ALMA MATER STUDIORUM - UNIVERSITÀ di BOLOGNA

Dottorato di ricerca in
Fisica

Ciclo XXV

Settore concorsuale di afferenza: 02/A1
Settore scientifico disciplinare: FIS/01

Evidence for a new boson in the search for the Standard Model Higgs particle decaying to four leptons at CMS

Dott. Marco Meneghelli

Relatore: Prof. Francesco L. Navarria

Correlatori: Dott. Francesca R. Cavallo
Dott. Paolo Giacomelli
Dott. Sylvie Braibant

Coordinatore dottorato: Dott. Fabio Ortolani

Esame finale anno 2013

CERN-THESIS-2013-009
20/03/2013

Contents

List of Figures

Preface

The Standard Model (SM) of elementary interactions, a quantum field theory based on symmetry principles, successfully describes the dynamics of the elementary particles and the fundamental interactions: the strong, the weak and the electromagnetic forces. The SM describes the matter as composed of twelve elementary particles, the fermions, that interact exchanging vector bosons, carriers of the fundamental interactions. The unification of the weak and electromagnetic forces occurs via a spontaneous symmetry breaking mechanism that gives rise to the mass of the bosons carriers of the weak force and predicts the existence of an elementary scalar boson, the Higgs particle. The SM has been tested with great precision in many experiments during the past decades: its predictive power has been verified up to the second perturbative order for the electroweak sector, the description of the strong interaction matches the experimental observations for energy scales spanning ten orders of magnitude and flavor physics has been extensively tested, revealing processes that violate some discrete symmetries. The Higgs boson remains nowadays still undetected, and its mass, a free parameter of the SM, unknown. Both direct and indirect searches took place, leading to exclude its existance in a large mass region and to individuate a preferred mass interval, where the boson is most likely to be found. The study of the electroweak symmetry breaking mechanism, and of the Higgs particle, considered to be the signature of the process, is one of the most important topics of modern particle physics.

The "Large Hadron Collider" (LHC), at the European Laboratory for Nuclear Research (CERN) in Geneva, is a proton-proton collider, designed to work at a center of mass energy of 14 TeV and with a peak instantaneous luminosity of 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Its operations with proton beams started in 2009. This thesis has been developed within the Compact Muon Solenoid (CMS) experiment at the LHC, whose main goal is the discovery of the Higgs boson, studying the outcome of high energy proton-proton collisions. One of the most important analyses of CMS is here presented: the search for the Higgs boson in the "four-lepton" channel. This is one of the analyses providing the best chances of discovery of the particle: it fully makes use of the excellent performances of CMS in terms of particle reconstruction and identification, particle detection efficiency and its capability to trigger hard

scattering events among the hadronic background. The dataset used for this study comprises data collected during 2011 (with LHC working at a center of mass energy of 7 TeV) and 2012 (at 8 TeV), for a statistics of 5 + 12 fb^{-1}.

The first chapter of the thesis presents the fundamental theory that underlies the physics at the LHC, the Standard Model. A description of the general strategy for the search of the Higgs boson at a hadron collider, the current bounds on its properties and the results nowadays available are reported.

The second chapter describes the CMS experiment at the LHC, highlighting the main features of the sub-detectors, driven by physical requirements. The performances during data taking periods (from 2010 to 2012) are summarized.

The third chapter is devoted to the description of methodologies used to reconstruct and identify the physical objects at CMS, using the information collected by the detection apparatus: muons, electrons, photons, hadron jets, taus and missing transverse energy from energy deposits, ionization tracks and energy imbalance. Particular emphasis is placed on lepton reconstruction (muons and electrons), since they are the key objects used in the analysis.

Chapters 4-6 are the core of the thesis: the search for the Higgs boson in the four-lepton channel is described. The analysis seeks a Higgs boson signal in the decay channel $H \rightarrow ZZ \rightarrow l^+l^-l^+l^-$, with $l = \mu, e$. The experimental signature of the events in this channel is the presence of four isolated leptons coming from the primary interaction vertex. The Higgs boson would be revealed by the presence of a narrow resonance peak, over the broader background, in the 4-lepton invariant mass spectrum. The signal rate is relatively small, but the experimental signature is very clean, permitting the peak reconstruction and the Higgs mass measurement with a resolution of a GeV/c^2 order. The signal over background ratio is locally of the order of one-to-one, anywhere in the mass range, the main backgrounds being the non resonant double Z production and a small contamination from $Z + jets$ and $t\bar{t}$ processes. In chapter 4 the analysis is approached studying the single Z decaying to a lepton pair, a channel that benefits of large statistics allowing a better understanding of the systematics associated to lepton objects. In chapter 5 the event selection criteria are described: their main aim is to maximize the signal over background ratio. The signal and background yields expected to survive the selection are evaluated, and techniques to control backgrounds are developed using data-driven approaches, in support of predictions from simulations. In chapter 6 the modeling of the statistical frame devoted to decide weather a signal is actually present is described. The expected results are exclusion limits for the SM Higgs boson if background only is observed. In case of a signal-like excess of data, its statistical significance is evaluated and its properties are studied.

Chapter 1

The Standard Model Higgs boson

In this first chapter the Standard Model of elementary interactions and the Higgs mechanism are briefly presented. A description of the general strategy for the search of the Higgs boson at a hadron collider and the currently available experimental results are reported.

Natural units are used in the following ($\hbar = c = 1$), unless specified.

1.1 The Standard Model of elementary interactions

The *Standard Model (SM)* [1][2][3] is the physical theory that currently best describes the behavior of elementary particles. It is a Quantum Field Theory (QFT) [4][5] built on symmetry principles[1]: it includes the QFT of the electroweak interaction (Glashow-Weinberg-Salam model, GWS) and of the strong interaction (Quantum Chromo Dynamics, QCD). Core of the theory is the ElectroWeak Symmetry Breaking (EWSB), via the Higgs mechanism, that gives rise to the mass of the vector bosons and predicts the existence of the Higgs boson. Three out of four of the fundamental forces are described by the theory: the strong, the weak and the electromagnetic (EM) force. The gravitational interaction is not taken into account, as it is not relevant at the scales of mass and distance typical in particle physics.

The SM describes matter as being composed of twelve elementary particles, the *fermions*, all having spin 1/2. The fermions can be divided into two main groups, *leptons* and *quarks*, and in three families, as reported in the following table:

[1]The reader is supposed to be familiar with the QFT formalism. For a QFT introduction see [4][5]

Fermions	1^{st} fam.	2^{nd} fam.	3^{rd} fam.	Interactions
Quarks	u	c	t	All
	d	s	b	
Leptons	ν_e	ν_μ	ν_τ	Weak
	e	μ	τ	Weak, EM

The interactions between particles are described in terms of exchange of *vector bosons*, spin 1 particles, carriers of the fundamental interactions:

Interaction	Boson	Range[cm]
EM	Photon (A or γ)	∞
Weak	$W^{\pm}$, Z	10^{-16}
Strong	Gluons (g)	10^{-13}

Each particle is the quantum associated to an elementary field $\Phi_i(x)$ in the Minkowsky space.

The SM is built around relativistic invariance: the fields (and, consequently, the associated particles) are classified depending on how they transform under a Lorentz transformation, the Lagrangian of the theory is constructed using the possible scalar combinations of the fields (up to terms of dimension 4 in mass, since the others are irrelevant in the QFT sense).

The theory is required to be locally invariant under transformations of the Lie group:

$$SU_c(3) \times SU_L(2) \times U_Y(1)$$

A theory of this type is called *gauge theory*. This requirement gives rise to the vector boson fields $V_i(x)$, as will be described in the following sections.

The SM Lagrangian can be divided into three parts: the bosonic, the fermionic and the Yukawa sectors:

$$L_{SM} = L_B + L_F + L_Y \tag{1.1}$$

It contains the description of elementary interactions, in terms of a quantum theory: the transition probabilities between quantum states are calculable from elementary assumptions: basically the most important predictable quantities are cross sections and decay rates in subnuclear processes. The three sectors will be briefly described in the next sections[2].

[2]Such description aims to be nothing more than an overview: it does not consider the problematics associated with the quantization of a non-abelian gauge theory: e.g., the renormalization issue, the necessary introduction of ghost fields, etc. For a complete review see [4][5]

1.1.1 The bosonic sector: the electroweak symmetry breaking and the Higgs mechanism

The bosonic sector is built around the complex scalar field (that has two components, and so four degrees of freedom):

$$\phi = \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix}$$

which is required to be locally invariant under the transformations of the Lie group:

$$SU_L(2) \times U_Y(1) \ni \Omega = e^{ig\alpha_a t^a} = e^{ig_L \alpha_L \ {}_b T^b + ig_Y \alpha_Y Y} \tag{1.2}$$

This group has four generators $(t^a)_{a=1,2,3,4}$: $(T^b)_{b=1,2,3}$ for $SU_L(2)$ and Y for $U_Y(1)$, and acts on the field ϕ:

$$\phi' = \Omega\phi$$

The group 1.2 is not simple, having two invariant sub-groups $SU_L(2)$ (L stands for "left") and $U_Y(1)$ (Y stands for "hypercharge"). The corresponding QFT has thereby two coupling constants g_L and g_Y, which are free parameters of the theory. The scalar field ϕ is a doublet of $SU_L(2)$ and has hypercharge 1/2.

Composed of Lorentz and gauge invariant terms, the bosonic part of L_{SM} can be written:

$$L_B = D^\mu \phi^\dagger D_\mu \phi - \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2 - \frac{1}{4} F_{a\mu\nu} F_a^{\mu\nu} \tag{1.3}$$

where D_μ is the covariant derivative, containing the gauge fields A_μ^a (= W^b,B):

$$D_\mu = \partial_\mu + ig_L T_b W_\mu^b + ig_Y Y B_\mu$$

$F_{a\mu\nu}$ is the tensor:

$$F_{a\mu\nu} = \partial_\mu A_{a\nu} - \partial_\nu A_{a\mu} - g f_a^{bc} A_{b\mu} A_{c\nu}$$

and f_a^{bc} are the structure constants of the group. Note that no mass terms are present in the Lagrangian: all the fields (ϕ and A_μ^a) are massless so far.

The potential

$$V(\phi^\dagger \phi) = \lambda(\phi^\dagger \phi)^2 + \mu^2 \phi^\dagger \phi$$

can have different properties depending on the values of its parameters, μ^2 and λ. λ must be positive in order to have a physical potential that grows asymptotically with fields, while μ^2 can be either positive or negative, leading to two different scenarios:

- $\mu^2 > 0$: all the bosonic fields are massless, the scalar field is quantized around the unique minimum of potential, that has the symmetry properties of the Lagrangian

- $\mu^2 < 0$: the potential has infinite minima for each field configuration satisfying

$$(\phi^1)^2 + (\phi^2)^2 = -\frac{\mu^2}{\lambda} = v^2$$

the scalar field can be quantized around an arbitrary minimum, that does not conserve the symmetry properties of the Lagrangian anymore. The typical choice for the minimum is

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix} \tag{1.4}$$

The symmetry is spontaneously broken, meaning that the Lagrangian keeps the symmetry properties while the vacuum state does not. The spontaneous symmetry breaking leaves a residual $U(1)$ symmetry still evident, called the electromagnetic symmetry group $U_{em}(1)$

As observed experimentally, the SM has a spontaneously broken symmetry, so $\mu^2 < 0$. The original symmetry $SU_L(2) \times U_Y(1)$ breaks down to $U_{em}(1)$:

$$SU_L(2) \times U_Y(1) \rightarrow U_{em}(1)$$

Three out of four Lie group dimensions are no more evident. The generator of the $U_{em}(1)$ symmetry group, called electromagnetic charge Q, is linked to the hypercharge group generator Y and to the third generator of $SU_L(2)$, T_3 by:

$$Q = Y + T_3$$

The kinetic term of 1.3, evaluated at the minimum 1.4, yields these relevant terms:

$$L_{mass} = \frac{1}{\sqrt{2}}(0 \quad v)(g_L T_b W_\mu^b + g_Y Y B_\mu)(g_L T^b W_b^\mu + g_Y Y B^\mu)\begin{pmatrix}0\\v\end{pmatrix}$$

Using the explicit form of the $SU_L(2) \times U_Y(1)$ generators, in the doublet representation:

$$T_b = \frac{\sigma^b}{2} \qquad Y = \frac{\mathbf{1}}{2}$$

where σ^b are the Pauli matrices, we can evaluate the matrix products explicitly:

$$L_{mass} = \frac{v^2}{8}[g_L^2(W_\mu^1)^2 + g_L^2(W_\mu^2)^2 + (-g_L^2 W_\mu^3 + g_Y B_\mu)^2] \tag{1.5}$$

As we can read from 1.5 (the quadratic terms in the Lagrangian are mass terms), three combinations of the fields A^a_μ acquire mass. They are the boson carriers of the weak interaction:

$$W^\pm_\mu = \frac{W^1_\mu \mp iW^2_\mu}{\sqrt{2}} \quad \text{with mass} \quad m_W = \sqrt{-\frac{\mu^2}{8\lambda}}\, g_L$$

and

$$Z_\mu = \cos\theta_W W^3_\mu - \sin\theta_W B_\mu \quad \text{with mass} \quad m_Z = \sqrt{-\frac{\mu^2}{8\lambda}(g_L^2 + g_Y^2)^2}$$

θ_W is the so called Weinberg angle, related to the coupling constants (and experimentally measured):

$$\tan\theta_W = \frac{g_Y}{g_L} \simeq 0.24 \text{ rad}$$

The fields $W^\pm$ and Z gain a degree of freedom, after the spontaneous symmetry breaking, "eating" a component of the scalar field ϕ: this process is known as the "Higgs mechanism".

The fourth combination of the fields A^a_μ, associated to the $U_{em}(1)$ group, keeps two degrees of freedom and remains massless. It is the photon:

$$A_\mu = \sin\theta_W W^3_\mu + \cos\theta_W B_\mu$$

The fourth real component of the field ϕ remains free after the spontaneous symmetry breaking. It is therefore a free field with an associated particle that, after the spontaneous symmetry breaking, acquires mass: the *Higgs boson H*. To investigate this question we can work in the unitarity gauge, parametrizing the field ϕ as follows:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$$

where $h(x)$ is a real scalar field, with null expectation value on vacuum. This parametrization makes explicit that ϕ has one degree of freedom after the symmetry breaking. The term containing the potential, in 1.3, takes the form:

$$V(h) = -\mu^2 h^2 + \sqrt{-\mu^2\lambda}\; h^3 + \frac{1}{4}\lambda h^4$$

The Higgs boson is the particle associated with the field $h(x)$, with mass:

$$m_H = \sqrt{-2\mu^2}$$

The expansion of the kinetic energy term in 1.3, in unitarity gauge, yields the terms 1.5, that contain the masses, plus additional terms involving the coupling between vector bosons and the Higgs field h:

$$L_K = \frac{1}{2}(\partial_\mu h)^2 + \left[m_W^2 W^{\mu+} W_\mu^- + \frac{1}{2} m_Z^2 Z^\mu Z_\mu\right]\left(1 + \frac{h}{v}\right)^2 \tag{1.6}$$

The couplings of the Higgs particle to the gauge vector bosons are proportional to the squared masses of the bosons.

The bosonic sector of the SM has four parameters: the two coupling constants g_L and g_Y, and the potential parameters μ and λ. A summary table of the SM bosons and their masses, as functions of L_B parameters is [8]:

Boson	**mass^2**	**measured mass [GeV/c^2]**
$W^\pm$	$-\frac{\mu^2 g_L^2}{8\lambda}$	80.385 ± 0.015
Z	$-\frac{\mu^2(g_L^2+g_Y^2)}{8\lambda}$	91.1876 ± 0.0021
A	0	0 ($< 10^{-26}$)
H	$-2\mu^2$	unknown

The parameters can be combined to form other four constants more easily measured by the experiments. Conventionally the four parameters of L_B are expressed using:

- *The EM fine-structure constant* (at low energy): $\alpha_{EM} = \frac{1}{137.035999679(94)}$
- *The Fermi constant*: $G_F = 1.1663787(6) \times 10^{-5}$ GeV^{-2}
- *The Z boson mass*: $m_Z = 91.1876(21)$ GeV/c^2
- *The Higgs boson mass*: unknown

1.1.2 The fermionic sector

The fermionic fields of the SM are Dirac four-component fields ψ_i. The chirality operator acts on the fields and divides them in left-handed and right-handed components:

$$\psi_{L,R} = \frac{\mathbf{1} \mp \gamma_5}{2}\psi \qquad \psi = \psi_L + \psi_R$$

Depending on their chirality, the fields transform differently under transformations of the $SU_L(2) \times U_Y(1)$ group. The left-handed components of fields of leptons (l_L) and quarks (q_L) of the three families are grouped into doublets of $SU_L(2)$:

$$\begin{pmatrix}\nu_{eL}\\ e_L\end{pmatrix} \quad \begin{pmatrix}\nu_{\mu L}\\ \mu_L\end{pmatrix} \quad \begin{pmatrix}\nu_{\tau L}\\ \tau_L\end{pmatrix} \qquad \begin{pmatrix}u_L\\ d_L\end{pmatrix} \quad \begin{pmatrix}c_L\\ s_L\end{pmatrix} \quad \begin{pmatrix}t_L\\ b_L\end{pmatrix}$$

The right-handed fields (l_R,q_R) are singlets of $SU_L(2)$, and do not transform under its action. Right-handed neutrinos are not included in the theory. Both right-handed and left-handed fields transform under the hypercharge group $U_Y(1)$. They carry both Y and T_3 charges. The following table summarizes charges of leptons l and quarks q:

Field	Y	T_3	Q
ν_L	-1/2	1/2	0
l_L	-1/2	-1/2	-1
l_R	-1	0	-1
q_L^u	1/6	1/2	2/3
q_L^d	1/6	-1/2	-1/3
q_R^u	2/3	0	2/3
q_R^d	-1/3	0	-1/3

The fermionic Lagrangian can be written (the sum over families is tacit):

$$\begin{aligned} L'_F &= i\bar{l}_L\gamma^\mu(\partial_\mu + ig_L\frac{\sigma^b}{2}W_{b\mu} + ig_Y(-\frac{1}{2})B_\mu)l_L + \\ & i\bar{l}_R\gamma^\mu(\partial_\mu + ig_Y(-1)B_\mu)l_R + \\ & i\bar{q}_L\gamma^\mu(\partial_\mu + ig_L\frac{\sigma^b}{2}W_{b\mu} + ig_Y\frac{1}{6}B_\mu)q_L + \\ & i\bar{q}^u_R\gamma^\mu(\partial_\mu + ig_Y\frac{2}{3}B_\mu)q^u_R + i\bar{q}^d_R\gamma^\mu(\partial_\mu + ig_Y(-\frac{1}{3})B_\mu)q^d_R \end{aligned} \tag{1.7}$$

This Lagrangian contains couplings between fermions and vector bosons. Consequently, after the electroweak spontaneous symmetry breaking, the coupling terms between fermions and massive bosons $W^\pm$,Z and A are:

- the electromagnetic interaction term:

$$L_{em} = -eA_\mu j^\mu_{em} \qquad j^\mu_{em} = -\bar{l}\gamma^\mu l + \frac{2}{3}\bar{q}^u\gamma^\mu q^u - \frac{1}{3}\bar{q}^d\gamma^\mu q^d$$

- the weak neutral-current interaction term:

$$L_Z = -\sqrt{g_L^2 + g_Y^2}Z_\mu j^\mu_Z \qquad j^\mu_Z = \bar{l}_L\gamma^\mu\frac{\sigma^3}{2}l_L + \bar{q}_L\gamma^\mu\frac{\sigma^3}{2}q_L - \sin\theta_W j^\mu_{em}$$

- the weak charged-current interaction term:

$$L_W = -\frac{g_L}{\sqrt{2}}(W^+_\mu j^{\mu-} + h.c.) \qquad j^{\mu-}_W = \bar{\nu}_L\gamma^\mu l_L + \bar{q}^u_L\gamma^\mu q^d_L$$

No new parameters are introduced in formula 1.7: the couplings between fermions and vector bosons depend on the 4 parameters of the bosonic sector of the SM. Note that no mass terms for fermions like $m(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L)$ are

permitted, since they would explicitly break the gauge symmetry: the right-handed and left-handed fermions belong to different $SU_L(2)$ representations and have different $U(1)$ charges: the fermions are massless in L'_F.

The *QCD* is introduced in the SM, postulating the $SU_c(3)$ local gauge invariance for the quark fields. The $SU_c(3)$ Lie group has 8 generators, each giving rise to a correspondent massless gluon field. The QCD Lagrangian can be written in a very compact way as:

$$L_{QCD} = \overline{\psi_q}(i\gamma^\mu D_\mu)\psi_q - \frac{1}{4}G_{a\mu\nu}G_a^{\mu\nu} \tag{1.8}$$

where D_μ is the covariant derivative, containing the 8 gluon gauge fields $(g_\mu^a)_{a=1,\dots,8}$:

$$D_\mu = \partial_\mu + ig_{QCD}G_a g_\mu^a$$

$G_{\mu\nu}$ is the tensor:

$$G_{a\mu\nu} = \partial_\mu g_{a\nu} - \partial_\nu g_{a\mu} - g_{QCD}f_a^{bc}g_{b\mu}g_{c\nu}$$

and f_a^{bc} are the structure constants of the group. The L_{QCD} charge is the so called color: each quark can express itself in three possible colors, the color index is tacit in formula 1.8.

The QCD is an asymptotically free theory, meaning that the running coupling constant $g_{QCD}(Q^2)$ becomes small in high energy-transfer processes: in QCD, a perturbative calculation of cross sections is possible only at high energy-transfer.

Adding the QCD Lagrangian, the formula for the fermionic sector of the SM is:

$$L_F = L'_F + L_{QCD}$$

1.1.3 The Yukawa sector

It is possible to write another term in the SM Lagrangian, which satisfies all the postulated symmetries. Thanks to this term, the masses of fermions are introduced in the SM. It is a sum of Yukawa-interaction terms that couple the fermions to the scalar field ϕ:

$$\begin{aligned} L_Y \;=\; - \sum_{i,j\in families} \Bigg(& \bar{l}_{iR}(g_L)_{ij}\phi^\dagger \begin{pmatrix} \nu_{jL} \\ l_{jL} \end{pmatrix} + \\ & + \bar{d}_{iR}(g_d)_{ij}\phi^\dagger \begin{pmatrix} u_{jL} \\ d_{jL} \end{pmatrix} + \bar{u}_{iR}(g_u)_{ij}\tilde{\phi}^\dagger \begin{pmatrix} u_{iL} \\ d_{iL} \end{pmatrix} \Bigg) + h.c. \end{aligned} \tag{1.9}$$

where $\tilde{\phi} = i \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \phi^*$ is the scalar doublet with hypercharge -1/2.

The Yukawa matrices $(g_f)_{ij}$ are diagonalized by bi-unitary transformations:

$$(g_f)_{ij}^{diag} = U_f^{\dagger}(g_f)_{ij}V_f \qquad \begin{cases} f_L & \rightarrow V_f f_L \\ f_R & \rightarrow U_f f_R \end{cases} \tag{1.10}$$

where V_f and U_f are unitary matrices. The 9 eigenvalues of $[(g_f)_{ij}]_{f\in\{fermions\}}$ are in one-to-one correspondence with fermion masses.

When the transformation 1.10 is inserted in 1.7, it gives rise, in the charged current terms, to a non-vanishing unitary matrix that mixes the families in the weak charged-current interactions, the CKM matrix:

$$V_{CKM} = V_u^{\dagger}V_d$$

This matrix describes the mixing of families in charged-current processes, so a down-like quark of a family can turn in an up-like quark of another family emitting a W boson, and vice-versa. The CKM matrix contains other 4 physical parameters: 3 "angles" and a "phase", which is responsible for the CP violation in some SM processes.

In formula 1.9, we can read the couplings of the Higgs boson to the SM fermions:

$$L_{Hf} = -m_f\bar{f}f\left(1+\frac{h}{v}\right)$$

Each coupling is proportional the corresponding fermion mass.

In total, the L_Y term of the SM Lagrangian has 13 free parameters: the 9 Yukawa coupling constants which are in one-to-one correspondence with fermion masses, and the 4 CKM matrix parameters.

The SM lets the masses of fermions have different values, since they are free parameters of the theory, but it does not explain why they are so different, spanning 6 orders of magnitude, from 0.5 MeV/c^2 for the electron to about 173 GeV/c^2 for the top quark. This fact is known as the flavor problem.

1.1.4 Experimental tests of the SM

The SM was tested [6][8] in many experiments both at low and high energy, and its free parameters measured (except for the Higgs mass).

The *low energy (i.e., low momentum transfer $Q^2 \ll M_Z^2$) electroweak observables* were measured in *neutrino scattering* experiments, extracting the weak couplings of SM fermions. *Electron-hadron deep inelastic scattering* experiments tested the Parity violation of the weak interactions. The physics of the τ *lepton* has been studied providing information on the weak couplings of leptons and hadrons.

The *physics of Z and W bosons* was tested both at hadron and lepton colliders. The Z and W bosons were discovered at the Sp$\bar{p}$S collider at

CERN in the '80s, but it was the LEP electron-positron collider that verified the Standard Model predictions on Z decay rates, cross sections, etc., up to the second order of radiative corrections, providing also an extremely precise measurement of the Z mass.

The experiments at the Tevatron collider at Fermilab studied *QCD*, *heavy flavor* and electroweak physics, discovered the top-quark in 1995, and measured its properties.

Precision flavor physics investigates flavor changing processes like the oscillation of K and B mesons, involving also CP violation. They were measured, e.g., in the BaBar experiment at the SLAC laboratory.

In general, almost all of the aspects of the Standard Model have been tested, including QCD, electroweak and heavy flavor physics. For a complete review see [8]. Globally, combining hundreds of observables, an impressive self-consistency of the SM, and in particular, of the spontaneously broken electroweak sector, is reached. To give an idea, the following table contains some (few) of the main SM observables, with the measured values compared to SM best fit predictions:

Quantity	**Meas. Value**	**SM prediction**
m_Z [GeV/c^2]	91.1876 ± 0.0021	91.1874 ± 0.0021
Γ_Z [GeV/c^2]	2.4952 ± 0.0023	2.4961 ± 0.0010
$\Gamma_{Z\to had}$ [GeV/c^2]	1.7444 ± 0.0020	1.7426±0.0010
$\Gamma_{Z\to l^+l^-}$ [MeV/c^2]	83.984 ± 0.086	84.005 ± 0.015
$\Gamma_{Z\to invisible}$ [MeV/c^2]	499.0 ± 1.5	501.69 ± 0.06
m_W [GeV/c^2]	80.420 ± 0.031	80.381 ± 0.014
$g_V^{\nu e}$	-0.040 ± 0.015	-0.0398 ± 0.0003
$g_A^{\nu e}$	-0.507 ± 0.014	-0.5064 ± 0.0001
m_t [GeV/c^2]	173.4 ± 1.0	173.5 ± 1.0
τ_τ [fs]	291.13 ± 0.43	290.75 ± 2.51

The last missing piece of the SM is the Higgs boson: until recently only constraints on its mass were available. The past and current searches for the Higgs boson are discussed in the next sections.

Finally, it should be noted that some experimental observations are in disagreement with SM predictions and suggest "new physics" beyond the SM. These are the *neutrino oscillations and masses*, the cosmological evidence of the existence of *Dark Matter and Dark Energy* in the universe and *matter-antimatter asymmetry*. Some theoretical problems need to be solved, for instance, the *hierarchy problem*, the *flavor problem* and the *grand unification*. Innumerable theoretical models were developed in the last decades to solve these problems, many predicting observable effects below the TeV scale. For a review see [7]. According to these theories the SM should be considered as an effective theory, valid at low energies, rather than a basic theory of nature.

1.2 Search for the Higgs boson at the LHC

The Higgs boson is the last missing piece of the SM and has not been observed experimentally despite the direct searches both at lepton and hadron colliders. Direct searches were performed at LEP [9] and Tevatron [10][11] leading to the exclusion of the SM Higgs boson, at 95% of Confidence Level (CL), in the mass range (see Fig. 1.1):

$$m_H < 114.4 \text{ GeV/c}^2 \quad \text{and} \quad m_H \in [156, 177] \text{ GeV/c}^2 \tag{1.11}$$

The electroweak fit of the LEP precise measurements at the Z boson res-

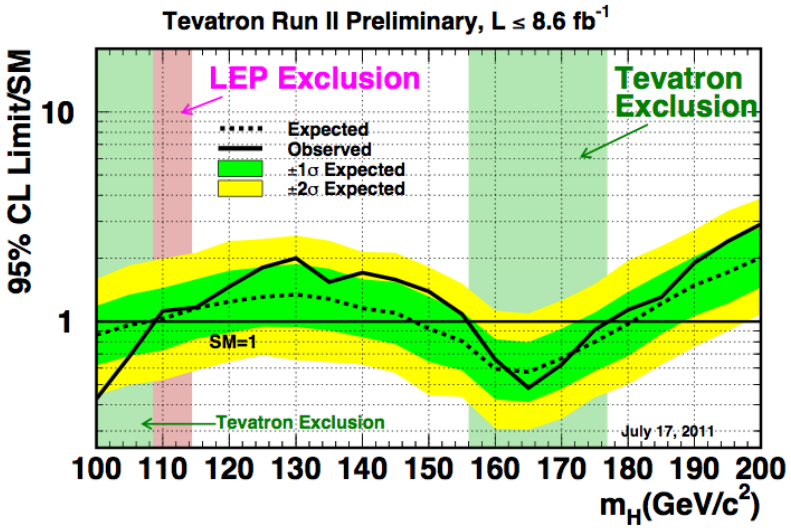


Figure 1.1: *Tevatron 95% CL upper limits on Higgs cross section, referred to the SM prediction. Tevatron and LEP exclusion ranges are highlighted.*

onance (that logarithmically depends on Higgs mass trough radiative corrections) [12], suggests a relatively low mass value for the Higgs, with the most probable value around 80 GeV/c^2 . More recent studies [13] combine the latest measurements of m_{top} and m_W and the most recent exclusions from Higgs searches at LEP and Tevatron. This combination provides a best fitted value of $m_H = 121^{+17}_{-6}$ GeV/c^2 and a 95% CL upper limit of 155 GeV/c^2 for the SM Higgs mass (see Fig. 1.2).

Given the excluded mass range and the electroweak fit results, the Higgs boson (if it exists) has most likely a mass between 114.4 GeV/c^2 and 155 GeV/c^2 , but theoretically the mass range from 177 GeV/c^2 to 1 TeV/c^2 is also open.

There also exist theoretical bounds on the Higgs mass. When m_H increases the amplitude for WW scattering process [19] becomes large. The

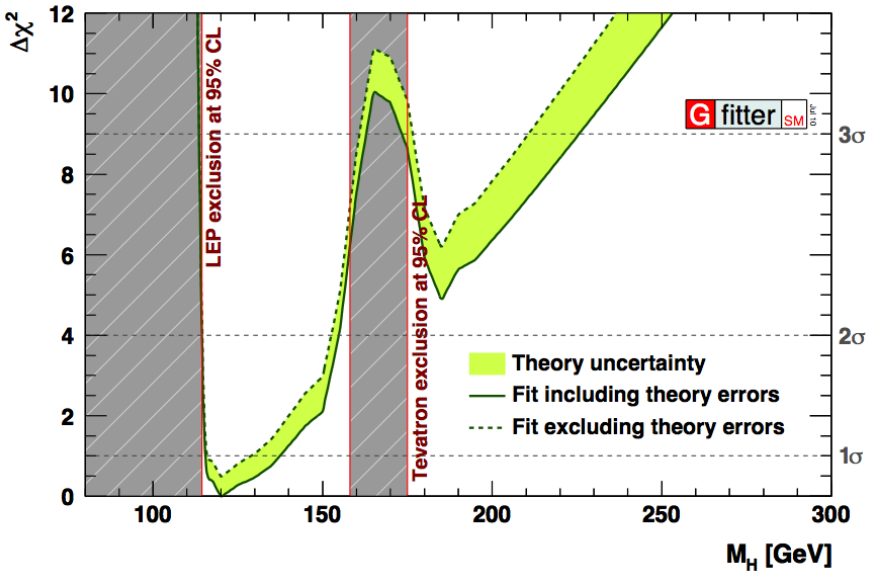


Figure 1.2: $\Delta\chi^2$ *as a function of* m_H *for a fit including the direct Higgs search results from LEP and Tevatron. The solid (dashed) line gives the results when including (ignoring) theoretical errors.*

unitarity limit is not exceeded only if

$$m_H < \left(\frac{8\sqrt{2}\pi}{3G_F}\right)^{\frac{1}{2}} \sim 1 \text{ TeV/c}^2$$

For masses around the TeV/c^2 the decay width of the SM Higgs becomes approximately equal to m_H and the Higgs can no longer be considered as a particle.

The Large Hadron Collider (LHC) [16][17] at CERN, a proton-proton (pp) collider that has worked at a center-of-mass energy of 7 TeV (8 TeV) in 2010-2011 (2012) and a luminosity of $\sim 5\times10^{33}$ cm^{-2}s^{-1} [3], is supposed to give the final answer about the Higgs existence. If the Higgs exists, and has a mass below 1 TeV/c^2, it will be discovered, otherwise it will be excluded definitely.

1.2.1 Higgs production

The Higgs production processes relevant at the LHC are, ordered by decreasing cross sections (see Fig. 1.3 for the corresponding Feynman diagrams):

[3] see section 2.1.2 for details on the LHC

- *gluon-gluon fusion:* $gg \to H$
- *vector boson fusion:* $qq \to Hqq$ via $WW, ZZ \to H$
- *associated production with a vector boson:* $q\bar{q} \to WH, ZH$
- *associated production with top quarks:* $gg, q\bar{q} \to t\bar{t}H$

The importance of these production processes reflects the preferred coupling of the Higgs boson to massive particles.

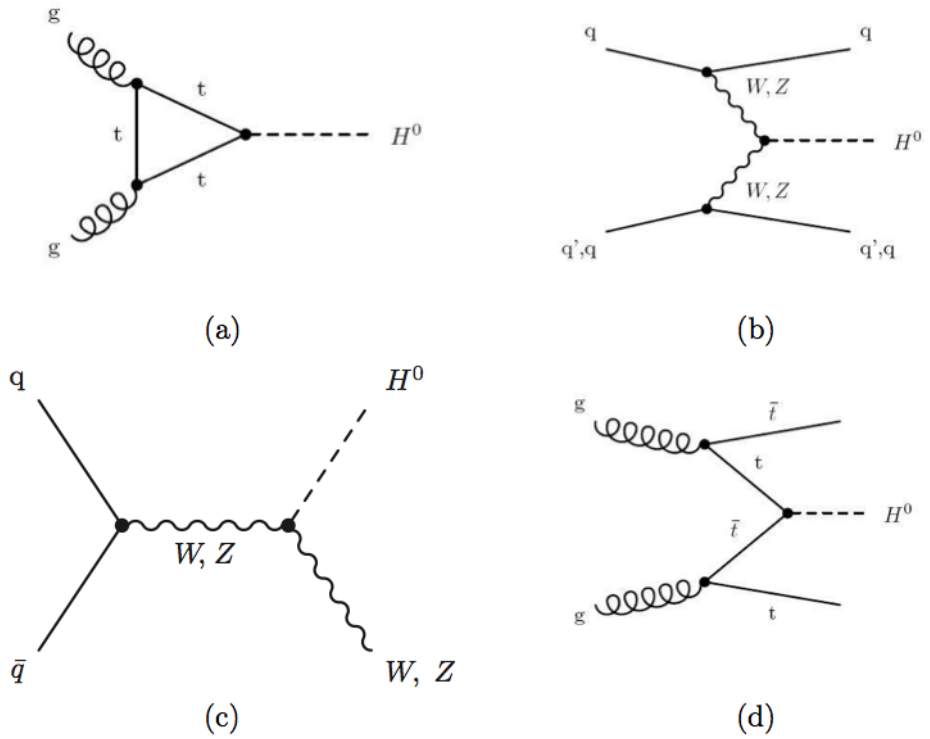


Figure 1.3: *Higgs production Feynman diagrams in proton-proton collisions: (a) gluon-gluon fusion; (b) VV fusion; (c) W and Z associated production (or Higgsstrahlung); (d) $t\bar{t}$ associated production.*

The numeric values of cross sections for these processes [18], which are mass-dependent, are shown in Fig. 1.4. The theoretical cross sections are calculated with perturbative QCD, using renormalized, scale-dependent, parton distribution functions:

$$\sigma(pp \to H) = \tag{1.12}$$

$$= \sum_{p_1 p_2} \int \mathrm{d}x_1 \mathrm{d}x_2 f_{p_1}(x_1, M^2) f_{p_2}(x_2, M^2) \times [\hat{\sigma}_0 + a\hat{\sigma}_1 + a^2\hat{\sigma}_2 + \ldots]_{p_1 p_2 \to H}$$

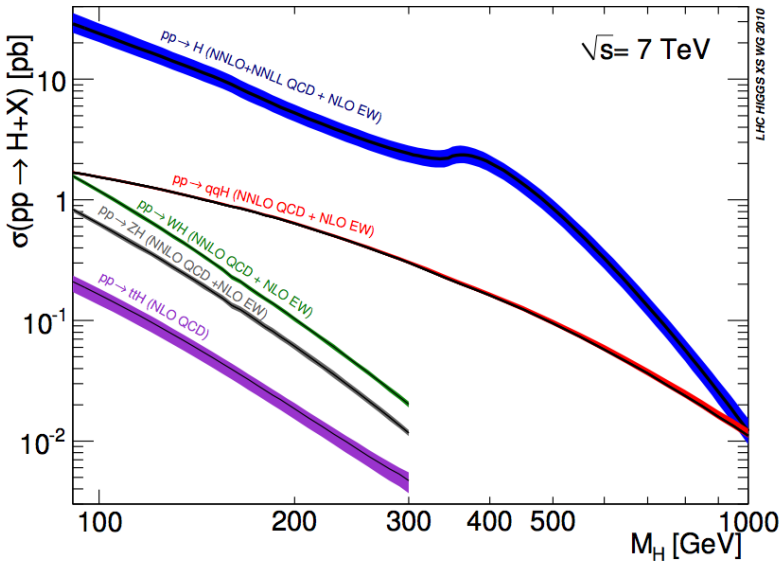


Figure 1.4: *Cross section for Higgs production at 7 TeV proton-proton collisions, as a function of the Higgs mass. The cross sections are given for the gluon fusion, vector boson fusion and associated production with a vector boson or a top pair.*

with $a = \alpha_S(M^2)/2\pi$, p_1, p_2 are the interacting partons, M^2 indicating the scale of the process; $\hat{\sigma}_j$ are the cross sections of the parton-level processes, at the j^{th} perturbative order. This perturbative expansion 1.12 is general [19], applicable to a wide variety of hard-scattering processes between hadrons. The zero-order term $\hat{\sigma}_0$ in 1.12 is called the Leading-Order (LO) in perturbation theory, the second $\hat{\sigma}_1$ Next-to-Leading-Order (NLO), then Next-to-Next-to-Leading-Order (NNLO) and so on. For the Higgs production, the cross sections are evaluated up to the NNLO for 7/8 TeV proton-proton collisions using the parton distribution functions described in [48][49]. The cross section for Higgs production is of the order of 10 pb for $m_H \sim 100$ GeV/c^2 , but it drops of 3 orders of magnitude as m_H approaches the TeV/c^2. At the LHC the Higgs production is copious if m_H is low, but it suffers from the large QCD background. If, otherwise, m_H is large, the smaller cross sections are compensated by a lower background. In general the discovery possibilities at the LHC are good in the entire mass range.

1.2.2 Higgs decay

The decay modes of the Higgs boson strongly depends on m_H. From the SM Lagrangian the decay amplitude for the Higgs decay are:

$$\Gamma(H \to f\bar{f}) = \frac{C_f G_F m_f^2 m_H}{4\pi\sqrt{2}} \left(1 - \frac{4m_f^2}{m_H^2}\right)^{3/2} \tag{1.13}$$

$$\Gamma(H \to W^+W^-) = \frac{G_F m_H^3}{8\pi\sqrt{2}} \left(1 - \frac{4m_W^2}{m_H^2}\right)^{1/2} \left(1 - \frac{4m_W^2}{m_H^2} + \frac{12m_W^4}{m_H^4}\right)$$

$$\Gamma(H \to ZZ) = \frac{G_F m_H^3 m_W^2}{16\pi\sqrt{2} m_Z^2} \left(1 - \frac{4m_Z^2}{m_H^2}\right)^{1/2} \left(1 - \frac{4m_Z^2}{m_H^2} + \frac{12m_Z^4}{m_H^4}\right)$$

where C_f is the color factor, $C_f = 3$ for quarks and $C_f = 1$ for leptons. Formulas 1.13 are valid above the two-particle mass threshold.

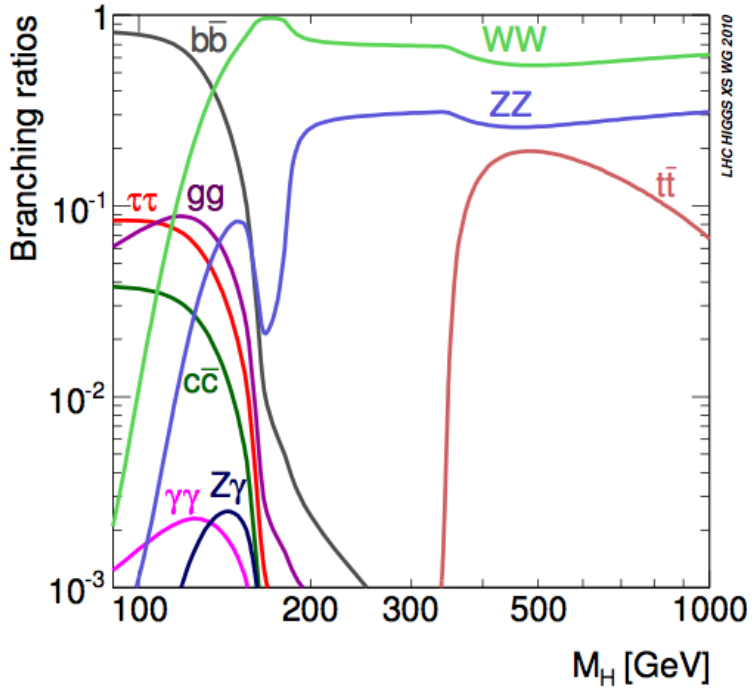


Figure 1.5: *Higgs boson decay branching ratios as functions of m_H.*

The Higgs decay branching ratios as functions of the Higgs mass are shown in Fig. 1.5:

- $m_H > 170$ GeV/c^2 : in the high mass region the decay to two massive vector boson (W,Z) dominates. They are themselves unstable, and so the final state would be composed of leptons and quark jets.

- $m_H < 150$ GeV/c^2 : in the low mass region the decay to a fermion pair is dominant. In particular the Higgs decays to $b\bar{b}$ and $\tau\tau$ have the highest branching ratios, since the Higgs coupling to fermions is proportional to their mass.

- $m_H \in [150, 170]$ GeV/c^2 : in the intermediate mass region both the fermionic and bosonic contributions are important

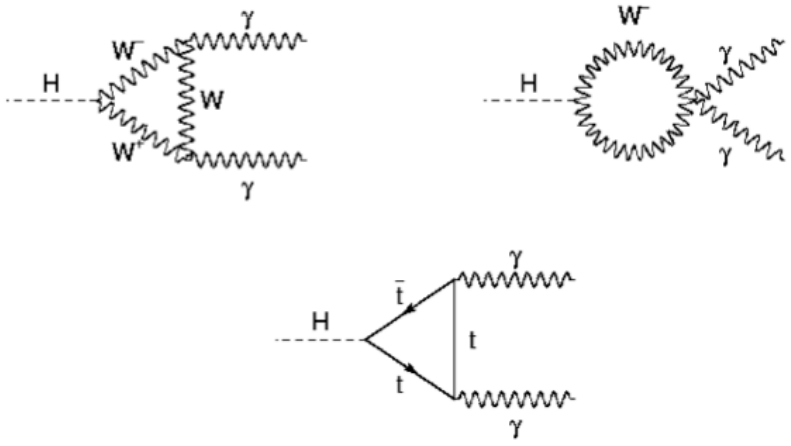


Figure 1.6: *Feynman diagrams for the Higgs boson decay to a photon pair.*

Certain rare decays like $H \to \gamma\gamma$, mediated by a top or W loop (see Fig. 1.6), are important. The formula that gives the decay amplitude is:

$$\Gamma(H \to \gamma\gamma) = \frac{\alpha^2 G_F m_H^3}{128\pi^3\sqrt{2}} \left| \sum_q e_q^2 I_q(m_q^2/m_H^2) + I_W(m_W^2/m_H^2) \right|^2 \qquad (1.14)$$

where I_q and I_W are dimensionless functions given by:

$$\begin{aligned} I_q(x) &= 4x[2 + (4x-1)F(x)] \\ I_W(x) &= -2[6x + 1 + 6x(2x-1)F(x)] \\ F(x) &= \theta(1-4x)\frac{1}{2}\left[\log\frac{1+\sqrt{1-4x}}{1-\sqrt{1-4x}} - i\pi\right]^2 + \\ &\quad -\theta(4x-1)2[\sin^{-1}(\sqrt{x}/2)]^2 \end{aligned}$$

This channel is important, despite the small branching ratio, especially for the Higgs search at hadron colliders, since the two-photon background is much smaller than the QCD one.

The Higgs decays that contain jets (especially light jets) in the final state are useless for a search at a hadron collider, given the huge QCD background. Thus rarer states containing leptons or photons in the final states are the most important ones.

1.2.3 Search strategies

Several analyses for the Higgs search at the LHC were studied [19][36]. Two channels are likely to provide the "best" chance for discovery [37], permitting also a very good mass reconstruction:

- $pp \to H \to ZZ \to l^+l^-l^+l^-$, for $m_H \gtrsim 120$ GeV/c^2
- $pp \to H \to \gamma\gamma$, for $m_H \lesssim 150$ GeV/c^2

Despite the small cross sections of these processes (the branching ratio is small in both), the signal-to-backround ratio is rather good. The main backgrounds are the irreducible ones, i.e., other SM processes that lead to the same particles in the final state: the QCD double photon production $pp \to \gamma\gamma$ and the non-resonant double Z production $pp \to ZZ \to l^+l^-l^+l^-$. Other "instrumental" backgrounds (i.e., to to detector specific features, like particle misidentification) can be reduced to an acceptable level by applying proper selection criteria to the reconstructed physical objects: the Higgs, in both analyses, would be revealed by the presence of a resonance peak above the broader background. The $H \to ZZ \to l^+l^-l^+l^-$ channel is the main argument of this thesis and will be extensively described in the next chapters.

Another important channels are:

- $H \to WW \to 2l2\nu$, for $m_H \in [130, 250]$ GeV/c^2
- $H \to b\bar{b}$, for $m_H \lesssim 135$ GeV/c^2 exploiting the associated production with vector bosons decaying to leptons: $pp \to H(\to b\bar{b})V(\to ll, l\nu, \nu\nu)$
- $H \to \tau\tau$, for $m_H \lesssim 150$ GeV/c^2

$H \to WW \to 2l2\nu$ is important in the intermediate mass region ($m_H \in [130, 250]$ GeV/c^2). It is a counting analysis (i.e., just sensitive to the final number of events, rather to the shape of a variable like the resonant invariant mass of the ZZ and $\gamma\gamma$ channels): having neutrinos in the final state the Higgs mass cannot be reconstructed.

The two channels $H \to b\bar{b}$ and $H \to \tau\tau$ can, for a low mass Higgs, significantly enhance the significance of a possible excess in $H \to \gamma\gamma$ and $H \to ZZ \to 4l$ analyses. The mass resolution in this case is very poor, of the order of 10-20 GeV/c^2 .

1.2.4 First results on the Higgs search at the LHC

With the data collected in 2010-2011 by the experiments ATLAS and CMS at the LHC (see next chapter for details), the SM Higgs boson was excluded in a wide mass range.

The results were obtained from the combination of the Higgs searches in the five decay modes listed in section 1.2.3: $\gamma\gamma$, $b\bar{b}$, $\tau\tau$, WW and ZZ.

CMS [14] has excluded the SM Higgs boson in the mass range [127,600] GeV/c^2 at 95% CL. The expected excluded mass range in the absence of signal was [118,543] GeV/c^2 . An excess of events above the expected SM background has been observed at low mass, making the observed limits weaker than expected. The largest excess, with a local significance of 3.1 standard deviations, is observed for a Higgs boson mass hypothesis of 124 GeV/c^2 . These results are summarized in Fig. 1.7, which shows the 95% CL upper limit on the Higgs cross section as a function of the hypothesized mass.

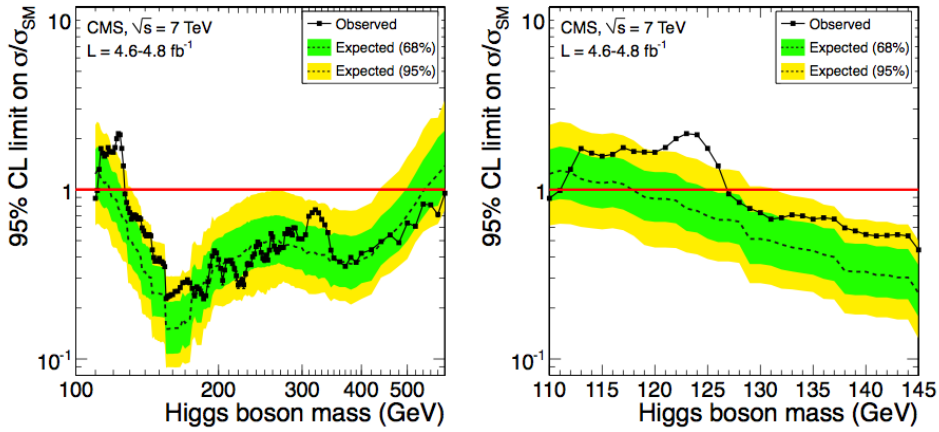


Figure 1.7: *CMS combined limit on Higgs cross section as a function of Higgs mass. 2011 data.*

According to the ATLAS results [15], the SM Higgs boson mass ranges of [111.4,116.6] GeV/c^2 , [119.4,122.1] GeV/c^2 , and [129.2,541] GeV/c^2 are excluded at 95% CL, while the range [120,560] GeV/c^2 was expected to be excluded in the absence of a signal. An excess of events was observed for a Higgs boson mass hypotheses around 126 GeV/c^2 with a local significance of 2.9 standard deviations.

These results dramatically restrict the allowed mass range for the SM Higgs boson, while the excess seen in both experiments at low mass suggests the possible presence of signal. To obtain a statistically solid result more data are needed.

Chapter 2

The Compact Muon Solenoid experiment at the LHC

The *Compact Muon Solenoid (CMS)* experiment at the LHC is designed to search for the Higgs boson and possible signals of "New Physics". In order to investigate efficiently a wide variety of sub-nuclear physical phenomena a versatile experimental structure is needed, as hermetic as possible. High performances are required in reconstruction and identification of physical objects: muons, electrons, photons and jets. The main features of CMS are a superconducting solenoidal magnet and a compact detection structure designed with particular attention to the muon, the tracking and the electromagnetic systems.

In the next sections the LHC is briefly presented, the CMS detection apparatus is described, highlighting the main features of the sub-detectors, driven by physical requirements. The performances during data taking periods (from 2010 to 2012) are summarized.

2.1 The LHC

2.1.1 Physics at the LHC

The constituents of matter and their interaction laws are described in the Standard Model, presented in the previous chapter. The least understood part of the SM is the Higgs sector and the related electroweak gauge spontaneous symmetry breaking. All the direct searches for the Higgs boson, before the LHC, have given negative results. The verification of the full Standard Model mechanism which includes the Higgs symmetry breaking mechanism with the discovery of the Higgs boson is indeed one of the main goals of physics today and this is the main goal of the LHC. The mass of the Higgs boson is not defined within the Standard Model theory, being one of the parameters of the theory and a high energy hadron machine is the

most suited to span the search for new particles in a unknown large range of mass.

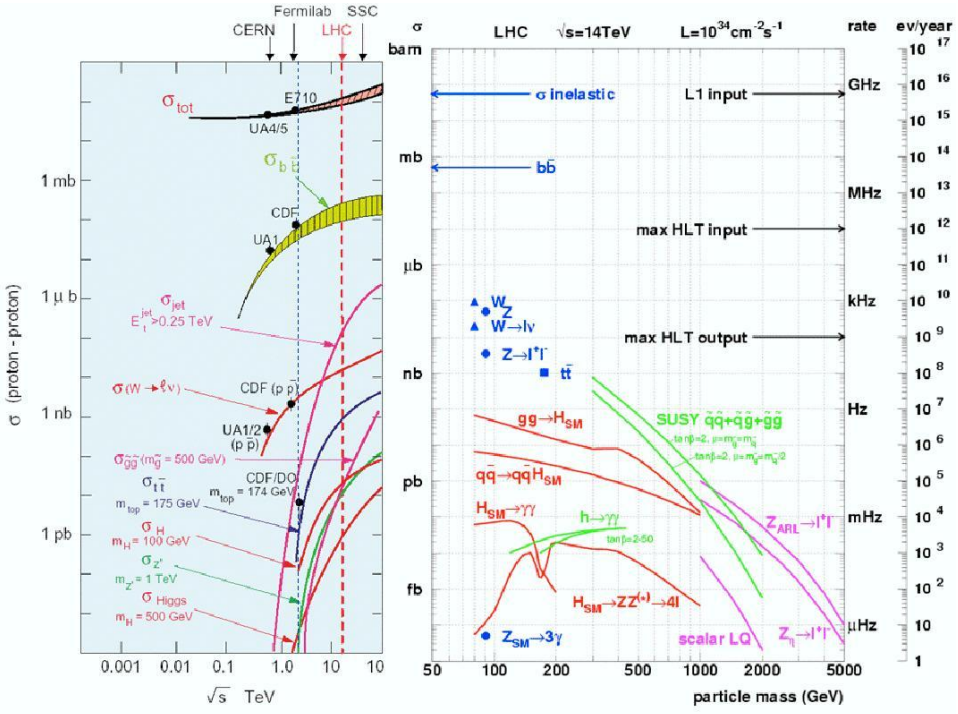


Figure 2.1: *Cross sections and rates for the main physical processes at the LHC.*

There are also many reasons, both theoretical and experimental, to search for "new physics" beyond the SM [7], as reported in section 1.1.4. Innumerable theoretical models have been developed in the last decades to solve these problems and one major aim of the LHC collider is to confirm or to disavow them experimentally: citing e.g., the *SuperSymmetry (SUSY)* that predicts a correspondence between bosons and fermions; for each SM particle a "partner" is predicted with opposite statistic behavior. Since we have never seen supersymmetric particles, SUSY must be broken in a way that lets the supersymmetric quanta acquire large masses (of the order of 100 GeV/c^2 but not larger than TeV/c^2 in order to allow the solution of the hierarchy problem). SUSY can also solve the Dark Matter problem with the introduction of a new discrete symmetry, the R-parity, that prevents the Lightest Supersymmetric Particle (LSP) from decaying to a SM particle. In order to be a Dark Matter candidate the LSP should be massive ($\sim$ 100 GeV) and neutral. Many SUSY models predict good candidates such as neutralinos or gravitinos.

The LHC is designed to collide protons with a center-of-mass energy of 14 TeV and a luminosity of 10^{34} cm^{-2}s^{-1}, in order to explore the TeV energy scale and search for new particles such as Higgs and SUSY ones. In Fig. 2.1 the predicted cross sections for the most relevant physical processes in pp collisions are reported as functions of the center of mass energy. Note that the ones relative to the most important "hard QCD" processes (like vector boson and Higgs production) grow faster with energy than the less important ones (like soft QCD background). Cross sections like those of Higgs production and top production are, at the energy of the LHC, orders of magnitude larger than at the Tevatron energy. We can also see in Fig. 2.1 the large amount of background processes: the total pp cross section at 14 TeV is:

$$\sigma(pp) = \sigma(pp)_{elastic} + \sigma(pp)_{inelastic} = 20 \text{ mb} + 80 \text{ mb} = 100 \text{ mb}$$

One of the most difficult challenges of the experiments at the LHC is to build an efficient trigger system to discriminate the interesting events from the background.

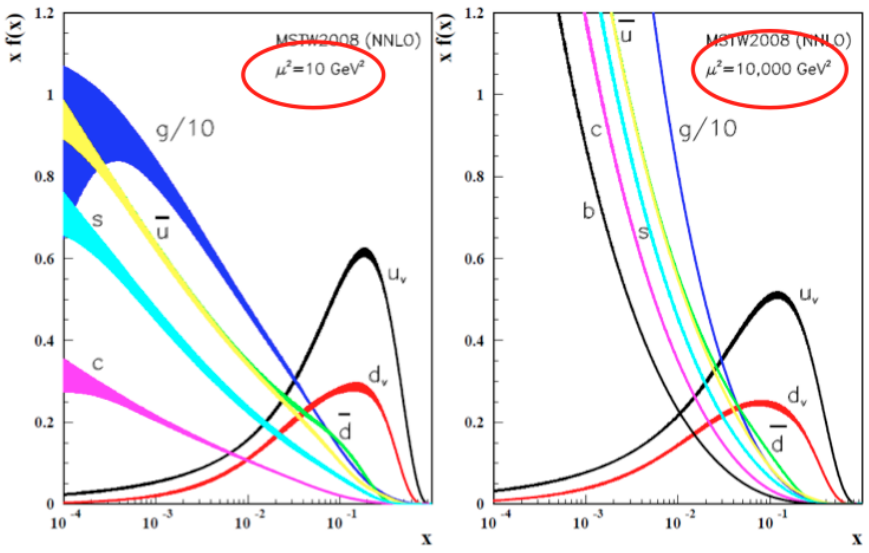


Figure 2.2: *The proton parton distribution functions, at 10 GeV (measured) and 10 TeV (extrapolated) scales.*

To explore a wide energy range a hadron collider is the proper tool. The hard collisions occur between quarks and/or gluons, the "partons" in the protons. If $\sqrt{s}$ is the energy in the pp center of mass, the real center of mass energy of each parton collisions is $\sqrt{x_1 x_2 s}$, where x_i is the fraction of momentum that each parton carries: the distributions of x_i for different types of partons are described by the Parton Distribution Functions (PDF),

widely explored by experiments at SLAC, SPS and HERA [48][49]. In Fig. 2.2 the parton distribution functions for the proton are reported, at 10 GeV (measured) and 10 TeV (extrapolated) scales: as shown, at high energy the gluon component dominates, so the LHC is de facto a gluon colliding machine. The collision energies can range between zero and $\sqrt{s}$, thus a large energy range is explored. Since the interactions occur between partons, the Lorentz-boost of the scattering is unknown: physical analyses must rely on relativistic-invariant quantities like transverse momenta and masses rather than on frame-dependent ones.

Besides the energy, the fundamental parameter of the collider is the luminosity, since the "interesting" processes like Higgs or SUSY ones are very rare. The statistics collected at Tevatron in 20 years, expressed in terms of integrated luminosity, is $\sim$10 fb^{-1}. The LHC is designed to collect up to 100 fb^{-1} per year.

2.1.2 The LHC

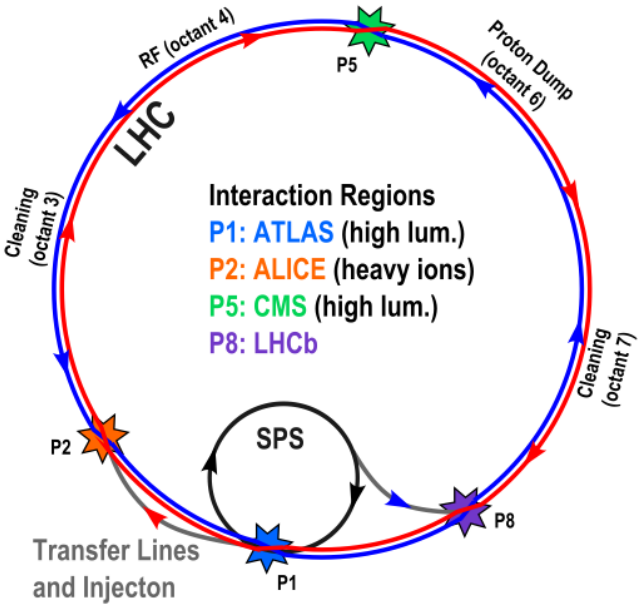


Figure 2.3: *The LHC, overall structure.*

The collider [24] (Fig. 2.3) has a length of 26.7 km; it is entirely made of superconducting magnets, with NbTi cables cooled down to 1.9 K. It is designed to accelerate two bunched proton beams in opposite direction up to an energy of 14 TeV, with a instantaneous luminosity of 10^{34}cm^{-2}s^{-1}, and to make them collide in four points around its circumference. The LHC

also collides heavy ions (like lead) for short periods in the year.

The following table summarizes the most relevant design parameters of the LHC.

Energy per proton	7 TeV
Dipole field	8.33 T
Design luminosity	10^{34} cm^{-2}s^{-1}
Bunch separation	25 ns
No. of bunches	2808
No. of particles per bunch	$1.15 \cdot 10^{11}$
β-value at IP	0.55 m
RMS beam radius at IP	16.7 μm
Luminosity lifetime	15 hr
Number of collisions/crossing	~ 20

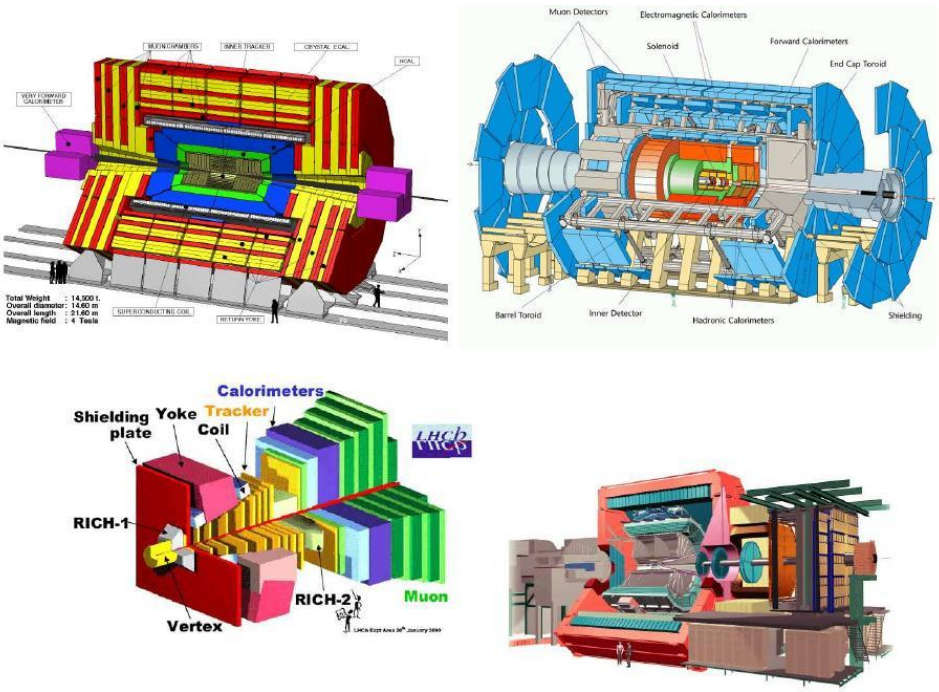


Figure 2.4: *The four experiments at the LHC. From top-left: CMS, ATLAS, LHCb, ALICE.*

Four experiments are built along the LHC circumference: *CMS* [20], *ATLAS (A Toroidal LHC ApparatuS)* [21], *ALICE (A Large Ion Collider Experiment)* [22], *LHCb (LHC Beauty experiment)* [23].

CMS and ATLAS are very similar in scope: they study the final states of produced in pp collisions and their goal is to find the Higgs boson and,

in general, to explore the TeV scale. ALICE is built to study the collisions of heavy ions to sudy phenomena like the quark-gluon plasma. LHCb is devoted to the flavor physics: it mostly studies processes involving b-quarks and c-quarks.

During 2010 and 2011 the LHC collided protons with a energy of 3.5 TeV per beam. It reached a maximum instantaneous luminosity of $5 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$, delivering an integrated luminosity of 5 fb^{-1}. In 2012 the energy per beam increased to 4 TeV, with a maximum instantaneous luminosity of $7 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$, for a total integrated luminosity of 20 fb^{-1}. The bunch separation, both in 2011 and 2012, was 50 ns, this leading to a higher *pile-up* of events in collisions (i.e. several inelastic pp collisions over imposed in the same bunch crossing) than the designed one.

2.2 The CMS detector

2.2.1 Physical requirements

The CMS design is physics-oriented [36].

The main goal of CMS is to search for the *Higgs boson* [42], whose decay channels strongly depend on its mass. Even if the hadronic decays are dominant in terms of cross section (especially in the low mass region) the channels involving leptons or photons in the final state are of extreme importance, being less contaminated by the huge QCD background of the LHC, and are de facto preferred in the CMS design. In the low mass interval $[114, 140]$ GeV/c^2 the two-photon decay $\text{H}\rightarrow \gamma\gamma$ is one of the main channels likely to yield a significant signal. For masses larger than about 120 GeV/c^2 , the ZZ decay with 4 leptons in the final state $\text{H}\rightarrow\text{ZZ}\rightarrow 2l2l'$ is the best channel to investigate. The $\text{H}\rightarrow\text{WW}\rightarrow 2l2\nu$ is important in the intermediate mass range $[120, 300]$ GeV/c^2 . The natural width of the Higgs boson in the intermediate mass region ($[114, 182]$ GeV/c^2) is only few MeV/c^2: the observed width of a Higgs signal will be dominated by the instrumental mass resolution, which must therefore be optimized.

The search for *Supersymmetric particles* is one of the goals of CMS [43][44]. The decays of SUSY particles, such as squarks and gluinos, involve cascades that, if R-parity is conserved, always contain the Lightest SUSY particle (LSP). The LSP is expected to interact only weakly, thus leading to significant E_t^{miss} in the final state. The rest of the cascade results in a abundance of leptons and jets. In the Gauge Mediated Symmetry Breaking (GMSB) the presence of hard isolated photons is expected.

Search for *massive vector bosons* as Z' lead to final states involving the presence of leptons from decays such as $Z' \rightarrow ee$ and $Z' \rightarrow \mu\mu$. Ways to distinguish between different models involve the measurement of the natural

width and the forward backward asymmetry, both of which require good momentum resolution at high p_T ($\delta_{p_T}/p_T < 0.1$ at $p_T = 1$TeV).

The LHC also allow precision studies of *QCD, electroweak* and *flavor physics* [39][40][41]. Testing the Standard Model at the LHC energy scale is the basic task of CMS. Extensive tests of QCD are possible through the measurement of jet and direct photon production with transverse energies up to 3-4 TeV and from cross sections which fall by 11 orders of magnitude. Similarly the electroweak sector is studied, with vector bosons produced in association with jets and heavy-flavor jets. Top quark is produced at the LHC with a large rate, giving the opportunity to test its Standard Model couplings and spin: this requires high efficiency in b-jet tagging. Flavor physics can give indications for physics beyond the Standard Model, providing complementary information with respect to direct searches. Searches for flavor changing neutral currents, lepton flavor violation through $\tau \to 3\mu$ or $\tau \to \mu\gamma$, precision measurements of $B_s^0 \to \mu\mu$ can open a window onto new physics.

The *detector requirements* to meet the physics program can be summarized as follows:

- high efficiency in muon identification and good momentum resolution over a wide range of momenta in the pseudo-rapidity (see section 2.2.2) region $|\eta| < 2.5$; good di-muon mass resolution ($\sim 1\%$ at 100 GeV/c^2); ability to determine unambiguously the charge of muons with $p < 1$ TeV/c;

- good charged particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ and b-jets, requiring pixel detectors close to the interaction region to identify primary and secondary vertices;

- good electromagnetic energy resolution, good di-photon and di-electron mass resolution ($\sim 1\%$ at 100 GeV/c^2), wide geometric coverage ($|\eta| < 2.5$), measurement of the direction of photons; π^0 rejection and efficient photon and lepton isolation at high luminosities;

- good E_t^{miss} and di-jet mass resolution, requiring hadron calorimeters with a large hermetic geometric coverage ($|\eta| < 5$) and with fine lateral segmentation: $\Delta\eta \times \Delta\phi < 0.1 \times 0.1$ (for η and ϕ definitions, see section 2.2.2);

- a high performant trigger system to reduce the event rate from 10^9 Hz to 10^2 Hz, the maximum storable in real-time on disks;

- a high granularity tracking detector to limit the effects of event pile-up: at the design luminosity about 20 inelastic interaction are expected per bunch crossing, with ~ 1000 tracks.

The CMS design meets these requirements. The detector structure is described in the following sections. The main features of CMS are a solenoidal magnet creating a 3.8 T field, a full silicon-based inner tracking system, an active scintillating crystal-based electromagnetic calorimeter and a complete muon detection system.

2.2.2 The detector

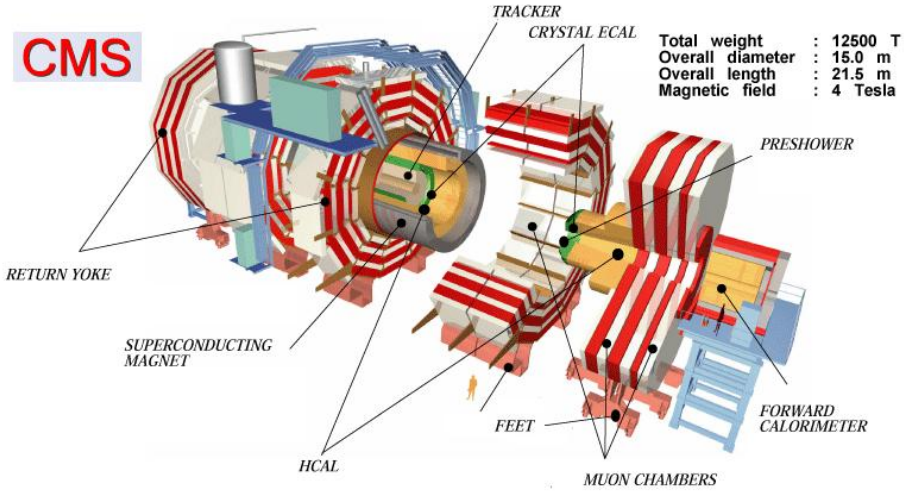


Figure 2.5: *The CMS overall structure.*

The CMS apparatus [24] presents the cylindrical structure shown in Fig. 2.5, with a central barrel closed by two end-caps. In order to maximize the acceptance for particles coming from pp interactions, the detector is almost hermetic: its hermeticity is limited only by the beam pipe that follows the cylinder axis. The overall dimensions of the CMS detector are a length of 28.7 m, a diameter of 15 m and a weight of 14,000 tons. The structure is built around the superconducting solenoid 7 m long with a diameter of 6 m: the magnetic field in the inner region is 3.8 T while externally the field lines are collected in the iron return yokes and the field has approximately the value of 1.7 T. Strong magnetic fields are needed in order to ensure large bending power to measure precisely the momentum of charged particles ($\delta_{p_T}/p_T < 10\%$ at $p \sim 1$ TeV/c). The sub-detectors that compose CMS are, from inside out: the silicon tracker, the electromagnetic and hadronic calorimeters, the magnet and, externally, the muon system.

The CMS conventional 3D *frame of reference* has its origin in the nominal beam interaction point, the x-axis pointing radially inward toward the center of the LHC, the y-axis pointing vertically upward; the z-axis points along the beam direction toward the Jura mountain from the LHC point 5. The

azimuthal angle ϕ is measured from the x-axis in the x-y plane (transverse plane). The polar angle θ is measured from the z-axis. Pseudo-rapidity is defined as $\eta = -\log\tan\frac{\theta}{2}$.

The tracker

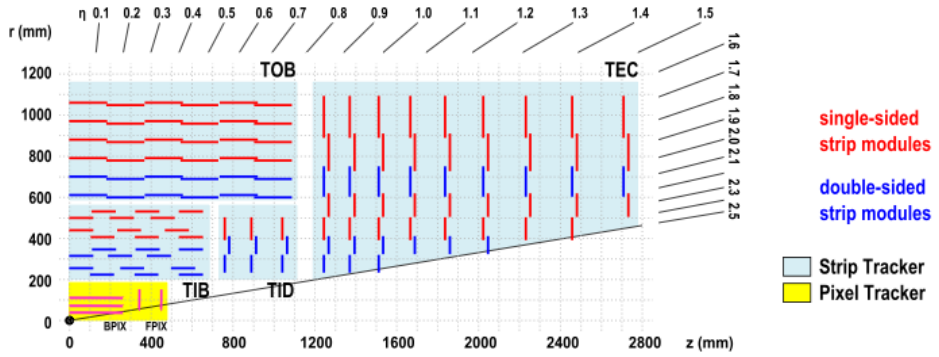


Figure 2.6: *The silicon Tracker structure.*

The tracker [29] is the detector closest to the interaction point: it is designed to efficiently reconstruct tracks of charged particles with high resolution in position and momentum, and to identify the primary and secondary vertices (i.e., the common origin of groups of tracks).

Entirely made of silicon semiconductor detectors, the tracker is divided into three regions delineated by considering the charged particle flux at various radii at high luminosity. Closest to the interaction vertex where the particle flux is the highest (10^7/s at $r \simeq 10$ cm) the PIXEL detectors are used; the size of a pixel is 100×150 μm^2. In the intermediate region ($20 < r < 55$ cm) the particle flux is low enough to allow the use of silicon micro-strip detectors with a minimum cell size of 10cm$\times$80 μm (Tracker Inner Barrel TIB and Tracker Internal Disks TID). In the outermost region (r > 55 cm) the particle flux has dropped sufficiently so that a larger-pitch silicon micro-strips with a maximum cell size of 25cm$\times$180μm (Tracker Outer Barrel TOB and Tracker End Caps TEC) can operate. The layout of the tracker is shown in Fig. 2.6. The outer radius extends to nearly 110 cm while the total length is approximately 540 cm. The pseudo-rapidity coverage is up to 2.5.

The hit spatial resolution in the pixels, monitored on real data [45], is of the order of 10 μm in the transverse $r\phi$ plane and 30 μm in z (here the resolution depends on the track angle). Two indicative plots are shown in Fig. 2.7 for pixels. The strip resolution depends on the strip size and ranges from 14 μm (for 80 μm sensor pitch) and 36 μm (for 183 μm pitch).

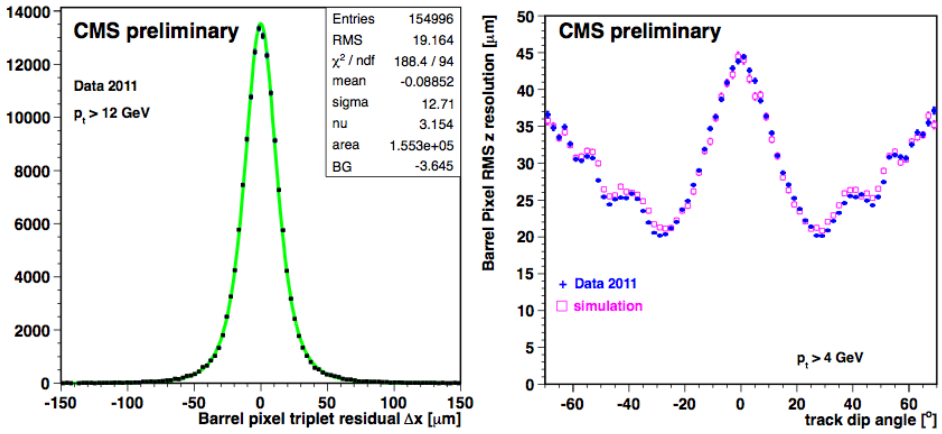


Figure 2.7: *Pixel barrel hit resolution measurements in the $r - \phi$ plane.*

The average hit finding efficiency is measured to be 99% for pixels and 99.7% for strips.

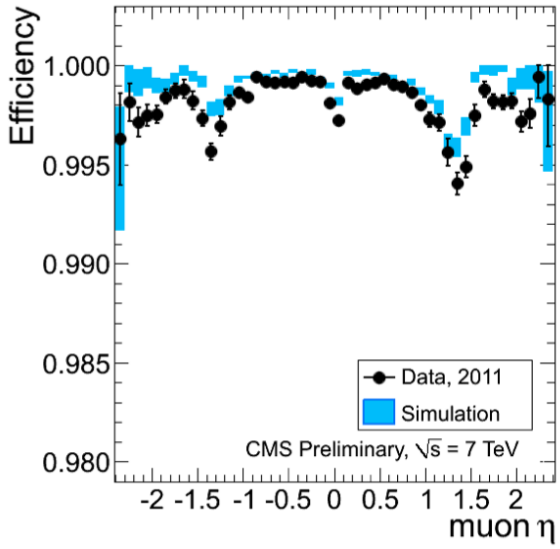


Figure 2.8: *Track finding efficiency for muons coming from the Z boson decay as a function of muon η.*

The tracks are reconstructed starting from seeds in the pixels, extrapolated with a Kalman filter technique. Track finding efficiency for muons coming from the Z boson decay is shown in Fig. 2.8 as a function of the muon η: the average value is above 99.7%.

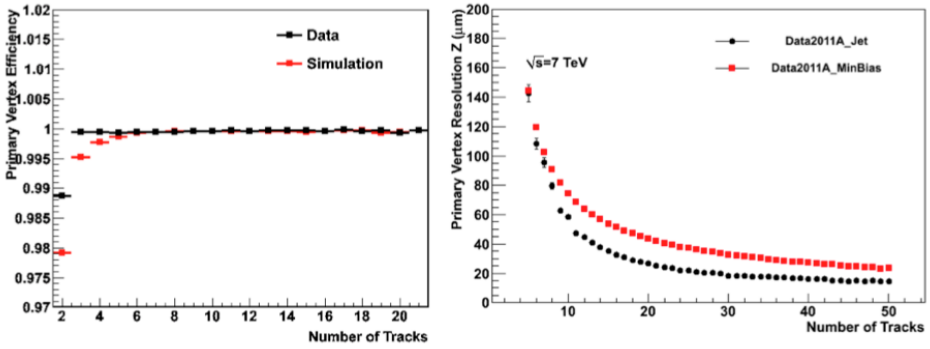


Figure 2.9: *Vertex finding efficiency (left) and resolution (right) as functions of the number of vertex tracks.*

The vertices are reconstructed with dedicated algorithms that select the tracks to be used, cluster them and fit the vertex position. In Fig. 2.9 primary vertex efficiency and resolution are shown as functions of the number of tracks, for data and simulation: the efficiency is close to unity if the vertex is built with at least three tracks, while the z-resolution ranges from 20 μm to 100 μm.

Electromagnetic calorimeter

The Electromagnetic Calorimeter (ECAL) detector has a crucial role in the physics program of CMS. Its design has been driven by the requirement of an excellent energy resolution [27].

It is a hermetic, homogeneous calorimeter with 61,200 lead tungstate ($PbWO_4$) crystals mounted in the central barrel, closed by 7,324 crystals in each of the two end-caps. The barrel section (EB) has an inner radius of 129 cm and covers the pseudo-rapidity range $|\eta| < 1.479$. The end-caps (EE) cover the rapidity range $1.479 < |\eta| < 3.0$ (see Fig. 2.10). A preshower detector (ES) is placed in front of the crystal calorimeter. Lead tungstate scintillating crystals have short radiation ($X_0 = 0.89$ cm) and Moliere (2.2 cm) lengths, are fast (80% of the light is emitted within 25 ns) and radiation hard (up to 10 Mrad). The relatively low light yield (30 γ/MeV) requires the use of photodetectors with an intrinsic gain that can operate in a 3.8 T magnetic field. Silicon avalanche photodiodes (APDs) are used as photodetectors. The usage of $PbWO_4$ crystals has allowed the design of a compact calorimeter inside the solenoid that is fast, has fine granularity, and is radiation resistant.

The energy resolution (measured in electron test beams) is parameterized

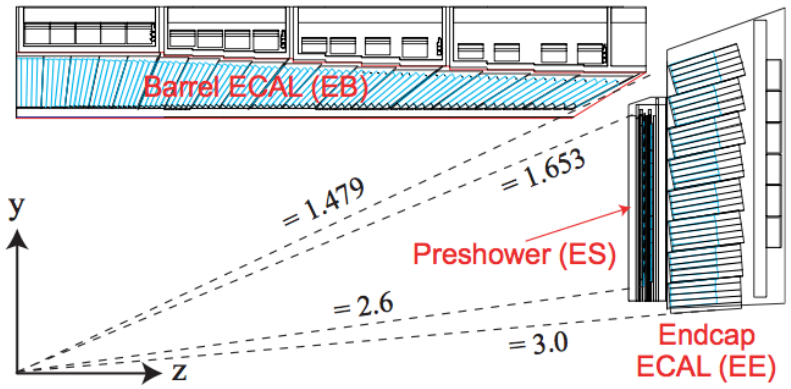


Figure 2.10: *The electromagnetic calorimeter structure.*

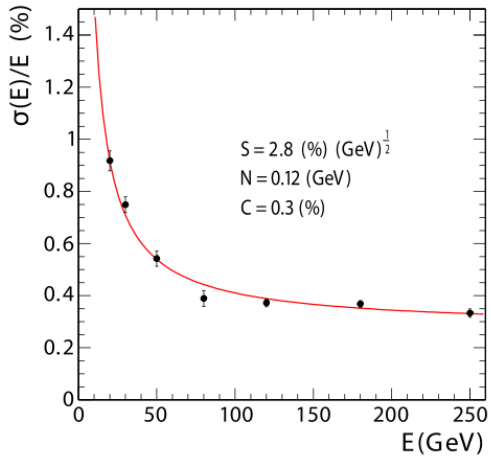


Figure 2.11: *ECAL energy resolution parametrized as a function of energy.*

as a function of energy, expressed in GeV (see Fig. 2.11):

$$\left(\frac{\sigma(E)}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2$$

where the first term (S) is stochastic (depending on event characteristics, especially on the electromagnetic shower development), the second is the noise term (N) (due to electronics and pile-up) and the last is a constant term (C) that depends on the non-uniformity of the longitudinal light collection, on the leakage of energy from the rear face of the crystals and on the accuracy of the detector inter-calibration constants.

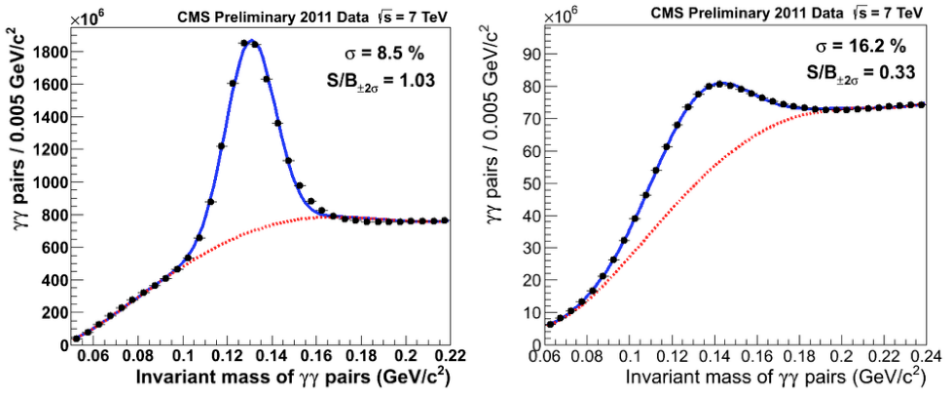


Figure 2.12: $\pi^0 \to \gamma\gamma$ *resonance for ECAL calibration, Barrel (left) and Endcap (right).*

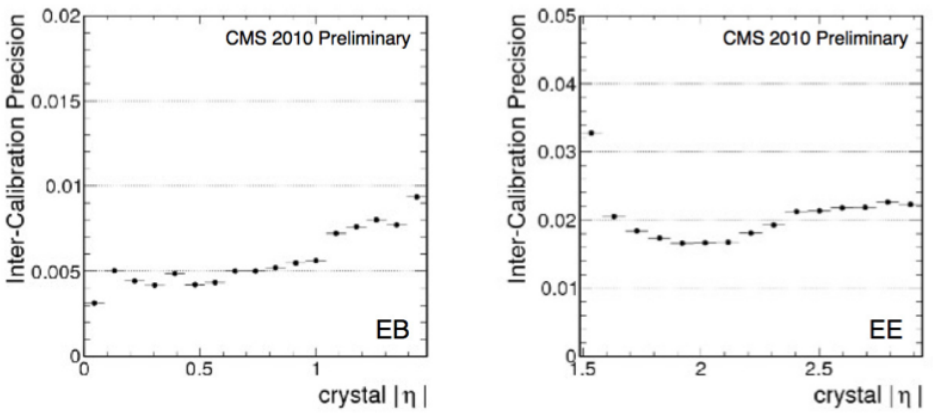


Figure 2.13: *Combined inter-calibration precision for the ECAL Barrel (left) and Endcap (right).*

The start-up calibration precision is improved using pp collision data, combining different methods [46]:

- the azimuthal symmetry method, exploiting the energy-flow invariance around the beam axis in events with minimal trigger requirements, to inter-calibrate crystals at the same pseudo-rapidity;
- π^0 and η resonances inter-calibration, using the invariant mass peaks of di-photon events from $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ candidates (see Fig. 2.12);
- isolated, high energy electrons from $W \to e\nu$ and $Z \to ee$ decays, comparing the energy measured in ECAL with the track momentum

measured in the CMS silicon Tracker;

- radiative Z decays, using $Z \to \mu\mu\gamma$, assuming the correct muon energy scale, to check the photon energy scale.

Fig. 2.13 shows the precision obtained combining all the different calibration strategies as a function of the crystal η.

The energy scale is tuned using $Z \to ee$ events from LHC collisions and others low-mass resonances like $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ (see Fig. 2.12).

Hadron calorimeter

The design of the hadron calorimeter (HCAL) [28] is strongly influenced by the choice of magnet parameters since most of the CMS calorimetry is located inside the magnet coil and surrounds the ECAL system. An important requirement of HCAL is to minimize the non-Gaussian tails in the energy resolution and to provide good containment and airtightness for the missing transverse energy E_T^{miss} measurement.

The barrel (HB) and endcap (HE) are joined hermetically with the barrel extending down to $|\eta| = 1.4$ and the end-cap covering the overlapping range $1.3 < |\eta| < 3.0$. The HCAL design maximizes material inside the magnet coil in terms of interaction lengths. Brass is the absorber material as it has a reasonably short interaction length, it is easy to machine and non-magnetic. Maximizing the amount of absorber before the magnet requires keeping to a minimum the amount of space devoted to the active medium: the tile/fiber, that consists of plastic scintillator tiles read out with embedded wavelength-shifting fibers. In order to improve the energy resolution, and measure energetic forward jets, hadron forward (HF) calorimeters in the end-caps and a layer of scintillators outside the coil, the hadron outer (HO) are added. The forward calorimeters [30] are located 11.2 m from the interaction point and extend the pseudo-rapidity coverage from $|\eta| = 2.9$ down to $|\eta| = 5$. Central shower containment in the region $|\eta| < 1.26$ is improved with an array of scintillators located outside the magnet in the outer barrel hadronic calorimeter (HO).

For gauging the performance of the HCAL, it is usual to look at jet and missing transverse energy resolutions. The granularity of the sampling in the 3 parts of the HCAL is chosen such that the jet energy resolution, as a function of E_T, is similar in all 3 parts. This is illustrated in Fig. 2.14, from simulated data.

Muon system

The required performance of the muon system is defined by the narrow states decaying into muons ($\eta, \rho, J/\psi, \Upsilon$, Z) and by the unambiguous determination of the charge of muons au to $p = 1$ TeV/c.

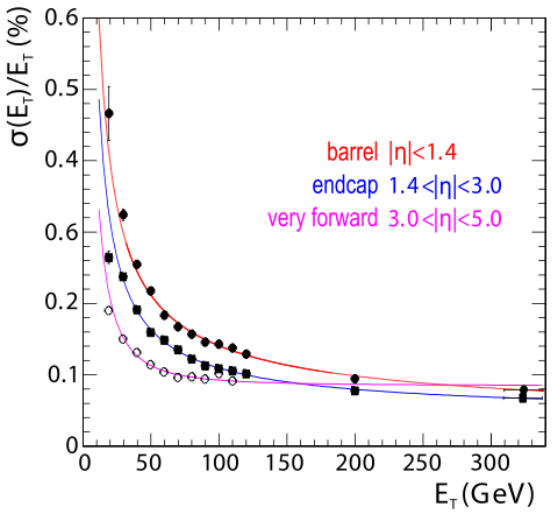


Figure 2.14: *Hadron calorimeter resolution as function of energy, from simulated data.*

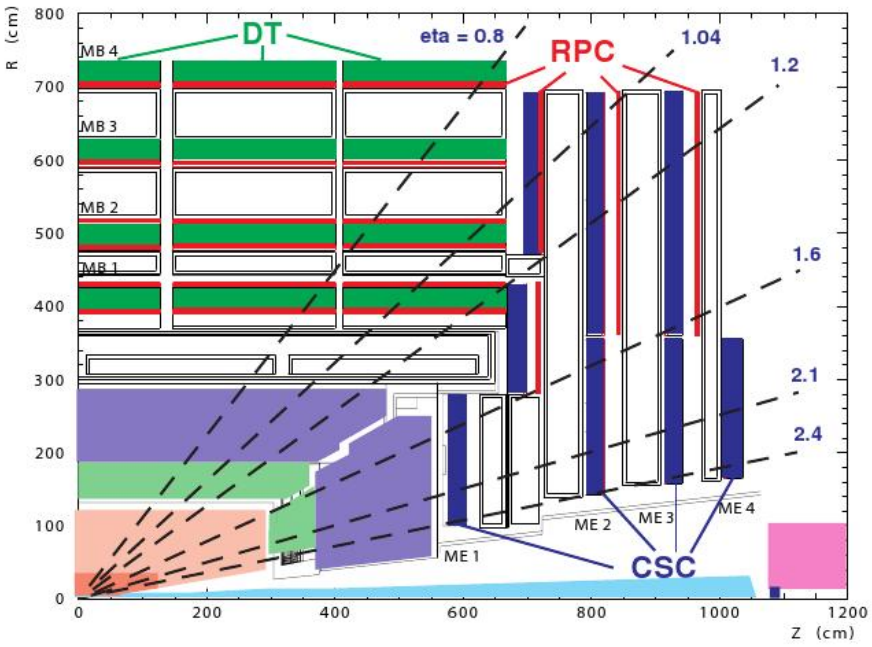


Figure 2.15: *The muon system structure.*

Centrally produced muons are measured in the inner tracker and in the barrel outer region. The detection system used outside the CMS magnet is

called the Muon system [26]. It is divided in a middle region, the barrel, and two end-caps (see the structure in Fig. 2.15). Three types of gaseous detectors are used in the muon system. The Drift Tubes (DT) in the barrel, the Cathode Strip Chambers (CSC) in the end-caps and the Resistive Plat Chambers (RPC) both in the barrel and in the end-caps. The DTs, the CSCs and the RPCs operate within the first level trigger system, providing two independent and complementary sources of information. They measure the position and direction of particles exploiting the return flux of the solenoid magnetic field. They perform momentum measurement independently from the tracker.

The barrel region is divided into five dodecahedral wheels: there are twelve sectors per wheel numbered from 1 to 12 starting from the sector in the positive x direction. Four parallelepiped DT chambers are located in each sector together with a variable number of RPCs (from two for the inner to zero for the outer). A sandwich of DTs and RPCs is called Muon Barrel stations and are indicated with MB1, MB2, MB3 and MB4 from the inner to the outer one. In the sectors 4 and 10, i.e., the vertical ones, there are two MB4 stations; conventionally the seconds of these stations are assigned to "sectors" 13 and 14. The DTs cover a pseudo-rapidity range of $|\eta| < 1.2$. The Muon Endcap system comprises four hundred and sixty eight CSCs in the two end-caps, divided in five disks each. A CSC is trapezoidal in shape and consists of six gaps, each gap having a plane of radial cathode strips and a plane of anode wires running almost perpendicularly to the strips. The signal on the wires is fast and is used in the Level-1 Trigger.

The stations in the barrel (and the disks in the end-caps) are separated by iron which collects the return magnetic flux. The iron has the double task of stopping particles debris of hadron shower escaping the hadron calorimeter and producing a non saturated (1.7 Tesla) field parallel to the beam line. It allows an almost field-less region for the DT chambers and yields the bending for transverse momentum measurement. Due to the calorimeter material in front of the first station, muons coming from the interaction region reach the first station if generated with momenta larger than 4-5 GeV/c. Muons can reach the last DT station if they have a momentum larger than about 7 GeV/c.

Tracks in the muon system are reconstructed in 3D with a hit resolution of the order of 250 μm in the rϕ plane and of 500 μm in z.

The local reconstruction efficiency (i.e., inside single stations) is close to 100%, with local inefficiencies for bad-working stations and at the edges of stations (see Fig. 2.16 for DTs).

2.2.3 Trigger and Data Acquisition

At the designed LHC luminosity (10^{34} cm^{-2}s^{-1}) a rate of 10^9 interactions every second is expected.

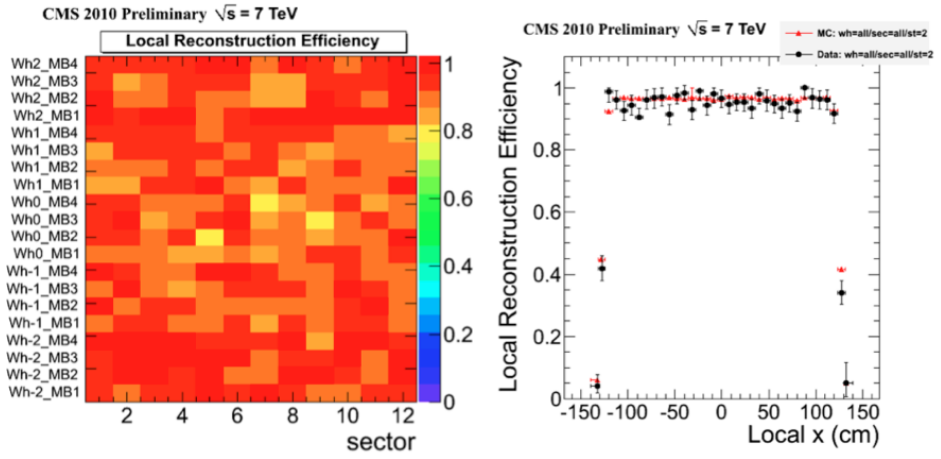


Figure 2.16: *DT local reconstruction efficiency.*

It is needed to reduce this rate by a factor of at least 10^7, i.e., down to 100 Hz, the maximum that can be archived by the on-line computer farm. A very efficient trigger system is necessary to select the most interesting proton-proton collision events. CMS trigger [24] is divided in two steps: the first, named *Level 1 trigger (L1)*, is based on custom electronics while the second, the *High Level Trigger (HLT)*, relies upon commercial processors. The hardware-based L1 trigger reduces the event rate from 1 GHz to 100 kHz while the software-based HLT brings the rate down to the final 100 Hz.

Level 1 trigger

A primary decision to retain an event (L1 accept) [47] has to be made within 3.5 μs with a frequency of 40 MHz.

Physics requirements on L1 are chosen to provide a high efficiency for the hard scattering physics to be studied at the LHC. This physics includes signals such as top decays, Higgs decays, W-W scattering, supersymmetry etc. The L1 trigger is based on the rough identification (at the level of detector electronics) of muons, electrons, photons and jets: the main requirement for a event in an acceptable pseudo-rapidity range is the presence of *high transverse momentum or energy*. Triggering is also important in the presence of considerable quantity of *missing transverse energy*. The trigger has to be inclusive and local. An initial selection of electrons, muons, photons and jets that relies on local information, is tied directly to their distinctive signatures, rather than on global topologies.

The L1 trigger involves the calorimetry and muon systems. In particular it is organized into three major subsystems: the L1 calorimeter trigger, the L1 muon trigger and the L1 global trigger. Each one is further organized into

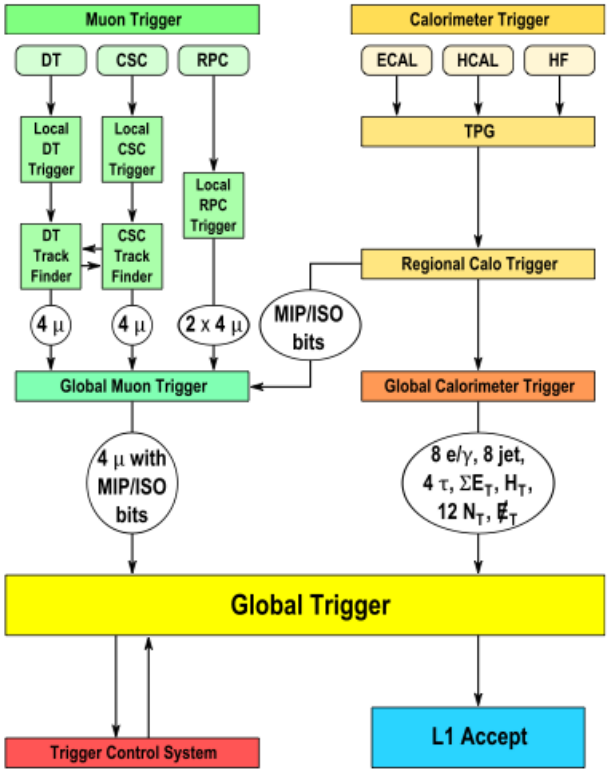


Figure 2.17: *Level-1 trigger scheme.*

sub-divisions, representing the sub-detectors. A scheme of the L1 trigger is reported in Fig. 2.17.

The High Level Trigger

The software-based High Level Trigger [47] takes data filtered from the L1 Trigger to perform a rough reconstruction of the events, and decide if an event should be kept for fine-reconstruction and data analysis, or rejected. This decision is based on the event suitability for inclusion in one of the various data-sets to be used for analysis. These data-sets are designed to be relevant for searches of top quark, Higgs boson, supersymmetry and other of the main physics topics for which the LHC was built. Typical datasets require high p_T associated to a HLT-reconstructed lepton or lepton pair, photon or photon-pair, jet or multi-jet event or the presence of a substantial amount of E_T^{miss}.

The Computing

The CMS data are available under several file formats, that contain a decreasing amount of information:

- RAW: the online HLT system creates "RAW" data events containing: the detector data, the L1 trigger result, the result of the HLT selections ("HLT trigger bits"), and some of the higher-level objects created during HLT processing.
- RECO: reconstructed objects (tracks, vertices, jets, electrons, muons, etc.) and reconstructed hits/clusters
- AOD: reconstructed objects (tracks, vertices, jets, electrons, muons, etc.), small quantities of very localized hit information.

The AOD format is the one used to perform physics analyses.

CMS relies on the "Grid" to analyze data, with informatics resources world-wide distributed. The Grid consists of several clusters of computers organized in hierarchial levels, with a primary "Tier-0" centre at CERN being supplemented by Tier-1 and Tier-2 centers at national laboratories and universities. Such a wide distributed calculus is needed in order to process the huge amount of CMS data and to run Monte Carlo simulations.

Chapter 3

Physical objects at CMS

The event reconstruction uses information collected by the detection apparatus in each proton-proton collision (reco-hits, energy deposits etc) to build the physical objects that left the original "trace" in the detector, and assign to these objects their kinematic properties like energy, momentum, position etc. The physical objects used in the physics analyses are (meta)stable particles like muons, electrons and photons or hadronic jets, formed by several hadronic particles, and missing transverse energy, a global variable associated to the event looking at the energy unbalance in the transverse plane.

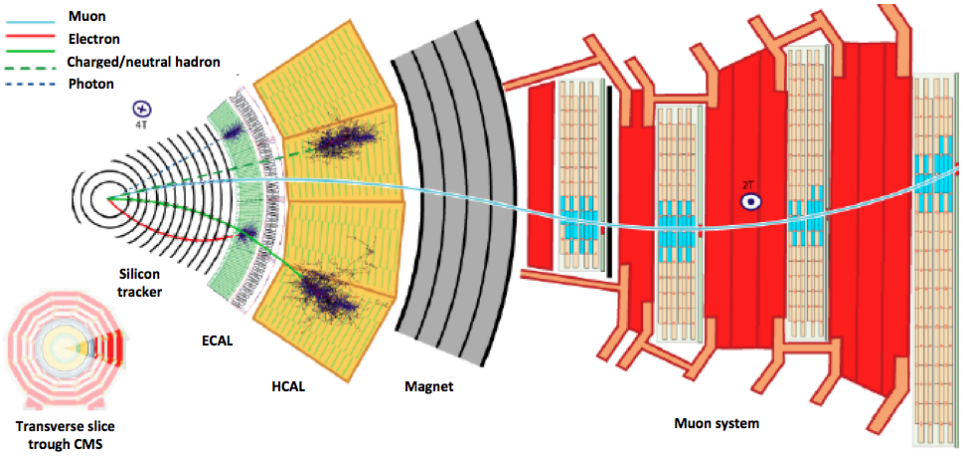


Figure 3.1: *A transverse section of CMS, showing the behavior of different particles when crossing the detection system.*

The behavior of different particles in the detector makes it possible to identify them (see Fig. 3.1):

- *Muons* cross the whole detector leaving little energy deposits in the calorimeters, they produce ionization in the inner tracker and in the

muon system (from this ionization, hits are reconstructed and tracks built);

- *Electrons* produce ionization in the inner tracker, and deposit all their energy in the electromagnetic calorimeter;
- *Photons* deposit all their energy in the electromagnetic calorimeter, without leaving tracks (since they are neutral);
- *Hadronic particles* (meta-stable mesons and baryons) are stopped in the hadron calorimeter leaving energy deposits and, if charged, produce ionization in the inner tracker.

To meet CMS goals reported in section 2.2.1, high efficiency in reconstruction and identification is required as well as high precision in assignment of kinematic properties.

The *luminosity measurement* is also crucial: the amount of statistics collected is expressed in units of integrated luminosity, and all physical results strongly depend on it.

The main CMS algorithms to measure luminosity and reconstruct objects are presented in the following sections. Particular emphasis is put on lepton reconstruction and identification, since leptons are used in the $H \rightarrow ZZ \rightarrow 4l$ analysis. The variables "isolation" and "significance of impact parameter" are presented, together with the treatment of dependence on the number of pile-up events.

3.1 Luminosity measurement

The *instantaneous luminosity* L at a collider machine, where one observes a rate of R for a physical process with cross section σ, is defined as:

$$L = \frac{R}{\sigma}$$

It depends on machine parameters, such as beam transverse dimensions, number of protons per bunch and number of bunches in the beams. It is a time dependent quantity $L(t)$ (since the beams degrade with time) that integrated over the data-taking period becomes the measurement of collected statistics, called *integrated luminosity*:

$$L_{int} = \int_{time} L(t)\ \mathrm{d}t$$

The real-time luminosity monitoring must be based on the measurement of a high, well known, cross section process using as little hardware as possible. One method of luminosity determination [31] is the "zero counting", where the luminosity is derived from the probability that a tower of the hadronic

forward calorimeter (HF) sees zero hits in a single collision. Another, more accurate method is "Pixel Cluster counting": there is a good correlation between the luminosity and the number of hit clusters in the Silicon Pixel detector. Due to their tiny size the individual pixels have a small occupancy and this allows to use this method also at very high luminosities.

The pixel-based method is used in CMS. The systematic error assigned to the luminosity collected during 2011 (2012) is 2.2% (4.4%[1]) [50].

3.2 The "Particle Flow" event reconstruction

The *particle-flow event reconstruction (PF)* [51][52] aims at reconstructing and identifying all stable particles in the event, i.e., electrons, muons, photons, charged hadrons and neutral hadrons, combining all CMS subdetectors for an optimal determination of their direction, energy and type (see Fig. 3.2). The list of individual particles is then used, as if it came

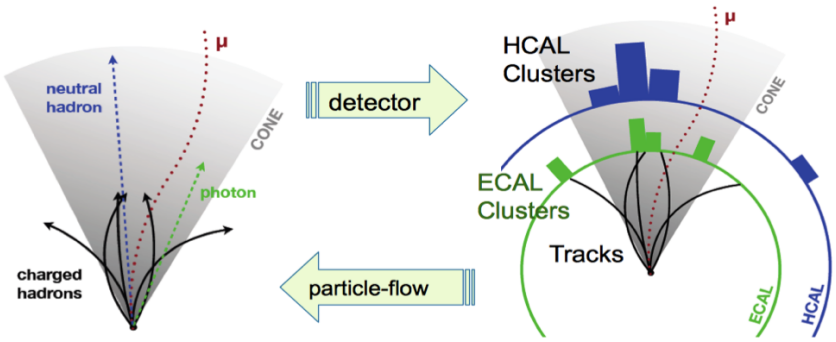


Figure 3.2: *The Particle-Flow concept pictorially.*

from a Monte-Carlo event generator, to build jets, to determine the missing transverse energy E_T^{miss}, to reconstruct and identify taus from their decay products, to quantify charged lepton isolation with respect to other particles, to tag b-jets, etc. Further details on PF algorithm are given in section 3.7.

The PF reconstruction is possible thanks to the CMS granularity (in particular in the tracker and ECAL) and the high magnetic field.

When used in analyses, PF has shown to perform very well in terms of efficiency and purity of samples, resolution of jet kinematics and E_T^{miss}. For this reason it is preferred to plain detector-based reconstruction and identification.

[1] In general, systematic uncertainties are higher in 2012 data. 2012 data are recent and less understood than 2011 ones; in addition the calibrations of detectors are preliminary in many cases.

3.3 Muon reconstruction and identification

In the standard CMS *muon* reconstruction [53][32][54] tracks are first reconstructed independently in the inner tracker (tracker track) and in the muon system (standalone-muon track). Then, the information is combined into a unique object, with two possibilities:

- *Global Muon Reconstruction (outside-in):* for each standalone-muon track, a matching tracker track is found by comparing the parameters of the two tracks, propagated onto a common surface, and a global-muon track is fitted combining the hits, using a Kalman-filter technique.

- *Tracker Muon Reconstruction (inside-out):* all tracker tracks with transverse momentum $p_T > 0.5$ GeV/c and momentum $p > 2.5$ GeV/c are considered as possible muon candidates and are extrapolated to the muon system taking into account the magnetic field, the average expected energy losses, and multiple scattering in the detector material. If at least one muon segment (i.e., a short track segment made of DT or CSC hits) matches the extrapolated track, the corresponding tracker track qualifies as a "Tracker Muon".

Tracker-muons are used to recover efficiency at low p_T, namely $p_T < 5$ GeV/c , given the weak penetrating power of low momentum muons.

Thanks to the high tracker-track efficiency and a very high efficiency of reconstructing segments in the muon system, about 99% of muons produced in pp collisions and having sufficiently high momentum ($p_T \gtrsim 5$ GeV/c) are reconstructed either as a Global Muon or a Tracker Muon.

The set of reconstructed muons contains a significant amount of misidentified (un-decayed) charged hadrons. In order to have a pure sample of muon candidates, identification requirements must be applied to the original collection. The selections typically used are [54]:

- *Soft muon selection:* the candidate must be a Tracker Muon, with the additional requirement that a segment in the muon system is 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. These additional requirements are optimized for low p_T (< 10 GeV/c) muons.

- *Tight muon selection:* the candidate must be reconstructed outside-in as a Global Muon, with the following additional requirements:
  - normalized χ^2 of the global-muon track fit < 10;
  - at least one muon chamber hit included in the final track fit;
  - matched to muon segments in at least two muon stations;

- more than 10 tracker hits (and at least one pixel hit);
- a transverse impact parameter $|d_{xy}| < 2$ mm.

- *Particle-Flow muon selection:* the selection criteria differ in strictness depending on whether the muon candidate is isolated or not, and whether its momentum is compatible with the energy deposition in the calorimeters assigned to the candidate by the particle-flow event reconstruction. The particle-flow muons are identified by the union of three different selections referred to as "isolated", "pf-tight" and "pf-loose". Reco muons are considered to be isolated if, in a cone of size $R = \sqrt{\phi^2 + \eta^2} = 0.3$ centered on the muon, the sum of the p_T of the tracks and of the transverse energy of the calorimeter hits is less than 10% of the muon p_T. After isolated muons are selected, the pf-tight and pf-loose muon selections are applied to the remaining reco muons. The pf-tight selection requires a minimum number of hits in the muon track and compatibility of the muon segment and calorimeter deposits as defined by a template based on simulation. Other reco muons, which have a track momentum significantly larger than the corresponding energy deposit in the calorimeter thereby making them incompatible with a charged hadron hypothesis, can be recovered by the pf-loose selection. In this selection, the requirements on the number of hits are relaxed, and the template-based compatibility requirement is replaced by a matching requirement of the track to hits in the muon stations.

The efficiency for the PF muon ID (the union of all the described selections) is calculated with the Tag&Probe technique [55]: it uses a muon pair coming from a heavy resonance (typically the Z), where one of the the muons is very well identified (tag) and the other is tested (probe), to extract the probe reconstruction and identification efficiency with a minimum biasing. The efficiency is shown in Fig. 3.3 as a function of the muon p_T, for 2011 (7 TeV) and 2012 (8 TeV) data and Monte Carlo (MC) simulation. The efficiencies are close to unity even for relatively low p_T, close to 5 GeV/c ; data and simulations are in good agreement, especially in the 7 TeV sample.

3.4 Electron reconstruction and identification

The standard algorithm for *electron* reconstruction in CMS is called *Gaussian Sum Filter (GSF)* [56][33]. It combines ECAL and tracker information. The energy deposits in the ECAL, are grouped in clusters and superclusters [58][59]. In the barrel region of ECAL, superclusters are formed by energy sums, clustered in a rectangle of 35 crystals in ϕ and 5 in η. The superclusters collect Bremsstrahlung photons emitted in the tracker volume. They

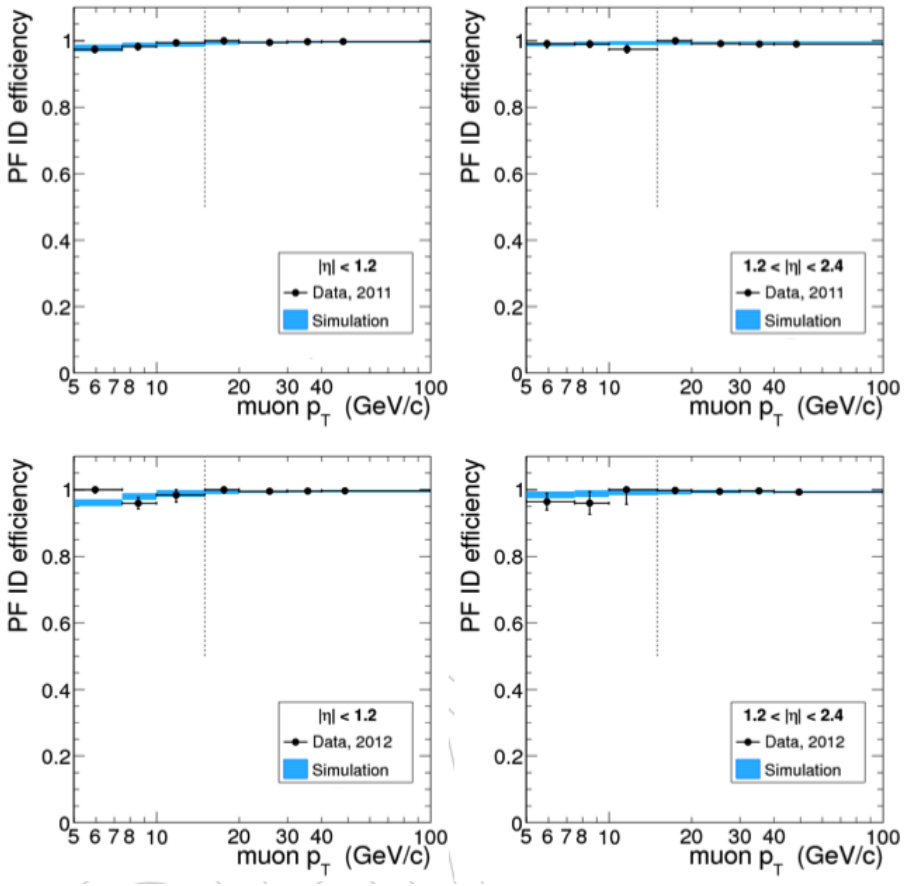


Figure 3.3: *PF muon ID efficiency, calculated with Tag&Probe technique, for 2011 (top) and 2012 data (bottom), in two different η regions (left and right).*

are used to search for hits in the innermost tracker layers, which seed electron tracks. Trajectories in the tracker volume are reconstructed using a dedicated modeling for electron energy loss (with a sum of gaussians) and fitted. A cleaning is performed to resolve ambiguous cases where several tracks are reconstructed due to the conversion of radiated photons in the tracker material. The four-momenta of electrons are obtained by taking angles from the associated GSF track, and the energy from the combination of tracker and ECAL information.

The reconstruction efficiency for isolated electrons is expected to be above 90% over the full ECAL acceptance. Integrated over the acceptance, the reconstruction efficiency for basic electron objects steeply rises to reach ∼90% at $p_T \sim 10$ GeV/c, and then more slowly to a plateau at ∼95% for $p_T \gtrsim 30$ GeV/c.

The purity of the sample of electron candidates (that is contaminated by hadrons) is enhanced by applying identification requirements to the basic collection of reconstructed electron objects. Electron candidates are selected using a multivariate technique: a Boosted Decision Tree (BDT) [57]. This BDT makes use of three main categories of variables: observables that match the information of the calorimeters and the tracker (including the pre-shower), purely-calorimetric and purely-tracking observables. The information carried out by these variables is "summarized" into a unique variable.

The complete list of BDT input variables is:

- track-ECAL matching variables:
  - E_{tot}/p_{in}, where E_{tot} is the supercluster energy and p_{in} the track momentum at the innermost track position;
  - E_e/p_{out}, where E_e is the energy of the cluster closest to the electron track extrapolation to ECAL and p_{out} the track momentum at the outermost track position;
  - $|\Delta\eta_{in}| = |\eta_{sc} - \eta_{in}^{extrap.}|$, where $|\eta_{sc}|$ is the energy weighted position in η of the supercluster and $\eta_{in}^{extrap.}$ is the η coordinate of the position of the closest approach to the supercluster position, extrapolating from the innermost track position and direction;
  - $|\Delta\phi_{in}| = |\phi_{sc} - \phi_{in}^{extrap.}|$, where $|\Delta\phi_{in}|$ is a quantity similar to the former one but in azimuthal coordinates;
  - $|\Delta\eta_{out}| = |\eta_e - \eta_{out}^{extrap.}|$, where η_e is the η position of the cluster closest to the electron track extrapolation to ECAL $\eta_{out}^{extrap.}$;
  - $1/E_{tot} - 1/p_{4-mom}$, which measures the deviation of the supercluster energy and electron momentum obtained by combining the tracker and ECAL information;
- shower shape variables:
  - $\sigma_{i\eta}$ the width of the ECAL cluster along the η direction computed in the 5×5 block of crystals centered on the highest energy crystal of the seed cluster;
  - $\sigma_{i\phi}$ as the former but in azimuthal coordinates;
  - $\eta - width$ supercluster η width;
  - $\phi - width$ supercluster ϕ width;
  - $(E_{5\times5} - E_{5\times1})/E_{5\times5}$: where $E_{5\times5}$ is the energy computed in block of crystals and $E_{5\times1}$ is the energy computed in the strip of crystals containing the cluster seed along ϕ;
  - $R_9 = E_{3\times3}/E_{tot}$ energy sum of 3×3 crystal centered on the most energetic, divided by the supercluster energy.

- To further improve the separation between electrons and charged hadrons, pure tracking observables are also used, both using the dedicated CMS electron tracking (GSF) or the standard Kalman-Filter tracks:
  - $f_{brem} = (p_{in} - p_{out})/p_{in}$ with the GSF track, which measures very well the Bremsstrahlung emission which helps in discriminating against charged-hadron particles;
  - χ^2_{GSF} ;
  - $hits_{KF}$;
  - χ^2_{KF} .

The training of the BDTs is performed on:

- background: a W+1-fake electron sample, taken from data;
- signal: a mixture of simulated $H \rightarrow ZZ \rightarrow 4e$ samples (with masses: 115-120-130-140 GeV/c^2).

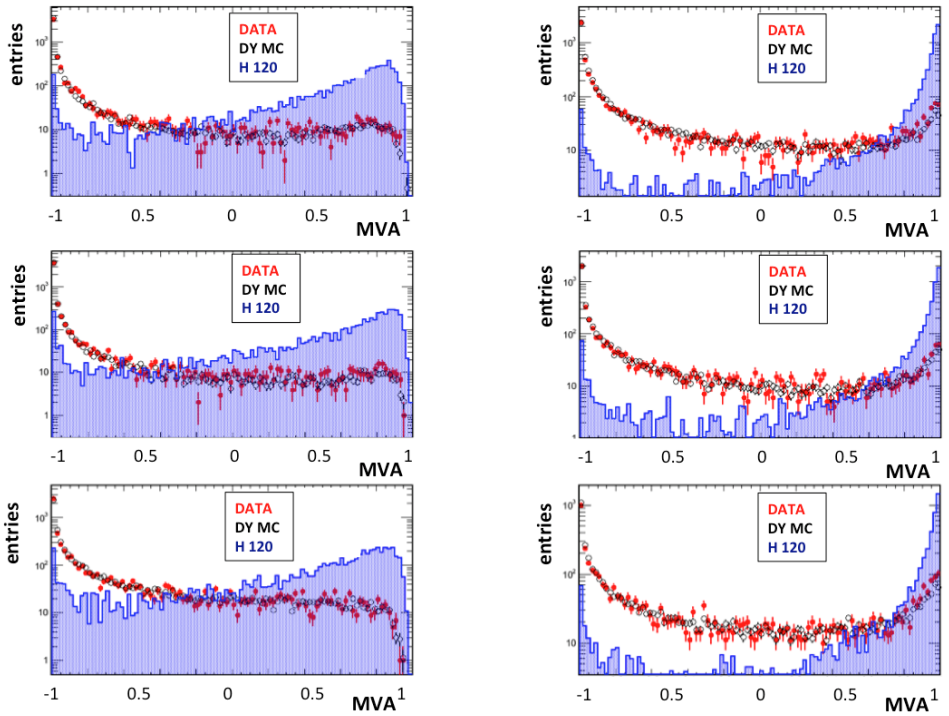


Figure 3.4: *Electron BDT output, one plot per (η,ϕ) bin: η growing from top-bottom, p_T growing from left-right.*

The training of the multivariate analysis is performed in three different bins in η and two in p_T in order to take into account the different material

budget in the tracker and the dependence of the electron observables on p_T. The output of the BDT is presented in Fig. 3.4. In each of the six p_T and η bins, the cut value on the BDT output was chosen so as to obtain the same background efficiency as the cut-based electron identification algorithm used in a previous analysis [62]. The BDT Working Points (WP) are reported in the following table:

BDT WP (>)	$p_T \in [5, 10]$ GeV/c	$p_T > 10$ GeV/c
$\eta < 0.8$	0.47	0.5
$\eta \in [0.8, 1.479]$	0.004	0.12
$\eta > 1.479$	0.295	0.6

The efficiency for electron reconstruction and BDT identification, calculated with the Tag&Probe technique [55], is shown in Fig. 3.5 as a function of the electron p_T, for 2011 (7 TeV) and 2012 (8 TeV) data and for MC simulation. The close-to-unity efficiency plateau is reached for momenta $\gtrsim 20$ GeV/c. At lower $p_T \sim 7$ GeV/c , the efficiencies decrease to 70-80%; data and simulations are in good agreement.

3.5 Photon reconstruction and identification

Photon candidates are reconstructed from energy deposits in the ECAL, superclusters [58][59]. This allows almost complete recovery of energy emitted via Bremsstrahlung when a photon converts [60][61], $\gamma \rightarrow ee$, and the electrons are bent by the 3.8 T field. In the end-cap region of the ECAL, superclusters are composed of one or more arrays of 5×5 crystals. Energy corrections are applied to the superclusters to take into account the interactions with the material in front of ECAL and shower containment. Electrons from conversions are also reconstructed using the supercluster as a seed for track finding, and information about the tracks and the fitted vertex are collected in a conversion object. Photon objects are then built with the relevant superclusters, assigning to the candidate the momentum in the location of the reconstructed primary vertex and taking into account the full information about the conversion (if present).

Photon energy calibration coincides with the ECAL calibration described in section 2.2.2.

Photons are identified and reconstructed by the particle-flow reconstruction with a specific clustering algorithm, efficient down to an energy of 230 MeV in the ECAL barrel and 600 MeV in the ECAL end-caps. The determination of photon energies and directions is verified in the data with $\pi^0 \rightarrow \gamma\gamma$ decays (see section 2.2.2), and is shown to be accurate, reliable, and in agreement with the predictions from simulation.

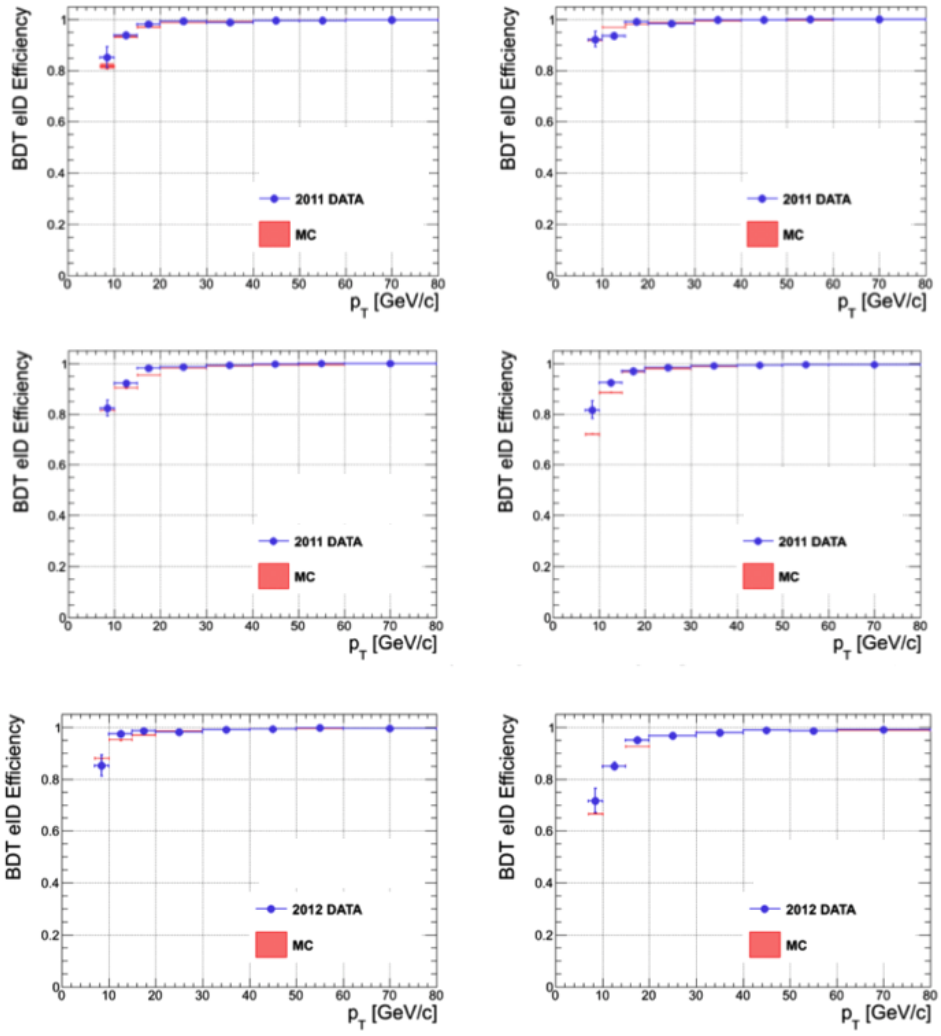


Figure 3.5: *BDT electron ID efficiency, calculated with Tag&Probe technique, for 2011 and 2012 data.*

Photons are required to be further away from the direction of any electron by 0.05 in η and by 2.0 rad in ϕ, to exclude photons included in electron superclusters.

The photon isolation is determined from the charged hadrons, photons and neutral hadrons identified by the particle-flow reconstruction in a cone of size $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.30$ around the photon direction. In this cone, all charged hadrons compatible as originating from the primary vertex and with a p_T larger than 200 MeV/c, all photons and neutral hadrons with

a p_T larger than 500 MeV/c are included in the sum for isolation computation. The absolute photon isolation is defined as the sum of the transverse momenta of all these contributions, called "iso-deposits". To discriminate against photons that are produced in pileup interactions, an additional isolation deposit is defined that corresponds to the charged particle sum from the vertices other than the primary vertex. Finally, the relative isolation is obtained by dividing the absolute isolation by the photon transverse momentum.

3.6 Prompt lepton selection

The discrimination between prompt leptons, e.g., coming from W or Z/γ^* bosons, and secondary leptons, e., coming from hadron decays or τ decay, relies mostly on two key variables: the significance of impact parameter and the isolation, that are briefly presented in the following paragraphs [64].

Significance of impact parameter

A lepton is considered to come from the event primary vertex if the *significance of its impact parameter (SIP)* is small enough . The Impact Parameter (IP_{3D}) is the distance of closest approach of the lepton track with respect to the reconstructed primary vertex, in 3D. When normalized to its significance, it gives the SIP_{3D}:

$$\mathrm{SIP}_{3D} = \frac{\mathrm{IP}_{3D}}{\sigma_{IP_{3D}}}$$

Primary leptons are selected requiring that $\mathrm{SIP}_{3D} < 4$.

Lepton isolation

The *lepton isolation* variable is evaluated performing the scalar sum of the transverse momenta of the particle flow candidates, reconstructed in a cone of $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.4$ around the lepton track, normalized to lepton p_T:

$$\mathrm{Iso}_{PF} = \frac{\sum_{\text{charged hadrons}} p_T \; + \; \sum_{\text{neutral hadrons}} p_T \; + \; \sum_{\text{photons}} p_T}{p_T^{\text{lepton}}} \qquad (3.1)$$

Some vetoes are posed to exclude the lepton itself from the computation. For electrons:

- barrel and end-cap: veto all the reconstructed particle-flow electrons and veto all the charged hadrons that share the same GSF track or the closest Combinatorial Track Finder (CTF) track with the electron

- end-cap: veto all the charged hadrons in a cone $\Delta R = 0.015$ and all the photons in a cone $\Delta R = 0.08$ around the electron

For muons:

- veto cone of $\Delta R = 0.01$ around the muon track

The cut on isolation is optimized for each analysis. For the $H \rightarrow ZZ \rightarrow 4l$ analysis, the cut is $\mathrm{Iso}_{PF} < 0.4$.

3.6.1 Pile-up dependence

The pile-up of collisions per bunch crossing has been increasing during the LHC operations (2010 - 2012), along with the instantaneous luminosity. The maximum number of pile-up events, averaged event by event, was 18 in 2011, while in 2012 it was 34 (in Fig. 3.6 an exceptional event with 78 reconstructed vertices is shown). With increasing pile-up the isolation variable tends to increase due to more energy entering the isolation cone. The charged particle-flow candidates, to be used for isolation computation, are first filtered through the algorithm"pfNoPileup" that relies on primary vertex association, while the neutral contribution is corrected with the "Fast-Jet" [66][67] algorithm, where the "fast jet energy density" (ρ) is used 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, $\overline{\rho}$ for each event, is taken. The correction to the neutral component of the isolation variable, the second term in formula 3.1, is then applied according to the formula :

$$\sum_{\text{n.h.}} p_T \quad \longrightarrow \quad \max\left(\sum_{\text{n.h.}} p_T - \overline{\rho}\cdot \mathrm{A}_{eff}(\eta)\ ,\ 0\ \text{GeV}\right)$$

where the effective area ($\mathrm{A}_{eff}(\eta)$) is defined as the ratio between the slopes (m_1 and m_2) of $\overline{\sum_{\text{n.h.}} p_T}$ and $\overline{\rho}$ as functions of the number of vertices N_{vtx}:

$$\overline{\sum_{\text{n.h.}} p_T}(\mathrm{N}_{vtx}) = m_1 \cdot \mathrm{N}_{vtx} + const.$$

$$\overline{\rho}(\mathrm{N}_{vtx}) = m_2 \cdot \mathrm{N}_{vtx} + const.$$

$$\mathrm{A}_{eff}(\eta) = \frac{m_1}{m_2}$$

In 2011, the energy density ρ was calculated using jets reconstructed with k_T algorithm (see section 3.7), taking as input all particles built with the particle-flow algorithm. In 2012, the same ρ definition was taken for muons. For electrons, all particles were considered, up to $|\eta| < 3$.

The effective areas $\mathrm{A}_{eff}(\eta)$ were computed, for muons and electrons, from 2011 and 2012 data, in bins of pseudo-rapidity, as reported in the following tables:

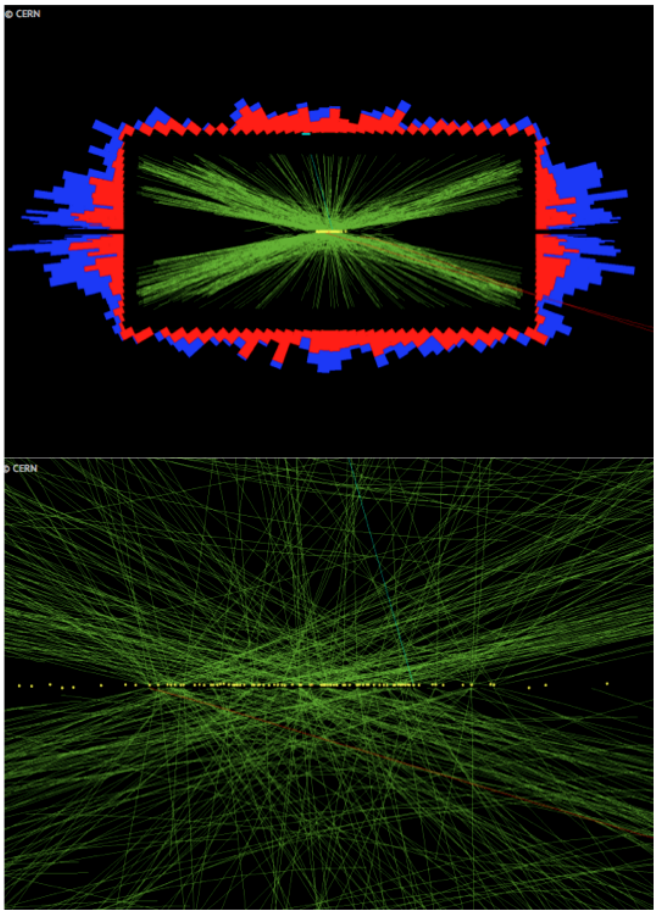


Figure 3.6: *An event with 78 reconstructed vertices.*

$\mathbf{A}_{eff}$ **MUONS**	2011	2012
$\|\eta\| < 1$	0.13	0.67
$\|\eta\| \in [1, 1.479]$	0.12	0.57
$\|\eta\| \in [1.479, 2]$	0.11	0.44
$\|\eta\| \in [2, 2.2]$	0.14	0.52
$\|\eta\| \in [2.2, 2.3]$	0.17	0.82
$\|\eta\| > 2.3$	0.19	0.66

$\mathbf{A}_{eff}$ **ELECTRONS**	2011	2012
$\lvert\eta\rvert < 1$	0.18	0.19
$\lvert\eta\rvert \in [1, 1.479]$	0.20	0.25
$\lvert\eta\rvert \in [1.479, 2]$	0.15	0.12
$\lvert\eta\rvert \in [2, 2.2]$	0.19	0.21
$\lvert\eta\rvert \in [2.2, 2.3]$	0.21	0.27
$\lvert\eta\rvert \in [2.3, 2.4]$	0.22	0.44
$\lvert\eta\rvert > 2.4$	0.29	0.52

The uncertainties on the effective areas is of the order of 1%.

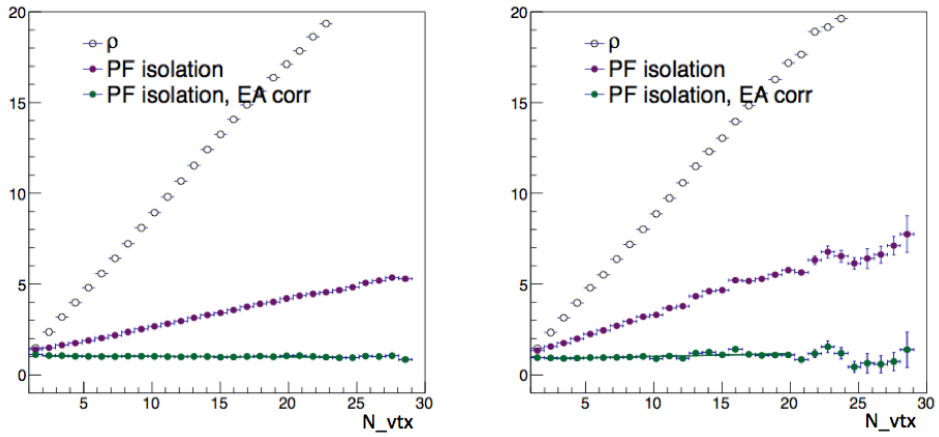


Figure 3.7: *ρ, ($Iso_{PF} \cdot p_T^{lepton}$) and ($Iso_{PF}^{corr} \cdot p_T^{lepton}$) as functions of N_{vtx} for identified electrons with a $Z \to ee$ selection in 2011 data, in two η bins: $|\eta| < 1.0$ (left), $2.2 < |\eta| < 2.3$ (right).*

After the FastJet correction the isolation variables show no dependence on the number of vertices, as shown in Fig. 3.7 for electrons.

3.6.2 Prompt lepton efficiencies

The cuts on isolation and SIP_{3D} ensure a high purity of the final lepton sample, while keeping a high efficiency for prompt lepton selection. The efficiencies, after SIP and isolation cuts are applied, are shown in Fig. 3.8, 3.9, for muons (and have to be compared with Fig. 3.3): the cut on SIP_{3D} still ensures a $> 99\%$ efficiency for muons with $p_T > 20$ GeV/c , the effects of isolation cut are more important at low p_T, here by a $\sim 5\%$. In Fig. 3.10 the case of electrons is shown (and has to be compared with Fig. 3.5): the efficiencies are lowered, by isolation and SIP cuts, of a $\sim 5\%$ for $p_T > 20$ GeV/c , and of a $\sim 10\%$ for p_T approaching 7 GeV/c .

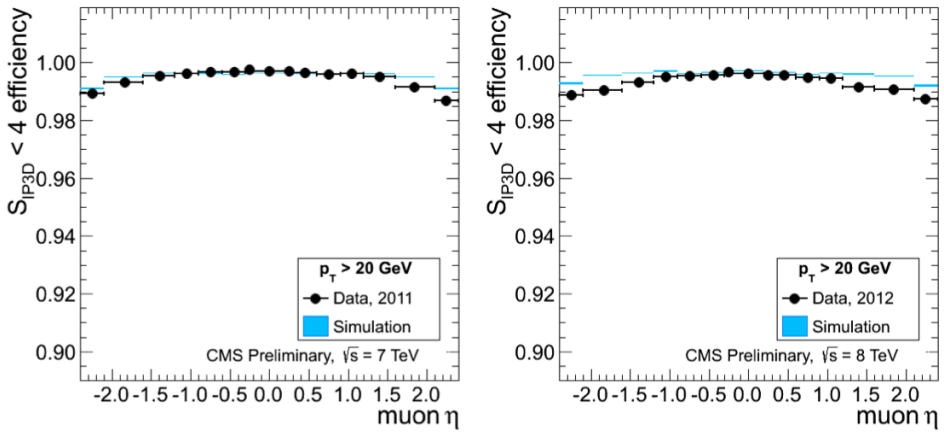


Figure 3.8: *Efficiency for the cut on the 3D impact parameter significance* $|SIP_{3D}| < 4$*, as function of muon pseudo-rapidity, for 2011 data (left) and 2012 data (right). Muons with* $p_T > 20$ *GeV and satisfying the Particle Flow identification requirements are used.*

3.7 Jet & E_T^{miss} reconstruction

The CMS standard *jet* reconstruction algorithms [33] use calorimeter information. Readout cells in HCAL are arranged in a tower pattern in η, ϕ space, projective to the nominal interaction point. The towers are used as input to several jet clustering algorithms. The energy associated with a tower is calculated as the sum of all contributing readout cells passing the online zero-suppression threshold and any additional offline software thresholds. Three calo-based jet reconstruction algorithms have been developed and studied for CMS: the *Iterative cone*, the *Midpoint cone* and the *Inclusive* k_T *jet* algorithms. They are documented in [34].

The *transverse energy balance* (E_T^{miss}) [33] can be measured with a good enough accuracy to help establish a physics signature involving one or more non-interacting particles.

The charged hadrons, neutral hadrons and photons are the basic constituents of hadronic jets. The *particle-flow algorithm* [51] exploits very efficiently the information coming from all sub-detectors: in particular the tracker has a very high resolution in p_T, the ECAL has a high granularity and an optimal energy resolution, while HCAL has a worse energy resolution and a lower granularity. The typical jet energy fractions carried by charged particles, photons and neutral hadrons are 65%, 25% and 10% respectively. These fractions ensure that 90% of the jet energy can be reconstructed with good precision by the particle-flow algorithm, both in value and direction, while only 10% of the energy is affected by the poor HCAL resolution. The reconstruction of jets starts from fundamental "elements", the charged-

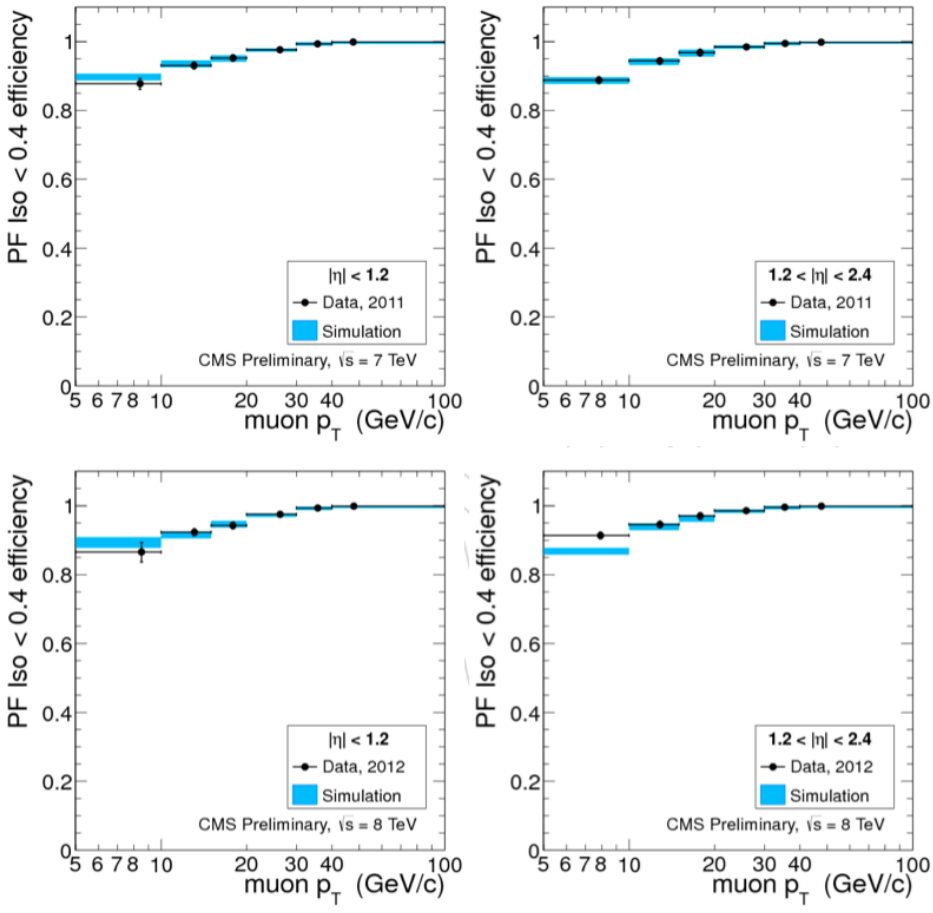


Figure 3.9: *Muon isolation efficiency for Particle Flow muons passing the impact parameter cut, measured with the Tag&Probe method in 2011 data (top) and 2012 data (bottom), as function of muon p_T, in the barrel (left) and end-caps (right).*

particle tracks and the calorimetric clusters, which are then topologically linked into "blocks". The core of the particle-flow algorithm interprets the blocks in terms of particles. The resulting list of reconstructed particles constitutes a global description of each event, available for subsequent physics analysis. The detailed procedure is explained in [51].

A snapshot of PF performances with jets can be seen in Fig. 3.11: the jet matching efficiency for PF jets and calo jets is shown as a function of p_T: the performances are comparable only above $p_T \sim 70$ GeV/c , for lower p_T values the performance of the PF algorithm is dramatically better. The p_T measurement resolution is also much better for PF jets, as shown in Fig. 3.12: the $(p_{rec} - p_{gen})/p_{gen}$ distribution for PF-jets is narrower than the calo-jet one, and less biased.

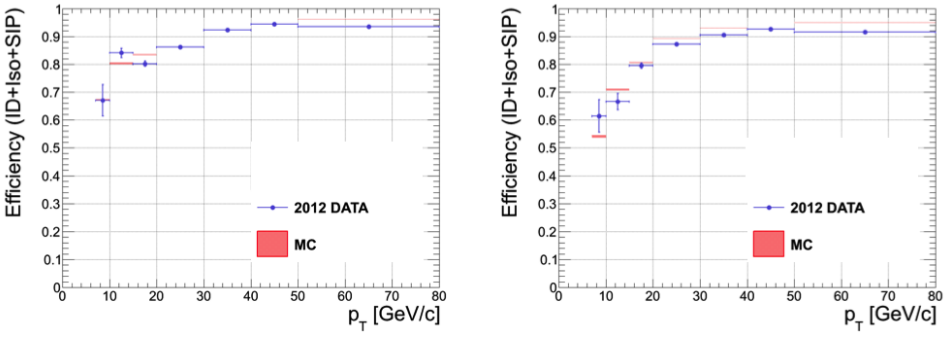
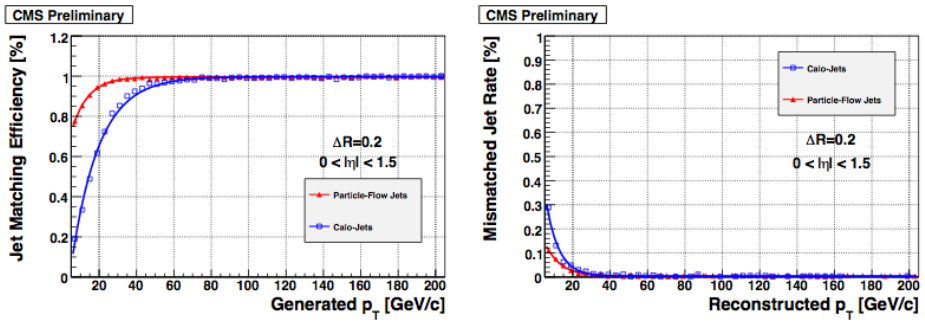
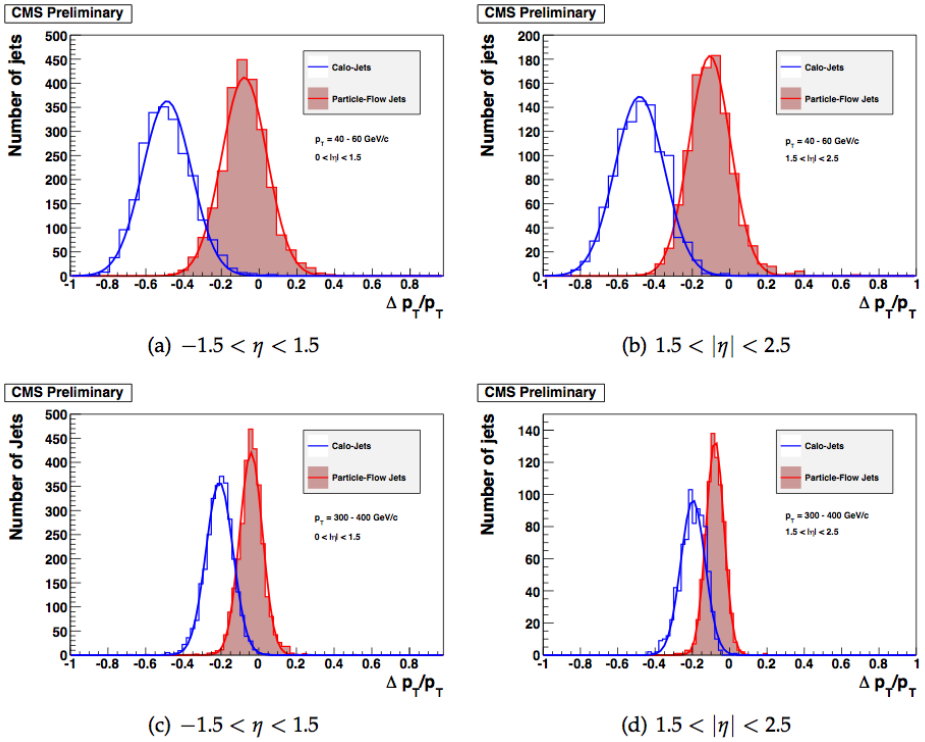


Figure 3.10: *Electron identification+isolation+$|SIP_{3D}|$ efficiencies (computed with the tag&probe method) as a function of the probe p_T , in two different η bins: $|\eta| < 1.442$ (left) and $1.442 < |\eta| < 2.5$ (right). Results are for 8 TeV data.*



(a) $\Delta R = 0.2$ (b) $\Delta R = 0.2$

Figure 3.11: *Jet matching efficiency (a) and mismatched jet rate (b), as obtained for calo-jets (open squares) and particle-flow jets (triangles) in the barrel, with a matching distance of 0.2. Efficiencies and fake rates are fit to exponential functions of p_T.*



(a) $-1.5 < \eta < 1.5$ (b) $1.5 < |\eta| < 2.5$

(c) $-1.5 < \eta < 1.5$ (d) $1.5 < |\eta| < 2.5$

Figure 3.12: *Distributions of* $(p_{rec} - p_{gen})/p_{gen}$ *for* p_{gen} *between 40 and 60 GeV/c (a,b) and between 300 and 400 GeV/c (c,d), as obtained from calo-jets (open histogram) and from particle-flow jets (solid histogram) pointing to the barrel (a,c) and to the end-caps (b,d). A Gaussian is used to fit the distributions.*

Chapter 4

Introduction to the $H \rightarrow ZZ \rightarrow 4l$ analysis

The $H \rightarrow ZZ \rightarrow 4l$ analysis [37][38][62][63][64][65], at low Higgs boson masses, relies on the selection of pp events containing four isolated prompt leptons, coming from two Z bosons, possibly off-mass-shell. With "leptons" we mean muons or electrons: the final state can occur in three sub-channels, 4μ, $4e$ and $2e2\mu$, that summed together form the inclusive 4-lepton channel. The Higgs boson would be revealed by the presence of a narrow resonance peak, over a broader background, in the 4-lepton invariant mass (m_{4l}) spectrum. The signal rate is relatively small, ranging from about 1 to 10 events per fb^{-1}, depending on the Higgs mass, but the experimental signature is very clean, permitting the peak reconstruction with a resolution of the order of few GeV/c^2 . The signal over background ratio is locally of the order of one-to-one, in the whole mass range. The main backgrounds are:

- the *irreducible* non resonant double Z production:

$$pp \rightarrow ZZ \rightarrow l_1^+ l_1^- l_2^+ l_2^- \qquad l_i = e, \mu, \tau$$

  The events from this SM process have the same characteristics as the signal ones: their rate cannot be reduced by applying cuts on lepton observables like isolation or impact parameter, or similar detector-related quantities. This is the dominant background in almost the whole mass range.

- the *reducible* backgrounds:

$$pp \rightarrow Z + jets \qquad pp \rightarrow t\bar{t} \rightarrow WW(\rightarrow 2l2\nu)b\bar{b}$$

  where *jets* indicates both heavy (b/c) or light jets. Leptons come from b/c hadron decays or from hadron misidentification. The rate of these processes is drastically reduced by applying cuts on isolation and

significance of the impact parameter of the leptons (see section 3.6), since half of the leptons in these events are contained in jets and/or come from secondary vertices. After the cuts are applied, only a tiny contribution at low invariant mass survives.

To give an idea, in Fig. 4.1 is reported the $m_{4\mu}$ spectrum, expected for 30 fb^{-1}, with the main backgrounds and some signals, for different Higgs mass hypotheses, superimposed.

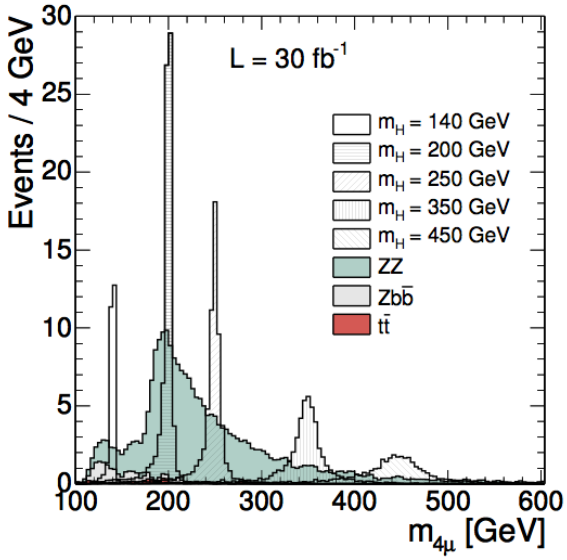


Figure 4.1: *Simulated data (4μ channel): $m_{4\mu}$ spectrum, expected for 30 fb^{-1} after a proper event selection, optimized for the $H \to ZZ \to 4l$ channel, is applied.*

The analysis is sensitive to the SM Higgs search in a wide mass range, from about $m_H \sim 115$ to 600 GeV/c^2 . However, given the excluded mass range (see section 1.2.4) and the current best fit to Higgs mass (see section 1.2), we specialize in a low mass analysis, in the range:

$$m_H \in [115, 150] \text{ GeV/c}^2$$

In this region, the Higgs peak in m_{4l} is very narrow, with a width of the order of 1 GeV/c^2 , dominated by the experimental resolution. Given the collected data, the analysis is designed to either find a Higgs signal, if present, or to exclude it. In case of a signal evidence, the Higgs mass and cross section should be measured. The determination of the spin-parity quantum numbers is expected to be inconclusive, with the present statistics.

The number of events that survives the $H \to ZZ \to 4l$ final selection is very small, since both Higgs and ZZ production are rare processes. It is very

useful to study the single Z boson first, with Z decaying into two leptons $Z \to ll$ ($l = e, \mu$), using the same object definition as for the $H \to ZZ \to 4l$ analysis. The single Z study benefits of a large statistics, since Z bosons are copiously produced at the LHC, and allows us to extract some systematics to be assigned to final event yields in the $H \to ZZ \to 4l$ analysis. This subject will be treated in sections 4.2 and 4.3.

4.1 Data and simulations

The data used in the analysis have been collected by CMS during the 2011-2012 period.

The LHC delivered 6.10 fb^{-1} with a center of mass energy of 7 TeV, and 23.26 fb^{-1} with a center of mass energy of 8 TeV.

Recorded CMS data undergo a certification process that involves all the subsystems: data are certified as good if all the subsystems showed a good behavior during the data taking period. This implies that the final amount of data, available for the physical analyses, is smaller than the LHC delivered one. The currently available statistics, for public analyses, is 5.05 fb^{-1} with a center of mass energy of 7 TeV and 11.93 fb^{-1} with a center of mass energy of 8 TeV (the latest public data are the ones delivered up to September 2012).

4.1.1 Experimental data

The $H \to ZZ \to 4l$ analysis uses primary datasets (PDs), produced centrally, which combine various collections of HLT. The detailed content of the PDs evolves in phase with the evolution of the trigger menu to cope with ever increasing instantaneous luminosity. For 2011 data, the analysis relies on the so-called *DoubleElectron* and *DoubleMuon* PDs [70]. These PDs are formed by a logic "OR" between various triggers with symmetric or asymmetric trigger thresholds (in p_T or E_T) for the two leptons, with or without additional identification and isolation requirements. In 2012, cross-triggers (with a muon and an electron) were added to recover few percent of efficiency in the $2e2\mu$ channel at low Higgs boson masses, forming the so-called *MuEG* PD. The trigger efficiency is evaluated from MC and an uncertainty of 1.5% is assigned to account for data/MC discrepancy [64].

The PDs and trigger paths used for this analysis are summarized in the following table:

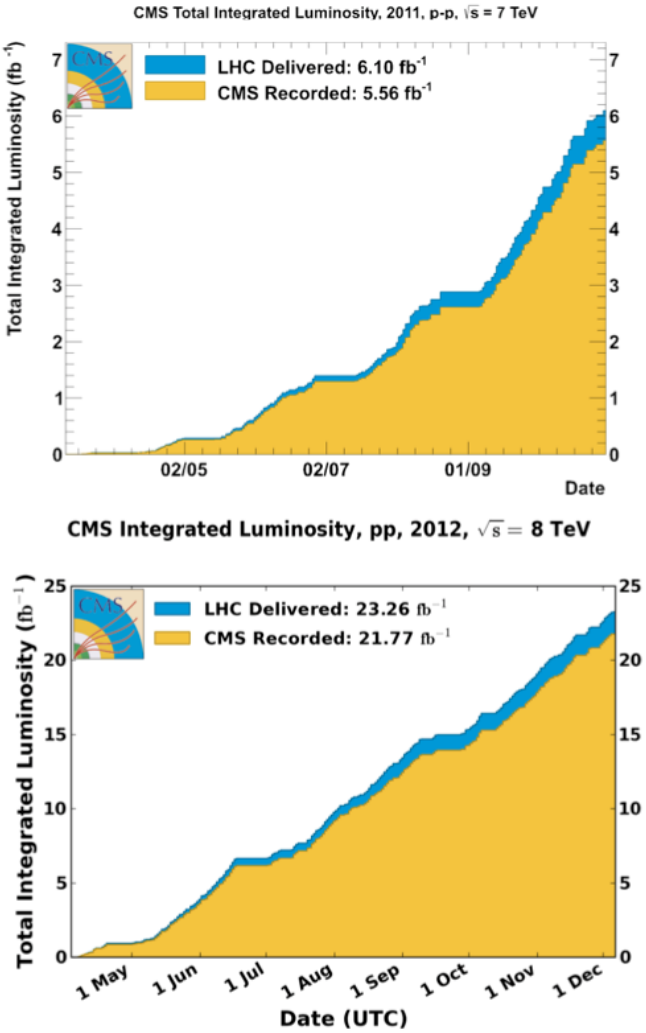


Figure 4.2: *Integrated luminosity delivered by LHC and recorded by CMS: in 2011 at 7 TeV (left) and 2012 at 8 TeV (right).*

2011	2012
Datasets	
/DoubleElectron/Run2011A-16Jan2012-v1 /DoubleMu/Run2011A-16Jan2012-v1 /DoubleElectron/Run2011B-16Jan2012-v1 /DoubleMu/Run2011B-16Jan2012-v1	/DoubleElectron/Run2012A-PromptReco-v1 /DoubleMu/Run2012A-PromptReco-v1 /DoubleElectron/Run2012B-PromptReco-v1 /DoubleMu/Run2012B-PromptReco-v1 /MuEG/Run2012A-PromptReco-v1 /MuEG/Run2012B-PromptReco-v1
Muon triggers	
HLT_DoubleMu7 OR HLT_Mu13_Mu8 OR HLT_Mu17_Mu8	HLT_Mu17_Mu8
Electron triggers	
HLT_Ele17_CaloTrk_Ele8_CaloTrk	HLT_Ele17_CaloTrk_Ele8_CaloTrk
Cross triggers	
	HLT_Mu17_TkMu8 OR HLT_Mu8_Ele17_CaloTrk OR HLT_Mu17_Ele8_CaloTrk
Integrated luminosity	
5.05 fb^{-1}	11.93 fb^{-1}

4.1.2 Simulated Samples

The SM Higgs boson signal samples and the background ones were obtained using detailed MC simulations.

The backgrounds include irreducible 4l contributions from di-boson production, via $q\bar{q} \rightarrow ZZ \rightarrow 4l$ and $gg \rightarrow ZZ \rightarrow 4l$, as well as instrumental backgrounds in which hadronic jets or secondary leptons from heavy meson decays are misidentified as primary leptons. The main sources of instrumental background contributions, are the $Z + jets$ production with $Z \rightarrow l^+l^-$ decays and the production of top quark pairs in the decay mode $t\bar{t} \rightarrow l^+l^-\nu\bar{\nu}b\bar{b}$.

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 [71].

The general multi-purpose MC event generator PYTHIA [74] is used to generate hard processes at leading order and for showering, hadronization, decays and addition of the underlying event. Event generation at higher orders makes use of the MadGraph (MadEvent) MC [75] event generators and of the POWHEG NLO generator [76], which is used for the Higgs boson signal and for the ZZ and $t\bar{t}$ background. Finally, the dedicated tool gg2zz [77] is used to generate the $gg \rightarrow ZZ$ contribution to the ZZ cross section. For the underlying event, the so-called "PYTHIA tune Z2" in 2011 and "PYTHIA tune Z2 star" in 2012, which rely on p_T-ordered showers are used.

In order to increase the statistics of some background samples, in addition to fully simulated events more events were produced using the f"Fast Simulation" (FASTSIM) [79] which, applying appropriate parametrizations, reduces the CPU time needed to simulate interactions within the detector. In particular, the Drell-Yan MC corresponds to ~ 10 fb^{-1} of full simulated events and to ~ 17 fb^{-1} of fast simulated ones.

The following table summarizes the Monte Carlo (MC) simulation datasets used for this analysis and the cross sections of the physical process:

Process	MC generator	$\sigma_{(N)NLO}$ @ 7 TeV	$\sigma_{(N)NLO}$ @ 8 TeV
Signal			
$gg \to H \to ZZ \to 4l$ $[m_H = 110 - 600$ GeV/c$^2]$	POWHEG	[1-20] fb	[1.2-25] fb
Irreducible background			
$q\bar{q} \to ZZ \to 4\mu(4e)$	POWHEG	15.34 fb	76.91 fb
$q\bar{q} \to ZZ \to 2e2\mu$	POWHEG	30.68 fb	176.7 fb
$q\bar{q} \to ZZ \to 2e(2\mu)2\tau$	POWHEG	30.68 fb	176.7 fb
$gg \to ZZ \to 2l2l'$	gg2zz	3.48 fb	4.47 fb
$gg \to ZZ \to 4l$	gg2zz	1.74 fb	2.24 fb
Reducible background			
Z + jets ($m_{ll} > 50$)	MadGraph	3048 pb	3503.7 pb
Z + jets ($m_{ll} > 50$)	MadGraph+FASTSIM	3048 pb	3503.7 pb
$t\bar{t} \to l^+l^-\nu\bar{\nu}b\bar{b}$	POWHEG	17.32 pb	23.64 pb

The cross sections for the $qq \to ZZ$ process at 8 TeV seem unnaturally high with respect to the 7 TeV ones: this is because the kinematical threshold on m_{Z_2} used in simulation was lowered from 12 GeV/c^2 to 4 GeV/c^2 .

All events, but the FASTSIM ones, were processed through a detailed simulation of the CMS detector based on GEANT4 [78], and were reconstructed with the same algorithms as those used for data.

4.2 Measurements at the Z boson resonance

4.2.1 The Drell-Yan process

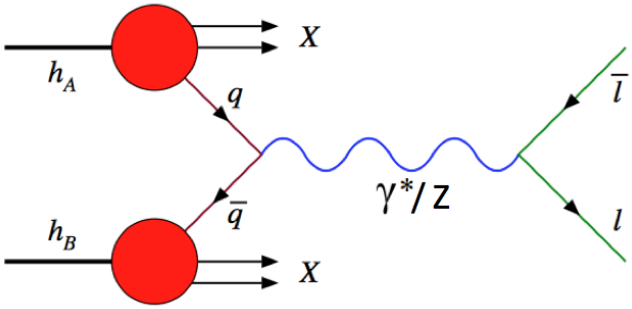


Figure 4.3: *First-order Feynman diagram for Drell-Yan process.*

The Drell-Yan process [68][69] is the paradigm for studying electroweak physics at a hadron collider. It is the lepton-pair production l^+l^- with large

invariant mass $M^2 = (p_{l^+} + p_{l^-})^2 \gg 1$ GeV/c^2 , by quark annihilation (see Fig. 4.3) via a vector boson (γ^* or Z).

The cross section for this process can be calculated using perturbative QCD, using renormalized, scale-dependent, parton distribution functions, extracted from deep inelastic scattering, and summing over quark-antiquark combinations:

$$\sigma(pp \to l^+l^- + X) = \tag{4.1}$$

$$= \sum_q \int \mathrm{d}x_1 \mathrm{d}x_2 f_q(x_1, M^2) f_{\bar{q}}(x_2, M^2) \times [\hat{\sigma}_0 + a\hat{\sigma}_1 + a^2\hat{\sigma}_2 + \dots]_{q\bar{q}\to l^+l^-}$$

with $a = \alpha_S(M^2)/2\pi$. The Drell-Yan process has been calculated up to the NNLO. In MC, the cross sections used are:

$$\sigma(pp(7TeV) \to l^+l^- + X \mid m_{ll} > 50GeV/c^2) = 3048 \text{ pb}$$

$$\sigma(pp(8TeV) \to l^+l^- + X \mid m_{ll} > 50GeV/c^2) = 3504 \text{ pb}$$

The width of the Z is small compared to its mass:

$$\Gamma_Z \sim 2.5 \text{ GeV/c}^2 \qquad m_Z \sim 91 \text{ GeV/c}^2$$

For this reason it is sufficient to consider the Z production as the production of an effective stable particle: we can multiply the production cross section by the two-lepton final-state branching ratio [19][4]. The shape of the 2-lepton mass distribution m_{ll} has a characteristic resonant shape, peaking at m_Z. At a MC-generator level it can be modeled by a relativistic Breit-Wigner:

$$f(m_{ll}, M_Z, \Gamma_Z) = \frac{2}{\pi} \frac{\Gamma_Z^2 M_Z^2}{(M_Z^2 - \Gamma_Z^2)^2 + M_Z^4(\Gamma_Z/M_Z)^2}$$

The reconstructed Z peak suffers from many resolution effects that smear it. A convolution between a Breit-Wigner and a gaussian (to describe the detector effects) describes the shape better. To account for the Drell-Yan γ^* low mass tail, another function needs to be included: an exponential, fading with mass growing, is commonly used. The resulting shape is shown in Fig. 4.4 for the simulated DY sample.

4.2.2 The $t\bar{t}$ process

A small contribution from the $t\bar{t}$ process is present when dealing with our 2-lepton selection (~ 0.5 %). The t quark decays to a b quark emitting a W, that can decay leptonically (see Fig. 4.5). If both Ws from both ts decay leptonically, this leads to a genuine prompt lepton pair in the final state. The $t\bar{t}$ process has been calculated up to the NLO. In MC, the cross sections used are:

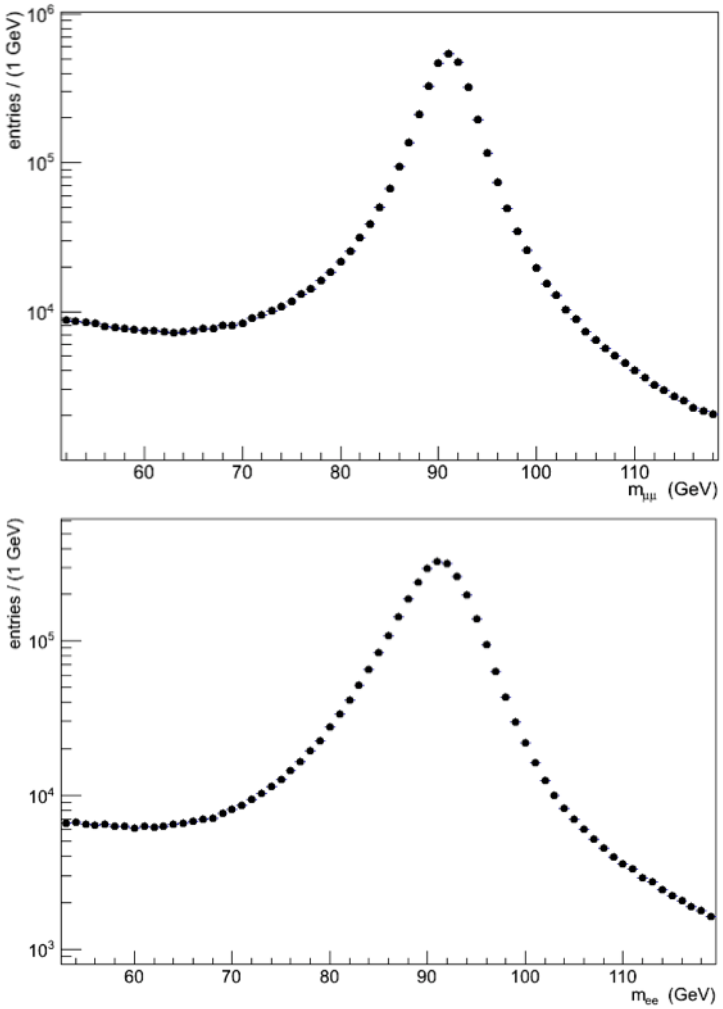


Figure 4.4: *Simulated data: $m_{\mu\mu}$ (top) and m_{ee} (bottom) shape at the reconstruction level, for the DY process.*

$$\sigma(pp(7TeV) \rightarrow t\bar{t} \rightarrow l^+l^-\nu\bar{\nu}b\bar{b}) = 17.32 \text{ pb}$$

$$\sigma(pp(8TeV) \rightarrow t\bar{t} \rightarrow l^+l^-\nu\bar{\nu}b\bar{b}) = 23.64 \text{ pb}$$

In this case the shape of the m_{ll} distribution is non resonant, spread over the acceptance mass range. The resulting shape of the $m_{\mu\mu}$ distribution is shown in Fig. 4.6, for a simulated $t\bar{t}$ sample. The shape is very similar for m_{ee}.

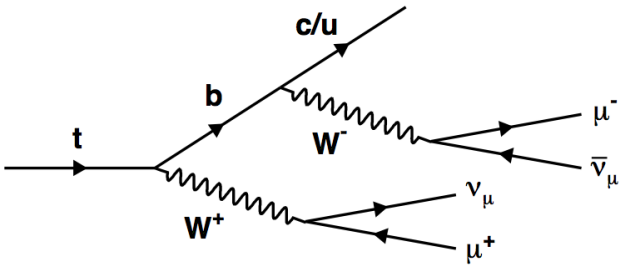


Figure 4.5: *Feynman diagram for top decay into W, b. In this case both the W and the b decay leptonically.*

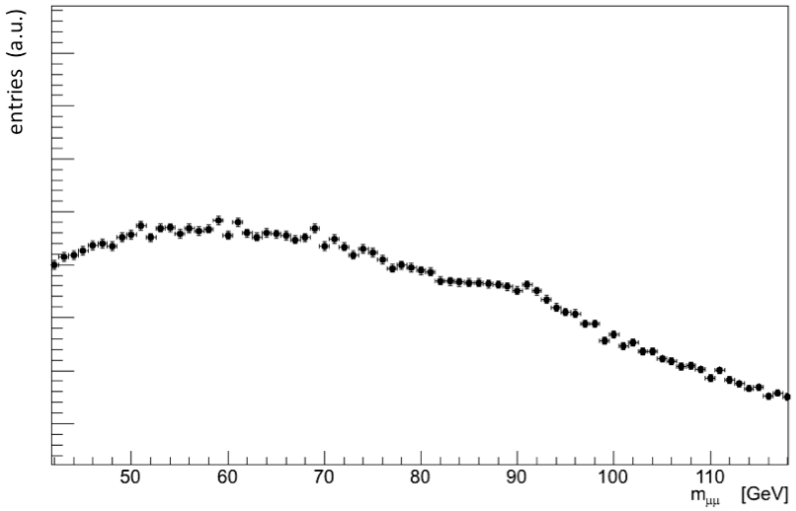


Figure 4.6: *Simulated data: $m_{\mu\mu}$ shape at reconstruction level, for the $t\bar{t}$ process.*

4.2.3 Z event selection and Final State Radiation recovery

Proton-proton collision events are required to pass a selection, to be chosen to build a Z candidate. The steps of the Z selection coincide with the first steps of the selection that will be applied on events when looking for $H \rightarrow 4l$ candidates (see section 5.1). We are basically looking for a pair of opposite-sign, isolated, prompt leptons with an invariant mass as close as possible to the nominal Z boson mass (91.19 GeV/c^2).

A Z decay into a lepton pair can be accompanied by a *Final State Radiation (FSR)* photon, $Z \rightarrow l^+l^-\gamma$. If the photon transverse momentum, ${p_T}^\gamma$, is required to exceed 2 GeV/c, about 8% (15%) of the decays into muons (electrons) is affected. As the photon emission is most often collinear

with one of the leptons, electron measured energies automatically include the energy of a large fraction of the emitted photons in the associated electromagnetic supercluster. On the other hand, muon measured momenta do not include the emitted photons. Final state radiation is therefore expected to degrade the Z mass resolution when measured with the sole muon pairs, and in turn degrade the Higgs boson mass resolution when measured with the four leptons momenta, especially in the 4μ and in the $2e2\mu$ final states and, to a lesser extent, in the 4e final state. It is also expected to reduce the efficiency of the lepton isolation cut when the emitted photon is in the lepton isolation cone. A recovery from FSR is done on Z candidates by associating FSR photons to the leptons.

The event selection consists of the following requirements on events. First a loose *skim* is applied on events, by requiring:

- at least one good primary vertex (PV) fulfilling the following criteria: high number of degrees of freedom ($N_{PV} > 4$), PV close to the nominal CMS interaction point, $z_{PV} < 24$ cm and $r_{PV} < 2$ cm;
- at least two reconstructed lepton candidates, either an electron basic track-supercluster object, with $p_T > 5$ GeV/c and $|\eta| < 2.5$, or a global muon object, or a tracker muon object, with $p_T > 3$ GeV/c and $|\eta| < 2.4$;
- the highest p_T leptons must have $p_{T,1} > 20$, $p_{T,2} > 10$ GeV/c;
- an invariant mass $m(ll) > 40$ GeV/c^2 , for a same-flavor lepton pair.

The events are then required to *have fired the triggers* reported in section 4.1.1:

- When running on data we get:
  - $Z \to \mu\mu$ channel from DoubleMu datasets;
  - $Z \to ee$ channel from DoubleElectron datasets.

Among the leptons reconstructed in the event, further conditions are required, to select prompt leptons, not contained in jets. A first *lepton loose selection* is applied: electrons within the geometrical acceptance of $|\eta| < 2.5$, with $p_T > 7$ GeV/c and having 0 or 1 expected missing inner hits, muons (global or tracker) satisfying $|\eta| < 2.4$, $p_T > 5$ GeV/c. Both electrons and muons should satisfy loose requirements on the transverse ($d_{xy} < 0.5$ cm) and longitudinal ($d_z < 1$ cm) impact parameter with respect to the primary vertex. Non-global tracker muons must be arbitrated, i.e. the segment in the muon system must be assigned to a unique track. In addition, it is required that $\Delta R > 0.02$ between the leptons.

Lepton cross cleaning: loose electrons which are within $\Delta R(\eta, \phi) < 0.05$ of a loose PF or global muon are removed.

Lepton tight selection:

- electrons should pass the electron identification criteria, described in section 3.4, muons should meet the Particle Flow Muon requirements (see section 3.3); non-global tracker muons must be arbitrated;
- cut on the significance of the impact parameter to the event vertex: $|SIP_{3D}| < 4$;
- the cut on isolation (Relative $Iso_{PF} < 0.4$) is postponed after FSR recovery in building Zs candidates (see next steps).

Photon selection (as described in section 3.5):

- PF photons from the particleFlow collection, plus PF photons created from the *ecalEnergy* of the muon PF candidates;
- $p_T > 2$ GeV/c, $|\eta| < 2.4$;
- photon cleaning: remove all PF photons that match with an electron according to ($|\Delta\phi| < 2$, $|\Delta\eta| < 0.05$) OR ($\Delta R < 0.15$) [For this veto we use electrons passing the same p_T , $|\eta|$, ID and SIP cuts used for the signal selection];
- for each photon, consider the closest lepton:
  - if $\Delta R(\gamma, l) < 0.07$, accept the photon if it has $p_T > 2$ GeV/c;
  - otherwise, if $\Delta R(\gamma, l) < 0.5$, accept the photon if it has $p_T > 4$ GeV/c and a PF relative isolation less than 1.0.

Building the Z candidate (nominal mass 91.188 GeV/c^2):

- di-lepton pairs (opposite charge and matching flavor e^+e^-, $\mu^+\mu^-$) are built from leptons passing ID and SIP cuts, but not yet isolation cut;
- for each photon attached that has either a lepton of the Z candidate as closest lepton, we test if the $ll + \gamma$ candidate satisfies the two criteria:
  - $4 < m(ll\gamma) < 100$ GeV/c^2 ;
  - $|m(ll\gamma) - 91.188| < |m(ll) - 91.188|$;

  this is then considered a FSR photon;
- if multiple photons satisfy these criteria, the best is selected, according to this logic:
  - if there's at least one photon with $p_T > 4$ GeV/c, the one with highest p_T is cosen;
  - if all photons have $p_T < 4$ GeV, the one that has the smallest ΔR to its closest lepton is chosen;

- if a FSR photon was selected, we exclude that photon from the photon isolation sum of the lepton of that Z if it was in the isolation cone and outside the isolation veto, to get the new isolation variable Iso'_{PF};
- if both leptons pass the isolation requirement, $Iso'_{PF} < 0.4$, we keep the Z candidate, including a possible associated photon $Z \to ll(\gamma)$;
- if there are no FSR photons associated to the lepton pair, we ask both leptons to pass the isolation requirement, $Iso_{PF} < 0.4$, and, if they do, we keep the Z candidate $Z \to ll$;
- the Z candidate with reconstructed mass $m(ll(\gamma))$ closest to the nominal Z boson mass is the final Z candidate: it must satisfy $51 < m_Z < 119$ GeV/c^2 .

Fig. 4.7 shows the effect of FSR recovery on $m_{\mu\mu}$ distribution, for $Z \to \mu\mu$: events with FSR are moved from the left tail of the mass distribution closer to the center of the resonance peak. Similar results hold for $Z \to ee$ candidates.

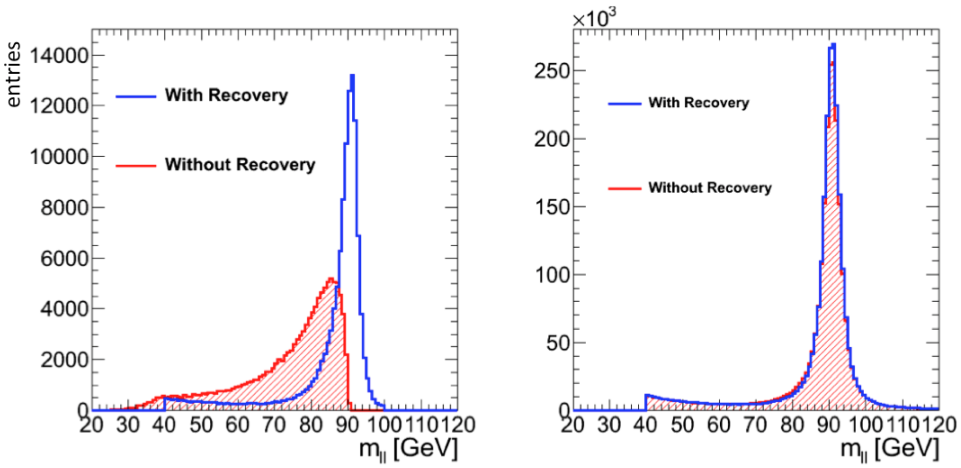


Figure 4.7: *2011 data: $Z \to \mu\mu$ mass spectrum with and without FSR recovery, for FSR affected Zs (left) and all Zs (right).*

4.3 Single Z results and systematic extraction

The selection described is applied to both MC and data (for the 8 TeV data, a sub-sample of 5.26 fb^{-1} is used), giving the m_{ll} distributions shown in Fig. 4.8, 4.9, and the event yields summarized in the following table:

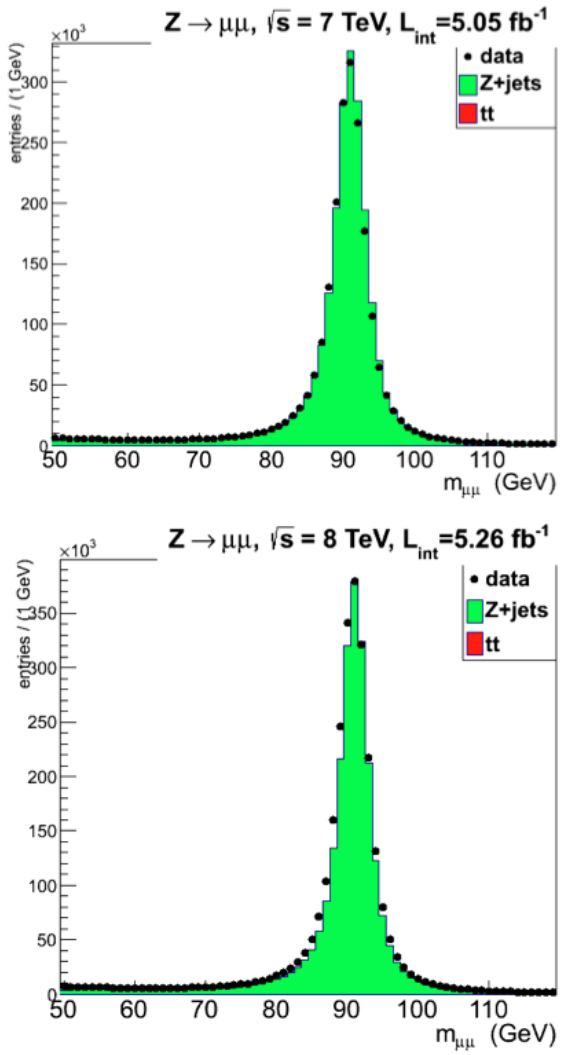


Figure 4.8: *$m_{\mu\mu}$ distributions for 7 TeV (top) and 8 TeV (bottom), data and MC.*

Observable	Data	Z+jets	$t\bar{t}$
$m_{\mu\mu}$ (7 TeV)	$2.18 \cdot 10^6$	$2.22 \cdot 10^6$	$3.3 \cdot 10^3$
m_{ee} (7 TeV)	$1.87 \cdot 10^6$	$1.88 \cdot 10^6$	$2.8 \cdot 10^3$
$m_{\mu\mu}$ (8 TeV)	$2.66 \cdot 10^6$	$2.43 \cdot 10^6$	$5.1 \cdot 10^3$
m_{ee} (8 TeV)	$1.98 \cdot 10^6$	$1.96 \cdot 10^6$	$4.4 \cdot 10^3$

The statistical uncertainties are very small in this case, since there are millions of entries both in data and MC, so the discrepancies between data and MC are due to systematic effects. MC simulation is not perfect: the main sources of systematic discrepancies are due to:

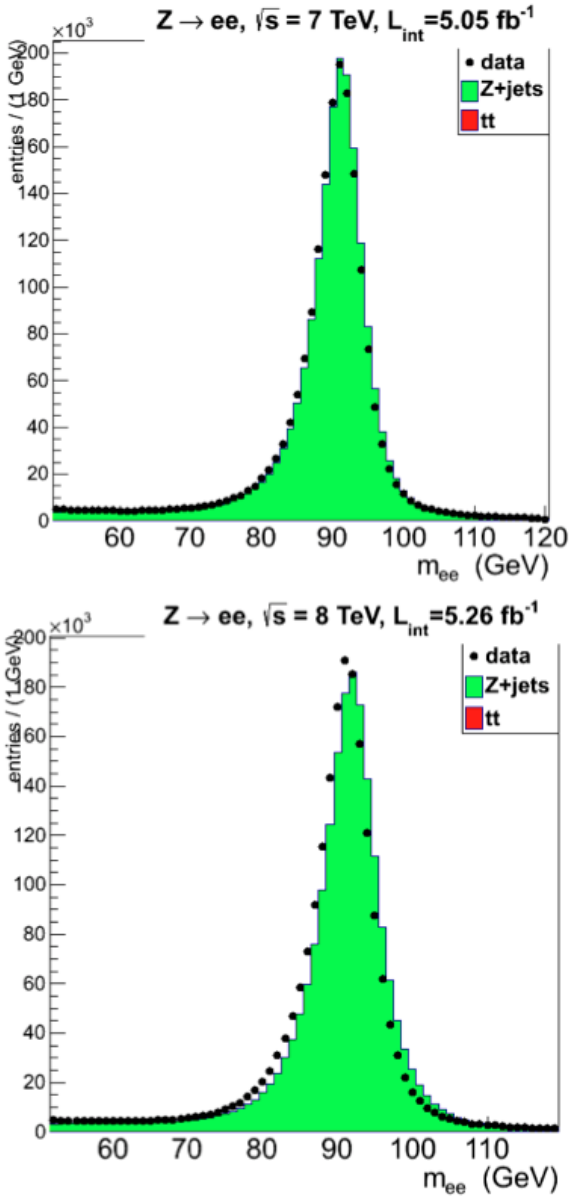


Figure 4.9: *m_{ee} distributions for 7 TeV (top) and 8 TeV (bottom), data and MC.*

- lepton ID, isolation and significance of impact parameter;
- trigger;
- integrated luminosity;
- detector calibration.

By looking at the data discrepancies with respect to the MC, we can estimate the size of these systematics, and extract MC-to-data correction quantities, to be used to correct the pure-MC predictions in the $H \to ZZ \to 4l$ analysis, therefore following a data-driven approach (see section 5.3).

To quantify the data to MC yield offset, we define *single-Z MC-to-data factors* c_Z^i, for each of the four distributions of Fig. 4.8, 4.9 as:

$$c_Z = \frac{N_Z^{data}}{N_Z^{MC}} \tag{4.2}$$

where $N_Z^{data/MC}$ are the number of events found in the distributions of Fig. 4.8, 4.9. The c_Z^i calculated from our data are:

$\mathbf{c}_Z^i$	**7 TeV**	**8 TeV**
Z$\to \mu\mu$	0.981	1.092
Z$\to ee$	0.992	1.007

The data-MC discrepancies in yield for $Z \to \mu\mu$ are of the order of 1% in the 7 TeV data, and of 5% in the 8 TeV data. For $Z \to ee$, the yields from data are as expected within the percent, but for 8 TeV data there is a clear shift to the left, probably due to the ECAL calibration (see Fig. 4.9): this is a detector calibration systematic effect, not included in the c_Z factors, as estimated above.

Chapter 5

The $H \rightarrow ZZ \rightarrow 4l$ analysis

In the $H \rightarrow ZZ \rightarrow 4l$ analysis, we look for events containing four isolated leptons, coming from the primary interaction vertex: the cuts on lepton isolation and significance of impact parameter allow us to minimize the contribution of reducible backgrounds, leaving us with an almost one-to-one signal over background ratio, at the end of the selection. The search is performed in three final state channels: 4μ, $4e$ or $2e2\mu$. The signal and background yields are evaluated separately for the three channels, and then summed together to form the 4-lepton inclusive channel.

The core of $H \rightarrow ZZ \rightarrow 4l$ analysis is described in this chapter. Firstly the tight selection, studied to maximize the signal over background ratio, is presented. Then we estimate how many signal and background events we expect at the end of the selection, using data-driven techniques in support of pure-MC predictions. The shapes of the m_{4l} distributions are studied and fitted with adequate functions. The systematics uncertainties on theoretical and instrumental quantities are also discussed.

5.1 Event selection

For the $H \rightarrow ZZ \rightarrow 4l$ analysis, an event selection was studied to maximize the sensitivity to SM Higgs discovery.

First a loose *skim* is applied on events, by requiring:

- at least one good primary vertex (PV) fulfilling the following criteria: high number of degrees of freedom ($N_{PV} > 4$), PV close to the nominal CMS interaction point, $z_{PV} < 24$ cm and $r_{PV} < 2$ cm;
- at least two reconstructed lepton candidates, either an electron basic track-supercluster object, with $p_T > 5$ GeV/c and $|\eta| < 2.5$, or a global muon object, or a tracker muon object, with $p_T > 3$ GeV/c and $|\eta| < 2.4$;
- the highest p_T leptons must have $p_{T,1} > 20$, $p_{T,2} > 10$ GeV/c;

- an invariant mass $m(ll) > 40$ GeV/c^2 , for a same-flavor lepton pair.

The events are then required to *have fired the triggers* reported in section 4.1.1:

- When running on data we get:
  - 4μ channel from DoubleMu datasets;
  - $4e$ channel from DoubleElectron datasets;
  - $2e2\mu$ channel from DoubleElectron and from DoubleMu, vetoing DoubleElectron triggers for 2011 data, while for 2012 we use in addition MuEG dataset vetoing DoubleElectron and DoubleMuon triggers.

Among the leptons reconstructed in the event, further conditions are required, to select prompt leptons not contained in jets. A first *lepton loose selection* is applied: electrons must lie within the geometrical acceptance of $|\eta| < 2.5$ and have a $p_T > 7$ GeV/c with 0 or 1 expected missing inner hits, muons (global or tracker) must satisfy $|\eta| < 2.4$, $p_T > 5$ GeV/c. Both electrons and muons should satisfy loose requirements on the transverse, $d_{xy} < 0.5$ cm, and longitudinal, $d_z < 1$ cm, impact parameter with respect to the primary vertex. Non-global tracker muons must be arbitrated. In addition, it is required that $\Delta R > 0.02$ between the leptons.

Lepton cross cleaning: loose electrons which are within $\Delta R(\eta, \phi) < 0.05$ of a loose PF or global muon are removed.

Lepton tight selection:

- electrons should pass the electron identification criteria, described in section 3.4, muons should meet the Particle Flow Muon requirements (see section 3.3); non-global tracker muons must be arbitrated;
- cut on the significance of the impact parameter to the event vertex: $|SIP_{3D}| < 4$;
- the cut on isolation (Relative $Iso_{PF} < 0.4$) is postponed after FSR recovery, in building Z candidates (see next steps).

Then follows this sequence of requirements:

- Z_1: a pair of good lepton candidates of opposite charge and matching flavor (e^+e^-, $\mu^+\mu^-$) with FSR recovery (as described in section 4.2.3) and cut on isolation of both leptons, after FSR recovery. The pair with reconstructed mass $m(ll(\gamma))$ closest to the nominal Z boson mass is retained and denoted Z_1. The selected Z_1 must satisfy $40 < m_{Z_1} < 120$ GeV/c^2 ;
- at least 4 good leptons;

- Z_2: another pair of good lepton candidates of opposite charge and matching flavor (e^+e^- or $\mu^+\mu^-$) with FSR recovery (as described in section 4.2.3) and cut on isolation of both leptons, after FSR recovery. If there is more than one combination, the pair formed by leptons with the highest p_T is retained and denoted Z_2. The selected Z_2 must satisfy $4 < m_{Z_2} < 120$ GeV/c^2 . In addition, it is required that $\Delta R > 0.02$ between Z_2 leptons and Z_1 ones;

- it is required that at least two leptons from the four selected have $p_{T,1} > 20$ GeV/c and $p_{T,2} > 10$ GeV/c;

- *QCD suppression*: for all the possible combinations of opposite-sign and same-flavor lepton pair, the mass must satisfy $m(ll) > 4$ GeV/c^2 ;

- *kinematic cuts:* $m_{4l} > 60$ GeV/c^2 and $m_{Z_2} > 12$ GeV/c^2 , to suppress some QCD resonances (like, e.g., the Υ).

An event that passes the full selection looks like the one sketched in Fig. 5.1: here the 4-leptons consist of two electrons and two muons ($Z_1 \to ee$ and $Z_2 \to \mu\mu$), with a mass of $m_{4l} = 125.8$ GeV/c^2 . In appendix A some other

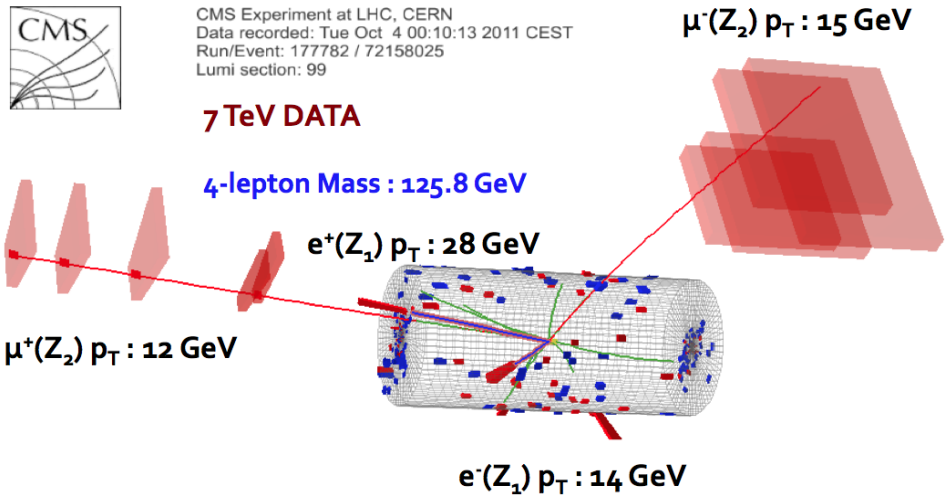


Figure 5.1: *An event that passes the selection (Run 177782 - Event 72158025): $Z_1 \to ee$, $Z_2 \to \mu\mu$, with a 4-lepton mass of* 125.8 *GeV/c^2 .*

events, passing the full selection, are displayed. In appendix B the full list of events passing the selection is reported, along with their main observables.

5.2 Systematic uncertainties

We can divide the systematic uncertainties into two main categories: *instrumental* and *theoretical* ones.

The instrumental uncertainties have already been discussed in the previous sections:

- integrated luminosity: 2.2% (4.4%) for 7 TeV (8 TeV) data, as reported in section 3.1
- trigger efficiency: 1.5%, as reported in section 4.1.1
- ID/isolation/SIP_{3D}: data/MC discrepancies are evaluated for $Z \to ll$ events, separately for $Z \to \mu\mu$ and $Z \to ee$. The quantities that embed these systematic effects are the c-factors, as reported in section 4.3. The data-MC yield discrepancies are of the order of 1% for 7 TeV data, and of 5% for 8 TeV data.

The theoretical cross sections, for both signal and background processes, are calculated with perturbative QCD, using renormalized, scale-dependent, parton distribution functions:

$$\sigma(pp \to X) = \tag{5.1}$$

$$= \sum_{p_1 p_2} \int \mathrm{d}x_1 \mathrm{d}x_2 f_{p_1}(x_1, M^2) f_{p_2}(x_2, M^2) \times [\hat{\sigma}_0 + a\hat{\sigma}_1 + a^2\hat{\sigma}_2 + \dots]_{p_1 p_2 \to X}$$

with $a = \alpha_S(M^2)/2\pi$ and p_1, p_2 denoting the partons, M^2 indicating the scale of the process; $\hat{\sigma}_j$ are the cross sections of the parton-level processes, at the j^{th} perturbative order.

The theoretical uncertainties mainly come from:

- $PDF + \alpha_S$, on the parton distribution functions $f_{p_i}(x_i, M^2)$ and in the coupling $\alpha_S(M^2)$
- QCD scale uncertainty, linked to the choice of QCD renormalization and factorization scales (μ_R and μ_F)

and are different for signal and background cross sections.

These systematic uncertainties, both instrumental and theoretical, are important when relying on MC simulations and on theoretical predictions of cross sections to calculate the expected signal and background yields. Concerning the backgrounds, data-driven techniques can help to eliminate of the systematic effects, as will be discussed in the following sections. The price to pay is, typically, that one gets statistical uncertainties in place of the systematic ones (due to the limited amount of data), but the swap is advantageous in many cases.

5.3 Signal yield

The cross section for the $pp \to H \to ZZ \to 4l$ ($l = \mu, e, \tau$) process at $\sqrt{s} =$ 7 TeV is shown in Fig. 5.2. It ranges from the fb order to about 20 fb, depending on the Higgs mass. At $\sqrt{s} = 8$ TeV, the cross section increases by 30% to 50% for $m_H = 115$ to 600 GeV/c^2 (see Fig. 5.3).

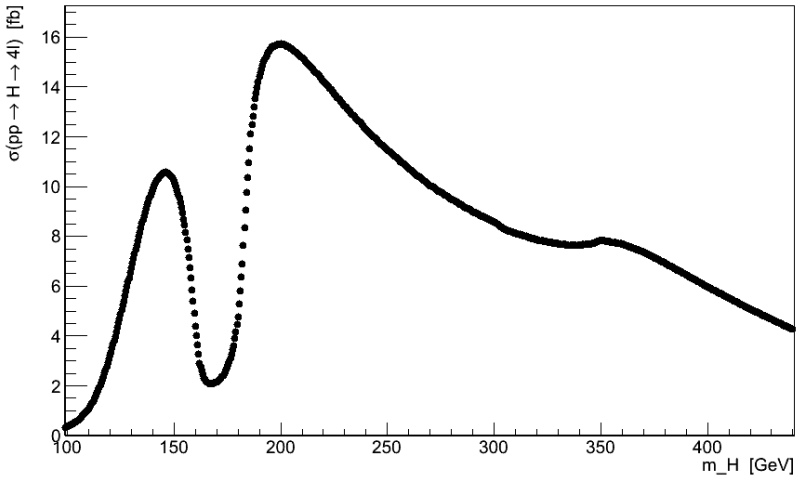


Figure 5.2: *Cross section for the $pp \to H \to ZZ \to 4l$ ($l = \mu, e, \tau$) process at 7 TeV as a function of the Higgs mass.*

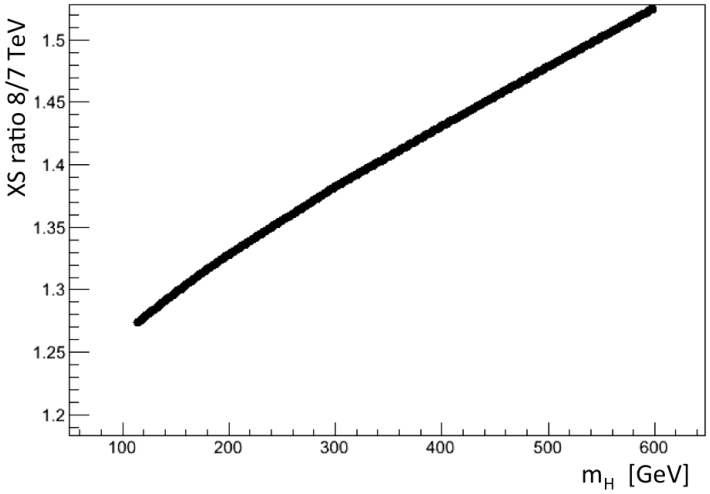


Figure 5.3: *Ratio between the cross sections at 8 TeV and 7 TeV, for the process $pp \to H \to ZZ \to 4l$ ($l = \mu, e, \tau$), as a function of the Higgs mass.*

Systematic errors on the signal total cross section, for each production mechanism, and for all Higgs boson masses are fully described in [71]. They come from $PDF+\alpha_S$ systematic errors (8-10 %) and from theoretical uncertainties, evaluated by varying QCD renormalization and factorization scales, μ_R and μ_F (8-10 %). According to Ref. [72], the $PDF+\alpha_S$ and QCD scale uncertainties are treated as uncorrelated. The uncertainty on $BR(H \to 4l)$ is 2% and m_H-independent. Combining the effects together, the uncertainty on the SM Higgs cross section is a function of the mass, but for $m_H \lesssim 180$ GeV/c^2 , it is almost constant and of the order of 15.5%.

The selection efficiency is evaluated from MC samples as a function of the Higgs mass. In Fig. 5.4 the efficiency times the geometrical acceptance and the branching ratio is shown for the 8 TeV $pp \to H \to 4l$ process, for the three sub-channels: it rises steeply for m_H ranging from 115 to ~ 180 GeV/c^2 , then it continues to rise slowly. Similar results are valid for the 7 TeV processes.

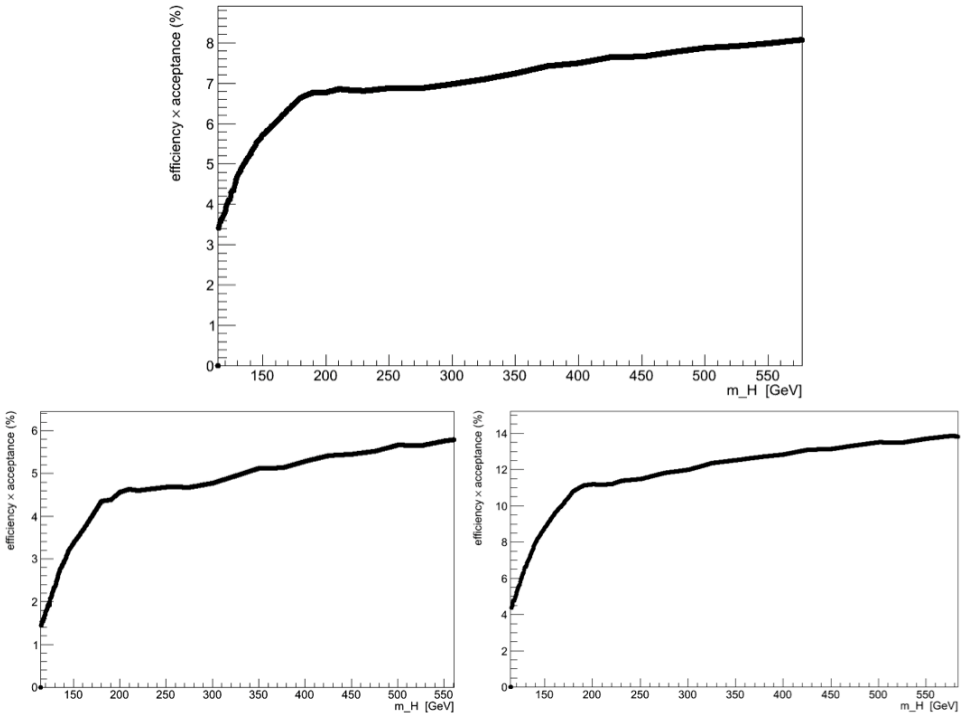


Figure 5.4: *MC: efficiency × acceptance × BR for the three sub-channels:* 4μ *(top),* $4e$ *(bottom-left),* $2e2\mu$ *(bottom-right).*

The efficiencies, both at 7 and 8 TeV, are corrected to account for the data-MC discrepancy, using the c-factors evaluated in section 4.3. The selection efficiency is mainly the efficiency of selecting two Z bosons, so we

can correct it as follows:

$$\epsilon_{4l}^{MC} = \epsilon_{Z_1^i}^{MC}\epsilon_{Z_2^j}^{MC} \longrightarrow \epsilon_{4l}^{data} = \epsilon_{Z_1^i}^{data}\epsilon_{Z_2^j}^{data} = \epsilon_{Z_1^i}^{MC} c_Z^i \; \epsilon_{Z_2^j}^{MC} c_Z^j$$

where the indexes i and j are used to indicate the possible final state of each Z: $Z \to \mu\mu$ or $Z \to ee$. The c-factor must be the one matching the corresponding final state.

The order of magnitude of the systematic uncertainty on efficiencies is the size of data-MC discrepancy, embedded, for a single Z selection, in the c-factors:

$$\delta(\epsilon_{Z_i}) = |1 - c_Z^i|$$

where $\delta(\epsilon)$ indicates the uncertainty in percentage.

$\delta(\epsilon_Z)$	**7 TeV**	**8 TeV**
Z$\to \mu\mu$	1.9%	9.2%
Z$\to ee$	0.8%	0.7%

To compute the systematic uncertainty to be assigned to the signal efficiencies, we simply propagate the errors:

$$\delta(\epsilon_{H\to 2l_i 2l_j}) = \sqrt{\delta^2(\epsilon_{Z\to l_i l_i}) + \delta^2(\epsilon_{Z\to l_j l_j})}$$

The numerical values of the uncertainties are summarized in this table:

$\delta(\epsilon_H)$	4μ **channel**	$4e$ **channel**	$2e2\mu$ **channel**
7 TeV	2.6%	1.1 %	2.1%
8 TeV	13%	1%	9.2%

The signal yields, expected for the available luminosity, are:

$$y_c^j(m_H) = \sigma_c^j(m_H)\; \epsilon_c^j(m_H)\; L_{int}^j$$

with c spanning over the sub-channels 4μ, $4e$ and $2e2\mu$ and j indicating 7 or 8 TeV. The yields are reported in Fig. 5.5 as a function of the Higgs mass. For a low mass Higgs we expect about 10 to 20 signal events, summing all three sub-channels, to pass the full selection.

5.3.1 Mass shape

The signal m_{4l} distribution, for a low mass Higgs, is a narrow resonance peak. Its width is dominated by the instrumental resolution. It can be fitted with a sum of a Student-t distribution [80] plus a Crystal-Ball function [81]: both peaking distributions positioned at the Higgs mass, the latter accounting for the radiative tails on the left side (see Fig. 5.6).

$$f_{H\to 4l}(m_{4l}) = r^H \cdot t(m_{4l}, m_0^H, w^H, \nu^H) + CB(m_{4l}, m_0^H, \sigma^H, \alpha^H, n^H)$$

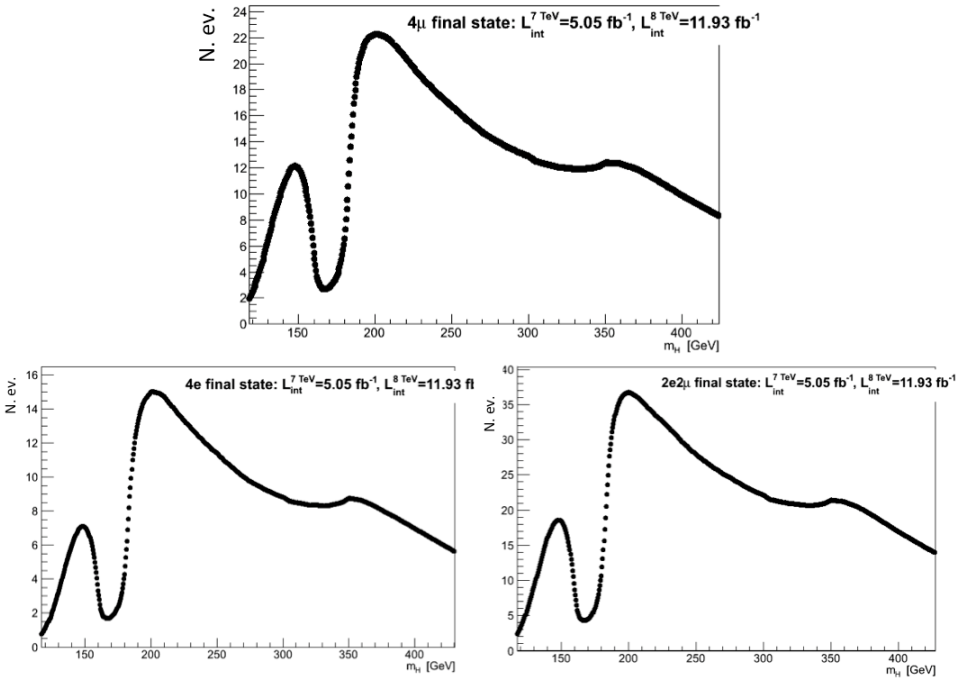


Figure 5.5: *Number of signal events expected (for the available statistics) as a function of the Higgs mass, for the three sub-channels:* 4μ *(top),* $4e$ *(bottom-left),* $2e2\mu$ *(bottom-right).*

In the above listed parameters, w^H and σ^H are the widths of the distributions, ν^H, α^H and n^H are parameters that regulate the relative importance of the tails. The fit well reproduces the signal shape for a low mass Higgs (namely for $m_H \lesssim 300$ GeV/c^2). Concerning the $f_{H\to 4l}$ parameters, some are pre-fixed externally, others are free and fit to the m_{4l} histograms, for each of the available MCs, at different m_H values. The free parameters are: the Student-t width w^H, the Crystal-Ball sigma σ^H and the ratio between the two functions r^H. Once the free parameters are fitted on the MCs, they are plotted as a function of the Higgs mass as shown in Fig. 5.7, and the trend is fitted, with polynomials. Not all parameters can be let free because they are correlated, and this would result in an irregular trend of the fitted values as functions of m_H. The following table summarizes the parameters used to model the signal m_{4l} shape:

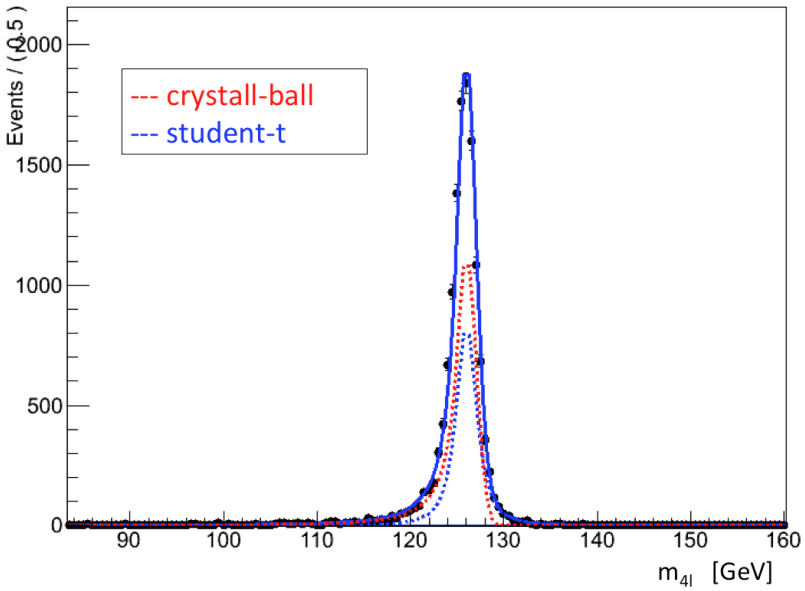


Figure 5.6: *Simulated data (4μ channel): signal fit ($m_H = 126$ GeV/c^2) with a Student-t (dashed blue) plus a Crystal-Ball function (dashed red).*

Signal parameters	4μ **channel**	$4e$ **channel**	$2e2\mu$ **channel**
m_0^H	m_H	m_H	m_H
w^H	$\text{pol2}^w_{4\mu}(m_H)$	$\text{pol2}^w_{4e}(m_H)$	$\text{pol2}^w_{2e2\mu}(m_H)$
ν^H	1.8	1.8	1.8
σ^H	$\text{pol2}^\sigma_{4\mu}(m_H)$	$\text{pol2}^\sigma_{4e}(m_H)$	$\text{pol2}^\sigma_{2e2\mu}(m_H)$
α^H	0.9	0.9	0.9
n^H	2.5	2.5	2.5
r^H	$\text{pol3}^r_{4\mu}(m_H)$	$\text{pol3}^r_{4e}(m_H)$	$\text{pol3}^r_{2e2\mu}(m_H)$

In Fig. 5.7, 5.6 the case of 4μ channel is shown, but very similar results hold for the $4e$ and $2e2\mu$ channels. This particular modeling of the signal mass shape is valid up to $m_H \sim 220$ GeV/c^2 .

5.4 Reducible background

The two relevant processes that can be referred to as "*reducible* background" are:

- the Drell-Yan (DY) process: $pp \rightarrow Z + jets$
- the double top production with leptons in the final state ($t\bar{t}$):
  $pp \rightarrow t\bar{t} \rightarrow WW(\rightarrow 2l2\nu)b\bar{b}$

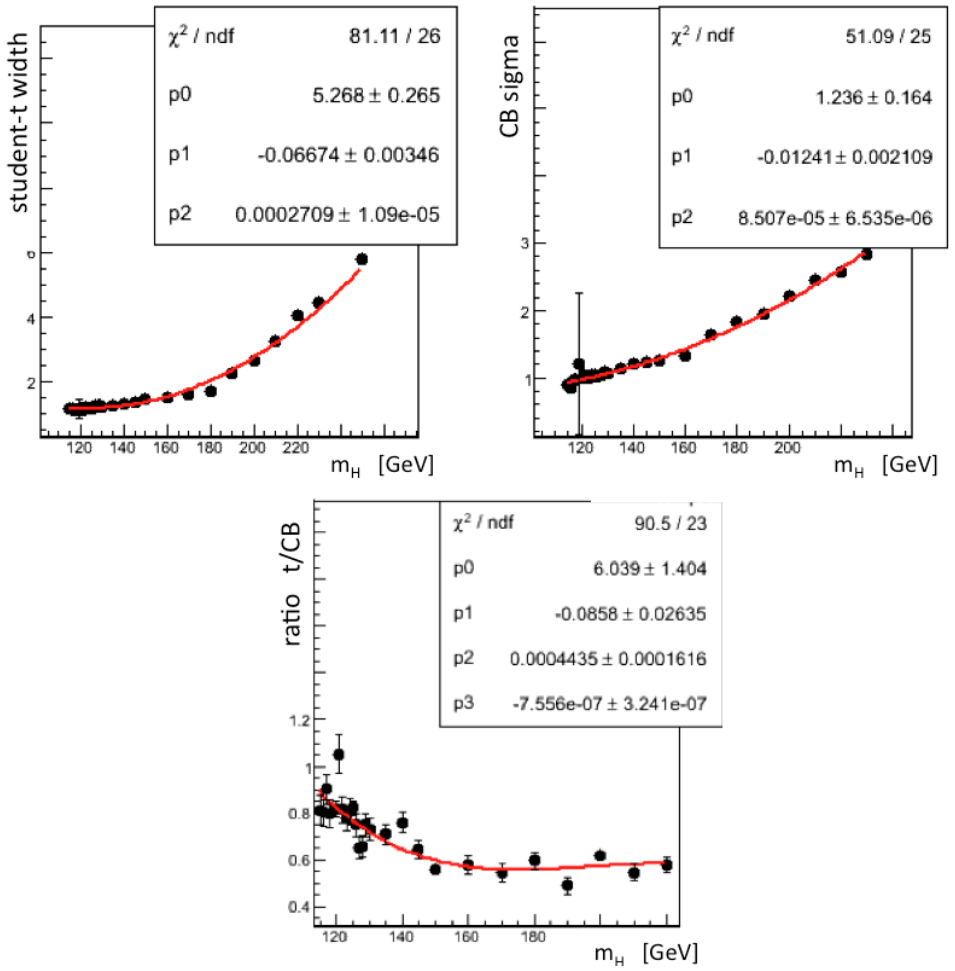


Figure 5.7: *Polynomial fit to the free parameters of signal-pdfs, as a function of the Higgs mass (4μ channel).*

The MCs, used to simulate these processes, and the relative cross sections, at 7 and 8 TeV, are listed in section 4.1.2.

These processes contain two genuine prompt leptons, coming from the $Z(\gamma^*)$ and WW decay. The other two leptons can come from a heavy quark decay, being in this case contained in a jet, or can be two fake leptons, e.g., pions misidentified as electrons. The yields of these processes are drastically reduced by the tight lepton ID requirements, and by applying cuts on the lepton isolation and on the significance of the impact parameter. However, a small contamination is still present after the event selection, especially at low masses.

A data-driven approach is used to estimate the reducible background yields:

- A *control region (CR)* is defined, in the selection criteria phase space. The selection remains the same as that described in section 5.1, but for the Z_2 lepton pair: at least one of its two leptons is required to have the Iso_{PF} or SIP_{3D} above the cut threshold. The *signal region (SR)* is the same as defined in 5.1.

- The *ratio* between the MC number of events in the SR (N_{SR}^{MC}) and in the CR (N_{CR}^{MC}) is taken, for both DY and $t\bar{t}$ processes.

- The *shapes of* m_{4l} *and* m_{Z_1} *distributions* are taken from MCs, for both DY and $t\bar{t}$ processes, by looking at the events in the CR.

- From data, the number of DY and $t\bar{t}$ events in the CR (N_{CR}^{data}) is evaluated fitting the m_{Z_1} distribution (which discriminates between the two processes), independently for the three sub-channels 4μ, $4e$ and $2e2\mu$.

- The number of events in the CR, estimated by the fit, is extrapolated to the SR, using the factors from MC:

$$N_{SR}^{j,c} = \left(\frac{N_{SR}^{j,c}}{N_{CR}^{j,c}}\right)^{MC} (N_{CR}^{j,c})^{data}$$

  where the index j indicates the subprocess (DY or $t\bar{t}$) and the index c the sub-channel (4μ, $4e$ and $2e2\mu$).

The following table contains the numbers found in MC and data, and the corresponding data-driven estimate of the reducible background yields, while the intermediate steps are sketched in Fig. 5.8-5.10:

Process	$\mathbf{N}_{SR}^{MC}$	$\mathbf{N}_{CR}^{MC}$	**Fitted** $\mathbf{N}_{CR}^{data}$	$\mathbf{N}_{SR}$ **(error %)**
DY (4μ)	3	166	181.9 ± 16.0	3.3 ± 1.9 (58.9%)
DY ($4e$)	1	18	36.3 ± 7.4	2.0 ± 2.1 (104.7%)
DY ($2e2\mu$)	2	187	157.3 ± 15.6	1.7 ± 1.2 (71.8%)
$t\bar{t}$ (4μ)	3	3027	145.4 ± 15.1	0.1 ± 0.1(58.7%)
$t\bar{t}$ ($4e$)	4	433	24.3 ± 6.6	0.2 ± 0.1(57.1%)
$t\bar{t}$ ($2e2\mu$)	14	3077	147.1 ± 15.5	0.7 ± 0.2(28.8%)

The uncertainty on the data-driven estimate is mainly statistical, due to the low statistics of MCs, that leads to a very limited amount of data in SR. Since the final expected yields are very low and distributed broadly in m_{4l} (with respect to the signal), the large uncertainties do not degrade the sensitivity of the analysis.

In Fig. 5.8, the fit to the m_{Z_1} distributions is shown, for data in CR. It allows to separate the number of DY and $t\bar{t}$ events in the CR, exploiting the fact that the m_{Z_1} distribution is very different in shape: a relativistic Breit-Wigner is used to model the peaked m_{Z_1} distribution of the DY process,

while a four-degree polynomial is used for the $t\bar{t}$, which has an almost flat shape.

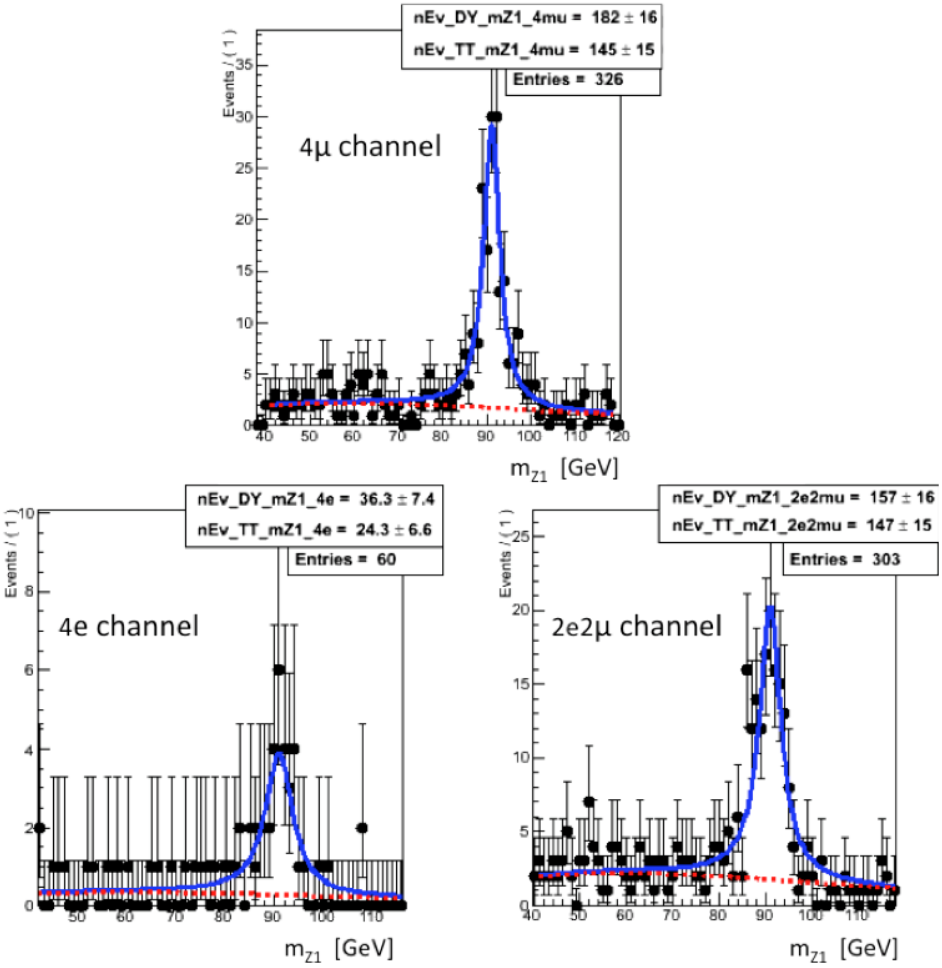


Figure 5.8: *Fit to m_{Z_1} distributions, for data in CR, in the three sub-channels. The number of DY and $t\bar{t}$ events are extracted exploiting the different shapes of DY and $t\bar{t}$ (dashed, in red) distributions.*

In Fig. 5.9 the shapes of the m_{4l} distributions are shown, taken from MC events in CR (only the 4μ channel is shown). A Landau function is used to model the DY distribution, while a convolution between a Landau and a Gauss function (centered at zero) is used for $t\bar{t}$. No systematic uncertainties are assumed on the shape of the m_{4l} distributions: they are considered negligible, given the low number of expected events after the selection with the available amount of data.

The distribution parameters are reported in the following tables:

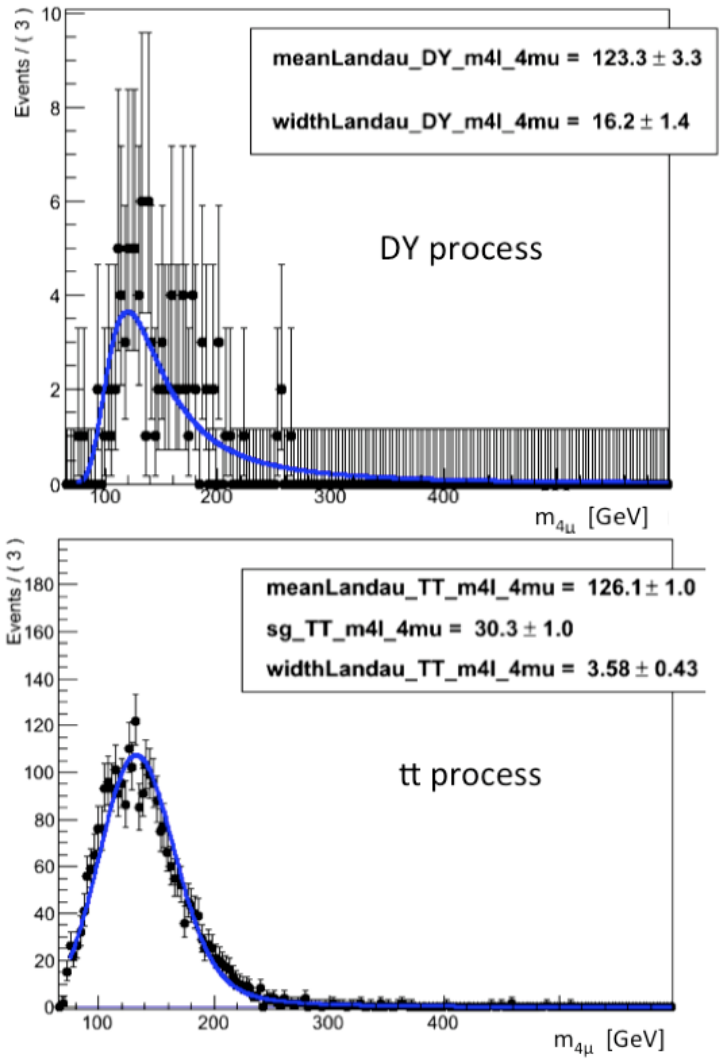


Figure 5.9: *Simulated data (4μ channel): fit to $m_{4\mu}$ distributions for the DY and $t\bar{t}$ processes.*

DY parameters	4μ **channel**	$4e$ **channel**	$2e2\mu$ **channel**
Landau MPV (GeV)	123.3	136.0	123.1
Landau width (GeV)	16.2	17.9	11.8

$t\bar{t}$ **parameters**	4μ **channel**	$4e$ **channel**	$2e2\mu$ **channel**
Landau MPV (GeV)	126.1	131.8	129.1
Landau width (GeV)	3.6	9.6	4.16
Gaussian σ (GeV)	30.3	25.0	28.5

As a summary, in Fig. 5.10, the m_{Z_1} and m_{4l} distributions, for data in CR, are shown: the three sub-channels are summed together and the separation between DY and $t\bar{t}$ components is highlighted.

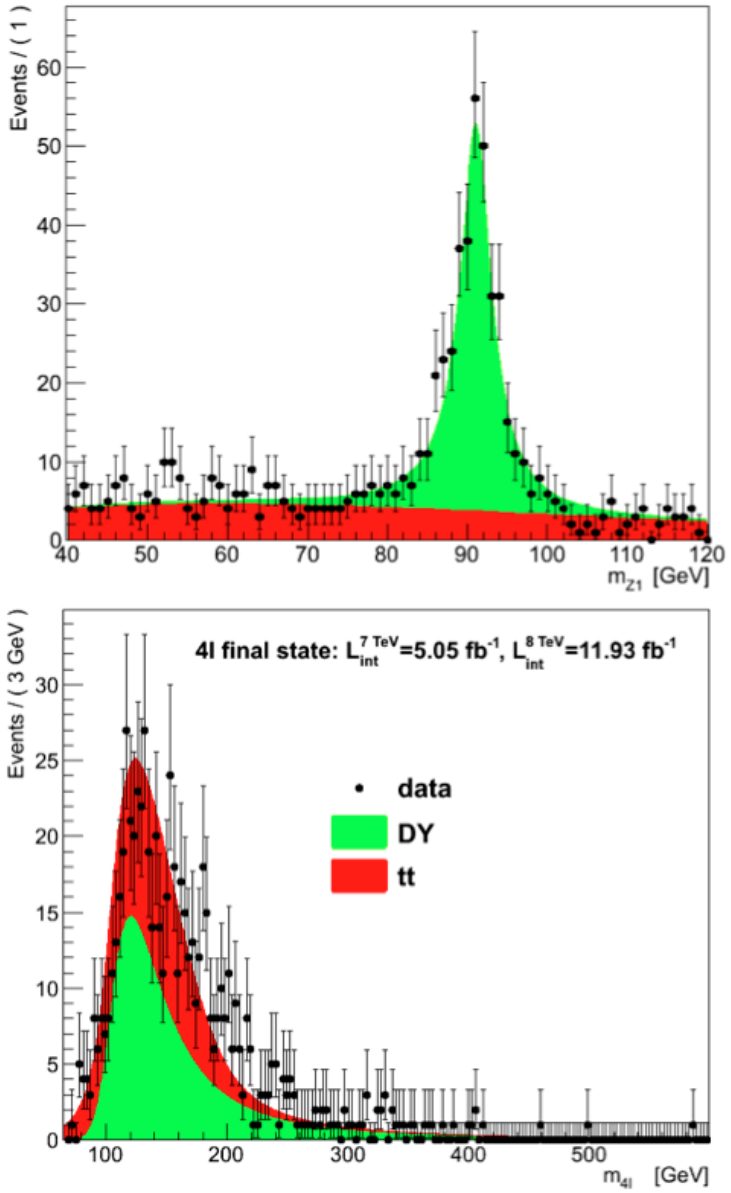


Figure 5.10: *m_{Z_1} and m_{4l} inclusive distributions, for data in the CR. The yields are fitted on m_{Z_1} (top) and the mass distribution are normalized according to the fit (bottom).*

5.5 Irreducible background

The *irreducible* double Z, non-resonant, production, with the Zs decaying to leptons:

$$pp \to ZZ(Z\gamma^*) \to l_1^+ l_1^- l_2^+ l_2^- \qquad l_i = e, \mu, \tau$$

can occur via quark or gluon fusion (see Fig. 5.11). We can single out three

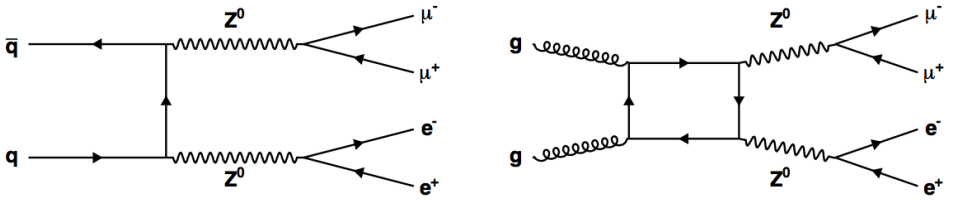


Figure 5.11: *Feynman diagrams for irreducible double Z, non-resonant, production, via quark fusion (left) and gluon fusion (right), with Zs decaying to leptons.*

relevant sub-processes :

- ZZ production via quark fusion: $q\bar{q} \to ZZ \to 4l$ with $l = \mu, e$
- ZZ production via gluon fusion: $gg \to ZZ \to 4l$ with $l = \mu, e$
- ZZ production via quark fusion, with a Z decaying to a τ pair: $q\bar{q} \to ZZ \to 2l2\tau$ with $l = \mu, e$

The MCs, used to simulate these processes, and the relative cross sections (calculated up to NNLO), at 7 and 8 TeV, are listed in section 4.1.2.

The uncertainties on the cross sections of these processes are:

- $PDF + \alpha_S$ uncertainties for $q\bar{q} \to ZZ \to 4l$ at NLO and $gg \to ZZ \to 4l$, evaluated using MCFM [73][64]: 6% for $q\bar{q} \to ZZ \to 4l$ and 12% for $gg \to ZZ \to 4l$.
- QCD scale systematic errors: the variations in the differential cross section $\mathrm{d}\sigma/\mathrm{d}m_{4l}$ are estimated as the renormalization and factorization scales are modified by a factor of two up and down from their default setting $\mu_R = \mu_F = m_Z$ [64]. The resulting values (averaged over m_{4l}) are 5% for $q\bar{q} \to ZZ \to 4l$ and 30% for $gg \to ZZ \to 4l$.

The selection efficiencies are obtained by applying the event selection on the ZZ MCs. The pure-MC estimate of the yield of each process is:

$$y_c^j(process) = \sigma_c^j(process)\ \epsilon_c^j(process)\ L_{int}^j$$

with c spanning over the sub-channels 4μ, $4e$ and $2e2\mu$ and j indicating 7 or 8 TeV. The expected yields for $L_{int} = 5.05$ fb^{-1} at 7 TeV and for $L_{int} = 11.93$ fb^{-1} at 8 TeV are:

7 TeV	4μ **channel**	$4e$ **channel**	$2e2\mu$ **channel**
$q\bar{q} \to ZZ \to 4l$	26.73	16.16	37.51
$gg \to ZZ \to 4l$	1.31	0.93	2.24
$q\bar{q} \to ZZ \to 2l2\tau$	0.4	0.3	0.7
8 TeV			
$q\bar{q} \to ZZ \to 4l$	70.8	39.25	94.2
$gg \to ZZ \to 4l$	3.99	2.84	6.79
$q\bar{q} \to ZZ \to 2l2\tau$	1.07	0.73	1.94

5.5.1 Mass shape

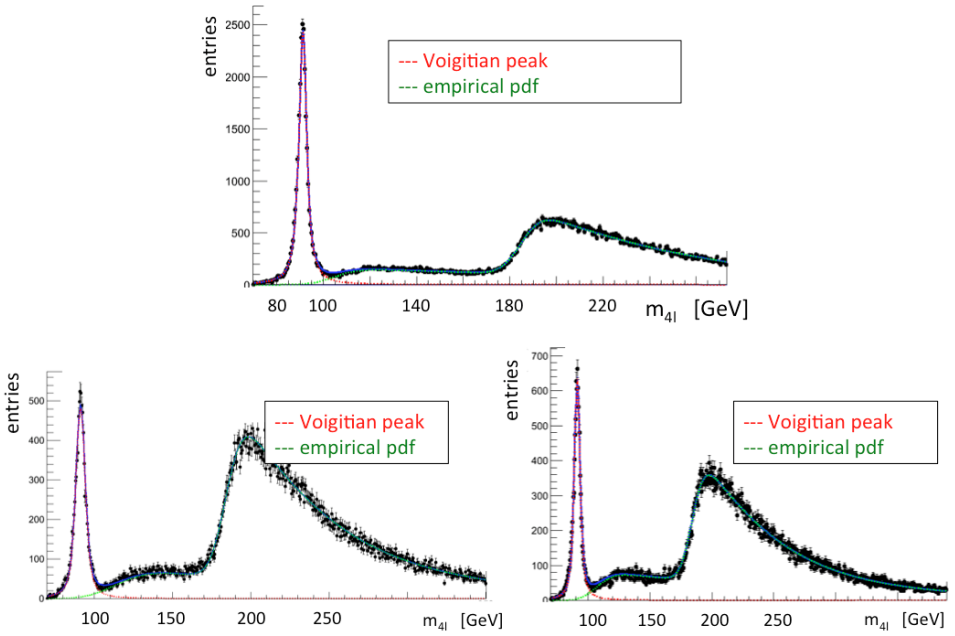


Figure 5.12: *Simulated data: m_{4l} distributions for the $q\bar{q} \to ZZ \to 4l$ process. 4μ (top), $4e$ (bottom-left), $2e2\mu$ (bottom-right). Fit: empirical pdf (green, dashed) + Voigtian function (red, dashed).*

The shapes of the m_{4l} distributions are taken from MCs, after having applied the full selection. No systematic uncertainties are assumed on the shape of the m_{4l} distributions: they are considered negligible for the analysis with respect to other sources of uncertainty, given the low number of events expected to pass the selection, with the available statistics.

The m_{4l} spectrum of the $q\bar{q} \to ZZ \to 4l$ background has the shape shown in Fig. 5.12, after the event selection is applied. Where the double on-shell Z production is accessible, for $m_{4l} \gtrsim 182$ GeV/c^2 , the rate of the process is higher. Below ~ 182 GeV/c^2 , one Z is off-shell (it can be indicated both

with Z or γ^*) and the process rate is lower. The process becomes resonant at 91.19 GeV/c^2 since the Feynman diagram in Fig. 5.13, that is one of the contribution to the ZZ(Zγ^*) process, gives rise to the resonance, called Z decay to four leptons, $Z \to 4l$. The Z peak is more evident in sub-channels containing muons, since the efficiency at low momentum is much higher for muons than for electrons. The m_{4l} spectrum is fitted with an "empirical"

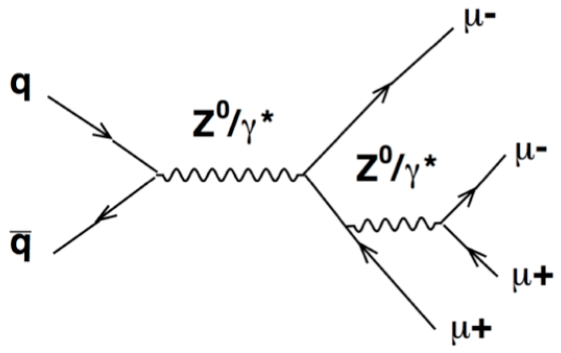


Figure 5.13: *Feynman diagram for the $Z \to 4\mu$ process.*

pdf, plus a Voigtian [82] function, for the Z peak:

$$f_{q\bar{q}\to ZZ}(m_{4l}, \vec{a}, \vec{b}, \vec{c}) = f_1(m_{4l}, \vec{a}) + f_2(m_{4l}, \vec{b}) + f_3(m_{4l}, \vec{c}) + r \cdot voig(m_{4l}, m_Z, \gamma, \sigma)$$

with

$$f_1(m_{4l}, \vec{a}) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{m_{4l} - a_1}{a_2}\right)\right)\frac{a_4}{1 + e^{(m_{4l}-a_1)/a_3}} \tag{5.2}$$

$$f_2(m_{4l}, \vec{b}) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{m_{4l} - b_1}{b_2}\right)\right)\left(\frac{b_4}{1 + e^{(m_{4l}-b_1)/b_3}} + \frac{b_6}{1 + e^{(m_{4l}-b_1)/b_5}}\right)$$

$$f_3(m_{4l}, \vec{c}) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{m_{4l} - c_1}{c_2}\right)\right)\frac{c_4}{1 + e^{(m_{4l}-c_1)/c_3}}$$

The values of the pdf parameters are reported in the following table, for the three sub-channels:

$f_{q\bar{q}\to ZZ}$ parameters	4μ channel	$4e$ channel	$2e2\mu$ channel
a_1 (GeV)	105.2	115.2	108.0
a_2 (GeV)	11.27	21.67	14.76
a_3 (GeV)	117.99	126.9	123.4
a_4	0.0456	0.035	0.038
b_1 (GeV)	185.4	185.2	185.0
b_2 (GeV)	9.57	11.7	11.0
b_3 (GeV)	3709	35.22	36.57
b_4	0.1038	0.133	0.109
b_5 (GeV)	57.88	63.13	60.56
b_6	0.0801	0.077	0.067
c_1 (GeV)	48.2	21	48
c_2 (GeV)	-5.5	-1.9687	-5.8
c_3 (GeV)	1344	88.84	930.9
c_4	0.085	0.029	0.472
r	0.22	0.11	0.13
m_Z (GeV)	91.19	91.19	91.19
γ (GeV)	3.37	3.74	3.50
σ (GeV)	0.68	2.17	1.52

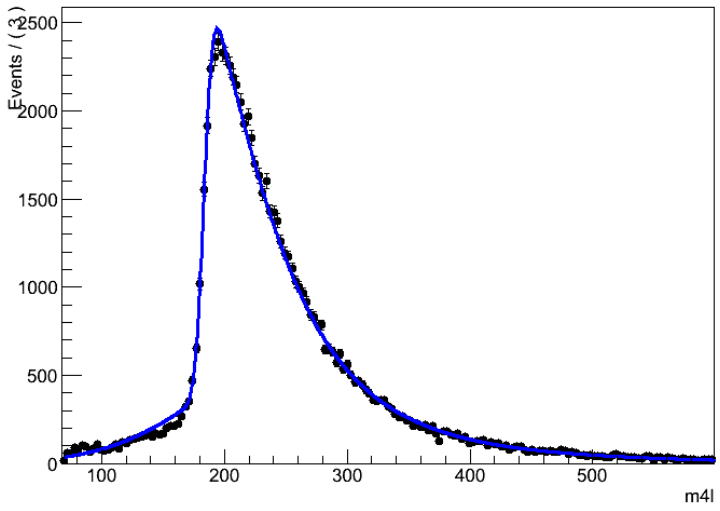


Figure 5.14: *Simulated data (4μ channel): $m_{4\mu}$ distribution for the $gg \to ZZ \to 4\mu$ process. Fit: empirical pdf.*

The m_{4l} spectrum of the $gg \to ZZ \to 4l$ background has the shape shown in Fig. 5.14 (the 4μ channel is shown, the shape is very similar in the $4e$ and $2e2\mu$ channels), after the event selection is applied. The m_{4l}

spectrum is fitted with an "empirical" pdf:

$$f_{gg\to ZZ}(m_{4l}, \vec{a}, \vec{b}) = f_1(m_{4l}, \vec{a}) + f_2(m_{4l}, \vec{b})$$

with the f_i functions of the same form as formulas 5.2. The values of the pdf parameters are reported in the following table, for the three sub-channels:

$f_{gg\to ZZ}$ **parameters**	4μ **channel**	$4e$ **channel**	$2e2\mu$ **channel**
a_1 (GeV)	223.8	177.9	155.2
a_2 (GeV)	108.6	70.38	46.3
a_3 (GeV)	213.7	113.1	111.2
a_4	0.106	0.041	0.0475
b_1 (GeV)	184	183.5	184.1
b_2 (GeV)	8.03	11.67	12.42
b_3 (GeV)	47.79	35.24	37.75
b_4	0.323	0.54	0.538
b_5 (GeV)	231.2	28.94	34.59
b_6	-0.104	-0.322	-0.324

For the $ZZ \to 2l2\tau$ ($l = \mu, e$) background, the m_{4l} spectrum is different from the previous ones: τ leptons decaying leptonically $\tau \to l\nu_l\nu_\tau$ are reconstructed as muons or electrons and part of their energy is lost due to the presence of neutrinos in the final state. The resulting spectrum has the shape shown in Fig. 5.15 (the 4μ channel is shown, the shape is very similar in the $4e$ and $2e2\mu$ channels), after the event selection has been applied: it is fitted with a Landau function convoluted with a gaussian:

$$f_{ZZ\to 2l2\tau}(m_{4l}, m_0, w, \sigma) = \text{gaus}(m_{4l}, 0, \sigma) \otimes \text{land}(m_{4l}, m_0, w)$$

The values of the pdf parameters are reported in the following table, for the three sub-channels:

$f_{ZZ\to 2l2\tau}$ **parameters**	4μ **channel**	$4e$ **channel**	$2e2\mu$ **channel**
Landau MPV m_0 (GeV)	146.6	149.7	149.4
Landau width w (GeV)	12.3	16.5	14.51
Gaussian σ (GeV)	16.4	11.7	14.3

5.5.2 Data-driven estimate of the ZZ yield

A data-driven (dd) estimate of the irreducible background yield can be done using as "normalization bands" (nb) the regions of the m_{4l} spectrum where the Higgs has been excluded (see section 1.2.4). Taking the pdf shapes of m_{4l} from MC, we rescale the pure-MC yields for irreducible ZZ backgrounds

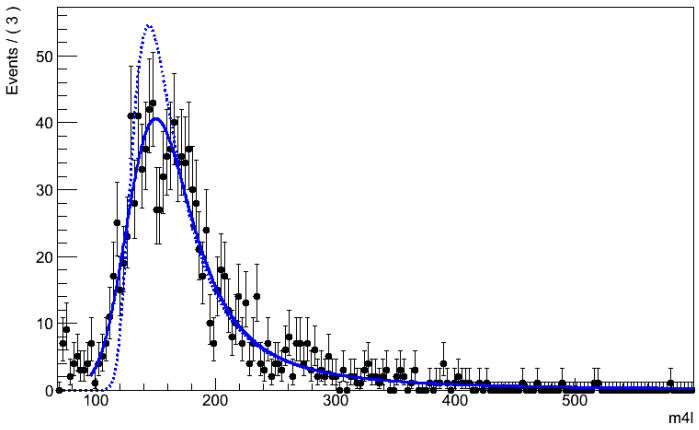


Figure 5.15: *Simulated data (4μ channel): m_{4l} distribution for the $q\bar{q} \to ZZ \to 2\mu 2\tau$ process. Fit: Landau function (dashed) convoluted with a Gaussian.*

N_i^{MC} (see the table in section 5.5), by looking at the data yield in the normalization bands. We define the normalization band as:

$$nb = \{m_{4l} \in [180, 600] \text{ GeV/c}^2 \}$$

The yield of the reducible background is almost null in this region. The number of events expected in the nb, from pure-MC prediction, is:

$$N_{nb}^{MC} = \sum_i N_i^{MC} \frac{\int_{nb} f_i(m_{4l}) \mathrm{d}m_{4l}}{\int f_i(m_{4l}) \mathrm{d}m_{4l}}$$

where the i-index sum is over the background sub-contributions ($qq \to ZZ \to 4l$, $gg \to ZZ \to 4l$ and the very small $ZZ \to 2l2\tau$ and reducible background) and f_i are their m_{4l} pdfs. In the three sub-channels these are the events found in the normalization band, compared to the MC expectations:

	4μ	$4e$	$2e2\mu$	**TOT**
N_{nb}^{MC}	70.36	48.91	109.18	228.45
N_{nb}^{data}	71	46	129	246

In Fig. 5.16 the corresponding m_{4l} distribution is shown. The formula to rescale the expected background yields, using data, is:

$$N_j^{dd} = \frac{N_j^{MC}}{N_{nb}^{MC}} N_{nb}^{data}$$

where the j index indicates the three irreducible backgrounds: $qq \to ZZ \to 4l$, $gg \to ZZ \to 4l$ and $ZZ \to 2l2\tau$. The data-rescaled predictions for the irreducible background, for $L_{int} = 5.05$ fb^{-1} at 7 TeV and $L_{int} = 11.93$ fb^{-1} at 8 TeV are:

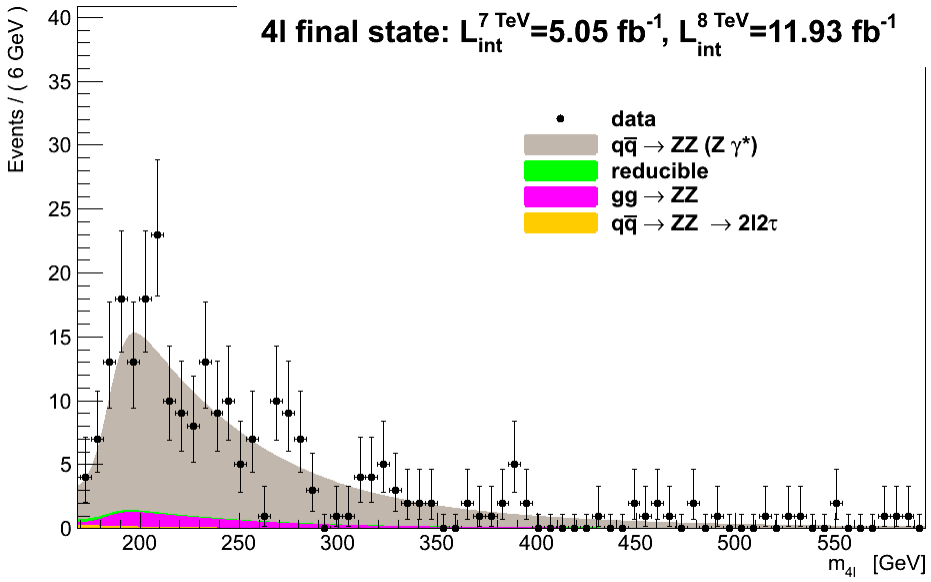


Figure 5.16: *m_{4l} distribution for data in the normalization band ($m_{4l} >$ 180 GeV/c^2), superimposed to the expected backgrounds.*

	4μ	$4e$	$2e2\mu$	**TOT**
$N^{dd}_{qq\to ZZ}$	98.4 ± 11.7	52.1 ± 7.7	155.6 ± 13.7	306.2 ± 19.6
$N^{dd}_{gg\to ZZ}$	5.3 ± 0.6	3.6 ± 0.5	10.7 ± 0.9	19.6 ± 1.2
$N^{dd}_{ZZ\to 2l2\tau}$	1.5 ± 0.2	1.0 ± 0.1	2.1 ± 0.2	4.6 ± 0.3

The uncertainty on the final yields, in each sub-channel, is purely statistical, and is of the order of 10%: it is competitive with the systematic uncertainty that would be assigned to the yields if the pure-MC estimate was used.

5.6 Final distributions

The m_{4l} distribution for data is shown if Fig. 5.17, superimposed to the expected backgrounds.

For $m_H > 180$ GeV/c^2 the background is normalized to data, by the dd estimate of ZZ, as described in section 5.5.2. The "interesting" mass range, below 180 GeV/c^2 , is zoomed in Fig. 5.18, and in Fig. 5.19 for the three sub-channels separately.

The number of events expected from background and observed in data (in the range $m_{4l} \in [110, 180]$) are reported in the following table:

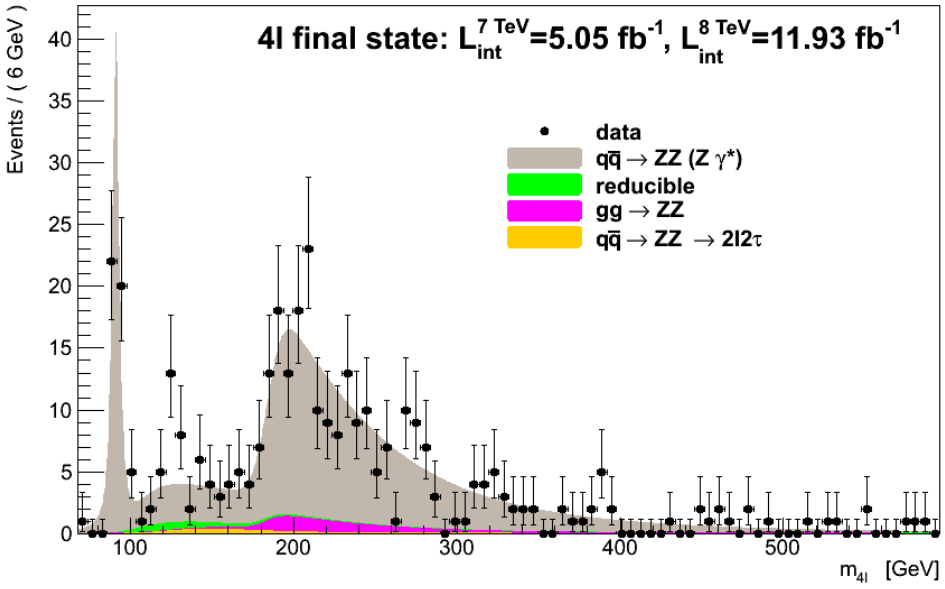


Figure 5.17: *m_{4l} distribution for data, superimposed to the expected backgrounds.*

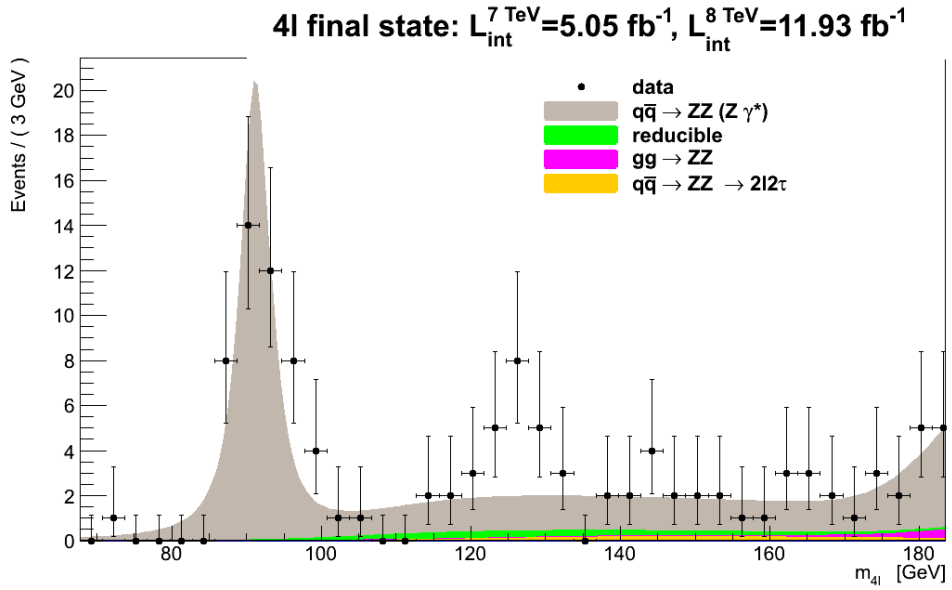


Figure 5.18: *m_{4l} distribution for data, in the range $m_{4l} \in [70, 180]$ GeV, superimposed to the expected backgrounds.*

$m_{4l} \in [110, 180]$ GeV	4μ	$4e$	$2e2\mu$	**TOT**
$N^{dd}_{qq\to ZZ}$	11.2	5.5	19.3	36.0
$N^{dd}_{gg\to ZZ}$	0.4	0.2	0.9	1.6
$N^{dd}_{ZZ\to 2l2\tau}$	0.9	0.5	1.8	3.3
N^{dd}_{red}	2.0	1.4	1.6	5.0
TOT bkg	14.5	7.7	23.7	45.8
Data	20	14	26	60

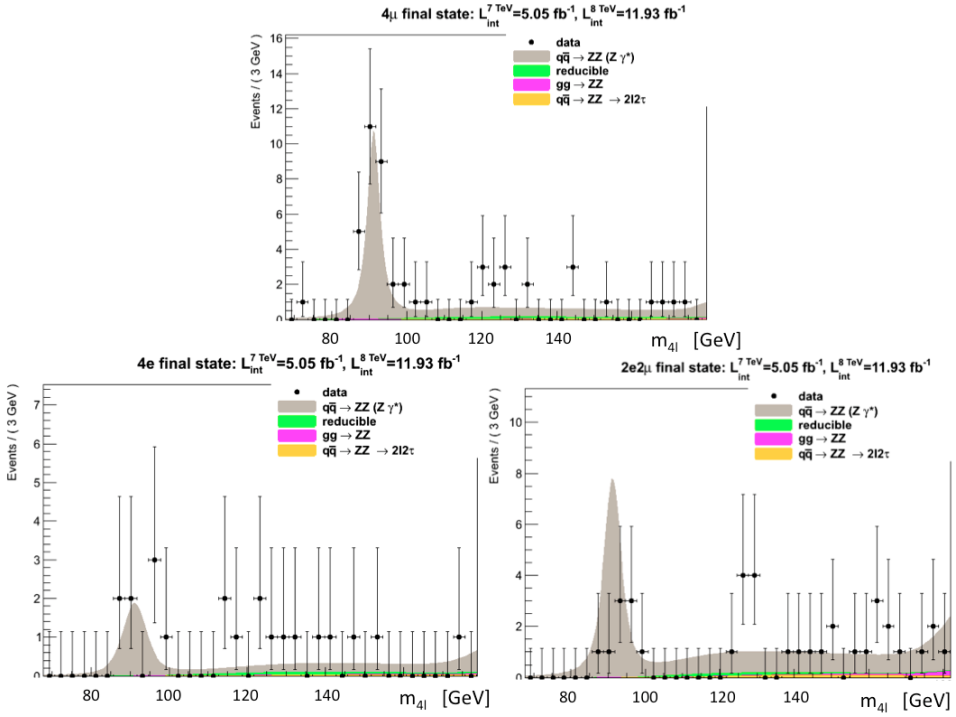


Figure 5.19: *m_{4l} distribution for data, for the three sub-channels: 4μ (top), $4e$ (bottom-left), $2e2\mu$ (bottom-right).*

There is an excess of events in data, peaking at $m_{4l} \sim 126$ GeV/c^2 , that suggests the possible presence of a signal. To evaluate the statistical significance of this excess the profile likelihood method is used, as described in the next chapter.

5.6.1 The $Z \to 4l$ peak

We can exploit the clear $Z \to 4l$ peak to estimate what is our capability to reconstruct a resonance peak, in the 4-lepton channel.

The number of $Z \to 4l$ events expected from MC and the yield observed in data are compatible and reported in the following table:

$m_{4l} \in [70, 100]$ GeV	4μ	$4e$	$2e2\mu$	**TOT**
MC	20.0	5.8	15.8	41.6
Data	30	8	9	47

To estimate how well we can fit the position of the peak, we let the m_Z parameter of the Voigtian to freely vary, and we fit the pdf function to the data, obtaining:

$$\hat{m_Z} = 91.72 \pm 0.57 \text{ GeV/c}^2$$

This value is very well compatible with the nominal Z mass value of 91.19 GeV/c^2 .

To estimate what is the experimental resolution, we let the σ parameter of the Voigtian to freely vary, keeping the Breit-Wigner γ fixed to the nominal Z width $\Gamma_Z = 2.5$ GeV/c^2 . We obtain an estimate of the m_{4l} resolution, as the fitted value of σ:

$$resolution = \hat{\sigma} = 2.27 \pm 0.57 \text{ GeV/c}^2$$

Which is compatible to what is expected from MC. The fit to $Z \rightarrow 4l$ peak is shown in Fig. 5.20. We can say that, with the information we have, we

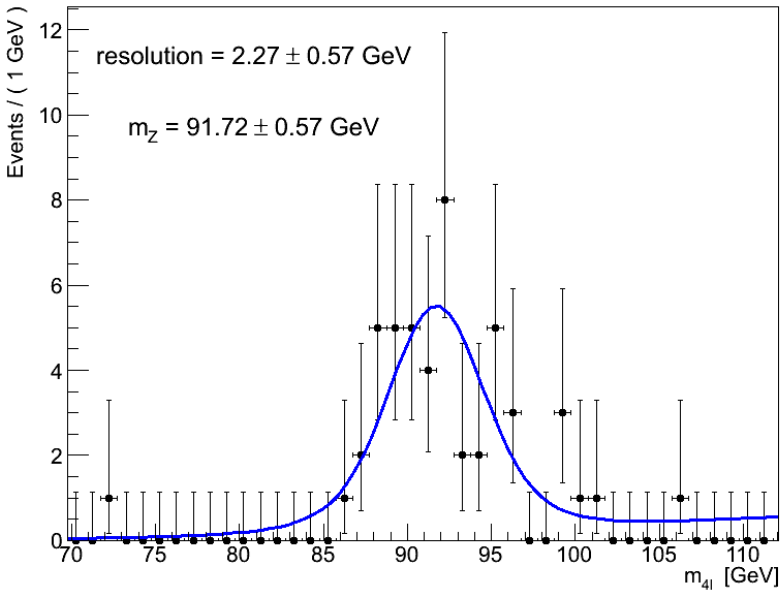


Figure 5.20: *Fit to the $Z \rightarrow 4l$ peak, with position (m_Z) and resolution (σ) as free parameters.*

are able to reconstruct a resonance peak in the m_{4l} distribution, and to fit its position correctly.

Chapter 6

Statistical analysis of data

Given the observed events, a detailed statistical analysis in necessary to establish whether this excess at low mass is compatible with a fluctuation of the background or has rather to be interpreted as signal. In this chapter, the modeling of the statistical frame devoted to discriminate between the background-only and the signal+background hypotheses is described. The low mass Higgs hypotheses are tested giving as output exclusion limits, for the SM Higgs (for the various mass values), and the statistical significance of the excess observed at low mass. At the mass where it is actually suggested the signal presence, the signal cross section is measured.

6.1 The Profile likelihood method

The statistical methodology used to test the signal presence hypothesis is described in the following [83].

The key ingredients are:

- a variable to observe in the data, called *observable*, and the probability distribution for it (possibly binned) under the signal+background ($s+b$) and background only (b) hypotheses:

$$P(obs|s+b) \qquad P(obs|b)$$

- a parameter of interest μ: called "signal strength modifier";
- the modeling of systematic uncertainties. A nuisance parameter θ_i is assigned to each source of systematic uncertainty. The expected signal and background yields are functions of these nuisance parameters, and are written as $\mu \cdot s(\theta)$ and $b(\theta)$, respectively. Most nuisance parameters are constrained by other measurements: they are included in the probability density functions $p_i(\tilde{\theta}_i|\theta_i)$, describing the probability to measure a value $\tilde{\theta}_i$ of the i-th nuisance parameter, given its true value θ_i.

We define the likelihood function L, given the data and the measured $\tilde{\theta}_i$:

$$L(data|\mu \cdot s(\theta) + b(\theta)) = P(data|\mu \cdot s(\theta) + b(\theta)) \cdot p(\tilde{\theta}|\theta)$$

if $P(data|\mu \cdot s(\theta) + b(\theta))$ is binned, it is a product of Poisson probabilities over all bins of the discriminant variable distribution.

In order to test the signal production hypothesis, we construct an appropriate test statistic q. The test statistic is a single number encompassing information on the observed data, expected signal, expected background, and the uncertainties associated with these expectations. It allows one to rank all possible experimental observations according to whether they are more consistent with b or with the $s + b$ hypotheses. In order to infer the presence or absence of a signal in the data, we compare the observed value of the test statistic with its distribution expected under the background-only $f(q|b)$ and under the signal+background hypotheses $f(q|s + b)$. The expected distributions are obtained by generating pseudo-datasets from the probability density functions $P(data|\mu \cdot s(\theta) + b(\theta))$ and $p(\tilde{\theta}|\theta)$. The values of the nuisance parameters θ, used for generating pseudo-datasets, are obtained from a fit to data, maximizing the likelihood L under the b or under the $s + b$ hypotheses.

6.1.1 Significance of an excess

In order to quantify the statistical significance of an excess over the background-only expectation, we define a test statistic q_0 as:

$$q_0 = -2 \log \frac{L(data|b(\hat{\theta}_0))}{L(data|\hat{\mu} \cdot s(\hat{\theta}) + b(\hat{\theta}))} \qquad \hat{\mu} \geq 0$$

where $\hat{\theta}_0$, $\hat{\theta}$, and $\hat{\mu}$ are the values of the parameters θ and μ that maximize the likelihoods in the numerator and denominator, and the subscript 0 indicates that the maximization in the numerator is done under the background-only hypothesis ($\mu = 0$). With this definition, a signal-like excess, i.e. $\hat{\mu} > 0$, corresponds to a positive value of q_0. In the absence of an excess, $\hat{\mu} = 0$, the likelihood ratio is equal to one, and q_0 is zero. This construction implies that a fit to the data is done, under the $s + b$ and b hypotheses.

In the asymptotic limit of a large number of events (large with respect to unity), the expected distributions of q_0 under the signal+background and under the background-only hypotheses are known analytically [84].

An excess can be quantified in terms of the p-value p_0, which is the probability to obtain a value of q_0 at least as large as the one observed in data, q_0^{obs}, under the background-only hypothesis:

$$p_0 = P(q_0 \geq q_0^{obs}|b)$$

We choose to transform the p-value in an equivalent quantity, the significance Z, via the Gaussian one-sided:

$$p_0 = \int_Z^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathrm{d}x$$

The significance is typically expressed in number of sigma σ: for Z = 3 we talk about a 3σ excess and "indication" of signal presence, for Z = 4 we talk about "evidence" of signal presence, for Z = 5 we talk about "discovery".

In Fig.6.1 is shown the typical scenario. The p-value is evaluated on the q_0 distribution under the background-only hypothesis $f(q_0|0)$ (left), and it is compared to the value expected in case of signal presence (right).

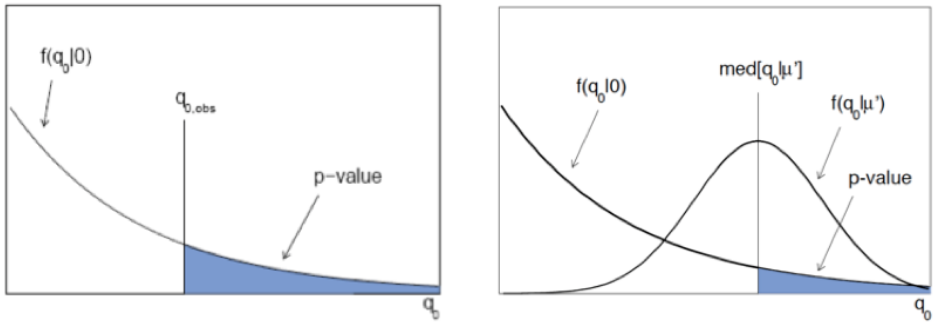


Figure 6.1: *Significance of an excess: the p-value is evaluated on the q_0 distribution under the background-only hypothesis (left), and it is compared to the value expected in case of signal presence (right).*

6.1.2 Exclusion limits

In order to set exclusion limits on a signal hypothesis, we define a test statistic q_μ, which depends on the hypothesized signal rate μ. The definition of q_μ makes use of a likelihood ratio similar to the one for q_0 , but uses instead the signal+background model in the numerator:

$$q_\mu = -2 \log \frac{L(data|\mu \cdot s(\hat{\theta}_\mu) + b(\hat{\theta}_\mu))}{L(data|\hat{\mu} \cdot s(\hat{\theta}) + b(\hat{\theta}))} \qquad 0 \leq \hat{\mu} < \mu$$

where the subscript in $\hat{\theta}_\mu$ indicates that, in this case, the maximization of the likelihood in the numerator is done under the hypothesis of a signal of strength μ. In order to force one-sided limits on the signal production rate, we constrain $\hat{\mu} < \mu$.

In the asymptotic limit of a large number of background events (large with respect to unity), the expected distributions of q_μ under the signal +

background and under the background-only hypotheses are known analytically [84].

For the calculation of the exclusion limit, we adopt the modified frequentist construction CL_s [86][87]. We define two tail probabilities associated with the observed data; namely, the probability to obtain a value for the test statistic q_μ larger than the observed value q_μ^{obs} for the signal+background $(\mu \cdot s + b)$ and for the background-only (b) hypotheses:

$$CL_{s+b} = P(q_\mu \geq q_\mu^{obs} | \mu \cdot s + b)$$

$$CL_b = P(q_\mu \geq q_\mu^{obs} | b)$$

the ratio gives CL_s:

$$CL_s = \frac{CL_{s+b}}{CL_b}$$

This method is studied to exclude the signal presence only if the analysis is sensitive to it. If $CL_s \leq \alpha$ for $\mu = 1$, we determine that the signal hypothesis is excluded at the $1-\alpha$ confidence level. To quote the upper limit on μ at the 95% confidence level, we adjust μ until we reach $CL_s = 0.05$. Equivalently, if the upper limit on μ goes below 1, the signal is excluded at the $1 - \alpha$ confidence level.

6.2 Building the model

In the $H \rightarrow ZZ \rightarrow 4l$ statistical analysis, the hypotheses to be tested are:

- the SM Higgs presence, for a given m_H, in addition to the other SM backgrounds
- the background-only presence

The observable is the four-lepton invariant mass m_{4l}. A *signal strength modifier* μ is introduced. It multiplies the expected SM Higgs boson cross section:

$$\sigma = \mu \cdot \sigma_{SM} \qquad \mu = \frac{\sigma}{\sigma_{SM}}$$

It is the parameter of interest.

To build the model, we use the expected signal and background yields y, and the pdf mass shapes f, as estimated in the previous chapter, to construct the probability distributions $P(m_{4l}|s+b)$ and $P(m_{4l}|b)$. The m_{4l} distribution expected in case of background-only is:

$$f_b(m_{4l}) = \sum_{i,c} y_{i,c} \; f_{i,c}(m_{4l})$$

the index i runs over all the backgrounds listed in the previous chapters, while c runs over the sub-channels. The m_{4l} distribution expected in case of signal+background is:

$$f_{s+b}(m_{4l};\mu) = \mu \sum_c y_c^S \; f_c^S(m_{4l}) + \sum_{i,c} y_{i,c} \; f_{i,c}(m_{4l})$$

In Fig. 6.2 the shape of the two distributions is shown. In case of a signal, a narrow resonance peak is expected, over the broader background.

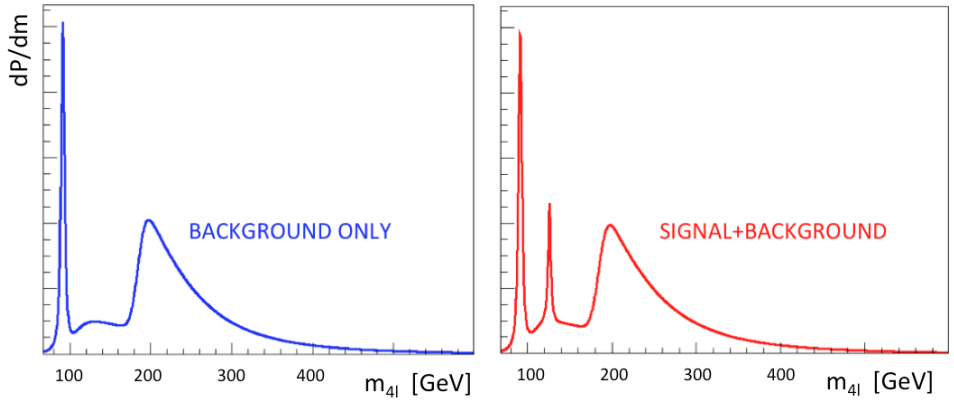


Figure 6.2: *Shape of mass distributions: Background-only (left) and Signal(m_H = 126)+Background (right).*

Nuisance parameters are used to model systematic uncertainties: these are present on the background yields, as evaluated in sections 5.4 and 5.5, and on the integrated luminosity and signal efficiencies. For each of these variables, x, a log-normal function is used to spread its nominal value:

$$x = x^{nom}(1+\delta_x)^{\beta_x} \qquad \beta_x \sim N(0,1)$$

δ_x is the percentage uncertainty on x^{nom}, β_x is the nuisance parameter associated to the variable x, on which we have uncertainty, distributed as a standard gaussian $N(0,1)$.

Taking the product of mass and nuisance distributions, we get the global pdfs:

$$f_B^{glob}(m_{4l};\vec{\beta_x}) = \left(\sum_{i,c} y_{i,c} \; f_{i,c}(m_{4l}) \right) \prod_{nuis} g(\beta_x)$$

$$f_{S+B}^{glob}(m_{4l};\mu,\vec{\beta_x}) = \left(\mu \sum_c y_c^S \; f_c^S(m_{4l}) + \sum_{i,c} y_{i,c} \; f_{i,c}(m_{4l}) \right) \prod_{nuis} g(\beta_x)$$

where $g(\beta_x)$ are standard gaussians $N(0,1)$. The free parameters of the global pdfs are, as indicated, the nuisance parameters β_x and the signal strength modifier μ.

6.3 Results

Given the relatively large amount of events (with respect to unity), we can use the asymptotic shapes for the discriminant q_0 and q_μ distributions (as reported in [84]).

In Fig. 6.3, the local p-value on the background-only hypothesis is shown, as a function of the Higgs mass, calculated with a 2 GeV/c^2 step. For a $m_H = 126$ GeV/c^2 we have an excess with a significance of 4.15 standard deviations (p-value = 1.67 10^{-5}). According to conventions, we can call this excess "evidence" for a boson with a mass value around 126 GeV/c^2 .

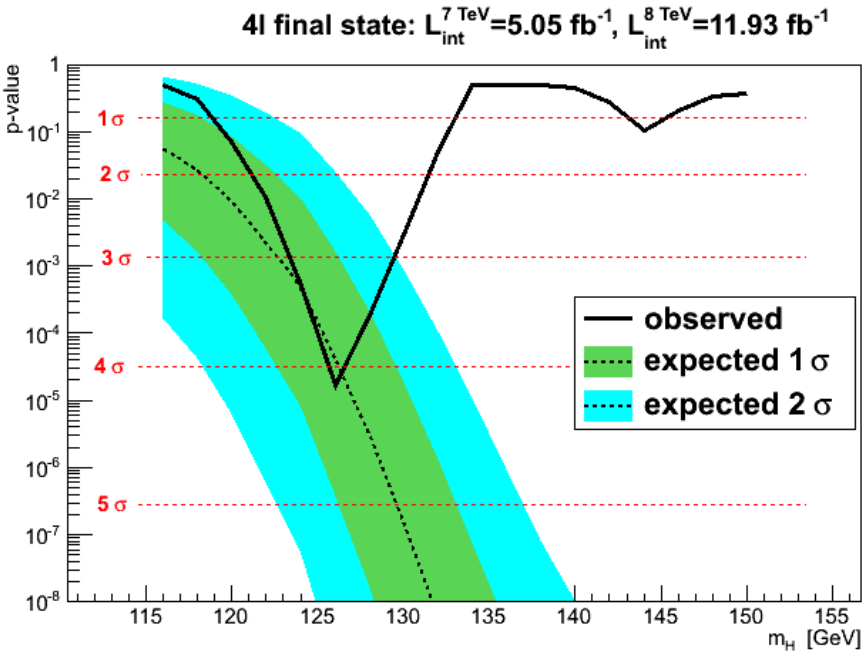


Figure 6.3: *Local p-value on the background-only hypothesis, as a function of the Higgs mass, compared to what expected for a SM Higgs signal.*

In Fig. 6.4 the 95% CL limit on σ/σ_{SM} as a function of the Higgs mass hypothesis is shown, calculated with a 2 GeV/c^2 step. The expected excluded interval, in case of background-only, is $m_H > 120$ GeV/c^2 , but the observed excess allows us to exclude the signal presence only for $m_H > 130$ GeV/c^2 .

In Fig. 6.5, CL_s is shown as a function of μ, for $m_H = 126$ GeV/c^2 : the observed limit, that is the value of μ for which CL_s equals 0.05, is much higher than the expected one in case of background-only, due to the excess of events observed in data.

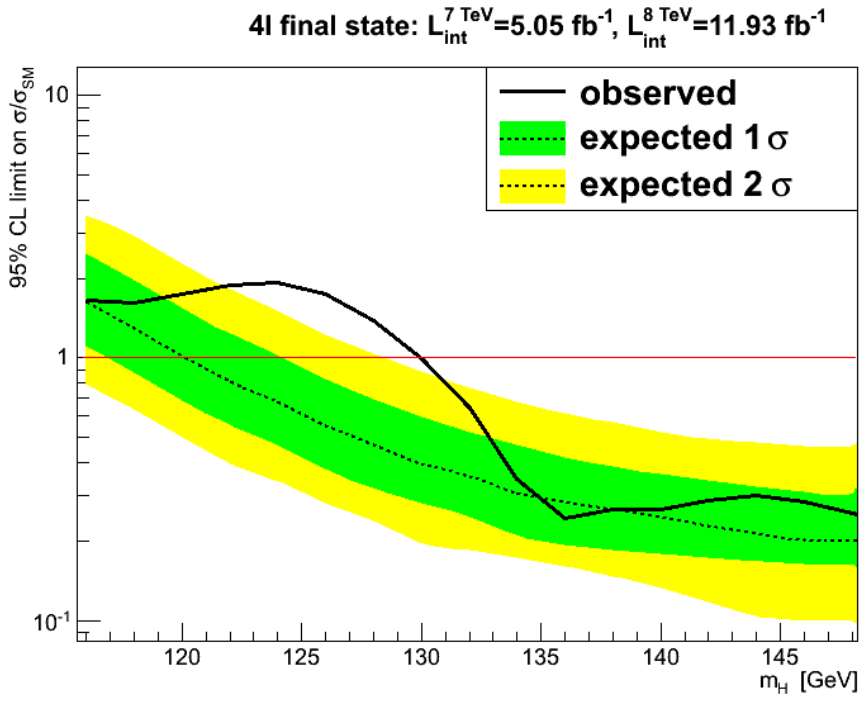


Figure 6.4: *95% CL limit on σ/σ_{SM}, as a function of the Higgs mass, compared to what expected for background-only.*

6.3.1 Mass measurement

To estimate the mass of this new boson, we use the graphical fit method [85]. Varying the Higgs mass hypothesis with a step of 0.05 GeV/c^2 , we fit the global signal+background pdf $f^{glob}_{S+B}(m_{4l};\mu,\beta_x)$ to data. Plotting the minimized value of the likelihood function as a function of the Higgs mass, we get the profile shown in Fig. 6.6. So the fit gives the best mass estimate:

$$\hat{m_H} = 126.2 \pm 0.55 \text{ GeV/c}^2$$

At $m_H = 126.2$ GeV/c^2 the local p-value on the background-only hypothesis is 1.55 10^{-5} with a significance of 4.17σ.

6.3.2 $H \to ZZ \to 4l$ cross section measurement

At $m_H = 126.2$ GeV/c^2 , the fit of $f^{glob}_{S+B}(m_{4l};\mu,\beta_x)$ to data gives, for the cross section modifier:

$$\hat{\mu} = 1.03 \pm 0.36$$

Its uncertainty already contains the systematics effects, since they are embedded in the model.

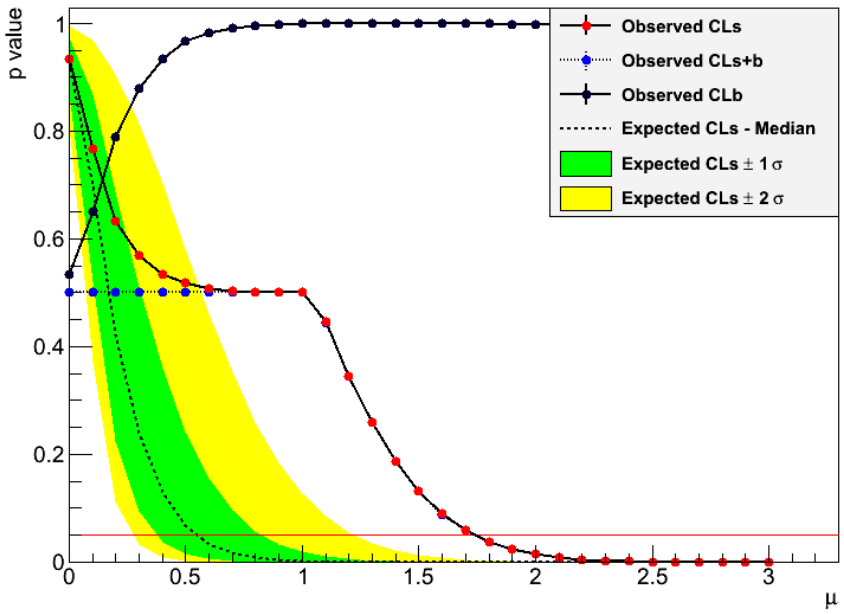


Figure 6.5: *CL_s as a function of μ for $m_H = 126$ GeV/c^2 , the observed limit, extracted with a scan of the interest parameter μ, is much higher than the expected one in case of background-only. 2 GeV/c^2 step.*

The m_{4l} distribution, with the signal superimposed over the background, is shown in Fig. 6.7

The number of events around the excess we see at 126 GeV/c^2 , expected from background, signal, and observed in data are reported in the following table:

$m_{4l} \in [110, 140]$ GeV	4μ	$4e$	$2e2\mu$	**TOT**
$N^{dd}_{qq\to ZZ}$	5.0	1.9	7.8	14.7
$N^{dd}_{gg\to ZZ}$	0.1	0.04	0.2	0.3
$N^{dd}_{ZZ\to 2l2\tau}$	0.3	0.2	0.6	1.1
N^{dd}_{red}	1.1	0.7	1.0	2.8
TOT bkg	6.6	2.8	9.6	18.9
H 126.2	4.5	2.1	6.2	12.8
Data	11	9	10	30

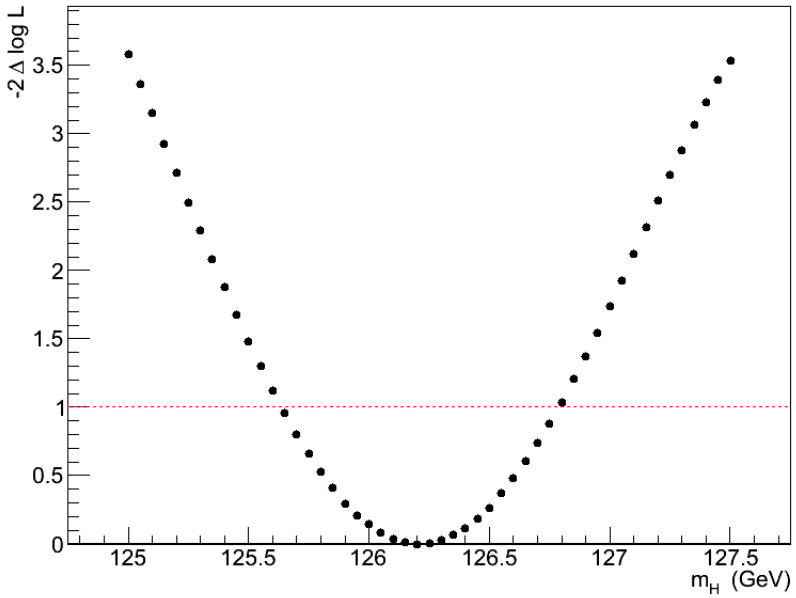


Figure 6.6: *Graphical fit to the Higgs mass: $-2\Delta \log L$ as a function of the mass hypothesis.*

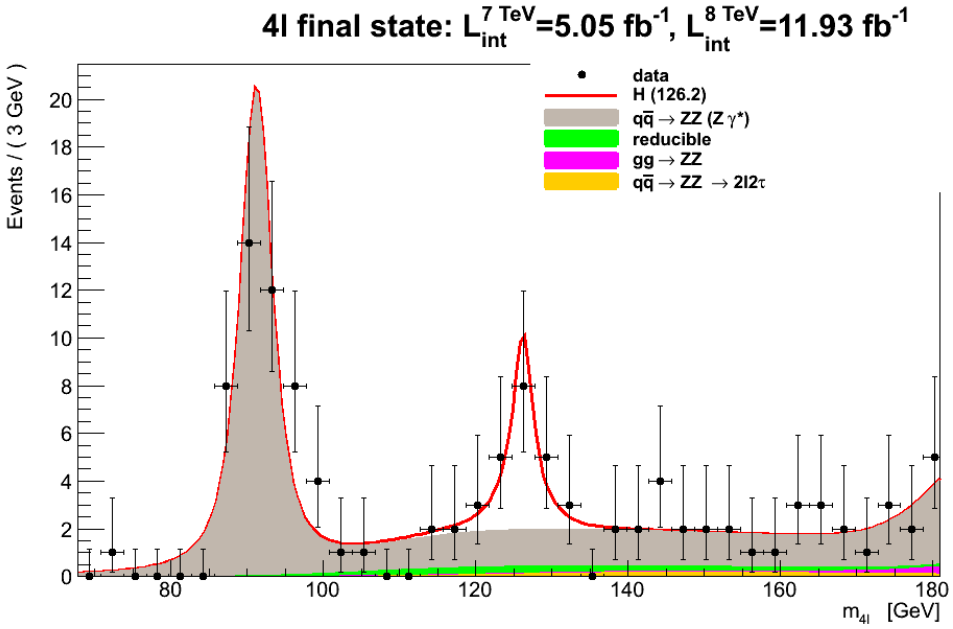


Figure 6.7: *The m_{4l} distribution of data, with background and signal superimposed.*

Conclusions

The search for the SM Higgs boson in the four-lepton channel at the CMS experiment has been presented, referring to 5+12 fb^{-1} of data at the LHC collider. The search makes use of the great detector performances in terms of efficiency-resolution and physical object reconstruction (particularly lepton reconstruction and identification). The background contributions have been studied using data-driven techniques and control regions, and have shown to be well understood. The search has lead to an evidence for a signal presence with a statistical significance of more than four standard deviations. The excess of data, with respect to the background-only predictions, indicates the presence of a new boson, with a mass of about 126 GeV/c^2 , decaying to two Z bosons, whose characteristics are compatible with the SM Higgs ones.

These results have to be contextualized in the announcement made by the CMS and ATLAS collaborations in July 2012: the discovery of a new particle, in the search for the SM Higgs boson, at a mass of about 125-126 GeV/c^2 [88][89]. The local significances of data excesses, over the backgrounds, were more than 5 standard deviations in both experiments (see Fig. 6.8). The results were obtained combining all the Higgs searches in the various sub-channels, the main analyses contributing to the discovery were $H \to \gamma\gamma$ and $H \to ZZ \to 4l$. The four-lepton analysis, presented in this thesis, has to be compared to the official CMS one [64]. The analyses are slightly different, since the event selection, the background modeling and control and the statistical analysis are carried on independently in the two approaches. Still, the final results, sketched in Fig. 6.9 and 6.10, are similar and compatible. The work presented in this thesis can be seen as a cross check to the official one.

Having observed a new particle, its properties have to be measured: particularly its mass, cross section, spin-parity and the couplings to other particles (BRs). This is needed to establish weather the new particle is actually the expected SM Higgs boson or just a similar one, possibly being a gate to "new physics" scenarios. The measurements of mass and cross section have already given positive results while the measurement of spin-parity requires more data. The couplings are measured by the signal yields in sub-channel searches: while $H \to \gamma\gamma$, $H \to ZZ \to 4l$ and $H \to WW$ have

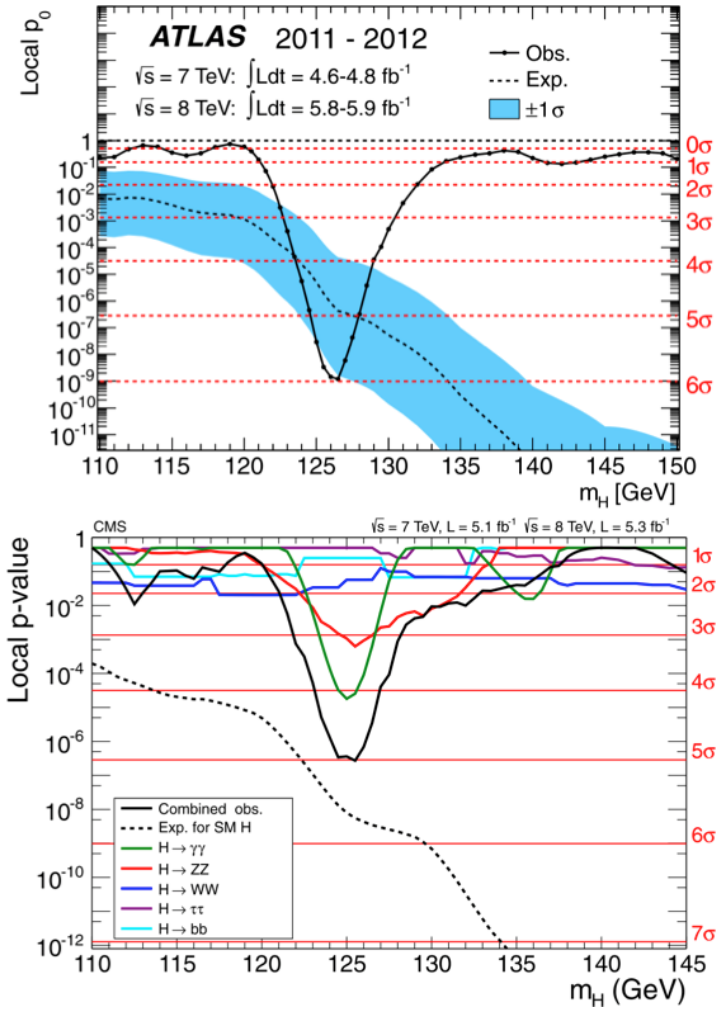


Figure 6.8: *CMS and ATLAS combined p-values on the background-only hypothesis. Both experiments see a discovery-level excess at a mass of about 125-126 GeV/c^2 .*

already fair individual estimates, more data are needed for the fermionic channels.

By the end of 2012, CMS has collected an integrated luminosity of about 21.8 fb^{-1}. These data should allow to have more precise results and the first hints concerning the measurement of the most challenging quantities.

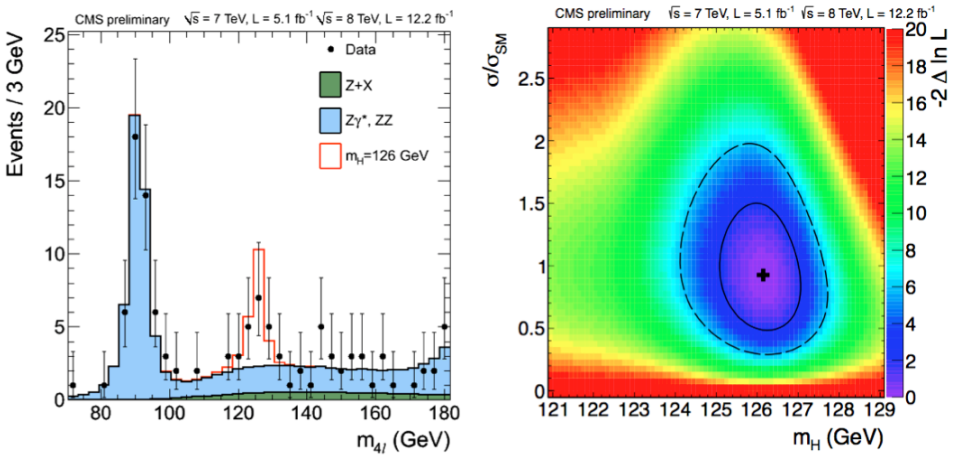


Figure 6.9: *Official results of the CMS $H \to ZZ \to 4l$ analysis. The mass distribution is shown on the left, the best fit to Higgs mass and cross section is shown on the right.*

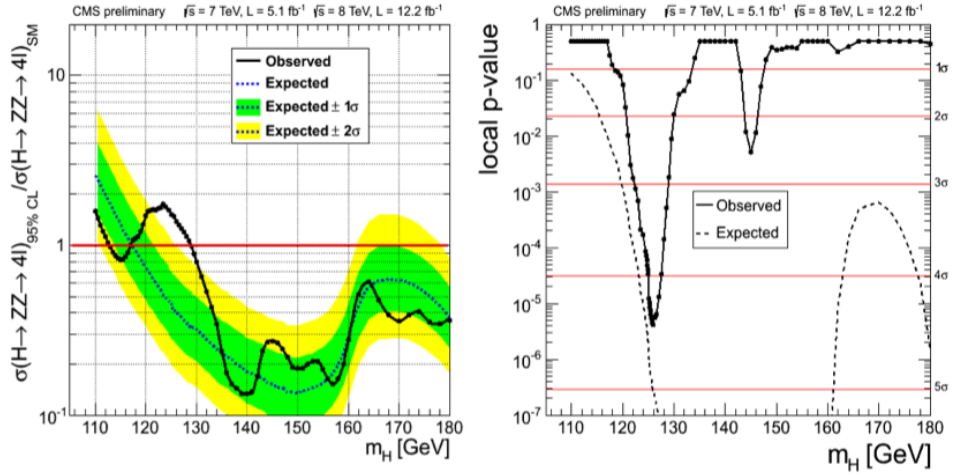


Figure 6.10: *Official results of the CMS $H \to ZZ \to 4l$ analysis. The limits of the signal cross section are shown on the left, the local p-value as a function of the Higgs mass is shown on the right.*

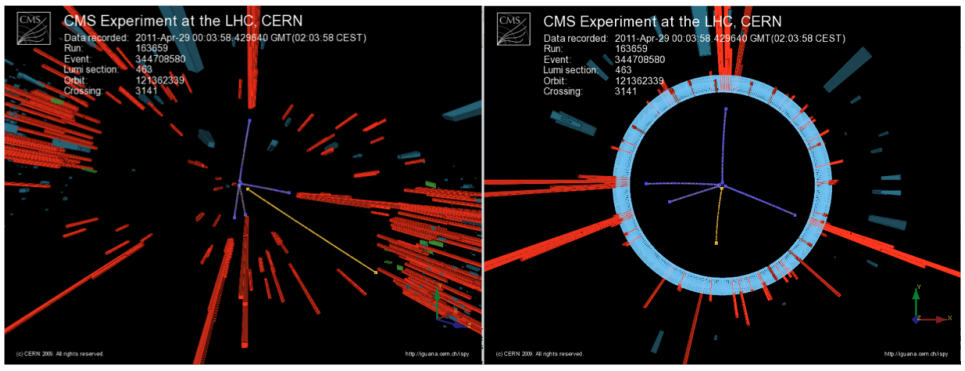
Appendix A

Event displays

Some events, passing the full analysis selection, are shown in the following. Their main characteristics are described.

7 TeV - 4e candidate

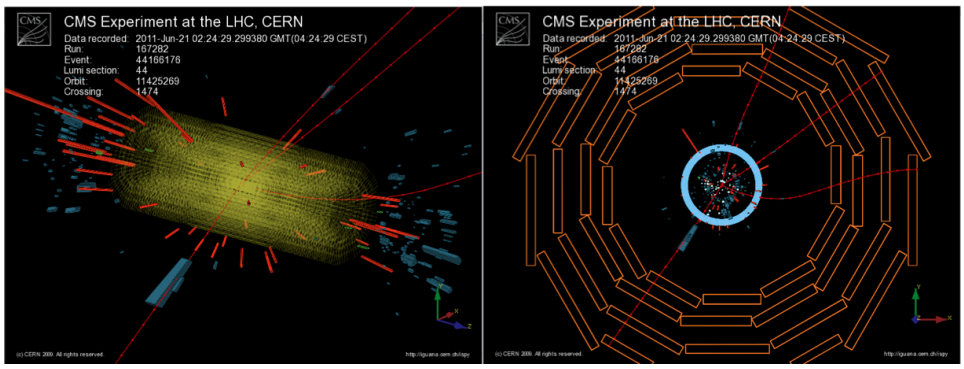
- $m_{4l} = 137.7$ GeV/c^2 $m_{Z_1} = 91.9$ GeV/c^2 $m_{Z_2} = 28.8$ GeV/c^2
- 4 electrons
- No extra electron nor muon
- No extra photon
- 2 extra tracks $p_T > 5$ GeV/c
- 4 extra jets $p_T > 10$ GeV/c
- 7 vertices in the event (all leptons come from the same primary vertex)



3D view (left) and $R-\phi$ view (right) of the event (Run 163659 - Event 344708580). The four electrons (in blue) are all in the ECAL barrel.

7 TeV - 4μ candidate

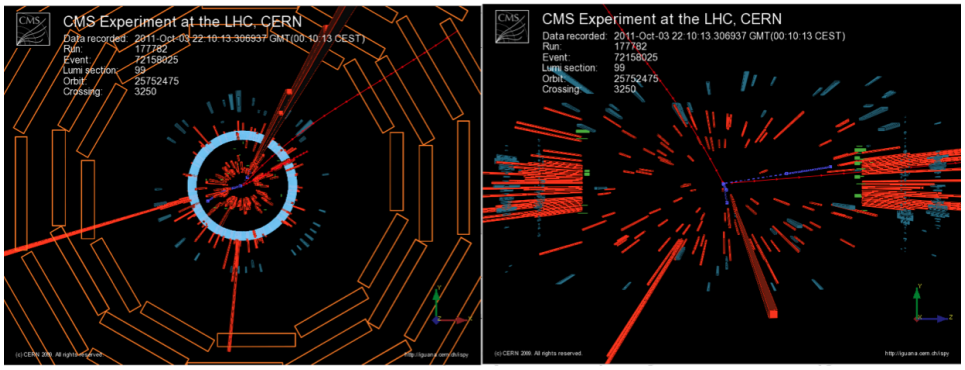
- $m_{4l} = 118.8$ GeV/c^2 $m_{Z_1} = 90.3$ GeV/c^2 $m_{Z_2} = 14.7$ GeV/c^2
- 4 muons
- No extra electron nor muon
- No extra photon
- No extra tracks with $p_T > 5$ GeV/c
- 1 extra jet, $p_T > 10$ GeV/c
- 2 vertices in the event (all leptons come from the same primary vertex)



3D view (left) and $R-\phi$ view (right) of the event (Run 167282 - Event 44166176). The four muons are in red. This is a very clean event with only two vertices.

7 TeV - $2e2\mu$ candidate

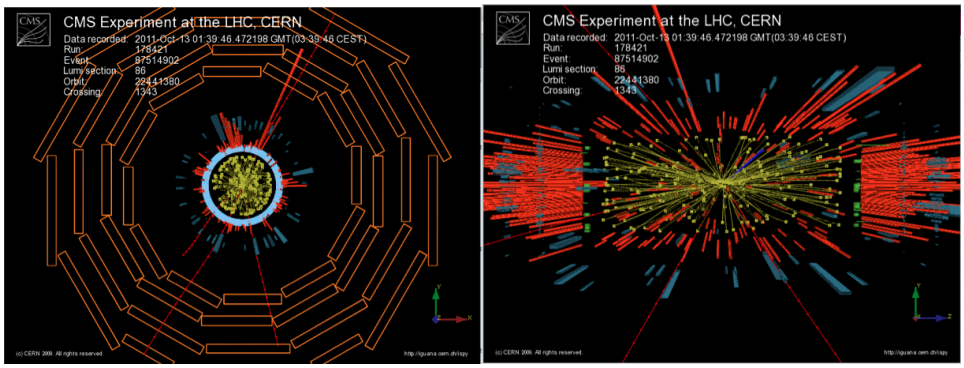
- $m_{4l} = 125.8$ GeV/c^2 $m_{Z_1} = 67.5$ GeV/c^2 $m_{Z_2} = 48.5$ GeV/c^2
- 2 electrons - 2 muons
- 1 extra electron, $p_T = 2.8$ GeV/c
- No extra photon
- No extra tracks with $p_T > 5$ GeV/c
- 6 extra jets with $p_T > 10$ GeV/c
- 12 vertices in the event (all leptons come from the same primary vertex)



$R-\phi$ view (left) and $R-Z$ view (right) of the event (Run 177782 - Event 72158025). Electrons are in blue, muons are in red.

7 TeV - 4μ candidate

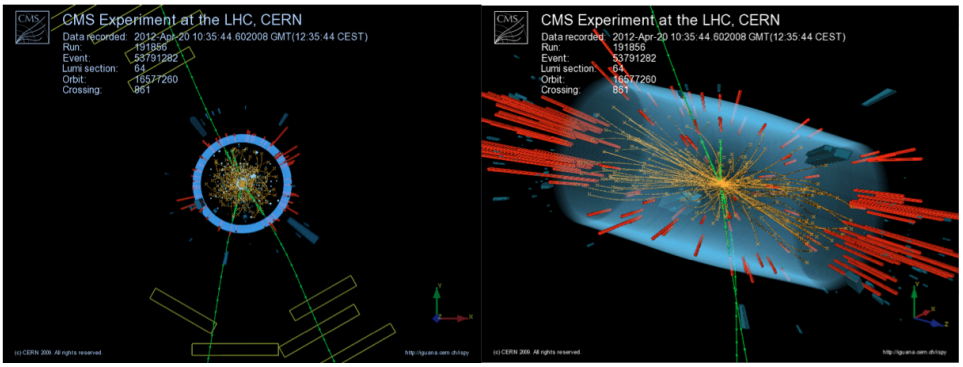
- $m_{4l} = 126.1$ GeV/c^2 $m_{Z_1} = 94.3$ GeV/c^2 $m_{Z_2} = 26.0$ GeV/c^2
- 4 muons
- 1 extra electron ($p_T > 6.7$ GeV/c) within a jet
- 1 extra photon, recovered through the FSR recovery algorithm
- 6 extra tracks with $p_T > 5$ GeV/c
- 18 extra jets with $p_T > 10$ GeV/c (leading one with ~ 27 GeV/c)
- 25 vertices in the event, 952 tracks (all leptons come from the same primary vertex)



$R-\phi$ view (left) and $R-Z$ view (right) of the event (Run 178421 - Event 87514902). This event has a very high vertex multiplicity (25 vertices, 952 tracks).

8 TeV - 4μ candidate

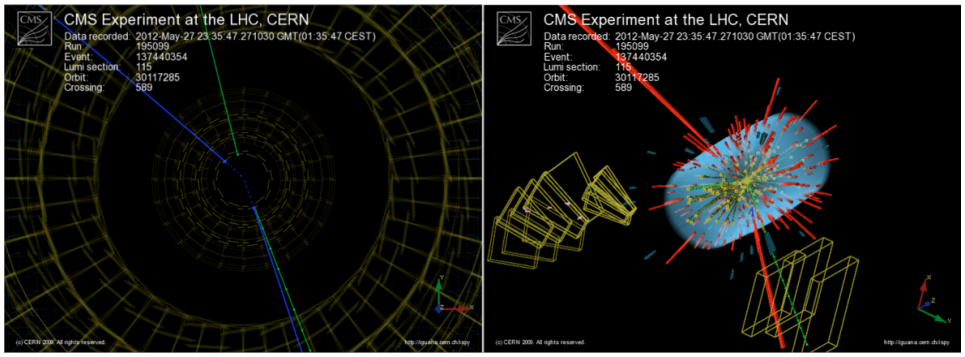
- $m_{4l} = 125.2$ GeV/c^2 $m_{Z_1} = 82.7$ GeV/c^2 $m_{Z_2} = 38.9$ GeV/c^2
- 4 muons
- No extra photon
- 2 extra tracks with $p_T > 5$ GeV/c
- 1 jet, with $p_T > 10$ GeV/c in the central region
- 11 vertices in the event (all leptons come from the same primary vertex)



$R-\phi$ view (left) and 3D view (right) of the event (Run 191856 - Event 53791282). All muons (in green) are central. One muon ($p_T = 22.3$ GeV) is only a tracker muon.

8 TeV - $2e2\mu$ candidate

- $m_{4l} = 127.0$ GeV/c^2 $m_{Z_1} = 75.9$ GeV/c^2 $m_{Z_2} = 29.2$ GeV/c^2
- 2 electrons - 2 muons
- No extra photon
- 3 extra tracks with $p_T > 4$ GeV/c
- 8 jets with $p_T > 10$ GeV/c
- 19 vertices in the event (all leptons come from the same primary vertex)



$R-\phi$ view (left) and 3D view (right) of the event (Run 195099 - Event 137440354). Muons are in green, electrons in blue.

Appendix B

Event list

7 TeV 4μ							
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
172992	1153485608	836	92.3974	69.4574	17.1781	0.0760517	1.13204
171369	160966858	150	218.87	90.1971	88.9331	0.232304	0.948184
167282	44166176	44	118.83	90.3153	14.7241	0.062304	1.44288
172952	559839432	466	231.938	91.1546	37.1512	0.356472	1.22842
172802	107360878	125	457.923	91.9017	85.0364	0.0551711	2.01946
167281	480301165	386	222.302	90.4201	54.8171	0.183197	1.11471
163817	155679852	174	144.912	91.303	34.8268	0	3.41878
166438	862270386	768	211.644	92.1058	15.0921	0.220267	1.51336
172819	298086610	220	92.8251	63.8447	13.3948	0.193693	1.95909
172163	191231387	128	198.821	92.1508	87.6951	0.108518	1.92409
166512	337493970	281	238.532	90.9722	93.1793	0.023534	1.62359
172208	66033190	75	308.561	87.6953	97.0167	0.0225639	2.31566
180250	905064541	496	88.7151	68.8727	13.5239	0.17237	2.06822
180250	45096064	28	453.408	91.6521	90.2473	0.339246	1.52684
178970	658883361	398	91.1787	72.7587	16.9671	0.107998	1.61631
176309	257489763	224	193.86	89.7658	86.3023	0.0198459	1.14212
177074	588602439	384	240.346	93.0132	87.6972	0.038652	2.78619
177790	657843813	527	237.879	90.9626	91.7448	0.0214508	1.57688
180076	456795917	271	91.1787	62.5269	17.24	0.125003	1.92663
177074	931848091	582	88.956	67.9472	15.7825	0.0328903	1.35294
178479	298608854	210	269.943	89.9294	89.7902	0.121533	1.57605
176799	35688265	24	193.42	90.5828	92.7722	0.106009	3.44632
178421	87514902	86	126.144	94.254	25.9949	0.343842	3.8152
177875	419370375	289	71.8107	52.2694	12.8962	0.154519	0.98267
178424	666626491	585	121.987	59.1591	31.6036	0.163328	2.72177
177790	222240677	168	280.363	91.6246	93.2306	0.100632	2.48329
179476	30532070	30	226.829	92.2916	92.3045	0.0718216	1.11403

7 TeV $4e$							
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
166438	78213037	79	214.528	95.0008	42.7092	0.0607279	3.93182
173243	16706390	12	284.934	90.825	93.1717	0.0526326	2.27029
171106	141954801	127	191.2	91.4006	91.8994	0.0227072	1.03303
172799	10347106	11	363.865	91.3144	81.6939	0.00522264	0.909415
163659	344708580	463	137.685	91.8968	28.7292	0.0103517	0.531106
163758	113529826	151	98.9094	68.9719	20.1743	0.103537	1.7296
173659	389185367	270	232.483	91.1148	89.2291	0.107357	1.16786
177318	270676815	169	369.891	93.0294	87.2499	0.0596805	1.26552
176201	562295642	354	132.459	85.9345	43.0253	0.0127831	1.99541
176548	403771114	231	328.787	94.2373	96.7053	0.0526677	2.91078
178162	10608364	10	223.069	88.6423	87.44	0.210849	1.65977
177875	148667118	133	366.257	90.3463	94.3917	0.251618	2.40101
178421	1610336854	1087	192.03	92.8716	60.0557	0.0169561	1.99858
180250	591651181	326	284.245	91.1332	93.3126	0.237798	0.758672
180076	79350642	46	196.479	92.2783	89.6616	0.0659627	3.70262

7 TeV							
$2e2\mu$	p.1						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
166408	917379387	724	256.837	88.1723	105.275	0.0206999	2.46947
163795	30998576	34	208.101	92.8129	82.3452	0.137923	2.37601
172868	933807102	689	314.899	92.2191	88.8583	0.0153019	0.721498
166950	1491724484	1373	191.969	91.3258	92.3887	0.163884	1.4354
172401	3729470	7	185.92	90.6852	92.4524	0.0149594	1.54442
167807	966824024	750	326.731	90.1922	95.3969	0.118519	1.73546
171050	591031316	459	95.5249	78.4279	16.1558	0.260451	1.58056
172949	1188043146	840	194.351	91.8215	88.7132	0.12291	1.74686
172620	218903169	242	130.308	92.2661	16.9626	0.22785	2.528
173657	65557571	85	388.564	91.5183	86.8527	0.0360944	2.56669
165633	394010457	303	243.677	91.2188	92.561	0.0600928	3.60426
166922	112725318	85	99.7349	55.3479	19.2241	0.187981	1.42911
163334	129514273	221	191.715	90.4188	93.9269	0.0563844	1.27604
163334	286336207	499	164.74	94.5133	66.7911	0.109265	0.396851
173692	2722114329	2066	190.009	90.4363	94.1232	0.0383779	0.421278
176468	215855118	128	325.315	94.8122	102.73	0.0449241	1.7481
178703	191352626	137	196.719	87.1381	84.404	0.0771812	1.40051
177782	72158025	99	125.788	67.5316	48.5022	0.132564	0.86952
175906	227517585	190	308.058	92.3998	94.2292	0.0980148	1.50149
175921	297753357	220	232.387	95.7479	85.2506	0.0434527	1.87407
175921	495614354	349	204.671	92.9164	96.3104	0.0199605	1.20521
175974	7526662	9	210.292	92.2295	97.5905	0.0462278	1.88047
178708	573962528	354	252.732	95.0648	80.5315	0.216788	2.20865

7 TeV							
$2e2\mu$	p.2						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
176886	1057019814	631	257.318	90.1327	92.2985	0.127768	3.9248
177139	290826062	183	192.18	94.8492	86.489	0.153594	1.62938
176304	418052877	300	205.803	91.7488	101.161	0.0560677	1.46232
176201	261184429	182	182.744	90.5773	16.9998	0.0343589	2.50237
176207	256888239	206	278.437	91.2834	94.2227	0.0768472	1.80222
179434	86225612	52	210.255	94.0808	94.5492	0.103344	1.33296
176886	427567024	260	533.634	93.723	95.1855	0.0247629	1.01411
177222	339499459	227	308.633	90.1481	86.7054	0.0862741	1.60151
178479	589085976	369	272.996	89.4718	86.3007	0.00328255	2.01243
176309	1340034258	950	130.062	75.5712	12.9278	0.0187917	2.03539
177449	58273256	68	94.9674	54.7275	17.069	0.249366	1.46972
178479	757111474	470	349.155	89.8074	93.4973	0.0896372	3.61432
178116	695859609	428	221.192	91.4945	86.7689	0.0566346	2.01787
178100	326364918	236	278.485	92.0812	86.556	0.0568085	0.805068
178116	709511403	437	341.155	94.8689	95.4206	0.0512062	1.49482
178731	248562036	192	319.8	89.3718	93.2017	0.254855	1.16374
178421	1450980155	973	178.449	89.6494	82.6185	0.0279448	0.507222
178786	277942410	197	386.656	92.7079	94.2537	0.0554758	1.608
178970	122998167	103	245.256	91.6288	90.4695	0.128039	1.6965
179452	1459855927	1056	162.042	93.2926	42.1631	0.0360797	2.27496
165970	275108397	236	142.429	91.6411	14.8522	0.0681013	1.50213
166554	395098004	333	89.9089	56.3183	19.5513	0.144305	2.23695
179563	1409064222	871	213.689	90.2735	87.3577	0.133583	1.50941
178866	140063742	82	151.525	85.8024	47.707	0.323359	1.97734

8 TeV							
4μ	p.1						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
191720	131384043	141	91.9899	71.818	17.1991	0.179625	1.82839
191830	480555885	318	167.508	91.3048	74.8821	0.137669	3.81964
191046	111444776	84	94.5718	60.6901	14.0374	0.177791	1.19655
191856	53791282	64	125.248	82.725	38.9256	0.0242782	1.21287
193575	246768035	318	248.042	88.3416	15.8889	0.188996	1.71183
196453	363484809	399	212.074	92.0222	81.1945	0.083446	1.00626
194224	400957190	261	190.556	90.7036	91.7236	0.151056	1.18478
194644	78891305	72	164.456	92.8001	29.1205	0.1969	1.07288
194704	372667387	415	229.315	90.6787	95.7087	0.0436739	1.53707
194789	164079659	118	241.158	81.0159	59.9726	0.777582	2.96813
194897	146471772	81	207.347	89.45	85.4072	0.0407862	2.62602
194912	829690349	503	88.3995	64.3396	20.8018	0.0925157	2.21238
195147	502821363	419	233.45	94.5241	103.085	0.0917785	1.79802
195251	147388276	80	215.427	91.0293	94.4109	0.0781044	1.66804
195304	128008670	90	206.152	89.3924	85.7474	0.070958	1.47127
195304	487153301	382	254.079	90.9386	34.0849	0.0501825	1.82312
190895	593623897	559	99.6613	81.2586	12.25	0.105735	1.70603
190895	836690441	839	386.773	91.9367	90.2944	0.30117	1.1873
195655	167570931	140	130.911	82.2002	20.0762	0.00698305	1.98692
195655	477289466	432	197.932	87.2823	77.2763	0.139971	0.773495
195658	338722570	364	243.958	90.5444	89.4043	0	1.24916
195774	152805866	78	344.476	91.3552	90.4697	0.0578741	1.74049
195774	174957822	90	197.159	93.8655	81.5737	0.0399603	1.11629
195774	549661559	312	173.692	89.5085	17.3657	0.035834	1.59134
195774	660605607	387	754.854	91.1455	85.5633	0.034326	1.58891
195950	630954116	667	179.783	88.8294	71.798	0.0470835	3.68985
195950	719208626	781	195.1	91.1218	95.0443	0.0775315	2.76655
195950	640913928	680	101.067	69.445	14.3959	0.350258	2.07799
196197	352383455	238	549.812	92.1352	93.8603	0.00540549	1.2393
196218	198713417	150	92.2948	57.3426	15.3516	0.12991	0.707581
196218	227768389	166	89.0455	63.7182	17.741	0.254555	2.08357
196218	858760563	590	89.3176	45.3615	30.9046	0.156866	1.34703
196239	687711552	766	450.99	89.1318	96.3451	0.116658	3.42546
196364	646351604	696	332.383	89.5868	94.556	0.0515274	2.54144
196364	574417012	606	90.5673	49.4733	29.2283	0.212214	2.4742

8 TeV							
4μ	p.2						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
198272	345644314	382	171.798	91.2719	62.0916	0.0570444	1.86083
198272	141178966	148	307.847	90.3873	94.0823	0.0276209	1.39283
198271	36542571	27	120.518	78.7551	18.3752	0.480543	1.38411
199008	115388710	118	183.644	85.5437	29.9389	0.0415734	1.70727
199021	99969333	107	86.4732	44.5341	16.1375	0.0415042	1.59651
199021	358003591	292	94.2805	57.6758	19.998	0.283486	1.44759
199021	919744647	782	90.1597	70.7346	14.0828	0.288566	1.20916
199021	1238510739	1083	279.405	92.0291	99.3147	0.0952423	0.795097
199319	770334833	580	269.155	91.1494	69.1775	0.119919	1.42987
199319	1203594102	970	125.474	58.3232	45.7291	0.214209	1.89666
199428	654887173	548	198.618	90.8848	88.9857	0.0430612	0.708787
198955	575302642	493	187.785	90.1213	88.87	0.0896476	1.11039
198955	1211003130	1146	144.7	83.3773	50.281	0.133378	1.6341
198969	797334672	659	95.8066	72.4312	12.4066	0.127262	1.27614
198969	809605982	672	199.064	91.1691	89.0712	0.0327752	0.998643
198941	262583111	217	132.56	87.6321	16.7763	0.318519	2.36361
198969	1103535297	937	463.202	92.1457	89.9052	0.0605815	1.31183
198487	664920764	580	92.0121	60.3863	24.6875	0.103385	1.8118
198487	1511453808	1522	446.515	94.7419	100.782	0.0532968	1.37796
199608	930210675	803	90.0111	68.2387	12.8421	0.171893	1.715
199409	402443918	303	206.538	91.7065	82.8884	0.0219534	0.857528
199752	133641846	104	232.52	88.0039	85.6695	0.104281	0.876604
199833	991082101	893	198.232	89.3331	89.0937	0.0808411	0.935994
199864	396825385	296	188.303	90.4327	86.0074	0.103632	1.40994
199876	331257910	289	239.431	93.1441	88.6665	0.133538	1.15121
199876	331969560	290	246.517	88.1171	96.3704	0.0966782	0.953324
199877	636036234	642	91.1158	55.6106	22.3201	0.0610033	1.45534
200042	250036559	300	252.045	97.2432	112.136	0.130168	2.31641
200091	1745216870	1655	314.077	91.9248	81.7329	0.123793	0.723118
200190	301478208	240	205.757	91.5826	87.0313	0.0743104	1.46596
200229	37968037	44	257.499	90.4122	82.8464	0.244098	1.59017
200244	95922345	59	187.492	88.7571	83.3705	0.103894	1.11872
200491	230077151	195	100.129	57.6672	14.7976	0.097241	1.4135
200600	1248257881	1019	204.161	91.4279	89.0637	0.0800996	1.81608

8 TeV							
4μ	p.3						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
200991	573405437	435	550.571	91.7384	82.2109	0.117896	0.883944
200992	468873972	572	318.67	91.7728	26.7197	0.176721	2.07215
200992	279759749	324	200.978	90.6735	92.5214	0.0800404	1.27583
201191	199183870	156	243.597	90.1076	94.2228	0.176597	0.813919
201191	491885761	319	87.7974	51.8211	32.9694	0.15106	2.61305
200525	547514345	410	191.07	91.2257	90.7336	0.105801	1.0426
200525	1066576272	866	210.674	92.615	72.7848	0.105397	1.10905
201278	1021783875	801	247.909	91.1682	97.1042	0.10753	1.62618
201625	546665360	369	144.773	92.6918	49.2993	0.0698798	1.93886
201707	552697757	426	87.1085	68.203	16.3415	0.17645	1.21672
201707	635670564	503	120.962	87.366	24.8341	0.28421	0.828304
201824	346074761	411	209.208	88.9816	94.0506	0.0160099	0.749805
202016	160415664	143	92.8167	57.1113	17.6491	0.160935	1.6895
201668	114003150	121	181.171	88.4212	87.9709	0.0174451	1.38201
202237	1087639087	777	213.014	91.3666	88.4088	0.298954	0.951913
202178	357059583	288	153.583	91.2239	40.8698	0.0106592	1.20233
202178	193770334	176	89.1192	46.2174	19.0431	0.218031	0.816504
202178	563981267	448	198.736	91.4398	92.5068	0.112208	0.930269
202178	1430970868	1286	122.59	77.5899	24.5533	0.115485	1.93978
202299	586836364	419	90.6099	51.8933	24.4651	0.169743	0.999058
202237	112735227	90	106.414	89.6274	15.5538	0.061653	1.39043
202478	933835054	901	266.941	92.9673	102.021	0.256914	1.3421
202504	177815904	158	390.443	90.3559	90.2371	0	0.69393
202504	974397273	780	96.1842	71.35	22.3191	0.237211	1.5914
202972	155730849	105	219.271	91.0357	90.7593	1.19355	1.07999
202973	681363400	713	118.308	60.792	36.3033	0.36509	1.75295
202328	726409783	493	92.6986	44.0241	21.1403	0.0200461	2.84785

8 TeV							
$4e$	p.1						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
193336	432667035	627	242.955	82.9929	76.2942	0.243685	0.379182
193575	400912970	523	122.432	74.2539	40.6877	0.156136	1.23431
193621	393672516	426	321.707	96.0635	78.9752	0.172548	0.819495
191062	330091192	303	203.708	91.0298	90.9494	0.101882	0.547438
195655	35364903	29	95.4814	60.5453	12.4079	0.0382573	0.818479
195656	78300349	85	237.322	95.0741	95.946	0.0851517	-0.745957
194119	168130224	190	197.992	90.8417	98.5973	0.0187516	1.71657
195948	450362044	269	272.088	93.5681	84.6304	0.0688937	1.99633
196027	123030498	79	115.711	83.5516	20.3851	0.301096	2.63709
196048	9047236	6	113.228	67.2617	39.1301	0.252104	2.00212
196349	234207734	169	180.157	89.7185	75.3684	0.0590785	1.44158
194153	93572313	100	200.105	85.7363	84.9557	0.100502	2.95923
194050	401484983	414	229.365	91.7564	91.7881	0.19655	1.0067
194480	109881951	122	192.479	91.0073	88.0927	0.111538	3.61638
194480	863682922	972	204.549	91.6522	96.4874	0.115219	1.27068
194076	340846024	323	147.345	78.8977	20.3548	0.199777	1.22356
196531	497827501	350	87.9303	58.8813	12.8075	0.0949314	1.58849
194108	818802354	859	260.285	96.3699	99.591	0.107155	1.46524
195304	1069824602	962	254.771	87.5653	97.2363	0.152643	0.898451
195378	155753218	148	207.419	91.6537	95.7703	0.0486384	1.56462
199021	232250171	198	215.868	93.2177	93.3818	0.921002	0.77622
198272	440202539	498	429.943	87.7507	105.769	0.308362	-0.213159
198212	15479119	27	90.8998	61.7081	20.6122	0.0404667	1.32312
198271	232701194	181	175.011	99.3066	49.8908	0.228861	0.943489
198485	174781164	164	280.491	88.1837	63.56	0.211383	1.90575
199569	111437792	90	153.222	78.6176	35.9391	0.0289552	1.29544
199608	813802208	697	240.091	91.6347	89.8635	0.0206644	1.98925
199703	217542176	244	141.855	79.145	46.1356	0.0546541	2.20507
199754	121019123	106	274.84	91.6252	78.7133	0.0808295	1.58899
199754	873005082	933	321.651	90.1891	92.9641	0.0239468	-0.160263
199876	92622577	75	130.338	87.7782	35.3037	0.0551849	0.827262
200091	1530537922	1391	233.695	91.3707	100.425	0.280142	-1.02006
200188	187920671	139	117.761	93.7181	14.9095	0.163845	2.47745

8 TeV							
$4e$	p.2						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
201191	1546558727	1183	179.403	98.4655	13.5471	0.0389404	1.50712
201196	38288110	44	187.392	95.8528	23.8193	0.386784	1.94599
201196	88225327	102	95.2831	62.2376	14.4276	0.330967	2.94477
200991	98100215	63	192.935	93.8509	85.236	0.14977	0.877434
200991	226487761	152	87.541	63.7198	21.8795	0.27877	1.101
200991	970724930	787	249.972	92.9551	81.88	0.0488409	1.71308
201278	1723198502	1517	88.8692	64.6668	16.187	0.0849974	1.75078
201707	805047482	656	125.501	74.5047	13.8888	0.0399108	1.83729
201727	43697222	83	315.19	88.5096	86.5763	0.100341	-0.126639
202016	191382222	173	468.421	91.9813	95.3339	0.118446	0.93565
202016	559670208	505	513.765	86.1981	84.7677	0.0298383	1.21773
202074	38166085	73	269.478	90.3952	78.7975	0.111268	1.00099
202178	487671954	386	211.139	92.2083	89.881	0.0326567	2.01163
202272	35886066	86	95.2452	62.2238	27.0858	0.347044	-0.86599
202299	421267699	304	124.656	93.5063	28.4389	0.0627645	1.32678
202478	1025978128	1008	188.249	90.5978	26.8653	0.214035	3.40209
202504	535251232	395	490.672	95.1801	76.2685	0.14909	2.60606
202333	235647099	212	188.03	91.1092	84.9082	0.182004	1.2065
202469	247836553	201	200.866	92.614	82.2496	0.031017	1.82954

8 TeV							
$2e2\mu$	p.1						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
191226	979037526	712	242.88	93.6789	83.1856	0.0904381	0.531931
191226	1730707882	1420	267.469	93.3165	94.6258	0.117227	1.56326
191226	1820521419	1520	201.023	91.1912	90.5632	0.919996	1.58107
191277	617265349	704	392.96	91.2622	86.4992	0.016511	0.938046
193541	306664497	480	180.612	90.242	85.3041	0.055983	1.81255
195398	53211301	66	274.627	89.5876	89.329	0.145753	1.0448
195530	215099909	137	392.295	91.2097	91.7577	0.0346135	1.40992
195552	793110394	553	588.935	97.3598	104.324	0.0320688	1.62995
195552	1487349718	1310	203.569	91.0302	34.6649	0.0422371	2.30518
195649	323790844	223	204.43	88.1118	81.5648	0.0886334	1.527
195658	59123305	71	205.768	91.8984	90.0043	0.0317992	0.518323
195757	244623214	148	188.608	91.5073	86.7011	0.171474	3.43001
195915	426559109	251	269.259	92.1192	84.9812	0.146851	2.32808
195915	990388348	690	122.537	52.3763	12.8923	0.322648	1.28279
195937	194955393	240	202.092	93.6296	67.643	0.119786	1.42941
196218	446438292	296	226.499	92.3805	92.5946	0.0142456	0.73752
196218	794370838	539	207.395	92.0651	92.4317	0.0676286	2.4416
196334	252537006	177	380.834	90.5431	93.4449	0.111904	3.37585
196452	1315227994	1016	275.775	87.6307	96.9881	0.0758028	0.95157
196453	1181788896	1337	201.27	91.2425	92.8963	0.0933394	1.08406
196453	1259337833	1427	235.796	90.4198	96.2609	0.0839039	1.18941
194314	243501633	180	230.875	90.5498	82.7169	0.190586	1.16733
194050	519488427	542	205.952	88.6187	104.804	0.0184696	3.10583
194533	677674570	496	333.975	83.8599	48.719	0.0297137	3.60163
194644	196674650	185	272.748	89.0551	88.3426	0.0255656	1.31992
194789	454236538	353	92.1003	74.6909	12.0847	0.172448	1.01778
194790	38426260	40	257.12	93.126	89.09	0.338801	1.24284
195099	137440354	115	126.924	75.8879	29.2406	0.0936533	1.22394
190895	481791479	446	234.743	92.6296	92.6412	0.104027	1.29982

8 TeV							
$2e2\mu$	p.2						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
190736	37061827	38	253.969	92.3989	94.4824	0.0128443	2.27844
190736	80889207	79	221.591	90.9662	90.4846	0.110873	1.38228
195113	622426000	532	574.158	91.0845	89.115	0.0639543	2.78394
195115	18955331	21	302.734	91.1428	95.1174	0.0476598	1.55823
195147	399538705	351	155.362	95.6406	16.1388	0.261031	1.22737
195147	607057396	493	213.273	92.1852	85.6702	0.185222	1.04347
195147	567117841	465	278.447	91.5429	89.8692	0.0181478	0.652115
195304	393582426	300	201.923	90.5094	94.3668	0.118629	2.58056
195378	372893489	296	247.354	94.2563	87.606	0.173904	2.86788
195378	906545322	757	227.8	82.9674	34.5931	0.116868	1.61572
194108	394007716	404	218.513	92.1918	87.9712	0.0467579	0.934887
195397	388415395	249	219.243	88.6571	36.4327	0.352437	0.7142
199428	448557875	357	267.329	99.5046	15.535	0.244619	2.29247
199428	483168057	391	205.333	91.4173	94.6079	0.120066	2.03362
198941	84544694	106	706.435	92.5404	89.149	0.155321	0.312533
198272	26252215	28	92.4004	44.2392	15.6735	0.0887897	1.74378
198212	221918655	337	233.855	92.7522	89.3306	0.0149505	0.958298
198063	167209826	241	225.42	91.7251	92.8388	0.033193	2.47978
198269	177095997	109	597.903	93.5015	87.2907	0.1457	1.01778
198271	631637861	540	130.093	93.7878	23.7979	0.152311	1.29079
199569	346269360	298	169.639	91.6516	36.3723	0.0198942	0.649673
199574	176859852	233	212.747	90.6859	93.8206	0.0281026	3.20442
199608	83178029	101	88.3362	55.7247	12.8662	0.266752	0.979178
199754	504370493	484	226.407	90.6383	95.2009	0.0663063	2.34349
199833	608612940	525	208.362	85.866	30.0917	0.325465	1.87973
199409	1023292406	878	193.614	91.1502	92.8746	0.0521602	0.750135
199833	933684749	833	209.408	93.8088	48.49	0.102575	1.98713
199877	30975991	27	242.589	96.1669	82.3808	0.0483285	1.2744
199877	244791672	224	268.966	87.994	85.3025	0.16246	1.08604
200075	451096187	377	477.144	92.5564	15.9483	0.22394	1.2443
200188	309219190	235	477.681	91.0585	88.4874	0.0171854	0.837584
200243	26763019	69	300.617	92.8844	87.5921	0.0675795	2.01394
200369	48304581	42	269.49	90.877	93.7051	0	1.80454

8 TeV							
$2e2\mu$	p.3						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
200466	153791279	226	125.68	89.6421	34.7116	0.0896414	1.54163
200473	124745478	163	138.481	95.8868	17.1568	0.156966	1.78755
201196	134355928	157	207.125	90.8727	93.6441	0.121044	2.14811
201196	340900299	419	186.465	89.7526	94.0915	0.0191909	2.43298
201202	290344628	316	683.687	90.827	92.8453	0.0452278	2.42712
201174	216745941	335	126.701	86.9969	21.0753	0.260649	2.3167
201097	385556884	292	151.32	96.0799	50.9891	0.259963	2.69193
201278	1072952721	844	274.31	92.7822	84.1885	0.0957463	2.0791
201278	2073195431	2000	217.345	90.3632	94.1846	0.0414625	3.22707
201613	9620116	6	526.05	91.4711	54.699	0.0313386	3.58552
201625	845122892	628	280.491	89.3233	86.7297	0.0673274	2.23325
201707	565245444	438	219.207	93.1661	94.7771	0.0413881	3.65389
201824	201387532	229	382.304	86.2708	60.307	0.125441	1.72142
202016	407076966	362	324.964	89.2263	88.8169	0.0303645	0.394656
201692	107217493	124	339.551	90.7011	90.6436	0.10242	1.99284
202178	739864564	604	186.396	90.904	90.7443	0.0697583	3.14065
202272	769186804	667	213.186	90.446	87.6955	0.0443886	0.222529
202054	319813054	394	182.435	87.1034	85.8245	0.0272659	2.06891
202060	954338141	791	161.8	97.0153	23.5089	0.335569	1.01023
202237	970333313	674	210.18	92.2634	89.0286	0.0651401	1.2437
202237	1409704558	1073	226.789	90.4942	84.5448	0.842756	3.61489
202237	1294338754	973	326.932	95.0329	95.6593	0.161279	2.8631
202299	340486709	251	374.147	89.5685	93.5562	0.152466	1.04475
202299	627007942	449	187.254	85.1208	14.2918	0.289666	1.77505
202093	53125147	75	247.398	92.8115	95.393	0.0569418	1.36549
202178	58141565	92	211.046	91.7888	92.7189	0.137479	1.7604
202178	412076062	328	286.272	89.3509	88.9224	0.0357667	1.25257

8 TeV							
$2e2\mu$	p.4						
run	**evt**	**LS**	$\mathbf{m}_{4l}$	$\mathbf{m}_{Z_1}$	$\mathbf{m}_{Z_2}$	$\mathbf{Iso}_{PF}$	$\mathbf{SIP}_{3D}$
202478	708573825	657	231.352	94.7351	95.0436	0.00406106	1.73563
202478	1037471026	1022	204.003	88.3262	85.1338	0.23289	1.03573
202504	1134909166	926	186.968	92.3663	88.144	0.0197251	2.22183
202504	1394353502	1191	233.992	91.017	88.0479	0.206373	1.31831
202973	779714361	839	578.68	90.674	94.2843	0.204019	2.84661
203002	630169854	469	185.143	87.8879	84.3866	0.113826	2.12575
202469	340674577	252	204.37	91.0633	93.8042	0.0466547	3.52614
193541	115301904	208	201.515	93.3065	93.4356	0.029164	1.48118
194051	6362525	7	146.501	86.1643	16.264	0.02721	2.81312
195165	306811048	235	231.207	92.6821	86.2618	0.0448413	1.02322
195397	1053831791	826	174.191	91.2635	80.8369	0.0681003	2.1088
195552	236444816	139	272.18	74.4404	47.4154	0.325729	2.04152
194119	3127300	4	163.932	89.0351	13.5401	0.398645	3.36962
195948	116169718	56	276.667	91.4367	85.6567	0.195984	1.80508
195950	254206496	240	191.729	118.644	14.2119	0.337061	1.30065
194050	59331872	90	96.3417	69.3812	14.2489	0.294265	2.50523
199008	562986689	447	144.803	89.9337	36.095	0.0761067	0.297342
198954	105011803	115	199.936	86.2476	66.7763	0.0212296	1.22901
199021	1464344864	1327	208.875	92.351	89.2671	0.10956	1.28163
199699	415865014	346	236.384	92.6513	95.9804	0.112399	0.899545
200091	1605749984	1481	129.93	92.3038	34.3452	0.193315	3.25934
199409	195577806	143	161.312	87.5802	53.1775	0.0353285	1.67997
199569	84956243	68	157.769	51.9022	14.0763	0.0507887	2.32299
201174	185135060	284	266.458	92.6508	95.9125	0.0453287	3.33096
200600	892658058	692	210.85	93.9753	86.4106	0.163083	1.20151
201278	1278946974	1033	234.156	93.8466	88.4984	0.126869	0.699264
201602	760006400	558	390.917	87.6443	67.9722	0	1.16498
202044	530312812	388	182.745	90.4881	87.9698	0.181254	2.24187
203002	1582982397	1360	92.081	45.7591	24.6261	0.255742	0.672671
194115	521997420	550	175.825	90.3478	49.485	0.156875	3.97785
194075	76117998	98	184.223	88.6837	93.7367	0.0118162	0.699705

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. 22 (1961) 579-588.

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. 19 (1967) 1264-1266.

[3] A. Salam, *Elementary Particle Physics*, N. Svartholm ed., Almquvist and Wiksell, Stockholm (1968).

[4] M.E. Peskin, D.V. Schroeder, *An Introduction To Quantum Field Theory*, Westview Press, 1995.

[5] S. Weinberg, *The Quantum Theory of Fields*, Cambridge University Press, vol. 1,2,3 (1995, 1996, 2000).

[6] P.S. Wells, *Experimental tests of the standard model*, Eur Phys J C 33 (2004) s01, s5-s20.

[7] J. Ellis, *Outstanding questions: Physics beyond the Standard Model*, Phil. Trans. R. Soc. A (2012) 818-830.

[8] *Particle Data Group* `http://pdg.lbl.gov/`

Phys Rev D86 (2012) 010001.

[9] ALEPH, DELPHI, L3 and OPAL Collaborations, *Search for the Standard Model Higgs Boson at LEP*, CERN-EP/2003-011, Phys. Lett. B 565 (2003) 61-75.

[10] CDF and DØ Collaborations, *Combined CDF and DØ Upper Limits on Standard Model Higgs-Boson Production with up to 8.6 fb^{-1} of Data*, (2011) arXiv:1107.5518v2.

[11] CDF and DØ Collaborations, *Updated Combination of CDF and DØ's Searches for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data*, (2012) arXiv:1207.0449.

[12] A. Hoecker at al., *Status of the global electroweak fit of the Standard Model*, arXiv:0909.0961v2, POS EPS-HEP2009 (2009) 366-380.

[13] M. Goebel, *Status of the global fit to electroweak precisions data*, arXiv:1012.1331v1.

[14] The CMS collaboration, *Combined results of searches for the standard model Higgs boson in pp collisions at* $\sqrt{s} = 7$ *TeV*, arXiv:1202.1488v1.

[15] The ATLAS collaboration, *Combined search for the Standard Model Higgs boson in pp collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Phisycal Review D 86, 032003 (2012).

[16] `http://lhc.web.cern.ch/lhc/lhc-designreport.html`

[17] L. Evans, P. Bryant, et. al., *LHC Machine*, JINST 3 (2008) S08001.

[18] `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt7TeV`

[19] R.K.Ellis et al., *QCD and Collider Physics*, Cambridge University Press.

[20] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1.

[21] ATLAS Technical Design Report, Vol I: *ATLAS DETECTOR AND PHYSICS PERFORMANCE*, CERN/LHCC 99-14, ATLAS TDR 14.

[22] `http://aliweb.cern.ch/Documents/TDR/index.html`

[23] *LHCb technical proposal*, CERN-LHCC-98-004.

[24] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 1.

[25] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 2.

[26] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 3.

[27] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 4.

[28] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 5.

[29] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 6.

[30] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 7.

[31] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 8.

[32] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 9.

[33] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 10.

[34] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 11.

[35] CMS Physics Technical Design report, Vol I: *Detector Performance and Software*, CERN/LHCC 2006-001, CMS TDR 8.1, Cap 12.

[36] CMS Physics Technical Design report, Vol II: *Physics Performance*, CERN/LHCC 2006-021, CMS TDR 8.2.

[37] CMS Physics Technical Design report, Vol II: *Physics Performance*, CERN/LHCC 2006-021, CMS TDR 8.2, Cap 2.

[38] CMS Physics Technical Design report, Vol II: *Physics Performance*, CERN/LHCC 2006-021, CMS TDR 8.2, Cap 3.

[39] CMS Physics Technical Design report, Vol II: *Physics Performance*, CERN/LHCC 2006-021, CMS TDR 8.2, Cap 7.

[40] CMS Physics Technical Design report, Vol II: *Physics Performance*, CERN/LHCC 2006-021, CMS TDR 8.2, Cap 8.

[41] CMS Physics Technical Design report, Vol II: *Physics Performance*, CERN/LHCC 2006-021, CMS TDR 8.2, Cap 9.

[42] CMS Physics Technical Design report, Vol II: *Physics Performance*, CERN/LHCC 2006-021, CMS TDR 8.2, Cap 10.

[43] CMS Physics Technical Design report, Vol II: *Physics Performance*, CERN/LHCC 2006-021, CMS TDR 8.2, Cap 11.

[44] CMS Physics Technical Design report, Vol II: *Physics Performance*, CERN/LHCC 2006-021, CMS TDR 8.2, Cap 13.

[45] CMS Collaboration, *CMS Tracking Performance Results from Early LHC Operation*, Eur. Phys. J. C 70 (2010) 1165-1192.

[46] P. Adzic et al., *Intercalibration of the barrel electromagnetic calorimeter of the CMS experiment at start-up*, J. Inst. 3 (2008) - P100007.

[47] CMS Technical Design report: *The Trigger and Data Acquisition project, Volume I – The Level-1 Trigger*, CERN/LHCC 2000-038, CMS TDR 6.1.

[48] D. R. Stump, *A new generation of CTEQ parton distribution functions with uncertainty analysis*, Prepared for 31st International Conference on High Energy Physics (ICHEP 2002), Amsterdam, The Netherlands, 24-31 Jul 2002.

[49] W. J. Stirling, A. D. Martin, R. G. Roberts, and R. S. Thorne, *MRST parton distributions* AIP Conf. Proc. 747 (2005) 16-21.

[50] CMS Collaboration, *Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update*, CMS PAS SMP-12-008.

[51] CMS Collaboration, *Particle-flow event reconstruction in CMS and performance for jets, taus, and missing E_T*, CMS Physics Analysis Summary CMS-PAS-PFT-09-001 (2009).

[52] CMS Collaboration, *Commissioning of the particle-flow event reconstruction with leptons from J/ψ and W decays at 7 TeV*, CMS Physics Analysis Summary CMS-PAS-PFT-10-003 (2010).

[53] CMS Collaboration, *Performance of CMS muon reconstruction in cosmic-ray events*, JINST 5 (2010) T03022.

[54] CMS Collaboration, *Performance of muon reconstruction and identification in pp collisions at $\sqrt{s} = 7$ TeV*, CMS PAS MUO-10-004.

[55] CMS Collaboration, *Measurement of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP 01 (2011) 0-80.

[56] S. Baffioni et al., *Electron Reconstruction in CMS*, CMS Analysis Note 2009/164 (2009).

[57] e/γ Physics Object Group, *Electron Selection Criteria in 2012*, CMS AN 2012/201 (2012).

[58] The CMS Collaboration, *Isolated photon reconstruction and identification at TeV*, CMS Physics Analysis Summary 2010/006 (2010).

[59] M. Anderson et al., *Review of clustering algorithms and energy corrections in ECAL*, (2010). IN-2010/008.

[60] H. Liu et al., *Conversion reconstruction with tracker-only seeded tracking CMS 900 GeV data*, CMS Analysis Note 2010/039 (2010).

[61] N. Marinelli, *Track finding and identification of converted photons*, CMS Note 2006/005 (2006).

[62] The CMS collaboration, *Search for the standard model Higgs boson in the decay channel $H \to ZZ \to 4l$ in pp collisions at $\sqrt{s} = 7$ TeV*, Physical Review Letters 108 (2012).

[63] N. Amapane et al., *Search for a Higgs boson in the decay channel $H \to ZZ^* \to 4l$*, CMS AN AN-11-387.

[64] N. Amapane et al., *Search for the standard model Higgs boson in the decay channel $H \to ZZ \to 4l$ in pp collisions*, CMS NOTE AN-12-141.

[65] The CMS collaboration, *Updated results on the new boson discovered in the search for the standard model Higgs boson in the $H \to ZZ \to 4l$ channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, CMS PAS HIG-12-041.

[66] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. B659 (2008) 119-126, arXiv:0707.1378.

[67] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP 04 (2008) 005, arXiv:0802.1188.

[68] S.D. Drell and T. Yan, *Partons and Their Applications at High Energies*, Ann. Phys., 66 (1971), 578.

[69] S.D. Drell, T.M. Yan, *Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies*,Phys. Rev. Lett., 25 (1970), 316.

`https://indico.cern.ch/getFile.py/access?contribId=2&resId=0&materialId=slides&confId=123724`

[70] M. Pierini, *talk at the Joint PVT/PO Meeting* (2011):

`https://twiki.cern.ch/twiki/bin/viewauth/CMS/PdwgMain`

[71] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, CERN Report CERN-2011-002, (2011).

[72] ATLAS and CMS Collaborations Collaboration, *LHC Higgs Combination Working Group report*, ATL-PHYS-PUB-2011-11, CMS NOTE-2011/005 (2011).

[73] C. W. J. Campbell, K. Ellis, *MCFM - Monte Carlo for FeMtobarn processes*, (2011).

[74] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP05 2198 [29] 2199 (2006) 026, arXiv:hep-ph/0603175.

[75] J. Alwall et al., *MadGraph/MadEvent v4: The New Web Generation*, JHEP 09 (2007) 028, arXiv:0706.2334.

[76] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP 11 (2007) 070, arXiv:0709.2092.

[77] T. Binoth, N. Kauer, and P. Mertsch, *Gluon-induced QCD corrections to* $pp \to ZZ \to 4l$, in Proceedings of the XVI Int. Workshop on Deep-Inelastic Scattering and Related Topics (DIS'07). 2008. arXiv:0807.0024.

[78] J. Allison et al., *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. 53 (2006) 270.

[79] D. Orbaker (on behalf of the Cms collaboration), *Fast simulation of the CMS detector*, 2010 J. Phys.: Conf. Ser. 219.

[80] R. V. Hogg, A. T. Craig, *Introduction to Mathematical Statistics*, (1978).

[81] J. E. Gaiser, *Appendix-F Charmonium Spectroscopy from Radiative Decays of the J/Psi and Psi-Prime*, Ph.D. Thesis, SLAC-R-255 (1982).

[82] J. J. Olivero, R. L. Longbothum, *Empirical fits to the Voigt line width: A brief review*, Journal of Quantitative Spectroscopy and Radiative Transfer 17 (2) (1977-02) 233-236.

[83] ATLAS and CMS Collaborations, LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011*, ATL-PHYS-PUB/CMS NOTE 2011-11, 2011/005, (2011).

[84] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C 71 (2011) 1-19, arXiv:1007.1727.

[85] G. Cowan, *Statistical Data Analysis*, Oxford Science Publications.

[86] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instrum. Meth. A 434 (1999) 435-443.

[87] A. Read, *Modified frequentist analysis of search results (the CLs method)*, Technical Report CERN-OPEN-2000-005, CERN, (2000).

[88] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B 716 (2012) 1-29.

[89] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B 716 (2012) 30-61.

Ringraziamenti

Per la supervisione e i consigli: Francesco, Francesca, Paolo e Sylvie.

Per il lavoro di squadra e l'immenso aiuto: Nicola.

Per le facilities e l'underlying work: CMS, LHC e il CERN.

Per il lavoro nelle DT: Luigi, Giuseppe, Carlo, Marc.