UNIVERSITÀ DEGLI STUDI MILANO BICOCCA

Facoltà di Scienze Matematiche Fisiche e Naturali

Scuola di Dottorato di Scienze
Corso di Dottorato di Ricerca in Fisica ed Astronomia

First measurement
of the WZ production cross section
with the CMS detector at the LHC

Coordinatore: prof. Giuseppe Chirico
Relatore: prof. Marco Paganoni

Tesi di dottorato di
Arabella Martelli
matricola 055006

Ciclo XXIV
Anno Accademico 2011-2012

ÉCOLE POLYTECHNIQUE

Thèse présentée pour obtenir le grade de

DOCTEUR EN SCIENCES

par

Arabella Martelli

First measurement
of the WZ production cross section
with the CMS detector at the LHC

Soutenue le 30/01/2012 devant le jury composé de:

Examinateur:	J.-C. Brient	École Polytechnique, Palaiseau
	G. Della Ricca	Universitá degli Studi di Trieste
Rapporteur:	G. Hamel de Monchenault	IRFU-SPP, CEA, Saclay
	S. Rahatlou	Universitá La Sapienza, Roma
Directeur de thèse:	C. Charlot	École Polytechnique, Palaiseau
	M. Paganoni	Universitá degli Studi Milano-Bicocca

Contents

Introduction

In this thesis it is presented the activity I performed during these three years of my PhD, working with the Compact Muon Solenoid detector and its first data. CMS is one of the two multipurpose experiments of the Large Hadron Collider, a high performing project developed to cover a very rich and ambitious physics program. To elucidate the nature of the electroweak symmetry breaking for which the Higgs mechanism is presumed to be responsible, to understand the origin of the particle masses, to look for physics beyond the Standard Model and to perform precision measurements of the SM are among the subjects of main interest.

These three years have been a very exciting period for research in the high energy physics field. In 2008, the CMS detector was operated for the first time in the 100 m underground cavern, during the cosmic ray data taking commissioning phase in preparation of the first LHC run. In November 2009 the LHC delivered its first proton-proton collisions at the centre-of-mass energies of 0.9 TeV and successively at 2.36 TeV, just exceeding that of the Tevatron (1.96 TeV) and making the LHC the highest energy collider ever operated. The LHC progressively optimized its performances, to regularly run in 2010 and 2011 at $\sqrt{s} =$7 TeV, delivering $\sim$5 fb^{-1} of integrated luminosity with 2011 operation, already allowing for a very rich research activity and promising fruitful results for the future.

In this context, the inclusive WZ diboson production cross section in pp collisions, $\sigma(pp \rightarrow WZ + X)$ at a centre-of-mass energy of 7 TeV has been measured, looking at the fully leptonic decay channel. The response of the electromagnetic calorimeter (ECAL) is studied through the measurement of the muon stopping power in the lead tungstate constituting the ECAL and electron objects are commissioned with the first data so to be used for any analyses.

The work here presented is organized in six Chapters, the first three to describe the phys-

ical context and the experimental apparatus as well as general concepts on the physics objects reconstruction, while the studies I performed in these three years are detailed in the last three Chapters.

In Chapter 1 the theoretical context and the physic motivations to study the WZ diboson production are presented. The guidelines of the Standard Model of the elementary particles are referred, with attention to the WZ production mechanism.

In Chapter 2 the experimental context is described, in terms of the LHC machine and physics in proton proton collisions. The CMS detector is presented with focus on its potentialities with respect to the studies described in this thesis.

Chapter 3 is dedicated to the physic objects, ingredients of the work performed. In particular the reconstruction algorithms are presented, as well as the elements to qualify the final candidates performances.

In Chapter 4 it is presented the analysis I worked at, to measure the muon stopping power in the lead tungstate ($PbWO_4$) of the electromagnetic calorimeter, dealing with cosmic ray muons in the momentum range from 5 GeV/c to 1 TeV/c recorded by CMS, while fully operating for the first time in the cavern.

Chapter 5 describes the detailed studies I performed to commission the electron object reconstruction, with first data from proton proton collisions. Particular attention was dedicated to the very first stage of the reconstruction process which is the finding of the track seed in the inner tracker layers and which plays a crucial role in the electron reconstruction efficiency.

Chapter 6 is dedicated to the analysis of the WZ physic channel, looked for in electrons and muons only final states. The event selection procedure I developed is described and the measured inclusive production cross section is presented.

The Physics context

Elementary particles and their interactions are today best described by the Standard Model of Particle Physics, a renormalizable quantum field theory, which is in agreement with all the available laboratory observations. High precision measurements carried on in low energy experiments, neutrino experiments as well as collider experiments (SLAC, LEP and the Tevatron) verified the SM predictions with excellent precision (order per mille) [1] [2]. Despite such impressive results, there are several reasons to believe that the SM is not the ultimate theory.

In this Chapter, the Physics picture, context of the WZ diboson search is presented.
At first, the Standard Model is briefly introduced, with focus on the electro-weak interaction. Then, the WZ production process is described and motivations for its study are given. In particular, its importance both within the context of the Standard Model and scenarios beyond is stressed. Existing experimental results are presented in the end.

1.1   The Standard Model of Interactions

One of the most profound insights in theoretical physics is that interactions are dictated by symmetry principles.

The Standard Model is a field theory based on the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group, accounting for three of the interactions between particles[1]. The $SU(3)_C$ symmetry of Quantum Chromodynamics [3], being related to the transformations in the color space, describes the interactions of quarks by means of eight massless colored gauge bosons, the gluons.

---

[1]Gravity is the fourth known fundamental interaction, but it is not currently included into an unified theory.

The $SU(2)_L \times U(1)_Y$ group defines the Electroweak Theory [4, 5, 6], as unification of the Weak Interactions with the Quantum Electrodynamics.

The Standard Model (SM) is currently widely accepted as the theoretical model describing the properties and interactions between the elementary particles.

1.1.1 The elementary particles and Interactions

Elementary particles are classified on the basis of conserved quantities which characterize their features, these are the quantum numbers (a quantified representation of the particles interaction properties).

Fermions with spin 1/2 and bosons with spin 1 are the two categories of elementary particles, the former composing matter and bosons mediating interactions between fermions. According to their properties, fermions are further divided into leptons and quarks and organized into families of identical structure and differing only for the particle mass.

In Tables 1.1 and 1.2, a classification of the elementary particles in the SM is given, in terms of $SU(3)_C \times SU(2)_L \times U(1)_Y$ interactions. The relevant quantum numbers are reported, in particular the electric charge (Q_{EM}), the weak hypercharge (Y), defined by the Gell-Mann and Nishijma relation $Q_{EM} = T_3 + Y/2$, and the third projection component of the weak isospin [7] (T).

$SU(3) \times SU(2) \times U(1)$ Mediator Bosons	EM charge	spin	mass (GeV)	Interaction
g	0	1	0	strong
γ	0	1	0	EM
W^+	+1	1	80.41 ± 0.10	
W^-	-1	1	80.41 ± 0.10	weak
Z^0	0	1	91.187 ± 0.007	

Table 1.1: The building blocks of the Standard Model (gauge bosons) [2]

Besides electric charge, the quarks also have a color charge. This is relevant for their strong interaction, which binds them together inside mesons and baryons (hadrons). Every quark appears in three different color states, belonging to a $SU(3)_C$ triplet, while the leptons are colorless $SU(3)_C$ singlets.

For each fermion, there exists a corresponding anti-particle, with exactly the same quantum numbers as the fermions, but charge reverted[2].

---

[2]All atoms in nature are made of fermions, from the first family

Fermion Generations			Q_{EM}	Y	T^3	Interaction
Leptons						
$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$ $\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}$		0	-1	$+1/2$	weak
			-1	-1	$-1/2$	weak, EM
e_R μ_R τ_R			-1	-2	0	weak, EM
Quarks						
$\begin{pmatrix} u_L \\ d'_L \end{pmatrix}$ $\begin{pmatrix} c_L \\ s'_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b'_L \end{pmatrix}$		$+2/3$	$+1/3$	$+1/2$	weak, EM, strong
			$-1/3$	$+1/3$	$-1/2$	weak, EM, strong
u_R c_R t_R			$+2/3$	$+4/3$	0	weak, EM, strong
d_R s_R b_R			$-1/3$	$-2/3$	0	weak, EM, strong

Table 1.2: The building blocks of the Standard Model (fermions) [2]

Fermionic fields

Fermions are described by quantum field operators, in the form of four components Dirac spinors ψ. By introducing the adjoint spinor $\bar{\psi} = \psi^\dagger \gamma^0$, the Lagrangian associated to a free fermion of mass m and the derived Dirac equation of motion are

$$\mathcal{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi \qquad (i\gamma^\mu \partial_\mu - m)\psi = 0 \qquad (1.1)$$

For further convenience, the Weyl spinor representation can be introduced. By projecting[3] the Dirac spinor into its left-handed (ψ_L) and right-handed (ψ_R) components, as

$$\psi = \psi_L + \psi_R = \begin{pmatrix} \psi'_L \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \psi'_R \end{pmatrix} \qquad (1.2)$$

$$\text{being} \quad \psi_L = P_L\psi = \frac{1}{2}\left(1 - \gamma^5\right)\psi, \qquad \psi_R = P_R\psi = \frac{1}{2}\left(1 + \gamma^5\right)\psi \qquad (1.3)$$

Bosonic fields

A free scalar boson of mass m is represented by a complex scalar field ϕ. Its dynamic is described by the Klein-Gordon Lagrangian, from which its equation of motion can be

---

[3]$\psi_L = P_L\psi = \frac{1}{2}\left(1 - \gamma^5\right)\psi$ and $\psi_R = P_R\psi = \frac{1}{2}\left(1 + \gamma^5\right)\psi$ are the projection operators used, for which the Weyl representation of the γ matrices is adopted, with $\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}$ and $\gamma^5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

derived

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - m^2 \phi^\dagger \phi, \qquad \left(\Box + m^2\right)\phi = 0 \qquad (1.4)$$

For a free vector exchange boson A_μ of mass m, the dynamic is instead described by the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m^2 A^\mu A_\mu \qquad (1.5)$$

Interactions and symmetries

The conservation of physics quantities is deeply connected to the symmetries of a physical system, concept expressed within the field theory by the *Noether's theorem* [8].

In particular, since any continuos transformation, preserving the invariance of the Lagrangian that describes a physical system, corresponds to a conserved current, each interaction is associated to a symmetry and can be described by the associated current.

Within the Standard Model, the electromagnetic theory is described by an $U(1)$ symmetry group. The corresponding conserved current is $J_\mu^{em} = \bar{\psi}\gamma_\mu Q\psi$, where Q is the generator of the symmetry group and represents the conserved charge associated to the current.

The weak charged currents are represented by $J_\mu^\pm = \bar{\psi}_L \gamma_\mu \tau_\pm \psi_L$ and couple the only ψ_L components, since left-handed fermions are doublets under the weak $SU(2)_L$ interaction, while the right-handed ones behave as singlets.

The J_μ^{NC} weak neutral current couples operators of opposite charge, both right-handed and left-handed, thus it cannot be directly identified with the third current of the $SU(2)_L$ triplet. Of course J_μ^{NC} cannot either be identified with the electromagnetic current.

The idea to preserve the $SU(2)_L$ symmetry group description of the weak interactions is to think to a $SU(2)_L \times U(1)_Y$ gauge theory, as the simplest unification of the parity violating weak force and the parity conserving electromagnetic force.

Basically, the idea is to involve the J_μ^{em} electromagnetic current and to introduce a weak hypercharge current J_μ^Y, described by the abelian factor $U(1)_Y$.

J_μ^{em} and J_μ^{NC} result from the combination of the two new orthogonal components, with J_μ^3 to complete the weak isospin triplet J_μ^i and $J_\mu^Y = \bar{\psi}\gamma_\mu Y\psi$. The hypercharge quantum numer Y is accordingly defined by the Gell-Mann and Nishijma relation $Q = T_3 + Y/2$.

1.1.2 The ElectroWeak Interaction

To describe theories with interactions between fields, local gauge symmetries have to be considered (see Appendix A.1 for a detailed example).

In general, the theoretical picture can be summarized as following. A Lagrangian symmetry is preserved locally under a gauged transformation G, if a covariant derivative $\mathcal{D}_\mu$ is used instead of the usual derivative ∂_μ. This can be done by introducing a set of real vector fields A^a_μ, such that $\mathcal{D}_\mu = \partial_\mu - igA^a_\mu T^a$, where g is the gauge coupling, and T^a are the generators of the group G. The interacting fields would directly emerge in the theory.

To express the kinetic term for such gauge fields a gauge tensor $F_{\mu\nu}$ antisymmetric in its two spatial indexes is introduced.

In the case of the electroweak interaction, the four fields corresponding to the $SU(2)_L \times U(1)_Y$ group are the three W^i_μ ($i = 1, 2, 3$) related to the $SU(2)_L$ symmetry and the B_μ corresponding to the $U(1)_Y$ group.

The covariant derivative

$$\mathcal{D}_\mu \equiv (\partial_\mu - ig\vec{T}\vec{W}_\mu - ig'\frac{Y}{2}B_\mu) \tag{1.6}$$

is introduced, where g and g' are the coupling constants related to $SU(2)_L$ and $U(1)_Y$, respectively. The electroweak Lagrangian includes kinematic terms for the gauge fields like

$$-\frac{1}{4}W^i_{\mu\nu}W^{\mu\nu}_i - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.7}$$

where the field strength tensors are[4]

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk}W_{\mu j}W_{\nu k} \tag{1.8}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.9}$$

and ϵ^{ijk} is a total antisymmetric tensor. Within this local gauge invariance formalism,

---

[4]The extra term in $W^i_{\mu\nu}$ is due to the non Abelian SU(2) group. In such a case, the local invariance of the gauge fields is guaranteed with an additional term of self-interaction, due to the non zero commutator, which is instead present for Abelian groups as U(1).

the electroweak Lagrangian is expressed in the form

$$\mathcal{L}_{YM} = \mathcal{L}_{kin} + \mathcal{L}_{charged} + \mathcal{L}_{neutral} \quad = \tag{1.10}$$

$$- \quad \frac{1}{4}W^i_{\mu\nu}W^{\mu\nu}_i - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}$$

$$+ \quad \bar{\psi}_L\gamma^\mu(\partial_\mu - ig\vec{T}^i\vec{W}^i_\mu - ig'\frac{Y}{2}B_\mu)\psi_L$$

$$+ \quad \bar{\psi}_R\gamma^\mu(\partial_\mu - ig'\frac{Y}{2}B_\mu)\psi_R$$

Once worked out, the above scripture suggests to perform a rotation by an angle θ_W in the neutral sector, to reveal the role of new physical vector fields Z_μ and A_μ (and also $W^\pm_\mu$) in terms of

$$Z_\mu = \cos\theta_W W^3_\mu - \sin\theta_W B_\mu \tag{1.11}$$

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W^3_\mu \tag{1.12}$$

$$\text{(and also)} \quad W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp W^2_\mu) \tag{1.13}$$

As outcome of such electro-weak sectors mixing

$$\mathcal{L}_{neutral} = \psi\gamma^\mu(g\sin\theta_W T_3 + \frac{Y}{2}g'\cos\theta_W)\psi A_\mu \tag{1.14}$$

$$+ \quad \psi\gamma^\mu(g\cos\theta_W T_3 - \frac{Y}{2}g'\sin\theta_W)\psi Z_\mu$$

$$J^{em}_\mu = -e(\bar{\psi}_R\gamma_\mu\psi_R + \bar{\psi}_L\gamma_\mu\psi_L) = e\bar{\psi}\gamma_\mu Q\psi \tag{1.15}$$

$$J^{NC}_\mu = \bar{\psi}\left(-\frac{ig}{\cos\theta_W}\right)\frac{\gamma_\mu}{2}(T^3 - 2\sin^2\theta_W Q - \gamma^5 T^3)\psi \tag{1.16}$$

$$- \quad \frac{ig}{\cos\theta_W}(J^3_\mu - \sin^2\theta_W J^{em}_\mu)$$

The electroweak unification can be represented by the relation 1.17, with the elementary charge e linked through the weak mixing angle θ_W to the coupling constants of the weak isospin and hypercharge

$$e = g\sin\theta_W = g'\cos\theta_W \tag{1.17}$$

The Triple Gauge Couplings

The Lagrangian in equation 1.10 can be expressed [9] in terms of the physical fields $W^\pm_\mu$,

Z_μ and A_μ at the first order as

$$\mathcal{L}_{YM} = \quad - \quad \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} - \frac{1}{2}W^+_{\mu\nu}W^{\mu\nu}_- \tag{1.18}$$
$$+ \quad ig\sin\theta_W(W^+_{\mu\nu}W^\mu_- A^\nu - W^-_{\mu\nu}W^\mu_+ A^\nu + F_{\mu\nu}W^\mu_+ W^\nu_-)$$
$$+ \quad ig\cos\theta_W(W^+_{\mu\nu}W^\mu_- Z^\nu - W^-_{\mu\nu}W^\mu_+ Z^\nu + Z_{\mu\nu}W^\mu_+ W^\nu_-) + ...$$

where
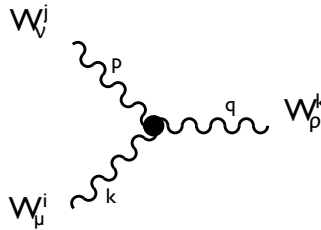
$$F_{\mu\nu} \quad = \quad \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.19}$$
$$Z_{\mu\nu} \quad = \quad \partial_\mu Z_\nu - \partial_\nu Z_\mu \tag{1.20}$$
$$W^\pm_{\mu\nu} \quad = \quad \partial_\mu W^\pm_\nu - \partial_\nu W^\pm_\mu \tag{1.21}$$

In the above parametrization, trilinear vertices appear for the W^+W^-Z and W^+W^-A fields. Such trilinear gauge couplings are exclusive of the $SU(2)_L$ gauge structure and come from its kinetic term.

It's the effect of the extra term in equation 1.8

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + gW^i_\mu \times W^i_\nu \tag{1.22}$$

to generate gauge field self-interactions. The $W^i_{\mu\nu}$ non Abelian fields act by themselves as sources and, at the tree level, this term leads to trilinear vertices as shown in Figure 1.1.

$$= ig\epsilon^{ijk}[g^{\mu\nu}(k-p)^\rho + g^{\nu\rho}(p-q)^\mu + g^{\rho\mu}(q-k)^\nu]$$

Figure 1.1: The general trilinear $W^1W^2W^3$ vertex and the corresponding Feynman rules as given by the self-coupling term of the Yang-Mills field.

The $W^1W^2W^3$ coupling reflects in the mixing between both the W^+, W^- bosons with a single V neutral boson field, with $V = \gamma, Z$

$$\mathcal{L}^{W^+W^-V}_{VEV} = \quad + \quad ie[A^\mu(W^+_{\mu\nu}W^\nu_- - W^-_{\mu\nu}W^\nu_+) + F_{\mu\nu}W^\mu_+ W^\nu_-) \tag{1.23}$$
$$+ \quad ie\cot\theta_W[Z^\mu(W^+_{\mu\nu}W^\nu_- - W^-_{\mu\nu}W^\nu_+) + Z_{\mu\nu}W^\mu_+ W^\nu_-)$$

1.2 The SM Spontaneous Symmetry Breaking

The theory presented up to now describes massless fermions and fields, while experimental results prove the fermions, and the $W^\pm$, Z bosons to be massive.

In order to preserve the symmetry of the theory built, mass terms need to be generated in a gauge invariant way. A brilliant solution to this problem is provided by the spontaneous symmetry breaking mechanism (see Appendix A.2 for a detailed treatment).

In brief, a symmetry is spontaneously broken when the vacuum expectation value of the transformed field operators (*vev*) is non zero, namely when the point of minimum in the potential term of the Lagrangian doesn't lie in zero. Essentially the symmetry breaking mechanism reveals the hidden symmetry of the theory.
With this mechanism, an additional degree of freedom emerges for each field spontaneously broken. Such an extra term corresponds to a massless scalar particle named as the *Goldstone boson*, which is in turn re-absorbed through the Higgs mechanism to provide a mass term for the broken field.

1.2.1 The Higgs field in the EW theory

With respect to the approach described in A.2.1 for a $U(1)$ symmetry group, in the slightly more complicated non abelian case of the Standard Model, masses need to be generated for the three gauge bosons $W^\pm$ and Z, while keeping the photon massless and preserving the QED as an exact symmetry.

Therefore, at least three degrees of freedom for the scalar fields are needed. The simplest choice is to introduce four real scalar fields ϕ_i, arranged into an $SU(2)$ doublet, with weak hypercharge $Y = 1$

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} \qquad \text{with} \qquad \begin{cases} \Phi^+ \equiv (\phi_1 + i\phi_2)/\sqrt{2} \\ \Phi^0 \equiv (\phi_3 + i\phi_4)/\sqrt{2} \end{cases} \tag{1.24}$$

To generate gauge boson masses, the correspondent Lagrangian

$$\mathcal{L}_H = (\mathcal{D}_\mu \Phi)^\dagger (\mathcal{D}_\mu \Phi) - V(\Phi^\dagger \Phi) = (\mathcal{D}_\mu \Phi)^\dagger (\mathcal{D}_\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 \tag{1.25}$$

with $\mu^2 < 0$ and $\lambda > 0$ has to be expanded around a proper vacuum expectation value.

The appropriate choice is

$$\Phi_0 \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \qquad \text{with} \qquad v = \sqrt{-\frac{\mu^2}{2\lambda}} \tag{1.26}$$

Any choice of Φ_0 which breaks a symmetry will inevitably generate a mass for the corresponding gauge boson. However, if the vacuum Φ_0 is still left invariant under some sub-group of gauge transformations, then the gauge bosons associated with this sub-group will remain massless. The choice of Φ_0 described in equation 1.26, with $T = \frac{1}{2}$, $T^3 = -\frac{1}{2}$ and $Y = 1$, breaks both $SU(2)$ and $U(1)_Y$ gauge symmetries.

At the same time since $Q = T^3 + Y/2 = 0$, Φ_0 is neutral and the $U(1)_{em}$ symmetry with generator Q remains unbroken, that is

$$\Phi_0 \to \Phi_0' = e^{i\alpha(x)Q}\Phi_0 = \Phi_0, \qquad \text{for any value of } \alpha(x) \tag{1.27}$$

The vacuum is thus invariant under $U(1)_{em}$ transformations, and the photon remains massless.

By following the schema previously sketched, the field Φ can be expressed in terms of four fields $\theta_{1,2,3}(x)$ and $H(x)$ results at first order

$$\Phi = e^{i\theta_a(x)\tau(x)^a/v} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + H) \end{pmatrix} \tag{1.28}$$

In particular the above parametrization can be further simplified, since the three θ_i fields can be gauged away by an $SU(2)$ transformation.

By exploiting the resulting *vev* form, the scalar potential reveals in

$$V = \frac{1}{2}(2\lambda v^2)H^2 + \lambda v H^3 + \frac{1}{4}\lambda H^4 \tag{1.29}$$

and the broken Lagrangian 1.25 can be expressed in function of the only physical fields

$W^\pm$, Z, A_μ (W^i, B) and H

$$\mathcal{L}_{H_{broken}} = \left\{ \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2}2v^2\lambda H^2 \right\} + \left\{ -\frac{1}{3!}6v\lambda H^3 - \frac{1}{4!}6\lambda H^4 \right\} \tag{1.30}$$

$$+ \left\{ \frac{1}{2}\frac{v^2 g^2}{4}W_\mu^{-\dagger}W^{-\mu} + \frac{1}{2}\frac{v^2 g^2}{4}W_\mu^{+\dagger}W^{+\mu} \right\}$$

$$+ \left\{ \frac{1}{2}\frac{v^2(g^2+g'^2)}{4}\left(\frac{gW_\mu^3 - g'B^\mu}{\sqrt{g^2+g'^2}}\right)^2 + 0\left(\frac{g'W_\mu^3 + gB^\mu}{\sqrt{g^2+g'^2}}\right)^2 \right\}$$

$$+ \left\{ \frac{1}{4}\left(2vH + H^2\right)\left[g^2 W_\mu^- W^{+\mu} + \frac{1}{2}(g^2+g'^2)\left(\frac{W_\mu^3 - g'B^\mu}{\sqrt{g^2+g'^2}}\right)^2 \right] \right\}$$

The term bilinear in the Higgs field provides $m_H = \sqrt{2\lambda v^2}$. Such a mass term is not predicted by the theory, since the coupling λ is a free parameter. The *vev* value can be expressed as a function of the Fermi constant $v = \sqrt{\frac{1}{G_F\sqrt{2}}} \simeq 246.22$ GeV.

1.2.2 Fermionic sector

The above Higgs mechanism can be exploited to generate mass terms for fermions. In particular, such masses represent the strength of the interaction between the Higgs doublet and the fermion fields, which is weighted by means of Yukawa couplings, specific for each charged lepton and quark

$$\mathcal{L}_{Yukawa} = \mathcal{L}_{lep} + \mathcal{L}_{quarks} \tag{1.31}$$

$$= -\sum_{l=e,\mu,\tau} G_l \left[(\bar{l}_L\Phi)l_R + \bar{l}_R(\Phi^\dagger l_L) \right]$$

$$- \sum_{q_u=u,c,t} G_{q_u}[\bar{q}_{uL}\Phi q_{uR} + h.c.] - \sum_{q_d=d,s,b} G_{q_d}[\bar{q}_{dL}\tilde{\Phi}q_{dR} + h.c.]$$

The l_L, q_{uL} and q_{dL} represent the left chiral weak isospin doublets of leptons and quarks, l_R denotes the right chiral lepton singlet and q_{uR} (q_{dR}) represent the up-like (down-like) right chiral quark singlets respectively.

The Yukawa Lagrangian developed around the *vev* 1.28 results

$$
\begin{aligned}
\mathcal{L}_{Yukawa}^{VEV} = \ & - \sum_{l=e,\mu,\tau} \frac{G_l}{\sqrt{2}} [v(\bar{l}_L l_R + \bar{l}_R l_L) + (\bar{l}_L l_R + \bar{l}_R l_L)H] && (1.32) \\
& - \sum_{q_u=u,c,t} \frac{G_{q_u}}{\sqrt{2}} [v(\bar{q}_{uL} q_{uR} + \bar{q}_{uR} q_{uL}) + (\bar{q}_{uL} q_{uR} + \bar{q}_{uR} q_{uL})H] \\
& - \sum_{q_d=d,s,b} \frac{G_{q_d}}{\sqrt{2}} [v(\bar{q}_{dL} q_{dR} + \bar{q}_{uR} q_{dL}) + (\bar{q}_{dL} q_{dR} + \bar{q}_{dR} q_{dL})H] = \\
& - \sum_{l=e,\mu,\tau} \frac{G_l}{\sqrt{2}} \left[v\bar{l}l + \bar{l}lH \right] \\
& - \sum_{q_u=u,c,t} \frac{G_{q_u}}{\sqrt{2}} \left[v\bar{q}_u q_u + \bar{q}_u q_u H \right] - \sum_{q_d=d,s,b} \frac{G_{q_d}}{\sqrt{2}} \left[v\bar{q}_d q_d + \bar{q}_d q_d H \right]
\end{aligned}
$$

where the G_l, G_{q_u}, G_{q_d} Yukawa couplings are chosen so as to generate the required lepton mass terms. Once again such parameters are not given by the model, but need to be determined experimentally.

$$
m_l = \frac{G_l v}{\sqrt{2}} \qquad m_{q_u} = \frac{G_{q_u} v}{\sqrt{2}} \qquad m_{q_d} = \frac{G_{q_d} v}{\sqrt{2}} \tag{1.33}
$$

An additional freedom for the mass eigenstates is given by the fact that the terms G_l, G_{q_u}, G_{q_d} do not need to be diagonal.

In the specific case of the quark sector, the weak eigenstates (d, s, b) do not correspond to the mass ones (d', s', b'), but are connected by a 3×3 Cabibbo-Kobayashi-Maskawa (CKM) matrix, in the form

$$
\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = V_{CKM} \begin{bmatrix} d \\ s \\ b \end{bmatrix} \qquad \text{where} \qquad V_{CKM} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix}
$$

In case the Higgs scalar boson is observed, given its explicit direct coupling to the fermions, an important test to verify the Higgs mechanism described in the SM would be the investigation of the Higgs decay amplitude into the daughter particles, which is indeed expected to be proportional to the squared mass of the decayed products.

1.2.3 Bosonic sector

As for fermions, the Higgs mechanism also works to provide masses for gauge bosons. By identifying in the broken Lagrangian 1.30 the relevant terms produced by the coupling to the Higgs field, the mass terms and the couplings for the gauge fields can be worked out

$$
\begin{aligned}
\mathcal{L}_{Bosons}^{VEV} &= (D_\mu \Phi)^\dagger (D_\mu \Phi) & (1.34) \\
&= (\Phi)^\dagger (g\vec{T}\vec{W}^\mu + g'\frac{Y}{2}B^\mu)(g\vec{T}\vec{W}_\mu + g'\frac{Y}{2}B_\mu)(\Phi) \\
&= \frac{1}{2}\partial^\mu H \partial_\mu H + [\frac{1}{4}g^2 W_\mu^+ W_-^\mu + \frac{1}{8}(g^2 + g'^2)Z^\mu Z_\mu](H+v)^2
\end{aligned}
$$

By picking up the terms which are bilinear in the new fields, it emerges that the $W^\pm$ and Z bosons have acquired masses, while the photon is still massless

$$
m_W^2 = \frac{1}{4}v^2 g^2, \qquad m_Z^2 = \frac{1}{4}(g^2 + g'^2)v^2, \qquad m_A = 0 \qquad (1.35)
$$

$$
\text{with} \qquad m_W = m_Z \cos(\theta_W) \quad \text{and} \quad \rho = \frac{m_W^2}{m_Z^2 \cos^2 \theta_W} \qquad (1.36)
$$

Through the measured values $m_Z = 91.1875 \pm 0.0021$ GeV/c^2 and $m_W = 80.399 \pm 0.025$ GeV/c^2 ($\sin\theta_W \simeq 0.23$), the ρ parameter is found in very good agreement (within a per mille precision) to its tree-level computation in the Standard Model ($\rho = 1$).

As for fermions, the eventual measurement of the Higgs decay amplitude into gauge bosons will be an important test of the SM Higgs mechanism.

1.3 The WZ diboson production

The $W^\pm Z$ final state can only be accessed through charged initial state processes. Its first observation [10] was performed at the Tevatron hadron collider, with the CDF detector, based on 1.1/fb of integrated luminosity from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV.

At hadron colliders, the dominant $W^\pm Z$ production mechanism is from quark-antiquark interactions, with higher order contributions from quark-gluon and gluon-gluon processes, this last being strongly suppressed by charge conservation. The leading-order (LO) Feynman diagrams for the $q\bar{q}' \to W^\pm Z$ process in the Standard Model are shown in Figure 1.2.

The most updated next-to-leading order cross section computation [11] for the WZ
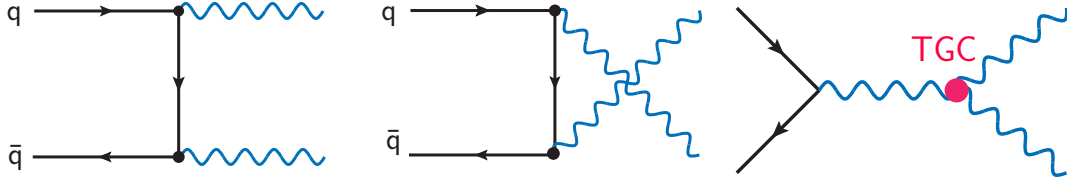
Figure 1.2: The Standard Model tree-level Feynman diagrams for $W^\pm Z$ production through the t,u and s-channel are shown. The s-channel diagram on the right, contains the WWZ trilinear coupling.

pair production at the LHC and Tevatron is reported in Table 1.3. This particular NLO calculation takes into account contributions from singly resonant diagrams, that can be significant when one of the bosons is off-shell, as well as both the contributions of a Z and a virtual photon, when considering the decay to charged leptons.

	Tevatron $p\bar{p}$	LHC pp	
	(1.96 TeV)	(7 TeV)	(14 TeV)
$\sigma_{NLO}(W^+Z)$ [pb]	1.73 ± 0.20	$11.88^{+0.65}_{-0.5}$	$31.50^{+1.23}_{-0.95}$
$\sigma_{NLO}(W^-Z)$ [pb]	1.73 ± 0.20	$6.69^{+0.37}_{-0.29}$	$20.32^{+0.79}_{-0.63}$
$\sigma_{NLO}(W^\pm Z)$ [pb]	3.46 ± 0.30	$18.57^{+0.75}_{-0.58}$	$51.82^{+1.46}_{-1.14}$

Table 1.3: *NLO* cross sections for WZ production at in $p\bar{p}$ and pp collisions[11]. Renormalisation and factorization scales are set equal to the average mass of the W and Z ($\mu_R = \mu_F = (M_W + M_Z)/2$). Upper and lower percentage deviations are obtained by varying the scales around the central scale by a factor of two. The vector boson are kept on-shell, with no decays included.

While the production cross section is expected to be the same for W^+Z and W^-Z at the Tevatron, this is not the case for the LHC, where the expected ratio of $\sigma_{NLO}(W^-Z)/\sigma_{NLO}(W^+Z)$ varies between 0.56 (for $\sqrt{s} =$7 TeV) and 0.65 (at 14 TeV) [11]. Indeed, in proton-proton collisions ($p \simeq uud$), the predominance of u quarks enhances the W^+Z production through $u\bar{d}'$ type interactions, with anti-quarks taken among sea-quarks.

The WZ process can be looked for in the fully leptonic ($l\nu + ll$), semi-leptonic ($jj + ll$, or $l\nu + jj$) or pure hadronic final state ($jj + jj$). By looking at the decay branching ratios of Zs and Ws (reported in Table 1.4), a search limited to a fully leptonic final state would profit from a 0.3% of the inclusive cross section (per single channel). At the same time, the big advantage of such a choice is to provide a very clean signature, which allows to gain in discriminating power between the signal and the hadronic/multi-jet backgrounds, dominating the production at the LHC.

Cross sections in Table 1.3 are plotted as a function of energy for both pp and $p\bar{p}$

$W^\pm$	$BR\%$		Z	$BR\%$
$e^+\nu_e$	10.75±0.13		e^+e^-	3.363 ± 0.004
$\mu^+\nu_\mu$	10.57±0.15		$\mu^+\mu^-$	3.366 ± 0.007
$\tau^+\nu_\tau$	11.25±0.20		$\mu^+\mu^-$	3.370 ± 0.008
$hadrons$	67.60±0.27		$hadrons$	69.91 ± 0.06
$l^+\nu_l$	10.80±0.09		l^+l^-	3.3658 ± 0.0023
			$\nu_l\bar{\nu}_l$	20.00 ± 0.06

Table 1.4: Decay branching ratios for the W and Z gauge bosons, from [2].

collisions on the left in Figure 1.3. On the right, the boson production rates are compared for pp collisions only, taking into account the decay branching ratios of the Ws and Zs into one species of leptons (see Tab. 1.4).
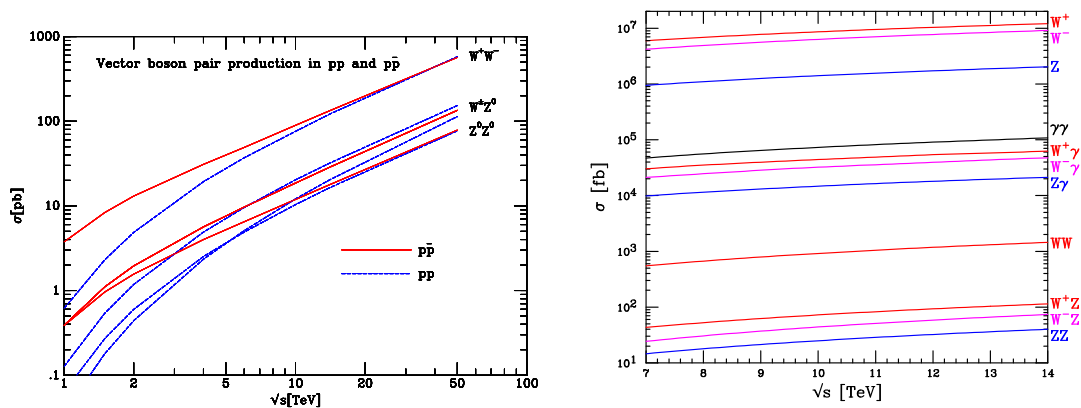
Figure 1.3: [11]On the left: energy dependence of vector boson pair production cross sections, in pp and $p\bar{p}$ collisions. On the right: NLO boson production in pp collisions, with decay branching ratios of the Ws and Zs into one species of leptons included.

With the CMS detector at the LHC, the first WZ event was observed with $\sim$30 pb^{-1} of integrated luminosity, looking for the only leptonic decays into electrons and muons ($\sigma(WZ) \times BR \simeq 0.07$ pb per single channel).

The study of diboson processes is valuable to test the Standard Model through the constraining of the SM parameters. At the same time, high precision measurements could yield information on New Physics processes.

1.3.1 WZ production beyond the Standard Model

Although the Standard Model has been a very successful theory, in very good agreement with most of the experimental results, it does not answer all the questions concerning

fundamental interactions. It does not incorporate the physics of general relativity and it does not provide an explanation for dark matter and dark energy and for the CP violation, responsible for the matter-antimatter asymmetry observed in our current universe.

Also, the existence of exactly three identical families of fermions whose masses span over a six order magnitude range (m_e =0.511 MeV to m_t =172 GeV) as well as the evidence of massive neutrinos, provided by the observed oscillations patterns, have no explanation within the SM.

The Standard Model is today considered an effective theory, the low energy limit (at the electroweak scale $\simeq 10^2$ GeV) of another more fundamental one (at the Planck scale $\simeq 10^{19}$ GeV). Nevertheless the Higgs boson, one of the fundamental pieces of the SM, is still unobserved and furthermore some technical aspects of the theory seem not natural, as the so-called hierarchy problem. In this effective theory, the radiative corrections to the Higgs mass are proportional to the scale of the new Physics, thus a very precise tuning of the parameters is needed for the self-consistency of the theory.

Given such a picture, it is widely believed that new Physics not described by the Standard Model should exist. Such new theories are considered if keeping in agreement with the current experimental data, they extend the SM, giving solution to some of the raised puzzles (see [12] for some references).

Beyond the Standard Model scenarios might account for new particles interacting with SM ones or very similar to the SM partners. Through direct production or loop effects, new particles should manifest modifying the peculiar Standard Model couplings, which are fixed by the specific $SU(2)_L \times U(1)_Y$ gauge.

Concerning the WZ channel, examples of common BSM models affecting its production are:

- Supersymmetry, with multi-lepton final states. The specific case of $MSSM$ [13] in particular accounts for $chargino + neutralino \rightarrow 3l + MET$;

- exotic WZ resonances [14], from heavy $W' \rightarrow WZ$ decays and Technicolor ρ_T (or a_T) $\rightarrow WZ$;

- Composite models [15], little Higgs phenomenology [16] and Higgsless models [17] as other examples;

Any deviation from the SM predictions would be evidence for New Physics. Furthermore, diboson events can be produced as decay products of the Higgs boson, and their polarization state might be of help in the determination of the Higgs spin and its separation from the diboson background continuum. WZ production is also a source of SM

background events for the same Higgs search, in the case of leptonic decays, when one of the leptons is missed.

A main effect of anomalous couplings in diboson production is the change in the cross section. As shown in Figure 1.4 for the W^-Z channel, non SM couplings enhance the weigh of each partonic process with the consequent growth of the $\sigma(WZ)$.
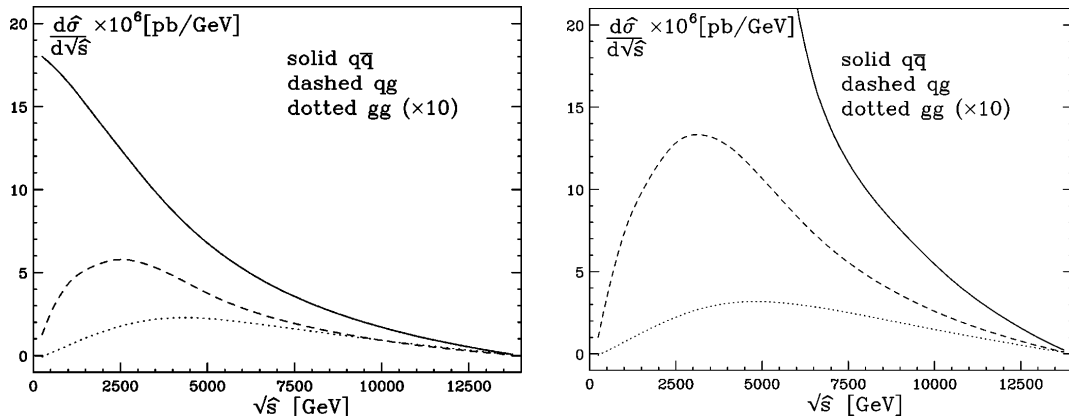
Figure 1.4: [18] Partonic cross sections for W^-Z: $\sigma(q\bar{q})$, $\sigma(qg)$, $\sigma(gg)$. On the left for standard SM couplings. On the right for $g_1^\gamma = 1$, $g_1^Z = 1.13$, $k_\gamma = 1.2$, $k_Z = 1.07$, $\lambda_\gamma = \lambda_Z = 0.1$, $\Lambda = 2$ TeV.

Also the kinematic of the final state products results consequently modified, in particular an hardest spectrum is expected for the gauge bosons produced with anomalous TGC. The example for the leading lepton p_T in W^-Z production is shown in Figure 1.5.

As well as through the production cross section measurement, diboson physics is a test of the Standard Model also through the direct measurement of the TGCs.
In this context, the WZ final state is of particular interest since it's the only production sensitive to the exclusive WWZ amplitude, without interference from $WW\gamma$.

1.3.2   The Triple Gauge Couplings formalism

As discussed in the previous Section (1.1.2), TGCs are a consequence of the non-Abelian nature of the electro-weak sector $SU(2)_L \times U(1)_Y$ and are univocally predicted.
So far, the precision measurements at previous experiments (mainly in the fermionic sector [20]) exclude only drastic modifications of such a simple gauge structure.

Experimental measurements of the TGCs can be sensitive to new phenomena at high energies, that would require more energy or luminosity to be observed directly.

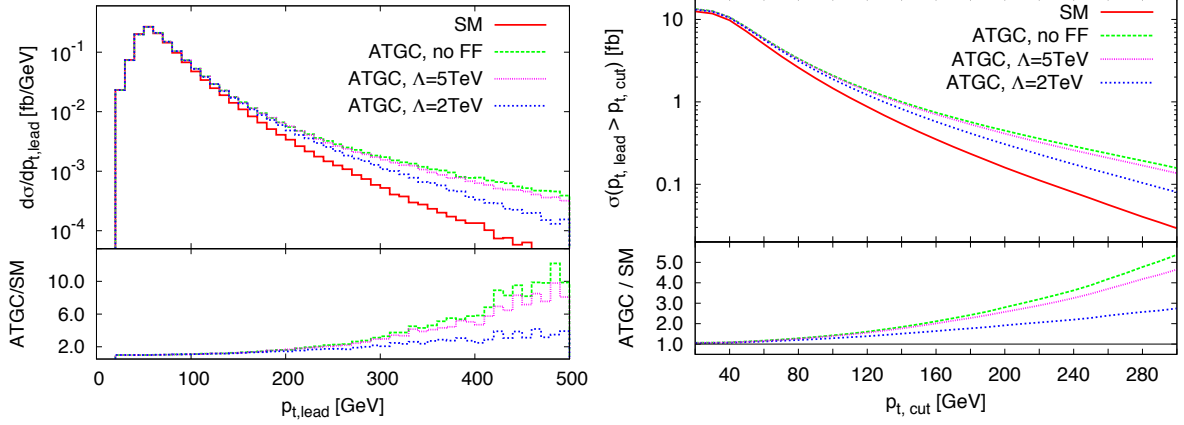Figure 1.5: [19] On the left: the p_T spectrum for the leading lepton in W^-Z production. On the right: the integrated cross section as a function of the cut on the leading lepton p_T. The results are shown using Standard Model couplings (red), as well as with anomalous couplings. The cases of no form factors, $\Lambda = 2$ TeV and $\Lambda = 5$ TeV are presented.

In this context, a general formalism can be used to describe the couplings of the charged to neutral vector bosons ($WW\gamma$, WWZ), obtained in the Standard Model electro-weak mixing, in the form of a most general effective Lagrangian [21]

$$
\begin{aligned}
\mathcal{L}_{WWV} = -ig_{WWV}[\quad &+ \quad g_1^V(W_{\mu\nu}^- W^{+\mu}V^\nu - W_{\mu\nu}^+ W^{-\mu}V_\nu) + k_V W_\mu^+ W_\nu^- V^{\mu\nu} \quad (1.37) \\
&+ \quad \frac{\lambda_V}{m_W^2} V^{\mu\nu} W_\mu^{+\rho} W_{\rho\mu}^- + ig_4^V W_\mu^- W_\nu^+ (\partial^\mu V^\nu - \partial^\nu V^\mu) \\
&+ \quad ig_5^V \epsilon^{\mu\nu\rho\sigma}[W_\mu^-(\partial_\rho W_\nu^+) - (\partial_\rho W_\mu^-)W_\nu^+]V_\sigma \\
&- \quad \frac{\tilde{k}_V}{2} W_\mu^- W_\nu^+ \epsilon^{\mu\nu\rho\sigma} V_{\rho\sigma} - \frac{\tilde{\lambda}_V}{2m_W^2} W_{\rho\mu}^- W_\nu^{+\mu} \epsilon^{\nu\rho\alpha\beta} V_{\alpha\beta}]
\end{aligned}
$$

where $V = Z^0, \gamma$, m_W is the nominal $W^\pm$ boson mass and the overall coupling strength g_{WWV} is fixed in the Standard Model $SU(2)_L \times U(1)_Y$ gauge to $g_{WW\gamma} = e$ and $g_{WWZ} = e\cot\theta_W$.

For the photon couplings ($V = \gamma$) in equation 1.37, the first three terms are related to the electric charge (Q_W), the magnetic dipole momentum (μ_W) and the electric quadrupole momentum (q_W) of the $W^\pm$ bosons through

$$
Q_W = eg_1^\gamma, \qquad \mu_W = \frac{e}{2m_W}(g_1^\gamma + k_\gamma + \lambda_\gamma) \qquad q_W = -\frac{e}{m_W^2}(k_\gamma - \lambda_\gamma) \qquad (1.38)
$$

The symmetry properties of the couplings in eq. 1.37 are reported in Table 1.5.

Within the Standard Model, at tree level, the couplings are set to $g_1^\gamma = g_1^Z = k_\gamma = k_Z = 1$, while all the others (referred to as anomalous couplings) vanish.

25

	$W^+W^-\gamma$	W^+W^-Z
C, P, T	$g_1^\gamma, k_\gamma, \lambda_\gamma$	g_1^Z, k_Z, λ_Z
CP	g_5^γ	g_5^Z
P	g_4^γ	g_4^Z
C	$\tilde{k}_\gamma, \tilde{\lambda}_\gamma$	$\tilde{k}_Z, \tilde{\lambda}_Z$

Table 1.5: Symmetry conservation properties of the charged TGC.

Customary parametrization

To allow a comparison between theory and experimental results, the TGCs are expressed in terms of the deviation from the Standard Model $\Delta g_1^V = g_1^V - 1$, $\Delta k_V = k_V - 1$, λ_V ($V = \gamma, Z$) with $\Delta k_Z = \Delta g_Z - \Delta k_\gamma tan^2\theta_W$.

Moreover, since the presence of anomalous couplings in the general Lagrangian parametrization (eq. 1.37) violate unitarity at high energies [8], experimental results (i.e. Tevatron ones) usually provide TGC measurements in function of $\hat{s}$,

$$\alpha(\hat{s}) = \frac{\alpha_0}{(1 + \frac{\hat{s}}{\Lambda_{NP}^2})^n} \tag{1.39}$$

where α_0 is the low-energy approximation of the coupling $\alpha(\hat{s})$, $\hat{s}$ is the square of the invariant mass of the diboson system and Λ_{NP} is the form factor scale, an energy at which new physics cancels divergences in the TGC vertex.

1.3.3 Existing experimental results

Currently the WZ production process has been observed both at Tevatron and LHC. In 2011 D0 and CDF provided updated high statistic results and both ATLAS and CMS measured for the first time this process at $\sqrt{s}$=7 TeV.

Only the analyses considering a fully leptonic signature are here reported. The kinematic ranges used to look for the $3l + MET$ final states are summarized in Table 1.6.
The latest measurements of the $\sigma(WZ)$ will be discussed at the end (Section 6.5.3) while presenting the result obtained in this thesis, together with the existing public results provided by ATLAS and CMS.
With respect to the above analyses, the available results on the triple gauge couplings measured through the exclusive WZ production are reported in Table 1.7. The presented 95% C.L. intervals for the measured TGCs are obtained by varying one of the couplings, while fixing the remaining ones to the SM values. These results represent the most accurate measurements on the charged couplings describing the WWZ vertex, obtained at

	m_Z (GeV/c^2)	$p_T^{\mu,e}(Z)$ (GeV/c)	$p_T^{\mu,e}(W)$ (GeV/c)	MET (GeV)	m_T^W (GeV/c^2)
CDF [22]	[76,106]	>10	>20	>25	-
D0 [23]	[60,120]	>15	>20	>20	-
ATLAS [24]	[66,116]	>15	> 20	>25	>20
CMS [25]	[60,120]	>10	>20	>30	-

Table 1.6: Reference analyses for the most updated measurement of the $\sigma(WZ)$ in $p\bar{p}$ collisions at 1.96 TeV (CDF and D0) and in pp collisions at 7 TeV (ATLAS and CMS). The analyses here listed are the ones looking for WZ in the fully leptonic final state. The kinematic region used to look for the $3l + MET$ signature are described.

hadron colliders.

The best available limits [26] on TGCs measurements for $WW\gamma$ and WWZ vertices are still the ones quoted by the LEP2 collaboration, with the data collected between 1997 and 2000 at $\sqrt{s} \simeq 209$ GeV. Combined results, obtained studying single$W$, singleγ and Wpair productions with ~ 700 pb^{-1} per experiment, are reported in Table 1.8.

With respect to the WWZ couplings, the D0 measurements (quoted with a per mille precision) are really closed to the precision of LEP2 results. Concerning $WW\gamma$ couplings, up today LEP2 measurements are at least one order of magnitude more constraining than those provided by any other collaboration at hadron colliders.

Coupling	Δg_1^Z	Δk_Z	λ_Z	$\mathcal{L}$ [fb^{-1}]
CDF	[-0.08, 0.20]	[-0.39, 0.90]	[-0.09, 0.11]	7.1
D0	[-0.053, 0.156]	[-0.376, 0.686]	[-0.075, 0.093]	4.1
ATLAS	[-0.20, 0.30]	[-0.9, 1.1]	[-0.17, 0.17]	1.02

Table 1.7: One-dimensional 95% C.L. limits on TGC parameters obtained from varying one of the couplings while fixing the remaining couplings to the SM values. A form factor scale of $\Lambda_{NP} =2$ TeV is used.

Coupling	Δg_1^Z	Δk_γ	λ_Z
LEP2	[-0.051, 0.034]	[-0.105, 0.069]	[-0.040, 0.026]

Table 1.8: Best available one-dimensional 95% C.L. limits on TGC parameters, obtained from varying one of the couplings while fixing the remaining couplings to the SM values. From LEP2 combined results [26].

The Experimental apparatus

The Large Hadron Collider [27] near Geneva, in Switzerland, is a two-ring superconducting hadron accelerator and collider, installed in the 26.7 Km tunnel which was constructed between 1984 and 1989 for the CERN LEP machine. It has been an unprecedented project, not only for its center-of-mass energy[1] and for its design luminosity[2] but also in terms of complexity, size of the experiments and human effort.

Out of the possible eight interaction regions of the collider, only four are equipped with underground caverns and detectors for physics research. ATLAS [28] and CMS [29], the two general purpose detectors, are located at the two high luminosity interaction points. LHCb [30], devoted to B-physics studies, and ALICE [31], dedicated to the study of heavy ion collisions, are both designed to operate instead at a reduced luminosity. A machine schematic is shown in Figure 2.1.

In the first part of this Chapter an overview of the machine is given. The second part is devoted to the CMS detector.

2.1 The Large Hadron Collider

The LHC is a proton-proton machine, designed to collide beams at the center-of-mass energy of $\sqrt{s} =$ 14 TeV, with a nominal luminosity of 10^{34} cm^{-2}s^{-1}.
The main motivation for choosing a hadron collider was to benefit from the varying center of mass energy in the parton interactions to explore physics in the TeV region. The LHC can provide parton-parton collisions up to energies of about 1 TeV and with a very high

---

[1]The currently operating value being order 3.5 times larger than at the highest machine previously operating (the Tevatron at Fermilab)

[2]$3 \cdot 10^{32}$ cm^{-2} s^{-1} was the highest luminosity achieved at the Tevatron proton-antiproton collider after the latest upgrades. The LHC design luminosity is two orders of magnitude higher.
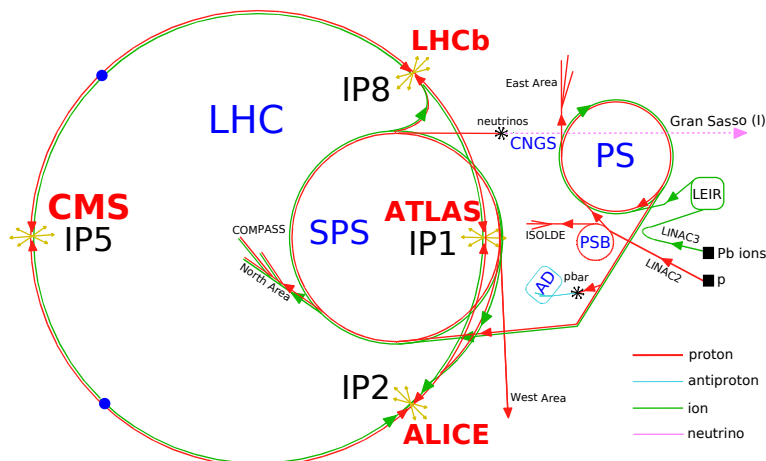
Figure 2.1: Schematic representation of the CERN accelerator complex.

collision rate, to favor the study of rare particles production.

The choice of pp beams, besides allowing to reach higher luminosities than those accessible at $p\bar{p}$ colliders and to profit from an easier production chain, is in support of the heavy ion program. Indeed, the LHC is also a heavy ion $Pb - Pb$ collider with a total energy up to 2.76 TeV per nucleon, with the purpose to investigate the quark gluon plasma.

To carry out its physic program, the LHC was designed with demanding conditions which represented a very ambitious technological challenge.

Two different vacuum chambers are used to circulate the two proton beams, in the 27 Km circumference of the tunnel, filled with about 1200 superconducting dipoles each providing a field of $\simeq$8.3 T, which represents the limiting factor to the achievable center-of-mass energy[3].

The LHC is currently running at a reduced energy of $\sqrt{s} =$7 TeV, with an instantaneous luminosity above 10^{33} cm^{-2} s^{-1}. Design and current parameters for the pp LHC operation are listed in Table 2.1.

2.1.1 LHC operation

In November 2009 the LHC delivered its first proton-proton collisions at the centre-of-mass energies of 0.9 TeV and successively at 2.36 TeV, for a total integrated luminosity of $\sim 10\mu b^{-1}$ and $\sim 4\mu b^{-1}$ respectively. This latter center-of-mass energy just exceeded that of the Tevatron (1.96 TeV), making the LHC the highest energy collider ever operated.

---

[3]From the equation p(TeV)=0.3 B(T) R(Km), where p is the beam momentum, B the magnetic field provided by the magnets of the machine and $R \simeq$4.3 Km is the radius of the LHC ring.

pp collisions		Design	October 2011
Centre-of-mass energy ($\sqrt{s}$) [TeV]		14	7
Luminosity ($\mathcal{L}$) [cm^{-2} s^{-1}]		10^{34}	$3.5 \cdot 10^{33\dagger}$
$\Delta t_{bunches}$ [ns]		25	50
Bunches per beam (n_b)		2808	$1380^\dagger$
Protons per bunch (N_p)		$1.15 \cdot 10^{11}$	$2.1 \cdot 10^{14\dagger}$
β-value at impact point (β^*) [m]		0.55	$1.6^{\dagger\dagger}$
beam σ_z [cm]		7.55	$5.7^{\dagger\dagger}$
beam σ^* [μm]		16.7	$\sim 25.^{\dagger\dagger}$

Table 2.1: The design and current machine parameters relevant for the LHC operation in pp collisions. $\dagger$ refers to the specific fill n. 2267 (30th October 2011). $\dagger\dagger$ refer to typical fills.

In 2010 the LHC upgraded its colliding energy at $\sqrt{s} =$7 TeV, providing in few months the first few pb^{-1} of integrated luminosity and approximately 40 pb^{-1} in the whole year. Starting from March 2011 the LHC kept colliding at the $\sqrt{s} =$7 TeV energy and progressively increased its instantaneous luminosity reaching the 10^{33} cm^{-2}s^{-1} in summer, already increased by a factor 5 in autumn.

The time evolution of the total integrated luminosity, during stable beams for pp running at $\sqrt{s} =$7 TeV, is shown in Figure 2.2, for 2010 (left) and 2011 (right) runs. The integrated luminosity delivered by the LHC is plotted in red and compared to the one recorded by CMS (in blue).
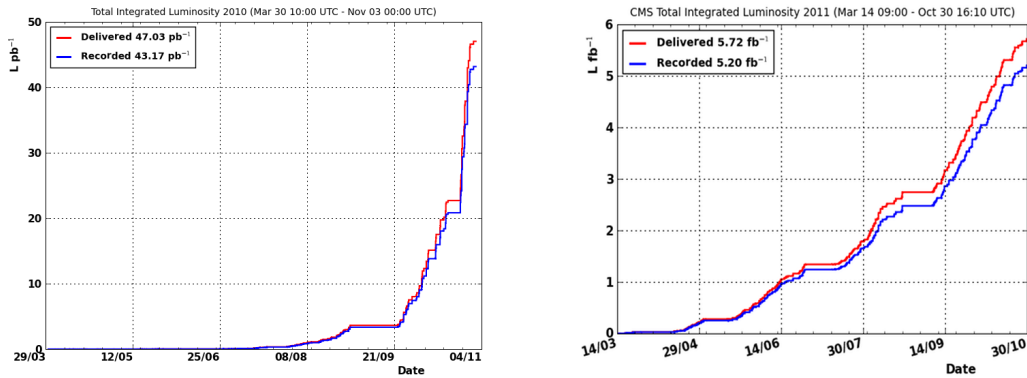
Figure 2.2: Total integrated luminosity vs. time for 2010 and 2011 stable beams operation.

Order 200/pb of integrated luminosity were collected by May and the first fb^{-1} by mid July. This 1.09/fb is used for the physic analysis presented in the last Chapter.

The LHC timeline foresees operation until 2012, most likely at $\sqrt{s}=$8 TeV, when a shutdown will allow further upgrades to enable collisions at higher energies.

2.1.2   Phenomenology of proton-proton collisions

Two main topologies of events are produced in pp collisions, depending on the energy transfer of the impinging partons, since monochromatic proton beams while colliding are essentially beams of partons (quarks and gluons) carrying a wide range of energy [32].

Large-distance collisions between the two incoming protons lead to a small momentum transfer ("soft collisions") and a suppression of particles scattering at large angle. The particles produced in the final state of such interactions have large longitudinal momentum, but small transverse momentum (p_T) relative to the beam line and most of the collision energy escapes down the beam pipe. The products of these soft interactions are called "minimum bias" events. They represent by far the majority of the pp collisions but they are not interesting for physics research purposes.

In case head-on collisions occur between two partons of the incoming protons, interacting at small distances, a large momentum can be transferred ("hard scattering"). In these conditions, final state particles can be produced at large angles with respect to the beam line (high p_T) with creation of massive particles.
However, these are rare events compared to the soft interactions. The total proton-proton cross section at 7 TeV is approximately 110 mb[4], out of which the contribution from the total inelastic processes is $\sim$70 mb[5] as measured by both ATLAS [33] and CMS [34] with 2010 data, while as an example, the production of a W boson through the annihilation of a quark-antiquark pair has a cross section of order 100 nb.

A comparison of the cross sections of the typical processes at the LHC is shown in Figure 2.3.

2.1.3   Experimental approach at the LHC

The LHC design value for proton-proton collisions of 40 MHz would provide $\sim$20 collisions per event, resulting in about 2000 charged particles every 25 ns. These reference values are still significative to understand how the detectors technologies had to be properly chosen in order to compete in such a high radiation environment: high granularity, excellent time resolution and radiation hard materials are fundamental requirements.

---

[4]A barn [b] is defined as 10^{-28} m^2

[5]The cross section from single diffractive (12 mb) and elastic (40 mb) scattering of the protons and diffractive events will not be seen by the detectors, since only inelastic scatterings give rise to particles at sufficient high angles with respect to the beam axis.
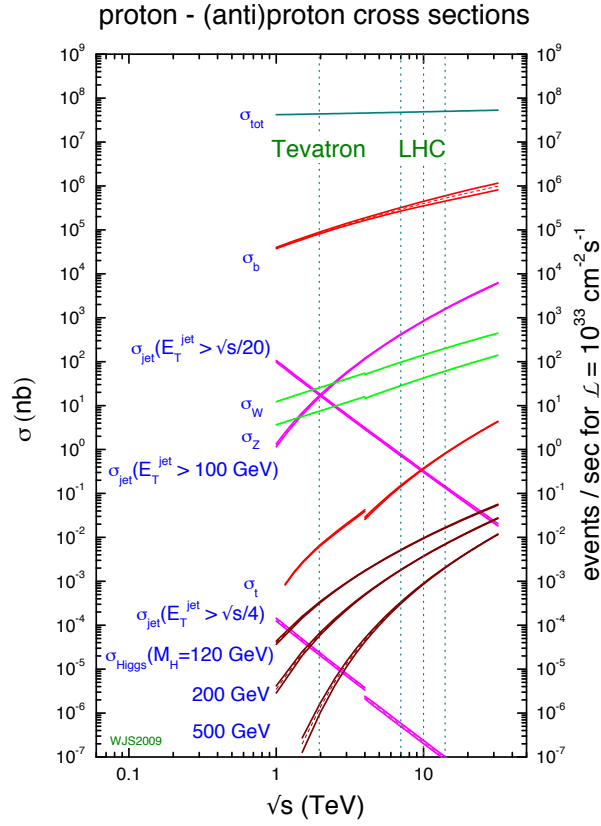
Figure 2.3: [35] Hadronic cross sections for different physic processes as a function of the centre-of-mass energy $\sqrt{s}$. Dotted lines mark the collision energies at $\sim$2 TeV (Tevatron), 7 TeV (LHC), 10 TeV and $\sim$14 TeV.

One of the most serious difficulties for the experimental operation at the LHC, with a major impact on the detector design, is the event "pile-up". Both "in-time pile-up", from multiple collisions produced in the same bunch crossing and filling the detectors acceptance, and "out-of-time pile-up", due to event overlapping between successive bunch crossing, are important contributions.

To reduce the pile-up a fast response time is mandatory for the LHC detectors, in order to limit the signal integration over too many bunch crossings. Sophisticated and highly-performing readout electronics is used, allowing for typical response times in the range of 20-50 ns (integration over 1-2 bunch crossings). The characteristics of the different final states could be afterwards successfully exploited to extract the events interesting for physics from the pile-up contribution.

Nevertheless a performing detector should be able to apply a clever and efficient selection on the events to be recorded, to reject most of the soft minimum bias products, while preserving high efficiency on interesting collision events.

Fine readout granularity is another fundamental requirement to enhance the discrimination from pile-up particles which traverse the highly segmented detectors, leading to possible confusion with objects from interesting physics events.

The LHC detectors finally have to be radiation resistant, because of the high flux of particles coming from the pp collisions.

Detectors at colliders are designed to give the best coverage of the solid angle at the interaction point. For the LHC, the natural reference frame consists of a right handed cartesian system, whose origin is fixed in the interaction vertex, with the x axis pointing towards the LHC centre, the z axis being parallel to the beam line and the y axis upward (see Figure 2.4). The coordinates triplet (r, ϕ, η) is a convenient and commonly used coordinate system which better suits the cylindric geometry: r is the distance from the z axis, ϕ is the azimuthal angle in the transverse plane and η is the pseudo-rapidity defined as $\eta = -ln(tan(\theta/2))$, where θ is the angle measured from the z axis in the $0, \pi$ range.
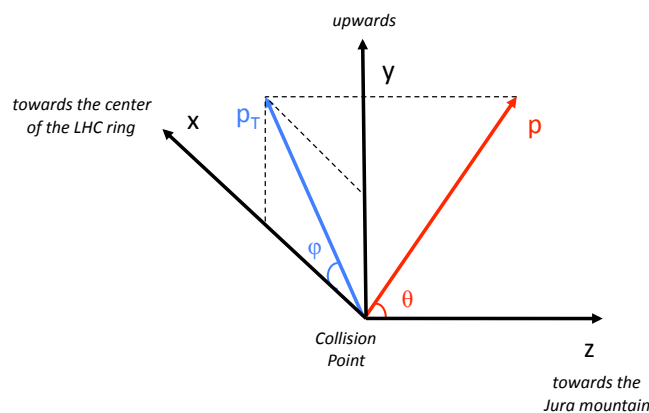
Figure 2.4: The CMS coordinate system.

Since at hadron colliders the energy balance of the events is known in the transverse plane only, the second reference frame is usually chosen to compute kinematic variables, such as the transverse momentum (p_T) or the Missing Transverse Energy (MET).

In the following, a description of the CMS detector is given, with focus on those subsystems of main importance for the work presented in this thesis.

2.2 The CMS detector

The Compact Muon Solenoid (CMS) is one of the two multi-purpose detectors of the LHC. Its layout has been mainly driven by the choice of a compact structure: a high magnetic field of 3.8 T, provided by a super-conducting solenoidal coil, a full silicon-based inner tracking system and an homogeneous scintillating crystal-based electromagnetic calorimeter are the main distinguishing features of CMS.

The CMS structure is the typical one for experiments at a collider: a cylindrical central section (the *barrel*) around the beam line, closed at its edges by two caps (the *endcaps*). The whole structure of ~ 22 m length and ~ 15 m in diameter is sketched in Figure 2.5.
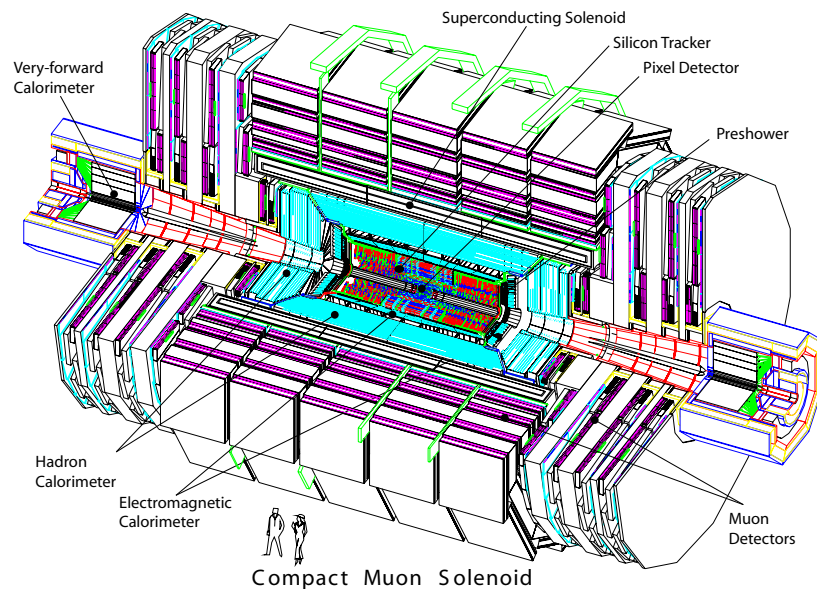
Figure 2.5: A perspective view of the CMS detector.

CMS is designed to provide high performances on leptons reconstruction and identification, with the intent to favor the detection of photons and leptons, characterized by a low yield but also by a low background. A relevant example is with respect to the Higgs boson search, in the di-photon channel or multi-leptons final state. The hermeticity of the electromagnetic calorimeter and the wide coverage of the muon system, together with the excellent momentum and energy resolution allow for a high efficiency detection, with leptons precisely measured over a very wide p_T range from a few GeV up to a few TeV. CMS also provides good reconstruction of the MET and di-jet mass, other key ingredients for high quality physics studies and precision measurements.

The CMS stratified structure presents different sub-detectors, positioned going from the interaction point outwards.

2.2.1 The tracking system

The CMS inner tracking system [36, 37] provides precise and efficient measurement of charged particles trajectories. It is a cylindric structure of 5.8 m length and a diameter of 2.5 m, surrounding the interaction point and dipped in the 3.8 T homogeneous magnetic field of the super-conductive solenoid.

The main challenge in the tracking system is to operate in the harsh LHC environment with competitive performances. Requirements on granularity, speed and radiation hardness lead to a tracker entirely based on silicon detector technology, with the crucial consequence of dealing with high power density detector electronics, which in turn requires efficient cooling. All this is in direct conflict with the aim of keeping to the minimum the amount of material, in order to limit multiple scattering, bremsstrahlung, photon conversion, nuclear interactions and in order to preserve the excellent performances of the electromagnetic calorimeter.

The layout of the CMS tracker presents two main sub-detectors: the Pixel Tracker, as innermost structure, and the Strip Tracker around. A schematic drawing of the whole structure is shown in Figure 2.6.
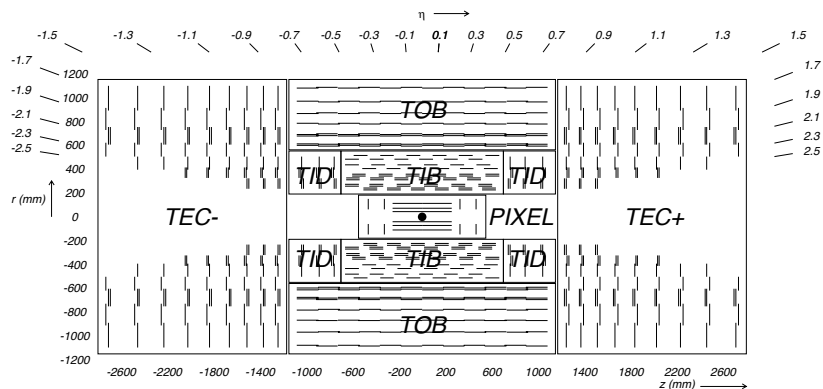
Figure 2.6: Schematic section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits.

The pixel sub-detector is the part of the tracking system closest to the interaction region, where the particle flux is the highest. It consists of three barrel layers (BPix) and two endcap disks (FPix). The 53 cm long BPix layers are located at mean radii of $\simeq 4.4$ cm, $\simeq 7.3$ cm and $\simeq 10.2$ cm. The FPix disks, extending from $\simeq 6$ to 15 cm in

radius, are placed on each side at $|z| = \pm 34.5$ and $|z| = \pm 46.5$ cm.

The spatial resolution is about 10 μm for measurements in the $r - \phi$ plane and about 20 μm for the z direction.

The arrangement of the three barrel layers and of the forward pixel disks on each side gives 3 tracking points over almost the full η range. Figure 2.7 shows the geometric arrangement and hit coverage, as a function of pseudo-rapidity η, for an interaction point at $z = 0$. In the high η region the 2 disk points are combined with the lowest possible radius point, from the innermost barrel layer.

Already in ideal conditions, such a configuration brings to important inefficiencies at high η (order 50% above 2.5). A further loss of coverage is expected for interactions displaced from $z=0$, which are likely to occur given the beam spot size σ_z=5.7 cm along the z axis.
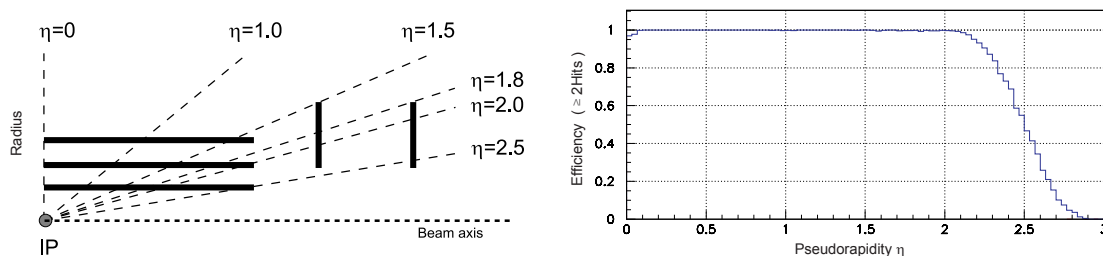

Figure 2.7: Geometrical layout of the pixel detector and hit coverage as a function of pseudo-rapidity.

The Silicon Strip sub-detector constitutes the outermost region of the tracking system. It is composed of two concentric sets of layers in the barrel part (TIB and TOB layers) and two blocks of forward disks (TID and TEC). The first three TID layers are placed on each side at $|z| = \pm 77.67$, $|z| = \pm 90.62$ and $|z| = \pm 103.57$cm, while the first three TEC disks are located at $|z| = \pm 127.46$, $|z| = \pm 141.46$ and $|z| = \pm 155.46$cm.

The single-point resolution is of the order of 30 μm in the $r - \phi$ plane and $\sim$300 μm in the z direction.

As previously mentioned, the CMS tracker involves both sensitive volumes and non-sensitive ones. Since the tracker requires a large amount of low-voltage power, a large amount of heat needs to be dissipated. Therefore a large part of the tracker material consists of electrical cables and cooling services. Other non-sensitive parts include support structures, electronics, the beam-pipe and the thermal screen outside the tracker.

The material inside the active volume varies from 0.4 X_0 at $\eta \approx 0$ to about 1.8 X_0 at $|\eta| \approx 1.4$, beyond which it falls to about 1 X_0 at $|\eta| \approx 2.5$.

In Figure 2.8, the sub-detectors and the different material contributions to the material budget are shown in units of radiation length.
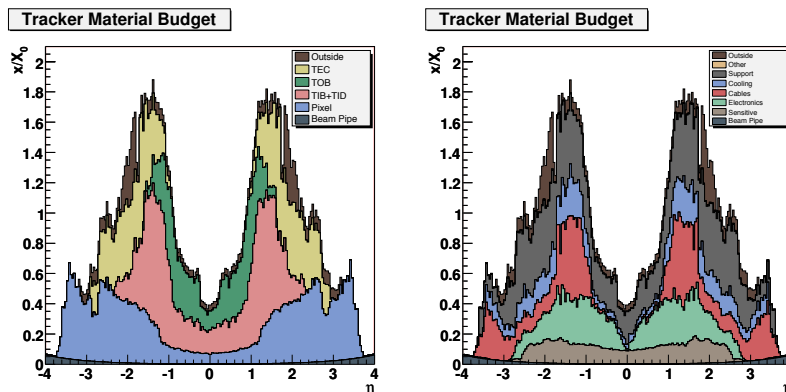
Figure 2.8: Material budget in units of radiation length as a function of pseudo-rapidity η, for the different sub-detectors (left) and broken down into the functional contributions (right).

The whole material budget is a real problem affecting in particular the measurement of electromagnetic particles. As an example for the electrons, it enhances the bremsstrahlung radiation, which has to be compensated by the use of dedicated reconstruction algorithms (3.2.2), able to deal with the energy lost. In the tracker material electrons radiate $\sim70\%$ of their energy by bremsstrahlung and photons have $\sim50\%$ probability to convert into e^+e^-.

2.2.2 The electromagnetic calorimeter ECAL

Just outside the tracker and surrounding it, the electromagnetic homogeneous calorimeter, made of scintillating crystals, provides precise identification and measurements on electrons and photons. The ECAL capability of identifying electromagnetic objects with excellent precision plays an essential role in the study of the electroweak symmetry breaking physics, through the high precision measurements of photons and electrons. A driving criteria in the ECAL design is indeed to detect the two-photons Higgs decay mode, for $m_H <150$ GeV/c^2, and the electrons from the $H \rightarrow ZZ^{(*)}$ and the $H \rightarrow WW$ decay chains, for 125 GeV/c$^2 < m_H <700$ GeV/c^2. The ECAL design [38] has been optimized to provide the cleanest signals and best energy resolution.

ECAL structure

The CMS ECAL is composed of 75848 lead tungstate ($PbWO_4$) crystals, organized in a modular structure. A view of the ECAL longitudinal section is given in Figure 2.9.

The central barrel (EB) is structured in both halves as 18 identical "super-modules", each covering 20° in ϕ and a pseudo-rapidity of $0 \leq |\eta| \leq 1.479$, with an innermost radius
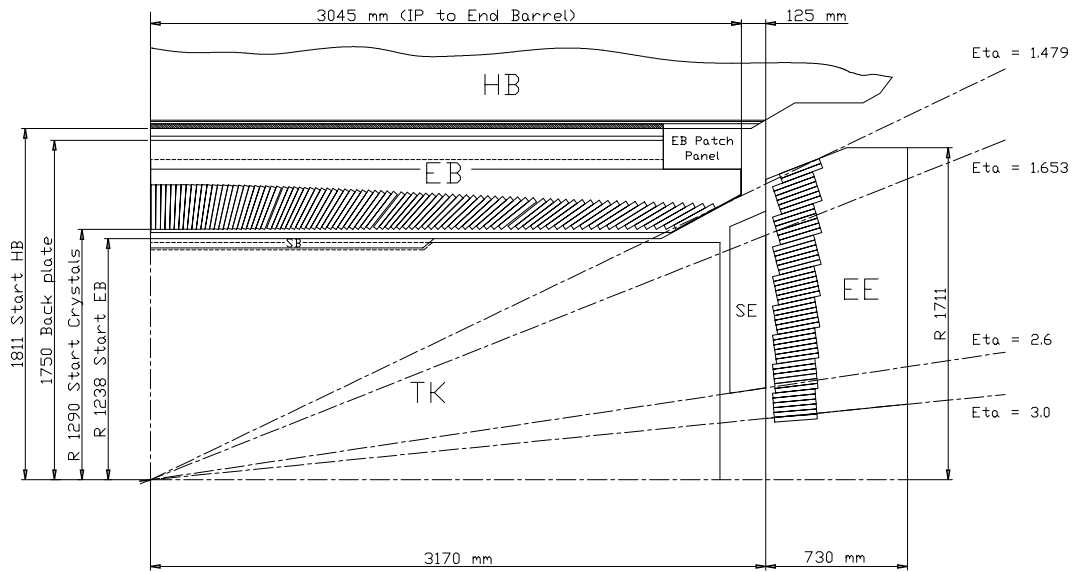
Figure 2.9: ECAL longitudinal section (one quadrant)

of 1.29 m. Blocks of 5×2 crystals compose sub-modules, which are in turn gathered into modules in the number of 40 or 50, surrounded by a metallic cross-plate. The 1700 crystals in the 85×20 modules structure form a super-module. Given the presence of 36 super-modules, the whole EB part consists of 61200 crystals.

Each crystal exposes a front face cross section of 22×22 mm^2 and a 23 cm of length, corresponding to 25.8 X_0. In order to avoid particles channeling, crystals are off-pointing by a 3° angle with respect to a straight line from the interaction point, both in η and ϕ. Blocks of 5×5 crystals are the basic elements of the ECAL data readout and are referred to as trigger towers.

The two endcaps (EE) are 3.17 m separated from the vertex and cover a pseudo-rapidity range of $1.479 \leq |\eta| \leq 3$, with precision energy measurements provided up to $|\eta| = 2.6$ and crystals being installed up to $|\eta| = 3$, in order to improve the energy-flow measurement in the forward direction. Each endcap is structured in two "Dees", semicircular aluminum plates composed by units of 5×5 crystals, known as "super-crystals".

The endcap crystals off-point from the nominal vertex position and are arranged into an $x - y$ grid. Crystals used here have a front face cross section of 28.62×28.62 mm^2 and are slightly shorter (22 cm) than those in the EB, given the support coverage provided by the preshower detector, which is a sampling structure made of lead and silicon strips of 3 X_0 in thickness.

The preshower detector is located just before the endcaps and its main purpose is to help identifying neutral pions, within the fiducial region $1.653 < |\eta| < 2.6$. It also improves

the position determination of electrons and photons, thanks to its superior granularity.

Lead Tungstate crystals were chosen as active detector. Their main peculiarities are:

- a short radiation length of $X_0 =$ 0.89 cm, resulting in short crystals (230 cm long), which fit in the compact structure;

- a short Moliere radius of 2.2 cm, allowing to contain the shower development into few crystals;

- a fast response, with 80% of the light emitted within 25 ns, allowing to stand the LHC high rate;

The crystals relatively low light yield (30γ/MeV) requires an amplification readout able to operate in the 3.8 T magnetic field. The different configuration of the magnetic field and the expected level of radiation, lead to the choice of avalanche photodiodes (APDs) in the EB and the one of vacuum phototriodes (VPTs) in the EE.

APDs have an active area of 5×5 mm^2 and they are glued in pair on the rear face of each crystal. They are operated at gain 50, at 18°C and they are read out in parallel. A very stable power supply system has been designed for the APDs, able to control the stability of the voltage at few tens of mV, thus minimizing the effect on the ECAL energy resolution at the level of per mille.
Given the crystals and APDs sensitivity to temperature fluctuations, a stable temperature with 0.01 °C precision is kept in the whole sub-modules volume.

Vacuum phototriodes have a single gain stage, a lower quantum efficiency and internal gain, if compared to APDs. These are settled by their larger surface coverage on the back face of the endcap crystals.

ECAL Readout

The readout of the ECAL crystals [39] is played in the front-end electronics, where the signal from 5 crystals is amplified, shaped and digitized in ADC[6] counts.
Here, the digitized data are used in blocks of 25 crystals to calculate the trigger primitives, subsequently transmitted to the L1 calorimeter trigger for each bunch crossing. Each trigger primitive consists in the summed transverse energy deposited in the 25 crystals tower with the additional information to characterize the lateral profile of the electromagnetic shower. The readout mechanical units [40], structured in sets of 5×5 crystals, correspond to the trigger tower blocks in the barrel, while in the endcaps these may extend over more

---

[6]An Analog to Digital Converter is used

than one supercrystal, in order to approximately follow the η, ϕ geometry of the HCAL.

At each event only a part of the ECAL channels can be read out, to avoid to exceed the about 100 kB per event allocated for the ECAL data.

A "selective reduction" of the data volume is performed by the Selective Processor [41], [42]. The strategy used consists in selecting those regions to be readout, relying on the energy deposited in the adjacent trigger towers. For the barrel case, the ECAL trigger towers are classified into 3 classes, by comparing the energy deposited in each one to 2 different thresholds, using the information from the Level-1 trigger primitives. Trigger towers result of high medium or low interest, if their energy is found above the higher threshold, between the two, or below the lower one.

A flexible use of this classification, in terms of threshold values and suppression readout levels of the channels within each class, allows to adapt to the data taking conditions.

The standard algorithm used in the simulation provides adequate data reduction even at high luminosity. High energy trigger towers ($E_T > 5$ GeV) active the full readout of its crystals and those in the neighbor towers (225 crystals for the EB case). In case of a medium energy ($E_T > 2.5$ GeV) trigger tower, only its crystals (25 in the EB case) are readout with no suppression, while low energy trigger towers, not closed to high energy ones, are readout with a "zero suppression" at about 3σ noise.

ECAL Calibration

The ECAL channels inter-calibration is needed, to correct for the important channel-to-channel variations[7]. It is indeed necessary to combine different crystals responses to provide a measurement of the energy deposited, since the shower generated by an electromagnetic particle impinging a crystal in its center is only partially contained in it. A typical electromagnetic cascade releases 98% of its energy in a 5×5 block of crystals and 81% is contained in the crystal of impact.

The error on the ECAL calibration contributes, together with light collection non uniformity, to the constant term C in the parametrization of the detector energy resolution:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2$$

where S is the stochastic term, taking into account the fluctuations in the shower lateral containment as well as contributions from photostatistics, and N is the noise term accounting for the overall electronic noise. The combination of the different contributions to the energy resolution is shown in Figure 2.10.

---

[7]Mainly due to the differences in light yield between crystals for the barrel (about 13% at construction) and the gain spread of the photodetectors in the endcaps (about 25%)
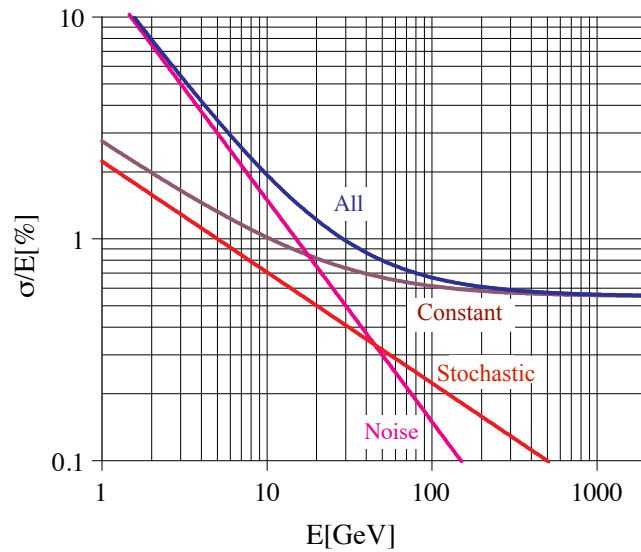
Figure 2.10: Different contributions to the energy resolution of the PbWO$_4$ calorimeter.

The most stringent requirements on the ECAL energy resolution are imposed by the necessity to measure the two photons of the SM $H \rightarrow \gamma\gamma$ decay channel. Such a measurement requires excellent performances for energies up to $\sim$100 GeV, where the constant C contribution results the dominant one (Fig. 2.10).

In particular a maximum C term of 0.5% can be allowed to guarantee the overall energy resolution demanded at high energies and this can be provided by controlling the contribution of the inter-calibration errors below 0.4%. The crystals inter-calibration is the fundamental procedure to keep the constant term at the desired level.

Several calibration techniques were used during the detector development, in accordance to the different progress of its construction.

A first pre-calibration took place before the installation in the CMS detector:

- before the crystals installation, laboratory measurements of light yield provided a first pre-calibration coefficient for each crystal (order 10-15% precision);

- during the test beam 9 ECAL super-modules were exposed to an electron beam of known energy (120 GeV) and direction. The data taken and their analysis provided a first set of inter-calibration coefficients with an inter-calibration accuracy of 0.3%;

- All the ECAL barrel super-modules were inter-calibrated in turn, being each exposed to cosmic rays for a period of about one week. The comparison of data with simulated cosmic ray events allowed to control the inter-calibration coefficients with an accuracy at the level of 1.5%.

The main method used consisted in selecting only those cosmic events releasing energy in a single crystal, by vetoing on the neighboring channels, thus guaranteeing that the cosmic muon has traversed one and only one crystal along its path.

Since only few super-modules were calibrated at the test beam, whereas all the super-modules were exposed to cosmic rays, the comparison between the sets of coefficients provided by the two methods validated the cosmics calibration technique at the level of few percent.

Using these techniques, it was possible to control the inter-calibration at the startup within 2% in the barrel region and at the level of 10-15% in the endcaps.

With the CMS detector fully operating and taking data, further long term inter-calibration procedures are performed directly in situ [43], over the entire detector.

- The $\phi symmetry$ [44], [45] exploits the energy flow invariance around the beam axis in minimum bias events, to provide inter-calibration coefficients in pseudo-rapidity regions;

- Data collected by dumping the LHC beam, in the collimators located about 150 m from the CMS detector, are used to improve the pre-calibration of the ECAL endcaps, providing a precision of about 1.6%, which is the best available and competitive with the the pre-calibration precision in most of the ECAL barrel. In these events, muons produced in the hadron cascade initiated by protons in the collimators reach the detector in a plain wide beam parallel to the beam axis, illuminating all the active channels.

- The reconstruction of $\pi^0 \to \gamma\gamma$ events is used, profiting from unconverted photons and imposing their invariant mass as a constraint [43];
  With higher integrated luminosity available, $Z \to e^+e^-$ events, producing energetic electrons correlated in different regions of the calorimeter, provide sufficient data for a global energy absolute calibration;
  The E/p measurement of isolated electrons, essentially from $W \to e\nu$ decays, are used to provide a local inter-calibration comparing the energy as measured in ECAL with the momentum estimated by the tracker;

- A light monitoring system allows to continuously monitor (at 20 min intervals) the optical transmission of each crystal near the scintillation spectrum peak ($\simeq 500$ nm) and cross-check it at a longer wavelength ($\simeq 700$ nm). Although the scintillation light yield of $PbWO_4$ crystals is not affected by radiation exposure, the crystals transparency is expected to decrease with the amount of radiation absorbed per unit of time. These effects taking place on a time scale of hours cause transparency changes of a few percent at the nominal luminosity. As an illustrative example in

Figure 2.11 the effect of transparency losses on the reconstructed π^0 invariant mass is shown before and after correction, over a time range of about a month of data taking.
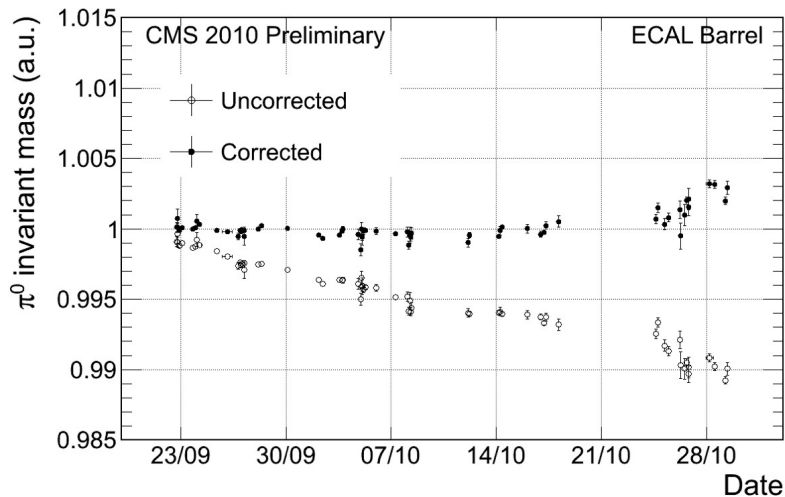

Figure 2.11: π^0 mass history plot before and after corrections to ECAL crystal energy due to transparency loss are applied. The invariant mass is normalized to unity at the start of the run period considered.

2.2.3 The hadron calorimeter HCAL

The CMS calorimetric system is complemented by the HCAL detector, with the purpose to measure energy of both charged and neutral hadrons, for the measurement of jets and missing transverse energy. A major constraint in CMS is the volume available between the outer extent of the ECAL and the inner one of the magnet coil.

To cope with the limited radial extension of the barrel part (HB), the outer hadron (HO) calorimeter is placed just outside the solenoid, to improve the central shower containment in the region $|\eta|$ <1.26. The HCAL coverage is extended along the η direction by the forward hadron calorimeter (HF) in the range 1.4< $|\eta|$ <3 and further complemented by a Cherenkov-based, radiation-hard technology down to $|\eta|$ =5.2. A longitudinal section of the HCAL is presented in Figure 2.12.

To maximize the material inside the magnetic coil in terms of interaction lengths in the HB, brass has been chosen as absorber material, having a reasonably short interaction length and being non-magnetic. Similarly, plastic scintillators tiles resulted as a good compromise to minimize the space devoted to the active medium. Overall 17 active plastic scintillators are located between the absorber plates, the innermost and outermost of which are made of stainless steel for structural strength.
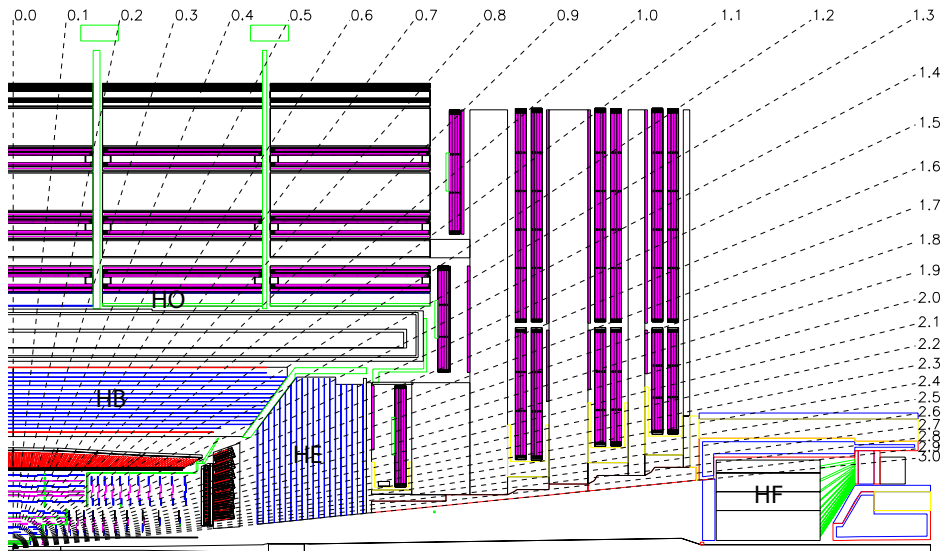
Figure 2.12: Longitudinal view of the CMS detector showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters.

The HO consists in an additional layer of scintillators, located just outside the solenoid which operates as additional absorber. The same brass absorber plates interleaved with active plastic scintillator layers are used in the HE.

HCAL towers have a size of $\Delta\eta \times \Delta\phi =$0.087×0.087 up to $|\eta|$ <1.74, increased to a size of $\Delta\eta \times \Delta\phi =$0.17×0.17 in the forward region. Each tower is instrumented with a wave length shifting fiber and the collected light is detected by a hybrid photodiode (HPD).

The HF calorimeters are placed 11.2 m from the interaction point and profit from radiation hard quartz fibers technology to provide fast collection of Cherenkov light, which is read out by photomultipliers. Besides ensuring good hermeticity, this sub-detector plays a crucial role in the determination of luminosity.

2.2.4 The magnetic field

The choice of a compact design for the CMS detector imposes a strong solenoidal magnetic field, in order to achieve the needed resolution on the muon momentum, which is required to be better than 10% for $p_T \simeq$1 TeV/c. Hence, a super-conductive solenoidal coil has been chosen to provide a 3.8 T uniform magnetic field.

Being the largest element in the whole CMS detector, the magnet constitutes the main support structure for the barrel detector components, hosting the tracker and both the calorimeters inside and the muon stations outside.

2.2.5 The muon system

One of the most characterizing detectors of the "Compact Muon Solenoid" is the muon system, located inside the magnet return yoke. Three types of gaseous detectors [46] have been installed to identify and measure muons, according to the different radiation environments, and to provide enough redundancy to cope with inefficiencies, in the measurements over the very large surface covered.

A sketch of the layout of the three muon detection systems is shown Figure 2.13.
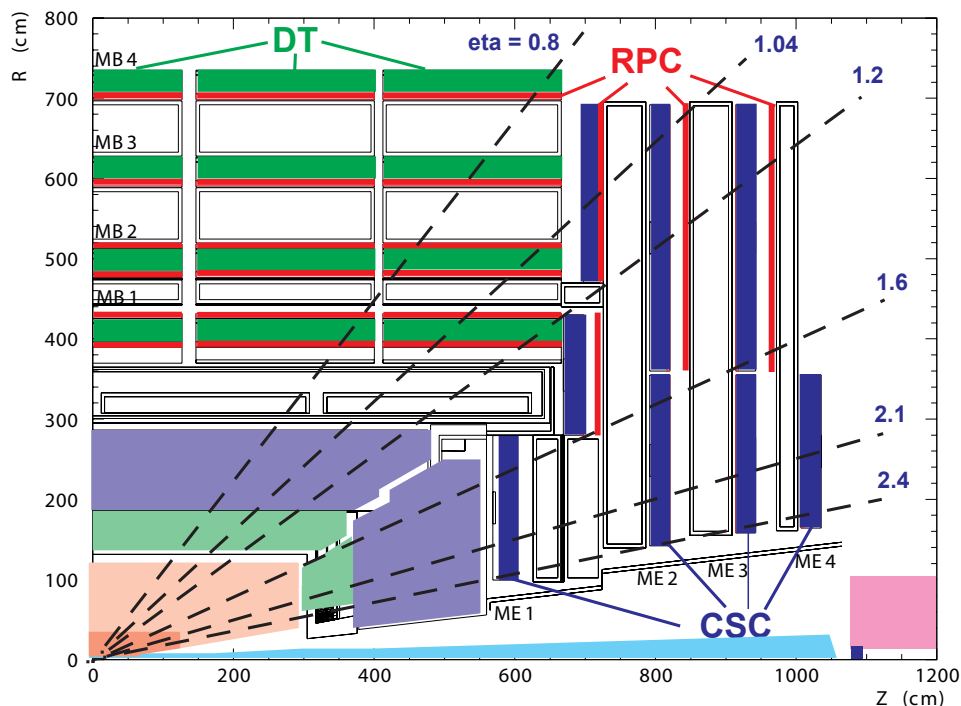
Figure 2.13: Layout of one quarter of the CMS muon system. RPC detectors providing accurate time information for triggering, DT and CSC detectors also allowing for accurate position measurement in the barrel and endcaps respectively.

In the barrel region, drift tube (DT) chambers with standard rectangular drift cells are used, since both the neutron-induced background and the muon rate are small and the 3.8 T magnetic field is uniform and mostly contained in the steel yoke. DT chambers cover the region $|\eta| < 1.2$ and are organized into 4 stations interspersed among the layers of the flux return plates. The first 3 stations contain 8 chambers each, in 2 groups of 4 and are separated as much as possible to achieve the best angular resolution, the two groups are indeed used to measure the muon coordinate in the $r - \phi$ bending plane and in the z direction. The fourth station does not contain z-measuring planes.

The drift cells of each chamber are offset by a half-cell width with respect to their neighbor to eliminate dead spots in the efficiency, providing the muon time measurement with excellent resolution.

In the 2 endcap regions, the muon rates and background levels are high and the magnetic field is less uniform. Cathode strip chambers (CSC) are used in this region covering the $|\eta|$ range between 0.9 and 2.4, thanks to their fast response time, fine segmentation, and radiation resistance. In each endcap 4 stations of CSCs are hosted, with chambers positioned perpendicular to the beam line and interspersed between the flux return plates. In each chamber, a precision measurement in the $r - \phi$ bending plane is provided by the cathode strips running radially outward, while the anode wires run approximately perpendicular to the strips and measure η and the beam-crossing time of a muon.

Both DT and CSC subsystems can trigger muons with good efficiency and high background rejection, independently from the rest of the detector. A complementary, dedicated trigger system consisting of resistive plate chambers (RPC) was added in both the barrel and endcap regions, providing a fast, independent and highly-segmented trigger with a sharp p_T threshold over a large portion of the rapidity range ($|\eta|$ <1.6). The RPCs are double-gap chambers, operated in avalanche mode to ensure good operation at high rates. They produce a fast response, with good time resolution but coarser position resolution than the DTs or CSCs and they help to resolve ambiguities in attempting to make tracks from multiple hits in a chamber.

The Physics objects

Signals as collected by CMS need to be reconstructed to provide the ingredients for physic analyses. A first online event selection is performed to reduce the total event rate and deal with an amount of data which can be further stored and processed. Such a task is performed by the trigger system in two main steps, Level-1 (L1) Trigger [47] and High-Level Trigger (HLT) [48] [49]. The full reconstruction is only performed after, for events passing the trigger selections.

In this Chapter, the criteria used to define the physical objects out of the raw data collected by CMS are presented, with focus on those particular elements of main interest in the context of this thesis work.

A brief overview on the trigger system is introduced in Section 3.1. Section 3.2 is dedicated to the algorithms used for muons, electrons, jets and MET reconstruction. In Section 3.3 the key-step of the data processing chain are outlined.

3.1 The trigger system

The trigger system at the LHC represents a very challenging task as being the first step of the event selection process. It has a crucial role, given the enormous range of production cross sections for the benchmark phenomena to be studied at LHC (see Figure 2.3), which spans over ~ 10 orders of magnitude.

The trigger system has to provide a huge reduction factor and at the same time must maintain high efficiency on the few interesting events among millions of background ones. It requires to work fast and to easily adapt to the different running conditions and physics targets.

Its highly flexible structure is provided by the combined action of the hardware system

made of largely programmable electronics (L1 Trigger) and the software system for the online event filter (HLT).

The design rate reduction of at least a factor 10^7 is obtained through the first L1 reduction to about 50 (100) kHz for the low (high) luminosity phase and the subsequent HLT operation, which reduces the output rate down to few 100 Hz.

3.1.1 The Level-1 Trigger

The L1 trigger decision to retain an event for further consideration or reject it has to be taken every 25 ns, while the feasible decision time for the L1 is $3.2\mu s$, corresponding to 128 beam crossings. Thus events are stored in pipelines while awaiting for the trigger decision. The very short time the L1 has to deal with makes it impossible to read all the raw data from the whole detector. The L1 system is therefore based on a hardware-electronic response involving the calorimetry and muon detectors only.

Its decision is based on the presence of local objects corresponding to photons, electrons, muons, jets and missing transverse energy, using information from calorimeters and muon systems in a given element of $\eta - \phi$ space and tested against several p_T or E_T thresholds. Only events accepted by the first level trigger are processed and further reduced by the HLT.

3.1.2 The High Level Trigger

The selected events to be written on tape are defined through the drastic rate reduction operated by the HLT. Its correct functioning is a key aspect for the CMS physics program, since the events rejected by the HLT are lost forever, except for very small samples retained for monitoring purposes.

The HLT aim is to save with high efficiency all the events found interesting for any physic analysis considered in CMS, minimizing at the same time the CPU requirements and discarding events as soon as possible. Therefore, it has access to the complete read-out data, with the possibility to perform complex calculations similar to those made in the offline software, and it works through sequential trigger levels, based on the L1 trigger regional output and on the presence of basic reconstructed objects in the calorimeters and muon chambers. As an example for electrons (muons), energy deposits in ECAL (muon chambers) are reconstructed using a simplified version of the algorithms used offline. Only events with energy deposits (momentum) above a programmable E_T (p_T) threshold are stored for further analysis.

With the rapidly evolving running conditions of the LHC (see Figure 2.2 for the LHC data taking history since 2010), the HLT menu needs to be updated accordingly, to keep the trigger rate under control while the instantaneous luminosity is increased by large factors. This is performed either by tightening the HLT criteria to select events or by applying a "prescale", sampling the triggered events with a 1/n reduction.

It's important to keep the HLT algorithms as close as possible to the reconstruction software used for the offline analysis. Besides reducing the complexity of the system, this ensures compatibility between the different levels of event selection. Any difference has indeed to be considered in the offline analysis, to minimize eventual biases.

3.2   The objects reconstruction

Again in order to quickly adapt to necessary changes, the software reconstruction is designed flexible, to combine measurements from the different sub-detectors, on the basis of the physics objects features.

Muons and electrons as standardly reconstructed in CMS are here described. These objects will be intensively referred to in the next Chapters, being main ingredients of the studies presented in this work. Also the reconstruction of jets and MET is presented in the following, with reference to the different algorithms used in CMS.

3.2.1   Muons

Muons are charged particles essentially detected in the tracker and muon systems. The muon reconstruction algorithm [39], [50] is seeded by the trigger information, which identifies a region of interest where the local reconstruction is performed. Three stages of reconstruction lead to objects of different level and their combination provides a robust and efficient final candidate. Muon tracks reconstructed in the muon spectrometer are identified as "standalone-muon tracks", while those supplemented by tracking detector measurements ("tracker tracks"[1]) are referred to as "global muons" or "tracker muons".

The standalone muon track (STA)

All muon layers participate to this reconstruction, which starts by building track segments in the innermost chambers. For this purpose, a straight line is fitted across the

---

[1]Tracks as reconstructed from the standard tracking algorithm and built with a Combinatorial Track Finder. The CTF is a Kalman Filter [51] based algorithm consisting of hit-based seeding followed by inside-out pattern recognition, cleaning, fitting and backward fitting

measurements reconstructed and combined within a chamber. In case of DT, up to 8 hits are available in the $r - \phi$ coordinate and 4 in the $r - z$ one, while in CSC up to 6 hits for planes are delivered.

The vectors (track position, momentum and direction) associated to each segment are used to seed the muon trajectory, working from inside out, using the Kalman-filter technique [51]. The predicted state vector at the next measurement surface is compared with the existing measurements and updated accordingly[2]. A suitable χ^2 cut is applied in order to reject bad hits, mostly due to showering, delta rays and pair production. Due to detector inefficiencies, geometrical cracks or hard showering, no matching hits (or segments) can be found and the search is continued in the next station iterating the procedure until the outermost surface of the muon system is reached.

Once the trajectory is built, a backward fit is performed, based on the same Kalman Filter algorithm, to evaluate the track parameters at the innermost muon station. Finally the track at the innermost muon station is propagated to the nominal interaction region and optionally further constrained with a vertex condition.

The global muon track

Global muons are reconstructed by fitting the standalone tracks to those selected in the tracker, within a region of interest defined by the propagation of the standalone tracks. This tracking is performed following an outside-in direction. Once two compatible tracks are found between standalone and tracker ones, the whole set of hits in the tracker and muon system is fitted and the final trajectory is extrapolated to the interaction region. Such an algorithm is efficient, provided a minimal number of hits and segments are measured in the muon system, conditions met for muon tracks with p_T above 6-7 GeV/c.

The tracker muon tracks

Tracker muons correspond to global tracks, reconstructed with the complementary approach of considering all the tracker tracks and identifying them as muons, if a compatible signature is found in the calorimeters (energy depositions) and in the muon system (compatible hits and segments). To reconstruct tracker muons an inside-out flow is followed, starting from the subset of the tracker tracks considered possible muon candidates ($p_T >$0.5 GeV/c and $p >$2.5 GeV/c) and extrapolated to the muon system, taking into account the muon energy loss in the material and the effect of multiple scattering. At least one muon segment (short track made of DT or CSC hits) matching the extrapolated track in position is required for the corresponding tracker track to be qualified as tracker muon. Tracker muons reconstruction provides a higher efficiency at low momentum (roughly

---

[2]At each step, the measured quantity taken into account is a reconstructed track segment in the barrel DT chambers, reconstructed hits in the RPCs and the individual reconstructed constituents (3D hits) of the segments in the CSC chambers.

$p < 5$ GeV/c) if compared with the global muon reconstruction which typically needs two or more segments in the muon chambers.

In case of failure for the tracker-muon system combination, only a standalone-muon track is found. This occurs only for about 1% of muons, thanks to the high tracker-track efficiency [52].

The Muon momentum

For standalone muons, the momentum precision is essentially determined by the measurement in the plane transverse to the muon bending angle, at the exit of the 3.8 T coil and taking the interaction point as the origin. The measurement uncertainty is dominated by multiple scattering in the material before the first muon station, above which the chamber spatial resolution becomes dominant.
For low p_T muons, the momentum is measured in the silicon tracker.

When a global muon is identified, the standalone muon reconstruction is combined with measurements from the silicon tracker, to obtain the best p_T resolution.
Figure 3.1 shows a comparison of the momentum resolution obtained from standalone and global reconstruction as a function of momentum. Two values of η have been chosen to illustrate the performances in the barrel and in the endcap regions.
The momentum resolution is dominated by the tracker performances up to ~ 200 GeV/c muon tracks, above which is it improved by the combination with standalone measurements.

3.2.2 Electrons

The energy released by electrons in ECAL is combined with measurements from the tracker to build the quantities associated to electron objects, whose reconstruction develops in three main steps. The energy is measured in the ECAL through the clustering algorithm, the following tracking step is performed by means of the seed finding and the trajectory building. As last step a preselection is applied, based on track-cluster compatibility criteria.

Clusters reconstruction in the ECAL

Electron reconstruction starts from the energy deposited in the electromagnetic calorimeter. Before reaching ECAL, electrons interact in the tracker material with a consequent
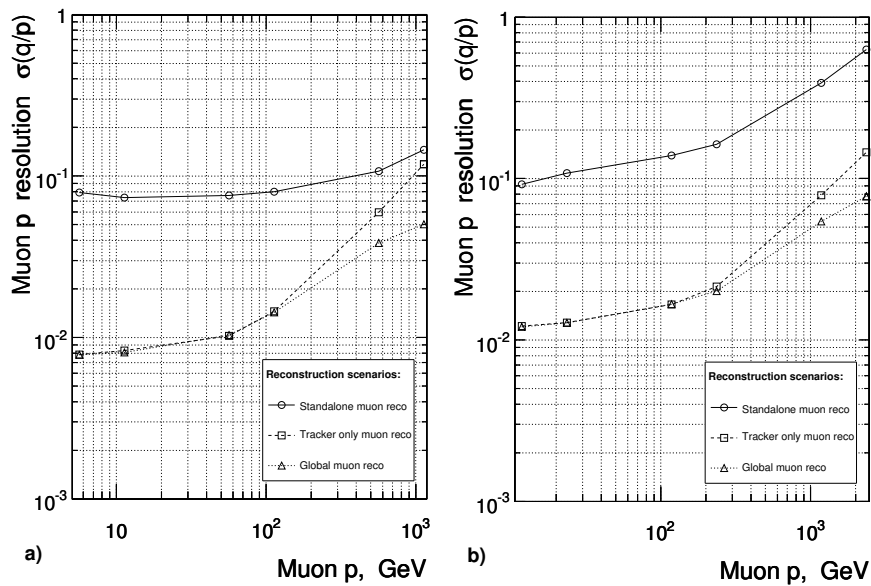
Figure 3.1: Resolution of $1/p$ versus p for standalone, global and tracker-only muon reconstruction. a) Barrel, $\eta = 0.5$ and b) Endcap, $\eta = 1.5$.

bremsstrahlung emission. Bending in the 3.8 T magnetic field, electrons reach the electromagnetic calorimeter where they generate showers involving several crystals. As a result of bremsstrahlung emission, electrons spread their energy along the ϕ direction, while approaching the ECAL surface.

To measure the electron initial energy, it is therefore necessary to collect the energy spread in ϕ due to photon emission, together with the energy deposited in the shower. This comes through a first *clustering* of neighboring crystals over a certain threshold into basic clusters, which are further grouped together into *superclusters* (SCs).

The clustering in the ECAL is performed through the "Hybrid" algorithm in the EB and the "Multi5×5" algorithm in the EE. In both cases the starting point consists in searching for crystals with a local maximum of deposited energy, called seeds.

The Hybrid Algorithm

Starting from a seed with transverse energy above a 1 GeV threshold, domains of 3×1 or 5×1 crystals along η are built, with their central crystal aligned in ϕ with the seed, thus constituting a cluster. Single clusters are assembled together around the most energetic one, to build superclusters. The maximum distance of domains in ϕ is ± 17 crystals (about ± 0.3 rad). Domains with an energy lower than $0.1\,\mathrm{GeV}$ are discarded in this process and only those consecutive clusters collecting on the whole an energy higher than 0.35 GeV are kept in the superclusters.

As shown in Figure 3.2, the hybrid supercluster is made up of a series of domains at

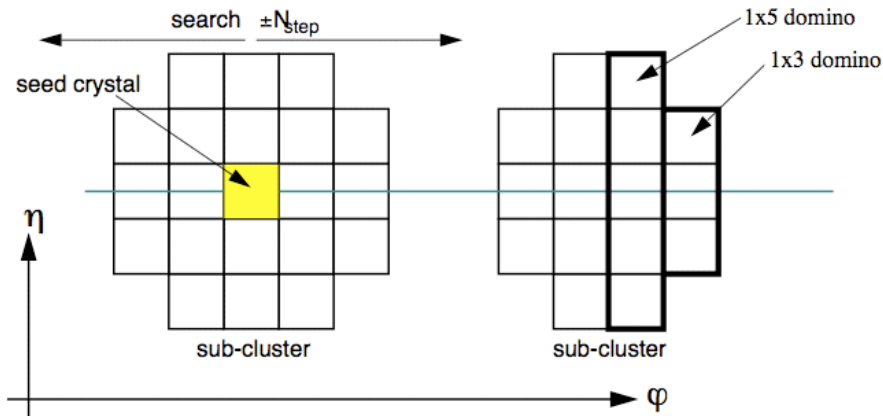constant η but spread in the ϕ-direction.

Figure 3.2: Schemes of the Hybrid clustering algorithm used in the ECAL barrel region.

The Multi5×5 Algorithm

The leading pattern of this clustering algorithm is driven by the layout of the crystals in the endcaps which does not follow an $\eta - \phi$ geometry.

Each crystal of transverse energy above the 0.18 GeV threshold seeds a cluster, provided it represents a local maximum in energy when compared to its four neighbors by side. A 5×5 matrix of crystals is built around the seed, including only those crystals not already belonging to another cluster.

To allow the energy of closely overlapping showers due to bremsstrahlung to be collected, the outer 16 crystals of the 5×5 matrix may seed a new cluster. For a final clustering with improved bremsstrahlung recovery, a rectangular window in η and extended in ϕ is created around the energy deposits with a transverse energy above 1 GeV.

In Figure 3.3 the collection of two overlapping energy deposits identified by the Multi5×5 algorithm and grouped into the same supercluster is shown.

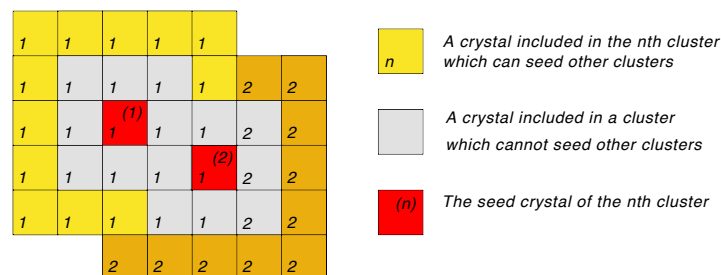

Figure 3.3: An illustration of two overlapping multi5×5 clusters.

The electrons and photons reconstructed in the endcap region $1.6 < |\eta| < 2.6$ deposit some fraction of their energy in the preshower just before reaching ECAL. Such energy is added to each endcap supercluster before further energy corrections are applied.

Energy corrections

The raw energy as measured in the ECAL is spoilt by detector effects and shower containment leakage. The measured energy of the electromagnetic objects results underestimated, which is mainly due to the interaction of electromagnetic particles with the tracker material, producing bremsstrahlung and photon conversions.

In addition to a better energy scale, the corrections allow to improve the resolution which is crucial for precision analyses involving photons or electrons. In the correction process, categories are also defined of "good" and "bad" electrons depending on the measured amount of bremsstrahlung.

The correct superclusters energy is computed as

$$E = F \times \sum G \times c_i \times A_i, \tag{3.1}$$

where A_i is the signal amplitude in ADC counts, G is a global scale calibration term and c_i is the crystal by crystal inter-calibration coefficient. F [53] is the factorization of three individual energy corrections:

- the $C_{EB}(\eta)$ factor which is applied only to EB clusters and compensates for lateral energy leakage, due to staggering;

- the f_{brem} factor which corrects for the response of the clustering algorithm to bremsstrahlung losses, as estimated from the cluster shape in ϕ;

- the $f(E_T, \eta)$ factor which parametrizes the remaining corrections, needed to account for energy containment losses.

Once the supercluster is built, its position is reconstructed in the ECAL, by means of a weighted sum of the energy depositions

$$x = \frac{\sum x_i \cdot W_i}{\sum W_i} \qquad \text{with} \quad W_i = W_0 + log\frac{E_i}{\sum E_j} \tag{3.2}$$

where x_i is the position of the ith crystal and W_i is the crystal weight. Given the ECAL crystals layout, the lateral position of the crystal axis depends on the depth, so the shower depth itself needs to be evaluated. This is parameterized as $A \cdot [B + log(E))]$,

where the parameters are different for electrons and photons, as photons penetrate deeper in the crystals before showering.

Electron track seed reconstruction

As described in the Section 3.2.1, the standard tracking in CMS is performed with a Kalman Filter algorithm. This one is suitable for muons, for which the multiple Coulomb scattering is the dominant material effect and is well modeled with Gaussian fluctuations. On the contrary a dedicated algorithm has been developed to reconstruct tracks which present asymmetric fluctuations of energy loss in the tracker material, as the case for electrons. The Gaussian Sum Filter algorithm [54] describes the bremsstrahlung energy loss probability distribution by a superimposition of several gaussians which model the Bethe-Heitler [55] function.

The entire electron track is built by performing an iterative propagation through the tracker layers [56], starting from the track seed found in the inner tracker and iteratively collecting hits up to the ECAL. At each step, compatible hits are looked for in the next available layer, by means of a propagation which takes into account the material budget between the two layers. The compatibility between the predicted and the measured hit is determined by means of a χ^2 test. If more than one compatible hit is found, the respective candidate trajectories are built and propagated in parallel. In order to avoid any efficiency losses at this stage, no strict cut is applied on the χ^2, but only the 5 best χ^2 candidate trajectories are kept for the following iteration. At each step a search path is abandoned when a hit is lost and only tracks with a minimum of five hits are kept.

With the GSF track fit, the change of curvature caused by the Bremsstrahlung photon emission is taken into account and different momentum estimates are obtained at the inner and outer track position. Such measurements at the inner (p_{in}) and outer (p_{out}) state allow to estimate the amount of energy loss in the tracker ($f(brem) = (p_{in} - p_{out})/p_{in}$), quantity which is widely exploited in the electron identification (Id) to help reducing the contamination from fakes (see Figure 3.4). Indeed as shown, the distribution is nearly flat for the signal, while it peaks at low $f(brem)$ values for the background, as expected from a fake electrons sample constituted by charged hadrons which do not radiate.

As a counterpart this algorithm is computing-time intensive and can only be run on a limited number of track seeds[3], resulting from a selection procedure.

Since the electron track seeding is the main source of efficiency loss, in CMS two different algorithms are used and combined to maximize the reconstruction efficiency.

---

[3]A seed represents a pair or triplet of measured hits in the inner tracker layers that is used to initiate the track finding.
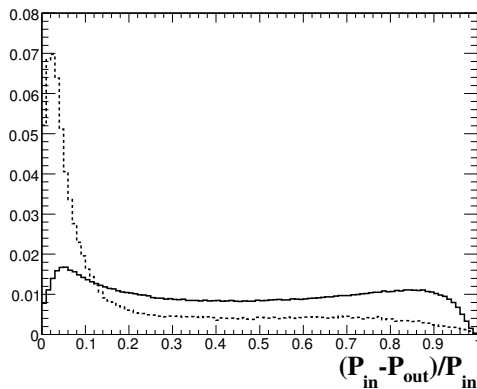
Figure 3.4: Electron bremsstrahlung fraction $f(brem)$ as measured from the normalized difference between the momentum estimate at the innermost and at the outermost track positions. For (solid line) electrons from $Z \rightarrow ee$ decays and (dashed line) background from a sample of QCD di-jet events with p^{hat} within 80-120 GeV/c.

The "ecal-driven" seeding profits from the supercluster position and energy measurements which are used to predict the position of the electron track in the innermost part of the tracker. This algorithm is optimized for isolated electrons in the p_T range relevant for Z or W decays, down to 4 GeV/c.

The "tracker-driven" algorithm starts from standard tracks as built for charged particles, extrapolated to the ECAL in order to be associated to bremsstrahlung clusters and retaining electron seeds as those of the tracks matching a cluster. It is more suitable for low p_T and not-isolated electrons [57].

In Figure 3.5 the resulting seeding efficiency as predicted by the simulation is shown in function of the generated electron η^e and p_T^e, for electrons from a sample of $Z \rightarrow ee$ decays. The separate contributions of each algorithm are also shown.

Although the tracker-driven seeding has been primarily developed and optimized for non isolated electrons, it brings additional efficiency on isolated electrons, in particular in the ECAL crack regions ($\eta \simeq 0$ and $|\eta| \simeq 1.5$) and at low p_T^e.

By combining with the tracker-driven seeds, the seeding efficiency is increased by 12.5% at 5 GeV/c, it is entirely dominated by the tracker-driven seeds below this value and it is improved at the 1-2% level at high p_T^e. Overall, the seeding efficiency for isolated electrons is $\simeq 95\%$ for $p_T^e = 10$ GeV/c and close to 100% for $p_T^e = 100$ GeV/c.

More details on the "ecal-driven" approach are given in Chapter 5, together with the description of the work done in the context of this thesis to commission the seeding algorithm with the first data.
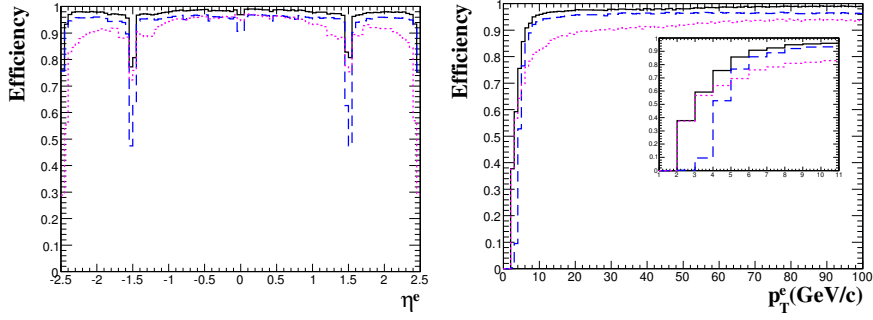
Figure 3.5: Expected electron seeding efficiency (solid line) as a function of generated electron η^e (left) and generated electron p_T^e, for a sample of electrons with uniform distribution in η^e and p_T^e for $p_T^e > 2GeV/c$. The individual contributions from the ecal-driven (dashed line) and from the tracker-driven seeding algorithms are also shown, as well as a zoom of the region $p_T^e <11$ GeV/c.

The seeds obtained with the tracker-driven and ecal-driven procedures are merged into a unique collection, keeping memory of their origin. The result is further filtered with a seed-cleaning to avoid the duplication of seeds found by both algorithms. The GSF tracking is finally run on the merged collection of seeds.

Electrons pre-selection

Electron candidates are built associating reconstructed GSF tracks to their corresponding superclusters, both for the ecal-driven and tracker-driven seeding methods.
Candidates are then preselected, relying on track-cluster matching criteria, in order to reduce the rate of jets faking electrons. The preselection is made very loose so to maximize the reconstruction efficiency and satisfy a large number of possible analyses.
For the ecal-driven electrons, $H/E^4<0.15$ GeV and $E_T >4$ GeV are already required at the seeding level. In addition to this, the track-cluster compatibility is ensured by means of the requirements $|\Delta\eta_{in}| <0.02$ and $|\Delta\phi_{in}| < 0.15$, where $\Delta\eta$ is measured as $|\eta_{sc} - \eta_{extrap.}|$ [5] and similarly for $\Delta\phi$.

The reconstruction efficiency as obtained from MC is shown in Figure 3.6 for a sample of $Z \to ee$ decays with uniform η^e and p_T^e distributions and $p_T^e >2$ GeV/c. The efficiencies are plotted as a function of generated electron η^e and p_T^e. The reconstructed electrons are required to match the generated ones in charge and in direction within a cone of size $\Delta R =0.05$. The efficiency is above $\simeq 90\%$ over the entire η range apart from the crack regions $|\eta| \simeq 1.5$ and $\eta \simeq 0$. The reconstruction efficiency rises steeply to reach $\simeq 90\%$ for

---

[4]Where H is the energy deposited in the HCAL towers in a cone of radius $\Delta R =0.15$ centered on the electromagnetic supercluster position and E is the energy of the electromagnetic supercluster.

[5]The *sc* label refers to the supercluster energy weighted position. The *extrap.* label refers to the position of closest approach to the supercluster, obtained by extrapolating the innermost track parameters.

$p_T^e \simeq 10$ GeV/c and then more slowly reaching a plateau of $\simeq 95\%$ for $p_T^e = 30$ GeV/c. The contribution of the ecal-driven and tracker-driven electrons is shown separately.
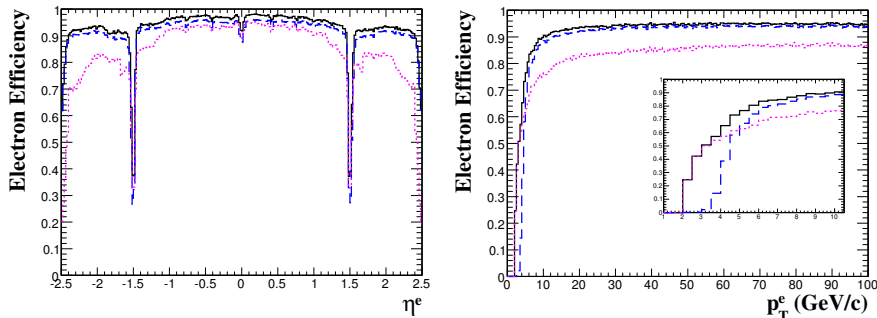

Figure 3.6: The reconstruction efficiency as obtained from MC after preselection (solid line) as a function of generated electron η^e (left) and generated electron p_T^e (right). A sample of di-electrons events with uniform distribution in η^e and p_T^e with $p_T^e > 2$ GeV/c is used. The individual contributions from ECAL seeded electrons (dashed line) and from tracker seeded electrons (dotted line) are also shown, as well as a zoom of the region $p_T^e < 10.5$ GeV/c.

Momentum determination

The electron momentum is best estimated if the energy measured in the ECAL is combined with the momentum provided by the tracker. In accordance to the respective sensitivity to bremsstrahlung induced effects, E and p are either combined or only one measurement is used.

The ECAL measurement in general dominates the resolution, however as can be expected, the tracker measurement is more effective at low energies as well as in those regions where the precision of the ECAL measurement is poor.

In Figure 3.7, the normalized effective rms of the momentum estimate for electrons in the ECAL barrel (left) and in the ECAL endcaps (right) is shown. The measurements provided by the ECAL and tracker alone as well as the combination of the two are reported.

3.2.3 Lepton identification and isolation

The ability to distinguish particles belonging to high interest physic processes, from fake signals and background decay products, is of primary importance for physics analyses. For this purpose, quality criteria are defined for each reconstructed particle. Identification (Id) and isolation (Iso) criteria are used. They are optimized depending on the kinematics and topology of the signal and associated backgrounds, to better cope with the specific demands of each physics analysis.

Pure leptonic final states are usually characterized by well defined signals isolated in the detector and satisfy high quality criteria. On the contrary leptons which are fakes
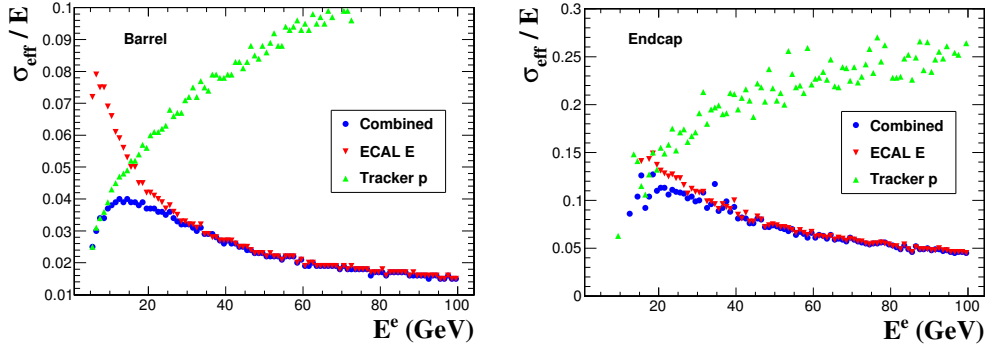
Figure 3.7: Effective resolution for the ECAL, the tracker and the combined momentum estimates as a function of the electron generated energy for electrons in the ECAL barrel (left) and electrons in the ECAL endcaps (right). Electrons are from a sample of di-electron events with uniformly distributed transverse momentum between 2 and 150 GeV/c. The precision is clearly improved by using the combined estimate with respect to the ECAL only measurement for energies below 25-30 GeV/c.

from jets or pions are surrounded by detector activity and fail some of the identification requirements.

Identification

Muon Id

The criteria used for muon identification are based on track quality parameters and the geometrical compatibility between the measurements from the different sub-detectors:

- the number of hits associated to the tracker track, standalone-muon track and global-muon track;

- the χ^2 of the various fits;

- the number of "good" hits (those having small residual with respect to the track);

- segments and hits in the muon system matched to the tracker track extrapolation;

- transverse impact parameter and related significance.

Electron Id

Track-cluster matching criteria as well as the energy in the HCAL just behind the super-cluster are used to identify electrons [58]:

- the cluster shower shape variable $\sigma_{i\eta i\eta}$[6] is particularly powerful to discriminate between electron energy depositions and showers developed by photons and fakes;

---

[6]Where $\sigma_{i\eta i\eta}^2 = \frac{\sum_i^{5\times5} w_i(\eta_i - \bar{\eta})^2}{\sum_i^{5\times5} w_i}$, with i running over all the crystals in a 5×5 block centered on the seed, η_i is the coordinate of the ith crystal, $\bar{\eta}_i$ is the energy weighted mean coordinate of the 5×5 block of crystals and w_i is the weight of the ith crystal.

- $\Delta\eta$ and $\Delta\phi$ matching angular distance between the supercluster and the track extrapolated at the ECAL surface;

- H/E to discard electron candidates if associated to important hadronic activity, behind the ECAL superclusters.

An additional criteria of fundamental importance to proper identify prompt electrons consists in the conversion rejection [59], based on the two following discriminant quantities:

- the number of missing expected hits before the innermost valid hit;

- the electron track distance from a candidate partner-track in terms of

    - Angular distance in the x, z plane (Dcot);

    - Two-dimensional linear distance in the $x - y$ plane (Dist).

Isolation

For both electrons and muons, energy deposits in the tracker and the calorimeters are used to compute isolation. For a given lepton, the sub-detector isolation consists in the overall amount of energy deposited within a cone around the lepton direction. A veto region is also used, not to take into account the energy contribution of the lepton itself.

- Relative track isolation
  $(RIso_{tk}^{e,mu}) = \sum_{tracks} p_T/p_T^{e,\mu}$ is the sum of the momentum of all reconstructed tracks in the defined region around the lepton, relative to the lepton momentum. A cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.3$ is usually used and casted around the lepton track direction to measure the activity due to charged contributions, considering only tracks with $p_T > 1$ GeV/c. A veto cone is used for muons, while for electrons the veto region includes also an η strip extended in ϕ to take into account the bremsstrahlung radiation;

- Relative calorimetric isolation
  $(RIso_{calo}^{e,mu}) = (\sum_{ECAL} E_T + \sum_{HCAL} H_T)/p_T^{e,\mu}$ is the sum of the energy deposits in the calorimeters, relative to the lepton p_T. The cone (usually 0.3 rad) is defined around the lepton direction as computed from the interaction point, to measure the activity due to neutral contributions. Calorimetry deposits associated to the electron footprint (electron case) or muon track (muon case) in ECAL are removed using different strategies for electrons and muons. A veto cone in HCAL is used for both leptons, while for electron isolation in ECAL, the veto cone is combined with an η strip extended in ϕ to take into account the bremsstrahlung radiation.

A unique variable is then computed as the relative isolation combining all the sub-detectors:

$$Iso^{e,\mu}/p_T^{e,\mu} = \left[\sum p_T(tracks) + \sum E_T(ECAL) + \sum H_T(HCAL) \right] /p_T^{e,\mu}, \qquad (3.3)$$

Other algorithms are available in CMS to qualify the reconstructed objects, making use of multi-variate analysis or categorized cuts. The criteria here reported are widely used and optimized for analysis dealing with vector bosons [60] leptonic decays. These criteria are the reference for the analysis described in Chapter 6.

3.2.4   Jet reconstruction

In a hadron collider the huge QCD cross section makes jets production by far the dominant physics process. It's important to provide jet reconstruction and identification tools able to meet the demands of the physics analysis, from the simple identification of hadronic activity up to the high performance reconstruction of hadronic objects, which can characterize the signature of specific physic processes.

Jet Clustering Algorithms

To allow for accurate comparisons between parton-level predictions from theory and hadron-level observations in collider experiments, well defined jet finding procedures are required. In the context of the high energy physics, two distinct types of algorithm are most used, a cone-based and a sequential-clustering one. In CMS both types of algorithm are implemented to reconstruct jets, the Iterative Cone [39] and SISCone [61], among the cone-based ones, and the k_t [62] [63], the *Cambridge/Aachen* [64], [65] and the $Anti - k_t$ [66] among the others.

The Iterative Cone algorithm is seeded by the most energetic ECAL+HCAL tower and further clusters all particles within a cone of radius R in the η, ϕ space around the seed axis into a trial jet. A new direction is computed from the energy weighted contributions of the constituents and a new cone is defined. The procedure is iterated, until both the jet energy and direction remain stable between two following steps. Jets constituents are then removed from the list of particles and the algorithm proceeds with the most energetic remaining particle as the next seed until no seed with energy above 1 GeV remains.

Despite of its simplicity, fast execution and good p_T resolution, the Iterative Cone is neither collinear nor infrared safe, like most commonly used cone-type algorithms.

The Seedless Infrared-Safe Cone was proposed combining a non-iterative approach, in-

frared and collinear safe, with a moderate computational execution time.

Sequential clustering algorithms are all based on successive pair-wise recombination of particles according to their reciprocal distances. All three algorithms are safe against soft emissions and collinear splitting, leading to a robust event interpretation in terms of partons and allowing for the application of the same algorithms in theoretical calculations for comparisons with experimental data. Very demanding CPU requirements are the major defects of these algorithms.

Two categories of jets as reconstructed in CMS are here reported. A third solution is provided within the particle flow reconstruction (Section 3.2.6).

Calorimeter Jet reconstruction (*CaloJets*)

For this category, the energy deposits in the electromagnetic and hadronic calorimeters are the input used, combined into unit of towers. The basic blocks are the HCAL cells and the geometrically corresponding matrix of 5×5 crystals just ahead, in the barrel region. A slightly more complex association is used in the endcap region, due to the particular geometry.
The energy deposits in a given cell are added to the tower estimate only if above a proper threshold (different for each detector region[67]), to reject electronic noise. Finally, towers are required to fulfill $E_T > 0.5$ GeV to be considered for further clustering.

Calorimeter Jet-Plus-Tracks reconstruction (*JPT*)

Calorimeter jets are further improved in terms of p_T resolution and response, profiting from the excellent performance of the CMS tracking system, thus leading to Jet-Plus-Tracks (*JPT*) objects [68].
Charged particle tracks are associated to the reconstructed *CaloJets*, relying on geometrical matching in $\eta - \phi$ between the jet axis in ECAL and the track momentum measured at the interaction vertex.

3.2.5 Missing Transverse Energy in CMS (*MET*)

Missing Transverse Energy is the main physics quantity used to infer the presence of particles invisible to the detector. Indeed at colliders, the momentum conservation for each event is a constraint that can be only exploited as projected in the plane transverse to the colliding beams. Neutral particles as neutrinos, those mis-reconstructed or falling outside the detector acceptance are revealed indirectly, looking for an unbalance in the event transverse energy, which is otherwise expected to be small, due to the hermeticity

of the CMS detector.

MET reconstruction performances are sensitive to detector miscalibration or inefficiencies, with the eventual effect of spoiling the discriminating power between undetected neutral particles and spurious detector/data taking effects.

As for jet reconstruction, here reported are two techniques, while a third is implemented in the context of the particle flow reconstruction (Section 3.2.6).

Calorimetric MET (*caloMET*)

For this algorithm, only ECAL and HCAL towers are used, MET being computed as the negative vectorial sum of the calorimetric energy depositions.

Such an estimate does not take into account muon energy, since muons are minimum ionizing particles that traverse the calorimeters almost unaffected (on average, they deposit only 2 GeV). Therefore muon contribution is added, by subtracting vectorially the muons momentum.

$$\overrightarrow{caloMET} = - \sum_{calotowers} \overrightarrow{E_T} - \sum_{muons} \overrightarrow{p_T^\mu} + \sum_{muons} \overrightarrow{E_T^\mu} \tag{3.4}$$

Further corrections can be applied to this *caloMET*, to account for Jet Energy Scale, un-instrumented regions of the detector and effects due to underlying events.

Track-Corrected MET (*tcMET*)

In this case, the MET as measured in the calorimeters is further corrected using the information from other sub-detectors. In particular measurements from the CMS tracker are exploited when possible.

An overall resolution improvement and a better description of the MET distribution in the tails result from replacing the well reconstructed tracks momenta to their energy measurements in the ECAL, while computing the MET. With exception for electrons and muons, the track momentum is used for all charged particles with p_T between 2 and 100 GeV/c [7]. As for *caloMET*, identified muons are corrected for in a similar way.

$$\overrightarrow{tcMET} = caloMET - \sum_{tracks} \overrightarrow{p_T^{track}} + \sum_{tracks} \overrightarrow{E_T^{track}} \tag{3.5}$$

A major goal of this algorithm is the better control of the tails in the MET distributions, for events with spurious contribution to the missing transverse energy.

---

[7]Tracks with p_T <2 GeV are fully compensated for assuming no response from the calorimeters (MET= 0). Tracks with p_T >100 GeV receive no correction, since at that energy the measurement of calorimetric energy for charged hadrons is already good enough not to gain anything from the tracker resolution.

Figure 3.8 shows the MET distributions measured in both $W \to e\nu$ and $W \to \mu\nu$ candidate events. The performances of the $caloMET$ and $tcMET$ reconstructions are compared to the $pfMET$ algorithm.
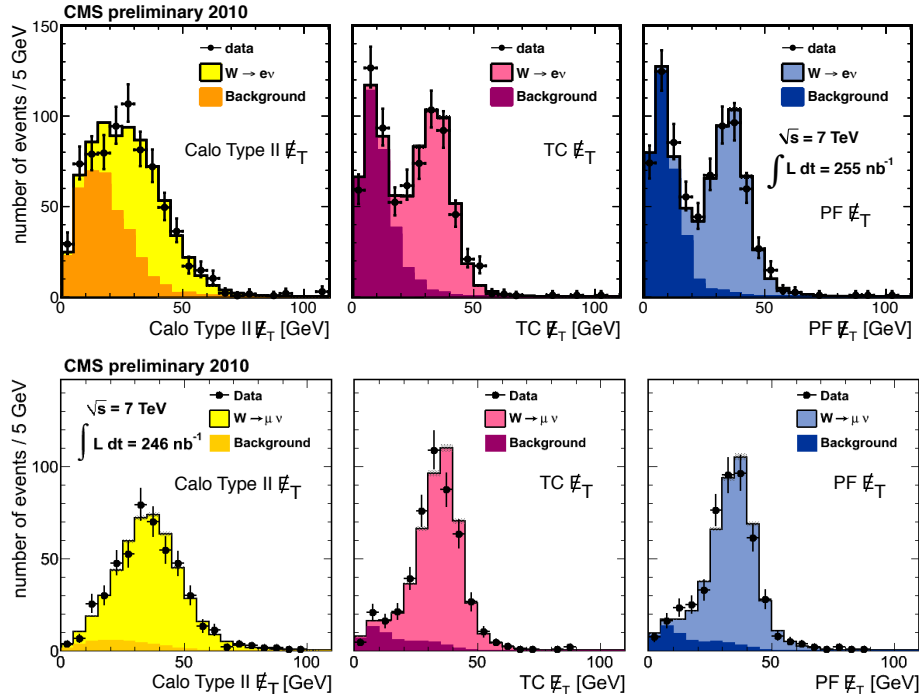

Figure 3.8: [69] MET distribution in $W \to e\nu$ (above) and $W \to \mu\nu$ (below) candidate events. The three columns correspond to the three MET algorithms (left to right) $caloMET$, $tcMET$ and $pfMET$. Both data (points) and simulation (histograms) results are shown.

3.2.6 Particle Flow Reconstruction ($pfJets$ and $pfMET$)

The key point of the particle flow (PF) algorithm [70] is to reconstruct, identify and calibrate each single particle in the event individually, by combining all CMS sub-systems. Charged particle tracks and clusters in the electromagnetic and hadronic calorimeters, as well as signals in the CMS preshower and the muon system are the elements to be connected to each other by a linking algorithm, to fully reconstruct each single particle, while getting rid of any possible double counting from different detectors. For each pair of elements in the event, the quality of the links performed is quantified on the basis of the distance between any two elements, connected directly or indirectly and producing a "block". Each block undergoes an identification to allow for the reconstruction of particles, available for subsequent physics analysis. A priority order in the object identification is followed, based on which detector systems contribute to a single block, muon (track + ECAL + HCAL + Muon System), electron (track + ECAL), photon (ECAL), charged hadron (track + ECAL + HCAL) or neutral hadron (HCAL) are reconstructed.

The particle flow provides an alternative to the standard reconstruction within CMS. Given its energy-flow scheme algorithm, the particle flow is particularly suited for (and preferred for, in CMS) reconstructing jets, *tau*-jets and MET.

The goal of the PF reconstruction is to produce fully calibrated inputs to the jet reconstruction, such that the resulting jets require little or no further a-posteriori energy corrections. In addition, it exploits the excellent tracking detector and the very good granularity of the ECAL to provide high resolution estimates on both energy and position. The combination from the different sub-detectors allows for a significant degree of redundancy making the reconstructed calorimetric objects less sensitive to the energy calibration.

For instance, the particle flow MET ($pfMET$) does not require any further correction for muons or tracks, since already considered in the inclusive reconstruction approach.

$$\overrightarrow{pfMET} = - \sum_{pfParticles} \overrightarrow{p_T} \tag{3.6}$$

3.3   Data processing

The efficiency of the data processing and reconstruction chain, as well as the capacity of delivering the output for studies within a short time from the data taking, plays a crucial role for a competitive physics research activity at the LHC.

The LHC experiments manage the data processing and delivery on a world-wide scale, through the Worldwide LHC Computing Grid (WLCG) [71], a distributed computing and data storage infrastructure. Given the large size of the CMS collaboration and the large volumes of data stored for analyses, it's through the Grid that the processed datasets are delivered to the storage points and that users can access them.

In CMS, the online software system presents a flexible structure, organized into large domains, data acquisition (DAQ) [72, 73], run control (RCMS) [74] and detector control system (DCS) [75]. It classifies the data taken in Primary Datasets (PDs), according to the trigger path activated in each event, in such a way to minimize any overlap.

The data processing is performed in a hierarchical architecture of computing centers (Tiers), allowing at the same time for the data distribution over the LHC Grid storage points.

Tier-0 centre at CERN receive the Primary Datasets in the `RAW` format, containing the full recorded information from the detector, plus a record of the trigger decision. Here,

data are processed towards a `RECO` format, for which detector filtering and correction of the digitized data are applied and object reconstruction and identification is performed, providing high-level physics objects, plus a full record of the reconstructed hits and clusters used for their production.

`RAW` and `RECO` data are then copied to Tier-1 centers, where they are stored and reprocessed when updated detector corrections and reconstruction algorithms are available. User access is only allowed at Tier-2, where most popular datasets are transferred in `AOD` (Analysis Object Data) format, containing a summary of the `RECO` level information.

The `ROOT` [76] data format is used to handle the events writing and to access the event information, which is organized in "Trees" providing the whole collections of the objects reconstructed in each event.

The CMS software `CMSSW` [77] is the framework used to process and analyze the datasets.

To allow for a comparison of physic measurements with expectations, Monte Carlo (MC) simulations are performed. Such event generation is then propagated through the detector by the detailed simulation package GEANT4 [78].

Finally, simulated events undergo the same reconstruction chain used to produce `ROOT` format file for the detected events.

ECAL energy scale validation with cosmic ray muons

During October-November 2008, the CMS Collaboration conducted a month-long data taking with comics known as the Cosmic Run At Four Tesla (CRAFT).
With all installed detector systems participating in taking data, CMS recorded 270 million cosmic ray muon events with the solenoid operating at a magnetic field intensity of 3.8 T.

The CRAFT data were exploited to test in situ the CMS magnet at the nominal current and to commission the experiment for an extended operation period [79], allowing to study the performance of the detectors and resulting particularly useful to validate its alignment and calibration.

In this Chapter it is presented the analysis on CRAFT data I worked on for the measurement of the muon stopping power in the lead tungstate ($PbWO_4$) of the electromagnetic calorimeter.
The context of the measurement is presented and a description of the experimental setup for the cosmic data taking is given. The measurement procedure and the instrumental effects are discussed afterwards and the results are presented in the last Section. Major attention is given on the importance of this measurement to test the ECAL energy scale, set with 120 GeV/c electron beam, down to the sub-GeV region at the level of few percents, using energy deposits of order 300 MeV associated with low-momentum muons.

4.1 Introduction

A sub-sample of about 250 million cosmic ray muon events, recorded by the CMS detector during the CRAFT runs, was exploited in order to assess the performances and commission the ECAL in preparation of the LHC beams.

It's in this context that the measurement of the dE/dx for muons in the ECAL was performed, as a further contribution to the crucial inter-calibration activity (see Section 2.2.2). The analysis developed in the measurement of the muon stopping power in lead tungstate ($PbWO_4$), for muons spanning a momentum ranging from 5 GeV/c to 1 TeV/c. The outcome of the measurement, which proved consistent with the expectations over the entire range, resulted of particular interest not only for calibration purposes [80], but brought also to the first experimental measurement of the muon critical energy in $PbWO_4$.

With this result, the ECAL energy scale, set with 120 GeV/c electrons at the test-beam is checked in the cavern at the level of few percents, using low-momentum muons energy deposits (order 300 MeV).

The analysis presented in this Chapter is relevant for the work presented in this thesis, where leptons and electromagnetic objects in particular are the main ingredients used. This analysis represents a detector level activity as a complement to the reconstruction studies performed on electrons (next Chapter) and the analysis part (last Chapter) where the WZ production cross section is measured taking advantage of the high performances on electrons and muons objects.

4.2 The operation setup

During CRAFT, the CMS detector installed 100 m underground was operated in cosmic data taking mode and thus tested in a configuration completely different from the design one. In particular, CMS is optimized to detect particles from collisions delivered by the LHC, traversing its structure from the beam pipe outward, and the reconstruction algorithms are accordingly designed to follow an inside-out track finding.

Nevertheless some modifications to the standard reconstruction code as well as a tuning of the detector operation parameters allowed for an efficient detection and reconstruction of cosmic-muon events.

4.2.1 The cosmic ray muons reconstruction

To deal with CRAFT data, dedicated muon reconstruction algorithms were therefore developed [81] in addition to the standard ones designed for pp collisions (see Section 3.2.1).

While the standard reconstruction typically yields two tracks ("2 legs") (one in each of the top and bottom halves of the detector) for a single cosmic muon, the dedicated cosmic-muon algorithms allow for a single fit of the track segments traversing the whole detector ("1 leg") and are optimized for muons not pointing to the nominal beam-interaction region.

The different objects reconstructed for a cosmic muon crossing CMS are displayed in Figure 4.1.
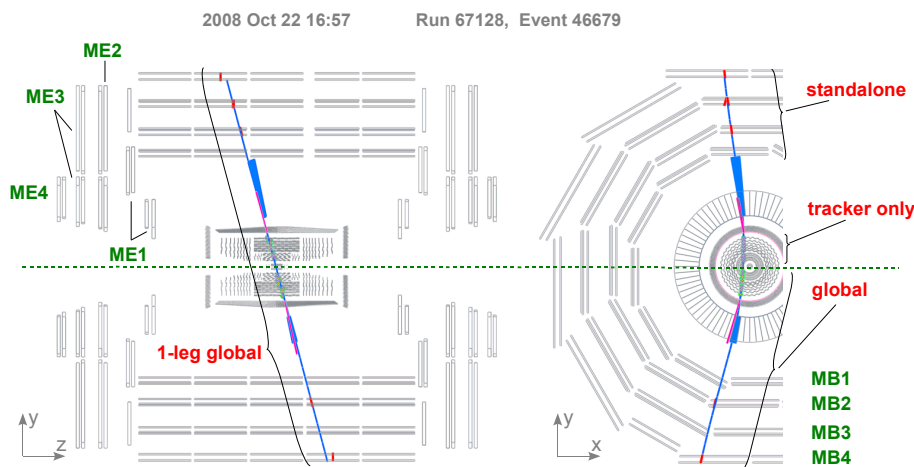

Figure 4.1: Event display of a cosmic muon crossing CMS: the side view (left) and a part of the transverse view (right). "MB" and "ME" labels indicate positions of the muon barrel and the muon endcap stations, respectively. The solid blue curve represents a 1-leg global muon reconstructed using silicon tracker and muon system hits in the whole detector. Small green circles indicate hits in the silicon tracker. Short red stubs correspond to fitted track segments in the muon system; as the z position is not measured in the outer barrel station, the segments in it are drawn at the z center of the wheel, with their directions perpendicular to the chamber. Energy deposits in the electromagnetic and hadron calorimeters are shown as (thin) magenta and (thick) blue bars, respectively.

Essentially it is enabled the reconstruction of tracks consisting of two standalone muon trunks at opposite sides of the detector and a single track that traverses the entire tracker sandwiched between them. The standard Combinatorial Kalman Finder[1] (CKF) and the specialized Cosmic Track Finder[2] (CosmicTF) [82], [83] were used to reconstruct tracker tracks. The overall variety of muon candidates available in CRAFT data can be summarized in:

---

[1]A specific configuration for cosmic-muon events provides single tracks (1-leg CosmicCKF) or two separate tracks in the two hemispheres of the detector (2-leg CosmicCKF).

[2]Expressly designed to reconstruct cosmic muons crossing the tracker as single tracks, with the possibility to split such tracks at the point of their closest approach to the nominal beam line into two separate track candidates, with each of the candidates refitted individually, leading to a pair of so-called "split tracks".

- LHC-like global muons, formed from CKF and cosmic standalone-muon tracks (STA), designed for the reconstruction of muons produced in beam collisions at the LHC. Only cosmic muons crossing the detector within few centimeters from the nominal interaction point are taken into account by this algorithm;

- 1-leg global muons, formed from 1-leg CosmicCKF tracker and STA standalone-muon tracks. These muons typically consist of a single track in the entire tracker sandwiched between two standalone-muon tracks and yield the best estimate of the muon parameters, profiting from the about 15 m long lever arm provided by the muon system;

- split global muons, each formed from a split tracker track and a cosmic STA standalone-muon track. A comparison between the pair of such tracks from a cosmic muon traversing the core of the detector provides a measurement of muon reconstruction performance (closed to 100% efficiency), since the two fits are performed independently. At the same time the splitting mechanism ensures that they originate from the same muon;

- 2-leg global muons, each formed from a 2-leg CosmicCKF tracker and a cosmic STA standalone-muon track. These tracks provide fully unbiased measurements of reconstruction performance, since the two 2-leg muon tracks typically found for each tracker-pointing cosmic muon are treated independently at all stages of the reconstruction. In this case, care must be taken to ensure that they were produced by the same muon.

4.2.2 The ECAL operation

The electromagnetic calorimeter has been designed for excellent energy resolution and is the main component of the CMS detector to identify and measure photons and electrons.

The CRAFT tests were the first opportunity to operate the ECAL installed in CMS for an extended period of time at 3.8 T, leading to a successful check of its performances [80]. Out of the 270 million cosmic ray events recorded by CMS, a total of 246 million were used in ECAL reconstruction and analysis, 158 million of which were taken with the nominal APD gain of 50 (G50), in order to study trigger performance, noise and the signatures of minimum ionizing particles (mip) in the same configuration as the one to be used for collision data. For the remaining commissioning runs, the APD gain was raised to 200 (G200), in order to study cosmic ray muon signatures in ECAL with higher efficiency, increasing the sensitivity to low-energy deposits and reducing the impact of the electronic noise. The noise per readout channel indeed resulted to be 1 ADC count independently

on the APD gain [80] and a *mip* muon traversing a crystal along its axis releases ~ 300 MeV (1 ADC$\simeq 9.3$ MeV at G200 in EB) thus leading to an increased signal to noise ratio for cosmic ray muons in G200.

To provide a good reconstruction efficiency down to low energy deposits associated with muons not pointing to the nominal interaction vertex, a dedicated algorithm has been developed to reconstruct the energy released by cosmic ray muons in ECAL.
Ad hoc thresholds were used in the online data reduction (see Section 2.2.2), with zero suppression threshold corresponding to about 20 MeV and the selective readout one to about 170 MeV.

For the energy reconstruction, the clustering algorithm can be seeded in CRAFT by both a crystal with an energy deposit of at least 15 ADC counts (139.5 MeV) or by a pair of adjacent crystals with at least 5 ADC counts (46.5 MeV). The cluster energy is then determined as the sum of all channels above 2 ADC counts (18.5 MeV) belonging to a 5×5 matrix of crystals. Contiguous clusters are subsequently merged and the resulting size in the ECAL depends on the impact angle of the muon at the calorimeter surface.

The cluster energy is obtained by applying the inter-calibration constants relative to the individual channels and an absolute energy scale factor. This last factor is defined as the response to 120 GeV/c electrons striking the centre of a 5×5 matrix of crystals in a reference region of ECAL [84] and it is extended over the entire ECAL, up to local containment effects. According to MC simulation, this containment factor corresponds to 97% of the electron energy, for an impact at the point of maximum response of the reference crystal matrix, while the energy deposits due to non-radiating relativistic particles are fully contained in a 5×5 crystal matrix.
Thus the energy scale defined at the test beam is scaled by 0.97 to measure the ECAL response to muons. Additional corrections, related to muon containment effects are properly considered and discussed in Section 4.4.2.

The calorimeter occupancy is shown in Figure 4.2 for the bottom half region, where some dead regions are present, corresponding to a super-module turned off during these tests and to some masked trigger towers. Given the angular distribution of the cosmic muon flux and the lower sensitivity of the ECAL endcaps to the energy released by a minimum ionizing particle, only the barrel part of the detector is considered.
The bin size is approximately corresponding to one crystal. The distribution shows a higher energy deposition in the positive η ECAL side, which is consistent with the geometric layout with the CMS shaft for the outside access located at $z < 0$ and the ECAL crystals off-pointing from the interaction vertex. The muons from the shaft not shielded

by the ground release more energy in the positive η side in ECAL, since crystals are there traversed along their axis.
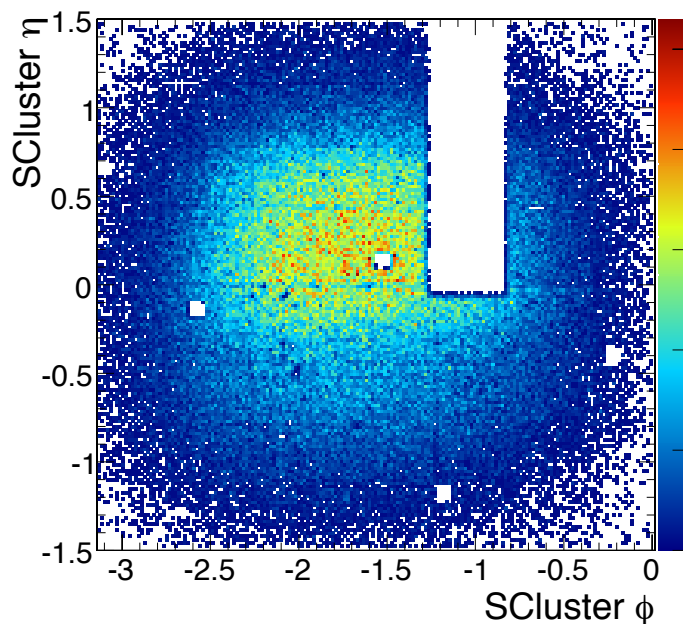

Figure 4.2: Occupancy plot showing the energy released by muons in the ECAL calorimeter, as a function of η and ϕ, for the bottom half and barrel region only.

4.3 Event selection and measurement procedure

By using the cosmic ray muons in the momentum range between 5 and 1000 GeV/c, the ΔE measured by ECAL has been used to estimate the muon stopping power in lead tungstate (PbWO$_4$).

4.3.1 The dE/dx approximation

The stopping power is experimentally determined by the measurement of the energy ΔE lost by a muon when traversing a known thickness Δx, which must be small in order that the approximation $f(E) = dE/dx \simeq \Delta E/\Delta x$ holds. Quantitatively $\Delta f/f$, where $\Delta f = f(E) - f(E - \Delta E)$, has to be smaller than the desired relative precision on the measurement of f.

$$\text{Since} \qquad \Delta f \sim \frac{df}{dE}\frac{dE}{dx}\Delta x, \qquad \text{then} \qquad \frac{\Delta f}{f} \sim \frac{df}{dE}\Delta x \qquad (4.1)$$

In the energy range considered, the stopping power for muons can be parametrized [85] as

$$f(E) = \left\langle -\frac{dE}{dx} \right\rangle = a(E) + b(E)E, \qquad (4.2)$$

where E is the total muon energy, x is the thickness of the traversed material, commonly measured in mass per unit surface, $a(E)$ is the stopping power due to collisions with atomic electrons and $b(E)$ is due to radiative processes such as bremsstrahlung, direct pair production, and photo-nuclear interactions [3].

The derivative of the stopping power with respect to E can be written as a sum of the contributions of collision and radiation processes, namely

$$\frac{\Delta f}{f} = \left[\left(\frac{df}{dE} \right)_{coll} + \left(\frac{df}{dE} \right)_{rad} \right] \Delta x \qquad (4.3)$$

In the context of this measurement Δx is typically 180 g/cm^2 (about 24.7X_0 in PbWO$_4$) and f_{rad} can be $b \cdot E$, with b determined by the value of f_{rad} at 1 TeV [2], leading to $(df/dE)_{rad} = b = 1.6 \times 10^{-5}$ cm^2/g.
This corresponds to a muon radiating in average 1/350 of its energy when traversing $\Delta x =$180 g/cm^2. Its average contribution to $\Delta f/f$ is then 1/350, which is negligible when compared to the precision of this measurement (order of percent).
The component of the derivative due to collision in the energy range of the present measurement is a monotonic decreasing function of E and its expected value at the lowest momentum (5 GeV/c) is 2.6×10^{-5} cm^2/g.
Thus $\Delta f/f$ is everywhere small when compared with other sources of uncertainty in the present measurement.
The above argument considers only the average energy loss, but it can be shown that the condition that a muon radiates on average a small fraction of its initial energy remains sufficient for a correct determination of the stopping power, even when the large event by event fluctuations in radiated energy associated with bremsstrahlung are considered, being $f(E)$ essentially a linear function of E in this energy range.

Thus for each muon, the momentum (p), its path length in ECAL (Δx) and the energy lost in the ECAL (ΔE) are the physical variables to be measured.
Collision and radiative processes, respectively dominating in the lowest and highest momenta, are properly taken into account in the analysis.

---

[3]At energies where radiative contributions are important, $a(E)$ and $b(E)$ are slowly varying functions of E.

4.3.2 The event selection

The muon momentum is measured from the track fit performed in the inner tracking system to profit from the high resolution of the silicon detectors and to perform the momentum measurement close to the inner surface of the ECAL crystals.

An initial sample of events is selected by requiring a single muon reconstructed in the whole detector, being associated to energy deposits both in the upper and in the lower half of the ECAL barrel and traversing the inner tracking system[4].

For a correct estimate, the muon momentum has to be measured upstream of the energy release in the calorimeter, so that only the energy deposited in the bottom half of ECAL are used and the E/P is required below 1 within the experimental accuracy.

The analysis is limited to muons whose momentum falls in the range from $p = 5$ GeV/c to $p = 1$ TeV/c. This interval is further divided in 25 bins of constant $\Delta \log(p)$ and events are rejected when the uncertainty on the momentum is larger than the momentum bin width ($\sigma_p/p <$0.275).

Finally muons are considered only if crossing the ECAL at an angle with respect to the crystal axis lower than 0.5 radians.

Overall the selection criteria resulted in a sample reduction of a factor 1.6, with a major effect due to the angle selection. For a proper treatment of containment effects, events with energy deposits above 500 MeV in the upper part of the ECAL barrel were vetoed (see Section 4.4.2), further reducing the sample by a factor 1.15.

The final sample consisted in 2.5×10^5 muons.

The muon spectrum as measured after the selections is shown in Figure 4.3 (a). The cut-off at low momentum is due to track reconstruction and extrapolation criteria. The spectrum of energy released and converted in crystals is shown in Figure 4.3 (b).

By fitting the tails of both the distributions with a $p_0 E^{p_1}$ function, the same spectral index ($p_1 = -2.592 \pm 0.009$ for the momentum spectrum and $p_1 = -2.605 \pm 0.016$ for the energy spectrum) was measured, as expected for bremsstrahlung photon radiation [86].

4.3.3 Track length and energy measurement in the ECAL

In order to estimate the muon path length in ECAL, each muon candidate is associated to a supercluster reconstructed in ECAL. This is performed through the geometrical matching between the supercluster barycenter position and the tracker track extrapolated on the ECAL inner surface, by requiring $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} <$0.1.

---

[4]A cut value is applied on both the transverse and longitudinal impact parameters ($|d0| <$100 cm, $|dz| <$100 cm).
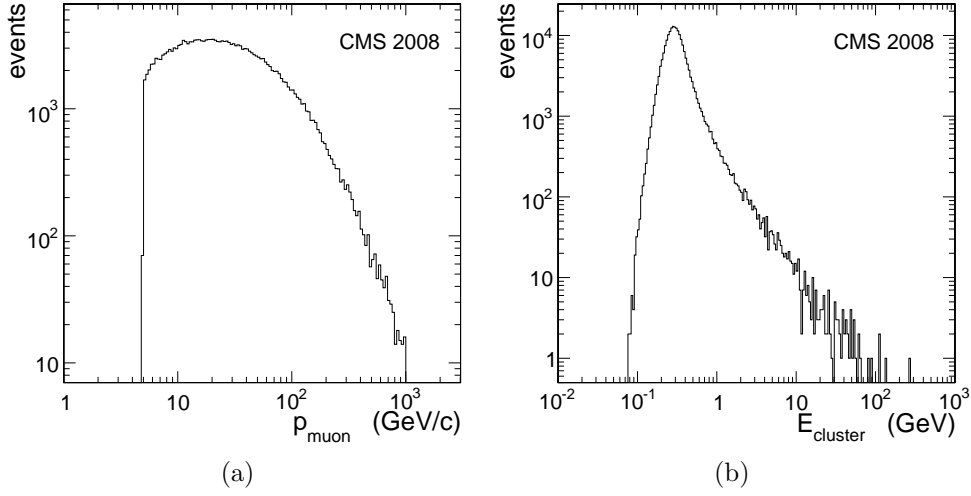
Figure 4.3: Momentum spectrum of the muons passing the selections (left) and spectrum of the reconstructed energy in the lower ECAL hemisphere (right). A logarithmic binning is used. (A m.i.p. muon traversing a crystal along its axis releases about 280 MeV on average.)

In Figure 4.4, $\Delta\eta$ is shown with respect to $\Delta\phi$ for all the events surviving the selection.
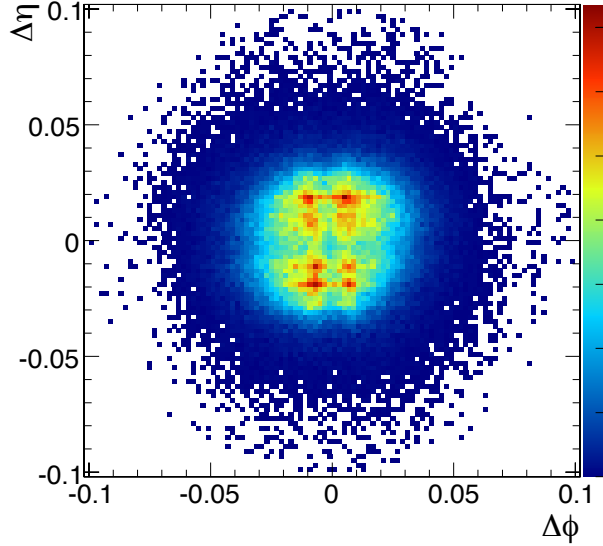

Figure 4.4: The radial distance between the barycenter of the energy deposition and the track position extrapolated at the calorimeter surface is shown in terms of $\Delta\eta$ vs $\Delta\phi$.

The muon track is then propagated inside the calorimeter, with a finite step procedure which takes into account the bending due to the magnetic field and the energy loss due to the material at each step. The track length is evaluated as the sum of the single steps propagated in the ECAL. In Figure 4.5, the obtained distribution is shown. The average

path length resulted 22.0 cm, with an *rms* spread of 2.6 cm, due to the angular distribution of the muon flux with respect to the crystal axis.
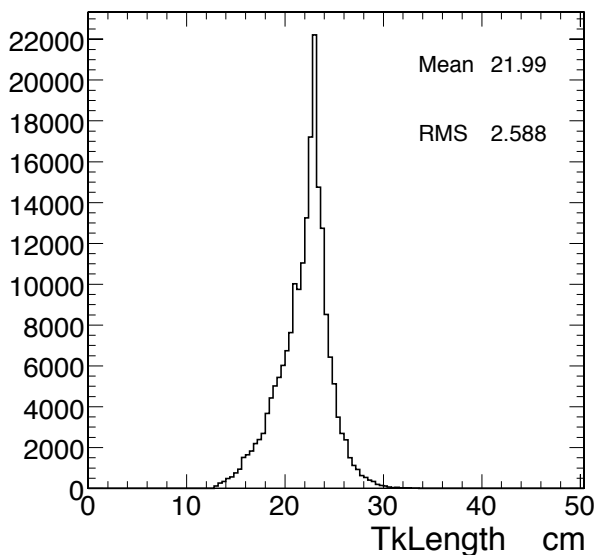

Figure 4.5: The average path length in ECAL as estimated from the track propagation.

In Figure 4.6 the distributions of $\Delta E/\Delta x$ as measured in ECAL are displayed, for a sub-sample of events with muon momentum below 10 GeV/c, where collision losses dominate (Fig. (a)), and for a complementary sample of events with muon momentum above 300 GeV/c, where sizable radiation losses are expected (Fig. (b)). The fraction of events with $dE/dx >$10 MeVcm^2/g is 1.3×10^{-3} and 0.08 in the left and right plots respectively.

4.4   Instrumental and containment effects

Multiple effects lead to systematic deviations of the measured deposited energy in the calorimeter from the energy lost by the muon, affecting the determination of dE/dx.

Instrumental effects and the physics processes involved in the muon energy transfer have to be considered differently depending on the contribution to the energy loss from the collision or irradiation regime.

In the following, the instrumental effects biasing the energy loss measurement in the low momentum range are discussed, together with the approach followed to treat the containment effects (see [87] for reference).
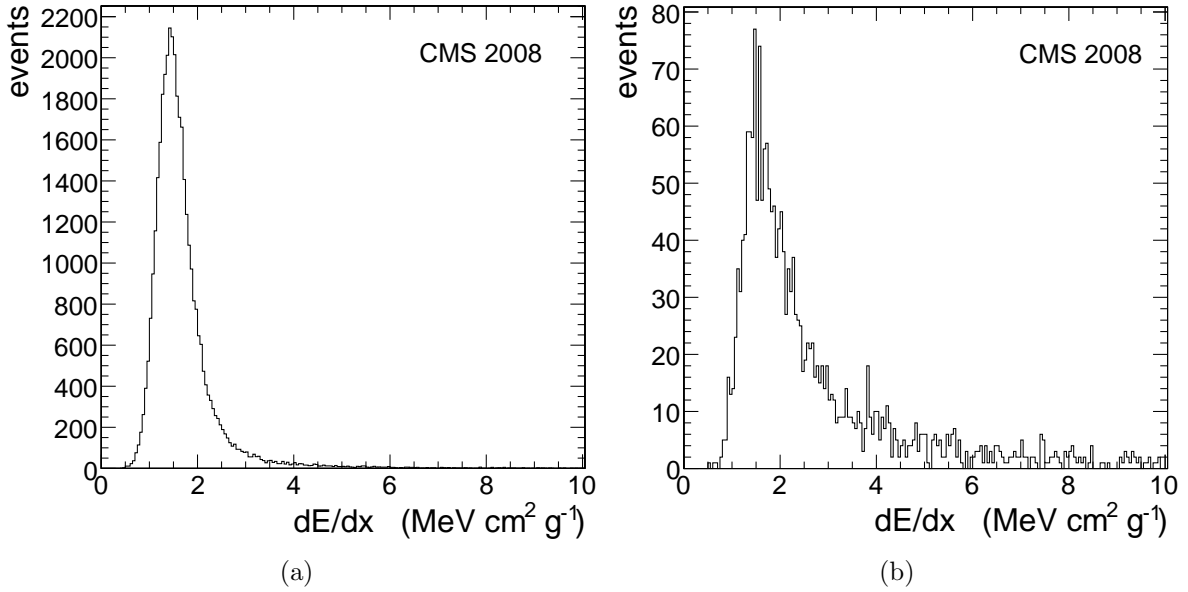
78

Figure 4.6: Measured distributions of $\Delta E/\Delta x$ in ECAL for muon momenta below 10 GeV/c (a) and for muon momenta above 300 GeV/c (b).

4.4.1 Instrumental effects

The relevant instrumental effects are related to the online data reduction and the energy reconstruction procedure, in terms of bias in the energy reconstruction introduced in presence of the particular thresholds used in the algorithms (Section 4.2.2).

For muons having a direction close to the crystal axis, the energy deposited in ECAL is likely to be above the selective readout (SR) threshold (90% of events with the angle between muon and crystal axis smaller than 0.1 radians). When such condition is met, a major bias arises from effects associated with the clustering threshold (18.6 MeV) and the noise spectrum ($\sigma_{noise} \sim 1$ ADC=9.3 MeV). The probability of noise fluctuation above the clustering threshold is measured in a dedicated analysis and a 14.7 MeV bias is estimated for muons at angles smaller than 0.1 radians in a 5×5 matrix of crystals (see Figure 4.7).

For muons at larger angles to the crystal axis, the average energy deposit per crystal decreases and both conditions for ZS readout (see Section 2.2.2) occur more frequently, with the higher ZS threshold reducing the noise bias, and the probability that the energy deposited in a single crystal is below the clustering threshold increases, thus contributing to a negative bias. The analysis of crystal multiplicity versus angle in experimental data indicates a constant plateau at small angles, followed by a monotonic increase at larger angles (Figure 4.8).

The bias is then expected to decrease with increasing angles, but its estimate depends on
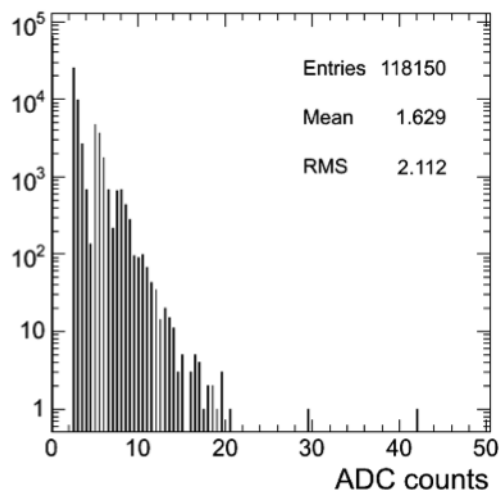
Figure 4.7: Under the hypothesis of selective readout, the noise contribution to the energy in a 5×5 matrix of crystals has been measured. The noise for each channel was simulated with a gaussian fluctuations of a 1 ADC σ_{noise} and only the contributions above the 18.6 MeV threshold were considered in the sum over the all 25 crystals. The result is reported in the spectrum shown. 1 ADC count is about 9.3 MeV.

several contributions and is less reliable than the estimate at small angles.
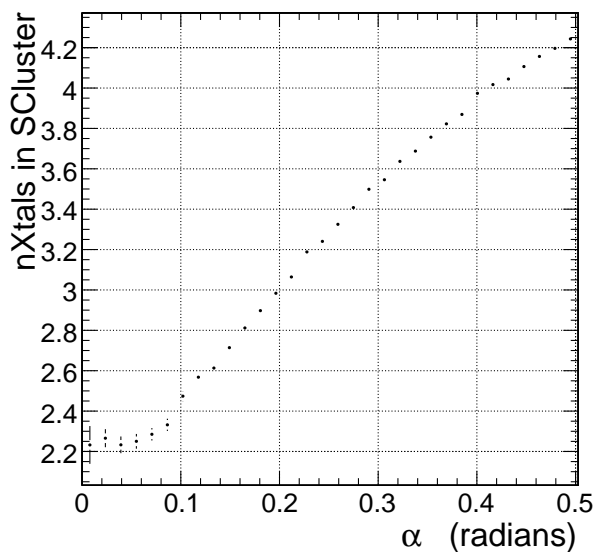
Figure 4.8: Crystal multiplicity versus angle from experimental data.

Considering the estimated offset to be negligible in the presence of the larger energy deposits associated with radiative processes, the angular dependence of the threshold effects has been further studied analyzing the dependence of the raw dE/dx on angle, for muons below 10 GeV/c. The result is shown in in Figure 4.9 and it is found consistent with the expectations. Nevertheless it is also consistent with a wider plateau, extending

up to 0.2 radians, or with a linear dependence of the bias on the angle, with no plateau. A correction is thus applied to the estimated collision component of the stopping power, relying on the fit to the data shown in Figure 4.9, assuming a plateau up to 0.1 radians and a decreasing linear trend at larger angles. Event by event, the measured dE/dx is normalized to the mean value from the fit in the plateau region and the 14.7 MeV bias is subtracted.

The cases of no plateau and of plateau up to 0.2 radians were considered as extreme cases in the evaluation of the systematic uncertainty, which is estimated to be 3.5 MeV, corresponding to about 1.2% of the average energy deposited by a muon in the collision regime.
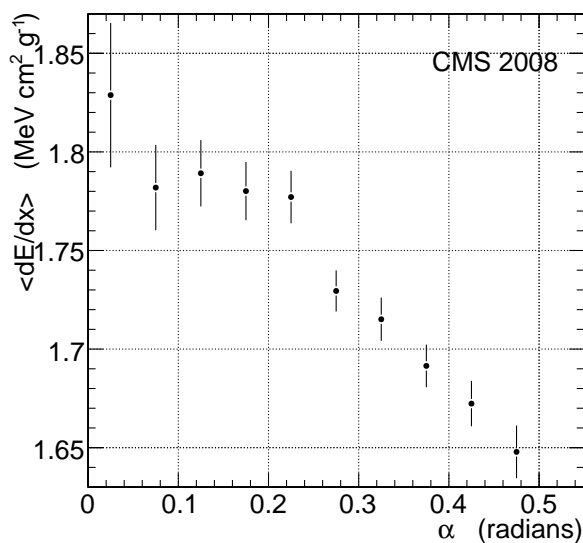

Figure 4.9: Dependence of the raw $\langle dE/dx \rangle$ on the angle α between the muon direction and the crystal axis, for muon momentum between 5 and 10 GeV/c. Vertical bars represent statistical errors.

4.4.2 The containment effect

The energy ΔE being considered in the theoretical approach is the energy lost by a muon along a given path in $PbWO_4$, while the first output of the measurement procedure is the energy deposited in a given volume of the calorimeter.

Since the energy lost by a muon is primarily transferred to particles having a finite range, there is in general a net energy flow through any closed surface, and the energy deposited in a finite volume is typically different from the energy lost by a muon traversing that volume.

If a cylinder volume with its axis along the muon direction is taken as a reference, secondaries are produced along the entire muon path, thus there is always a non-negligible flow

of energy carried by secondaries entering and leaving the two ends of the cylinder, while the energy flow across the lateral surface can be made negligible by choosing a sufficiently large radius.

In a homogeneous material an equilibrium is approximately reached between the energy flowing in through the front surface and the energy flowing out of the rear surface, when the front surface is preceded by an amount of material at least corresponding to the energy weighted average range of secondaries. Under these equilibrium conditions, the energy deposited in the volume is equal to the energy lost by the muon.

For δ electrons a condition near to equilibrium is reached with a preceding medium thickness of the order of 10 g/cm^2, while at least 20 radiation lengths are required in case of photon radiation.

Net containment corrections are thus expected to be small or negligible for collision losses, while sizable for radiative losses.

4.4.3 The containment corrections

A dedicated simulation based on the GEANT4 package [78] has been performed including bremsstrahlung, direct pair production and photo-nuclear interactions, to provide a quantitative estimate of the containment corrections [87], in terms of a difference between the energy lost and the energy deposited in the ECAL, relying on a simplified description of the detector and materials surrounding it.

Given the setup of the ECAL detector relevant for the CRAFT data taking and the presence of the strong magnetic field, material separated by more than a few tens of centimeters from the crystals is not effective in producing secondaries flowing in the measurement volume and the effect of the upper half of the ECAL was ignored.

To further reinforce the validity of this approximation only muons producing a signal of less than 500 MeV in the upper part of the ECAL were selected, thus removing most of those emitting a hard photon. Moreover, the energy deposited in the lower half of ECAL by muons releasing less than 500 MeV in the upper ECAL hemisphere (the sample here selected) was compared with the complementary sample, as a further check. The weighted average of the ratio of the dE/dx measured in the two samples in seven momentum bins, in the muon momentum region between 145 GeV/c and 1000 GeV/c, where radiative losses are relevant, is 0.95 ± 0.09, thus the complementary configurations are found compatible within the precision of the comparison. Given the modest statistical precision of this result, the veto in the event selection was conservatively applied, thus leading the PbWO$_4$ homogeneous region to be only preceded by order 10 g/cm^2 of effective nearby material.

As previously mentioned, energy leakage and effects of upstream and downstream materials are quantitatively different for collisions and radiative processes and are treated separately.

According to the simulation, the energy carried by secondaries flowing out of the rear detector surface is on average 3% of the energy lost in the crystals for a muon of momentum 15 GeV/c. This represents an upper limit to containment corrections in the low-momentum region, where collision losses dominate. By comparing the deposits in the upper and lower hemispheres of the ECAL barrel, for muons in the momentum range 5-10 GeV/c, no difference in dE/dx is found with a sensitivity better than 1% [80]. Net containment corrections are thus neglected for collision losses and 1% systematic uncertainty on the null correction is assumed.

To quantify the corrections for energy containment in radiative processes, dedicated simulations of two extreme cases were considered. No material in front of ECAL was assumed to represent the upper limit to the containment correction, where only rear losses are quantified. The case of the whole tracker material budget concentrated just in front of ECAL was used to give a lower limit to the correction, as it maximally overestimates the energy flow through the front face due to upstream material. The results of the two simulations are used to quote the correction for the energy containment (as average value) and the systematic uncertainty (as half of the difference between the two values). The computed result amounts to $(28.5\pm5)\%$ of the average energy lost at 1 TeV/c, while it reduces to $(14.5\pm2.5)\%$ at 170 GeV/c. Such precision is sufficient over the entire momentum range for the present analysis. The correction for non containment of radiated secondaries is applied over the whole momentum spectrum to the expected contribution of radiative processes to the stopping power.

4.5 Statistical analysis and results

In Figure 4.10 (a) the measured muon energy loss after the corrections is compared to the expected values as a function of the muon momentum [85] and in Figure 4.10 (b) the ratio between the experimental measurement and the expectation is shown. The two regions identified in the plots represent the minimum interval containing 68% of the probability distribution centered around the expected result (region delimited by the continuous curves) and the expected 68% C.L. probability central interval (grey shaded area).

Statistical errors have been derived with a numerical technique based on the simulation
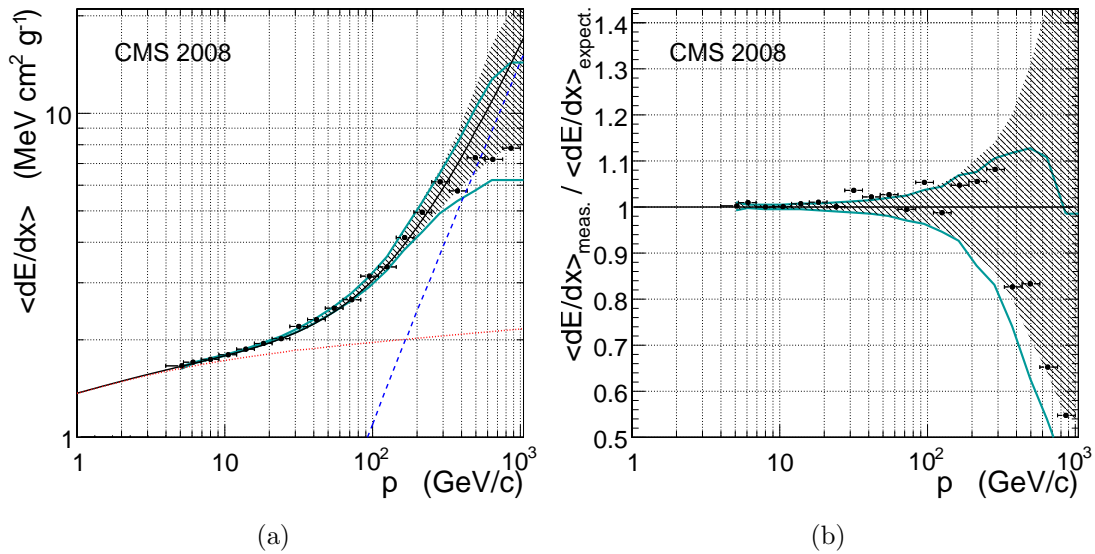
Figure 4.10: (a) Muon stopping power measured in $PbWO_4$ (dots) as a function of muon momentum compared to expectations [85] (continuous black line). The expected contributions from collision and radiative processes are shown as well (red dotted line and blue dashed line respectively). (b) Ratio of the measured and the expected values of the muon stopping power, as a function of muon momentum. In both figures, the shaded grey area indicates the expected 68% probability central interval, while the continuous cyan curves delimit the minimum interval containing 68% of the probability around the expected result [87].

of the expected result with GEANT4, to properly take into account the characteristics of radiative losses [87]. These correspond indeed to rare processes with high energy releases and generate probability distribution functions (*pdf*) of the energy loss per single event presenting long tails at high energy (see Figure 4.6).

In this condition, the average dE/dx of a sample of N events has a slow convergence to a normal distribution and the "statistical variance" of the sample cannot provide a reliable estimate of the statistical error on the measurement.

This is particularly relevant where radiation losses are important ($p > 170$ GeV/c) and the "sample mean" and the "sample variance" are highly correlated and sensitive to rare events with very large energy release. For the high momentum bins, due to the increasing probability of large fluctuations in radiative energy losses and the decreasing size of the sample statistics, the "experimental error" ($rms/\sqrt{N}$) is not a good estimator for the consistency of the "sample mean" with the expected value. This trend is particularly marked in the highest momentum bin of Fig. 4.10, where the expectation value lies outside the 68% probability interval.

GEANT4 was used to simulate 10000 pseudo-experiments per momentum bin, each with the same statistics as the actual bin population. For each bin, the expected *pdf* of the

mean was obtained as the distribution of the mean values of the stopping power from the different experiments. For each *pdf*, two 68% probability intervals for the expected result were derived: the central interval, obtained by discarding 16% of the result on both tails of the *pdf*, and the interval of minimum width containing 68% of the result.

4.5.1 Results

The curve

$$(dE/dx)_{meas} = \alpha \left[\left(\frac{dE}{dx} \right)_{coll} + \beta \times \left(\frac{dE}{dx} \right)_{rad} \right] \qquad (4.4)$$

where *coll* and *rad* label the predicted energy losses in $PbWO_4$ due to ionization and radiative processes (bremsstrahlung, pair and photo-nuclear production) respectively [2], is fitted to experimental stopping power data. The parameters α and β account for the overall energy scale and for the relative contribution of radiation to collision losses. With the adopted parameterization, β allows for the measurement of the critical energy, without depending on the overall energy scale. The fit result is

$$\begin{aligned} \alpha &= 1.004^{+0.002}_{-0.003}(stat.) \pm 0.019(syst.) \\ \beta &= 1.07^{+0.05}_{-0.04}(stat.) \pm 0.06(syst.) \end{aligned} \qquad (4.5)$$

The muon critical energy

A muon critical energy of $160^{+5}_{-6}(stat.)\pm8(syst.)$ GeV is derived, in agreement with the computed value of 169.5 GeV for $PbWO_4$ [2], assuming as definition of muon critical energy the energy at which the collision loss rate equals the radiative loss rate, including bremsstrahlung, direct pair production and photo-nuclear interactions in the radiative processes.

The systematic uncertainty includes a contribution of 4.5 GeV from the uncertainty of the containment corrections, dominated by the limited knowledge of the correction for radiation losses, and a contribution of 6 GeV due to the stability of the result against event selections. In the latter case, sizable contributions come from the variation of the acceptance of the muon angle with respect to the crystal axis, which was varied between 0.3 and 0.7 radians, and from the requirement that E/p be lower than 1.

The ECAL energy scale

As previously discussed, the ECAL energy scale was set from test beam measurements, using a 120 GeV/c electrons beam.

Before the installation in the cavern, the ECAL scale linearity has been checked with test-beam operations and with cosmic muons data.

At the ECAL-HCAL combined test-beam the ECAL scale was controlled better than 1% in the energy range 2-100 GeV, using electron beams with energies in the range 2-9 GeV/c and positron beams with energies between 9 and 100 GeV/c [88].

At the ECAL dedicated test-beam, where 9 super-modules were exposed to electrons in the momentum range 15-250 GeV/c, the maximum deviation observed over the 20-180 GeV range is of the order of 0.2%.

All the ECAL barrel super-modules were exposed to cosmic muons in the dedicated test bench prepared in the CERN North Area [84], with the purpose to extend the high precision measurements performed on very few super-modules at the test-beams to the whole EB. In particular, the ECAL energy scale set with 120 GeV/c electrons was compared to the one measured with exposing the single super-modules to cosmic muons, giving a difference smaller than 0.4% [89].

In this analysis, the ECAL energy scale is measured after installation in the cavern 100 m underground, with the whole CMS operating to take data.

With respect to the α parameter reported in 4.5, the ECAL energy scale is measured consistent with unity within a 1.9% error, which is completely dominated by the systematic uncertainties[5].

The result here presented is mainly determined by the precision of the measurements in the muon momentum region below 20 GeV/c, where radiation losses are marginal and the typical energy released by muons is about 300 MeV.

This result allows to test the ECAL energy scale set with 120 GeV/c electrons at the test-beam, in the sub-GeV region while operating in the CMS cavern, with a precision of 1.9%.

---

[5]The energy scale dependence on the angle and on the clustering, as well as the uncertainty on the containment corrections.

Commissioning of the electron seeding with the first data

With the LHC changing its operating conditions and entering the high energy collision era, a continuous monitoring of the detector response as well as the physics objects performances is required. The commissioning activity is fundamental to identify and correct for eventual effects not foreseen in the simulated studies and thus to provide solid elements to carry on reliable physics analyses. A first purpose is to validate the modeling of the detector with data, through a detailed comparison between data and simulation. At the same time the reconstruction algorithms are validated, by checking the performances of the basic ingredients which lead to the final physics objects.

In this context, my primary activity was focused on the commissioning of the ecal-driven electrons as reconstructed with the CMS software, with particular attention to the very first stage of the reconstruction process which is the seed finding in the inner tracker layers.

In the following, the ecal-driven seed finding algorithm is presented (Section 5.1) and the work I did to commission the algorithm is discussed, from the initial studies performed on the first collision samples of minimum bias events (Section 5.2), moving to those obtained with samples of W-tagged events (Section 5.3).

5.1 The ecal-driven seeding

Within the electron reconstruction (described in Section 3.2.2), the seed finding algorithm represents the first connection step between the energy as measured in ECAL and the trajectory as built from the interaction point up to the calorimeter.
Seeds are indeed pre-tracks made of a pair (or triplet) of measured hits in the inner tracker

layers, found compatible with the electron supercluster.

The key concept of the specific ecal-driven strategy is that the barycenter position of the electron supercluster reconstructed in ECAL lies on the helix of the initial electron trajectory and does not depend on the bremsstrahlung photon emission.

In this way, the inner trajectory can be predicted by back-propagating the helix parameters through the magnetic field towards the innermost detector region, where radiation is unlikely to have occurred, and seeds for electron tracks can be identified by looking for the inner hits measured in the tracker, if found in a region compatible with the expectation.

5.1.1   Search windows and measured hits

The electron seed finding algorithm starts with all the trajectory seeds built from tracker hits pairs and triplets to select those that may correspond to an electron track. The selection is based on the tracker hits constituting the seeds, which are required to match the measured ECAL supercluster. The seed finding algorithm starts from the search of the first hit, namely the innermost measured hit. For this purpose an helix is drawn in the 3.8 T magnetic field from the barycenter of the supercluster in ECAL to the nominal $(0, 0, 0)$ interaction vertex, using the measured transverse energy to determine the curvature. This is done for each charge hypothesis since the electron charge information gets available only later on in the reconstruction process[1].

Once a first hit compatible with the extrapolated trajectory is found, it is used together with the supercluster information to update the z vertex coordinate. Then starting from the revised vertex, a second helix is built passing through the first measurement and reaching the supercluster, by assuming the same charge hypothesis which led to the first hit.

Figure 5.1 illustrates the strategy used by the seed finding algorithm to select trajectory seeds suitable for electrons, projected in the r_T, ϕ plane.

First and second hits are looked for in search windows defined in the z, ϕ (barrel region) or r_T, ϕ (forward region) planes around the extrapolated positions.

In Figure 5.2 the geometric definition of the expected windows is shown with respect to the projection in the z, r_T plane. A quantitative formulation of the expected windows width is given in 5.1.

---

[1]Three different algorithms are used in CMS to determine the electron charge [90] which rely on the GSF-track charge, the general track charge and the charge as determined by the supercluster position.
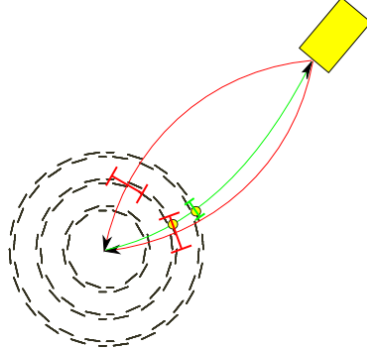
Figure 5.1: Filtering of trajectory seeds by the matching of their hits in the barrel layers, by using the supercluster and the vertex constraints. The two red trajectories are used in the first hit search. The green trajectory connects the vertex with the supercluster position, passing through the first measured hit and is used in the second hit search.
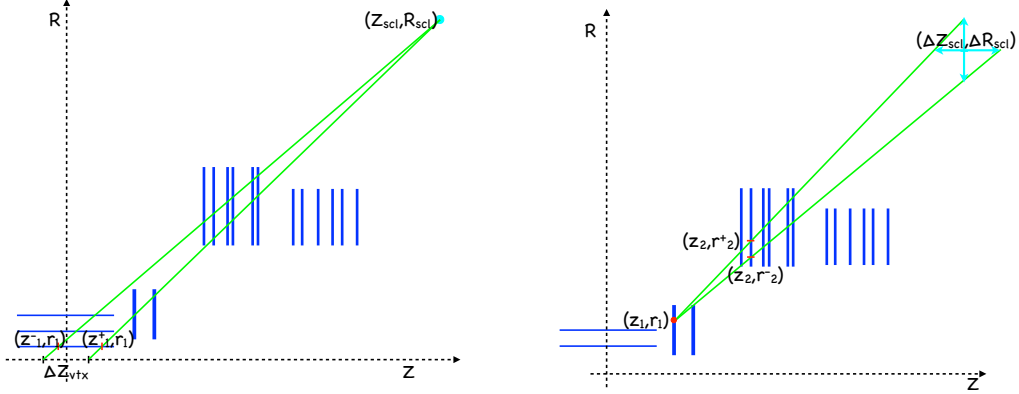

Figure 5.2: On the left: search windows to look for 1st hits (z_1, r_1). To compute the ΔZ_1 expected windows in the z, r_T projection, the vertex uncertainty (ΔZ_{vtx}) and the supercluster position (Z_{scl}, R_{scl}) are the constraints used. On the right: search windows to look for 2nd hits (z_2, r_2). To compute the ΔR_2 expected windows in the z, r_T projection, the supercluster uncertainty (ΔR_{scl}) and the 1st measured hit (z_1, r_1) are the constraints used in this case.

$$\Delta Z_1 = \Delta Z_{vtx}(1 - r_1/R_{scl}) \quad \text{and} \quad \Delta R_1 = R_{Scl} \cdot \Delta Z_{vtx} \frac{Z_{scl} - z_1}{Z_{scl}^2 + (\Delta Z_{vtx}/2)^2} \quad (5.1)$$

$$\Delta Z_2 = \Delta Z_{scl} \frac{r_2 - r_1}{R_{scl} - r_1} \quad \text{and} \quad \Delta R_2 = \Delta R_{scl} \frac{z_2 - z_1}{Z_{scl} - z_1}$$

Combinations of first and second hits are looked for in the barrel pixel layers (BPix), the forward pixel disks (FPix) and TEC layers [91][90], so to improve the coverage in the forward region.

To limit the number of fake reconstructed electrons, only the superclusters with $E_T >4$ GeV are considered and $H/E <0.15$ GeV is required. Moreover, the first ϕ window is made $1/E_T$ dependent within the range $5 < E_T < 35$ GeV.

To look for the first hits, the search windows in ϕ are defined asymmetric as a function of the charge hypothesis, in order to account for the curvature uncertainty due to bremsstrahlung effects in the E measurement, and in z are defined loose to account for the beam spot[2] extension (σ_z) along the z axis.

Smaller windows are instead used to look for the second hit, in particular in the z, r_T plane.

The search area optimized in the simulation were found overall well suited for the reconstruction of prompt and isolated electrons in real data. Within the studies presented in this Chapter, the ϕ window size for the second hit was increased, to prevent from a small loss of efficiency in the forward region.

The current search windows, used to look for the first and second hits of the reconstructed trajectory seeds, are presented in Table 5.1 and Table 5.2 respectively. The windows widths are quoted both in the ϕ and z projection (r_T instead of z in the forward region).

	δz (BPix)	δr_T (FPix or TEC)	$\delta\phi$ (pos. charge)	$\delta\phi$ (neg. charge)
5 GeV/c	$\pm 5\sigma_z$	$\pm 5\sigma_z$	[-0.075;0.155] rad	[-0.155;0.075] rad
10 GeV/c	$\pm 5\sigma_z$	$\pm 5\sigma_z$	[-0.046;0.096] rad	[-0.096;0.046] rad
35 GeV/c	$\pm 5\sigma_z$	$\pm 5\sigma_z$	[-0.026;0.054] rad	[-0.054;0.026] rad

Table 5.1: Definition of the seed search windows for the first hit. Outside the range $5< E_T <35$ GeV the ϕ windows are fixed to the values corresponding to $E_T =5$ GeV or $E_T =35$ GeV. σ_z is the beam spot width along the z axis.

δz (BPix)	δr_T (FPix)	δr_T (TEC)	$\delta\phi$ (BPIX)	$\delta\phi$ (FPIX or TEC)
± 0.9 mm	± 1.5 mm	± 2 mm	± 4 mrad	± 6 mrad

Table 5.2: Definition of the seed search windows for the second hit.

The same seed finding algorithm here described is applied at the HLT [93] to allow for an efficient filtering of the genuine signatures from jets faking electrons. Also the improvements achieved with the studies performed on the offline reconstruction presented in this Chapter were updated in the online algorithms so to keep the HLT algorithms as close as possible to the ones used offline.

---

[2]The beam spot is measured in CMS as the mean value of the vertex position distribution averaged over ~1000 events, using the only pixel-tracks [92]. Its size as determined from the fit resulted to be 5.7 cm along the z axis, in the considered period of data taking.

The residual distances between the measured and the extrapolated positions ($\Delta\phi$, Δr or Δz)

$$
\begin{aligned}
\Delta\phi_1 &= \phi_{1^{st}hit}^{meas.} - \phi_{1^{st}hit}^{extr.} & \Delta z_1 = z_{1^{st}hit}^{meas.} - z_{1^{st}hit}^{extr.} & \quad \Delta r_1 = r_{1^{st}hit}^{meas.} - r_{1^{st}hit}^{extr.} \\
\Delta\phi_2 &= \phi_{2^{nd}hit}^{meas.} - \phi_{2^{nd}hit}^{extr.} & \Delta z_2 = z_{2^{nd}hit}^{meas.} - z_{2^{nd}hit}^{extr.} & \quad \Delta r_2 = r_{2^{nd}hit}^{meas.} - r_{2^{nd}hit}^{extr.}
\end{aligned}
\tag{5.2}
$$

are good estimators of the seed finding performances. Such quantities are computed for each measured hit and are stored in each ecal-driven seed object by the electron reconstruction algorithm.

5.1.2 The seed cleaning

The obtained ecal-driven seeds collection is merged with the tracker-driven one (Section 3.2.2), in order to run the GSF tracking on the resulting unique collection. As previously mentioned, to prevent the tracking algorithm from running on the same eventual duplicated input seeds, the merged seeding collection is filtered with a seed-cleaning. Essentially among the seeds saved, only those not sharing hits with previous ones are taken into account in the next steps of the reconstruction. If the same hits are measured for the two different charge hypotheses, the second seed is removed and replaced by the first one and the information that the seed has been found by both algorithms is stored. In some cases, such a cleaning appeared to lead to a wrong association of the angular distance $\Delta\phi$ (eq. 5.3) to the hit saved. $\Delta\phi = \phi^{meas.} - \phi^{extr.}$ is indeed computed for each hit and charge hypothesis and the expected hit position depends on the charge assumed in the extrapolation (information which is not stored). In any case the proper functioning of the overall electron reconstruction algorithm is not affected by this possible wrong assignment of the $\Delta\phi$ window to the measured hit, since the stored angular distances are only used for monitoring purposes. Thus only the monitoring distributions are affected by the loss of the charge hypothesis information. This particular effect will be recalled and commented in Section 5.3.2, in the context of the validation studies here presented.

5.2 The electron seeding commissioning with minimum bias events

5.2.1 First studies at 900 GeV

At the end of 2009 at the beginning of its operation, the Large Hadron Collider provided each experiment proton-proton collisions at the centre-of-mass energies of 0.9 and 2.36

TeV, for a total integrated luminosity of $\sim 10\mu b^{-1}$ and $\sim 4\mu b^{-1}$ respectively. Data collected at $\sqrt{s}$ =900 GeV were first exploited for the commissioning of the electromagnetic physics objects.

For comparisons with real data, about 10M Monte Carlo minimum bias events were generated and reconstructed for both 900 GeV and 2.36 TeV centre-of-mass energies. Reconstructed electron candidates in such minimum bias events are expected to mainly origin from "fakes" from charge hadrons, with a signal from "real" electrons mostly constituted by pairs from converted γ from π^0 decays.

Figure 5.3 presents the expected η and p_T distributions of the reconstructed electron candidates compared to data. The MC distributions are normalized to the total number of electrons in the data. The contribution of the ecal-seeded electrons is shown in the top figures, while in the bottom ones it is shown the MC expectation from the matching with generator level particles (electron, γ,π^0). A cone size of ΔR =0.15 is used for this.

With reference to Figure 5.3 a), the majority of electron candidates in the very low p_T range is tracker-seeded and the lower part of the spectrum is explained by the threshold at 2 GeV/c in the track reconstruction algorithm. In the same distribution, the threshold for the superclusters E_T in the reconstruction of the ecal-driven electrons is also clearly visible. Below such 4 GeV/c value the tracker-driven electrons entirely dominate the distribution of the reconstructed candidates, while above the majority of the candidates is ecal-driven.
The η distribution rises in the ECAL endcaps except for the ecal-driven electrons, which are particularly sensitive to the larger fake rate in the barrel region starting from $|\eta| > 1$.

The distributions of the distance of closest approach to the beam spot both in the transverse plane and along the z axis are presented in Fig. 5.4, showing the different contributions from the MC expectation. The large tails in the distribution of the impact parameter in the transverse plane are due to conversions.

The electron candidates in data and Monte Carlo, are essentially background from fake and real electrons from conversions. Nevertheless such sample allowed to assess at a first level the general quality and the proper functioning of the algorithms and the modeling of the detector response in the simulation [94].

To focus on the performances of the ecal-driven seed finding (described in the first Section 5.1), the residual distributions of the differences between the measured and the expected hit positions on the tracker layers (defined in equation 5.3) are looked at.
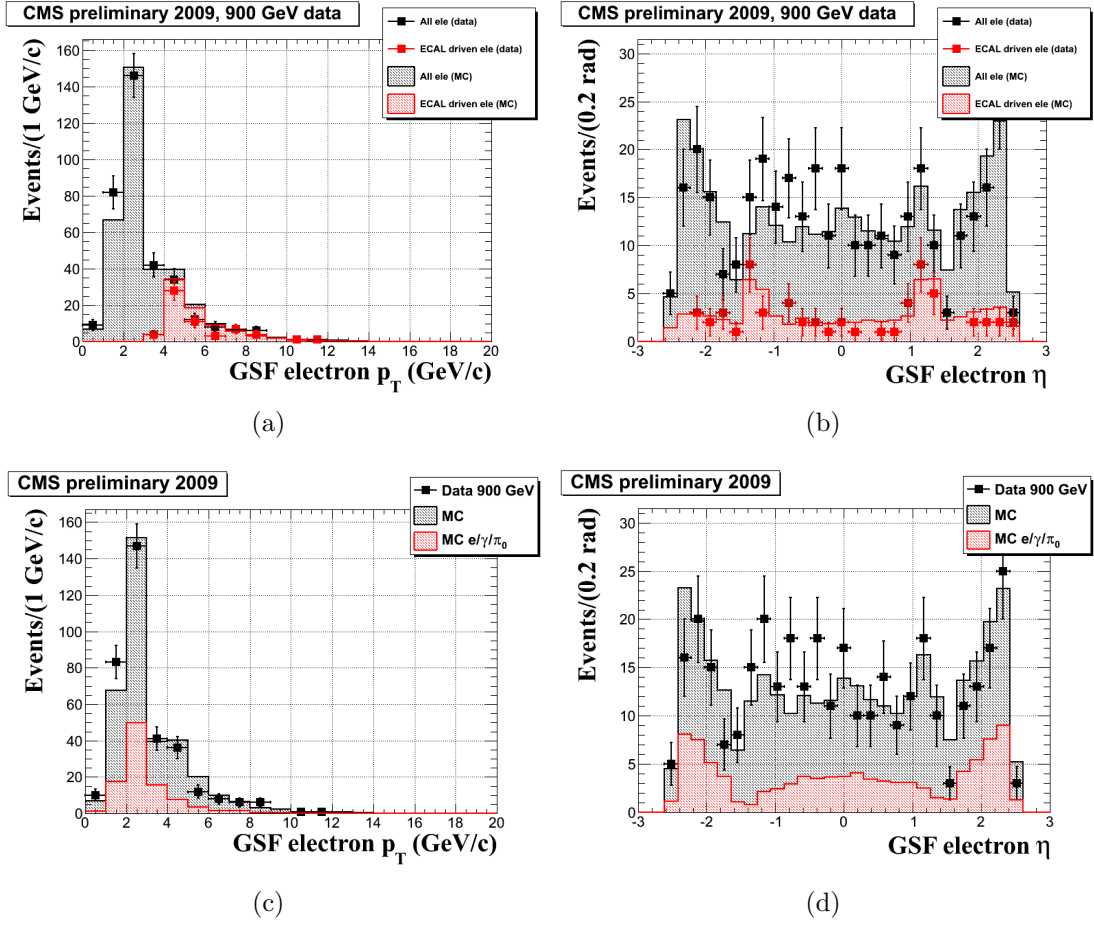
Figure 5.3: Distributions of (a,c) electron transverse momentum and (b,d) electron pseudo-rapidity. MC distributions are normalized to the total number of electrons in the data. In (a, b) the contributions of all and ecal-driven electrons are shown, while the MC expectation for electron candidates matched to either a generated electron, a γ or a π^0 is shown in(c,d).

In order not to bias the tails of such distributions, in the original processing of the data the seeding search windows used in the electron reconstruction were made very large. Results obtained with such processing of the data will be referred to as "loosened".

As previously described, the matching windows used to look for the first hits of the seeds are wide enough to guarantee high efficiency, with margin for some contamination from fakes. For the commissioning purpose the interesting residual distributions are those relative to the second hits, because the tighter cuts select higher purity electron candidates, showing eventual problems with clearer evidence and thus allowing for a prompt intervention.

In the following, among the residual distributions defined in equation 5.3, only the ones
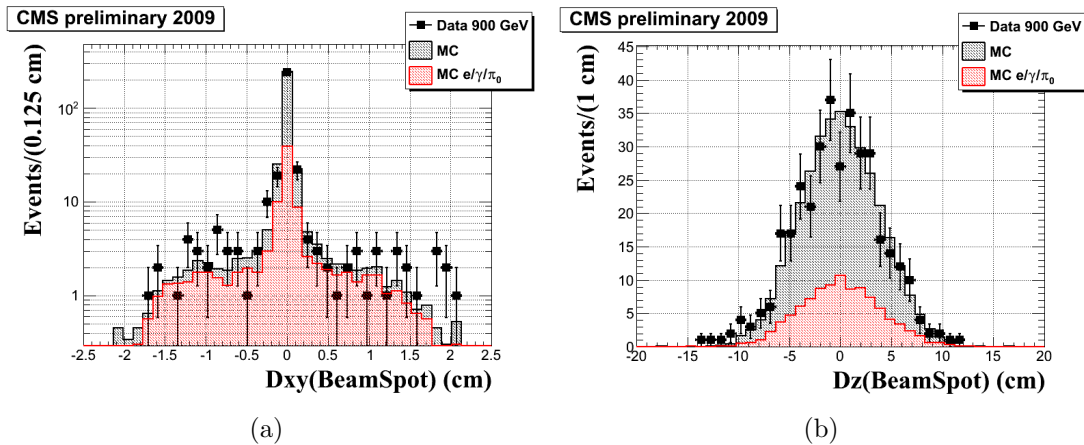
Figure 5.4: Distributions of (a) the distance of closest approach to the beam spot in the transverse plane and (b) along the z axis. The Monte Carlo expectation for electron candidates matched with a true generated electron or a γ or a π^0 is visible.

measured for the second hits will be considered and are reported here for convenience.

$$\Delta\phi_2 \;\; = \;\; \phi_{2^{nd}hit}^{meas.} - \phi_{2^{nd}hit}^{extr.} \quad \Delta z_2 = z_{2^{nd}hit}^{meas.} - z_{2^{nd}hit}^{extr.} \quad \Delta r_2 = r_{2^{nd}hit}^{meas.} - r_{2^{nd}hit}^{extr.} \quad (5.3)$$

Figure 5.5 presents the distributions obtained with the "loosened" reconstruction, for both $\phi - z$ and $\phi - r_T$ coordinates, separately for barrel pixels and forward layers (forward pixels and TEC layers). The Monte Carlo distributions are normalized to the total number of electrons in the data.

An overall good agreement is visible in the distributions, with the flat trend in the tails and very large peaks due to the important contributions from fakes and conversions.

Similar distributions for the ecal-driven seeding parameters as obtained using the full reprocessing of the electron reconstruction are presented in Fig. 5.6 for the barrel pixel case only due to the very low statistic.

5.2.2   Initial studies with 7 TeV collisions

In April 2010, the LHC started its collision operations at $\sqrt{s} = 7$ TeV providing in few months the first few pb^{-1} of integrated luminosity used by CMS to finalize the electron commissioning.

With about 2/nb of integrated luminosity from minimum bias collisions, a first comparison was carried on for variables sensitive to the track seeding algorithm. Some discrepancies
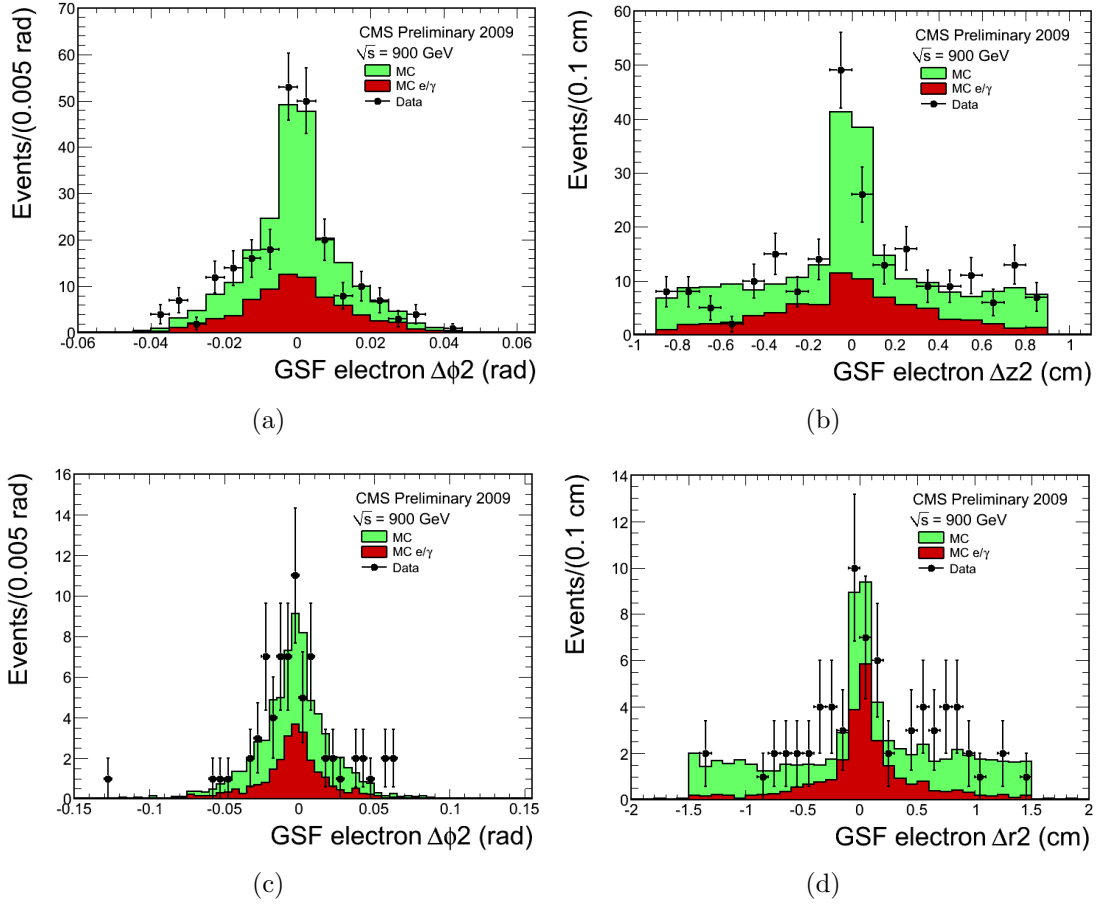
Figure 5.5: $\Delta\phi_2$ and Δz_2 (or Δr_2) residuals measured for the second hit of ecal-driven electrons reconstructed in the "loosened" configuration: (a) pixel barrel, ϕ coordinate (b) pixel barrel, z coordinate (c) pixel forward and TEC, ϕ coordinate (d) pixel forward and TEC, r_T coordinate. The expected contribution from the matching with generator level particles is also shown.

between the simulated collision parameters and the real ones were found and corrected for.

The standard reconstruction was applied to both data and the simulated minimum bias sample. Also in this case the MC sample is mostly constituted from "fakes" from charged hadrons (61.5%) and from electrons from conversions (33.9%). "Real" electrons (~4.6%) are composed from heavy flavors (34.9% of Ds, 46.7% of Bs), Dalitz decays (13.5%) and few J/Ψs. The electrons selected cover the p_T range above 5 GeV/c.

In figure 5.7 the residual distributions for the second seeding hits in the pixel barrel and forward disks are presented with data points over-imposed to the shape from MC. In the simulation, the reconstructed electrons are matched within a $\Delta R < 0.05$ cone to the particles at the generator level and the different contributions are visible with different colors.
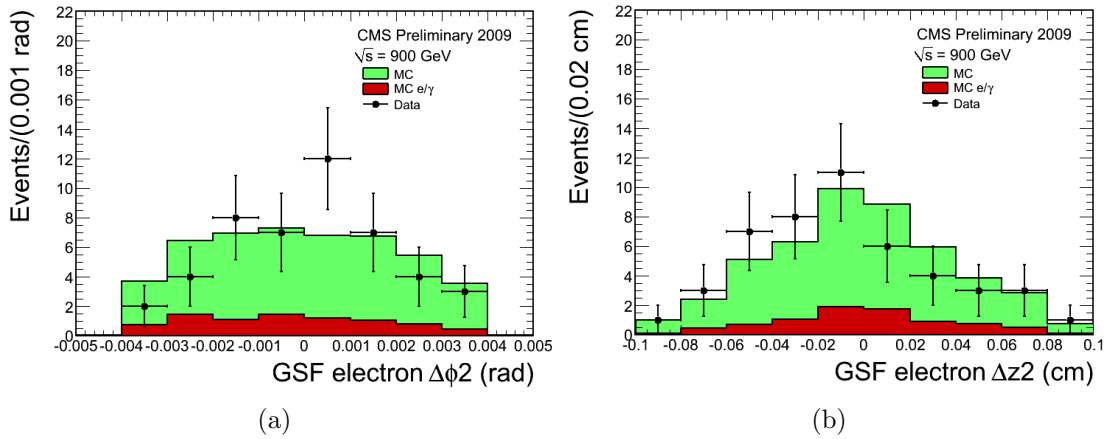
Figure 5.6: $\Delta\phi_2$ and Δz_2 (or Δr_2) residuals measured for the second hit of ecal-driven electrons reconstructed with the standard configuration: (a) pixel barrel, ϕ coordinate (b) pixel barrel, z coordinate. The expected contribution from the matching with generator level particles is also shown.
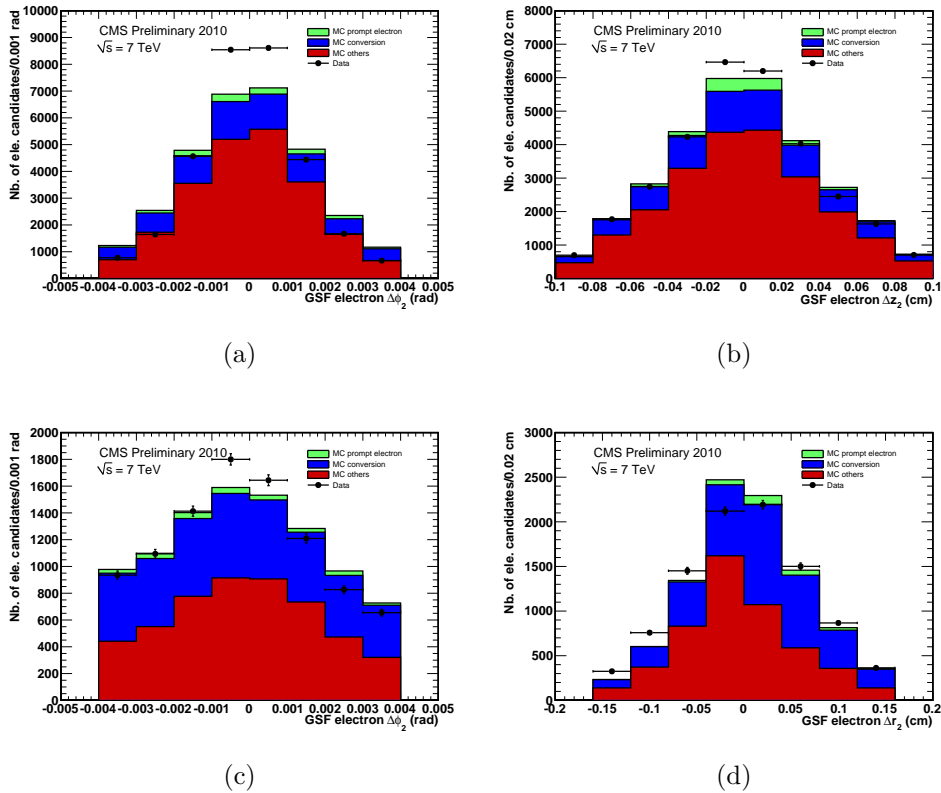

Figure 5.7: $\Delta\phi_2$ and Δz_2 (or Δr_2) residuals measured for the second hit of ecal-driven electrons: (a) pixel barrel, ϕ coordinate, (b) pixel barrel, z coordinate, (c) pixel forward, ϕ coordinate and (d) pixel forward, r_T coordinate. Expected contributions from electron candidates matched to a generated electron, a γ conversion and to other generated particles are shown. A $p_T > 5$ GeV/c cut is applied on the reconstructed electrons.

A bad agreement between data and MC is visible for the seeds belonging to the barrel part and a clear asymmetry emerges in the Δr_2 distribution in the forward pixel disks.

The spotted MC data discrepancies originated from an inaccurate simulation of the real conditions, which undergo a quick evolution.
With respect to this, a first basic variable checked is the beam spot (BS), as being a first input to the seed-finding algorithm. The BS is measured in CMS as the mean value of the vertex position distribution averaged over ~1000 events, using the only pixel-tracks [92]. As documented in [90], an inaccurate beam spot modeling brings to a visible inefficiency in the final electron reconstruction (order 15% drop in efficiency for a BS offset in the y direction of 1 mm).

In Figure 5.8, the simulated beam spot widths in the transverse (x,y) and longitudinal (z) planes are compared to the ones measured from data. The simulated distributions appear indeed broader, explaining part of the disagreements noticed and at the same time proving how the LHC started its collisions in better working conditions than expected.
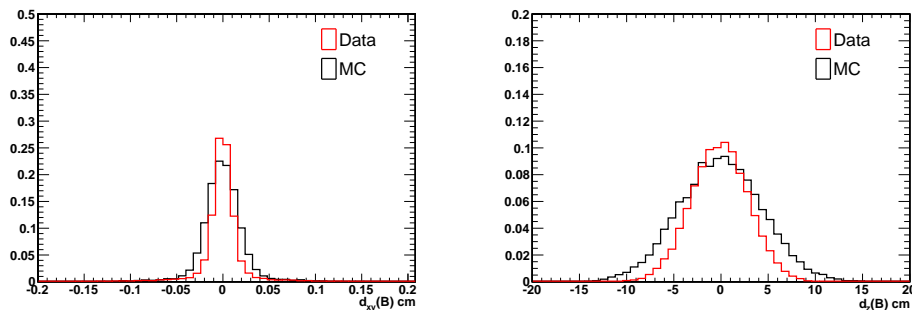

Figure 5.8: Beam spot distribution width both in the transverse (left) and longitudinal (right) plane, for data and simulation. The distributions are shown normalized to unity.

Despite the low statistic and the high contribution from fake electrons, this initial data/MC comparison (Fig 5.7) gave us a glimpse on necessary adjustments to be made to the seed reconstruction.

The excess of data in the left tail of the Δr_2 distribution in the pixel forward as well as the asymmetry in the $\Delta \phi_2$ are the aspects discussed in the following.

5.3 The electron seeding commissioning with W tagged events

With the increasing statistic at $\sqrt{s}$=7 TeV, the rediscovery at CMS of the first W bosons allowed the commissioning activity to be performed with prompt and isolated electrons of interest provided by W decays.

Starting with almost 200/nb of integrated luminosity available by July 2010 and used for a first set of commissioning results [95], the full validation of the electron objects was possible with an integrated luminosity of $\sim$3 pb^{-1} and the final electron track seeding was commissioned in the barrel region for p_T >25 GeV/c with 14.5 pb^{-1} of 2010 data.

For comparison with real data, simulated samples of W events based on the official CMS Monte Carlo production were used, with support provided by a private reprocessing to optimize and verify the required corrections when needed. Background events from QCD have been simulated and reconstructed as well.

To validate the seeding windows size for "good quality" high p_T electrons, a W event selection was applied to both data and MC and the selected electron candidates from $W \rightarrow e\nu$ were further required to satisfy some stringent conditions on the isolation and identification.

5.3.1 Sensitivity to the detector alignment

At the beginning of the data taking, the ECAL endcaps were found few millimeters misaligned with respect to the tracker, leading to strong differences in the track-cluster geometrical matching variables and in particular $\Delta\eta$.

The seeding algorithm proved really sensitive to such initial misalignment, due to the track-cluster matching on the innermost tracker layers which is performed while looking for ecal-driven seeds.

The reference distributions are shown in Figure 5.9, in the comparison between the data points and the gray dashed histograms.

The distributions present the difference between the measured and the predicted hit position in the second layer of pixels for electrons with ecal-driven seeds. The data correspond to an integrated luminosity of $\mathcal{L}_{int} = 198\, nb^{-1}$, and the distributions from MC are normalized to the total number of events in the data. The background from QCD di-jet events is negligible.

The misalignment effect is clearly visible in the forward region, in particular for the

$r_T - z$ projection, where the data shape is completely deformed with respect to the ideal alignment condition simulated in the Monte Carlo.

First quantitative measurements of such misalignment were provided by preshower studies [96] and were used to produce a low statistic MC sample with the misaligned detector conditions as measured in the data, to allow for a prompt feed-back from the commissioning activity.

Once the measured ECAL misalignment is taken into account (blue histograms, on the same plot in Figure 5.9), the agreement with data points is much better, although at this stage the tune is still not perfect.

The alignment constants were updated afterwards based on the analysis of $\sim$2.7 pb^{-1}, by looking at the track-cluster geometrical matching variables $\Delta\eta$ and $\Delta\phi$ distributions of electron candidates from tagged W events [97].
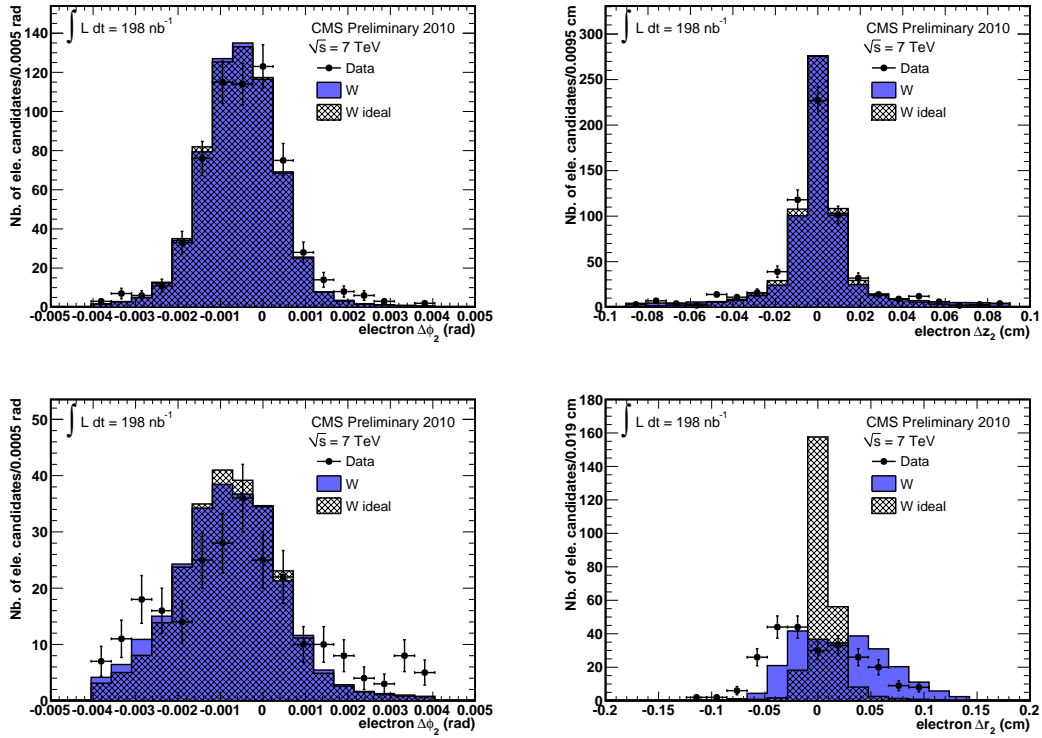

Figure 5.9: $\Delta\phi_2$ and Δz_2 (or Δr_2) residuals measured for the second hit of ecal-driven electrons in BPix (top) and forward region (bottom), for the $r_T - \phi$ (left) and $r_T - z$ (right) projection. The MC distributions are shown for a perfect alignment (gray dashed histogram) and for the alignment as determined from the initial preshower measurement (blue filled histogram). The simulation distributions are normalized to the number of events in the data sample.

A further problem was highlighted in the seeding distributions looking at the $r_T - \phi$ plane. A bias was observed in the residual distributions for the ϕ coordinate in both data and simulation, thus suggesting an issue in the reconstruction algorithm.

This effect, clearly visible in the central detector part, is further enhanced moving to the forward region (Fig. 5.9 top left compared to bottom left).

Such an offset was identified as due to the displacement of the beam spot with respect to the origin of the CMS reference frame [98], which was not taken into account when determining the ϕ angle between the measured and the extrapolated hit positions.

In Figure 5.10 the seeding distributions are compared between two MC simulations, using either (0,0) or the Beam Spot transverse coordinate as reference position to evaluate the ϕ angle. The bias is corrected by using the proper displaced BS position in the transverse plane and the clear improvement obtained in the $r_T - \phi$ projection is visible also for the distributions of the $r_T - z$ windows.

The effect of the updated reference frame for the angle computation is clearly visible in Figure 5.11, where the two compared trends refer to the output of the standard electron reconstruction algorithm in CMS, before and after implementing the proper corrections[3].

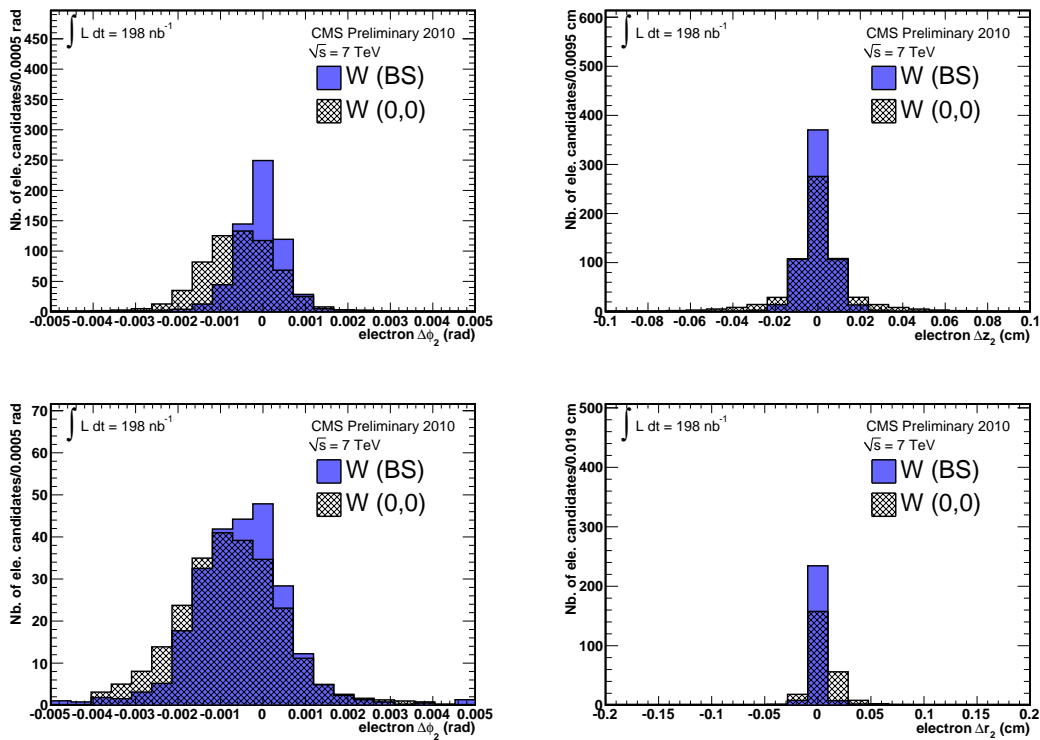Figure 5.10: $\Delta\phi_2$ and Δz_2 (or Δr_2) residuals measured for the second hit of ecal-driven electrons in BPix (top) and forward region (bottom), for the $r_T - \phi$ (left) and $r_T - z$ (right) projection. To measure the ϕ (r_T and z) angles, both the (0,0) reference (gray dashed histogram) and the beam spot position (blue filled histogram) are used.

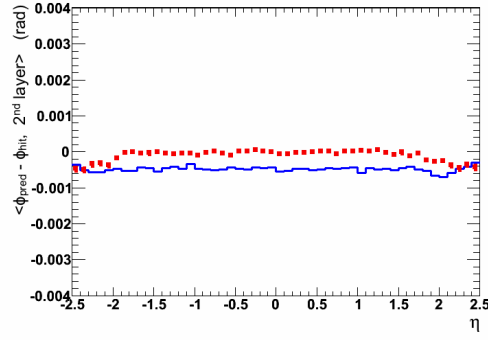---

[3]CMSSW_3_8_0 versus CMSSW_3_8_1

Figure 5.11: For the $\Delta\phi_2$ residual measured for the second hit of ecal-driven electrons, the average value is shown as a function of the electron η. The red trend corresponds to the updated algorithm (BS used as reference) and is compared with the inaccurate old one in blue (0,0 as reference) [99]. Simulated samples of single electrons with p_T =35 GeV/c are used.

5.3.2 Residual $\Delta\phi$ bias in the forward region

The improved performances of the seeding algorithm (Figure 5.11) still show in the forward region an average offset of the $\Delta\phi_2$ window, which is instead completely canceled overall the $|\eta| <2$ region.

To understand such trend, the $\Delta\phi_2$ differences have been looked at for different sub-detectors hosting the second seeding hits (see Figure 5.12) and for both the electron and positron reconstructed candidates (Figure 5.13).

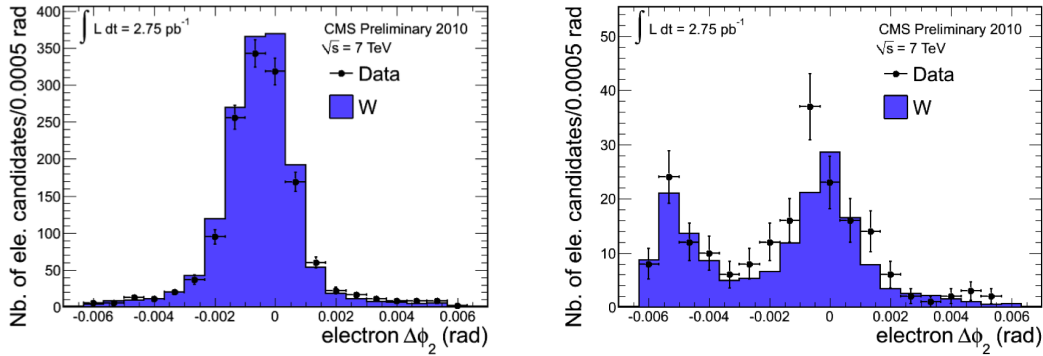Figure 5.12: $\Delta\phi_2$ residuals measured for the second hit of ecal-driven electrons in FPix (left) and TEC layers (right).

The offset clearly enhanced moving in the forward region is seen associated to the positive charge hypothesis (Figure 5.13, on the right) used in the reconstruction algorithm, while the distributions are well centered in zero and symmetric for negative charges (Figure 5.13, on the left).
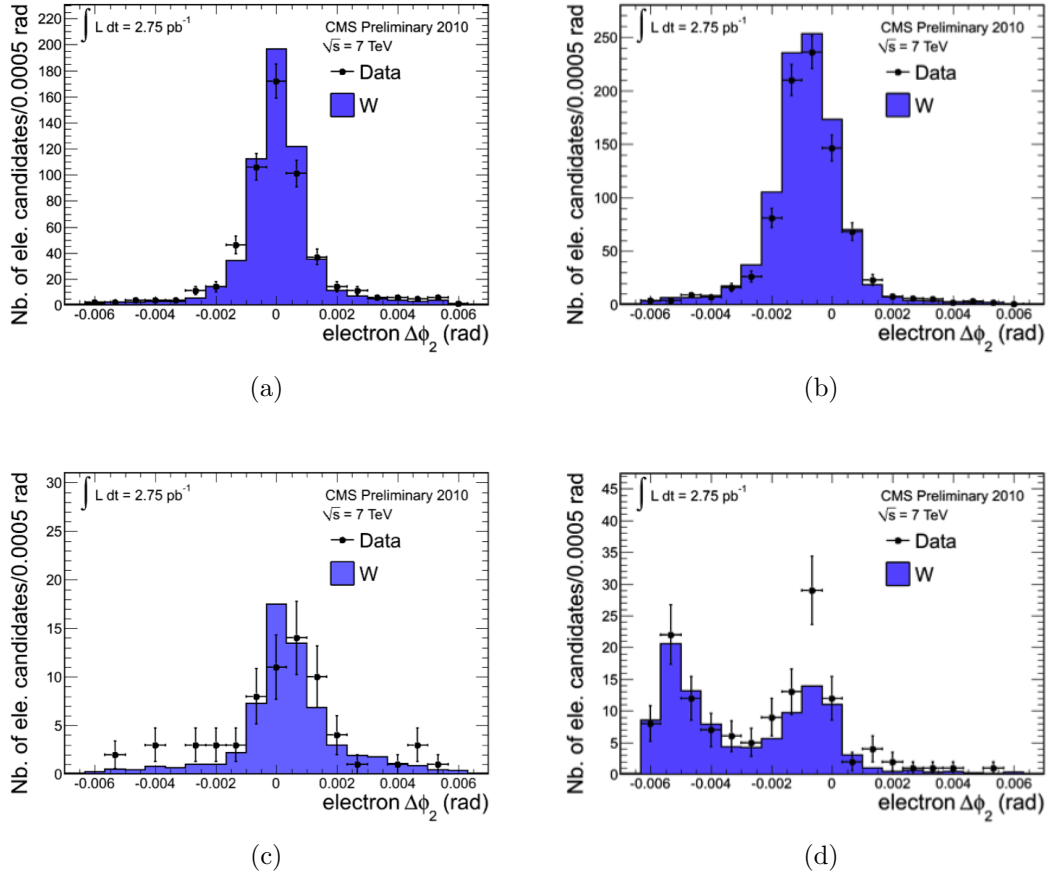
(a)

(b)

(c)

(d)

Figure 5.13: $\Delta\phi_2$ residuals measured for the second hit of ecal-driven electrons in FPix (a,b) and TEC layers (c,d). Distributions are shown for electrons candidates (a,c) and positrons (b,d).

By further investigating the residual window $\Delta\phi_2 = \phi_2^{meas.} - \phi_2^{extr.}$, it emerged that the offset was due to the wrong $\phi_2^{extr.}$ coordinated used in measuring the distance. In particular when exactly the same hit was found, independently from the charge hypothesis followed, the $\phi_2^{extr.}$ considered was always the one corresponding to a negative charge.

From a careful check of the whole electron reconstruction chain, it emerged that this problem originated from the "seed cleaning" (Section 5.1.2) and it is due to the replacement of a seed with the identical one, found under a different charge hypothesis. Indeed with this operation, also the $\Delta\phi_2$ associated to the hits are replaced, but these windows are not charge-independent. In particular, since the negative charge hypothesis is the first tested while looking for seeds, it happened that the $\Delta\phi_2$ measured for the e^+ extrapolation was replaced with the one obtained with the negative charge assumption.

For a proper evaluation of the seeding windows, the residual which should be taken into account is the one found with the same charge hypothesis as the one of the reconstructed

candidate. This condition is adopted in the following.

Through the commissioning activity initial problems were found affecting the electron reconstruction performances. Essentially with order 3 pb^{-1} of integrated luminosity of data, the spotted issues were understood solved and tested.

The final validation was performed taking into account both signal and QCD background processes in the simulation and compared to data, looking at $\sim$14.5 pb^{-1} out of the 36 pb^{-1} recorded by CMS in 2010.

5.3.3 Final seed-finding validation with 14^{-1}pb

The CMS centrally produced primary datasets were used and further selected with the official W/Z skim, requiring:

- at least an electron[4] with p_T >20 GeV/c;

- $MET > 12$ GeV together with one electron found, to select W events;

- $M(e^+e^-)$ >40 GeV/c^2, to select events with a Z boson.

To provide a pure sample of electrons from W events, a further selection was applied, rejecting QCD and pileup events:

- $MET > 25$ GeV;

- no jet[5] with p_T >15 GeV/c and $|\eta| < 2.4$ in the events;

- exactly one reconstructed vertex;

- the electron $p_T > 25$ GeV/c.

Finally, the electron-MET system was required to satisfy:

- $\Delta\phi(SC, MET) > \pi/2$;

- $40 < M_T(SC, MET) <$100 GeV/c^2, where M_T is the W transverse mass[6] .

---

[4]With reference to Section 3.2.3, loose isolation is required (looser than working point (WP) 85% for EB and WP 95% for EE) and loose identification on $\sigma_{i\eta i\eta}$ and H/E only.

[5]A jet cleaning is applied, requiring that no electrons are reconstructed within a $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.30$ from the jet

[6]$M_T = \sqrt{2 \cdot MET \cdot E_T \cdot (1 - cos(\Delta\Phi_{l,MET}))}$, where l is the selected electron and E_T its transverse energy.

The electron candidates were further required to satisfy the isolation and identification criteria (see Section 3.2.3) corresponding to a tight selection, while no cut was applied on the track-supercluster $\Delta\eta$ and $\Delta\phi$ matching variables in order to produce unbiased distributions for the studied seeding residuals.

The same skim selection as used for data was applied to the simulated events of signal and background.

In Table 5.3 the number of events in the data sample passing the different steps of the W event selection are reported.

selections	events passing
Skim	65680
exactly one vertex	15970
$MET > 25$ GeV/c	12971
Jet Veto	11234
$40 < M_T < 100$ GeV/c^2	10116
tight electron Id	9693
tight electron Iso	9314
tight electron conversion rejection	8625

Table 5.3: Number of events surviving the different steps of the W selection for the data sample.

Figure 5.14 presents the kinematic distributions p_T and η of the selected electron candidates. The expectation from simulation is also visible, normalized to the integrated luminosity of the data sample.

In Figure 5.15 the MET of the selected events and the transverse mass M_T of the W candidates are shown. Also in this case, the expectation from simulation is shown, normalized to the integrated luminosity of the data sample.

Very few QCD events are found surviving the selection. A reasonable agreement is found with the simulation although some slight excess is observed for the MC in particular for electron candidates in the ECAL endcaps, as well as a slightly different shape in the data for the p_T and M_T spectra. This latter is likely due to the imperfect calibration, since the used sample included only a preliminary version of the laser transparency corrections.

The contributions to the different combinations of first-second seeding hits in the tracker layers for the selected electron candidates are quoted in Table 5.4.

The number of electrons reconstructed with both seeding hits in the forward region was checked to be strongly suppressed by the applied selection and in particular the conversion veto.

(a)

(b)

(c)

(d)

Figure 5.14: Kinematic distributions of the selected electron candidates in linear (top) and logarithmic (bottom) scale: (a,c) transverse momentum and (b,d) pseudo-rapidity. The MC distribution is also shown, normalized to the integrated luminosity of the data sample. Very few *QCD* events survive the selection.

layers combination	number of electron candidates		
	EB + EE	EB	EE
BPix-BPix	7105	5765	1340
BPix-FPix	1478	9	1469
FPix-FPix	11	0	11
FPix-TEC	31	0	31
TEC-TEC	0	0	0

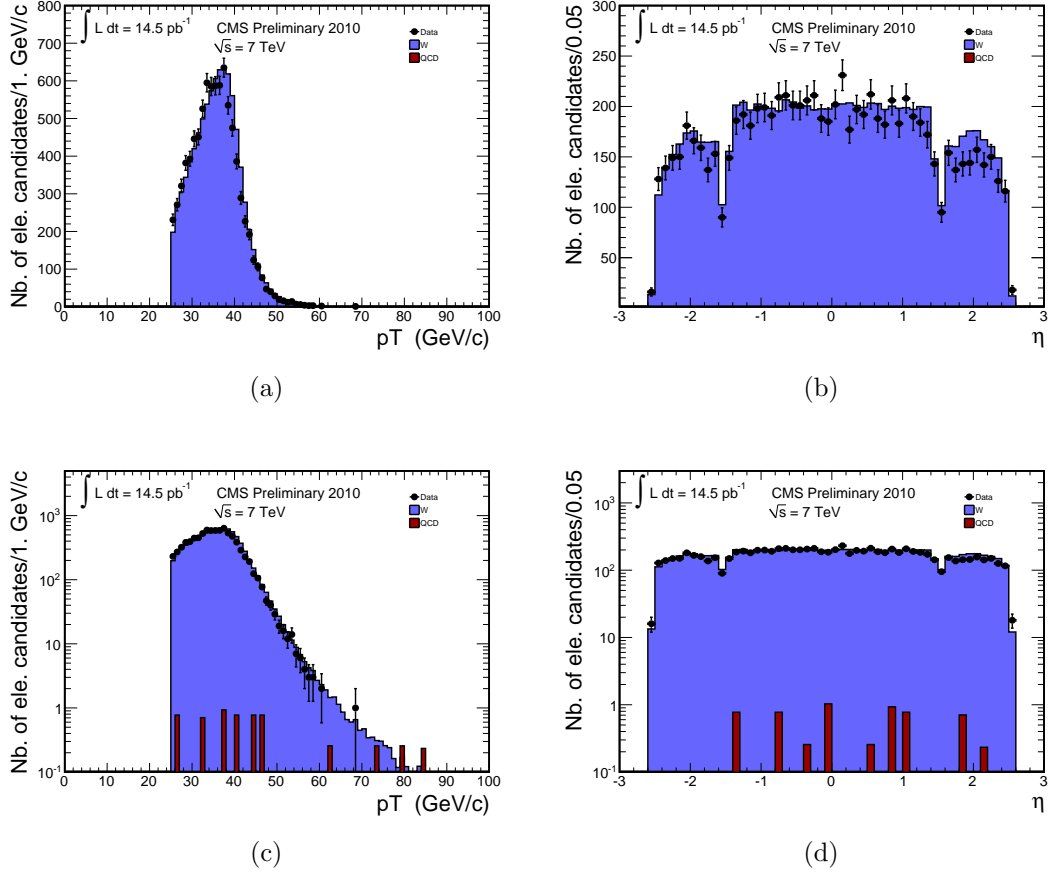Table 5.4: Contributions of first-second seeding layers for the selected electron candidates.
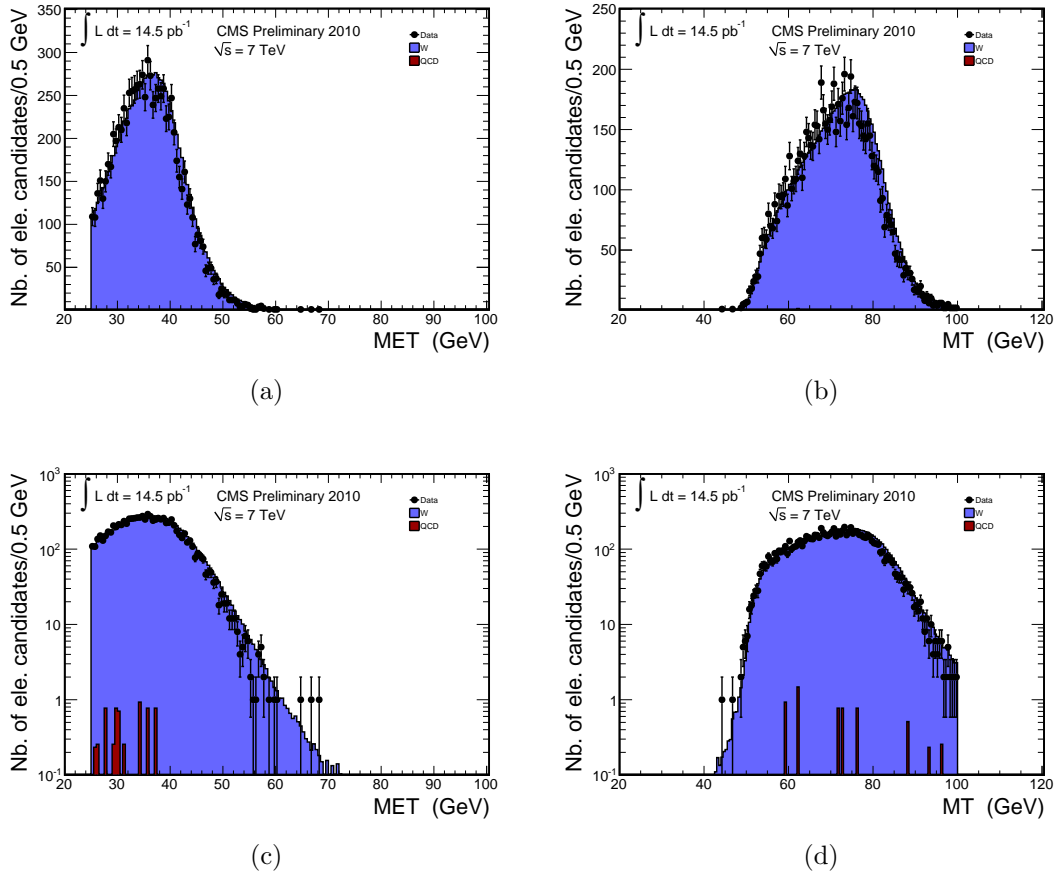
Figure 5.15: Kinematic distributions of the selected events in linear (top) and logarithmic (bottom) scale: (a,c) missing transverse energy, and (b,d) transverse mass of the (lepton, MET) system. The MC distribution is also shown, normalized to the integrated luminosity of data sample. Very few QCD events survive the selection.

The seeding distributions of the ecal-seeded electrons, showing the difference between the measured and predicted hit position in the second seeding layers, are presented in Figures 5.16, 5.17 for combinations involving the pixel layers. The data correspond to an integrated luminosity of $\mathcal{L}int \sim 14.5\ pb^{-1}$.

The distributions are presented for both ϕ and z coordinates and for the BPix-BPix and BPix-FPix combinations of tracker seeding layers for the first two hits. The statistics does not allow to probe the FPix-FPix, FPix-TEC and TEC-TEC combinations.

The distributions are shown in logarithmic scale in Figure 5.16 and in linear scale in Figure 5.17.

Since the residual differences are evaluated at the seeding level for both charge hypotheses (Section 5.3.2), the one corresponding to the reconstructed electron charge is selected, when producing these distributions for each electron (Section 5.3.2).

The MC expectations are also shown, normalized to the number of entries of each distribution in the data. Compared to what is shown in Fig. 5.14, 5.15, the QCD contribution is here smaller (the 10 QCD events are here reduced to 4), as can be expected for background by the requirement of charge agreement between the considered electron and the residual difference computed at the seeding level.
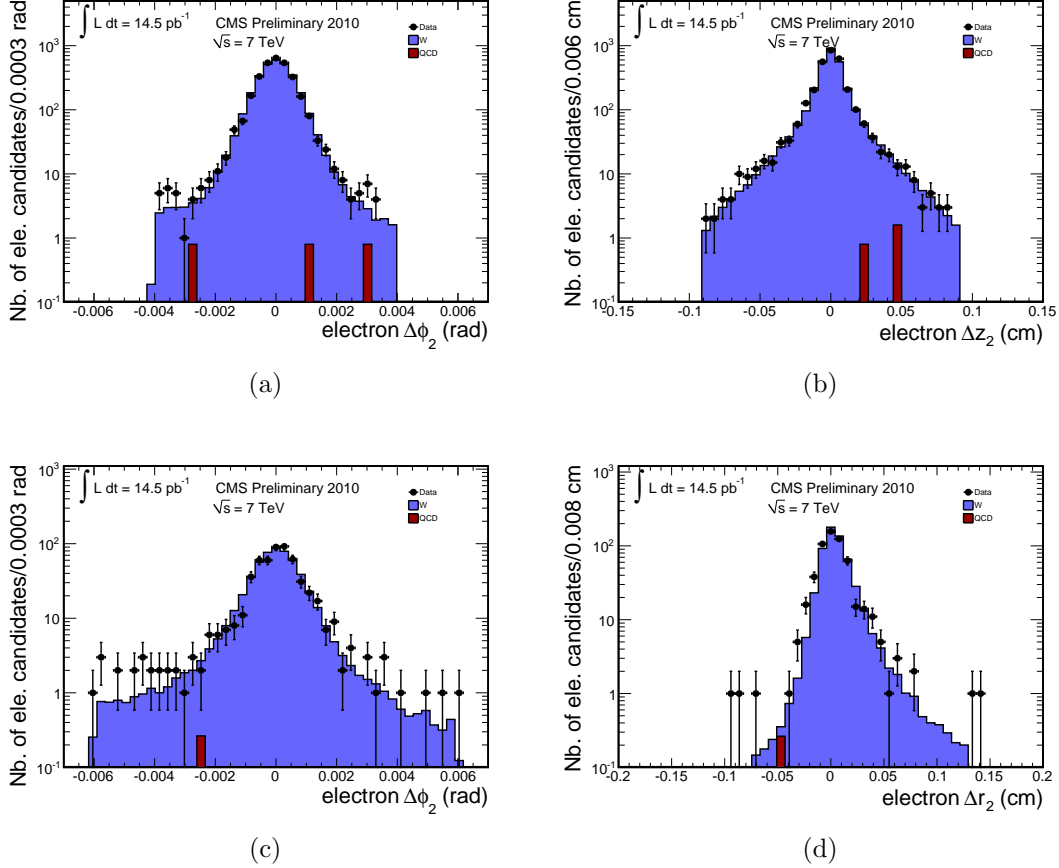

(a)   (b)

(c)   (d)

Figure 5.16: $\Delta\phi_2$, Δz_2 (or Δr_2) residuals measured for the second hit of ecal-driven electron candidates from W events: BPix-BPix (top) and BPix-FPix (bottom) combinations of tracker layers for the two hits are considered. The MC expectation is normalized to the number of entries in the data for each distribution. Distributions are shown in logarithmic scale.

The distributions are seen unbiased and a good agreement is found with the simulation. A slight shift is observed for the Δr_2 BPix-FPix distribution, indicating some possible residual misalignment between the ECAL and tracker detectors, indeed improved alignment constants were introduced in the reconstruction later on. Nevertheless the situation is much improved with respect to the initial data taking conditions as can bee seen comparing this result with the distribution shown in Figure 5.9.
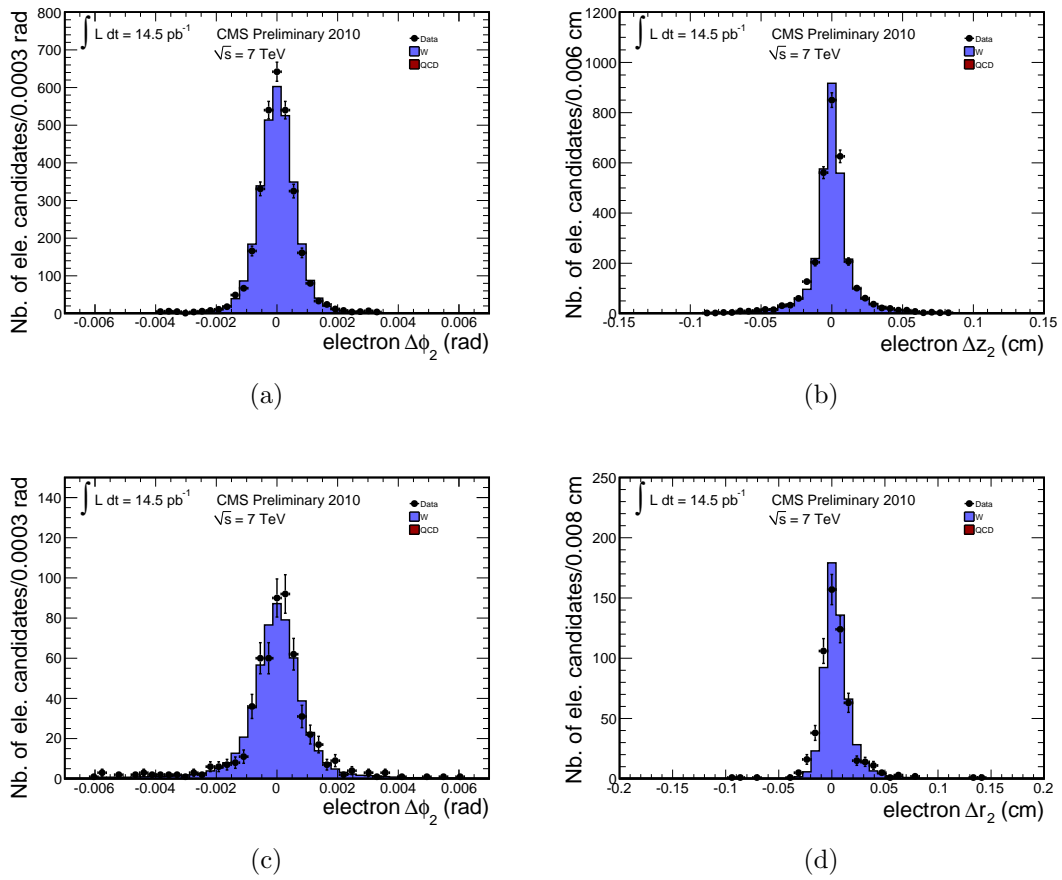
(a)

(b)

(c)

(d)

Figure 5.17: Same as Figure 5.16, in linear scale.

The measurement of the WZ production cross section

This Chapter is devoted to the study of the electro-weak associated production WZ and to the measurement of its production cross section in proton proton collisions at $\sqrt{s}$ =7 TeV.

The inclusive cross section as predicted by the Standard Model for $pp \rightarrow WZ + X$ at $\sqrt{s}$ =7 TeV is $18.57^{+0.75}_{-0.58}$ pb (NLO) [11]. The choice of looking for the decay channel $WZ \rightarrow l'\nu l\bar{l}$, with $l, l' = e, \mu$ reduces the effective production rate per single channel to $\sigma(WZ \rightarrow l'\nu l\bar{l})$=0.067 pb, corresponding to a branching ratio of 1.43%, but brings the advantage to deal with the cleanest signature, a crucial point in the hadronic environment of the LHC. The analysis presented here allowed to observe the first WZ event in CMS with less than 36/pb of data, in the final state $ee\mu + MET$ (details on the event presented in Appendix B.1).

Already with about 200/pb of data, corresponding to few months of data taking in 2011, a preliminary measurement of the WZ cross section was possible. The analysis and results here presented profit from the 1.09 fb^{-1} integrated luminosity collected by CMS up to summer 2011.

In the first Section, the relevant elements exploited in the analysis are presented, while the WZ signal extraction strategy and the background treatment are described in Section 6.2 and 6.3, respectively.

The procedure followed to measure the WZ cross section is the subject of the last Section, with a result of $\sigma(WZ) = 19.11^{+3.30}_{-2.53}$(stat.)$\pm 1.10$(syst.)$\pm 1.15$(lumi.)pb, in agreement with the NLO expectation from theory.

6.1 The WZ event topology

The leptonic WZ decay signature consists of a couple of same-flavor, opposite-charge, isolated leptons, whose invariant mass is compatible with the one of the Z boson, together with a third high p_T, isolated lepton and a significant missing transverse energy (MET) associated to the escaping neutrino. All the production processes with at least three leptons in the final state are background to such a signature, the most dangerous being those closer to the specific WZ kinematic requirements. A classification can be made in terms of instrumental and physical backgrounds.

The only physical source is represented by ZZ events when one of the leptons is either non reconstructed or falling outside the detector acceptance.

The instrumental sources are

- $QCD, W + jets, Z + jets, WW + jets$ events (with 1 to 3 jets faking leptons);

- $Zbb, t\bar{t}$ events (with leptons from heavy flavor jets);

- $Z\gamma$ (with electron-conversion for the e-channels);

6.1.1 Data and Monte Carlo samples

The data samples used were recorded by the CMS detector up to July 2011, selecting the certified run ranges for good detector operation and corresponding to an overall luminosity of 1.09fb^{-1}. The primary datasets chosen are a combination of several High Level Trigger (HLT) paths, reconstructed with the `CMSSW_4_2_X` software version.

The simulated signal and backgrounds were produced with different Monte Carlo event generators and then digitized and reconstructed through a detailed detector simulation based on the GEANT4 package [78], using the same `CMSSW_4_2_X` software version.

The WZ signal sample was produced with `PYTHIA` [100], by enabling the W and Z bosons to decay into leptons only (electrons, muons and taus). Since only $W \to l\nu$ and $Z \to l^+l^-$ with $l = e, \mu$ are considered in this analysis, the contribution of leptonic decays of τ's from W or Z is taken as a background and subtracted.

The background processes were produced using `PYTHIA`, `MADGRAPH` [101] and `powheg` [102] generators.

The full list of simulated signal and background samples is presented in Table 6.1.

A robust strategy to extract the signal of interest is to proceed gradually in defining the phase space for the final state, until it dominates over background sources. At the same time, this approach allows to preserve the control regions in the earlier stages of the analysis, to perform data driven estimate and control of the backgrounds. This goal

Process	MC generator	$\sigma_{\text{NLO}} \times BR$ (pb)	Events
$WZ \to 3l\nu$	`pythia`	0.594	1097759
$ZZ \to 4l$	`pythia`	0.095	1096075
$WW \to 2l2\nu$	`pythia`	4.88	210667
$Z + jets \to ll$	`madgraph`	3048	36277961
$t\bar{t} \to 2l2\nu 2b$	`powheg`	17.32†	10339374
$Zbb \to ll$	`madgraph`	23.5	7614970
$Z\gamma \to ee\gamma$	`madgraph`	47.60	323881
$Z\gamma \to \mu\mu\gamma$	`madgraph`	46.79	321534
$Z\gamma \to \tau\tau\gamma$	`madgraph`	42.85	326396
$QCD multi\text{-}jet$ $(b/c \to e + X,\ e\text{-enriched},\ \mu)$	`pythia`	-	-

Table 6.1: Signal and background processes simulated and considered for the analysis. $\dagger$ refers to NNLO computation.

can be met with a cut based procedure. It is here organized in two main steps (see Section 6.2), with a preselection to strongly reduce the instrumental backgrounds at first and then a main selection to extract the signal and to optimize the expected significance.

6.1.2 The object selection

The criteria used to select the physics objects of interest for the $WZ \to 3l\nu$ decay chain, essentially electrons, muons and MET, are of fundamental importance for a performing cut based analysis.

The lepton isolation

Electrons and muons as reconstructed in CMS are further selected on the basis of few discriminating variables (defined in Section 3.2.3), in terms of identification, isolation and conversion rejection (the latter only for the electrons).
Threshold cuts on the above criteria have been tuned together, so to define selection working points (WP) of different efficiency on the signal and rejection on the backgrounds[103]. Such kind of selections are referred to as "simple cuts".

Within the WZ analysis, two different working points are used. A loose selection (WP95) required for leptons from the Z decays and a tight one (WP80) for those from W decays. In the identification of a Z boson, the mass constraint applied with high precision on the same flavor opposite charge leptons is already a powerful requirement to reduce the backgrounds, while for W decays, a tighter requirement on the final state lepton is the only tool to reduce the very high background from multi-jet events.
The variables used for both the WP selections are listed in Tables 6.2 and 6.3, for electrons and muons respectively.

	loose selection (WP95)		tight selection (WP80)	
Electron ID	EB	EE	EB	EE
$\sigma_{i\eta i\eta}$	0.012	0.031	0.001	0.031
$|\Delta\eta_{\text{in}}|$	0.007	0.011	0.005	0.006
$|\Delta\phi_{\text{in}}|$	0.8	0.7	0.027	0.021
Isolation	EB	EE	EB	EE
Combined relative isolation	0.15	0.1	0.07	0.06
Conversion rejection	EB	EE	EB	EE
Expected missing inner hits	0	0	0	0
Dist	NA	NA	0.02	0.02
Dcot	NA	NA	0.02	0.02

Table 6.2: Selections used in the "simple cuts" for electron candidates in EB and EE. Loose selections are used for $Z \to ee$ leptons, the tight ones for e from W decays. Variables are defined in Section 3.2.3.

	loose selection (WP95)	tight selection (WP80)
Muon ID		
Number of pixel hits	>0	>0
Number of tracker hits	>10	>10
Normalized χ^2	<10	<10
Number of muon hits	>0	>0
Number of chambers with matched segments	>1	>1
$dxy(PV)$	< 0.2cm	< 0.2cm
Isolation		
Combined Relative Isolation	<0.15	<0.1

Table 6.3: Selections used in the "simple cuts" for muon candidates. Loose selections are used for $Z \to \mu\mu$ leptons, the tight ones for μ from W decays. Variables are defined in Section 3.2.3.

The isolation distributions for the leptons after the preselection[1], are shown in Figure 6.1. Such variables are the ones used to identify the WZ signal events in the analysis here presented and do not correspond to the standard ones described in Section 3.2.3. The dedicated algorithm used to compute the variables shown in Figure 6.1 is discussed in Section 6.1.3.

Eventual differences between data and simulation can bring to important bias in the events selection. A major issue characterizing the whole 2011 LHC operation is the pile-up increase.

To make sure to deal with simulated events well reproducing data, a pile-up re-weighting procedure is applied to all the variables used in the analysis.

As an example, the number of the reconstructed vertices per event in MC is shown in Figure 6.2, once rescaled to the pile-up in data.

---

[1]Essentially triggered events containing at least three leptons (see Section 6.2.1).
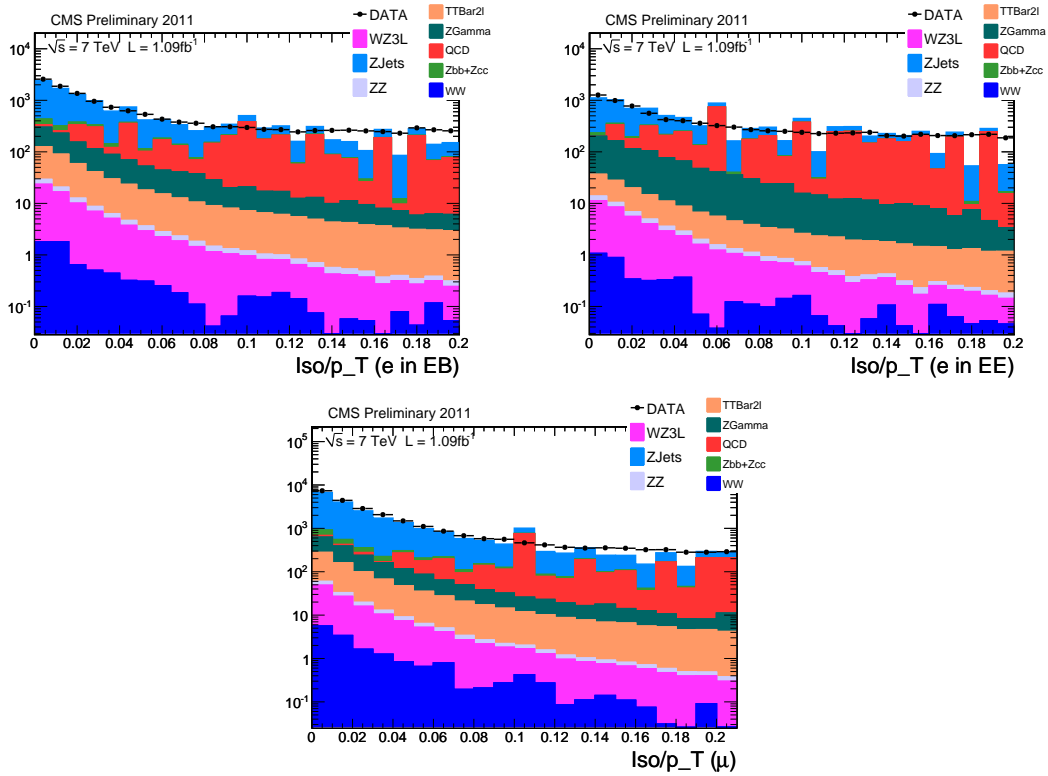
Figure 6.1: Combined Relative Isolation for e in EB (top-left), e in EE (top-right), μ (bottom). The simulated distributions are normalized to the luminosity in data (1.09/fb).

Residual pile-up effects on the variables important for the $WZ \to 3l\nu$ analysis are further investigated in Section 6.1.3.
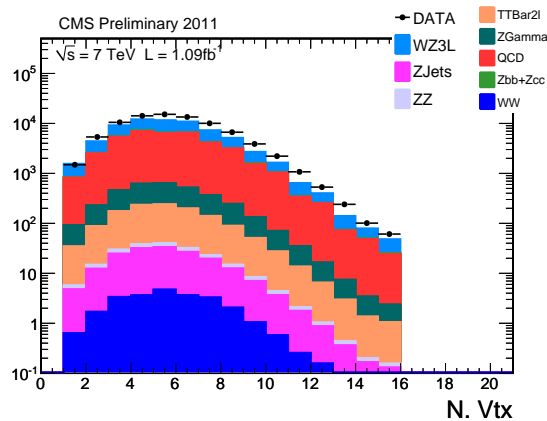

Figure 6.2: Number of the reconstructed vertices per event in MC, rescaled to pile-up in data. The data distribution over-imposed corresponds to the data taking up to summer 2011. Simulated distributions are normalized to the luminosity in data (1.09/fb).

6.1.3   Pile-up effects

With reference to Figure 6.2, the increase in pile-up brought in summer 2011 to an average of six vertices reconstructed per event in data, with contributions up to 16 vertices.

Multiple overlapping interactions lead to an enhanced detector occupancy, almost saturated by the particles produced from different vertices. As a consequence, the number of jets is much increased as well as the density of the reconstructed tracks and the mean energy deposited in the detector. Calorimetric measurements result particularly sensitive to such conditions, with biases on the objects isolation and identification.

The pile-up effect on the electrons, muons and MET performances are discussed in the following.

Pile-up effects on MET

For the simulated signal and background processes considered in the WZ analysis, the fraction of events above a MET threshold is shown in Figure 6.3 with respect to the number of reconstructed vertices per event.
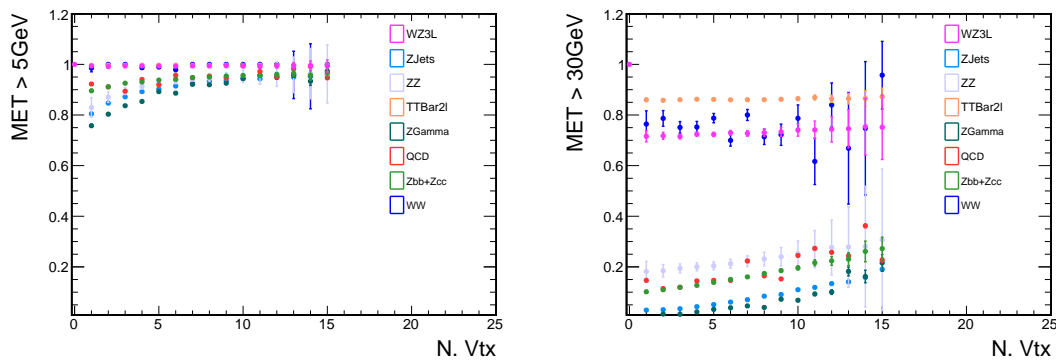

Figure 6.3: Fraction of events surviving the MET >5 GeV (left) and MET >30 GeV (right) cut, with respect to the number of reconstructed vertices per event.

The raising trend in both plots is expected with the enhanced calorimetric activity. This phenomenon only affects the spurious sources of MET, leaving the efficiency flat for genuine contributions as for WZ, $t\bar{t}$ and WW here looked at.

Such a behavior could be crucial at higher pile-up, where the increased fraction of background events surviving the MET cut becomes comparable to the one for signal, which instead keeps constant. Over the pile-up range considered, the above consideration is particularly important for low thresholds of the MET cut (MET >5 GeV in Fig. 6.3,

on the left), while tighter cuts ($MET >30$ GeV in Fig. 6.3, on the right) still allow for a good discrimination between spurious and genuine MET signals.

In the WZ analysis, the MET cut is well above the mean contribution from pile-up fluctuations, since it is optimized to look for the neutrino from W decay. Residual effects are nevertheless taken into account as systematic uncertainties.

Pile-up effects on leptons

As for MET, the fraction of leptons surviving the isolation cut applied in the analysis is shown versus the number of vertices per event in Figures 6.4 and 6.5 for electrons and muons respectively.
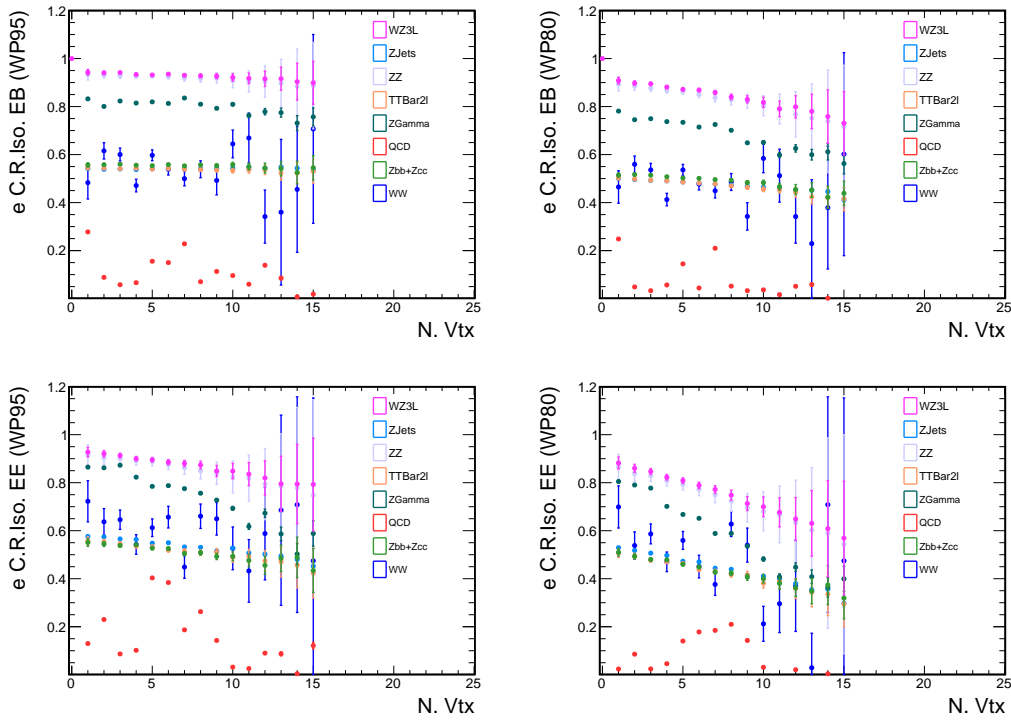
Figure 6.4: Fraction of electrons surviving the isolation cuts: loose in EB (top-left), tight in EB (top-right), loose in EE (bottom-left), loose in EB (bottom-right), with respect to the number of vertices.

Similar plots are shown for the cut on the H/E variable[2] in Figure 6.6, for electrons in the endcap region. The reference cuts for these plots correspond to the loose selection (plots on the left) and tight selection (plots on the right) used in the analysis (Tables 6.2 and 6.3).

---

[2]Among the electron Id variables, H/E is the only one affected by pile-up and only for electrons in the endcap region.

Figure 6.5: Fraction of muons surviving the isolation cuts: loose (left), tight (right), with respect to the number of vertices.
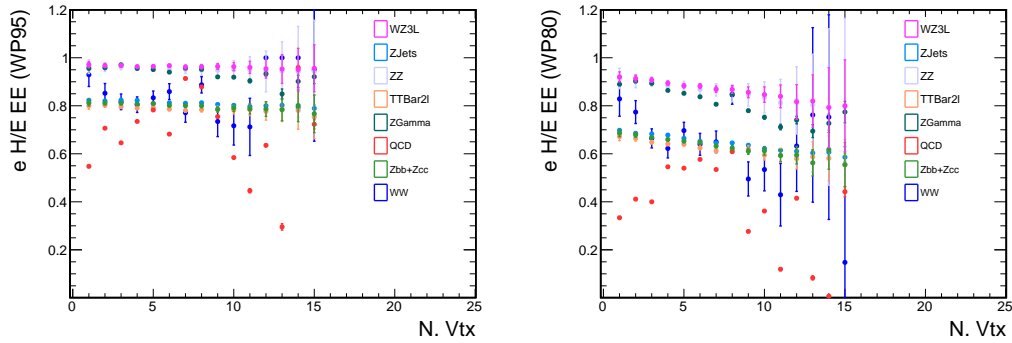
Figure 6.6: Fraction of electrons surviving the Id. cut on H/E: loose (left), tight (right), with respect to the number of vertices.

The decreasing trend at high pile-up is emphasized when tight cuts are applied (plots on the right), since loose requirements allow instead for bigger energy depositions in the calorimeters. Over the whole pile-up range looked at, the signal efficiency for the variables considered is reduced by about 10% for muons and up to $\sim$30% for electrons in the endcaps, for tight cuts.

The affected variables need to be corrected for the pile-up dependency, to provide a solid unbiased analysis with respect to the selections applied. The method chosen for this purpose is described in Section 6.1.3. The effect of the pile-up correction on the leptons isolation is shown in Figure 6.7 for electrons and in Figure 6.8 for muons, proving flat trends with respect to the number of vertices per event.
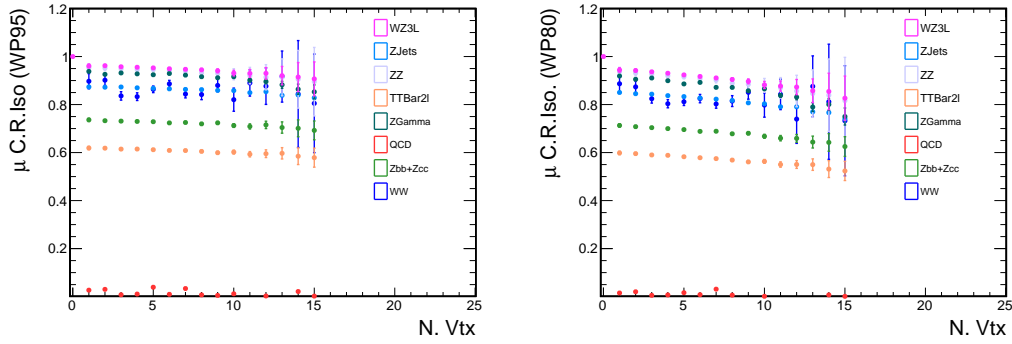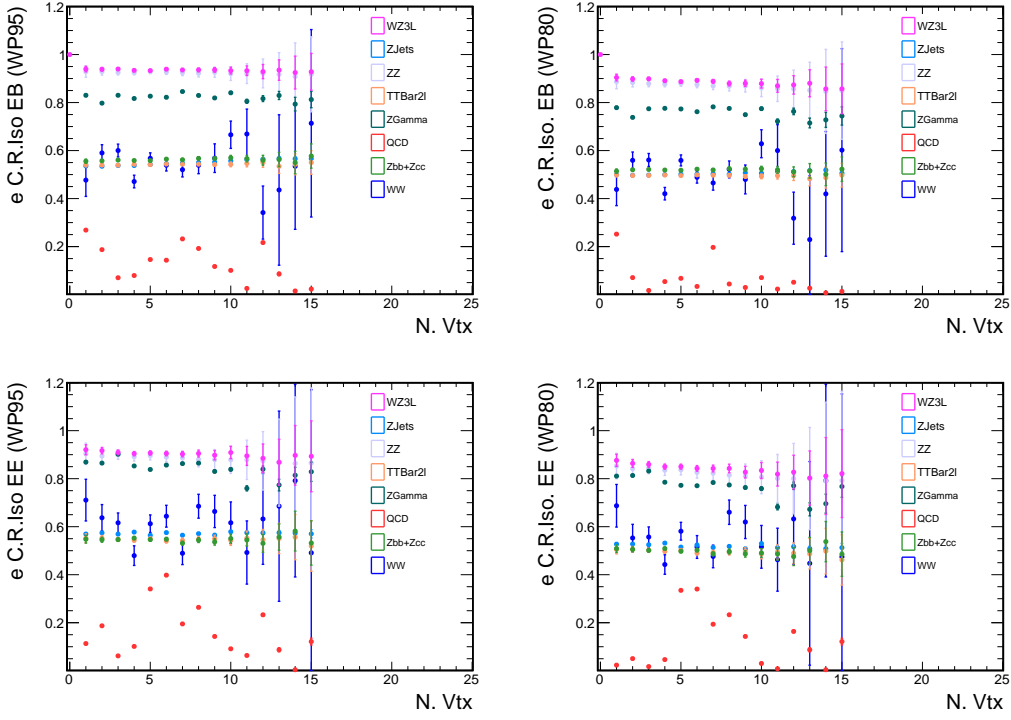
Figure 6.7: Fraction of electrons surviving the Iso. cut (after pile-up correction): loose in EB (top-left), tight in EB (top-right), loose in EE (bottom-left), tight in EE (bottom-right), with respect to the number of vertices.
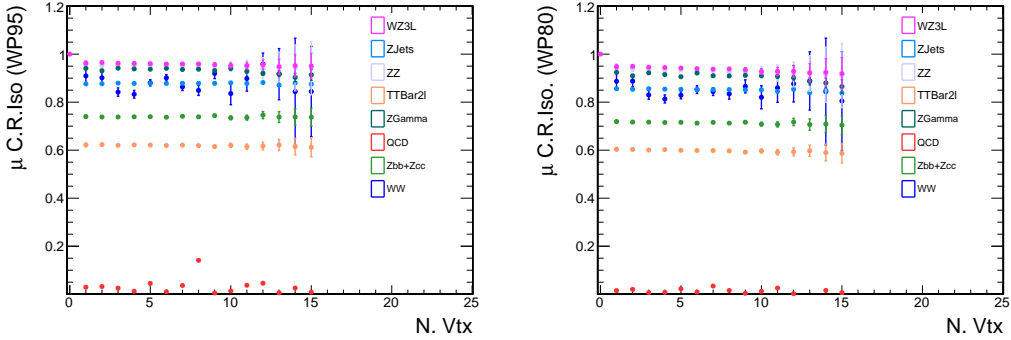
Figure 6.8: Fraction of muons surviving the Iso. cut (after pile-up correction): loose (left), tight (right), with respect to the number of vertices.

Pile-up corrections

The procedure adopted takes advantage of the FastJet correction algorithm [104][105], which provides a method to calculate the energy density (ρ) per unit area due to the contribution of pile-up and underlying events.

Since lepton isolations are computed within fixed cones, corresponding to areas as well, the isolation variable computed in a cone of radius R around the lepton can be thus

corrected as

$$Iso_{corrected} = Iso - \rho \cdot \pi R^2 \qquad (6.1)$$

However this solution cannot be used for the H/E variable, since H/E accounts for the calorimetric energy deposits, in a cone of radius 0.15 radians around the lepton direction, which is too small to integrate the ρ contribution without suffering from fluctuations. Since a tight cut on H/E is of great help in the background rejection, but at the same time it's important to provide a stable and efficient selection, the required pile-up correction should be suitable for both the loose and tight selections used in the WZ analysis.

Because the electron isolation in HCAL is computed with an inner veto cone ($R =0.15$), to provide complementary information with respect to H/E, the optimal solution is to define a new HCAL isolation for electrons, computed in the full cone $0. < \Delta R < 0.3$ radians. The resulting variable has the H/E discriminating power and is solid and stable against pile-up subtraction at the same time. H/E is then removed from the identification requirements and the isolation variables for both electrons and muons can be thus corrected as in 6.1.

Moreover, rather than assuming a fixed geometrical area of 0.3 rad. corresponding to the value used to compute the leptons isolation, an "effective area" A can be considered, not to suffer from any threshold and footprint removal applied differently between the isolation algorithm and the FastJet one.
A is computed empirically as the value making the isolation variable to be flat over the pile-up range considered.

By representing the isolation trend as a discretized variable function of the number of vertices (Vtx_i), Iso_{Vtx_i} can be made flat over the increasing number of vertices per event Vtx_i $(i, i+1, ...)$, following the prescription in equation 6.1.

$$
\begin{aligned}
Iso^{corrected}_{Vtx_i} &= Iso^{corrected}_{Vtx_{i+1}} \qquad \text{corresponds to} \\
Iso_{Vtx_i} - \rho_{Vtx_i} \cdot \pi R^2 &= Iso_{Vtx_{i+1}} - \rho_{Vtx_{i+1}} \cdot \pi R^2 \\
\text{from which it can be obtained } \pi R^2 &= \frac{Iso_{Vtx_{i+1}} - Iso_{Vtx_i}}{\rho_{Vtx_{i+1}} - \rho_{Vtx_i}} = A \qquad (6.2)
\end{aligned}
$$

To better identify eventual systematic effects in specific detector regions, the isolation computed for electrons and muons has been looked at separately for objects in the barrel

and in the forward region, and by comparing the different sub-detectors isolations separately in Figure 6.9. The effect is observed to be stronger in the calorimeters (ECAL, HCAL) and quite feeble in the tracking system, since the tracks considered while computing track-isolation are required to belong to the same vertex.

(a)

(b)

(c)
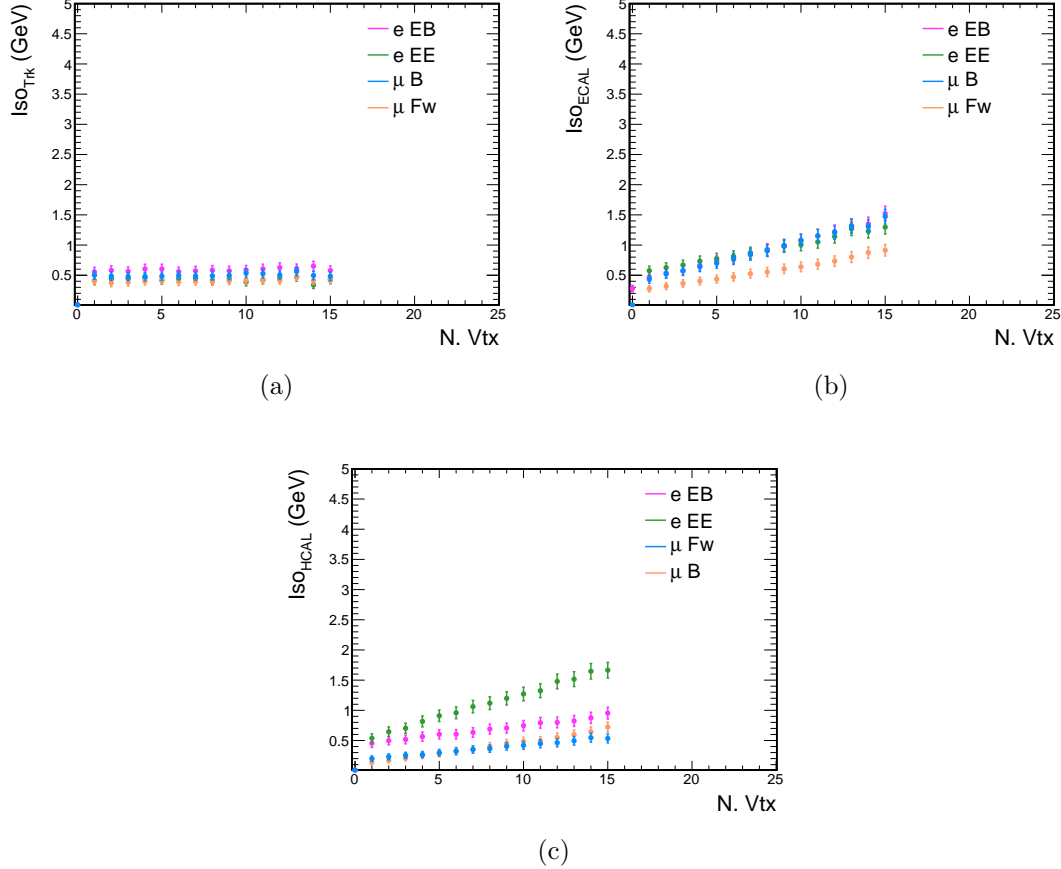
Figure 6.9: Mean value of the isolation variable as a function of the number of reconstructed primary vertices, for the tracker isolation (a), the ECAL isolation (b) and the HCAL isolation is shown: magenta (electrons in EB), green (electrons in EE), blue (muon in barrel region), orange (muon in forward region).

For simplicity a correction common to the sub-detectors is optimized. The combined isolation variable is plotted in Figure 6.10 for electrons and muons in the central and forward detector regions. A linear fit is over-imposed on each variable, to evidence the rising trends.

In Table 6.4, the computed effective areas for electrons and muons in the barrel and forward detector regions are reported. Such values are used in the pile-up subtraction procedure, applied event by event and the resulting combined isolation variables are shown in Figure 6.11.

Figure 6.10: Mean value of the combined isolation variable as a function of the number of reconstructed primary vertices is shown: magenta (electrons in EB), green (electrons in EE), blue (muon in barrel region), orange (muon in forward region).

Electrons	Barrel region	Forward region
$\pi * R^2_{eff}$	0.129	0.164
Muons	Barrel region	Forward region
$\pi * R^2_{eff}$	0.12	0.105

Table 6.4: Effective areas computed for the combined electron (top) and muon (bottom) isolations.

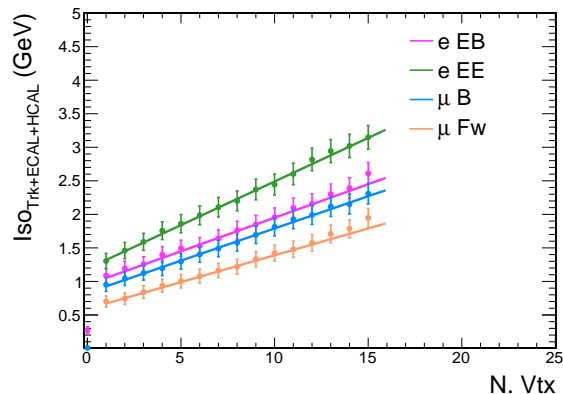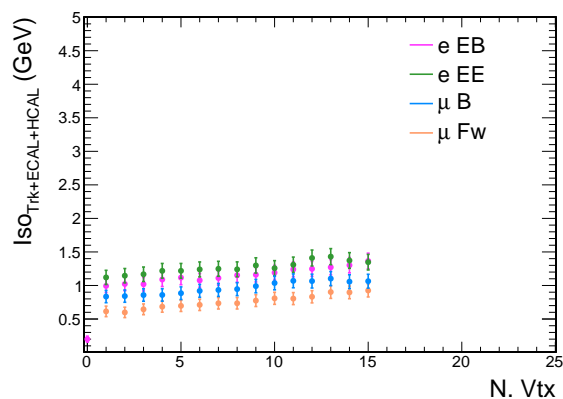Figure 6.11: Mean value of the combined isolation variable as a function of the number of reconstructed primary vertices, after correcting for pile-up dependency: magenta for electrons in EB, green for electrons in EE), blue for muon in barrel region and orange for muon in forward region.

6.2 The WZ event selection

As previously mentioned, the strategy followed to extract the WZ signal relies on cut based criteria to select those events satisfying the topologic and kinematic requirements of the final state signature.

6.2.1 The preselection

A first filtering of the events[3] is performed with a preselection, which commonly gathers the event selections related to the signal topology.

At first, only the events triggered by the High Level Trigger paths of interest are considered. In particular, only the double HLT listed here below are used, since they are efficient enough not to need further single[4] or crossed[5] triggers to cover inefficiencies.

- For the $Z \to ee$ channel:

    - `HLT_Ele17_CaloIdL_CaloIsoVL_Ele8_CaloIdL_CaloIsoVL_v*`

    This is a double electron trigger which requires two HLT electrons, one with $E_T > 17$ GeV, a second with $E_T > 8$ GeV. In addition loose identification and isolation requirements[6] are required on both legs. At Level-1 this HLT path is seeded by the trigger `L1_SingleEG12` i.e. a trigger which requires one electromagnetic object with $E_T > 12$ GeV.

- For the $Z \to \mu\mu$ channel, consistently with the changes in the HLT menu during the data taking in 2011, two different paths are applied in the inner and following periods respectively:

    - `HLT_Double_Mu7_v*`

    This is a double muon trigger which is based on L3 (combined muon system and tracker information) muon reconstruction and requires two L3 muons at HLT with $p_T > 7$ GeV. At Level-1 this HLT path is seeded by the trigger `L1_DoubleMu3` i.e. a double muon trigger which requires $p_T > 3$ GeV. This HLT path is only used until "prescaled" (see Section 3.1).

    - `HLT_Mu13_Mu8_v*`

    This is a double muon trigger which is based on L3 (combined muon system and tracker information) muon reconstruction and requires two L3 muons at HLT with $p_T > 8$ and 13 GeV. At Level-1 this HLT path is seeded by the trigger `L1_DoubleMu3` i.e. a double muon trigger which requires $p_T > 3$ GeV.

---

[3]A very beginning skimming of the Datasets is previously performed, to reduce the event size to deal with. It consists in rejecting the events with less than three leptons.

[4]HLT based on the presence of a single lepton identified to trigger the event.

[5]HLT based on the presence of an electron-muon pair identified to trigger the event.

[6]Looser or comparable with WP95 selections applied in the analysis.

The motivation for using such double lepton HLT paths is to deal with events triggered by the presence of a Z boson. This allows to have unbiased control samples to measure the backgrounds in the WZ analysis, where the 3rd lepton selection is used as discriminating cut, after the Z candidate is identified. For this reason, it's important to make sure that the Z candidate is reconstructed out of the leptons triggering the event, condition obtained by imposing that the leptons from the Z decay in the offline analysis are matched to the HLT objects, for both data and simulated samples.

In case of muons, the trigger objects can be directly matched to the offline reconstructed leptons in terms of geometrical cones, provided the correct p_T thresholds are applied, while the direct comparison between online and offline objects is not in general possible for electrons. Following the procedure described in [106], a set of minimal cuts is applied on offline reconstructed electrons for both data and MC, so to insure to be on the efficiency plateau of the double electron trigger.

The preselection looks for events:

- which have fired the `DoubleElectron` or `DoubleMuon` trigger (see above);

- which contain at least 3 leptons having $p_T >$ 10 GeV/c;

- for which the two same-flavor, opposite-charge leptons matching[7] the HLT objects have an invariant mass above 40 GeV/c^2;

- surviving a loose vertex consistency requirement, in terms of $|d_{xy}{}^8| < 0.5$ and $|d_z{}^9| < 0.5$ for the selected leptons.

For each preselection cut, the event counts are reported in Table 6.5, for both data and simulated samples. At each step, the number of events is shown for each considered channel and summing them together (combined).

The preselection strongly reduces the instrumental backgrounds, in particular the 30% of the QCD events survive, while keeping the efficiency on the signal at the level of 87%. At this stage, some excess is found in data with respect to MC expectation, which is likely related to a lack of statistic in the simulation of QCD events. For instance dedicated enriched QCD samples are used for the electron population, but not for muons. Moreover such a discrepancy is canceled proceeding further with the steps in the analysis, where there are no QCD events surviving the selections.

---

[7]For electrons it is in the sense of the minimal set of cuts.
[8]Track distance of closest approach to the primary vertex in the transverse plane.
[9]Track distance of closest approach to the primary vertex in the longitudinal plane.

channel	DATA	WW	WZ	ZZ	$t\bar{t}$	$Z\gamma$	QCD	Zbb	$Z + jets$	S+B(MC)
Preselection: Skim										
combined	235794	70.08	183.48	42.50	1952.32	4863.61	113501	2944.54	87330.6	210888.23
$ee+l$	110895	31.69	86.52	16.49	793.63	2299.44	51887.9	1328.1	38229.5	94673.27
$\mu\mu+l$	124911	38.39	96.96	26.	1158.69	2564.17	61613.2	1616.44	49101.1	116214.96
Preselection: HLT requirement + 3 leptons $p_T > 20,10,10$ GeV/c										
combined	91652	26.64	162.84	37.74	1185.01	2315.85	40277.4	1173.8	30625.6	80051.98
eee	15377	6.39	39.51	9.80	217.1	976.47	4195.44	272.92	8037.74	13755.37
$ee\mu$	10562	3.79	37.40	4.72	224.79	79.21	5174.55	205.03	2915.34	8644.82
$\mu\mu e$	32869	13.17	45.55	14.45	524.04	1158.81	12822.9	479.88	15806.8	30865.60
$\mu\mu\mu$	38532	3.84	44.21	10.26	357.23	127.80	21514.9	274.24	4453.72	26786.19
Preselection: HLT requirement + 3 leptons $p_T > 20,10,10$ GeV/c + Vtx compatibility										
combined	84978	25.12	159.36	36.89	1137.56	2207.03	36347.7	1131.61	29311.8	74216.71
eee	14086	5.96	38.55	9.57	209.27	925.49	3613.75	263.16	7663.85	12729.6
$ee\mu$	9423	3.54	36.63	4.60	214.95	74.08	4324.97	196.64	2724.96	7580.36
$\mu\mu e$	31209	12.53	44.57	14.13	503.82	1111.81	11914.	464.42	15278.7	29343.99
$\mu\mu\mu$	35415	3.56	43.26	10.02	340.40	119.55	19589.3	262.90	4193.76	24562.76

Table 6.5: Observed and expected signal and background yields at the different step of the preselection. The numbers correspond to an integrated luminosity of 1.09/fb.

The l^+l^- invariant mass spectrum at this stage of the analysis is shown in Figure 6.12.
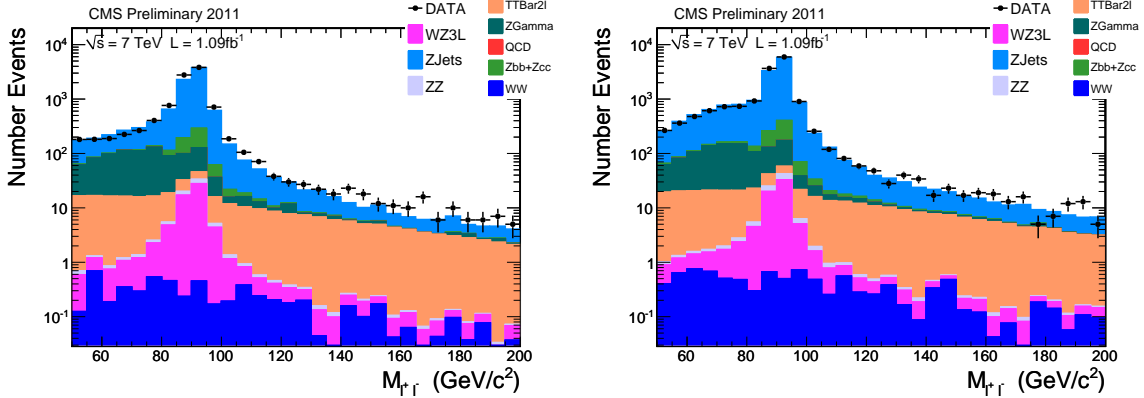
Figure 6.12: Invariant mass distribution for e^+e^- (left) and $\mu^+\mu^-$ muons (right) at the preselection level. The simulated distributions are normalized to the luminosity in data (1.09/fb).

6.2.2 The Z candidate selection

The first main selection consists in looking for the Z candidate, which is built out of a sample of leptons, provided they

- satisfy loose criteria (WP95) on the identification and isolation (see Section. 6.1.2);

- are matched to a trigger object from the selected HLT paths;

- have same-flavor and opposite charge;

- have[10] $p_T > 20$, 10 GeV/c (for the $Z \to ee$) and $p_T > 15$, 15 GeV/c (for the $Z \to \mu\mu$);

- have an invariant mass in the [60, 120] GeV/c^2 range;

In case several matching combinations are found, the best candidate is chosen as the closest to the nominal Z boson mass [2]. Particular attention is paid to build the Z candidate out of the HLT matched leptons, to avoid biases on the eventual third lepton in terms of identification or isolation, which are loosely applied in the on-line selection and are used in the offline analysis to discriminate the signal from the remaining backgrounds, $Z + jets$ in primis.

For the events surviving the Z candidate selection, the $Z+jets$ background completely dominates, with still important contamination from $Z\gamma$, Zbb and $t\bar{t}$, while QCD events are completely removed. The different contributions are shown in Figure 6.13 as a function of the MET in the event.

Figure 6.13: Missing transverse energy for the ee (left) and $\mu\mu$ channel (right), after the Z candidate is selected. The simulated distributions are normalized to the luminosity in data (1.09/fb).

6.2.3 The W candidate selection

A lepton from a W boson leptonic decay is then looked for out of the remaining leptons requiring $p_T > 20$ GeV/c and tight (WP80) criteria on both its identification and isolation. In case several candidates are found, the one with the highest p_T is selected.

---

[10]The p_T thresholds are constrained by the matching requirement with the HLT objects.

Events are further rejected if a second Z boson candidate is found, to reduce the contamination from ZZ background. For this purpose a second Z is looked for, by applying the same requirements as used to select the first candidate, except for the matching of the leptons with the HLT objects. The same-flavor constraint applied on the two Z leptons is also relaxed. This allows a better discrimination against eventual electron-muon combinations, from a $Z \to \tau\tau$ decay and contributes with a further 2% reduction on the ZZ event yield, after the 2nd Z veto.

The missing transverse energy and the di-lepton invariant mass at this stage of the selection are shown in Figures 6.14 and 6.15 respectively.

By requiring both the Z and W boson candidates, the WZ signal starts to emerge at high MET and the well defined background contamination, mainly from $Z + jets$, ZZ and $Z\gamma$, is essentially concentrated in the low MET region.
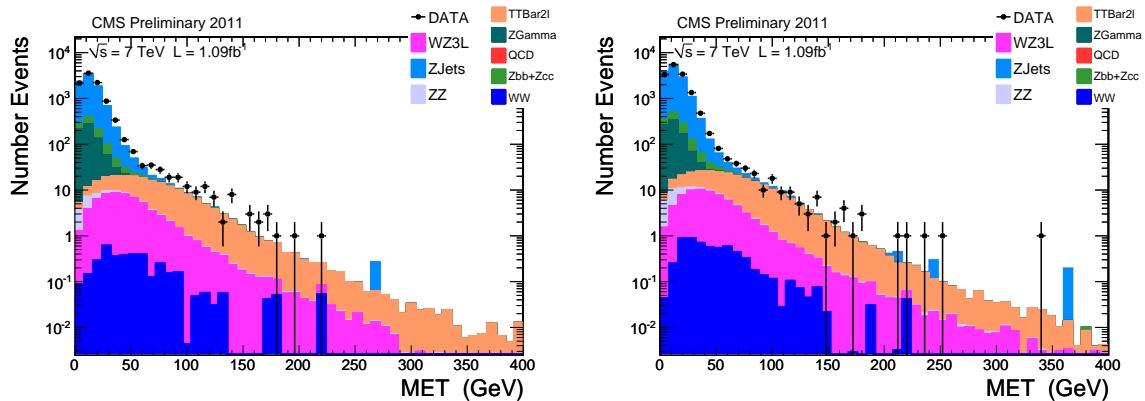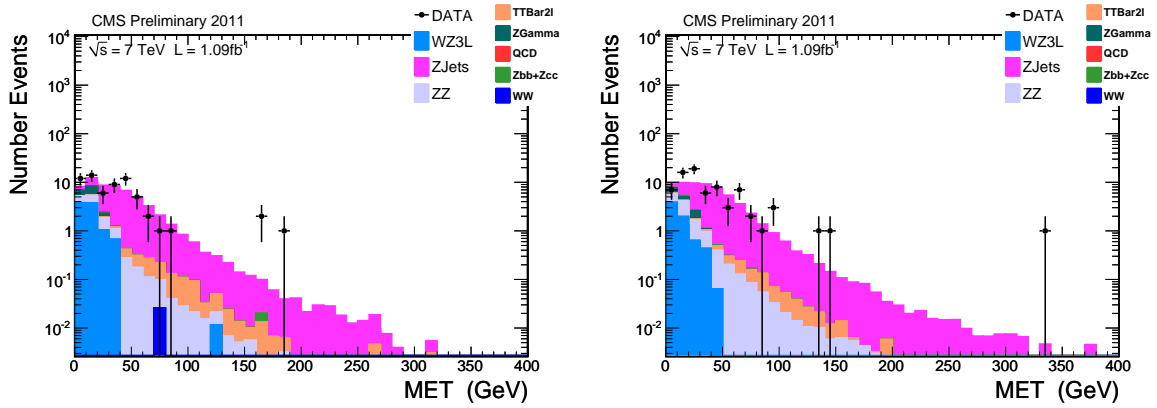

Figure 6.14: Missing transverse energy for the ee (left) and $\mu\mu$ channel (right), after the Z and W candidates are selected. The simulated distributions are normalized to the luminosity in data (1.09/fb).

The W candidate selection is finalized, looking for the high MET produced in presence of a high p_T neutrino. The cut on the missing transverse energy at 30 GeV preserves the 70% of the signal, which is extracted with a very small contamination from backgrounds ($S/B \simeq 8.7$).

For those events selected by the full cut sequence, the invariant mass of the Z candidate is shown in Figure 6.16, for both $Z \to ee$ (left) and $Z \to \mu\mu$ (right) channels.

For each cut applied in the main selection, the event yield is reported in Table 6.6, for both data and simulated samples. At each step, the number of events is shown for each considered channel and by summing them together (combined). The data event yield for the four channels combined is compared to the expectation from MC in Figure 6.17.

Both in Table 6.6 and Figure 6.17, the simulated events are rescaled to 1.09 fb^{-1} of inte-
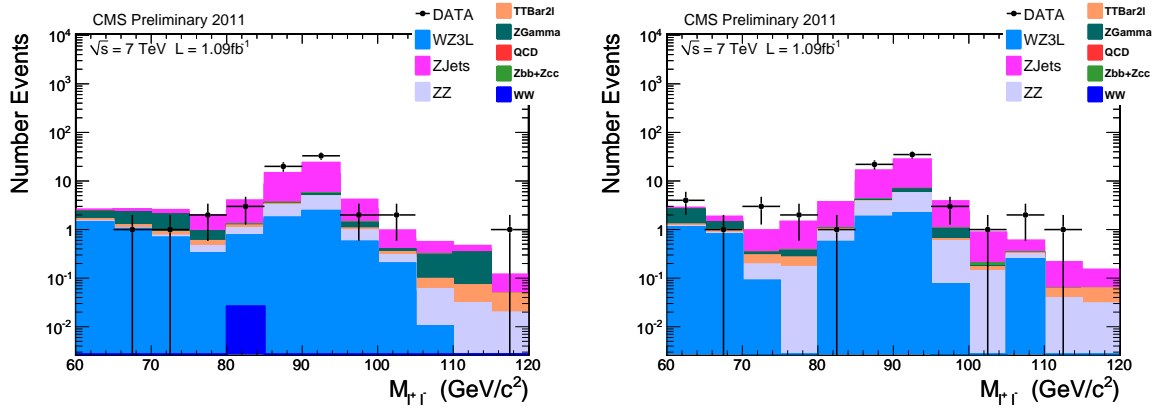
Figure 6.15: Di-lepton invariant mass for the ee (left) and $\mu\mu$ channel (right), after the Z and W candidates are selected. The simulated distributions are normalized to the luminosity in data (1.09/fb).
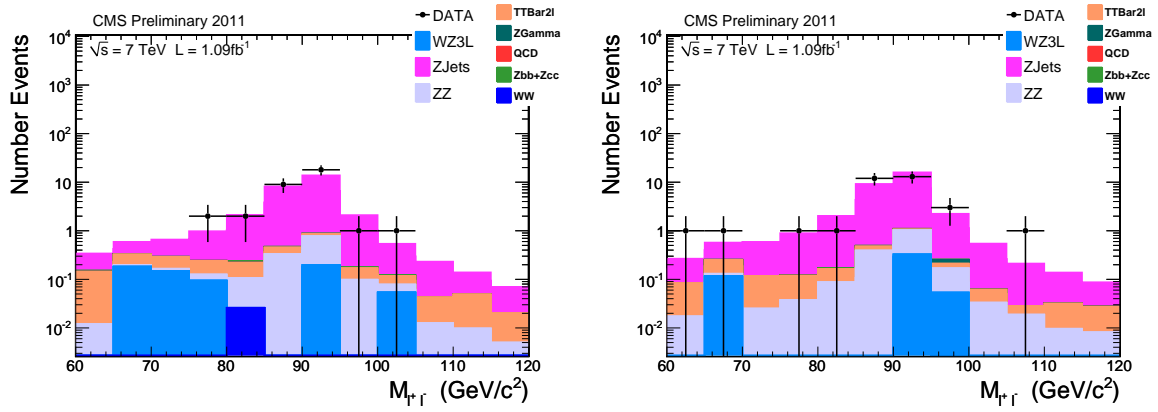
Figure 6.16: Z candidate invariant mass for the ee (left) and $\mu\mu$ channel (right), at the end of the selection. The simulated distributions are normalized to the luminosity in data (1.09/fb).

grated luminosity.

The signal event yield is well balanced among the four channels, with a higher efficiency in presence of muons and in particular for the tight selection on the 3rd lepton from W decay (l^+l^-e vs $l^+l^-\mu$ channels).

The highest background contamination is for the three electrons channel, with a 38% contribution to the background event yield from $Z + jets$, which results the most dangerous process to be controlled.

The $ee\mu$ and $\mu\mu\mu$ channels profit from the high efficiency in identifying muons, especially in case of tight selections, but suffer from a bad rejection power against spurious signals, with ZZ in particular contributing more than 60% to the backgrounds in the only $\mu\mu\mu$ channel.

126

channel	DATA	WW	WZ	ZZ	$t\bar{t}$	$Z\gamma$	QCD	Zbb	$Z+jets$	S+B(MC)
Preselection: HLT requirement + 3 leptons $p_T > 20,10,10$ GeV/c + Vtx compatibility										
combined	84978	25.12	159.36	36.89	1137.56	2207.03	36347.7	1131.61	29311.8	74216.71
eee	14086	5.96	38.55	9.57	209.27	925.49	3613.75	263.16	7663.85	12729.6
$ee\mu$	9423	3.54	36.63	4.60	214.95	74.08	4324.97	196.64	2724.96	7580.36
$\mu\mu e$	31209	12.53	44.57	14.13	503.82	1111.81	11914.	464.42	15278.7	29343.99
$\mu\mu\mu$	35415	3.56	43.26	10.02	340.40	119.55	19589.3	262.90	4193.76	24562.76
step 4: Z selection										
combined	24035	9.88	128.06	29.83	346.1	1458.72	0	838.31	21152.2	23963.16
$Z \to ee$	9547	3.82	59.69	11.23	149.43	646.01	0	346.61	7962.58	9179.38
$Z \to \mu\mu$	14493	6.07	68.36	18.60	196.66	646.01	0	491.69	13189.7	14783.78
step 5: 3rd lepton selection										
combined	152	0.05	83.10	20.26	2.15	8.94	0.	1.47	16.41	132.38
eee	36	0.02	17.11	4.79	0.45	4.65	0.	0.4	8.89	36.32
$ee\mu$	34	0.03	21.99	2.21	0.84	0.16	0.	0.34	1.	26.57
$\mu\mu e$	41	0.	19.15	7.09	0.4	4.13	0.	0.27	4.67	35.71
$\mu\mu\mu$	41	0.	24.84	6.17	0.47	0.	0.	0.46	1.83	33.78
step 6: $2^{nd}Z$ veto										
combined	140	0.05	83.09	11.98	2.15	8.94	0.	1.47	16.41	124.09
eee	36	0.02	17.11	3.1	0.45	4.65	0.	0.4	8.89	34.62
$ee\mu$	29	0.03	21.99	2.0	0.83	0.16	0.	0.34	1.	26.36
$\mu\mu e$	35	0.	19.15	3.29	0.4	4.13	0.	0.27	4.67	31.92
$\mu\mu\mu$	40	0.	24.84	3.58	0.47	0.	0.	0.46	1.84	31.19
step 7: MET cut										
combined	66	0.05	58.18	2.84	1.81	0.32	0.	0.18	1.47	64.86
eee	16	0.02	11.95	0.44	0.38	0.16	0.	0.03	0.64	13.63
$ee\mu$	17	0.03	15.50	0.85	0.7	0.	0.	0.06	0.18	17.33
$\mu\mu e$	13	0.	13.39	0.45	0.33	0.16	0.	0.04	0.37	14.73
$\mu\mu\mu$	20	0.	17.34	1.10	0.4	0.	0.	0.05	0.27	19.17

Table 6.6: Observed and expected signal and background yields at the different step of the selection. The numbers correspond to an integrated luminosity of 1.09/fb.

Figure 6.17: Number of events at each step of the event selection. The simulated distributions are normalized to the luminosity in data (1.09/fb).

127

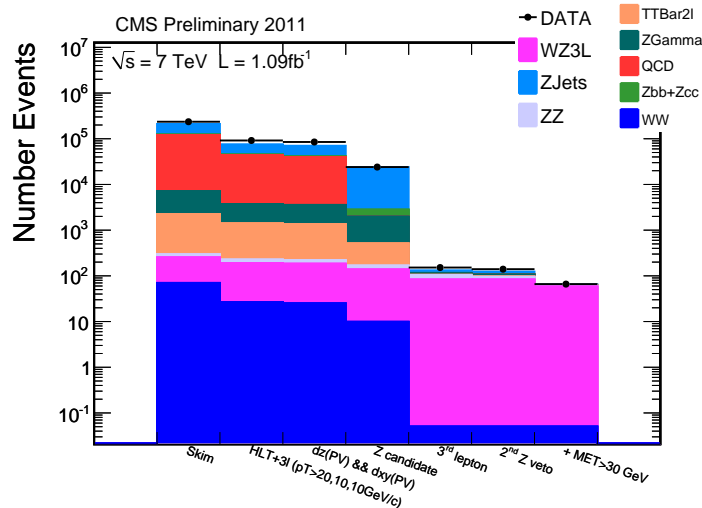Such an important contamination from ZZ is expected, given the crucial role of the MET cut in the final background rejection. Although transverse energy is well balanced in a ZZ event, in case a muon falls outside the detector acceptance ($|\eta| > 2.4$ for tracker and muon chambers) it crosses the detector undetected, generating a spurious MET. This is not the case for electrons, since the calorimetric coverage is provided over a wider range, almost down to the beam pipe ($|\eta| \simeq 3$ for ECAL and HCAL and $|\eta| \simeq 5$ for the HF).

The other processes give a negligible contamination, with $Z\gamma$ contributing to the eee and $\mu\mu e$ channels in case a final state photon converts and mimics an electron.

The QCD process, while being the dominant background up to the Z candidate selection, is completely killed afterwards. Although multi-jet processes are generally not perfectly modeled by the simulation and only a data driven approach can provide a reliable estimate of their event yield, in the context of this analysis QCD can be fully taken from the simulation, since it's safely not a background to the three high p_T leptons signature of the final state.

For the events at the end of the selection, the distribution of the Z candidate invariant mass is shown in Figure 6.18, for each of the four channels considered. The same distributions are shown in Figure 6.19 for the MET variable, before applying the cut on this variable.

6.3 The backgrounds estimation

A good control of the background processes, with eventual bias being covered by the systematic uncertainties, allows to deal with a reliable analysis monitored through the progressive event reduction towards the signal extraction.

The MC modeling can be safely trusted for those backgrounds, which negligibly contribute to the final state selection, as well as for the diboson processes, which are well reproduced in the simulation. The leptonic final states in particular are generated with processes well tested in the MC generation and can be safely exploited.

The event yield for ZZ, $Z\gamma$, Zbb and WW is estimated from the MC. A 20% systematic uncertainty is quoted as precision on the electroweak background processes. For ZZ and $Z\gamma$ a 7.5% and 13% uncertainty is assumed respectively, being the precision on the cross section as measured by CMS.
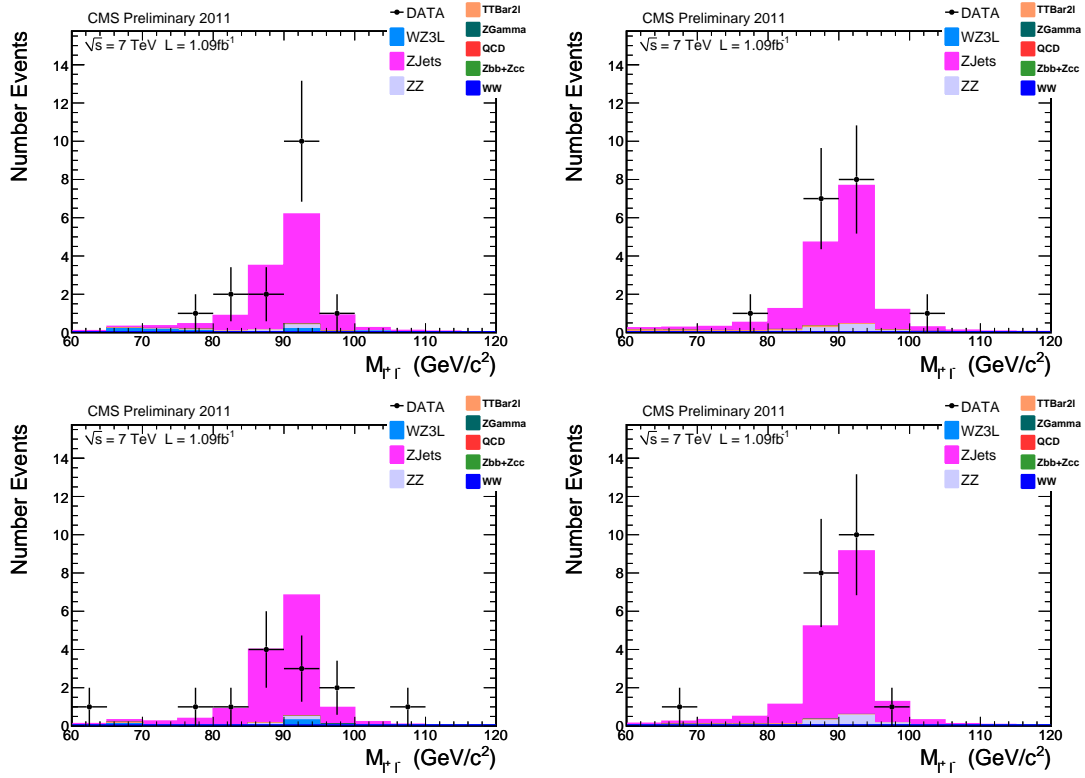
Figure 6.18: The Z candidate invariant mass is shown for the events at the end of the selection. The distributions are shown for each of the four final states considered, eee (top left), $ee\mu$ (top right), $\mu\mu e$ (bottom left), $\mu\mu\mu$ (bottom right). The simulated distributions are normalized to the luminosity in data (1.09/fb).

Good performances are being reached also in the simulation of the hadronic final states, greatly improved during the last years of hadron colliders experience. Nevertheless the event yield estimate through data driven techniques is of further support to control the jet production processes.

Concerning the $t\bar{t}$ background, the combinatorics produced by the $WbWb$ decays can lead to the same final state topology $3l + MET$. In particular leptons from W present very similar features to the signal ones and those from $b - jet$s can pass the identification and isolation criteria and have a high enough p_T to be selected.

6.3.1 $Z + jets$ data driven approach

To cross-check the $Z + jets$ MC estimate, the event yield in data is compared to the one obtained in the simulation in a control region dominated by $Z + jets$, and the event reduction is monitored along the main cuts of the selection chain.
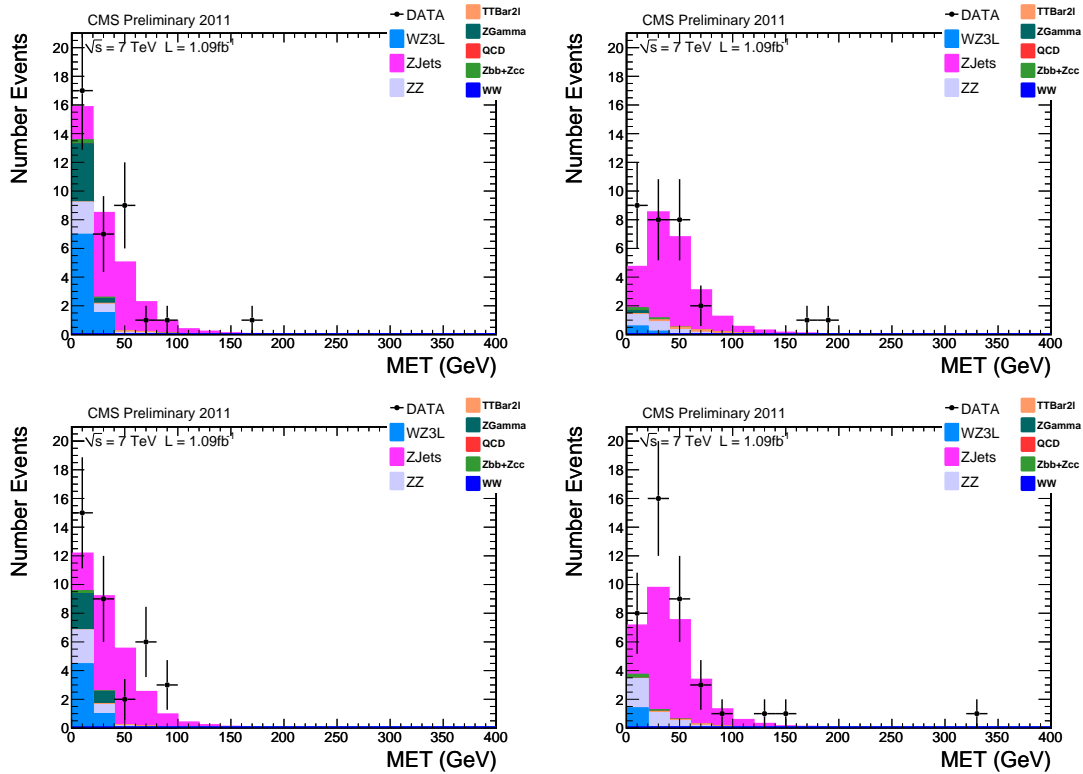
Figure 6.19: The event MET distribution is shown before the cut, for each of the four final states considered: eee (top left), $ee\mu$ (top right), $\mu\mu e$ (bottom left), $\mu\mu\mu$ (bottom right). The simulated distributions are normalized to the luminosity in data (1.09/fb).

A first control region enriched in $Z+jets$ is obtained by requiring a di-lepton pair satisfying the Z candidate selection and inverting the MET cut (MET <30 GeV required). Figure 6.20 presents the corresponding MET distributions for data and MC for both the ee and $\mu\mu$ channels.

The $Z+jets$ background is seen to entirely dominate the control region, with contributions of 88% and 91% to the total MC event yield in $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ channels respectively. In this control region, data and simulation are in very good agreement both in terms of shape modeling and event yield.

The $Z+jets$ data over MC ratio can be estimated as $(data - MC_{Z+jets\ subtracted})/Z+jets$ and it is measured 1.06 (for the ee channel) and 1 (for the $\mu\mu$ channel).

Moving closer to the signal phase space, another control region is defined by asking for the 3rd lepton, but keeping the Id and isolation requirements relaxed. The conversion rejection is instead maintained. Figure 6.21 presents the corresponding MET distributions for data and MC for both the ee and $\mu\mu$ channels.
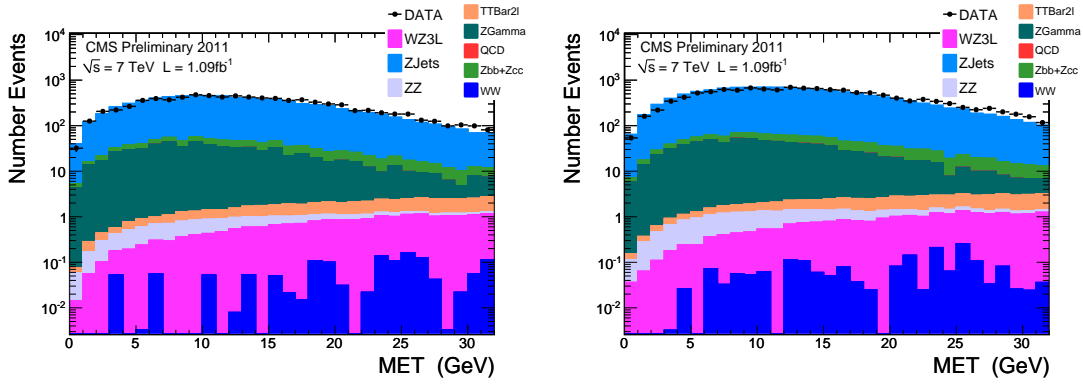
Figure 6.20: MET distributions in the ee and $\mu\mu$ channels after the Z candidate selection and for MET <30 GeV. The simulated distributions are normalized to the luminosity in data (1.09/fb).
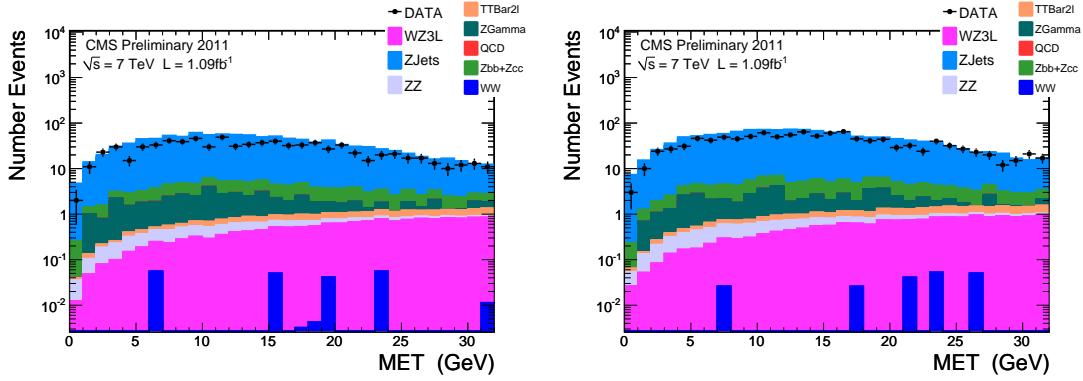
Figure 6.21: MET distributions in the ee and $\mu\mu$ channels after the 3rd lepton selection (relaxing Id and Iso, while keeping the conversion rejection). The simulated distributions are normalized to the luminosity in data (1.09/fb).

The $Z + jets$ background is again entirely dominating the control region, with a contribution to the MC of 92% in ee and 93% in $\mu\mu$, fully consistent with the values found in the previous case.

The WZ signal contamination in this region is order 0.7% and 0.6% for $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ respectively. Again data and MC are in good agreement in this control region and the $Z + jets$ data over MC ratio expressed as $(data - MC_{Z+jets\ subtracted})/Z + jets$ is 0.92 (ee) and 0.98 ($\mu\mu$).

The $Z + jets$ event yield is controlled in data within $\sim$10%. To take into account eventual differences between the MET shape for data and MC, while moving from the control to the signal region, a 20% systematic uncertainty is conservatively assigned.

Further data driven cross-checks

Within the context of the CMS community, a dedicated working group finalized an analysis for the WZ search [107]. In such analysis, both $t\bar{t}$ and $Z + jets$ backgrounds were measured from data. Such estimates are here considered as a further compatibility cross-check of the MC event yield listed in Table 6.6, since the selection strategy followed in [107] is really closed to the one developed and presented in this analysis. In particular from this comparison, a systematic error on the $t\bar{t}$ MC event yield is assigned.

$t\bar{t}$ from control region

The data/MC ratio measured in a $t\bar{t}$ enriched control region (CR) is used to rescale the MC $t\bar{t}$ background in the signal region (SR), obtaining a data driven $t\bar{t}_{SR}$ estimate. The defined control region corresponds to the signal extraction cuts but

- reverting the isolation cut on the less isolated lepton (greater than 1.5);

- cutting on the b-tag discriminant[11] value (greater than 2.5);

The MC and data-driven $t\bar{t}$ measured in the signal region are reported in Table 6.7.

Channels	MC $t\bar{t}_{SR}$	Data $t\bar{t}_{SR}$
eee	0.45±0.04	0.48±0.25
$ee\mu$	0.73±0.04	0.61±0.21
$\mu\mu e$	0.24±0.04	0.37±0.10
$\mu\mu\mu$	0.38±0.04	0.24±0.09

Table 6.7: From [107], $t\bar{t}$ background yield from MC and estimated in data through a control region. Event counts correspond to 1.09/fb integrated luminosity.

The MC and data-driven $t\bar{t}$ event yield just shown are in agreement within order 2σ. Moreover the $t\bar{t}$ MC event yield shown in Table 6.6 is compatible with both these estimates (Table 6.7), conservatively within 50%.

Thus, a 50% systematic uncertainty is assigned to the $t\bar{t}$ event yield in this analysis.

$Z + jets$ and $t\bar{t}$ control with the Matrix Method

The "matrix method" was used to measure the combined contribution of $Z + jets$ and $t\bar{t}$ backgrounds entirely from data.

The measured yield in data for the signal-like and background-like events is given in Table 6.8 in the first and second column respectively. The event counts for the signal-like

---

[11]Discriminating variable provided by the b-tagging Track Counting High Efficiency algorithm [108] applied to the jet with highest p_T in the event

events based on the MC information $TrueN_{lep}$ is reported in the last column.

Type	$\epsilon_{\text{tight}} \cdot N_{\text{lep}}$	$P_{\text{fake}} \cdot N_{\text{jet}}$	$TrueN_{lep}$
eee	20.24 ± 4.76	1.76 ± 0.67	14.47 ± 3.80
$ee\mu$	17.46 ± 4.56	2.54 ± 0.86	17.49 ± 4.18
$\mu\mu e$	11.40 ± 3.67	1.60 ± 0.58	13.95 ± 3.73
$\mu\mu\mu$	17.82 ± 4.54	2.18 ± 0.76	18.56 ± 4.31

Table 6.8: From [107], signal-like and background-like ($t\bar{t}$ and $Z + jets$) event yield as measured with the "matrix method". The signal-like contribution measured in the MC is also reported. Event counts correspond to 1.09/fb integrated luminosity.

From a closure test between the signal event yield in Table 6.6 and the signal-like estimate from MC (Table 6.8 last column), the "matrix method" and the MC expectation are found in agreement within ∼25%.

By comparing the $(Z+jets) + (t\bar{t})$ event yield from MC (Table 6.6), to the $P_{\text{fake}} \cdot N_{\text{jet}}$ values measured with the "matrix method", the two estimates are found well compatible within the total uncertainties quoted.

6.3.2 WZ to taus background

$WZ \rightarrow 3l\nu$ represents a background for all final states when W or Z bosons decay to τ leptons.

This background cross section is considered proportional to the WZ cross section itself and it is therefore subtracted from the observed signal yield as a fraction determined from MC. In Table 6.9, the inclusive leptonic WZ event yield is reported together with the contribution from τ channels.

Sample	eee	$ee\mu$	$\mu\mu e$	$\mu\mu\mu$
$WZ \rightarrow 3l\nu(l = e, \mu, \tau)$	11.95	14.50	13.39	17.34
$WZ \rightarrow 3l\nu(l = e, \mu, \tau) - WZ \rightarrow 3l\nu(l = e, \mu)$	0.84	1.05	0.88	1.16

Table 6.9: Expected yield for WZ MC sample in leptonic channels and considering τ decays of W or Z bosons, obtained after all selection steps. The numbers correspond to an integrated luminosity of 1.09/fb.

6.4 The cross section measurement

For a given process, it's production cross section is measured in a fiducial region, representing the geometric and kinematic acceptance of the signature looked for.

The production cross section extrapolated to a wider acceptance is given by:

$$\sigma = \frac{N_{sig}}{A \cdot \epsilon \cdot \mathcal{L}} \tag{6.3}$$

where N_{sig} is the number of observed signal events, A represents the fiducial acceptance of the signature looked for with respect to the acceptance of the extrapolation, ϵ is the selection efficiency for the events in the acceptance and $\mathcal{L}$ is the integrated luminosity.

The value of A is affected by the PDF and other theoretical uncertainties, while the value of ϵ is susceptible to errors from triggering and reconstruction.

In order to control the uncertainties on ϵ, the efficiencies obtained from the simulation (ϵ_{sim}) are corrected with factors (ρ) computed as the efficiency ratios $\rho = \epsilon_{data}/\epsilon_{sim}$, derived by measuring ϵ_{data} and ϵ_{sim} in the same way on simulation and data respectively (see Section 6.4.1).

$A \times \epsilon$ in equation 6.3 can be thus replaced by $A \times \epsilon_{sim} \times \rho$, where $A \times \epsilon_{sim}$ is the fraction of the WZ events decaying in electron and muon only final states generated in the kinematic range M_Z within $[60,120]\text{GeV/c}^2$, $p_{TW} > 20$ GeV/c.

The values of $A \cdot \epsilon_{sim}$ and ρ are given in Table 6.10 for each of the four channels. The correction factors ρ computed for the efficiency on the WZ event selection are measured in Section 6.4.1.

Channel	$A \cdot \epsilon_{sim}$	ρ
eee	$0.154 \pm 0.001(stat)$	0.97 ± 0.08
$ee\mu$	$0.200 \pm 0.001(stat)$	1.00 ± 0.06
$\mu\mu e$	$0.173 \pm 0.001(stat)$	0.94 ± 0.05
$\mu\mu\mu$	$0.224 \pm 0.001(stat)$	0.97 ± 0.05

Table 6.10: $A \cdot \epsilon_{sim}$ and ρ_{eff} per channel

Moreover, the number of signal events N_{sig} is not directly measured, but it is obtained by subtracting the estimated number of background events N_{bkg} from the observed number of selected WZ candidate events N_{obs}.

WZ to taus background is also taken into account and subtracted as a fraction of N_{sig} itself, where f_τ is the fraction of WZ to taus events estimated from MC and given in Table 6.9. In such a way this background is subtracted as a fraction of the same WZ cross section that is being measured and does not depend on the WZ MC cross section.

Equation 6.3 can therefore be rewritten as

$$\sigma = \frac{(N_{obs} - N_{back}) \cdot (1 - f_\tau)}{A \cdot \epsilon_{sim} \cdot \rho \cdot \mathcal{L}},$$

(6.4)

6.4.1 Efficiencies measurement from data

To measure the inclusive selection efficiencies for electrons and muons, a "tag-and-probe" technique [109] is used.

The idea is to combine the requirements of a mass constraint for a pair of basic objects (i.e. tracks for muons, or clusters of calorimetry cells for electrons) with a tight selection (order WP60) applied on one leg (the "tag"), so to ensure sufficient purity. The other leg (the "probe") is used to measure the efficiency of a given reconstruction algorithm or identification criterium.

The inclusive efficiency is defined as the ratio of the number of passing probes to the total number of probes before the cut. It can be factorized in a series of terms, each corresponding to different stages of the selection and that can be measured independently:

$$\epsilon = \epsilon_{HLT} \cdot \epsilon_{Reco|Clustering} \cdot \epsilon_{Id|Reco} \cdot \epsilon_{Iso|Id}$$

where each term represents the efficiency for the probe to pass a given selection or reconstruction step, provided that it fulfills the criteria for the previous one.

For each factor to be measured, the distributions obtained in data for passing and failing probes are fitted with a proper modeling of the signal and the background shapes[12], while a simple event counting can be performed on MC [107].

For each ϵ_{data} the systematic uncertainty is determined by varying the background fit modeling [13].

Electron reconstruction

The results are presented in Table 6.11 for SC with $E_T >$20 GeV in different pseudo-rapidity bins. A very good agreement is found between data and MC, the efficiency assumed for electrons to be 100% is confirmed by MC within few per mille.

The above data/MC (ρ) factors, measured in pseudo-rapidity bins, are normalized over the W and Z lepton pseudo-rapidity distribution of selected signal MC events, resulting in ρ (ele RECO) = 0.997± 0.009.

As discussed in Section 6.2.1, an ad hoc selection is used to emulate the High Level Trigger for electrons, providing a close to 100% efficiency for signal-like events. A negligi-

---

[12]For $Z \rightarrow ee$ the function used is a Voigtian = (Lorentzian*Gauss)*Crystal Ball. For $Z \rightarrow \mu\mu$ the function used is a Gauss*Crystal Ball. An Exponential shape is used to fit the background.
[13]Exponential*Erf and polynomial shapes are used as alternative.

	data	MC	data/MC
$0 < |\eta| < 0.78$	$97.0 \pm 0.11 (stat.) \pm 0.98 (syst.)$	$98.08^{+0.03}_{-0.04}$	0.999 ± 0.010 (stat. + syst.)
$0.78 < |\eta| < 1.444$	$97.35^{+0.17}_{-0.19}(stat.) \pm 1.22 (syst.)$	97.6 ± 0.04	0.997 ± 0.012 (stat. + syst.)
$1.444 < |\eta| < 1.566$	$94.18 \pm 0.57 (stat.) \pm 0.12 (syst.)$	$95.74^{+0.17}_{-0.18}$	0.983 ± 0.006 (stat. + syst.)
$1.566 < |\eta| < 2.0$	$95.24 \pm 0.25 (stat.) \pm 0.14 (syst.)$	96.0 ± 0.07	0.992 ± 0.003 (stat. + syst.)
$2.0 < |\eta| < 2.5$	$93.66^{+0.31}_{-0.32}(stat.) \pm 0.14 (syst.)$	93.48 ± 0.1	1.002 ± 0.004 (stat. + syst.)

Table 6.11: Electron reconstruction efficiency from simultaneous fits of passing and failing samples for different η regions on data and MC. A Z mass window of $60 < m(e, SC) < 120$ GeV/c^2 and a WP60 selection for the tag is used. Efficiencies are in %. Systematic uncertainties are added in quadrature in the given data/MC ratios.

ble overall data/MC discrepancy in the trigger efficiency for the signal and the irreducible backgrounds is assumed and a systematic uncertainty of 1.5% is assigned.

Muon reconstruction

The efficiency to reconstruct a muon in the inner tracker has been measured and found to be 99% or higher in the whole tracker acceptance, in good agreement with the expectations from simulations [52]. Results are reported in table 6.12.

	data	MC	data/MC
ϵ_{trk}	0.989 ± 0.001	0.995 ± 0.001	0.994 ± 0.001
ϵ_{sta}	0.981 ± 0.002	0.984 ± 0.001	0.997 ± 0.002
$\epsilon_{HLT(13\ GeV)}$	0.962 ± 0.001	0.976 ± 0.001	0.986 ± 0.002
$\epsilon_{HLT(8\ GeV)}$	0.962 ± 0.001	0.976 ± 0.001	0.986 ± 0.002

Table 6.12: Efficiency values obtained from the "tag-and-probe" fits on muons, for the reconstruction and HLT steps. The errors quoted are purely statistical, a 0.5% systematic error is assigned on each efficiency measurement when considering the analysis results.

Electron selection

By following the above procedure, efficiencies are determined for the electron identification and isolation. In these cases the probes are reconstructed GSF electrons passing respectively the identification and the isolation criteria.

Efficiencies are determined for both the loose (WP95) and the tight (WP80) selection used in the analysis and are reported in Table 6.13. Since in the analysis, loose requirements are used for the Z candidate selection together with a p_T threshold of $p_T > 10$ GeV/c, the efficiencies are determined in this case for $p_T > 10$ GeV/c. In the case of the tight (WP80) selection, the efficiencies are determined for $p_T > 20$ GeV/c.

In Figure 6.22 the fit of the M_{ll} distributions for the case of the electron Id efficiency measurement and for the tight (WP80) selection is shown.
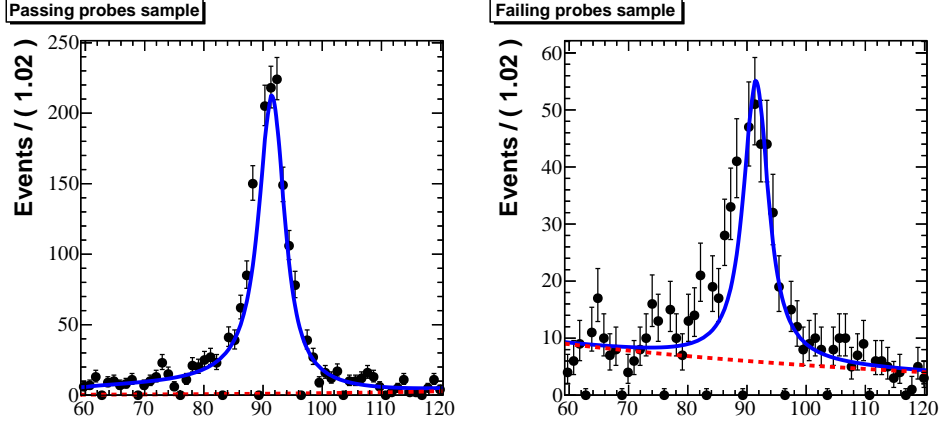

Figure 6.22: M_{ll} distribution for passing and failing probes used for to measure the electron efficiency for WP80-Id selection.

	data	MC	data/MC
WP95 Id	0.941± 0.012	0.937	1.004 ± 0.016 (stat. + syst.)
WP80 Id	0.839± 0.019	0.848	0.990 ± 0.025 (stat. + syst.)
WP95 Iso	0.902± 0.014	0.902	1.000 ± 0.019 (stat. + syst.)
WP80 Iso	0.904± 0.014	0.919	0.984± 0.019 (stat. + syst.)

Table 6.13: Electron identification and isolation efficiency from simultaneous fits of passing and failing samples for different WPs. A Z mass window of $60 < m(e, SC) < 120$ GeV/c^2 and a WP60 selection for the tag is used. Systematic uncertainties are added in quadrature in the given data/MC ratios.

Muon selection

The measured efficiencies for the muon isolation and identification selections are listed in Table 6.14.

	data	MC	data/MC
ϵ_{Id}	0.975 ± 0.001	0.977 ± 0.001	0.998 ± 0.001
ϵ_{Iso}	0.966 ± 0.001	0.957 ± 0.001	1.010 ± 0.002

Table 6.14: Efficiency values obtained from the tag and probe fits, for the Id and Iso requirements. The errors quoted are purely statistical, a 0.5% systematic error is assigned on each efficiency measurement when considering the analysis results.

6.4.2 The ρ_{WZ} correction factor

The data/MC (ρ) factors measured for the single steps and summarized above are combined, by considering the same criteria used while selecting the W and Z leptons in the analysis.

The obtained ρ_W and ρ_Z are in turn combined to extract the ρ_{WZ} correction factor corresponding to the WZ event selection. The results are summarized in Table 6.15 for each of the four channels in the final state.

Channel	$\rho_{eff}(WZevent)$	$\rho_{eff}(Zlepton)$	$\rho_{eff}(Wlepton)$
eee	0.97 ± 0.08	1.00 ± 0.03	0.97 ± 0.03
$ee\mu$	1.00 ± 0.06	1.00 ± 0.03	1.00 ± 0.01
$\mu\mu e$	0.94 ± 0.05	0.98 ± 0.01	0.97 ± 0.03
$\mu\mu\mu$	0.97 ± 0.05	0.98 ± 0.01	1.00 ± 0.01

Table 6.15: data/MC (ρ) efficiency computed for the Z and W boson selection and combined to measure the WZ event selection efficiency ratio.

The measured uncertainties on the data/MC efficiency ratios (ρ_{WZ}) are taken as systematic uncertainty.

6.4.3 Scale and resolution effects

Eventual biases on the scale and resolution of the ingredients used in the WZ analysis can significantly affect the result. As in [107], such parameters are estimated and taken into account, since contributing to the systematic precision on the final measurement.

Electron energy scale

The electron momentum scale can be controlled using Z boson decays to electrons. Fits to the Z line-shape selecting electrons with $E_T > 25$ GeV show differences in scale between data and MC of 0.1% (0.2%) in the ECAL barrel (ECAL endcaps) [110]. The electron momentum scale uncertainty also receives other contributions, mainly from the ECAL laser corrections and linearity. Overall, the ECAL energy scale is known to 0.6% in EB and to 1.5% in EE.

A 2% electron momentum scale uncertainty is conservatively assumed and propagated through the analysis chain, reflecting in a maximum of 1.7% systematic error, for the three electron channel.

Muon energy scale

From 2010 studies on J/ψ and Z resonances [111, 112], the absolute momentum scale is measured for muons within 1% in both data and MC. Such uncertainty is propagated

through the analysis chain and results a negligible effect.

MET resolution and scale

The MET resolution depends essentially on how well the event is measured and it is not affected by the neutrino itself.

By following the method described in [113], the resolution on the MET is measured in events where the role of the neutrino is played by a well measured object that is removed from the event.

For instance in events with a reconstructed $Z \to ll$ decay, the transverse hadronic recoil vector $\vec{u}$ can be defined, where $\vec{q}_T$ is the Z boson momentum, as

$$\vec{u} = -\vec{MET} - \vec{q}_T \tag{6.5}$$

With perfect resolution, $\vec{u} = -\vec{q}_T$. The recoil vector can be decomposed in the components u_1 and u_2, parallel and orthogonal to $\vec{q}_T$ respectively, with the mean values of $|u_i|$ measuring the MET response and the spread of u_i reflecting the MET resolution.

Both $|u_i|$ and σ_{u_i} are fitted as a function of p_{TZ}, with a linear and quadratic function [107] respectively and the same procedure is applied to data and MC.
The measured differences between data and MC estimates are taken as systematic uncertainties and propagated through the analysis chain.
The resulting systematic error on each of the four final states is $\sim 0.3\%$ for the MET scale and $\sim 0.5\%$ for the MET resolution.

Theoretical uncertainties

As reported in [107] the theoretical systematic uncertainties, both from PDF and from higher order QCD (NLO) contributions are considered and their effect on the signal acceptance is studied.

The systematic uncertainty from the PDF is estimated at 0.93%, by using `LHAPDF` and PDF re-weighting methods with different PDF runs through MC generation [114] and by studying the variation on the signal acceptance by means of the `CTEQ66` PDF library.

The acceptance used for the cross section measurement is determined on signal sample produced with `PYTHIA`, where higher order QCD corrections are not taken into account. By comparing the generator acceptance computed with `MCFM` [115] at Born level and at NLO, the effect of higher order corrections (difference between the two values) is taken

as systematic uncertainty and it amounts to 2.5%.

6.4.4 Systematic uncertainties summary

With reference to [107], the systematic uncertainties relevant for the $\sigma(WZ)$ measurement here presented are summarized in Table 6.16.

Also the effect of charge mis-identification on both electrons and muons has been considered and found negligible within the context of this analysis.

Source	Systematic uncertainty	eee	$ee\mu$	$\mu\mu e$	$\mu\mu\mu$
		Effect on $\mathcal{F} = A \cdot \epsilon_{sim}$			
Electron energy scale	2%	1.7%	0.25%	0.9%	n/a
Muon p_T scale	1%	n/a	0.5%	0.2%	0.9%
MET Resolution		0.5%	0.5%	0.5%	0.5%
MET Scale		0.3%	0.2%	0.1%	0.1%
Pileup		3.1%	0.8%	1.6%	1.6%
PDF	1.0%	1.0%	1.0%	1.0%	1.0%
NLO effect	2.5%	2.5%	2.5%	2.5%	2.5%
Total uncertainty on $\mathcal{F} = A \cdot \epsilon_{sim}$		4.5%	2.9%	3.3%	3.3%
		Effect on ρ_{WZ}			
Electron trigger	1.5%	1.5%	1.5%	n/a	n/a
Electron reconstruction	0.9%	2.7%	1.8%	0.9%	n/a
Electron ID and isolation	2.5%(WP95), 3.2%(WP80)	5.9%	5.0%	3.2%	n/a
Muon trigger	0.54%	n/a	n/a	1.08%	1.08%
Muon reconstruction	0.74%	n/a	0.74%	1.48%	2.22%
Muon ID and isolation	0.74%	n/a	0.74%	1.48%	1.94%
Total uncertainty on ρ_{eff}		6.7%	5.6%	4.2%	3.6%
Background estimation		Effect on WZ yield			
ZZ	7.5%	0.2%	0.5%	0.3%	0.5%
$Z\gamma$	13%	0.2%	0.%	0.2%	0.%
$other$	20%	0.1%	0.1%	0.07%	0.06%
$Z + jets$	20%	1.0%	0.3%	0.7%	0.3%
$t\bar{t}$	50%	1.4%	2.4%	1.5%	1.2%
		Effect on $\mathcal{L}$			
Luminosity	6.0%	6.0%	6.0%	6.0%	6.0%

Table 6.16: Summary of systematic uncertainties for the WZ $\to 3\ell$ cross section measurement.

6.5 Results

6.5.1 Single channel cross section

The WZ cross section measured in the fiducial region and extrapolated to the total acceptance can be computed for each of the final states considered as

$$\sigma(pp \to WZ + X) \times \mathcal{BR}(W \to \ell\nu) \times \mathcal{BR}(Z \to \ell\ell) = \frac{(N_{obs} - N_{back}) \cdot (1 - f_\tau)}{A \cdot \epsilon_{sim} \cdot \rho \cdot \mathcal{L}}, \quad (6.6)$$

The value measured in $\int \mathcal{L}\, dt = 1.09 fb^{-1}$ of integrated luminosity in the four channels eee, $ee\mu$, $\mu\mu e$ and $\mu\mu\mu$ are reported in Table 6.17.

channel	N_{obs}	N_{bkg}	$\sigma \times BR$ (pb)
$eee\nu$	16	2.69	$0.081^{+0.029}_{-0.022}(stat) \pm 0.007(syst) \pm 0.005(lumi)$
$ee\mu\nu$	17	2.85	$0.065^{+0.022}_{-0.017}(stat) \pm 0.004(syst) \pm 0.004(lumi)$
$\mu\mu e\nu$	13	2.12	$0.061^{+0.025}_{-0.018}(stat) \pm 0.003(syst) \pm 0.004(lumi)$
$\mu\mu\mu\nu$	20	3.04	$0.072^{+0.022}_{-0.017}(stat) \pm 0.004(syst) \pm 0.004(lumi)$

Table 6.17: Number of observed WZ candidate events in the individual final states. The statistical errors account for Poisson 68% C.L. intervals. The background event yield is listed for each channel, accounting also for WZ to *taus* contribution. The cross sections computed correspond to $\int \mathcal{L}\, dt = 1.09 fb^{-1}$.

6.5.2 Cross Section Combination

The inclusive $\sigma(WZ)$ measurement is obtained as the combination of the different N_{sig_i} observations performed on the subset of ith decay channels, assuming that N_{sig_i} is distributed following a Poisson distribution of mean $= \sigma \times BR_i(A_i\epsilon_i\mathcal{L})$, where the branching ratio of each channel BR_i is known

$$\sigma \times BR_i = \frac{N_{sig_i}}{A_i \cdot \epsilon_i \cdot \mathcal{L}} \tag{6.7}$$

This assumption is particularly suitable when dealing with small statistical samples, whose fluctuations are well described by Poisson *pdf*.

Through the Maximum Likelihood method the best unbiased estimator of the combination can be computed. It corresponds to

$$\sigma = \frac{\sum_i N_{sig_i}}{\mathcal{L} \cdot \sum_i (BR_i A_i \epsilon_i)} \tag{6.8}$$

By following this approach, the inclusive WZ production cross section extrapolated to the full acceptance is

$$\sigma(WZ) = 19.11^{+3.30}_{-2.53}(\text{stat.}) \pm 1.10(\text{syst.}) \pm 1.15(\text{lumi.})\text{pb} \tag{6.9}$$

where the statistical errors correspond to Poisson 68% C.L. intervals. The systematic errors quoted account for the correlation between the systematic uncertainties, namely

- among all channels:
    - The PDF and NLO uncertainties on the acceptance of the selection criteria;

141

- The uncertainty on MET energy scale and on pile-up;

- The uncertainty on cross section of subtracted backgrounds;

- among the channels with electrons:

    - The electron reconstruction, identification, isolation, charge-id and energy scale uncertainties;

    - The electron HLT efficiency uncertainty ($Z \to e^+e^-$ channels only);

- among the channels with muons:

    - The muon reconstruction, identification, isolation and energy scale uncertainties;

    - The muon HLT matching uncertainty ($Z \to \mu^+\mu^-$) channels only;

6.5.3   Discussion

This measurement relies on the observation performed in the kinematic range $60< M_Z <120$ GeV/c^2, $p_{TW} >20$ GeV/c in the e, μ channels only, in 1.09 fb^{-1} of integrated luminosity of data, corresponding to the LHC collisions delivered up to July 2011.

The result presented corresponds to the inclusive WZ production cross section, extrapolated to full acceptance, thus the expectation from the Standard Model next-to-leading order calculation [11] of $18.57^{+0.75}_{-0.58}$pb can be taken as reference value for the comparison with theory.

The quoted value 6.9 has a precision dominated by the statistic uncertainty (15%), while its systematic uncertainty accounts for 5.7%. It is fully compatible with the NLO Standard Model expectation, within the 4% uncertainty on the theory estimate.

Other existing measurements

As already mentioned at the end of the first Chapter (1.3.3), the inclusive $\sigma(WZ)$ has been measured for both $p\bar{p}$ collisions at $\sqrt{s} =1.9$ TeV and pp collisions at $\sqrt{s} =7$ TeV, by the collaborations working at Tevatron and LHC respectively.
In summer 2011 a measurement at $\sqrt{s}=7$ TeV has been published by the ATLAS collaboration based on order 1 fb^{-1} of integrated luminosity.
CMS also published its first result for the summer 2011 conferences, based on an analysis which is very close to the one presented in this thesis, though with some differences that will be further discussed in the following.

The latest measurements of $\sigma(WZ)$ are reported in Table 6.18, together with the luminosity of data analyzed and the corresponding NLO Standard Model expectation from theory.

	Measured $\sigma(WZ)$ [pb]	$\mathcal{L}$ [fb^{-1}]	Theory $\sigma_{\mathrm{NLO}}(WZ)$ [pb]
CDF [22]	$3.9^{+0.8}_{-0.7}$(stat+sys)	7.1	3.46 ± 0.21
D0 [23]	$3.89^{+1.02}_{-0.85}$(stat+sys)± 0.31(lumi.)	4.1	3.25 ± 0.19
ATLAS [24]	$20.5^{+3.1}_{-2.8}$(stat.)$^{+1.4}_{-1.3}$(syst.)$^{+0.9}_{-0.8}$(lumi.)	1.02	$17.3^{+1.3}_{-0.8}$
CMS [25]	17.0 ± 2.4(stat.)± 1.1(syst.)± 1.0(lumi.)	1.01	17.53 ± 0.55

Table 6.18: WZ inclusive production cross section measured in $p\bar{p}$ collisions at 1.96 TeV (CDF and D0) and in pp collisions at 7 TeV (ATLAS and CMS).

Tevatron experiments measured $\sigma(WZ)$, profiting from order 5 times more statistics than the LHC measurements. CDF and D0 results quote a total precision of 21% and 27% respectively. Within the experimental errors both measurements are in good agreement with the expectation from theory, which is accurate at a $\sim$6% level.

First measurements provided by ATLAS and CMS in 2011 used order 1/fb of integrated luminosity and agree with the Standard Model expectation within $\sim$1σ.
ATLAS result quotes a precision of 18%, with a major contribution (15%) from the statistic uncertainty and a systematic precision of the order 7%.
CMS result has a statistical uncertainty of 14% and a systematic error of 6.5%.
The CMS public analysis is commented further in the following, considering similarities and differences with respect to the one presented in this thesis work.

Measurements with the CMS detector

The analysis presented in this thesis has been a main contribution to the CMS public result [25] in which I was strongly involved. Hence not surprisingly, the CMS public analysis uses a similar cut based selection. Nevertheless, newer MC samples made available in summer 2011 are used here. Moreover, different approaches are followed to estimate some of the backgrounds and to compute the inclusive cross section.
In addition, the two results quoted for the WZ cross section are not directly comparable, since the CMS public one (Table 6.18) is extrapolated to the $60< M_{ll} <120$ GeV/c^2 region, while the one presented in this Chapter (equation 6.9) is extrapolated to full acceptance. However it's still possible to comment on their performances, comparing the uncertainties on the results.

In [25] both $t\bar{t}$ and $Z + jets$ contributions were measured with a fully data driven technique, by means of the "Matrix-Method".

The event yield obtained channel by channel in both the analyses is compared in Table 6.19, in terms of the number of observed signal events (N_{obs}), the number of background events estimated either from MC or data driven techniques (N_{bkg}) and the WZ event yield from MC.

	Chapter 6.5 results			CMS public analysis [25]		
channel	N_{obs}	N_{bkg}	WZ_{MC}	N_{obs}	N_{bkg}	WZ_{MC}
$eee\nu$	16	2.69	11.95	22	2.98	14.47
$ee\mu\nu$	17	2.85	15.50	20	3.63	17.40
$\mu\mu e\nu$	13	2.12	13.39	13	2.03	13.95
$\mu\mu\mu\nu$	20	3.04	17.34	20	3.15	18.56

Table 6.19: Channel by channel event yield for observed WZ candidates, backgrounds and MC signal events, corresponding to $\int \mathcal{L}\, dt = 1.09 fb^{-1}$. N_{bkg} includes the WZ to $taus$ contribution. The comparison is presented for the WZ analysis search described in this thesis (on the left) and the CMS public one (on the right).

The selection used in this thesis work appears to be slightly tighter, therefore leading to less observed events, particularly for the $Z \rightarrow ee$ channels, but also to smaller background event yield.

To compute the inclusive production cross section, the estimate presented in Chapter 6.5 relies on the Maximum Likelihood method assuming that the number of signal events in each channel (N_i) follows a Poisson distribution of mean $= \sigma \times BR_i(A_i\epsilon_i\mathcal{L})$, while in [25] $\sigma(WZ \rightarrow 3l)$ is computed as weighted mean of the cross sections measured in each leptonic channel. The latter can be expressed as a linear combination of the channel by channel measured σ_i, with weighting factors (w_i), under the constraint $\sum_{i=1}^{4} w_i = 1$, as

$$\sigma(WZ \rightarrow 3l) = w_1 \cdot \sigma_{WZ \rightarrow eee\nu} + w_2 \cdot \sigma_{WZ \rightarrow ee\mu\nu} + w_3 \cdot \sigma_{WZ \rightarrow \mu\mu e\nu} + w_4 \cdot \sigma_{WZ \rightarrow \mu\mu\mu\nu} \quad (6.10)$$

Both results are quoted with errors that include the contribution of the correlated systematic uncertainties, namely through the covariance matrix (Σ).

Σ can be obtained [116] as a function of the error matrix E defined as

$$E = \begin{bmatrix} \sigma_1^2 & \sigma_{12}^{corr}\sigma_{21}^{corr} & \sigma_{13}^{corr}\sigma_{31}^{corr} & \sigma_{14}^{corr}\sigma_{41}^{corr} \\ \sigma_{21}^{corr}\sigma_{12}^{corr} & \sigma_2^2 & \sigma_{23}^{corr}\sigma_{32}^{corr} & \sigma_{24}^{corr}\sigma_{42}^{corr} \\ \sigma_{31}^{corr}\sigma_{13}^{corr} & \sigma_{32}^{corr}\sigma_{23}^{corr} & \sigma_3^2 & \sigma_{34}^{corr}\sigma_{43}^{corr} \\ \sigma_{41}^{corr}\sigma_1^{corr} & \sigma_{42}^{corr}\sigma_{24}^{corr} & \sigma_{43}^{corr}\sigma_{34}^{corr} & \sigma_4^2 \end{bmatrix}$$

where the σ_i^2 account for the statistic and uncorrelated systematic variances for each WZ cross section measured in the channel i, while off-diagonal terms σ_{ij}^2 are the correlated uncertainties for the measurement in the channels i and j.
The estimate of the cross section through the weighted mean depends on the correlated uncertainties.

For a better comparison between the two CMS analyses, the weighted mean approach is also used here, obtaining the following inclusive cross section ($\sigma(WZ)$), computed from the $\sigma(WZ \to 3l\nu)$ by using the respective W and Z boson branching ratios for leptonic decay modes [2]

$$\sigma(WZ) = 18.87 \pm 2.61(\text{stat.}) \pm 1.04(\text{syst.}) \pm 1.13(\text{lumi.})\text{pb} \qquad (6.11)$$

The result with the weighted mean method (6.11) is found comparable with the one obtained using the likelihood (6.9). The systematic error for the measurement here presented (5.7%) is slightly lower than the one quoted in the CMS public result (6.5%) as the combined effect of dealing with less background (see Table 6.19) though taking into account more correlated uncertainties.

A graphical comparison of the inclusive WZ production cross section in pp and $p\bar{p}$ collisions is given in Figure 6.23 for theoretical computations and most updated measurements.

Outlook

At the end of 2011 pp collision operation, CMS recorded order 5 fb^{-1} of integrated luminosity. With such statistic, the precision on the measurement would be reduced by order a factor 2, still making the statistical accuracy the dominant uncertainty, but bringing it at the same order of magnitude of the systematic uncertainty and precision on the theory. During the last months of data taking, also the uncertainty on the luminosity has been reduced to order 4%.
With a factor of 5 more statistic, systematic uncertainties will become dominant. The precision on the efficiencies measured from data would be improved, thus turning into relevant contributions the estimate of the backgrounds and pile-up, which should be controlled with higher accuracy.

With order 5 fb^{-1} of data, also the measurement of the WWZ triple gauge coupling would be possible and already interesting. Indeed, CMS should be able to measure these parameters with a precision comparable to the one provided by the experiments at Tevatron, if not already better.
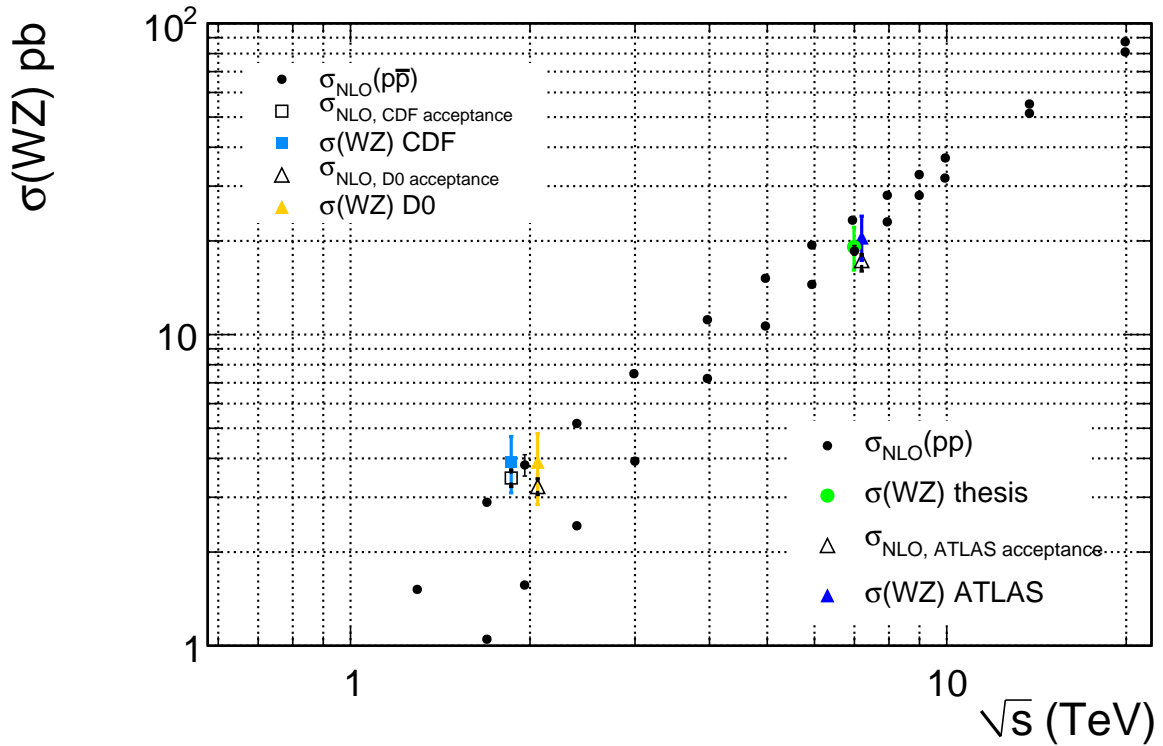
Figure 6.23: Black dots: inclusive WZ production NLO cross section in pp and $p\bar{p}$ collisions from [11]. Measurements from CDF [22], D0 [23], ATLAS [24] already presented in Table 6.18 are shown with colored squares and triangles. The corresponding theoretical NLO cross sections quoted in the CDF, D0 and ATLAS papers are reported for comparison, shown with black empty symbols. The measured values are shown shifted with respect to the x axis, to avoid the overlapping between different results. CDF and D0 measured $\sigma(WZ)$ in $p\bar{p}$ collisions at $\sqrt{s}$=1.96 TeV, CMS and ATLAS in pp collisions at $\sqrt{s}$=7 TeV. The $\sigma(WZ)$ measured with CMS in this thesis is shown with a green circle, to be compared with black dots for $\sigma_{\mathrm{NLO}}(pp)$.

With 2012 data taking, likely at the centre-of-mass energy increased to 8 TeV, the higher WZ production and higher integrated luminosity would make it possible to measure the TGCs providing the most accurate results, to be compared with LEP ones. Tevatron potentialities are indeed already overcome by the LHC and anyway its operation will not continue in the future.

Conclusions

The commissioning activity and the physic studies I performed during the first three years of the CMS data taking were presented in this thesis.

I had the possibility to contribute to detector oriented activities, following the evolution of the LHC operation and the CMS data taking. I studied in detail the electron reconstruction algorithm and carefully validated its performances with first collision data. From the physics analysis point of view, I could profit from the first 1/fb of data from pp collisions at 7 TeV to study the WZ electroweak production, in the direction to further contribute in the understanding of the Standard Model of elementary particles.

In particular, during the data taking exercise in Fall 2008, cosmic ray events were collected by the CMS detector to test in situ the CMS magnet at the nominal current and to commission the experiment for an extended operation period.
Within this context, the muon stopping power in the lead tungstate ($PbWO_4$) of the electromagnetic calorimeter was measured. The outcome of the measurement, which proved consistent with the expectations over the entire energy range, resulted of particular interest for calibration purposes and brought the first experimental measurement of the muon critical energy in $PbWO_4$. With this result, the ECAL energy scale, set with 120 GeV/c electrons at the test-beam is validated in situ at the level of few percents, using energy deposits of order 100 MeV associated with low-momentum muons.

Moreover, with the LHC entering the high energy collision era in 2009, it was important to commission the physics objects, to provide solid elements fundamental to carry out reliable physics analyses. A detailed comparison between data and simulation allows to validate the detector response, at the same time it's important to control the reconstruction algorithms and monitor the performances of physics objects with data.

In this context the commissioning of the ecal-driven electrons as reconstructed in the CMS detector was addressed, with particular attention to the very first stage of the reconstruction process which is the finding of the track seed in the inner tracker layers. This study resulted useful, for instance to account for real condition effects not implemented in the simulation. The electron seed reconstruction was fully validated after few months of data taking, with order 14 pb^{-1} of data, profiting from an electron sample for physics studies from W decays.

Finally, the last part of the thesis is dedicated to the study of the WZ diboson production, in the decay channel $l\nu l\bar{l}$, with $l = e, \mu$. This electroweak process, was observed for the first time in $p\bar{p}$ collisions at the Tevatron accelerator with 1 fb^{-1} of integrated luminosity. Its inclusive production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV is of the order of 18 pb(NLO). The first event was observed in CMS with order 30pb^{-1} of integrated luminosity and the 1/fb registered by CMS up to July 2011, allowed for the first measurement of the WZ production in the decay channel chosen.

Diboson physics provides important tests of the Standard Model, through precision measurements such as the production cross section itself and the trilinear gauge coupling. In this particular context, the WZ is the only process to allow for the measurement of the WWZ vertex alone without any other amplitude contribution. At the same time anomalous couplings in the WWZ production could be searched for in this channel, being also a benchmark for some BSM scenarios.

For this study, the event selection was defined so to provide a robust signal extraction, relying on a cut based selection and focusing on the leptonic signature which provides a clean signal, a key element for analysis at hadron colliders. The analysis here presented leads to the measurement of the inclusive WZ production cross section, extrapolated to the full acceptance. The result obtained profits from 1.09 fb^{-1} of data from pp collisions at 7 TeV and is in agreement with the expectation from the SM.

This measurement represents an important step towards the study of diboson physics and leads the way towards the multiple interesting multiboson measurements that lie ahead.

Already with 2011 operations, the LHC potentialities in the field of electroweak measurements overtook Tevatron ones and with 2012 the full investigation of high energy physics will be through the pp LHC data. The high statistic at high center-of-mass energy (likely 8 TeV in 2012) will allow for high precision electroweak measurements able to further constraint the Standard Model.

Appendix

A.1 Local gauge invariance

To take into account interacting theories, local gauge symmetries have to be considered. In the following, the explicit example for the a $U(1)$ symmetry group is presented.

Let's consider the local $U(1)$ symmetry and the Lagrangian for a free fermion of mass m

$$\psi \to e^{i\alpha(x)}\psi \qquad \text{and} \qquad \mathcal{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi \tag{A.1}$$

if the derivative ∂_μ is replaced by the so-called *covariant derivative*

$$\partial_\mu \to \mathcal{D}_\mu \equiv \partial_\mu - ieA_\mu(x) \tag{A.2}$$

where A_μ is a new vector gauge field, transforming as $A_\mu(x) \to A_\mu(x) + \dfrac{1}{e}\partial_\mu\alpha(x)$, the object $\mathcal{D}_\mu\psi$ behaves indeed as the field ψ, $\mathcal{D}_\mu\psi \to e^{i\alpha(x)}\mathcal{D}_\mu\psi$, under the $U(1)$ phase rotation.
The resulting Lagrangian preserves invariance under the local gauge transformation

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu\psi A_\mu \tag{A.3}$$

As a result, the interaction term between the Dirac spinor ψ and the vector gauge field A_μ emerges spontaneously preserving the gauge invariance of the theory.
To complete the Lagrangian with a kinetic term for the A_μ field, the tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ can be introduced.

It must be stressed that A_μ is a massless field, since the addition of a mass term $\dfrac{1}{2}m^2 A_\mu A^\mu$ is prohibited by gauge invariance.

A.2 Spontaneously Broken Symmetry

The appearance of a massless particles in the theory, for every continuous symmetry spontaneously broken, is a general result known as *Goldstone's theorem.* The number of new massless particles (Goldstone bosons) in the theory is related to the degrees of freedom of the symmetry group and it corresponds to the number of broken symmetries.

For a rotation in N dimensions, described by $N(N-1)/2$ parameters, each of them corresponding to a continuous symmetry, after a spontaneous breakdown of the $O(N)$ symmetry to an $O(N-1)$, there are still $(N-1)(N-2)/2$ unbroken symmetries and the number of massless bosons is then $N(N-1)/2 - (N-1)(N-2)/2 = N - 1$.

As an example, the case of a complex scalar theory invariant under a local $U(1)$ symmetry is here considered. It can be represented through either a doublet of complex fields $(\phi, \phi^\dagger)$ or two real fields (ϕ_1, ϕ_2). The corresponding gauge invariant Lagrangian is

$$
\begin{aligned}
\mathcal{L}_{start} &= (\mathcal{D}_\mu \phi)^\dagger \mathcal{D}^\mu \phi - V(\phi^\dagger \phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \\
&= \frac{1}{2}(\mathcal{D}_\mu \phi_1)^2 + \frac{1}{2}(\mathcal{D}^\mu \phi_2)^2 - \frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) - \lambda(\phi_1^2 + \phi_2^2)^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}
\end{aligned}
\tag{A.4}
$$

where the effective potential $V(\phi^\dagger \phi)$ is chosen in the particular form

$$
V(\phi^\dagger \phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2, \qquad \lambda > 0
\tag{A.5}
$$

Depending on the sign of μ^2, two different configurations are realized, as sketched in figure A.1.

If $\mu^2 > 0$, the symmetry is exact and there exists a unique vacuum state for the theory, at $< \phi >= 0$. On the other hand, if $\mu^2 < 0$, which means that μ can not longer be interpreted as a mass for the field ϕ, the vacuum state is infinitely degenerate for all the configurations satisfying

$$
|\phi| = vev \equiv \sqrt{-\frac{\mu^2}{2\lambda}}
\tag{A.6}
$$

The $U(1)$ symmetry is spontaneously broken, if a particular solution is chosen, as an example the ground state

$$
\phi_1(x) = v + \eta(x), \quad \phi_2(x) = \xi(x)
\tag{A.7}
$$

$$
\phi = \sqrt{\frac{1}{2}}(v + \eta(x) + i\xi(x)), \quad \phi^\dagger = \sqrt{\frac{1}{2}}(v + \eta(x) - i\xi(x))
\tag{A.8}
$$

By expanding the Lagrangian in A.4 up to some interaction terms that are omitted

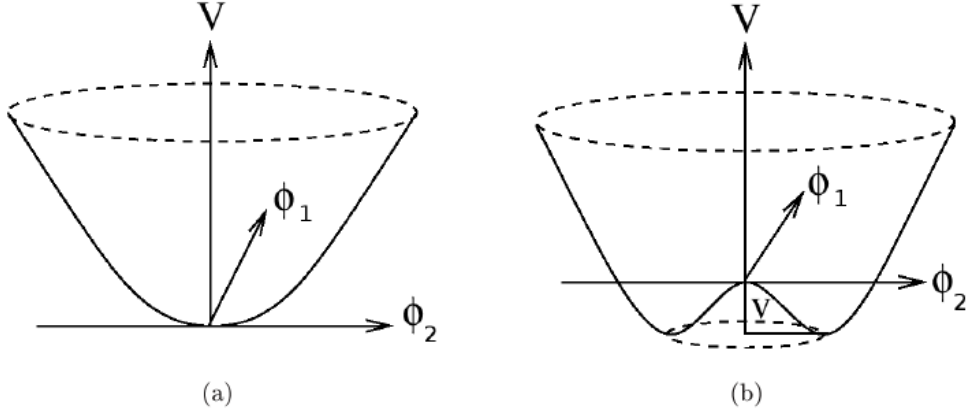Figure A.1: The potential $V(\phi) = \mu^2\phi^2 + \lambda\phi^4$ for $\mu^2 > 0$ (a) and for $\mu^2 < 0$ (b). The parameter λ is assumed to be positive

for simplicity, the broken Lagrangian on the chosen *vev* presents interesting terms

$$\mathcal{L}_{broken} = \frac{1}{2}(\partial_\mu\xi)^2 + \frac{1}{2}(\partial^\mu\eta)^2 - v^2\lambda\eta^2 + \frac{1}{2}e^2v^2A_\mu A^\mu - evA_\mu\partial^\mu\xi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \qquad (A.9)$$

The third term in $(-v^2\lambda\eta^2)$ has the form of a mass for the η field, with $m_\eta = \sqrt{2\lambda v^2}$. The first term represents the kinetic energy of the ξ field, to which no mass term corresponds. The theory essentially contains also a massless scalar field, which is known as a Goldstone boson.

It is possible to get the flavor of this effect, by looking at the shapes of the potential in Figure A.1. The mass term for η is a consequence of the restoring force against radial oscillations, while the symmetry under $U(1)$ rotations that the Lagrangian still exhibits means that there is no resistance to excitations along the ξ direction. The potential in the tangent ξ direction is flat, implying a massless mode.

The mass for the η field has been "revealed", by expanding the Lagrangian around the stable vacuum state *vev*.

Experimental evidence needs to be considered with respect the presented formalism, the existence of (massive) weak gauge bosons for instance and that the massless scalar particles, as predicted by the spontaneous global symmetry breaking, have never been detected.

It was Higgs' insight [117] that when a gauge theory is combined with a spontaneous local symmetry breaking model, the two problems solve themselves rather elegantly. By looking at the example just shown (equation A.4), already interesting hints not expected phenomenologically emerge, as the off-diagonal structure $A_\mu\partial^\mu\xi$ and the further massive

vector term $A_\mu A^\mu$.

A.2.1 The Higgs Boson

Higgs had found that when scalar fields are coupled to the gauge theory, the Goldstone bosons are eaten by the gauge bosons. These latter become massive with their longitudinal polarization being provided by the absorbed Goldstone boson.

By referring to the previous example, in the beginning (equation A.4) the theory had four degrees of freedom, two for the complex scalar field ϕ and two for the massless electromagnetic field A_μ.

After the spontaneous breaking of the local gauge symmetry, it apparently shows five degrees of freedom (equation A.9), one for ϕ_1, one for ϕ_2 and three for the massive photon A_μ. Therefore, there must be a field which is not physical and indeed, in $\mathcal{L}_{broken}$ there is the bilinear term $A_\mu \partial^\mu \xi$ which is not welcome.

A clue is to notice that at first order, the field ϕ in A.7 becomes

$$\phi = \frac{1}{\sqrt{2}}(v + \eta + i\xi) \simeq \frac{1}{\sqrt{2}}(v + \eta)e^{i\xi/v} \tag{A.10}$$

which suggests to substitute in the original Lagrangian (equation A.4) a different set of real fields h, θ, A_μ, where

$$\phi \to \frac{1}{\sqrt{2}}(v + h(x))e^{i\theta(x)/x}, \qquad A_\mu \to A_\mu + \frac{1}{ev}\partial_\mu \theta \tag{A.11}$$

With this particular choice of gauge, with $\theta(x)$ chosen so that h is real, only the physical particles are left in the Lagrangian

$$\mathcal{L}'_{broken} = \frac{1}{2}(\partial_\mu h)^2 - \lambda v^2 h^2 + \frac{1}{2}e^2 v^2 A_\mu^2 + \frac{1}{2}e^2 A_\mu^2 h^2 + kinetic\,term + ... \tag{A.12}$$

The Goldstone boson actually does not appear in the theory. That is, the apparent extra degree of freedom is spurious and it corresponds to the freedom to make a gauge transformation.

The A_μ field (with two degrees of freedom) absorbed the would be Goldstone boson (with one degree of freedom) and became massive (i.e. with three degrees of freedom).

The massive scalar particle h is called the Higgs boson.

Appendix

B.1   First WZ event observed in CMS

The event selection described in Chapter 6.2, although not fully optimized by November 2010, allowed to identify the first WZ candidate event, already in the 36^{-1}pb of integrated luminosity recorded by CMS in 2010.
It was observed in the $e^+e^-\mu^+$ decay channel.

The kinematic features of this event are listed in the following Tables, with details on leptons in Tab. B.1, details on bosons in Tab. B.2 and on the whole event listed in Tab B.3. Few event displays are also shown.

Leptons kinematic	e^+	e^-	μ^+
energy (GeV)	54.94	104.40	144.53
p_T (GeV/c)	54.78	104.35	52.07
η	-0.076	0.033	1.680
ϕ	2.318	-2.671	0.401
$d_z(PV)$ (mm)	-0.0065	0.091	0.0456
$d_{xy}(PV)$ (mm)	0.0059	0.0281	0.0099
E/p	0.981	1.111	-
matches	-	-	2

Table B.1

Boson candidates	Z	W
Invariant mass (GeV/c^2)	91.52	-
Invariant transverse mass ($pfMT$) (GeV/c^2)	-	60.98

Table B.2

Missing Transverse Energy	pf	tc	$calo$
MET (GeV)	101.25	98.79	90.46
ϕ	-0.4657	-0.4641	-0.3409
Global event features			
$p_T(ee\mu)$ (GeV/c)	83.37		
$\Delta\phi(Z,\mu)$ (rad.)	2.79		
$\Delta\phi(pfMET,\mu)$ (rad.)	0.867		
$pfJet$ with $p_T >$10 GeV/c	1 reconstructed		
p_T (GeV/c)	18.1		
η	-1.326		
ϕ (rad.)	1.737		

Table B.3
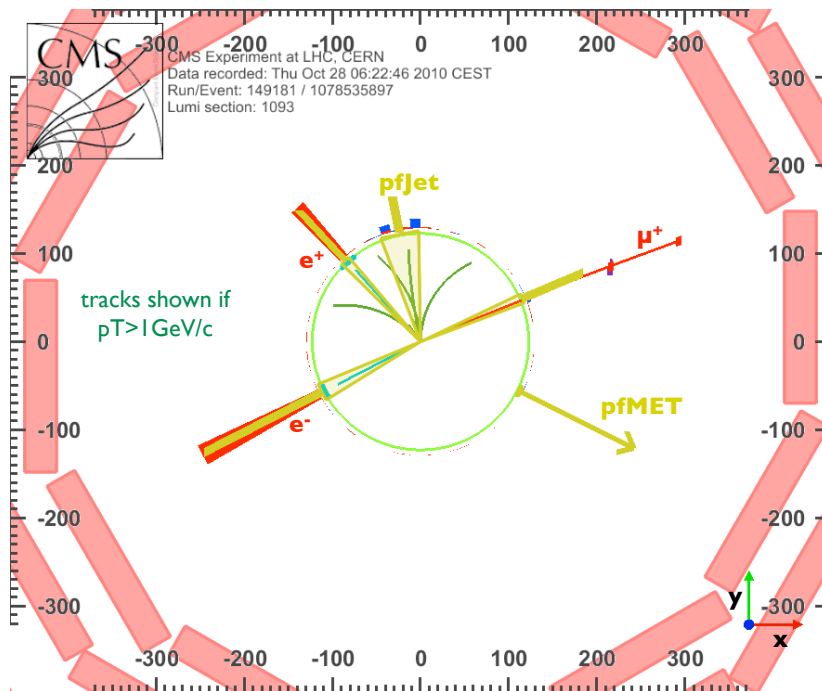
Figure B.1
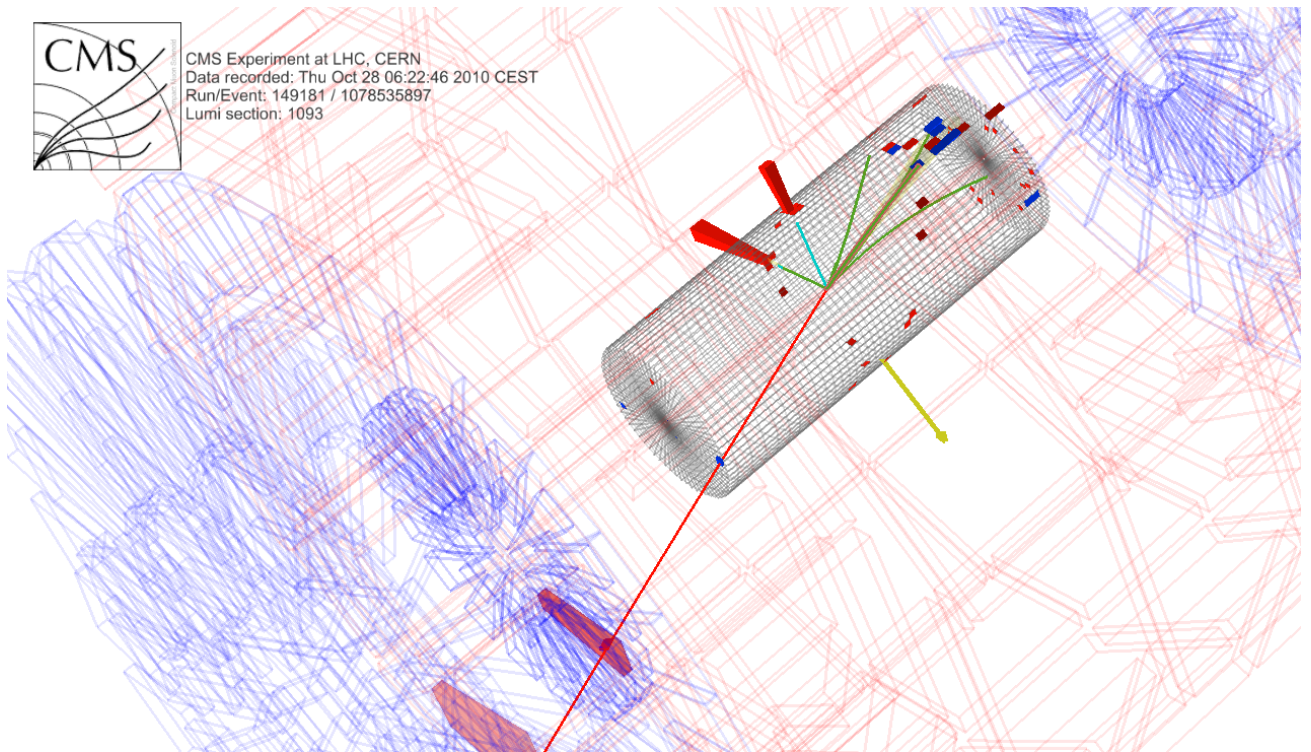
Figure B.2
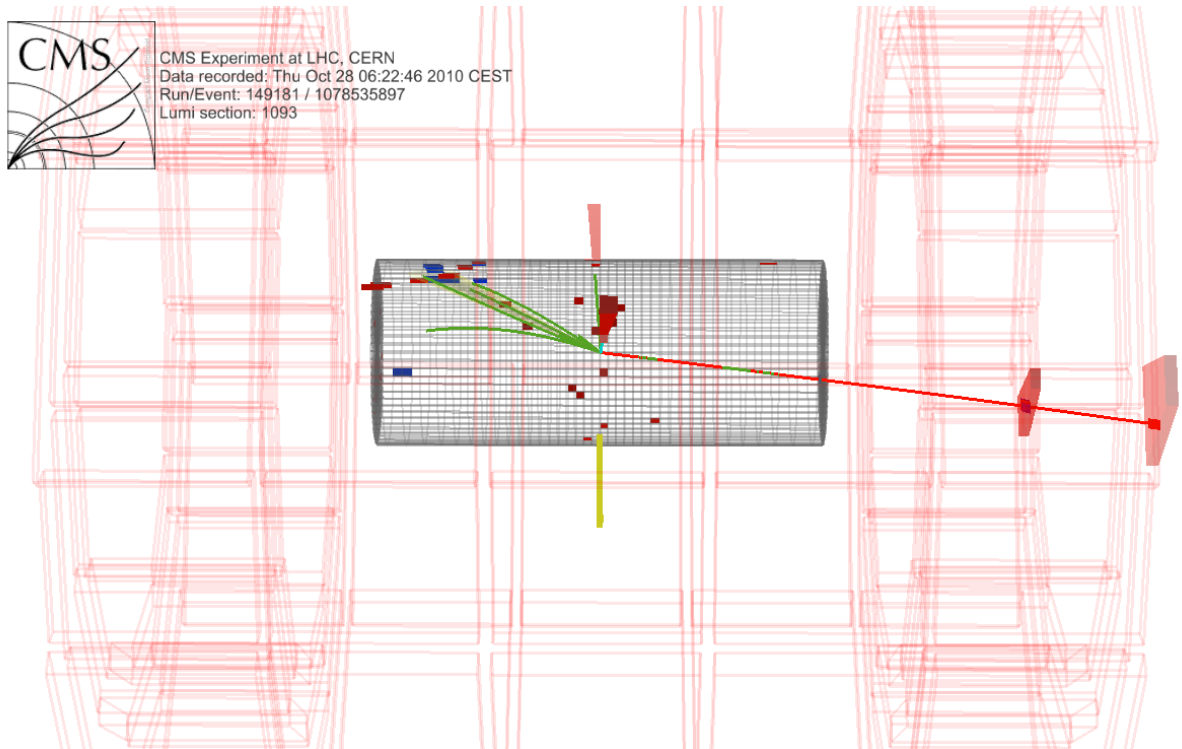
Figure B.3

Bibliography

[1] ALEPH, CDF, D0 et al. Precision Electroweak Measurements and Constraints on the Standard Model. `http://arxiv.org/abs/0811.4682`, 2008.

[2] K. Nakamura et al. (Particle Data Group). Review of particle physics. *J. Phys.*, G37:075021, 2010.

[3] F.J. Yndurain. *The Theory of Quarks and Gluon Interactions*. Springer-Verlag, 1993.

[4] S. Weinberg. Phys.Rev.Lett. (19):1264, 1967.

[5] S.L. Glashow. Nucl.Phys. (B-22):579, 1961.

[6] A. Salam, J.C. Ward. Phys. Lett. 168(13), 1964.

[7] Francis Halzen, Alan D. Martin. *QUARKS AND LEPTONS: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc., 1984.

[8] M. Peskin and D. Schroeder. *An Introduction to Quantum Field Theory*. Westview, 1995.

[9] G. Ridolfi. An Introduction to the Standard Model of ElectroWeak Interactions. `http://www.ge.infn.it/~ridolfi/notes/smcom.ps`.

[10] CDF Collaboration. Observation of the WZ Production. `http://arxiv.org/abs/hep-ex/0702027`.

[11] John M. Campbell *et al.* Vector boson pair production at the LHC. `http://dx.doi.org/10.1007/JHEP07(2011)018`, 2011. 10.1007/JHEP07(2011)018.

[12] P. Langacker. *The standard model and beyond*. Taylor and Francis, Boca Raton, FL, 2010.

[13] H. Baer *et al.* Trileptons from chargino-neutralino production at the CERN LHC. `http://arxiv.org/pdf/hep-ph/9404212v1`, 1994.

[14] CMS Collaboration. Search for W' (or techni-rho) to WZ. *CMS PAS*, EXO-11-041, 2011.

[15] K. Agashe, R. Contino, A. Pomarol. The Minimal Composite Higgs Model. 2005.

[16] M. Perelstein. Little Higgs Models and Their Phenomenology. `http://arxiv.org/pdf/hep-ph/0512128v1`, 2008.

[17] C. Csáki, C. Grojean, L. Pilo, J. Terning. Towards a Realistic Model of Higgsless Electroweak Symmetry Breaking. *Phys. Rev. Lett.*, 92:101802, Mar 2004.

[18] Adamson, K. L. and de Florian, D. and Signer, A. Gluon induced contributions to WZ and $W\gamma$ production at NNLO. `http://link.aps.org/doi/10.1103/PhysRevD.65.094041`, May 2002.

[19] T. Melia, *et al.* W^+W, WZ and ZZ production in the POWHEG BOX. `http://xxx.lanl.gov/`

abs/1107.5051v3, 2011.

[20] G. Altarelli, R. Kleiss, C. Verzegnassi. *Z physics at LEP 1.* CERN, European Organization for Nuclear Research in Geneva, 1989.

[21] K. Hagiwara, R.d. Peccei, D. Zeppenfeld. Probing the weak boson sector in $e^+e^- \rightarrow W^+W^-$. *Nuclear Physics*, B282(253-307), 1987.

[22] CDF Collaboration. Measured WZ Diboson Cross Section with 7.1/fb at CDF. `http://www-cdf.fnal.gov/physics/ewk/2011/wz/PublicPages/WZwebpage.html`.

[23] D0 Collaboration. Measurement of the $WZ \rightarrow l\nu ll$ cross section and limits on anomalous triple gauge couplings in $p\bar{p}$ collisions at $\sqrt{s}$ =7 TeV. `http://www.sciencedirect.com/science/article/pii/S037026931001258X`, 2011.

[24] ATLAS Collaboration. Measurement of the WZ production cross section and limits on anomalous triple gauge couplings in proton-proton collisions at $\sqrt{s}$ =7 TeV with the ATLAS detector. `http://arxiv.org/pdf/1111.5570`, 2011.

[25] CMS Collaboration. Measurement of the WW, WZ and ZZ cross sections at CMS. `http://cdsweb.cern.ch/record/1370067`, 2011.

[26] R. Brunelière. Triple and quartic gauge couplings at LEP 2. *The European Physical Journal C - Particles and Fields*, 33:s709–s713, 2004. 10.1140/epjcd/s2004-03-1625-3.

[27] LHC collaboration. LHC Machine. `http://iopscience.iop.org/1748-0221/3/08/S08001`, 2008.

[28] ATLAS collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. `http://iopscience.iop.org/1748-0221/3/08/S08003`, 2008.

[29] CMS collaboration. The CMS experiment at the CERN LHC. `http://iopscience.iop.org/1748-0221/3/08/S08004`, 2008.

[30] LHCb collaboration. The LHCb Detector at the LHC. `http://iopscience.iop.org/1748-0221/3/08/S08005`, 2008.

[31] ALICE collaboration. The ALICE experiment at the CERN LHC. `http://iopscience.iop.org/1748-0221/3/08/S08002`, 2008.

[32] Gianotti, F. Collider physics: LHC. `http://cdsweb.cern.ch/record/458489/files/p219.pdf`, 2000.

[33] ATLAS Collaboration. Measurement of the Inelastic Proton-Proton Cross-Section at $\sqrt{s}$=7 TeV with the ATLAS Detector. `http://arxiv.org/pdf/1104.0326v1`, 2011.

[34] CMS Collaboration. Measurement of the inelastic pp cross section at $\sqrt{s}$=7 TeV with the CMS detector. *CMS PAS*, FWD-11-001, 2011.

[35] J. Stirling. `http://projects.hepforge.org/mstwpdf/plots/plots.html`.

[36] Karimki, V. The CMS tracker system project: Technical Design Report. `http://cdsweb.cern.ch/record/368412`, 1997.

[37] CMS Collaboration. The CMS tracker: addendum to the Technical Design Report. `http://cdsweb.cern.ch/record/490194`, 2000.

[38] CMS Collaboration. CMS ECAL Technical Design Report. `http://cms-ecal.web.cern.ch/cms-ecal/ECAL_TDR`, Dec-15 1997.

[39] CMS Collaboration. CMS Physics TDR Volume 1: Detector Performance and Software. `http://cdsweb.cern.ch/record/922757`, 2006.

[40] CMS Collaboration. Data filtering in the readout of the CMS Electromagnetic Calorimeter. `http:`

//iopscience.iop.org/1748-0221/3/02/P02011, 2008.

[41] R. Alemany *et al.* Overview of the ecal off-detector electronics of the cms experiment. *Nuclear Science Symposium Conference Record, 2004 IEEE*, 2:1053–1057, 2004.

[42] N. Almeida *et al.* The selective read-out processor for the cms electromagnetic calorimeter. *Nuclear Science Symposium Conference Record, 2004 IEEE*, 3:1721–1725, 2004.

[43] CMS Collaboration. ECAL 2010 performance results. *CMS DP*, DP-11-008, 2011.

[44] CMS Collaboration. Electromagnetic calorimeter calibration with 7 TeV data. *CMS PAS*, EGM-10-003, 2010.

[45] CMS Collaboration, editor. *Commissioning, Performance and Calibration of Crystals of the CMS Electromagnetic Calorimeter*, Paris, 2010. $35^{t}h$ International Conference on High Energy Physics (ICHEP2010).

[46] CMS Collaboration. *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[47] Dasu, S. and others. CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems. http://cdsweb.cern.ch/record/706847, 2000.

[48] Cittolin, Sergio and Rcz, Attila and Sphicas, Paris. CMS trigger and data-acquisition project. Technical design report, vol. 2: Data-acquisition and high-level trigger. http://cdsweb.cern.ch/record/578006, 2002.

[49] Adam, W. and others. The CMS high level trigger. http://arxiv.org/pdf/hep-ex/0512077v1, 2006.

[50] CMS Collaboration. Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV. *CMS PAS*, MUO-10-002, 2010.

[51] R. and Fruhwirth. Application of kalman filtering to track and vertex fitting. http://www.sciencedirect.com/science/article/pii/0168900287908874, 1987.

[52] CMS Collaboration. Measurement of Tracking Efficiency. *CMS PAS*, TRK-10-002, 2010.

[53] C. Seez *et al.* Addendum to ECAL calibration and the definition of calibrated RecHits:details, B-field, and the cluster correction scheme. 1.1, November 2007.

[54] W. Wdam *et al.* Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC. 31, 2005.

[55] K. Gopala, B. Rudraswamy, P. Venkataramaiah, and H. Sanjeeviah. Thick-target bremsstrahlung spectra generated by the β particles of $^{90}-^{90}$y and ^{99}Tc. http://link.aps.org/doi/10.1103/PhysRevA.34.5126, Dec 1986.

[56] C. Charlot. La reconstruction des electrons dans CMS. http://cmsfrance.in2p3.fr/HabilitationCCharlot.pdf, janvier 2007.

[57] F. Beaudette *et al.* Electron Reconstruction within the Particle Flow Algorithm. *CMS AN*, AN-10/034, 2010.

[58] S. Baffioni *et al.* Electron Identification in CMS. *CMS AN*, 2009/178, 2009.

[59] D. Barge *et al.* Study of photon conversion rejection at CMS. *CMS AN*, 2009/159, 2009.

[60] https://twiki.cern.ch/twiki/bin/view/CMS/EWKVBTF.

[61] G.P.Salam *et al.* A practical seedless infrared-safe cone jet algorithm. *JHEP*, 05:086, 2007.

[62] S.Catani *et al.* Longitudinally invariant K_t clustering algorithms for hadron hadron collisions. *Nucl. Phys. B*, 406:187–224, 1993.

[63] S.D. Ellis *et al.* Successive Combination Jet Algorithm For Hadron Collisions. *Phys. Rev. D*, 48:3160, 1993.

[64] Y.L. Dokshitzer *et al.* Better jet clustering algorithms. *JHEP*, page 9708:001, 1997.

[65] M. Wobisch *et al.* Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering. 1999.

[66] M. Cacciari *et al.* The anti-k_t jet clustering algorithm. *JHEP*, page 0804:063, 2008.

[67] R. Demina *et al.* Calorimeter Energy Thresholds for Jet Reconstruction in CMS. *CMS AN*, 2006/020, 2006.

[68] CMS Collaboration. Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS. *CMS PAS*, JME-09-002, 2009.

[69] CMS Collaboration. CMS MET Performance in Events Containing Electroweak Bosons from pp Collisions at $\sqrt{s}$ =7 TeV. *CMS-PAS*, JME-10-005, 2010.

[70] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Peformance for Jets, E_T^{miss}, and Taus. *CMS PAS*, PFT-09-001, 2009.

[71] C. Eck *et al.* LHC computing Grid: Technical Design Report. `http://lcg.web.cern.ch/lcg/TDR/`, 20 June 2005.

[72] H. Sakulin. DAQ Operations and Plans. 2-3 November 2010. Run Coordination Workshop.

[73] J. Gutleber, S. Murray, and L. Orsini. Towards a homogeneous architecture for high-energy physics data acquisition systems. *Comput. Phys. Commun.*, 153:155–163, 2003.

[74] G. Bauer *et al.* The run control and monitoring system of the CMS experiment. *Journal of Physics: Conference Series*, 119, 2008.

[75] P.C. Burkimsher. JCOP Experience with a Commercial SCADA Product, PVSS. In *ICALEPCS 2003 Contributions to the Proceedings: 9th International Conference on Accelerator and Large Experiment Physics Control Systems*, Gyeongju, Korea, 2003.

[76] `http://root.cern.ch/drupal/`.

[77] `https://cmsdoc.cern.ch/cms/cpt/Software/html/General`.

[78] S. Agostinelli *et al.* Geant4 a simulation toolkit. `http://www.sciencedirect.com/science/article/pii/S0168900203013688`, 2003.

[79] CMS Collaboration. Commissioning of the CMS experiment and the cosmic run at four tesla. `http://stacks.iop.org/1748-0221/5/i=03/a=T03001`, 2010.

[80] CMS Collaboration. Performance and operation of the CMS electromagnetic calorimeter. `http://iopscience.iop.org/1748-0221/5/03/T03010`, 2010.

[81] CMS Collaboration. CMS data processing workflows during an extended cosmic ray run. `http://stacks.iop.org/1748-0221/5/i=03/a=T03006`, 2010.

[82] CMS Collaboration. Commissioning and performance of the CMS silicon strip tracker with cosmic ray muons. `http://stacks.iop.org/1748-0221/5/i=03/a=T03008`, 2010.

[83] CMS collaboration. Stand-alone Cosmic Muon Reconstruction Before Installation of the CMS Silicon Strip Tracker. `http://stacks.iop.org/1748-0221/4/i=05/a=P05004`, 2009.

[84] P. Adzic *et al.* Intercalibration of the barrel electromagnetic calorimeter of the CMS experiment at start-up. `http://iopscience.iop.org/1748-0221/3/10/P10007`, 2008.

[85] Donald E. Groom *et al.* Muon Stopping Power and range tables 10 MeV to 100 TeV. `http://www.sciencedirect.com/science/article/pii/S0092640X01908617`, 2001.

[86] B. Rossi. High-energy cosmic rays, 1959.

[87] CMS Collaboration. Measurement of the muon stopping power in lead tungstate. `http://iopscience.iop.org/1748-0221/5/03/P03007`, 2010.

[88] F. Cavallari *et al.* Linearity of the CMS Electromagnetic Calorimeter from H2 test beam data. *CMS DN*, 2007/019, 2007.

[89] R. Arcidiacono. Studies of the CMS Electromagnetic Calorimeter performance in the electron test beam. *CMS CR*, 2008/099, 2008.

[90] W. Adam *et al.* Electron Reconstruction in CMS. *CMS AN*, AN-09/164, 2009.

[91] A. Martelli. Optimisation of electron reconstruction in the CMS experiment to improve the discovery of the Higgs boson in the H→ZZ* → 4e channel, July 2008.

[92] T. Miao *et al.* Beam Position Determination using Tracks. *CMS AN*, 2007/021, 2007.

[93] Michael B. Anderson *et al.* Electron and Photon High-Level Trigger in CMS. *CMS AN*, AN-10/066, 2010.

[94] CMS Collaboration. Electromagnetic physics objects commissioning with first LHC data. *CMS PAS*, EGM-10-001, 2010.

[95] CMS Collaboration. Electron Reconstruction and Identification at $\sqrt{s} = 7\,TeV$. *CMS PAS*, EGM-10-004, 2010.

[96] The CMS Collaboration. Electromagnetic calorimeter commissioning and first results with 7 TeV data. *CMS AN*, 2010/2012, 2010.

[97] A. Massironi *et al.* ECAL alignment with electrons. *CMS AN*, 2011/064, 2011.

[98] CMS Collaboration. Tracking and Vertexing Results from First Collisions. *CMS AN*, 2010/001, 2010.

[99] `http://cmsdoc.cern.ch/cms/Physics/egamma/www/validation/381/Electrons/vs380new/RelValSingleElectronPt35_Mc/index.html`.

[100] T. Sjostrand et al. PYTHIA 6.4 physics and manual. *JHEP*, 2006(05):026, 2006.

[101] J. Alwall *et al.* MadGraph/MadEvent v4: The New Web Generation. `http://iopscience.iop.org/1126-6708/2007/09/028`, 2007.

[102] `http://powhegbox.mib.infn.it`.

[103] N. Adam *et al.* Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{(s)} = 7\,TeV$. *CMS AN*, 2010/116, 2010.

[104] M. Cacciari and G. P. Salam. Pileup subtraction using jet areas. *Phys. Lett. B*, 659:119126, (2008).

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

[106] N. Amapane *et al.* Search for a Standard Model Higgs boson produced in the decay channel H → ZZ$^{(*)}$ → 4l. *CMS AN*, 2011/123, 2011.

[107] A. Adiguzel *et al.* Measurement of the WZ cross section at the CMS Experiment. *CMS AN*, 2011/259, 2011.

[108] CMS Collaboration. Algorithms for b Jet identification in CMS. *CMS PAS*, BTV-09-001, 2010.

[109] G. Daskalakis *et al.* Measuring Electron Efficiencies at CMS with Early Data. *CMS AN*, 2007/019, 2007.

[110] S. Nourbakhsh *et al.* Energy scale of CMS Electromagnetic Calorimeter with 2010 Data. *CMS AN*, 2011/038, 2011.

[111] CMS Collaboration. Performance of muon identification in pp collisions at sqrt(s) = 7 TeV. *CMS PAS*, MUO-10-002, 2010.

[112] CMS Collaboration. Measurement of momentum scale and resolution of the CMS detector using low-mass resonances and cosmic ray muons. *CMS PAS*, TRK-10-004, 2010.

[113] G. Bauer *et al.* Modeling of $W \to l\nu \, MET$ with Boson Recoil. *CMS AN*, 2010/332, 2010.

[114] Campbell, John M. and Huston, J. W. and Stirling, W. J. Hard Interactions of Quarks and Gluons: A Primer for LHC Physics. *Rept. Prog. Phys.*, 70:89, 2007.

[115] Campbell, John M. and Ellis, R. Keith. An update on vector boson pair production at hadron colliders. *Phys. Rev.*, D60:113006, 1999.

[116] Lyons *et al.* How to combine correlated estimates of a single physical quantity. *Nucl. Instrum. Meth.*, A270:110, 1988.

[117] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

Aknowledgements

ABSTRACT

In this thesis work, the first measurement of the WZ cross section with the CMS detector at the LHC using the leptonic modes is presented. The response of the electromagnetic calorimeter (ECAL) is studied through the measurement of the muon stopping power in the lead tungstate of the ECAL. Electron objects are commissioned with the first data so to be used for any analyses.

With the cosmic ray data, the muon stopping power in the lead tungstate of the electromagnetic calorimeter was measured over the momentum range 5-1000 GeV/c and found consistent with expectations. Such a result allowed to extend the validity of the ECAL energy scale, previously set with 120 GeV/c electron beam, down to the sub-GeV region consistently with $1.004^{+0.002}_{-0.003}$(stat.)±0.019(syst.).

With the LHC starting operation, commissioning studies were dedicated to the electron reconstruction, with particular attention to the very first stage of the tracking in the inner tracker layers. The algorithms were accurately optimized, with an electron sample from W boson decay and fully validated with order 14 pb^{-1} of data.

The physics measurement is the cross section of the WZ associated production in pp collisions at $\sqrt{s} =7$TeV. The clean leptonic signature allowed for an efficient signal extraction in each considered channel and a good background rejection. The first event was observed with 36pb^{-1} of data. With 1.09fb^{-1}, the WZ cross section was measured for the first time at $\sqrt{s} =7$TeV $\sigma(WZ) = 19.11^{+3.30}_{-2.53}$(stat.)$\pm1.10$(syst.)$\pm1.15$(lumi.)pb and found consistent with the SM expectation $(18.57^{+0.75}_{-0.58}$pb NLO).

RÉSUMÉ

Cette thèse présente la première mesure de production du WZ avec le détecteur CMS au LHC en utilisant les modes leptoniques. La réponse du calorimètre électromagnétique (ECAL) est étudiée par la mesure du pouvoir d'arrêt de muons dans le tungstate de plomb d'ECAL. Les électrons sont validés avec les premières données pour être utilisable dans toutes les analyses.

Avec les données de rayons cosmiques, le pouvoir d'arrêt de muons traversants le tungstate de plomb du calorimètre électromagnétique a été mesuré pour une gamme d'impulsion 5-1000 GeV/c. Le résultat compatible avec l'attendu a permis de valider l'échelle d'énergie d'ECAL, déterminée auparavant avec un faisceau d'électrons de 120 GeV/c, dans la région du sub-GeV en accord avec $1.004^{+0.002}_{-0.003}$(stat.)±0.019(syst.).

Les données des premières collisions de LHC ont permis la vérification des algorithmes de reconstruction des électrons, en particulier pour la détermination des pré-traces dans la partie la plus interne du trajectographe. Les algorithmes ont été optimisés, sur des électrons issus de désintégrations de bosons W et entièrement validés avec 14 pb^{-1} de données environ.

La mesure physique est la section efficace de production des bosons associés WZ dans les collisions proton-proton à $\sqrt{s} =7$TeV. La signature claire de la désintégration en leptons permet l'extraction efficace du signal et la réjection du bruit de fond, pour chaque canal considéré. Le premier évènement a été observé avec 36pb^{-1} de données. Avec 1.09fb^{-1}, la section de production de WZ a été mesurée pour la première fois à $\sqrt{s} =7$TeV $\sigma(WZ) = 19.11^{+3.30}_{-2.53}$(stat.)$\pm1.10$(syst.)$\pm1.15$(lumi.)pb et trouvée en accord avec la prédiction du Model Standard $(18.57^{+0.75}_{-0.58}$pb NLO).