Università degli Studi di Torino

Scuola di Dottorato in Scienza ed Alta Tecnologia
Indirizzo di Fisica ed Astrofisica

Search for the Standard Model Higgs boson in the decay channel $\mathrm{H} \to \mathrm{ZZ}^{(*)} \to 4\ell$ with the CMS Experiment

Coordinatore
Prof. Guido Boffetta

Tutore
Dr. Nicola Amapane

Relatore
Dr. Chiara Mariotti

Candidato
Cristina Botta

Ciclo XXIV

Date of dissertation

20^{th} December 2011

Candidate

Cristina Botta

Tutor

Dr. Nicola Amapane
Università degli Studi di Torino

Degree board

Prof. Gian Piero Passarino
Università degli Studi di Torino

Prof. W. Murray
Particle Physics-Rutherford Appleton Laboratory (Oxford) and CERN

Prof. Albert De Roeck
University of Antwerp (Belgium) and CERN

Ph.D. coordinator

Prof. Guido Boffetta
Università degli Studi di Torino

Acknowledgements

Grazie a chi mi ha insegnato tutto questo.
A Chiara, Nicola, Yves, Riccardo, Giovanni, Ivan.

Grazie per il lavoro di squadra.
A Nicola, Roberto, Christophe, Claude, Ivica, Stephanie, Mario, Nhan.

Grazie per il fondamentale sostegno.
Ad Alessandra e Chiara, al gruppo CMS Torino.

Grazie alle mie famiglie.
Quella di Savigliano, di Parigi, di St Genis.

Contents

Acknowledgements **i**

Introduction **1**

1 Standard Model Higgs Boson **5**
1.1 The Standard Model of Elementary Particles 5
1.2 The Electroweak Theory 7
1.3 The Higgs Mechanism 9
1.3.1 Vector Boson Masses and Couplings 12
1.3.2 Fermion Masses and Couplings 12
1.4 Higgs boson mass 13
1.4.1 Theoretical bounds on m_{H} 14
1.4.2 Experimental bounds on m_{H} from LEP and Tevatron 15
1.5 Motivation for the LHC 17

2 The CMS Detector at the LHC **19**
2.1 The Large Hadron Collider 19
2.1.1 The design concept of LHC 20
2.1.2 Definition of Kinematic Variables 22
2.1.3 Phenomenology of Proton-Proton Collisions 22
2.2 The CMS Experiment 24
2.2.1 The Magnet 27
2.2.2 The Tracker 28
2.2.3 The Electromagnetic Calorimeter 31
2.2.4 The Hadronic Calorimeter 33
2.2.5 The Muon System 35
2.3 The CMS Trigger System 39
2.3.1 The Level-1 Trigger 41
2.3.2 The High Level Trigger 47

3 Search for the SM Higgs in the $ZZ^{(*)} \to 4\ell$ decay channel with CMS **51**
3.1 Higgs Phenomenology at the LHC 51
3.1.1 Higgs Production 52

3.1.2 Higgs Decay 54
3.2 Higgs Search at CMS 55
3.3 The 4ℓ final state 59
3.4 Monte Carlo Datasets and Cross Sections 61
3.5 Experimental Datasets and Integrated Luminosity 67

4 Physics Object: Muons **69**
4.1 Muon Reconstruction 70
4.1.1 Local Segments Reconstruction 70
4.1.2 Stand-Alone Muon Reconstruction 72
4.1.3 Global Muon Reconstruction 73
4.1.4 Tracker Muon Reconstruction 75
4.2 Muon quality selection 77
4.3 Muon Isolation 79
4.4 Muon High Level Trigger 81
4.4.1 Muon Trigger Path 81
4.4.2 Muon reconstruction in the HLT 82
4.4.3 Performance of the Muon HLT 86
4.5 Muon Measurements with data 89
4.5.1 Simulation and data sample 91
4.5.2 Reconstruction and Identification efficiency 94
4.5.3 Muon Isolation efficiency 101
4.5.4 Compatibility with a common vertex 105
4.5.5 Trigger efficiency 107
4.5.6 Trigger rejection rates 112
4.5.7 Momentum Scale 112

5 Physics Object: Electrons **117**
5.1 Electron Reconstruction 117
5.2 Electron Identification 118
5.3 Electron Isolation 119
5.4 Electron High Level Trigger 120
5.5 Electron Measurements with data 121

6 Event Selection and Kinematics **123**
6.1 Minimum Lepton Requirements 123
6.2 Preselection and choice of the 4ℓ candidate 124
6.3 Background Reduction Cuts 127
6.4 Trigger Requirements 133
6.5 Selection Performance 135

7 Systematic Uncertainties and Background Estimation from Data 143
7.1 Signal systematic uncertainties 143
7.1.1 Instrumental uncertainties 144
7.2 Background Evaluation and Control 146
7.2.1 Evalutation of the $ZZ^{(*)}$ continuum 147
7.2.2 Inclusive instrumental and reducible backgrounds estimation 152
7.2.3 Evalutation of the $Zb\bar{b}/c\bar{c}\to 4\ell$ and $t\bar{t} \to 4\ell$ backgrounds 159

8 Results 167
8.1 Mass Distributions and Kinematics 167
8.2 Measurement of the $ZZ \to 4\ell$ Cross Section 174
8.3 Mass Measurement Uncertainties 175
8.4 Interpretation of the Results 177

Summary 185

A Event Properties 187

B Event Displays. Low Mass Candidates 191

Bibliography 201

Introduction

The Standard Model (SM) of particle physics predicts the existence of a unique physical Higgs scalar boson associated to the spontaneous electroweak symmetry breaking, whose mass is a free parameter of the theory. The Higgs boson is the only still unconfirmed element of the SM and it is regarded as the responsible for the masses of all known elementary particles. Giving an answer to the fundamental question about its existence is a matter of the highest priority in the field of particle physics.
The Large Hadron Collider (LHC) delivered the first collision at a center of mass energy of 7 TeV, the highest ever reached by a particle collider, on 30 March 2010. On the past 30 October, the 2011 proton-proton collision operations were stopped after LHC delivered an integrated luminosity of 5.7 fb^{-1}, meeting the goal of the first data-taking period, whose aim was to provide the LHC experiments with enough data to constrain the Higgs boson mass over the allowed range.
The inclusive production of SM Higgs bosons followed by the decay $\text{H}\rightarrow\text{ZZ}^{(*)} \rightarrow \ell^{\pm}\ell^{\mp}\ell'^{\pm}\ell'^{\mp}$with $\ell, \ell' = e$ or μ, in short $\text{H}\rightarrow 4\ell$, is expected to be one of the most important channels for the Higgs boson discovery or exclusion over a wide range of possible m_{H} values at the Compact Muon Soleneoid (CMS) experiment. Its clear experimental signature with four isolated leptons arising from the same interaction vertex is highly favored in the search of these very rare events given the challenging hadronic environment of high-pileup pp collisions.

A search for a SM Higgs boson in the four-lepton decay channel, $\text{H}\rightarrow 4\ell$, is presented in this thesis. The analysis is designed for a hypothetical Higgs boson with a mass in the range $110 < m_{\text{H}} < 600$ GeV/c^2.
The search for a Higgs signal has to cope with a reducible background contribution from the continuum $ZZ^{(*)}$ production via $q\bar{q}$ and gluon-induced processes. Reducible backgrounds derive from the $Zb\bar{b}$ and $t\bar{t}$ processes, with Ws undergoing leptonic decays, and secondary leptons produced within b-jets, as well as from instrumental backgrounds such as Z+jets or WZ+jet(s) where jets are misidentified as leptons. The contributions from the reducible and instrumental backgrounds are eliminated relying almost solely on the measurements of leptons.

Therefore a robust detection of muons over the full acceptance of the CMS detector and within a large momentum range, as well as excellent muon trigger capabilities are fundamental ingredients of this analysis. In particular, preserving the highest possible efficiency for muon trigger and reconstruction even at low momentum while ensuring sufficient discrimination against hadronic jets faking muons and against muons from decay-in-flight is mandatory for the search of the very rare $H\rightarrow 4\mu$ events.

It is therefore to the muon trigger and reconstruction performances that I devoted the first two years of my Ph.D.
When I joined the CMS Collaboration, the LHC was scheduled to start physics operation in one year, and the experiment was in a very active preparatory phase. I have been firstly involved in the muon trigger, with the goal to optimize the algorithms for efficiency, characterize their performace on simulations, and develop monitoring tools to be used on data. When the assembling phase of the detector was completed, and the first proton beams were circulated, the LHC high-energy operation was delayed by one year due to a major technical incident. The collaboration decided to profit of this delay to engage in a full detector commissioning phase with cosmic rays. This exercise has provided detailed insight into the performance of the CMS muon reconstruction and trigger algorithms and the experience gained was invaluable in the preparation for data from LHC collisions. The muon trigger, reconstruction and identification algorithms were adapted to detect muons coming from the top of the detector instead than from the beamline. Cosmic muons arriving in the CMS cavern have a very wide transverse momentum spectrum, opening the possibility to study in detail the performances of the hardware and software in different ranges of muon energy. Specifically, I was involved in the development of methods to measure for the first time the reconstruction and trigger efficiency with data, with an adapted tag-and-probe technique using the half-track of a reconstructed cosmic muon in the top half of the detector as a tag.
As soon as LHC provided enough collisions, I started to use the first $J/\psi\rightarrow\mu\mu$ events and, soon after, the first $Z\rightarrow\mu\mu$ to measure the efficiency of muon trigger, reconstruction, identification, and isolation criteria with more sophisticated tag-and-probe algorithms, as a function of the relevant variables. In particular I used J/ψ events to study the trigger turn-on-curves and Z decays to explore the muon momentum ranges above 10 GeV/c. I devoted particular attention to the performance of the hardware and software muon trigger, investigating and addressing early issues which were spotted by these measurements.
I am now performing these studies in the contest of the $H\rightarrow 4\ell$ analysis to provide precise measurements of the muons efficiencies (cf. Chapter 4). The comparison between the muon efficiencies in data and in simulation

is of particular interest for the H$\rightarrow 4\ell$ analysis, which requires an estimate of the number of expected signal events. Discrepancies between data and simulation, as well as systematic uncertainties on measured efficiencies have to be taken into account and propagated to the final event selection.
Another study that I performed on early collision data is the measurement of the muon trigger rate as a function of the transverse momentum thresholds. This measurement shows a very good agreement with the expectations, meaning that the simulation reproduces the actual trigger efficiency for muons from different sources. This means that the simulation can be used to predict muon trigger rates. That is essential to decide the trigger thresholds in order to face the increase of rate with istantaneous luminosity. Several such updates to the trigger menus happened since the start of the 2010 data taking, as the istantaneous luminosity increased by more than one order of magnitude since then.
In the context of the H$\rightarrow 4\ell$ analysis, which relies on double lepton triggers, I have been working on the definition of the proper lepton trigger strategy to ensure the highest possible efficiency (97-99% depending on the Higgs boson mass) while tightening the lepton identification requirements and increasing transverse momentum thresholds in order to keep the lepton trigger rates as low as requested by the data acquisition system.

Besides the lepton measurements, which are the basis of the H$\rightarrow 4\ell$ analysis, I have been one of the core developers of the H$\rightarrow 4\ell$ search at CMS.
In particular I have been working on the definition of the event selection strategy (cf. Chapter 6) in order to keep the highest possible signal efficiency while rejecting backgrounds; for this purpose I searched for the best topological and kinematical variables to distinguish the muons originated by decays of the Higgs boson from background muons in $Zb\bar{b}$ and $t\bar{t}$ events.
In order to have a precise estimation of the rate of background events that survive the event selection, we can extract it from data with data-driven methods, instead of simply using simulations; I have been developing a strategy to define phase space regions dominated by $Zb\bar{b} \rightarrow 4\ell$ and $t\bar{t} \rightarrow 4\ell$ events, and to extrapolate the observed rate back to the signal phase space region. Finally I worked on the estimation of the systematic uncertainties induced by the muon measurements that need to be considered in the final interpretation of the results (cf. Chapter 7).
The analysis and the results obtained with $\mathcal{L}$= 1.13$\pm$0.07 fb^{-1} have been presented for the first time at the EPS HEP 2011 conference [1] [2]. Afterwords an update with $\mathcal{L}$= 1.66$\pm$0.07 fb^{-1} has been presented at the Lepton-Photon 2011 conference [3] [4].
In this thesis the analysis strategy and the results with the full 2010-2011 integrated luminosiy $\mathcal{L}$= 4.71$\pm$0.21 fb^{-1} will be presented. Supporting documentation can be find in [5].

Chapter 1

Standard Model Higgs Boson

The fundamental components of matter and their interactions are nowadays accurately described by the Standard Model of Particle Physics (SM) [6, 7, 8, 9], which is based upon two separate Quantum Field Theories, describing the electroweak interaction (Glashow-Weinberg-Salam model or GWS) and the strong interaction (Quantum Chromo-Dynamics or QCD). In this chapter, a short overview of the SM (Section 1.1) and of the Electroweak Theory (Section 1.2) is given, focusing on the ElectroWeak Symmetry Breaking (EWSB), the Higgs mechanism and the Higgs boson properties (Section 1.3). In Section 1.4 theoretical constraints on the Higgs boson mass, a free parameter of the model, as well as those from direct searches at previous collider experiments, are presented. Finally, in light of what previously discussed, in Section 1.5 the LHC Physics Program is presented.

1.1 The Standard Model of Elementary Particles

The SM describes the matter as composed by twelve elementary particles, the *fermions*, all having half-integer spin. Fermions can be divided into two main groups, *leptons* and *quarks*, whose classification is given in Table 1.1. Quarks are subject to both strong and electroweak interactions and do not exist as free states, but only as constituents of a wide class of particles, the *hadrons*, such as protons and neutrons. Leptons, instead, only interact through electromagnetic and weak forces.
In the SM, the interactions between particles are described in terms of the exchange of *bosons*, integer-spin particles which are carriers of the fundamental interactions. The main characteristics of bosons and of the corresponding interactions are summarised in Table 1.2. The *gravitational* interaction is not taken into account, as it is not relevant at the typical mass and distance scales of particle physics.

Table 1.1: *Classification of the three families of fundamental fermions.*

Fermions	1^{st} fam.	2^{nd} fam.	3^{rd} fam.	Charge	Interactions
Quarks	u	c	t	$+2/3$	All
	d	s	b	$-1/3$	
Leptons	e	μ	τ	-1	Weak, E.M.
	ν_e	ν_μ	ν_τ	0	Weak

Table 1.2: *Classification of the three families of fundamental fermions.*

	Electromagnetic	Weak	Strong
Quantum	Photon (γ)	$W^{\pm}$, Z	Gluons
Mass [GeV/c^2]	0	80, 90	0
Coupling constant	$\alpha(Q^2 = 0) \approx \frac{1}{137}$	$\frac{G_F}{(\hbar c)^3} \approx 1.2 \cdot 10^{-5}$ GeV^{-2}	$\alpha_s(m_Z) \approx 0.1$
Range [cm]	∞	10^{-16}	10^{-13}

All this complex fenomenology naturally comes out from the elegant mathematical formalism of a QFT; in particular the SM of particle physics is a perturbatively renormalizable QFT. The renormalization which is a necessary condition for a perturbative theory is guaranteed by the *local gauge simmetry* of the SM lagrangian.
According to Noether's theorem a conservation law must derive from the local invariance of the Lagrangian. The SM is built on the search of the proper local simmetry to describe the nature of the fundamental interactions between particles. From this common procedure the unification of the description of the electromagnetic, weak and strong forces in a unique mathematical formalism follows. Historically this path started from the Quantum Electrodynamics (QED) with the simmetry group $U(1)_{EM}$. Then followed the unification of the electromagnetic and weak forces in the Electroweak Theory described by the Glashow Weinberg and Salam model (GWS) and based on the local simmetry $SU(2)_I \otimes U(1)_Y$. Finally the Quantum Chromodynamics, to describe the strong interaction with the conservation of the *color* quantum number, is built in analogy to the GWS model as a gauge theory for the $SU(3)_C$ group. The SM is therefore a QFT with local gauge $SU(2)_I \otimes U(1)_Y \otimes SU(3)_C$.
Despite this simmetry predicts with precision and accuracy the fenomenology of particle interactions, it is broken by the mass terms of the Lagrangian; a necessary ingrediend of the SM is the therefore a Spontaneous Simmetry Breaking mechanism that allows to introduce the mass terms in a local gauge invariant Lagrangian. In the next sections, the Electroweak

Theory and the EWSB, with its simplest realization, the Higgs mechanism, will be described in some details.

1.2 The Electroweak Theory

From a historical point of view, the starting point for the study of electroweak interactions is Fermi's theory of muon decay [10], which is based on an effective four-fermion Lagrangian:

$$\mathcal{L} = -\frac{4G_F}{\sqrt{2}}\bar{\nu}_\mu\gamma^\alpha\frac{1-\gamma_5}{2}\mu\bar{e}\gamma_\alpha\frac{1-\gamma_5}{2}\nu_e\,, \tag{1.1}$$

where G_F is the *Fermi coupling constant* reported in Table 1.2. Equation 1.1 represents a "point-like" interaction, with only one vertex and without any intermediate boson exchanged. It is usually referred to as $V-A$ interaction, being formed by a *vectorial* and an *axial* component. The term $\frac{1}{2}(1-\gamma_5)$ that appears in it is the negative helicity projector. Only the negative helicity (*left-handed*) component of fermions takes part to this interaction.

Fermi's Lagrangian is not renormalisable and it results in a non-unitary scattering matrix. Both problems of renormalisability and unitarity are overcome, as already said, describing the weak interaction by a *gauge* theory, i.e. requiring its Lagrangian to be invariant under local transformations generated by the elements of some Lie group (*gauge transformations*). The specific group of local invariance (*gauge group*) is to be determined by the phenomenological properties of the interaction and of the particles involved. In particular, the resulting Lagrangian must reduce to Equation 1.1 in the low energy limit. A detailed derivation of this Lagrangian is not provided in this work, but the results are summarised in the following.
A gauge theory for weak interactions is conceived as an extension of the theory of electromagnetic interaction, the Quantum Electro-Dynamics or QED, which is based on the gauge group $U(1)_{EM}$, associated to the conserved quantum number Q (*electric charge*). In this case, the condition of local invariance under the $U(1)_{EM}$ group leads to the existence of a massless vector boson, the *photon*.
A theory reproducing both the electromagnetic and weak interaction phenomenology is achieved by extending the gauge symmetry to the group $SU(2)_I \otimes U(1)_Y$. In this sense, the weak and electromagnetic interactions are said to be unified. The generators of $SU(2)_I$ are the three components of the *weak isospin* operator, $t^a = \frac{1}{2}\tau^a$, where τ^a are the Pauli matrices. The generator of $U(1)_Y$ is the *weak hypercharge* Y operator. The corresponding

quantum numbers satisfy the following relation

$$Q = I_3 + \frac{Y}{2},$$

where I_3 is the third component of the weak isospin (eigenvalue of t^3). Fermions can be divided in doublets of negative-helicity (*left-handed*) particles and singlets of positive-helicity (*right-handed*) particles, as follows:

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

where $\ell = e, \mu, \tau$, $u = u, c, t$ and $d = d, s, b$. Neutrinos have no *right* component, as their mass is taken as null. In Table **??**, I_3, Y and Q quantum numbers of all fermions are reported.

Table 1.3: *Isospin (I_3), hypercharge (Y) and electric charge (Q) of all fermions.*

	I_3	Y	Q
$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}$	$\begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}$	$\begin{pmatrix} 2/3 \\ -1/3 \end{pmatrix}$
u_R, d_R	0, 0	$4/3$, $-2/3$	$2/3$, $-1/3$
$\begin{pmatrix} \nu_{\ell,L} \\ \ell_L \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$
ℓ_R	0	-2	-1

As well as for QED, the requirement of local gauge invariance with respect to the $SU(2)_I \otimes U(1)_Y$ group introduces now four massless vector fields (*gauge fields*), $W_\mu^{1,2,3}$ and B_μ, which couple to fermions with two different coupling constants, g and g'. Note that B_μ does not represent the photon field, because it arises from the $U(1)_Y$ group of hypercharge, instead of $U(1)_{EM}$ group of electric charge. The gauge-invariant Lagrangian for fermion fields can be written as follows:

$$\mathcal{L} = \overline{\Psi}_L \gamma^\mu \left(i\partial_\mu + g t_a \mathrm{W}_\mu^a - \tfrac{1}{2} g' Y \mathrm{B}_\mu \right) \Psi_L + \overline{\psi}_R \gamma^\mu \left(i\partial_\mu - \tfrac{1}{2} g' Y \mathrm{B}_\mu \right) \psi_R \tag{1.3}$$

where

$$\Psi_L = \begin{pmatrix} \psi_L^1 \\ \psi_L^2 \end{pmatrix}$$

and where Ψ_L and ψ_R are summed over all the possibilities in Equation 1.2. As already stated, $W_\mu^{1,2,3}$ and B_μ do not represent physical fields, which are given instead by linear combinations of the four mentioned fields: the charged bosons W^+ and W^- correspond to[1]

$$\mathrm{W}_\mu^\pm = \sqrt{\frac{1}{2}} \left(\mathrm{W}_\mu^1 \mp i \mathrm{W}_\mu^2 \right), \tag{1.4}$$

[1] In the following, a different notation will be also used: $W_\mu^{(-)} = W_\mu$, $W_\mu^{(+)} = W_\mu^\dagger$.

while the neutral bosons γ and Z correspond to

$$A_\mu = B_\mu \cos\theta_W + W^3_\mu \sin\theta_W \quad (1.5)$$

$$Z_\mu = -B_\mu \sin\theta_W + W^3_\mu \cos\theta_W, \quad (1.6)$$

obtained by mixing the neutral fields W^3 and B- with a rotation defined by the *Weinberg angle* θ_W. In terms of the fields in Equations 1.4 to 1.6, the interaction term between gauge fields and fermions, taken from the Lagrangian in Equation 1.3, becomes

$$\mathcal{L}_{int} = \frac{1}{2\sqrt{2}} g (J^+_\alpha W^{(+)\alpha} + J^-_\alpha W^{(-)\alpha}) + \frac{1}{2}\sqrt{g'^2 + g^2} J^Z_\alpha Z^\alpha - e J^{EM}_\alpha A^\alpha, \quad (1.7)$$

where J^{EM} is the electromagnetic current coupling to the photon field, while J^+, J^- and J^Z are the three weak isospin currents. It is found that

$$J^Z_\alpha = J^3_\alpha - 2\sin^2\theta_W \cdot J^{EM}_\alpha.$$

A- can then be identified with the photon field and, requiring the coupling terms to be equal, one obtains

$$g \sin\theta_W = g' \cos\theta_W = e \quad (1.8)$$

which represents the electroweak unification. The GWS model thus predicts the existence of two charged gauge fields, which only couple to left-handed fermions, and two neutral gauge fields, which interact with both left- and right-handed components.

1.3 The Higgs Mechanism

In order to correctly reproduce the phenomenology of weak interactions, both fermion and gauge boson fields must acquire mass, in agreement with experimental results. Up to this point, however, all particles are considered massless: in the electroweak Lagrangian, in fact, a mass term for the gauge bosons would violate gauge invariance, which is needed to ensure the renormalisability of the theory. Explicit mass terms for fermions, instead, would not violate gauge invariance, but in the GWS model the Lagrangian is also required to preserve the invariance under *chirality* transformations, and this is achieved only with massless fermions. Masses are thus introduced with the *Higgs mechanism* [11], which allows fermions and $W^\pm$, Z bosons to be massive, while keeping the photon massless. Such mechanism is accomplished by means of a doublet of complex scalar fields,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^1 + i\phi^2 \\ \phi^3 + i\phi^4 \end{pmatrix}, \quad (1.9)$$

which is introduced in the electroweak Lagrangian within the term

$$\mathcal{L}_{EWSB} = (D^{\mu}\phi)^{\dagger}(D_{\mu}\phi) + V(\phi^{\dagger}\phi), \tag{1.10}$$

where $D_{\mu} = \partial_{\mu} - igt_a W^a_{\mu} + \frac{i}{2}g' Y B_{\mu}$ is the covariant derivative. The Lagrangian in Equation 1.10 is invariant under $SU(2)_I \otimes U(1)_Y$ transformations, since the kinetic part is written in terms of covariant derivatives and the potential V only depends on the product $\phi^{\dagger}\phi$. The ϕ field is characterised by the following quantum numbers:

	I_3	Y	Q
$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$	$\begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Writing the potential term as follows (see also Figure 1.1 for a graphical representation)

$$V(\phi^{\dagger}\phi) = -\mu^2\phi^{\dagger}\phi - \lambda(\phi^{\dagger}\phi)^2, \tag{1.11}$$

with $\mu^2 < 0$ and $\lambda > 0$, it results to have a minimum for

$$\phi^{\dagger}\phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}. \tag{1.12}$$

This minimum is not found for a single value of ϕ, but for a manifold of non-zero values. The choice of (ϕ^+, ϕ^0) corresponding to the ground state, i.e. the lowest energy state or *vacuum*, is arbitrary, and the chosen point is not invariant under rotations in the (ϕ^+, ϕ^0) plane: this is referred to as *spontaneous symmetry breaking*. If one chooses to fix the ground state on the ϕ^0 axis, the vacuum expectation value of the ϕ field is

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v^2 = -\frac{\mu^2}{\lambda}. \tag{1.13}$$

The ϕ field can thus be rewritten in a generic gauge, in terms of its vacuum expectation value:

$$\phi = \frac{1}{\sqrt{2}} e^{\frac{i}{v}\phi^a t_a}\begin{pmatrix} 0 \\ \mathrm{H} + v \end{pmatrix}, \quad a = 1,2,3\,,$$

where the three fields ϕ^a and the fourth $\phi^4 = \mathrm{H} + v$ are called *Goldstone fields*. Being scalar and massless, they introduce four new degrees of freedom, in addition to the six degrees due to the transverse polarisations of the massless vector bosons $\mathrm{W}^{\pm}$ and Z. The unitary gauge is fixed by the transformation

$$\phi' = e^{-\frac{i}{v}\phi^a t_a}\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \mathrm{H} + v \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \phi^4 \end{pmatrix}.$$

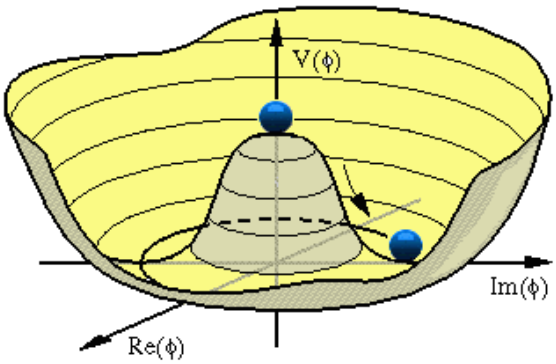


Figure 1.1: *Shape of the Higgs potential of Equation 1.11.*

The remaining field, the *Higgs field*, has now a zero expectation value. Rewriting the Lagrangian in Equation 1.10 with the ϕ field in the unitary gauge, $\mathcal{L}_{EWSB}$ results from the sum of three terms:

$$\mathcal{L}_{EWSB} = \mathcal{L}_{\mathrm{H}} + \mathcal{L}_{\mathrm{HW}} + \mathcal{L}_{\mathrm{HZ}}, \tag{1.14}$$

where the three terms can be written as follows, using the approximation $V \sim \mu^2 \mathrm{H}^2 + \text{const}$ and neglecting higher order terms:

$$\begin{aligned} \mathcal{L}_{\mathrm{H}} &= \frac{1}{2}\partial_\alpha \mathrm{H} \partial^\alpha \mathrm{H} + \mu^2 \mathrm{H}^2 \\ \mathcal{L}_{\mathrm{HW}} &= \frac{1}{4} v^2 g^2 \mathrm{W}_\alpha \mathrm{W}^{\dagger\alpha} + \frac{1}{2} v g^2 \mathrm{H} \mathrm{W}_\alpha \mathrm{W}^{\dagger\alpha} \qquad (1.15) \\ &= m_{\mathrm{W}}^2 \mathrm{W}_\alpha \mathrm{W}^{\dagger\alpha} + g_{\mathrm{HW}} \mathrm{H} \mathrm{W}_\alpha \mathrm{W}^{\dagger\alpha} \\ \mathcal{L}_{\mathrm{HZ}} &= \frac{1}{8} v^2 (g^2 + g'^2) \mathrm{Z}_\alpha \mathrm{Z}^\alpha + \frac{1}{4} v (g^2 + g'^2) \mathrm{H} \mathrm{Z}_\alpha \mathrm{Z}^\alpha \qquad (1.16) \\ &= \frac{1}{2} m_{\mathrm{Z}}^2 \mathrm{Z}_\alpha \mathrm{Z}^\alpha + \frac{1}{2} g_{\mathrm{HZ}} \mathrm{H} \mathrm{Z}_\alpha \mathrm{Z}^\alpha \end{aligned}$$

Equations 1.15 and 1.16 now contain mass terms for fields $\mathrm{W}^\pm$ and Z: each of the three gauge bosons has acquired mass and an additional degree of freedom, corresponding to the longitudinal polarisation. At the same time, three of the four Goldstone bosons have disappeared from the Lagrangian $\mathcal{L}_{EWSB}$, thus preserving the total number of degrees of freedom: the degrees related to the missing Goldstone bosons have become the longitudinal degrees of the vector bosons. Only the H scalar field is still present and has acquired mass itself: it is the Higgs field.

Summarising, the Higgs mechanism is used to introduce the weak boson masses, without explicitly breaking the gauge invariance, thus preserving the renormalisability of the theory. When a symmetry is "spontaneously" broken, in fact, it is not properly eliminated: it is rather "hidden" by the

choice of the ground state. It can be shown that the minimum of the Higgs field is still invariant under the $U(1)_{EM}$ group. Hence, the electromagnetic symmetry is unbroken and the photon does not couple to the Higgs boson and remains massless.

1.3.1 Vector Boson Masses and Couplings

Equations 1.15 and 1.16 show that the masses of vector bosons $W^{\pm}$ and Z are related to the parameter v, characteristic of the EWSB, and to the electroweak coupling constants:

$$\begin{cases} m_W = \frac{1}{2} v g \\ m_Z = \frac{1}{2} v \sqrt{g^2 + g'^2} \end{cases} \quad \rightarrow \quad \frac{m_W}{m_Z} = \frac{g}{\sqrt{g^2 + g'^2}} = \cos\theta_W . \tag{1.17}$$

Also the couplings of vector bosons to the Higgs can be obtained from Equations 1.15 and 1.16, and are found to depend on the square of m_W and m_Z:

$$g_{HW} = \frac{1}{2} v g^2 = \frac{2}{v} m_W^2 \tag{1.18}$$

$$g_{HZ} = \frac{1}{2} v (g^2 + g'^2) = \frac{2}{v} m_Z^2 . \tag{1.19}$$

A relation between the decay ratios of the Higgs boson to a W pair and to a Z pair can be derived from Equations 1.18 and 1.19:

$$\frac{BR(H \rightarrow W^+W^-)}{BR(H \rightarrow ZZ)} = \left(\frac{g_{HW}}{\frac{1}{2} g_{HZ}} \right)^2 = 4 \left(\frac{m_W^2}{m_Z^2} \right)^2 \simeq 2.4 .$$

Finally, the EWSB energy scale can be determined from the relation between the v parameter and the Fermi constant G_F:

$$v = \left(\frac{1}{\sqrt{2} G_F} \right)^{\frac{1}{2}} \simeq 246 \text{ GeV}. \tag{1.20}$$

1.3.2 Fermion Masses and Couplings

The Higgs mechanism is also used to generate the fermion masses, by introducing in the SM Lagrangian an $SU(2)_I \otimes U(1)_Y$ invariant term, called *Yukawa term*, which represents the interaction between the Higgs and the fermion fields. Since ϕ is an isodoublet, while the fermions are divided in left-handed doublet and right-handed singlet, the Yukawa terms (one for each fermion generation) must have the following expression for leptons:

$$\mathcal{L}_\ell = -G_{H\ell} \cdot \bar{l}_\ell \phi \ell_R + \bar{\ell}_R \phi^\dagger l_\ell .$$

In the unitary gauge, the first component of ϕ is zero, therefore a mass term will arise from the Yukawa Lagrangian only for the second component of l_ℓ: this correctly reproduces the fact that neutrino is (approximately) massless.

$$\mathcal{L}_\ell = -\frac{G_{\mathrm{H}\ell}}{\sqrt{2}} v \bar{\ell}\ell - \frac{G_{\mathrm{H}\ell}}{\sqrt{2}} \mathrm{H}\bar{\ell}\ell \,. \tag{1.21}$$

As far as the quark fields are concerned, the *down* quarks (d, s, b) are treated in the same way as leptons; *up* quarks (u, c, t), instead, must couple to the charge-conjugate of ϕ

$$\phi^c = -i\tau_2 \phi^* = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^3 - i\phi^4 \\ -\phi^1 + i\phi^4 \end{pmatrix}$$

which becomes in the unitary gauge

$$\phi^c = \frac{1}{\sqrt{2}} \begin{pmatrix} \eta + v \\ 0 \end{pmatrix}$$

Therefore, the Yukawa Lagrangian is

$$\mathcal{L}_Y = -G_{\mathrm{H}\ell} \overline{L}_L \phi \ell_R - G_{\mathrm{H}d} \overline{Q}_L \phi d_R - G_{\mathrm{H}u} \overline{Q}_L \phi^c u_R + h.c.\,. \tag{1.22}$$

From Equation 1.21, the mass of a fermion (apart from neutrinos) and its coupling constant to the Higgs boson are found to be

$$m_f = \frac{G_{\mathrm{H}f}}{\sqrt{2}} v \tag{1.23}$$

$$g_{\mathrm{H}f} = \frac{G_{\mathrm{H}f}}{\sqrt{2}} = \frac{m_f}{v} \,. \tag{1.24}$$

Being $G_{\mathrm{H}f}$ free parameters, the mass of the fermions cannot be predicted by the theory.

1.4 Higgs boson mass

The Higgs boson mass is the only yet unknown free parameter of the SM. The Higgs in fact has never been observed experimentally and its mass cannot be predicted by the SM. It depends on the parameters v and λ, but while the former can be estimated by its relation with the constant G_F of Fermi's theory, the latter is characteristic of the field ϕ and cannot be determined other than measuring the Higgs mass itself. However theoretical indications exists and experimental constraints, from direct and indirect searches at other colliders, narrow the possible range.

1.4.1 Theoretical bounds on m_{H}

An upper bound on the Higgs mass comes from unitarity of the scattering matrix. Considering the elastic scattering of longitudinally polarized Z bosons:

$$\mathrm{Z}_L\mathrm{Z}_L \rightarrow \mathrm{Z}_L\mathrm{Z}_L \tag{1.25}$$

the unitarity bound on the corresponding amplitude, in the limit $s \ll m_Z^2$, implies

$$m_H < \sqrt{\frac{16\pi}{3}} v \sim 1 \text{ TeV}/c^2 \tag{1.26}$$

Slightly more restrictive bounds ~ 800 GeV$/c^2$ can be found considering other scattering processes, such as $\mathrm{Z}_L\mathrm{W}_L \rightarrow \mathrm{Z}_L\mathrm{W}_L$. Of course this argument is valid in the limit of the conventional scenario in which the Higgs sector is within the perturbative regime. For masses ~ 1.4 TeV$/c^2$ the decay width of the SM Higgs into a pair of gauge bosons becomes approximately equal to m_H and the Higgs can no longer be considered as a particle. Therefore for large m_H the perturbative approach is not valid anymore and not-perturbative tecniques are required.

Moreover, as described in [12] and [13], the SM must address two problematic situations to be valid up to the reduced Planck scale $M_P \sim 2 \times 10^{18}$ GeV by which some new physics associated with quantum gravity must surely appear. Constraints to the Higgs boson mass can be found by imposing the energy scale Λ up to which the SM is valid, before the perturbation theory breaks down and non-SM phenomena emerge. The upper limit is obtained requiring that the running quartic coupling of Higgs potential λ remains finite up to the scale Λ (*triviality* bound). This limit is shown with the upper pair of lines in Figure 1.2 [12]. A lower limit is found instead by requiring that λ remains positive after the inclusion of radiative corrections, at least up to Λ: this implies that the Higgs potential is bounded from below, i.e. the minimum of such potential is an absolute minimum (*vacuum stability* bound). Looser constraints are found by requiring such minimum to be local, instead of absolute (*metastability* bounds). These lower limits are shown, with their uncertainties, with shaded bands in Figure 1.2.

If the validity of the SM is assumed up to the Plank scale ($\Lambda \sim 10^{19}$ GeV), the allowed Higgs boson mass range is between 130 and 190 GeV, while for $\Lambda \sim 1$ TeV the Higgs mass can be up to 700 GeV$/c^2$. Part of this region is already excluded by by direct experimental search at Tevatron as will be shown in the next section.

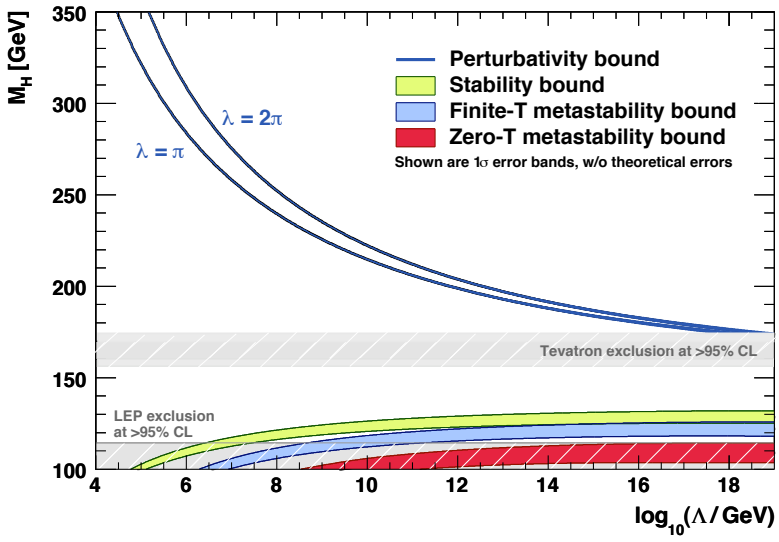


Figure 1.2: *The scale Λ at which the renormalization-group equations drive the SM Higgs coupling non-perturbative, and the scale Λ at which these equations create an instability in the electroweak vacuum ($\lambda < 0$) [12]. The width of the bands indicates the errors induced by the uncertainties in m_t and α_s (added quadratically). The* triviality *bound is given for $\lambda = \pi$ and $\lambda = 2\pi$. Their difference indicates the size of the theoretical uncertainty in this bound. The absolute vacuum stability bound is displayed by the shaded green band, while the less restrictive metastability scenarios are shown with the shaded blue and red bands. The theoretical uncertainties in these bounds have been ignored in the plot. The grey areas indicate the LEP and Tevatron exclusion domains that will be discussed in Section 1.4.2.*

1.4.2 Experimental bounds on m_{H} from LEP and Tevatron

Bounds on the Higgs mass are provided by measurements at LEP [14][15], SLC [16] and Tevatron [17].
Direct searches at LEP-II allowed to set a lower limit of 114.4 GeV/c^2 (95% C.L.) on the Higgs boson mass [18], while recent results from the Tevatron experiments exclude the mass range of 156 to 177 GeV/c^2 (95% C.L.) [19]. The Tevatron searches are currently least sensitive for masses around $m_{\mathrm{H}} \simeq 130$ GeV/c^2 and the mass range $m_{\mathrm{H}} \gg 2 \times M_{\mathrm{Z}}$ remains there unexplored.
Moreover, constraints on the Higgs boson mass can be extracted indirectly from the measurement of other electroweak observables, which have a logarithmic dependence on m_{H} through the radiative corrections [20]. All the precision electroweak measurements performed by the four LEP experiments and by SLD, CDF and D∅ [21][22] have been combined together

and fitted, assuming the SM as the correct theory and using the Higgs mass as free parameter. The result of this procedure is summarised in Figure 1.3, where the $\Delta\chi^2$ of the fit, defined as $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$, is plotted as a function of m_{H}. The black, solid curve is the result of the fit, while the

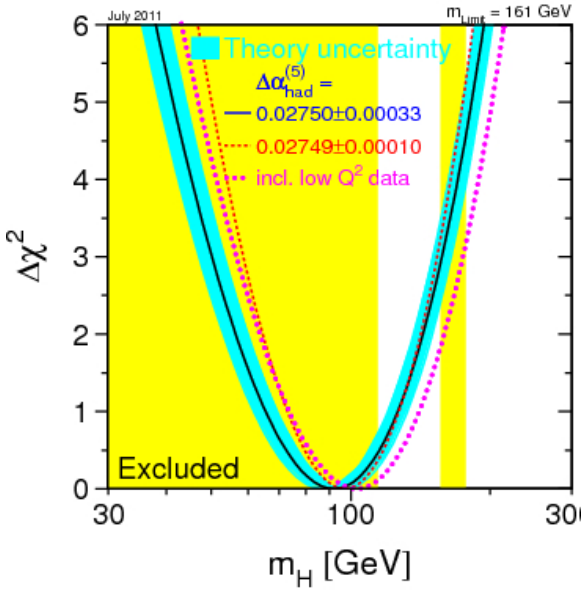


Figure 1.3: $\Delta\chi^2$ *of the fit to the electroweak precision measurements performed at LEP, SLC and Tevatron as a function of the Higgs boson mass (July 2011) [21]. The black, solid line represents the result of the fit, and the blue, shaded band is the theoretical error from unknown higher-order corrections. The yellow area represents the regions excluded by LEP-II and Tevatron.*

blue band represents the theoretical uncertainty due to unknown higher order corrections. The yellow area shows the regions excluded by LEP-II and Tevatron measurements. The Tevatron exclusion, in particular, is derived by the results shown in Figure 1.4. The indirectly measured value of the Higgs boson mass, corresponding to the minimum of the curve in Figure 1.3, is $m_{\text{H}} = 92^{+34}_{-26}$ GeV/c^2, where the errors represent the experimental uncertainty at 68% C.L. derived from the black line, thus not taking the theoretical uncertainty into account. An upper limit of 161 GeV/c^2 can also be set at 95% C.L., including the theoretical uncertainty. This limit increases to 185 GeV/c^2 when including the direct search limit of 114.4 GeV/c^2. These results are model-dependent, as the loop corrections take into account only contributions from known physics, and are thus well-grounded only within the SM theory.

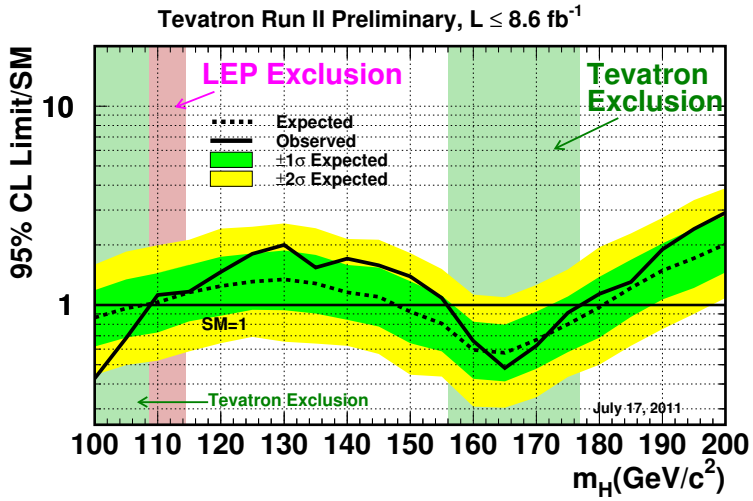


Figure 1.4: *95% C.L. upper limits on the Higgs boson production cross section as a function of the Higgs boson mass (July 2011) [19]. The solid (dashed) line represents the observed (expected) ratio between the excluded cross section and the SM one. The green and yellow bands indicate, respectively, the 68% and 95% probability regions for a fluctuation of the expected limit. The regions in which the line is below 1 represent the mass ranges where the SM Higgs boson can be exluded at 95% C.L. These results are obtained combining CDF and DØ data, for a Tevatron integrated luminosity of 8.2 fb^{-1}.*

1.5 Motivation for the LHC

From what discussed up to now, the discovery of the mechanism that gives origin to the mass of all known particles requires a machine able to span the energy range from about 100 GeV to 1-2 TeV. The main goal of the LHC project is certainly the search for the SM Higgs boson and the study of its properties, or the investigation of some alternative EWSB mechanisms of which the Higgs is the simplest realization in the SM. Other physics motivations for the LHC are summarised in the following.

Physics beyond the SM. There are strong motivations to think that the SM is not the ultimate theory of particle interaction, but only a well tested low energy approximation of some more fundamental theory. One reason is the presence in the SM of "too many" free parameters for a fundamental theory: there are at least 19 of them, 7 more if non-vanishing neutrino masses are assumed, and most of these are introduced by the Higgs mechanism.

A stronger reason is that, unless some fine-tuned cancellation takes place, the Higgs boson mass suffers from a divergent radiative correction proportional to a high-energy cut-off. New physics such as the *Supersymmetry* (SUSY), can introduce physical reasons for a cancellation of the Higgs mass divergent correction at a cost of introduc-

ing even more free parameters than the SM.
Moreover, *Grand Unification Theories* (GUTs) predict that the running coupling constants of the three fundamental interactions, extrapolated to a very high energy scale (10^{16} GeV), unify to a single value: this does not occur, unless some new phenomena arise at an intermediate scale. SUSY, e.g., predicts the existence of a number of new particles at the TeV scale, which would thus be accessible at the LHC: particles with masses up to 3-5 TeV/c^2 can be observed, depending on the total integrated luminosity.

Precision Measurements. The LHC is a factory of heavy particles, such as W and Z bosons and t and b quarks: the huge amount of events that can be collected, thanks to the high luminosity and centre-of-mass energy, allows all sorts of precision measurements: W mass, WWγ and WWZ triple gauge couplings, mass and decay properties of t quarks, measurement of the strong coupling constant α_S, quantities related to CP violation, and many others.

Heavy ion physics. When running as a heavy ion ($^{208}Pb^{82+}$) collider, the LHC allows to study the phase transition from hadronic matter to a plasma of deconfined quarks and gluons, the *quark-gluon plasma* or QGP.

Besides all existing models and motivations, the LHC is open to any new and unexpected phenomenon or particle that may show up at this unexplored energy scale.
Two of the main LHC experiments, CMS (Compact Muon Solendoid) and ATLAS (A ToroidaL ApparatuS), are designed to address these fundamental questions. In the next chapter an introduction to the world largest particle accelerator and an extensive description of the CMS detector are provided.

Chapter 2

The CMS Detector at the LHC

2.1 The Large Hadron Collider

The LHC [23] is the world's major particle accelerator; it is an unprecedent machine in terms of energy, luminosity, size and complexity of experiments, cost, and involvement of human resources.
On 23 November 2009, the accelerator produced the first proton-proton collisions. After few pilot runs at energies of 450 GeV and 1.18 TeV per beam, the energy was ramped up to 3.5 TeV and, on 30 March 2010, the first collisions at a centre-of-mass energy of 7 TeV, the highest ever reached at a particle collider, were recorded by the four experiments: ALICE, ATLAS, CMS and LHCb. With this energy, about 47 pb^{-1} of integrated luminosity were delivered during 2010 (see Figure 2.1), with a maximum instantaneous luminosity of $2 \cdot 10^{32}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, obtained with start-up conditions: 368 bunches per beam with collisions every 150 ns.
The 2011 operations have been successful. From March to October the machine operations have been interrupted only for short periods in order to prepare and test further increases of the luminosity. On 26 October 2011 the highest instantaneous luminosity has been reached with a peak value of $3.5 \cdot 10^{33}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ with 1331 colliding bunches per beam and collisions every 50 ns.
The total integrated luminosity delivered by LHC in 2011 is 5.73 fb^{-1}; 5.22 fb^{-1} have been recorded by CMS (see Figure 2.2).
In the next years, the LHC will progressively increase its energy and instantaneous luminosity, reaching eventually the design values of 14 TeV, seven times the highest energy reached so far at Tevatron, and $10^{34}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, about two orders of magnitude more than the luminosity of any previous machines.
In the context of the Heavy Ion LHC program, the first lead ion beams were circulated in the LHC in November 2010. In one month, between November and December, about 8 $\mu\mathrm{b}^{-1}$ of Pb-Pb collisions were delivered,

at a centre-of-mass energy of 574 TeV (7 TeV per proton pair, 2.76 TeV per nucleon pair), the highest ever touched in heavy ion experiments. The 2011 Heavy Ion Run is currently in preparation.
The design lifespan of LHC is of 10 years.

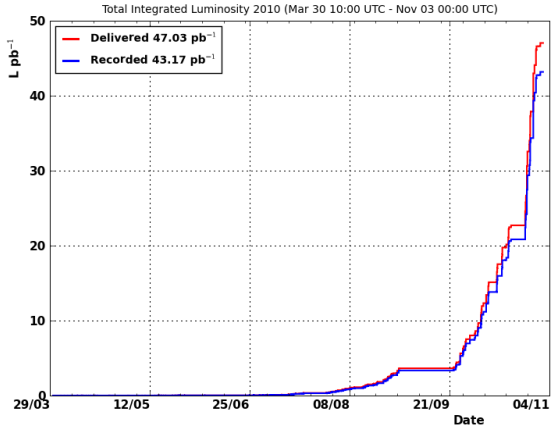


Figure 2.1: *Integrated luminosity delivered by the LHC (in red) and recorded by CMS (in blue) in 2010 proton-proton collisions.*

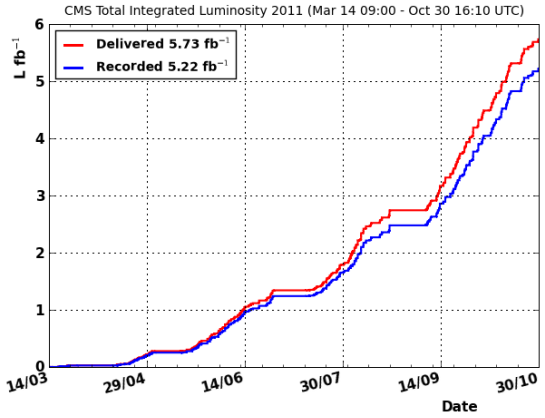


Figure 2.2: *Integrated luminosity delivered by the LHC (in red) and recorded by CMS (in blue) in 2011 proton-proton collisions.*

2.1.1 The design concept of LHC

The LHC is placed in the already existent 26.7 km long LEP tunnel, situated at a depth of about 100 m underground on the French-Swiss border. The main design characteristics of the machine are listed in Table 2.1.

Table 2.1: *LHC design parameters for p-p and Pb-Pb collisions.*

Parameter	p-p	Pb-Pb
Circumference [km]	26.659	
Beam radius at interaction point [μm]	15	
Dipole peak field [T]	8.3	
Design centre-of-mass energy [TeV]	14	1148
Design Luminosity [$\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$]	10^{34}	$2 \cdot 10^{27}$
Luminosity lifetime [h]	10	4.2
Number of particles per bunch	$1.1 \cdot 10^{11}$	$\sim 8 \cdot 10^{7}$
Number of bunches	2808	608
Bunch length [mm]	53	75
Time between collisions [ns]	24.95	$124.75 \cdot 10^{3}$
Bunch crossing rate [MHz]	40.08	0.008

Since collisions occur between particles of the same charge, two separate acceleration cavities with two different magnetic field configurations are required. The bending power needed to keep the beam circulating is the limiting factor to the achievable centre of mass energy. In fact, from the equation

$$p\,[\mathrm{TeV}/c] = 0.3 \cdot B\,[\mathrm{T}] \cdot \rho\,[\mathrm{km}]\,,$$

where p is the beam momentum, B the magnetic field and ρ the radius of the ring, one can deduce that the magnetic field required for $p = 7$ TeV/c is about 5.4 T. In practice, since it is not possible to fill the whole machine with magnets, the needed power is obtained by using about 1200 superconducting dipoles operating at 1.9 K, each providing a magnetic field of about 8.3 T. Boosts are given by 400 MHz superconducting radiofrequency cavities with a voltage ranging between 8 and 16 MV. The channels for the two beams acceleration are inserted in a single cryostat. The event rate R of a process with cross section σ is given by

$$R \;=\; \mathcal{L} \cdot \sigma\,,$$

where $\mathcal{L}$ is the instantaneous luminosity of the machine, defined as the number of collisions per unit time and cross-sectional area of the beams [24]:

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi\sigma_x\sigma_y}\,;$$

σ_x and σ_y characterise the Gaussian transverse beam profiles in the horizontal and vertical directions.

2.1.2 Definition of Kinematic Variables

Before proceeding, it is useful to introduce some general definitions and variables commonly used to describe the phenomenology at hadron colliders. The reference frame used in this work is the standard CMS reference frame (Figure 2.4) which is a Cartesian system centered in the interaction point and with the z axis tangent to the beam line. The x axis is chosen to be horizontal and pointing towards the centre of the ring, and the y axis is vertical and pointing upwards. The direction of the z axis, i.e. the direction of the beam, is referred to as *longitudinal*. The x-y plane, orthogonal to the beam line, is called *transverse* plane. Based on these definitions, the momentum of a particle can be divided in two components: the *longitudinal momentum* p_z and the *transverse momentum* p_{T}, defined as

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}\,. \tag{2.1}$$

The *rapidity* of a particle of energy E is defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}\,. \tag{2.2}$$

Rapidity has the property of being *additive* under Lorentz boosts along the z direction, i.e. it is simply shifted by a constant when subjected to such transformations. For high-energy particles, rapidity can be approximated by *pseudorapidity*

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right)\,, \tag{2.3}$$

which only depends on the polar angle θ of the particle momentum, i.e. its angle with respect to the z axis.

2.1.3 Phenomenology of Proton-Proton Collisions

At the nominal centre of mass energy of 14 TeV the total inelastic proton-proton cross section σ_{pp} is about 55 mb. These events include two classes of interaction:

- soft collisions: large-distance collisions between two incoming protons, in which only a small momentum is transferred; particle scattering at large angle is thus suppressed, and the final state particles have small transverse momentum, $\langle p_{\mathrm{T}} \rangle \simeq 500$ MeV/c, so that most of them escape down the beam pipe;
- hard collisions: since protons are not elementary particles, occasionally collisions with high transferred p_{T} occur between two of their constituents (*partons*, i.e. quarks and gluons). These interactions represent the interesting physics events, where massive particles may

be created. The rate of hard interactions, though, is several orders of magnitude lower than that of soft interactions.

In hard interactions, the *effective* centre-of-mass energy $\sqrt{\hat{s}}$, i.e. the centre-of-mass energy of the two interacting partons, is proportional to the fractions x_a and x_b of proton energy carried by the two involved partons:

$$\sqrt{\hat{s}} = \sqrt{x_a\, x_b\, s}\,,$$

where $\sqrt{s}$ is the centre-of-mass energy of the proton beams. The distributions of the fractional momentum of partons inside the protons are called *parton distribution functions* (PDFs). They are different for each type of parton and are functions of the exchanged momentum, Q^2. At high Q^2, the contribution of gluons and sea quarks increases with respect to that of valence quarks. PDFs are measured in Deep Inelastic Scattering (DIS) experiments and different models are available. In Figure 2.3 the CTEQ6M PDFs [25] are shown for two different values of Q^2.

In reconstructed the event kinematic no constraints can be applied on the

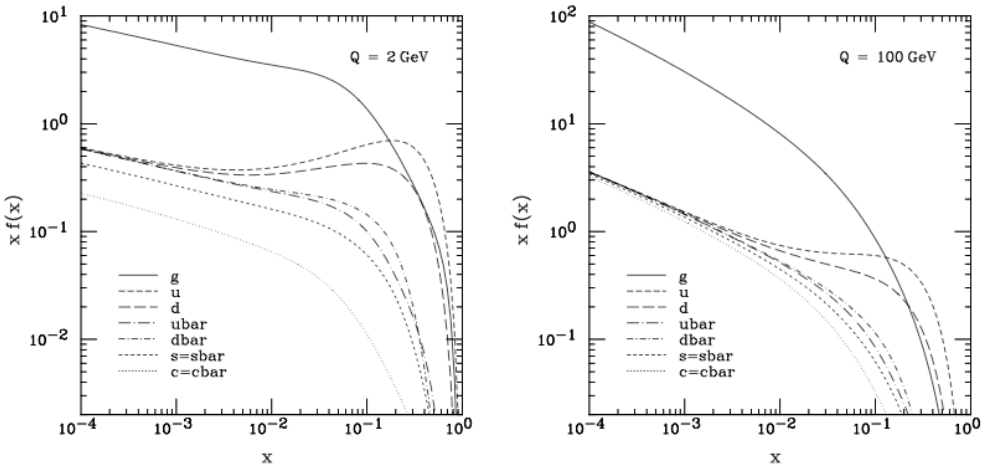


Figure 2.3: *CTEQ6M PDFs for Q=20 GeV/c (left) and Q=100 GeV/c (right) [25].*

total momentum of final state particles since the two interacting partons have variable and unknown momenta. Assuming that the transverse momenta of partons are negligible, though, the total transverse momentum in the final state must be zero. The longitudinal momentum, instead, remains unconstrained.

Another important consequence is that the centre of mass of the interaction may be boosted along the beam direction. For this reason, it is necessary to use quantities which have invariance properties under boosts along this direction, such as the transverse momentum (Equation 2.1) and the rapidity (Equation 2.2). Given the high energy of the particles under study,

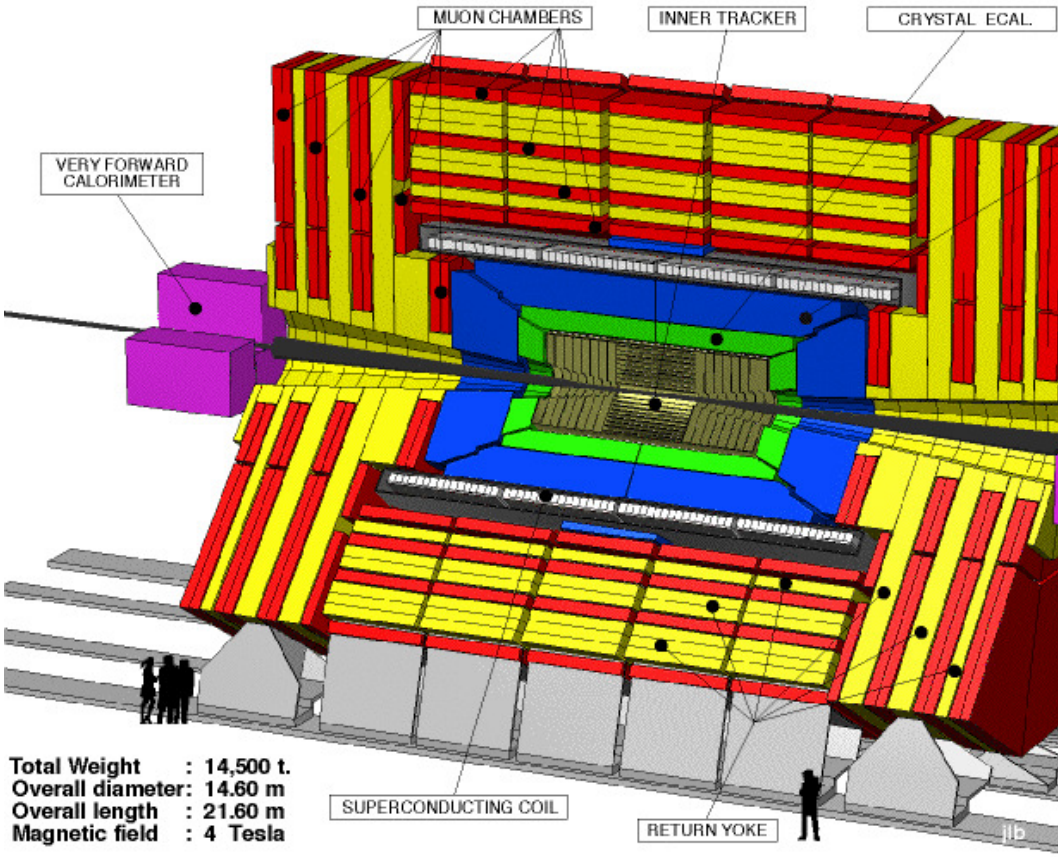
rapidity is usually replaced by pseudorapidity (Equation 2.3). Particles produced in soft collisions are mostly distributed at high rapidity. However, the soft interaction rate is so large that the residual tail at high-p_T is comparable with the hard interaction rate, and constitutes a background to high p_T signal events.

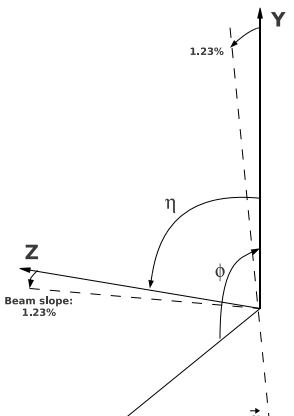
2.2 The CMS Experiment

The Compact Muon Solenoid (CMS) [26][27] is one of the two general-purpose detectors which operates at LHC. Its physics goals range from the search for the Higgs boson to the searches for new physics beyond the SM and to precision measurements of already known physics processes. To achieve these goals, excellent lepton reconstruction and particle identification are required. The main features of the CMS detector are the 3.8 T superconducting solenoid, which allows a compact design with a strong magnetic field, a robust and redundant muon system, a good electromagnetic calorimeter, and a high-quality tracking system. The overall structure of CMS consists of several cylindrical layers coaxial to the beam axis (the *barrel layers*), closed at both ends by detector disks orthogonal to the beam direction (the *endcaps*), to ensure optimal hermeticity. The overall length is 21.6 m, the diameter 14.6 m and the total weight about 14 500 t. Schematic views of the CMS detector are shown in Figures 2.5 and 2.6.
The CMS coordinate system (see Figure 2.4) used to describe the detector geometry is a right-handed Cartesian frame, with the z axis coincident with the beam direction, the z axis pointing to the centre of the LHC ring and the y axis directed upwards. Because of the cylindrical symmetry of the CMS design, the reconstruction algorithms use a cylindrical coordinate system (r,ϕ, η), r being the distance from the z axis, ϕ the azimuthal coordinate with respect to the x axis, and η the pseudorapidity defined by Equation 2.3.

As already mentioned, the core of the apparatus is the solenoid, which contains, from inside out, the following detectors:

- the *tracker*, made of a silicon *pixel detector* in the inner region, closest to the beam, and of silicon *microstrip detectors* in the outer region, used to reconstruct charged particle tracks and primary and secondary interaction vertices;
- the *electromagnetic calorimeter* (ECAL), which allows for precise measurement of electron and photon energies; it is made of lead tungstate ($PbWO_4$) scintillating crystals, both in the barrel and in the endcaps, and is complemented by a forward *preshower detector*;


MUON CHAMBERS
INNER TRACKER
CRYSTAL ECAL.
VERY FORWARD CALORIMETER
Total Weight : 14,500 t.
Overall diameter: 14.60 m
Overall length : 21.60 m
Magnetic field : 4 Tesla
SUPERCONDUCTING COIL
RETURN YOKE


Y
1.23%
η
Z
ϕ
Beam slope:
1.23%

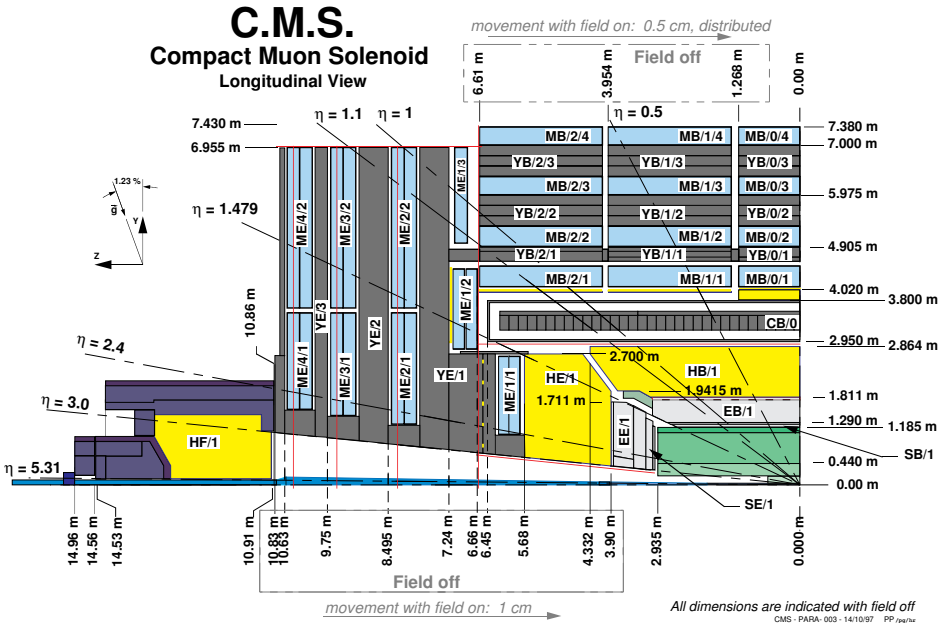


Figure 2.5: *Longitudinal view of one quarter of the CMS detector.*

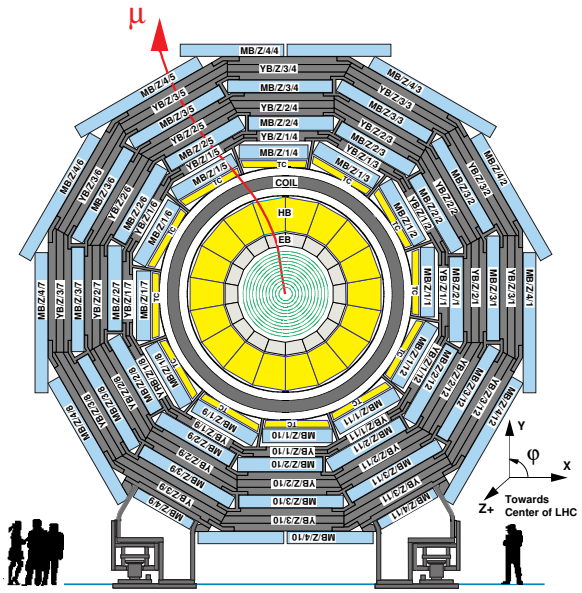


Figure 2.6: *Transverse view of the CMS barrel region.*

- the *hadron calorimeter* (HCAL), used for jet direction and transverse energy measurements, extended in the forward region with the *"very forward calorimeter"*.

Outside the magnet coil, the iron return yoke of the magnet hosts the *muon spectrometer*, used for reconstruction of muon tracks: *drift tubes* (DT) in the barrel and *cathode strip chambers* (CSC) in the endcaps, both complemented by *resistive plate chambers* (RPC), to ensure redundancy and robustness to the muon trigger.

2.2.1 The Magnet

The CMS magnet [28] is a 13 m long superconducting solenoid, the largest ever built. It is able to generate a uniform magnetic field of 4 T in the inner region, storing about 2.5 GJ of energy (Figure 2.7).
It operates at a temperature of 4 K, ensured by a sophisticated helium cooling system. At such temperature, the flat Ni Tb cable becomes superconducting, allowing a 20 kA current to flow without appreciable loss. The whole magnet is then contained in an enormous vacuum cylinder, which isolates it from the external environment.
Outside, an iron structure composed by five barrel layers and three disks for each endcap constitutes the iron yoke, needed to bridge the return magnetic flux, which otherwise would get lost, disturbing the surrounding equipment.
The CMS magnet provides a large bending power, allowing a precise measurement of the transverse momentum of charged particles. A further and independent p_T measurement outside the solenoid is possible thanks to the bending power in the iron yoke.

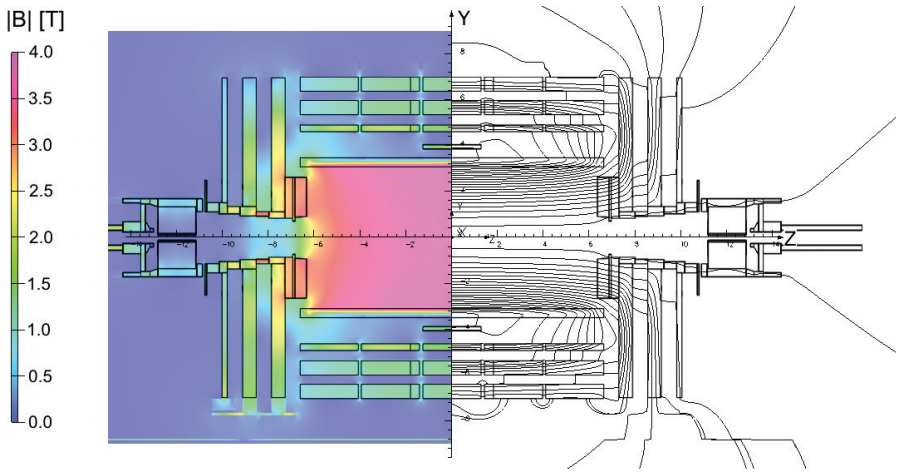


Figure 2.7: *Layout of the magnetic field of CMS [29].*

2.2.2 The Tracker

The tracker [30] (see Figure 2.8), which is the innermost subdetector and the closest to the interaction point, is dedicated to track and vertex reconstruction. It extends in the region $|\eta| < 2.5$, $r < 120$ cm, $|z| < 270$ cm, and it is completely based on silicon detectors, covering a surface of 210 m^2, the largest ever designed for detectors of this kind. The CMS tracker has to satisfy some important physics requirements:

- an efficient reconstruction of isolated lepton tracks: in all the pseudorapidity coverage the efficiency is close to 100% up to very low p_T ($\sim$ 500 MeV/c);
- a good lepton momentum resolution: for $|\eta| < 2$, $\sigma(p_T)/p_T < 4\%$ for single muons in a large range of p_T (up to $\sim$ 100 GeV/c);
- reconstruction of interaction vertex and identification of secondary vertices (Figure 2.9): this task is essential for tagging and reconstruction of b-jets.

In order to fulfill all these tasks and perform a good pattern recognition, two main properties have driven the tracker design:

- low cell occupancy: this requires high granularity detectors, especially those closer to the interaction point, and fast primary charge collection, obtained by using thin detectors and overdepleting the silicon bulks;
- large hit redundancy: ten layers of silicon detectors provide many measured hits (12-14) per track.

These properties allow for a high tracking efficiency and a low rate of fake tracks.
Moreover, several material budget constraints are imposed by the necessity to minimise electron *bremsstrahlung* and hadronic interactions, not to degrade tracking and ECAL performances. The region with the highest amount of material is the transition region between barrel and endcap ($1 < |\eta| < 2$), due to the high density of cables. Finally, both pixel and microstrip detectors have to be kept at a working temperature of -10 °C for the whole tracker volume, in order to limit the radiation damage to silicon sensors, due to the high flux of hadrons and backscattered neutrons[1].

Pixel Detector

The pixel detector (Figure 2.10) consists of three barrel layers and two endcap disks for each side. The barrel layers, 53 cm long, are placed

[1] Backscattering of neutrons from nuclear interactions in the material of ECAL.

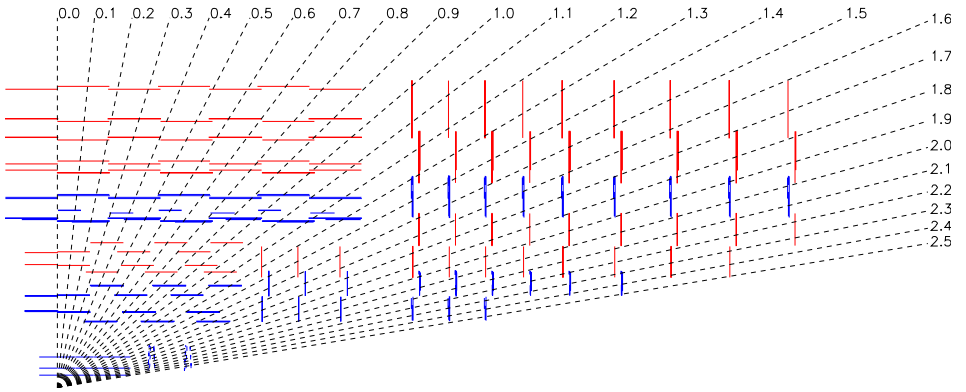


Figure 2.8: *Longitudinal view of one quarter of the silicon tracker, including the pixel detector.*

at r = 4.4, 7.3 and 10.2 cm. The first layer will be replaced by an outer layer at r = 13 cm during the high luminosity phase, to reduce the radiation damage. The two disks of each endcap consist of 24 blades, arranged in a turbine-like shape, having the inner radius of 6 cm and the outer of 15 cm. The total area covered with pixels is about 0.92 m^2. The inner detector provides at least two hits for tracks originating within $2\sigma_z$ from the nominal interaction vertex, in the pseudorapidity region $|\eta| < 2.2$. Due to the high density of tracks, 100×150 μm^2 pixels are used to ensure low cell occupancy. A spatial resolution of about 10 μm in the r-ϕ plane and 15 μm in the z coordinate can be achieved in the barrel, about 15 μm and 20 μm respectively in the endcaps.

Silicon Microstrip Detector

The silicon microstrip detector is divided in two main regions. The inner region is made of 4 barrel layers (*tracker inner barrel* or TIB) and 3 disks at each side (*tracker inner disks* or TIDs). The outer system, instead, consists of 6 barrel layers (*tracker outer barrel* or TOB) and 9 disks for each endcap (*tracker endcaps* or TECs). It covers a radial region between 20 and 120 cm and the pseudorapidity region $|\eta| < 2.5$.
All four regions (TIB, TID, TOB, TEC) are provided with both single-sided and double-sided microstrip modules. The strips are oriented along the z direction in the barrel and along the r coordinate in the endcaps. The microstrip detector is designed to provide a spatial resolution of about 40-60 μm in the r-ϕ plane and about 500 μm along z. The occupancy is lower than 1%.

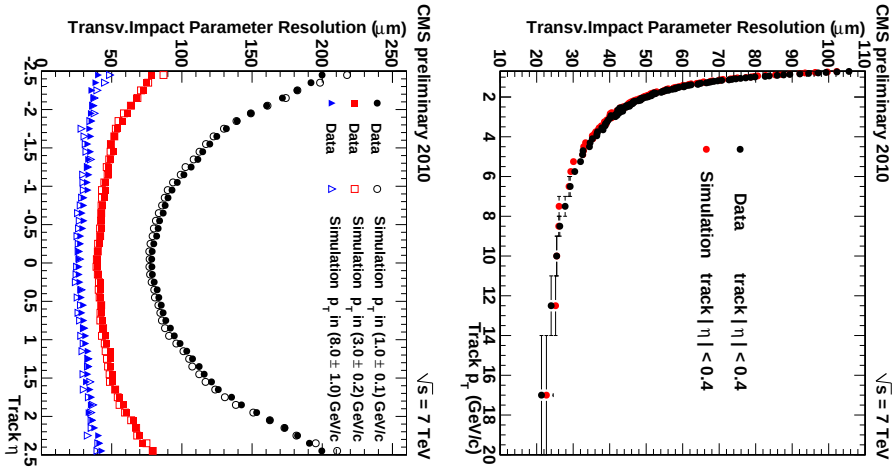


Figure 2.9: *Transverse impact parameter resolution vs. $|\eta|$, in data and simulation, for different muon momenta (right). Transverse impact parameter resolution vs. p_T, in data and simulation, for $|\eta| < 0.4$ (left), [31].*

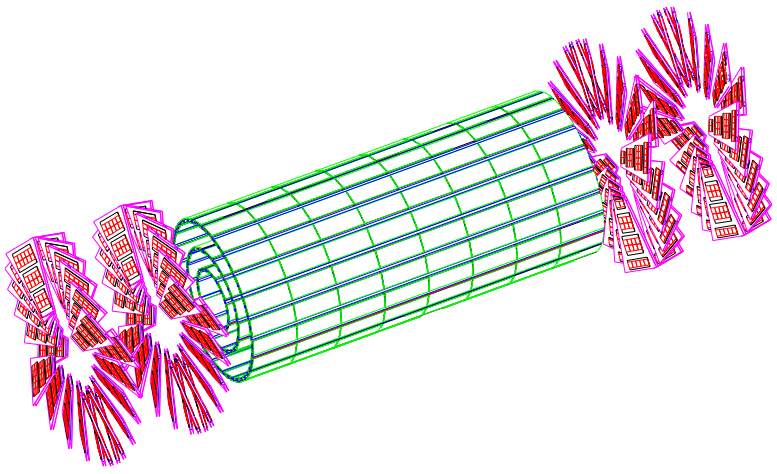


Figure 2.10: *The pixel detector. The barrel section and the two disks of the endcaps are visible.*

2.2.3 The Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) [32] is made of 74 848 lead tungstate ($PbWO_4$) crystals, chosen because of their excellent energy resolution.
The lead tungstate is characterised by a high density (8.28 g/cm^3) and a short radiation length (X_0 = 0.89 cm), so the calorimeter is very compact and can be placed inside the magnetic coil. Furthermore, it has a small Molière radius (2.2 cm), which gives the ECAL a very fine granularity, needed because of the high particle density produced at the LHC. Moreover, these crystals are characterised by a very short scintillation decay time, which allows to collect about 80% of the light within 25 ns, so that they can be used at the crossing rate of 40 MHz.
The ECAL barrel covers the central rapidity region ($|\eta| < 1.48$) and the two ECAL endcaps extend the coverage up to $|\eta| = 3$ (Figure 2.11).

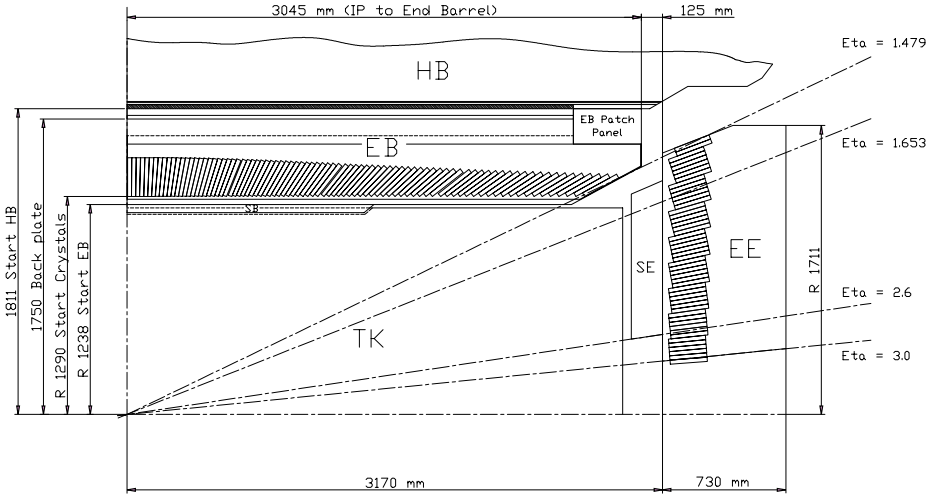


Figure 2.11: *Longitudinal view of one quarter of the CMS ECAL.*

The crystals have a trapezoidal shape and are arranged in a η-ϕ grid in the barrel and a x-y grid in the endcaps. The barrel crystals have a front face area of 2.2 $\times$ 2.2 cm^2 (thus matching the Molière radius), 23 cm length and are positioned at r = 1.29 m. Hence the total depth of ECAL barrel is 25.8 X_0, and the transverse granularity is $\Delta\eta \times \Delta\phi$ = 0.0175 $\times$ 0.0175 rad. In the endcaps, crystals have 2.47 $\times$ 2.47 cm^2 front face, 22 cm length (corresponding to 24.7 X_0) and are positioned at z = 3.17 m. Both in the barrel and in the endcaps, the crystals are tilted of about 3° in η and in ϕ, thus giving the structure a geometry slightly off-pointing from the interaction region, in order to improve the hermeticity of the detector.

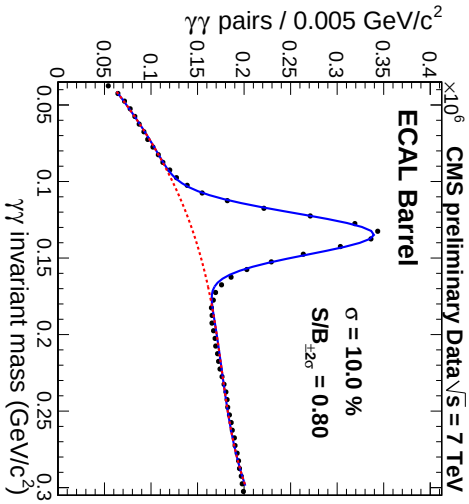


Figure 2.12: *Di-photon invariant mass spectrum reconstructed by ECAL with about 250 nb^{-1} of data at 7 TeV. The π^0 peak is visible. The mass resolution is of the order of 10%.*

For trigger purposes, the ECAL crystals are grouped together into 68 *trigger towers*, whose boundaries line up with the subdivisions of the HCAL. In the endcaps, a *preshower* device with higher granularity, consisting of two lead radiators and two planes of silicon strip detectors, are used to distinguish between showers initiated by neutral pions and photons, or charged pions and electrons. The relatively low light yield of the crystals (about 30 γ/MeV) requires photodetectors with intrinsic high gain that can operate in a magnetic field. Silicon Avalanche Photodiodes (APDs) and Vacuum Phototriodes (VPTs) are used to collect the scintillation light in the barrel and in the endcaps, respectively. The energy resolution of a calorimeter can be parametrised as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{a}{\sqrt{E}}\right)^2 + \left(\frac{\sigma_n}{E}\right)^2 + c^2$$

where a is called stochastic term and includes the effects of fluctuations in the number of photo-electrons, as well as in the shower containment; σ_n is the noise from the electronics and pile-up; and c is a constant term related to the calibration of the calorimeter. The values of the three constants measured on test beams are reported in Table 2.2. The different contributions are shown in Figure 2.13.

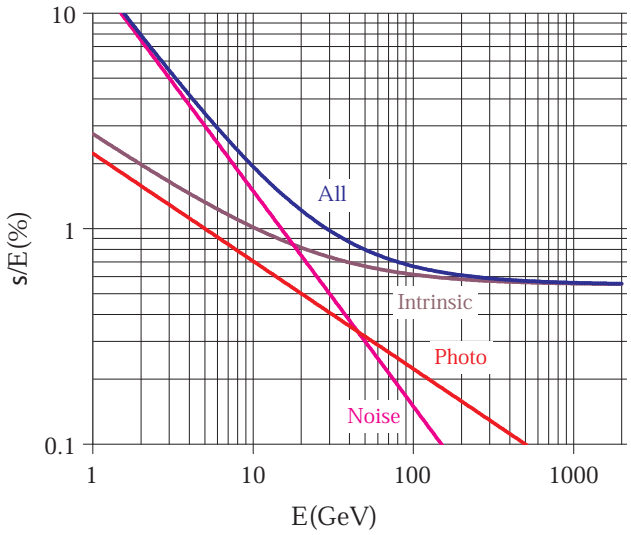


Figure 2.13: *Different contributions to the energy resolution of the ECAL. The curve labelled "intrinsic" includes the shower containment and a constant term of 0.55%.*

Contribution	Barrel (η =0)	Endcap ($\|\eta\|$ =2)
Stochastic term	$2.7\%/\sqrt{E}$	$5.7\%/\sqrt{E}$
Constant term	0.55%	0.55%
Noise (high luminosity)	0.155 GeV	0.205 GeV
Noise (low luminosity)	0.210 GeV	0.245 GeV

Table 2.2: *Contribution to the energy resolution of ECAL (the energy E is expressed in GeV).*

2.2.4 The Hadronic Calorimeter

The hadronic calorimeter (HCAL) [33] plays an essential role in the identification and measurement of hadrons and neutrinos by measuring the energy and direction of jets and missing transverse energy flow in the event. One of the main design requirements for the HCAL is therefore a high hermeticity. In particular, the HCAL angular coverage must include the very forward region, since the identification of forward jets is very important for the rejection of many backgrounds.
The CMS HCAL is thus subdivided in four regions, which provide a good segmentation, a moderate energy resolution and a full angular coverage up to $|\eta| = 5$. The *barrel hadronic calorimeter* (HB) surrounds the electromagnetic calorimeter and covers the central pseudorapidity region up to $|\eta| = 1.3$. The end regions are covered up to $|\eta| = 3$ by the two *endcap hadron calorimeters* (HE). The HB and HE are located inside the solenoid

magnet. To satisfy the hermeticity requirements, then, two *forward hadronic calorimeters* (HF) surround the beam pipe at $|z| = 11$ m, extending the pseudorapidity coverage up to $|\eta| = 5$. Finally, an array of scintillators located outside the magnet, which is referred to as the *outer hadronic calorimeter* (HO), is used to improve the central shower containment.
The overall HCAL is assembled with essentially no uninstrumented cracks or dead areas. Even the gap between the HB and the HE, through which the services and cables of the ECAL and the tracker pass, is inclined at 53° and points away from the centre of the detector.
The HB and HE are sampling calorimeters with active plastic scintillators interleaved with brass plates. This absorber material has been chosen because of its reasonably short interaction length. Moreover, it is non-magnetic. The read-out system is composed of wavelength-shifting fibres. Apart from the first layer, which is read out separately, all the other layers are read out together in towers of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ rad.
The HCAL depth, in terms of nuclear absorption length, goes from $5.15\,\lambda_0$ at $\eta = 0$ to $9.1\,\lambda_0$ at $|\eta| = 1.3$, and is $10.5\,\lambda_0$ in the endcaps. In the barrel region, such depth is not enough to contain the full shower. This is the reason why an additional "tail catcher", the HO, is placed outside the magnet.
The energy resolution (epressed in GeV) is

$$\sigma_E/E \sim 65\%\sqrt{E} \oplus 5\%$$

in the barrel,

$$\sigma_E/E \sim 85\%\sqrt{E} \oplus 5\%$$

in the endcaps, and

$$\sigma_E/E \sim 100\%\sqrt{E} \oplus 5\%$$

in the very forward calorimeter.

2.2.5 The Muon System

The muon spectrometer [34] is placed outside the magnet and is embedded in the iron return yoke, so that the 1.8 T average magnetic field can be used as bending field to enhance muon p_T measurements. It covers the pseudorapidity region $|\eta| < 2.4$ and it performs three main tasks: triggering on muons, identifying muons, and assisting the tracker in measuring the momentum and charge of high-p_T muons.
The spectrometer is composed by 3 types of gaseous particle detectors (see Figure 2.14): Drift Tube (DT) chambers and Cathode Strip Chambers (CSC) are used in the regions of $|\eta| < 1.2$ and $0.9 < |\eta| < 2.4$, respectively, and are complemented by a system of Resistive Plate Chambers (RPC) covering the range of $|\eta| < 1.6$. The use of these different technologies defines three regions in the detector, referred to as barrel ($|\eta| < 0.9$), overlap ($0.9 < |\eta| < 1.2$), and endcap ($|\eta| > 1.2$). The reason for these different technologies lies in the different particle rates and occupancies, both higher in the endcaps, and in the intensity of the stray magnetic field, which is lower in the barrel.

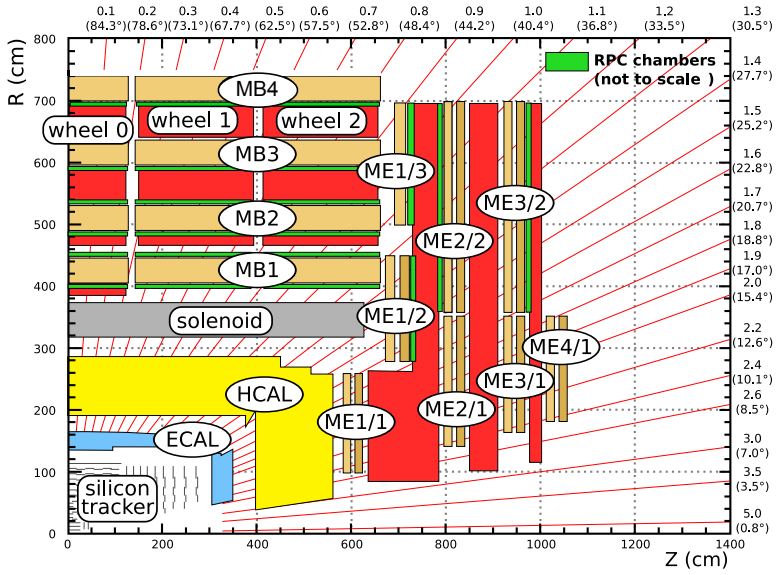


Figure 2.14: *An r–z cross-section of a quadrant of the CMS detector with the axis parallel to the beam (z) running horizontally and radius (r) increasing upward. The interaction region is at the lower left corner. Shown are the locations of the various muon stations and the steel disks (red areas). The four drift tube (DT) stations are labeled MB ("muon barrel") and the cathode strip chambers (CSC) are labeled ME ("muon endcap"). Resistive plate chambers (RPC, in green) are in both the barrel and the endcaps of CMS.*

Drift Tubes Chambers

In the barrel region, the expected occupancy is low (< 10 Hz/cm^2), allowing for the use of drift tubes as detection element. In Figure 2.15, a section of a drift tube cell is illustrated.

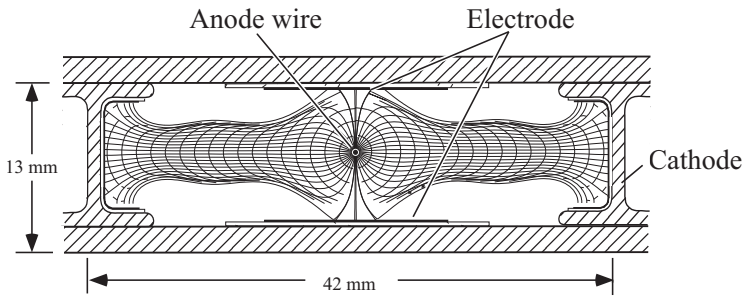


Figure 2.15: *Section of a drift tube cell.*

The DT chamber design is very redundant: each chamber is made up of twelve layers of DTs, grouped in three independent subunits, called *superlayers* (SLs). Two of them have the anode wires parallel to the beam axis, to measure the transverse coordinate *r-ϕ*; the remaining SL is placed between the other two and is orthogonal to them, in order to determine the longitudinal coordinate *z*. The structure of a DT chamber is shown inFigure 2.16.

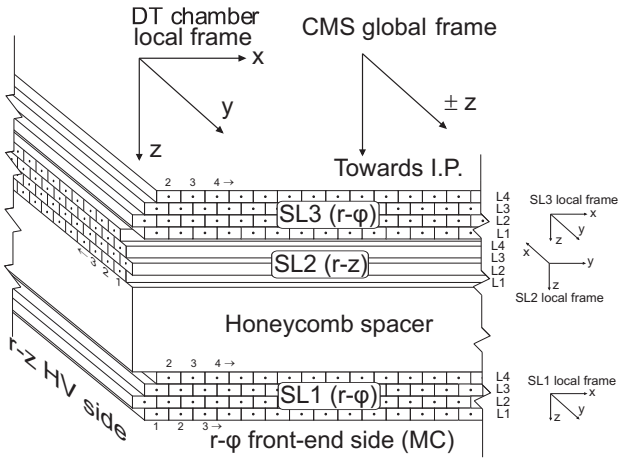


Figure 2.16: *Schematic view of a DT chamber, also showing the different orientation of the SLs.*

The DTs are made of parallel aluminium plates, with cells obtained with

perpendicular "I"-shaped aluminium cathodes. The anodes are 50 μm diameter steel wires placed between the cathodes. Field shaping is improved with two biased and insulated strips glued in correspondence to the wire. The internal volume is filled with a gas mixture of Ar (85%) and CO_2 (15%) at atmospheric pressure, because this gas is non-flammable and can be safely used in underground operations in large volumes, as required in CMS. The single hit position resolution is about 260 μm at nominal voltage values, with a single cell efficiency close to 100%. The angular resolution of the full chamber is about 1.8 mrad in the bending plane [35].

Cathode Strip Chambers

Due to the larger occupancy of the endcap regions, from few Hz/cm^2 to more that 100 Hz/cm^2, and the intense and non-uniform magnetic field, cathode strip chambers have been chosen in this region. The CSCs are multiwire proportional chambers with one cathode plane segmented in strips running orthogonal to the wires. An avalanche developed on a wire induces a distributed charge on the cathode plane. The orthogonal orientation of the cathode strips with respect to the wires allows the determination of two coordinates from a single detector plane (see Figure 2.17). Each chamber is formed by six trapezoidal layers, with strips in the radial direction for a precise measurement of the azimuthal coordinate ϕ.

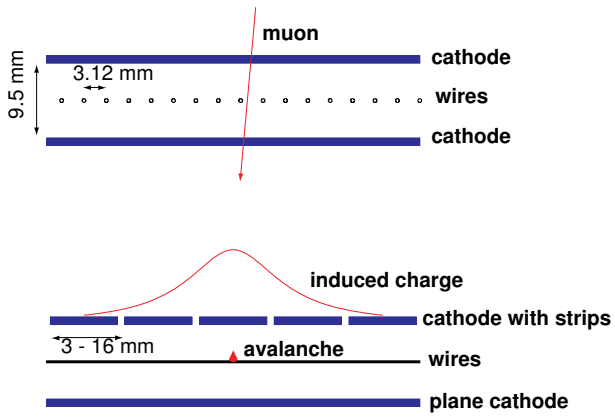


Figure 2.17: *Orthogonal sections of a cathode strip chamber.*

Wires have a resolution of about 0.5 cm, strips about 50 μm. The full chamber spatial resolution varies from about 50 μm in the first CSC station to about 250 μm in the fourth [36].

Resistive Plate Chambers

Resistive plate chambers are installed both in the barrel and the endcap regions. They have a limited spatial resolution, but an excellent time resolution, thanks to their fast response of about 3 ns. This feature makes the RPCs suitable for trigger purposes: they are used as a dedicated trigger subsystem, mainly for unambiguous bunch crossing identification. The RPCs (Figure 2.18) used in CMS are "double-gap" RPCs, made of four bakelite planes forming two coupled gaps 2 mm thick, filled with a $C_2 H_2 F_4$ and $C_4 H_{10}$ gas mixture. They operate in avalanche mode: a moderate electric field across the gap allows to sustain a higher rate, but a robust front-end signal amplification is needed, since the gas multiplication is reduced.

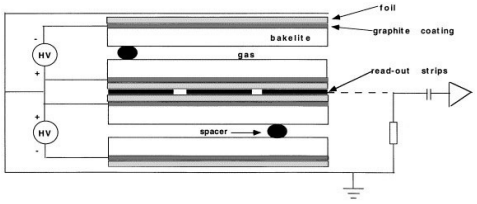


Figure 2.18: *Section of a double gap resistive plate chamber.*

2.3 The CMS Trigger System

To reach the nominal luminosity of 10^{34} cm^{-2} s^{-1}, LHC will operate at a bunch crossing rate of 40 MHz with 80% of the bunches filled. At the current luminosity of $3 \cdot 10^{33}$ cm^{-2} s^{-1} LHC is operating at a bunch crossing rate of 20 MHz with 47% colliding bunches. Such a high luminosity has the drawback that several interactions overlap in the same bunch crossing. This effect is referred to as *pile-up*, and also accounts for the overlap in the detector of signals from different bunch crossings, due to the limited speed of detector response and read-out. At the nominal luminosity 17 inelastic non-diffractive interactions per bunch crossing are foreseen; with the current luminosity a mean of 10 reconstructed primary vertices per collision are observed at CMS.
In total CMS has $\sim 10^8$ data channels that are checked each bunch crossing; the data size per event is $\sim$ 1 MB.
Technical difficulties in handling, storing and processing this extremely large amounts of data impose a reduction factor on the rate of events that can be written to permanent storage. As a consequence one of the most important and difficult aspects of the experiment is the Trigger, the real-time selection and recording of the useful events.
Given that, at the nominal energy, the total inelastic cross-section at LHC is expected to be σ_T = 55 mb and that the most interesting processes that the CMS physics program is expected to cover (see Figure 2.19) have a cross-section of the order of hundreds of pb, the CMS trigger is designed to perform a data reduction from 32 MHz down to $\sim$ 100 Hz. The selection of the events is based on their physics content, so that the online algorithms must have a level of sophistication comparable to that of the off-line reconstruction preserving the efficiency to identify physics objects as high as possible.
On the other hand, the time available to accept or reject an event is extremely limited, given the bunch crossing time of 25 ns. In such a short time interval, it is impossible to read out all raw data from the detector. For this reason, CMS adopts a multi-level trigger design, where each step of the selection uses only part of the available data. In this way, higher trigger levels have to process fewer events and can use more refined algorithms to perform a detailed reconstruction.
The criteria for the selection of the events must be as inclusive as possible for unexpected new phenomena that may appear. Moreover the selection must accept data samples to be used for the calculation of efficiency of the online and offline reconstruction algorithms and filters; for the same reason the events selected must be tagged to indicate the reason for their selection.
The CMS trigger is structured in two physical levels: the *L1 Trigger* (L1) [37] and the *High Level Trigger* (HLT) [38]. The L1 is based on custom-made

hardware and uses only coarsely segmented data from calorimeters and muon detectors, while all the high-resolution data is held in pipeline memories in the front-end electronics. The HLT, instead, is software implemented in a single processor farm, and it is organised in several logical levels of increasing complexity, each accessing more data than the previous one.
The following sections briefly describe the design of the L1-Trigger and the general structure of the HLT.

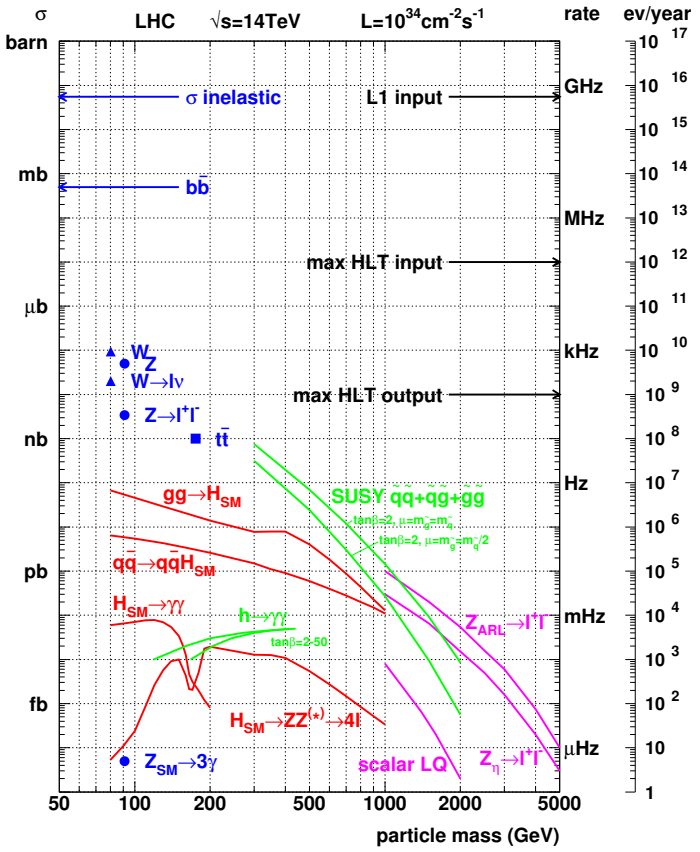


Figure 2.19: *Cross sections and event rates at the LHC design luminosity as a function of particle masses. The L1 and HLT input and output rates are indicated.*

2.3.1 The Level-1 Trigger

The Level-1 trigger is implemented on custom-built programmable hardware. It runs dead-time free and has to take an *accept-reject* decision for each bunch crossing, i.e. every 25 ns. This is achieved with a synchronous pipelined structure of processing elements, each taking less than 25 ns to complete. At every bunch crossing, each processing element sends its results to the next element and receives a new event to analyze. During this process, the full detector data are stored in pipeline memories, whose depth is technically limited to 128 bunch crossings. The Level-1 decision is therefore taken after a fixed time of 3.2 μs. This time must include also the transmission time between the detector and the counting room (a cable path of up to 90 m each way) and, in the case of DT detectors, the electron drift times (up to 400 ns). The time available for calculations can be no more than 1 μs.
As fast processing is needed, the data used by the L1-Trigger system have coarser granularity and lower resolution than the full information from the front-end electronics. In brief, the calorimeters and the muon subdetectors provide trigger primitives in the form of local energy deposits in calorimeter trigger towers and track segments or hits in muon chambers (DT, RPC, CSC, HCAL, ECAL, HF). Regional and global processors identify trigger objects: muon candidates from the L1 Muon Trigger, electron/photon and jet candidates, total and missing transverse energy sums from the L1 Calorimeter Trigger. A full set of trigger primitives are produced every 25 ns. Objects are ranked and sorted. They form the basis for trigger decisions taken by the final L1 stage, the Global Trigger (GT), according to programmable algorithms. The Trigger Control System (TCS) determines if the subdetectors are ready to read out the event, and if the data acquisition (DAQ) system is ready to receive it. Data from trigger primitives, regional energy sums, muon candidates from each sub-detector, and final trigger objects are sent in parallel to the DAQ for each accepted event. Control and monitoring of the L1 trigger operation are performed centrally by dedicated software. A schematic representation of the L1 trigger is provided in Figure 2.20.
With this complex system the L1 Trigger has to reduce the event rate to about 100 kHz. The performance of the L1-Trigger measured during the CMS commissioning with cosmic muons and with first LHC beams is described in Ref. [39].

The L1 Calorimeter Trigger

Trigger Primitives. For triggering purposes the barrel and endcap calorimeters are subdivided in trigger towers. The pattern of energy deposited in

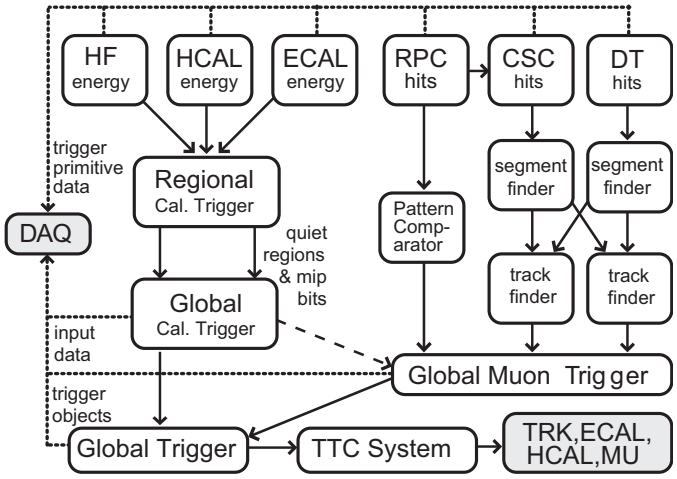


Figure 2.20: *Structure of the Level-1 Trigger system [37].*

those towers is analyzed to identify electron/photon and jet candidates, and the tower energies are summed to obtain the candidate transverse energy (E_T).
A trigger primitive is generated for each trigger tower in the ECAL and HCAL detectors, up to $|\eta| = 3.0$. The towers have the same segmentation in both the ECAL and HCAL. Their size is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ in the barrel and in the endcaps up to $|\eta| = 1.8$. For $|\eta| > 1.8$, the tower segmentation in η increases to $\Delta\eta = 0.1 - 0.35$. Trigger primitives from the forward region, which covers the range $3.0 < |\eta| < 5.0$, are used for jet and energy sum triggers only. A single trigger primitive is generated for each HF trigger region, which are equal to $3\eta \times 2\phi$ readout towers, and are of constant size; $\Delta\eta \times \Delta\phi = 0.5 \times 0.349$.
In ECAL the E_T estimate for the trigger tower is transferred to the Regional Calorimeter Trigger (RCT) together with an additional fine-grain veto bit that is set if the highest two adjacent strips in the tower contain less than 90% of the total E_T. This gives some indication of the lateral shower shape, and can be used to reject L1 electron/photon (e/γ) candidates that result from physical jets. A fine-grain bit, used by dedicated minimum bias triggers, is set for each HF trigger region if one or more of the 6 readout towers entering the sum has E_T above a programmable threshold and it is also sent to the RCT.

RCT. The RCT receives the ECAL and HCAL trigger primitives in 18 electronics crates, each covering one half of the detector in z and 40° in ϕ. The Electron Identification Cards then identify e/γ candidates up to $|\eta| \sim 2.5$, using a sliding window algorithm based on 3×3 trigger towers, with the central tower of the 3×3 window required to have greater

E_T than its neighbours. The resulting candidates are classified as isolated or non-isolated, according to the ECAL trigger primitive fine-grain veto information, and to the ratio of HCAL to ECAL E_T, calculated in the RCT. The E_T of the e/γ candidate is taken as the sum of the transverse energy in the central tower and its highest-E_T neighbour, and a coarse position is assigned as the centre of the 4×4 tower region in which the candidate is contained. Each RCT crate transmits up to four isolated and four non-isolated e/γ candidates to the Global Calorimeter Trigger (GCT).
In addition, the Receiver Cards sum the ECAL and HCAL tower E_T values over non-overlapping 4×4 towers (for barrel and endcaps) and forward these region sums via the Jet Summary card to the GCT. For each region, the RCT sends a τ-veto bit to the GCT, which indicates that the tower energy is spread out over multiple towers, rather than contained in a small number of contiguous towers, and hence is not consistent with a τ-lepton decay. The HF trigger regions are forwarded directly to the GCT without processing.

GCT. The GCT hardware [40] has been completely redesigned since the L1 Trigger Technical Design Report, to take advantage of new technologies and improve the robustness of this complex system. The e/γ candidates received are sorted based on E_T, and the highest four isolated and non-isolated candidates are forwarded to the GT. The region sums are processed, finding jets and computing their E_T. Jet candidates are identified using a 3×3 sliding window of trigger regions (equivalent to 12×12 trigger towers, or 1.05×1.05 in $\eta \times \phi$). The jet-finder algorithm is described in detail in [41]. After jets are found, Look-Up Tables (LUTs) are used to apply a programmable η-dependent jet energy scale correction. Jets found with $|\eta| > 3.0$ are classified as forward jets. Those found with $|\eta| < 3.0$ are classified as central or τ, depending on the OR of the nine τ-veto bits associated with the 9 regions in the 3×3 window. The GCT also calculates total and missing E_T from the trigger regions, and total and missing H_T. The total H_T is the scalar sum of E_T identified in jets, and missing E_T is the corresponding vector sum in the $x - y$ plane. Finally, minimum-bias trigger quantities are formed by summing E_T in rings around the beampipe in the HF calorimeter (for $4 < |\eta| < 4.5$ and $4.5 < |\eta| < 5$), and by counting fine-grain bits set for each HF trigger tower.
The four highest E_T jets in each of the central, τ and forward categories are sent to the GT, along with E_T^{total}, E_T^{miss}, H_T^{total}, H_T^{miss} and the minimum-bias quantities. The GCT transmits all input and output data to the DAQ for each triggered event, to be used for diagnostics, monitoring and to define define HLT regions of interest.

The L1 Muon Trigger

The Muon Trigger has the task to identify muons, reconstruct their position and transverse momentum and provide bunch crossing assignment with high purity and efficiency. Its scheme relies on 2 independent and complementary triggering technologies: one based on the precise tracking detectors in the barrel and endcaps (DTs, CSCs), and the other based on the RPCs. The tracking detectors provide good space and time resolution, while the RPC system provides excellent timing with somewhat lower resolution; this is needed to reinforce the measurement of the correct beam-crossing time at the highest LHC luminosities.
Each of the L1 muon trigger subsystems has its own trigger logic. The DT and CSC share trigger information in the overlap region, enabling each of the three muon subdetectors to deliver its own list of up to four muon candidates, ranked and sorted according to decreasing reconstruction quality and transverse momentum, to the Global Muon Trigger (GMT) [42]. This then combines them and forwards up to four candidates to the GT.

DT Trigger. The on-chamber electronics produces trigger primitives, consisting of information about identified track segments. It is directly connected to the read-out electronics, and performs a straight segment fit within a superlayer using at least three hits out of the four layers of drift cells. In $r-z$ superlayers, only segments pointing to the interaction point are selected. The segments reconstructed in the two $r-\phi$ superlayers in each chamber are matched by the Track Correlator (TRACO), that improves the angular resolution thanks to the bigger lever arm. Trigger primitives from a given muon sector are sent to the SectorCollector electronics, located outside of the detector. The signals from each station are synchronised, coded, and forwarded through high-speed optical links [43] to the Drift Tube Track Finder (DTTF) located in the underground counting room adjacent to the detector. The DTTF also receives trigger primitives from the CSCs for the barrel-endcap overlap region. The DTTF system performs a matching between trigger primitives received from the DT stations and assigns a quality code as well as ϕ, η, charge and transverse momentum (p_{T}) values to the reconstructed muon track. The track matching is based on extrapolation.
The Sector Collector and the DTTF also read out their input and output data for several time samples around the triggered event for diagnostics and monitoring.

CSC Trigger. In each chamber, the track segment position, angle and bunch crossing, are first determined separately in the nearly orthogonal anode and cathode views. The cathode readout is optimised to measure the ϕ-coordinate, while the anode readout is optimised to identify

the bunch crossing. The front-end electronics boards reconstruct track segments using pattern-recognition firmware based on pattern templates. These templates require track segments in cathode as well as anode views to point towards the interaction point, with an angular acceptance, of the order one radian, which depends on the station. The track segments from the cathode and anode readout from each chamber are finally combined into 3-dimensional local tracks, which are the CSC trigger primitives. The trigger primitives are collected by the Muon Port Cards, which sort them and send up to three candidates to the CSC Track Finder (CSCTF) via optical fibres. The CSCTF matches trigger primitives to form complete tracks and determine their p_T, η, ϕ and charge. The CSCTF functionality is described in more detail in Ref. [44]. The CSCTF also receives trigger primitives from the DT system for the barrel-endcap overlap region. The CSCTF is optimized to cope with the non-axial magnetic field present in the endcap region. Thus, the algorithms of the CSCTF are inherently 3-dimensional to achieve maximum background rejection, in particular for low momentum tracks.
In addition to the p_T, η, ϕ, and charge, each track identified by the CSCTF carries a quality code. This quality code is used along with the p_T to sort the candidates; the highest-ranking four are sent to the Global Muon Trigger. The quality code is a two-bit word that is used to indicate the expected coarse p_T resolution. Quality 3 (high p_T resolution) refers to a three- or four-segment track with one of the segments in ME1. Quality 2 (medium p_T resolution) refers to a 2-segment coincidence with one of the segments in ME1. Quality 1 (low p_T resolution) refers to any other 2-segment coincidence. Quality 3 candidates, with $5 < p_T < 35$ GeV/c, are expected to have about 20% resolution in p_T, while Quality 2 are expected to have about 30%. Very poor resolution is expected for Quality 1 candidates.
This is true for the pseudorapidity region up to 2.1; in the $2.1 < \eta < 2.4$ the ME1/1 chambers have coarser readout granularity due to constraints in cabling of the detector that results in a worse p_T resolution of the CSC candidates that comes from this region. In particular this is causing a triple p_T measurement ambiguity of the candidate; the choice of retaining the highest p_T one is currently done to save efficiency despite it affects the trigger rate.
In addition to identifying muons originating from the interaction point, the CSCTF identifies tracks from "halo muons", coming from the interaction of the LHC beam with the gas particles in the beam pipe or with the beam pipe itself.

RPC Trigger. The RPC trigger is based on the spatial and temporal coincidence of hits in several layers. The Pattern Comparator trigger logic [45] compares signals from all four muon stations to predefined hit pat-

terns in order to find muon candidates. Muon p_T, charge, η, and ϕ are assigned according to the matched pattern. The algorithm requires a minimum number of hit planes, which varies with the p_T and location of the muon. Either 4/6 (four out of six), 4/5, 3/4 or 3/3 hit layers are minimally required. A quality value, encoded in two bits, reflects the number of hit layers.
Analog signals from the chambers are digitized by Front End Boards, then zero-suppressed and assigned to the proper bunch crossing. They are then sent via optical links to Trigger Boards located in the underground counting room. Each of the 84 Trigger Boards can produce up to four muon candidates for every bunch crossing. The Final Sorter Board sorts the candidates by quality and p_{T}, and sends up to eight muon candidates, four from the barrel and four from the endcaps, to the GMT.

Global Muon Trigger (GMT). The Global Muon Trigger receives up to 4 candidates from each of the DTTF and CSCTF and up to 8 candidates (4 in the barrel, and 4 in the endcap) from the RPC trigger. Look-up tables (LUTs) are used to combine candidates identified by more than one subdetector, and to assign a quality code based on the number of subdetectors involved, as well as on the quality of the track candidates, as assigned by the track-finders; these quality codes and their definitions are listed in Table 2.4. The GMT returns the reconstructed p_{T} at the interaction point of each of the four best muons in discrete momentum bins, with the highest bin at 140 GeV/c. These p_{T} bins are listed in tTable 2.3.
It also returns the global position of the hit which made the L1 decision at the second muon station (MB2 in the barrel, ME2 in the endcaps). This set of information will be called *L1 Muon particle*. The four highest quality and p_{T} muon candidates are forwarded to the GT.

Table 2.3: *Possible values of L1 Muon Particle discrete momentum bins.*

P_t code	0	1	2	3	4	5	6	7
Low edge of p_t bin $/GeV/c$	/	0.0	1.5	2.0	2.5	3.0	3.5	4.0
P_t code	8	9	10	11	12	13	14	15
Low edge of p_t bin $/GeV/c$	4.5	5.0	6.0	7.0	8.0	10	12	14
P_t code	16	17	18	19	20	21	22	23
Low edge of p_t bin $/GeV/c$	16	18	20	25	30	35	40	45
P_t code	24	25	26	27	28	29	30	31
Low edge of p_t bin $/GeV/c$	50	60	70	80	90	100	120	140

Table 2.4: *Possible values of L1 Muon Particle quality bit and the corresponding definitions; Quality 2 is not used in single and double muon trigger paths, Quality 3 is not used in single muon trigger path, Quality 4 is currently not used.*

Q	Definition
0	no muon
1	beam halo muon
2	RPC unmatched to DT or CSC with Q=1,2,3 in particular η regions
3	CSC unmatched to RPC with Q=1 and $\eta < 2.1$
3	CSC unmatched to RPC with Q=2 and ($\eta < 1.5$ or $\eta > 1.8$)
4	very low quality category (not used)
5	RPC unmatched to DT or CSC
6	DT or CSC unmatched to RPC
7	matched DT-RPC or CSC-RPC

The L1 Global Trigger

The main task of the Global Trigger is to reject or to accept events for readout and further processing by the high-level trigger. Before performing trigger algorithm calculations, it has to first receive and synchronise the muon and calorimeter input data. The data are then transmitted to the Global Trigger Logic (GTL) board. This unit is programmed to provide a menu of up to 128 algorithms, which can transform logical combinations of L1 trigger objects (muons, jets, e/γ, calorimeter transverse energy sum, etc.) with selection criteria (energy/momentum thresholds, quality bit, etc.) into decision bits. These bits can be enabled to contribute to a final OR of decisions which determines whether the data are read out; combinations and thresholds are programmable and depend on the istantaneous luminosity scenario. The list of enabled triggers is called, in the CMS jargon, L1 Trigger Menu; for example in the menu for $\mathcal{L}$=3 $\cdot$ 10^{33} cm^{-2} s^{-1}, the lowest p_T threshold allowed for a selection criteria based on single muon objects with quality > 3 is set to 16 GeV/c (L1_SingleMu16).

2.3.2 The High Level Trigger

The subsequent data reduction factor of $\mathcal{O}$(1000) is obtained by the software-implemented HLT.
The HLT processes all events accepted by the Level-1 trigger in a single processor farm; the reconstruction and selection in the HLT takes place in steps and the selection is optimized step by step in order to reject events as quickly as possible. The basic idea is to reconstruct those part of each physics object that can be used for selection while minimizing the overall CPU usage.
As an example, the reconstruction of a muon includes the reconstruction of

a track in the muon spectrometer, the search of a matching tracker track, and the global track reconstruction; at the end on each step a set of selection criteria results in the rejection of a significant fraction of the events accepted by the previous step. The rate of events that need to be processed through the remaining algorithms is decreased reducing the required CPU time. Reconstruction and selection are therefore closely intertwined in the online environment of the filter farm.
As a convention the term Level-2 Trigger is used to refer to algorithms and requirements of the first selection step in the HLT process; tipically, a Level-2 trigger, which has the full Level-1 rate as input, uses only information from the calorimeter and muon detectors. In contrast, Level-3 Trigger refers to a further selection step that includes the reconstruction of full tracks in the tracker. Because of the high number of channels, the complex pattern recognition and higher combinatorics, track reconstruction is a process that demands large amount of CPU time; thus tracker-tracks are used only for the L2 pre-selected events.

To minimize the CPU time required by the HLT, a key feature of the algorithms is to reconstruct the information in the CMS detector only partially. In many cases the decision on whether en event should be accepted by the HLT involves the reconstruction of quantities in only a limited region of the detector. As an example, for an event accepted by the Level-1 trigger in the inclusive muon stream, only the parts of the muon chambers indicated by the Level-1 trigger results, and the corresponding region in the tracker, need to be considered for the validation of the muon.
In the case of the HLT, the reconstruction of physics object is driven by the corresponding candidates identified by the Level-1 trigger. This approach leads to significant CPU savings.

A list of reconstruction algorithms and filters for one or more physics objects that starts any time a L1 bit is enabled is called *H*LT Path. For each luminosity scenario the set of trigger paths that, if enabled, contribute to a final OR of decisions which determines whether to reject or store an event, is called, in the CMS jargon, HLT Menu. A single trigger path can require the presence of one or more physics objects of a particular type that pass specific kinematic thresholds, and it can also mix different type of physics objects (*cross triggers*). The rate of expected events for a particular set of trigger paths given a luminosity scenario is studied on simulation; with the increase of the luminosity the requirements on the trigger paths are becoming more and more stringent due to rate constraints, and the selection algorithms more complex to cope with the requests of each physics analysis. Trigger paths with lower threshold than those necessary to face the event rate are kept in the HLT Menu with a "prescale" factor applied; this is done to have on storage the data sample required to measure and

monitor the efficiency of the higher threshold trigger paths.
A more detailed description of the Muon and Electron HLT Paths is given in Section 4.4 and Section 5.4.

Chapter 3

Search for the SM Higgs in the $ZZ^{(*)} \to 4\ell$ decay channel with CMS

As described in Chapter 1, the Standard Model predicts the existence of a single physical Higgs scalar boson associated to the spontaneous electroweak symmetry breaking mechanism. The mass m_H of this scalar boson is a free parameter of the theory. In the following sections the Higgs phenomenology at the LHC collider will be described. As it will be shown, the inclusive production of SM Higgs bosons followed by the decay $H \to ZZ^{(*)} \to \ell^{\pm}\ell^{\mp}\ell'^{\pm}\ell'^{\mp}$ with $\ell, \ell' = e$ or μ, in short $H \to 4\ell$, is expected to be one of the most important channel for the Higgs boson discovery or exclusion over a wide range of possible m_H values.

3.1 Higgs Phenomenology at the LHC

The Higgs physics at LHC is different than in previous colliders due mainly to the larger center of mass energy of the pp collisions. Particle production at LHC reach the mass range of few TeV/c^2 in a Feynman x region where the gluon density is much larger than the quark density; for Higgs physics the LHC can be thought as a "gluon-collider".
The experiments at the LHC will search for the Higgs boson within a mass range going from 100 GeV/c^2 to about 1 TeV/c^2. In this section, the main Higgs production and decay processes are described analyzing, for each mass region, the most promising channels to look at for the Higgs exclusion or discovery.
While the Higgs boson mass is not predicted by the theory, its couplings to fermions and bosons are predicted to be proportional to the corresponding particle masses (for fermions) or squared masses (for bosons), as in

Equations 1.18, 1.19 and 1.24. For this reason, the Higgs boson production and decay are dominated by channels involving heavy particles, mainly the $W^{\pm}$ and Z bosons and the third generation fermions. As for the remaining gauge bosons, the Higgs does not couple to photons and gluons at tree level, but only by one-loop graphs where the main contribution is given by $q\bar{q}$ loops for the $gg \to H$ channel and by W^+W^- and $q\bar{q}$ loops for the $\gamma\gamma \to H$ channel.

3.1.1 Higgs Production

The main processes contributing to the Higgs boson production at a proton-proton collider are represented by the Feynman diagrams in Figure 3.1, and the corresponding cross sections are shown in Figure 3.2, for centre-of-mass energies of 7 and 14 TeV [46]. The former is the energy provided by the LHC during the 2010-2011 runs, the latter is the LHC design energy that will be gradually reached in the next years. The total production cross section at 7 TeV is up to one order of magnitude lower than at 14 TeV as shown in Figure 3.2.

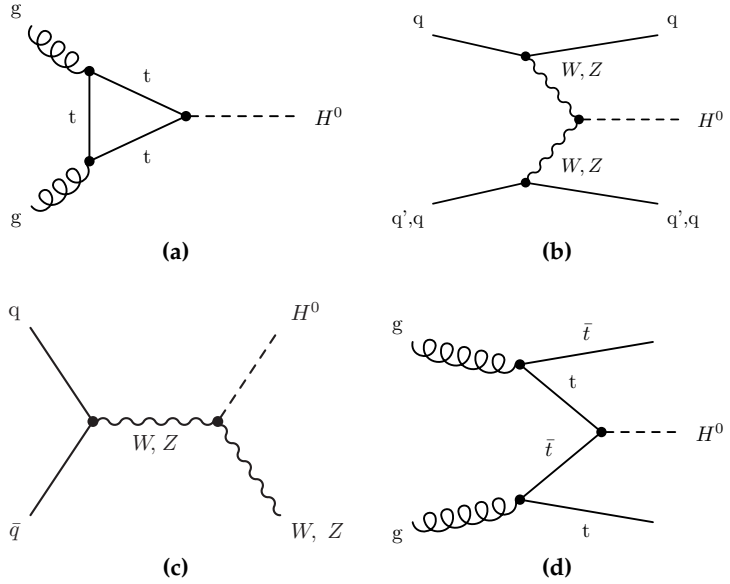


Figure 3.1: *Higgs production mechanisms at tree level in proton-proton collisions: (a) gluon-gluon fusion, (b) vector boson fusion, (c) W and Z associated production (or Higgsstrahlung), and (d) $t\bar{t}$ associated production.*

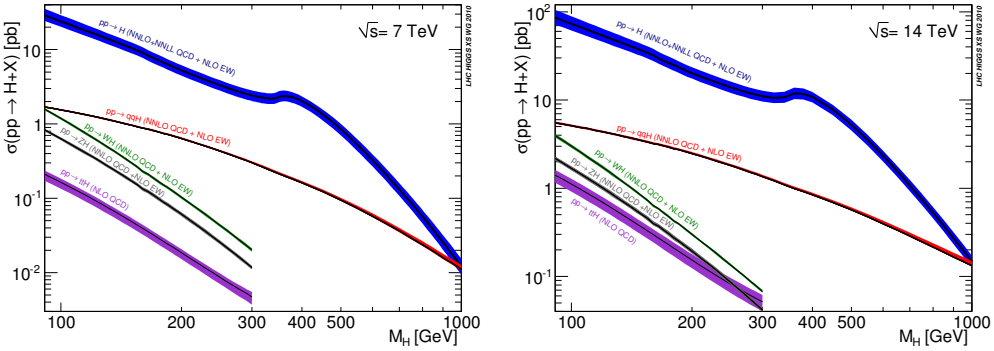


Figure 3.2: *Cross sections for the different Higgs boson production channels, as functions of the Higgs boson mass, at (left) 7 TeV and (right) 14 TeV LHC centre-of-mass energy [46].*

Gluon-Gluon Fusion

The gluon-gluon fusion is the dominating mechanism for the Higgs boson production at the LHC over the whole mass range, because of the high luminosity of gluons. The process is shown in Figure 3.1a with a t quark-loop, which is the main contribution, due to the large coupling constant $g_{\mathrm{H}t}$ (Equation 1.24). The latest results in the computation of the cross section for this process, shown in Figure 3.2 and used in the analysis presented in this thesis, include next-to-next-to-leading order (NNLO) QCD contributions, complemented with next-to-next-to-leading log (NNLL) resummation, and next-to-leading order (NLO) electroweak corrections. An uncertainty of 15-20% on the calulation of this cross section is assumed, mostly depending on the choice of the *parton density functions* (PDFs, see Section 2.1.3) and on the uncalculated higher-order QCD radiative corrections (Section 7.1).

Vector Boson Fusion

The vector boson fusion (VBF, Figure 3.1a) is the second contribution to the Higgs boson production cross section. It is about one order of magnitude lower than gg-fusion for a large range of m_{H} values, and the two processes become comparable only for masses of the order of 1 TeV$/c^2$. Nevertheless, this channel is very interesting because of its clear experimental signature: the presence of two spectator jets with high invariant mass in the forward region provides a powerful tool to tag the signal events and discriminate the backgrounds, thus improving the signal to background ratio, despite the low cross section. Also for this process, NNLO QCD and NLO EW calculations are available. The uncertainties are in general lower than for the gluon fusion mode, of the order of 10%.

Associated Production

In the *Higgsstrahlung* process (Figure 3.1c), the Higgs boson is produced in association with a $W^{\pm}$ or Z boson, which can be used to tag the event and is several orders of magnitude lower than those of gg-fusion and VBF. The cross section for this process is known at the NNLO QCD and NLO EW level. The inclusion of these available contributions increases the LO cross section by about 20-25%.

The last process, illustrated in Figure 3.1d, is the associated production of a Higgs boson with a $t\bar{t}$ pair. Also for this process, the cross section is orders of magnitude lower than those of gluon and vector boson fusion. The presence of the $t\bar{t}$ pair in the final state can provide a good experimental signature. For this cross section, NLO QCD calculations are available.

3.1.2 Higgs Decay

Different strategy for the Higgs boson identification can be developed depending on its mass, as not just the production but also the decaying modes change dramatically across the possible mass range. In Figure 3.3 the branching ratios are shown as a function of the Higgs boson mass.
Fermion decay modes dominate the branching ratio in the low mass region (up to about 150 GeV/c^2). In particular, the channel $H \to b\bar{b}$ gives the largest contribution, since the b quark is the heaviest fermion available. When the decay channels into vector boson pairs open up, they quickly dominate. A peak in the $H \to W^+W^-$ decay is visible around 160 GeV/c^2, when the production of two on-shell W bosons becomes possible and the production of a real ZZ pair is still not allowed. At high masses, above 350 GeV/c^2, also $t\bar{t}$ pairs can be produced.
The total width of the Higgs boson resonance, given by the sum of the partial widths of all possible decay channels, is shown in Figure 3.3 as a function of m_H. Below the $2m_W$ threshold, the Higgs boson width is of the order of the MeV/c^2. In this low mass region, the Higgs boson width is smaller than the experimental mass resolution. Then it rapidly increases, but remains below 1 GeV/c^2 up to $m_H \sim 200$ GeV/c^2. In the high mass region, $m_H > 2m_Z$, the total Higgs boson width is dominated by the W^+W^- and ZZ partial widths, which can be written as follows:

$$\Gamma(H \to W^+W^-) = \frac{g^2}{64\pi}\frac{m_H^3}{m_W^2}\sqrt{1-x_W}\left(1 - x_W + \frac{3}{4}x_W^2\right) \quad (3.1)$$

$$\Gamma(H \to ZZ) = \frac{g^2}{128\pi}\frac{m_H^3}{m_W^2}\sqrt{1-x_Z}\left(1 - x_Z + \frac{3}{4}x_Z^2\right) \quad (3.2)$$

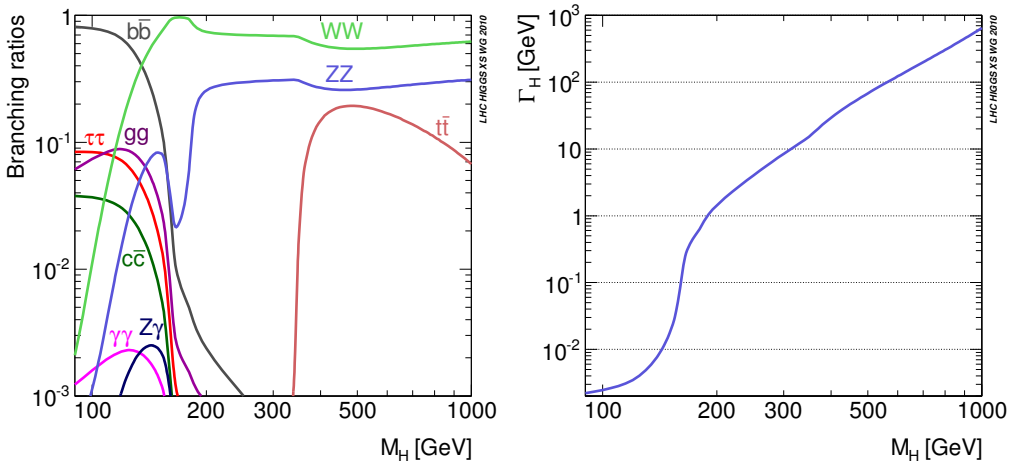


Figure 3.3: *Branching ratio of different Higgs boson decay channels as a function of the Higgs boson mass (right), and total decay width of the Higgs boson as a function of its mass (left) [46].*

where

$$x_W = \frac{4m_W^2}{m_H^2}\,, \qquad x_W = \frac{4m_Z^2}{m_H^2}\,.$$

As the mass grows, x_W, $x_Z \to 0$ and the leading term in Equations 3.1 and 3.2 grows proportionally to m_H^3. Summing over the W^+W^- and ZZ channels, the Higgs resonance width in the high mass region can be written as

$$\Gamma(H \to VV) = \frac{3}{32\pi}\frac{m_H^3}{v^2}\,. \tag{3.3}$$

From Equation 3.3, it can be noted that $\Gamma_H \simeq m_H$ for $m_H \simeq 1.4$ TeV/c^2. If m_H is larger than about 1 TeV/c^2, therefore, it becomes experimentally very problematic to separate the Higgs resonance from the VV continuum. Actually, being the resonance width larger than its own mass, the Higgs boson cannot be properly considered as a particle any more as already discussed in Section 1.4.1.

3.2 Higgs Search at CMS

The most promising decay channels for the Higgs boson exclusion or discovery do not only depend on the corresponding branching ratios, but also on the capability of experimentally detecting the signal while rejecting the backgrounds. In the following these channels will be illustrated for different mass ranges. Given the expected scenario for the first running period of the LHC, i.e. few fb^{-1} of pp collision data collected at 7 TeV, the analyses have been mainly optimized for the best sensitivity in the Higgs

boson exclusion limit. At the design center of mass energy, the LHC plans to collect several hundreds of fb^{-1} of data; if the Higgs exists, the LHC will become a Higgs factory and all its properties will be explored. This farther scenario will not be discussed in the following.
To understand the sensitivity of each channel is useful to multiply production cross section by the branching ratio; these values are compared in Figure 3.4 for the main channels that can be explored at the LHC experiments on the basis of the final state reconstruction efficiency and the signal-over-background ratio.
Figure 3.5 shows the expected exclusion limits as a function of the Higgs boson mass, for the channels and integrated luminosity analyzed by the CMS Collaboration up to the 2011 summer conferences (i.e. between 1.1 and 1.7 fb^{-1}). Here, and in the rest of the present work, the statistic procedure used to determine exclusion limits is the modified frequentist construction (often referred to as CLs or hybrid frequentist-bayesian) [47, 48]. To fully define the method, one needs a choice of the test statistics and of how nuisance parameters are treated in its construction and in generating pseudo-experiment. In this work, the prescription prepared by the LHC Higgs Combination Group [49] has been followed. More details will be provided for the $H\to 4\ell$ case in Section 8.4.
The strategy for the search of the SM Higgs boson at CMS in the 100 GeV/c^2-1 TeV/c^2 mass range, can be described dividing this range into three mass regions. For each of these the most important channels are presented.

Low Mass Region

For $m_H < 135$ GeV/c^2, despite the branching ratio favors the $b\bar{b}$ decay, the process $gg \to H \to b\bar{b}$ is experimentally very difficult to explore because of the enormous background from di-jet production. The $b\bar{b}$ decay has been already explored at CMS in the search for the SM Higgs produced in association with a W or a Z boson (*Higgsstrahlung*) and the following final states have been examined: $W(\mu\nu)H$, $W(e\nu)H$, $Z(\mu\mu)H$, $Z(ee)H$, $Z(\nu\nu)H$. In these cases the event has a clear signature despite the lower cross section x BR; the expected sensitivity in the exclusion limit with an early integrated luminosity of 1.1 fb^{-1} is shown in Figure 3.5. The expected upper limit at 115 GeV/c^2 is found to be 5.7 times the standard model expectation [51].
Similar results are obtained exploring the $\tau\tau$ decay; also in this case, the Higgs boson signal produced via the gluon-gluon fusion process is overwhelmed by the Drell-Yan production of τ-pairs in the mass region of interest. To overcome this the analysis relies on the VBF production of Higgs bosons; the two forward "tagging" jets from the incoming quarks which radiate the vector bosons provide means for distinguishing the Higgs boson production from SM background processes. Three indipendent τ pair

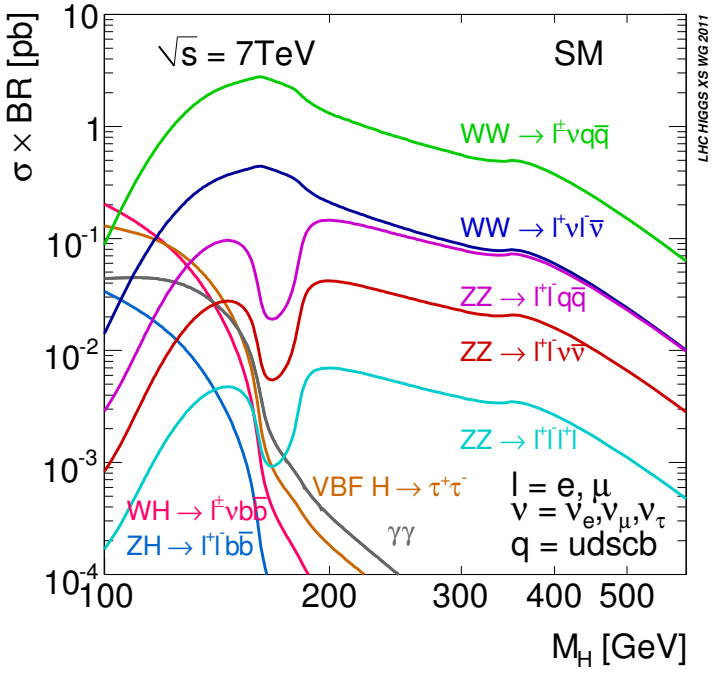


Figure 3.4: *Standard Model Higgs boson production cross section times branching ratio as a function of the Higgs boson mass.*

final states where one or both taus decay leptonically are studied: $e\tau_h$, $\mu\tau_h$, $e\mu$ where the symbol τ_h is used to indicate a reconstructed hadronic decay of a τ [52].
For the Higgs boson search at low mass, the H $\rightarrow \gamma\gamma$ channel seems to be the most promising. Despite its small branching ratio, $\sim$ 0.2% in this region of interest, the H $\rightarrow \gamma\gamma$ provides a clean final-state topology with an effective mass peak that is reconstructed with great precision. The dominant sources of background are the irreducible direct di-photon production and the reducible pp$\rightarrow \gamma$ + jet and pp$\rightarrow$ jet + jet. The expected signal rate is at least an order of magnitude smaller than the SM background rate [53].
From Figure 3.4 it is clear that, despite not favored by the branching ratio, the H$\rightarrow W^+W^- \rightarrow 2l2\nu$ channel can play a role even in this low mass region. The analysis of this channel is performed in CMS dividing the events into three categories according to the event jet multiplicity: H + 0 jets, H + 1 jet, and H + 2 jets. W^+W- candidates, with both W bosons decaying leptonically, are selected in final states consisting of 2 isolated, high-p_T, oppositely-charged leptons (electrons or muons) and large missing transverse energy due to the undetected neutrinos. All Higgs production mechanisms are considered as part of the signal. As shown in Figure 3.5 this

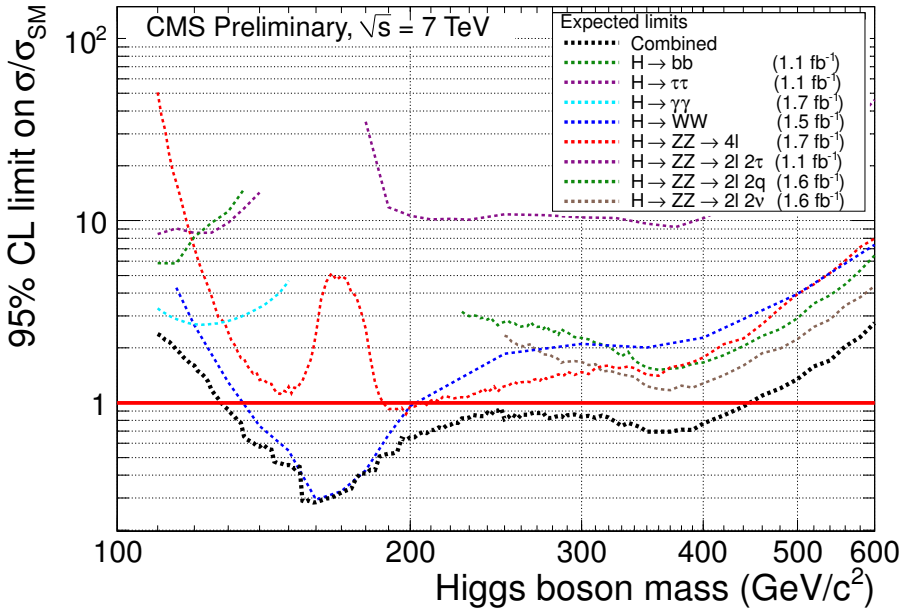


Figure 3.5: *The median expected 95%C.L. upper limits on the signal strength modifier $\mu = \sigma/\sigma_{SM}$ as a function of the SM Higgs boson mass in the range 110-600 GeV/c^2 for the eight principal Higgs boson search channels and their combination for an integrated luminosity between 1.1 and 1.7 fb^{-1}, assuming the background-only hypothesis [50].*

channel is expected to almost reach the exclusion limit in this low mass region with the early integrated luminosity of 1.5 fb^{-1}. As the Higgs mass cannot be reconstructed due to the undetected neutrinos this channel is much more powerful for the exclusion than for the discovery [54].

Among the ZZ channels that are explored at CMS, despite the $\sigma \times BR$ (Figure 3.4) is higher for the $2\ell 2q$ and $2\ell 2\nu$ final states ($\ell, \ell' = e$ or μ), the $H \to ZZ^{(*)} \to 4\ell$ channel, is the only one that can play a role in the low mass region, due to its clear experimental signature that can be reconstructed with high efficiency and very low background contamination. The four-lepton mass resolution is very good, of the order of 1%, thus can contribute to the upper limit for the Higgs boson exclusion as well as to a hypothetical discovery. The expected sensitivity in the low mass region, as shown in Figure 3.4 varies significantly from 115 to 135 GeV/c^2 with an upper limit from 13 to 2 times the SM expectations with 1.7 fb^{-1} of integrated luminosity. The analysis of this channel at CMS is the subject of this thesis, thus the expected and observed sensitivity over the full Higgs boson mass range will be presented in detail in Chapter 8. From these results will be possible to conclude the important role this channel is playing even at low mass as, due to the very low background, its exclusion limit

almost scale linearly with the luminosity increase.

Intermediate Mass Region

For mass values between 140 and 180 GeV/c^2, the Higgs boson decays into WW^* and ZZ^* open up and their branching ratios quickly increase, so the best channels in this mass region are $H \rightarrow WW^{(*)} \rightarrow 2\ell 2\nu$ and $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$. The branching ratio of $H \rightarrow WW^{(*)}$ is higher, because of the stronger coupling of the Higgs boson to charged current than to neutral current. Moreover, this decay mode becomes particularly important in the mass region between $2m_W$ and $2m_Z$, where the Higgs boson can decay into two real W's and not yet into two real Z's, and its branching ratio is close to one. This directly reflects in a reduced sensitivity of the 4ℓ channel for the Higgs boson exclusion around m_H=160 GeV/c^2. Despite the lower branching ratio the 4ℓ channel remains the best candidate for a discovery in this mass range for the reasons previously discussed.

High Mass Region

This region corresponds to mass values above the $2m_Z$ threshold, where the Higgs boson can decay into a real ZZ pair. Though the $H \rightarrow ZZ$ partial width is lower than the $H \rightarrow WW$ one, the $H \rightarrow ZZ$ channels play the major roles. The 4ℓ decay is the "golden channel" for a high mass Higgs boson discovery. Moreover for masses higher then 250 GeV/c^2 the $H \rightarrow ZZ \rightarrow 2\ell 2\nu$ and $H \rightarrow ZZ \rightarrow 2\ell 2q$ can provide an important contribution to the combined sensitivity to the Higgs boson exclusion. In particular for an early integrated luminosity of 1.6 fb^{-1} the expected upper limit between 250 to 600 GeV/c^2 is found to vary between 2 to 4 times the standard model for the $2\ell 2\nu$ final state and between 2 to 6 times for the $2\ell 2q$ final state [55] [56].
The worsening of limits at high masses is caused by the decreasing signal cross section.

3.3 The 4ℓ final state

In the following 4ℓ final state signal and background processes are described in more detail.
In the analysis of the H$\rightarrow$ 4ℓ channel that will be presented in this thesis, all Higgs boson production mechanisms are considered as part of the signal. As described in Section 3.1.1, the main diagrams associated to the signal are those described in Figure 3.6. The final states considered are: 4μ, 4e, 2μ2e.

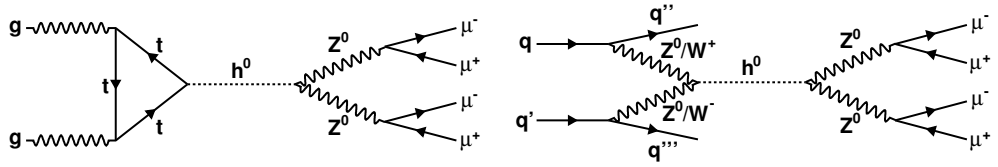


Figure 3.6: *Lowest order signal diagrams: with the Higgs boson produced via gluon-gluon fusion (left) and via vector boson fusion (right).*

Four-leptons events from non resonant di-boson production constitute the main source of background events. In the following this background will be called for simplicity ZZ and will be considered in the category of *irreducible* backgrounds, as the event topology and kinematic is very similar to those of signal events. The lowest order production mechanism is the one represented in Figure 3.7 (left), $q\bar{q} \to ZZ^*/Z\gamma^*$. The gluon-induced ZZ background, although technically of NNLO compared to the first order Z-pair production, amounts to a non-negligible fraction of the total irreducible background (see Section 3.4) at masses above the $2m_Z$ threshold. The associated diagram is represented in Figure 3.7 (right).

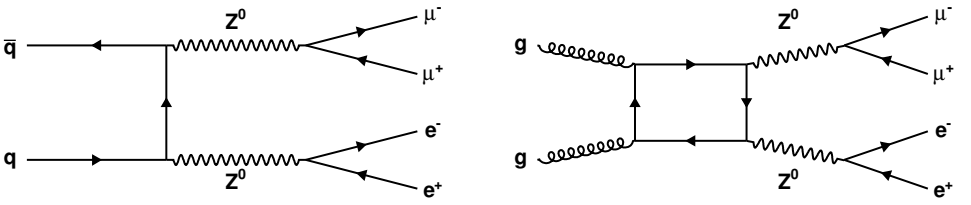


Figure 3.7: *Lowest order diagrams for the $q\bar{q} \to ZZ^*/Z\gamma^*$ process (left) and for the $gg \to ZZ^*/Z\gamma^*$ process (right).*

In the category of *reducible* backgrounds, those with final state leptons coming from b (c) decays are the most important; the main source of background events of this type are the $Zb\bar{b}$ (and $Zc\bar{c}$) associated production with $Z \to \ell^+\ell^-$ decays, and the production of top quark pairs in the decay mode $t\bar{t} \to WbW\bar{b} \to \ell^+\ell^-\nu\bar{\nu}b\bar{b}$ (Figure 3.8). B mesons can decay semi-leptonically in three different ways: with a direct decay ($b \to \ell$, BR$\sim$10.7%), with a cascade decay ($b \to c \to \ell$, BR$\sim$8%), with a "wrong sign" cascade decay ($b \to \bar{c} \to \ell$, BR$\sim$1.6%). Thus these processes with two b decaying leptonically can lead to 4ℓ final state events. These backgrounds are called reducible as the experimental signature of leptons from b decay can be separated from that of leptons from W and Z decays as will be discussed in Section 6.3.

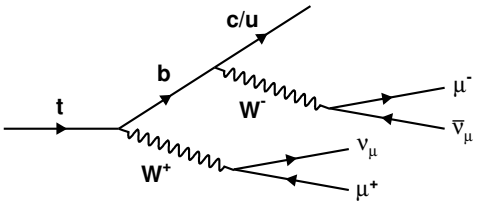


Figure 3.8: *Most probable decay chain with which a quark top can produce two opposite sign muons.*

The category of *instrumental* backgrounds is finally use to indicate background events with final state leptons from mis-identication of other particles, i.e. QCD multi-jets and Z/W+light jets processes where leptons mainly comes from jets faking leptons. More precisely this is the general case for electrons, while reconstructed muons in these processes, in addition to those from the Z and W decay, mainly comes from decay in fight of light primary hadrons, as will be explained in Section 4.2.

3.4 Monte Carlo Datasets and Cross Sections

SM Higgs boson signal samples, as well as samples for a large variety of electroweak and QCD-induced SM background processes, have been obtained using detailed Monte Carlo (MC) simulations. Generated events were processed through the full GEANT [57] detector simulation, trigger emulation and event reconstruction chain. A realistic misalignment scenario, based on the knowledge of positions of different elements of the muon and inner tracker systems at the time of the Monte Carlo production, was used as input for the reconstruction.
In this section, a detailed descriptions of the MC samples that have been used for the optimization of the event selection strategy prior to the experimental data that will be described in (Chapter 6) is presented. These sample are further used in the analysis for the comparisons with the measurements (Section 6.5), the evaluation of acceptance corrections and systematics (Section 7.1), and for the background evaluation procedure where measurements in a "background control" region are extrapolated to the "signal" region (Section 7.2).
The MC samples used for the specific performance studies on muon reconstruction and trigger will be described in Chapter 4.
Table 3.1 summarizes the Monte Carlo simulation datasets used for the $H\rightarrow 4\ell$ analysis. In CMS jargon these MC samples are part of the "Sum-

mer11" production. In all these samples a pile-up scenario is simulated; it is characterized by a flat distribution from 0 to 10 interactions per event with a tail above 10 to higher values corresponding to a poisson distribution with a mean of 10 interactions; this scenario has been designed in order to be able to perform a proper event re-weighting based on the real-data scenario.
All the signal and background cross sections are re-weighted to NLO. In the case of Higgs production via the gluon fusion mechanism, the most recent NNLO+NNLL calculations of the cross sections are included [46].
The general multi-purpose Monte Carlo event generator PYTHIA [58] is used for several processes including QCD multijet production. It serves either to generate a given hard process at leading order (LO), or, in cases where the hard processes are generated at higher orders, only for the showering, hadronization, decays, and for adding the underlying event. This is the case for the MadGraph (MadEvent) Monte Carlo [59] event generators which are used to generate multi-parton amplitudes and events for some important background processes. This is also the case for the POWHEG NLO generator [60] which is used for the Higgs boson signal and for the ZZ and $t\bar{t}$ background. For the latter the $t\bar{t}$ decays are handled, exceptionally, within POWHEG. Finally, this is also the case for the dedicated tool GG2ZZ [61] used to generate the $gg \to ZZ$ contribution to the ZZ cross section. For the underlying event, the so-called "PYTHIA tune Z2" which relies on p_T-ordered showers is used. For the parton density functions in the colliding protons the CTEQ6M PDF set is used.
More details on the event generators and the background samples used in this analysis, as well as indications on the procedures used to re-weight the MC events where needed, are provided in the following sections, where ℓ is to be intended as e, μ or τ.

Signal: $H \to ZZ^{(*)} \to 4\ell$

The Higgs boson samples used in the current analysis are generated with POWHEG [60] which incorporates NLO gluon fusion and vector boson fusion. The Higgs boson widths are taken from Ref. [46]. Additional samples with WH, ZH and $t\bar{t}H$ associated production are produced with PYTHIA. The Higgs boson is forced to decay to two Z-bosons, which are allowed to be off-shell, and both Z-bosons are forced to decay via $Z \to 2\ell$. Generator level events are re-weighted according to the total cross section $\sigma(pp \to H)$ which comprises the gluon fusion contribution up to NNLO and NNLL taken from Ref. [62, 63, 46, 64, 65, 66, 67, 68, 69, 70] and the weak-boson fusion contribution at NNLO computed in Ref. [46, 71, 72, 73, 74, 75]. The total cross section is scaled by the $BR(H \to 4\ell)$ [46, 76, 77, 78, 79]. Figure 3.9 shows the $H \to 4\ell$ cross-section as a function of the Higgs mass m_H for $\sqrt{s} = 7$ TeV.

A total of 27 Monte Carlo samples with different mass hypothesis were produced in the range 115 to 600 GeV/c^2.

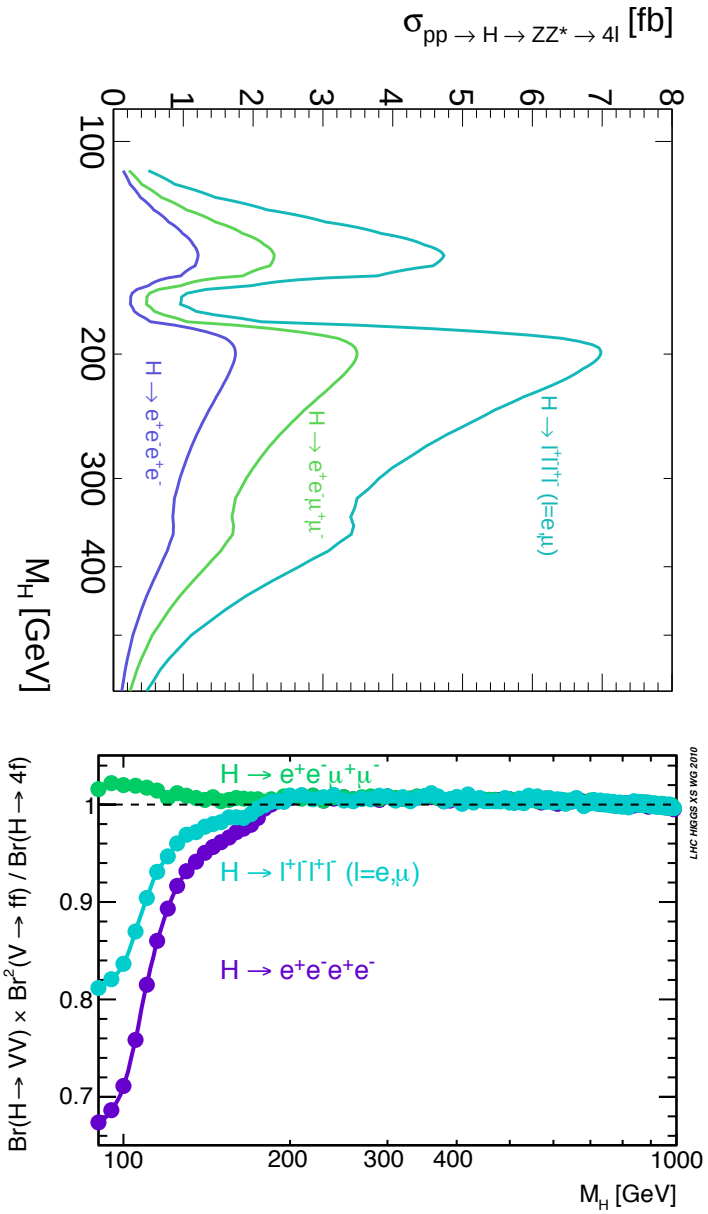


Figure 3.9: *Cross-section for SM Higgs in* $H \to 4\ell$, $H \to 2e2mu$ *and* $H \to 4e$ *(or* 4μ*) as a function of* m_H *in pp collisions at* $\sqrt{s} = 7$ *TeV (top). Cross-section enhancement due to the interference of amplitudes with permutations of identical leptons originating from different Z-bosons, as a function of* m_H *(bottom) [5].*

In comparison to $\sigma(pp \to H) \cdot BR(H \to ZZ^{(*)} \to 2e2\mu)$, the 4μ and $4e$

channel cross-sections are enhanced in the case of off-shell Z boson due to an interference of amplitudes with permutations of identical leptons originating from different Z-bosons, as shown in Figure 3.9 (bottom). This is correctly taken into account by Prophecy4f [46, 76, 77].
The POWHEG MC program used to simulate the $gg \to H$ process results in a Higgs Boson p_T spectrum that differs significantly from the best theoretical calculation which is available at NNLL+NLO. A theoretical estimate of this p_T spectrum is computed using the HqT [80] program, which implements such NNLL+NLO calculation. A re-weighting procedure has been studied to be applied to the simulated events (see Section 7.1). But the effect is very small for this analysis in which no direct constraints are imposed on the transverse momentum of the 4ℓ system, or on the hadronic recoil against this system (e.g. no jet veto or missing transverse momentum cut).
In the analysis presented in this thesis, signal efficiencies and mass shape distributions are determined on gluon fusion simulated samples, and assumed to be the same for the other Higgs production mechanisms.

Background: $q\bar{q} \to ZZ^{(*)} \to 4\ell$

For the current analysis we use the samples $q\bar{q} \to ZZ^{(*)} \to 4l$ produced with POWHEG, than include the complete NLO simulation, interfaced to PYTHIA for showering, hadronization, decays and the underlying event. The samples are divided in $4e(\mu)(\tau)$, $2e2\mu$, $2e(2\mu)2\tau$ final states.

Background: $gg \to ZZ^{(*)} \to 4\ell$

A full NNLO calculation for the ZZ production which would also take these gluon-induced diagrams into account is not available. Therefore the contributions are estimated by using the dedicated tool gg2ZZ [61], which computes the $gg \to ZZ$ at LO, which is of order α_s^2, compared to α_s^0 for the LO $q\bar{q} \to ZZ$. The hard scattering $gg \to ZZ^{(*)} \to 4\ell$ events are then showered and hadronized using PYTHIA.
The gg2ZZ tools provide the functionality to compute the cross-section after applying a cut on the mimimally generated invariant mass of the same-flavour lepton pairs (which can be interpreted as the Z/γ invariant mass) $m_{\ell\ell}^{\min} = 10\,\mathrm{GeV/c^2}$. This number is computed by using the LO PDF set CTEQ6L1, and the central renormalization and factorization scales $\mu_R = \mu_F = m_Z$, where $m_Z = 91.188\,\mathrm{GeV/c^2}$ is the nominal Z-boson mass. To estimate the accuracy of this number the renormalization and factorization scales were varied in the range $\mu \in [\mu_0/2, 2\,\mu_0]$; therefore an error of $^{+28\,\%}_{-20\,\%}$ is computed. The large uncertainty is expected, since the calculation is only LO, and only at NLO the scale dependencies start to be reduced. It is thus very hard to estimate the accuracy of the convergence

of the perturbative series (which contains only the first coefficient here), thus an uncertainty of $\pm 50\,\%$ on this number is assumed.
The gg2ZZ generator gives the contribution for final states with unlike flavours of the lepton pairs, but it was also used to estimate the like-flavour background. This is an approximation which is only strictly valid when $m_{4\ell} \geq 2\,m_Z$. Below this threshold the relative amount of like-flavour events increases compared to unlike-flavour events.
The total cross section is 3.48 fb for events with a given different flavour lepton pairs in the final state. The differential cross-section for $gg \to ZZ^{(*)}$ as a function of the four lepton invariant mass for different flavour lepton pairs was provided in Ref [81].

Background: Z+jets$\to 2\ell$+jets

Z+jets$\to$ 2ℓ+jets samples was generated with MadGraph, with a statistics of $\approx$ 40M events representing an equivalent integrated luminosity above $\mathcal{O}(10)$ fb^{-1}. This sample correspond to an inclusive Z+jets sample, where at matrix element (ME) level the generated process is pp$\to$ Z +N_p (p=partons) with N_p=0,1,2,3,4.
Both light ($q = d, u, s$) and heavy-flavor ($q = c, b$) jets are included in the sample. A generation cut on two-lepton invariant mass of $m_{2\ell} > 50$ GeV/c^2 is imposed in the simulation. A total NNLO cross section of 3048 pb is used; as the LO cross section for this sample is 2321 pb. Therefore a K factor of 3048 / 2321 = 1.31 when using this sample has been applied.

In this work, to separate the contribution from heavy-flavor jets from light-flavor jets, (from now on referred to as the Z$b\bar{b}/c\bar{c}$ and Z+light jets samples) the Madgraph Z+jets sample was partitioned using a filter selecting events with two b or two c quarks produced in the final state. The b and c quarks can be produced in the hard-scattering (Madgraph ME) or in the evolution of the showering (PYTHIA) via gluon splitting.
The filter efficiency is 6% for b quarks and 28% for c quarks. Studies on the comparison between $Z + N_p$ ME processes with $b(c)$ produced in the showering and $Z + bb(cc) + N_p$ ME processes have shown that the expected cross section is 2(4) time larger in the first case.

Process	**MC generator**	$\sigma_{(N)NLO}$
Higgs boson $H \to ZZ \to 4\ell$		
$gg \to H$	POWHEG	[1-20] fb
$VV \to H$	POWHEG	[0.2-2] fb
WH; ZH; $t\bar{t}H$	PYTHIA	[0.01-0.05] fb
ZZ continuum		
$q\bar{q} \to ZZ \to 4e(4\mu)(4\tau)$	POWHEG	15.34 fb
$q\bar{q} \to ZZ \to 2e2\mu$	POWHEG	30.68 fb
$q\bar{q} \to ZZ \to 2e(2\mu)2\tau$	POWHEG	30.68 fb
$gg \to ZZ \to 2\ell 2\ell'$	gg2ZZ	3.48 fb
$gg \to ZZ \to 4\ell$	gg2ZZ	1.74 fb
Other di-bosons		
$WW \to 2\ell 2\nu$	PYTHIA	4.88 pb
$WZ \to 3\ell\nu$	PYTHIA	0.595 pb
$t\bar{t}$ and single t		
$t\bar{t} \to \ell^+\ell^-\nu\bar{\nu}b\bar{b}$	POWHEG	17.32 pb
t (s-channel)	POWHEG	3.19 pb
$\bar{t}$ (s-channel)	POWHEG	1.44 pb
t (t-channel)	POWHEG	41.92 pb
$\bar{t}$ (t-channel)	POWHEG	22.65 pb
$\bar{t}$ (tW-channel)	POWHEG	7.87 pb
Z/W + jets ($q = d, u, s, c, b$)		
W + jets	MadGraph	31314 pb
Z + jets	MadGraph	3048 pb
QCD inclusive multi-jets, binned $\hat{p}_T^{\min}$		
$b, c \to e + X$	PYTHIA	
EM-enriched	PYTHIA	
MU-enriched	PYTHIA	

Table 3.1: *Monte Carlo simulation datasets used for the signal and background processes; Z stands for Z, Z^*, γ^*; ℓ means e, μ or τ; V stands for W and Z; $\hat{p}_T$ is the transverse momentum for $2 \to 2$ hard processes in the rest frame of the hard interaction.*

3.5 Experimental Datasets and Integrated Luminosity

The data samples used in the analysis that will be presented in this work are those recorded by the CMS experiment during 2010 for the run range from 136033 to 149442, during period "A" of 2011 for the run range from 160329 to 173692, and during period "B" of 2011 for the run range from 175860 to 180252.
A detailed description of the main changes in data-taking conditions from January to October of this year when the 2011 run was terminated, are presented in Section 4.5.1.
The CMS standard selection of runs and luminosity sections is applied, which requires high quality data with a good functioning of the different sub-detectors. Thus, similar detector operation conditions are imposed for the validation of the data to be used for the analysis of the $4e$, 4μ and $2e2\mu$ channels.
Of the total integrated luminosity $\mathcal{L}$ of 4711 pb^{-1}, a sample corresponding to $\mathcal{L}$=36 pb^{-1} was collected in 2010, $\mathcal{L} = 2195$ pb^{-1} in the period 2011A and $\mathcal{L} = 2516$ pb^{-1} in the period 2011B. The absolute pp luminosity is known with a precision of 4.5%.
The analysis relies on primary datasets (PDs) produced by the CMS central data-reconstruction; each primary dataset is characterized by the list of trigger paths that contribute to the final OR of decision that determines in which of the possible primary datasets an event has to be stored. The detailed content of the PDs evolves in phase with the evolution of the trigger menu to cope with ever increasing instantaneous luminosity. For the 2010 data, the analysis relies on the so-called "EG" and "MU" PDs for the data taking with instantaneous luminosities L in the range $10^{29} - 10^{31}$ cm$^{-2}s^{-1}$, and on "Electron" and "MU" PDs for $L > 10^{31}$ cm$^{-2}s^{-1}$. For the 2011 data, the analysis relies on the so-called "DoubleElectron" and "DoubleMuon" PDs. These latter PDs are formed by a "OR" between the available *DoubleMuon* (Section 4.4) and *DoubleElectron* (Section 5.4) triggers in the HLT Menu.

A "skimming" common to all 4ℓ channels was applied on PDs for data reduction. It requires:

- at least 2 reconstructed lepton candidates, either an electron basic track-supercluster object (Section 5.1) or a Global Muon object, or a Tracker Muon (Section **??**) object;
- $p_{T,1} > 20$ GeV/c, $p_{T,2} > 10\,(7)$ GeV/c for electron (muon) objects;
- an invariant mass $M_{1,2} > 40$ GeV/c^2.

The skimming requirements on a pair of leptons are less stringent than the requirements imposed for at least one pair of leptons in the H$\rightarrow 4\ell$

analysis, so that the skim should not entail additional efficiency losses for signal events.

Chapter 4

Physics Object: Muons

The reconstruction of the SM Higgs boson in the decay chain $H \to 4\ell$ imposes high-performance lepton reconstruction, identification and isolation as well as excellent lepton energy-momentum measurements. The identification of isolated leptons emerging from the event primary vertex allows for a drastic reduction of QCD-induced sources of misidentified leptons or non-prompt leptons coming from hadron decays. The precise energy-momentum measurements translate in a precision Higgs boson mass measurement $m_{4\ell}$, the most discriminating observable for the Higgs boson search. With four leptons in the final state, and in view of the modest fraction of the total production cross-section observable in the 4ℓ channels, a very high lepton reconstruction efficiency is mandatory. For Higgs bosons with masses $m_H < 2m_Z$, one lepton pair at least couples to an off-shell Z^* boson. The softest lepton in that pair typically has $p_T^\ell < 10$ GeV/c for masses $m_H < 140$ GeV/c^2. Preserving the highest possible reconstruction efficiency while ensuring sufficient discrimination against hadronic jets faking leptons (mainly for the electron case) and against leptons from decay-in-flight of light hadrons (mainly for the muon case) is especially challenging for the reconstruction of leptons at very low p_T^ℓ.

The robust detection of muons over the full acceptance of the CMS detector in a condition of very high rate and background is obtained with a muon system that allows an efficient and pure identification of muons and with the inner tracker that provides a very precise measurement of their properties. Excellent muon momentum resolution and trigger capability are enabled by the high-field solenoidal magnet and its flux return yoke. The latter also serves as a hadron absorber to facilitate the identification of muons.
In this chapter the description of the muon reconstruction algorithms (Section 4.1), of the muon quality selection requirements (Section 4.2), of the muon isolation criteria (Section 4.3), of the muon trigger logic (Section 4.4)

are presented.
The study of their performances on data and simulation was a preliminary part of the work presented in this thesis; the methods used for the measurements and the results obtained with 2011 collision data are reported (Section 4.5). The performance of the reconstruction and trigger algorithms is first studied on single-muon Monte Carlo samples generated at different transverse momentum values. These samples are produced with "ideal conditions", i.e. using the CMS design geometry and considering perfect detector alignment and calibration. Muons have a flat pseudorapidity (η) distribution, between -2.5 and 2.5.
The expected and real performances of the reconstruction, identification, isolation and trigger are then compared using the data and simulated samples described Section 4.5.

4.1 Muon Reconstruction

In the CMS standard muon reconstruction the high-level muon physics objects are reconstructed in a multi-faceted way, with the final collection being composed of three different muon types, i.e. *Stand-alone*, *Global* and *Tracker* muons. The reconstruction in the muon spectrometer starts with the reconstruction of hit positions in the DT, CSC and RPC subsystems. Hits within each DT and CSC chamber are then matched to form ÒsegmentsÓ (track stubs). The segments are collected and matched to generate seeds that are used as a starting point for the actual track fit of DT, CSC and RPC hits. The result is a reconstructed track in the muon spectrometer, and is called ÒStand-alone MuonÓ. Stand-alone muon tracks are then matched with tracker tracks to generate ÒGlobal MuonÓ tracks, featuring the full CMS resolution. ÒTracker MuonsÓ are muon objects reconstructed with an algorithm that starts from a silicon tracker track and looks for compatible segments in the muon chambers.
In the following, the most important features of the reconstructions algorithms are described; a more complete and detailed description is available at [82], [83].

4.1.1 Local Segments Reconstruction

Local reconstruction is the reconstruction of basic hits and segments in individual muon chambers, starting from the raw data from the detector read-out. The results are track segments in the DTs and CSCs and individual points in the RPCs.
Local reconstruction in the DTs [84] begins with the reconstruction of

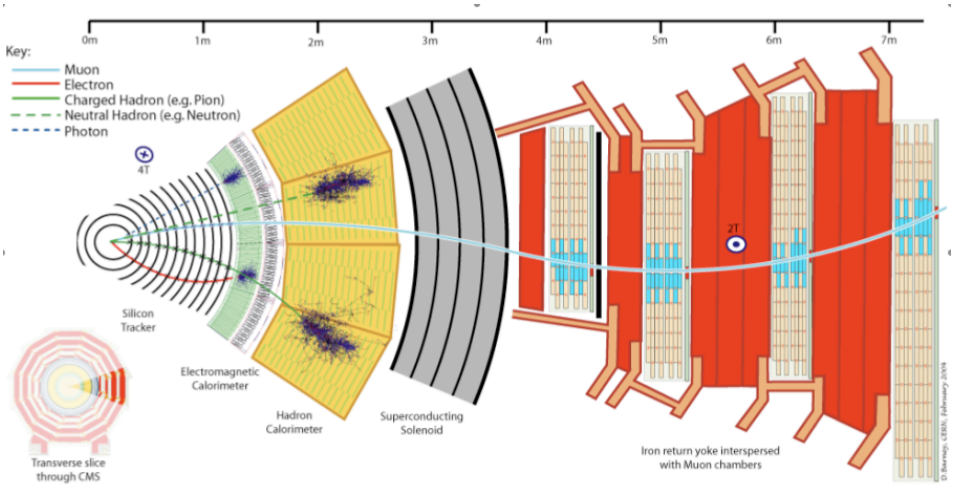


Figure 4.1: *Schematic representation of a reconstructed muon track crossing the CMS detector.*

mono-dimensional hits in individual drift cells. The only information contained in these hits is their distance from the anode, with an intrinsic left/right ambiguity and without any information about their position along the wire. The cell hits are the starting point for the reconstruction of segments in the r-ϕ and r-z projections separately. These two-dimensional segments still do not provide any information about the coordinate along the sense wires, but they allow the measurement of the track angle in the measurement plane (orthogonal to the wires). The direction and position of the muon crossing the chamber are obtained combining the two projections. The resulting three-dimensional segments have an angular resolution of about 0.7 mrad in ϕ and about 6 mrad in θ with a position resolution up to 200 μm [85].
Each CSC plane measures a point in two dimensions. The radial coordinate r is measured by the wires, the azimuthal coordinate ϕ by the strips. To obtain a precise measurement, the charge distribution of a cluster of three neighbouring strips is assigned a position according to abulated positions values that are pre-determined by fits of charge distributions. The hits in a chamber are used to fit a three-dimensional straight line segment (made of up to six points). The position resolution of segments varies from about 50 μm in the first CSC station to about 250 μm in the fourth [36]. The directional resolution varies with the chamber type, with an average of about 40-50 mrad in ϕ, slightly worse in θ.
The RPCs are characterized by an excellent time resolution, of the order of few nanoseconds, while their spatial resolution is limited by the strip pitch. In each chamber, the two coordinates (ϕ and z in the barrel, ϕ and r in the endcaps) and their uncertainties are obtained by clusterising

the strips which fired and computing their centroid, assuming a uniform charge distribution on the area of the strips. The resolution on ϕ is around 1 cm, while the orthogonal coordinate is only constrained by the strip length [86].

4.1.2 Stand-Alone Muon Reconstruction

Seeding. The reconstruction of a track in the muon spectrometer starts from an initial state, called *seed*, estimated from DT and CSC segments. The seeding algorithm searches for a pattern of segments in the stations using rough geometrical criteria and produces a set of initial states which are the starting point for the reconstruction of muon tracks. The p_T of the seed candidate is estimated with a parametrization based on either the $\Delta\phi$ that can be the bending angle of the segment with respect to the vertex direction (barrel region) or on the difference in the ϕ coordinate between two of the segments.

Pattern Recognition and Fit. The track is then extended using the Kalman filter technique [87]. In the CMS implementation of this recursive algorithm, the pattern recognition is performed layer by layer, while the trajectory parameters are updated. Once all hits have been collected, a final fitting step (*smoothing*) can be applied, updating the trajectory state at the location of all intermediate hits with the information from all the collected measurements, thus obtaining the optimal track parameters.
The algorithm is flexible enough to allow for different possible strategies: the filter can be applied in either direction, from the innermost layer towards the outermost or viceversa (*forward* or *backward*). In the current configuration the forward and backward filter are consecutively applied to remove a possible bias from the initial seed.
In the forward filter the parameters of the seed state are propagated to the innermost compatible muon chamber. Here the most compatible measurement is searched for on a χ^2 basis, estimating the *incremental* χ^2 ($\Delta\chi^2$) given by the inclusion in the fit of a given track segment (or individual hit, in the case of RPCs). If this measurement it is considered compatible with the track (according to a $\Delta\chi^2$ threshold) the track parameters are updated and the track is propagated to the next reachable chamber; the same procedure is repeated until no more chambers are reachable. The forward filtering step is considered successful if compatible measurements are found in at least two muon chambers, at least one of them being a DT or CSC chamber. Otherwise the reconstruction algorithm is stopped and no track is produced.
The outcome of the forward filter can be affected by a significant bias from the initial seed state. For this reason, the track parameters obtained at the last update of the forward pattern recognition are used as input to a sec-

ond filtering step, this time navigating inward (backward filter). Finally, to further reduce the bias from the seed it is possible to *refit* the hits collected during the backward filter a configurable number of times, each time using the result of the previous fit as a starting state. In the current configuration the refit is applied once.

Trajectory cleaning. The trajectory building algorithm is run for each seed. If the seeding algorithm fails to merge all the track segments from the same muon, several seeds can be built from a single muon, giving rise to duplicates of the same tracks. These duplicates, called *ghosts*, usually share a fraction of their hits. In order to remove them, all the track candidates that share at least one hit are compared with each other and only the best candidate is kept, according to criteria based on χ^2, p_T and number of hits.

In order to improve the momentum resolution of tracks, the beam spot position is used to constrain the stand-alone track parameters. The resolution of a stand-alone track is expected to be between 10 and 15% up to $p_T = 100\ \mathrm{GeV}/c$ and $\sim 30\%$ at $1\ \mathrm{TeV}/c$. A small bias in q/p_T, coming from the seed initial state, is still present (1-3%), although reduced by about one order of magnitude with respect to the seed.

4.1.3 Global Muon Reconstruction

The reconstruction of Global Muon tracks begins after the completion of the reconstruction of the stand-alone tracks and the inner tracker tracks. Each stand-alone tracks is matched to a compatible tracker track and a fit of all the available measurements is performed.

Track Reconstruction in the Inner Tracker. As in the muon system, the first step of the track reconstruction [88] is finding a *seed*, which is the starting point for the pattern recognition. Seeds are built using two or three consecutive hits, in the pixel and/or in the strip detector. The pattern recognition is then performed layer by layer, with an iterative technique based on the Kalman filter, similarly to that used in the muon spectrometer alone. The algorithm ends with a final fit of the collected hits, followed by the suppression of ghost tracks.

Track Matching. The process of identifying the tracker track to be combined with a given stand-alone muon track is referred to as *track matching* and consists of two steps. The first step consists in the definition of a *region of interest* (ROI) in the track parameter space that roughly corresponds to the stand-alone muon track, and to select the subset of tracker tracks inside this ROI. The second step is to iterate over these tracks, apply-

ing more stringent spatial and momentum matching criteria to choose the best tracker track to combine with the stand-alone muon. The ROI is defined using the stand-alone track parameters and assuming that the muon comes from the interaction point. The matching is performed by comparing the parameters of the stand-alone track with those of all tracker tracks in the ROI. This is best done by propagating the tracks onto a common reference surface, e.g. the detector layer of the innermost stand-alone track hit taking into account the magnetic field, the average expected energy losses and the multiple scattering in the detector materials. In order to maximise the matching efficiency, several criteria are applied in sequence. First, the tracks with a χ^2 compatibility with the stand-alone track below a fixed threshold are chosen. If all tracks fail this cut, then the positions of the stand-alone and inner tracks on the reference plane are compared. If also this criterion fails and no pair is found within a fixed cut, the matching is attempted by comparing the track directions at the interaction point, with a very loose cut applied. If all criteria fail, the reconstruction is stopped and no global track is produced. The matching algorithm can select more than one tracker track for a given stand-alone. In this case, all matched tracks proceed in the reconstruction chain.

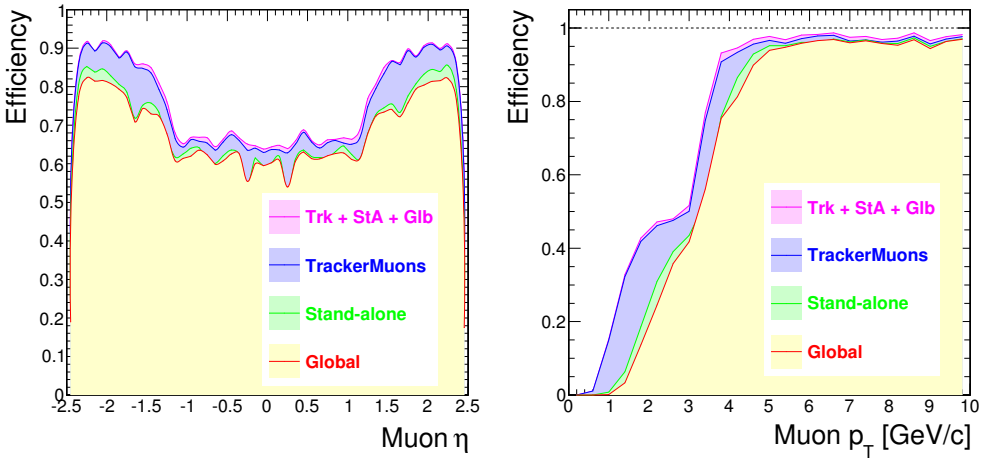
Global Fit. The last step consists in fitting a global track using all hits belonging to the matching tracker and stand-alone tracks. The global refit algorithm attempts a fit for each tracker-stand-alone track pair. If more than one global track is produced for a given stand-alone, the one with the best χ^2 is chosen. Thus, for each stand-alone muon there is a maximum of one global muon that will be reconstructed.
The reconstruction ends with the association of energy deposits in the calorimeters to the global tracks. The efficiency for stand-alone track reconstruction is mainly driven by detector acceptance discontinuities. A further inefficiency appears for global tracks with p_{T} below 5 GeV/c. At low p_{T} muons easily stop in the yoke without crossing all muon stations, especially in the barrel region, and stand-alone tracks are reconstructed with a relatively low number of hits and with poorer momentum resolution. This makes the matching with the tracker tracks more difficult and less efficient.
The resolution of global tracks is found to be between 1 and 2% up to p_{T} = 100 GeV/c and slightly above 6% at 1 TeV/c. At large transverse momenta ($p_{\mathrm{T}} > 200$ GeV/c), the global-muon fit improve the momentum resolution compared to the tracker track only. The q/p_{T} bias is of the order of few permil up to 1 TeV/c.

4.1.4 Tracker Muon Reconstruction

The global muon track reconstruction starts from the muon system and combines stand-alone muon tracks with tracks reconstructed in the inner tracker. However, a large fraction of muons with transverse momentum below ~ 6 GeV/c does not leave enough hits in the muon spectrometer to be reconstructed as stand-alone muons. Moreover, some muons can escape in the gap between the wheels. A complementary approach has therefore been designed to improve the muon reconstruction efficiency at low p_{T}.
In the tracker muon reconstruction approach all tracks with $p_{\mathrm{T}} > 0.5$ GeV/c and $p > 2.5$ GeV/c are considered as possible muon candidates and are extrapolated to the muon system (as described for the *t*rack matching in Section 4.1.3). For each crossed or nearly crossed chamber the algorithm stores the following variables:

- the distance between the propagated track and the nearest chamber edge, in both the chamber local x and y directions, with a conventional sign (negative inside the active volume, positive outside the active volume), and the corresponding 1σ uncertainty;
- the position (x, y) and slope $(dx/dz, dy/dz)$ of the extrapolated track in the local chamber coordinates and thecorresponding 1σ uncertainties;
- the segments in the chamber that are near to the propagated track ("associated segments").

Muons identified by at least one associated segment are called *tracker muons*. Since each track is treated individually, if two or more tracks are close to each other, it is possible that the same segment or set of segments is associated to more than one track, resulting in duplicate tracker muons. This ambiguity is resolved by the *arbitration* algorithm, which assigns segments to tracks by looking at the best Δx or $\Delta R = \sqrt{\Delta x^2 + \Delta y^2}$ match.
Different physics analyses can further select the tracker muons using the variables listed above, in order to balance the purity and efficiency of the muon identification.
Figure 4.2 shows the efficiency of the different muon reconstruction algorithms described so far, for transverse momenta below 10 GeV/c. We can clearly see that at low momentum (roughly $p_{\mathrm{T}} < 5$ GeV/c) the tracker muon reconstruction is more efficient then the global muon reconstruction, since it requires only a single segment in the muon system, whereas global muon reconstruction is designed to have high efficiency for muons penetrating through more then one muon station. The turn-on curve for tracker-muon efficiency in Figure 4.2b shows a step at 3 GeV/c, where the contribution of the muon barrel begins: in the central region ($|\eta| < 1.2$)



(a) **(b)**

Figure 4.2: *Muon reconstruction efficiencies (a) vs. η and (b) vs. p_T, for all the reconstruction algorithms described so far, for simulated muons with a flat η distribution between -2.5 and 2.5 and a flat p_T distribution between 0 and 10 GeV/c [83].*

muons with lower transverse momentum do not reach the spectrometer, because of the energy loss in the inner detectors and the bending in the strong magnetic field. The corresponding inefficiency is also visible in Figure 4.2a.
The plateau efficiency remains flat up to very high p_T values (~ 1 TeV/c) where it starts to decrease. CMS has developed specialized algorithms for high-p_T muon reconstruction and momentum assignment. As the muon passes through the iron of the magnet return yoke radiative processes can alter the muon's trajectory. The energy loss as well as the extra hits produced in the muon chambers from the resulting electromagnetic showers can effect the track fit so that the resulting estimate of the muon's momentum at the production vertex is significantly different from the true value. Therefore, several different strategies on how to include information from the muon system have been developed and studied using cosmic rays. The description of the dedicated high-p_T reconstruction algorithms can be found elsewhere [89]; in this thesis only the standard muon reconstruction has been studied and used.

In summary, the majority of muons (with sufficient momentum) from collisions are reconstructed either as a *Global Muon* or a *Tracker Muon*, and very often as both. However, if both approaches fail (this occurs only for about 1% of muons from collisions, thanks to the high tracker-track efficiency [90]) and only a standalone-muon track is found, the muon falls into

a third category of muon candidates, *Stand Alone Muons*. The muons in this category have worse momentum resolution and larger contamination from cosmic rays than the previous two muon categories, and are usually not used in physics analyses. The results of these three algorithms are merged into a single collection of muon candidates, each one containing information from the standalone, tracker, and global muon reconstruction, where available. Candidates found both by the Global Muon and the Tracker Muon approaches that share the same tracker track are merged into a single candidate. Similarly, standalone-muon tracks not included in a Global Muon are merged with Tracker Muons that share muon segments.
The final muon momentum is assigned based on the result of the global fit if both the global and tracker-only fit estimate the muon p_{T} to be above 200 $\sim$ GeV/c and if the results of the two fits for q/p agree to within two sigma of the tracker-only fit. Otherwise the estimate from the tracker-only fit is chosen.

4.2 Muon quality selection

The combination of different algorithms provides a robust and efficient muon reconstruction. A given physics analysis can achieve the desired balance between identification efficiency and purity by applying a selection based on various muon identification variables that can help in distinguish among the several sources of reconstructed muons. The low p_{T} range, (0-30 GeV/c), is mainly dominated by semileptonic decays of heavy-flavour hadrons; this contribution is accompanied by a high rate of muons coming from decay-in-flight of light primary hadrons (π and K) and of fake muons due to hadrons not fully contained in the calorimeter (punch-through).
The relative weight of these contributions depends the muon selection criteria, which in most physics analyses are tuned to reduce them in favor of prompt muons such as the decays of W and Z bosons or of promptly produced quarkonia states.
In order to reduce the punch-through contribution, selection variables related to the penetration depth into the muon system are used; decay-in-flight muons as well as mis-reconstructed tracks and secondary products form the interaction with the material are suppressed with the transverse and longitudinal impact parameter variables ($|d_{xy}|$, $|d_z|$) evaluated with respect to the primary vertex[1]; requirements on $|d_z|$ are also useful to reject products from other primary vertices due to pile-up interactions. Tight

[1]Due to pile-up, multiple vertices can be present in a single bunch crossing. The reconstructed vertex with the largest sum of of the square of the transverse momenta of the tracks in the track cluster is assumed to be the primary vertex of the hard scattering event.

$|d_{xy}|$, $|d_z|$ if too tight, can reject products of secondary vertices such as muons from b decays that are signal in some physics analysis.
In the following, three muon identification algorithms are presented; their performances have been measured on data and are described in Section 4.5.2.

- *Soft Muon selection:* This selection requires the candidate to be a Tracker Muon, with the additional requirement that a segment must be matched in both x and y coordinates with the extrapolated tracker track, so that the pull for local x and y is less than 3. Segments that form a better match with a different tracker track are not considered. These additional requirements are optimized for $p_T < 10$ GeV/c muons. This selection is presently used in quarkonia and B-physics analyses in CMS.

- *Tight Muon selection:* For this selection, the candidate must be reconstructed as a Global Muon, with the following additional requirements: normalized χ^2 of the global muon track fit less than 10; at least one muon chamber hit included in the global muon track fit; match to muon segments in at least two muon stations (this implies that the muon is also reconstructed inside-out as a Tracker Muon); its corresponding tracker track including more than 10 silicon tracker hits of which at least one pixel hit. Cuts on the transverse impact parameter $|d_{xy}| < 2$ mm and on the longitudinal impact parameter $|d_z| < 1$ mm with respect to the primary vertex are imposed. With this selection, the rate of muons from decays-in-flight is significantly reduced, at the price of a few percent loss in efficiency for prompt muons such as those from W and Z decays (Section 4.5.2). The muon selection presently used in CMS electroweak analyses is similar to the Tight Muon selection.

- *Global Muon selection:* For this selection, the candidate must be reconstructed as a Global Muon, with just the additional requirement that its corresponding tracker track must include more than 10 silicon tracker hits. Cuts on the transverse impact parameter $|d_{xy}| < 5$ mm and longitudinal impact parameter $|d_z| < 5$ mm with respect to the primary vertex are imposed. With this selection the high efficiency of the muon reconstruction is preserved while ensuring a good measurement of the muon momentum. This selection has been chosen for the H$\rightarrow 4\ell$ analysis: the very low signal rate imposes a requirement of the highest possible selection efficiency. Moreover the background from QCD is negligible due to the request of four isolated muons.
  As shown in Figure 4.2 the Global Muon reconstruction is almost reaching the efficiency plateau for muons with $p_T \sim 5$ GeV/c; this is the minimum p_T cut applied on the muons in the H$\rightarrow 4\ell$ analysis.

4.3 Muon Isolation

The requirement that a muon is an isolated particle in the event, meaning that the energy flow in its vicinity is below a certain threshold, can effectively help discriminating muons from decays of $W^{\pm}$ and Z^0 bosons from muons produced as a result of QCD processes. Two different isolation criteria are here considered:

- Tracker relative isolation (R_{Iso}^{Tk}). The isolation variable is defined in this case as the scalar sum of the p_T of all tracks reconstructed in the inner tracker whose direction has a distance from the muon track direction
  $\Delta R \equiv \sqrt{(\Delta\varphi)^2 + (\Delta\eta)^2} < 0.3$. The p_T of the muon track itself is not included in the sum. For the muon to be considered isolated, the ratio of the p_T sum to the muon track p_T is required to be below a certain threshold. Track directions and values of p_T are computed at the point of closest approach to the nominal center of the detector.

- Tracker plus calorimeters (combined) relative isolation (R_{Iso}). The discriminating variable is similar to R_{Iso}^{Tk}, but the numerator of the ratio also includes the sum of energies measured in ECAL and HCAL towers found within a cone of radius $\Delta R < 0.3$ centered on the muon track direction. Deposits in towers around the muon crossing point in the calorimeters are vetoed.

The selection criteria for tracker tracks and energy deposits in ECAL and HCAL to be used in isolation cones and veto regions are specified in Table 5.2.

Table 4.1: *List of parameters for muon isolation. ΔR_v is the radius of the veto cone, p_T is the transverse momentum of the tracks in the cone, E is the energy deposited in each ECAL Tower within the cone, ΔZ, Δr are the minimum distances of a track to the cone apex in the longitudinal and in the radial direction, respectively.*

Type	ΔR	Deposits	Veto region	Thresholds
Tracker	0.3	CTF tracks	$\Delta R_v < 0.015$	$p_T > 1.0$ GeV/c, $\Delta Z < 0.2$ cm, $\Delta r < 0.1$ cm
ECAL	0.3	Towers	$\Delta R_v < 0.07$	$E > 0.25$ GeV
HCAL	0.3	Towers	$\Delta R_v < 0.1$	

Each of these algorithms has features that can suit the requirements of different analyses. For example, the tracker plus calorimeters relative isolation with a threshold of 0.15 is the algorithm currently chosen for the measurement of the $W^{\pm}$ and Z^0 cross sections [91]. Searches for heavy resonances decaying into muon pairs have instead adopted the tracker

relative isolation given that high-energy muons are expected to radiate a significant amount of energy in the calorimeters. For completeness other analyses are using the a third algorithms based on the Particle-flow approach whose description it is beyond the scope of this work.
As described in Section 6.3 the R_{iso} variable is used for the H$\rightarrow 4\ell$ analysis with a cut on the sum of this variable for the 2 least isolated leptons.

Isolation variables are the most pile-up sensitive variables in the H$\rightarrow 4\ell$ analysis. Pile-up causes the mean energy deposited in the detector to increase, leading to a rise of the mean isolation values. Thus, the efficiency of a cut on isolation variables depends on pile-up conditions. The effect is observed to be large in calorimeters (ECAL, HCAL) and quite feeble in the tracking system, mostly due to the requirement that the tracks contributing to the isolation cone originate from a common vertex.
Therefore, in order to have a pile-up robust analysis, the isolation variable has to be corrected. In the context of H$\rightarrow 4\ell$ analysis among several correction methods, the one using `FastJet` [92, 93] energy density (ρ) in the event has been chosen to estimate the mean pile-up contribution within the isolation cone of a lepton. A ρ variable is defined for each jet in a given event and the median of the ρ distribution for each event is taken. The correction to the isolation variable is then applied according to the formula

$$\Sigma \text{ Iso}_{\text{corrected}} = \Sigma \text{ Iso} - \rho \cdot A \tag{4.1}$$

where A is the area of the cone in the (η,ϕ) space. It has the dimension of an angle, since ρ is given in $1/(\Delta\eta\Delta\phi)$ units. Rather than computing a geometrical area, an effective area is considered, to avoid dealing with different thresholds in the isolation and `FastJet` algorithms. The effective area is defined as the ratio of the slope obtained with a linear fit of Σ Iso(N_{vtx}) to the one from a linear fit of $\rho(N_{vtx})$.
The performance of the ρ corrected R_{Iso} variable are presented in Section 4.5.3.

4.4 Muon High Level Trigger

The logic structure of a HLT muon trigger path and the HLT muon reconstruction strategy are presented in Section 4.4.1 and Section 4.4.2 respectively. In the context of the work presented in this thesis, studies to optimize the trigger algorithms for efficiency and characterize their performace on simulations are described in Section 4.4.3 after a brief description of the algorithms themselves.

4.4.1 Muon Trigger Path

Each trigger path (Section 2.3.2) is a chain consisting of two different types of modules: modules of the HLT reconstruction chain and modules of the HLT filter logic, where the trigger requirements are applied.
The reconstruction chain is common to all trigger paths containing muons. The CMS software framework guarantees that only one instance of each reconstruction producer is run, independently of the number of muon trigger paths present in the HLT. Even if only a given subset of muon candidate may be relevant for a particular trigger path, the reconstruction chain will process all muon candidates (up to 4) selected by the Level-1 Global Muon Trigger (cf. Section 2.3.1). This approach enormously simplifies the trigger logic and facilitates the addition of new trigger paths.
The HLT muon reconstruction starts only if the L1 trigger bit decision for at least one of the paths containing muons is positive. The reconstruction chain continues if at least one of the trigger paths with muons requires it. Conversely, it stops at a given point if all muon trigger paths do not require the subsequent reconstruction steps. The usual case is a negative decision for all trigger paths if no muon passes some thresholds in the initial selection steps.
In particular, this means that a muon candidate failing to pass the L1 filter of a given path may still be reconstructed at L2 and L3, if the event passes to the next trigger levels because of another trigger requirement; such a candidate ("volunteer") is however not considered for the trigger path which failed at L1.

The main steps of the HLT muon reconstruction are:

- L2 Seeding: the parameters of the L1 muon candidates are converted into seeds for the L2 reconstruction;
- L2 Reconstruction: using muon system information, the stand-alone reconstruction is performed;
- L2 Isolation: a separate module computes the isolation of each muon candidate using calorimeter information;

- L3 Reconstruction: the tracker reconstruction is performed in the silicon tracker and with a global fit the reconstructed track is merged with the L2-track; to keep execution time low, the tracker reconstruction is regional (i.e. the pattern recognition and track fitting are performed only in a small slice of the tracker);
- L3 Isolation: the isolation for the L3-candidate is computed using information from nearby tracks reconstructed only with pixel hits.

Filters modules can be interleaved between the previous pieces of reconstruction. These filters are specific for each trigger path. Their role is to stop the trigger sequence if muon candidates do not satisfy some given conditions. In addition, they provide as output a reduced list of candidates that should be considered in subsequent filter steps. For instance the filters used in particular single and double muon trigger paths are described in table Table 4.2.

Table 4.2: *Thresholds and requirements for the basic muon triggers. The absolute value of the impact parameter, $|d0|$, is calculated with respect to the beam spot. The track p_T sum is computed over pixel tracks only. q is the L1 quality bit.*

Bit Name	L1	L2	L2 Iso	L3	L3 Iso
HLT_Mu5	$p_T \geq 3$ $q > 3$	$p_T \geq 4$	N/A	$p_T \geq 5$ $\|d0\| \leq 2$	N/A
HLT_IsoMu9	$p_T \geq 7$ $q > 3$	$p_T \geq 7$	CaloIso≤ 4	$p_T \geq 9$ $\|d0\| \leq 2$	PixelIso≤ 1
HLT_DoubleMu7	$p_T \geq 5$ $q > 3$	$p_T \geq 5$	N/A	$p_T \geq 7$ $\|d0\| \leq 2$	N/A

4.4.2 Muon reconstruction in the HLT

In this section muon reconstruction is described in the context of the HLT, stressing in particular the differences with the off-line reconstruction. For a more detailed description see [82].

Level-2 Muon Seeding

The L2 muon reconstruction starts from an initial seed state. Unlike the off-line case, where seeds are obtained combining segments in the muon system, in the on-line reconstruction the L1 muon particle are used as *external seeds*, with a significant reduction of computing time. The full muon reconstruction is then performed on a regional basis, only where a L1 candidate is found. For each L1 muon candidate promoted by the

GMT to the HLT, a L2 seed is built. An initial state is created from the position and momentum of the L1 candidate and fixed errors are assigned to all the parameters. Finally, the seed state is propagated from the second muon station (Section 2.3.1) to the innermost compatible muon chamber.

Level-2 Muon Trajectory Building

Starting from each L2 seed, the reconstruction of L2 tracks proceeds exactly as in the off-line case, described in Section 4.1.2. The local reconstruction and trajectory building are performed only in those regions of the muon system where a L1 particle/L2 seed was built, in order to comply with the time requirements of the trigger. Once the trajectories are built, ghost suppression is applied and a beam spot constraint is imposed. Each trajectory preserves a link to its own seed, hence to the corresponding L1 particle. This is particularly important to allow the suppression of volunteers.
A notable difference from the off-line reconstruction is present in the application of ghost suppression. This peculiar behaviour was expressly introduced to address a specific problem concerning volunteers, arising at cleaning level. A single muon can generate more than one L1 particle ("ghosts"), e.g. if the GMT fails to merge candidates from different subdetectors. In this case, several tracks with shared hits will be produced, and the cleaner is designed to keep only one. Suppose e.g. that two L1 candidates are ghosts of the same muon and only one of them passes a certain trigger filter. If the track selected by the cleaner is the one seeded by the failing L1 object, this track is taken as a volunteer and rejected by the L2 filter; this results in an artificial inefficiency that has been found in the context of the present study and cured before before the starting of the data taking. In the current configuration, the cleaner produces a map of links between each L2 track and the seeds of all its duplicates. In such a way, the L2 filter can accept a L2 track if any of the corresponding L1 candidates has passed the L1 filter. The corresponding improvement in the efficiency is shown in Figure 4.4.

Level-2 Isolation

The same algorithm that is used off-line to calculate the calorimeter isolation is used to apply isolation requirements at L2. The calorimeter towers are used as input to the isolation algorithm. The tower energies above thresholds are summed in a cone of $\Delta R < 0.24$ centered on the direction of the L2 muon track taken at the interaction region. To remove the contribution of the muon the sum of energy in a veto cone is subtracted. The veto cone has a ΔR of 0.07 in ECAL (0.1 in HCAL) and is centered around the state of the L2 muon propagated to the inner ECAL surface. In the

sum, the ECAL and HCAL deposits are weighted with factors reported in Table 4.3. The table also summarizes the thresholds applied to individual deposits.
A L2 Muon is tagged as isolated if the weighted sum is lower than an η-dependent threshold that varies between 2 and 4 GeV/c^2.

Table 4.3: *Configurable parameters for L2 calorimeter deposit calculation.*

Parameter name	Value	Parameter description
`Threshold_E`	0.2	Threshold for Ecal tower
`Threshold_H`	0.5	Threshold for Hcal tower
`Weight_E`	1.5	Weight for summing the Ecal energy
`Weight_H`	1.0	Weight for summing the Hcal energy
`DR_Max`	0.24	cone for deposit calculation
`DR_Veto_E`	0.07	veto cone in Ecal
`DR_Veto_H`	0.1	veto cone in Hcal

Level-3 Muon Seeding

After the completion of Level-2 muon reconstruction and Level-2 filtering steps, the algorithm proceeds to reconstruct Level-3 muon candidates. In the HLT environment, the full tracker reconstruction cannot be performed because it is too CPU intensive. Therefore, track reconstruction is done in small regions of the central tracker corresponding to likely muon candidates. This restrictive reconstruction is accomplished by a proper selection of trajectory seeds which limits the track reconstruction to a region consistent with the L2 muon, i.e. the ROI described in Section 4.1.3.
In the following paragraphs, different L3 seeding algorithms are described. The *hit-based seeds* use combinations of hits found on the tracker layers to estimate the initial position and direction. *State-based seeds*, instead, use a trajectory state, without information from the tracker measurements, to find the initial position and direction.

Hit-Based Seeding. The inside-out hit-based seed option (*IOHit*) selects pairs or triplets of hits in the pixel detector, as in the off-line seeding. The inner hit is required to be in the ROI. In the case of L2 muons with $|\eta| > 2$, the hit-pairs can be a combination of pixel and strip layers. In the outside-in hit-based seed option (*OIHit*), the L2 trajectory state is propagated to the outer tracker bound. From here, compatible measurements are looked for in TOB and TEC layers and used to update the predicted state from L2.

State-Based Seeding. The track parameters obtained from the L2 mea-

surement are propagated to the first compatible layer of the pixel detector (inside-out state-based seed or *IOState*) or to the outermost layer of TOB and TEC (outside-in state-based seed or *OIState*). The state given at the innermost (or outermost) layer of the tracker is used to proceed with pattern recognition from the inside-out (or outside-in).

Cascade algorithm. Each of the L3 seeding algorithms described above performs differently in different parts of the detector and has advantages and disadvantages. For this reason, a combined seeding sequence is used in order to benefit from the advantages of each one, while minimising the disadvantages. Three of the four algorithms, the OIState, OIHit and IOHit, are run in a sequence, starting with the fastest (OIState) and finishing with the slowest (IOHit). In order to save CPU time, the slower algorithms are never called if the faster algorithms have a favourable outcome: if a L3 muon is successfully reconstructed from the seed, then the sequence for that L2 muon is stopped; otherwise, the sequence continues to the next seed generator and L3 reconstruction module. The improvement in efficiency obtained with this algorithm will be shown in Figure 4.6. The *OIState* algorithm has been chosen for the beginning of the data taking (2011). The more sophisticated cascade algorithm was adopted as the default after validation on 2011 data, which was carried on in the context of the work presented in this thesis.

Level-3 Muon Trajectory Building

Once the L3 seeding is completed, tracks are reconstructed in the tracker with the same procedure described in Section 4.1.3, only inside the ROI determined by each L2 muon. The best tracker tracks to be combined with a given L2 muon are then selected, following the same matching criteria as in the off-line reconstruction. Finally, for each L2-tracker match, a global fit is performed, using the whole set of hits in the tracker and muon system. If more than one global L3 track is built from the same L2 muon, only the one with the best χ^2 is kept. Thus, for each L2 muon, there is a maximum of one global L3 muon that is reconstructed.

Level-3 Isolation

The p_{T} of the selected pixel tracks is summed in a cone of ΔR less than 0.24 centered on the L3 muon track projected to a cylinder of radius 6 cm to be less sensitive to mis-alignment. A veto cone of ΔR 0.01 is opened around the leading p_{T} track providing that its p_{T} is greater than 2 GeV and its ΔR with respect to the L3 track is less than 0.025, otherwise the veto cone is centered on the L3 track.

The selection applied to the scalar p_T sum described above is eta dependent with thresholds that vary from 0.8 to 1.2 GeV/c^2.

4.4.3 Performance of the Muon HLT

The efficiency of each step of the muon trigger chain has been measured on simulation. The trigger efficiency is measured with respect to off-line reconstructed muons as the interest of any physics analysis is to know how efficient is the trigger in selecting muons that can be reconstructed offline.
Therefore, for trigger efficiency to be evaluated, trigger objects need to be matched to the muon reconstructed offline. L1 muon particle are matched to offline muons by position, extrapolating the muon's tracker track to the muon system (ME2, MB2) where the L1 position variables are reported (cf. Section 2.3.1). The ΔR cone for the matching is 0.5; the offline muon is matched to the L1 muon particles with the highest p_T within the cone. This procedure has been proved to be 99.9% efficient on simulated single muon events. The HLT muon objects (L2 and L3 muons) are matched to the offline reconstructed ones by direction at the vertex. The cone size for the L2 matching is 0.3, for L3 is 0.1; the matching efficiency is higher than 99.9%.

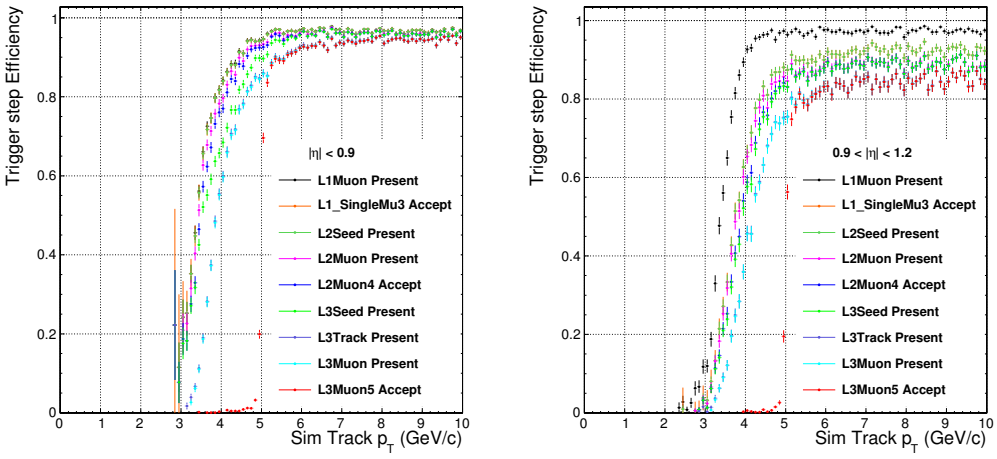


Figure 4.3: *Reconstruction and filtering efficiency for each step of the HLT_Mu5 trigger path for Global Muons with p_T (0-10) GeV/c and pseudorapidity regions: $|\eta| < 0.9$ (Barrel) (left) and $0.9 < |\eta| < 1.2$ (Overlap) (right).*

In Figure 4.3 and Figure 4.4 the efficiencies for each step of the HLT_Mu5 path are presented for low p_T (0-10 GeV/c) reconstructed Global Muons

in three pseudorapidity regions: Barrel, Overlap, Endcap. The efficiency plateau is reached for muons with $p_T \geq 7$ GeV/c and the efficiency value is $\sim$ 95%, $\sim$ 87%, $\sim$ 90% for the three regions respectively. The turn-on region of the distributions gets sharper going from the L1 filter step to the L3 filter step due to much better momentum resolution. The improvement obtained with the improved L2 cleaning strategy described in Section 4.4.2 is showed for the Endcap region (where the issue was more evident due to the higher rate of L1 ghosts) in Figure 4.4.

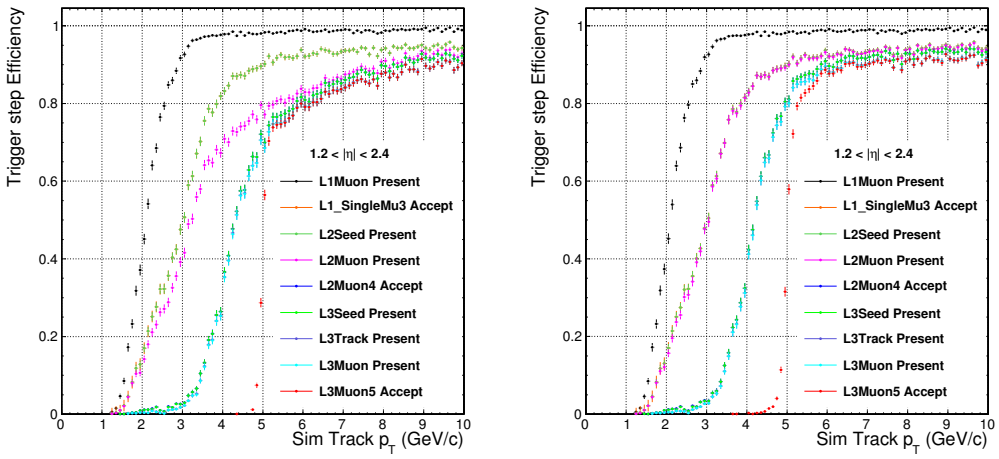


Figure 4.4: *Reconstruction and filtering efficiency for each step of the HLT_Mu5 trigger path for Global Muons with p_T (0-10) GeV/c and pseudorapidity region $1.2 < |\eta| < 2.4$ (Endcap) before the improvement in the "volunteer suppression" logic of the L2 algorithm (left) and after the improvement (right).*

In Figure 4.5 and Figure 4.6 the efficiencies for each step of the HLT_Mu5 path are presented for high p_T (200-1000 GeV/c) reconstructed Global Muons in the three regions. The plateau efficiency slightly decreases at very high momenta. The improvement obtained moving from the *OIState* algorithm to the *Cascade* algorithmSection 4.4.2 is showed for the Overlap region (where the improvement was more evident) in Figure 4.6.

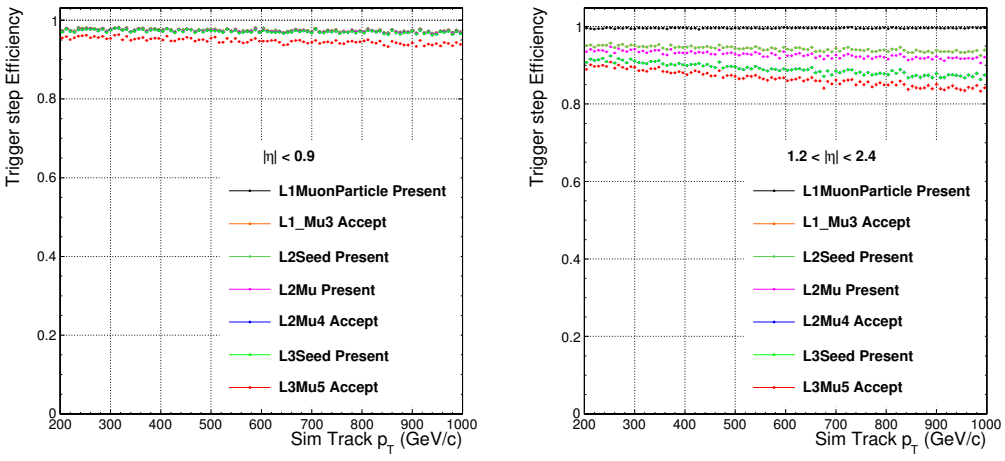


Figure 4.5: *Reconstruction and filtering efficiency for each step of the HLT_Mu5 trigger path for Global Muons with p_T (200-1000) GeV/c and pseudorapidity regions: $|\eta| < 0.9$ (Barrel) (left) and $1.2 < |\eta| < 2.4$ (Endcap) (right).*

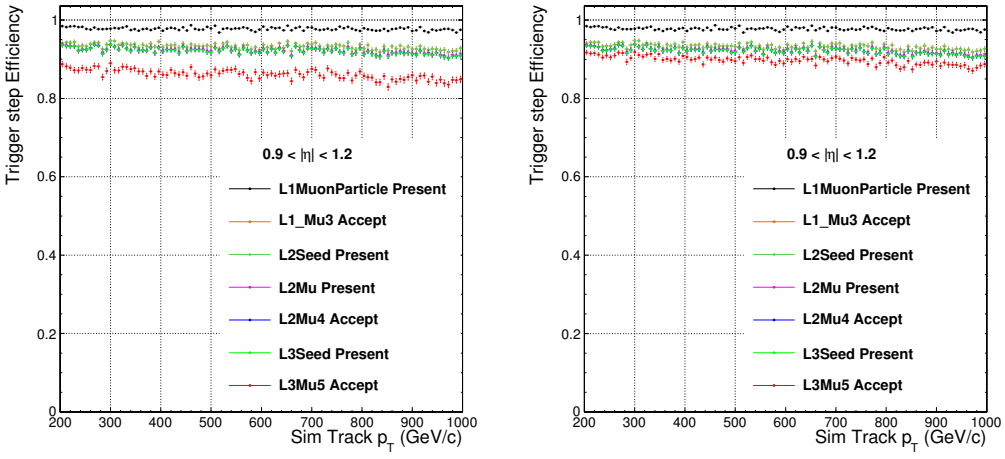


Figure 4.6: *Reconstruction and filtering efficiency for each step of the HLT_Mu5 trigger path for Global Muons with p_T (200-1000) GeV/c and pseudorapidity region $0.9 < |\eta| < 1.2$ (Overlap) when using the* OIState *algorithm at L3 (left) and when using the current default* Cascade *algorithm (right).*

4.5 Muon Measurements with data

In this section the performance of the CMS muon reconstruction, identification and trigger measured on 2011 LHC collision data are discussed.
Muon reconstruction in CMS and the response of various subdetectors to muons have been previously studied in great details using muons from cosmic rays [94] [95] in the commissioning phase of the detector (2008-2009). The first studies on 60 nb^{-1} of 2010 proton-proton collision data were reported in Ref. [96] and more detailed ones with the full 2010 luminosity, 36 pb^{-1}, in Ref. [97]. The work presented in the following has been done in the contest of the $H\rightarrow 4\ell$ analysis. The comparison between the muon efficiencies in data and in simulation is of particular importance for the analysis which requires an estimate of the number of expected signal events. Discrepancies between data and simulation, as well as systematic effects on the efficiency measurements presented in the following sections will be taken into account and propagated to the final event selection of the analysis (Section 7.1).
A collision event in which two muons were reconstructed involving all main CMS detectors in shown in Figure 4.7.

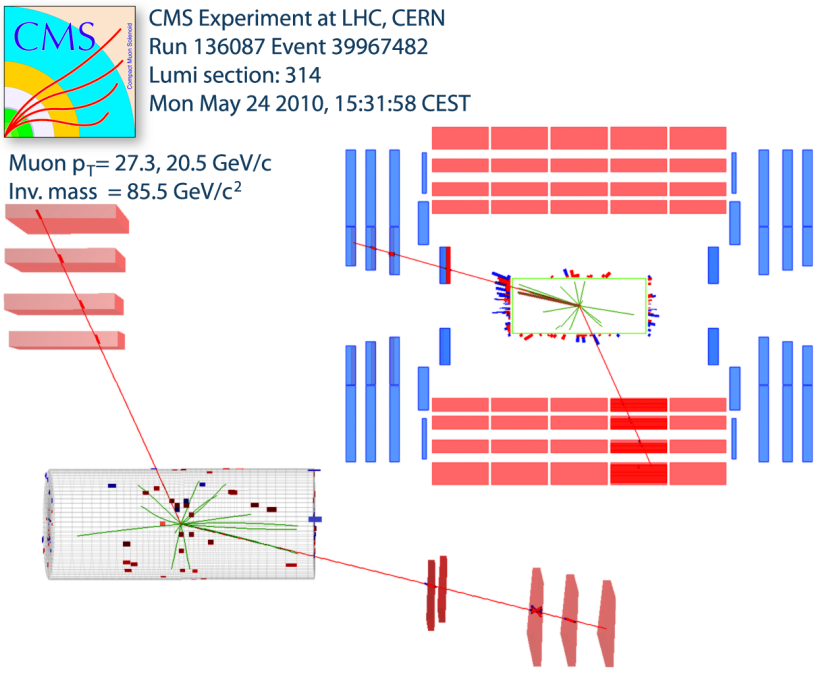


Figure 4.7: *Display of a collision event in which two opposite-sign muons were reconstructed; one muon was identified by the DTs and RPCs, while the other one by the CSCs. Only tracks with transverse momentum $p_T > 1$ GeV/c are shown.*

Throughout this chapter, efficiencies are defined in a relative manner, such that the efficiency of the whole muon reconstruction-identification -isolation-triggering chain can be calculated as the product of the individual factors:

$$\epsilon = \epsilon_{\text{RECO|track}} \times \epsilon_{\text{ID|RECO}} \times \epsilon_{\text{ISO|ID}} \times \epsilon_{\text{TRIGGER1Leg|ISO}} \tag{4.2}$$

where each term represents the efficiency for the probe to pass a given selection or reconstruction step, given that it passes the criteria for the previous one.
In Section 4.5.2, muon identification efficiencies are defined relative to reconstructed tracker tracks. In Section 4.5.3, isolation efficiencies are calculated on a sample of identified muons. In Section 4.5.5, trigger efficiencies are defined relative to muons identified offline and passing isolation criteria.
These efficiencies are evaluated on an exclusive sample of muons from J/ψ and Z decays by applying a tag-and-probe technique.

The tag-and-probe method

Using this technique it is possible to obtain almost unbiased estimates of the efficiencies of the different stages of muon trigger and offline reconstruction. Events are selected with strict selection requirements on one muon (the *tag* muon), and with a more relaxed selection on the other muon (the *probe* muon), such that the selection applied to the probe muon does not bias the efficiency that one wants to measure. The fraction of probe muons which pass the selection under study gives the measurement of its efficiency.
In the efficiency measurement it is important to subtract the combinatorial background of tag-probe pairs not coming from the resonance under study, where the probe is for example, in the case of muon identification efficiency, a charged hadron track. The subtraction is done by performing a simultaneous fit to the invariant mass spectra for passing and failing probes with identical signal shape and appropriate background shapes: the efficiency is then computed from the normalization of the signal shapes in the two spectra. For the J/ψ tag-and-probe the lineshape model is based on a Crystal Ball function for the resonance and an exponential function for the background. For the Z tag-and-probe a Voigtian function is used to model the resonance shape and an exponential function to model the background.
The uncertainty on the fitted efficiency is determined from the profiled likelihood function. In the estimation, several parameters of the fit are left floating: normalizations of signal and background, efficiency of the background and parameters controlling the shapes of the signal and background. In this way, the uncertainty band includes naturally the contribu-

tions from the background subtraction procedure; when the background is large, like for J/ψ, these contributions completely dominate the pure statistical uncertainty.

4.5.1 Simulation and data sample

The 2011 data have been collected with different LHC running conditions corresponding to different values of instantaneous luminosity; as a consequence several instances of HLT Menu were used for the online selection in order to address the evolving rate constraints.
The increase of the instantaneous luminosity has been accompanied by the increase of the number of pile-up interactions that translates in more reconstructed primary vertices per event.
For the performance studies that will be presented in this chapter it is useful to subdivide the analyzed data into five different periods that correspond to significant changes in pile-up conditions as well as in the muon trigger configuration (see Figure 4.8). The distribution of the number of reconstructed primary vertices per event for each of the 5 periods, which are listed in table Table 4.4, is shown in Figure 4.9.

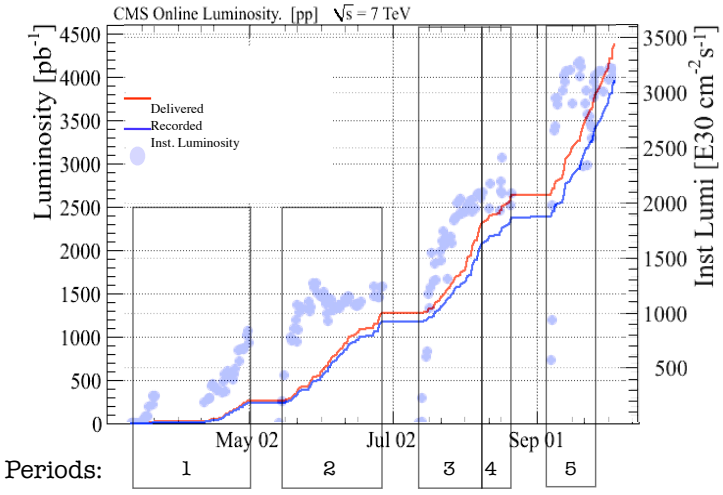


Figure 4.8: *Schematic representation of 2011 data partition into five periods. The data collected in periods 1, 2, 3, 4 corresponds to the 2011-A run described in Section 3.5, while period 5 is a part of 2011-B run.*

The trigger paths characterized by the request of muon objects in the event

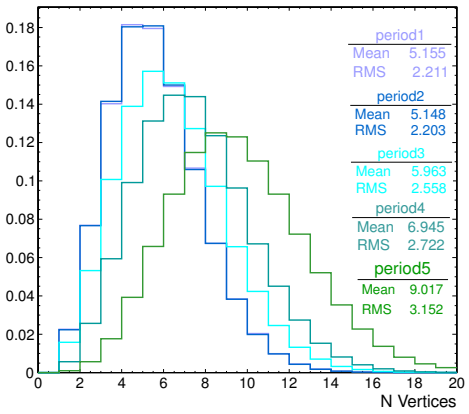


Figure 4.9: *Distribution of the number of primary vertices reconstructed per event in the five data-taking periods. Events are selected with the* `HLT_IsoMu24` *trigger path).*

Period	Average Ist. Luminosity	Run Range	Int. Luminosity
2011A - 1	$\mathcal{L}$=5 · 10^{32} cm^{-2} s^{-1}	160404-163869	204.2 pb^{-1}
2011A - 2	$\mathcal{L}$=1 · 10^{33} cm^{-2} s^{-1}	165088-167913	886.6 pb^{-1}
2011A - 3	$\mathcal{L}$=2 · 10^{33} cm^{-2} s^{-1}	170249-173198	781 pb^{-1}
2011A - 4	$\mathcal{L}$=3 · 10^{33} cm^{-2} s^{-1}	173236-173692	253 pb^{-1}
2011B - 5	$\mathcal{L}$=3 · 10^{33} cm^{-2} s^{-1}	175860-177053	652.4 pb^{-1}

Table 4.4: *Periods of data taking.*

can be subdivided in to three groups; the paths in each group contribute to the final OR of decisions that determines whether an event is stored in one of the following PDs: *SingleMuon* dataset, *DoubleMuon* dataset, *MuOnia* dataset. In the following the main features of the trigger paths that "seed" each PD are described and their usage in the studies presented through this chapter is explained.

- *SingleMuon triggers.* These triggers only require the presence of one L3 muon in the event; as a consequence the lowest un-prescaled trigger path without isolation requirements in the $3 \cdot 10^{33}$ cm^{-2} s^{-1} HLT Menu has a p_T threshold of 40 GeV/c (`HLT_Mu40`) that can goes down to 24 GeV/c if the isolation requirements are applied (`HLT_IsoMu24`). The events used to measure the muon identification, isolation and trigger efficiency with the Z tag-and-probe technique (Section 4.5.2) are selected with the `HLT_IsoMu24` trigger.

- *DoubleMuon triggers.* These triggers require the presence of two L3 Muons in the event; the rate constraints are thus less stringent and the p_T thresholds of the lowest unprescaled path in the $3 \cdot 10^{33}$ cm^{-2} s^{-1}

HLT Menu are 13 and 8 GeV/c for the two muons. The efficiency of this trigger (`HLT_DoubleMu13Mu8`) is discussed in Section 4.5.5. A tighter version of this trigger, `HLT_DoubleMu17Mu8`, which was needed for the for the very last period of data taking ($3.5 \cdot 10^{33}$ cm^{-2} s^{-1}) is the trigger used for the H$\rightarrow 4\ell$ analysis as will be described in Section 6.4.

- *MuOnia triggers.* For analysis involving measurements of low mass resonances, double muon triggers with lower p_{T} thresholds but with the additional request of opposite-charge L3 muons yielding an invariant mass in a restricted range have been deployed. The contribution of these triggers is clearly shown in the di-muon invariant mass spectrum obtained with 1.1 fb^{-1} of 2011 data in Figure 4.10. Moreover, in order to measure the muon identification efficiency in the low p_{T} range (which is not accessible with Z events) and as well the efficiency of these low p_{T} double muon triggers, a specialized high-level trigger was implemented: muon-plus-track trigger. This trigger selects events in which a L3 Muon can be paired to an inner tracker track of opposite charge yielding an invariant mass close to that of the J/ψ peak. In order to sample evenly the efficiency turn-on curve, multiple instances of the trigger have been deployed with different thresholds on the transverse momentum of the silicon tracker track. These triggers have been used in Section 4.5.2 to measure the muon identification efficiency with the J/ψ tag-and-probe method.

In order to compare the results obtained in data to predicted ones, a number of simulated MC samples (cf. Section 3.4) were used (in CMS jargon they are part of the "Summer11" MC production).
Samples of QCD, in which only events containing at least one muon with transverse momentum greater than given thresholds are selected at generation level, and $t\bar{t}$ events were generated using PYTHIA 6 [99] with the Z2 tune [100]. Samples of prompt J/ψ as well as J/ψ originating from the decays of B hadrons were generated with PYTHIA interfaced to EVTGEN [101]. Inclusive W and Z samples and non-resonant Drell-Yan were produced using the POWHEG [60] event generator, interfaced with PYTHIA for the simulation of showering and hadronization processes. For W and Z samples with a given number of jets, the MADGRAPH [102] event generator was used, combined with PYTHIA for showering and hadronization. In all these samples a pile-up scenario is simulated; it is characterized by a flat distribution from 0 to 10 interactions per event with a tail above 10 to higher values corresponding to a poisson distribution with a mean of 10 interactions; this scenario has been designed in order to be able to perform a proper event re-weighting based on the actual-data scenario. A re-weighting procedure has not been applied for the muon performances studies presented in this chapter (while it is properly taken into account in

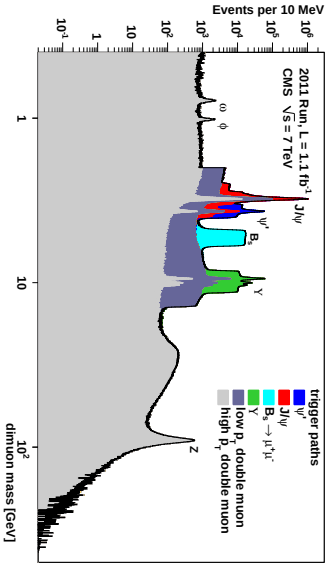


Figure 4.10: *Yield of the different double muon trigger path as a function of the di-muon invariant mass with 2011 data collected by early July, corresponding to an integrated luminosity of 1.1 fb^{-1} obtained from overlapping several double muon trigger paths [98].*

the final results of the H$\rightarrow$ 4ℓ analysis); proper care is then needed while comparing data/mc variables that shows a dependence on the pile-up scenario.

4.5.2 Reconstruction and Identification efficiency

When measuring muon identification efficiencies with the tag-and-probe technique, the probes are tracks reconstructed using only the inner tracker, so there is no bias from the muon subdetectors. The efficiency to reconstruct a muon in the inner tracker has been measured separately and found to be 99% or higher in the whole tracker acceptance, in good agreement with the expectation from simulations [91, 103].

For the Z resonance, an unbiased sample of di-muon pairs can be collected efficiently using SingleMuon triggers. In particular, in the following, the Z tag-and probe is performed on events selected by the `HLT_IsoMu24`, with a couple of tag and probe objects that yield to an invariant mass in the (70-130) GeV/c^2 range and with the tag muon being a Tight Muon (Section 4.2), isolated ($R_{Iso} < 0.1$) and matched with the L3 Muon of the `HLT_IsoMu24` path; in this way the bias of the trigger request on the probe measurements is removed. With this event selection the muon identifica-

tion efficiency has been measured in the p_T range (20-100) GeV/c.
The
For the measurement with J/ψ decays, the tag muon is requested to be reconstructed as a Global muon and a Tracker muon, the tag-and-probe couple is requested to yield to an invariant mass in the (2.8-3.4) GeV/c^2 range and finally, to further reduce the background, the distance between the tag and probe tracks at the point of closest approach is required to be less then 1 mm.
The (prescaled) muon-plus-track triggers are used to collect events; in particular two instances of the same trigger have been used to sample two different p_T ranges. The `HLT_Mu5_Track2` path has been used to measure the muon identification efficiency in the p_T range (2-7) GeV/c while the `HLT_Mu7_Track7` for the p_T range (7-20) GeV/c.
These triggers do not bias the efficiencies related to the muon system, as L3 muon object of the muon-plus-track trigger is matched with the tag muon and not to the probe. However, they introduce a small positive bias in the efficiency for the muon identification selections that include quality requirements on the muon tracker track. This bias was found to be $(0.7 \pm 0.1)\%$ in the barrel and $(0.3 \pm 0.2)\%$ in the endcaps on 2010 data [97]. These biases are not corrected for as they cancel out in data-to-simulation efficiency ratios that are used in the analysis presented in the next chapters.
Under certain kinematic configurations muons from J/ψ decays can be close to each other in the muon system. This introduces unwanted correlations in the measurement and can result in inefficiencies for some muon identification algorithms. In order to obtain an unbiased measurement of single-muon efficiencies, a separation requirement has been applied to the tag-probe pairs: the extrapolated impact points of the two muon tracks on the surface of the first muon station must have an angular separation $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.5$. The impact of the same requirement on $Z \to \mu^+\mu^-$ events is small: only 0.2% of the $Z \to \mu^+\mu^-$ events fail the separation criterion above [97]).
Efficiency of ID requirements including $|d_{xy}|$ and $|d_z|$ cannot be measured with J/ψ tag-and-probe because of the significant contamination of non prompt muons coming from J/ψ from B hadrons, expecially for high p_T J/ψ muons (10-20 GeV/c). These ID requirements are therefore only considered while measuring efficiencies with Z tag-and-probe.

Results

Figure 4.11 shows the muon identification efficiency given that a tracker track exists for three selection types: Soft, Global, Tight Muons. The efficiency is reported with respect to the probe p_T in two pseudorapidity regions, barrel and endcap, and it is measured using $J/\psi \to \mu^+\mu^-$ events

for $p_T < 20$ GeV/c and $Z \to \mu^+\mu^-$ events for $p_T > 20$ GeV/c. The results on the data from Period 1 are compared with the ones extracted applying the same procedure on simulated events.
For comparisons with $Z \to \mu^+\mu^-$ events, an unweighted sample of simulated events corresponding to an integrated luminosity of 2.5 fb^{-1} has been used, consisting of $Z \to \mu^+\mu^-$, W+jets, and $t\bar{t}$ samples (cf. Section 4.5.1). For studies at the J/ψ peak, separate samples of prompt J/$\psi \to \mu^+\mu^-$ and $B \to J/\psi + X \to \mu^+\mu^- + X$ corresponding to 13 pb^{-1} have been used. Simulation of the background processes from QCD events has not been included for either the J/ψ as the Z case, as it would have not been feasible to simulate a proper amount of inclusive muon-plus-track and single muon trigger events.

In general a good agreement is found between tag-and-probe results in data and in simulation, see Figure 4.11. The main discrepancy is around efficiency turn-ons, where the efficiency in data is systematically higher than in the simulation. Moreover the efficiency of the Soft selection is in general higher in data then in simulation. This is due to the slightly more conservative uncertainties assumed in the muon identification on data which make the identification requirements slightly more efficient [97].
The plateau is reached at $p_T \sim 5/3$ GeV/c for Soft Muons (Barrel/Endcap), $\sim 6/4$ GeV/c for Global Muons (Barrel/Endcap), $\sim 9/7$ GeV/c for Tight Muons (Barrel/Endcap).
The dependency of the plateau efficiency on the pseudorapidity has been measured using $Z \to \mu^+\mu^-$ events and is shown in Figure 4.12. In order to estimate the effect of pile-up on the muon identification performance, the measurement of the efficiency at the plateau has been performed as function of the number of reconstructed primary vertices N_{vtx}. The simulation does not show any dependence with respect to N_{vtx}, while in data a very little dependance can be observed. Moreover, independently of the amount of pile-up, a degradation of the muon reconstruction performances localized in the endcap region for the Global and Tight selection is observed in Period 5; the reasons of this efficiency loss are currently under investigation.
As already introduced, these measurements are needed to properly correct the simulation with data-to-simulations correction factors; the corrections that will be used in the H$\to 4\ell$ analysis, as described in Section 7.1, are those presented in Table 4.5 and Table 4.6 where the efficiency computed in several p_T bins are reported for different η regions according to the natural separation suggested by the muon spectrometer layout (Section 2.2.5). These correction factors are reported for Global Muon considering data collected in Period 1, 2, 3, 4 all together and data collected in Period 5.

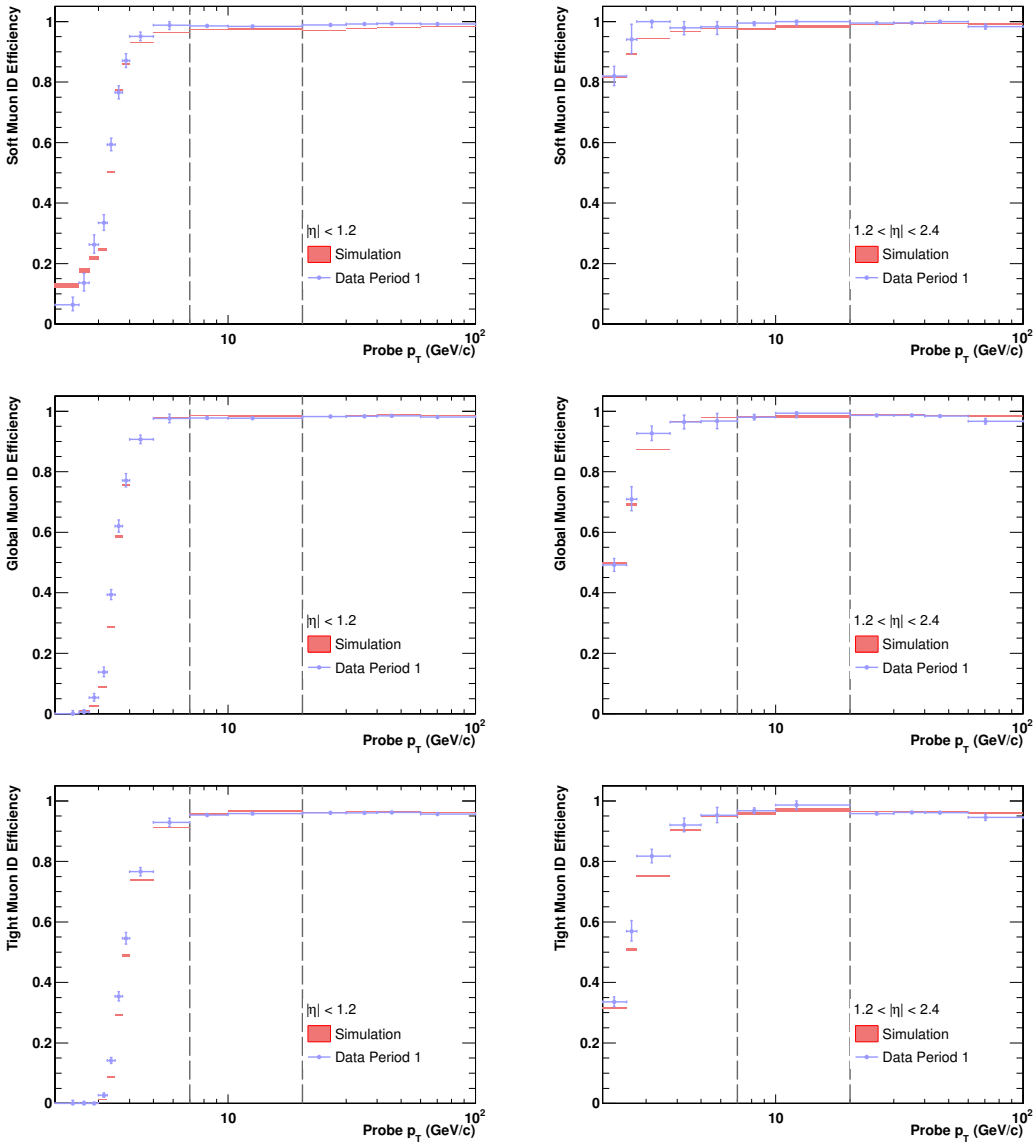


Figure 4.11: *Tag-and-probe results for the muon identification efficiency in data Period 1 compared to simulation. Given that a tracker track exists, the plots show the efficiency as a function of muon p_T for Soft Muons (top), Global Muons (middle) and Tight Muons (bottom) in the $|\eta| < 1.2$ (Barrel) (left) and $1.2 < |\eta| < 2.4$ (Endcap) (right) regions. The measurement is done using $J/\psi \to \mu^+\mu^-$ events for $p_T < 20$ GeV/c and $Z \to \mu^+\mu^-$ events for $p_T > 20$ GeV/c. In tange covered with J/ψ the $|d_{xy}|$ and $|d_z|$ cuts have been removed from the muon identification types. The dotted lines at 7 and 20 GeV/c subdivide the events used for the measurement in three subsets depending on the trigger used to select them: HLT_Mu5_Track2, HLT_Mu7_Track7, HLT_IsoMu24.*

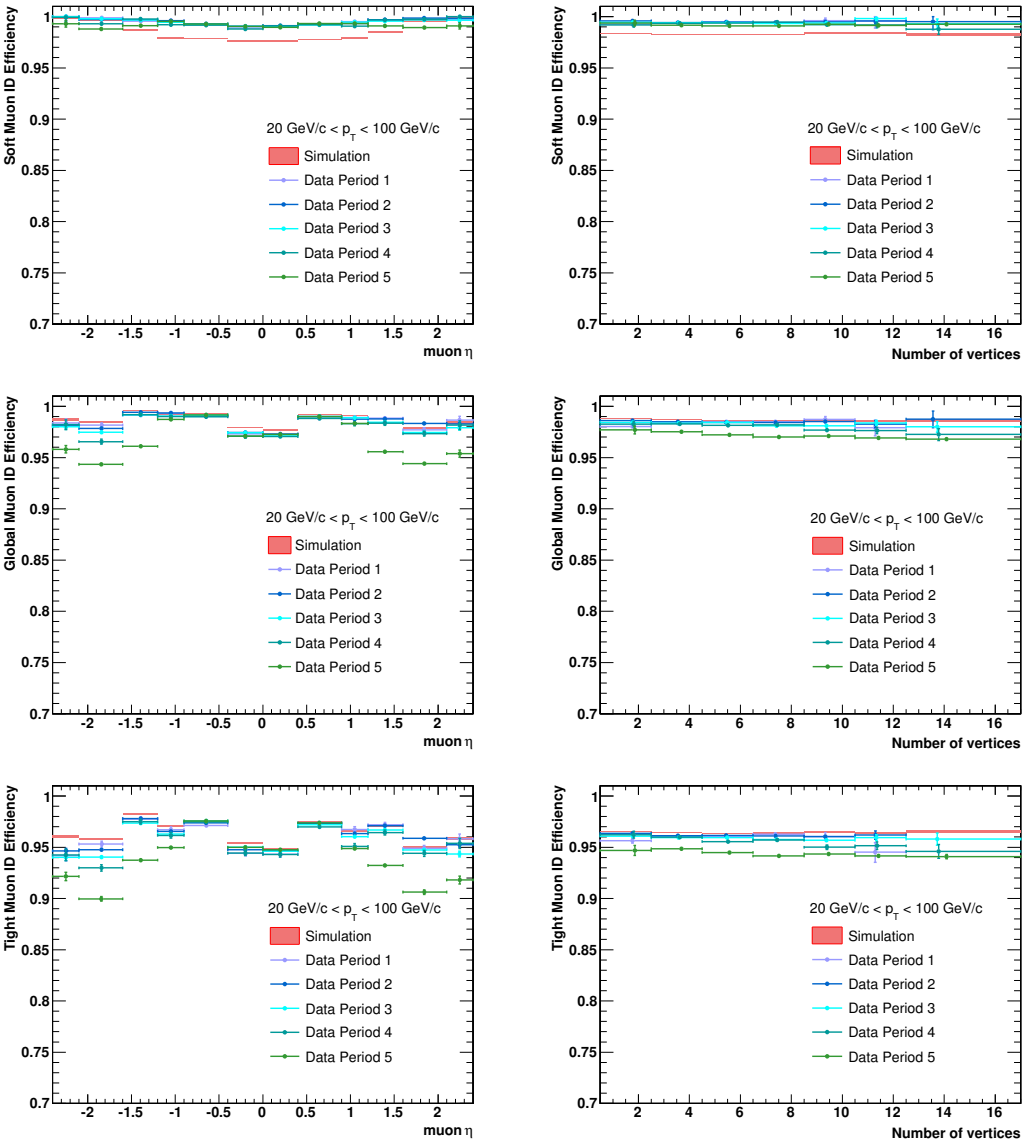


Figure 4.12: *Tag-and-probe results for the muon identification efficiency for all the 5 data periods compared to simulation. Given that a tracker track exists, the plots show the efficiency as a function of muon η (left) and N_{vtx} (right) for Soft Muons (top), Global Muons (middle) and Tight Muons (bottom) The measurement is done using $Z \to \mu^+\mu^-$ events for probe muons with $p_T > 20$ GeV/c.*

Systematic uncertainties

The systematic uncertainties on the measured efficiency coming from the lineshape modelling and background subtraction have been studied on simulated events and on data; a detailed description of the different sources

Table 4.5: *Global Muon identification efficiencies for single muons in Period 1 to 4, measured with the tag-and-probe technique on data, and data/MC discrepancy obtained applying the method on MC. All measurements are obtained using Z decays, except below 20 GeV/c, where J/ψ muon decays are used. The first uncertainty quoted in the last column derives from the uncertainty on the fitted efficiencies in data and simulation; the second one is from the additional systematic uncertainties described in this section.*

p_T Range (GeV/c)	$\|\eta\|$ Coverage	MC (%)	Data (%)	Data/MC Ratio
5-7	0-1.2	0.980	0.977	$0.996 \pm 0.007 \pm 0.010$
5-7	1.2-2.4	0.978	0.996	$1.018 \pm 0.013 \pm 0.014$
7-20	0-1.2	0.985	0.977	$0.991 \pm 0.003 \pm 0.010$
7-20	1.2-2.4	0.982	0.979	$0.997 \pm 0.005 \pm 0.014$
20-30	0-0.9	0.980	0.976	$0.997 \pm 0.002 \pm 0.002$
20-30	0.9-1.2	0.988	0.989	$1.001 \pm 0.003 \pm 0.002$
20-30	1.2-1.6	0.990	0.989	$0.999 \pm 0.003 \pm 0.004$
20-30	1.6-2.1	0.988	0.986	$0.998 \pm 0.002 \pm 0.004$
20-30	2.1-2.4	0.989	0.986	$0.997 \pm 0.004 \pm 0.004$
30-50	0-0.9	0.987	0.983	$0.996 \pm 0.001 \pm 0.002$
30-50	0.9-1.2	0.993	0.990	$0.997 \pm 0.001 \pm 0.002$
30-50	1.2-1.6	0.992	0.990	$0.997 \pm 0.001 \pm 0.004$
30-50	1.6-2.1	0.981	0.977	$0.995 \pm 0.001 \pm 0.004$
30-50	2.1-2.4	0.986	0.982	$0.995 \pm 0.001 \pm 0.004$
> 50	0-0.9	0.986	0.982	$0.996 \pm 0.001 \pm 0.002$
> 50	0.9-1.2	0.993	0.992	$0.999 \pm 0.002 \pm 0.002$
> 50	1.2-1.6	0.991	0.987	$0.996 \pm 0.002 \pm 0.004$
> 50	1.6-2.1	0.975	0.968	$0.992 \pm 0.003 \pm 0.004$
> 50	2.1-2.4	0.976	0.966	$0.990 \pm 0.006 \pm 0.004$

of systematics can be found in [97]. The final systematic values quoted in table Table 4.5, Table 4.6 have been obtained with the following checks. In simulation, the efficiencies obtained applying the tag-and-probe method have been compared with the ones computed by simple counting of the passing and failing probes in events where the reconstructed dimuon is matched to a simulated $J/\psi \to \mu^+\mu^-$ or $Z \to \mu^+\mu^-$ decay.
In data, the efficiencies for J/ψ events have been computed with the tag-and-probe method also with a different lineshape model using a simple Gaussian instead of a Crystal Ball function to model the resonance, and a a quadratic polynomial instead of an exponential for the background. Moreover, the residual correlation effects between the two muons have been studied by changing the separation criteria.

Table 4.6: *Global Muon identification efficiencies for single muons in Period 5, measured with the tag-and-probe technique on data, and data/MC discrepancy obtained applying the method on MC. All measurements are obtained using Z decays, except below 20 GeV/c, where J/ψ muon decays are used. The first uncertainty quoted in the last column derives from the uncertainty on the fitted efficiencies in data and simulation; the second one is from the additional systematic uncertainties described in this section.*

p_T Range (GeV/c)	$\|\eta\|$ Coverage	MC (%)	Data (%)	Data/MC Ratio
5-7	0-1.2	0.980	0.960	0.979 ± 0.021 ± 0.010
5-7	1.2-2.4	0.978	0.881	0.901 ± 0.035 ± 0.014
7-20	0-1.2	0.985	0.978	0.993 ± 0.009 ± 0.010
7-20	1.2-2.4	0.982	0.938	0.956 ± 0.018 ± 0.014
20-30	0-0.9	0.980	0.979	0.999 ± 0.003 ± 0.002
20-30	0.9-1.2	0.988	0.984	0.996 ± 0.006 ± 0.002
20-30	1.2-1.6	0.990	0.954	0.963 ± 0.005 ± 0.004
20-30	1.6-2.1	0.988	0.957	0.969 ± 0.005 ± 0.004
20-30	2.1-2.4	0.989	0.960	0.971 ± 0.009 ± 0.004
30-50	0-0.9	0.987	0.983	0.996 ± 0.001 ± 0.002
30-50	0.9-1.2	0.993	0.985	0.992 ± 0.001 ± 0.002
30-50	1.2-1.6	0.993	0.960	0.967 ± 0.001 ± 0.004
30-50	1.6-2.1	0.981	0.944	0.962 ± 0.002 ± 0.004
30-50	2.1-2.4	0.986	0.959	0.972 ± 0.003 ± 0.004
>50	0-0.9	0.986	0.979	0.993 ± 0.002 ± 0.002
>50	0.9-1.2	0.993	0.989	0.996 ± 0.004 ± 0.002
>50	1.2-1.6	0.991	0.955	0.963 ± 0.005 ± 0.004
>50	1.6-2.1	0.976	0.931	0.955 ± 0.006 ± 0.004
>50	2.1-2.4	0.976	0.928	0.950 ± 0.013 ± 0.004

Muon Identification Probability for particles other then muons

Together with the identification efficiency on prompt isolated muons it is useful to known, for each identification algorithm, the probability to identify as a muon particles other then muons (kaons, pions, protons) which can be source of punch-through (mainly from protons) and decay-in-flight (manly from kaons and pions).
With a study performed on 2010 data and described in details in Ref. [97] this probability has been estimated directly from data. In particular the fraction of events in which the hadron tracks are identied as a Soft Muon, Global Muon, or Tight Muon has been measured as a function of several relevant track parameters. It is interesting for the analysis that will be presented in the following chapters to note that the misidentification probability increases with the momentum of the hadron track and it reaches a

plateau value that is lower than 1% for Soft and Global Muon while it is lower that 0.1% per hadron track for Tight Muon identification.

4.5.3 Muon Isolation efficiency

The efficiency of the Isolation criteria (R_{Iso}^{Tk}, R_{Iso}) described in Section 4.3 with respect to prompt Global muons from Z decay has been measured in data and simulation with the Z tag-and-probe technique. As the minimum p_{T} cut in the H$\rightarrow 4\mu$ channel is 5 GeV/c, the p_{T} range considered for the following measurements is (5-100) GeV/c; clearly the statistic uncertainty in the lowest p_{T} bins is larger due to the p_{T} spectrum of Z decays. Figure 4.13 shows the efficiency as a function of the threshold on the isolation variables. The variable R_{Iso} is sensible to the number of pile-up interactions in the event. In order to account for differences in the amount of pile-up in data and MC, the isolation efficiency has been computed in different bins of the number of reconstructed primary vertices N_{vtx}. A general good data-simuation agreement is found for each N_{vtx} bin.
In Figure 4.14 the efficiencies of the isolation requirements $R_{Iso}^{Tk} < 0.7$ and

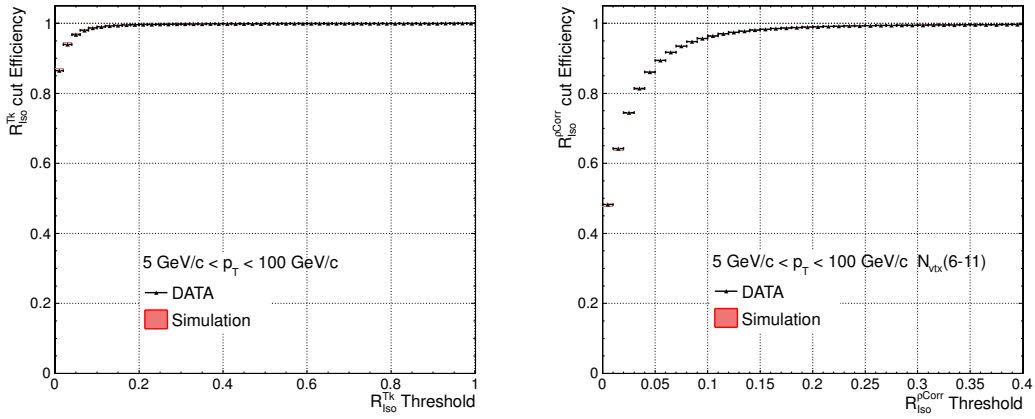


Figure 4.13: *Efficiency of the isolation algorithms for Global Muons with p_{T} (5-100) GeV/c from Z decay as a function of the threshold on the corresponding isolation variable. The R_{Iso} variable includes the ρ correction described in Section 4.3.*

$R_{Iso} < 0.15$ are presented as a function of the number of reconstructed primary vertices. The R_{Iso} variable is considered here before the ρ correction showing the important dependence on the pile-up conditions.
The effect of the ρ correction on the R_{Iso} efficiency is shown in Figure 4.15. The N_{vtx} dependence is considerably reduced even if not neglected. A good agreement between data and simulation is found both before and after the correction for all the five periods of data-taking.

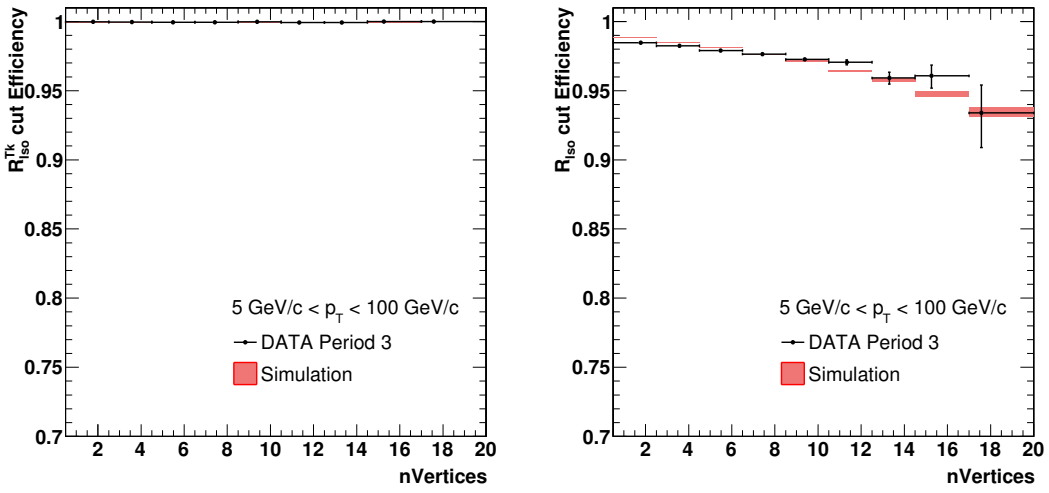


Figure 4.14: *Efficiency of the isolation requirements $R^{Tk}_{Iso} < 0.7$ (left), $R_{Iso} < 0.15$ (right) for Global Muon with p_T (5-100) GeV/c from Z decay as a function of the number of reconstructed primary vertices. The distributions refer to data collected in Period 3 but the trends are similar in all the periods.*

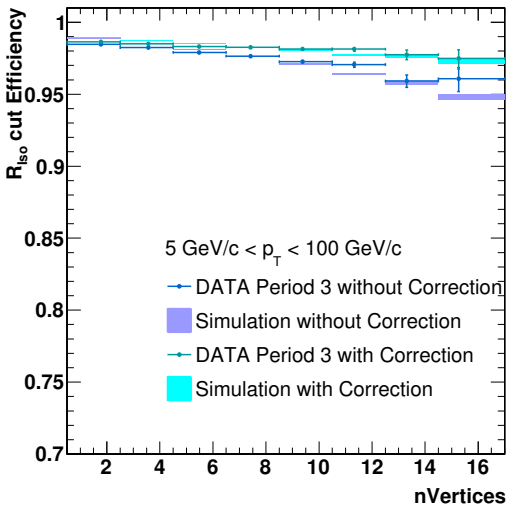


Figure 4.15: *Efficiency of the isolation requirements $R_{Iso} < 0.15$ for Global Muon with p_T (5-100) GeV/c from Z decay as a function of the number of reconstructed primary vertices. The distribution before and after the ρ correction is shown for data Period 3 and compared to simulation.*

In order to monitor the performances of the isolation algorithms throughout the data taking periods and check whether they remained stable, the distribution of the R_{Iso} (ρ corrected) efficiency as a function of p_T and pseu-

dorapidity of Global Muons in each of the five periods has been compared subdividing the events in three different pile-up categories according to the number of reconstructed primary vertices; no significant differences have been found (Figure 4.16).

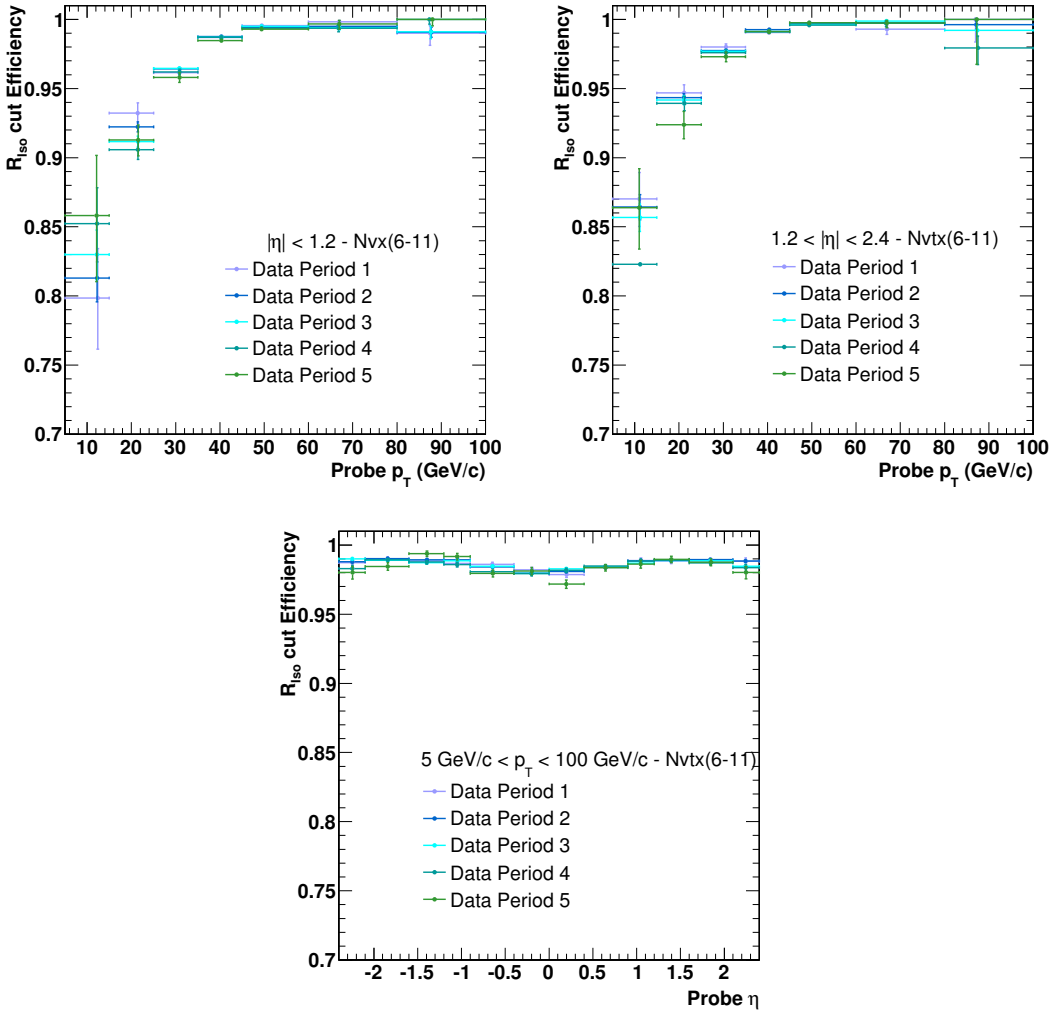


Figure 4.16: *Efficiency of the isolation requirement $R_{Iso} < 0.15$ for Global Muon from Z decay for $6 < N_{vtx} < 11$; the comparison between the five data-taking periods is shown as a function of the Global Muon η and p_T.*

Therefore, isolation efficiency as a function of the p_T and η of the probe Global Muon were estimated considering together all data of the five periods; these correction factors are reported in Table **??**. A further subdivision in N_{vtx} bins is needed as a residual pile-up dependence is still present even after the ρ correction is applied. The absolute efficiencies as a func-

tion of the Global Muon p_T and η are shown for data and simulation in Figure 4.17 for one bin of N_{vtx}.

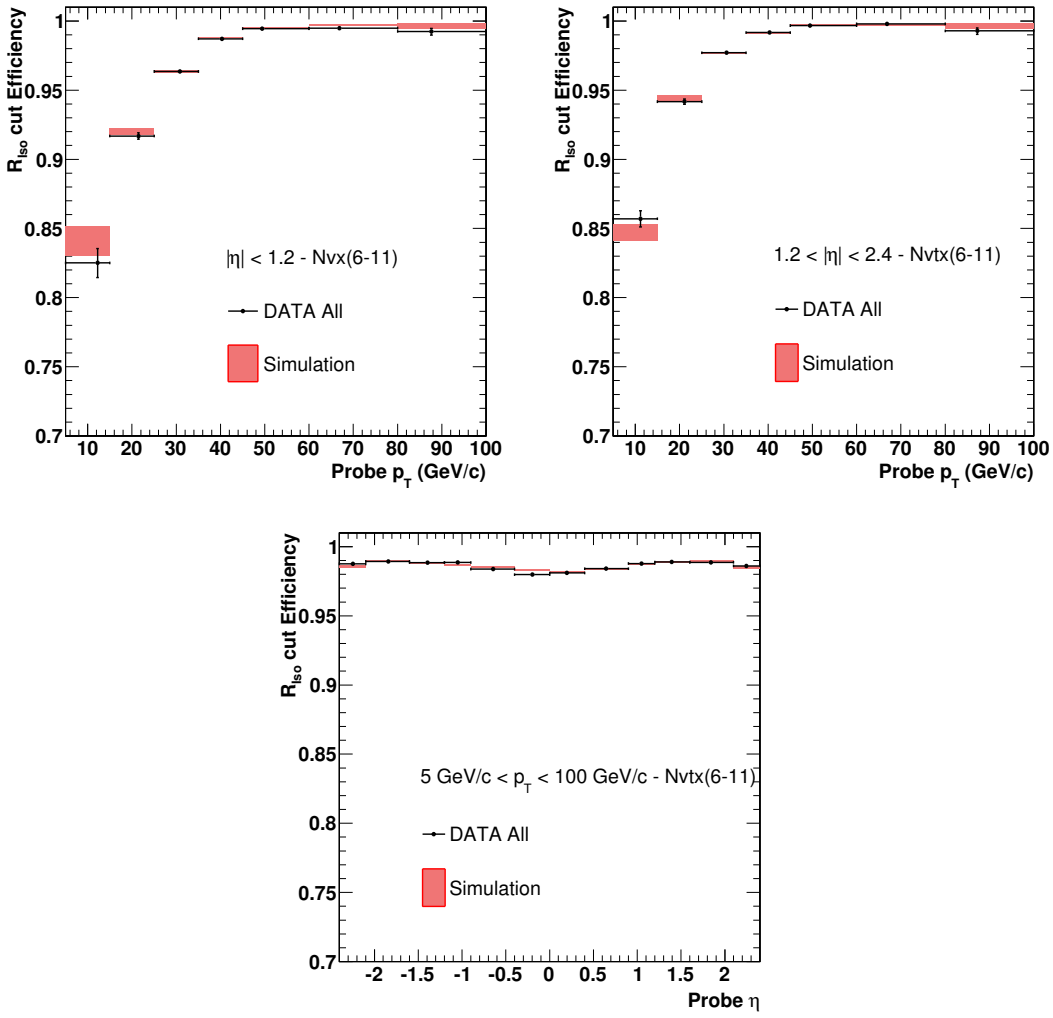


Figure 4.17: *Efficiency of the isolation requirements $R_{Iso} < 0.15$ for Global Muon from Z decay for events with number of reconstructed vertices: $6 < N_{vtx} < 11$; the comparison between data and simulation is shown as a function of the Global Muon p_T and and η.*

Systematic uncertainties

Systematic uncertainties on measured isolation efficiencies are derived in a similar way as for reconstruction and identification efficiencies. The efficiencies obtained applying the tag-and-probe method to the simulated samples of muons have been compared with the ones computed by simple counting of the passing and failing probes in events where the re-

constructed dimuon is matched to a simulated $Z \rightarrow \mu^+\mu^-$ decay. The results are in good agreement: the difference in the absolute efficiencies is smaller than 0.1%. On data, the efficiencies for $Z \rightarrow \mu^+\mu^-$ events have been computed using a quadratic polynomial instead of an exponential for the background, the results differ by 0.05%. These differences were taken as systematic uncertainty on the measured efficiencies.

4.5.4 Compatibility with a common vertex

On top of the $|d_{xy}|$, $|d_z|$ cuts embedded in the Muon ID definition, in the $H \rightarrow ZZ \rightarrow 4\ell$ analysis selection, a tighter cut on the compatibility of the leptons with the primary vertex is applied to reject muons from b hadrons decays. For each lepton the compatibility with this vertex can be estimated by calculating the impact parameter significance, denoted by $\text{SIP}_{3D} = \frac{|\text{IP3D}|}{\sigma_{IP3D}}$, where $|IP3D|$ is the absolute value of the tridimensional lepton impact parameter with respect to the primary vertex, and σ_{IP3D} the associated uncertainty.
Figure 4.18 shows the efficiency as a function of the threshold on the SIP variable for isolated Global Muons from Z decay: the $H \rightarrow 4\ell$ analysis working point is $SIP_{3D} < 4$. For this very loose cut, the data-MC agreement is good.

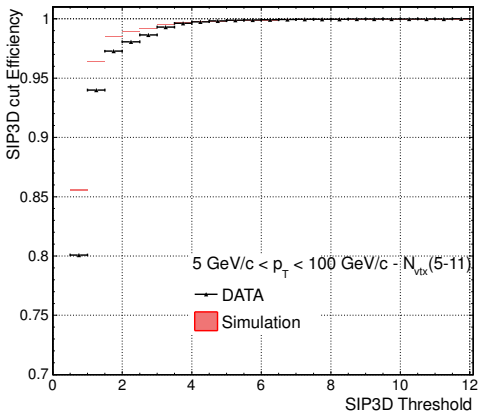


Figure 4.18: *Efficiency of the SIP_{3D} discriminant for Global Muon with p_T (5-100) GeV/c from Z decay as a function of the threshold on the SIP_{3D} variable.*

Table 4.7: *Isolation efficiencies for $R_{Iso} < 0.15$ computed on Global Muons from Z decay, and data/MC ratios. The results are reported in several p_T and η bins and for specific ranges of N_{vtx}. The first uncertainty quoted in the last column derives from the uncertainty on the fitted efficiencies in data and simulation; the second one represents the additional systematic uncertainties described in this section.*

p_T Range (GeV/c)	$\|\eta\|$ Coverage	N_{vxt}	Data (%)	MC (%)	Data/MC Ratio
5-15	0-1.2	0-5	0.882	0.860	0.975 ± 0.018 ± 0.001
15-25	0-1.2	0-5	0.940	0.928	0.987 ±0.004 ± 0.001
25-35	0-1.2	0-5	0.973	0.970	0.997 ±0.001 ± 0.001
35-45	0-1.2	0-5	0.992	0.989	0.996 ±0.000 ± 0.001
45-60	0-1.2	0-5	0.997	0.996	0.998 ±0.000 ± 0.001
60-80	0-1.2	0-5	0.998	0.996	0.997 ±0.001± 0.001
>80	0-1.2	0-5	0.997	0.996	0.999 ±0.002± 0.001
5-15	1.2-2.4	0-5	0.911	0.883	0.969 ±0.009± 0.001
15-25	1.2-2.4	0-5	0.961	0.957	0.996 ±0.003 ± 0.001
25-35	1.2-2.4	0-5	0.982	0.982	1.001 ±0.001 ± 0.001
35-45	1.2-2.4	0-5	0.994	0.993	0.999 ±0.000 ± 0.001
45-60	1.2-2.4	0-5	0.998	0.997	0.999 ±0.000 ± 0.001
60-80	1.2-2.4	0-5	0.998	0.997	0.999 ±0.001 ± 0.001
>80	1.2-2.4	0-5	1.000	0.997	0.997 ±0.002± 0.001
15-25	0-1.2	6-11	0.920	0.917	0.997 ±0.004 ± 0.001
25-35	0-1.2	6-11	0.964	0.963	1.000 ±0.001 ± 0.001
35-45	0-1.2	6-11	0.988	0.987	0.999 ±0.000 ± 0.001
45-60	0-1.2	6-11	0.995	0.994	0.999 ±0.000 ± 0.001
60-80	0-1.2	6-11	0.997	0.995	0.997 ±0.001 ± 0.001
>80	0-1.2	6-11	0.996	0.992	0.996 ±0.003 ± 0.001
5-15	1.2-2.4	6-11	0.847	0.857	1.011 ±0.010 ± 0.001
15-25	1.2-2.4	6-11	0.944	0.942	0.997 ±0.003 ± 0.001
25-35	1.2-2.4	6-11	0.977	0.977	1.000 ±0.001 ± 0.001
35-45	1.2-2.4	6-11	0.991	0.992	1.001 ±0.000 ± 0.001
45-60	1.2-2.4	6-11	0.997	0.997	0.999 ±0.000 ± 0.001
60-80	1.2-2.4	6-11	0.997	0.998	1.000 ±0.001 ± 0.001
>80	1.2-2.4	6-11	0.997	0.993	0.996 ±0.002 ± 0.001
5-15	0-1.2	12-20	0.804	0.741	0.921 ±0.065 ± 0.001
15-25	0-1.2	12-20	0.888	0.889	1.001 ±0.013 ± 0.001
25-35	0-1.2	12-20	0.951	0.957	1.006 ±0.004± 0.001
35-45	0-1.2	12-20	0.982	0.983	1.001 ±0.002 ± 0.001
45-60	0-1.2	12-20	0.993	0.994	1.001 ±0.001 ± 0.001
60-80	0-1.2	12-20	0.997	0.997	1.000 ±0.002 ± 0.001
>80	0-1.2	12-20	0.995	1.000	1.005 ±0.003 ± 0.001
5-15	1.2-2.4	12-20	0.773	0.824	1.066 ±0.040 ± 0.001
15-25	1.2-2.4	12-20	0.920	0.925	1.005 ±0.010± 0.001
25-35	1.2-2.4	12-20	0.965	0.970	1.004 ±0.004 ± 0.001
35-45	1.2-2.4	12-20	0.987	0.991	1.004 ±0.002 ± 0.001
45-60	1.2-2.4	12-20	0.995	0.996	1.002 ±0.001 ± 0.001
60-80	1.2-2.4	12-20	0.996	0.996	1.000 ±0.004 ± 0.001
>80	1.2-2.4	12-20	1.000	1.000	1.000 ±0.000 ± 0.001

4.5.5 Trigger efficiency

In this subsection the trigger efficiency with respect to the offline Global Muon selection, measured with the Z tag-and-probe, is discussed. Efficiencies on data are compared to the ones evaluated using the MC samples described in Section 4.5.1. The only difference with respect to what previously described, considering the selection requirements on the tag and probe muons, is that the probe muons are required to be isolated ($R_{Iso} < 0.15$). The following measurements therefore probe how efficient is the trigger in selecting muons that can be reconstructed offline as isolated Global Muons.
The measurements have been performed for the lowest un-prescaled DoubleMuon and SingleMuon isolated trigger paths of the $3 \cdot 10^{33}$ cm^{-2} s^{-1} HLT Menu: `HLT_DoubleMu13Mu8` and `HLT_IsoMu24` respectively. Only a tighter version of the DoubleMuon trigger, `HLT_DoubleMu17Mu8`, remains unprescaled in the very last part of 2011 data-taking (with $3.5 \cdot 10^{33}$ cm^{-2} s^{-1} instantaneous luminosity) and this is the trigger path used in the H$\rightarrow 4\ell$ analysis (Section 6.4).
For the trigger efficiency to be evaluated, trigger objects (L1 and HLT) need to be matched to the isolated Global Muon reconstructed offline; this is done as described in Section 4.4.2.

The efficiency has been measured separately for the five data periods, as changes in the trigger configuration occurred at the end of each period. At the end of Period 1 improvements in the GMT made the L1 efficiency increase in the $0.9 < |\eta| < 1.2$ region and changes in the CSCTF logic to better deal with the readout ambiguity of ME1/1 chambers in the L1 p_T assignement (Section 2.3.1) improved the L1 efficiency in the $|\eta| > 1.6$ region. At the end of Period 2, in order to face the increasing rate of fake muons reconstructed by the HLT algorithms in the overlap and endcap regions, tighter quality cuts had to be applied on the L2 Muons of SingleMuon trigger paths thus reducing the HLT efficiency in those pseudorapidity regions. At the end of Period 3 tighter thresholds were needed at L1 to cope with the increased rate; as a consequence the threshold on the L1 seed of the `HLT_DoubleMu13Mu8` changes from 3 GeV/c to 3.5 GeV/c and the threshold on the L1 seed of `HLT_IsoMu24` changes from 12 GeV/c to 16 GeV/c. Finally at the end of Period 4 a new improvement in the GMT logic of the p_T assignment was deployed.

Double Muon Trigger

The `HLT_DoubleMu13Mu8` is seeded at L1 by the `L1_DoubleMu3` trigger bit that requests two L1 Muon Particles with $p_T > 3$ GeV/c and Quality$\geq$3. The HLT level requires one L3 Muon with $p_T > 13$ GeV/c (that must have previously pass the L2 Muon step with $p_T > 7$ GeV/c) and one L3 Muon with $p_T > 8$ GeV/c. In the following the L1, HLT and L1+HLT efficiencies for the "Mu13 leg" of the double muon trigger are studied.
In Figure 4.19 the L1, HLT, L1+HLT efficiencies are presented as a function of the number of reconstructed primary vertices, showing a small but not negligible dependence of the HLT reconstruction algorithms to the pile-up scenario. The L1+HLT data-simuation discrepancy, which is of the order of 1-3% is period dependent as expected but it is not dependent on the pile-up conditions.
In Figure 4.19 the L1, HLT, L1+HLT efficiency at the p_T plateau are studied as a function of η of the isolated Global Muons and a comparison between the five periods is shown. The changes in the hardware and software configurations are well visible in the L1 efficiency, which is not affected by pile-up. In the HLT efficiency distribution curve it can be noted that the degradation due to pile-up is mainly present in the endcap regions, where rate and occupancy are higher.
Subdividing the events in bins of N_{vtx} to deal with the different pile-up scenario between data and simulation, the L1+HLT efficiencies are studied as a function of p_T of the isolated Global Muons and compared with simulation. In Figure 4.20 the comparison between the data in Period 5 and simulation is presented; he plateau of the L1+HLT efficiency is reached at $\sim$ 16 GeV, the turn on curve is steep due to the good L3 p_T resolution and it is well reproduced in simulation. The data-mc discrepancy at the plateau is of the order of 2% in the barrel and 3% in the endcaps.
The data-to-simulation ratios obtained comparing simulation and data (all periods together) have been computed at the p_T plateau, (20-100) GeV/c, for several η regions corresponding to main trigger detector discontinuities: $|\eta| < 0.9$ (where DT and RPC are available for the trigger),
$0.9 < |\eta| < 1.2$ (where the DT RPC and CSC are available for the trigger), $1.2 < |\eta| < 1.6$ (CSC and RPC), $1.6 < |\eta| < 2.1$ (CSC only) and $2.1 < |\eta| < 2.4$ (where CSC only can trigger and the ME1/1 chambers are involved). These numbers are reported for three different N_{vtx} ranges in Table 4.8.

Single Muon Trigger

The `HLT_IsoMu24` is seeded at L1 by the `L1_SingleMu12(16)` trigger bit, which requests one L1 Muon Particle with $p_T > 12(16)$ GeV/c and Qual-

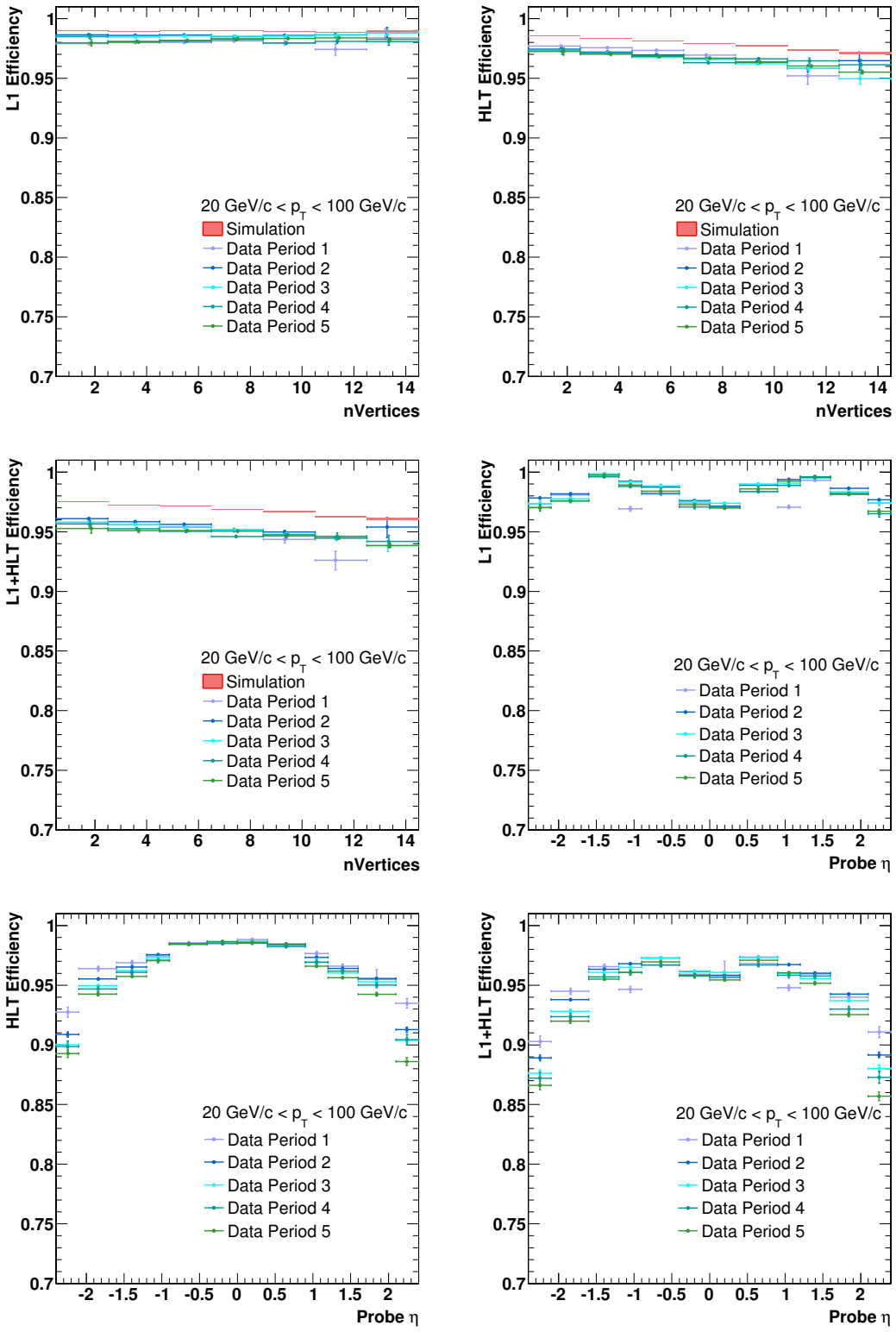


Figure 4.19: *L1, HLT and L1+HLT efficiencies for the "Mu13" leg of the DoubleMuon trigger with respect to N_{vtx} and η of the isolated Global muon. The efficiencies measured in the five data periods are compared to simulation.*

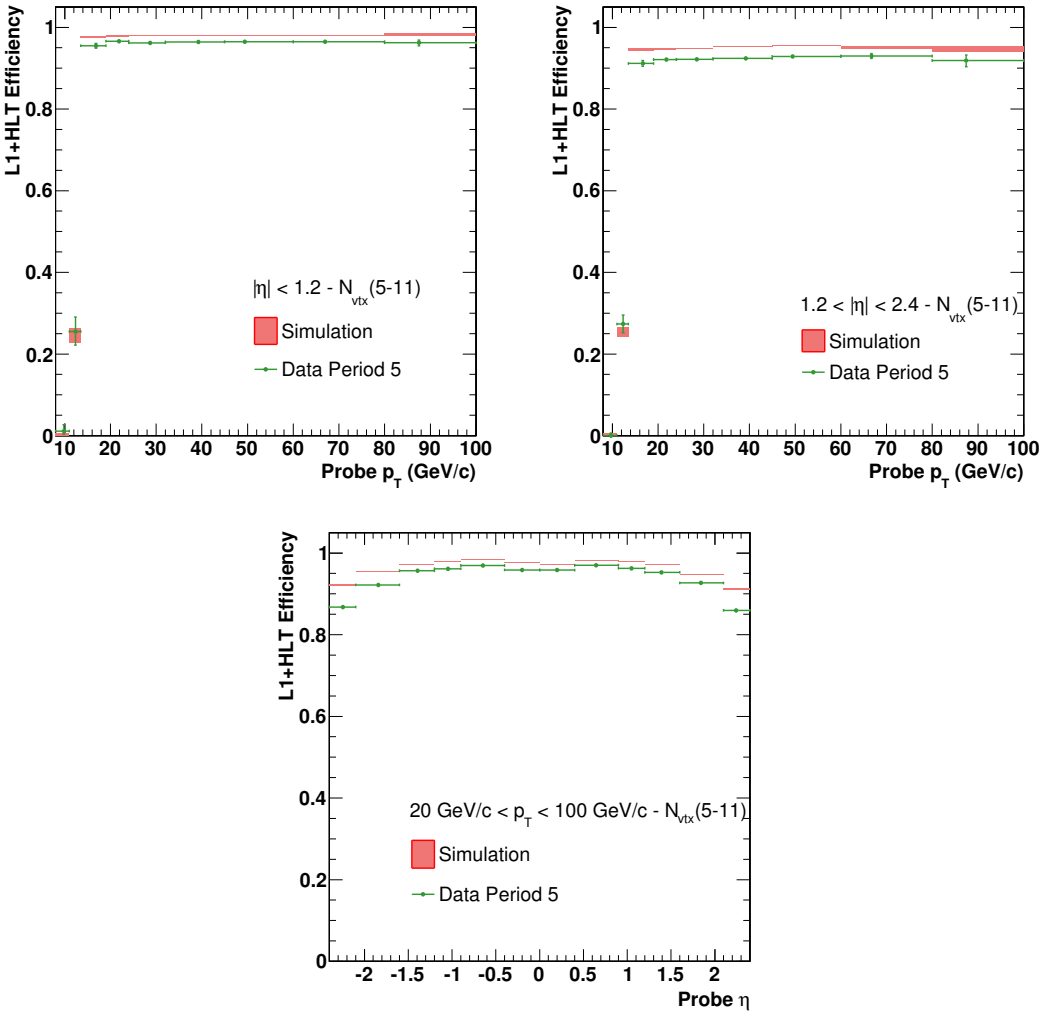


Figure 4.20: *L1+HLT efficiencies for the "Mu13" leg of the DoubleMuon trigger with respect to p_T of the isolated Global Muon for data Period 5 and simulation.*

Table 4.8: *L1+HLT efficiencies with respect to isolated Global Muon for the "Mu13" leg of the DoubleMuon trigger measured at the p_T plateau in several η regions and three different N_{vtx} ranges. The first uncertainty quoted in the last column derives from the uncertainty on the fitted efficiencies in data and simulation; the second one represents the additional systematic uncertainties described in this section.*

$\|\eta\|$ Coverage	N_{vxt}	Data (%)	MC (%)	Data/MC Ratio
0-0.9	0-5	0.981	0.967	$0.986 \pm 0.000 \pm 0.001$
0.9-1.2	0-5	0.984	0.968	$0.984 \pm 0.001 \pm 0.001$
1.2-1.6	0-5	0.977	0.964	$0.986 \pm 0.001 \pm 0.001$
1.6-2.1	0-5	0.961	0.942	$0.980 \pm 0.001 \pm 0.001$
2.1-2.4	0-5	0.935	0.902	$0.966 \pm 0.002 \pm 0.001$
0-0.9	6-11	0.980	0.966	$0.986 \pm 0.001 \pm 0.001$
0.9-1.2	6-11	0.980	0.960	$0.981 \pm 0.001 \pm 0.001$
1.2-1.6	6-11	0.972	0.958	$0.985 \pm 0.001 \pm 0.001$
1.6-2.1	6-11	0.952	0.934	$0.982 \pm 0.001 \pm 0.001$
2.1-2.4	6-11	0.918	0.880	$0.959 \pm 0.002 \pm 0.001$
0-0.9	12-20	0.978	0.960	$0.982 \pm 0.002 \pm 0.001$
0.9-1.2	12-20	0.964	0.946	$0.981 \pm 0.003 \pm 0.001$
1.2-1.6	12-20	0.963	0.946	$0.982 \pm 0.003 \pm 0.001$
1.6-2.1	12-20	0.942	0.923	$0.980 \pm 0.004 \pm 0.001$
2.1-2.4	12-20	0.879	0.842	$0.957 \pm 0.008 \pm 0.001$

ity > 3. The HLT level requires one L3 Muon with $p_T > 24$ GeV/c (that must have previously pass the L2 Muon step with $p_T > 16$ GeV/c). In Figure 4.21 the L1+HLT efficiencies are presented as a function of the number of reconstructed primary vertices, the isolated Global Muon p_T and η; a comparison between the five periods is shown. These distributions show a significant dependence of the HLT algorithms to the pile-up scenario and this is mainly due to the isolation requirements at L2 that involve the ECAL and HCAL deposits (see Section 4.5.3). The plateau efficiency of this trigger in the last data taking period is of the order of 85% for $|\eta| < 0.9$ and 75% for $|\eta| < 0.9$ mainly because of the tighter quality cuts at L1 with respect to the L1 seed of the DoubleMuon trigger and because of the L2 quality cuts applied on SingleMuon triggers in order to suppress the trigger rate.
In Figure 4.21 the L1 and HLT efficiency at the p_T plateau are also studied separately as a function of η of the isolated Global Muons. The changes in the hardware and software configurations in the L1 and HLT configurations are well identified in the results.

Systematic uncertainties

Similar conclusions as those reported in Section 4.5.3 can be drawn with the same kind of studies presented there.

4.5.6 Trigger rejection rates

The main task of the trigger is to reduce the rate of events to be recorded, while keeping high efficiency for the physics signal events that are to be studied. The previous sections focused on the trigger efficiency; in the following the fraction of minimum-bias events which is rejected as a function of trigger p_{T} threshold at Level-1, Level-2, and Level-3 is presented and compared to the prediction from minimum-bias simulation.
Only for this study of trigger rates, a data sample collected in 2010 with the highly prescaled "L1 pass-through" trigger, that only request a L1 Muon Particle with Quality > 3 and $p_{\mathrm{T}} > 7$ has been used.
The rates of accepted triggers above unprescaled p_{T} thresholds have been evaluated on a data sample that, just for this study, comes from 2010 collision and that has been collected with the highly prescaled "L1 pass-through" trigger, that only request a L1 Muon Particle with Quality > 3 and $p_{\mathrm{T}} > 7$. This sample corresponds to an integrated luminosity of 26.5 nb^{-1}. The resulting cross-sections are shown in Fig. 4.22.

To reproduce the trigger rate accurately, the simulation must reproduce the correct sample composition, the correct trigger efficiency, and also correctly describe the resolution for muons from different sources including the resolution tails. The current level of agreement between the Level-1 and HLT trigger performance in data and the Monte Carlo simulation is generally good. These and other results demonstrate that the trigger simulation is a useful tool to predict muon trigger rates and confirm that the performance of the CMS muon trigger system matches the design expectation.

4.5.7 Momentum Scale

The early data collected by CMS in 2010 are most valuable towards studying detector performance, calibrating detector subsystems and reconstruction tools. In particular the determination of the transverse momentum is highly sensitive to the precise alignment of the silicon sensors of the tracker and of the muon chambers, to the material composition and distribution inside the tracking volume, and to the detailed map of the magnetic field inside and outside the solenoid volume.
The momentum scale and resolution of muons is studied with different approaches on different p_T ranges: in the low and medium p_{T} ranges

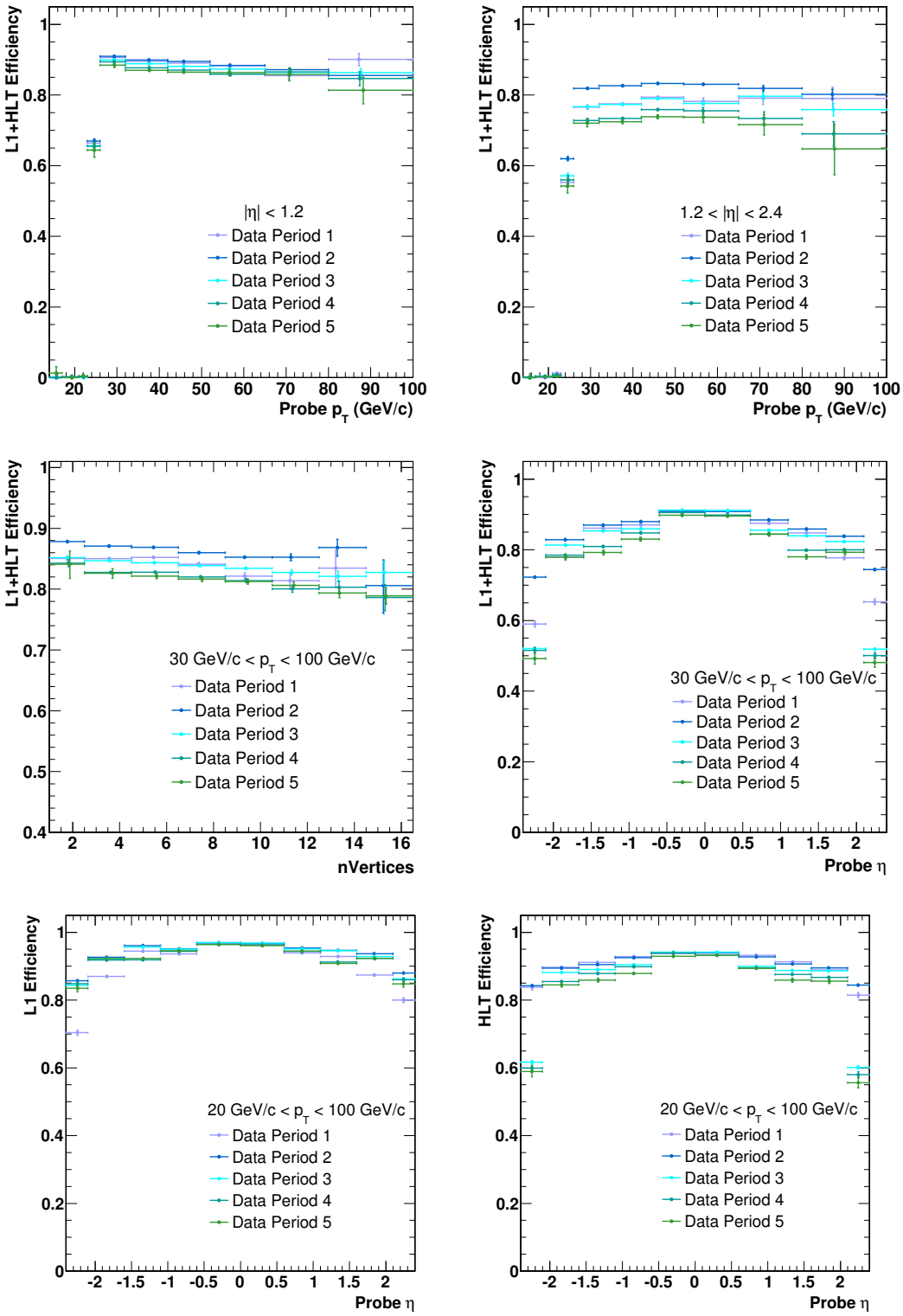


Figure 4.21: *L1+HLT efficiencies for the Single Muon Isolated trigger with respect to N_{vtx} and p_T, η of the isolated Global Muon for all the data periods. L1 and HLT efficiencies are also shown separately for the p_T plateau efficiency as a function of η of the isolated Global Muon (bottom row).*

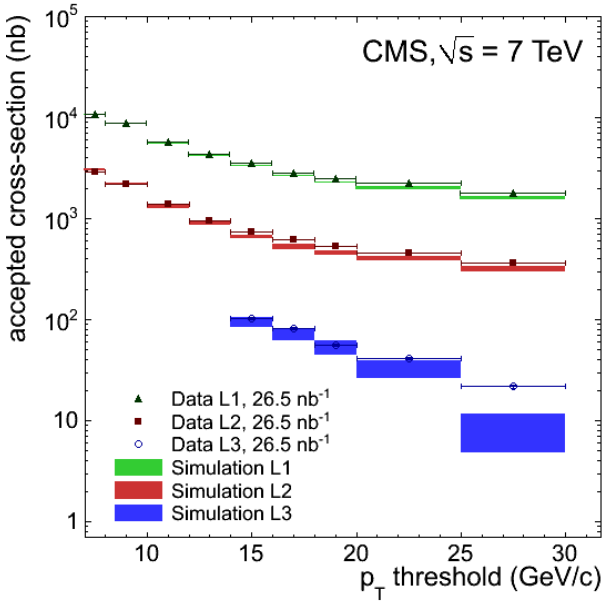


Figure 4.22: *The accepted cross-section of events as a function of trigger p_T threshold for the actual Level-1, Level-2, and Level-3 trigger objects processed online in data, compared to the emulated Level-1 and HLT trigger in simulation.*

([0-100] GeV/c) the mass constraint of di-muon decays from well known resonances J/Ψ and Z is used for the calibration. The MuScleFit (Muon momentum Scale calibration Fit) is one of the algorithm [104] used in CMS to extract absolute measurements of the momentum scale, making use of a reference model based on our best knowledge of J/Ψ and Z production at LHC.

Analyzing the mass distribution of the Z boson as a function of muon kinematics is possible to discover biases on the muon p_T in the intermediate range [20-100] GeV/c which is the interested one for the H→. The strategy exploits the position of the Z peak as returned by Voigtian fits in bins of muon η and ϕ, separately for positive and negative muons.

MuScleFit performs an unbinned likelihood fit with a reference model to correct the momentum scale, using ansatz function modeling the bias as input. Using data collected during 2010, a sinusoidal bias in the azimuthal angle on the track curvature was found and it was shown to be fully recovered by the calibration procedures. The amplitude of the azimuthal correction for a $p_T = M_Z/2$ is $\Delta(p_T) = 0.233 \pm 0.021$ (stat.)$\pm$0.023(syst.) GeV. An additional smaller bias inferior to 1% was also found versus the pseudorapidity and was also corrected, setting the final estimate of the bias found on the muon momentum scale in 2010 data to be of the order of 0.5% of the muon momentum [97].

The control of identification, isolation and trigger efficiency, as well as of the momentum scale of reconstructed muons is one of the most important requirements of the $H\rightarrow 4\ell$ analysis. In this chapter it has been shown with dedicated studies that the muon related efficiencies can be measured on data with good precision and that a general agreement with the expectation is found.
Data-to-simulation efficiency ratios have been provided in η, p_T, and N_{vtx} bins and also for different data-taking periods when necessary. These ratios can be used to correct the simulation on an event-by-event basis. In particular this will be done for the muon identification correction factors on the signal and background MC sample used for the $H\rightarrow 4\ell$ analysis. Moreover uncertainties on the correction factors will be propagated to obtain a systematic uncertainty on the final expected signal and background yields.
Additional uncertainties will be added for the data-to-simulation isolation and trigger discrepancies. A more detailed discussion on all the systematic uncertainties induced by muon measurements will be presented in Section 7.1.

Chapter 5

Physics Object: Electrons

A complete description of the CMS electron reconstruction and trigger is beyond the scope of this work; in this chapter, only the details needed for a complete comprehension of the H$\rightarrow$ 4ℓ analysis performances are presented. The electron reconstruction (cf. Section 5.1), identification (cf. Section 5.2), isolation (cf. Section 5.3) criteria as well as the electron trigger requirements (cf. Section 5.4) presented are those used in the 4ℓ analysis event selection (cf. Section 6.1).
These criteria are chosen as the best compromise between preserving the efficiency as high as possible down to low p_{T} and reducing the contamination of fake eletrons coming from the misidentification of hadronic jets. The methods to measure the identification, isolation and trigger efficiencies on real data are briefly discussed in Section 5.5.

5.1 Electron Reconstruction

The electron reconstruction combines ECAL and tracker information. Electron candidates are reconstructed from clusters of energy deposits in the ECAL, which are then matched to hits in the silicon tracker.
The standard CMS electron reconstruction algorithm is considered in this work [105, 106, 107]. The energy deposited in the electromagnetic calorimeter (ECAL) is measured in clusters of clusters (superclusters) which collect bremsstrahlung photons emitted in the tracker volume. Superclusters are used to search for hits in the innermost tracker layers which are used to seed electron tracks. This procedure is complemented by a tracker-driven approach allowing to improve the reconstruction efficiency at low p_T. Trajectories in the tracker volume are reconstructed using a dedicated modeling of the electron energy loss and fitted with a Gaussian Sum Filter. A cleaning is performed to resolve ambiguous cases where several tracks are reconstructed due to the conversion of radiated photons in the tracker material. Electron candidates are preselected using loose cuts on track-cluster

matching observables so to preserve the highest possible efficiency while removing part of the QCD background. The reconstruction efficiency for isolated electrons is expected to be above $\approx$ 90% over the full ECAL acceptance, apart from some narrow "crack" regions. Integrated over the acceptance, the reconstruction efficiency for basic electron objects steeply rises to reach $\approx$ 90% at $p_T = 10$ GeV/c, and then more slowly to reach a plateau of $\approx$ 95% for $p_T^e = 30$ GeV/c.
The four-momenta for an electrons is obtained by taking angles from the associated Gaussian Sum Filter (GSF) track, and the energy form a combination of tracker and ECAL information [105]. The information from the track is measured at the distance-of-closest approach to the associated primary vertex. The electron tracks are not re-fitted to the common vertex.

5.2 Electron Identification

In the H $\rightarrow$ ZZ $\rightarrow 4\ell$ analysis the electron candidates are required to have p_{T} larger than 7 GeV/c and a reconstructed $|\eta| < 2.5$. In addition, electrons are selected among the basic collection of reconstructed candidates using the "Cut-in-Category" technique. The electron candidates are separated into categories according to observables that are sensitive to the amount of bremsstrahlung [106]. These variables are the fraction of radiated energy as measured from the innermost and outermost state of the electron track (f_{brem}) and the ratio E/p between the supercluster energy and the measured track momentum at the vertex. Three categories are defined to separate electrons with quite different measurement characteristics and purity: "brem", "lowbrem" and "badtrack", two others categories are defined to separate electrons in "crack" and "pure tracker-driven", all of them (except the last one) are split into barrel and endcap, leading to 9 categories. The cuts are optimized to give the best signal to background ratio (s/b) for single electrons. As the shape of most discriminating variables strongly depends on the transverse energy (E_T) of the electron, cuts are made E_T-dependent. The cuts are defined for the following variables:

- $|\Delta\eta_{\mathrm{in}}| = |\eta_{\mathrm{sc}} - \eta_{\mathrm{in}}^{\mathrm{extrap.}}|$, where η_{sc} is the energy weighted position in η of the supercluster and $\eta_{\mathrm{in}}^{\mathrm{extrap.}}$ is the η coordinate of the position of closest approach to the supercluster position, extrapolating from the innermost track position and direction;

- $|\Delta\phi_{\mathrm{in}}| = |\phi_{\mathrm{sc}} - \phi_{\mathrm{in}}^{\mathrm{extrap.}}|$, where $\Delta\phi_{\mathrm{in}}$ is a quantity similar to the preceeding one but in azimuthal coordinates;

- $E_{\mathrm{seed}}/p_{\mathrm{in}}$, where E_{seed} is the seed cluster energy and p_{in} the track momentum at the innermost track position;

- H/E: ratio of energy deposited in the Hadronic Calorimeter directly behind the ECAL cluster (H) and the energy of the electron supercluster (E);

- $\sigma_{i\eta i\eta}$: supercluster η width taken from cluster shape covariance.

The cut values are listed in Table 5.1. In addition, all electrons must have no more than one expected missing inner hits in order to reject conversions, and transverse impact parameter $|d_{xy}| < 5$ mm and longitudinal impact parameter $|d_z| < 5$ mm (as described for the Global Muon selection in Section 4.2). Henceforward eID designates electrons selected according to the technique and cuts just described.

Table 5.1: *Upper thresholds used in the electron identification for electrons categories in the barrel (EB) and in the endcaps (EE). Where a range is specified the cuts are made E_T-dependent between E_T^{min}=10 GeV/c^2 and E_T^{max}=40 GeV/c^2.*

	$\|\Delta\eta_{\text{in}}\|$ [E_T^{min}-E_T^{max}]	$\|\Delta\phi_{\text{in}}\|$ [E_T^{min}-E_T^{max}]	$E_{\text{seed}}/p_{\text{in}}$	H/E [E_T^{min}-E_T^{max}]	$\sigma_{i\eta i\eta}$ [E_T^{min}-E_T^{max}]
"brem" EB	[8.92-9.23]$\times10^{-3}$	[0.063-0.069]	0.65	[0.171-0.222]	[1.16-1.27]$\times10^{-2}$
"lowbrem" EB	[3.96-3.77]$\times10^{-3}$	[0.153-0.233]	0.97	[0.049-0.052]	[1.07-1.08]$\times10^{-2}$
"badtrack" EB	[8.50-8.70]$\times10^{-3}$	[0.290-0.296]	0.91	[0.146-0.147]	[1.08-1.13]$\times10^{-2}$
"crack" EB	[13.4-13.9]$\times10^{-3}$	[0.077-0.086]	0.78	[0.364-0.357]	[3.49-4.19]$\times10^{-2}$
"brem" EE	[6.27-5.60]$\times10^{-3}$	[0.181-0.185]	0.37	[0.049-0.042]	[2.89-2.81]$\times10^{-2}$
"lowbrem" EE	[10.5-9.40]$\times10^{-3}$	[0.234-0.276]	0.70	[0.145-0.145]	[3.08-3.02]$\times10^{-2}$
"badtrack" EE	[11.2-10.7]$\times10^{-3}$	[0.342-0.334]	0.33	[0.429-0.326]	[0.99-0.98]$\times10^{-2}$
"crack" EE	[30.9-62.0]$\times10^{-3}$	[0.393-0.353]	0.97	[0.420-0.380]	[3.37-4.28]$\times10^{-2}$
"tracker-driven"	[18.8-4.10]$\times10^{-3}$	[0.284-0.290]	0.59	[0.399-0.132]	[4.40-2.98]$\times10^{-2}$

5.3 Electron Isolation

The requirement that an electron is an isolated particle in the event can be performed with the same algorithms as for muons.

- Tracker relative isolation (R_{Iso}^{Tk}). This algorithm calculates the scalar sum of the p_{T} of all tracks reconstructed in the inner tracker whose direction has a distance from the electron track direction $\Delta R \equiv \sqrt{(\Delta\varphi)^2 + (\Delta\eta)^2} < 0.3$. The p_{T} of the electron track itself is not included in the sum. For the electron to be considered isolated, the ratio of the p_{T} sum to the electron track p_{T} is required to be below a certain threshold. Track directions and values of p_{T} are computed at the point of closest approach to the nominal center of the detector.

- Tracker plus calorimeters (combined) relative isolation (R_{Iso}). The discriminating variable is similar to R_{Iso}^{Tk}, but the numerator of the ratio also includes the sum of energies measured in ECAL and HCAL towers found within a cone of radius $\Delta R < 0.3$ whose axis is taken as the ECAL or HCAL supercluster centroid viewed from the electron vertex taken at the nominal center of the detector.

The selection of tracker tracks, energy deposits in ECAL and HCAL used in isolation cones and veto regions are specified in Table 5.2. The same

Table 5.2: *List of parameters for electron isolation. $\Delta\eta$ is the difference in pseudo-rapidity with respect to the direction of the cone axis, ΔR is the radius of the veto cone, p_T is the transverse momentum of the tracks in the cone, E is the energy deposited in each ECAL rechit within the cone, $E_T = E \cdot \sin(\theta)$ is the transverse energy, ΔZ, Δr are the minimum distances from a track to the cone apex in the longitudinal and in the radial direction, respectively.*

Electron				
Type	ΔR	Deposits	Veto region	Thresholds
Tracker	0.3	CTF tracks	$\Delta\eta < 0.015$	$p_T > 0.7$ GeV/c, $\Delta Z < 0.2$ cm
ECAL	0.3	RecHits	$\Delta\eta < 1.5$ crys. $\Delta R < 1.5$ crys.	$E_T > 0.08$ GeV (EB) $E > 0.1$ GeV (EE)
HCAL	0.3	Towers	$\Delta R < 1.5$	

procedure as for muon isolation (Section 4.5.3) is used to correct the R_{Iso} variable from the effect of the pile-up [5].

5.4 Electron High Level Trigger

While the high purity of the HLT muon reconstruction allows to keep the p_T threshold very low, the contamination from fake electrons implies the need of calorimetry- and tracker-based electron ID and isolation to keep both the E_T threshold and the rate low.
A HLT electron candidate is built requiring a supercluster with E_T above a certain threshold matching an electromagnetic L1 candidate and a hit in the pixel layers of the CMS detector compatible with an electron trajectory. More requirements are added to select good electron candidates: CaloId, CaloIso, TkId and TkIso and are listed in table 5.3. The double electron trigger path that remains un-prescaled up the $3 \cdot 10^{33}$ cm^{-2} s^{-1} HLT Menu is `HLT_Ele17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL`
seeded at L1 by the `L1_DoubleEG_12_5` path.
This trigger is used in the H$\rightarrow 4\ell$ analysis as described Section 6.4 and requires two HLT electron objects with E_T thresholds 17, 8 GeV/c.

Table 5.3: *Requirements on HLT electron candidates. Values in parentheses correspond to endcaps requirements. L=Loose, T= Tight, VL=Very loose.*

Name	Cuts
CalIdL	$H/E < 0.15(0.10)$ $\sigma_{i\eta i\eta} < 0.014\ (0.035)$
CalIdT	$H/E < 0.10(0.075)$ $\sigma_{i\eta i\eta} < 0.011\ (0.031)$
CalIsoVL	$ECalIso/E_T < 0.2\ (0.2)$ $HCalIso/E_T < 0.2\ (0.2)$
TkIdVL	$d\eta < 0.01\ (0.01)$ $d\phi < 0.15\ (0.10)$
TkIsoVL	$TrkIso/E_T < 0.2\ (0.2)$

5.5 Electron Measurements with data

Also for electrons the complete control of reconstruction, identification, isolation and trigger efficiency from data has been performed using the Z tag-and-probe technique down to the extreme edge of the p_T domain that can be controlled with Z events (the minimum p_T requirement for electron in the 4ℓ analysis is 7 GeV). For this purpose the events are selected with the trigger: `HLT_Ele17_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_SC8_Mass30_v8` which requires tight identification and isolation for one electron plus only a super cluster in the events, yielding an invariant mass greater than 30 GeV/c^2. In the tag and probe algorithms the tag electron is then matched to the tight HLT electron object to avoid biasing the efficiency measurements.
By using appropriate definitions for probes, as described in Section 4.5, the overall efficiency per electron can be factorized in a series of terms, that can be measured independently:

$$\epsilon = \epsilon_{\text{RECO|clustering}} \times \epsilon_{\text{ID|RECO}} \times \epsilon_{\text{ISO|ID}} \times \epsilon_{\text{TRIGGER1Leg|ISO}} \qquad (5.1)$$

where each terms represents the efficiency for the probe to pass a given selection or reconstruction step, given that it passes the criteria for the previous one.
The clustering efficiency for electrons is assumed to be 100% efficient. This is checked to be the case in the MC within few per mille in the p_T range of interest for this analysis.
The details of the method and the results (efficiencies and data/mc correction factors) computed in η, p_T, N_{vtx} bins can be found in Ref. [5].

Electron charge mis-identification has been measured on 2010 data using

Z events and a charge mis-ID of 0.004 ± 0.001 (0.028 ± 0.003) was measured in the ECAL barrel (ECAL endcaps) in very good agreement with the simulation [108]. No significant p_T dependency was observed in the range of on-shell Z boson decays, also in agreement with the expectation.

In order to correct the electron momentum scale, calibration procedure of the ECAL response to the Z peak are performed on data. The correction function extracted with this procedure are then used to reprocess the electron reconstruction in data. On top of the re-reconstruction of 2011 data with the most update corrections the scale uncertainty on electrons have been estimated to be 0.31% (0.38%) in the ECAL barrel (ECAL endcap) [5].

The control of identification, isolation and trigger efficiency, as well as of the energy scale of reconstructed electrons, is one of the most important requirements of the H$\rightarrow 4\ell$ analysis. It has been shown [5] that electron related efficiencies can be measured on data with good precision and that a general agreement with the expectation is found.
Data-to-simulation efficiency ratios have been provided in η, p_{T}, and N_{vtx} bins, and also for different data-taking periods when necessary. These ratios can be used to correct the simulation on an event-by-event basis. In particular this will be done for the electron identification correction factors on the signal and background MC sample used for the H$\rightarrow 4\ell$ analysis. Moreover uncertainties on the correction factors will be propagated to obtain a systematic uncertainty on the final expected signal and background yields.
Additional uncertainties will be added for the data-to-simulation isolation and trigger discrepancies. A more detailed discussion on all the systematic uncertainties induced by electron measurements will be presented in Section 7.1.

Chapter 6

Event Selection and Kinematics

An optimal analysis in the 4ℓ channel (which includes the 4μ, $4e$, $2\mu 2e$ final states) must preserve the highest possible reconstruction efficiency for the Higgs boson signal while eliminating the contributions from the reducible and instrumental backgrounds (Section 3.3). This is achieved in this work relying solely on the signature of leptons. The analysis aims at the highest possible lepton reconstruction, identification and isolation efficiencies, that are compatible with a negligible reducible and instrumental background, in a transverse momentum and pseudorapidity acceptance of $p_T^e > 7$ GeV/c and $|\eta^e| < 2.5$ for electrons and $p_T^\mu > 5$ GeV/c and $|\eta^\mu| < 2.4$ for muons. The analysis strategy derives from previous detailed prospective analysis [109, 110, 111, 112, 113, 114] adapted to the expected conditions for 2010-2011 runs [115, 116] [81]. It relies on a simple sequence of cuts for the lepton identification and isolation, kinematic selection, and specific background suppression requirements (Section 6.1, Section 6.2, Section 6.3), as well as on methods relying on data for the evaluation of experimental and background systematic uncertainties (Chapter 7). The selection cuts allows for a drastic reduction of instrumental background rates. The analysis is designed for the observation of a SM-like Higgs boson in the mass range from 110 to 600 GeV/c^2. The same selection requirements and analysis are used for the first measurement of the diboson production in the $ZZ \to 4\ell$ channel.

6.1 Minimum Lepton Requirements

On the basis of the studies on lepton performances on simulation and data presented in Chapter 4 and Chapter 5, choices have been made to define the "leptons candidates" that will be used in the analysis presented in the next sections. In this work lepton candidates are defined as objects

selected as described in Section 4.2 and Section 5.2 which also satisfy a loose relative tracker isolation cut (R_{Iso}^{Tk}). Therefore a lepton candidate must satisfy these conditions:

- muons: $p_T^{\mu} > 5$ GeV/c, $|\eta^{\mu}| < 2.4$, Global Muon selection.
- electrons: $p_T^{e} > 7$ GeV/c, $|\eta^{e}| < 2.5$, eID.
- any lepton: $R_{Iso}^{Tk} < 0.7$.

The effect of the p_{T} and η cuts on the acceptance for simulated signal samples is shown in Figure 6.1 for the three final states.

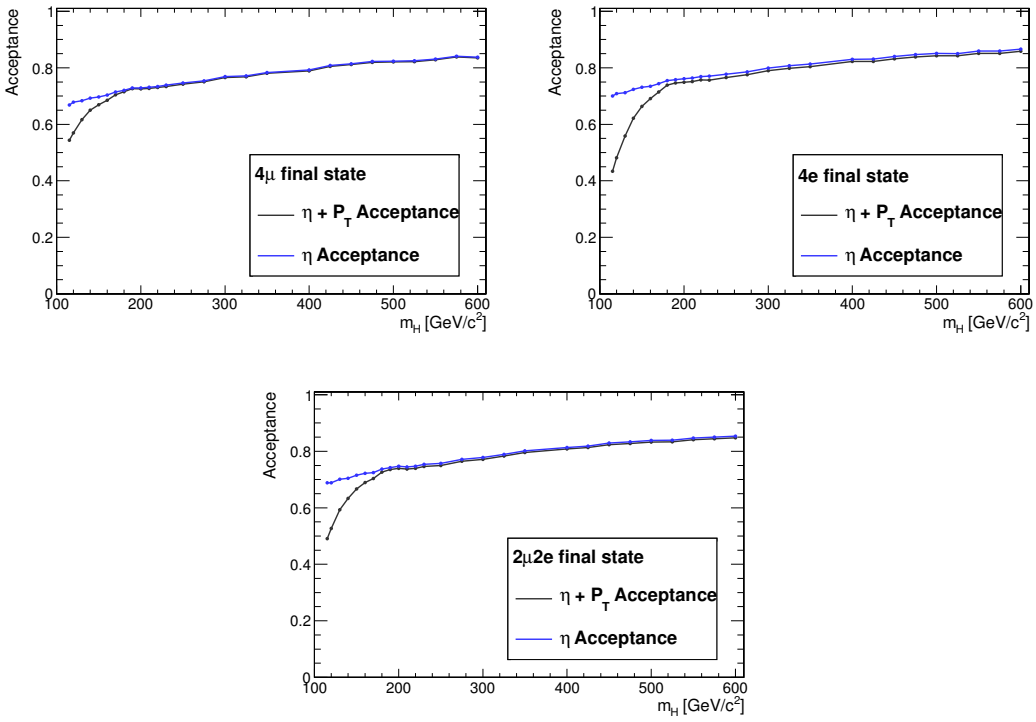


Figure 6.1: *p_{T} and η acceptance at generator level as a function of the simulated Higgs boson mass. The lepton p_{T} is understood here to be taken after final-state QED radiation.*

6.2 Preselection and choice of the 4ℓ candidate

The objective of the pre-selection is to reduce the contribution of "instrumental" backgrounds with jets faking leptons (mainly in case of electrons) and leptons from decay-in-flight of light hadrons (in case of muons), as described in Section 3.3. This is to be done while preserving the maximal

signal efficiency and the phase space to define control regions for the evaluation of background systematics.
By reducing the number of these extra leptons in signal-like events, the pre-selection allows to better solve the problem of combinatorial ambiguities caused by the presence of more than four leptons in the event.
The pre-selection is considered successful if the QCD multi-jets and Z/W + light jets contributions are brought to a level comparable to, or below, the contribution of the three main backgrounds in this analysis, namely the reducible $t\bar{t}$, $Zb\bar{b}$ and the irreducible $ZZ^{(*)}$. Moreover it should essentially only reject events which would fail the criteria imposed by the final selection (Section 6.3). It consists of a selection of the final set of four leptons through minimal identification and transverse momentum requirements, loose isolation, and kinematics.
The preselection is applied to events that have fired the relevant electron and muon triggers, consistently in data and MC (see Section 6.4) and consists in the following requirements:

1. *First Z:* a pair of lepton candidates of opposite charge and matching flavour (e^+e^-, $\mu^+\mu^-$) satisfying $m_{1,2} > 50$ GeV/c^2, $p_{T,1} > 20$ GeV/c and $p_{T,2} > 10$ GeV/c; the pair with reconstructed mass closest to the nominal Z boson mass is retained and denoted Z_1; the sum of the combined relative isolation for the two leptons of the Z_1 should satisfy $R_{iso,j} + R_{iso,i} < 0.35$; the significance of the impact parameter to the event vertex, SIP_{3D}, is required to satisfy $|SIP_{3D} = \frac{IP}{\sigma_{IP}}| < 4$ for each lepton of the Z_1.

2. *Three or more leptons:* at least another lepton candidate of any flavour or charge.

3. *Four or more leptons and a matching pair:* a fourth lepton candidate with the flavour of the third lepton candidate from the previous step, and with opposite charge.

4. *Choice of the "best 4ℓ" and Z_1, Z_2 assignments:* retain a second lepton pair, denoted Z_2, among all the remaining $\ell^+\ell^-$ combinations with $m_{Z_2} > 12$ GeV/c^2 and such that the reconstructed four-lepton invariant mass satisfies $m_{4\ell} > m_{4\ell}^{min}$. For the $4e$ and 4μ final states, at least three of the four combinations of opposite sign pairs must satisfy $m_{\ell\ell} > 12$ GeV/c^2. If more than one Z_2 combination satisfies all the criteria, the one built from leptons of highest p_T is chosen.

In the first step the first Z is chosen and all the selection requirements of the analysis, which are motivated in more detail in Section 6.3, are applied on its leptons. The p_T requirements ensure that the selected leptons are on the plateau of efficiency for the Trigger paths used (Section 6.4). The

control samples for the Z+jets, $Zb\bar{b}$ and $t\bar{t}$ backgrounds are obtained as a subset of the event sample left after this step (requiring 2 other "leptons" satisfying only loose identification criteria). The second step allows for a control of the three-lepton event rates which include mainly events from WZ di-boson production. The choice of the best combination of four leptons with $m_{4\ell} > m_{4\ell}^{min}$ completes the four steps. The value of $m_{4\ell}^{min}$ is currently set at 100 GeV/c^2.

The choice of the best 4ℓ candidate and the assignment of the leptons pairs to the Z1 and Z2 are delicate points of this analysis. Doing this at the preselection level was found to preserve the signal efficiency while providing an early rejection of background events. In particular the choice of the Z1 as the best lepton pair combination according to a mass criteria and the request that three of the four combinations of opposite sign pairs in same-flavor events must satisfy $m_{\ell\ell} > 12$ GeV/c^2 can discard events with a Z and a converted photon or events with low mass resonances decaying into leptons pairs that could survive due to wrong lepton pairing.
There are conceptually two aspects to be distinguished in these last steps: the choice of the four leptons among those available in the event and the assignment of the leptons pairs to the Z1 and Z2 bosons. For the $ZZ^{(*)}$ background and the Higgs bosons signal, it is found that when the four leptons coming from the Z decay are within the detector eta acceptance, they are selected correctly in almost 100% of the cases. In Figure 6.2 the signal efficiency for the preselection steps, relative to events within the eta acceptance at generator level, is shown for two Higgs boson masses. The same efficiency, with the additional requirement that the candidate composed by the "correct" four reconstructed leptons is retained, is also shown. "Correct" leptons are defined as those matched to the generator-level leptons deriving from the Higgs boson decay. The criteria for matching generated to reconstructed leptons are the following:

- $\Delta R < 0.5$ and $\Delta p_T / p_T < 0.5$ between the gen-object and the reco-object.
- Reco-object and gen-object must have the same charge.
- the gen-object p_{T} are intended after final-state QED radiation. As the electron reconstruction is very efficient in recollecting photons, in case the matching is unsuccessful for a given electron it is attemped with the generated electrons before radiation in the MC history.

It can be concluded that if the event is in the η acceptance at generator level, the right candidate is lost in only 0.5% of the cases. This includes cases where one of the leptons is not reconstructed or is below the p_{T} cuts, but another random lepton from the event is chosen.

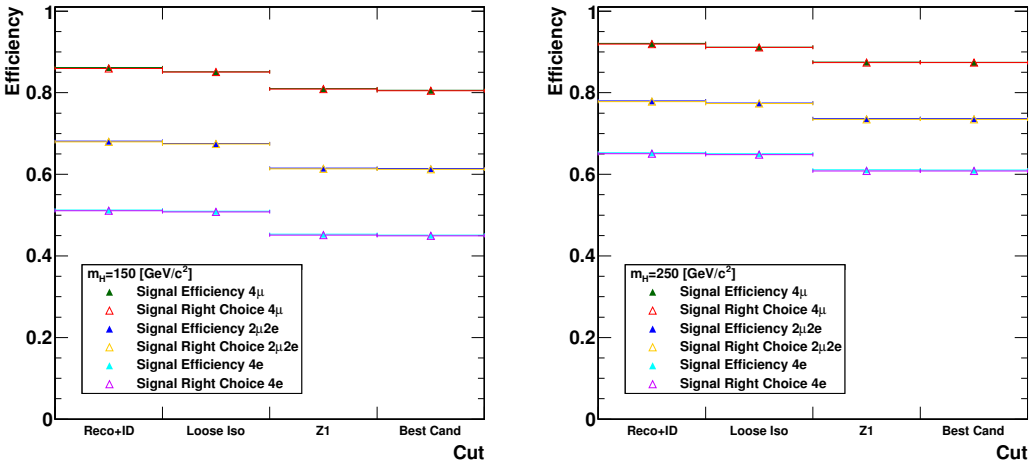


Figure 6.2: *Efficiency of the preselection steps for simulated signal samples with m_H=150 GeV/c^2 (left) and m_H=250 GeV/c^2 (right) for the three final states 4μ, 4e, 2e2μ. The efficiency is computed with respect to generated events in the η acceptance and it is compared with what is obtained if the additional requirement that the right 4 leptons candidate is preserved ("Signal Right Choice").*

For 4μ and 4e events, once the correct four leptons are chosen, the efficiency of the algorithm to find the correct pairing of the leptons to the two Z bosons (according to the MC history) is shown in Figure 6.3 as a function of the simulated Higgs boson mass. This efficiency has a dip at $m_{\mathrm{H}} \sim 180\ \mathrm{GeV}/c^2$ where the two Z bosons are forced to have similar mass, so that the choice becomes arbitrary; the decrease at low mass is due to cases where both Zs are off-shell, while the algorithm is designed to favour combinations with one of the two Z on-shell (it can be noted that the MC history pairing is purely conventional in this case).
In fact, only a fraction of these wrong assignments contribute to an efficiency loss. This fraction, i.e. how often the event is discarded by the kinematical cuts of the selection (cf. Section 6.3) after the choice of the best candidate, but would have passed if the right Z1/Z2 combination was taken, is shown Figure 6.3. These results will be commented later on.

6.3 Background Reduction Cuts

The subsequent steps further suppress the reducible backgrounds from $Zb\bar{b}/c\bar{c}$, $t\bar{t}$, and the remaining WZ + jet(s), and define the phase space for the Higgs boson signal. The focus is put on rejecting leptons coming from the decays of the b quarks.

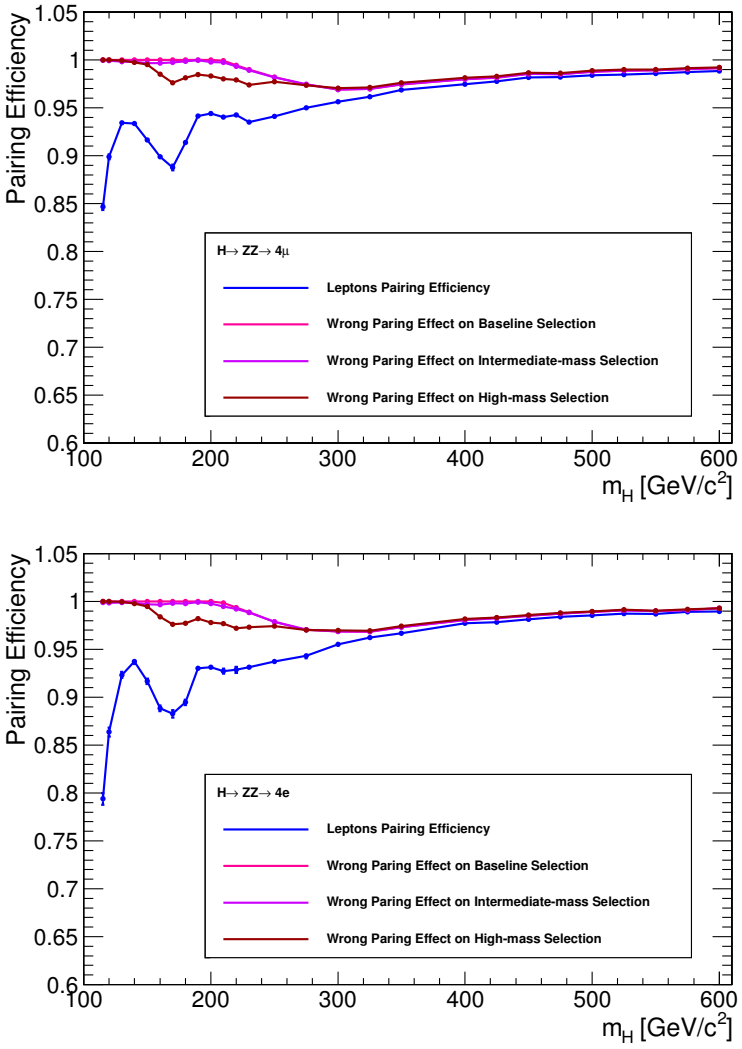


Figure 6.3: *Efficiency of right Z1, Z2 lepton pairing of the best candidate assignment logic as a function of the simulated Higgs mass; the effect of the loss of efficiency on top of the complete analysis selection criteria (Section 6.3) due to wrong lepton paring is also shown ("Wrong pairing effect").*

Such leptons are likely to be accompanied by hadronic products from the fragmentation and decay processes initiated in the b-quark jets. Moreover, because of the long lifetime of b-hadrons, they are likely to have a large impact parameter with respect to the primary vertex. Thus, lepton isolation and lepton impact parameter measurements allow for a powerful rejection.
The selection which is applied on top of the pre-selection requirements for the SM Higgs boson search with hypothetical mass in the range $m_{\mathrm{H}} > 100$ GeV/c^2 contains three different requirements which are summarized here.

1. *Relative isolation for selected leptons:* for any combination of two leptons i and j, irrespective of flavour or charge, the sum of the combined relative isolation must be $(R_{iso,j} + R_{iso,i}) < 0.35$.

2. *Impact parameter for selected leptons:* the significance of the impact parameter to the event vertex, $\mathrm{SIP_{3D}}$, is required to satisfy $\mathrm{SIP_{3D}} = |\frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}| < 4$ for each lepton, where IP is the lepton impact parameter in three dimensions with respect to the primary interaction vertex, and σ_{IP} the associated uncertainty.

3. *Z and $Z^{(*)}$ kinematics:* $m_{\mathrm{Z1}}^{min} < m_{\mathrm{Z_1}} < 120$ GeV/c^2 and $m_{\mathrm{Z2}}^{min} < m_{\mathrm{Z_2}} < 120$ GeV/c^2, where m_{Z2}^{min} and m_{Z2}^{min} are defined below.

Concerning the isolation requirement, among the several observables investigated, the one that showed the highest discriminating power, in terms of the best background rejection for a high signal acceptance, is the sum of relative isolation for selected lepton (R_{Iso}). The signal efficiency for a cut on $R_{iso,j} + R_{iso,i} < 0.35$ is 97% with respect to preselected events with m_H=150 GeV/c^2 that decays in 4μ, while it is 20% and 10% for $\mathrm{Zb\bar{b}}$ and $\mathrm{t\bar{t}}$ respectively.
The rejection power of the isolation cut after pre-selection is illustrated in Figure 6.11 that will be commented in Section 6.5.
Lepton impact parameter significance is used to build observables that can provide a relevant background rejection; the best criterion found for this channel is to require $\mathrm{SIP_{3D}}| < 4$ for all the leptons. The signal efficiency for a cut on $\mathrm{SIP_{3D}} < 4$ of each lepton is 99% with respect to preselected events with m_H=150 GeV/c^2 that decays in 4μ, while it is 20% and 10% for $\mathrm{Zb\bar{b}}$ and $\mathrm{t\bar{t}}$ respectively.
The rejection power of the impact parameter cut after pre-selection is illustrated in Figure 6.11, which will be commented in Section 6.5.
While these requirements are sufficient to eliminate the leptons from heavily boosted b-quark jets in $t\bar{t}$ events, the b-quark jets in $Zb\bar{b}$ events are in general less collimated in the detector and lead to leptons with a softer p_T spectrum. In order to best preserve the signal detection efficiency while

acting on low p_T^ℓ lepton candidates to suppress the $Zb\bar{b}$ background, the isolation criteria for a pair of leptons (generally corresponding to the pair at lowest $m_{\ell^-\ell^+}$) can be made p_T^ℓ dependent. Such criteria, which have been extensively studied in [115] is not currently used in the analysis.
Three sets of kinematic cuts have been introduced to maximize the sensitivity in different ranges of Higgs boson mass hypothesis. A **baseline** analysis is defined by requiring $m_{Z2}^{min} \equiv 12$ GeV/c^2 and $m_{Z1}^{min} \equiv 50$ GeV/c^2. This provides a best sensitivity for masses $m_H < 130$ GeV/c^2. A **intermediate-mass** analysis is defined by requiring $m_{Z2}^{min} \equiv 20$ GeV/c^2 and $m_{Z1}^{min} \equiv 60$ GeV/c^2. Finally, a **high-mass** analysis is defined by requiring $m_{Z2}^{min} \equiv 60$ GeV/c^2 and $m_{Z1}^{min} \equiv 60$ GeV/c^2.
Each definition selects a subset of the events of the previous one. The enlarged phase space of the baseline selection for the Higgs boson signal is needed at very low masses given the very small cross section $\times$ branching ratio, at the price of a larger background (see Figure 6.4). The increased acceptance for the signal becomes small ($< 10\%$ in relative compared to the baseline selection) for masses above ≈ 130 GeV/cc where reducing the phase space to better suppress the background become advantageous (see Figure 6.5). For Higgs boson masses above $\approx 2 \times m_Z$, a further restriction the phase space of the pair of Z boson can be made without significant loss of acceptance for the signal, with the benefit of a slight reduction of the ZZ^* background.

Coming back to the pairing efficiency discussed in Section 6.2, Figure 6.3 shows that the wrong pairing has a negligible impact on efficiency with the baseline and intermediate selection at low mass, while a decrease in efficiency is seen for $m_H < 200$ GeV/c^2 and it is explained by more frequent cases where one Z goes above 120 GeV/c^2 due to wrong pairing. The difference between baseline and high-mass efficiency curve is expected due to the different Z2 mass lower cut.

The signal detection efficiencies from MC for a 4ℓ system within the η acceptance as a function of the Higgs boson mass hypothesis are shown in Figure 6.6. The overall efficiency of the baseline selection is evaluated to be rising from about 72%/42%/54% at $m_H = 190$ GeV/c^2 to about 82%/59%/71% at $m_H = 400$ GeV/c^2 for the 4μ / $4e$ / $2\mu 2e$ final states.

The fraction of selected candidates on the signal samples that are made by at least one lepton not coming from the $H \to ZZ \to 4\ell$ decay is shown in Figure 6.7 for the baseline selection. The contamination mainly comes from events with τ in the final state with a per-mille contribution of three other mutually exclusive cases:

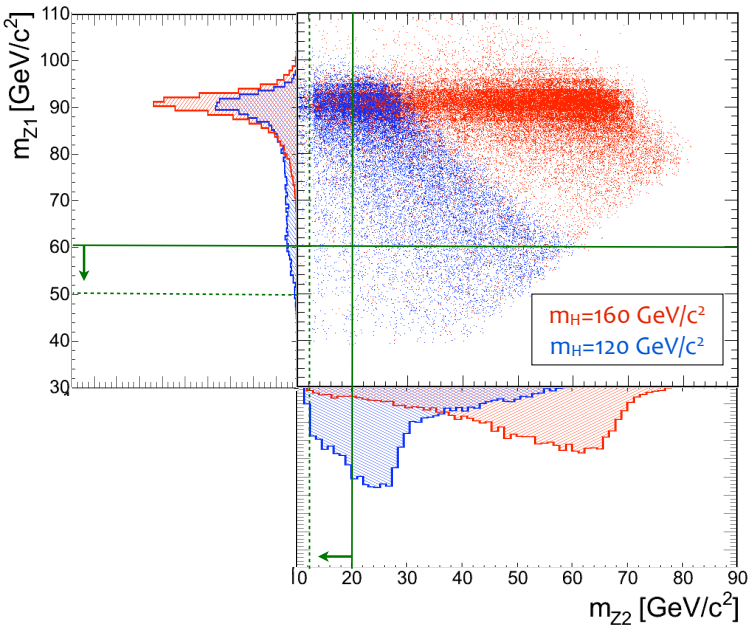


Figure 6.4: *Distribution of m_{Z1} versus m_{Z2} for two different hypotheses of Higgs boson mass: 120 and 160 GeV/c^2. The figure shows the gain of efficiency at low Higgs boson masses while relaxing the lower cuts on m_{Z1} and m_{Z2}.*

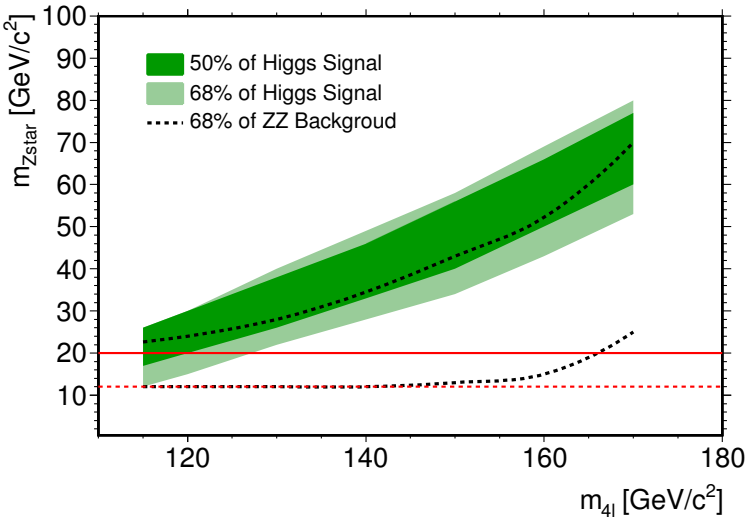


Figure 6.5: *Distribution of m_{Z2} versus m_{4l} for different Higgs boson mass hypotheses and for the ZZ^* background; the figure shows that relaxing the lower cut on m_{Z2} is not convenient for $m_H < 130$ GeV/c^2.*

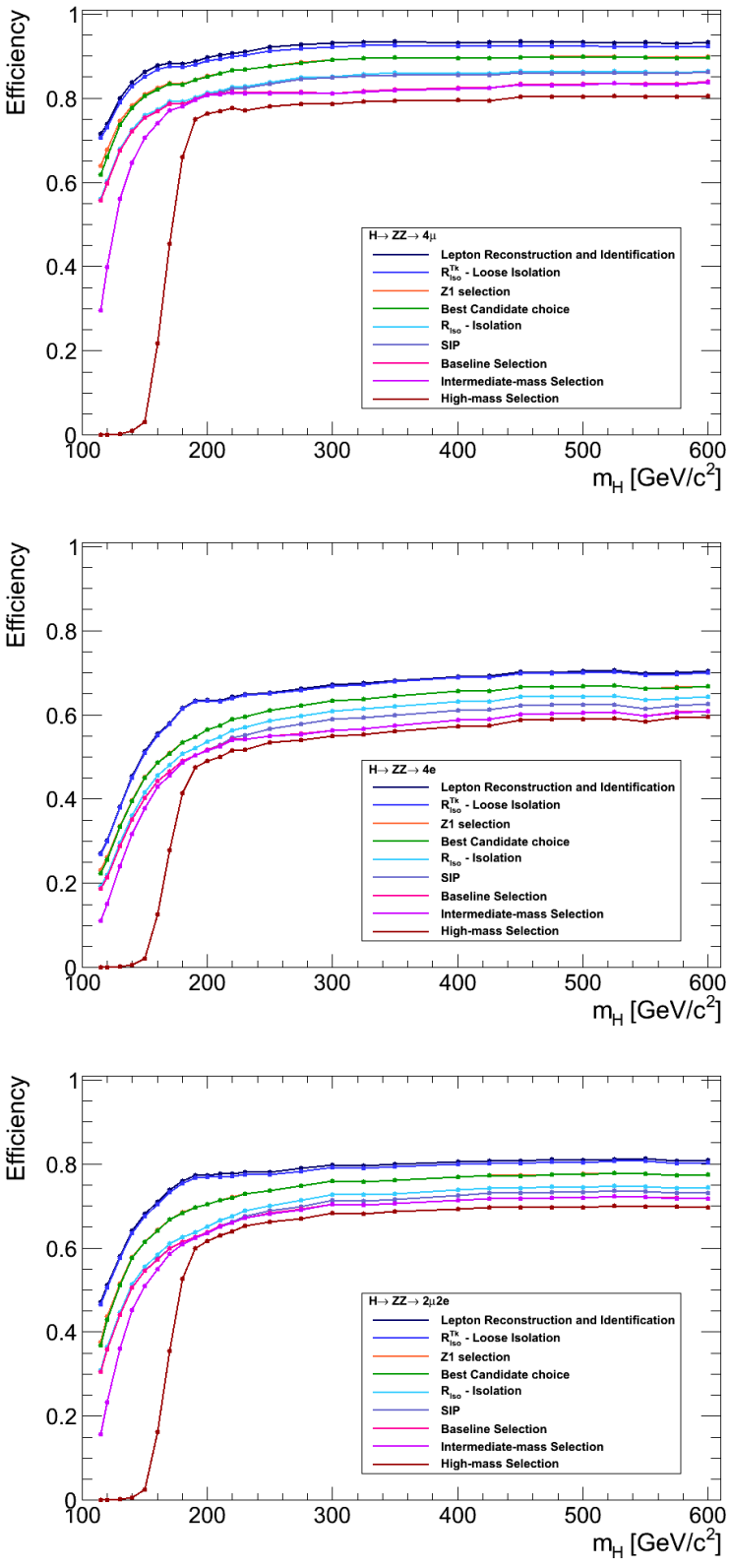


Figure 6.6: *Signal efficiency for each step of the selection as a function of the simulated m_H; the efficiency is evaluated with respect to generated events in η acceptance.*

- *Wrong leptons choice*: the "correct" leptons (defined according to the definition in Section 6.2) were available, but at least one was not chosen by the "best candidate choice algorithm".
- *Right leptons in acceptance but not reconstructed*: all the generated leptons are in the η, p_{T} acceptance, but at least one was not reconstructed.
- *Right leptons not in the acceptance*: at least one of the 4 generated leptons is out of the η, p_{T} acceptance.

The first case is a failure of the best candidate algorithm, in the last two cases the best candidate algorithm picks another background lepton instead of one that is missing.

6.4 Trigger Requirements

The trigger efficiency is expected to be very high (typically close to 100%) within the acceptance for this analysis as defined by the selection cuts described above.
The basic strategy for the analysis is to use unprescaled double-lepton paths, and to rely on the fact that, among the four leptons of the different possible final states ($4e$, 4μ and $2e2\mu$), at least two leptons are present in every reconstructed event with $p_{T,1}$ and $p_{T,2}$ comfortably above the respective thresholds for the double-lepton trigger (see Figure 6.8).

To achieve a very high trigger efficiency it is important to maintain reasonably low thresholds on the lepton triggers, and ensure careful monitoring of the trigger performances within the detector acceptance.
This basic strategy fits well for the 2011 data taking given trigger menus established for instantaneous luminosities up to $5 \cdot 10^{33}$ cm^{-2} s^{-1}. With the future increasing luminosities, the basic set of double-lepton triggers could be complemented by triple-electron paths in the $4e$ channel, and by the double- and triple-lepton cross trigger paths for the $2e2\mu$ channel.
A small dependence of efficiency on m_{H} derives from the fact that additional leptons are available in the event, besides those which are explicitly required to be above the trigger threshold. For this reason, the efficiency rises up to about $m_{\mathrm{H}} \simeq 2 \times m_{\mathrm{Z}}$ where it is compatible with 100 %.

The double-lepton trigger paths that have been presented in Section 4.5.5 and in Section 5.4 are used to selects the events for the three channels:

- 4μ channel: `HLT_DoubleMu17Mu8`
- 4e channel: `HLT_Ele17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL`

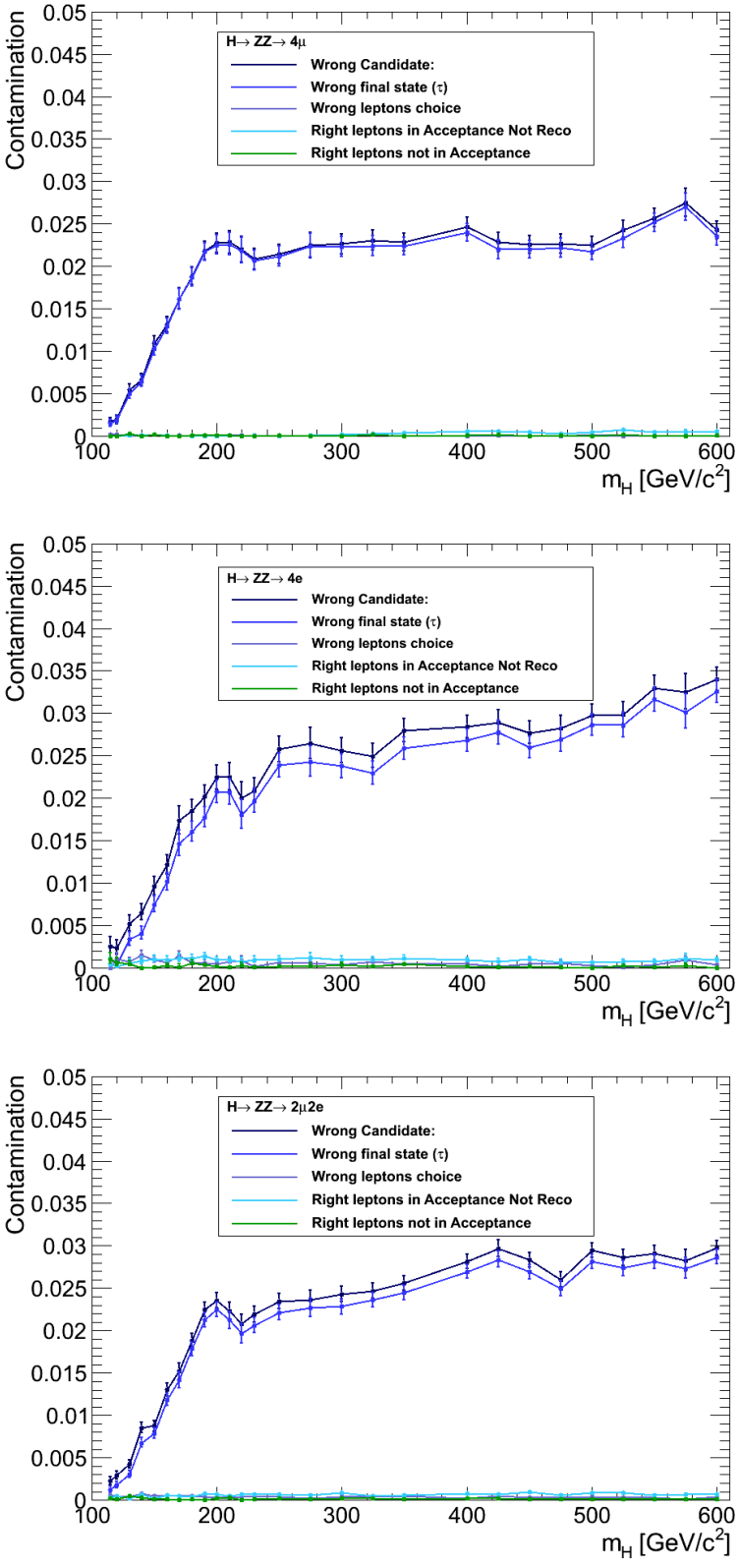


Figure 6.7: *Effect of "signal contamination": how often a reconstructed 4ℓ candidate does not match the generated 4ℓ candidate. The different sources of contamination are shown; the only significant one is the contamination from τs with electrons or muons as decay product.*

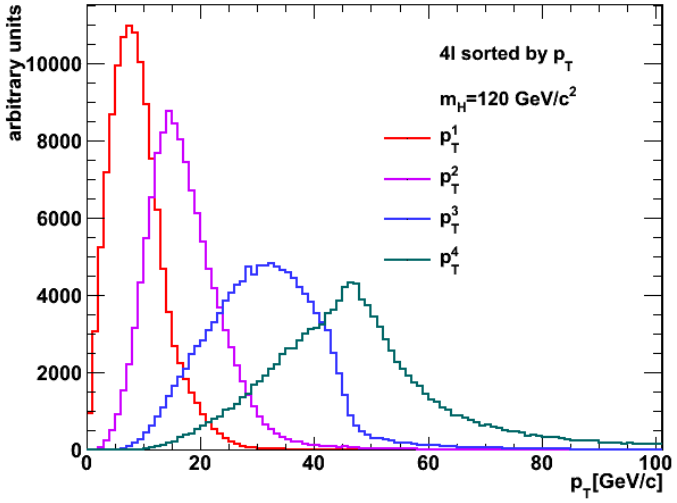


Figure 6.8: *p_T at generator level of the four muons from an Higgs of mass 120 GeV/c^2*

- $2\mu 2e$ channel: or of the above triggers.

The expected efficiencies on our signals for these paths have been evaluated on MC and, at the end of the analysis, are all higher then 98.7% as reported in table Table 6.1 and shown, for all Higgs mass hypothesis, in Figure 6.9.

Table 6.1: *Signal efficiencies for the chosen double-lepton triggers for the 4μ, $2e2\mu$, and $4e$ channels and for four different Higgs boson masses.*

Channel	$m_H = 120$	$m_H = 130$	$m_H = 150$	$m_H = 200$
4μ	99.64 ± 0.06	99.75 ± 0.04	99.90 ± 0.02	99.97 ± 0.01
$2e2\mu$	98.39 ± 0.10	98.99 ± 0.07	99.51 ± 0.04	99.77 ± 0.03
$4e$	99.70 ± 0.09	99.79 ± 0.06	99.81 ± 0.05	99.90 ± 0.03

6.5 Selection Performance

In the following the performances of the selection will be presented through data-simulation comparisons at different steps of the event selection; the data and simulated samples are those described in Section 3.5 and Section 3.4. The total integrated luminosity is 4.71 fb^{-1}.
In Figure 6.10 a first comparison is shown for the step 1 of the selection (Section 6.2). The best $Z \to e^+e^-$ or $Z \to \mu^+\mu^-$ in the event are selected

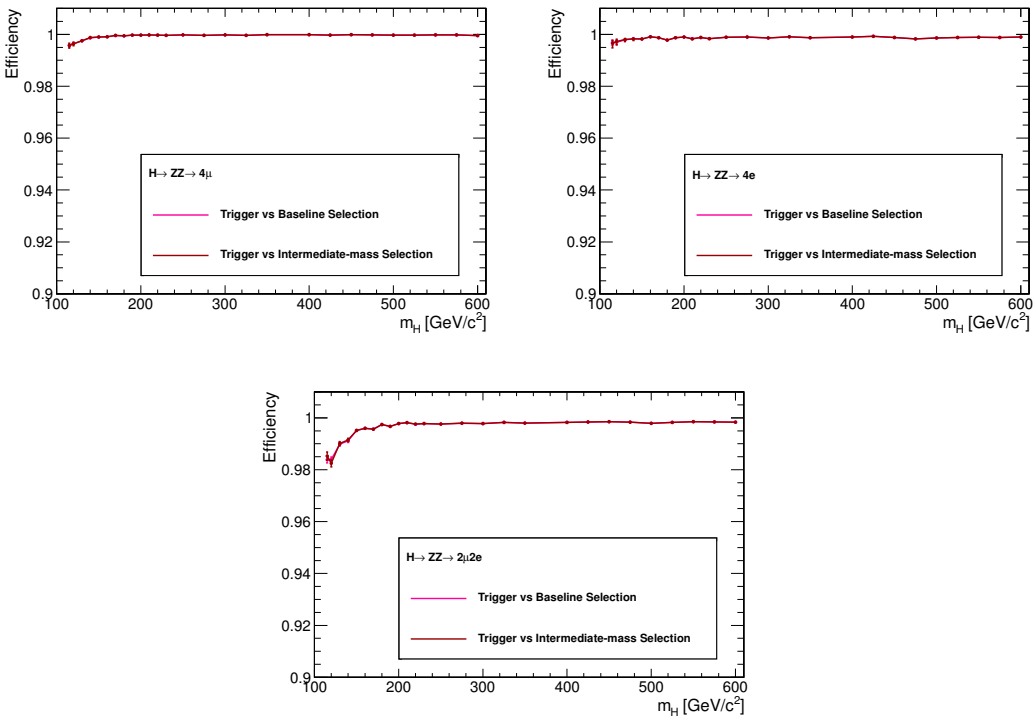


Figure 6.9: *Trigger efficiency on top of the final selection for the three final states evaluated on simulated Higgs boson samples as a function of m_H.*

from the DoubleElectron and DoubleMuon data streams. At this stage the inclusive production of single Z bosons overwhelmingly dominates the event rate.

Electron energy scale corrections, computed with dedicated studies on independent samples of $Z \to e^+e^-$ events both for 2011A and 2011B datasets have been used to correct the scale in data and in simulation; a very good agreement is obtained. The $Z \to \mu^+\mu^-$ peak is what is obtained without any muon energy scale correction applied on top of the central data reprocessing.

Going to 4μ, $4e$, $2\mu 2e$ events the low data event rate prevent from a significative data-to-simulation comparison. For this reason the control of the 4ℓ events rate is performed completely relaxing the flavour and charge requirements on the additional pair of leptons (the highest p_T ones) in $Z_1 + 2\ell$ events: the "extended phase-space" selection. In Figure 6.11: here the distribution of the reconstructed m_{Z1}, of the reconstructed mass of the additional leptons m_{Zll}, of the reconstructed four-lepton invariant mass $m_{4\ell}$, of the SIP_{3D} distribution for the lowest- and highest-SIP_{3D} additional leptons and the distribution of the sum of the R_{Iso} variable for the additional leptons are shown. In these distributions the normalization of the

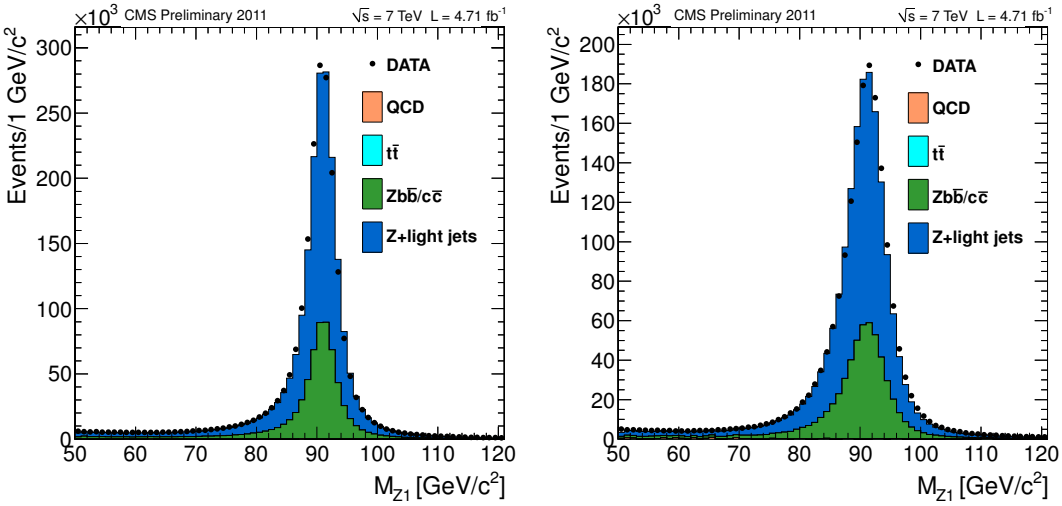


Figure 6.10: *Comparison between data and MC for the reconstructed mass m_{Z1}; the $Z \to \mu^+\mu^-$ peak is on the left and the $Z \to e^+e^-$ one on the right. The samples correspond to an integrated luminosity of $\mathcal{L} = 4.71\, fb^{-1}$.*

dominant backgrounds $t\bar{t}$ and $Zb\bar{b}/c\bar{c}$ is multiplied by scale factors obtained with dedicated control regions to extract from data the rate of these reducible backgrounds. These data-driven estimates will be presented in Section 7.2.3.
While in data the shapes of the variables that will be used for the following steps of the selection (SIP, R_{Iso}, m_{Z_2}) are well reproduced in the MC, a small excess is still observed especially in the low m_{Z2} region dominated by the reducible backgrounds $t\bar{t}$, $Zb\bar{b}/c\bar{c}$ and Z +light jets.

Figure 6.12 shows a comparison between data and MC expectation after the full pre-selection (i.e. 4ℓ candidate) after the choice of the 4ℓ for the reconstructed mass m_{Z1}, the reconstructed mass m_{Z2} and the reconstructed four-lepton invariant mass, for the $4e$ and 4μ and $2e2\mu$ channels respectively. The sample of 4ℓ events after the pre-selection contains reducible background from $t\bar{t}$ and $Zb\bar{b}/c\bar{c}$, and a possible contribution from the Z +light jets instrumental background, in particular in the $4e$ channel. A small contribution of WZ + light jet(s) background also survives, coming from events with a least one fake lepton. In this analysis, such a background is not worrying. First of all the production cross-section $\times$ branching ratio for the WZ$\to$ 3ℓ is something easily measured by CMS and furthermore with a rather clean and distinct signal topology. Moreover the production cross-section $\times$ branching fraction is only about $10\times$ larger than that of $ZZ \to 4\ell$ ($\ell = e, \mu$). Thus, the combination of the isolation and identification criteria that is applied on the additional fake lepton

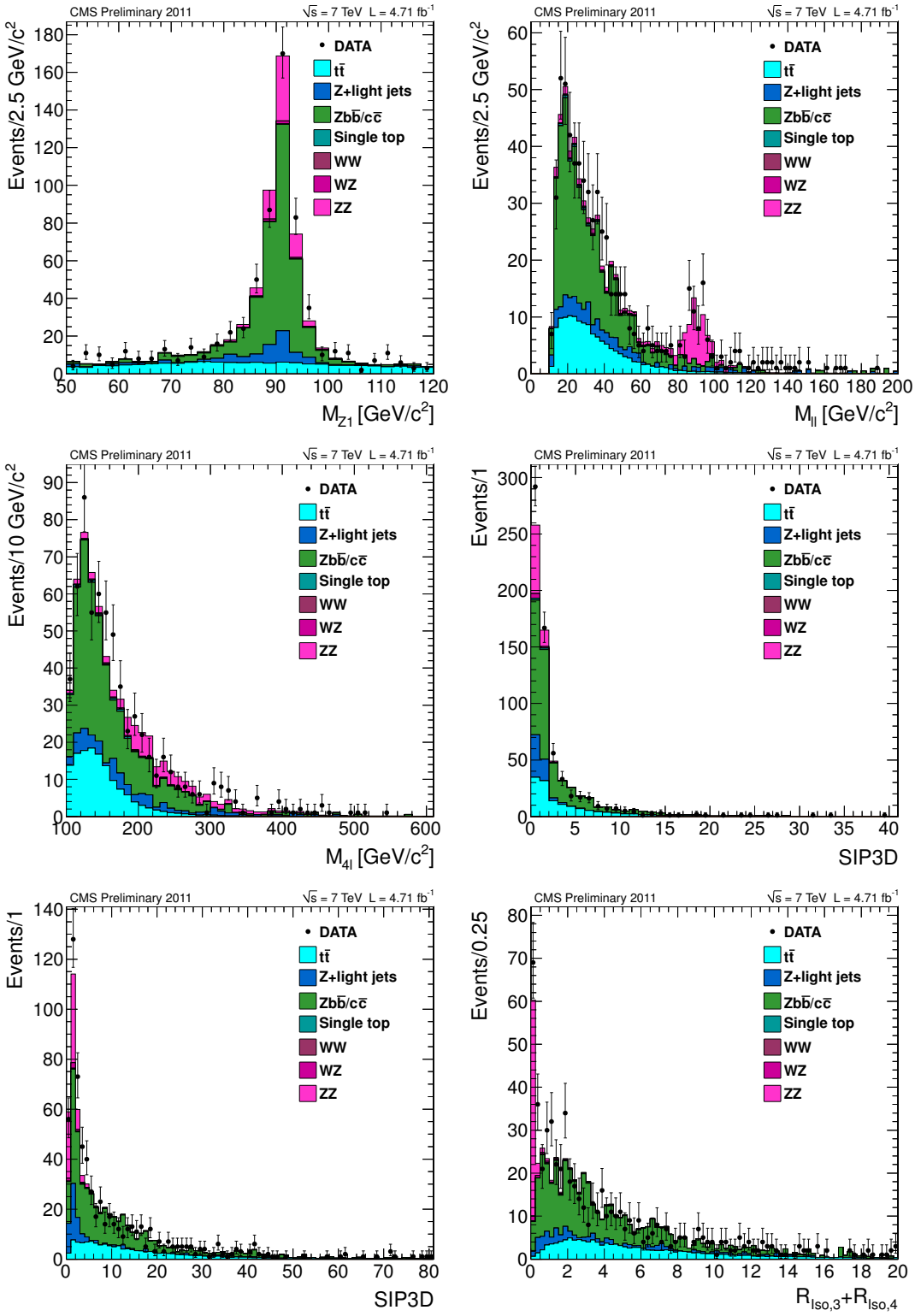


Figure 6.11: *"Extended phase-space" selection: $Z_1 + 2\ell$ events. Comparison between data and MC for the reconstructed mass m_{Z1}, the reconstructed mass of the additional leptons m_{Zll}, the reconstructed four-lepton invariant mass $m_{4\ell}$, the SIP_{3D} distribution for the lowest- and highest- SIP_{3D} additional leptons and the distribution of the sum of the R_{Iso} variable for the additional leptons. The samples correspond to an integrated luminosity of $\mathcal{L} = 4.71\ fb^{-1}$.*

candidate guaranties that this background remains small in the 4ℓ phase space.
Considering the other backgrounds, as already said, data-driven methods to control their actual rate from data are provided and described in Chapter 7.

The events yields as a function of the selection steps are shown in Figure 6.13 for the 4μ, 4e channels and in Figure 6.14 for the $2e2\mu$ channels. The MC and data yields are summarized in Table 6.2. A general good agreement is found. The properties of each of the events surviving the baseline, intermediate and high mass selection are discussed in detail in Chapter 8.

Table 6.2: *Event yields in the (a) 4μ, (b) 4e and (c) $2\mu 2e$ channel for the trigger and the seven event selection steps (see text), with steps three and four regrouped as "Presel." for the choice of the best four leptons and Z_1, Z_2 assignments. The samples correspond to an integrated luminosity of $\mathcal{L} = 4.71\ fb^{-1}$.*

(a)

Cut	QCD	tt	Z+jets	Zbb/cc	WZ	ZZ	h200	Total±	Data
HLT	1.09e+06	1.17e+04	1.52e+06	7.09e+05	688	176	33.7	3.37e+06±1.14e+04	3.62e+06
Z1	908	5.25e+03	1.43e+06	6.66e+05	582	138	28.8	2.11e+06±1.02e+03	2.13e+06
Z1 + l	-	290	1.31e+03	3.9e+03	176	47.4	9.69	5.76e+03±46	7.03e+03
Presel.	-	3.31	0.79	13.1	0.20	21.9	5.26	39±2	49
Iso	-	0.08	-	0.79	0.06	20.9	5.02	21.9±0.6	32
IP	-	0.02	-	-	0.04	20.6	4.93	20.67±0.07	26
baseline	-	0.02	-	-	0.03	19.7	4.93	19.78±0.07	23
intermediate	-	0.02	-	-	0.03	18.1	4.91	18.18±0.07	21
high-mass	-	-	-	-	0.005	15.2	4.55	15.24±0.04	14

(b)

Cut	QCD	tt	Z+jets	Zbb/cc	WZ	ZZ	h200	Total±	Data
HLT	1.12e+06	7.14e+03	1.42e+06	6.56e+05	652	164	30.9	3.25e+06±1.69e+04	4.02e+06
Z1	1.77e+04	4.65e+03	1.24e+06	5.76e+05	510	121	25.4	1.84e+06±2.3e+03	1.89e+06
Z1 + l	-	126	2.7e+03	2.1e+03	136	38	8.09	5.13e+03±44	6.12e+03
Presel.	-	0.70	0.79	1.98	0.23	14.2	3.62	18±1	25
Iso	-	0.05	-	0.40	0.11	13.5	3.43	14.1±0.4	13
IP	-	0.02	-	-	0.08	13	3.3	13.09±0.04	12
baseline	-	0.02	-	-	0.06	12.4	3.3	12.43±0.04	12
intermediate	-	0.02	-	-	0.05	11.9	3.28	11.95±0.04	12
high-mass	-	0.02	-	-	0.01	10.4	3.05	10.45±0.04	9

(c)

Cut	QCD	tt	Z+jets	Zbb/cc	WZ	ZZ	h200	Total±	Data
HLT	1.43e+06	1.97e+04	2.96e+06	1.37e+06	1.35e+03	317	56.3	5.87e+06±1.86e+04	8.01e+06
Z1	1.8e+04	9.97e+03	2.69e+06	1.25e+06	1.1e+03	237	48.1	3.97e+06±2.5e+03	4.05e+06
Z1 + l	40.7	637	5e+03	6.27e+03	298	81.5	18	1.24e+04±78	1.44e+04
Presel.	-	5.7	1.19	18.6	0.46	35.1	8.76	62±3	65
Iso	-	0.14	0.40	3.57	0.23	33	8.09	37±1	43
IP	-	0.09	-	1.59	0.15	32.1	7.84	34.0±0.8	39
baseline	-	0.05	-	1.59	0.15	30.9	7.84	32.6±0.8	37
intermediate	-	0.05	-	1.19	0.12	29	7.82	30.4±0.7	33
high-mass	-	0.02	-	0.40	0.06	25.4	7.42	25.8±0.4	30

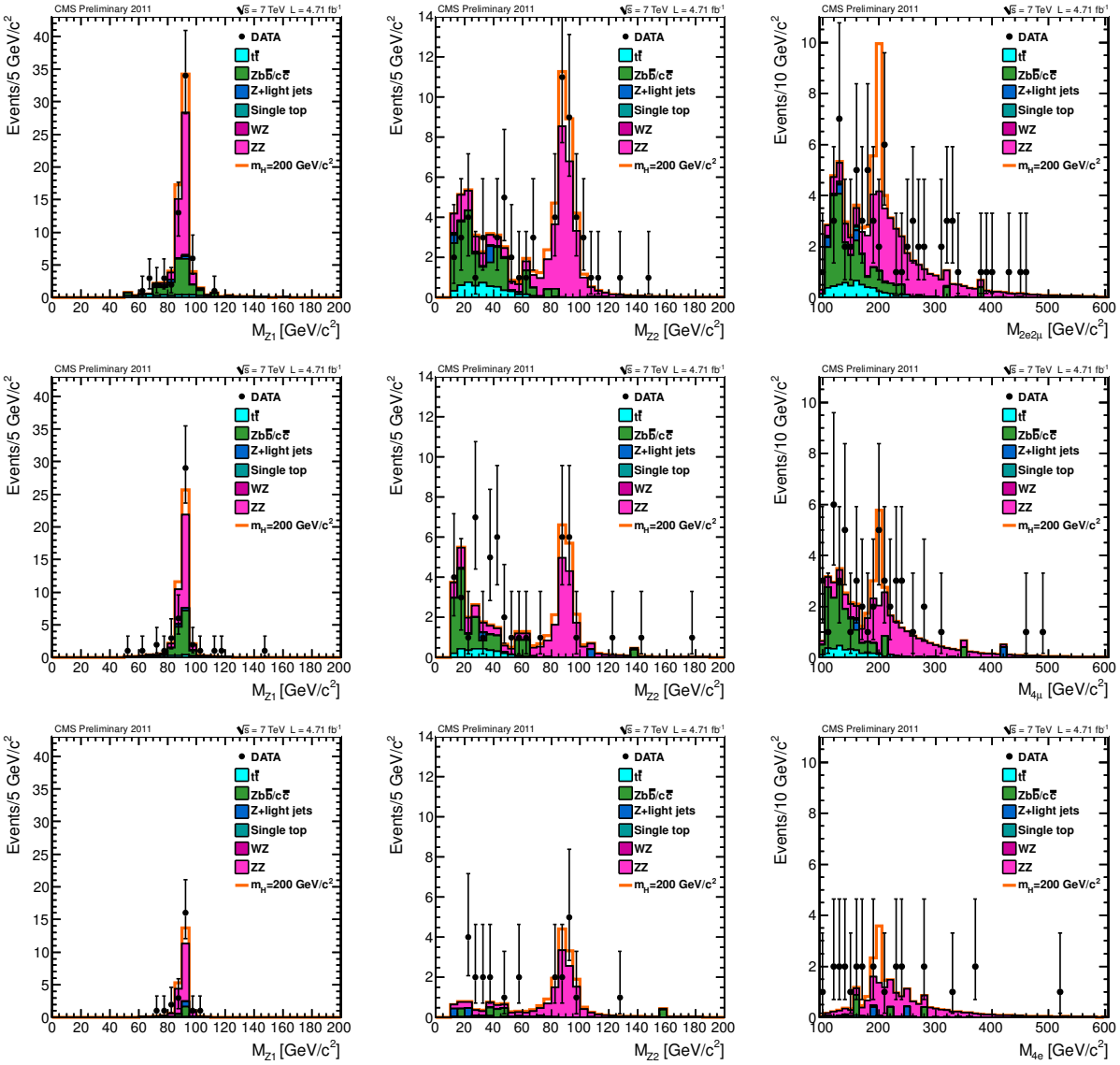


Figure 6.12: *Data-MC comparison of mass spectra after the selection of the 4ℓ candidate in the 2μ2e, 4μ and 4e channel. The comparison is shown for the reconstructed mass m_{Z_1}, the reconstructed mass m_{Z_2} and the reconstructed four-lepton invariant mass m_{ZZ}. The samples correspond to an integrated luminosity of $\mathcal{L} = 4.71\ fb^{-1}$.*

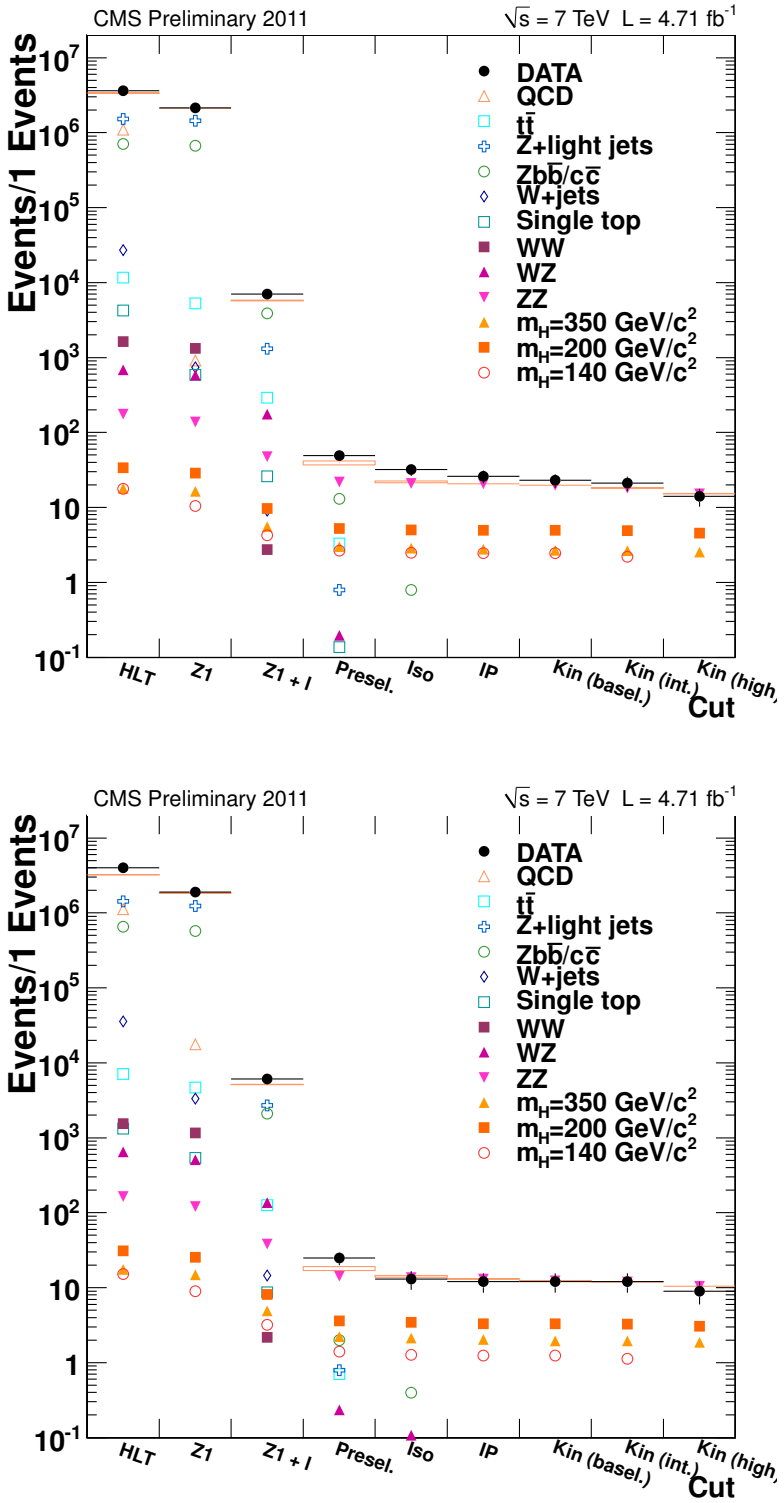


Figure 6.13: *Events yields as a function of the selection steps for the 4μ channel (top) and 4e channel (bottom). The samples correspond to an integrated luminosity of* $\mathcal{L} = 4.71\ fb^{-1}$.

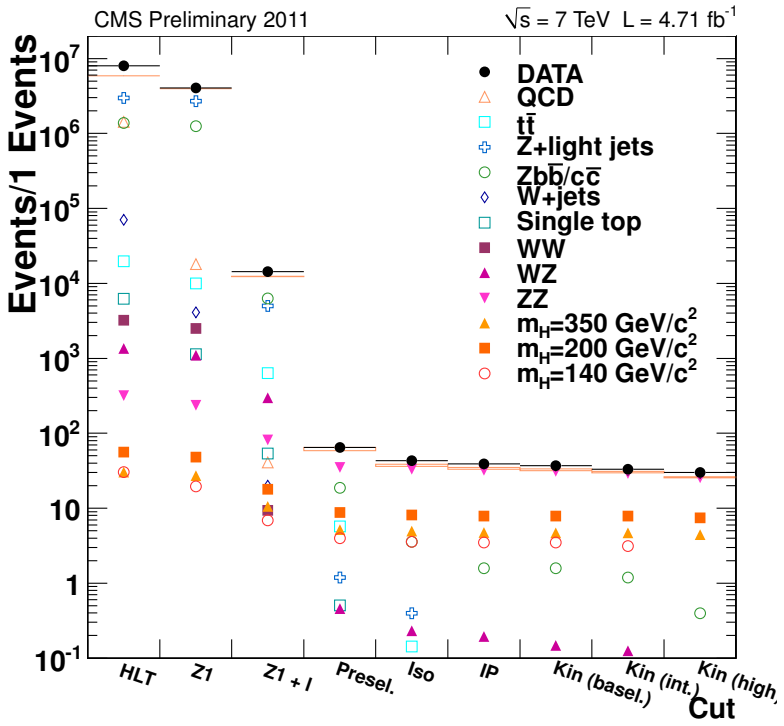


Figure 6.14: *Events yields as a function of the selection steps for the 2μ2e channel. The samples correspond to an integrated luminosity of* $\mathcal{L} = 4.71\ fb^{-1}$.

Chapter 7

Systematic Uncertainties and Background Estimation from Data

In this chapter the sources of systematic uncertainties affecting the final results as well as the methods used to estimate them are discussed (Section 7.1). Different sources of systematic uncertainties are considered: theoretical and instrumental. The effect of these uncertainties is propagated on the expected signal yields, $N^{\mathrm{H}}_{\mathrm{expected}}$, (i.e. normalization) and on the functional form that represents the event density model for the signal, $\rho^{\mathrm{H}}(m)$, (i.e. shape).
Data-driven methods to estimate the expected reducible and irreducible background yields, $N^{\mathrm{B}}_{\mathrm{expected}}$, as well as the systematics uncertainties related to these methods are then presented (Section 7.2). The $\rho^{\mathrm{B}}(m)$ functional form for the backgrounds that will be used together with the data-driven normalization for the final interpretation of the results (Section 8.4) are described throughout the chapter.

7.1 Signal systematic uncertainties

Theoretical uncertainties

Systematic uncertainties on the signal total cross section for each production mechanism and for all Higgs boson masses are fully defined elsewhere [46]. They consist in from PDF+α_s systematic errors computed using three set of PDF (CT10, MSTW08 and NNPDF) following the PDF4LHC recomendation and from theoretical uncertainties evaluated by varying QCD renormalization and factorization scales (μ_R and μ_F). The PDF+α_s and QCD scale uncertainties are treated as uncorrelated [117].
The uncertainty on $BR(H \to 4l)$ is taken to be 2% [76, 77] and assumed to

be m_H-independent.
When the Higgs boson total width Γ_H becomes very large, there are additional uncertainties related to the theoretical treatment of running Higgs width and due to non-negligible effects of the signal-background interference between $gg \to H \to ZZ$ and $gg \to ZZ$. Following the prescription given in Ref. [117], we add one more uncertainty on the Higgs boson cross sections for all sub-channels, $150\% \times (m_H/\text{TeV})^3$, intended to cover all systematic errors specific to high mass Higgs bosons.
Depending on the Higgs boson mass, the lepton kinematic cuts restrict the signal acceptance to $\mathcal{A} \sim 0.6 - 0.9$ [118]. The acceptance uncertainties $\delta\mathcal{A}/\mathcal{A}$ are evaluated by using MCFM and varying renormalization and factorization scales by a factor of two up and down. It has been shown [5] that the acceptance differences are very small (0.1-0.2%) and, therefore, can be neglected.
To estimate the effect of the harder Higgs p_T spectrum in POWHEG than in theoretical calculation at NNLL+NLO, Higgs boson events in MC have been re-weighted to make their p_T spectrum match the one obtained in HqT. The relative change in the H$\to 4\ell$ acceptance arising from the lepton kinematical cuts used in the analysis was found to be O(1%) [5], which is much smaller than the theoretical errors on the $gg \to H$ cross section, O(10%). Thus, this correction is neglected in the $H \to 4\ell$ search.

Table 7.1: *Summary of the magnitude of theoretical and phenomenological systematic uncertainties in percent for* $H \to ZZ^{(*)} \to 4\ell$ *process. Uncertainties are common to all* 4ℓ *channels.*

Source of uncertainties	uncertainties for different processes				
	ggH	VBF	WH	ZH	ttH
gg partonic luminosity	8				8-10
$qq/q\bar{q}$ partonic luminosity		2-7	3-4	3-5	
QCD scale uncert. for $gg \to H$	5-2				
QCD scale uncert. for VBF qqH		0-3			
QCD scale uncert. for VH			0-1	1-2	
QCD scale uncert. for ttH					3-1
4ℓ-acceptance for $gg \to H$	negl.	negl.	negl.	negl.	negl.
Wide Higgs uncertainties		1+1.5(m_H/1 TeV/c^2)			
Uncertainty on $BR(H \to 4\ell)$	2	2	2	2	2

7.1.1 Instrumental uncertainties

The uncertainty on the luminosity measurement is estimated as 4.5% [119].

The observed data-to-simulation discrepancy in the lepton reconstruction and identification efficiencies measured with the tag and probe data-driven techniques described in Section 4.5.2 for muons and in [5] for electrons

is used to correct the simulation on an event-by-event basis. The uncertainties on this efficiency correction are propagated independently to obtain a systematic uncertainty on the final yields and to shape parameters. Table 7.2 reports the range of yield uncertainties found for the different Higgs boson mass hypotheses.
As discussed in Section 4.5.5 the data-to-simulation discrepancies per trigger leg are of the order of 2%. Smaller discrepancies are found for the electron trigger [5]. However, since up to four leptons are available for the trigger, the overall trigger efficiency for events passing the off-line selection is very close to 100% as shown in Section 6.4, and the data/MC discrepancy per trigger leg is suppressed. Therefore the overall data/MC discrepancy in trigger efficiency for the signal is assumed to be negligible and a conservative systematic uncertainty of 1.5% on the signal normalization is assigned.

Good data-to-simulation agreement is found in isolation efficiencies as a function of η and p_{T} for a fixed isolation cut, as discussed in Section 4.5.3 for muons and in [5] for electrons. Since the isolation cut is applied on the sum of the isolation values of pairs of leptons, a systematic uncertainty on the efficiency of this cut in MC cannot be properly determined from the discrepancy of efficiencies for a fixed isolation cut. A systematic uncertainty is therefore estimated by considering the cut on the sum as a variable cut on the worst-isolated lepton of the pair, and propagating the largest data-to-simulation discrepancy observed while varying the cut in the full range $[0.0, 0.35]$ for several N_{vtx} and p_{T} bins, as illustrated for muons in Figure 4.13. The systematic on the normalization is computed for each Higgs mass hypothesis and the average results are shown in Table 7.2.

As shown in Figure 4.18 for muons, a very good data-to-simulation agreement is found in the SIP_{3D} cut efficiency at the chosen threshold SIP_{3D}=4. Therefore no systematic uncertainties have been propagated to the signal normalization and shape.

The measurement of the absolute muon and electron momentum/energy scale exploiting resonances allows an accurate calibration [120, 121]. The uncertainties for muons/electrons energy scale quoted in Section 4.5.7 and Section 5.5 are further propagated through the shape of the expected signal reconstructed mass distributions.

Table 7.2: *Summary of the magnitude of instrumental systematic uncertainties in percent for* $H \to ZZ^{(*)} \to 4\ell$*. The instrumental systematic uncertainties for all five Higgs boson production mechanisms are assumed to be same. The uncertainties assigned for the lepton reconstruction, identification and isolation apply to the event yields. The uncertainty assigned to the electron/muon scale is further propagated through the shape of the expected signal.*

Source of uncertainties	Uncertainties for different channels		
	$4e$	4μ	$2\mu 2e$
Luminosity	4.5	4.5	4.5
Trigger	1.5	1.5	1.5
electron reco/ID	3.8-1	-	2-0.5
muon reco/ID	-	2-0.8	1.2-0.4
electron isolation	2	-	1
muon isolation	-	1	1
electron E_T scale (error on E_T scale)	0.3-0.4	-	0.3-0.4
muon p_T scale (error on p_T scale)	-	0.5	0.5

7.2 Background Evaluation and Control

The total number of signal-like events surviving the baseline selection is relatively small for the current integrated luminosity, as was mentioned in Section 6.5. The small number of observed events precludes a precise evaluation of the background in a relevant narrow signal-like mass window only from the measurement of nearby side-bands. The analysis thus relies on other methods, based on experimental data, for the control of the background and the evaluation of associated systematic uncertainties.
According to the event yields evaluated from MC simulations and presented in Table 6.2, the background is overwhelmingly composed of the $ZZ^{(*)}$ continuum. Only a small contamination remains from the reducible and instrumental backgrounds. The $t\bar{t}$ and WZ backgrounds appear negligible, i.e. they both represent $\ll 1\%$ of the total background rate expected for the baseline selection. The MC event yields in Table 6.2 do not allow to conclude on the situation for Z+light jets, and $Zb\bar{b}/c\bar{c}$ backgrounds as at the end of the selection we run our of MC statistics. These backgrounds must be evaluated from data. In particular the small contamination from Z+light jets, and $Zb\bar{b}/c\bar{c}$ which is expected to be concentrated mostly at low $m_{4\ell}$ is considered.
The typical procedure to evaluate background from data, consists of choosing a wide background control region outside the signal phase space which gets populated by relaxing the event selection, and verifying that the event rates change according to the expectation from simulation. If a specific background contribution has to be determined, the corresponding control region must be chosen carefully since any of the other reducible backgrounds might rapidly become dominant if the event selection is re-

laxed, thus making the extrapolation to the signal phase space difficult.
The number of events $N^{\mathrm{B}}_{\mathrm{expect}}$ from a given background B expected in the signal region in a mass range from m_1 to m_2 can be written as:

$$N^{\mathrm{B}}_{\mathrm{expect}}[m_1, m_2] = N^{\mathrm{B}}_{\mathrm{control}} \times \left(\frac{A^{\mathrm{B}}_{\mathrm{signal}}}{A^{\mathrm{B}}_{\mathrm{control}}} \right) \times \int_{m_1}^{m_2} \rho^{\mathrm{B}}(m) dm \qquad (7.1)$$

where $N^{\mathrm{B}}_{\mathrm{control}}$ is the background rate in the control region, $A^{\mathrm{B}}_{\mathrm{signal}}$ and $A^{\mathrm{B}}_{\mathrm{control}}$ are the analysis acceptance in the "signal"-like and "background"-like regions respectively, and $\rho^{\mathrm{B}}(m)$ is the event density model for the background. The ratios between $A^{\mathrm{B}}_{\mathrm{signal}}$ and $A^{\mathrm{B}}_{\mathrm{control}}$ terms are called in the following "transfer-factors" α. The mass range considered in this analysis is the full mass range $[m_1, m_2] = [100, 600]$ GeV/c^2 covered by the baseline, intermediate and high mass selection. Detailed discussions on the definition of background control regions, extrapolations to the signal region, assumptions made on the shape of $\rho^{\mathrm{B}}(m)$ for a given background B in the signal region, and propagation of systematic errors on the various background contributions are presented in the following sections.
The question of the control of the irreducible contribution of the $ZZ^{(*)}$ continuum is discussed in Section 7.2.1. Detailed discussions on the control of the instrumental and reducible backgrounds (definition of background control regions, assumptions made on the shape of $\rho^{\mathrm{B}}(m)$ in the signal region, propagation of systematic error) are presented in Section 7.2.2 and Section 7.2.3. The background control method presented in Section 7.2.2 is inclusive and cover all the Z+X final states. The Section 7.2.3 demonstrates the possibility to further disentangle the $Zb\bar{b}/c\bar{c}$ and $t\bar{t}$ in a specific background control region, thus allowing an evaluation of the relative contributions of Z+light jets and $Zb\bar{b}/c\bar{c}$ to the different 4ℓ final states in the signal region.

7.2.1 Evalutation of the $ZZ^{(*)}$ continuum

The N^{ZZ}_{expect} for the $ZZ^{(*)}$ di-boson continuum can be directly estimated from MC prediction. Nevertheless, for the present available integrated luminosities, a data-driven method that can be used is the normalization to the measured Z rate. The method relies on the measurement of the inclusive single Z production which is used to predict the total ZZ rate within the acceptance defined by this analysis, making use of the ratio of the theoretical cross-sections for Z and ZZ production and of the ratio of the reconstruction and selection efficiencies for the 2ℓ and 4ℓ final states. In this section a comparison between the MC expectation and the data-driven one is presented.

With the ZZ/Z method, it is possible to evaluate the expected $ZZ^{(*)} \to 4\mu$, $ZZ^{(*)} \to 4e$, $ZZ^{(*)} \to 2\mu 2e$ contributions in the signal region. The contamination of $ZZ^{(*)}$ events with τ decays (included in the MC samples used in the analysis and listed in Section 3.4) will be taken from simulation.

The $ZZ^{(*)} \to 4\ell$ control region is defined here by the observed single Z inclusive rate for $Z \to \ell\ell$ so that one has for the N^{ZZ}_{control} in Eq. 7.1:

$$N^{''ZZ''}_{\text{control}} \equiv N^{Z\to\ell\ell}_{\text{obs}}. \tag{7.2}$$

The ratio of acceptance to go from the control to the signal region is then given by combining the theoretical cross-sections and the selection efficiencies as obtained from MC simulation.

$$(A^{ZZ}_{\text{signal}}/A^{ZZ}_{\text{control}}) \equiv (A^{ZZ}_{\text{signal}}/A^{Z\to\ell\ell}) = R^{\sigma}_{\text{theory}} \times R^{\epsilon}_{\text{MC}} \tag{7.3}$$

where

$$R^{\sigma}_{\text{theory}} = R^{\sigma}_{\text{theory},q\bar{q}} + R^{\sigma}_{\text{theory},gg} = \frac{\sigma^{q\bar{q}\to ZZ\to 4\ell}_{\text{NLO}}}{\sigma^{pp\to Z\to 2\ell}_{\text{NNLO}}} + \frac{\sigma^{gg\to ZZ\to 4\ell}_{\text{LO}}}{\sigma^{pp\to Z\to 2\ell}_{\text{NNLO}}} \tag{7.4}$$

and

$$R^{\epsilon}_{\text{MC}} = \frac{\epsilon^{ZZ\to 4\ell}_{\text{MC}}}{\epsilon^{Z\to 2\ell}_{\text{MC}}} \tag{7.5}$$

This ZZ/Z method has the following advantages: the statistical uncertainty on $N^{''ZZ''}_{\text{control}}$ can be considered negligible and the instrumental uncertainties on R^{ϵ}_{MC} partially cancel out as will be described in the follow. Theoretical uncertainties, however, do not cancel in the ratio ($R^{\sigma}_{\text{theory}}$), as one would naively expect [122].
The comparison between the overall uncertainty (theoretical and instrumental) on N^{ZZ}_{expect} estimated with the ZZ/Z method and with simulation will be presented in the following.

Theoretical calculation and uncertainties

The ratio of $ZZ^{(*)}$ and Z cross-sections is defined by $R^{\sigma}_{\text{theory}}$ in Eq. 7.4, and it is computed separately for the $q\bar{q}$ annihilation ($R^{\sigma}_{\text{theory},q\bar{q}}$) and gluon-gluon fusion ($R^{\sigma}_{\text{theory},gg}$) [5].
In the computation a cut on the di-lepton invariant mass to be greater than 12 GeV/c^2 is applied for each Z boson, to match the generator level cut applied in MC samples. The obtained R^{σ}_{theory} for each final states are the following:

$$\begin{array}{llll} R^{\sigma,4e}_{\text{theory},q\bar{q}} & = 1.270 \times 10^{-5} & R^{\sigma,4e}_{\text{theory},gg} & = 5.709 \times 10^{-7} \\ R^{\sigma,4\mu}_{\text{theory},q\bar{q}} & = 1.270 \times 10^{-5} & R^{\sigma,4\mu}_{\text{theory},gg} & = 5.709 \times 10^{-7} \\ R^{\sigma,2e2\mu}_{\text{theory},q\bar{q}} & = 2.404 \times 10^{-5} & R^{\sigma,2e2\mu}_{\text{theory},gg} & = 1.142 \times 10^{-6} \end{array} \tag{7.6}$$

The theoretical uncertainties on the $gg \to ZZ^{(*)} \to 4\ell$, $q\bar{q} \to ZZ^{(*)} \to 4\ell$ cross sections that affect the estimation of N^{ZZ}_{expect} from simulation are fully defined elsewhere [122]. They come from PDF+α_s systematic errors computed using three set of PDF (CT10, MSTW08 and NNPDF) following the PDF4LHC recommendation, and from theoretical uncertainties evaluated by varying QCD renormalization and factorization scales (μ_R and μ_F). The obtained four-lepton mass-dependent systematic uncertainties are summarized in Table 7.3.
Detailed studies carried out to estimate the average uncertainties on the $R^{\sigma}_{\text{theory},q\bar{q}}$ and $R^{\sigma}_{\text{theory},gg}$ that affect the estimation of N^{ZZ}_{expect} with the ZZ/Z method, show that the average PDF+α_s and QCD scale uncertainties are of the same order as those in Table 7.3 [122].

Table 7.3: *Summary of the magnitude of theoretical and phenomenological systematic uncertainties in percent for $gg \to ZZ^{(*)} \to 4\ell$, $q\bar{q} \to ZZ^{(*)} \to 4\ell$, and $pp \to Z \to 2\ell$ processes.*

Source of uncertainties	uncertainties for different processes	
	$q\bar{q} \to ZZ^{(*)} \to 4\ell$	$gg \to ZZ^{(*)} \to 4\ell$
gg partonic luminosity	-	10
$qq/q\bar{q}$ partonic luminosity	5	-
QCD scale uncert.	2-6	20-45

Selection efficiencies and instrumental uncertainties

The ratio of $ZZ^{(*)}$ and Z selection efficiencies is given by Eq. 7.5 (R^{ϵ}_{MC}). The efficiency ϵ_{MC} is defined as the ratio between events passing all selection criteria and the number of generated events with the same generator-level cut described above. These terms are computed on the corresponding MC samples described in section Section 3.4. At reconstruction level, the selection of $ZZ^{(*)}$ events fulfills all the criteria defined in Chapter 6. The selection of Z events is obtained at the first step of the event selection (Z1). The following results are obtained for the $q\bar{q}$ annihilation and gluon-gluon fusion separately and for 4μ, $4e$, $2\mu 2e$ reconstructed final states. To take into account in the computation of N^{ZZ}_{expect} also the contribution of 4μ, $4e$, $2\mu 2e$ events with leptons coming from Z leptonic τ decays, the $ZZ^{(*)}$ selection efficiencies, computed with respect to the corresponding final state at generator level, also include the contributions of these "τ events".

For the selection of $ZZ^{(*)}$ events the efficiencies are:

	baseline	intermediate	high-mass
$\epsilon_{MC,q\bar{q}}^{ZZ\to 4e}$	0.188	0.181	0.160
$\epsilon_{MC,gg}^{ZZ\to 4e}$	0.282	0.279	0.263
$\epsilon_{MC,q\bar{q}}^{ZZ\to 4\mu}$	0.303	0.277	0.234
$\epsilon_{MC,gg}^{ZZ\to 4\mu}$	0.414	0.405	0.363
$\epsilon_{MC,q\bar{q}}^{ZZ\to 2\mu 2e}$	0.252	0.236	0.206
$\epsilon_{MC,gg}^{ZZ\to 2\mu 2e}$	0.345	0.339	0.313

(7.7)

and for the selection of Z events the efficiencies are:

	baseline	intermediate	high-mass
$\epsilon_{MC}^{Z\to 2e}$	0.375	0.367	0.380
$\epsilon_{MC}^{Z\to 2\mu}$	0.434	0.425	0.425

(7.8)

Finally the ratio of $ZZ^{(*)}$ and Z selection efficiencies are:

	baseline	intermediate	high-mass
$R_{MC,q\bar{q}}^{\epsilon,4e}$	0.504	0.494	0.435
$R_{MC,gg}^{\epsilon,4e}$	0.753	0.760	0.711
$R_{MC,q\bar{q}}^{\epsilon,4\mu}$	0.698	0.654	0.553
$R_{MC,gg}^{\epsilon,4\mu}$	0.954	0.957	0.856
$R_{MC,q\bar{q}}^{\epsilon,2e2\mu}$	0.624	0.598	0.521
$R_{MC,gg}^{\epsilon,2e2\mu}$	0.855	0.860	0.794

(7.9)

It has been checked that for the $2e2\mu$ denominator taking either the arithmetic or geometrical mean between $Z \to ee$ and $Z \to \mu\mu$ efficiencies it gives the same results.

The selection efficiencies are affected by the instrumental uncertainties already discussed in Section 7.1.1. These uncertainties partially cancel out when considering the ratio R_{MC}^{ϵ}: the first one is the luminosity uncertainty. To compute the effect of lepton identification data/MC corrections on the selection efficiency ratio and their uncertainties have been propagated event-by-event considering only the leptons from the Z2. The obtained efficiency correction factor is applied to the final estimated yield, together with the corresponding systematic uncertainty.
For the uncertainties on trigger efficiency the systematic is larger for the efficiency ratio R_{MC}^{ϵ} with respect to that one for the $ZZ^{(*)} \to 4\ell$ selection efficiency. The Z and ZZ data events are being selected with Double Muon

or Double Electron triggers but in the first case only up to two leptons are available for the trigger.
For the uncertainty on lepton isolation the same procedure as for the signal MC samples has been used to compute a systematic on N^{ZZ}_{expect} estimated with the simulation. The assumption that this uncertainty affect in the same way the efficiency ratio ZZ/Z is made as usually the worse isolated leptons, on which the selection cut is applied (Section 6.3), are those from the Z2.

A summary of the instrumental uncertainties that affect the $ZZ^{(*)} \to 4\ell$ selection efficiency, and therefore the N^{ZZ}_{expect} estimated with the simulation, is shown in table Table 7.4. A summary of the of the instrumental uncertainties that affect the efficiency ratio $ZZ^{(*)} \to 4\ell/Z \to 2\ell$, and therefore the N^{ZZ}_{expect} estimated with the ZZ/Z method, is reported in the same table.

Table 7.4: *Summary of the magnitude of instrumental systematic uncertainties in percent for $ZZ^{(*)} \to 4\ell$ and for the $ZZ^{(*)} \to 4\ell/Z \to 2\ell$ ratio. The instrumental systematic uncertainties are assumed to be the same on $gg \to ZZ^{(*)} \to 4\ell$, $q\bar{q} \to ZZ^{(*)} \to 4\ell$. The uncertainties assigned for the lepton reconstruction, identification and isolation apply to the event yields. The uncertainty assigned to the electron/muon scale is further propagated through the shape of the expected signal and background reconstructed mass distributions.*

Source of uncertainties	Error for different processes					
	$qqZZ/ggZZ \to 4\ell$			$ZZ^{(*)} \to 4\ell/Z \to 2\ell$		
	$4e$	4μ	$2\mu 2e$	$4e$	4μ	$2\mu 2e$
Luminosity	4.5	4.5	4.5	-	-	-
Trigger	1.5	1.5	1.5	2	2	2
electron reco/ID	1.8	-	1.1	1.3	-	0.2
muon reco/ID	-	1.0	0.5	-	0.8	0.6
electron isolation	2	-	1	2	-	1
muon isolation	-	1	1	-	1	1
electron E_T scale (error on E_T scale)	0.3-0.4	-	0.3-0.4	0.3-0.4	-	0.3-0.4
muon p_T scale (error on p_T scale)	-	0.5	0.5	-	0.5	0.5

Measurement of Z rates from data

To extract the $Z \to \mu\mu$, $Z \to ee$ rates from data a fit to the Z1 distribution showed in Figure 6.14 has been performed. A Breit-Wigner function convoluted with a Crystal Ball function for the Z peak and an exponential function for the background have been used. Event counts for an integrated luminosity of 4.71 fb^{-1} are: $N^{Z\to^{--}}_{obs} \simeq 2.095(2.041) \times 10^6$ and $N^{Z\to ee}_{obs} \simeq 1.832(1.791) \times 10^6$ for the baseline (intermediate-mass/high-mass) selection. Statistical uncertainties for such a large number of events are negligible, while for systematic uncertainties we assume a value of 2%,

based on the estimate of the amount of background events in the selection used for the Z cross section measurement in CMS.

Results

The N^{ZZ}_{expect} and relative uncertainties extracted with the ZZ/Z method and directly from MC for an integrated luminosity of 4.71 fb^{-1} in the signal region in a mass range from 100 to 600 GeV/c^2 with the baseline, intermediate and high-mass selections are compared in Table 7.5 for the $q\bar{q}$ annihilation and and for the gluon-gluon fusion separately. The results are in agreement and the overall systematic uncertainty (theoretical and instrumental) on N^{ZZ}_{expect} is eventually of the same order. For the final interpretation of the results (Section 8.4) the MC estimate is used. The reason for this choice is that the mass-dependent computation of the theoretical uncertainties on the ZZ cross section are more accurate that the inclusive theoretical uncertainty on the cross section ratio ZZ/Z reported in
The event density $\rho^{ZZ}(m)$ (see Eq. 7.1) is obtained by a fit to the reconstructed spectrum $dN/dM_{4\ell}$ obtained from MC simulation for the phase space of either the "baseline" or "intermediate-mass" or "high-mass" selection. Two mass shapes are determined, separately for the $q\bar{q} \to ZZ^{(*)} \to 4\ell$ (Section 3.4) and $gg \to ZZ^{(*)} \to 4\ell$ (Section 3.4) contributions. The shape of the NLO 4*l* invariant mass distribution is obtain fitting the samples produced with POWHEG, with an empirical functional form [116]. The shape of the $gg \to ZZ$ 4*l* invariant mass distribution is obtained fitting the samples produced with gg2ZZ generator. The ZZ background shape fits are shown in Figure 7.1 [5].

7.2.2 Inclusive instrumental and reducible backgrounds estimation

The inclusive method here presented allows for an inclusive measurement of all the main instrumental and reducible backgrounds, that remain after the first step of the selection(Section 6.2), at the same time. This background sample is dominated by events with two prompt and isolated signal-like leptons and two additional ones from jets faking leptons or from decay of light- or heavy-flavour hadrons. The control sample is obtained as subsets of the events that satisfy the *First Z* step, requiring an additional pair of reconstructed leptons of same sign (to avoid signal contamination) and same flavour (SS-SF: $e^{\pm}e^{\pm}, \mu^{\pm}\mu^{\pm}$). The SS-SF leptons are requested to pass SIP_{3D} cut while no identification or isolation requirements are imposed. The reconstructed four-lepton invariant mass is required to satisfy $m_{4\ell} > 100$ GeV/c^2 and at least three of the four combinations of $\ell\ell$ pairs must satisfy $m_{\ell\ell} > 12$ GeV/c^2 (as in the best candidate choice). The SS-SF leptons invariant mass m_{Z_2} is required to satisfy the

Table 7.5: *Number of ZZ background events and relative uncertainties in the signal region in a mass range from 100 to 600 GeV/c^2, estimated from normalization to the measured Z rate and from simulation, for baseline, intermediate and high-mass event selections.*

baseline			
	channel	Normalization to Z rate	MC model simulation
$q\bar{q}$	$N^{ZZ\to 4\mu}$	18.2 ± 1.6	18.0 ± 1.5
	$N^{ZZ\to 4e}$	11.7 ± 1.1	11.5 ± 1.0
	$N^{ZZ\to 2\mu 2e}$	29.0 ± 2.5	28.4 ± 2.4
gg	$N^{ZZ\to 4\mu}$	1.12 ± 0.34	1. 11 ± 0.34
	$N^{ZZ\to 4e}$	0.79 ± 0.24	0.77 ± 0.24
	$N^{ZZ\to 2\mu 2e}$	1.8 ± 0.58	1.85 ± 0.56
intermediate-mass			
	channel	Normalization to Z rate	MC model simulation
$q\bar{q}$	$N^{ZZ\to 4\mu}$	16.7 ± 1.5	16.4 ± 1.4
	$N^{ZZ\to 4e}$	11.2 ± 1.0	11.0 ± 1.0
	$N^{ZZ\to 2\mu 2e}$	27.1 ± 2.4	26.6 ± 2.2
gg	$N^{ZZ\to 4\mu}$	1.10 ± 0.33	1.08 ± 0.33
	$N^{ZZ\to 4e}$	0.77 ± 0.23	0.76 ± 0.23
	$N^{ZZ\to 2\mu 2e}$	1.85 ± 0.56	1.82 ± 0.55
high-mass			
	channel	Normalization to Z rate	MC model simulation
$q\bar{q}$	$N^{ZZ\to 4\mu}$	14.1 ±1.3	13.8 ±1.2
	$N^{ZZ\to 4e}$	9.8 ±0.9	9.6 ±0.8
	$N^{ZZ\to 2\mu 2e}$	23.6 ±2.1	23.2 ±1.9
gg	$N^{ZZ\to 4\mu}$	0.98 ±0.30	0.96 ± 0.29
	$N^{ZZ\to 4e}$	0.72 ±0.22	0.71 ± 0.22
	$N^{ZZ\to 2\mu 2e}$	1.71 ± 0.52	1.68 ± 0.51

baseline, intermediate-mass or the high-mass selections. From this set of events the inclusive number of reducible background in the signal region is obtained taking into account the probability for the two additional leptons to pass the isolation and identification analysis cuts obtained form a "fake rate measurement". The fake rate is intended to estimate the probability that a "fake-lepton" can be selected as signal-like lepton. In this context the expression "fake leptons" will be used to indicate all the reconstructed leptons that are not coming from prompt isolated signal like leptons (from Z and W). In case of reconstructed electrons this category is mainly dominated by jets faking leptons, while in case of muons the contributions of real muons from decay-in-ßight of light primary hadrons and from the semileptonic decays of heavy-flavour hadrons is more important then the contribution of fake muons from the mis-identication of

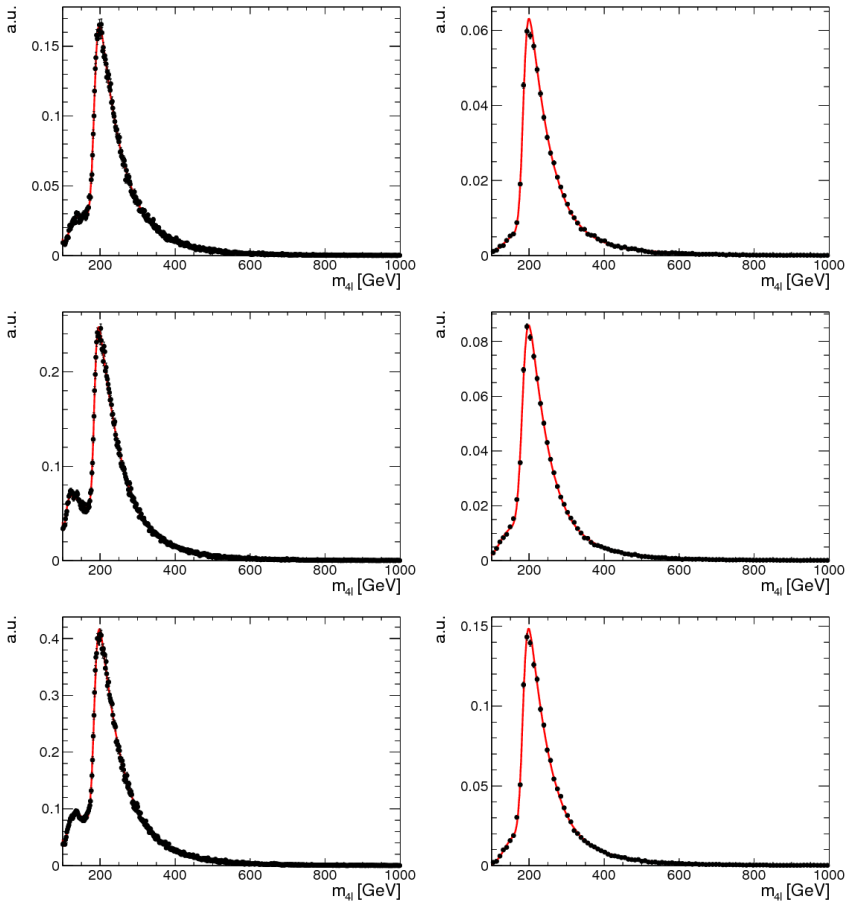


Figure 7.1: *NLO ZZ (left) and $gg \rightarrow ZZ$ (right) background shape for 4e (top), 4μ (middle), and 2μ2e (bottom) final states.*

other particles, as explained in Section 4.2.

Fake rate measurement

The fake rate measurement is obtained from an independent sample of Z_1 plus exactly one lepton (*"fakeable"* object) and where contamination from WZ events is suppressed requiring that the imbalance on the measured energy deposition in the transverse plane be below 25 GeV. The *"fakeable"* object is defined as an electron or muon with relaxed ID requirement and passing $|SIP_{3D}| < 4$. In particular a fekeable electron is a reconstructed electron without additional ID requirements (Section 5.1) and a fekeable muon is a reconstructed muon satisfying the GlobalMuon or TrackerMuon reconstruction (Section 4.1.3, Section 4.1.4) without additional ID requirements.

The fake rate is computed as the ratio of *"fakeable"* object passing the identification and isolation criteria (described in Section 6.1, Section 6.3) over the total number of *"fakeable"* objects. It is computed for each lepton flavour as a function of p_{T}, for barrel and endcaps separately and for various isolation cuts ($R_{iso} < 0.35$, $R_{iso} < 0.175$ and $0.175 < R_{iso} < 0.35$).
From the MC prediction it can be predicted that in the sample used to compute the fake rate measurements the ratio between Z+light jets and $Zb\bar{b}/c\bar{c}$ contributions is 1.7 (1.7) for electrons in the barrel (endcap) and and 1.04 (1.3) for muons in the barrel (endcap), while the overall contamination of real leptons coming from di-boson processes (WZ/ZZ) is $\ll 1\%$.
The obtained per-lepton fake rates are shown in Figure 7.2 for electrons and Figure 7.3 for muons for 2011A data. A comparison with the per-lepton fake rates measured with only the 2011B data is shown in [5]; a good agreement is found.

Extraction to the signal region

Starting from the control sample previously described and using the fake rate measurement, the final reducible background prediction in the signal region is given by the following expression:

$$\begin{aligned} N^{\mathrm{Z+X}}_{\mathrm{expect}} =& N^{\mathrm{DATA}} \times (\frac{\mathrm{OS}}{\mathrm{SS}})^{\mathrm{MC}} \times \\ & \Big[\epsilon_1(p_T,\eta)_{|R_{iso}<0.175} \times \epsilon_2(p_T,\eta)_{|R_{iso}<0.175} + \\ & \frac{1}{2} \times \epsilon_1(p_T,\eta)_{|R_{iso}<0.175} \times \epsilon_2(p_T,\eta)_{|0.175<R_{iso}<0.35} + \\ & \frac{1}{2} \times \epsilon_1(p_T,\eta)_{|0.175<R_{iso}<0.35} \times \epsilon_2(p_T,\eta)_{|R_{iso}<0.175}\Big] \end{aligned} \tag{7.10}$$

where:

- N^{DATA} is the number of events in the control region,
- $(\frac{\mathrm{OS}}{\mathrm{SS}})^{\mathrm{MC}}$ is a correction factor between opposite sign and same sign control samples obtained form the MC. It is estimated as 0.93(1.28,0.94) for the $4e(4\mu, 2e2\mu)$ final states.
- $\epsilon_i(p_{\mathrm{T}},\eta)_{|R_{iso}<k}$ is the fake rate probability for each of additional pair of reconstructed leptons ($i = 1, 2$) in function of p_{T} and η and for a defined isolation cut (k).

The signal region is a triangle in relative isolation plane $(R_{iso,1}, R_{iso,2})$ (see Figure 7.4). The expression above corresponds to an independent data-driven estimate of the background for each of the shaded areas in the figure. In particular, the first term in parenthesis of the equation 7.10 represents the probability that the two additional pair of reconstructed

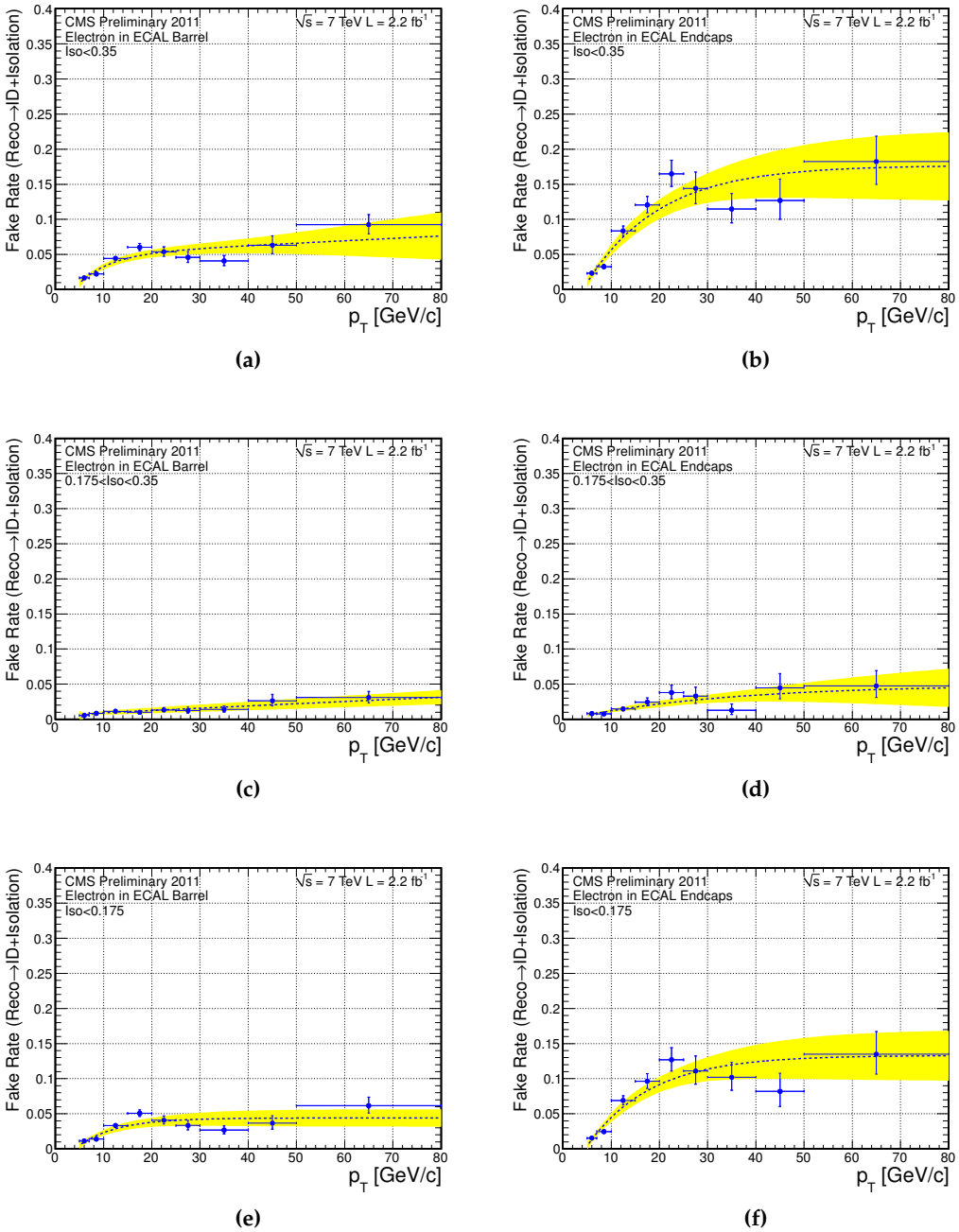


Figure 7.2: *Fake rate measurement from data per electron as a function of* p_{T} *for* $R_{iso} < 0.35$ *in the barrel (a) and endcap (b), for* $0.175 < R_{iso} < 0.35$ *in the barrel (c) and endcap (d), and for* $R_{iso} < 0.175$ *in the barrel (e) and in the endcap (f). Parametrizations are superimposed (dashed blue line) with statiscal uncertainties shown as shaded yellow band.*

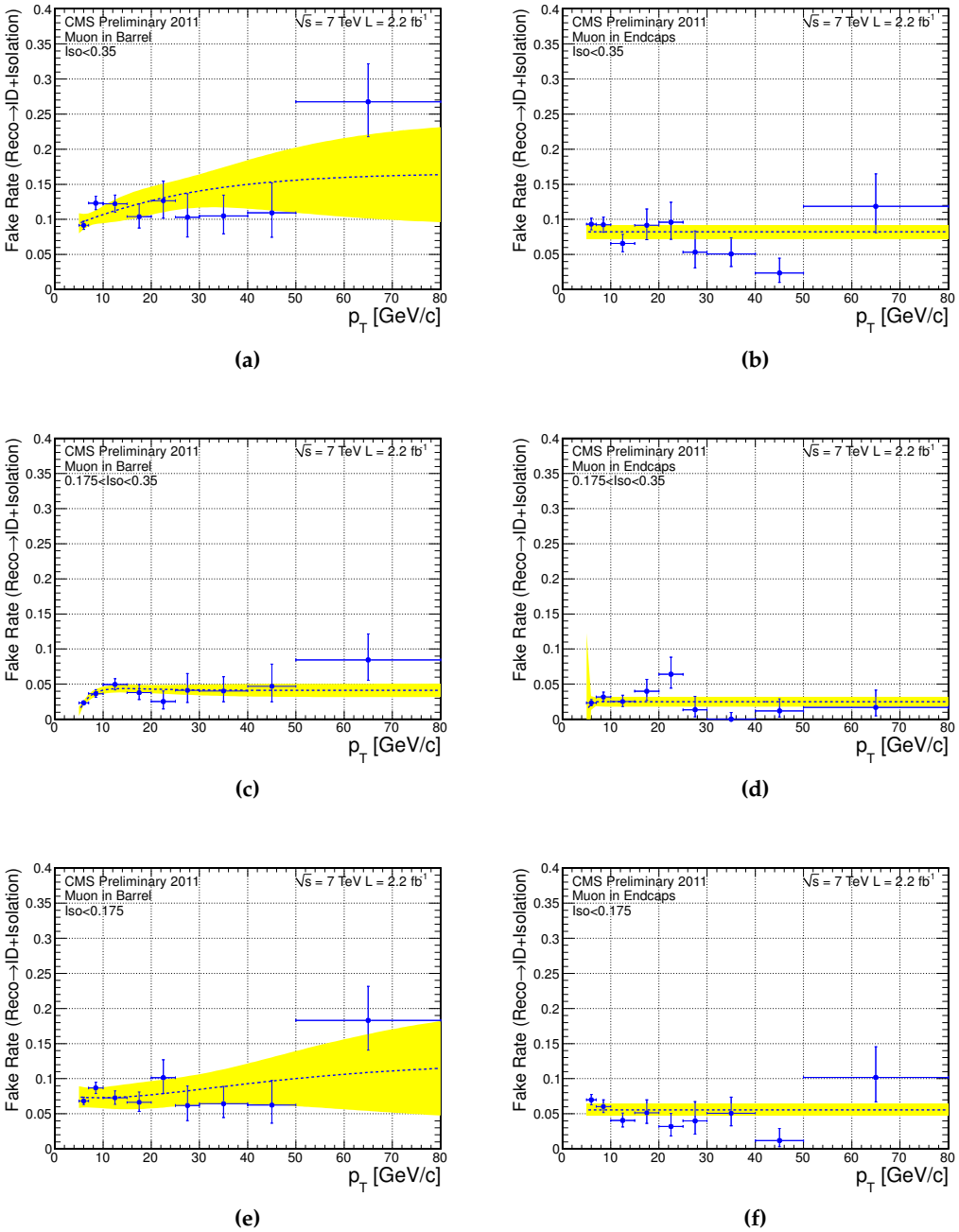


Figure 7.3: *Fake rate measurement from data per muon as a function of* p_{T} *for* $R_{iso} < 0.35$ *in the barrel (a) and endcap (b), for* $0.175 < R_{iso} < 0.35$ *in the barrel (c) and endcap (d), and for* $R_{iso} < 0.175$ *in the barrel (e) and in the endcap (f).Parameterizations are superimposed (dashed blue line) with statiscal uncertainties shown as shaded yellow band.*

leptons ($i = 1, 2$) fall in the squared labelled by A and the last two terms times the factor $\frac{1}{2}$ represent the probability that they fall in the two small triangles labelled by B_1 and B_2.

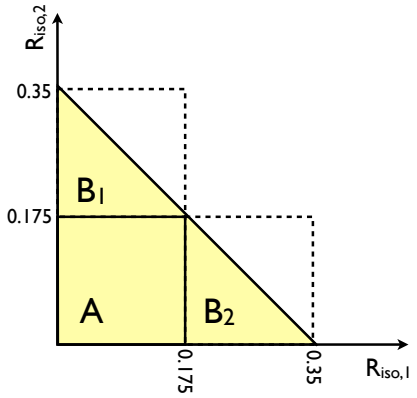


Figure 7.4: *Schematic representation of the signal region in plane $(R_{iso,1}, R_{iso,2})$ for the two least isolated leptons.*

With 4.71 fb^{-1} of collected integrated luminosity, the number of expected Z+X events and the relative systematic and statistical errors in the signal region in a mass range from $m_1 = 100$ GeV/c^2 to $m_2 = 600$ GeV/c^2 for the baseline, intermediate-mass and high-mass selections are listed in Table 7.6. The statistical error quoted represents the number of events in the control region, while the systematic one is extracted varying the paramterization of the fake rates by ± 1 sigma (yellow bands on Figure 7.2 and Figure 7.3) and inflated by 10% according to account for the WZ contribution, as will be described in the following.
For the specific case of the $2\mu 2e$ final state with baseline selection, a comparison between data and simulation in the control sample, Z1 plus an additional pair of same-sign leptons with same flavour (SS-SF), is shown in Figure 7.5 (top). In the same figure (bottom) the data-simulation comparison is also shown for a sample selected in the same way but asking for additional leptons with opposite-sign (OS-SF). A general good agreement both in shapes and rates, can be appreciated. The differences in rates in the simulation between the two regions are used to compute the correction factor in equation 7.10 for the final data-driven estimation.
Due to lack of statistic in the reducible backgrounds simulation at the end of the event selection, the event density $\rho^{Z+X}(m)$ is obtained by a fit to the reconstructed spectrum $dN/dM_{4\ell}$ obtained from MC simulation in the signal-like OS-SF control region; the shapes for the reducible background Z+X that will be used in the final interpretation of the results, obtained with the recipe just described, are shown in Figure 7.6 for the baseline selection.

Table 7.6: *The number of events from Z+X expected and the relative systematic and statistical errors in the signal region in a mass range from $m_1 = 100$ GeV/c^2 to $m_2 = 600$ GeV/c^2, for baseline, intermediate-mass and high-mass analyses.*

	baseline
$N^{Z+X\to 4e}$	1.67 ± 0.05 (3.2%) (stat., 952 events)$\pm$0.50(30.2%)(syst.)
$N^{Z+X\to 4\mu}$	1.13 ± 0.09 (8.3%) (stat., 143 events)$\pm$0.46(40.6%)(syst.)
$N^{Z+X\to 2e2\mu}$	2.71 ± 0.08 (2.9%) (stat., 1215 events)$\pm$0.88(32.6%)(syst.)
	intermediate-mass
$N^{Z+X\to 4e}$	1.45 ± 0.05 (3.7%) (stat., 746 events)$\pm$0.3(20.4%)(syst.)
$N^{Z+X\to 4\mu}$	0.81 ± 0.08 (9.8%) (stat., 103 events)$\pm$0.26(31.7%)(syst.)
$N^{Z+X\to 2e2\mu}$	2.22 ± 0.08 (3.3%) (stat., 934 events)$\pm$0.54(22.5%)(syst.)
	high-mass
$N^{Z+X\to 4e}$	0.47 ± 0.04 (8.4%) (stat., 143 events)$\pm$0.11(22%)(syst.)
$N^{Z+X\to 4\mu}$	0.22 ± 0.03 (20.8%)(stat., 23 events)$\pm$0.06(35.7%)(syst.)
$N^{Z+X\to 2e2\mu}$	0.65 ± 0.05 (7.6%) (stat., 175 events)$\pm$0.16(23.5%)(syst.)

As described in more detail in [5], a full MC closure test was also performed starting from the fake rate measurement on simulation. Due to lack of statistics in MC samples, only the $2e2mu$ channel could be tested. For this purpose, the fake rate was computed and parameterized using MC samples, the same way as in the data. The estimated reducible background extracted using formula 7.10 is compared to the rates estimated applying the full selection on MC samples. The two methods agree fairly, given the large statistical errors.
This method may not be accurate in estimating the reducible background containing three real prompt leptons and one jet faking lepton like WZ. This background is included in the control region (Z1+2 SS-SF leptons) but its contribution is underestimated when fake rates are used to extrapolate to the signal region. The fake rate is actually computed on a sample that is highly dominated by fake leptons and leptons from decay of light- and heavy-flavour hadrons. On the other hand, studies on MC show that the overall contribution of WZ to the total reducible background is small (between 5 and 10%). To the total reducible background estimated with the Z+X method is assigned an additional systematic uncertainty of 10%.

7.2.3 Evalutation of the $Zb\bar{b}/c\bar{c}\to 4\ell$ and $t\bar{t}\to 4\ell$ backgrounds

A contamination of $Zb\bar{b}/c\bar{c}$ and $t\bar{t}$ events remains after pre-selection (Section 6.2), and a small amount of such events will survive the final selection (Section 6.3), especially in the low m_H range. These contaminations are estimated inclusively with the Z+X method, but a dedicated strategy to control these processes independently and compute their contributions in

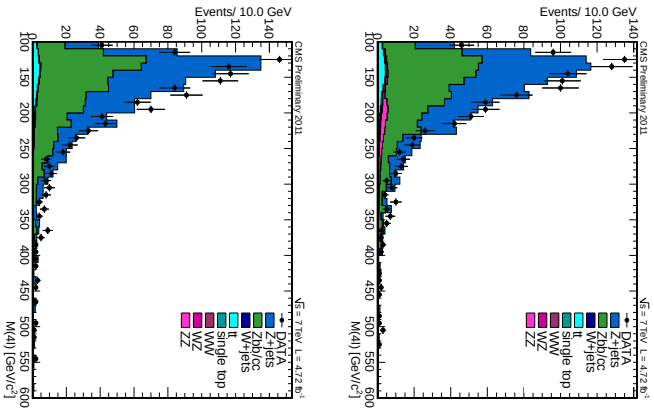


Figure 7.5: *Comparison between data and MC of the reconstructed four-lepton invariant mass $m_{4\ell}$ in the Z+X background control region. SS-SF control sample (top), OS-SF control sample (bottom). The samples correspond to an integrated luminosity of $\mathcal{L} = 4.71\, fb^{-1}$.*

the signal-region is presented in this section.

For the measurement of the $Zb\bar{b}/c\bar{c}$ and $t\bar{t}$ rates, a four-lepton background control region is defined in the following manner: on top of the "extended phase-space selection" (defined with the request of the Z1 plus two additional leptons without flavour or charge requirements and no vertex or isolation constraints, as described in Section 6.5) the request that the two additional highest-p_{T} leptons must have SIP3D > 5 is applied. The extended phase-space selection is needed to increase the event rate while the SIP_{3D} cut is needed to ensure a negligible contribution of the other reducible backgrounds (Z+light jets) in the control region, and to define a signal-free phase-space.
A comparison of the simulation expectation with data in the control region is shown for the m_{Z1} distribution in Figure 7.7a, Figure 7.7c, Figure 7.7d. As can be inferred from these plots, the $Zb\bar{b}/c\bar{c}$ and $t\bar{t}$ contributions can be easily disentangled from the distribution of the reconstructed invariant mass of the Z candidate; the current statistics is sufficient to allow for a fit of the resonant and non resonant contributions. An unbinned fit has been performed using a Breit-Wigner function convoluted with a Crystal Ball function for the Z peak and using Chebychev Polynomials for the $t\bar{t}$ spectrum (Figure 7.7b).

With 4.71 fb^{-1} the control region contains 235 events, out of which 167 belong to the Z1+$\mu\mu'$ category, 58 to the Z1+$e\mu$ category and 10 events

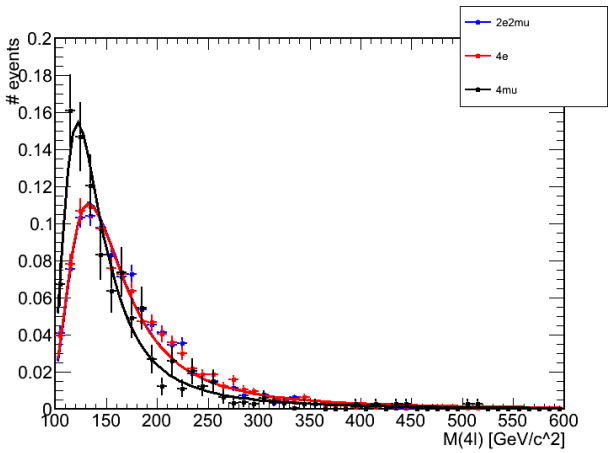


Figure 7.6: *The shape for Z+X reducible background for the 4e, 4μ and 2e2μ final state.*

are observed in the Z1+ee' category. The results of the fit procedure is the following: 151±17 $Zb\bar{b}/c\bar{c}$ events, 84±15 $t\bar{t}$ events. Comparing with simulation we derive data/MC scale factors in the control region of 1.16±0.14 for $Zb\bar{b}/c\bar{c}$ and 1.06±0.22 for $t\bar{t}$. These correspond to the factors that have been applied for the control of the Z1+2ℓ extended phase-space in Section 6.5.

To extract the numbers of $Zb\bar{b}/c\bar{c}$ and $t\bar{t}$ events in the three signal region 4μ, 4e, 2μ2e several transition factors α need to be applied.
The proportion of Z1+$\mu\mu'$, Z1+$e\mu$, Z1+ee' events is taken from data and applied to the results of the fit: this high SIP_{3D} lepton region is mainly populated by muons with respect to electrons and the relative proportions are 71%, 25%, 4%.
The ratio between Z1+ opposite-sign (OS) (Figure 7.7d) leptons events with respect to Z1+ same-sign (SS) (Figure 7.7c) leptons events is again taken from data and the factor obtained is 2.5. Then, assuming that the Z1 leg is equally distributed to two-muons and two-electrons final state, statistical transfer factor from Z1+$\mu\mu'$ to $\mu^+\mu^-\mu^+\mu^- = 4\mu$ and from Z1+ee' to $e^+e^-e^+e^- = 4e$ are equal to 1/2(2.5/(1+2.5). As an estimate of the final number of events in the $\mu^+\mu^-e^+e^- = 2e2\mu$ category we take the sum of 4μ and 4e categories. The overall factor that takes into account the reduction from e.g. $Zb\bar{b}/c\bar{c}\rightarrow 4\ell$ to $Zb\bar{b}/c\bar{c}\rightarrow 4\mu$ events will be called $\alpha_{4\mu}$.
The second transfer factor α accounts for the kinematic criteria on the m_{Z2} and $m_{4\ell}$ that are included in the best candidate choice (Section 6.2); this reduction factor α_{kin1} is taken from data. The events of the control region that pass these requirements are shown in the Figure 7.8.

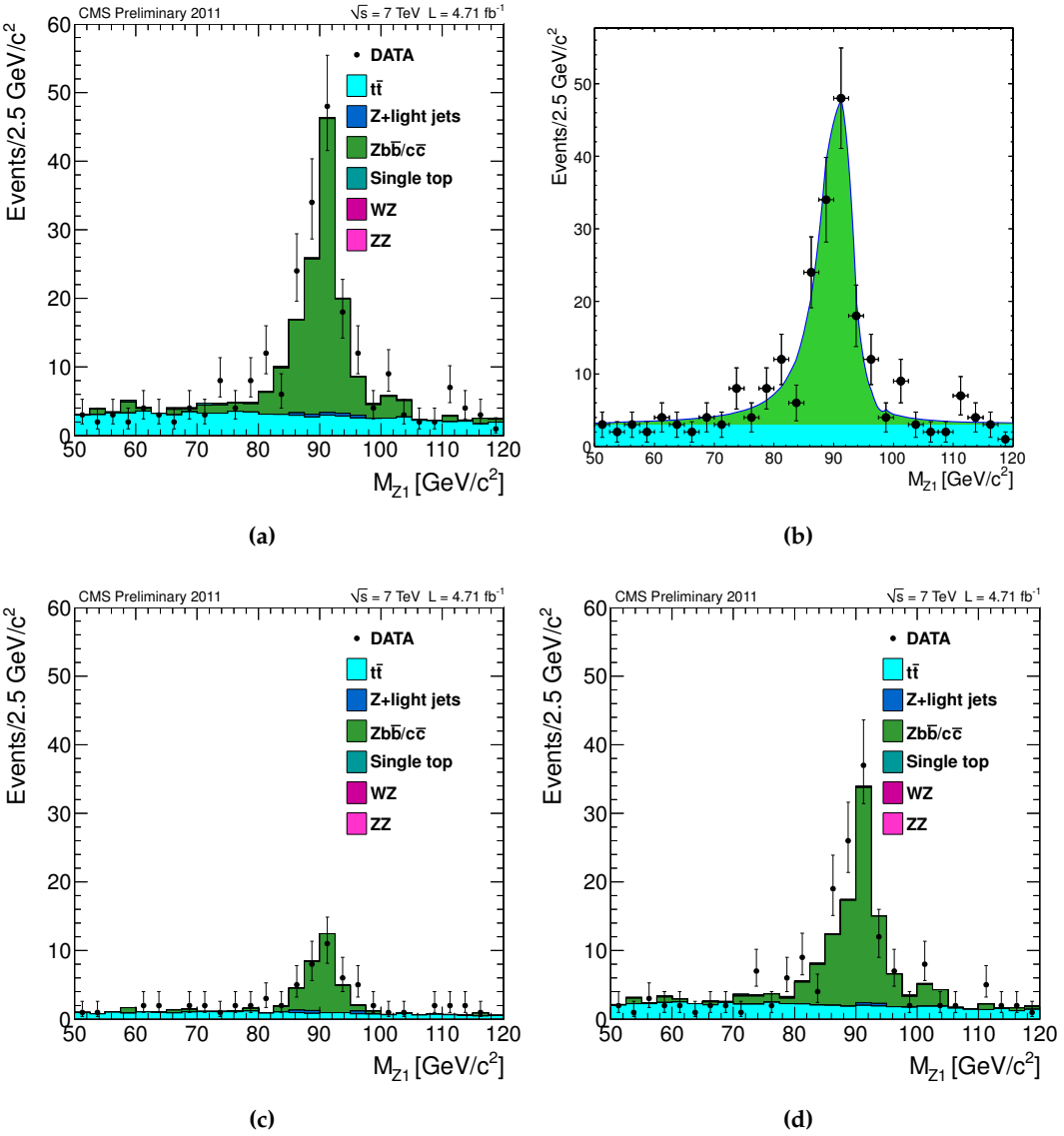


Figure 7.7: *Distribution of the Z1 mass in the Z1+ 2ℓ background control region for the $Zb\bar{b}/c\bar{c}$ and $t\bar{t}$ (a). The results are shown for data (points with statistical uncertainties) and stacked simulation expectation for the backgrounds (shaded histograms). A fit to the data is performed (b) using a Breit-Wigner function convoluted with a Crystal Ball function forthe Z1 peak and a Chebychev Polynomials for the $t\bar{t}$ spectrum. The same distribution is shown for the Z1+SS (same sign leptons) events (c) and for Z1+OS (opposite sign leptons) (d).*

In order to determine transfer factors for the SIP_{3D} and R_{Iso} cuts as described in Section 6.3, the functional shape of the SIP_{3D} and R_{Iso} distribu-

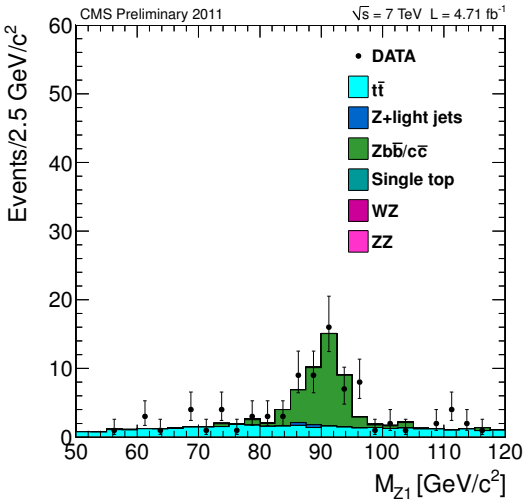


Figure 7.8: *Distributions of the Z1 mass in the Z1+ 2ℓ background control region for the $Zb\bar{b}/c\bar{c}$ and $t\bar{t}$ for events with m_{Z2} and m4ℓ in the signal region. The results are shown for data (points with statistical uncertainties) and stacked Monte Carlo expectation for the backgrounds (shaded histograms) for the full statistics corresponding to $\mathcal{L} = 4.71\, fb^{-1}$ of integrated luminosity.*

tions for the Z2 leptons of the $t\bar{t}$ and $Zb\bar{b}/c\bar{c}$ backgrounds as obtained from the simulation is assumed relying on the fact that they well reproduce the data shapes as shown in Figure 6.11. From these shapes the acceptance ratios α: $\alpha_{SIP_{3D}} = A_{SIP_{3D}} < 4/A_{SIP_{3D}} > 5$ and $\alpha_{R_{Iso}} = (R_{iso,j} + R_{iso,i} < 0.35)/(allR_{Iso,i,j})$ are calculated.
The final factor needed is α_{kin2} that takes into account the selection on m_{Z2} that differs in the "baseline", "intermediate-mass", "high-mass" selection. The total number of $t\bar{t}$ and $Zb\bar{b}/c\bar{c}$ events in the full $m_{4\ell}$ specra of 4μ signal-like region can thus be estimated as follows:

$$N_{expect}^{Zb\bar{b}/c\bar{c}\longrightarrow 4\mu} = N_{control-fit}^{Zb\bar{b}/c\bar{c}\longrightarrow 4\ell} \times \alpha_{4\mu} \times \alpha_{kin1} \times \alpha_{SIP_{3D}} \times \alpha_{R_{Iso}} \times \alpha_{kin2}$$

Following this strategy the results obtained for $Zb\bar{b}/c\bar{c}$ and $t\bar{t}$ backgrounds in the the 4μ, 4e, $2\mu 2e$ final states have been evaluated for the three kinematic selection (baseline, intermediate and high mass) and are reported in table Table 7.7. The related uncertainties take into account the limited data statistic in the control region as well as the statistical uncertainties on the transfer factors α taken from simulation.
The event density $\rho^{Zb\bar{b}/c\bar{c}}(m)$ (see Equation 7.1) is obtained from a fit to the reconstructed spectrum $dN/dM_{4\ell}$ obtained from MC simulation for $Zb\bar{b}/c\bar{c}$; after applying the full selection (going into the phase space of either the "baseline" or "intermediate-mass" or "high-mass" selection) the MC statistic is very poor, therefore the functional form for the fit for

$\rho^{Zb\bar{b}/c\bar{c}}(m)$ (Landau function, top of Figure 7.9) is extracted from the $m_{4\ell}$ distribution in the "enlarged phase-space" of Figure 6.11. The same procedure has been followed to extract the $t\bar{t}$ background functional form (polynomial function, bottom of Figure 7.9).

Table 7.7: *Number of events for $t\bar{t}$ and $Zb\bar{b}/c\bar{c}$ backgrounds estimated from the control region with inverted SIP_{3D}, relaxed isolation, charge, and flavour requirements for two leptons. They correspond to an integrated luminosity of $\mathcal{L} = 4.71\, fb^{-1}$.*

	baseline	intermediate-mass	high-mass
$N^{Zb\bar{b}/c\bar{c}\to 4\mu}$	0.7 ± 0.3	0.5 ± 0.2	0.08 ± 0.04
$N^{Zb\bar{b}/c\bar{c}\to 4e}$	0.04 ± 0.03	0.04 ± 0.03	0.008 ± 0.007
$N^{Zb\bar{b}/c\bar{c}\to 2\mu 2e}$	0.7 ± 0.3	0.5 ± 0.2	0.09 ± 0.04
$N^{t\bar{t}\to 4\mu}$	0.017 ± 0.002	0.013 ± 0.002	-
$N^{t\bar{t}\to 4e}$	0.03 ± 0.01	0.02 ± 0.01	-
$N^{t\bar{t}\to 2\mu 2e}$	0.04 ± 0.01	0.04 ± 0.01	-

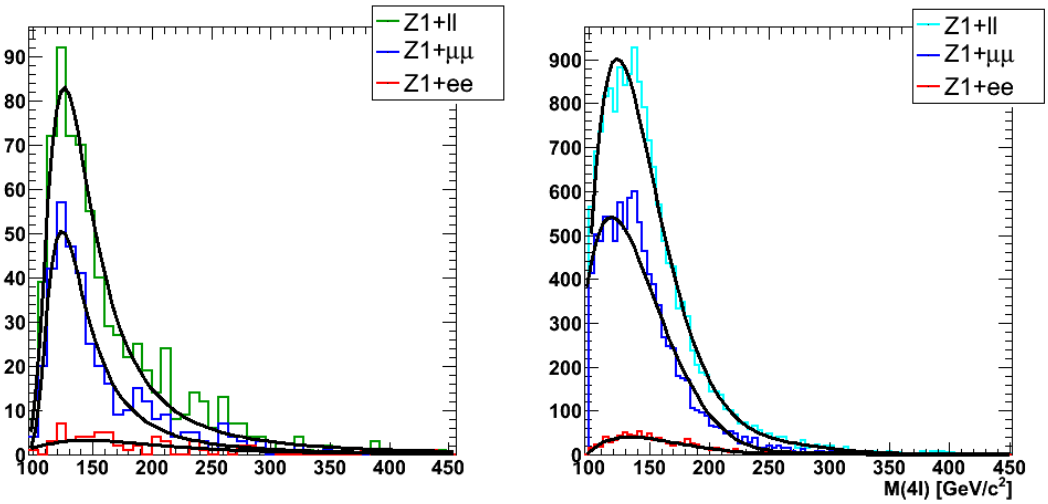


Figure 7.9: *The shapes for $Zb\bar{b}/c\bar{c}$ (top) and $t\bar{t}$ (bottom) reducible backgrounds fitted in the extended-phase space selections for : Z1+$\ell\ell$, Z1+$\mu\mu$, Z1+ee.*

Chapter 8

Results

8.1 Mass Distributions and Kinematics

The reconstructed four-lepton invariant mass distribution obtained in the 4μ, $4e$ and $2e2\mu$ channels with the baseline selection is shown in Figure 8.1 for the data, and compared to expectations from the SM backgrounds. Also shown is the measurement and comparison with data for the sum of the three 4ℓ channels. The same distributions for the high-mass selection are shown in Figure 8.2.
The Z+X reducible background distribution is obtained combining the rate normalization from data-driven method and knowledge on shape taken from the MC samples, as described in Section 7.2. Its distribution shows that the reducible and instrumental backgrounds rates are very small.
The ZZ and signal normalization and shapes are directly taken from MC samples.
The number of events observed, as well as the background rates in the signal region within a mass range from $m_1 = 100$ GeV/c^2 to $m_2 = 600$ GeV/c^2, are reported for each final state in Table 8.1 for the baseline selection.
The observed distribution is found to be compatible with the expectation from SM continuum production of $ZZ^{(*)}$ pairs. Seventy-two candidate events are selected with the baseline selection. Of these candidates, twenty-three are found in the 4μ channel, twelve in the $4e$ channel, and thirty-seven in the $2e2\mu$ channel. Fifty-two satisfy the high mass selection, fourteen 4μ, eight $4e$ and thirty $2e2\mu$ events, and contribute to the signal phase space for the measurement of the ZZ cross section that will be discussed in Section 8.2.
A zoom on the low mass range ($m_{\mathrm{H}} < 160$ GeV/c^2) is shown in Fig. 8.3 for the combination of the three channels. Thirteen of the candidates, five 4μ, three 4e and five $2\mu2e$, have been observed in this region while 9.5 ± 1.3 background events are expected.

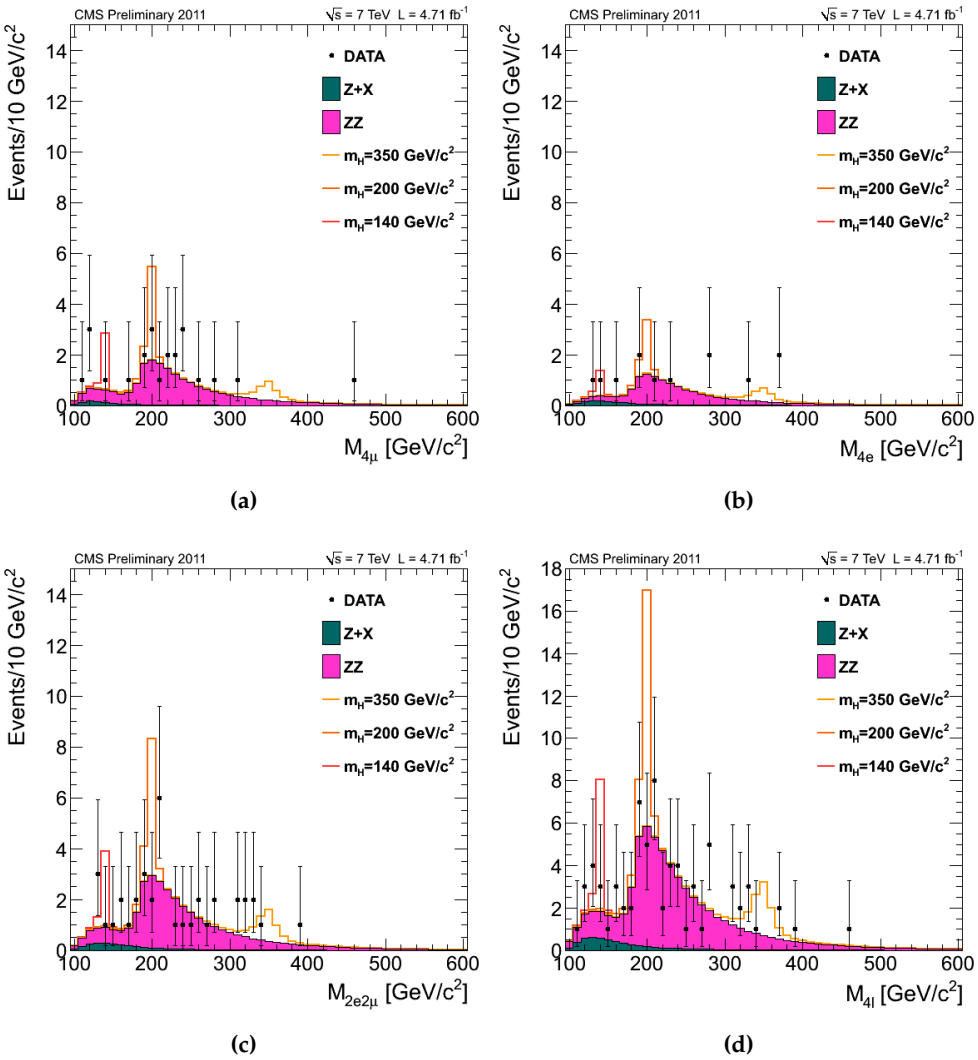


Figure 8.1: *Distribution of the four-lepton reconstructed mass for the baseline selection in the (a) 4e, (b) 4μ, (c) 2e2μ, and (d) the sum of the 4ℓ channels. The samples correspond to an integrated luminosity of* $\mathcal{L} = 4.71\ fb^{-1}$.

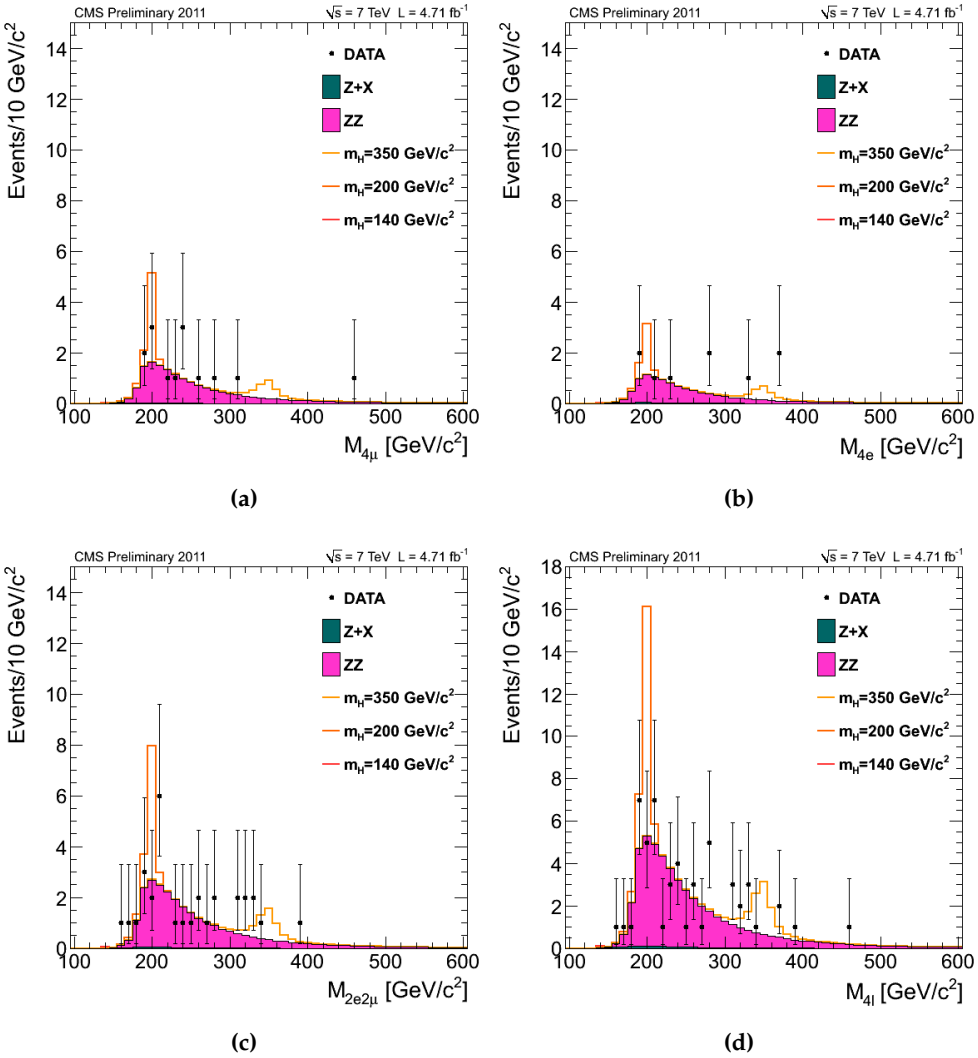


Figure 8.2: *Distribution of the four-lepton reconstructed mass for the high-mass selection in the (a) 4e, (b) 4μ, (c) 2e2μ, and (d) the sum of the 4ℓ channels. The samples correspond to an integrated luminosity of* $\mathcal{L} = 4.71\ fb^{-1}$.

Table 8.1: *Number of events observed, background and signal rates for each final state in a mass range from $m_1 = 100$ GeV/c^2 to $m_2 = 600$ GeV/c^2 for the baseline selection. For Z+X background the data driven estimation is used. For the backgrounds the average (theoretical and instrumental) systematic on the expected number of events are reported.*

Baseline	$4e$	4μ	$2e2\mu$
ZZ	12.27 ± 1.16	19.11 ± 1.75	30.25 ± 2.78
Z+X	1.67 ± 0.55	1.13 ± 0.55	2.71 ± 0.96
All background	13.94 ± 1.28	20.24 ± 1.83	32.96 ± 2.94
$m_H = 120$ GeV/c^2	0.25	0.62	0.68
$m_H = 140$ GeV/c^2	1.32	2.48	3.37
$m_H = 350$ GeV/c^2	1.95	2.61	4.64
Observed	12	23	37

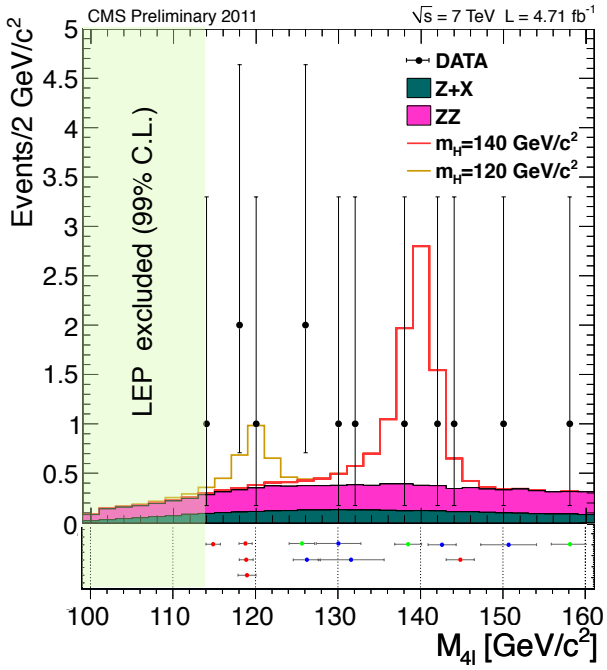


Figure 8.3: *Distribution of the four-lepton reconstructed mass for the sum of the 4ℓ channels in the low-mass domain with $m_H < 160$ GeV/c^2. The LEP exclusion limit at 99% C.L. is shown. On the bottom of the distribution an un-binned representation of the data with the event-by-event mass uncertainties, that will be introduced in Section 8.3, is presented. The results are presented for an integrated luminosity of 4.71 fb^{-1}.*

The correlation between the four-lepton reconstructed mass, the transverse momentum of the four-lepton system, the reconstructed mass of the second (Z2) and first (Z1) lepton pairs are shown in Figure 8.4, 8.5 and 8.6.

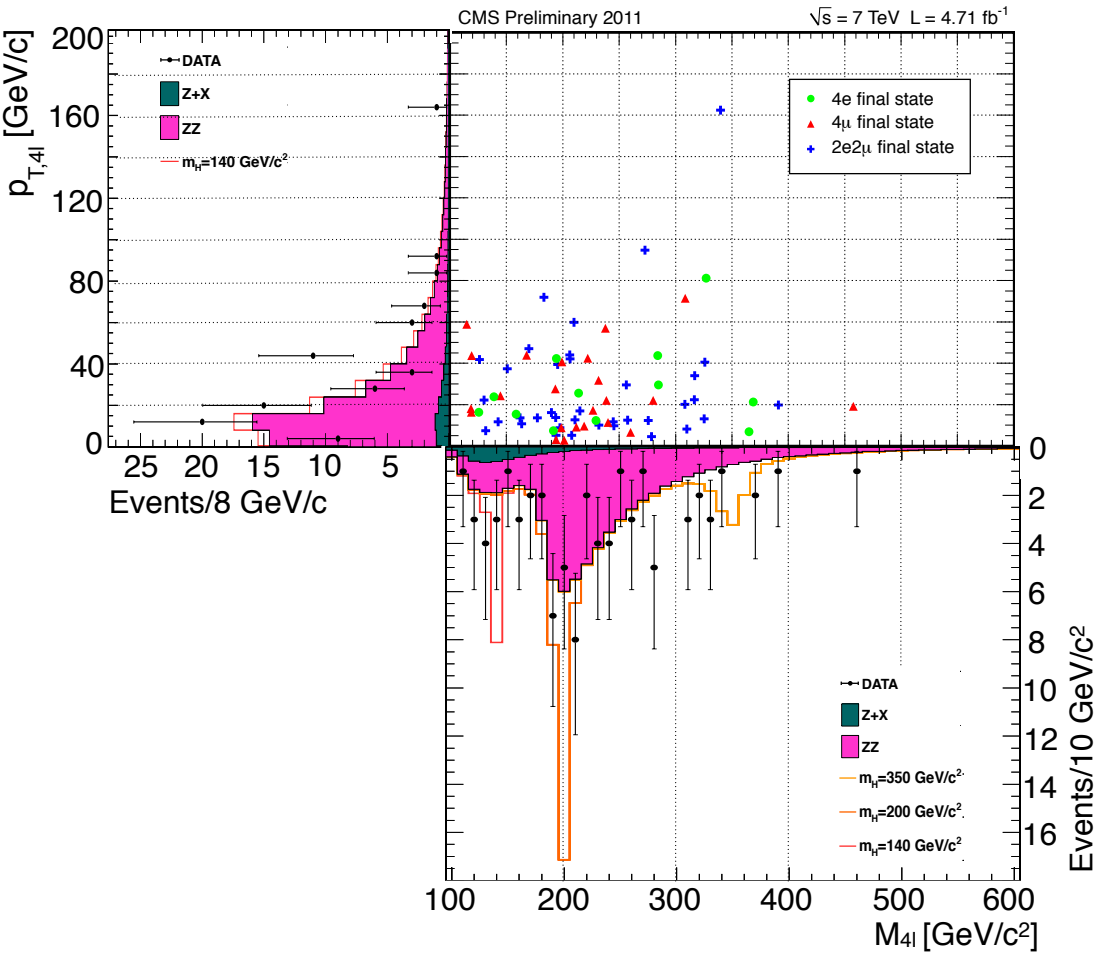


Figure 8.4: *Distribution of transverse momentum of four-lepton system ($p_{T,4l}$) versus the four-lepton reconstructed mass ($M_{4\ell}$) for the 4e (green circle), 4μ (green triangle) and 2e2μ (blue cross) final states and their projections. In the projections points represent the data, shaded histograms represent the signal and background expectations. The results are presented for an integrated luminosity of 4.71 fb^{-1}.*

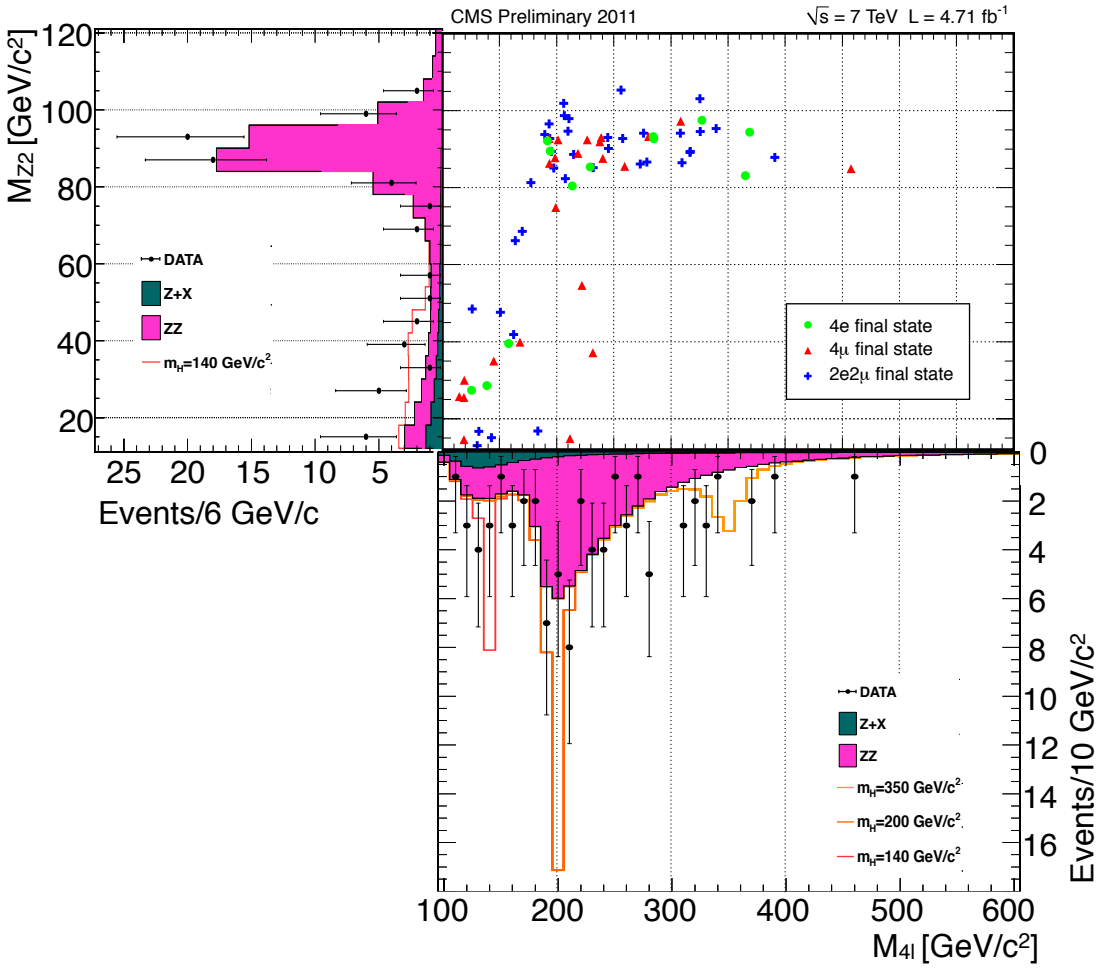


Figure 8.5: *Distribution of the reconstructed mass of the second lepton pair (M_{Z2l}) versus the four-lepton reconstructed mass ($M_{4\ell}$) for the 4e (green circle), 4μ (green triangle) and 2e2μ (blue cross) final states and their projections. In the projections points represent the data, shaded histograms represent the signal and background expectations. The results are presented for an integrated luminosity of 4.71 fb^{-1}.*

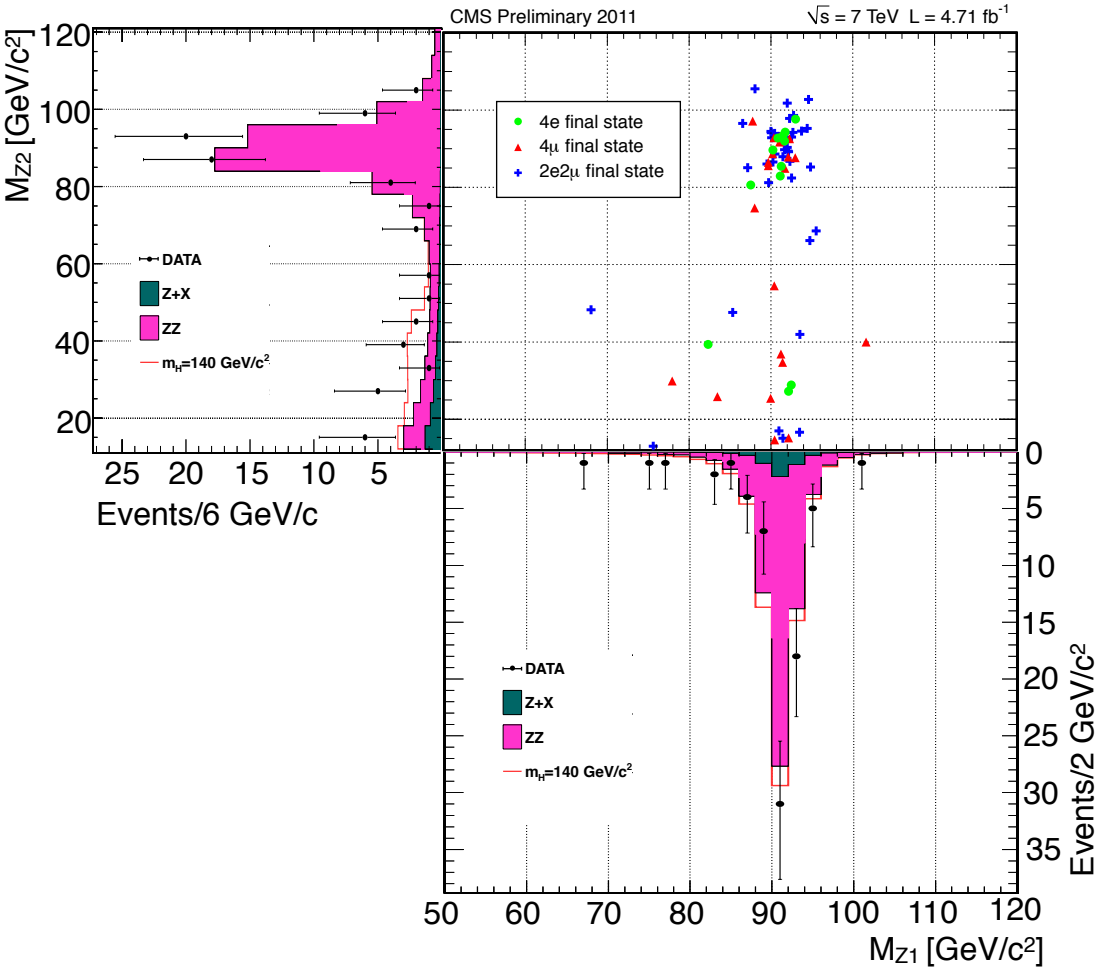


Figure 8.6: *Distribution of the reconstructed mass of the second lepton pair (M_{Z2l}) versus the reconstructed mass of the first lepton pair (M_{Z1}) for the 4e (green circle), 4μ (green triangle) and 2e2μ (blue cross) final states and their projections. In the projections points represent the data, shaded histograms represent the signal and background expectations. The results are presented for an integrated luminosity of 4.71 fb^{-1}.*

8.2 Measurement of the $ZZ \to 4\ell$ Cross Section

The high-mass event selection which imposes the presence of two lepton pairs with invariant masses in the range $60 < m_{\ell^+\ell^-} < 120$ GeV/c^2 can be used to provide a measurement of the total cross section $\sigma(pp \to ZZ + X) \times \mathcal{B}(ZZ \to 4\ell)$. The corresponding mass-spectra are shown in Figure 8.2.
This measured cross section is obtained via:

$$\sigma(pp \to ZZ + X) \times \mathcal{B}(ZZ \to 4\ell) = \frac{\sum_{i_{ch}}(N_{\text{obs}}(i_{\text{ch}}) - N_{\text{back}}(i_{\text{ch}}))}{A \times \epsilon_{(ZZ\to 4\ell)\times\mathcal{L}}} \quad (8.1)$$

where i_{ch} means 4e, 4μ, 2e2μ. A is the p_T and η acceptance at generator level evaluated with respect to generated events with $60 < m_{Z1} < 120$ GeV/c^2 and $60 < m_{Z2} < 120$ GeV/c^2. As described in Section 6.1 the acceptance is defined by the faction of events with four leptons satisfying $p_T > 7(e), 5(\mu)$ within $|\eta^e| < 2.5$ and $|\eta^\mu| < 2.4$. The lepton p_T is understood here to be taken after final-state QED radiation. Finally the efficiency $\epsilon_{ZZ\to 4\ell}$ is the ratio between the events satisfying the high-mass selection with respect to those within the acceptance and with $60 < m_{Z1} < 120$ GeV/c^2 and $60 < m_{Z2} < 120$ GeV/c^2 at generator level.
To compute acceptance and efficiency factors the $q\bar{q} \to ZZ$ and $gg \to ZZ$ MC samples described in Section 3.4 are used.
The total cross section for a pair of Z bosons in the mass range $60 < m_Z < 120$ GeV/c^2 is found to be:

$$\sigma(pp \to ZZ + X) \times \mathcal{B}(ZZ \to 4\ell) = 28.1^{+4.5}_{-3.9}(\text{stat.}) \pm 1.0(\text{syst.}) \pm 1.3(\text{lumi.})\ \text{fb}$$

The measured cross section agrees within about one standard deviation with the expectation from the SM [123] which predicts 27.9 ± 1.9 fb. In Table 8.2 the number of observed events in data, the estimated number of reducible and instrumental background events from the Z+X data-driven method, the expected number of background events from ZZ processes with one of the two Z decaying into a pair of τ leptons, and the expected number of $ZZ \to 4\ell$ events are reported for each channel.

The statistical uncertainty is set as the 68.27% confidence interval for a Poisson distribution. To the systematic uncertainty three terms contribute. First of all the the effect on the selection efficiencies of the uncertainties on muon and electron measurements, reported in Table 7.4 are considered. Additional systematics affecting the estimated number of instrumental and reducible backgrounds come from the limited amount of data in the background control regions which propagates to the background evaluation in the signal region, and from the uncertainty on the extrapolation factor from the background control to the signal control regions; these

Table 8.2: *Number of observed events, number of expected events from reducible and irreducible backgrounds and from ZZ events with one of the two Z decaying into a pair of τ leptons, and expected $ZZ \to 4\ell$ events. The event selection is the high-mass selection and the integrated luminosity considered is 4.71 fb^{-1}.*

Final state	N_{obs}	$N_{back}(Z+X)$	$N_{back}(ZZ_\tau)$	$N_{exp} ZZ \to 4\ell$
4μ	14	0.47 ± 0.26	0.031 ± 0.003	14.7 ± 1.3
4e	8	0.22 ± 0.07	0.026 ± 0.002	10.3 ± 1.0
$2e2\mu$	30	0.65 ± 0.20	0.057 ± 0.005	24.8 ± 2.3
Total	52	1.34 ± 0.53	0.115 ± 0.010	49.78 ± 4.6

uncertainties are reported in Table 7.6. Finally for the QCD scale and the PDF+α_s uncertainties affecting the acceptance terms the recommendation of the LHC Higgs Cross Section working group [46].

8.3 Mass Measurement Uncertainties

The precision on the estimation of the m_H and m_{Z*} masses can vary significantly on on an event-by-event basis depending on the lepton p_T^ℓ and η_ℓ. In the case of electrons, energy measurement uncertainties can furthermore vary significantly depending on the "category" of the reconstructed electron object, i.e. depending if the electrons initiate a shower early in the tracker volume ("showering" electrons) or reach the ECAL surface largely unperturbed ("golden" electrons). Overall, uncertainties on the measured m_H and m_{Z*} can vary by a factor up to three for the same initial mass m_H. The different behaviour for electrons and muons is partly dealt on average by the fact that the analysis is carried separately for the three different final states, $4e$, 4μ and $2e2\mu$. To take into account the precision of the measurement on an event-by-event basis within each of the final states, one can either create sub-category of events, distinguishing for instance events containing central or forward leptons and "golden" or showering electrons, or fully treat the problem by propagating the uncertainties evaluated on individual lepton legs on an event-by-event basis. The latter is more powerful as it can improve on the significance on an eventual discovery depending on the clustering in mass of a handful of events. Therefore this strategy has been implemented.
Event-by-event mass errors are evaluated starting from the errors on the individual lepton momenta. For muons the full error matrix as obtained from the muon track fit is used. For electrons the estimated error on the momentum magnitude as obtained from the combination of the ECAL and tracker measurement is used, neglecting the uncertainty on the track direction from the GSF fit. These lepton momentum measurement errors

are then propagated to the 4l mass error and to the Z1 and Z2 mass errors using an analytical error propagation including all correlations.
The agreement between data and MC in the distributions of the estimated muon and electron uncertainties has been studied. A very good agreement as been found [5].
To verify the accuracy of the m_{4l} uncertainty estimate, the MC pull distributions have been studied. Pulls are defined as the difference between reconstructed and generated m_{4l}, divided by the estimated uncertainty. Figure 8.7 shows the pull distribution for the 4μ final state for a particular value of m_H and the width of the pull distribution, obtained from a Gaussian fit of the core as a function of the m_{4l} reconstructed mass. The m_{4l} uncertainties are found to be underestimated by 15-20%. This is attributed to final state radiation, which causes the left tail of the pull distribution, to the residual lepton momentum scale biases, which determine an additional smearing of the m_{4l} resolution, as well as to the modelling of the individual lepton momentum uncertainties. Studies of the resolution on the Z peak show a similar behavior in data and in MC. Therefore, the event-by-event mass uncertainties are scaled to take into account the average pull width, separately for the three final states. Figure 8.8 shows the final distribution of m_{4l} uncertainties as a function of m_{4l} for data events passing the baseline selection.

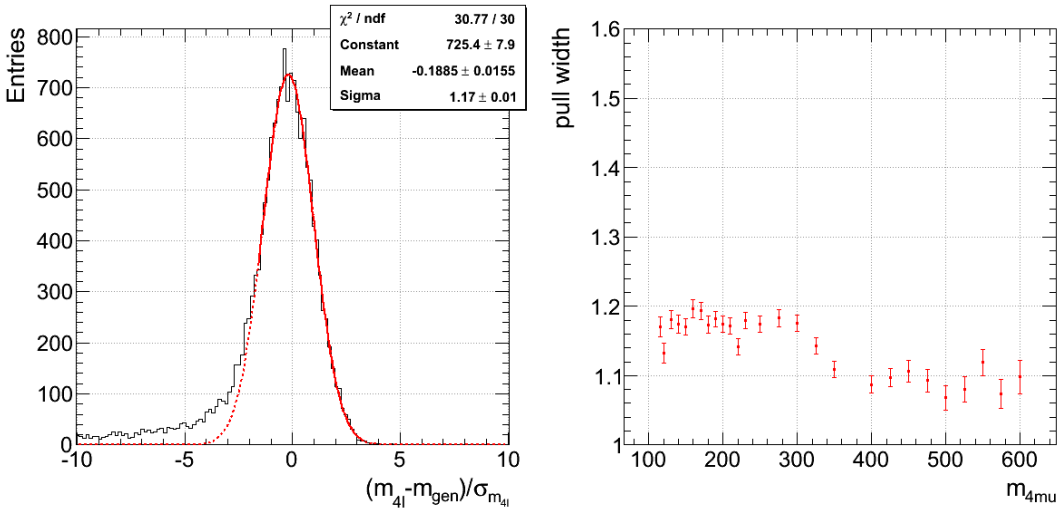


Figure 8.7: *MC pull distribution of the 4l invariant mass for the 4μ channel at a Higgs mass hypothesis of 150 GeV/c^2 (left), σ of the MC pull distributions as a function of reconstructed mass of the 4ℓ system (right).*

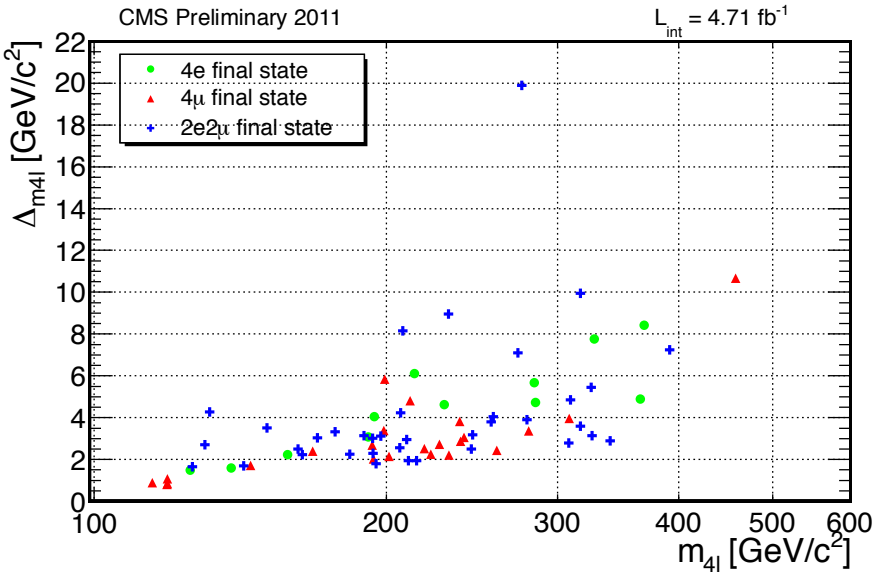


Figure 8.8: *Errors on the m_{4l} measurement as a function of m_{4l} in data.*

8.4 Interpretation of the Results

In order to quantify the sensitivity of the experiment to the presence of a Higgs boson signal a mass shape method is used.
As already anticipated in Section 3.3, as a prime method for reporting limits in this work, the modified frequentist construction (often referred to as CL_s or hybrid frequentist-bayesian) [47, 48] has been used. To fully define the method, one needs to make a choice of the test statistic and how one would treat nuisance parameters in the construction of the test statistic and in generating pseudo-data. In this work, the prescription prepared by the LHC Higgs Combination Group [49] has been followed.
For each value of the signal strength r ($r = \sigma/\sigma_{SM}$), a test statistics is defined starting from the likelihood ratio $2\ln Q = 2\ln(\mathcal{L}_{s+b}(r)/\mathcal{L}_b)$ where $\mathcal{L}_{s+b}(r)$ is likelihood for the signal-plus-background hypothesis, and $\mathcal{L}_b$ is the one for the background-only hypothesis. The predicted distribution of the test statistics in the two cases is determined from toy Monte Carlo, performing a bayesian pseudo-integration on the systematical uncertainties. The observed value of the test statistics is then compared with the prediction, yielding the p-values given the background or the signal-plus-background hypotheses. The two p-values are combined to construct the pseudo p-value $CL_s = CL_{s+b}/(1 - CL_b)$. The upper limit on r at 95% confidence level is defined as the value of r for which $CL_s = 0.05$.
To quantify an excess of events, the test statistic q_0, defined as follows has

been used:

$$q_0 \;=\; -2\ln\frac{\mathcal{L}(\text{data}|0,\hat{\hat{\theta}}_0)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})} \quad \text{and } \hat{\mu} \geq 0. \tag{8.2}$$

This test statistic is known to have a half χ^2 distribution for one degree of freedom, which allows us to evaluate significances (Z) and p-values (p_0) from the following asymptotic formula, derived from the asymptotic properties of the test statistic based on the profile likelihood ratio [124]:

$$Z = \sqrt{q_0^{\text{obs}}}, \tag{8.3}$$

$$p_0 \;=\; P(q_0 \geq q_0^{\text{obs}}) \;=\; \int_Z^\infty \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx \;=\; \frac{1}{2}\left[1-\text{erf}\left(Z/\sqrt{2}\right)\right] \tag{8.4}$$

where q_0^{obs} is the observed test statistic calculated for $\mu = 0$ and with only one constraint $0 \leq \hat{\mu}$, which ensures that data deficits are not counted on an equal footing with data excesses. The approximation has been tested and works well for the range of expected background and signal yields.

Inputs to the Exclusion Limits

The exclusion limits for a SM-like Higgs Boson are computed for a large number of mass points in the mass range from 110 to 600 GeV/c^2. The choice of the intervals in between Higgs mass hypotheses is driven by either detector resolution of the resonance or its natural width depending on which is larger. For the masses from 110 -160 GeV/c^2, the limits are computed every 1 GeV/c^2; for the masses from 160-290 GeV/c^2, the limits are computed every 2 GeV/c^2; for the masses from 290-350 GeV/c^2, the limits are computed every 5 GeV/c^2; for the masses from 350-400 GeV/c^2, the limits are computed every 10 GeV/c^2; for the masses from 400-600 GeV/c^2, the limits are computed every 20 GeV/c^2. The choice of Higgs mass points allows for good sensitivity of the exclusion limits in the full mass range. Due to the large number of mass points and availability of simulated signal samples, we must interpolate the mass shapes of signal hypotheses where no simulation exists.

The mass shape method uses the unbinned $4l$ invariant mass distribution for the hypothesized signal shape and expected background shape as a discriminator.

Background contributions

As extensively discussed in Chapter 7, the contributing background processes consist of the irreducible electroweak ZZ production and reducible contributions coming from Z+jets, $Zb\bar{b}$, and $t\bar{t}$ that are accounted for in

the data driven estimation of Z+X background events presented in Section 7.2.2. The expected background yields derived respectively from simulation and from data-driven techniques are given in Table 7.5 and Table 7.6.
Two mass shapes of the irreducible ZZ background are determined from MC, separately for the NLO and $gg \to ZZ$ contributions and have been shown in Figure 7.1. The shapes for the reducible backgrounds are also determined from simulation: they are determined from the $Z + X$ control region using the OS-SF control samples as shown in Figure 7.6.

Signal contributions

The signal shape is determined using 17 simulated samples covering the full mass range. The shapes for each simulated sample are fit using a function obtained as a convolution between a Breit-Wigner-like probability density function to describe the theoretical resonance line shape, convoluted with a Crystal-Ball function to describe the detector effects.
The Higgs boson theoretical line shape is described by the functional form [118]:

$$BW(m_{H*}|m_H) = \frac{1}{\mathcal{N}} \frac{\Gamma_{gg}(m_{H*})\Gamma_{ZZ}(m_{H*})}{(m^2_{H*} - m^2_H)^2 + m^2_{H*}\Gamma_{\text{tot}}(m_{H*})} * L'_{gg}(m_{H*}) * m_{H*} \tag{8.5}$$

Here, m_{H*} is the $4l$ invariant mass while m_H is the mass of the Higgs hypothesis. The partial Higgs widths [46] are given by Γ and $L'_{gg}(m_{H*})$ is the gluon partonic luminosity function. The motivation for choosing the shape analysis is to cover the line shape in the full mass region. In particular, the running width and gluon partonic luminosity function is important in the high mass region above approximately 300 GeV/c^2. Further, the choice of signal theoretical line shape is flexible enough to encompass theoretical uncertainties on the line shape in the high mass region.
The parameters of the Crystal-Ball function are interpolated for the Higgs boson mass points where there is no simulated sample available.
The fits for a few representative mass points are given in Fig. 8.9. The yields of the signal contribution are taken from MC after applying the baseline or high-mass selections, opportunely corrected with data-to-simulation scale factors (Section 7.1).
The observed and mean expected 95% CL upper limits on Higgs $\sigma(pp \to \mathrm{H} + X) \times \mathcal{B}(\mathrm{ZZ} \to 4\ell)$ from an analysis based on the shape of the mass distributions are shown in Figure 8.10. The limits are made using a $\mathrm{CL_s}$ approach, for the expected ratios to the SM. The bands represent the 1σ and 2σ probability intervals around the expected limit. We account for systematic uncertainties in the form of nuisance parameters with a log-normal probability density function. The estimation of the backgrounds

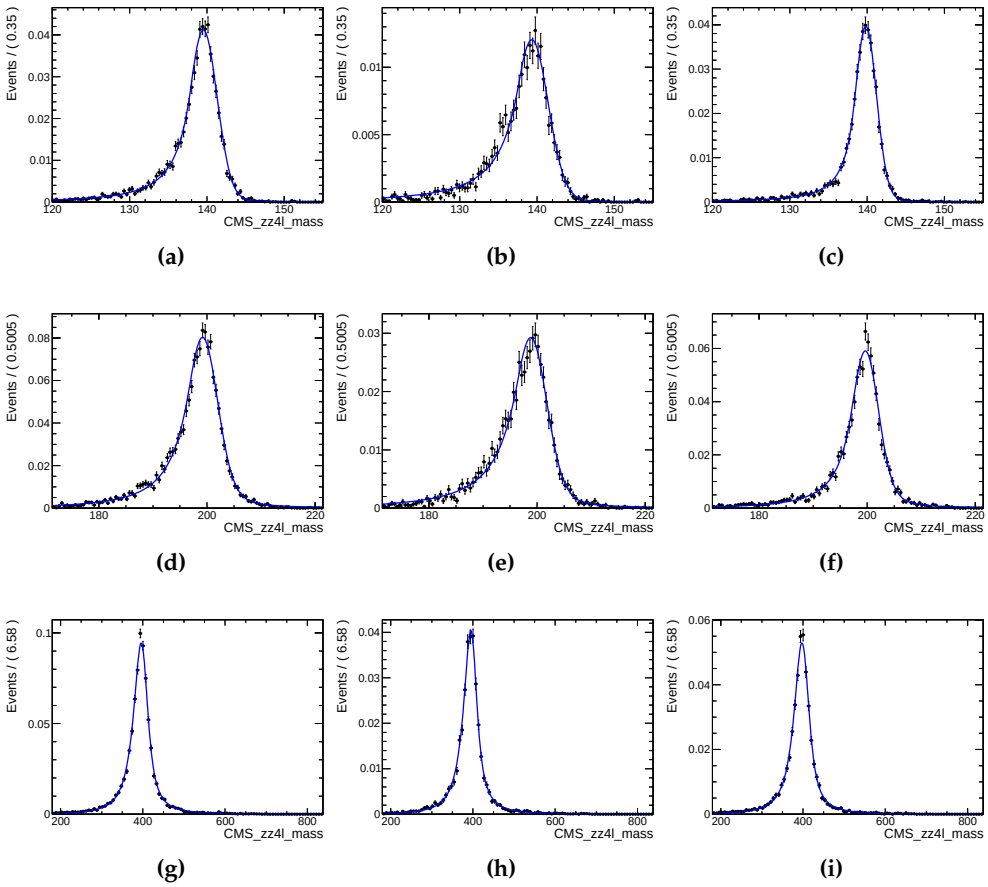


Figure 8.9: *Fits of simulated data in 3 channels for 3 representative mass points. The 2e2μ channel is on the left, 4e channel in the middle, and 4μ channel on the right. The 120 GeV/c^2 mass point is on top, 200 GeV/c^2 mass point is in the middle, and 400 GeV/c^2 mass point is on the bottom.*

follows the strategies described in Section 7.2.

The upper limits exclude the Standard Model Higgs boson at 95% CL in the ranges $134 < m_H < 158$ GeV/c^2, $180 < m_H < 305$ GeV/c^2 and $340 < m_H < 460$ GeV/c^2. The exclusion limits extend at high mass beyond the sensitivity of previous collider experiments. The expected limits reflect the dependence of the branching ratio $\mathcal{B}(H \to ZZ)$ on the Higgs boson mass. The worsening of the limits at high mass arise from the decreasing cross section for signal. Thanks to the excellent mass resolution and low background the structure in the measured limits follows the fluctuations of the number of observed events.

The significance of the local fluctuations with respect to the Standard Model expectation as a function of the Higgs boson mass is shown in

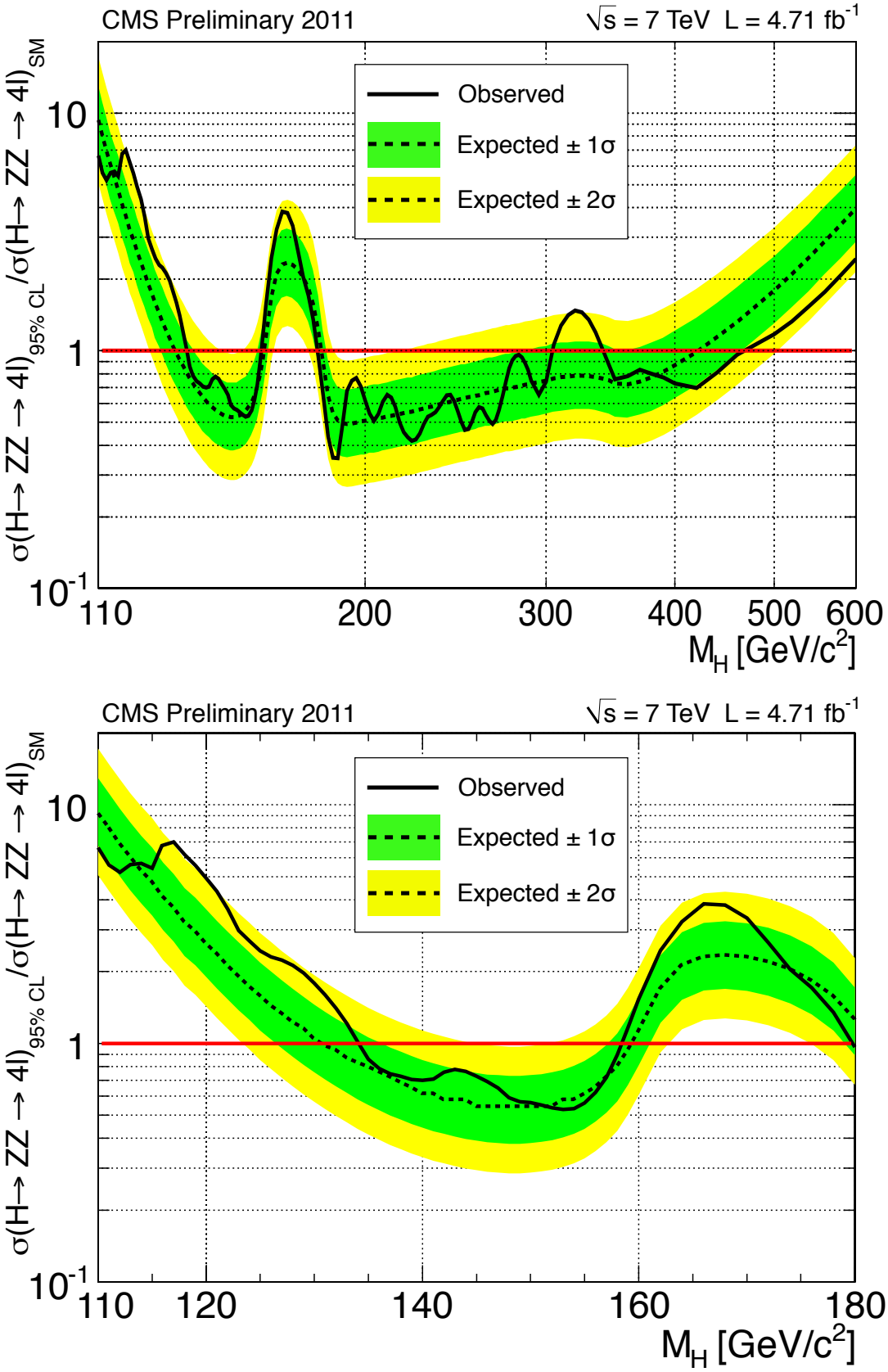


Figure 8.10: *The mean expected and the observed upper limits at 95% C.L. on $\sigma(pp \to H + X) \times \mathcal{B}(ZZ \to 4\ell)$ for a Higgs boson (top) in the mass range 110-600 GeV/c^2, (bottom) zoom in the low mass range (110-180 GeV/c^2), for an integrated luminosity of 4.71 fb^{-1} using the CL_s approach. The results are obtained using a shape analysis method.*

Figure 8.11 for the combination of the three channels. Also shown is the significance computed by taking into account the event-by-event uncertainties discussed in Section 8.3, instead of the average resolution for each final state. The most significant fluctuations are observed for masses around $\sim$ 120 GeV/c^2 and $\sim$ 320 GeV/c^2. The fluctuation around $\sim$ 120 GeV/c^2 corresponds to about 2.5σ (2.7σ) not including (including) event-by-event mass uncertainties. The global significance is below 1.0σ (about 1.6σ) when the look-elswhere effect [47, 48] is accounted for over the full mass range (for the range $100 < m_{4l} < 160$ GeV/c^2).
The expected significance of local fluctuations in the hypothesis of SM Higgs boson for any given mass point is also shown in Figure 8.11 for the low mass region. In the low mass region that remains not excluded, i.e. $114.4 < m_H < 133$ GeV/c^2, thus still possibly consistent with expectations of the Standard Model, the expected significance for the SM Higgs boson with the present luminosity is always below 3σ.

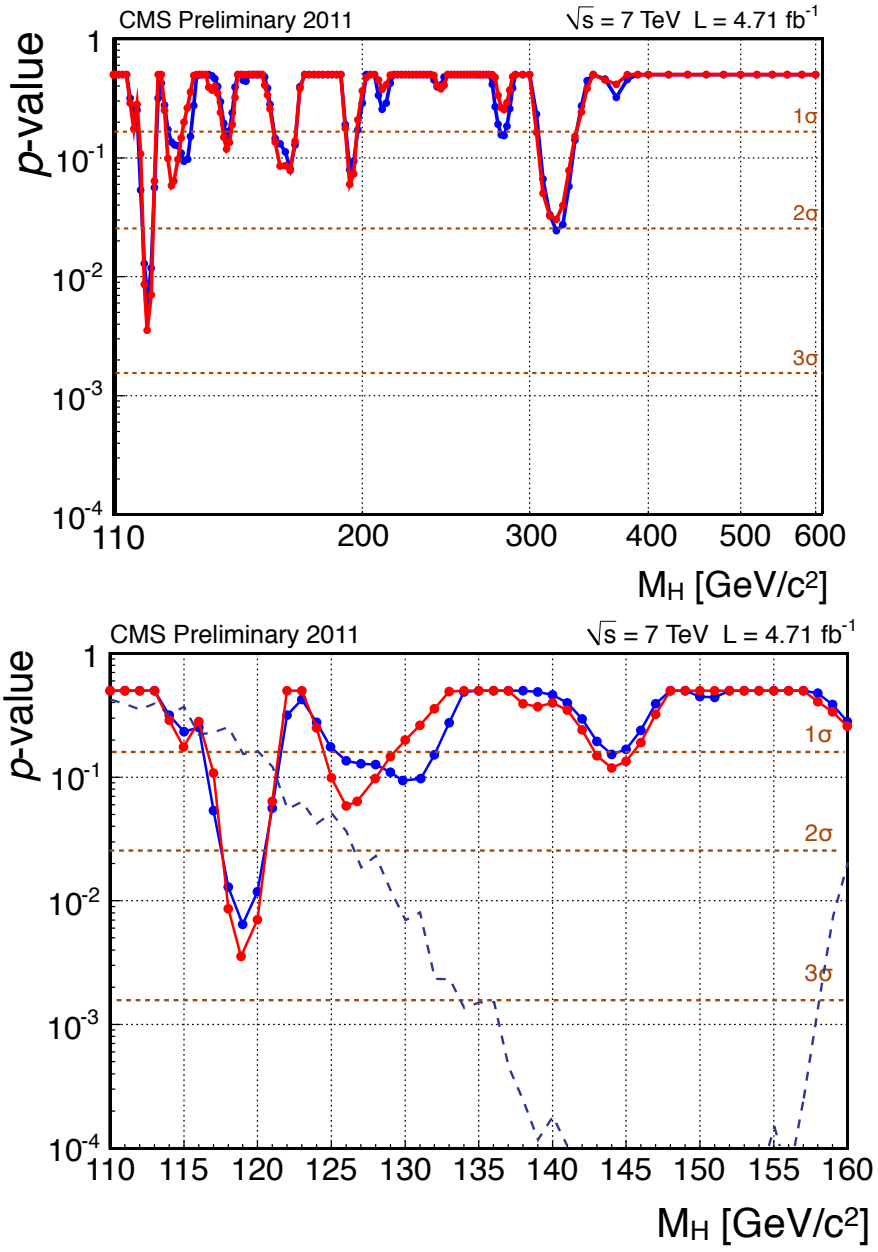


Figure 8.11: *Significance of the local fluctuations with respect to the standard model expectation as a function of the Higgs boson mass for an integrated luminosity of 4.71 fb^{-1}, with the default average (blue) and event-by-event (red) resolutions, in the mass range 110-600 GeV/c^2 (top), and in the low mass range (110-160 GeV/c^2) (bottom). The expected significance at any given mass point, should a SM Higgs boson exist with that mass, is shown with a dashed line.*

Summary

In this thesis I presented a search for the Standard Model Higgs boson produced in pp collisions at $\sqrt{s} = 7$ TeV and decaying in the final state $ZZ^{(*)} \to 4\ell$ with $\ell = e, \mu$. The analysis relies solely on lepton signatures. In particular, I first discussed the preparatory work I did to characterize in detail the muon trigger, reconstruction and identification performances, and highlighted their impact on the obtained physics results.
I then illustrated the choice of the event selection, consisting of simple sequential cuts on lepton reconstruction, identification and isolation variables, and of a set of kinematic cuts.
Also, I presented a study to control the instrumental background from Z+jets and the reducible backgrounds from $Zb\bar{b}/c\bar{c}$ and $t\bar{t}$, with misidentified primary leptons, proving that their contribution is negligible over most of the mass range, with a small contamination remaining at low mass.
The measured m_{4l} invariant mass distribution is found to be compatible with the expectation from SM continuum production of $ZZ^{(*)}$ pairs. A total of 52 four-lepton candidates has been observed for the high-mass selection compared to an expectation of 51.3 ± 4.6 events from SM background. This selection is used to provide a measurement of the total cross section $\sigma(pp \to ZZ + X) \times \mathcal{B}(ZZ \to 4\ell) = 28.1^{+4.6}_{-4.0}(\text{stat.}) \pm 1.2(\text{syst.}) \pm 1.3(\text{lumi.})$ fb. The measurement agrees within one standard deviation with the SM prediction of 27.9 ± 1.9 fb.
In the baseline selection, 72 four-lepton candidates have been observed, 12 in $4e$, 23 in 4μ and 37 in $2e2\mu$ final states, while 67.1 ± 6.0 events are expected from standard model background processes. Thirteen of the candidates are observed within $100 < m_{4\ell} < 160$ GeV/c^2 while 9.5 ± 1.3 background events are expected.
Excesses are observed for masses near 120 GeV/c^2 and 320 GeV/c^2. The most significant excess near 120 GeV/c^2 corresponds to about 2.5σ (2.7σ) significance not including (including) candidate mass uncertainties. The significance is less than 1.0σ (about 1.6σ) when the look-elsewhere effect is accounted for over the full mass range (for the range $100 < m_{4\ell} < 160$ GeV/c^2).

In summary, a major fraction of the mass range $110 < m_H < 600$ GeV/c^2 is excluded at 95% CL and only the regions $114.4 < m_H < 133$ GeV/c^2 and $173 < m_H < 178$ GeV/c^2 remain possibly consistent with expectations of the standard model.

The result of the H$\rightarrow 4\ell$ analysis is one of the inputs to the "grand-combination" of results of searches for a SM Higgs boson in five decay modes ($\gamma\gamma$, $b\bar{b}$, $\tau\tau$, ZZ and WW) with 4.7 fb^{-1} in CMS. With this amount of data the expected exclusion range in the absence of the Standard Model Higgs boson is 117-543 GeV/c^2 at 95% C.L. The observed data allow the CMS Collaboration to exclude the SM Higgs boson mass range 127-600 GeV/c^2 at 95% C.L. At 99% C.L. the exclusion limit is 128-525 GeV/c^2. The excess of events observed for hypothesised Higgs boson masses at the low end of the explored range makes the observed limits weaker than the expected ones. To ascertain the origin of this excess, more data are required [117].

Appendix A

Event Properties

In the following the properties of the candidate four-lepton combination are listed for each event.

Table A.1: *Properties of the four-lepton combinations satisfying the baseline selection for the Higgs boson search.*

Event	Channel	m_{Z_1}	m_{Z_2} (GeV/c^2)	$m_{4\ell}$	$\Delta m_{4\ell}$	$p_{T,4\ell}$ (GeV/c)	$y_{4\ell}$
A	4μ	91.37	92.60	201.18	2.15	2.86	0.18
B	4μ	101.53	40.04	167.99	2.40	43.74	1.45
C	$2e2\mu$	94.51	66.05	163.84	2.24	10.54	-0.54
D	$4e$	92.50	28.87	138.58	1.63	23.98	0.39
E	$2e2\mu$	92.40	82.35	207.85	8.15	5.02	1.83
F	4μ	91.30	34.83	144.91	1.71	24.13	-0.36
G	$2e2\mu$	91.22	93.19	244.58	2.50	11.98	-0.47
R-A	$2e2\mu$	91.64	14.92	142.62	1.70	11.51	0.90
H	$2e2\mu$	88.09	105.28	256.49	3.80	29.33	-1.21
I	$4e$	87.53	80.69	213.62	6.34	25.08	0.06
R-B	4μ	92.11	15.09	211.64	4.80	9.09	0.08
J	4μ	90.97	93.18	238.53	2.87	22.00	0.26
K	$2e2\mu$	92.39	92.80	193.51	3.01	13.95	0.82
L	4μ	90.42	54.82	222.30	2.26	42.30	-0.64
R-C	4μ	90.31	14.72	118.83	0.81	16.41	0.13
M	4μ	77.80	29.68	119.03	1.08	43.93	0.58
N	$4e$	92.31	27.22	125.66	1.55	16.09	0.07
O	$2e2\mu$	90.19	94.77	325.56	3.13	40.91	-0.43
P	$4e$	91.70	92.21	191.59	3.18	7.50	-0.34
Q	4μ	90.20	88.93	218.87	2.52	9.83	0.79
R	4μ	92.15	87.70	198.82	3.38	8.67	1.23
S	4μ	87.70	97.02	308.56	3.97	71.11	0.34
R-D	$2e2\mu$	93.54	16.96	131.61	4.25	7.30	1.17
T	$4e$	91.29	82.98	365.34	5.07	6.65	0.52
U	4μ	91.90	85.04	457.92	10.66	19.02	-0.52
V	4μ	90.22	25.36	118.94	0.85	18.15	-0.34
W	$2e2\mu$	92.22	89.76	316.71	3.60	22.021	-0.59
X	$2e2\mu$	91.82	89.42	195.23	1.79	39.27	-0.31
Y	4μ	91.15	37.15	231.94	2.22	31.74	-0.40
Z	$4e$	90.77	92.95	284.83	4.90	29.47	1.50
AA	$2e2\mu$	91.52	87.83	391.35	7.24	19.63	1.10
AB	$4e$	90.97	85.24	229.23	4.79	12.35	-1.39
AC	$2e2\mu$	90.44	93.78	189.63	3.14	16.18	0.81
AD	$2e2\mu$	92.40	94.24	308.13	2.80	20.41	0.002
AE	$2e2\mu$	94.96	85.25	231.65	8.96	10.09	1.27
AF	$2e2\mu$	92.92	98.72	206.86	4.22	42.32	-1.242
AG	$2e2\mu$	92.23	98.01	210.75	1.95	12.52	-0.05
R-E	$2e2\mu$	91.03	17.00	183.25	2.25	71.99	-0.49
AH	$2e2\mu$	89.92	94.22	275.80	19.8991	12.34	-1.64
AI	$4e$	82.42	39.50	158.22	2.30	15.05	-0.06

Event	Channel	m_{Z_1}	m_{Z_2} (GeV/c^2)	$m_{4\ell}$	$\Delta m_{4\ell}$	$p_{T,4\ell}$ (GeV/c)	$y_{4\ell}$
AJ	$2e2\mu$	91.75	101.76	206.42	2.56	44.27	-0.47
AK	4μ	89.77	86.30	193.86	2.03	3.35	-1.15
R-F	$2e2\mu$	75.59	12.93	130.08	2.70	22.43	-1.66
AL	$2e2\mu$	94.58	102.73	325.09	5.45	13.28	-0.77
AM	$4e$	92.88	97.47	327.47	8.04	81.21	0.27
AN	4μ	90.58	92.77	193.42	2.68	27.80	1.61
AO	$2e2\mu$	90.13	92.65	257.81	4.046	12.06	0.02
AP	4μ	93.01	87.70	240.35	3.06	11.21	-0.06
AQ	$2e2\mu$	86.49	96.33	193.72	2.30	5.18	-0.65
AR	$2e2\mu$	90.15	86.39	309.60	4.85	8.35	1.023
AS	$2e2\mu$	67.94	48.50	126.24	1.65	41.76	1.27
AT	4μ	91.62	93.23	280.36	3.36	21.88	0.37
AU	4μ	90.96	91.74	237.88	3.81	56.69	-0.91
AV	$4e$	91.72	94.19	368.73	8.73	21.11	0.90
AW	$2e2\mu$	92.35	86.56	278.95	3.90	4.26	-0.05
AX	$2e2\mu$	94.44	95.42	339.84	2.89	162.16	-0.93
AY	4μ	83.35	26.00	114.84	0.88	58.81	-0.50
AZ	$2e2\mu$	89.65	81.37	177.02	3.33	13.54	1.52
BA	4μ	89.79	85.76	259.93	2.43	6.57	-0.96
BB	$2e2\mu$	89.47	85.95	273.20	7.10	94.78	-1.09
BC	$2e2\mu$	87.14	85.02	197.38	3.12	8.77	0.84
BD	$2e2\mu$	92.03	89.37	316.96	9.95	33.72	-1.61
BE	$2e2\mu$	85.191	47.707	150.691	3.5113	37.484	0.803
BF	4μ	88.05	74.84	199.21	5.85	40.91	1.46
BG	$2e2\mu$	91.74	90.47	245.41	3.19	9.55	-0.70
BH	$2e2\mu$	93.70	94.55	209.85	2.96	59.86	0.08
BI	$2e2\mu$	93.43	42.16	162.15	2.49	13.51	-0.96
BJ	4μ	92.29	92.30	226.83	2.72	17.20	-0.75
BK	$2e2\mu$	95.69	68.62	169.81	3.03	46.99	-1.32
BL	$2e2\mu$	90.27	88.46	214.83	1.94	16.80	0.54
BM	$4e$	90.55	89.61	194.30	4.20	42.11	-1.29
BN	$4e$	91.57	93.10	284.11	5.87	43.88	-0.95

Appendix B

Event Displays. Low Mass Candidates

In the following the features of the low mass events ($m_{4l} < 140$ GeV/c^2) are presented.

Candidate R-C ($m_{4\ell} = 138.6\,\text{GeV}/c^2, \Delta m_{4\ell} = 1.12\,\text{GeV}/c^2, m_{Z_1} = 92.5\,\text{GeV}/c^2, m_{Z_2} = 28.9\,\text{GeV}/c^2$).

- 4 electrons
- No extra electron nor muon
- No extra photon
- 2 extra tracks $p_T > 5\,\text{GeV}/c$
- 4 extra jets $p_T > 10\,\text{GeV}/c$
- MET (PF): 9.95 GeV
- 7 vertices in the event (all leptons come from the same primary vertex)

Lepton	p_T	η	ϕ	charge	R_{iso}	$\lvert\text{SIP}_{3D}\rvert$
electron	43.6	0.17	-0.37	+	0.01	1.36
electron	52.6	0.58	-2.84	-	0.01	0.17
electron	26.8	0.16	1.58	+	0.02	2.70
electron	13.8	0.79	3.07	-	0.00	0.53

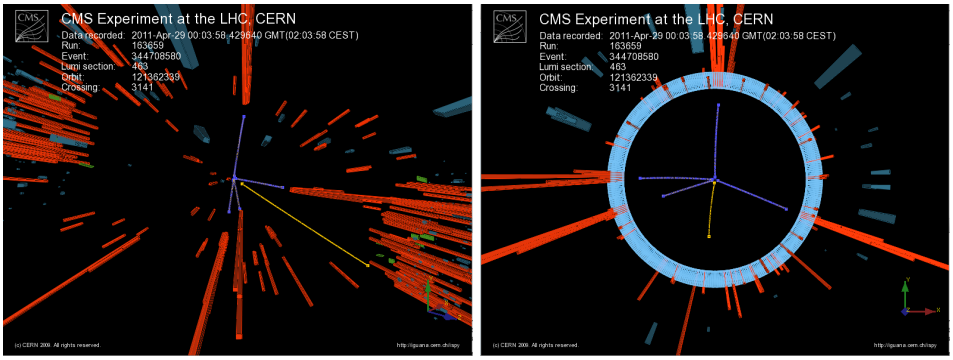


Figure B.1: *(a) 3D view and (b) R-Phi view of the event D (Run # 163659 - Event # 344708580). The four electrons (in blue) are all in the ECAL barrel. The electrons with p_T = 26.8 GeV/c and 13.8 GeV/c at ϕ = 1.58 and 3.07 rad respectively and constituting the Z2 candidate are showering with clear double cluster patterns. There are two other tracks with $p_T > 5$ GeV/c close to each other at $\phi \sim -1.7$ rad and in the forward region, well away from the electrons. One is reconstructed as a GSF track (in yellow) and has several nearby low p_T tracks, including a possible conversion leg. Possibly a tau jet.*

Candidate R-C ($m_{4\ell} = 118.8\,\mathrm{GeV}/c^2, \Delta m_{4\ell} = 0.71\,\mathrm{GeV}/c^2, m_{Z_1} = 90.3\,\mathrm{GeV}/c^2, m_{Z_2} = 14.7\,\mathrm{GeV}/c^2$).

- 4 muons
- No extra electron nor muon
- No extra photon
- No extra tracks $p_T > 5\,\mathrm{GeV}/c$
- 1 extra jets $p_T > 10\,\mathrm{GeV}/c$ (central, $p_T = 18\,\mathrm{GeV}/c$)
- MET (PF): 16.3 GeV
- 2 vertices in the event (all leptons come from the same primary vertex)

Lepton	p_T	η	ϕ	charge	R_{iso}	$\lvert\mathrm{SIP}_{3D}\rvert$
muon	51.5	-0.32	-2.13	+	0.07	0.66
muon	36.6	0.41	0.54	-	0.00	1.44
muon	16.4	0.68	1.37	+	0.07	0.66
muon	5.2	0.77	-0.48	-	0.00	1.29

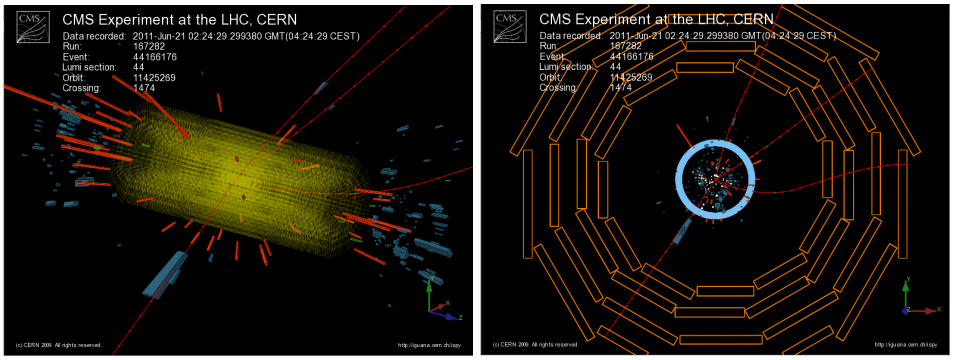


Figure B.2: *(a) 3D view and (b) R-Phi view of the event R-C (Run # 167282 - Event # 44166176). Very clean event with only two vertices. The two muons (consituting the Z1 candidate are close to back-to-back in the transverse plane ($\phi = -2.13$ rad and $\phi = 0.54$ rad). The Z2 candidate is constituted by the two lowest p_T muons, one of which is just above the threshold with a p_T of 5.2 GeV/c. Muons are in red.*

Candidate M ($m_{4\ell} = 119.0\,\text{GeV}/c^2, \Delta m_{4\ell} = 0.94\,\text{GeV}/c^2, m_{Z_1} = 77.8\,\text{GeV}/c^2, m_{Z_2} = 29.7\,\text{GeV}/c^2$).

- 4 muons
- 1 extra electron in the forward region, $p_T = 11.7\,\text{GeV}/c$
- 1 extra isolated photon E_T = 12.9 GeV, η = 1.06, ϕ = -2.30
- 3 extra tracks $p_T > 5\,\text{GeV}/c$ (one corresponding to the extra electron)
- No extra jets $p_T > 10\,\text{GeV}/c$
- MET (PF): 21.8 GeV
- 6 vertices in the event (all leptons come from the same primary vertex)

Lepton	p_T	η	ϕ	charge	R_{iso}	$\lvert\text{SIP}_{3D}\rvert$
muon	66.9	0.41	-0.45	+	0.00	0.30
muon	21.0	1.32	-2.80	-	0.00	0.71
muon	15.7	1.05	2.44	+	0.35	0.25
muon	6.0	-0.96	-0.30	-	0.00	0.62

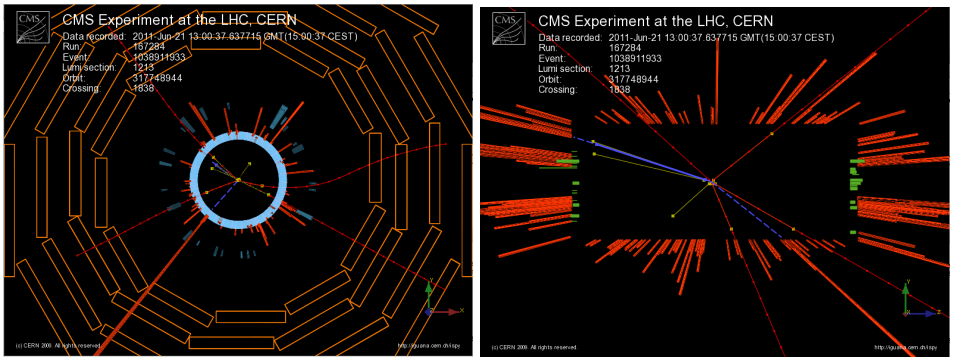


Figure B.3: *(a) R-Phi view and (b) R-Z view of the event M (Run # 167284 - Event # 1038911933). The isolated photon (dashed blue line) is selected by the FSR recovery algorithm. It has a p_T of 12.9 GeV/c and is at $\Delta R \sim 0.57$ from the muon of 21.0 GeV/c of the Z1 candidate. The extra electron in the forward region is not isolated with several surrounding tracks with $p_T > 5$ GeV/c (in yellow) and energy deposits in the calorimeters. The muon of 15.7 GeV/c at $\phi = 2.44$ rad is just below the isolation threshold used in the selection.*

Candidate N ($m_{4\ell} = 125.7\,\text{GeV}/c^2, \Delta m_{4\ell} = 1.07\,\text{GeV}/c^2, m_{Z_1} = 92.3\,\text{GeV}/c^2, m_{Z_2} = 27.2\,\text{GeV}/c^2$).

- 4 electrons
- No extra electron nor muon
- No extra photon
- No extra tracks $p_T > 5\,\text{GeV}/c$
- No extra jets $p_T > 10\,\text{GeV}/c$
- MET (PF): 35 GeV
- 4 vertices in the event (all leptons come from the same primary vertex)

Lepton	p_T	η	ϕ	charge	R_{iso}	$\lvert\text{SIP}_{3D}\rvert$
electron	46.2	-0.03	-2.95	+	0.02	1.07
electron	46.4	-0.02	0.32	-	0.02	2.31
electron	23.6	0.52	-2.18	+	0.01	2.07
electron	7.3	-0.28	1.58	-	0.00	0.71

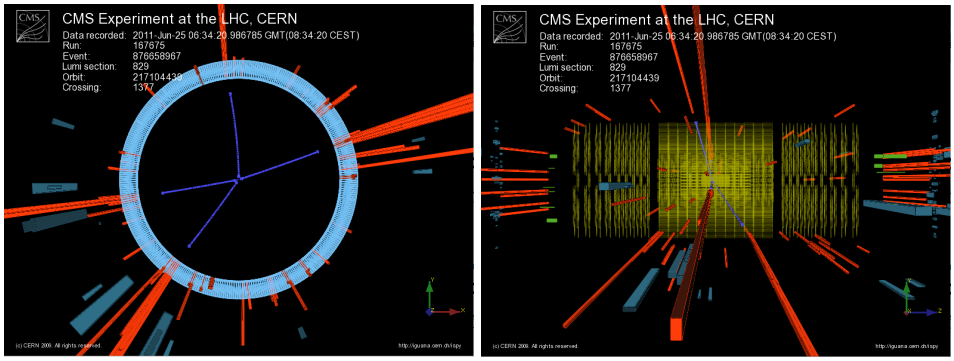


Figure B.4: *(a) R-Phi view and (b) R-Z view of the event N (Run # 167675 - Event # 876658967). A very nice low mass four electrons event. All electrons (in blue) are in the ECAL barrel, two are 'showering' and two are 'golden', all with excellent values for the ID variables and well isolated. The event is very clean with only four vertices and no other tracks with $p_T > 5$ GeV/c. The Z1 candidate is formed by the two electrons at $\eta \sim 0$ and $\phi = -2.95$ and 0.32 rad (back-to-back configuration). MET is on the high side. The 4th electron is just above the p_T threshold used in the selection.*

Candidate R-D ($m_{4\ell} = 131.6\,\text{GeV}/c^2, \Delta m_{4\ell} = 3.27\,\text{GeV}/c^2, m_{Z_1} = 93.5\,\text{GeV}/c^2, m_{Z_2} = 17.0\,\text{GeV}/c^2$).

- 2 electrons - 2 muons
- 1 extra electron ($p_T = 2.4\,\text{GeV}/c$)
- No extra photon
- No extra tracks $p_T > 5\,\text{GeV}/c$
- No extra jets $p_T > 10\,\text{GeV}/c$
- MET (PF): 9.0 GeV
- 3 vertices in the event (all leptons come from the same primary vertex)

Lepton	p_T	η	ϕ	charge	R_{iso}	$\lvert\text{SIP}_{3D}\rvert$
electron	56.4	1.60	-2.32	+	0.01	2.53
electron	37.1	1.02	1.23	-	0.00	0.97
muon	10.9	-0.23	0.11	+	0.00	1.75
muon	9.0	1.28	0.50	-	0.21	1.95

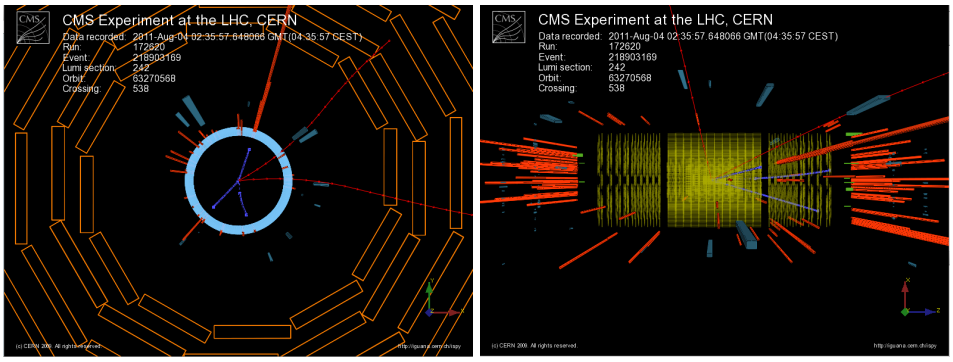


Figure B.5: *(a) R-Phi view and (b) R-Z view of the event R-D (Run # 172620 - Event # 218903169). The Z1 candidate is constituted by the two well isolated electrons (in blue) at $\phi = -2.32$ and 1.23 rad (back-to-back configuration). The leading electron is a 'crack' electron in the transition region between the ECAL barrel and the positive endcap. Its $\sigma_{i\eta i\eta}$ is not very good due to this particular location, but the electron passes the ID thanks to the loose requirements applied. The other electron is a 'golden' in the ECAL barrel. The two muons (in red) constituting the low mass pair have p_T around 10 GeV/c. There is an extra reconstructed electron of very low p_T (2.4 GeV/c) in the forward region ($\eta = 1.7$) and not starting from the primary vertex..*

Candidate V ($m_{4\ell} = 118.9\,\mathrm{GeV}/c^2, \Delta m_{4\ell} = 0.74\,\mathrm{GeV}/c^2, m_{Z_1} = 90.2\,\mathrm{GeV}/c^2$, $m_{Z_2} = 25.4\,\mathrm{GeV}/c^2$).

- 4 muons
- No extra electron nor muon
- No extra photon
- No extra tracks $p_T > 5\,\mathrm{GeV}/c$
- No extra jets $p_T > 10\,\mathrm{GeV}/c$
- MET (PF): 8.2 GeV
- 3 vertices in the event, 101 tracks (all leptons come from the same primary vertex)

Lepton	p_T	η	ϕ	charge	R_{iso}	$\lvert\mathrm{SIP}_{3D}\rvert$
muon	58.4	-0.47	1.71	-	0.00	0.44
muon	34.6	-0.25	-1.56	+	0.00	1.39
muon	5.7	-1.67	-0.77	-	0.08	1.07
muon	5.4	1.35	-2.52	+	0.20	0.40

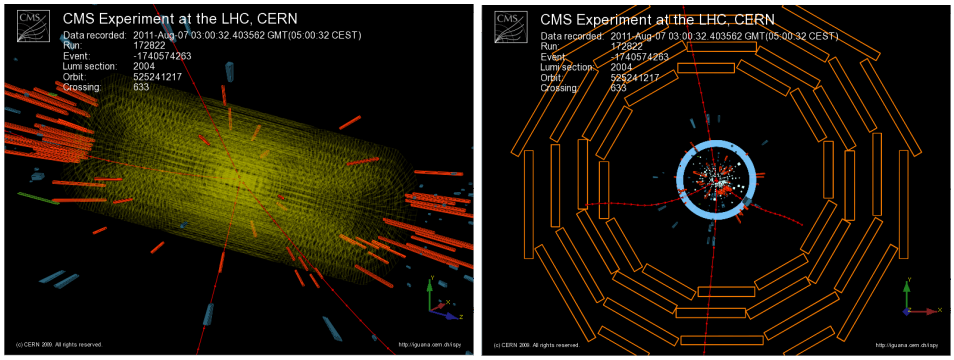


Figure B.6: *(a) 3D view and (b) R-Phi view of the event V (Run # 172822 - Event # 2554393033). Very clean event with low vertex multiplicity. The two muons constituting the Z1 candidate are in the central part and in a back-to-back configuration. The muons forming the Z2 pair have very low p_T (5.65 and 5.44 GeV/c), just above the selection threshold. Muons are in red.*

Candidate AS ($m_{4\ell} = 126.2\,\text{GeV}/c^2, \Delta m_{4\ell} = 1.27\,\text{GeV}/c^2, m_{Z_1} = 67.9\,\text{GeV}/c^2, m_{Z_2} = 48.5\,\text{GeV}/c^2$).

- 2 electrons - 2 muons
- 1 extra electron ($p_T = 2.8\,\text{GeV}/c$)
- No extra photons
- No extra tracks $p_T > 5\,\text{GeV}/c$
- 6 extra jets $p_T > 10\,\text{GeV}/c$ (mostly $|\eta| > 2.5$)
- MET (PF): 3.2 GeV
- 12 vertices in the event (all leptons come from the same primary vertex)

Lepton	p_T	η	ϕ	charge	R_{iso}	$\lvert\text{SIP}_{3D}\rvert$
electron	28.2	2.44	1.04	-	0.00	0.54
electron	14.1	0.13	-2.84	+	0.28	3.44
muon	14.7	-0.37	0.82	+	0.03	0.26
muon	11.9	2.35	0.42	-	0.03	0.87

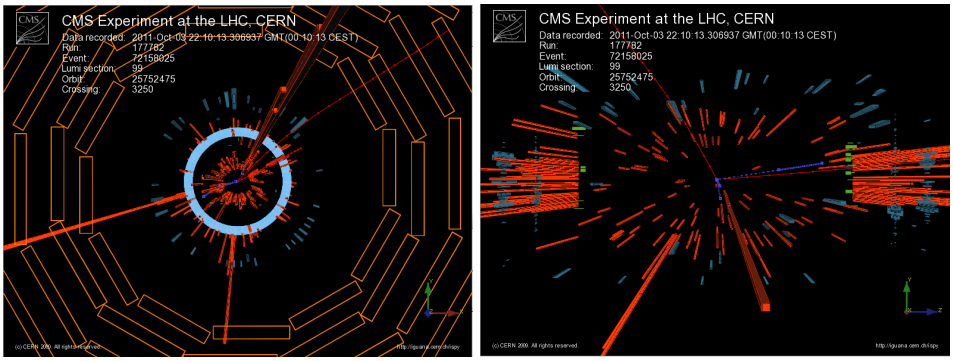


Figure B.7: *(a) R-Phi view and (b) R-Z view of the event AS (Run # 177875 - Event # 148667118). The leading electron of 28.2 GeV/c constituting the Z1 candidate is very close to the acceptance boarder with $\eta = 2.438$. It is ECAL driven only and only starts in the TEC layers with one missing expected hit in the innermost layers. The sub-leading electron consituting the Z1 candidate as a rather high relative isolation and a SIP_{3D} close to the cut value. One of the two muons is also close to the acceptance boarder with an η of 2.35. Overall not a very nice candidate. Electrons are in blue and muons in red.*

Candidate AY ($m_{4\ell} = 114.8\,\text{GeV}/c^2, \Delta m_{4\ell} = 0.77\,\text{GeV}/c^2, m_{Z_1} = 83.4\,\text{GeV}/c^2, m_{Z_2} = 26.0\,\text{GeV}/c^2$).

- 4 muons
- 1 extra electron ($p_T > 6.7\,\text{GeV}/c$) within a jet
- 1 extra isolated SC with E_T = 7.6 GeV
- 6 extra tracks $p_T > 5\,\text{GeV}/c$ (one corresponding to the extra electron within the jet)
- 18 extra jets $p_T > 10\,\text{GeV}/c$ (leading one with $\sim 27\,\text{GeV}/c$)
- MET (PF): 11.6 GeV
- 25 vertices in the event, 952 tracks (all leptons come from the same primary vertex)

Lepton	p_T	η	ϕ	charge	R_{iso}	$\lvert\text{SIP}_{3D}\rvert$
muon	67.4	-0.53	-1.39	+	-0.01	0.35
muon	27.7	-0.73	1.18	-	0.06	0.61
muon	5.8	-1.74	-2.53	+	0.07	3.82
muon	13.9	0.54	-1.90	-	0.10	0.85

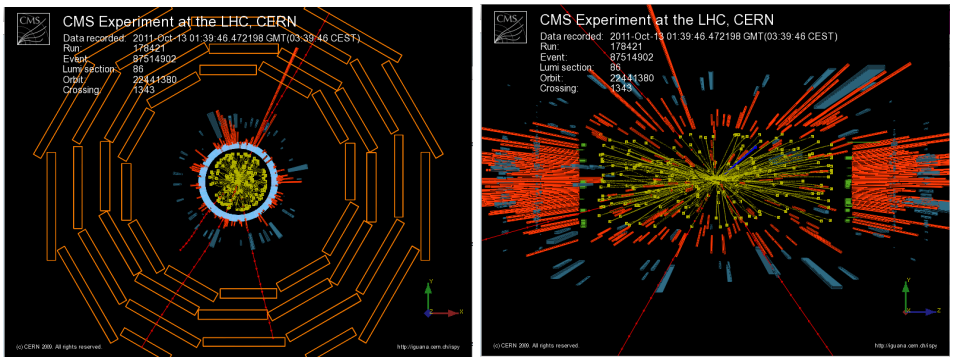


Figure B.8: *(a) R-Phi view and (b) R-Z view of the event AY (Run # 178421 -Event # 87514902). Very high vertex multiplicity (25 vertices, 952 tracks). All tracks with $p_T > 0.8$ GeV/c are shown in yellow on both views. There is one isolated extra SC of $E_T = 7.6$ GeV, close to the muon at $\phi = 1.2$ rad ($\Delta dR = 0.34$). There is one extra electron of 6.7 GeV/c, well away from the four muons and clearly within a jet. The muons are in red.*

Bibliography

[1] N. Amapane *et al.*, "Search for a standard model higgs boson in the decay channel $H \to ZZ^{(*)} \to 4l$ - version v7," *CMS AN 2011/123*.

[2] The CMS Collaboration, "Search for a standard model higgs boson in the decay channel $H \to ZZ^{(*)} \to 4l$," *CMS Physics Analysis Summary*, vol. http://cdsweb.cern.ch/record/1369822, 2011.

[3] N. Amapane *et al.*, "Search for a standard model higgs boson in the decay channel $H \to ZZ^{(*)} \to 4l$ - version v9," *CMS AN*, vol. 2011/123, 2011.

[4] The CMS Collaboration, "Search for a standard model higgs boson produced in the decay channel $H \to ZZ^{(*)} \to 4l$," *CMS Physics Analysis Summary HIG-11-015*.

[5] N. Amapane *et al.*, "Search for a standard model higgs boson in the decay channel $H \to ZZ^{(*)} \to 4l$," *CMS AN*, vol. 2011/387, 2011.

[6] S. L. Glashow, "Partial Symmetries of Weak Interactions," *Nucl. Phys.*, vol. 22, pp. 579–588, 1961.

[7] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.

[8] A. Salam, "Elementary Particle Physics," *Almquvist and Wiksell, Stockholm*, 1968.

[9] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak Interactions with Lepton-Hadron Symmetry," *Phys. Rev.*, vol. D2, pp. 1285–1292, 1970.

[10] E. Fermi, "Trends to a theory of Beta radiation," *Nuovo Cimento*, vol. 11, pp. 1–19, 1934.

[11] S. Dawson, "Introduction to electroweak symmetry breaking," 1998. `arXiv:9901280`.

[12] J. Ellis, J. Espinosa, G. Giudice, A. Hoecker, and A. Riotto, "The Probable Fate of the Standard Model," 2009. `arXiv:0906.0954`.

[13] G. Ridolfi, "Search for the Higgs boson: Theoretical perspectives," 2001. arXiv:0106300.

[14] *LEP design report*. CERN-LEP-84-01, 1984.

[15] P. Proudlock, "LEP200 Design report," Tech. Rep. SL-PC-Tech-Note-91-87, CERN, 1991.

[16] *SLC design handbook. Stanford Linear Collider. Design report*. Stanford, CA: SLAC, 1984.

[17] Tevatron I Group, "Design Report Tevatron 1 Project," 1984. FERMILAB-DESIGN-1984-01.

[18] The LEP Working Group for Higgs boson searches, "Search for the standard model Higgs boson at LEP," *Phys. Lett.*, vol. B565, pp. 61–75, 2003.

[19] CDF and D∅ Collaborations, "Combined CDF and D∅ Upper Limits on Standard Model Higgs-Boson Production with up to 8.6 fb^{-1} of Data ," 2011. arXiv:1107.5518v2.

[20] The LEP Collaborations, the LEP Electroweak Working Group, "A Combination of preliminary electroweak measurements and constraints on the standard model," 2006. arXiv:0612034 [hep-ex].

[21] The LEP Electroweak Working Group. http://lepewwg.web.cern.ch/LEPEWWG/.

[22] Tevatron Electroweak Working Group. http://tevewwg.fnal.gov/.

[23] "LHC Design Report." CERN 2004-003, 2004.

[24] K. Nakamura *et al.*, "Review of particle physics," *J. Phys.*, vol. G37, p. 075021, 2010.

[25] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP*.

[26] R. Adolphi *et al.*, "The CMS experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[27] The CMS Collaboration, "The CMS Physics Technical Design Report, Volume I: Detector performance and software," Tech. Rep. CERN/LHCC 2006-001 and CMS TDR 8.1, CERN, 2006.

[28] The CMS Collaboration, "The Magnet Project Technical Design Report," Tech. Rep. CERN/LHCC 97-10 and CMS TDR 1, CERN, 1997.

[29] The CMS Collaboration, "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays," *JINST*, vol. 5, p. T03021, 2010.

[30] The CMS Collaboration, "The Tracker Project Technical Design Report," Tech. Rep. CERN/LHCC 98-6 and CMS TDR 5 (1998), CERN, 1998.

[31] The CMS Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions." CMS PAS TRK-10-005, 2010.

[32] The CMS Collaboration, "The Electromagnetic Calorimeter Project Technical Design Report," Tech. Rep. CERN/LHCC 97-33 and CMS TDR 4, CERN, 1997.

[33] The CMS Collaboration, "The Hadronic Calorimeter Technical Design Report," Tech. Rep. CERN/LHCC 97-31 and CMS TDR 2, CERN, 1997.

[34] The CMS Collaboration, "The Muon Project Technical Design Report," Tech. Rep. CERN/LHCC 97-32 and CMS TDR 2, CERN, 1997.

[35] S. Chatrchyan *et al.*, "Performance of the CMS Drift Tube Chambers with Cosmic Rays," *JINST*, vol. 5, p. T03015, 2010.

[36] S. Chatrchyan *et al.*, "Performance of the CMS Cathode Strip Chambers with Cosmic Rays," *JINST*, vol. 5, p. T03018, 2010.

[37] The CMS Collaboration, "The TriDAS Project Technical Design Report, Volume I: The Level-1 Trigger," Tech. Rep. CERN/LHCC 2000-038 and CMS TDR 6.1, CERN, 2000.

[38] W. Adam *et al.*, "The CMS high level trigger," *Eur. Phys. J.*, vol. C46, pp. 605–667, 2006.

[39] S. Chatrchyan *et al.*, "Performance of the CMS Level-1 Trigger during Commissioning with Cosmic Ray Muons and LHC beams," *JINST*, vol. 5, p. T03002. 49 p, Dec 2009.

[40] M. Stettler *et al.*, "The CMS Global Calorimeter Trigger hardware design," in *Proceedings of the 12th Workshop on Electronics for LHC and Future Experiments*, 2006.

[41] G. Iles *et al.*, "Revised CMS global calorimeter trigger functionality and algorithms," in *Proceedings of the 12th Workshop on Electronics for LHC and Future Experiments*, 2006.

[42] H. Sakulin and A. Taurok, "Implementation and test of the first-level Global Muon Trigger of the CMS experiment," in *Proceedings of the 11th Workshop on Electronics for LHC and Future Experiments*, 2005.

[43] L. Guiducci *et al.*, "DT Sector Collector electronics design and construction," in *Proceedings of the Topical Workshop on Electronics for Particle Physics*, 2007.

[44] D. Acosta *et al.*, "A 3-D track-finding processor for the CMS level-1 muon trigger," in *Proceedings of the Conference for Computing in High-Energy and Nuclear Physics*, 2003.

[45] RD5 Collaboration, "Pattern Comparator Trigger (PACT) for the muon system of the CMS experiment," *Nucl. Instrum. Meth.*, vol. A370, pp. 389–395, 1996.

[46] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables," *CERN-2011-002*, CERN, Geneva, 2011.

[47] T. Junk, "Confidence level computation for combining searches with small statistics," *Nucl.Instrum.Meth.*, vol. A434, pp. 435–443, 1999.

[48] A. L. Read, "Modified frequentist analysis of search results (the CLs method)," *CERN Yellow Report*, vol. CERN-2000-005, p. 81, 2000.

[49] ATLAS Collaboration, CMS Collaboration, and LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011," *ATL-PHYS-PUB-2011-818, CMS NOTE-2011/005*, August 2011.

[50] The CMS Collaboration, "Search for standard model higgs boson in pp collisions at $\sqrt{s} = 7tev$ and integrated luminosity up to $1.7fb^{-1}$," *CMS Physics Analysis Summary*, vol. CMS-PAS-HIG-11-022, 2011.

[51] The CMS Collaboration, "Search for the Standard Model Higgs Boson decaying to bottom quarks and produced in association with a W or a Z boson," *CMS Physics Analysis Summary*, vol. CMS-PAS-HIG-11-012, 2011.

[52] The CMS Collaboration, "Search for neutral higgs bosons decaying to tau pairs in pp collisions at $\sqrt{s} = 7tev$," *CMS Physics Analysis Summary*, vol. CMS-PAS-HIG-11-020, 2011.

[53] The CMS Collaboration, "Search for a higgs boson decaying into two photons in the cms detector," *CMS Physics Analysis Summary*, vol. CMS-PAS-HIG-11-021, 2011.

[54] The CMS Collaboration, "Search for the higgs boson decaying to w^+w^- in the fully leptonic final state," *CMS Physics Analysis Summary*, vol. CMS-PAS-HIG-11-014, 2011.

[55] The CMS Collaboration, "Search for the standard model higgs boson in the decay channel H $\rightarrow$ ZZ $\rightarrow 2\ell 2q$ at cms," *CMS Physics Analysis Summary*, vol. CMS-PAS-HIG-11-017, 2011.

[56] The CMS Collaboration, "Search for the higgs boson in the H $\rightarrow$ ZZ $\rightarrow 2\ell 2\nu$ channel in pp collisions at $\sqrt{s} = 7tev$," *CMS Physics Analysis Summary*, vol. CMS-PAS-HIG-11-016, 2011.

[57] The Geant4 Collaboration, "GEANT – Detector Description and Simulation Tool."
`http://geant4.web.cern.ch/geant4/`.

[58] S. M. T. Sjostrand and P. Z. Skands, "Pythia 6.4 physics and manual," *JHEP*, vol. 0605, 2006.

[59] J. Alwall *et al.*, "MadGraph/MadEvent v4: The New Web Generation," *JHEP*, vol. 09, p. 028, 2007.

[60] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP*, vol. 11, p. 070, 2007.

[61] T. Binoth, N. Kauer, and P. Mertsch, "Gluon-induced QCD corrections to $pp \rightarrow ZZ \rightarrow \ell\bar{\ell}\ell'\bar{\ell}'$," 2008.

[62] C. Anastasiou, R. Boughezal, and F. Petriello, "Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion," *JHEP*, vol. 04, p. 003, 2009.

[63] D. de Florian and M. Grazzini, "Higgs production through gluon fusion: updated cross sections at the Tevatron and the LHC," *Phys. Lett.*, vol. B674, pp. 291–294, 2009.

[64] S. Dawson, "Radiative corrections to Higgs boson production," *Nucl. Phys.*, vol. B359, pp. 283–300, 1991.

[65] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas, "Higgs boson production at the LHC," *Nucl. Phys.*, vol. B453, pp. 17–82, 1995.

[66] R. V. Harlander and W. B. Kilgore, "Next-to-next-to-leading order Higgs production at hadron colliders," *Phys. Rev. Lett.*, vol. 88, p. 201801, 2002.

[67] C. Anastasiou and K. Melnikov, "Higgs boson production at hadron colliders in NNLO QCD," *Nucl. Phys.*, vol. B646, pp. 220–256, 2002.

[68] V. Ravindran, J. Smith, and W. L. van Neerven, "NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions," *Nucl. Phys.*, vol. B665, pp. 325–366, 2003.

[69] S. Catani, D. de Florian, M. Grazzini, and P. Nason, "Soft-gluon resummation for Higgs boson production at hadron colliders," *JHEP*, vol. 07, p. 028, 2003.

[70] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, "NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders," *Phys. Lett.*, vol. B670, pp. 12–17, 2008.

[71] M. Ciccolini, A. Denner, and S. . Dittmaier, "Strong and electroweak corrections to the produc tion of Higgs + 2-jets via weak interactions at the LHC," *Phys. Rev. Lett.*, vol. 99, p. 161803, 2007.

[72] M. Ciccolini, A. Denner, and S. e. Dittmaier, "Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC," *Phys. Rev.*, vol. D77, p. 013002, 2008.

[73] T. Figy, C. Oleari, and D. Zeppenfeld, "Next-to-leading order jet distributions for Higgs boson production via weak-boson fusion," *Phys. Rev.*, vol. D68, p. 073005, 2003.

[74] K. Arnold *et al.*, "VBFNLO: A parton level Monte Carlo for processes with electroweak bosons," *Comput. Phys. Commun.*, vol. 180, pp. 1661–1670, 2009.

[75] P. Bolzoni, F. Maltoni, and M. Moch, Sven-Olaf a nd Zaro, "Higgs production via vector-boson fusion at NNLO in QC D," *Phys. Rev. Lett.*, vol. 105, p. 011801, 2010.

[76] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, "Precise predictions for the Higgs-boson decay H $\to$ WW/ZZ $\to$ 4 leptons," *Phys. Rev.*, vol. D74, p. 013004, 2006.

[77] A. Bredenstein, A. Denner, S. Dittmaier, and M. Weber, "Radiative corrections to the semileptonic and hadronic Higgs-boson decays H $\to$W W / Z Z$\to$ 4 fermions," *JHEP*, vol. 0702, p. 080, 2007.

[78] A. Djouadi, J. Kalinowski, M. Muhlleitner, and M. Spira, "An update of the program HDECAY," in *The Les Houches 2009 workshop on TeV colliders: The tools and Monte Carlo working group summary report*, 2010.

[79] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, "NNLO Computational Techniques: the Cases $H \to \gamma\gamma$ and $H \to gg$," *Nucl. Phys.*, vol. B811, pp. 182–273, 2009.

[80] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, "Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC," *Nucl.Phys.*, vol. B737, pp. 73–120, 2006.

[81] S. Baffioni *et al.*, "Search strategy for the higgs boson in the $zz^{(*)}$ decay channel at $\sqrt{s} = 10$ with the cms experiment'," *CMS AN*, vol. 2010/237, 2010.

[82] G. Abbiendi *et al.*, "Muon Reconstruction in the CMS Detector." CMS AN-2008/097, July 2009.

[83] D. Trocino, *Muon Reconstruction and Momentum Scale Calibration and Their Application to Standard Model Higgs Searches with the CMS Experiment* (Chapter 3). PhD thesis, CMS TS-2011/042, Università degli Studi di Torino, 2011. Defended on 27 Jan 2011.

[84] N. Amapane *et al.*, "Local Muon Reconstruction in the Drift Tube Detectors." CMS AN-2009/008, June 2009.

[85] The CMS Collaboration, "Performance of the CMS muon detector using pp collisions at $\sqrt{s}$ = 7 TeV." CMS PAS MUO-11-001, In Preparation.

[86] S. Chatrchyan *et al.*, "Performance Study of the CMS Barrel Resistive Plate Chambers with Cosmic Rays," *JINST*, vol. 5, p. T03017, 2010.

[87] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting," *Nucl. Instrum. Meth.*, vol. A262, pp. 444–450, 1987.

[88] V. Khachatryan *et al.*, "CMS Tracking Performance Results from early LHC Operation," *Eur. Phys. J.*, vol. C70, pp. 1165–1192, 2010.

[89] S. Chatrchyan *et al.*, "Performance of CMS Muon Reconstruction in Cosmic-Ray Events," *JINST*, vol. 5, p. T03022, 2010.

[90] The CMS Collaboration, "Measurement of tracking efficiency," *CMS Physics Analysis Summary*, vol. CMS-PAS-TRK-10-002, 2010.

[91] The CMS Collaboration, "Measurements of inclusive w and z cross sections in pp collisions at $\sqrt{s}$ = 7 tev," *J. High Energy Phys.*, vol. 01, p. 080, 2011.

[92] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas," *Phys. Lett.*, vol. B659, pp. 119–126, 2008.

[93] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets," *JHEP*, vol. 04, p. 005, 2008.

[94] The CMS Collaboration, "Commissioning of the CMS experiment and the cosmic run at four tesla," *JINST*, vol. 5, p. T03001, 2010.

[95] S. Chatrchyan *et al.*, "Performance of CMS Muon Reconstruction in Cosmic-Ray Events," *JINST*, vol. 5, p. T03022, 2010.

[96] The CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ tev," *CMS Physics Analysis Summary*, vol. CMS-PAS-MUO-10-002, 2010.

[97] The CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s}$ = 7 TeV." CMS PAS MUO-10-004, 2010.

[98] E. Woehri *et al.*, "BPH Trigger CMS Twiki." `https://espace.cern.ch/cms-quarkonia/trigger-bph/default.aspx`.

[99] T. Sjostrand, S. Mrenna, and P. Skands, "Pythia 6.4 physics and manual," *JHEP*, vol. 05, p. 026, 2006.

[100] S. Chatrchyan *et al.*, "Measurement of the Underlying Event Activity at the LHC with $\sqrt{s}$ = 7 TeV and Comparison with $\sqrt{s}$ = 0.9 TeV," *submitted to JHEP*, 2011.

[101] D. Lange, "The $t\bar{t}$ evtgen particle decay simulation package," *Nucl. Instrum. Meth.*, vol. A462, pp. 152–155, 2001.

[102] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, "MadGraph/MadEvent v4: the new web generation," *JHEP*, vol. 09, p. 028, 2007.

[103] The CMS Collaboration, "Measurement of tracking efficiency," *CMS PAS*, vol. TRK-10-002, 2010.

[104] S. Bolognesi *et al.*, "Calibration of track momentum using dimuon resonances in CMS." CMS AN-2010/059, March 2010.

[105] S. Baffioni *et al.*, "Electron reconstruction in cms," *CMS Analysis Note*, vol. 2009/164, 2009.

[106] The CMS Collaboration, "Electron reconstruction and identification at $\sqrt{s} = 7$ tev," *CMS Physics Analysis Summary*, vol. CMS-PAS-EGM-10-004, 2010.

[107] The CMS Collaboration, "Electron commissioning results at at $\sqrt{s} =$ 7 tev," *CMS Detector Performance Summaries*, vol. 2011/003, 2011.

[108] S. Baffioni *et al.*, "Electron charge determination commissioning from 7 tev data," *CMS AN*, vol. 2010/468, 2010.

[109] M. Aldaya *et al.*, "Discovery potential and search strategy for the standard model Higgs boson in the $H \to ZZ^{\star} \to 4\mu$ decay channel using a mass-independent analysis," *CMS Note*, vol. 2006/106, 2006.

[110] M. Aldaya *et al.*, "A method for determining the mass of the standard model higgs boson using the $h \to 4\ell$ decay channel," *CMS Note*, vol. 2006/107, 2006.

[111] S. Baffioni *et al.*, "Discovery potential for the SM Higgs boson in the H$\to$ $ZZ^{(*)} \to 4e$ decay channel," *J. Phys. G*, vol. 34, pp. N23–N46, 2007.

[112] S. Abdullin *et al.*, "Search strategy for the standard model higgs boson in the h $\to$ zz $\to 4\mu$ decay channel," *CMS Note*, vol. 2006/122, 2006.

[113] D. Futyan *et al.*, "Search for the standard model higgs boson in the two-electron and two-muon final state with cms," *CMS Note*, vol. 2006/136, 2006.

[114] G. Bayatian *et al.*, "CMS technical design report, volume II: Physics performance," *J.Phys.G*, vol. G34, pp. 995–1579, 2007.

[115] The CMS Collaboration, "Search strategy for the higgs boson in the zz decay channel with the cms experiment," *CMS Physics Analysis Summary*, vol. CMS-PAS-HIG-10-003, 2008.

[116] S. Baffioni *et al.*, "Search strategy for the higgs boson in the zz decay channel with the cms experiment," *CMS AN*, vol. 2008/050, 2008.

[117] The CMS Collaboration, "ATLAS+CMS combination of summer 11 results,"

[118] P. Avery *et al.*, "A complete model of the four-lepton invariant mass distributions for H $\to$ ZZ $\to$ 4l and ZZ $\to$4l events," *CMS Analysis Note*, vol. 2011/202, 2011.

[119] The CMS Collaboration, "Absolute calibration of the luminosity measurement at cms," *CMS Detector Performance Summaries*, vol. CMS-DP-2011-002, 2011.

[120] The CMS Collaboration, "Performance of muon reconstruction and identification in pp collisions at $\sqrt{s} = 7$ tev," *CMS Physics Analysis Summary*, vol. CMS-PAS-MUO-10-004, in preparation.

[121] S. Nourbakhsh *et al.*, "Energy scale of cms electromagnetic calorimeter with 2010 data," *CMS Analysis Note*, vol. 2011/038, 2011.

[122] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), "Handbook of LHC Higgs cross sections: 2. Differential Distributions," *to be published*.

[123] C. W. John Campbell, Keith Ellis, "Mcfm - monte carlo for femtobarn processes." 2011.

[124] G. Cowan, K. Cranmer, E. Gross and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J.*, vol. C71, pp. 1–19, 2011.