample (top row) and in the simulation (bottom row). Left-hand plots refer to events with $0 < \eta_{\mu^-} < 0.8$ and $0.8 < \eta_{\mu^+} < 2.4$ while for the right-hand plots $-2.4 < \eta_{\mu^-} < -0.8$ and $0.8 < \eta_{\mu^+} < 2.4$.

3.3.2 Fit strategy

Scale ansatz function

Since at low-moderate transverse momentum the muon chambers are not expected to contribute much to the p_{T} measurement, the track is taken as the one reconstructed in the silicon tracker (*tracker track*).
In the data collected in 2011-2012, the bias in the assignment of the momentum is mainly related to geometrical effects, e.g. deformations of the tracker geometry used in the reconstruction still present after the alignment procedure. For this reason, scale corrections have been implemented as functions of the curvature of the track, defined as:

$$k = \frac{q}{p_{\mathrm{T}}}$$

We modeled the corrections with an analytical function defined in five bins of the muon pseudorapidity:

$$k_{\mathrm{corr}} = (1 + p_0)\left(k_{\mathrm{raw}} - \frac{\delta}{2} - C_j\left(\varphi, \eta\right)\right) \tag{3.1}$$

where j is an index running on the η bin: [-2.4,-2.1], [-2.1,-1.5], [-1.5,1.5], [1.5,2.1] and [2.1,2.4].
The terms in equation (3.1) are:

- $k_{\mathrm{raw(corr)}}$ is the curvature before (after) the correction
- p_0 is a *global scale* term accounting for effects like the inaccurate knowledge of the magnetic field
- δ represents an absolute bias in the curvature, e.g. a bias on the transverse momentum of the track different for negative and positive muons
- $C_j\left(\varphi, \eta\right)$ accounts for residual misalignment effects in each of the five η bins. The functional form:

  $$C_j\left(\varphi, \eta\right) = a_{1,j}\sin(\varphi + \varphi_{1,j}) + a_{2,j}\sin(2\varphi + \varphi_{2,j}) + b_j(\eta - \eta_{0,j}) + b_{0,j} \tag{3.2}$$

  was choosen to model the weak modes more frequently found in the post-alignment geometry, namely the *sagitta* (described by $a_{1,j}$), the *twist* (described by b_j) and the elliptical deformations (described by $a_{2,j}$) [1].

Resolution ansatz function

The resolution on the p_{T} was modeled as a sum in quadrature of two terms:

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = q_0 p_{\mathrm{T}} \oplus q_j \tag{3.3}$$

[1] The three effects described, often referred to as "weak modes", correspond to the following parametric deformations: $r\Delta\varphi = c_s\cos\varphi$ (sagitta), $\Delta r = r(1 - c_e\cos2\varphi)$ (elliptical) and $\Delta\varphi = c_t z$ (twist).

where the parameters q_j, representing the contribution of the multiple Coulomb scattering to the resolution, have been computed in 12 equally spaced η bins.

Likelihood minimization

The MuScleFit algorithm allows the user to run several iterations, giving for each iteration the possibility to decide which set of parameters to fit and which instead to keep fixed. Taking advantage of this feature, a multi-step strategy has been developed, mainly to cope with the large number of parameters which sum up from the parameterization of scale and resolution ansatz functions. The steps are summarized in the following.

- **Step 1**: in the very first iteration only the resolution parameters are determined, while the scale remains untouched. This preliminary fit is performed to retrieve the information of the p_{T} resolution before any correction is applied to the scale, which is crucial to have a reliable estimation of the mass resolution in the likelihood used in the following iterations
- **Step 2**: in the second step the correction to the scale is determined using the resolution extracted previously to compute the likelihood
- **Step 3**: in the third step the scale parameters are re-fitted (2nd iteration) with starting values being the ones from the previous iteration. If the scale fit is converging, the correction to the scale extracted in this iteration should be very small, possibly compatible with the identity
- **Step 4**: in this step the "global scale" term described above is fitted, while all the other parameters in the likelihood are kept fixed to the values determined previously
- **Step 5**: in this very last stage the resolution is determined again to spot how it changed after the scale calibration (usually the p_{T} resolution is improved after the scale fit).

3.3.3 Datasets and selection

Since the momentum scale is heavily affected by the tracker geometry used in the reconstruction, a new calibration is needed everytime a new alignment of the tracker becomes available. The difference in the measurement of the transverse momentum can be very large using different tracker alignment conditions, as it will be highlighted later.
In CMS the MC generated event samples are obtained from a full simulation of the detector based on GEANT [27]. The alignment of the detector, as well as other conditions like dead channels, is realistically simulated in the reconstruction. This procedure results in a biased measurement of the momentum also in Monte Carlo event samples, which thus have to be calibrated as well as in the data samples.
As it was mentioned previously, the calibration campaign of the momentum scale for the physics analyses during the Run 1 of the LHC was performed using $Z \to \mu\mu$

events. In addition, samples with J/ψ and Υ decaying to dimuons have been used to validate the corrections.
The complete list of all the datasets, both for collision and simulated data, are listed in table 3.1. To compute corrections, datasets corresponding to different data taking periods are merged according to the center-of-mass energy of the collisions and the conditions used in the reconstruction: for instance, samples corresponding to the 2011A and 2011B periods are merged together into one single dataset. The 2012D period is an exception to this rule, since in that dataset a small degradation in the muon reconstruction has been noticed, so a separate calibration is performed.
Events have been simulated with the MadGraph event generator [28] and the POWHEG NLO α_{S} generator [29] for $Z \to \mu\mu$ events: cross-checks have been performed and demonstrated that the result of the calibration is not sensitive to the generator used, so we used both depending on the availability. The J/ψ and Υ events have been generated with PYTHIA [30].
To ensure a low contamination, reconstructed muons from Z decays have been further selected according to criteria which have been developed specifically for vector boson decays. Those include cuts on acceptance, topology, isolation and quality of the track measurement.
Also for J/ψ and Υ events an offline selection has been applied requiring a minimal p_{T} threshold and muons to be identified with the PF identification [31].
The details of the muon selection are listed in table 3.2 where the following definitions are used:

- d_{xy} and d_z are the *impact parameter* of the muon track calculated in the transverse plane and along the $z-\mathrm{axis}$ respectively (the reference frame of CMS is described in 2.3.1);
- $I^{\mathrm{rel}}_{\mathrm{comb}}$ is the isolation in a cone of $\Delta R = 0.3$ around the muon's track, computed with the sum of tracks' p_{T} and the energy deposits in the calorimeters, divided by the muon's p_{T};
- $I^{\mathrm{rel}}_{\mathrm{PF}} = \dfrac{\sum^{\mathrm{ch,had}} p_{\mathrm{T}} + \max\left(\sum^{\mathrm{neutral,had}} E_T + \sum^{\mathrm{neutral,had}} E_T - \Delta\beta, 0\right)}{p_{\mathrm{T}}}$
  is the "Particle Flow" isolation computed in a cone of $\Delta R = 0.3$ around the muon's track. The sums are running over the *particle flow* candidates [31] and $\Delta\beta$ is a correction to the isolation energy due to pile-up interactions.

In addition to the per-muon selection, the reconstructed dimuon mass must lie between 60-120 GeV, 2-4 GeV and 8-12 GeV for Z, J/ψ and Υ events respectively.

3.3.4 Extra-smearing

Despite the sensible improvement observed in the data after the corrections, the resolution on the muon momentum in Monte Carlo samples can be up to 15% better than in the data. Since this difference is a source of systematic uncertainty for many analyses, a procedure to match the data and MC resolution is mandatory. This matching is improved by adding an extra gaussian smearing to the muons' curvature in the MC:

$$k_{\mathrm{smeared}} = k + |k| \cdot G(0, s)$$

Table 3.1: Datasets corresponding to the different samples of data and simulated events used for the study of the muon momentum scale at center of mass energy of $\sqrt{s}$=7 TeV and $\sqrt{s}$=8 TeV. “Label” is how the samples are indicated later.

$\sqrt{s}$	**Type**	**Label**	**Dataset name**
7 TeV	data	2011_DATA_44X	/DoubleMu/Run2011A-08Nov2011-v1,
	data		/DoubleMu/Run2011B-19Nov2011-v1
	data		/MuOnia/Run2011A-08Nov2011-v1,
	data		/MuOnia/Run2011B-19Nov2011-v1
	data	2011_DATA_42X	/DoubleMu/Run2011A-May10ReReco-v1,
	data		/DoubleMu/Run2011A-05Aug2011-v1,
	data		/DoubleMu/Run2011A-PromptReco-v4,
	data		/DoubleMu/Run2011A-PromptReco-v6,
	data		/DoubleMu/Run2011B-PromptReco-v1
	data		/MuOnia/Run2011A-05Aug2011-v1,
	data		/MuOnia/Run2011B-PromptReco-v1
	Monte Carlo	2011_MC_44X	/DYJetsToLL_TuneZ2_M-50_7TeV-madgraph-tauola/Fall11-PU_S6_START44_V9B-v1
	Monte Carlo		/JPsiToMuMu_2MuPEtaFilter_7TeV-pythia6-evtgen/Fall11-PU_S6_START44_V9B-v1
	Monte Carlo		/Upsilon1SToMuMu_2MuEtaFilter_tuneD6T_7TeV-pythia6-evtgen/Fall11-PU_S6_START44_V9B-v1
	Monte Carlo	2011_MC_42X	/DYToMuMu_M-20_CT10_TuneZ2_7TeV-powheg-pythia/Fall11-PU_S6_START42_V14B-v1
	Monte Carlo		/JPsiToMuMu_2MuPEtaFilter_7TeV-pythia6-evtgen/Summer11-PU_S4_START42_V11-v2
	Monte Carlo		/Upsilon1SToMuMu_2MuEtaFilter_tuneD6T_7TeV-pythia6-evtgen/Fall11-PU_S6_START42_V14B-v1
8 TeV	data	2012ABC_DATA_53X_Prompt	/DoubleMu/Run2012A-13Jul2012-v1,
	data		/DoubleMu/Run2012B-13Jul2012-v4,
	data		/DoubleMu/Run2012C-PromptReco-v1,
	data		/DoubleMu/Run2012C-PromptReco-v2
	data	2012D_DATA_53X_Prompt	/DoubleMu/Run2012D-PromptReco-v1
	data	2012ABC_DATA_53X_ReReco	/DoubleMu/Run2012A-22Jan2013-v1,
	data		/DoubleMuParked/Run2012B-22Jan2013-v1,
	data		/DoubleMuParked/Run2012C-22Jan2013-v1
	data	2012D_DATA_53X_ReReco	/DoubleMuParked/Run2012D-22Jan2013-v1
	Monte Carlo	2012_MC_53X	/DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball/DR53X-PU_S10_START53_V7A-v1
	Monte Carlo		/Upsilon1SToMuMu_2MuPtEtaFilter_tuneD6T_8TeV-pythia6-evtgen/Summer12_DR53X-PU_S10_START53_V7A-v1
	Monte Carlo		/JPsiToMuMu_2MuPtEtaFilter_tuneD6T_8TeV-pythia6-evtgen/Summer12_DR53X-PU_S10_START53_V7A-v2

Table 3.2: Selection of reconstructed muons.

$Z \to \mu\mu$ selection		
$\sqrt{s}$=7 TeV	Kinematics and topology:	$p_{\rm T} >$ 20 GeV, $\lvert\eta\rvert$ <2.4, $d_{xy} <$ 0.2 cm
	ID:	Global muon ID
	Isolation:	$I^{\rm rel}_{\rm comb} < 0.15 p_{\rm T}$
	Track quality:	$N_{\rm hits}$ tracker > 10
		$N_{\rm hits}$ pixel > 0
		$N_{\rm hits}$ muons > 0
		$\chi^2/{\rm ndof} < 10$
		$N_{\rm mu-stations}$ matched > 1
$\sqrt{s}$=8 TeV	Kinematics and topology:	$p_{\rm T} >$ 20 GeV, $\lvert\eta\rvert$ <2.4, $d_{xy} <$ 0.2 cm, d_z <5 cm
	ID:	PF muon ID
	Isolation:	$I^{\rm rel}_{\rm PF} < 0.12 p_{\rm T}$
	Track quality:	$N_{\rm tk-layers}$ with measurements > 5
		$N_{\rm hits}$ pixel > 0
		$N_{\rm mu-stations}$ matched > 1
$J/\psi \to \mu\mu$ selection		
$\sqrt{s}$=7 and 8 TeV	Kinematics:	$p_{\rm T} >$ 5 GeV
	ID:	PF muon ID
$\Upsilon \to \mu\mu$ selection		
$\sqrt{s}$=7 and 8 TeV	Kinematics:	$p_{\rm T} >$ 5 GeV
	ID:	PF muon ID

The difference in the p_{T} resolution depends in general on η and p_{T} of the muons. For this reason, the information on the resolution extracted from the fit (σ_{MC} and σ_{data}), as described in 3.3.2, is used to extract s as a continuous function:

$$s \equiv s(p_{\mathrm{T}}, \eta) = \gamma \frac{\sqrt{\sigma_{\mathrm{data}}^2 - \sigma_{\mathrm{MC}}^2}}{p_{\mathrm{T}}}$$

γ is a coefficient, of order unity, which is empirically determined.

3.4 Results

The fitted parameters of the scale function described in section 3.3.2, for the samples listed in table 3.1 are shown in table 3.3.

The scale function is validated using the mass and the experimental width of the resonance. The scale is primarily inspected by looking at the measured Z mass ($M_Z^{\rm fit}$) for different values of η and ϕ of the muon ("local validation"). Each bin is filled with events in which at least one of the two muons has η, φ lying in the interval of that bin: thus each events fills twice the mass spectra used in the validation.
An interesting observable which is highly sensitive to the residual misalignment of the detector is $M_Z^{\rm fit}$ vs. $\Delta\eta = \eta_{\mu^+} - \eta_{\mu^-}$. A η-dependent rotation of the tracker in the transverse plane, often referred to as a "twist", would result infact in a linear bias of the $p_{\rm T}$ as a function of η.
The "peak" value is extracted in each bin by means of a fit to the lineshape, in which the signal is modeled by a Breit-Wigner convoluted with a Crystal Ball function and the background by an exponential.
The results of such a validation are shown in figures 3.3 - 3.9. The effect of the corrections is clearly visible, in particular it is evident how the "φ-bias" present before the corrections is recovered thanks to the harmonic terms in the scale function, as it is described in 3.3.2. Another effect which is clearly visible in the "2011_DATA_42X" sample in figure 3.3e is believed to be due to a "twist" deformation of the tracker, resulting in a linear bias as a function of $\Delta\eta$. Here the η-dependent terms in the scale function helps to recover a flat profile.
To fully appreciate the correlation of the bias in the mass measurement between pseudorapity and azimuthal angle of the muons, two-dimensional "thermal maps" are also produced. Grids of 16×20 bins in the (ϕ, η) plane are used with each bin being filled with some 4000-20000 events. These validations are presented in figures 3.10 - 3.16. As a general comment, wherever the maps show a non-local, smooth bias, the analytic approach in the scale parameterization used in the MuScleFit algorithm works very well. However, for some samples kinematic regions of the muon exists where the bias is not fully recovered after the scale fit and the maps show "spikes" where the reconstructed mass is off up to ~ 1 GeV from the nominal Z mass, especially in the regions of larger muons' psudorapidities.
This is believed to be a conspiration of different causes:

- the scale parameterization implemented in the fit may be missing some "global" misalignment effects (e.g. higher order deformations);
- there may be "local" residual misalignments in some modules which the scale function cannot recover;
- there may be other causes of bias in the momentum measurement beyond the misalignment, which are not considered in the scale function, like an imperfect parameterization of the energy loss of the muon traversing the detector.

The *sagitta* deformation, i.e. a global shift of the tracker sensors in the transverse plane, is probably the most important feature that affects the track measurement in CMS. This deformation, in a large statistical sample which is ϕ-symmetric, results in

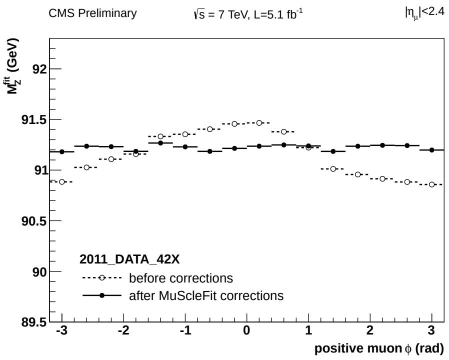
a "smearing" of the dimuon mass around a central scale, rather than a global shift, given that it is sinusoidal shaped and integrated over $[-\pi, +\pi]$. The benefit of the corrections can be appreciated also by looking at the dimuon mass resolution, here estimated as the width of the Crystal Ball function, which is shown in figure 3.17 as a function of the η of the positive muon. The resolution improves after the corrections, up to 30% in the outermost η bins.
The effect on the resolution in the MC samples is in general smaller, given that before any correction it is already better with respect to the collision data. Anyway, on the simulation an artificial smearing is applied after the scale corrections, as explained in 3.3.4, in order to match the resolution observed in the data.

After both momentum corrections and smearing are applied, it is interesting to study the absolute scale and the mass resolution by means of a fit to the lineshape using all the events in the dataset, in order to have a hint on the overall improvement on an inclusive sample of $Z \to \mu\mu$ events. This study is particularly interesting when data and MC are compared. For this comparison the two types of datasets have been considered together according to the release of the CMSSW software ("42X", "44X", "53X") used to reconstruct the events, which includes consistent sets of conditions for the detector for real and simulated data. In this sense, the smearing procedure for each MC sample is optimized to match the data resolution in the corresponding dataset for the collisions [2].
The results of these fits are summarized in figures 3.18 - 3.21. The raw muon momentum measurement results in a mass scale mismatch of up to 0.2 % at the Z peak, as it is summarized in the plots. After MuScleFit corrections are applied this shift is practically compatible with zero, within the uncertainties from the fit. The initial mismatch in the resolution between data and MC can be seen in the discrepancy in the values of σ_{CB}, which is as large as 10 %. After the corrections to the momentum resolution applied on simulation, this discrepancy is reduced at the level of 0.5 % or better. Both the absolute scale and the resolution will be discussed further in section 3.5. The combined effect of the absolute scale correction and the smearing is clearly visible in the data/MC ratio. It has to be remarked that the value of the Z boson mass coming from the fit is estimated using the pole of the BW distribution and may have a bias with respect to the nominal value ($M_Z = 91.188$ GeV) due to the signal model chosen in the pdf. Still the comparison between different samples is believed to be unbiased, since possible systematic effects due to the choice of the model should cancel out.

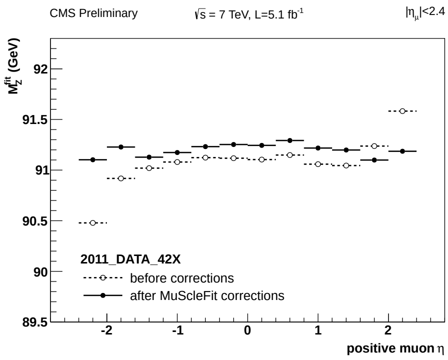
[2] notice that, for this purpose, two different smearings are applied on the "MC_2012_53X" sample depending on the data sample which is being used

Table 3.3: Values of the fitted parameters for the scale correction (equation 3.1 and 3.2) for data and simulation samples at 7 TeV and 8 TeV. The five sections in the lower part of the table correspond to the five η bins [-2.4,-2.1], [-2.1,-1.5], [-1.5,+1.5], [+1.5,+2.1] and [+2.1,+2.4]. The values of the b_0 parameters are fixed to guarantee the continuity of the correction function at the boundaries between the η bins.

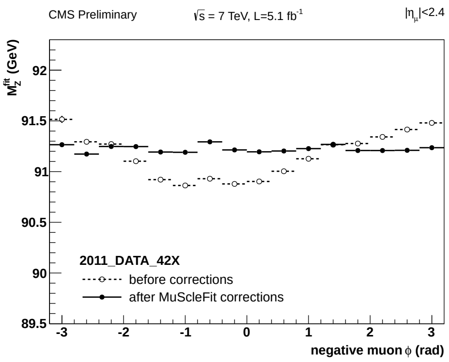
Sample	2012_MC_53X_smearReReco	2011_MC_44X	2011_DATA_42X	2011_DATA_44X	2012ABC_DATA_ReReco_53X	2012D_DATA_ReReco_53X	2011_MC_42X
p_0	0.00000	0.00000	-0.00121	-0.00122	-0.00139	-0.00135	0.00000
δ	0.00005	0.00004	0.00005	0.00004	0.00004	0.00004	0.00004
a_1	0.00027	0.00022	0.00145	0.00080	0.00028	0.00025	0.00112
ϕ_1	0.14179	0.28482	1.00365	1.33254	1.21602	1.21894	-0.00828
a_2	0.00023	0.00030	0.00053	0.00041	0.00019	0.00017	0.00091
ϕ_2	-1.71046	-1.68476	1.32513	1.79848	1.78374	1.93410	-1.99991
b	-0.00000	-0.00007	-0.00025	-0.00013	-0.00004	-0.00005	0.00001
b_0	-0.00002	-0.00001	0.00029	0.00010	0.00003	0.00004	-0.00011
a_1	0.00001	0.00006	0.00039	0.00009	0.00002	0.00000	0.00021
ϕ_1	-1.04015	-1.13170	0.97164	1.17708	0.26269	0.15297	-0.94322
b	0.00004	0.00003	-0.00043	-0.00024	-0.00008	-0.00009	0.00018
b_0	0.00000	0.00000	0.00003	-0.00004	-0.00002	-0.00002	-0.00000
a_1	0.00007	0.00007	0.00015	0.00015	0.00007	0.00007	0.00019
ϕ_1	-1.64733	-1.75023	-1.52230	-1.30574	-1.24722	-1.39464	-1.77400
b	-0.00000	-0.00000	-0.00002	0.00003	0.00001	0.00001	0.00000
a_1	0.00003	0.00013	0.00007	0.00001	0.00002	0.00001	0.00037
ϕ_1	-1.68410	-1.40495	1.29260	0.89885	0.21788	1.17093	-1.23903
b	0.00000	-0.00000	-0.00054	-0.00018	-0.00003	-0.00003	0.00001
b_0	-0.00000	-0.00000	-0.00003	0.00004	0.00002	0.00002	0.00000
a_1	0.00018	0.00014	0.00096	0.00058	0.00021	0.00032	0.00087
ϕ_1	-0.94791	-1.42615	1.58460	1.85334	1.94164	1.89407	-1.40062
a_2	0.00012	0.00001	0.00053	0.00028	0.00012	0.00008	0.00083
ϕ_2	0.38600	0.78290	-0.32991	-0.84138	-1.10183	-1.23759	2.70902
b	-0.00000	-0.00001	-0.00076	-0.00020	-0.00008	-0.00010	-0.00006
b_0	-0.00000	-0.00000	-0.00035	-0.00006	-0.00001	-0.00000	0.00001



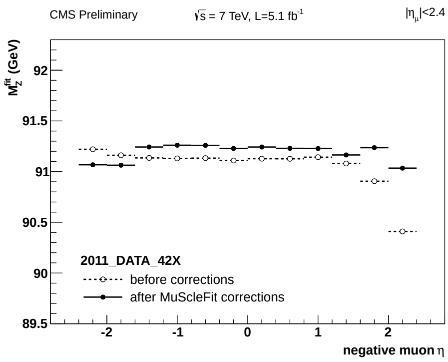
(a) M_Z^{fit} vs. φ_{μ^+}



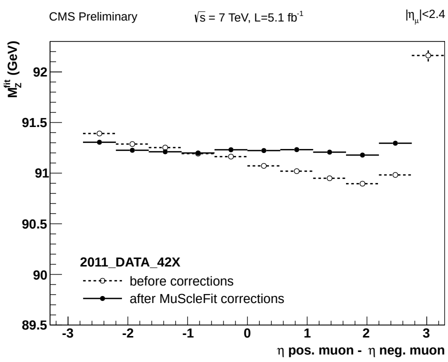
(b) M_Z^{fit} vs. η_{μ^+}



(c) M_Z^{fit} vs. φ_{μ^-}

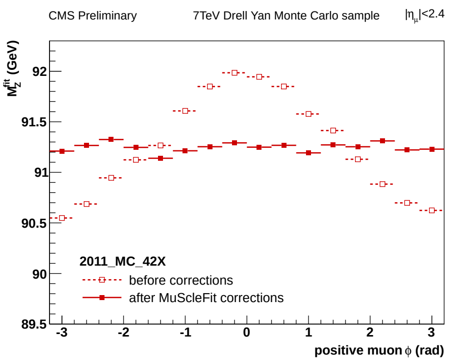


(d) M_Z^{fit} vs. η_{μ^-}

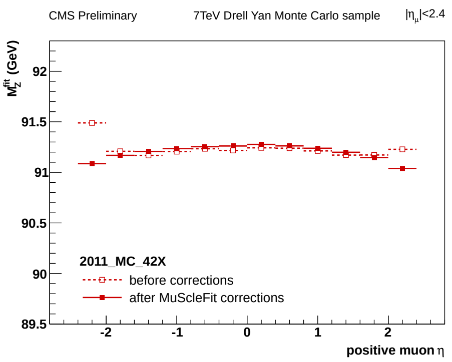


(e) M_Z^{fit} vs. $\Delta\eta$

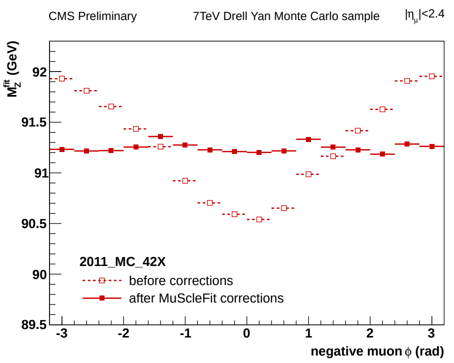
Figure 3.3: Local validation of the scale correction performed on the "2011_DATA_42X" sample.



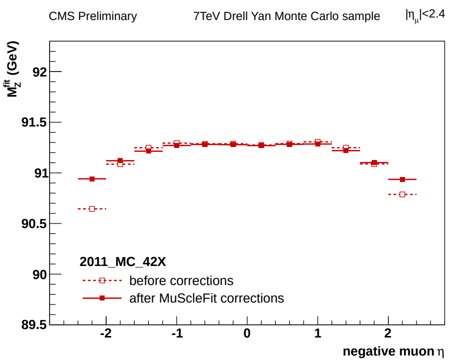
(a) M_Z^{fit} vs. φ_{μ^+}



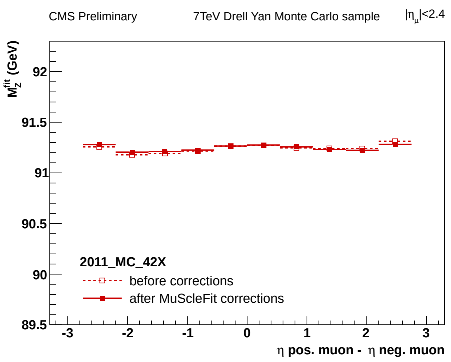
(b) M_Z^{fit} vs. η_{μ^+}



(c) M_Z^{fit} vs. φ_{μ^-}

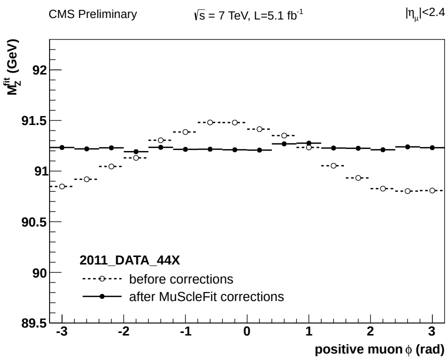


(d) M_Z^{fit} vs. η_{μ^-}

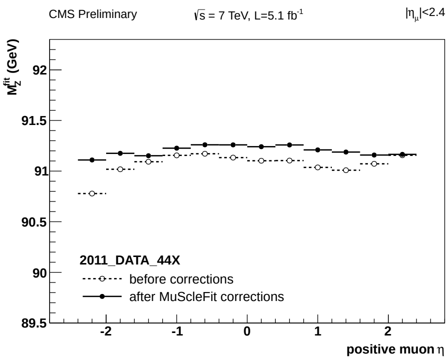


(e) M_Z^{fit} vs. $\Delta\eta$

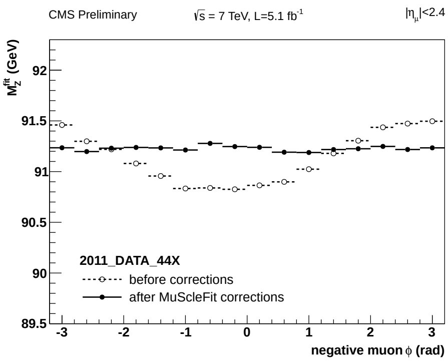
Figure 3.4: Local validation of the scale correction performed on the "2011_MC_42X" sample.



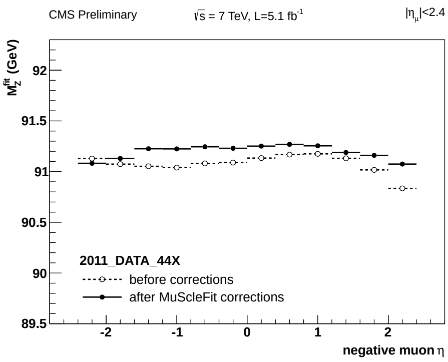
(a) $M_Z^{\rm fit}$ vs. φ_{μ^+}



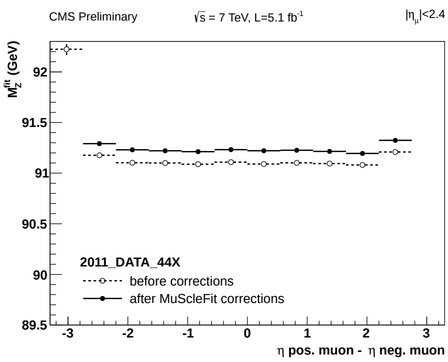
(b) $M_Z^{\rm fit}$ vs. η_{μ^+}



(c) $M_Z^{\rm fit}$ vs. φ_{μ^-}

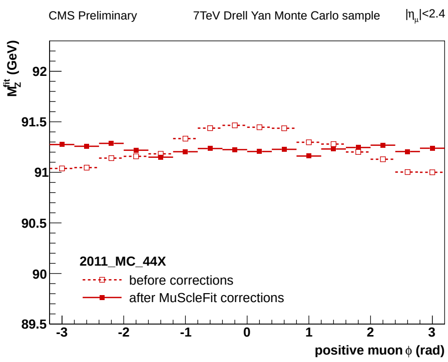


(d) $M_Z^{\rm fit}$ vs. η_{μ^-}

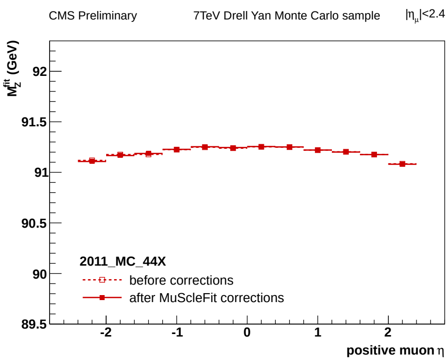


(e) $M_Z^{\rm fit}$ vs. $\Delta\eta$

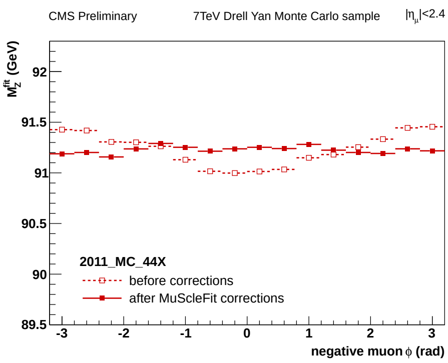
Figure 3.5: Local validation of the scale correction performed on the "2011_DATA_44X" sample.



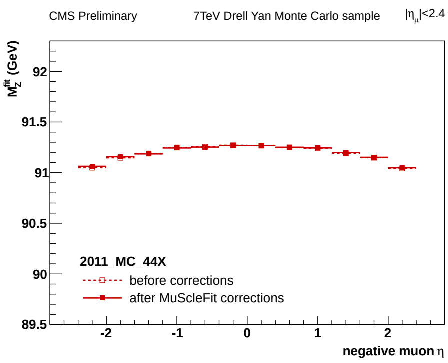
(a) M_Z^{fit} vs. φ_{μ^+}



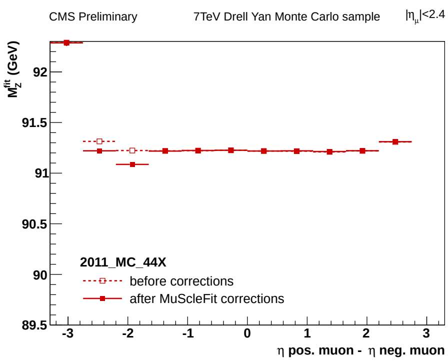
(b) M_Z^{fit} vs. η_{μ^+}



(c) M_Z^{fit} vs. φ_{μ^-}

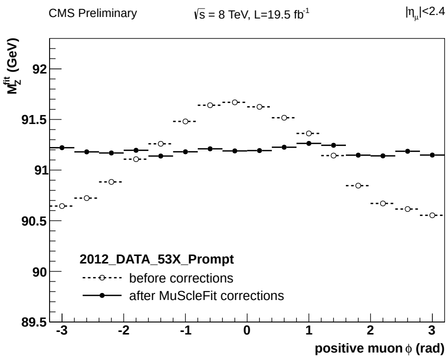


(d) M_Z^{fit} vs. η_{μ^-}

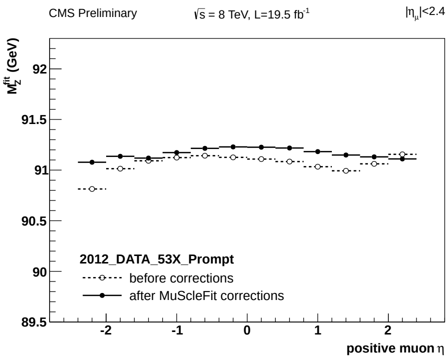


(e) M_Z^{fit} vs. $\Delta\eta$

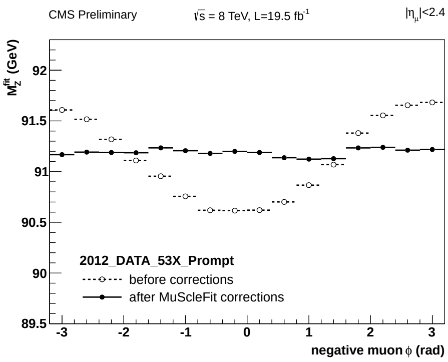
Figure 3.6: Local validation of the scale correction performed on the "2011_MC_44X" sample.



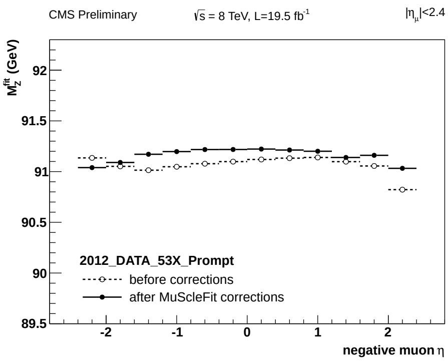
(a) M_Z^{fit} vs. φ_{μ^+}



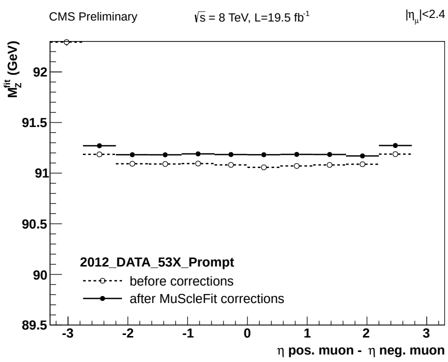
(b) M_Z^{fit} vs. η_{μ^+}



(c) M_Z^{fit} vs. φ_{μ^-}

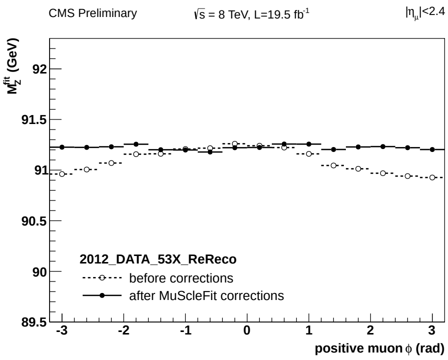


(d) M_Z^{fit} vs. η_{μ^-}

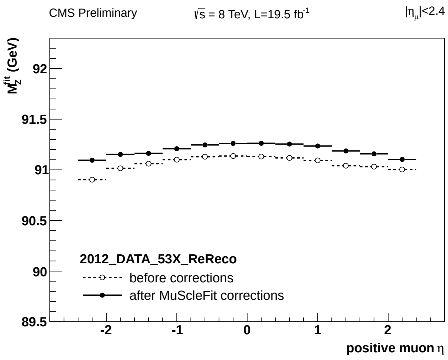


(e) M_Z^{fit} vs. $\Delta\eta$

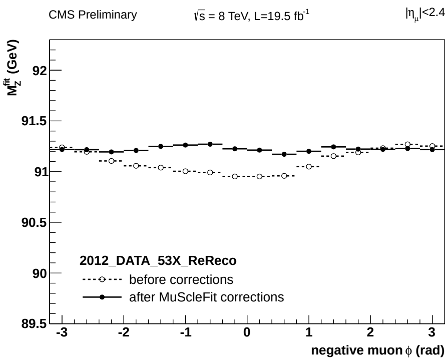
Figure 3.7: Local validation of the scale correction performed on the "2012_DATA_53X_Prompt" sample.



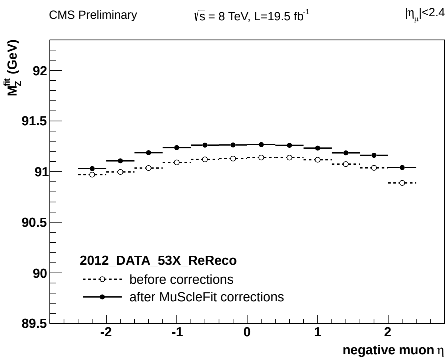
(a) M_Z^{fit} vs. φ_{μ^+}



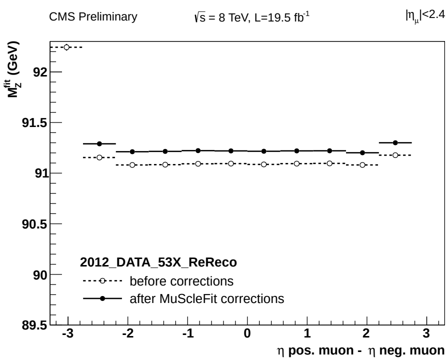
(b) M_Z^{fit} vs. η_{μ^+}



(c) M_Z^{fit} vs. φ_{μ^-}

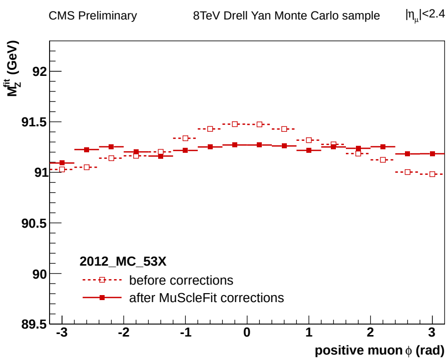


(d) M_Z^{fit} vs. η_{μ^-}

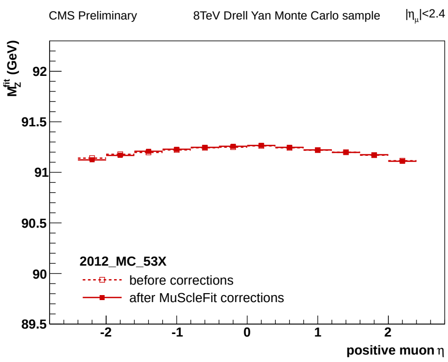


(e) M_Z^{fit} vs. $\Delta\eta$

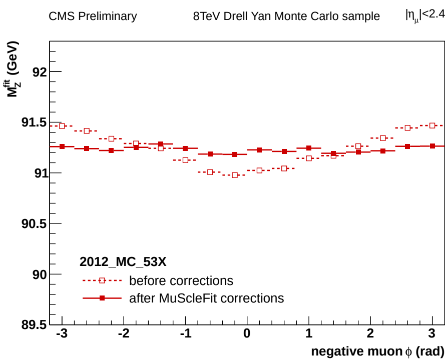
Figure 3.8: Local validation of the scale correction performed on the "2012_DATA_53X_ReReco" sample.



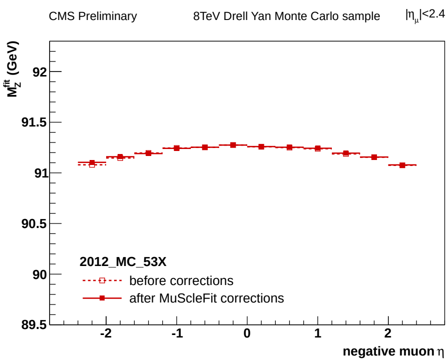
(a) M_Z^{fit} vs. φ_{μ^+}



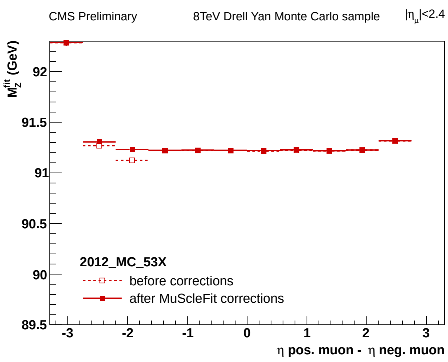
(b) M_Z^{fit} vs. η_{μ^+}



(c) M_Z^{fit} vs. φ_{μ^-}

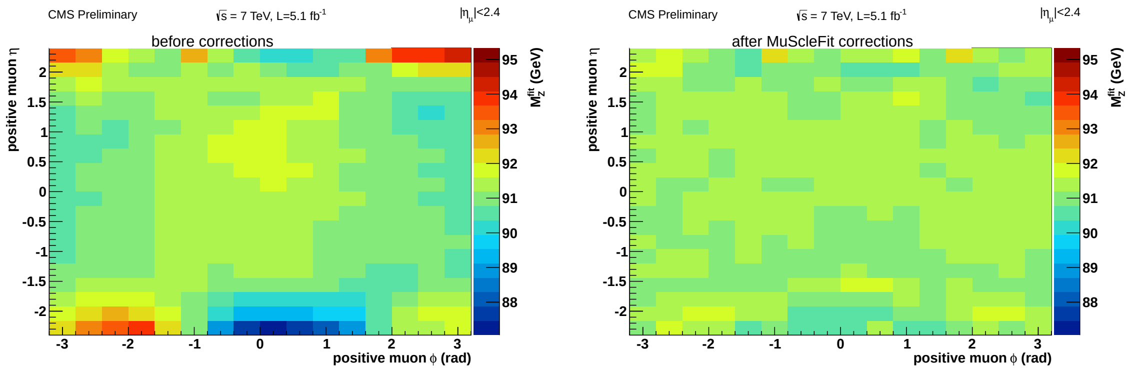


(d) M_Z^{fit} vs. η_{μ^-}

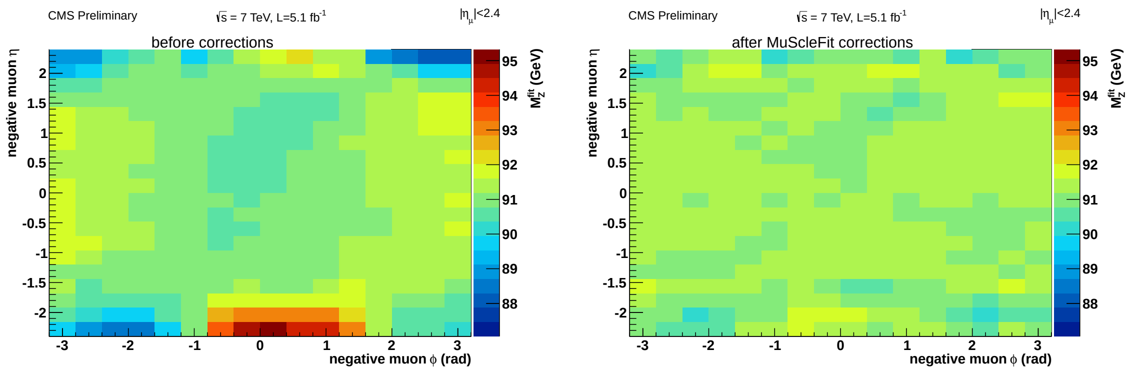


(e) M_Z^{fit} vs. $\Delta\eta$

Figure 3.9: Local validation of the scale correction performed on the "2012_MC_53X" sample.

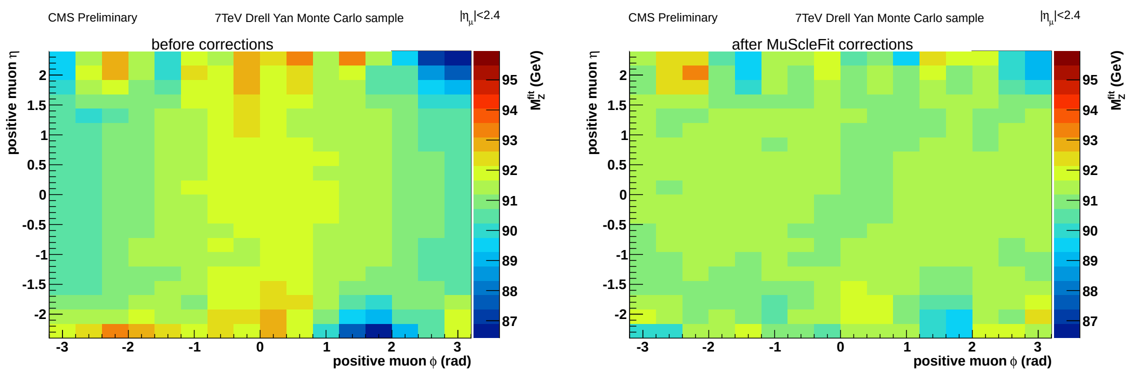


(a)

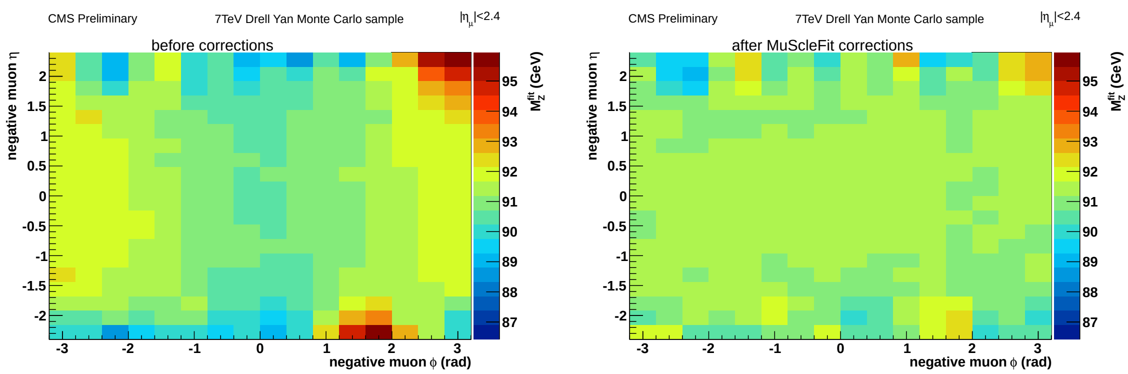


(b)

Figure 3.10: "Thermal maps" showing the reconstructed mass M_Z^{fit} with the azimuthal angle (φ) and the pseudorapidity (η) of the muon, for positive 3.10a and negative 3.10b muons, before (left) and after (right) corrections, for the "DATA_2011_42X" sample.

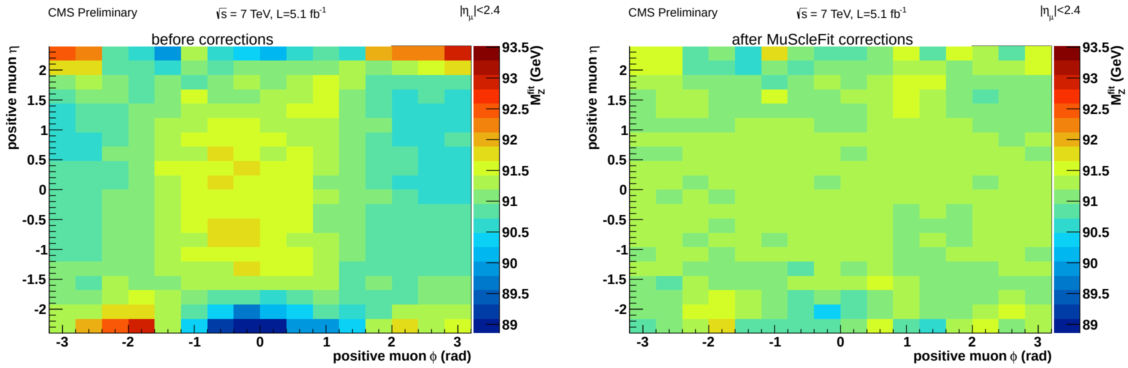


(a)

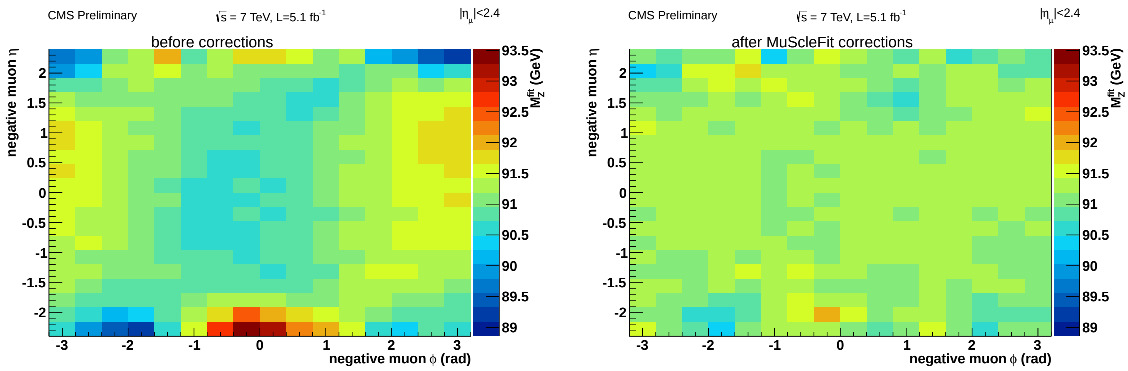


(b)

Figure 3.11: "Thermal maps" showing the reconstructed mass M_Z^{fit} with the azimuthal angle (φ) and the pseudorapidity (η) of the muon, for positive 3.11a and negative 3.11b muons, before (left) and after (right) corrections, for the "MC_2011_42X" sample.

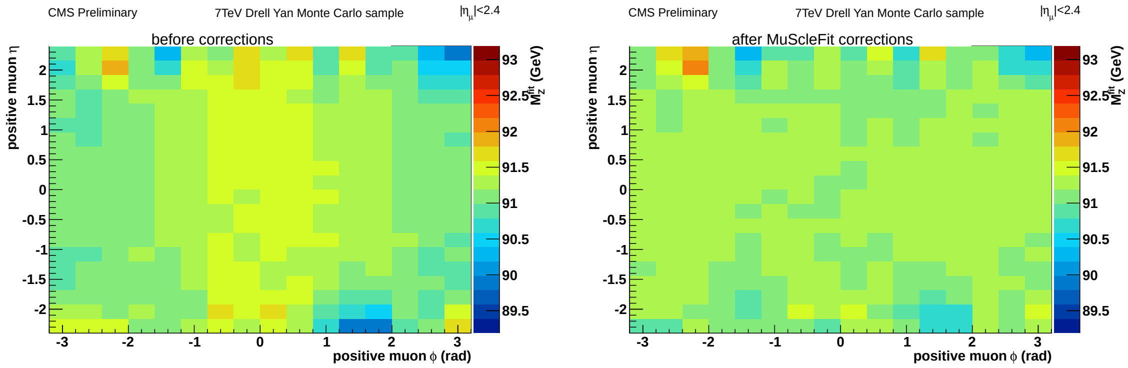


(a)

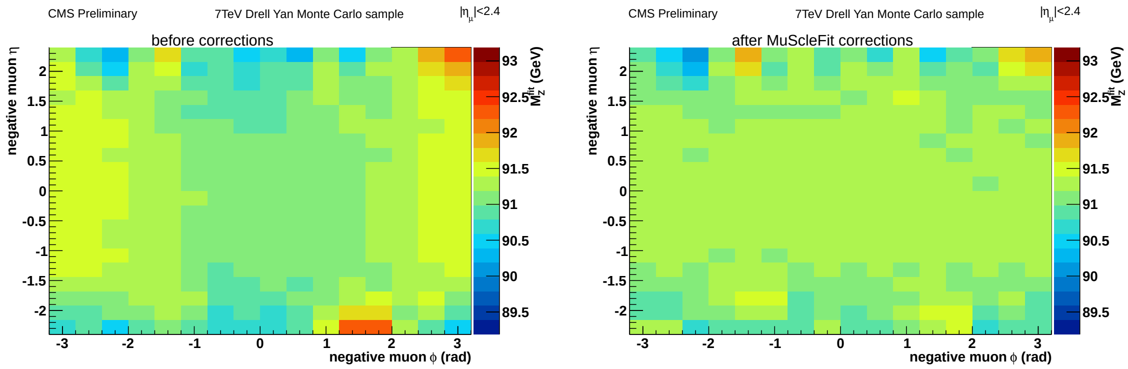


(b)

Figure 3.12: "Thermal maps" showing the reconstructed mass M_Z^{fit} with the azimuthal angle (φ) and the pseudorapidity (η) of the muon, for positive 3.12a and negative 3.12b muons, before (left) and after (right) corrections, for the "DATA_2011_44X" sample.

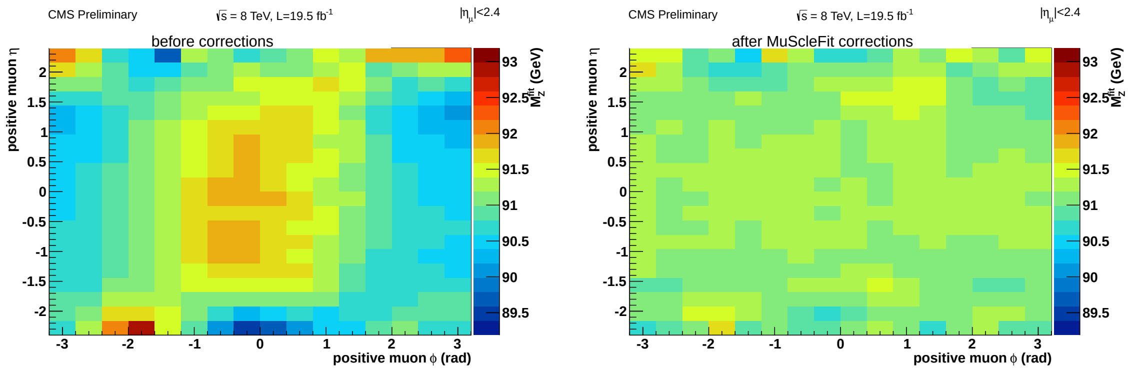


(a)

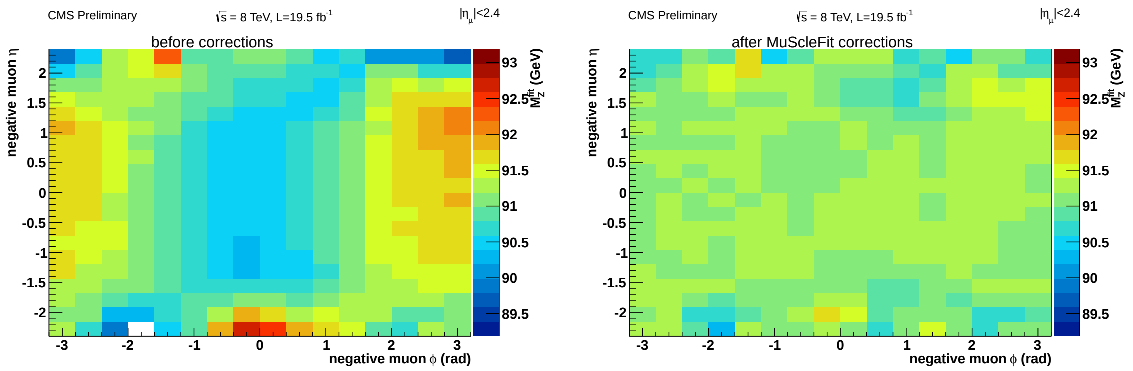


(b)

Figure 3.13: "Thermal maps" showing the reconstructed mass M_Z^{fit} with the azimuthal angle (φ) and the pseudorapidity (η) of the muon, for positive 3.13a and negative 3.13b muons, before (left) and after (right) corrections, for the "MC_2011_44X" sample.

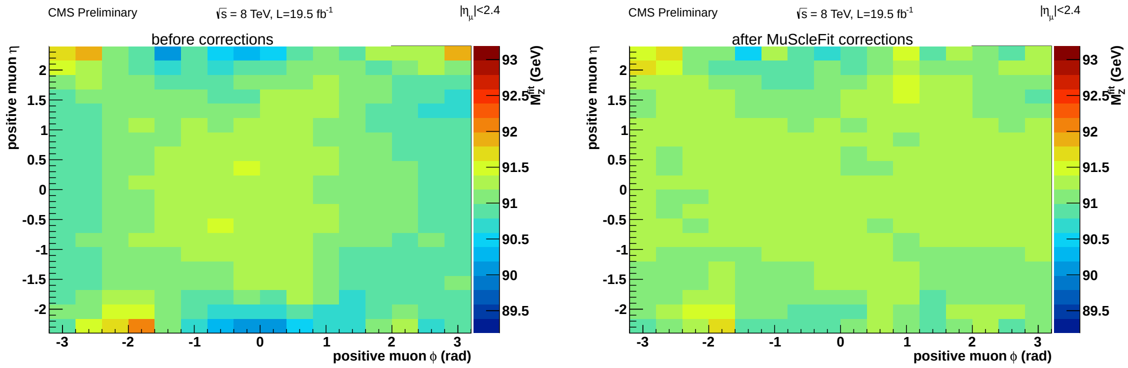


(a)

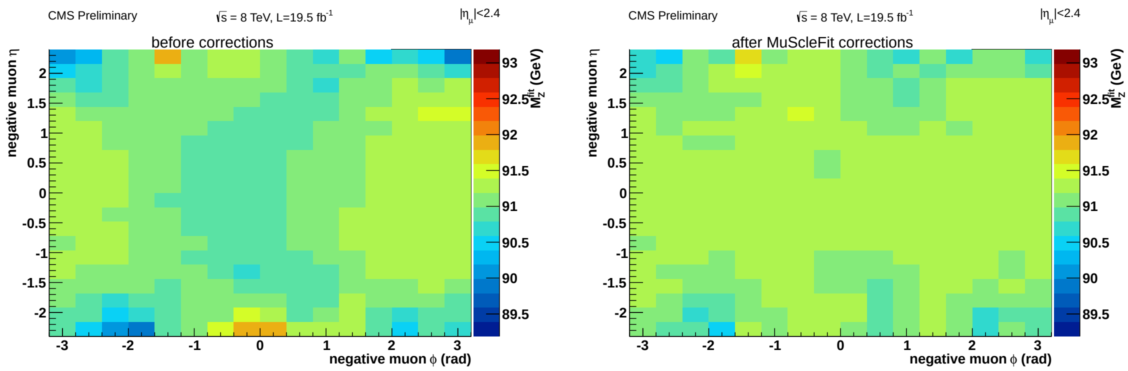


(b)

Figure 3.14: "Thermal maps" showing the reconstructed mass M_Z^{fit} with the azimuthal angle (φ) and the pseudorapidity (η) of the muon, for positive 3.14a and negative 3.14b muons, before (left) and after (right) corrections, for the "DATA_2012_53X_Prompt" sample.

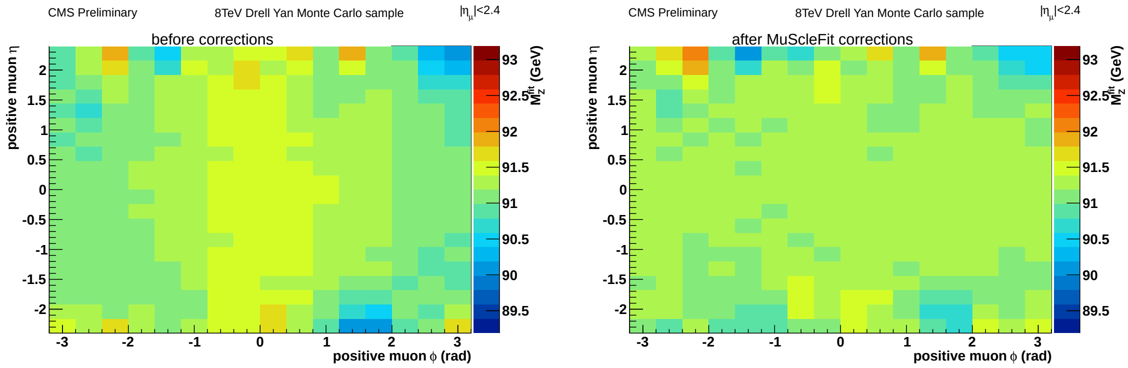


(a)

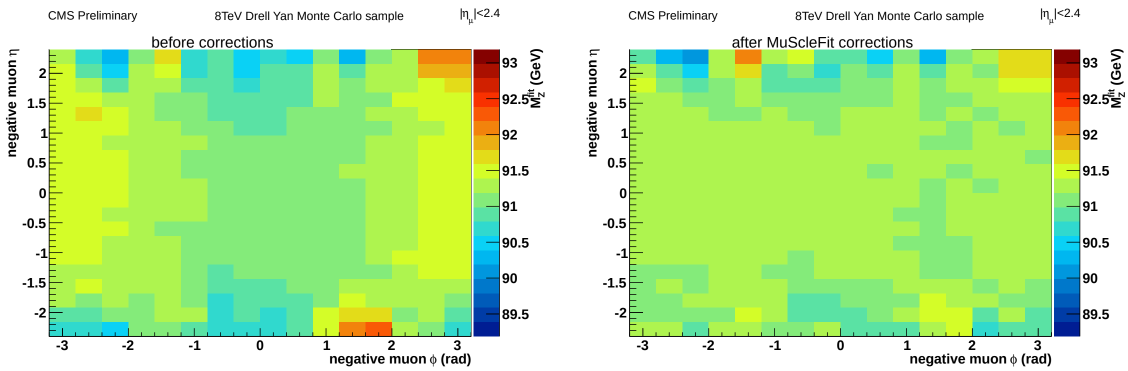


(b)

Figure 3.15: “Thermal maps” showing the reconstructed mass M_Z^{fit} with the azimuthal angle (φ) and the pseudorapidity (η) of the muon, for positive 3.15a and negative 3.15b muons, before (left) and after (right) corrections, for the “DATA_2012_53X_ReReco” sample.

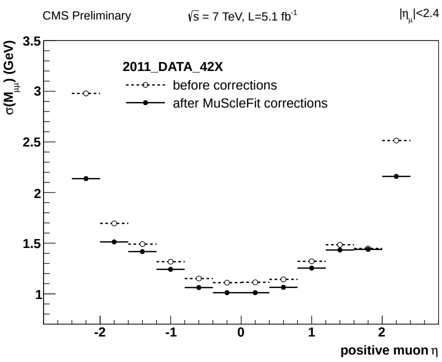


(a)

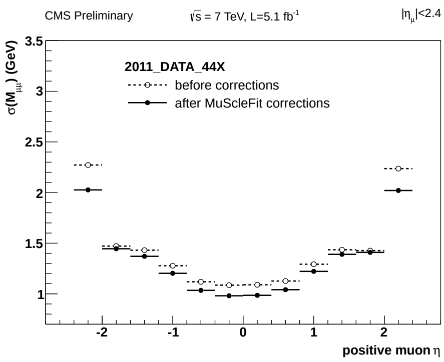


(b)

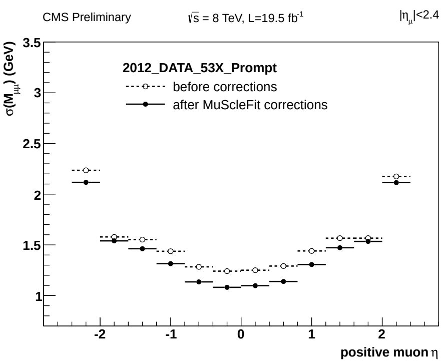
Figure 3.16: "Thermal maps" showing the reconstructed mass M_Z^{fit} with the azimuthal angle (φ) and the pseudorapidity (η) of the muon, for positive 3.16a and negative 3.16b muons, before (left) and after (right) corrections, for the "MC_2012_53X" sample.



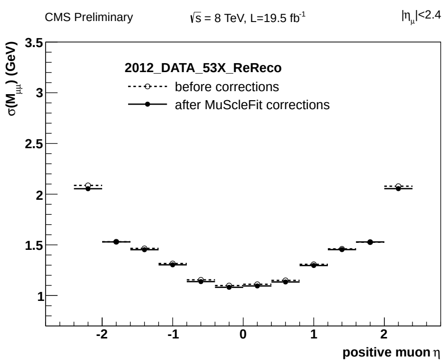
(a) 2011_DATA_42X



(b) 2011_DATA_44X



(c) 2012_DATA_53X_Prompt



(d) 2012_DATA_53X_ReReco

Figure 3.17: The mass resolution as a function of the η of the positive muon, for the 4 different data samples that are listed in table 3.1, shown before (dashed line) and after (solid line) the corrections.

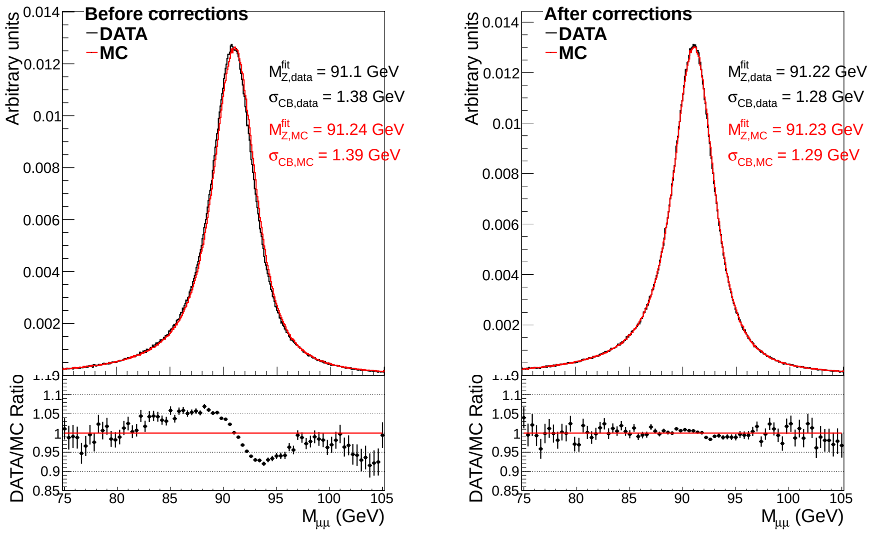


Figure 3.18: The dimuon spectrum in selected $Z \to \mu\mu$ events for data and MC with the results of the fit, before (left) and after (right) the corrections for the 2011 data and MC reconstructed with the 42X release of CMSSW. The fitted background has been subtracted from both data and MC.

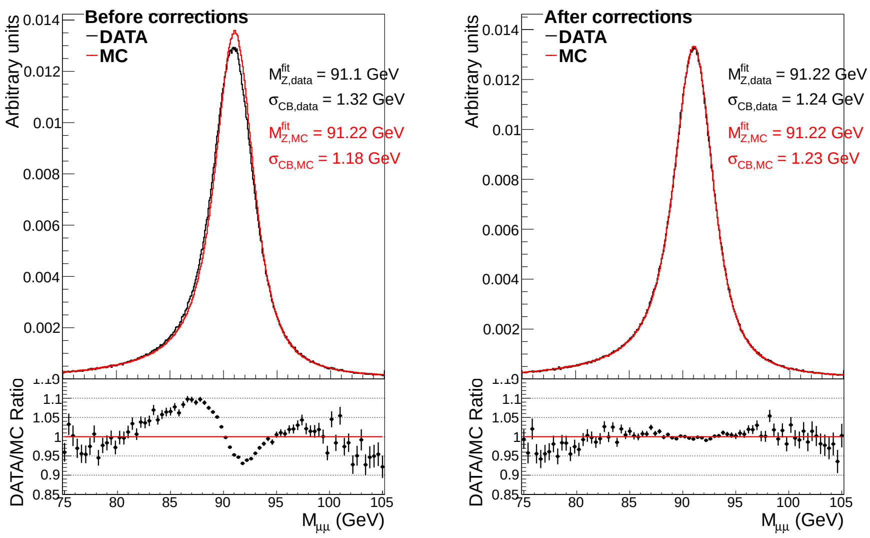


Figure 3.19: The dimuon spectrum in selected $Z \to \mu\mu$ events for data and MC with the results of the fit, before (left) and after (right) the corrections for the 2011 data and MC reconstructed with the 44X release of CMSSW. The fitted background has been subtracted from both data and MC.

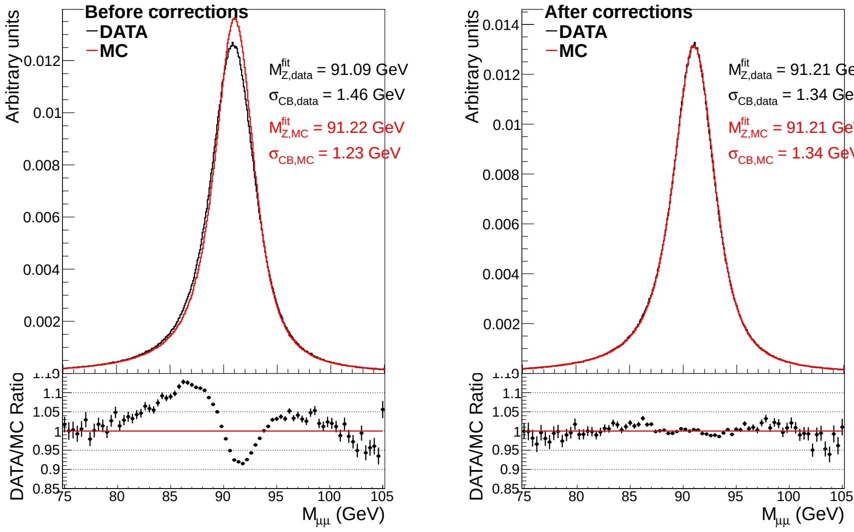


Figure 3.20: The dimuon spectrum in selected $Z \to \mu\mu$ events for data and MC with the results of the fit, before (left) and after (right) the corrections for the 2012 data and MC using the Prompt reconstruction conditions. The fitted background has been subtracted from both data and MC.

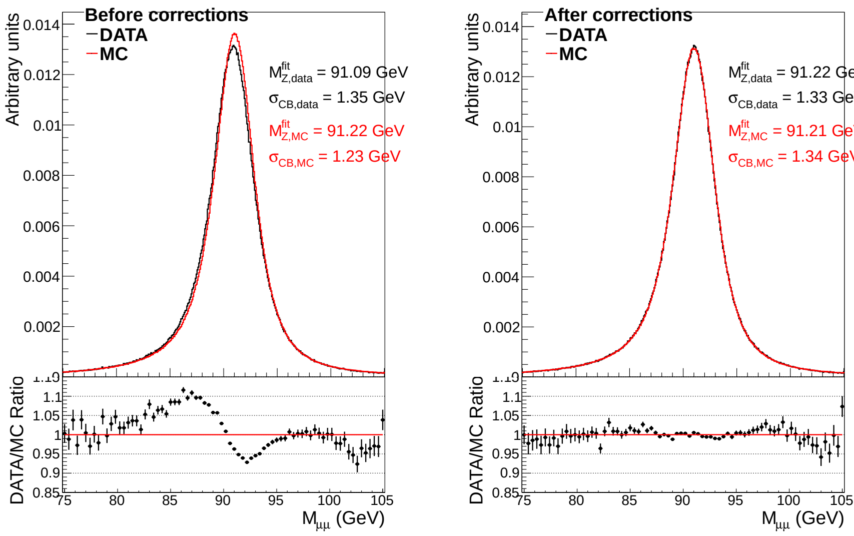


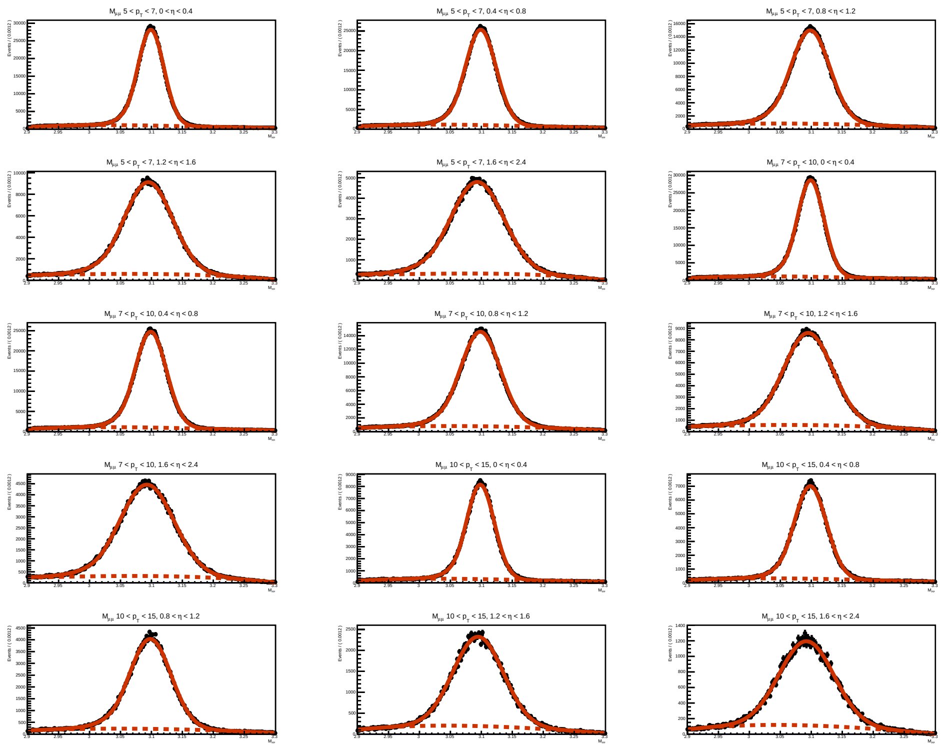
Figure 3.21: The dimuon spectrum in selected $Z \to \mu\mu$ events for data and MC with the results of the fit, before (left) and after (right) the corrections for the 2012 data and MC using the ReReco reconstruction conditions. The fitted background has been subtracted from both data and MC.

3.5 Systematic uncertainty on the muon momentum scale and resolution

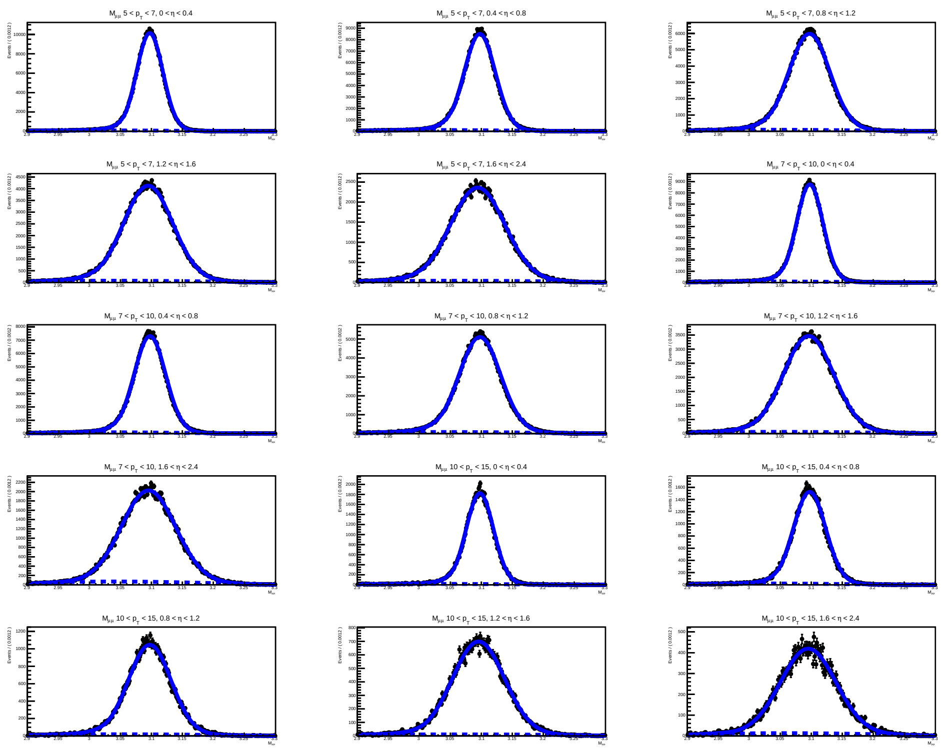
In order to assess the systematic uncertainty on the momentum scale of the muons a more complete validation is needed, to test the p_{T} measurement in CMS over the largest possible range and in events with different topologies. For this purpose, along with the Z, also the J/ψ and $\Upsilon(1S)$ resonances have been used. Spectra of dimuon mass have been filled in different $|\eta|$ and p_{T} bins, where, as for the validation shown in section 3.4, at least one muon is required to fulfill the kinematic requirement, and no cut is applied to the other one.
An unbiased procedure to extract the mass of the resonance has been developed for the three resonances: the pdfs chosen for the signal and the background are summarized in table 3.4. For the Z a slightly different shape of the background has been used, with respect to the one in 3.4, which optimizes the fits at high $p_{\mathrm{T}}s$. Examples of fits are in figures 3.22 - 3.27.

Table 3.4: Probability density functions (pdf's) used to extract the mass of the J/ψ, $\Upsilon(1S)$ and Z resonances from the spectrum of the dimuon invariant mass x. $CB(x)$ is the Crystal Ball, $BW(x)$ is the Breit-Wigner, "Bern(x)" are the Bernstein polynomials.

Resonance		Fit Range [GeV]	Signal pdf	Background pdf
J/ψ	DATA	[2.9,3.3]	$CB(x,\mu,\sigma)$	3rd order Bern. pol.
	MC	[2.9,3.3]	$CB(x,\mu,\sigma)$	3rd order Bern. pol.
$\Upsilon(1S)$	DATA	[8.7,11.0]	$CB_(x,\mu_1,\sigma)+CB(x,\mu_2,\sigma)$ $+CB(x,\mu_3,\sigma)$	4th order Bern. pol.
	MC	[8.8,10.0]	$CB_(x,\mu,\sigma)$	4th order Bern. pol.
Z	DATA	[75,105]	$BW(x,\mu,\Gamma_Z)\otimes CB(x,0,\sigma)$	$\exp(a_0+a_1x+a_2x^2)$
	MC	[75,105]	$BW(x,\mu,\Gamma_Z)\otimes CB(x,0,\sigma)$	$\exp(a_0+a_1x+a_2x^2)$

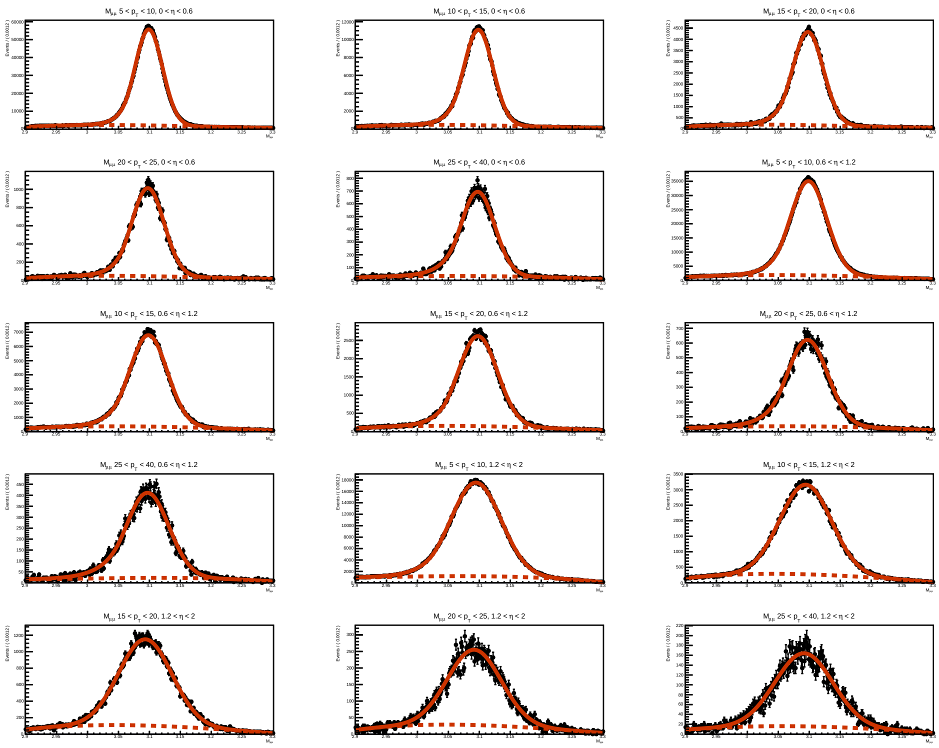


(a) DATA

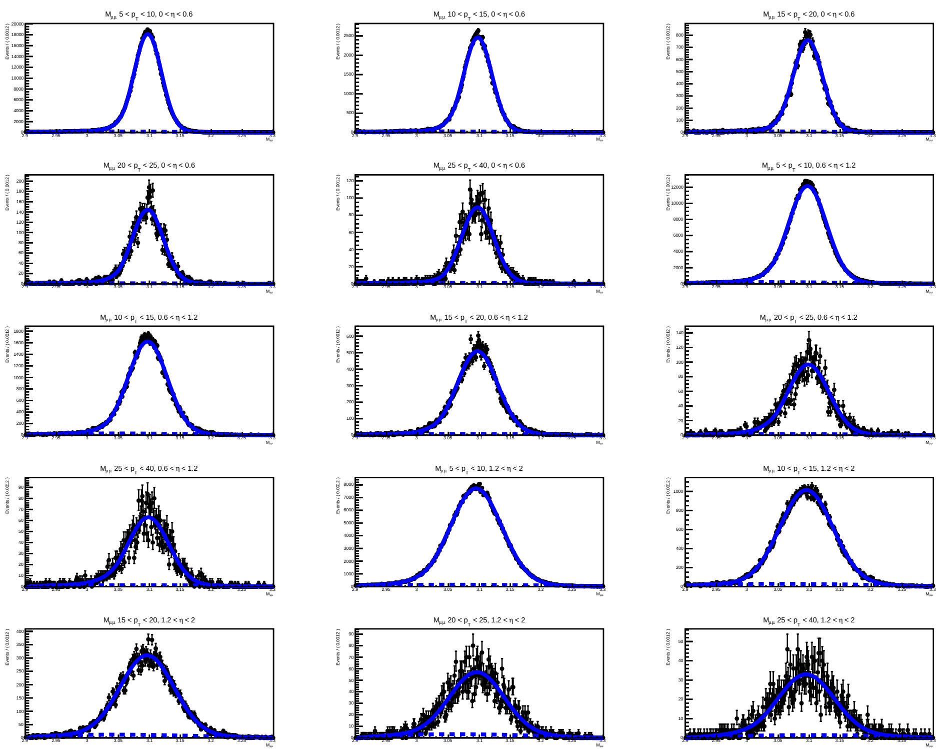


(b) MC

Figure 3.22: Example of fits to the dimuon spectrum in the mass region around the J/ψ resonance, for different $|\eta|$ bins and in three categories for muon p_{T}, in data (3.23a) and MC (3.23b), for the datasets "2012_DATA_53X_ReReco" and "2012_MC_53X" respectively.

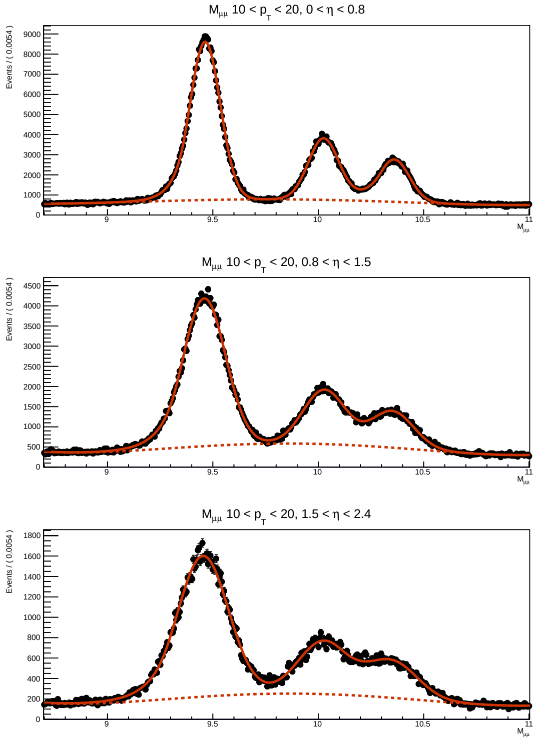


(a) DATA

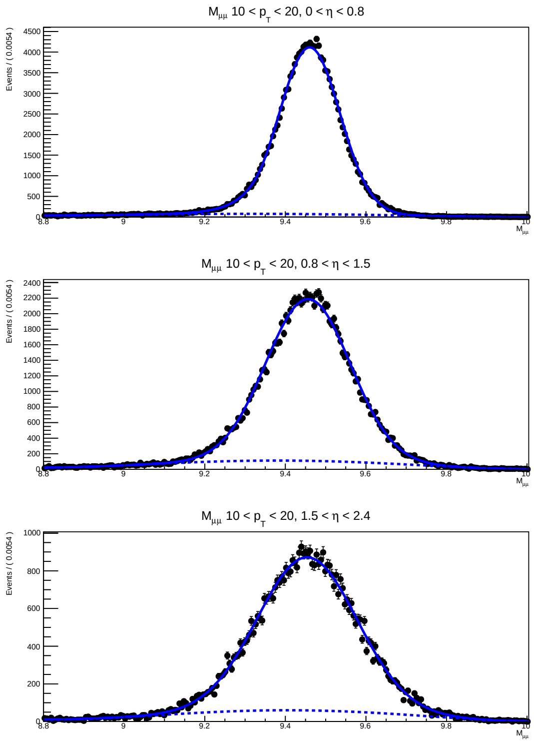


(b) MC

Figure 3.23: Example of fits to the dimuon spectrum in the mass region around the J/ψ resonance, for different p_{T} bins and in three categories for muon $|\eta|$, in data (3.23a) and MC (3.23b), for the datasets "2012_DATA_53X_ReReco" and "2012_MC_53X" respectively.

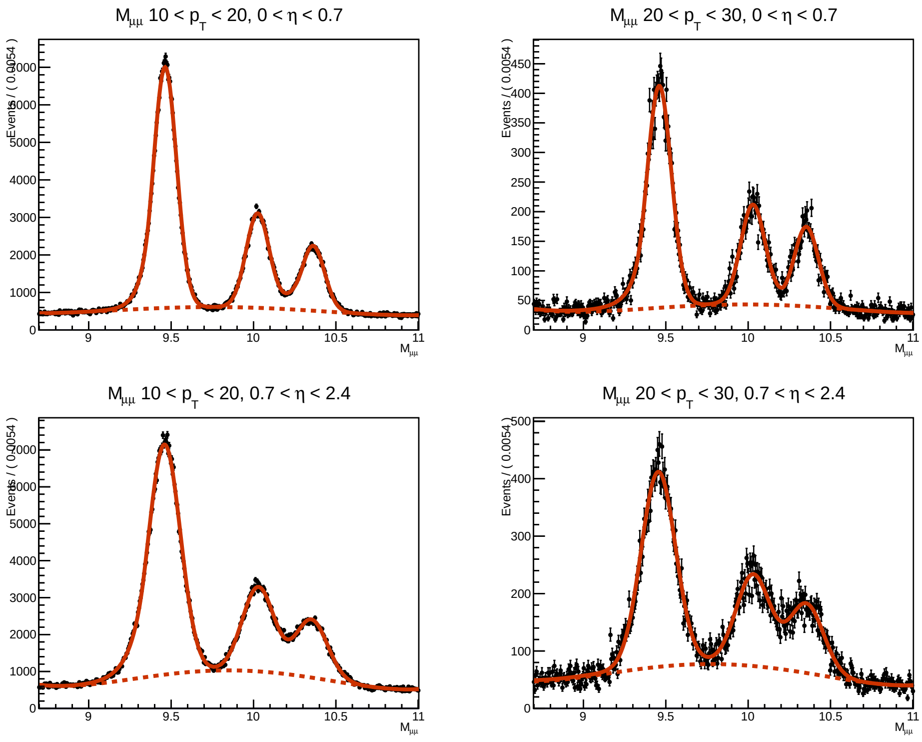


(a) DATA

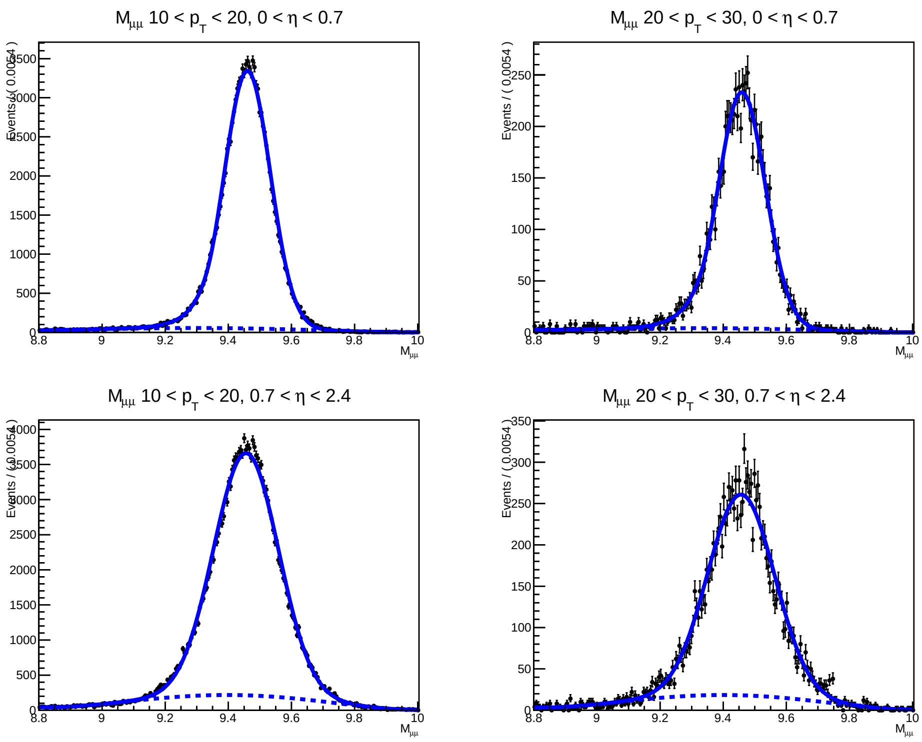


(b) MC

Figure 3.24: Example of fits to the dimuon spectrum in the mass region around the Υ resonance, for different $|\eta|$ bins, in data (3.25a) and MC (3.25b), for the datasets "2012_DATA_53X_ReReco" and "2012_MC_53X" respectively.

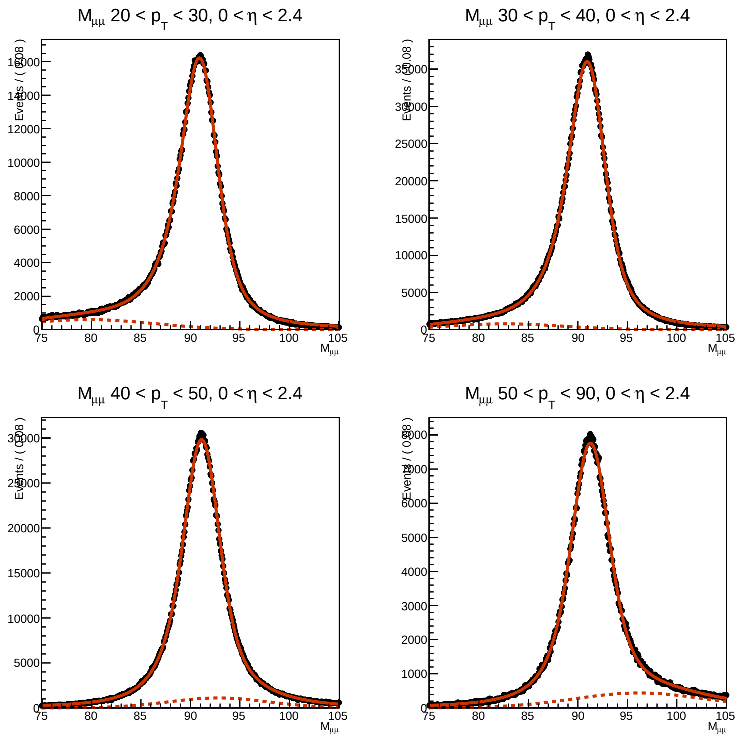


(a) DATA

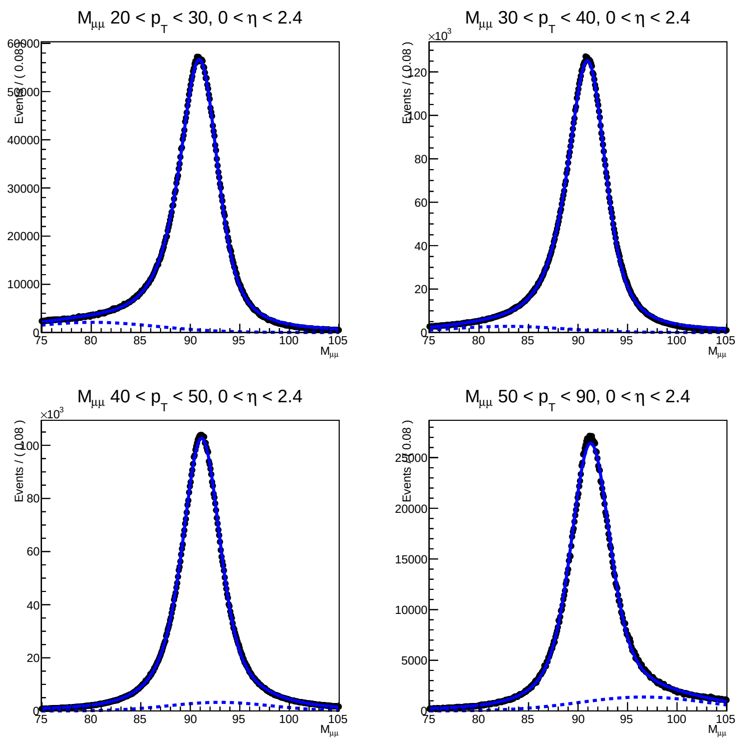


(b) MC

Figure 3.25: Example of fits to the dimuon spectrum in the mass region around the Υ resonance, for different p_{T} bins in two different $|\eta|$ categories, in data (3.25a) and MC (3.25b), for the datasets "2012_DATA_53X_ReReco" and "2012_MC_53X" respectively.

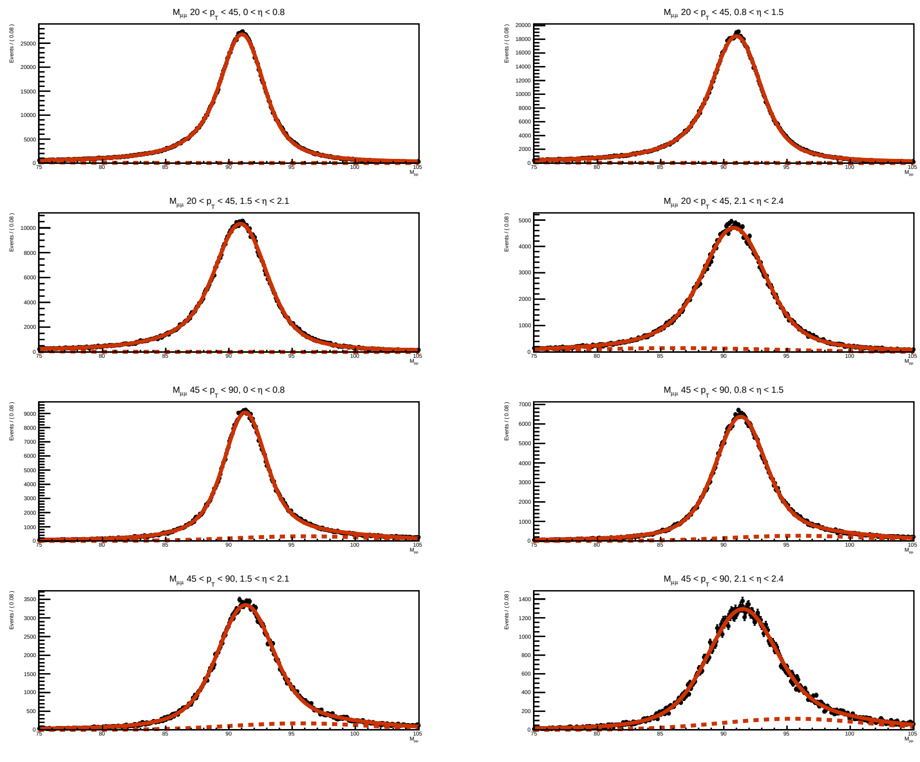


(a) DATA

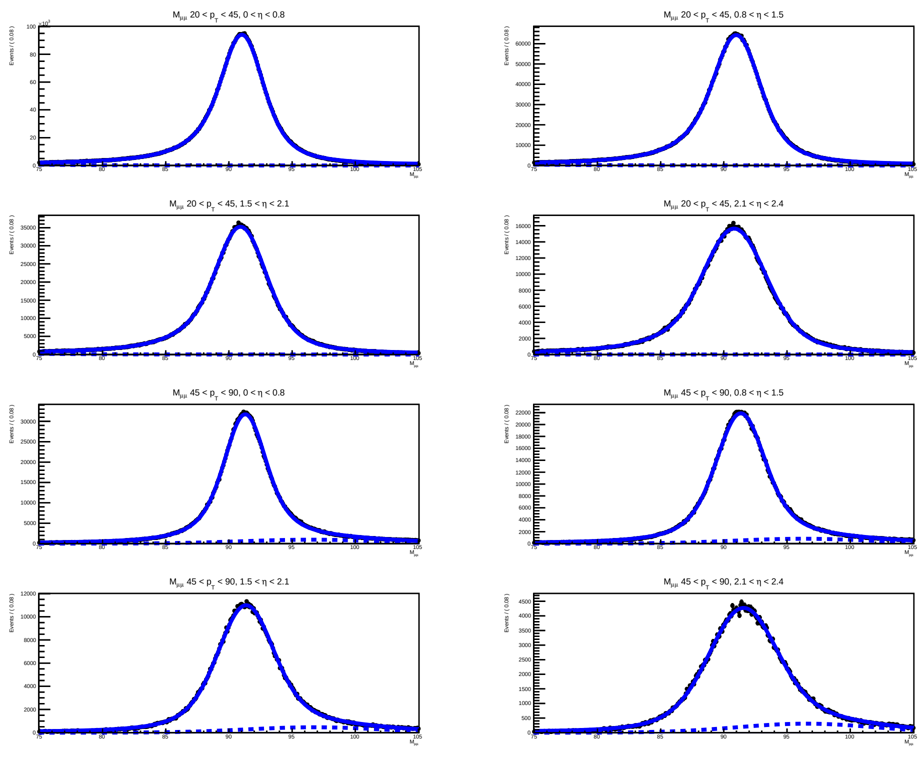


(b) MC

Figure 3.26: Example of fits to the dimuon spectrum in the mass region around the Z resonance, for different p_{T} bins, in data (3.27a) and MC (3.27b), for the datasets "2012_DATA_53X_ReReco" and "2012_MC_53X" respectively.



(a) DATA



(b) MC

Figure 3.27: Example of fits to the dimuon spectrum in the mass region around the Z resonance, for different $|\eta|$ bins and in two categories for muon p_{T}, in data (3.27a) and MC (3.27b), for the datasets "2012_DATA_53X_ReReco" and "2012_MC_53X" respectively.

There are two common use cases for defining a systematic error on the momentum scale:

- *Case I*: analyses where the spectrum in the data is compared with a model known from the simulation. Here the systematic error is indicated by the residual discrepancy between the post-correction spectrum in the data and in the simulation. Under the hypothesis that the systematic uncertainty is fully correlated between the two muons, the scale uncertainty is defined as:

  $$\frac{\delta p_{\mathrm{T}}^{\mathrm{data-MC}}}{p_{\mathrm{T}}} = \frac{M_{\mathrm{data}} - M_{\mathrm{MC}}}{M_{\mathrm{PDG}}} \tag{3.4}$$

  where M_{PDG} are the world average reported by the Particle Data Group (PDG) [32]: $M_{Z,\mathrm{PDG}} = 91.188\,\mathrm{GeV}$, $M_{\Upsilon(1S),\mathrm{PDG}} = 9.460\,\mathrm{GeV}$ and $M_{J/\psi,\mathrm{PDG}} = 3.097\,\mathrm{GeV}$.

  The fractional difference on the mass, after MuScleFit corrections, is shown as a function of $|\eta|$ or p_{T} of one of the two muons in figure 3.28 - 3.31. It is everywhere well within ±0.10%, for all the samples, with the largest departure from 0 being that for the point in the large $|\eta|$ and p_T bins. Conservatively we assign a ±0.05% uncertainty everywhere apart from this point where a ±0.10% uncertainty is assumed.

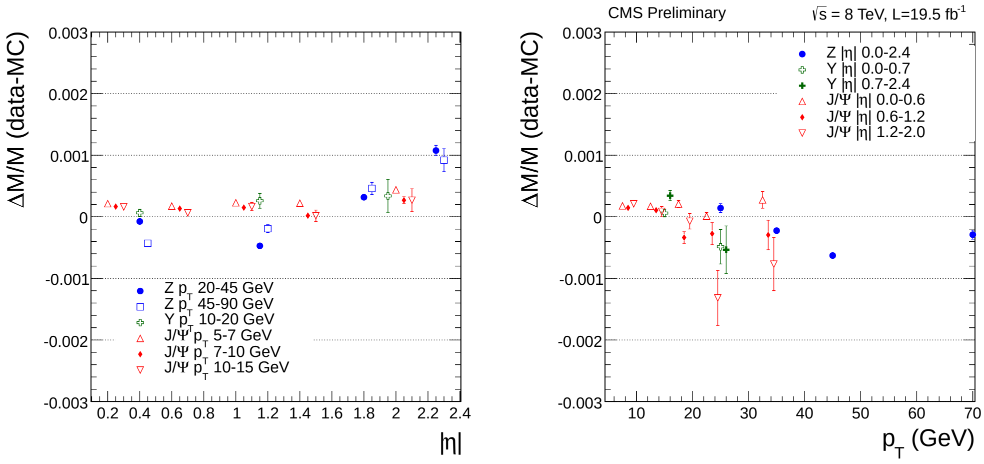
- *Case II*: analyses aiming to perform an absolute measurement of the momentum of the muon, e.g. not relying on a model for the expected spectrum. In this case, values of the observables built from the measured momenta, namely invariant masses, differing from standard references are indications of systematic errors. With respect to the previous case, the systematic error on p_{T} has an additional contribution that is evaluated using the simulation and taking as estimator the discrepancy between the values of the mass of the Z boson extracted from the spectra of the dimuons and the world average reported by the PDG:

  $$\frac{\delta p_{\mathrm{T}}^{\mathrm{abs}}}{p_{\mathrm{T}}} = \frac{M_{\mathrm{MC}} - M_{\mathrm{PDG}}}{M_{\mathrm{PDG}}} \tag{3.5}$$

  Again, the above relation between momentum and masses holds under the hypothesis that the systematic errors are fully correlated between the two muons.

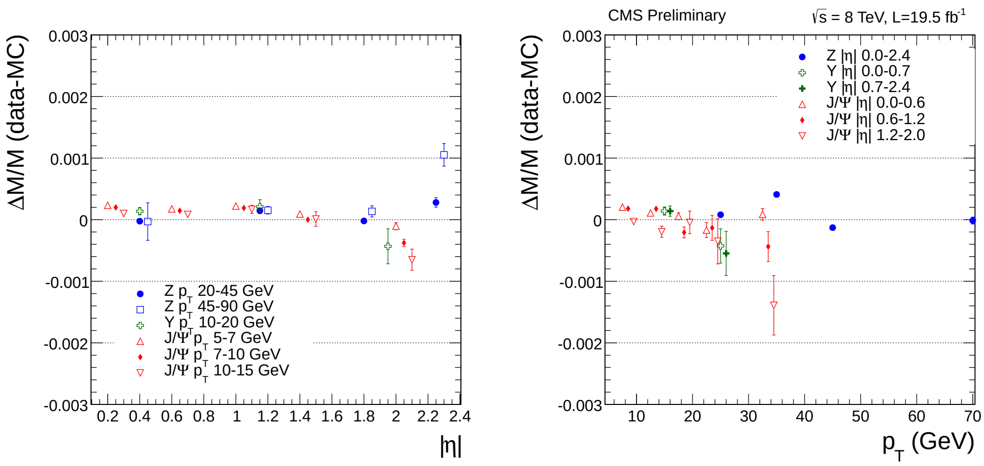
  We checked the behaviour of the fractional mass difference defined in equation 3.5 against p_{T} and $|\eta|$ of one of the two muons using simulated events reconstructed in realistic conditions (figure 3.32 - 3.35). Discrepancies with the PDG value are visible, especially at high p_{T} and high $|\eta|$. They have already been spotted discussing the "local validation" in section 3.4, where possible causes of such remaining bias were also presented. Considering also a possible bias in the mass measurement itself, which would affect such an absolute measurement, this contribution to the uncertainty on the p_{T} scale is estimated conservatevely in order to cover the maximum variation seen in all the points.

The summary of the systematics uncertainty associated to the muon momentum scale is presented in table 3.5. The values listed there refer explicitely to the 8 TeV MC reconstructed with the 53X release and its comparison with the data in the ReReco stream, as they represent so far the most accurate reconstruction of the events recorded by the CMS experiment. However, as one can easily infer by comparing figures 3.28, 3.29 and 3.30 with figure 3.31 and figures 3.32, 3.33 and 3.34 with figure 3.35, the same values are valid also for the other samples that have been studied in this section.



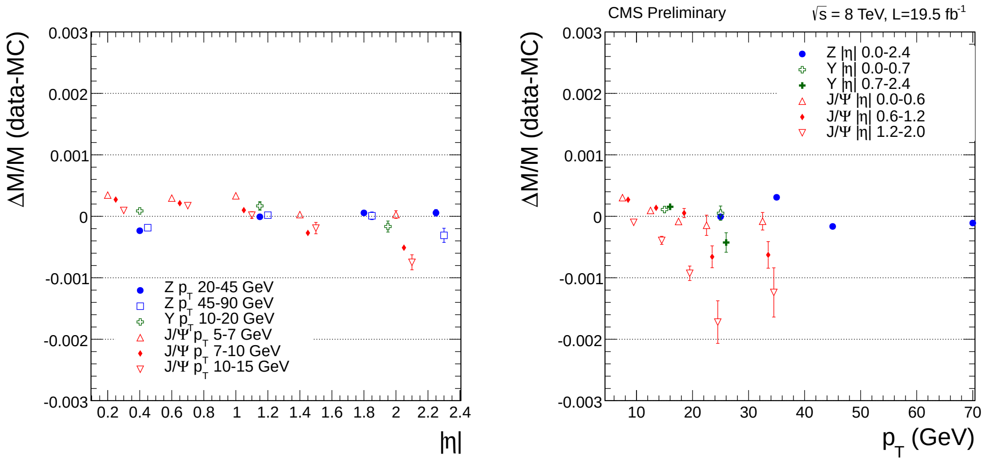
(a) (b)

Figure 3.28: The fractional difference between data and MC on the fitted resonance mass, as a function of muon $|\eta|$ 3.28a and p_{T} 3.28b, for the 2011 data and simulated samples reconstructed with the 42X release.



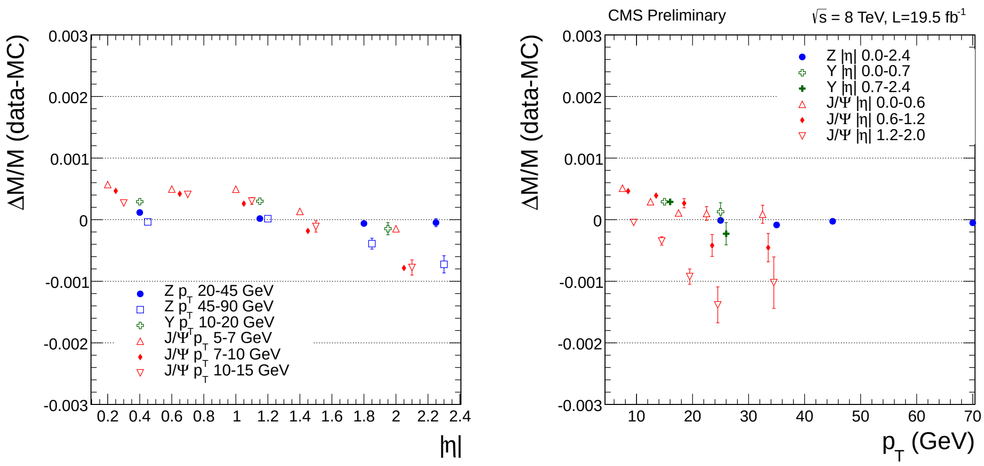
(a) (b)

Figure 3.29: The fractional difference between data and MC on the fitted resonance mass, as a function of muon $|\eta|$ 3.29a and p_{T} 3.29b, for the 2011 data and simulated samples reconstructed with the 44X release.



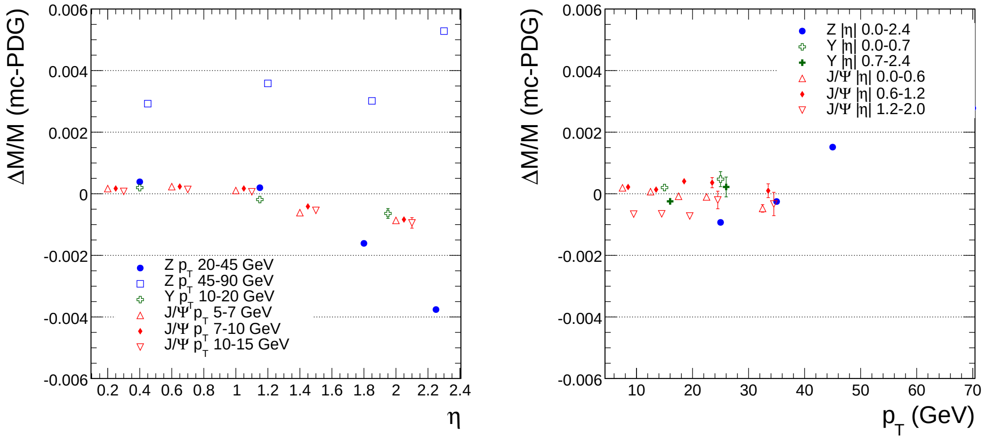
(a) (b)

Figure 3.30: The fractional difference between data and MC on the fitted resonance mass, as a function of muon $|\eta|$ 3.30a and p_{T} 3.30b, for the 2012 data in the Prompt reconstruction stream and the simulated sample in the 53X release.



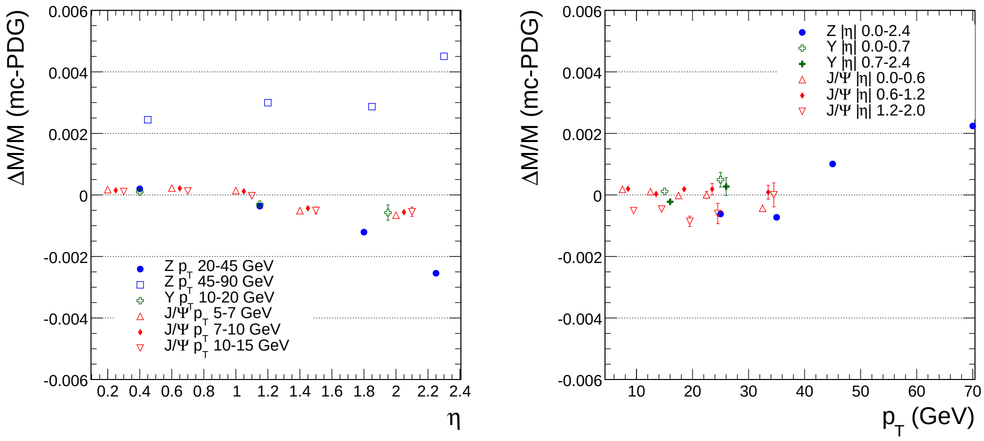
(a) (b)

Figure 3.31: The fractional difference between data and MC on the fitted resonance mass, as a function of muon $|\eta|$ 3.31a and p_{T} 3.31b, for the 2012 data in the ReReco reconstruction stream and the simulated sample in the 53X release.



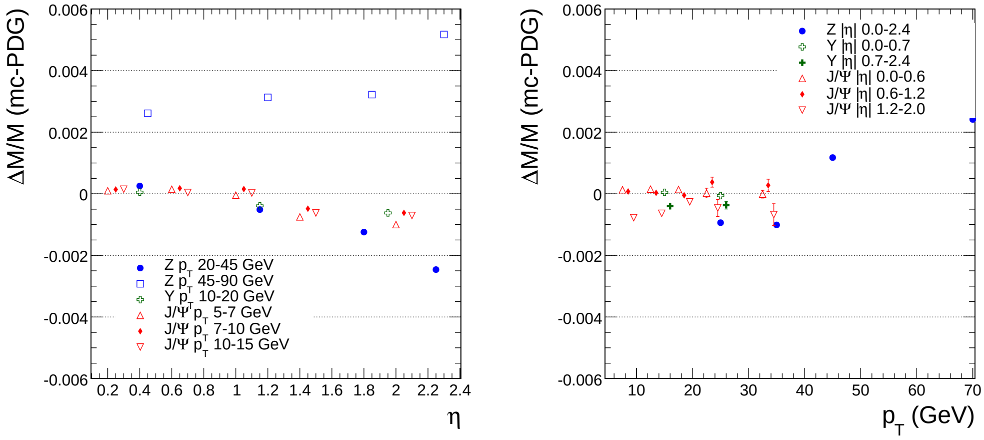
(a) (b)

Figure 3.32: The fractional difference between the MC and the nominal value (PDG) on the fitted resonance mass, as a function of muon $|\eta|$ 3.32a and p_{T} 3.32b, for the 2011 7TeV simulated sample reconstructed with the 42X release.



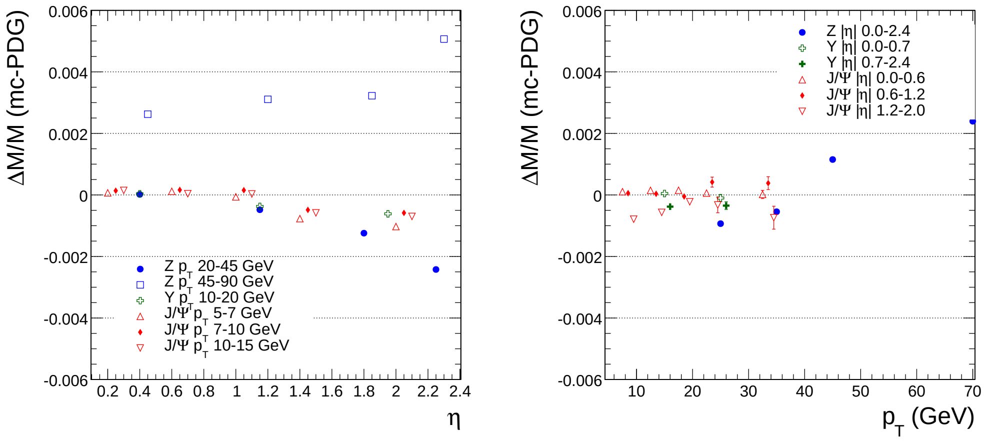
(a) (b)

Figure 3.33: The fractional difference between the MC and the nominal value (PDG) on the fitted resonance mass, as a function of muon $|\eta|$ 3.33a and p_{T} 3.33b, for the 2011 7 TeV simulated sample reconstructed with the 44X release.



(a) (b)

Figure 3.34: The fractional difference between the MC and the nominal value (PDG) on the fitted resonance mass, as a function of muon $|\eta|$ 3.34a and p_{T} 3.34b, for the 2012 7TeV simulated sample reconstructed with the 53X release.



(a) (b)

Figure 3.35: The fractional difference between the MC and the nominal value (PDG) on the fitted resonance mass, as a function of muon $|\eta|$ 3.35a and p_{T} 3.35b, for the 2012 8TeV simulated sample reconstructed with the 53X release.

Table 3.5: Estimated systematic uncertainties on the momentum scale of muons, post-correction, for the 2012 ReReco data. Second column corresponds to *Case I* systematics, while the sum of second and third columns corresponds to *Case II* systematics described in the text.

	DATA - MC	MC - PDG
p_T <45 GeV, 0< $\lvert\eta\rvert$ <1.5	0.05%	0.05%
p_T <45 GeV, 1.5< $\lvert\eta\rvert$ <2.4	0.10%	0.20%
p_T >45 GeV, 0< $\lvert\eta\rvert$ <2.0	0.05%	0.30%
p_T >45 GeV, 2.0< $\lvert\eta\rvert$ <2.4	0.10%	0.50%

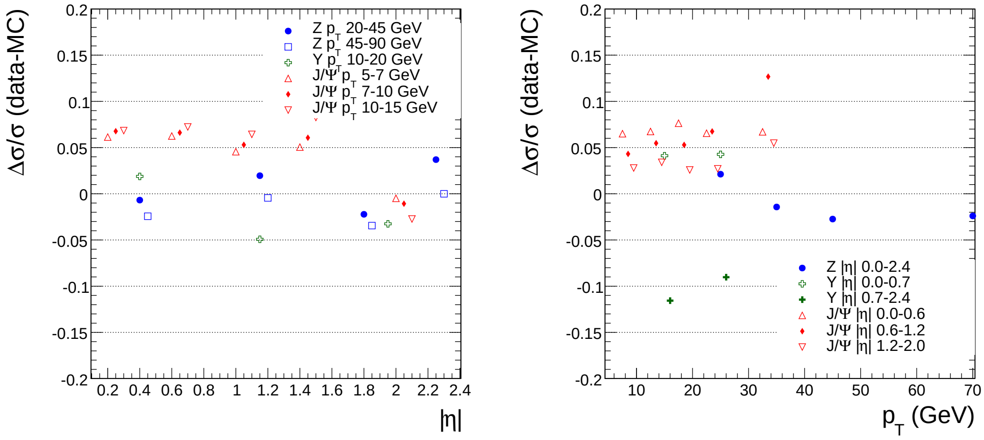
For what concerns the systematic uncertainty on the resolution on the p_{T} of a single muon, a common approach is to rely on the simulation, studying the spread of the distributions of the reconstructed momentum compared to the generated one. In this context, discrepancies between MC and data are usually taken into account as a source of systematic uncertainty. In case the uncertainties are correlated between the muons, one can make use of the following approximation:

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} \approx \frac{\sigma_M}{M} \tag{3.6}$$

The systematic uncertainty to the single muon resolution is defined as the fractional difference between mass resolution in data and MC:

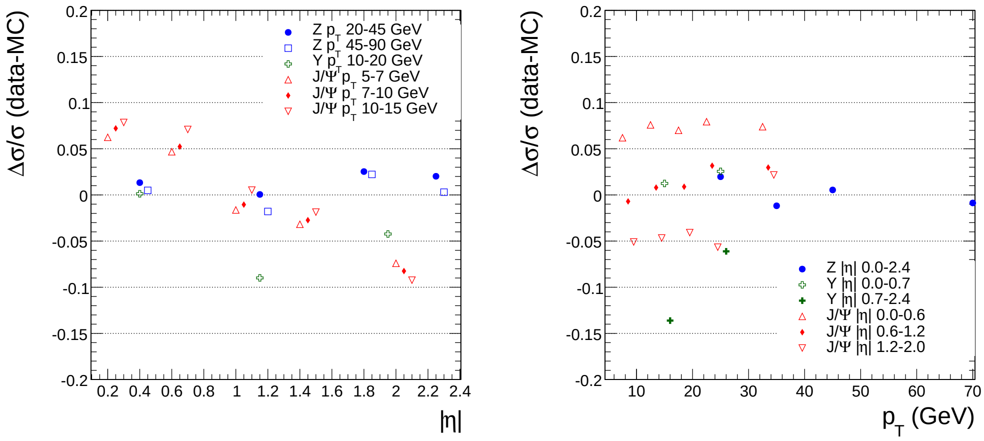
$$\frac{\delta\sigma_{p_{\mathrm{T}}}}{\sigma_{p_{\mathrm{T}}}} = \frac{\sigma_{\mathrm{data}} - \sigma_{\mathrm{MC}}}{\sigma_{\mathrm{MC}}} \tag{3.7}$$

The quantity defined in equation 3.7 is studied, after resolution corrections, against the p_{T} and $|\eta|$ of the muon, using the same binning as for the investigation of the scale systematics. The validation is shown in figures 3.36 - 3.39. With the exception of a few points for the low mass resonances (J/ψ, Υ), MC resolution always matches the one in the data at the level of ~5% or better. In particular, being the parameterization of the p_{T} resolution used in the smearing extracted from the Z samples, one can notice that the Z points exhibit in general a better data to MC agreement. This behaviour suggests the possibility to improve the smearing at low p_{T}s by extracting a dedicated p_{T} resolution parameterization from a sample enriched in low p_{T} muons.



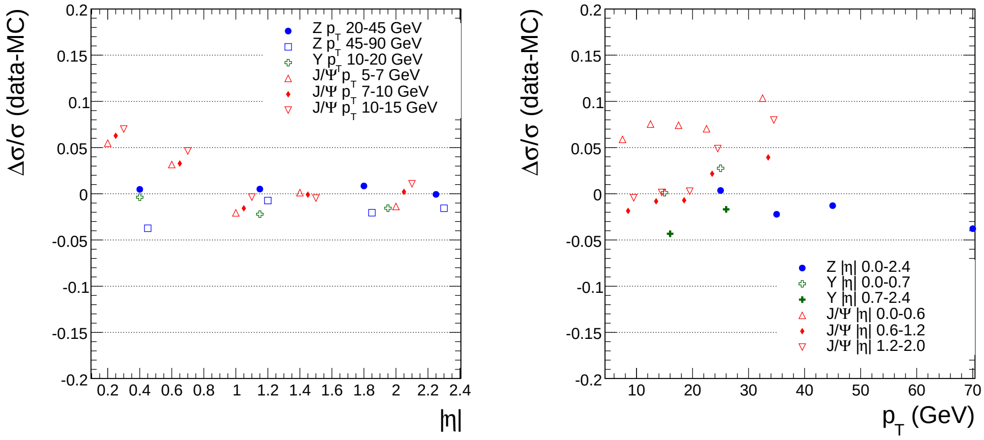
(a) (b)

Figure 3.36: The fractional difference between data and MC on the fitted mass resolution, as a function of muon $|\eta|$ 3.36a and p_{T} 3.36b, for the 2011 data and simulated samples reconstructed with the 42X release.



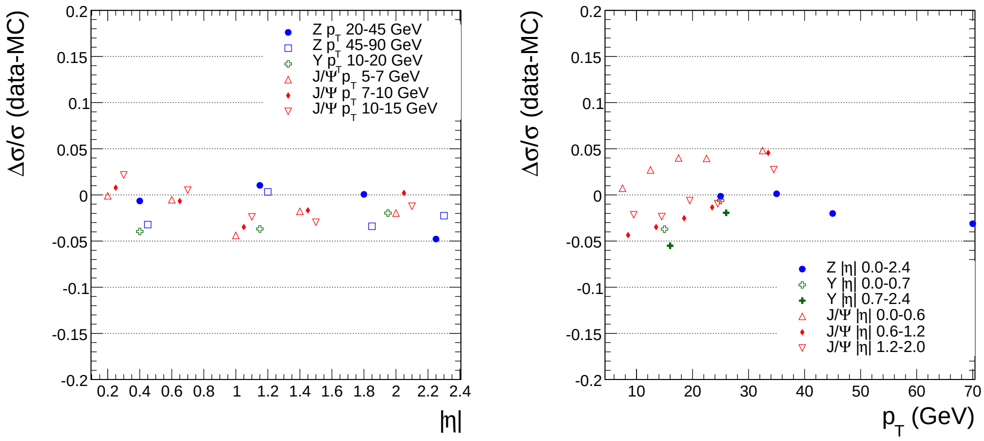
(a) (b)

Figure 3.37: The fractional difference between data and MC on the fitted mass resolution, as a function of muon $|\eta|$ 3.37a and p_{T} 3.37b, for the 2011 data and simulated samples reconstructed with the 44X release.



(a) (b)

Figure 3.38: The fractional difference between data and MC on the fitted mass resolution, as a function of muon $|\eta|$ 3.38a and p_{T} 3.38b, for the 2012 data in the Prompt reconstruction stream and the simulated sample in the 53X release.



(a) (b)

Figure 3.39: The fractional difference between data and MC on the fitted mass resolution, as a function of muon $|\eta|$ 3.39a and p_{T} 3.39b, for the 2012 data in the ReReco reconstruction stream and the simulated sample in the 53X release.

3.6 Impact of muon momentum corrections on physics analyses

Muon momentum corrections are applied in many physics analyses in CMS which rely on the precise measurement of the muon momentum.
In the searches for new resonances (like the SM Higgs), the main benefit comes from the improvement in the mass resolution which enhances the local significance of the excess driven by the signal: the narrower the peak, the more discriminating power against the continuum background is achieved.
In precision measurements involving muons, like the mass of the W boson, the estimation of the momentum scale and its uncertainty are crucial tasks which require dedicated efforts and calibrations.

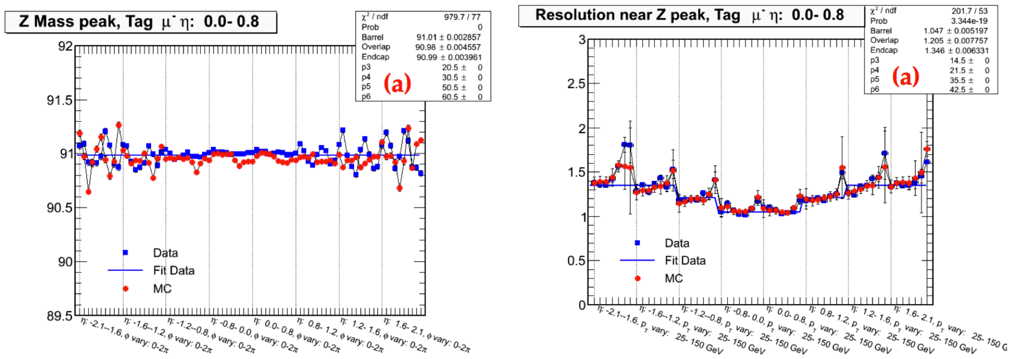
In the following two examples chosen among the SM Higgs searches are outlined, both using MuScleFit corrections for the final results.

3.6.1 Search for a SM Higgs boson in the $\mu^+\mu^-$ final state [33]

The search for a low-mass Higgs boson decaying to a muon pair is one of the most suitable case to apply Z-derived corrections, like the one provided with the MuScleFit package, since the phase space of the muon is pretty much the same, i.e. moderate-high p_{T}. Both scale and resolution corrections have been applied in CMS to extract the final results, using 2011 and 2012 data.
A validation has been performed using a sample enriched in $Z \rightarrow \mu\mu$ decays, divided according to different bins either for the "probe" muon and the "tag" muon. An example of such validation is shown in figure 3.40 for 2012 data and MC.
The muon momentum resolution has been computed in the MC and corrected by a factor to take into account discrepancies between data and MC. This factor, measured in (η, p_{T}) bins, in most cases is found to be compatible with one after MuScleFit resolution corrections are applied.
Finally, the effect of the muon momentum corrections has been checked on the final results and found to improve the expected limits by $\sim$5%.

3.6.2 Measurement of the properties of a Higgs boson in the four-lepton final state [15]

The search for a Higgs boson decaying to a pair of Zs in a four-lepton final state, with leptons being muons and electrons, is the most sensitive signature due to the low background contamination. It has been one of the most important channels leading to the discovery of a new boson with mass $\sim$ 125 GeV.
Thanks to the fully reconstructed final state, it is, together with $H \rightarrow \gamma\gamma$, one of the main decay in which the mass can be measured. The corrections to the momentum scale and resolution, both for muons and electrons, were crucial in this measurement. The scale systematics was investigated with a validation similar to the one described in section 3.5, in order to assess the residual data-MC discrepancy. As an additional



(a) (b)

Figure 3.40: The Z mass (3.40a) and the dimuon mass resolution (3.40b) extracted from a fit to the dimuon spectrum around the Z resonance, for negative tag muon in the region $|\eta| < 0.8$ and varying η, ϕ for the probe muon. Blue points are from data samples, red points from simulation.

check, Z decays into 4 leptons have been studied with the same selection used in the Higgs boson analysis. This process is a *standard candle* because its peak in the 4ℓ mass distribution can be used for a direct validation of the understanding of lepton scale and resolution in the phase space just next to the decay of the newly discovered Higgs boson, with five times more events, according to the SM.
Both results are shown in figure 3.41 for the muons, after MuScleFit corrections. The residual mismatch between data and MC in the momentum scale is well within $\sim 0.1\%$ for any value of the muon p_{T}, as can be seen in figure 3.41 (left). This result is in agreement with the validation shown in figure 3.31b. From figure 3.41 (right) one can sees how the residual fitted scale in the data is compatible with zero, and the mass resolution matches the one in the MC within the errors.

The mass measurement in the 4ℓ channel was performed with a simultaneous fit to the signal-background discriminant $\mathcal{D}_{\mathrm{bkg}}^{\mathrm{kin}}$, the invariant mass of the four-lepton system $m_{4\ell}$ and the per-event mass errors $\mathcal{D}_{\mathrm{mass}}$, where $\mathcal{D}_{\mathrm{mass}} \equiv \sigma_{m_{4\ell}}/m$ and $m = m_{4\ell}$ for backgrounds and data and $m = M_H$ for signal.
The final result is:

$$M_H = 125.6\,^{+0.5}_{-0.4}\,(\mathrm{stat})\ \pm 0.2\,(\mathrm{syst.})\ \mathrm{GeV} \tag{3.8}$$

In the 4μ channel the contribution to the uncertainty coming from the scale systematic is 0.07%, while the systematic on the resolution is negligible.

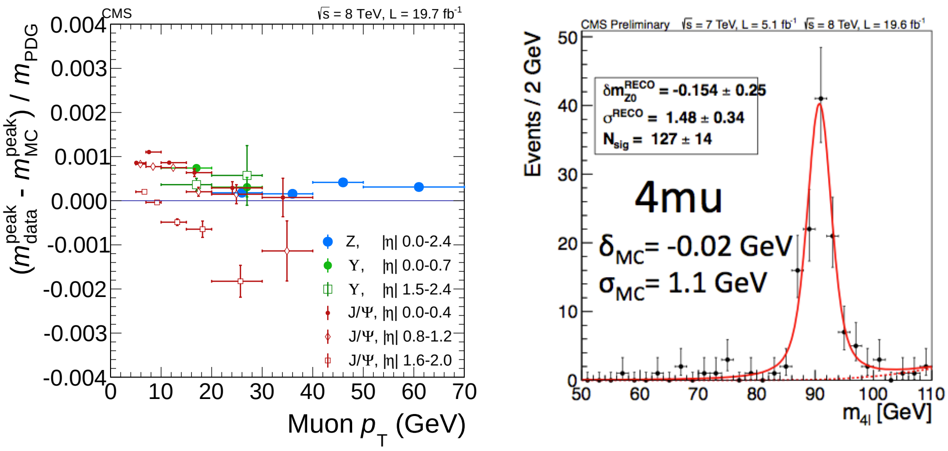


Figure 3.41: Validation of the muon momentum scale in the $H \to ZZ \to 4\ell$ analysis. Figure on the left shows the discrepancy in the scale between data and MC measured in events with dimuon decays of the Z, J/ψ and Υ. Figure on the right shows the fit to the 4μ mass spectrum in the region of the Z.

Chapter 4

Observation of a Higgs boson in the $H \to ZZ \to 4\ell$ channel and measurement of its properties

Abstract

The channel in which a Higgs boson decays to a pair of Z bosons, which subsequently decay into a fully leptonic final state provides an extremely clean final state, a low background rate and a complete reconstruction of the Higgs kinematics. For these reasons it has been a "riding horse" towards the discovery of a new Higgs-like boson with mass around 125 GeV, announced by the ATLAS and CMS collaborations on 14th of July 2012 [11, 12]. In this chapter, the analysis with the full statistics of the Run 1 of the LHC is presented, including the reconstruction of the physics objects like leptons and jets, the event selection, the signal efficiency and the statistical procedure applied to extract the results.

4.1 Introduction

The analysis in the $H \to ZZ^{(*)} \to 4\ell$ channel, where $\ell = e, \mu$, has the advantage that the four-momentum of each object of the final state daughters of the Higgs is precisely reconstructed. This feature, together with the low contamination of backgrounds coming from SM processes, makes it the most sensitive to the search of the Higgs boson over a large mass range. While most of the effort has been focused on measuring the properties of the new boson discovered at a mass of 125 GeV, like the mass, the width and the tensor structure, this process has been demonstrated to be powerful also for the search of other Higgs-like states.

This chapter is organized as follows. In section 4.2 the datasets used in the analysis are presented. In 4.3 the reconstruction of the physics objects (leptons, photons, jets) is described. In 4.4 the event selection and categorization is introduced. In section 4.5 the techniques to estimate the yields of background processes are explained in details. The kinematic discriminants used in the analysis are described in section 4.6. Section 4.7 presents the yields and kinematic distributions of selected events. The details of the fit and its inputs are discussed in section 4.8. In section 4.9 the main sources of systematic uncertainties are listed. The final section, 4.10, is devoted to the presentation of the results, namely: the exclusion limits in the whole mass range, the significance of the excess at low mass, the signal strength, the measurement of the mass, width and spin-parity of the new resonance.

4.2 Datasets

4.2.1 Experimental data

The data samples used in this analysis have been collected by the CMS detector during 2011 and 2012: they correspond to an integrated luminosity of $\mathcal{L} = 5.1\,\mathrm{fb}^{-1}$ at 7 TeV and $\mathcal{L} = 19.7\,\mathrm{fb}^{-1}$ at 8 TeV respectively. A standard selection is applied which requires high quality data with a good functioning of the sub-detectors.
The analysis relies on primary datasets selected by the combination of different High Level Triggers (HLT). These include *DoubleMu* and *DoubleEle* triggers, which requires two muons or electrons, the cross trigger *MuEG*, requiring one muon and one electron, which was introduced only in 2012 to improve the efficiency in the $2e2\mu$ channel, and finally *TripleEle* triggers requiring three electrons.
The 4μ candidates are selected from the DoubleMu datasets, the $4e$ from the DoubleEle and the TripleEle datasets, while the $2e2\mu$ candidates come from the logical OR of all the datasets.

4.2.2 Simulated samples

The Monte Carlo (MC) samples are used to optimize the selection, evaluate acceptance corrections, in the background estimation and to assess the systematic uncertainties. The SM Higgs boson signal samples, as well as many other SM background processes, have been generated with programs based on the state-of-the-art theoreti-

cal calculations. Simulated events are then processed through a detailed simulation of the CMS detector based on GEANT4 [27] and are reconstructed using the same algorithms used for the data. They are used to optimize the selection, evaluate acceptance corrections, in the background estimation and to assess systematic uncertainties.
The most important MC samples used in the analysis are listed in table 4.1.
For the signal, samples are generated for several mass points between 110-1000 GeV, and for the main production modes. The dominant background is due to the SM ZZ or $Z\gamma^*$ productions, via $q\bar{q}$ annihilation or gluon fusion, which will be referred to as ZZ in the following. Smaller contributions arise from Z+jets and $t\bar{t}$ events, where there are two additional leptons coming from heavy quark decays or from the misidentification of jets. Other di-boson production, like WZ, WW may give small contributions in the earliest stages of the selection.
The simulation includes overlapping pp interactions (pileup) matching the observed number of interactions per LHC beam crossing observed in the data. The average number of measured pileup interactions are 9 and 21 for 7 TeV and 8 TeV datasets respectively.

Table 4.1: The main MC samples used in the analysis, with their cross sections for 7 TeV and 8 TeV.

Process	**MC generator**	$\sigma_{(N)NLO}$	
		7 TeV	8 TeV
SM Higgs $H \to ZZ \to 4\ell$			
$gg \to H$ (ggH)	POWHEG	[1-20] fb	[1.2-25] fb
$qq \to qqH$ (VBF)	POWHEG	[0.2-2] fb	[0.3-2.5] fb
ZZ continuum			
$q\bar{q} \to ZZ \to 4\ell$	POWHEG	66.09 fb	76.91 fb
$q\bar{q} \to ZZ \to 2\ell 2\ell'$	POWHEG	152 fb	177 fb
$gg \to ZZ \to 4\ell$	gg2VV	1.74 fb	4.81 fb
$gg \to ZZ \to 2\ell 2\ell'$	gg2VV	3.48 fb	12.03 fb
$t\bar{t}$			
$t\bar{t} \to \ell^+\ell^-\nu\bar{\nu}b\bar{b}$	POWHEG	17.32 pb	23.64 pb
$Z/W + \text{jets}$			
W+jets	MadGraph	31314 pb	36257 pb
Z+jets, $m_{\ell\ell} > 50\,\text{GeV}$	MadGraph	3048 pb	3503 pb
Z+jets, $10 < m_{\ell\ell} < 50\,\text{GeV}$	MadGraph	805 pb	915 pb

4.3 Physics Objects

4.3.1 Leptons

Electron reconstruction and identification

Electron candidates are required to be within the geometrical acceptance defined by $p_{\mathrm{T}} > 7\,\mathrm{GeV}$ and $|\eta| < 2.5$.
The electron reconstruction in CMS [34, 35, 36, 37] combines the informations of the ECAL and the tracker measurements. In the ECAL, the electron candidates are formed from arrays of energy clusters in the ϕ direction, called *superclusters*, which recover the bremsstrahlung photons emitted in the tracker and part of the collinear final state radiation (FSR). Superclusters are also used to identify hits in the track layers in order to initiate the reconstruction of electron tracks. At low p_{T} of the electron and in the transition between the ECAL barrel (EB) and the ECAL endcap (EE), a complementary approach based on the track seeds is also used to improve the reconstruction efficiency. Both type of tracks, either the one reconstructed with the "outside-in" and the "inside-out" approaches, are built using the Gaussian Sum Filter (GSF) algorithm, which accounts for the energy loss due to bremsstrahlung. Electron candidates are selected with loose cuts on the track-supercluster matching variables, in order to keep the highest possible efficiency and at the same time reject part of the background ("fake" electrons). Additional requirements are implemented to remove electrons coming from photon conversions inside the tracker.

The energy measured in individual crystals is corrected with several procedures, as described in [38, 39].
The supercluster energy is further corrected for the imperfect containment of the electromagnetic shower of the clustering algorithm, the electron energy not deposited in the ECAL and leakages due to gaps between crystals or modules, using a regression technique based on a Boosted Decision Tree (BDT). This correction brings an improvement of the order of 30 % on the mass resolution, with respect to a more traditional approach based on a parameterized energy response obtained from MC, as estimated in simulation with $H \to 4e$ decays.

The magnitude of the electron momentum is estimated combining the track momentum estimate and the corrected supercluster energy, using a multivariate regression function, which takes as inputs the two measurements, their uncertainties and the amount of bremsstrahlung.
The precision in the determination of the electron momentum is dominated by the tracker for $p_{\mathrm{T}} < 15\,\mathrm{GeV}$ and by the ECAL at higher momenta, as it is shown in figure 4.1 (right). In figure 4.1 (left) the effect of the energy corrections and tracker-ECAL combination is shown in the simulation, in a sample of $H \to 4e$ decays.

Muon reconstruction and identification

Muon candidates are require to have $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ and $|\eta| < 2.4$.
The muon reconstruction combines the information from both the tracker and the muon system. Tracks in the two detectors are matched either starting from the segments in the muon system ("outside-in") or starting from the track reconstructed

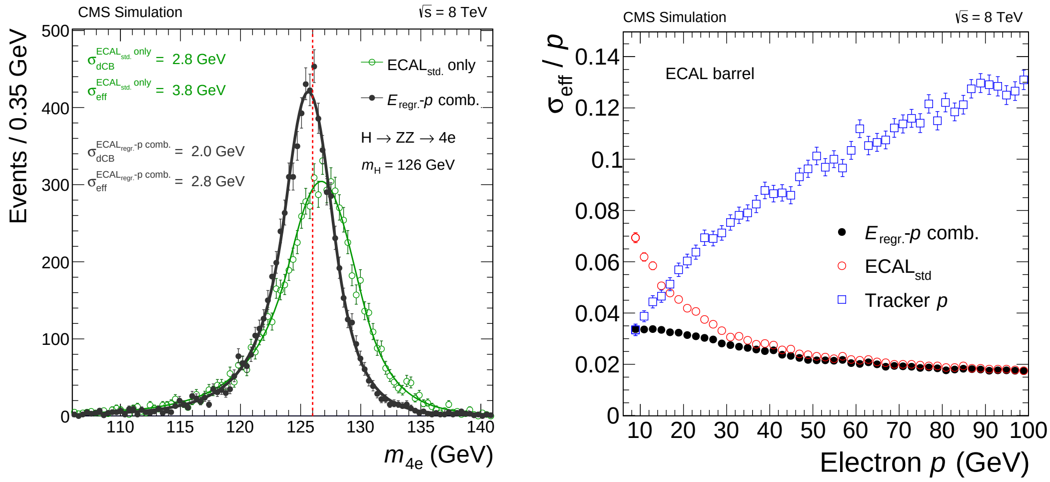


Figure 4.1: (left) Mass distribution in simulation for a sample of $H \to 4e$ decays, with $m_H = 126\,\mathrm{GeV}$, comparing the ECAL-only electron momentum estimation (green points) and the tracker-ECAL combination as used in the analysis (black points). (right) The effective relative electron momentum resolution as a function of the electron momentum, for electron in the EB.

in the inner tracker ("inside-out"): in the case where the two can be combined in a single fit, the reconstructed muon is called a *global muon*. Soft muons may have not enough energy to pass through the whole muon detector and leave signal only in one or two muon stations, where a station is made by multiple detection planes between two iron layers. Nevertheless, if a track in the inner tracker is matched to at least one segment a *tracker muon* is reconstructed. At large transverse momenta, i.e. $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, the global-muon fit improves the momentum resolution compared to the tracker-only. The analysis described here makes use of both tracker muons and global muons.

The muons are required to satisfy the *Particle Flow* (PF) muon identification [31], which applies minimal requirements on the track components in the muon system and takes into account a matching with small energy deposits in the calorimeters.

The knowledge of the true position of the tracker modules in the detector affects sensibly the measurement of the p_{T} of the muons, as it has been highlighted in chapter 3. A statistical accuracy of better than $\leq 10\mu m$ is achieved on the position of the tracker sensors thanks to the alignment procedure, which is carried on mainly with cosmic and collision tracks, making use of vertex information and resonance mass constraints. The alignment of the muon system contributes to a lesser extent to the momentum measurement.

The resolution for the muons used in the analysis is dominated by the multiple coulomb scattering in the detector material and is about 2% in the barrel and 6% in the endcaps.

Lepton isolation and vertex compatibility

Lepton candidates are required to be well isolated from other particles in the event, in order to discriminate leptons originating from decays of the Z boson, which constitutes the signal, from leptons coming from other processes, which are typically close to jets. The lepton isolation is computed relative to its transverse momentum and is defined by:

$$R^{\ell}_{\mathrm{Iso}} = \frac{\left(\sum p_{\mathrm{T}}^{\mathrm{charged}} + \max\left[0, \sum p_{\mathrm{T}}^{\mathrm{neutral}} + \sum p_{\mathrm{T}}^{\gamma} - p_{\mathrm{T}}^{\mathrm{PU}}\right]\right)}{p_{\mathrm{T}}^{\ell}} \tag{4.1}$$

In equation 4.1 the sums extend to all particle candidates which are reconstructed by the PF algorithm as charged hadrons, neutral hadrons and photons, inside a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$. Photons which are identified as FSR (see section 4.3.2) are removed from the sum.
The leptons candidates are required to satisfy $R^{\ell}_{\mathrm{Iso}} < 0.4$: this cut has been chosen to maximize the discovery potential.
The term $p_{\mathrm{T}}^{\mathrm{PU}}$ represents the contribution of neutral hadrons coming from pile up interactions. For electrons, it is estimated using the FASTJET [40, 41, 42] technique and is defined by:

$$p_{\mathrm{T}}^{\mathrm{PU}} \equiv \rho \times A_{\mathrm{eff}} \tag{4.2}$$

In equation 4.2:

- ρ is the median of the energy-density distribution in the area of any jet in the event reconstructed by the k_{T} clustering algorithm with a distance parameter $D = 0.6$, with $p_{\mathrm{T}}^{\mathrm{jet}} > 3\,\mathrm{GeV}$ and $|\eta^{\mathrm{jet}}| < 2.5$
- A_{eff} is the geometrical area of the isolation cone corrected for the different PU depositions as a function of the electron pseudorapidity.

For the muons the contribution to the isolation from PU is defined as:

$$p_{\mathrm{T}}^{\mathrm{PU}} \equiv \sum_i 0.5 \times p_{\mathrm{T}}^{\mathrm{PU,i}}, \tag{4.3}$$

where the sum is extended to all the charged hadrons originating from PU vertices and 0.5 is to correct for the different fraction of charged and neutral hadrons in the isolation cone.

Leptons are also required to originate from the primary vertex, in order to suppress the contribution of non-prompt leptons from in-flight decays of hadrons. This is achieved by requiring $\mathrm{SIP}_{3\mathrm{D}} < 4$, where we define the significance of the impact parameter $\mathrm{SIP}_{3\mathrm{D}}$ as the ratio between the 3-dimensional impact parameter of the lepton track with respect to the primary vertex divided by its uncertainty.

Lepton efficiency

The efficiency of the reconstruction, identification and selection ($SIP_{3\mathrm{D}}$ and isolation) of the leptons is measured using the "tag and probe" method [43]. This technique makes use of inclusive samples of $Z \to \mu\mu$ (for $p_{\mathrm{T}}^{\ell} > 15\,\mathrm{GeV}$) and $J/\psi \to \mu\mu$ (for $p_{\mathrm{T}}^{\ell} < 15\,\mathrm{GeV}$) decays. These samples are obtained by combining a mass constraint

from a pair of basic objects, like tracks for muons or clusters in the ECAL for electrons, with a "tight" lepton selection applied on one of the two objects, called the "tag", in order to ensure a sufficient purity. The other leg (the "probe") is used to measure the efficiency of a certain cut, defined as the ratio of the number of probes passing that cut over the total number of probes. The yields are computed by means of a fit to the mass spectrum around the resonance.
The same measurement is done both in data and in simulation, and scale factors are extracted in bins of lepton p_{T} and η and used in the analysis to match the efficiency in the simulation to the one measured in the data. In general data and MC agree at the level of 3% or better, with the exception of electron in the range $p_{\mathrm{T}}^{e} < 15\,\mathrm{GeV}$.
For the muons, the tracking efficiency is also considered in the scale factors. This correction has always been negligible with respect to the others, except for the highest PU data-taking periods, where a loss in efficiency has been measured due to the different tolerances used in the tracking algorithm.
The electron and muon efficiencies measured in the 8 TeV dataset are shown in figure 4.2.

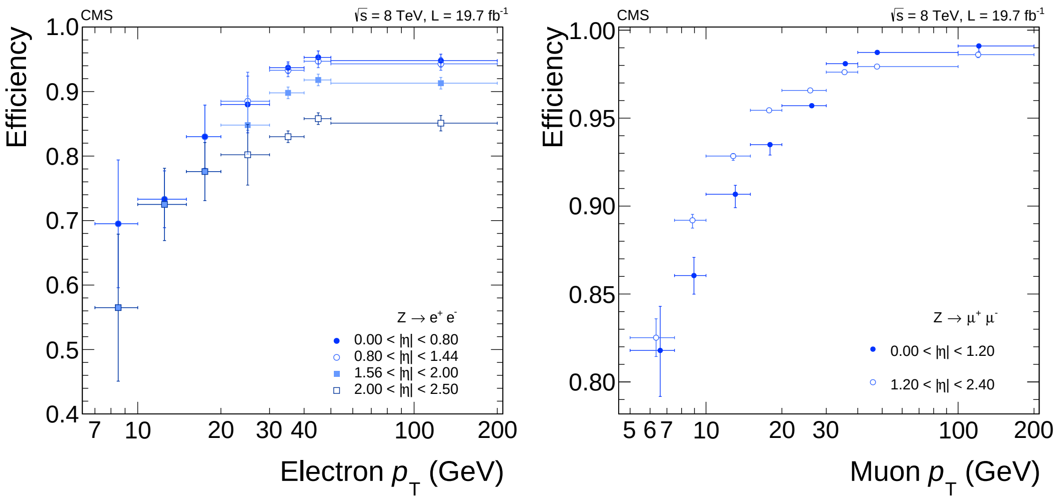


Figure 4.2: The efficiency for the reconstruction, identification and selection of electrons (left) and muons (right) is shown as a function of the lepton p_{T}, for the 8 TeV data.

Lepton momentum scale and resolution

The measurement of the lepton momentum is one of the most important tasks in this analysis, which strongly relies on the accurate determination of the invariant mass and the p_{T} of the 4-lepton system.
The approach used is different for muons and electrons, as their reconstruction involves different subdetectors. In both cases the scale and resolution are extracted from the simulation, which is based on the best knowledge of the detector. In particular, the response of the ECAL and the alignment of the tracker and muon detectors are realistically simulated in the reconstruction of the MC samples. Nevertheless, discrepancies between data and simulation in the scale and resolution remain and therefore calibrations are implemented for muons and electrons. The residual difference after

calibration is taken into account as a systematic uncertainty.

For the electrons, the scale calibration is performed in two steps. First, a set of corrections is extracted by comparing the Z boson mass shift in data and MC, in different categories depending on the electron pseudorapidity and the amount of bremsstrahlung. Also a time dependency is considered, to take into account the loss of transparency in the crystals. A "linearity correction" to the momentum scale is applied in a second step, binned in the p_{T} of the electron, to account for the p_{T}-dependent difference between data and MC.
The energy of the electron is then smeared using a gaussian factor to match the resolution observed in the data.

For the muons, a comprehensive and detailed description of the calibration of the scale and the resolution was given in chapter 3.
The muon p_{T} scale and resolution are extracted with a likelihood fit to the dimuon mass distribution around the Z resonance, which uses a reference model convolved with a gaussian for the signal and an exponential for the background. Corrections for the scale are extracted as analytic functions in both data and MC. The resolution corrections are implemented as gaussian smearing factors to the muon momentum, calculated using the information extracted from the fit.

The lepton momentum scale after the calibration is studied using dimuon decays of J/ψ, Υ and Z divided in several bins based on the lepton $|\eta|$ and p_{T}, in order to cover the full momentum range relevant in the $H \to ZZ \to 4\ell$ analysis. In each bin the position of the resonance peak is extracted from a fit using a Breit-Wigner (BW) function convolved with a Crystal Ball (CB) for the signal model and polynomials for the background. The offset between the mass measured in the data and in the MC is shown in figure 4.3.
For the electrons, the residual mismatch is within 0.2% in the barrel and up to 0.3% in the endcaps. For the muons, the agreement between data and MC is within 0.1% in almost all the momentum range, with the exception of the most boosted and forward categories in the J/ψ events, which is anyway an atypical configuration for the muons used in the analysis and therefore is not considered relevant.

Similarly, the width of the peak due to instrumental resolution (σ_{eff}) is compared between data and simulation. The agreement is well within 3% for the electrons and 5% for the muons, as shown in figure 4.4.

4.3.2 Final state radiation

Events with a Z boson decaying to leptons may be accompanied by the emission of final state radiation, i.e. photons which are emitted almost collinear to the lepton direction and carrying part of its momentum. An algorithm to identify such events has been developed, in order to fully reconstruct the $Z \to \ell^+\ell^-\gamma$ decay and remove the photon from the isolation cone of the corresponding lepton. The criteria to select FSR photon candidates have been tuned mainly to keep low the contamination from background photons.
The low energy photons are reconstructed in CMS with the PF algorithm [44], their energies and directions are studied in data and simulation using $\pi^0 \to \gamma\gamma$ events.

The final state radiation has usually a harder spectrum with respect to other

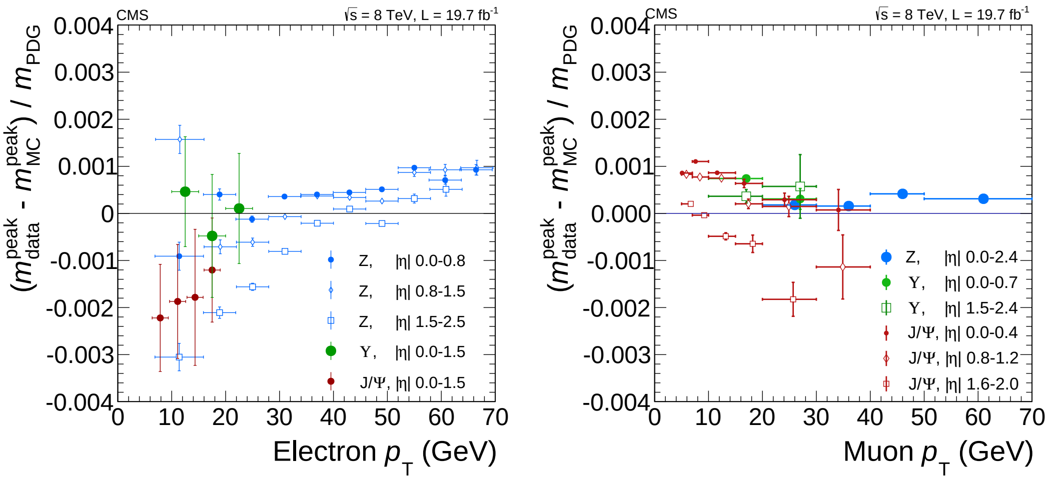


Figure 4.3: Relative difference in the peak position of the resonance between data and simulation in dilepton J/ψ, Υ and Z events, as a function of the lepton p_{T}, for electrons (left) and muons (right).

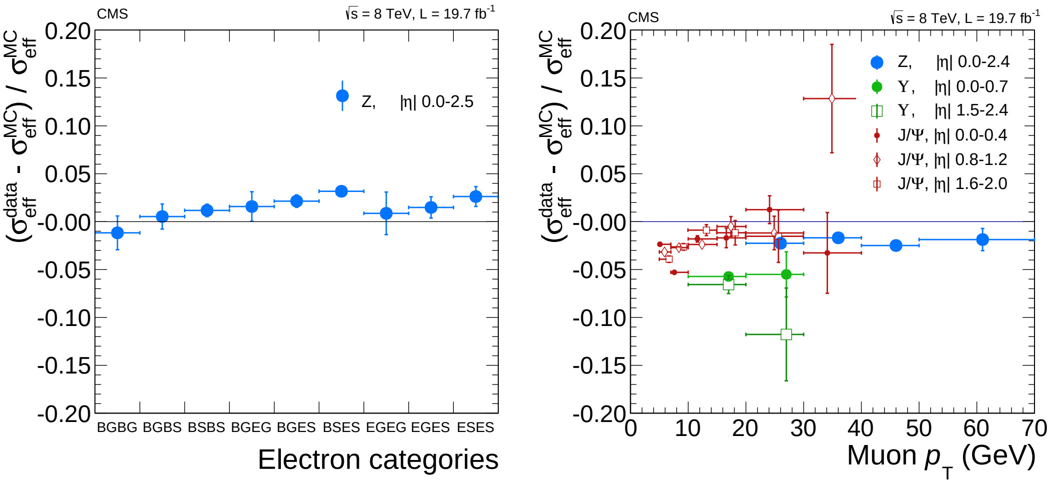


Figure 4.4: Relative difference in the instrumental width of the resonance between data and simulation in dilepton J/ψ, Υ and Z events, as a function of the lepton p_{T}, for electrons (left) and muons (right).

sources of spurious photons, like PU interactions or initial-state radiation and its direction is almost collinear to the direction of the lepton from which is emitted. Therefore the FSR photon candidate is required to have $p_{\mathrm{T}}^{\gamma} > 2$ GeV if it is found within a cone $\Delta R < 0.07$ around the lepton. If $0.07 < \Delta R < 0.5$ then the photon must have $p_{\mathrm{T}}^{\gamma} > 4$ GeV and be isolated. The isolation variable $R_{\mathrm{iso}}^{\gamma}$ is defined in a similar way as in equation 4.1 except for the PU neutral contribution, which is not removed here, and isolation cone which has size $\Delta R = 0.3$. Isolated photons satisfy $R_{\mathrm{iso}}^{\gamma} < 1$.

In case there are multiple FSR photon candidates associated to a Z boson, only the highest-p_{T} one is considered, if it has $p_{\mathrm{T}}^{\gamma} > 4$ GeV, otherwise the one closest to any Z daughter lepton is chosen. When forming the Z candidates (details on the selection are in section 4.4), FSR photons are added to the candidate only if $M_{\ell^+\ell^-\gamma}$ is closer to the nominal Z mass with respect to $M_{\ell^+\ell^-}$. The photon is anyway discarded if $M_{\ell^+\ell^-\gamma} > 100$ GeV. With the selection described above, about 1.5%, 4.6% and 9% of the events in the $H \to 4e, 2e2\mu, 4\mu$ MC samples are affected by the FSR recovery algorithm. The lower rate in the electron channels is expected because part of the collinear photon emission is already included in the supercluster energy. The rates have been measured also in data selecting inclusive Z decays and have been found in agreement with the simulation. A gain of $\sim$ 3%, 2% and 1% in the efficiency is expected for the selection of $H \to 4\mu, 2e2\mu, 4e$ events, mainly thanks to the removal of the FSR photon from the isolation cone of the lepton. The four-momentum of the photon is added to the one of the lepton for the computation of kinematic variables.

4.3.3 Jets

Jets are used in the analysis to discriminate between the different production mechanisms, in particular gluon fusion from vector boson fusion (VBF) or associated production VH, where V is a Z or W boson decaying hadronically.
Jets are reconstructed with the anti-k_{T} clustering algorithm [45], with distance parameter $D = 0.5$, using all PF candidates in the event. Jet energy corrections are applied as a function of the jet η and p_{T} [46]. In particular, a correction is applied to subtract the contribution from PU, based on the jet-area method [40, 41]. Jet candidates are required to have $p_{\mathrm{T}}^{\mathrm{jet}} > 30$ GeV, $|\eta^{\mathrm{jet}}| < 4.7$ and to be far from each of the four leptons of the Higgs candidate and from photon identified as FSR (see 4.3.2), requiring $\Delta R^{\ell,\mathrm{jet}} > 0.5$.

In order to discard jets coming from PU interactions, within the tracker acceptance, for the charged particles, only PF candidates compatible with the primary vertex are used in the reconstruction of jets. In the full acceptance, a multivariate technique is used to discriminate between PU jets and jets arising from the main interaction, which takes as input the jet shapes, multiplicity of charged and neutral components and the fraction of momentum carried by the hardest component [47].

4.4 Event selection and categorization

4.4.1 Best candidate selection

The event selection is based on the reconstruction of four well-isolated leptons compatible with the primary vertex and the recovery of FSR described above. In the first step of the selection flow, a Z boson candidate is formed with pairs of leptons of opposite charge and same flavour. Among all such pairs, the one with mass closest to the nominal Z-boson mass is retained and denoted Z_1. Its mass is required to satisfy $40 < m_{Z_1} < 120$ GeV.
Then, with the remaining leptons, additional $\ell^+\ell^-$ pairs are formed in order to build the so-called Z_2. If more than one Z_2 candidates are available, the one with the highest scalar sum of leptons' p_T is retained. The chosen Z_2 is required to have $12 < m_{Z_2} < 120$ GeV. For $m_\mathrm{H} > 180$ GeV both Zs are on-shell, thus this cut has been chosen to maximize the efficiency for the search for a low mass Higgs boson, while keeping under control the background coming from $Z +$ low mass resonances. Among the four leptons of the Higgs candidate, at least one is required to have $p_\mathrm{T} > 20$ GeV and another one to have $p_\mathrm{T} > 10$ GeV. These thresholds ensure that leptons are on the "plateau" of the trigger efficiency.
To further reduce the contamination of non-prompt leptons, any opposite sign lepton pairs (irrespective of the flavour) is required to have $m_{\ell^+\ell^-} > 4$ GeV.
Finally, the phase space for the search of the SM Higgs boson is reduced by applying $m_{4\ell} > 100$ GeV. For the measurements of the single-resonant $Z \to 4\ell$, used as standard candle, the phase space is defined by $m_{4\ell} > 70$ GeV.

4.4.2 Signal efficiency

The selection efficiency for the signal is estimated in the simulation separately for each final state and for each production mechanism. It is defined as the ratio of the number of events reconstructed and selected over the number of events generated. As described in section 4.3.1, the simulation has been corrected for the lepton efficiency by applying a weight to each event according to the p_T and η of each lepton.
The efficiency has been studied as a function of the Higgs mass hypothesis and interpolated between the points corresponding to the available full simulated MC samples.
The results are shown in figure 4.5 for the three different final states, for 7 and 8 TeV and for the different production mechanisms. For the associated production (WH, ZH and ttH) samples only up to 200 GeV have been produced, since the contribution to the total cross section becomes negligible above [1].
The selection efficiency increases rapidly up to $m_H \simeq 2m_Z$, where both the Z bosons become on-shell, then flattens. The additional smooth rise at high masses is due to the increased acceptance. This effect can be inspected by looking at the efficiency within the acceptance, as it shown in figure 4.6 for the gluon fusion process at 8 TeV: here the rise at high masses disappears while the steep rise at low mass is still present due to Z bosons being more and more on-shell.
The plots in figure 4.5 show a good agreement between the different production

[1] Cross section for the three processes have been computed up to 300 GeV.

mechanisms. They also reflect the higher efficiency in reconstructing muons with respect to electrons in CMS at this p_{T} regime. Some differences are present between 7 TeV and 8 TeV samples and are due to different reconstruction and identification of the physics objects and different PU conditions between the two datasets.

4.4.3 Event categorization

For a Higgs boson with a mass of 125 GeV about 88% of the cross section comes from the gluon fusion process, $\approx$ 7% comes from VBF while the associated production mechanisms contribute to a even lesser extent. The separation of these processes is important for the study of the couplings of the Higgs with the SM particles, since gluon fusion and ttH involve only couplings with fermions while VBF and VH only with vector bosons.
The most significant observables to discriminate among these different processes are the jet multiplicity, the jet topology and the Higgs p_{T}. The VBF process is indeed expected to give a typical signature in the final state, with two jets at a large rapidity gap. Jets may arise also from the VH and ttH processes, where the vector boson or the top decays hadronically. The Higgs p_{T} is expected to have a harder spectrum in the VBF and VH processes with respect to the ggH and ttH (see figure 4.7 left). For this reason, in order to improve the sensitivity to the couplings, the selected events are grouped according to the following categories:

- **0/1 jet**: events with less than two reconstructed jets
- **di-jet**: events with at least two reconstructed jets

where the jets are selected as described in section 4.3.3.
In the 0/1 jet category the transverse momentum of the four-lepton system ($p_{\mathrm{T}}^{4\ell}$) is used to distinguish VBF and VH from gluon fusion. In the di-jet category a linear Fisher discriminant $\mathcal{D}_{\mathrm{jet}}$ is defined combining two VBF-sensitive variables, the pseudorapidity difference among the two jets ($\Delta\eta_{jj}$) and the invariant mass of the two jet system (m_{jj}). The coefficients of the linear combination are tuned to maximize the separation power between the gluon fusion and the VBF/VH processes.
In figure 4.7 the $p_{\mathrm{T}}^{4\ell}$ and the $\mathcal{D}_{\mathrm{jet}}$ variables are shown in the simulation for the different processes.

In order to assess the predicted contribution of each in process in the two categories, the fraction of di-jet events ("di-jet ratio" in the following) is estimated in the simulation for each production mechanism in a similar fashion as for the efficiency (see section 4.4.2). The dependence on the Higgs mass is investigated by using the available MC samples as points to interpolate with a smooth function.
Results are shown in figure 4.8. The di-jet ratio does not depend on the lepton final state, as expected. There is a mild dependence on m_H in the gluon fusion and the vector boson fusion samples, which is due to the higher order corrections implemented in those generator, which depend on the virtuality of the process. The fraction of di-jet events is instead very stable in the samples corresponding to the associated production, which are generated at leading order (LO). In particular in the ttH samples more then 97% of the events contain two reconstructed jets.

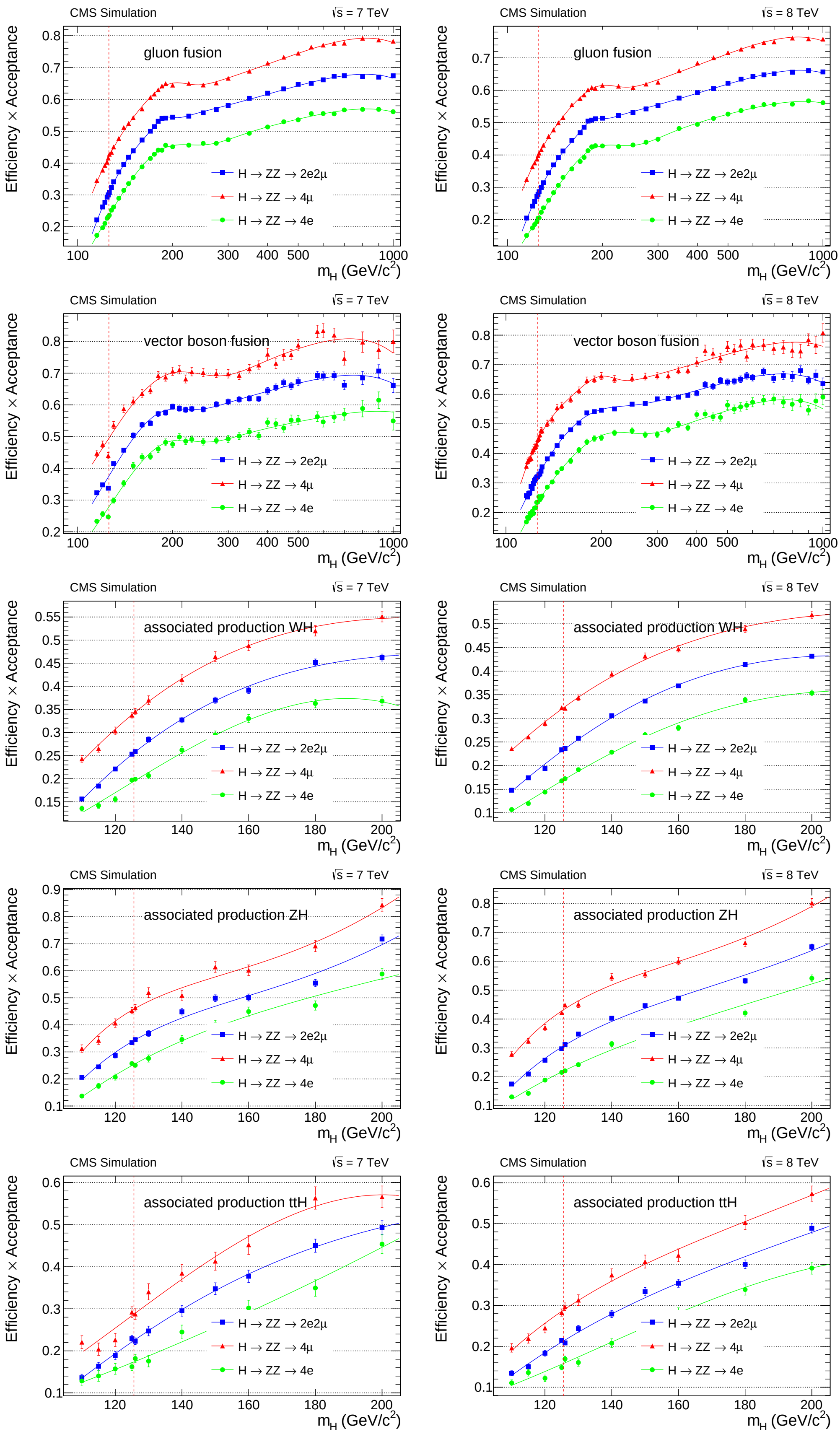


Figure 4.5: The selection efficiency as a function of the Higgs boson mass hypothesis, for 7 TeV (left) and 8 TeV (right) samples, for the different production mechanisms. The three final states, i.e. 4μ (red), $4e$ (green) and $2e2\mu$ (blue) are overlayed. The curves represent the smooth interpolation which is fitted in the points corresponding to the MC samples.

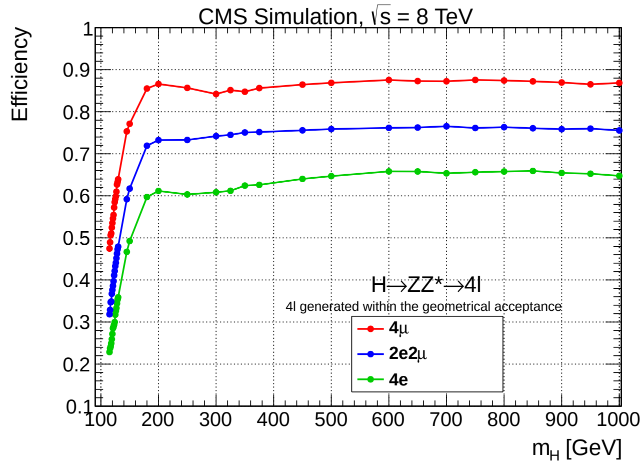


Figure 4.6: The selection efficiency within the geometrical acceptance for the gluon fusion process at 8 TeV.

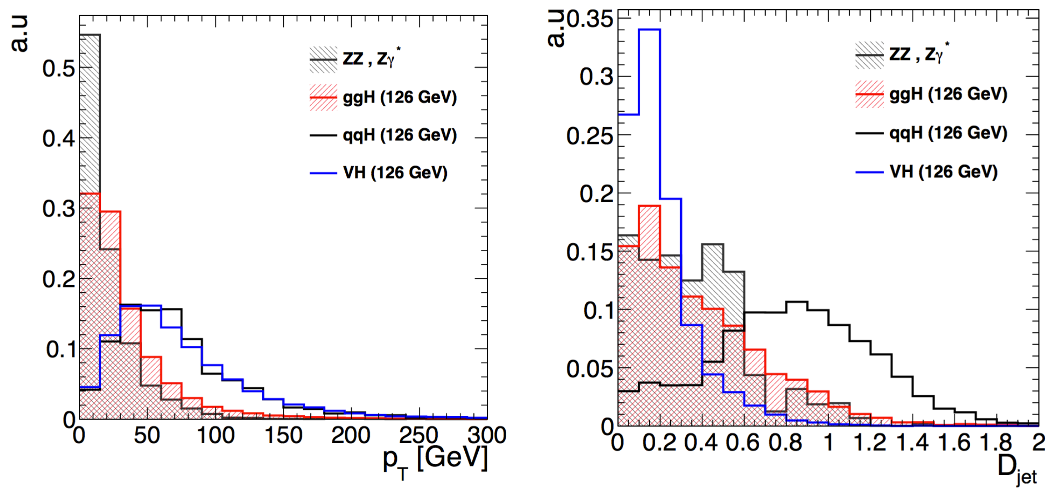


Figure 4.7: Comparison of the shape of the transverse momentum of the four-lepton system in the 0/1 jet category for different processes (left). Comparison of the shape of the Fisher discriminant in the di-jet category for different processes (right).

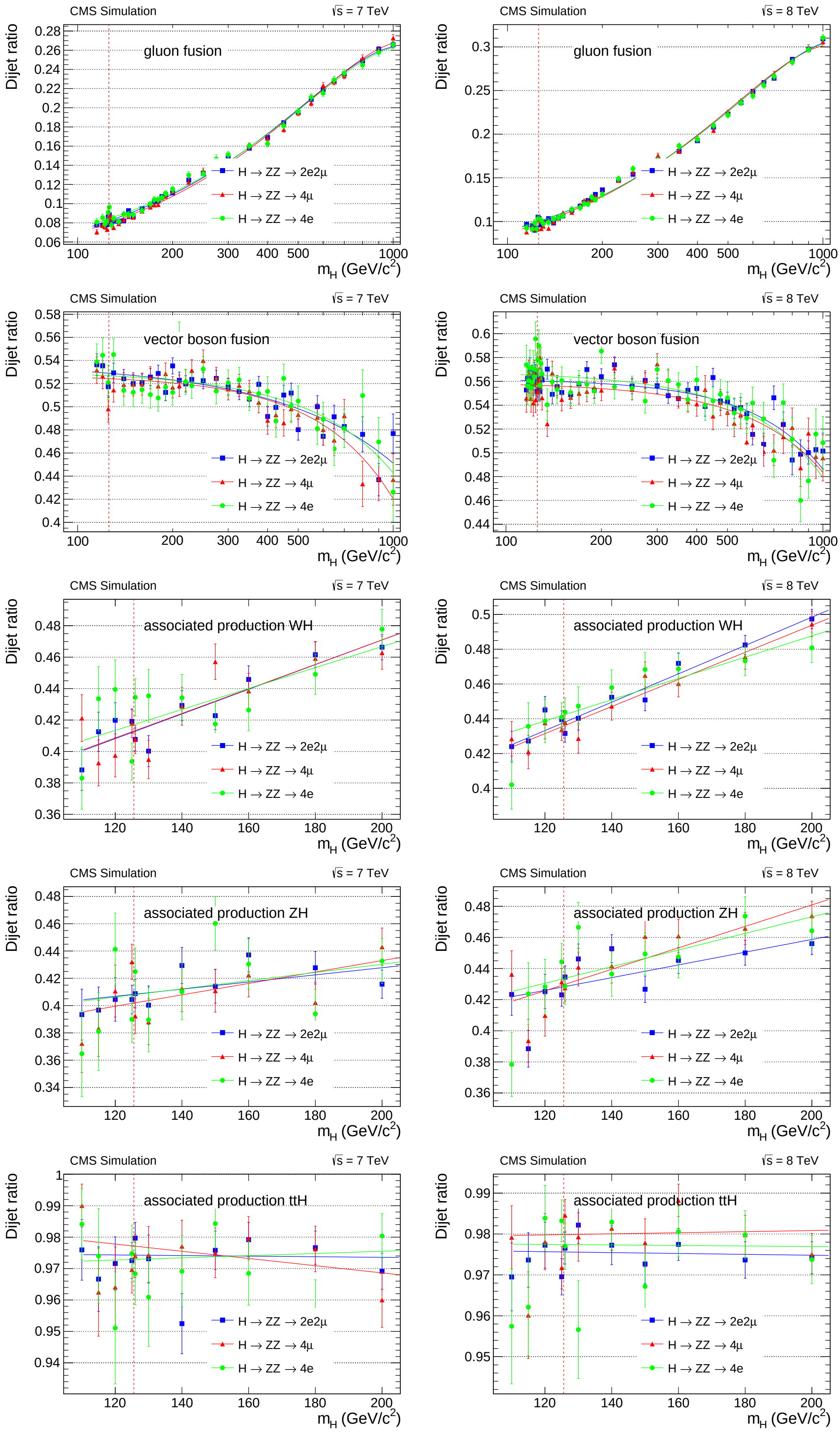


Figure 4.8: The di-jet ratio as a function of the Higgs boson mass hypothesis, for 7 TeV (left) and 8 TeV (right) samples, for the different production mechanisms. The three final states, i.e. 4μ (red), $4e$ (green) and $2e2\mu$ (blue) are overlayed. The curves represent the smooth interpolation which is fitted in the points corresponding to the MC samples.

4.5 Background estimation

The dominant background in the $H \to ZZ \to 4\ell$ search comes from the irreducible ZZ production through $q\bar{q}$ annihilation or gluon fusion. Other sources of background arise from the $Z + \text{jets}$, $t\bar{t}$, $WZ + \text{jets}$ processes, where additional leptons may come either from hadron decays or from the misidentification of jets ("fake" leptons).

4.5.1 Irreducible background

The irreducible background coming from the SM ZZ production is evaluated in the simulation. The NLO (LO) cross sections for the $q\bar{q} \to ZZ$ ($gg \to ZZ$) processes are calculated with MCFM [48, 49, 50]. The relative contribution of the gg component ranges from 2% to 6% in the full range of $m_{4\ell}$. The shape is modeled with an empirical function whose parameters are fitted in the MC samples. An example of such fits is shown in figure 4.9 for the 4μ final state and the 8 TeV MC: the shapes for the other final states are very similar.

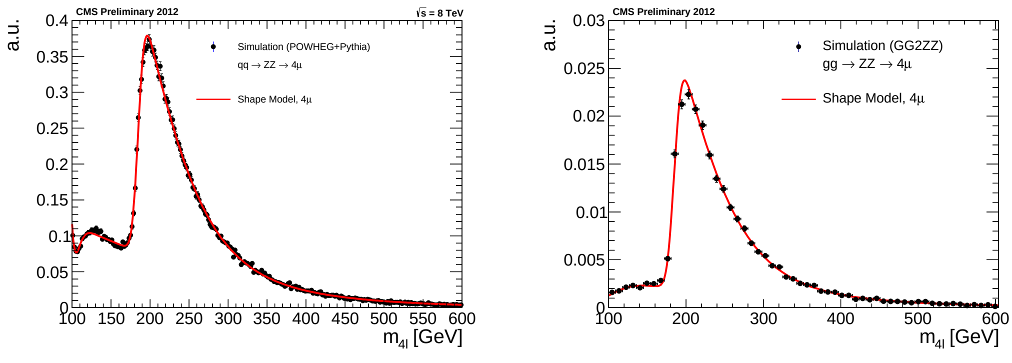


Figure 4.9: The shapes of the $q\bar{q} \to ZZ$ (left) and $gg \to ZZ$ (right) irreducible background for the 4μ final state in the 8 TeV simulation. The curve in red is the empirical model fitted to the MC prediction.

The uncertainties on the shape from experimental sources are negligible compared to the normalization uncertainties and thus are neglected.

The irreducible background from double-parton interaction $Z +$ DY process has been calculated using PYTHIA 6.4.14 [30] and found to be negligible.

4.5.2 Reducible background

The reducible background is dominated by Z + jets events with jets misidentified as leptons and will be denoted $Z+X$ in the following. Two independent methods have been developed to estimate the $Z+X$ contribution in the signal region using control regions in data. The results of the two are combined to get the final estimate.

Lepton misidentification probability

Both methods make use of the lepton misidentification probability $f(p_{\mathrm{T}}^{\ell}, \eta^{\ell})$, or "fake rate", to extrapolate the background yields from the control region to the signal region. This probability is defined as the fraction of non-signal leptons passing the lepton selection of the analysis.

The fake rate is estimated in a sample composed by $Z_1 + 1\ell_{\mathrm{loose}}$ events, where beside to the opposite sign, same flavour pair which forms the Z_1, exactly one additional lepton is reconstructed fulfilling all the selection requirements described in 4.3.1 (electrons) and 4.3.1 (muons) but for the identification and the isolation cuts. The $Z_1 + 1\ell_{\mathrm{loose}}$ sample is defined in a slightly different way in the two methods. In the method using the same-sign (SS) control region, the cut on the m_{Z_1} is less stringent, thus the corresponding sample is enriched in FSR events resulting in a larger fake rate estimation for the electrons, due to photon conversion. This can be taken into account by exploiting the linear dependency of the fake rate as a function of the fraction of *loose* electrons with one missing hits in the tracker: once the average of this observable is measured in the sample where the fake rate is applied, the fake rate itself can be corrected accordingly.
The lepton misidentification probability ranges from 1% to 15% and has a mild dependence on the pseudorapidities for the electrons.

Method with opposite-sign (OS) leptons

In this method a control region is defined by selecting events with the Z_1 plus two OS same-flavour leptons. Two categories of such events are defined:

- **2P2F control region**: the two additional leptons are both required to fail identification and isolation.
  This sample is meant to be used to estimate the contribution in the signal region of events with only two genuine leptons, such as Z + jets or $t\bar{t}$.
- **3P1F control region**: only one of the additional leptons is required to fail identification and isolation.
  This sample is meant to be used to estimate the contribution in the signal region of events with three genuine leptons, such as WZ + jets.

The number of reducible background events in the signal region is given by the following formula:

$$N_{\mathrm{exp}}^{Z+X} = \frac{f_{\mathrm{loose}}}{1-f_{\mathrm{loose}}}\left(N_{\mathrm{3P1F}} - N_{\mathrm{2P2F}\to\mathrm{3P1F}} - N_{\mathrm{3P1F}}^{ZZ}\right) + \frac{f_{\mathrm{loose},1}}{1-f_{\mathrm{loose},1}}\frac{f_{\mathrm{loose},2}}{1-f_{\mathrm{loose},2}}N_{\mathrm{2P2F}} \tag{4.4}$$

In equation 4.4, the following quantities appear:

- N_{3P1F} and N_{2P2F} are the number of events in the 3P1F and 2P2F control regions respectively
- f_{loose} and $f_{\text{loose},1(2)}$ are the fake rate of the failing lepton in the 3P1F sample and of the first (second) failing lepton in the 2P2F sample.
- $N_{\text{2P2F}\to\text{3P1F}}$ is the contamination of 2P2F-like processes in the 3P1F control region, and is estimated by:

$$N_{\text{2P2F}\to\text{3P1F}} = \left(\frac{f_{\text{loose},1}}{1-f_{\text{loose},1}} + \frac{f_{\text{loose},2}}{1-f_{\text{loose},2}}\right) N_{\text{2P2F}}$$

- N_{3P1F}^{ZZ} is the number of ZZ events in the 3P1F control region and is taken from the simulation

Method with same-sign (SS) leptons

In this method, the control region is defined with the reconstruction of the Z_1 candidate plus two additional leptons of the same flavour and charge. As explained in section 4.5.2, the fake rate used in this control region is corrected for the presence of FSR photon conversions into electrons.
The number of expected $Z+X$ background events in the signal region is estimated by:

$$N_{\text{exp}}^{Z+X} = r_{\text{OS/SS}} \times N_{\text{CR}}^{\text{data}} \times \tilde{f}_3 \tilde{f}_4 \tag{4.5}$$

In equation 4.5 $N_{\text{CR}}^{\text{data}}$ is the number of events in the same-sign control region, $\tilde{f}_{3(4)}$ is the corrected fake rate for the third (fourth) lepton and $r_{\text{OS/SS}}$ is the ratio of OS over SS events, estimated in the simulation from the ratio of the events in the 2P2F OS and 2P2F SS control regions and found to be close to unity.

Combination of the two methods

The two methods give results which are compatible within the uncertainties and are combined to get the final prediction for the $Z+X$ yield in the signal region.
The $m_{4\ell}$ shape is taken from the OS method, by fitting separately the 2P2F/3P1F-like processes, using an empirical function built from Landau and exponential distributions. The result of the fit compared to the data is shown in figure 4.10.

The two methods have been further validated with events that pass the full analysis selection except for the Z_2 being formed out of $e^{\pm}\mu^{\mp}, e^{\pm}e^{\pm}, \mu^{\pm}\mu^{\pm}$ pairs. The prediction for the reducible background has been found to be in good agreement with the data in this control region, as shown in figure 4.10.

The dominant uncertainty in the $Z+X$ background estimation comes from the limited number of events in the control regions and in the fake rate samples. Also the difference in the composition of the fake rate sample with respect to the control regions is taken into account, as well as shape uncertainties obtained by varying the

function to model the background. The total uncertanty is estimated to be 20%, 25% and 40% for the $4e$, $2e2\mu$ and 4μ channel respectively.
The final yield obtained from the combination are given in section 4.7.

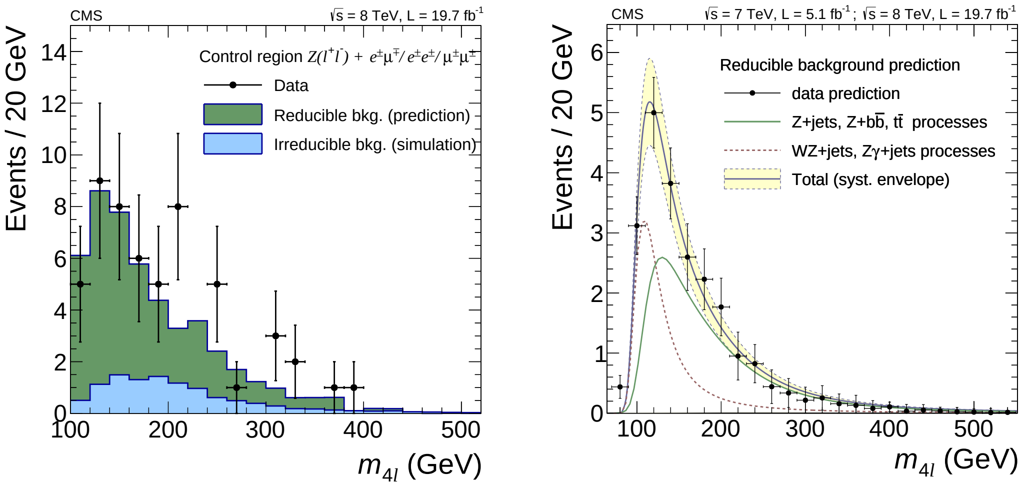


Figure 4.10: (left) The data in the SS control sample compared to the reducible background prediction (dark green) and the expected ZZ contribution. (right) The reducible background model for the three channels together compared to the data in the control region. The 2P2F-like and 3P1F-like contributions are shown separately

The reducible background contribution to the total background is about 9%, 6% and 3% (42%, 28% and 14%) in the $4e$, $2e2\mu$ and 4μ channel respectively for the mass range $100 < m_{4\ell} < 1000$ GeV ($121.5 < m_{4\ell} < 130.5$ GeV).

4.6 Kinematic discriminants

The CMS detector is able to reconstruct precisely the four-momentum of each of the Higgs daughter particle in the $H \to ZZ \to 4\ell$ decay mode. Thus, it is possible to exploit also the angular correlation between the decay products and the masses of the two Z bosons, in addition to the four-lepton invariant mass, to further discriminate the signal from the background.
The kinematic properties of the Higgs production and decay have been extensively studied in the literature [51, 52]. Five angles $\vec{\Omega} \equiv (\theta^*, \Phi_1, \theta_1, \theta_2, \Phi)$ and two masses m_{Z_1}, m_{Z_2} fully describe the kinematic of the 4ℓ system in the center-of-mass frame. In addition to these seven observables, the transverse momentum $p_{\mathrm{T}}^{4\ell}$ and the rapidity $y^{4\ell}$ are needed to characterize the system in the lab frame. The $p_{\mathrm{T}}^{4\ell}$ is already used as an independent discriminator, as explained in section 4.4.3. The rapidity has been shown to have a limited discriminator power and is not used.

A matrix-element likelihood approach, already introduced in Refs. [11, 14], is used to construct a kinematic discriminant based on the decay observables. The definition of such discriminant is based on the conditional event probability $\mathcal{P}^{\mathrm{kin}}_{\mathrm{bkg(sig)}}(m_{Z_1}, m_{Z_2}, \vec{\Omega}|m_{4\ell})$ computed from LO matrix element squared for background (signal) processes. For the signal $m_{\mathrm{H}} = m_{4\ell}$. The kinematic discriminant is defined by:

$$\mathcal{D}_{\mathrm{bkg}}^{\mathrm{kin}} = \frac{\mathcal{P}_{\mathrm{sig}}^{\mathrm{kin}}}{\mathcal{P}_{\mathrm{sig}}^{\mathrm{kin}} + \mathcal{P}_{\mathrm{bkg}}^{\mathrm{kin}}} \tag{4.6}$$

In the definition of equation 4.6 no information about the four-lepton invariant mass is used, and hence $\mathcal{D}_{\mathrm{bkg}}^{\mathrm{kin}}$ can be used as a second independent discriminating variable. The $\mathcal{P}_i$ are normalized for each $m_{4\ell}$ value in such a way that the probabilities $P(\mathcal{D} > 0.5|\mathrm{H}) = P(\mathcal{D} < 0.5|\mathrm{bkg})$.
Detector acceptance effects have found to have negligible impact, since they almost totally cancel in the ratio of equation 4.6.

The analysis makes use of the MELA (Matrix Element Likelihood Approach) framework [52,53], with the matrix elements (ME) for the signal taken from JHUGEN [51,52] and for the $q\bar{q} \to ZZ$ background from MCFM. These ME have been tested against the MEKD [54] framework, which is based on MADGRAPH [28] and FEYNRULES [55].
Also Bayesian neural network (BNN) and boosted decision tree (BDT) frameworks have been implemented as cross-checks and demonstrated to give similar performances as the ME approach.

4.7 Yields and kinematic distributions

The distribution of the reconstructed four-lepton invariant mass for the event passing the selection, for the 4μ, $4e$ and $2e2\mu$ final states combined and for the 7 TeV and 8 TeV datasets is shown in figure 4.11.
While for $m_{4\ell} \gtrsim 2m_Z$ the data are in good agreement with the expectation from the background processes, a clear excess is seen at low mass in the region around $m_{4\ell} \approx 126$ GeV. This result confirms, using the full dataset available in the Run 1 of the LHC, the evidence for a narrow resonance compatible with the SM Higgs boson with a mass of 126 GeV, that was already reported in Refs [11, 12, 13, 14].
The $Z \to 4\ell$ single-resonant peak is observed in agreement with the expectation.

In table 4.2 the observed events and the expected background rates in the full mass range are listed, together with the expected signal yields for two m_{H} hypotheses. Given the observation of an excess in the region around 126 GeV, in table 4.3 the breakdown by channels of the yields is presented integrating over the range $121.5 < m_{4\ell} < 130.5$ GeV. In table 4.4, for the same mass range, the contribution of each production mechanisms is shown in the two jet-based event categories.

In figure 4.12 the distribution of the selected events in the $(m_{4\ell}, \mathcal{D}_{\mathrm{bkg}}^{\mathrm{kin}})$ plane is shown, compared to the expectation for the background-only and background plus signal hypothesis. A clear excess of events with high rank of $\mathcal{D}_{\mathrm{bkg}}^{\mathrm{kin}}$ in a mass range around 126 GeV is seen, which is compatible with the presence of a SM Higgs boson signal of that mass.

The transverse momentum of the four-lepton system in the 0/1-jet category and the $\mathcal{D}_{\mathrm{jet}}$ in the dijet category are compared to the expectation from the simulation,

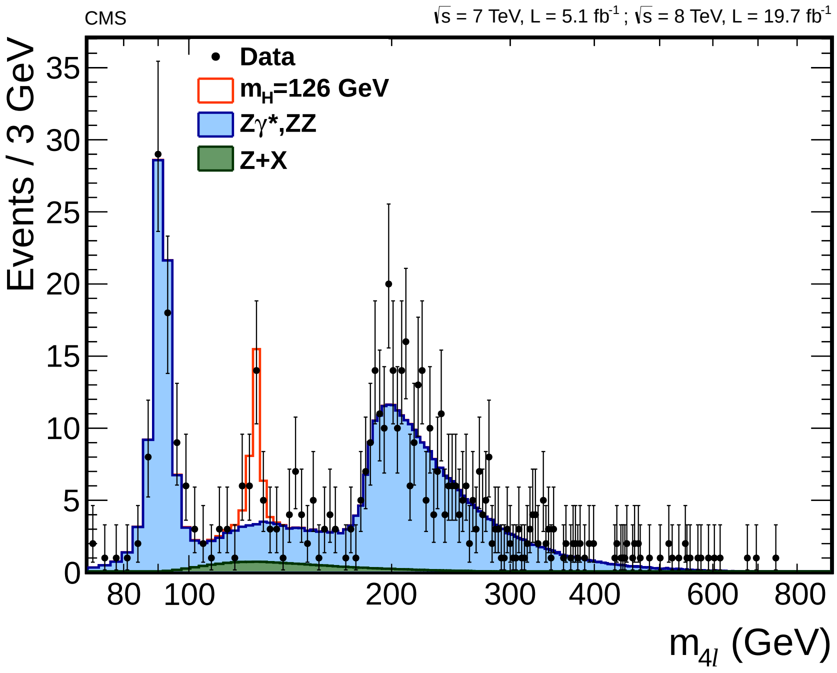


Figure 4.11: The distribution of the four-lepton invariant mass for the events passing the selection, for all the three final states and 7 and 8 TeV data combined.

Table 4.2: The number of observed events compared to the expected yields of the background processes and two signals for each final state, in the range $m_{4\ell} > 100$ GeV, for the 7 and 8 TeV datasets combined. The uncertainty includes both statistical and systematic sources.

Channel	$4e$	$2e2\mu$	4μ	all
ZZ background	77 ± 10	191 ± 25	119 ± 15	387 ± 31
Z+X background	7.4 ± 1.5	11.5 ± 2.9	3.6 ± 1.5	22.6 ± 3.6
All backgrounds	85 ± 11	202 ± 25	123 ± 15	410 ± 31
$m_{\mathrm{H}} = 500$ GeV	5.2 ± 0.6	12.2 ± 1.4	7.1 ± 0.8	24.5 ± 1.7
$m_{\mathrm{H}} = 800$ GeV	0.7 ± 0.1	1.6 ± 0.2	0.9 ± 0.1	3.1 ± 0.2
Observed	89	247	134	470

Table 4.3: The number of observed events and the expected yields of the background processes and the SM Higgs boson signal at $m_{\mathrm{H}} = 126$ GeV for each final state, in the range $121.5 < m_{4\ell} < 130.5$ GeV for the 7 and 8 TeV datasets combined. The uncertainty includes both statistical and systematic sources.

Channel	$4e$	$2e2\mu$	4μ	all
ZZ background	1.1 $\pm$ 0.1	3.2 $\pm$ 0.2	2.5 $\pm$ 0.2	6.8 $\pm$ 0.3
Z+X background	0.8 $\pm$ 0.2	1.3 $\pm$ 0.3	0.4 $\pm$ 0.2	2.6 $\pm$ 0.4
All backgrounds	1.9 $\pm$ 0.2	4.6 $\pm$ 0.4	2.9 $\pm$ 0.2	9.4 $\pm$ 0.5
$m_{\mathrm{H}} = 125$ GeV	3.0 $\pm$ 0.4	7.9 $\pm$ 1.0	6.4 $\pm$ 0.7	17.3 $\pm$ 1.3
$m_{\mathrm{H}} = 126$ GeV	3.4 $\pm$ 0.5	9.0 $\pm$ 1.1	7.2 $\pm$ 0.8	19.6 $\pm$ 1.5
Observed	4	13	8	25

Table 4.4: The number of observed events and the expected yields of the background processes and the SM Higgs boson signal at $m_{\mathrm{H}} = 126$ GeV for each jet category, in the range $121.5 < m_{4\ell} < 130.5$ GeV for the 7 and 8 TeV datasets combined. The signal yield is further split in the different production mechanisms. The uncertainty includes both statistical and systematic sources.

Category	0/1-jet	Dijet
ZZ background	6.4 $\pm$ 0.3	0.38 $\pm$ 0.02
Z+X background	2.0 $\pm$ 0.3	0.5 $\pm$ 0.1
All backgrounds	8.5 $\pm$ 0.5	0.9 $\pm$ 0.1
ggH	15.4 $\pm$ 1.2	1.6 $\pm$ 0.3
ttH	-	0.08 $\pm$ 0.01
VBF	0.70 $\pm$ 0.03	0.87 $\pm$ 0.07
WH	0.28 $\pm$ 0.01	0.21 $\pm$ 0.01
ZH	0.21 $\pm$ 0.01	0.16 $\pm$ 0.01
All signal, $m_{\mathrm{H}} = 126$ GeV	16.6 $\pm$ 1.3	3.0 $\pm$ 0.4
Observed	20	5

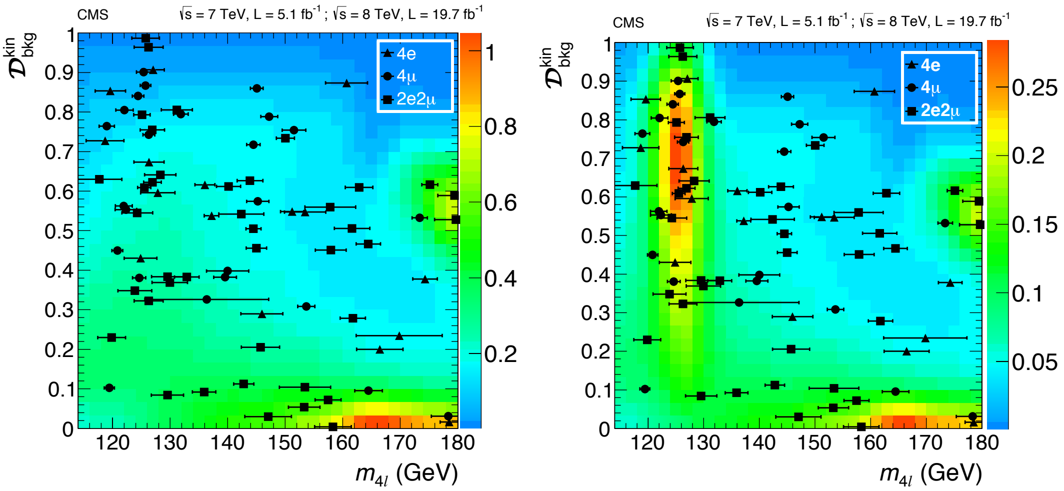


Figure 4.12: The distribution of the selected events in the $(m_{4\ell}, \mathcal{D}_{\rm bkg})$ plane. The colors correspond to the expected relative density of yields (in arbitrary units) for the background-only (left) and the signal ($m_{\rm H} = 126$ GeV) plus background hypothesis (right). Horizontal bars represent the measured mass uncertainties.

for events with an invariant mass between 121.5 and 130.5 GeV in figure 4.13. A good agreement is seen in the 0/1-jet category, while more statistics is needed to draw any conclusion for the dijet category.

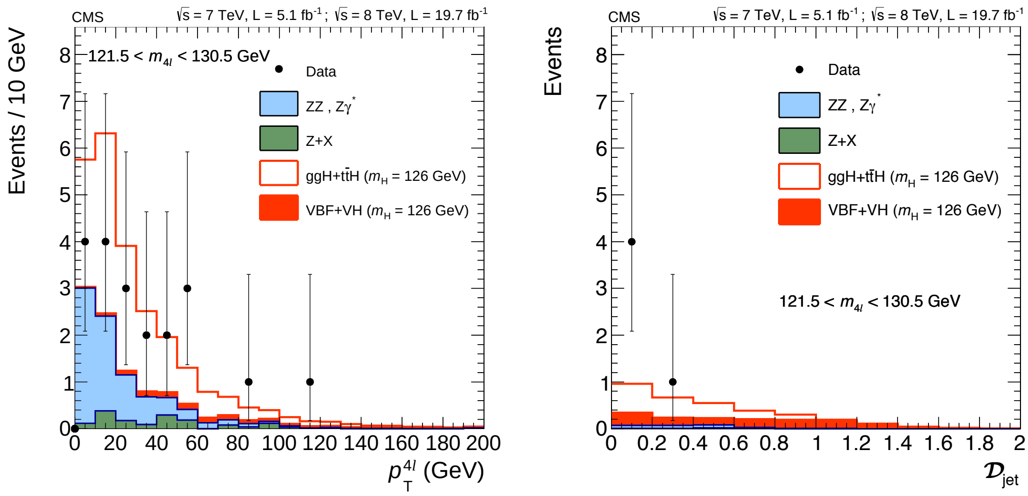


Figure 4.13: The distribution of the transverse momentum of the four-lepton system in the 0/1-jet category (left) and the one of the $\mathcal{D}_{\rm jet}$ discriminant in the dijet category (right), for events with $121.5 < m_{4\ell} < 130.5$ GeV.

4.8 Combined fit model and inputs

The properties measured in the $H \to ZZ \to 4\ell$ channel, such as the signal strength, the mass and the width of the newly discovered resonance, are extracted by means

of an unbinned likelihood fit to the selected events in the data.
The fit model includes the probabilty density function (pdf) for five signal components (ggH, VBF, WH, ZH, ttH) and three background processes ($qq \to ZZ$, $gg \to ZZ$ and Z+X). The normalization of each of these processes is incorporated in the fit model as a nuisance parameter, constrained by ancillary measurements in the simulation or in control regions in data, and is profiled during the process of the maximization of the likelihood.
In order to take into account experimental and theoretical systematic uncertainties the shapes are varied between alternative ones.

The different likelihoods for each measurement are defined below. The "dimension" of the likelihood refers to the number of observables used in the fit.

1. In order to estimate the exclusion limits on the cross section, the signal strength and the signal significance the following 3D likelihoods are defined as:

$$\mathcal{L}_{\rm 3D}^{\mu,0/1-\rm jet}(m_{4\ell}, \mathcal{D}_{bkg}, p_{\rm T}^{4\ell}) = \mathcal{P}(m_{4\ell}|m_{\rm H},\Gamma)\mathcal{P}(\mathcal{D}_{\rm bkg}^{\rm kin}|m_{4\ell}) \times \mathcal{P}(p_{\rm T}^{4\ell}|m_{4\ell}) \quad (4.7)$$

$$\mathcal{L}_{\rm 3D}^{\mu,\rm dijet}(m_{4\ell}, \mathcal{D}_{bkg}, \mathcal{D}_{\rm jet}) = \mathcal{P}(m_{4\ell}|m_{\rm H},\Gamma)\mathcal{P}(\mathcal{D}_{\rm bkg}^{\rm kin}|m_{4\ell}) \times \mathcal{P}(\mathcal{D}_{\rm jet}|m_{4\ell}) \quad (4.8)$$

   In this equation the first term is the one-dimensional pdf function of $m_{4\ell}$, the second probability includes the kinematic discriminant introduced in section 4.6 and the third dimension makes use of the $p_{\rm T}$ of the four-lepton candidate in the 0/1-jet category and of the Fisher discriminant in the dijet one, as explained in section 4.4.3.

2. For the measurement of the mass and the width of the resonance the following likelihood is used:

$$\mathcal{L}_{\rm 3D}^{m,\Gamma} \equiv \mathcal{L}_{\rm 3D}^{m,\Gamma}(m_{4\ell}, \mathcal{D}_m, \mathcal{D}_{\rm bkg}^{\rm kin}) = \mathcal{P}(m_{4\ell}|m_{\rm H},\Gamma,\mathcal{D}_m)\mathcal{P}(\mathcal{D}_m|m_{4\ell}) \times \mathcal{P}(\mathcal{D}_{\rm bkg}^{\rm kin}|m_{4\ell}) \quad (4.9)$$

   In equation 4.9 the *per-event mass uncertainties* $\mathcal{D}_m$ are used to improve the sensitivity in the mass and width measurements. They represent the expected uncertainty on the measured four-lepton invariant mass based on the kinematics and on the category of the leptons of the candidate. Their parameterization is given as a function of $m_{\rm H}$ for the signal and $m_{4\ell}$ for the backgrounds. Further details can be found in Ref. [15].

The signal lineshape is modeled, for masses $m_{\rm H} < 400$ GeV, with a relativistic Breit-Wigner (BW) convolved with a double-sided Crystall Ball, to account for the instrumental resolution. Examples of such pdf for a signal of $m_{\rm H} = 126$ GeV are given in figure 4.14.
For masses above 400 GeV a slightly different parameterization has been adopted to better describe a resonance with a huge intrinsic width (up to $\approx$ 600 GeV at $m_{\rm H} = 1$ TeV). In this model, the Γ parameter of the BW is left floating in the fit and required to be always larger than the experimental resolution $\sigma_{\rm dCB}$.

For the ZZ and Z+X background the pdf $\mathcal{P}(m_{4\ell})$ is modeled with the empirical functions described in sections 4.5.1 and 4.5.2 respectively.

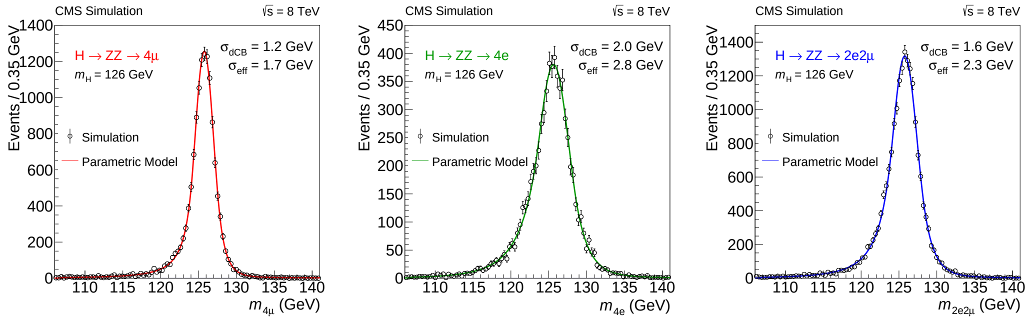


Figure 4.14: The parameterization of the signal lineshape at $m_{\mathrm{H}} = 126$ GeV for the three final states: $4e$ (left), $2e2\mu$ (center) and 4μ (right). The parametric model is superimposed on the points of the MC sample to which it is fitted.

The probabilities $\mathcal{P}(\mathcal{D}_{\mathrm{bkg}}^{\mathrm{kin}}|m_{4\ell})$, $\mathcal{P}(p_{\mathrm{T}}^{4\ell}|m_{4\ell})$, $\mathcal{P}(\mathcal{D}_{\mathrm{jet}}|m_{4\ell})$ and $\mathcal{P}(\mathcal{D}_{m}|m_{4\ell})$ are described as 2D conditional templates, in which the normalization is set to 1 for each $m_{4\ell}$ bin.
The templates for the signal and the irreducible background are built from simulation, while for the reducible background it is modeled using the control regions in data. Experimental uncertainties have a small effect on the shapes of $\mathcal{D}_{\mathrm{bkg}}^{\mathrm{kin}}$ and are incorporated as alternative distributions. The difference in the shape between the irreducible and reducible backgrounds is assigned to the latter as systematic uncertainty.
Several other systematics are considered for the $p_{\mathrm{T}}^{4\ell}$ templates: using alternative sets of parton distribution functions, varying the QCD scale, the resummation scale (only gluon fusion) and the quark mass effects. For the associated production, the LO prediction from PYTHIA is used, and effects from higher order corrections are taken into account as systematic uncertainty.
For the $\mathcal{D}_{\mathrm{jet}}$ the effects of using different generators and underlying event tunes are investigated and taken into account as additional sources of systematic uncertainty.

4.9 Systematic Uncertainties

The systematic uncertainties considered in the analysis can be divided in the following categories.

- $\boldsymbol{\alpha_{\mathrm{S}}}$**+PDF**: correspond to the variations of sets of Parton Distribution Functions (PDF). They are about 3% and 7% for qq and gg initiated processes respectively.
- **Renormalization and factorization scales**: they are varied to assess the impact of contribution from missing higher orders. The estimated uncertainty is about 7% and $\lesssim 1\%$ ($\approx 3\%$ and 24%) for gg and qq initiated processes, for the signal (ZZ background).
- **Signal acceptance**: it is evaluated using MCFM and estimated to be 2%.
- **Branching ratio $\boldsymbol{H \to ZZ}$**: 2% following the recommendation in [56].
- **Luminosity**: 2.6% at 8 TeV.

- **Lepton efficiency**: it ranges from $\approx 2\%$ to 11% depending on the final state.
- **Control region**: it is mainly due to the limited statistic in the control region and affects the normalization of the reducible background. Also the effect of the difference in the composition of the control sample and the sample in which the fake rate is computed is taken into account, as explained in section 4.5.2.
- **Di-jet category**: a 30% normalization uncertainty is assigned to the $gg \to H + 2\ \mathrm{jets}$ process and 20% to the VBF production cross section in the dijet category.
- **Lepton scale and resolution**: the eletron scale systematic on the four-lepton mass is evaluated to be 0.3% in the $4e$ channel and 0.1% in the $2e2\mu$ channel. The muon momentum scale accounts for a 0.1% systematic in the 4μ final state.
  The energy resolution uncertainty is 20% affecting the expected width of the signal and is estimated from the maximum discrepancy observed between data and MC.
- **Other shape uncertainties**: instrumental effects (jet energy scale and resolution) as well as theoretical uncertainties ($p_{\mathrm{T}}^{4\ell}$) affecting the shapes are included as shape variations.

4.10 Results

The results of the analysis, which are described in this section, include the outcome of the search for SM-like Higgs boson in the range 100-1000 GeV along with the estimation of the local significance of the excess seen at low mass, the measurement of its mass and width, and the test of different spin-parity scenarios.

4.10.1 Exclusion limits and local signal significance

The selected events are examined for 187 Higgs boson mass hypotheses in the range $100 < m_{\mathrm{H}} < 1000\ \mathrm{GeV}$. The step chosen for such mass points reflects both the intrinsic width of a SM Higgs boson, which is very small at low masses ($\Gamma_{\mathrm{H}} \approx 4\ \mathrm{MeV}$ at 125 GeV) but becomes huge at high masses ($\Gamma_{\mathrm{H}} \approx 600\ \mathrm{GeV}$ at 1 TeV), and the expected resolution on the four-lepton mass system.
The 3D likelihood model defined in equation 4.7 and 4.8 is used to set upper limits on the ratio of the production cross section to the SM prediction $\mu \equiv \sigma/\sigma_{\mathrm{SM}}$. The modified frequentist method CL_{S} [57, 58, 59] is adopted, and the Bayesian approach is used as cross-check, which gives similar results. The results presented in this section makes use of the asymptotic formulae described in Ref. [60]. More details about the statistical procedure used to extract the limits are given in Appendix A.

In figure 4.15 the 95% confidence level (C.L.) exclusion limits are shown as a function of the hypothetical Higgs mass. A SM-like Higgs boson is excluded in this channel for masses between 114.5-119.0 GeV and 129.5-832.0 GeV, for an expected exclusion range of 115-740 GeV. This result reflects both the excess seen at low mass and the slight deficit of events at high invariant masses.

The local p-value, i.e. the significance of an excess with respect to the background-only hypothesis, is shown in figure 4.15 (right) for the low mass range. The minimum is reached at 125.7 GeV, in the vicinity of the mass of the newly discovered particle, and corresponds to a local significance of about 6.8σ (6.7σ expected). As a cross-check, 1D ($m_{4\ell}$) and 2D ($m_{4\ell}$,$\mathcal{D}^{\rm kin}_{\rm bkg}$) likelihood models are also used, giving a significance of 5.0σ (5.6σ expected) and 6.9σ (6.6σ expected) respectively.
No other significant excess is observed in the whole mass range.

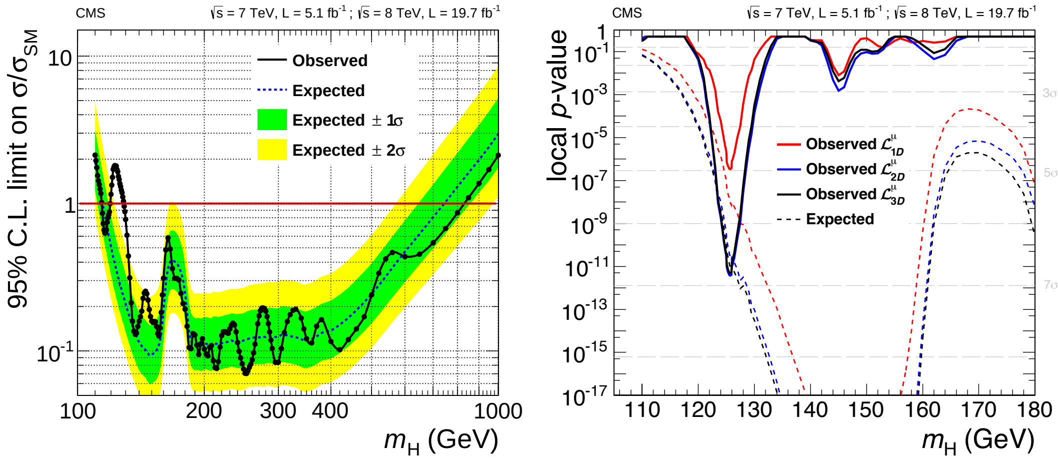


Figure 4.15: (left) The exclusion limits on μ as a function of the Higgs boson mass. The green and yellow bands represent the 1 and 2 σ uncertainty ranges for the expectation (background-only hypothesis). (right) The local p-value as a function of the Higgs mass for the three different likelihood models (low mass range). The horizontal dashed lines show the significance of the p-value in terms of number of standard deviations σ.

4.10.2 Mass and width

The measurement of the mass and of the width of the newly discovered resonance makes use of the definition of equation 4.9 for the likelihood. The per-event mass errors defined there are expected to reduce the uncertainty on the mass and width determination by 10%, with respect of the use of the average resolution. This large improvement is driven by the wide spread of the mass uncertainty, which strongly depends on the $p_{\rm T}$ and η of the leptons. This feature is also enahanced by the low number of events in this channel.

Figure 4.16 (left) shows the likelihood scan versus the resonance mass with the breakdown by channels. The cross section is left floating in the fit.
The impact of the statistical uncertainty is assessed by keeping all the nuisance parameters fixed to their best fit values. The main systematic affecting the result is the lepton scale.
The three channels are in agreement within the uncertainties. The measured mass

from the combination is:

$$m_{\mathrm{H}} = 125.6 \pm 0.4\ (\mathrm{stat.})\ \pm 0.2\ (\mathrm{syst.})\ \mathrm{GeV}$$

In figure 4.16 (right) the scan of the likelihood versus the width of the resonance is shown. The data are compatible with the hypothesis of a narrow resonance and an upper limit of 3.4 GeV on the width is put at 95% C.L., for an expected limit of 2.8 GeV.

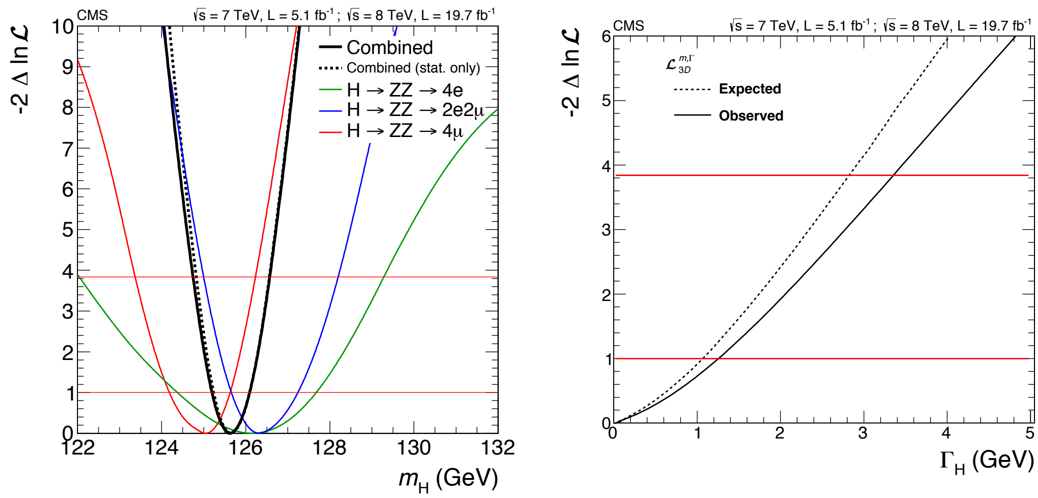


Figure 4.16: The likelihood scan versus the mass (left) and width (right) of the new resonance. The red lines at $-2\Delta\ln\mathcal{L} = 1$ and 3.84 represent the 68% and 95% C.L. respectively.

4.10.3 Signal strength

The signal strength modifier μ, or simply signal strength, is defined as the ratio of the measured cross section to the one expected from the SM.
The maximum likelihood fit is performed using the model defined in equation 4.7 and 4.8. The signal strength measured from the combination of all the channels and categories is, at $m_{\mathrm{H}} = 125.6$ GeV:

$$\mu = 0.93^{+0.26}_{-0.23}\ (\mathrm{stat.})\ ^{+0.13}_{-0.09}\ (\mathrm{syst.})$$

The expected value is $\mu = 1.00^{+0.31}_{-0.26}$, hence the total observed uncertainty agrees very well with the expected one.

The signal strength is shown separately for the two jet categories in figure 4.17 (left) and is $0.83^{+0.31}_{-0.25}$ in the 0/1-jet category and $1.45^{+0.89}_{-0.62}$ in the dijet one. In both cases the measured value is compatible with the SM expectation and with the inclusive measurement.

The jet multiplicity categorization and the introduction of variables sensitive to the production mechanism like $\mathcal{D}_{\mathrm{jet}}$ and $p_{\mathrm{T}}^{4\ell}$ makes it possible to assess the contribution

of fermion-mediated (ggH and ttH) and vector boson-mediated processes (VBF and VH) separately.
Two different signal strength modifiers are introduced, $\mu_{\mathrm{ggH,t\bar{t}H}}$ and $\mu_{\mathrm{VBF,VH}}$, and, assuming a mass of 125.6 GeV, 68% and 95% 2D contours are extracted in the plane $(\mu_{\mathrm{ggH,t\bar{t}H}}, \mu_{\mathrm{VBF,VH}})$ profiling all the other nuisance parameters.
Figure 4.17 (right) shows the result of the scan. The values that maximize the likelihood are:

$$\mu_{\mathrm{ggH,t\bar{t}H}} = 0.80^{+0.46}_{-0.36}$$
$$\mu_{\mathrm{VBF,VH}} = 1.7^{+2.2}_{-2.1}$$

Both values are compatible with unity, that is both fermion and vector boson couplings are compatible with the hypothesis of the SM Higgs boson. Given that $\mu_{\mathrm{VBF,VH}}$ is also compatible with 0, no conclusion can be drawn yet about the existence of the VBF and VH processes. Anyway, given that the coupling of the Higgs with the Z boson enters directly in the decay, they must be non null and there is no reason to believe that they are not present also on the production side.

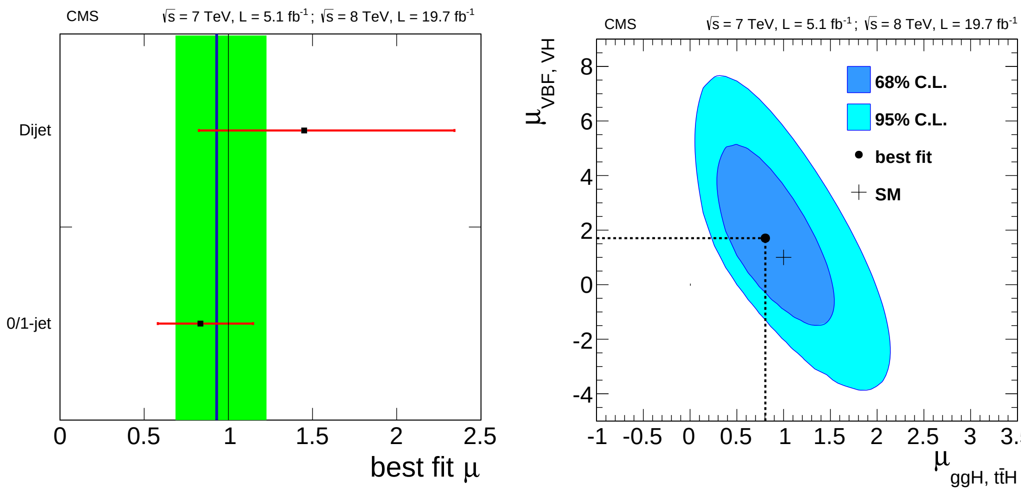


Figure 4.17: (left) The signal strength in the two jet categories. The blue vertical line is the combined value and the green band its uncertainty. The black vertical line is the value expected for a SM Higg boson. (right) The 68% and 95% C.L. contours of the signal strength associated to the fermion mediated processes ($\mu_{\mathrm{ggH,t\bar{t}H}}$) and vector boson-mediated processes ($\mu_{\mathrm{VBF,VH}}$).

4.10.4 Spin and Parity

The spin and parity quantum numbers of the newly discovered resonance are investigated using the events selected in the range $106 < m_{4\ell} < 141$ GeV, testing exotic scenarios for the tensor structure of the resonance against the hypothesis of a SM Higgs boson, which is a scalar [61, 62, 63, 64, 65, 66, 67, 68, 69, 70].
For such measurements, a different likelihood is defined, using the $\mathcal{D}_{\mathrm{bkg}}$ variable to discriminate against the background and $\mathcal{D}_{J^P}$, which is built to distinguish the J^P model under investigation against the SM scalar hypothesis 0^+. The definition of $\mathcal{D}_{J^P}$

is similar to equation 4.6 but for the presence of the alternative signal probability instead of the background one. The 2D probability density functions $\mathcal{P}(\mathcal{D}_{J^P}, \mathcal{D}_{\rm bkg})$ are described with templates constructed in a similar fashion as for $\mathcal{P}(\mathcal{D}_{\rm bkg}, m_{4\ell})$ (see section 4.8).
Twelve different spin-parity models are tested, defined in table 4.5.

Table 4.5: The spin-parity model tested in the analysis.

J^P	Production	Description
		Production dependent models
1^-	$q\bar{q} \to X$	Exotic vector
1^+	$q\bar{q} \to X$	Exotic pseudovector
$2^+_{\rm m}$	$gg \to X$	Graviton-like boson with minimal couplings
$2^+_{\rm m}$	$q\bar{q} \to X$	Graviton-like boson with minimal couplings
$2^+_{\rm b}$	$gg \to X$	Graviton-like boson with SM in the bulk
$2^+_{\rm h}$	$gg \to X$	Tensor with higher dimension operators
$2^-_{\rm h}$	$gg \to X$	Pseudotensor with higher dimension operators
		Production independent models
0^-	any	Pseudoscalar
$0^+_{\rm h}$	any	Non-SM scalar with higher dimension operator
1^-	any	Exotic pseudovector
1^+	any	Exotic vector
$2^+_{\rm m}$	any	Graviton-like boson with minimal couplings

The test statistic is defined as $q = -2\ln(\mathcal{L}_{J^P}/\mathcal{L}_{0^+})$, where $\mathcal{L}_{J^P}$ $(\mathcal{L}_{0^+})$ is the likelihood evaluated for the J^P (0^+) hypothesis. The consistency of the observed statistics with the J^P hypothesis is calculated in terms of standard deviation using a one-sided Gaussian integral. The $CL_{\rm S}$ criterion is used to infer the exclusion at a given confidence level, using the formula:

$$CL_{\rm S} = P(q \geq q_{\rm obs}|J^P + {\rm bkg})/P(q \geq q_{\rm obs}|0^+ + {\rm bkg})$$

The pseudoscalar and the spin-1 models are excluded with 99% C.L. or higher. The spin-2 model are excluded at 95% or higher C.L., while the $0^+_{\rm h}$ model is disfavoured by data with a $CL_{\rm S} = 4.5\%$. The results are also summarized graphically in figure 4.18.

The most general spin-0 boson $H \to ZZ$ decay amplitude is made of 3 contributions, whose coefficients are named $a_{1,2,3}$. For a SM Higgs boson the tree-level a_1 amplitude dominates, while the pseudoscalar is dominated by the a_3 coupling.
In addition to the pure J^P states tested against the SM Higgs boson hypothesis, also the phenomenology of a mixture of CP-even and CP-odd states can be investigated. For this purpose, the following continous quantity is defined:

$$f_{a3} = \frac{|a_3|^2\sigma_3^2}{|a_1|^2\sigma_1^2 + |a_2|^2\sigma_2^2 + |a_3|^2\sigma_3^2} \tag{4.10}$$

where σ_i corresponds to the effective cross section that one gets setting $a_j = 0, i \neq j$. In the SM, $\sigma_1/\sigma_3 = 6.36$.

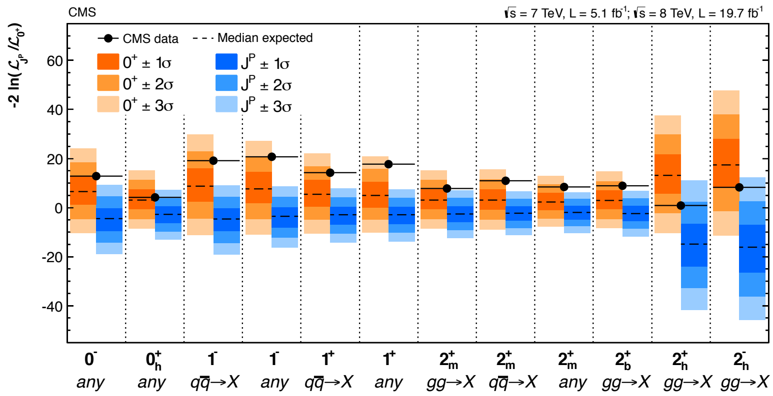


Figure 4.18: Summary of the expected distribution of the test statistics and the observed value for the twelve spin-parity models under study, for the SM Higgs (0^+) and alternative hypothesis (J^P).

The f_{a3} parameter is fitted to the data and an upper limit of $f_{a3} < 0.47$ at 95% C.L. is obtained. By assuming $a_2 = 0$, this limit can be translated into $|a_3/a_1| < 2.4$.

Chapter 5

Study of the interference between the Higgs signal and the ZZ continuum background with LO generators

Abstract

The interference of the Higgs signal with the background is a genuine quantistic effect that occurs at the amplitude level. It has to be taken into account in the search for a high mass Higgs-like state, where the width of the resonance becomes huge and the shape of the signal is significantly affected. In this chapter the study of the interference using leading order (LO) generators is described and a novel method to parameterize and include it in the signal lineshape is introduced.

5.1 Introduction

In the context of the search for a high mass Higgs boson in the $H \to ZZ \to 4\ell$ channel, the main processes that contribute to the total cross section are the gluon-fusion (ggH) and the vector boson fusion (VBF). Both these processes interfere at the amplitude level with the $gg \to ZZ$ and $qq' \to ZZqq'$ backgrounds respectively, which have the same initial and final state. This interference has a negligible impact on the total signal rate, but nonetheless distorts the shape of the signal: this effect becomes more and more important when going to high Higgs masses, due to the width becoming larger and larger.

In this thesis, the interference between the Higgs signal and the continuum background has been studied using LO generators. The results are described in section 5.2 for the gluon-fusion process and in section 5.3 for the VBF process. In section 5.4 a novel method to account for the interference in the signal lineshape is discussed.

5.2 The interference in the ggH channel

The interference in the gluon fusion channel is studied using the LO generator GG2VV [71,72] version 3.1.5. The use of a LO generator is mandatory since no calculation of the $gg \to ZZ$ continuum process is available at higher order. A detailed comparison of GG2VV against MCFM [48, 49, 50] has been carried on in the context of the measurement of the Higgs width from the off-shell cross section (see Ref. [21]). The two programs have been found to be in fair agreement when using the same settings. For the purpose of studying the interference at high Higgs masses, several samples of events have been produced, for $m_H > 400\,\mathrm{GeV}$ and for different values of a universal real scale factor C' for the couplings ($C' = 1$ in the SM). This choice is motivated by the search for BSM Higgs-like states at high mass, described in chapter 6. Each sample is generated with a statistics corresponding to an integrated luminosity of about $1M\mathrm{fb}^{-1}$ and required approximately three days to be produced with one hundred jobs running in parallel and producing files in the Les Houches Event (LHE) format.
All the samples have been generated using the CTEQ6L1 LO pdf and with a "running" renormalization and factorization scales set to $m_{4\ell}/2$. The following cuts have been applied at generator level:

- $p_{\mathrm{T}}^{\ell} > 5\,\mathrm{GeV}$ and $|\eta^{\ell}| < 2.7$ for electrons and muons
- $m_{4\ell} > 100\,\mathrm{GeV}$ which defines the phase space of the Higgs search in the $H \to ZZ \to 4\ell$ analysis (see 4.4)
- $p_{\mathrm{T}}^{\ell\ell} > 1\,\mathrm{GeV}$ which is included in the generator settings for computational stability and has a negligible effect for the range of Higgs masses under study

The cuts listed above are applied to mimic the lepton selection of the $H \to ZZ \to 4\ell$ analysis while trying to be inclusive enough to keep a good efficiency in the sampling during the generation process.

The GG2VV generator allows to compute the cross section for the signal-only (S), background-only (B) amplitudes and the process where the signal, background and the interference are all present at the same time (S+B+I). The three amplitudes are compared in figure 5.1 for different Higgs masses and C' values.
In some of the plots in figure 5.1 it can be seen that the interference is destructive above the pole mass and contructive below: this feature will be more evident in section 5.4.

5.3 The interference in the VBF channel

In order to study the interference in the VBF channel, the PHANTOM generator [73] version 1.2.3 has been used. It allows to generate $qq' \to ZZqq'$ up to α^6_{EWK} and therefore includes also the interference between a Higgs produced in the VBF channel with the background process which proceeds through the "fermion box" diagram.
As for the gluon-fusion, the interference has been studied generating several samples corresponding to different masses, scaling of the couplings and of the width of the Higgs. Each sample is generated with a statistics corresponding to an integrated luminosity of about $1M\mathrm{fb}^{-1}$ and required approximately three days to be produced with one hundred jobs running in parallel and producing files in the Les Houches Event (LHE) format.
In the generation of samples with PHANTOM, the CTEQ6L1 pdf sets has been used with a "running" renormalization and factorization scales set to $m_{4\ell}/\sqrt{2}$. The following cuts are applied at generator level:

- $p^{\ell}_{\mathrm{T}} > 5\,\mathrm{GeV}$ and $|\eta^{\ell}| < 2.7$ for electrons and muons
- $m_{4\ell} > 100\,\mathrm{GeV}$ which defines the phase space of the Higgs search in the $H \to ZZ \to 4\ell$ (see 4.4)
- $p^{j}_{\mathrm{T}} > 10\,\mathrm{GeV}$, $|\eta^{j}| < 6.5$ and $m_{jj} > 30\,\mathrm{GeV}$ for jets

As in the case of the gluon fusion, the selection cuts mimic the one of the analysis at reconstructed level.

In PHANTOM, processes corresponding to signal only or background only amplitudes cannot be calculated, because they spoil the regularization of the continuum processes at high masses. In order to access the mass distribution of each single contribution, the following relation is used:

$$P_{\mathrm{tot}} = \mu \times P_S + \sqrt{\mu} \times P_I + P_B \tag{5.1}$$

In equation 5.1 μ is the signal strength, P_{tot} is the total probability and $P_{S,B,I}$ are the signal, background, interference contributions, each with its appropriate scaling.

One can construct a three-equation linear system (for every value of the Higgs mass and width), each equation corresponding to the relation 5.1 evaluated at a certain signal strength. The signal, background and interference probabilities (P_S, P_B and P_I) can be then extracted by inverting the following system:

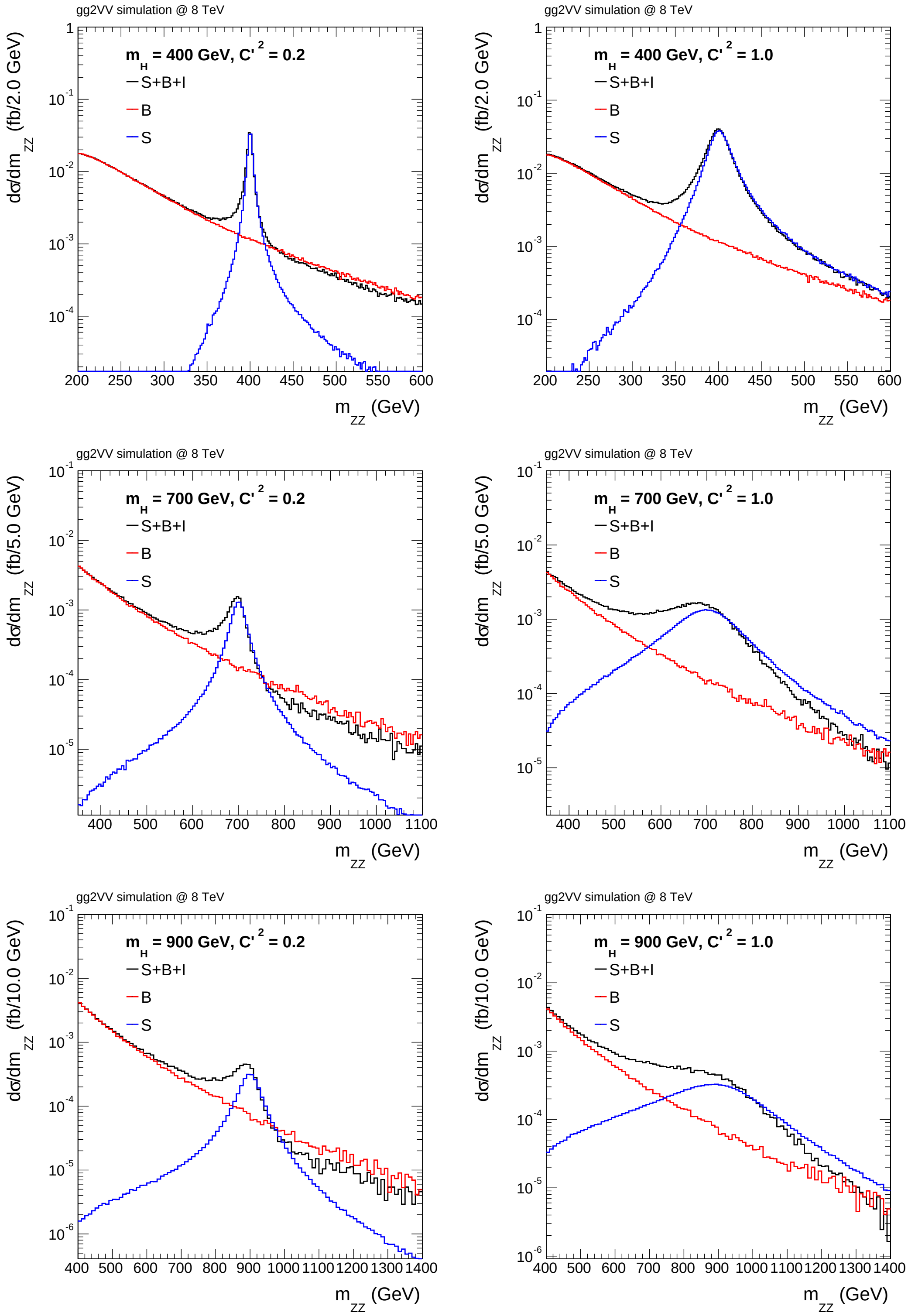


Figure 5.1: The generator level differential cross section obtained with the LO program GG2VV, for different Higgs masses and width scale factor C'^2 ($C' = 1$ corresponds to the SM). The contributions of the Higgs signal, the continuum background and the sum of the two plus their interference are overlayed.

$$\begin{pmatrix} P_{\mu_1} \\ P_{\mu_2} \\ P_{\mu_3} \end{pmatrix} = \begin{pmatrix} \mu_1 & \sqrt{\mu_1} & 1 \\ \mu_2 & \sqrt{\mu_2} & 1 \\ \mu_3 & \sqrt{\mu_3} & 1 \end{pmatrix} \begin{pmatrix} P_S \\ P_I \\ P_B \end{pmatrix} \tag{5.2}$$

A workaround used in PHANTOM to generate the background only distribution is to set the Higgs mass to a very large value, i.e. $\approx$ 10 TeV: in this way one expects the high mass part of the spectrum to resemble to the non-regularized non-Higgs m_{ZZ} distribution. The comparison between the background only distribution obtained with this workaround and the one obtained using equation 5.2 is show in figure 5.2. The two distributions agree very well within the statistical uncertainties. The one extracted with the 10 TeV Higgs is used in the following as background only probability. This choice presents some advantages: the background sample is generated only once with higher statistics, the system in 5.2 is reduced to two equations and the number of samples to be produced is reduced by a factor 2/3.

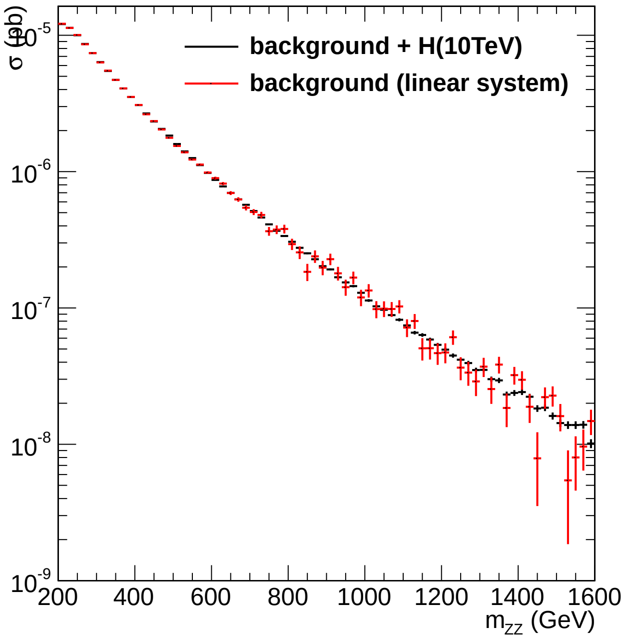


Figure 5.2: The comparison between the PHANTOM generation with the Higgs mass set to 10 TeV (black points) and the distribution P_B defined in equation 5.2 (red points), for the high mass part of the m_{ZZ} spectrum.

Examples of differential m_{ZZ} cross sections obtained with PHANTOM showing the different contributions are shown in figure 5.3 for different values of the Higgs mass and width scale factor C'^2.

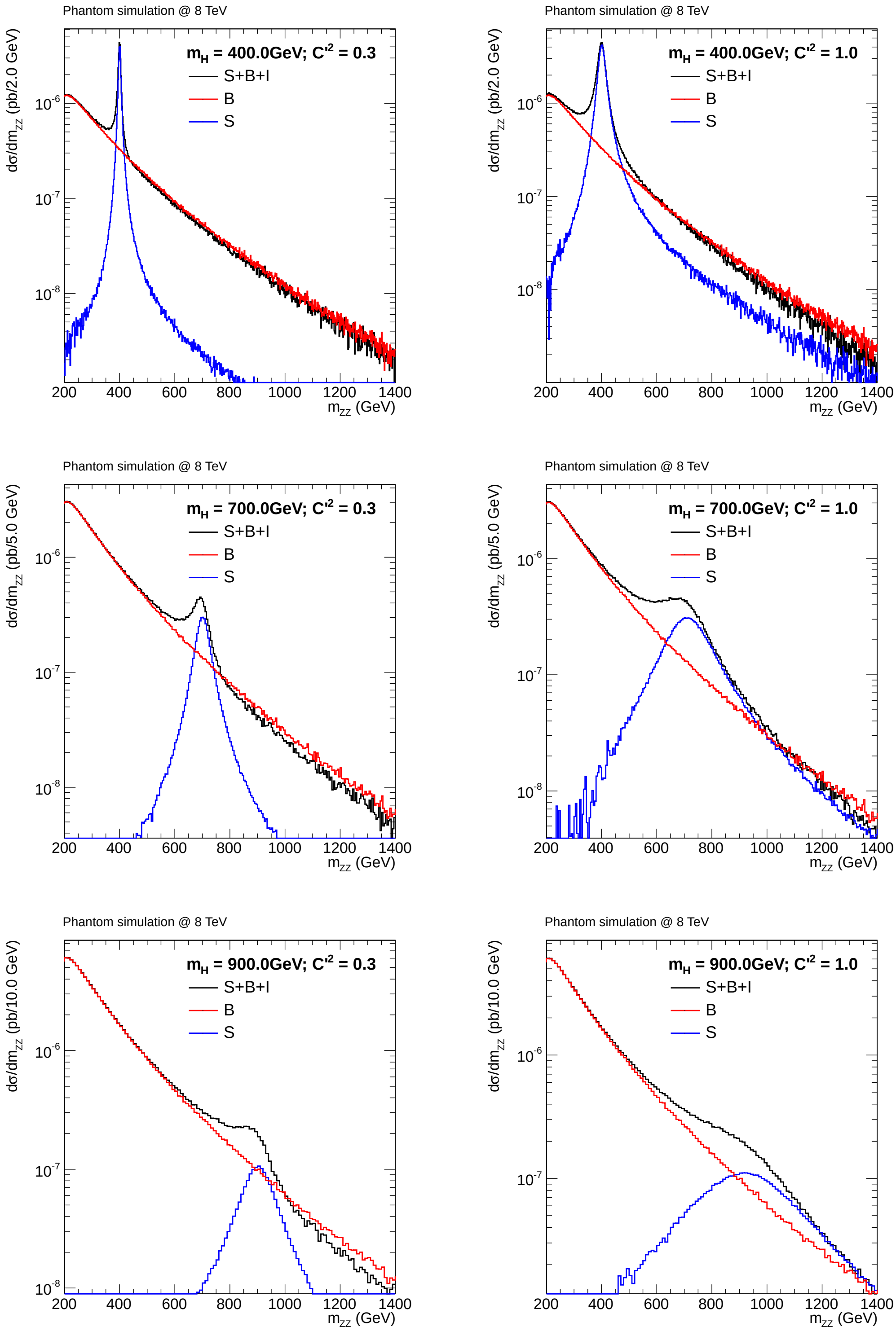


Figure 5.3: The generator level differential cross section obtained with the LO program PHANTOM, for different Higgs masses and width scale factor C'^2 ($C' = 1$ corresponds to the SM). The contributions of the Higgs signal, the continuum background and the sum of the two plus their interference are overlayed.

5.4 The extraction and parameterization of the interference

In sections 5.2 and 5.3 the use of LO generators to compute differential cross-section for the gluon-fusion and the vector boson fusion channels has been discussed and the procedure to obtain the unphysical cross sections corresponding to the interference only amplitude has been described. In this section the procedure to incorporate the interference in the signal lineshape, which is used as a signal model in the fit, is explained using the samples generated with the GG2VV and PHANTOM programs.

The choice to assign the contribution of the interference to the signal is somewhat arbitrary: a more complete approach would be to describe the interference separately in the fit to extract the yields, and make it scale as it is predicted by the model. However, the approximation described in the following implements the exact normalization for the interference with respect to the signal for $\mu = 1$ and it is believed to be enough accurate for the purpose of the search for a heavy Higgs-like state, described in chapter 6.

Figure 5.4 shows examples of interference patterns extracted from PHANTOM. It is contructive below the pole mass ($m_{ZZ} < m_H$) and destructive above ($m_{ZZ} > m_H$), and tends to zero far from the pole. The shape is not exactly symmetrical, but the effect of the interference on the total cross section is negligible, as discussed in Ref. [74]. Figure 5.4 shows also how the effect becomes more and more relevant as the mass of the Higgs increases: for $m_H < 400\,\mathrm{GeV}$ it is negligible and no correction to the signal lineshape is implemented, while at very high masses (figure 5.4 right) the interference spoils the signal shape significantly. The uncertainties in the low mass tail reflect the nature of the method used to extract the shape. In that part of the spectrum, infact, $\sigma_{S+B+I} \approx \sigma_B >> \sigma_S$, σ_I, thus extracting the interference is a *fine tuning* procedure, that is obtaining a small yield by subtracting two large yields.

An analytical parametrization is used to describe the shape of the interference, which is built from simple assumptions. If the signal part in the amplitude is described with a real pole and the background is modeled with a simple exponential, which is a good approximation in the region $m_{ZZ} \gtrapprox 300\,\mathrm{GeV}$, we can write:

$$|S+B|^2 \approx \left| C_{\mathrm{S}} \cdot \frac{1}{m_{ZZ}^2 - m_H^2} + C_{\mathrm{B}} \cdot e^{-\alpha m_{ZZ}} \right|^2 = |S|^2 + |B|^2 + I \qquad (5.3)$$

In equation 5.3 the real pole can be removed using the complex mass scheme [75], namely with the substitution $m_H^2 \to m_H^2 - i m_H \Gamma$, leading to the following expression for I:

$$I(m_{ZZ}|m_H, \Gamma, \alpha) = Re\left(\frac{e^{-\alpha m_{ZZ}}}{m_{ZZ}^2 - m_H^2 + i m_H \Gamma} \right) \qquad (5.4)$$

The signal only shape is described, at high mass, by a *modified Breit-Wigner* formula, which has been found to be more effective in modeling a resonance with a

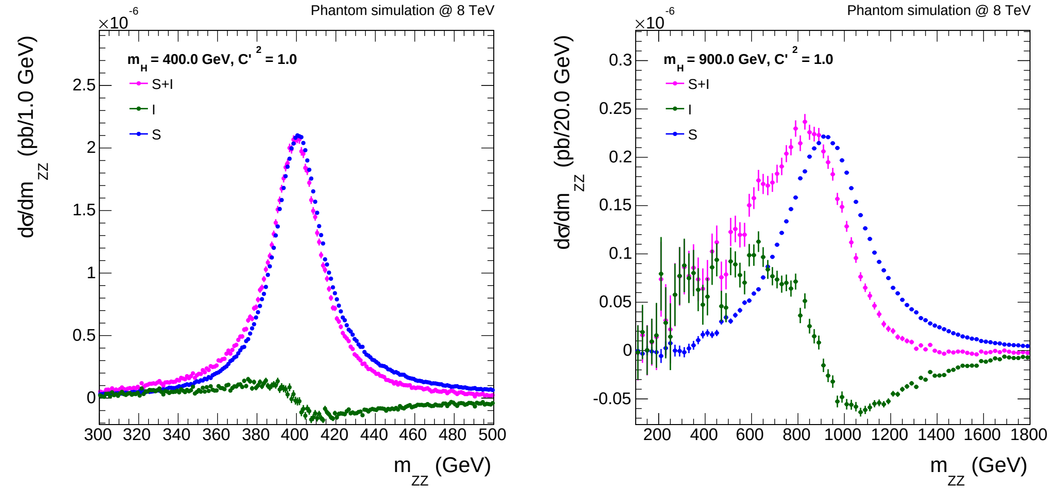


Figure 5.4: The m_{ZZ} distribution for the signal only (blue), interference only (green) and signal plus interference (black) for two Higgs masses, from a PHANTOM simulation at 8 TeV.

large intrinsic width:

$$S(m_{ZZ}|m_H, \Gamma) = \frac{m_{ZZ}}{(m_{ZZ}^2 - m_H^2)^2 + m_H^2\Gamma^2} \tag{5.5}$$

Examples of interference only fits are shown in figure 5.5. From these plots is clear how the uncertainty on the interference contribution becomes large at low virtualities, due to the reasons mentioned above. This is true in particular for the very high Higgs boson masses $m_H \gtrsim 800\,\text{GeV}$, where some fits are suboptimal.

In figure 5.6 examples of signal only fits are shown, for different masses and width scale factors. The model written in equation 5.5 works very well for every mass and width of the Higgs.

Finally, in order to obtain a model which describes the signal shape corrected for the interference, the signal only and the interference only parameterizations, defined in equation 5.5 and 5.4 respectively, are combined:

$$S_{\text{corr}}(m_{ZZ}|m_H, \Gamma, \alpha, r) = S(m_{ZZ}|m_H, \Gamma) - r \cdot I(m_{ZZ}|m_H, \Gamma, \alpha) \tag{5.6}$$

The parameter r is fitted on the $S+I$ distribution, keeping all the other parameters fixed.
The pdf defined by 5.6 is forced to be greater or equal to zero in the whole mass range in which is defined, to avoid complications coming from defining a negative-valued pdf and negative weights for the MC samples. This approximation is largely covered by the systematics assigned to the signal shape, as described in section 6.2.1.

The results of the fit to the distributions of the signal corrected for the interference are shown in figure 5.7.

The parameters of the analytical model described above are interpolated among

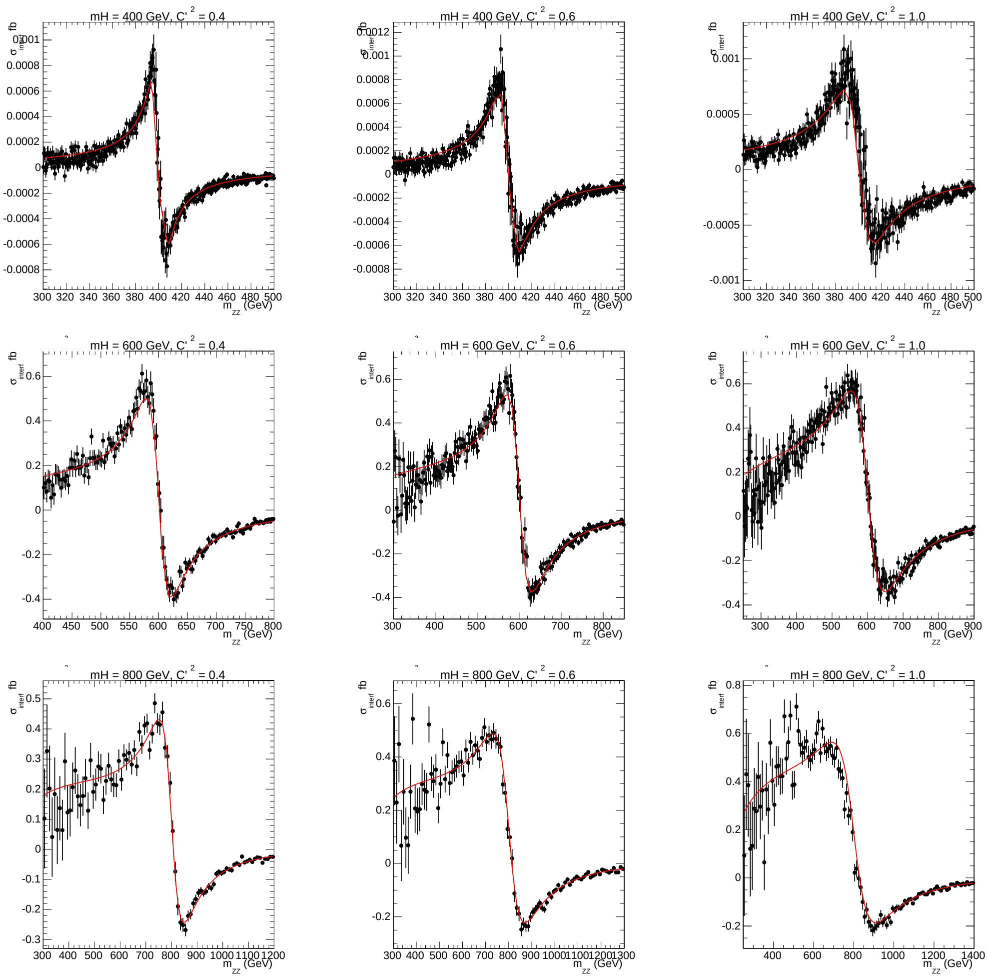


Figure 5.5: Fits to the shape of the interference in the gluon fusion samples produced with the GG2VV generator at 8 TeV, for different values of the mass (from top to bottom) and width scale factor (from left to right).

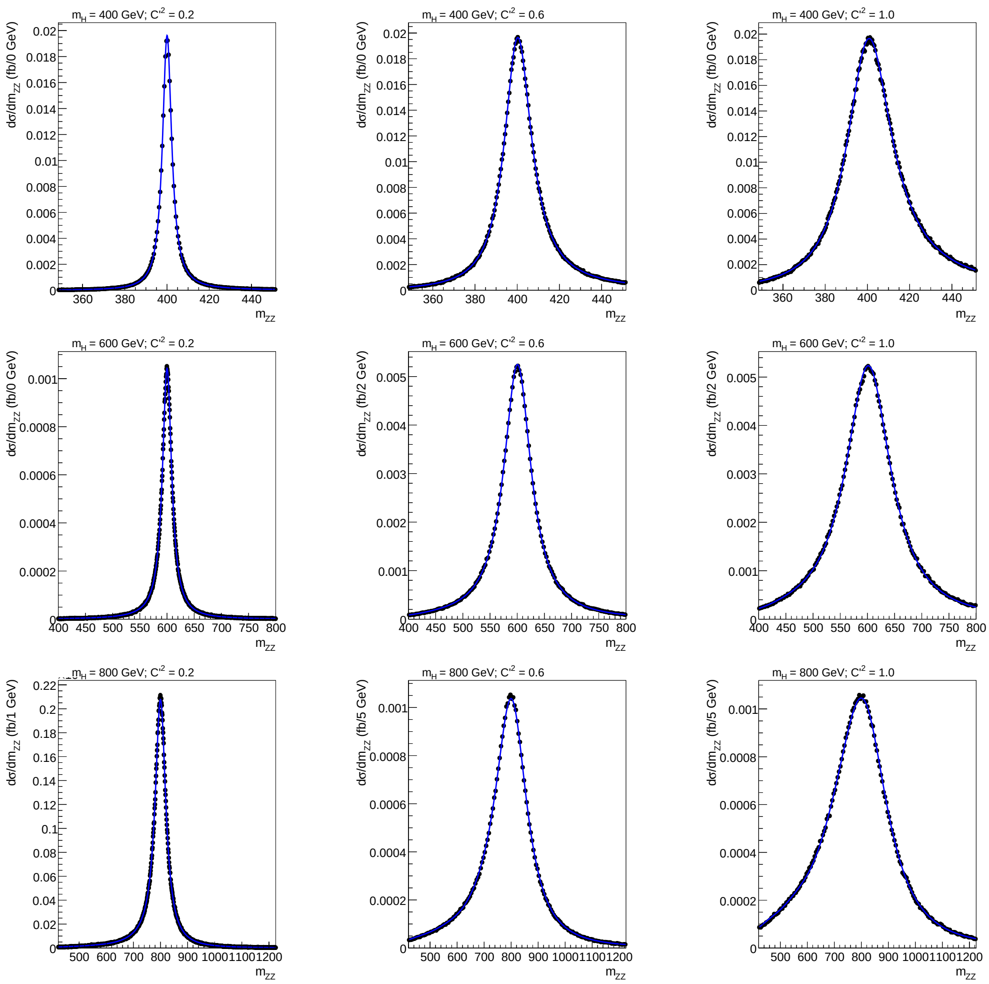


Figure 5.6: Fits to the shape of the signal (not corrected for the interference) in the gluon fusion samples produced with the GG2VV generator at 8 TeV, for different values of the mass (from top to bottom) and width scale factor (from left to right).

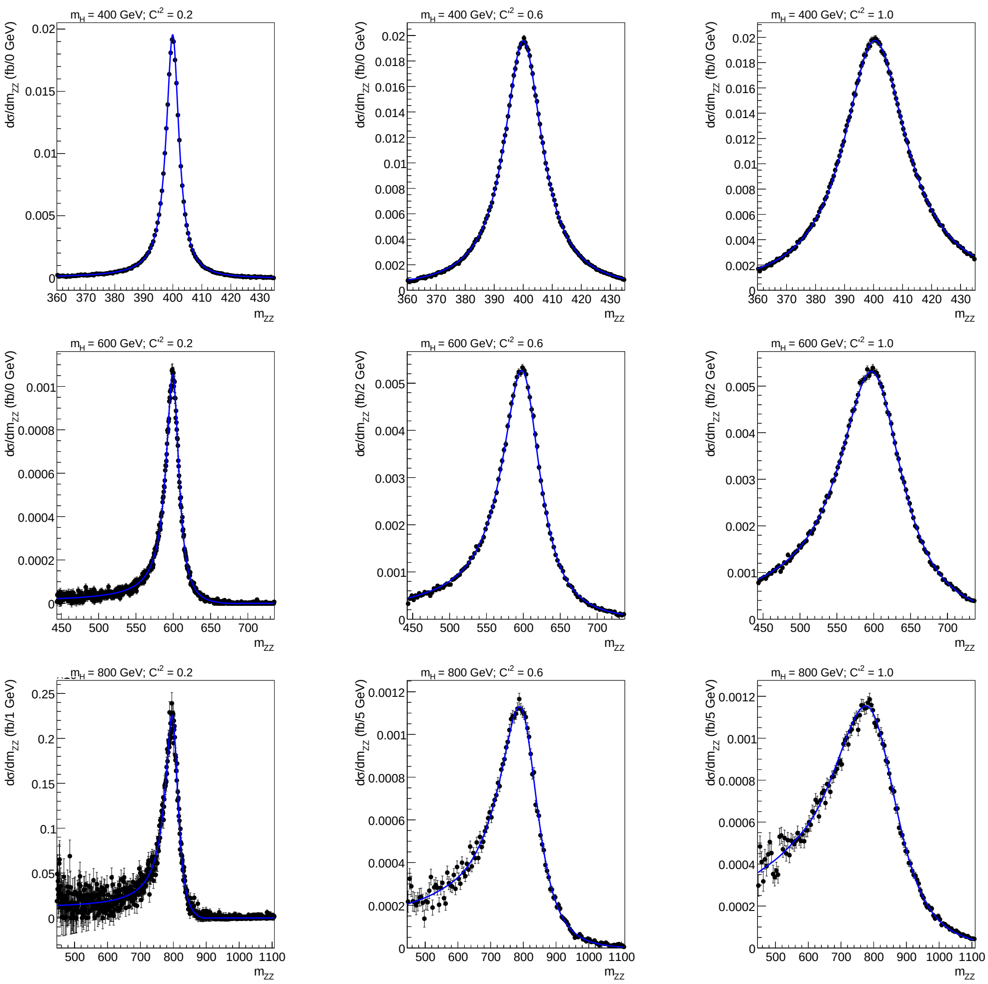


Figure 5.7: Fits to the shape of the signal corrected for the interference in the gluon fusion samples produced with the GG2VV generator at 8 TeV, for different values of mass (from top to bottom) and width scale factor (from left to right).

the discrete points, corresponding to the samples produced with the GG2VV and PHANTOM generators, to get the signal lineshape of a heavy Higgs boson for a generic mass and coupling scale factor. This probability density function has been implemented in the framework of the RooFit package [76] and is used in chapter 6 in the context of the search for a high mass Higgs-like state.

Chapter 6

Search for a high mass Higgs boson

Abstract

The observation of a 125 GeV Higgs boson is consistent with the theoretical constraints of the SM and its measured properties are in good agreement with the SM predictions. However, there is a possibility that the newly discovered particle is part of an extended scalar sector and is only partially responsible for the electroweak symmetry breaking. A heavy Higgs state is invoked in many models of new physics predicting an extended scalar sector. The simplest class of these models includes an additional singlet scalar field which interacts with the SM particles through the mixing with the SM Higgs. The search for a heavy electroweak singlet state decaying to a pair of Z bosons, in the four-lepton final state, is presented in this chapter. Preliminary results of the combination with other ZZ and WW final states is also shown.

6.1 Introduction

There are several reasons to discard the hypothesis of a heavy Higgs in the SM.
For example there exist theoretical constraints which set upper limits on the mass of a SM Higgs boson. The unitarization of the longitudinal W/Z scattering amplitude, for instance, is achieved with a Higgs with mass $m_h \lesssim 700\,\mathrm{GeV}$. Similar constraints are obtained from the requirement that no Landau pole appears in the Higgs self coupling.
The experimental results of the direct searches in CMS currently exclude the SM Higgs boson in the mass range 127-710 GeV [77]. Preliminary results from CMS using the full statistics of the Run 1, presented in 6.3, show that the excluded range can be extended up to 1000 GeV.
With this in mind, the search for the SM-like Higgs boson at high mass is not so meaningful anymore, since the 125 GeV Higgs takes already the role of restoring the unitarization of the SM at high four-lepton invariant masses. For this reason, the focus of this chapter is on the exclusion of a Higgs with non-SM couplings.

Many BSM theories, however, invoke the existence of a heavy neutral Higgs state: the Two Higgs Doublet Model (2HDM) is probably the most popular of them, which includes the Minimal Supersymmetric Standard Model as a Type II 2HDM. However, the simplest realization of an extended scalar sector is performed in the electroweak singlet (EWS) models, described in 1.3.2. The EWS phenomenology has

been investigated in CMS for a singlet mass in the range $145\,\mathrm{GeV} < m_H < 1000\,\mathrm{GeV}$ using the final states which are more sensitive to a signal in this "high mass" region, namely the WW and ZZ decay channels.
In section 6.2 the results in the $H \to ZZ^{(*)} \to 4\ell$ channel are discussed. In section 6.3 the preliminary results of the CMS combination of the high mass analyses are presented. This results of the combination are undergoing internal review in the collaboration and the submission of the paper is foreseen for the early 2015.

6.2 BSM intepretation of the high mass data in the $H \to ZZ^{(*)} \to 4\ell$ channel

The high mass data ($m_H \geq 145\,\mathrm{GeV}$) in the four-lepton channel are interpreted in the EWK singlet models and the results are presented in this section.
The baseline object reconstruction and event selection, as well as most of the statistical tools, are the ones of the "Legacy" SM analysis, described in chapter 4. For the purpose of this analysis, some changes with respect to the Legacy strategy are implemented, which are summarized here:

- the signal lineshape is revisited, in order to implement the effect of the interference with the background and the scaling of the width of the resonance for the BSM interpretation;
- the yield and shape of the $qq' \to qq'ZZ \to 4\ell$ background process are estimated and accounted for in the fit;
- the Fisher discriminant used in the 2-jet category, described in section 4.4.3, is replaced by the VBF MELA discriminant, described in section 6.2.3;
- the MINLO [78] generator is used to describe the gluon fusion Higgs production in association with jets and POWHEG v1.5, implementing the Complex Pole Scheme (CPS) [79], is used for both the gluon fusion and vector boson fusion processes;
- the SM Higgs cross sections and branching ratios are updated to the last available calculations [74].

6.2.1 Signal lineshape

The MC samples used in the analysis are described in section 4.2.2. The Higgs signal samples produced with POWHEG are reweighted to include the effect of the interference with the background, calculated with the GG2VV and PHANTOM generators, as explained in chapter 5. In particular, the unbinned model described in section 5.4 is used to extract m_{ZZ}-dependent event weights: this smoothes out the statistical uncertainties in the tails and allows to extract the weights for arbitrary values of the parameters of the model, i.e. m_H, C' and $\mathrm{BR}_{\mathrm{new}}$. It has to be noticed that these weights are also used for the other ZZ channels included in the combination. The interference weights are applied only for Higgs masses above 400 GeV, where the effect is sizeable, while the reweighting for the scaling of the resonance width is applied

for any Higgs mass.
The weights are defined as:

$$w(m_{ZZ}) = \frac{S_{\text{corr}}(m_{ZZ})}{S(m_{ZZ})} \tag{6.1}$$

where in the equation above, S_{corr} and S are defined in 5.6 and 5.5 respectively.

Examples of signal reweighted shapes are given in figures 6.1 and 6.2 for the gluon fusion and the vector boson fusion respectively.

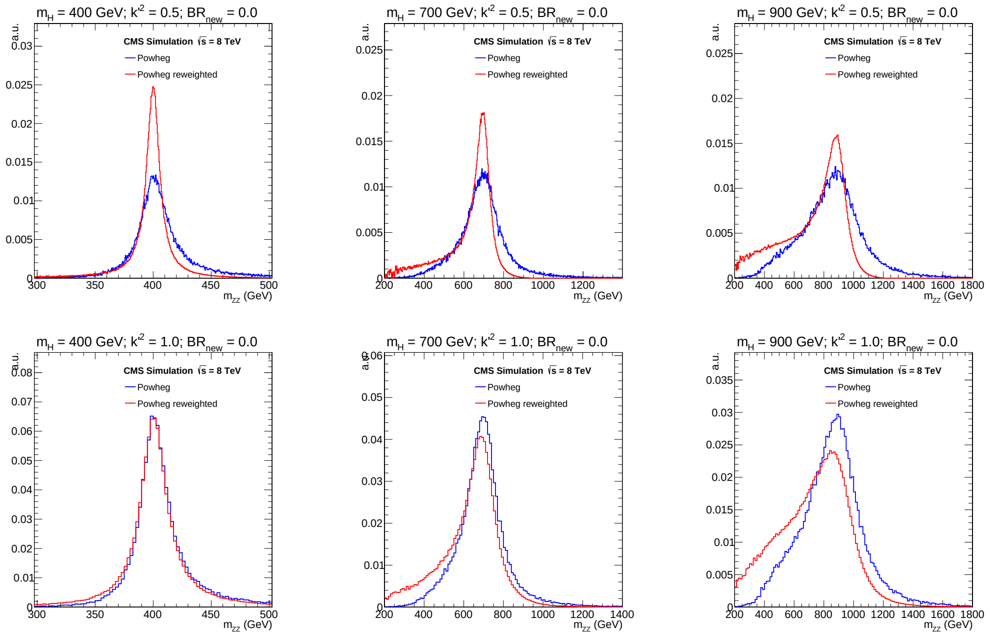


Figure 6.1: The effect of the interference reweighting on the 8 TeV gluon fusion Higgs signal MC samples, for $C'^2 = 0.5$ (top) and for the SM (bottom), i.e. $C'^2 = 1$.

The $m_{4\ell}$ unbinned signal model is the convolution of the pdf defined in 5.6, which accounts for the theoretical shape corrected for the interference with the background, with a double-sided Crystal Ball (CB) function which describes the experimental resolution. Since the resolution depends only on m_H but not on C', the CB parameters are derived for one value of C' and interpolated with polynomials between the Higgs mass points.
In figure 6.3 and 6.4 the pdfs of the signal model are compared with the reweighted MC, for the gluon fusion and the vector boson fusion signal shape respectively.

An additional source of uncertainty in the shape of the signal comes from the calculation of the interference. While writing this document, the background in the gluon fusion channel is known only at LO. Since the NNLO K-factors are known for the signal, a question arises whether these large corrections (factor of two) should be included in the calculation of the interference.
Three approaches have been proposed in Ref. [74] (S is the signal, B the background and I the interference):

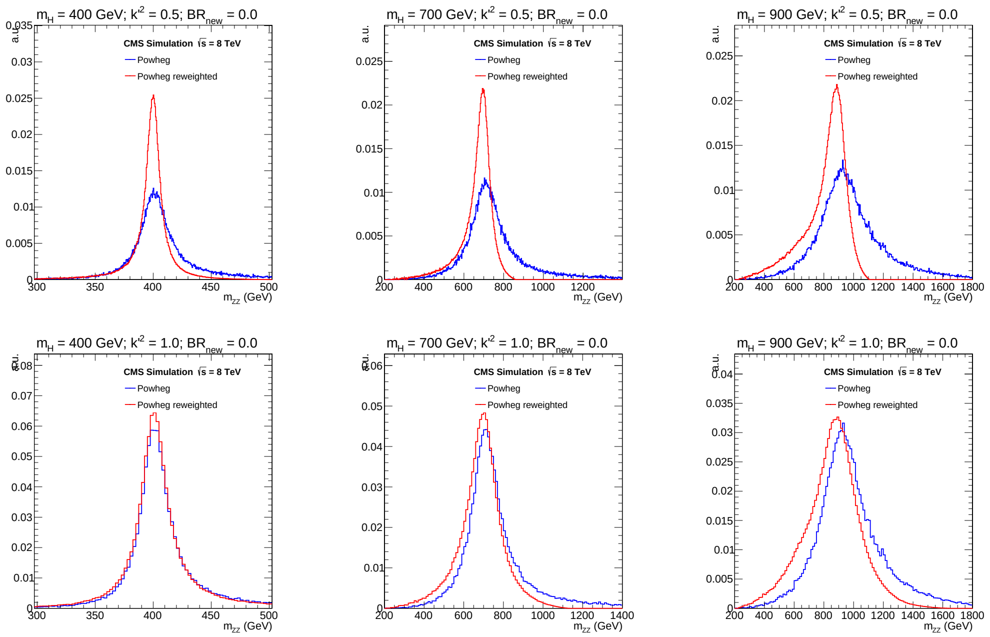


Figure 6.2: The effect of the interference reweighting on the 8 TeV vector boson fusion Higgs signal MC samples, for $C'^2 = 0.5$ (top) and for the SM (bottom), i.e. $C'^2 = 1$.

- *additive approach*

$$\frac{d\sigma^{\mathrm{NNLO}}_{\mathrm{eff}}}{dM_{ZZ}} = \frac{d\sigma^{\mathrm{NNLO}}}{dM_{ZZ}}(S) + \frac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(I) + \frac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(B)$$

- *multiplicative* approach

$$\frac{d\sigma^{\mathrm{NNLO}}_{\mathrm{eff}}}{dM_{ZZ}} = K_D\left[\frac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(S) + \frac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(I)\right] + \frac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(B),$$

$$K_D = \frac{\dfrac{d\sigma^{\mathrm{NNLO}}}{dM_{ZZ}}(S)}{\dfrac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(S)}$$

- *intermediate* approach

$$\frac{d\sigma^{\mathrm{NNLO}}_{\mathrm{eff}}}{dM_{ZZ}} = K_D\frac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(S) + \sqrt{K_D^{gg}}\frac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(I) + \frac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(B),$$

$$K_D^{gg} = \frac{\dfrac{d\sigma^{\mathrm{NNLO}}}{dM_{ZZ}}(gg \to H(g) \to ZZ(g))}{\dfrac{d\sigma^{\mathrm{LO}}}{dM_{ZZ}}(gg \to H \to ZZ)},$$

$$K_D = K_D^{gg} + K_D^{\mathrm{rest}}$$

In the equations above, K_D are the full NNLO K-factors while K_D^{gg} are the

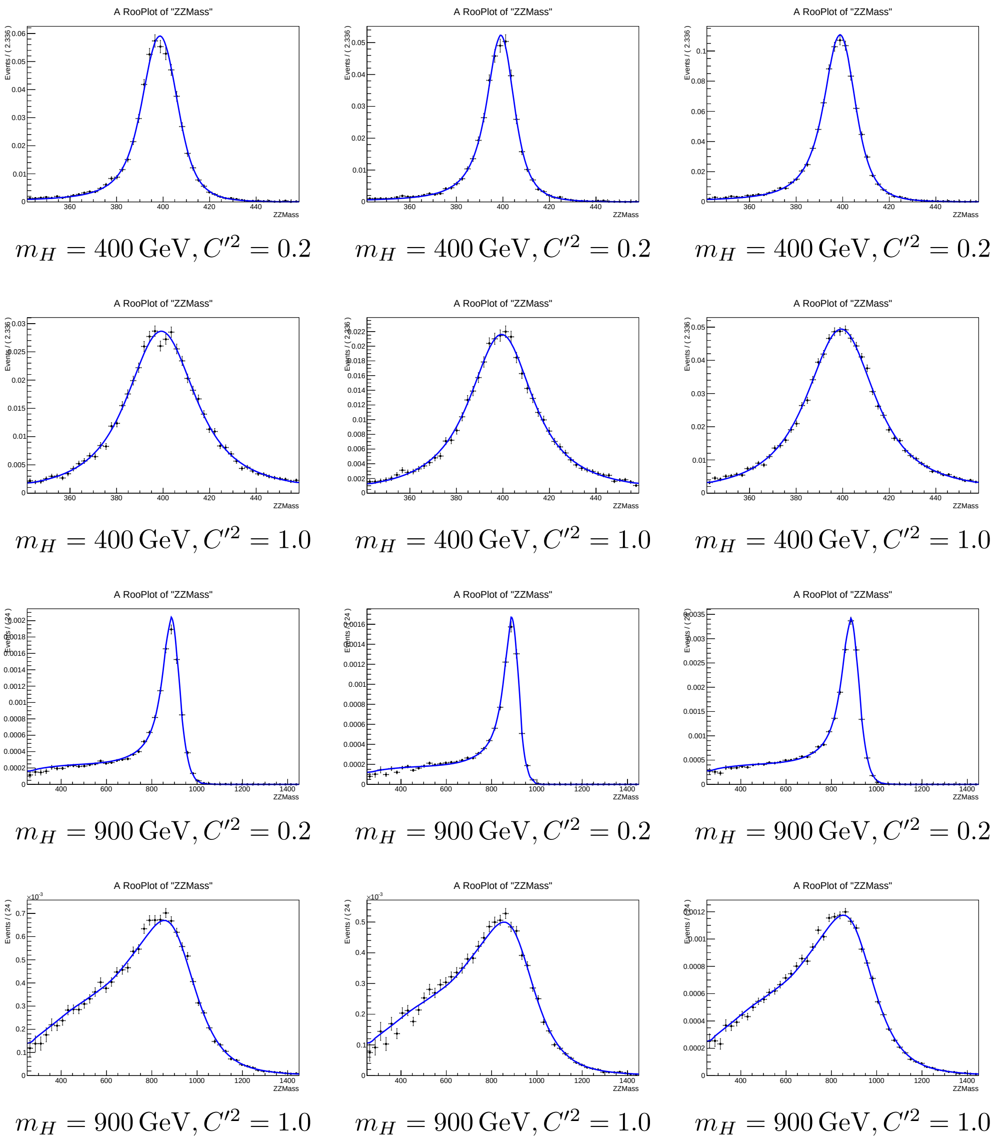
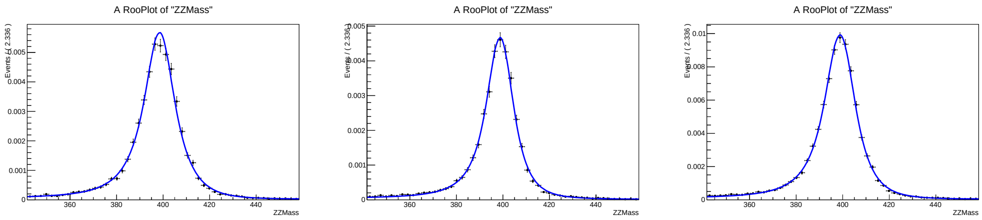
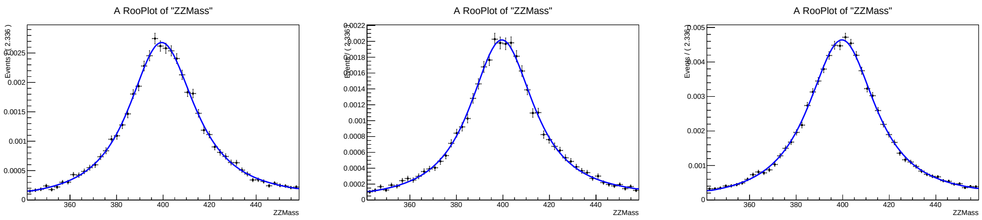


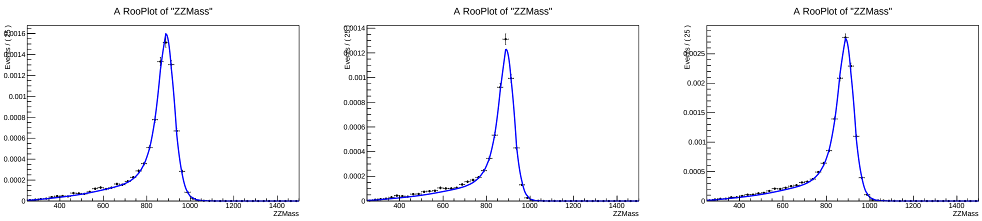
Figure 6.3: Examples of $m_{4\ell}$ shape models for the gluon fusion signal (8 TeV), for the 4μ (left), $4e$ (center) and $2e2\mu$ (right) final states, for two masses and two different values of C'.



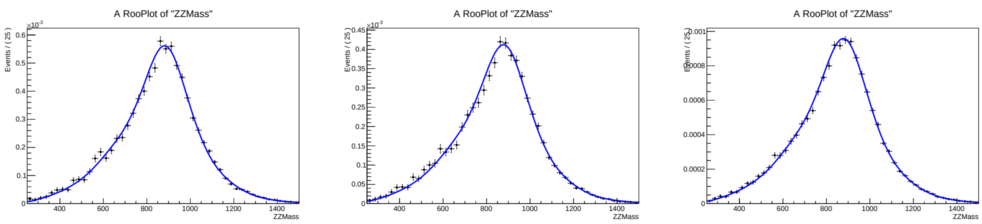
$m_H = 400\,\mathrm{GeV}, C'^2 = 0.2$ $m_H = 400\,\mathrm{GeV}, C'^2 = 0.2$ $m_H = 400\,\mathrm{GeV}, C'^2 = 0.2$



$m_H = 400\,\mathrm{GeV}, C'^2 = 1.0$ $m_H = 400\,\mathrm{GeV}, C'^2 = 1.0$ $m_H = 400\,\mathrm{GeV}, C'^2 = 1.0$



$m_H = 900\,\mathrm{GeV}, C'^2 = 0.2$ $m_H = 900\,\mathrm{GeV}, C'^2 = 0.2$ $m_H = 900\,\mathrm{GeV}, C'^2 = 0.2$



$m_H = 900\,\mathrm{GeV}, C'^2 = 1.0$ $m_H = 900\,\mathrm{GeV}, C'^2 = 1.0$ $m_H = 900\,\mathrm{GeV}, C'^2 = 1.0$

Figure 6.4: Examples of $m_{4\ell}$ shape models for the vector boson fusion signal (8 TeV), for the 4μ (left), $4e$ (center) and $2e2\mu$ (right) final states, for two masses and two different values of C'.

K-factors calculated using only the amplitudes involving gluons.

The intermediate approach has been used for the nominal shape, while the additive and multiplicative are used to build shape variations which are taken into account as systematic uncertainty. The difference between the three different approaches is practically small at moderate masses, but it becomes large when m_H approaches 1 TeV. In figure 6.5 examples of the envelope of the three approaches are shown.
The same uncertainty is assigned also in the vector boson fusion channel, where there is no specific prescription.

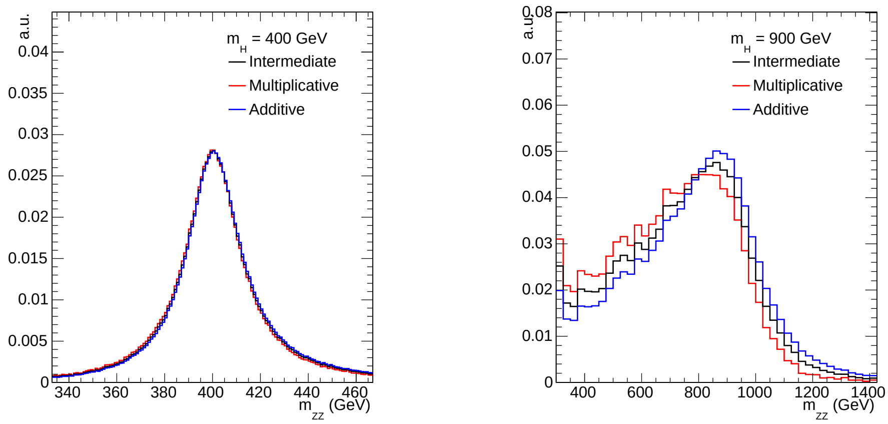


Figure 6.5: The comparison of the three different approaches for the calculation of the interference in the shape of the signal, with the GG2VV generator at $\sqrt{s} = 8\,\mathrm{TeV}$ and for different masses.

6.2.2 $qq' \to qq'ZZ$ background estimation

The contribution of the subleading $qq' \to qq'ZZ \to 4\ell$ background has been included in the high mass analysis, with respect to the Legacy analysis, thanks to the availability of the MC PHANTOM samples. The prediction from the simulation is rescaled with a 6% constant K-factor to match the NLO cross section. It has to be noticed that the rate of this process is negligible in the 0/1-jet category, with respect to the other dominant backgrounds ($qq \to ZZ$, $gg \to ZZ$, $Z + X$). In the 2-jet category it accounts for about 8% of the total background.

The $m_{4\ell}$ shape of the $qq' \to qq'ZZ$ process is shown in figure 6.6. It is derived with an empirical parameterization similar to the one used for the $qq \to ZZ$ and $gg \to ZZ$ processes (see 4.5.1). The shape of the $\mathcal{D}_{\mathrm{bkg}}^{\mathrm{kin}}$ is found to be very similar to the $qq \to ZZ$ background, while specific shapes for the p_{T}^{H} and $\mathcal{D}_{\mathrm{jet}}$ are derived from simulation.

6.2.3 Matrix element discriminant in the 2-jet category

In agreement with the Legacy analysis, events are categorized according to the jet multiplicity into the 0/1-jet and 2-jet category (see 4.4.3).

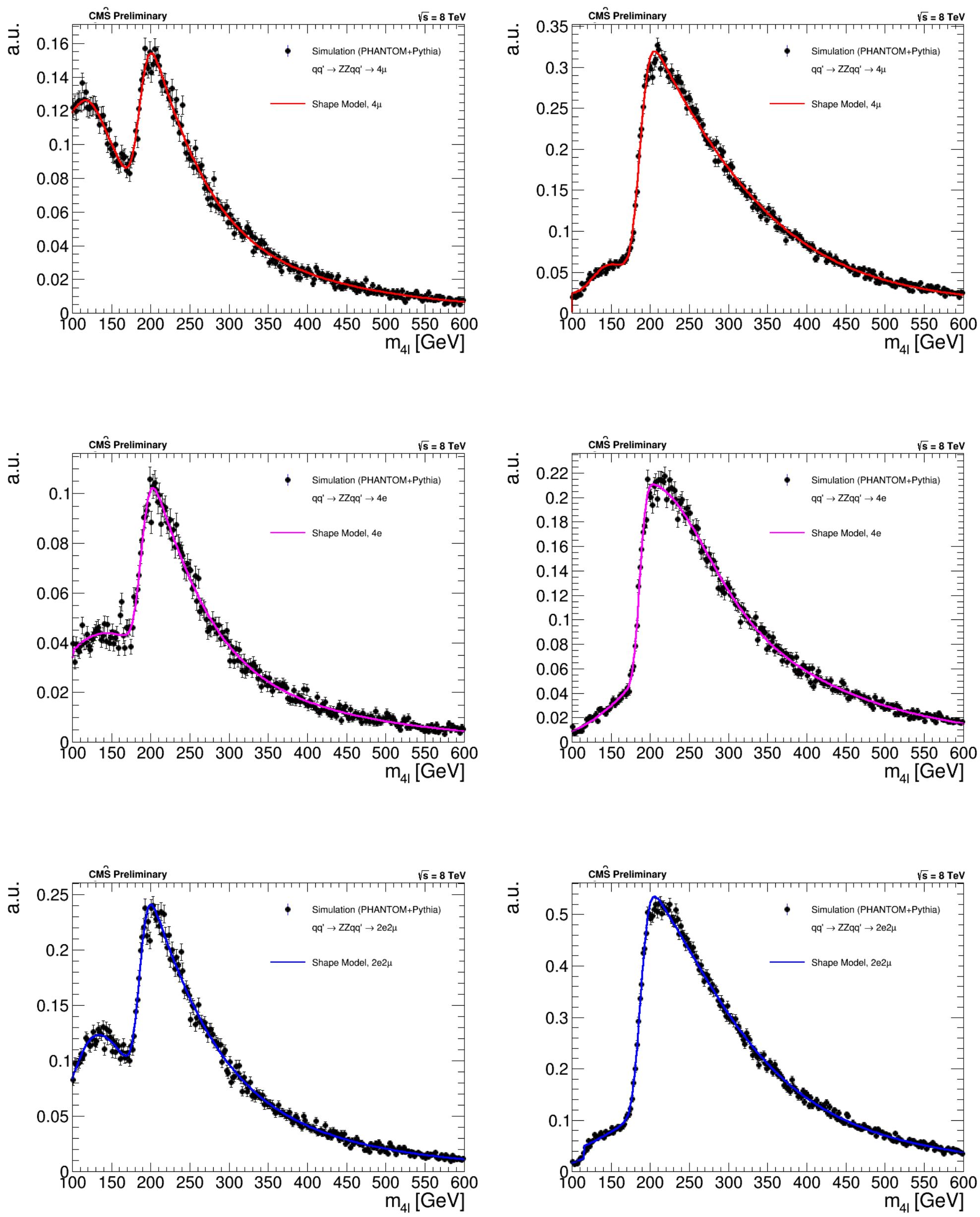


Figure 6.6: Model of $qq' \to qq'ZZ$ background ($\sqrt{s} = 8\,\mathrm{TeV}$) for 4μ (top), $4e$ (center) and $2e2\mu$ (bottom) channels, in the 0,1 jet category (left) and ≥ 2 jet category (right).

The procedure remains unchanged in the 0/1-jet category, where the transverse momentum of the four-lepton system is used to discriminate between the different production mechanisms. In the 2-jet category, instead, the Fisher discriminant is replaced with a matrix element based discriminant which uses the full kinematic information of the Higgs candidate plus the two associated jets. This discriminant is implemented in the MELA package and uses ME calculated with the JHUGEN program. As for the Fisher discriminant, it is designed to separate the VBF and the gluon fusion plus two jets processes.
Similarly to equation 4.6, we define:

$$\mathcal{D}_{\text{jet}} = \frac{\mathcal{P}_{\text{VBF}}}{\mathcal{P}_{\text{VBF}} + \mathcal{P}_{\text{ggH+jj}}} \tag{6.2}$$

where $\mathcal{P}_{\text{VBF}}$ and $\mathcal{P}_{\text{ggH+jj}}$ are the probabilities for the VBF and the gluon fusion processes respectively.
As for the Fisher discriminant, the templates are built from MC or control regions for the reducible background. For the signal, templates are filled with events coming from all the MC samples available corresponding to different mass hypotheses. The two-dimensional templates in the $(m_{4\ell}, \mathcal{D}_{\text{jet}})$ plane for the gluon fusion and the vector boson fusion are shown in figure 6.7 together with their projections in the $\mathcal{D}_{\text{jet}}$ variable. The peak at very small values of $\mathcal{D}_{\text{jet}}$ for the VBF process corresponds to events in which the two selected jets are not the ones coming from the fragmentation of the two associated quarks of the VBF amplitude.
The systematic effect of using different MC generators and tunings, as well variations induced by different PDF sets and factorization/renormalization scale and jet energy scale uncertainties are investigated and found to affect very little the shape of the discriminant. Conservatively, the maximum variation is taken as systematic uncertainty.

6.2.4 Effect of h(125) in the BSM high mass search

In the EWK singlet models there are two physical Higgs-like states, one corresponding to the SM doublet field and one to the singlet field of the "hidden sector". In the search described in this chapter, the SM Higgs is identified with the newly discovered resonance with mass 125 GeV, $h(125)$, and the search is focused on a high mass singlet state.
The presence of the 125 GeV boson is used to put an indirect constraint on the C' parameter. From the measurement of the signal strength in the $H \to ZZ^{(*)} \to 4\ell$ channel (4.10.3) we obtain $C'^2 < 0.61$ at 95% C.L. Using the most recent CMS combination [80] for the signal strength available in this moment, we get $C'^2 < 0.26$ at 95% C.L.

Given the very good resolution in the four-lepton channel, the $h(125)$ on-shell cross section is negligible in the mass range of interest of this search, i.e. $m_H > 145\,\text{GeV}$. However, it is well known that the off-shell cross section and the interference of the low mass signal with the background are sizeable at high ZZ invariant masses [20, 75]. Figure 6.8 shows the ratio between the cross section of a 125 GeV Higgs boson plus background plus interference to the background only cross section,

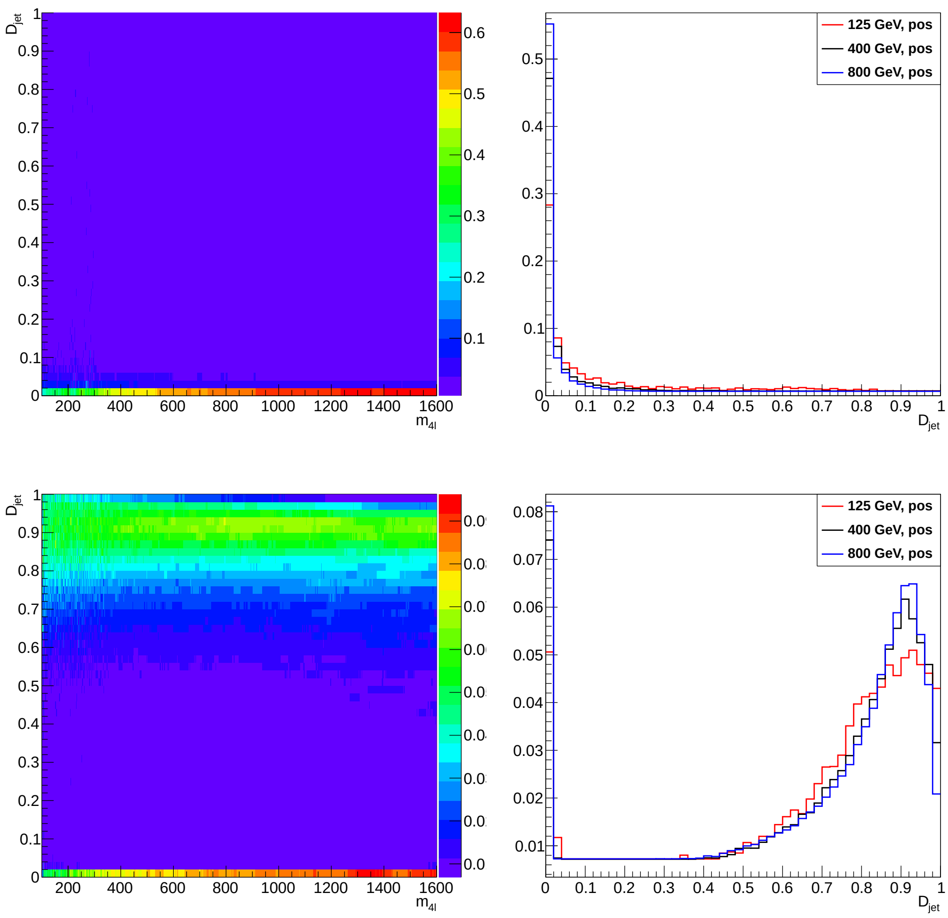
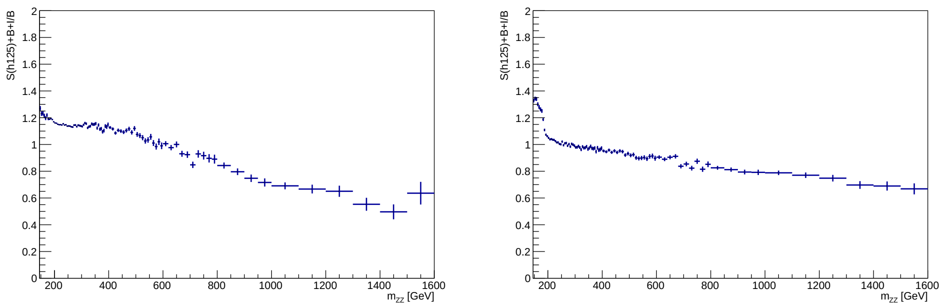


Figure 6.7: The two-dimensional templates in the $m_{4\ell}, \mathcal{D}_{\text{jet}}$ variables (left) and their projections in the $\mathcal{D}_{\text{jet}}$ variable (right), for the gluon fusion (top) and the vector boson fusion (bottom) processes.

for the gluon fusion and VBF channels at 8 TeV. It can be noticed that up to 300-400 GeV there is a positive contribution coming from the off-shell cross section, while above the negative interference starts to dominate. The presence of the low mass Higgs boson is considered as a missing contribution in the $gg \to ZZ$ and $qq' \to qq'ZZ$ backgrounds. This is taken into account as a systematic uncertainty assigned to the background rates, based on figure 6.8. This is believed to be a conservative approach for different reasons. First of all it is extracted for the SM ($C'^2 = 1$) case, where there is the largest off-shell signal plus interference contribution. Secondly, at very high masses, where the negative interference dominates, a complete description of the contribution of the Higgs at 125 GeV would result in a smaller background and thus a higher sensitivity to the search for the singlet resonance.



ggH channel | VBF channel

Figure 6.8: The ratio of the signal plus background plus iterference cross section for a 125 GeV Higgs boson over the background only cross section, for the gluon fusion channel (left) and VBF channel (right) at $\sqrt{s} = 8\,\mathrm{TeV}$.

6.2.5 Results

The observed number of events and their properties are interpreted in the context of the EWS model, where the relevant parameters are:

- $\boldsymbol{m_H}$ is the mass of the EWS state. It is assumed that $m_H > m_h$ where h is the discovered boson with mass around 125 GeV. This parameter is scanned over the range $145\,\mathrm{GeV} < m_H < 1000\,\mathrm{GeV}$;
- $\boldsymbol{C'^2}$ is the singlet state contribution to the electroweak symmetry breaking and regulates its signal strength as in equation 1.25. This parameter is investigated in the range 0.01 to 1 (SM);
- $\mathbf{BR_{new}}$ is the contribution to the branching ratio of the singlet due to new decay channels not present in the SM, like $H \to hh$. The scan is done in the range from $\mathrm{BR_{new}} = 0$ to $\mathrm{BR_{new}} = 0.5$.

The search covers the region in which $\Gamma_H \leq \Gamma_{\mathrm{SM}}$.

In figure 6.9 the 95% C.L. observed upper limits on the signal strength, defined as the ratio of the measured cross section to the expectation from the model, are shown as a function of the singlet mass, for two different C'^2 values and for $\mathrm{BR_{new}} = 0$. The bands corresponding to 1-2 sigmas uncertainty of the expected limit are also shown. In figure 6.10 the same kind of plot is shown for several values of C' and two different values of $\mathrm{BR_{new}}$. The uncertainty bands are removed here to improve the readability of the plot.
The mass hypotheses for which the limit is smaller than 1 are excluded at 95% C.L.

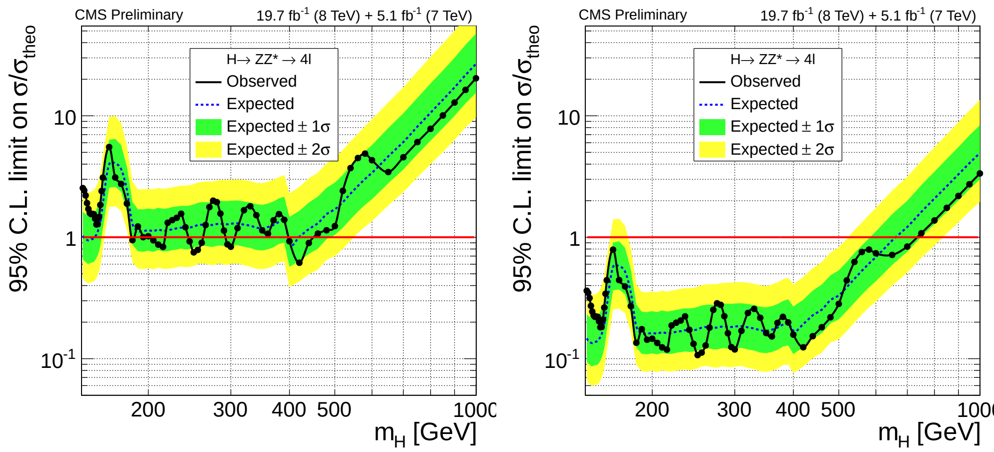
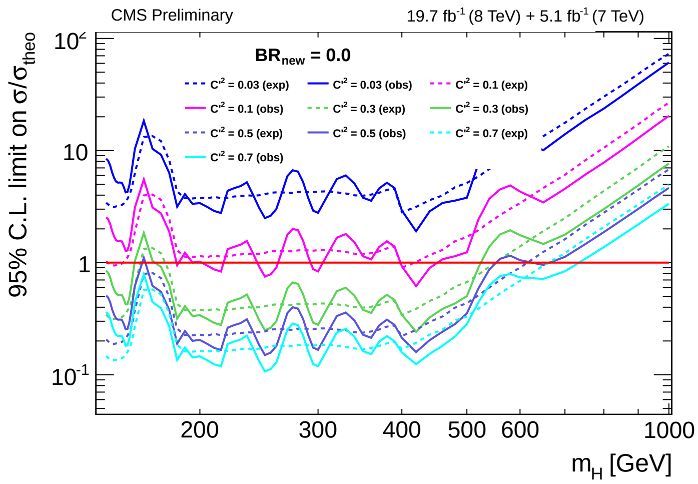


Figure 6.9: The 95% C.L. upper limits on the signal strength as a function of the singlet mass, for $C'^2 = 0.1$ (left) and $C'^2 = 0.7$ (right), for $\mathrm{BR_{new}} = 0$. Expected limits are in dashed blue line, observed limits in solid black line. The green (yellow) band represents the 1 sigma (2 sigma) uncertainty.

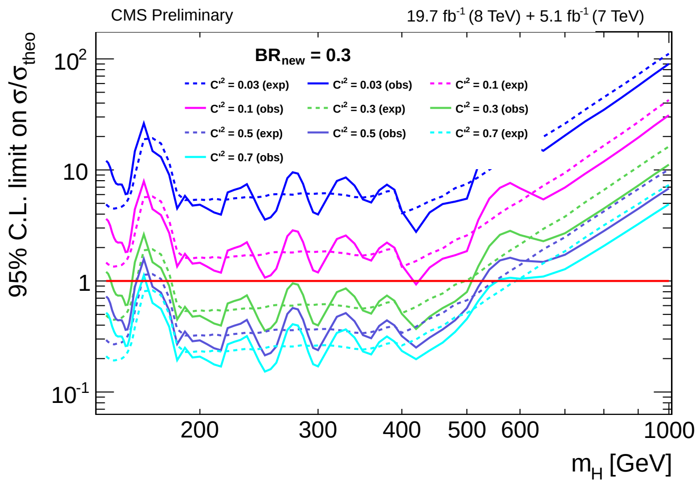
No significant excess is seen in the whole mass range under investigation: some fluctuations are visible, always well within the 1-sigma uncertainty. The largest excess is at a mass around 145 GeV, corresponding to a local significance still below the 3-sigma level. A slight deficit of events is seen at very high masses ($m_H > 600\,\mathrm{GeV}$), with no event observed above 750 GeV. We also observe the typical dependence of the expected limit (and the observed limit as well) in the region 160-200 GeV, where there is a steep decrease of the branching ratio of the ZZ decay in favour of the WW mode with both Ws on-shell. In the vicinity of the $2m_Z$ threshold, the branching ratio, and also the sensitivity of the ZZ decay channel, rises again and reaches a plateau at about 200 GeV.
From figure 6.10 is already clear that values of $C'^2 \lesssim 0.1$ (for any $\mathrm{BR_{new}}$) are not excluded by this search.

Figure 6.11 shows the exclusion limits in the $(C'^2, \mathrm{BR_{new}})$ plane, while in figure 6.12 the exclusion is shown in the (m_H, C'^2) plane.

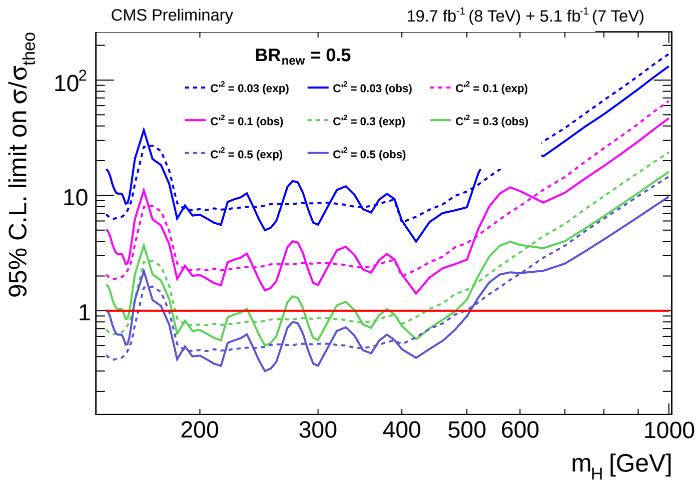
From these plots we see how the direct search is indeed able to exclude a large part of the parameter space of the model which is not covered by the indirect limit, coming from the measurement of the SM Higgs at 125 GeV, for singlet masses below about 700 GeV. For larger values of the heavy Higgs, the direct search looses sensitivity due to the small expected number of signal events. From about 820 GeV and above, no region in the $(C'^2, \mathrm{BR_{new}})$ plane is constrained by the direct search.



$\mathrm{BR}_{\mathrm{new}} = 0$

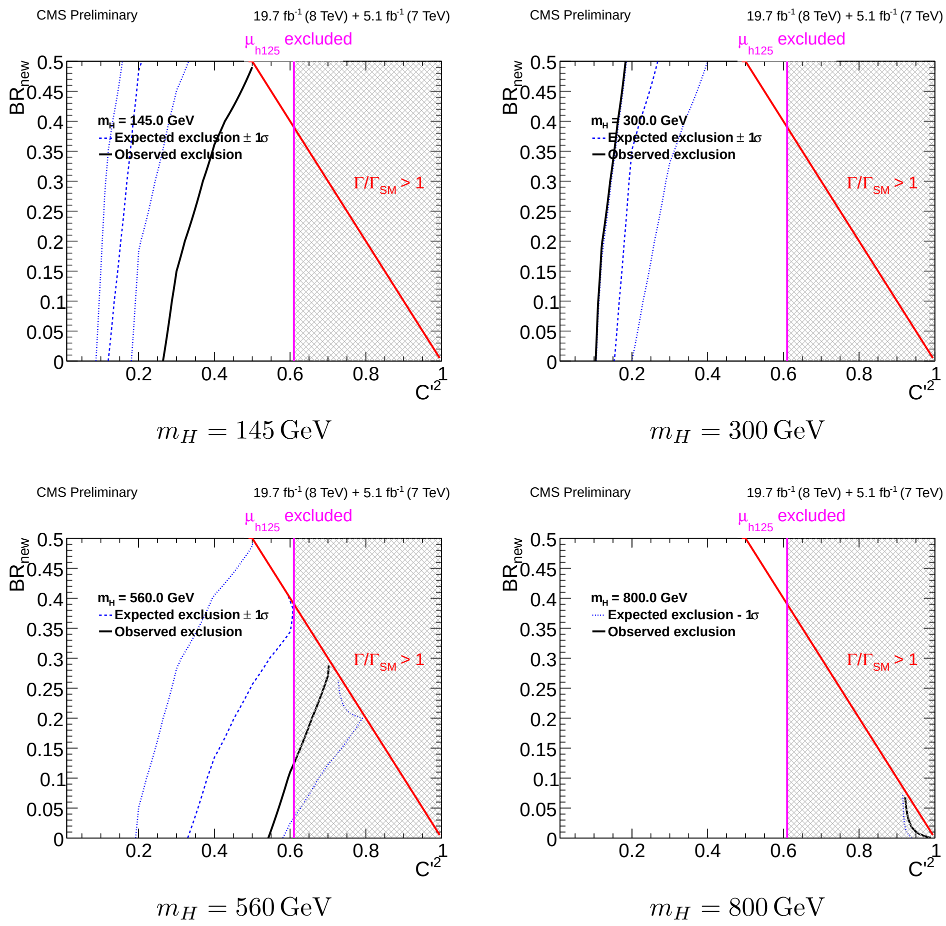


$\mathrm{BR}_{\mathrm{new}} = 0.3$



$\mathrm{BR}_{\mathrm{new}} = 0.5$

Figure 6.10: The 95% C.L. upper limits on the signal strength as a function of the singlet mass, for $\mathrm{BR}_{\mathrm{new}} = 0.0$ (top), $\mathrm{BR}_{\mathrm{new}} = 0.3$ (middle) and $\mathrm{BR}_{\mathrm{new}} = 0.5$ (bottom), for different values of C'^2. Expected limits are in dashed lines, observed limits in solid lines.



$m_H = 145\,\mathrm{GeV}$ $m_H = 300\,\mathrm{GeV}$

$m_H = 560\,\mathrm{GeV}$ $m_H = 800\,\mathrm{GeV}$

Figure 6.11: The exclusion in the $(C'^2, \mathrm{BR_{new}})$ parameter space for four different hypothesis of singlet mass. The grey band represents the area excluded by the measurement of the signal strength of $h(125)$.

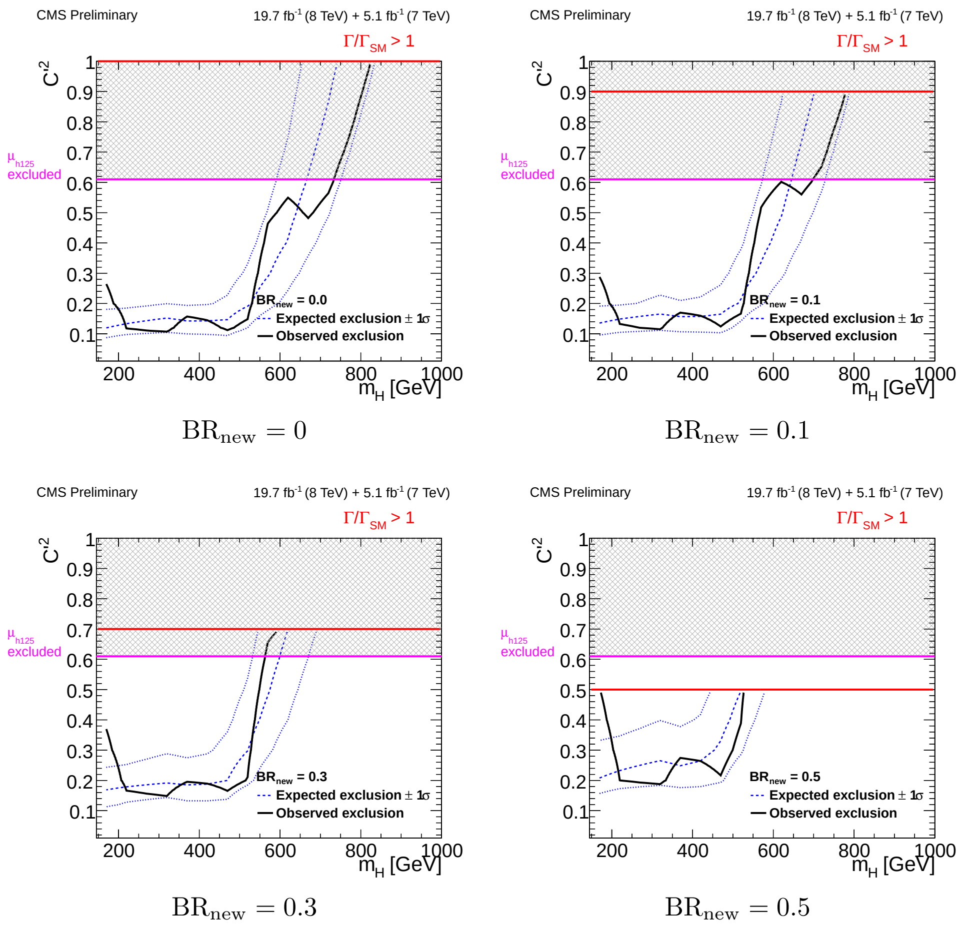


Figure 6.12: The exclusion in the (m_H, C'^2) parameter space for four different $\mathrm{BR}_{\mathrm{new}}$ values. The grey band represents the area excluded by the measurement of the signal strength of $h(125)$.

6.3 The combination of the WW and ZZ channels

The data in the $H \to 4\ell$ channel are combined with other WW and ZZ final states to search for an additional Higgs-like boson in the mass range from 145 GeV to 1 TeV. After the discovery of a low mass Higgs, with properties measured in good agreement with the prediction of the SM, the focus is on the interpretation of the data in the EWS scenario. Nevertheless, exclusion limits for the SM Higgs boson at high mass are also estimated.

Several final states are investigated, which are briefly described in the following.

- $H \to WW \to \ell\nu\ell\nu$
  In this channel the heavy Higgs decays in two W bosons, both of which decay leptonically, resulting in a signature with two isolated, high-p_{T} leptons (electrons and muons) and missing transverse energy. The analysis is described in details in Ref. [81].

- $H \to WW \to \ell\nu qq$
  In this channel one of the two heavy Higgs daughters is required to decay leptonically, in order to trigger the event, while the other decays hadronically. This results in a larger branching ratio with respect to the leptonic decay, but also a larger background coming from the non-resonant SM W+jets production. The final state thus contains one isolated, high-p_{T} electron or muon with $E_{\mathrm{T}}^{\mathrm{miss}}$ in association with a pair of jets.
  For $m_H > 600\,\mathrm{GeV}$, when the heavy Higgs daughters have large transverse momenta, the jets originating from the hadronization of the final state quarks are usually very close to each other. For this reason, the analysis makes use of special techniques to identify the W boson whose decay products are reconstructed in a single jet (merged category).

- $H \to ZZ \to 2\ell 2\ell'$
  A heavy Higgs boson is searched in the ZZ decay mode using fully leptonic final states, where one of the Z ($Z \to \ell\ell$) is required to decay to dielectron and dimuon and the other ($Z \to \ell'\ell'$) can decay to electrons, muons or taus. The final state is composed by four well-isolate leptons with high transverse momentum, with the main background coming from the non-resonant ZZ production. More details about the analysis using the $2\ell 2\ell$ final state are in chapter 4 as well as in Ref. [15].
  It has to be remarked that in the $2\ell 2\tau$ final state the sub-channels in which both taus decay to muons or to electrons are removed to avoid any possible overlap with the $2\ell 2\ell$ analysis.

- $H \to ZZ \to 2\ell 2\nu$
  In this channel one Z decays leptonically (muons and electrons) and the other decays to neutrinos, thus a large missing transverse energy is required in the final state together with two well-isolated leptons with high p_{T}. The analysis is described in detail in Ref. [82].

- $H \to ZZ \to 2\ell 2q$
  This channel, with one Z decaying hadronically and the other leptonically

(muons and electrons), has the largest branching ratio of all the ZZ channels, but has also a large background from the Z+jets production. As in the $WW \to \ell\nu qq$ channel, a dedicated strategy is used to account for the fact that at very high masses the two jets originating from a boosted Z boson may be reconstructed in a single "merged" jet.

Figure 6.13 shows the combined result for the SM interpretation of the search with the breakdown of the different channels.

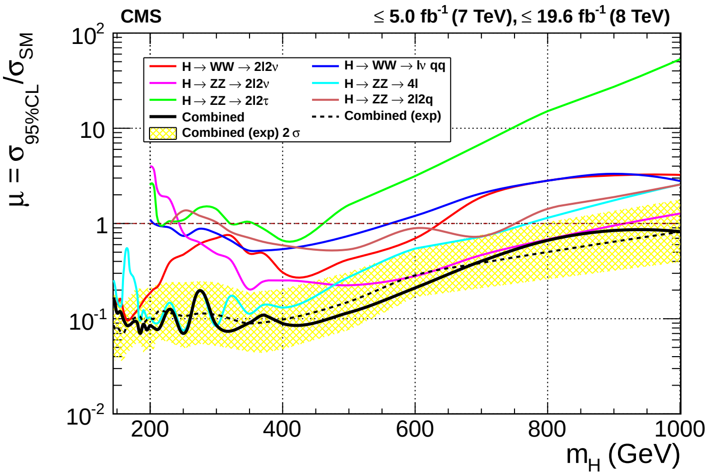


Figure 6.13: The 95% C.L. upper limits on the signal strength of a SM Higgs boson, as a function of the hypothesized mass. The contribution of each channel is shown with different colors. The red solid (dashed) line is the observed (expected) limit corresponding to the combination of all the channels.

No significant excess is seen in the whole mass range under investigation and a SM-like Higgs boson is excluded in the whole range of the search, i.e. 145-1000 GeV, thus extending the mass region excluded by CMS in the previous search [77]. From figure 6.13, it can be noticed that the dominant channel up to about 180 GeV is the $WW \to \ell\nu\ell\nu$, which has a larger branching ratio than the $ZZ \to 2\ell 2\ell$ channel. In the region 180-500 GeV, the $ZZ \to 2\ell 2\ell$ is the most sensitive analysis, while for masses larger than 500 GeV the $ZZ \to 2\ell 2\nu$ is the dominant channel, thanks to the larger branching ratio.

The exclusion limits for the EWS model in the parameter space (m_H, C'^2) are shown, for different values of $\mathrm{BR}_{\mathrm{new}}$, in figure 6.14. The same limits are also shown for the ZZ and WW channels separately.
Both for the SM and BSM case, the ZZ channels exclude a larger portion of the parameter space with respect to the WW channels, due to the presence of the $ZZ \to 2\ell 2\ell$ final state, which is the most sensitive analysis in the search for a high mass Higgs.

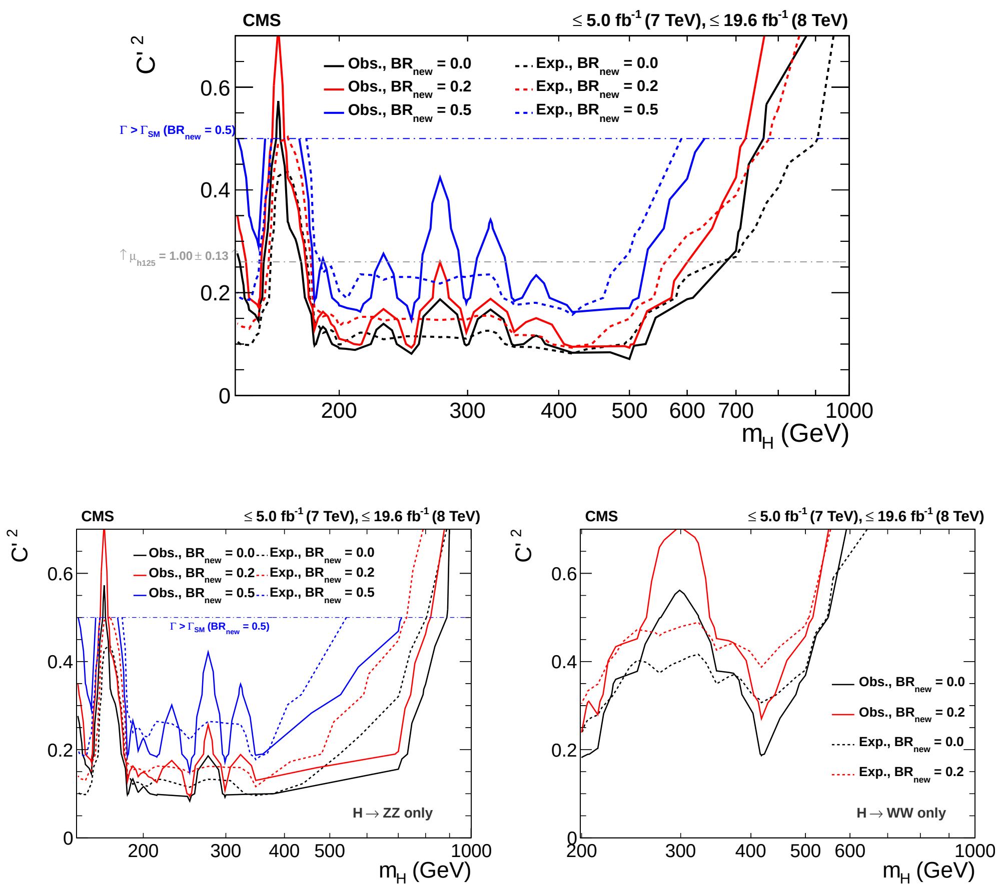


Figure 6.14: The 95% C.L. upper limits on the signal strength of a SM Higgs boson, as a function of the hypothesized mass and the C'^2 parameter, for all channels participating to the combination (top), ZZ channels only (left) and WW channels only (right). Different values of $\mathrm{BR}_{\mathrm{new}}$ are investigated.

Summary

In this thesis I presented the investigation of the Higgs sector in the $H \to 4\ell$ channel using the data of the Run 1 of the LHC collected by the CMS experiment.
This channel has been a "leading horse" in the path to the discovery of a new particle with mass around 125 GeV, which is compatible with the long sought Higgs boson of the SM. I have had the priceless pleasure to participate to this exceptional discovery, which alone is enough to define the Run 1 as a success.
Its properties, such as the mass, the width, the couplings and the spin-parity, have been studied in detail in the four lepton channel, using the best knowledge of the detector and the full dataset available. This "legacy" analysis has been published and it is documented in this thesis (chapter 4), as it is one of my main contributions to the collaboration in these three years.
As a crucial ingredient of the $H \to 4\ell$ analysis, but also part of a more general task in CMS, I developed corrections to the muon momentum scale and resolution using $Z \to \mu^+\mu^-$ events. The details of the algorithm used and the validation of the results have been presented in chapter 3, together with examples of the impact of such corrections on the physics analyses with muons in CMS.
With the discovery of a Higgs boson at low mass, the question arises whether this particle is the only responsible for the electroweak symmetry breaking or if it is part of a non-minimal scalar sector, which would possibly imply new physics. The search for an additional Higgs singlet state at high mass has been described in chapter 6: in the absence of any significant excess, I derived exclusion limits in the parameter space of the model.
The effect of the interference of the signal with the background is crucial in the heavy Higgs phenomenology, when the intrisic width of the resonance becomes large. For this reason I carried on a dedicated study using leading order generators, which is presented in chpater 5.

The LHC has performed incredibly well in the Run 1, delivering collisions at an unprecendented energy and luminosity. Thanks to the outstanding performances of both accelerator and detectors, and the effort of thousands of scientists, a Higgs has been discovered, which sets a milestone in the history of particle physics and, at the same time, shades light on the future research in this field.
In particular, the performances in the $H \to 4\ell$ channel, which is the main character of the story that I have told in these pages, have exceeded any optimistic expectation. The end of the story, however, is yet to be written...
With the upcoming Run 2, the energy and peak luminosity will further increase, pushing to new limits the sensitivity of the experiments to the discovery of new particles. Moreover, the accuracy in the measurement of the couplings of the Higgs, which is

still limited by statistics in many channels, will sensibly improve and is expected to reach about 10% in many measurements.
With these perspectives, the future runs of the LHC look like a great opportunity for the scientific community to give answers to the open questions of physics.

Appendix A

Statistical procedure for the calculation of the upper limits

The statistical procedure used in the extraction of the limits is known as *modified frequentist method*, oftet referred to as CL_s and it is outlined in the following.

The expected yields for the signal and background processes are denoted as $s(\theta)$ and $b(\theta)$ respectively, where θ represents the set of *nuisance parameters* which are used to implement the experimental uncertainties affecting the measurement.
The first step is to construct the likelihood function:

$$\mathcal{L}(\mathrm{data}|\mu,\theta) = \mathrm{Poisson}(\mathrm{data}|\mu\cdot s(\theta)+b(\theta))\cdot p(\tilde{\theta}|\theta) \tag{A.1}$$

In the above equation, "data" may refer to the actual observation or to *pseudo data*, generated to assess the distribution of the test statistic, as will be discussed later. $p(\tilde{\theta}|\theta)$ is the probability density function (pdf) of the nuisance parameter θ, where $\tilde{\theta}$ is the default value.
In the case of a binned model for the observable under scrutiny, the poissonian term in equation A.1 is a product over the bins of Poisson probabilities:

$$\mathrm{Poisson}(\mathrm{data}|\mu\cdot s(\theta)+b(\theta)) = \prod_i \frac{(\mu s_i+b_i)^{n_i}}{n_i!}e^{-(\mu s_i+b_i)} \tag{A.2}$$

where, in the equation above, n_i is the number of events observed in the bin i, s_i and b_i are the expected number of events from signal and background processes.
For an unbinned model describing N events in the data sample we have:

$$\mathrm{Poisson}(\mathrm{data}|\mu\cdot s(\theta)+b(\theta)) = N^{-1}\prod_i (\mu S f_s(x_i)+B f_b(x_i))\cdot e^{-(\mu S+B)} \tag{A.3}$$

In equation A.3 S and B are the total expected yields for the signal and backgroud processes respectively, $f_{s(b)}(x_i)$ are the pdfs for the signal (background) processes describing the observable x.

To assess the compatibility of the data with the signal + background hypothesis, where the signal is eventually scaled by μ, the following test statistic q_μ is defined:

$$q_\mu = -2\ln Q = -2\ln\frac{\mathcal{L}(\mathrm{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\mathrm{data}|\hat{\mu},\hat{\theta})},\ \hat{\mu} \leq \mu \tag{A.4}$$

The quantity Q is known as the *profile likelihood ratio*. In equation A.4, $\hat{\theta}_\mu$ is the conditional likelihood estimator for the nuisances, given the observation "data" and the signal strength μ, while $\hat{\mu}, \hat{\theta}$ are the values the gives the global maximum of the likelihood.
The condition $\hat{\mu} \leq \mu$ is imposed to obtained a one-sided limit, meaning that upward fluctuations, if any, will not be considered as against the signal hypothesis μ.

The pdf $f(q_\mu|\mu,\hat{\theta}_\mu^{\mathrm{obs}})$ and $f(q_\mu|0,\hat{\theta}_0^{\mathrm{obs}})$ are constructed by tossing many Monte Carlo pseudo experiments, where $\hat{\theta}_\mu^{\mathrm{obs}}$ and $\hat{\theta}_0^{\mathrm{obs}}$ are the values of the nuisances which best describe the observations under the signal+background (μ) and background only ($\mu = 0$) hypothesis. The nuisance parameters are kept fixed in the generation of pseudo-data but are free to float in the fit to determine the test statistic of equation A.4.
Despite being the most reliable method, the generation of pseudo experiments may be very time consuming, especially in the case of complicated models, like the one used in the $H \to 4\ell$ analysis and described in section 4.8. However, it has been demonstrated [60] that simplified formulas exist for $f(q_\mu|\mu,\hat{\theta}_\mu^{\mathrm{obs}})$ which allows to speed up the calculation of the upper limits and their uncertainties. They are based on the properties of the *Asimov dataset*, [1] which is the dataset with the expected signal, background and the nominal value for the nuisance parameters, i.e. with all the fluctuations set to 0.
With this philosophy in mind, one can use a *pseudo-Asimov* technique, in which instead of defining a dataset with one entry per bin i, with weight equal to the number of expected events ν_i (Asimov dataset), one generates from the pdfs N entries each with weight equal to ν_i/N (N is usually about 100 times the number of expected events).
The pseudo-Asimov method has been demonstrated to be lessed biased with respect to the Asimov in case of small number of events, but is affected by the intrinsic statistical uncertainty related to the sampling of the distributions.

Given the observed value for the test statistic q_μ^{obs}, two p-values are computed, associated to the compatibility of the observations with the signal+background and background only hypothesis, namely p_μ and p_b:

$$p_\mu = P(q_\mu \geq q_\mu^{\mathrm{obs}}|\mathrm{signal+background}) = \int_{q_\mu^{\mathrm{obs}}}^{\infty} f(q_\mu|\mu,\hat{\theta}_\mu^{\mathrm{obs}})dq_\mu \tag{A.5}$$

$$1 - p_b = P(q_\mu \geq q_\mu^{\mathrm{obs}}|\mathrm{background}) = \int_{q_0^{\mathrm{obs}}}^{\infty} f(q_\mu|0,\hat{\theta}_0^{\mathrm{obs}})dq_\mu \tag{A.6}$$

[1]The name is inspired by the tale *Franchise* by Isaac Asimov. In it, during elections, the electorate was replaced by the single most representitive voter.

From equations A.5 and A.6 one constructs the following quantity:

$$\mathrm{CL}_s(\mu) = \frac{p_\mu}{1 - p_b} \tag{A.7}$$

Fixing $\mu = 1$ in equation A.7 is used to infer whether the Higgs-like resonance H is excluded at a confidence level (C.L.) α if $\mathrm{CL}_s \leq \alpha$.
For the purposes of quoting 95% C.L. limits on a certain μ, which has become a well established convention in high energy physics, one adjusts μ until $\mathrm{CL}_s = 0.05$.

Based on Ref. [60], the 95% C.L. upper limit is given by:

$$\mu_{\mathrm{up}} = \hat{\mu} + \sigma\Phi^{-1}(0.95) \tag{A.8}$$

where Φ is the cumulative distribution of the standard (normal) Gaussian and σ is defined as:

$$\sigma = \frac{\mu^2}{q_{\mu,\mathrm{PA}}} \tag{A.9}$$

In the equation above $q_{\mu,\mathrm{PA}}$ is the test statistic evaluated using the pseudo-Asimov dataset.
The median expected limit is:

$$\mathrm{med}[\mu_{up}] = \sigma\Phi^{-1}(0.95), \tag{A.10}$$

and the $\pm N\sigma$ uncertainty bands are defined as:

$$\mathrm{band}_{N\sigma} = \sigma \cdot (\Phi^{-1}(0.95) \pm N) \tag{A.11}$$

References

[1] S. L. Glashow, "Partial-symmetries of weak interactions", *Nuclear Physics* **22** (1961), no. 4, 579 – 588, doi:http://dx.doi.org/10.1016/0029-5582(61)90469-2.

[2] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266, doi:10.1103/PhysRevLett.19.1264.

[3] A. Salam, "Weak and Electromagnetic Interactions", *Conf.Proc.* **C680519** (1968) 367–377.

[4] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Non-Abelian Gauge Theories", *Phys. Rev. Lett.* **30** (Jun, 1973) 1343–1346, doi:10.1103/PhysRevLett.30.1343.

[5] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?", *Phys. Rev. Lett.* **30** (Jun, 1973) 1346–1349, doi:10.1103/PhysRevLett.30.1346.

[6] G. Degrassi, S. Di Vita, J. Elias-Miro et al., "Higgs mass and vacuum stability in the Standard Model at NNLO", *JHEP* **1208** (2012) 098, doi:10.1007/JHEP08(2012)098, arXiv:1205.6497.

[7] J. M. No and M. Ramsey-Musolf, "Probing the Higgs Portal at the LHC Through Resonant di-Higgs Production", *Phys.Rev.* **D89** (2014), no. 9, 095031, doi:10.1103/PhysRevD.89.095031, arXiv:1310.6035.

[8] V. V. Khoze, C. McCabe, and G. Ro, "Higgs vacuum stability from the dark matter portal", *JHEP* **1408** (2014) 026, doi:10.1007/JHEP08(2014)026, arXiv:1403.4953.

[9] M. Bowen, Y. Cui, and J. D. Wells, "Narrow trans-TeV Higgs bosons and H -¿ hh decays: Two LHC search paths for a hidden sector Higgs boson", *JHEP* **0703** (2007) 036, doi:10.1088/1126-6708/2007/03/036, arXiv:hep-ph/0701035.

[10] C. Englert, Y. Soreq, and M. Spannowsky, "Off-Shell Higgs Coupling Measurements in BSM scenarios", arXiv:1410.5440.

[11] S. Chatrchyan et al., "Observation of a new boson at a mass of 125 GeV with the {CMS} experiment at the {LHC}", *Physics Letters B* **716** (2012), no. 1, 30 – 61, doi:http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[12] G. A. et al., "Observation of a new particle in the search for the Standard Model Higgs boson with the {ATLAS} detector at the {LHC}", *Physics Letters B* **716** (2012), no. 1, 1 – 29, doi:http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[13] I. Chatrchyan, S. et al.", "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$=7 and 8 TeV", *Journal of High Energy Physics* **2013** (2013), no. 6, doi:10.1007/JHEP06(2013)081.

[14] CMS Collaboration Collaboration, "Study of the Mass and Spin-Parity of the Higgs Boson Candidate via Its Decays to Z Boson Pairs", *Phys. Rev. Lett.* **110** (Feb, 2013) 081803, doi:10.1103/PhysRevLett.110.081803.

[15] CMS Collaboration Collaboration, "Measurement of the properties of a Higgs boson in the four-lepton final state", *Phys.Rev.* **D89** (2014) 092007, doi:10.1103/PhysRevD.89.092007, arXiv:1312.5353.

[16] CMS Collaboration Collaboration, "Observation of the diphoton decay of the Higgs boson and measurement of its properties", *Eur.Phys.J.* **C74** (2014), no. 10, 3076, doi:10.1140/epjc/s10052-014-3076-z, arXiv:1407.0558.

[17] CMS Collaboration Collaboration, "Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the standard model", Technical Report CMS-PAS-HIG-14-009, CERN, Geneva, (2014).

[18] T. A. collaboration, "Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data",.

[19] LHC Higgs Cross Section Working Group Collaboration, "LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle", arXiv:1209.0040.

[20] F. Caola and K. Melnikov, "Constraining the Higgs boson width with ZZ production at the LHC", *Phys. Rev. D* **88** (Sep, 2013) 054024, doi:10.1103/PhysRevD.88.054024.

[21] CMS Collaboration Collaboration, "Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs", *Phys.Lett.* **B736** (2014) 64, doi:10.1016/j.physletb.2014.06.077, arXiv:1405.3455.

[22] "Determination of the off-shell Higgs boson signal strength in the high-mass ZZ final state with the ATLAS detector", Technical Report ATLAS-CONF-2014-042, CERN, Geneva, (Jul, 2014).

[23] CMS Collaboration Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[24] S. Bolognesi, M. A. Borgia, R. Castello et al., "Calibration of track momentum using dimuon resonances in CMS", CMS Analysis Note 2010/059, (2010).

[25] CMS Collaboration Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[26] S. Dittmaier and M. Huber, "Radiative corrections to the neutral-current Drell-Yan process in the Standard Model and its minimal supersymmetric extension", *JHEP* **1001** (2010) 060, doi:10.1007/JHEP01(2010)060, arXiv:0911.2329.

[27] T. G. Collaboration, "GEANT - Detector Description and Simulation Tool".

[28] J. Alwall, P. Demin, S. de Visscher et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP* **0709** (2007) 028, doi:10.1088/1126-6708/2007/09/028, arXiv:0706.2334.

[29] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **0711** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[30] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **0605** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[31] CMS Collaboration Collaboration, "Particle-flow commissioning with muons and electrons from J/Psi and W events at 7 TeV",.

[32] Particle Data Group Collaboration, "Review of Particle Physics (RPP)", *Phys.Rev.* **D86** (2012) 010001, doi:10.1103/PhysRevD.86.010001.

[33] CMS Collaboration Collaboration, "Search for the standard model Higgs boson in the dimuon decay channel in pp collisions at sqrt(s)= 7 and 8 TeV", Technical Report CMS-PAS-HIG-13-007, CERN, Geneva, (2013).

[34] S. Baffioni, C. Charlot, F. Ferri et al., "Electron reconstruction in CMS", *Eur.Phys.J.* **C49** (2007) 1099–1116, doi:10.1140/epjc/s10052-006-0175-5.

[35] CMS Collaboration Collaboration, "Electron reconstruction and identification at sqrt(s) = 7 TeV", Technical Report CMS-PAS-EGM-10-004, CERN, Geneva, (2010).

[36] CMS Collaboration Collaboration, "Electron commissioning results at $\sqrt{s}$=7 TeV",.

[37] CMS Collaboration Collaboration, "Electron performance with 19.6 fb^{-1} of data collected at $\sqrt{s} = 8$ TeV with the CMS detector.",.

[38] CMS Collaboration Collaboration, "Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV", *J. Instrum.* **8** (Jun, 2013) P09009. 51 p. Comments: Submitted to JINST.

[39] CMS Collaboration Collaboration, "ECAL Performance Plots, 2012 Data",.

[40] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Physics Letters B* **659** (2008), no. 1–2, 119 – 126, doi:http://dx.doi.org/10.1016/j.physletb.2007.09.077.

[41] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets", *JHEP* **0804** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188.

[42] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur.Phys.J.* **C72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[43] CMS Collaboration Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV", *JHEP* **1101** (2011) 080, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.

[44] CMS Collaboration Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", Technical Report CMS-PAS-PFT-10-002, CERN, Geneva, (2010).

[45] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm", *JHEP* **0804** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[46] CMS Collaboration Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[47] CMS Collaboration Collaboration, "Pileup Jet Identification", Technical Report CMS-PAS-JME-13-005, CERN, Geneva, (2013).

[48] J. M. Campbell and R. Ellis, "{MCFM} for the Tevatron and the {LHC}", *Nuclear Physics B - Proceedings Supplements* **205–206** (2010), no. 0, 10 – 15, doi:http://dx.doi.org/10.1016/j.nuclphysbps.2010.08.011. Loops and Legs in Quantum Field Theory Proceedings of the 10th {DESY} Workshop on Elementary Particle Theory.

[49] J. M. Campbell and R. K. Ellis, "Update on vector boson pair production at hadron colliders", *Phys. Rev. D* **60** (Nov, 1999) 113006, doi:10.1103/PhysRevD.60.113006.

[50] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", *JHEP* **1107** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.

[51] Y. Gao, A. V. Gritsan, Z. Guo et al., "Spin determination of single-produced resonances at hadron colliders", *Phys. Rev. D* **81** (Apr, 2010) 075022, doi:10.1103/PhysRevD.81.075022.

[52] S. Bolognesi, Y. Gao, A. V. Gritsan et al., "Spin and parity of a single-produced resonance at the LHC", *Phys. Rev. D* **86** (Nov, 2012) 095031, doi:10.1103/PhysRevD.86.095031.

[53] I. Anderson, S. Bolognesi, F. Caola et al., "Constraining anomalous HVV interactions at proton and lepton colliders", *Phys.Rev.* **D89** (2014) 035007, doi:10.1103/PhysRevD.89.035007, arXiv:1309.4819.

[54] P. Avery, D. Bourilkov, M. Chen et al., "Precision studies of the Higgs boson decay channel $H \to ZZ \to 4\ell$ with MEKD", *Phys. Rev. D* **87** (Mar, 2013) 055006, doi:10.1103/PhysRevD.87.055006.

[55] N. D. Christensen and C. Duhr, "FeynRules – Feynman rules made easy", *Computer Physics Communications* **180** (2009), no. 9, 1614 – 1641, doi:http://dx.doi.org/10.1016/j.cpc.2009.02.018.

[56] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables", doi:10.5170/CERN-2011-002, arXiv:1101.0593.

[57] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, "Procedure for the LHC Higgs boson search combination in Summer 2011", Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, (Aug, 2011).

[58] T. Junk, "Confidence level computation for combining searches with small statistics", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **434** (1999), no. 2–3, 435 – 443, doi:http://dx.doi.org/10.1016/S0168-9002(99)00498-2.

[59] A. L. Read, "Presentation of search results: the CL s technique", *Journal of Physics G: Nuclear and Particle Physics* **28** (2002), no. 10, 2693.

[60] G. Cowan, K. Cranmer, E. Gross et al., "Asymptotic formulae for likelihood-based tests of new physics", *The European Physical Journal C* **71** (2011), no. 2, doi:10.1140/epjc/s10052-011-1554-0.

[61] Y. Gao, A. V. Gritsan, Z. Guo et al., "Spin determination of single-produced resonances at hadron colliders", *Phys.Rev.* **D81** (2010) 075022, doi:10.1103/PhysRevD.81.075022, arXiv:1001.3396.

[62] S. Bolognesi, Y. Gao, A. V. Gritsan et al., "On the spin and parity of a single-produced resonance at the LHC", *Phys.Rev.* **D86** (2012) 095031, doi:10.1103/PhysRevD.86.095031, arXiv:1208.4018.

[63] Y. Chen, N. Tran, and R. Vega-Morales, "Scrutinizing the Higgs Signal and Background in the $2e2\mu$ Golden Channel", *JHEP* **1301** (2013) 182, doi:10.1007/JHEP01(2013)182, arXiv:1211.1959.

[64] A. Soni and R. Xu, "Probing CP violation via Higgs decays to four leptons", *Phys.Rev.* **D48** (1993) 5259–5263, doi:10.1103/PhysRevD.48.5259, arXiv:hep-ph/9301225.

[65] S. Choi, D. Miller, M. Muhlleitner et al., "Identifying the Higgs spin and parity in decays to Z pairs", *Phys.Lett.* **B553** (2003) 61–71, doi:10.1016/S0370-2693(02)03191-X, arXiv:hep-ph/0210077.

[66] B. Allanach, K. Odagiri, M. Palmer et al., "Exploring small extra dimensions at the large hadron collider", *JHEP* **0212** (2002) 039, doi:10.1088/1126-6708/2002/12/039, arXiv:hep-ph/0211205.

[67] C. Buszello, I. Fleck, P. Marquard et al., "Prospective analysis of spin- and CP-sensitive variables in H -¿ Z Z -¿ l(1)+ l(1)- l(2)+ l(2)- at the LHC", *Eur.Phys.J.* **C32** (2004) 209–219, doi:10.1140/epjc/s2003-01392-0, arXiv:hep-ph/0212396.

[68] R. M. Godbole, D. Miller, and M. M. Muhlleitner, "Aspects of CP violation in the H ZZ coupling at the LHC", *JHEP* **0712** (2007) 031, doi:10.1088/1126-6708/2007/12/031, arXiv:0708.0458.

[69] W.-Y. Keung, I. Low, and J. Shu, "Landau-Yang Theorem and Decays of a Z' Boson into Two Z Bosons", *Phys.Rev.Lett.* **101** (2008) 091802, doi:10.1103/PhysRevLett.101.091802, arXiv:0806.2864.

[70] O. Antipin and A. Soni, "Towards establishing the spin of warped gravitons", *JHEP* **0810** (2008) 018, doi:10.1088/1126-6708/2008/10/018, arXiv:0806.3427.

[71] N. Kauer and G. Passarino, "Inadequacy of zero-width approximation for a light Higgs boson signal", *Journal of High Energy Physics* **2012** (2012), no. 8, doi:10.1007/JHEP08(2012)116.

[72] T. Binoth, N. Kauer, and P. Mertsch, "Gluon-induced QCD corrections to $pp \to ZZ \to \ell\bar{\ell}\ell'\bar{\ell}'$", doi:10.3360/dis.2008.142, arXiv:0807.0024.

[73] A. Ballestrero, A. Belhouari, G. Bevilacqua et al., "PHANTOM: A Monte Carlo event generator for six parton final states at high energy colliders", *Comput.Phys.Commun.* **180** (2009) 401–417, doi:10.1016/j.cpc.2008.10.005, arXiv:0801.3359.

[74] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", doi:10.5170/CERN-2013-004, arXiv:1307.1347.

[75] N. Kauer and G. Passarino, "Inadequacy of zero-width approximation for a light Higgs boson signal", *Journal of High Energy Physics* **2012** (2012), no. 8, doi:10.1007/JHEP08(2012)116.

[76] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling", *eConf* **C0303241** (2003) MOLT007, arXiv:physics/0306116.

[77] CMS Collaboration Collaboration, "Search for a standard-model-like Higgs boson with a mass in the range 145 to 1000 GeV at the LHC", *Eur.Phys.J.* **C73** (2013) 2469, doi:10.1140/epjc/s10052-013-2469-8, arXiv:1304.0213.

[78] K. Hamilton, P. Nason, and G. Zanderighi, "MINLO: Multi-Scale Improved NLO", *JHEP* **1210** (2012) 155, doi:10.1007/JHEP10(2012)155, arXiv:1206.3572.

[79] S. Goria, G. Passarino, and D. Rosco, "The Higgs Boson Lineshape", *Nucl.Phys.* **B864** (2012) 530–579, doi:10.1016/j.nuclphysb.2012.07.006, arXiv:1112.5517.

[80] CMS Collaboration Collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", arXiv:1412.8662.

[81] S. e. a. Chatrchyan, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states", *Journal of High Energy Physics* **2014** (2014), no. 1, doi:10.1007/JHEP01(2014)096.

[82] CMS Collaboration Collaboration, "Search for the standard model Higgs boson in the H to ZZ to 2 $\ell 2\nu$ channel in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **1203** (2012) 040, doi:10.1007/JHEP03(2012)040, arXiv:1202.3478.