Sapienza Università di Roma
Dottorato di Ricerca in Fisica
Scuola di dottorato "Vito Volterra"

Studies on the dielectron spectrum with the first data of the CMS experiment at the Large Hadron Collider

Thesis submitted to obtain the degree of
"Dottore di Ricerca" - PhilosophiæDoctor
PhD in Physics - XXII cycle - October 2009

by

Alessandro Palma

Program Coordinator
Prof. Enzo Marinari

Thesis Advisors
Prof. Egidio Longo
Dr. Riccardo Paramatti
Dr. Paolo Meridiani

Be Water, My Friend

Bruce Lee

Contents

List of Figures

List of Tables

Acknowledgements

This work is the result of an extremely intense period of my life; I would have never achieved this without the constant, precious support of some people.

This thesis is dedicated to Riccardo Paramatti, whose great charismatic support and trust have motivated me even in incredibly confused and tricky times; transparent to flattering, "C ya in 10-years time" then :)

Thanks to Egidio Longo, Marcella Diemoz, Paolo Meridiani and everybody from CMS Rome for giving me trust and the possibility of learning so much; thanks to Sebastiano Albergo from CMS Catania.

Thanks to Georgios Daskalakis for integrating my work into the CMS EWK-electron Working Group at CERN.

Thanks to Andrea Marano, Antonello Cammisecra and Ivan Cascone for making this work possible even in faraway contexts.

No words would have been written without Gaia. Thank you, you know who you are.

Thanks to Dima, Giulio, Luca, Francy, Carlo, Lan, Andrea, Mirco and Timpo for being the best friends I could have ever dreamt of; I am lucky and honoured to walk by your side :)

Thanks to Marina, Placido, Pancrazio, Antonio, Alessandra and all you mates from Sicily for giving me an extra-great time; thanks to Savia for the amazing riceballs.

Thanks to my mother and my sister.

Introduction

The Large Hadron Collider is a proton-proton collider that is going to operate at CERN from winter 2009. It will provide collisions at unprecedently high centre-of-mass energies in the range 7-14 TeV, and it will allow to both make precision tests of the Standard Model of particle physics and explore the realm of the New Physics that many theories place in the TeV range.
The Compact Muon Solenoid (CMS) is a general-purpose experiment that will allow to make measurement of the collision products with an ambitious scientific programme spanning from the precise measurement of the W and the top quark mass to the search for the Higgs boson, supersymmetric particles and mini-black holes.
At the LHC startup, a number of techniques will be put in place in order to precisely calibrate the many subsystems that make up the detector, like the tracking system or the electromagnetic calorimeter.

The first part of this work describes how to use $Z \to e^+e^-$ events in order to calibrate the CMS electromagnetic calorimeter, which makes use of scintillating crystals in order to precisely measure the energy of electrons and photons coming from the proton-proton interactions. Using the very precise knowledge of the Z mass coming from LEP experiments, it is possible to set the absolute scale of the calorimeter as well as calibrating regions of the calorimeter with various topologies, and finely correct the calorimeter reponse to electrons. Focus is put on the first weeks of data taking.

The second part of this work concentrates on the misidentification of the electric charge of electrons/positrons in CMS. It will be shown how it is possible to extract the charge misidentification rate from the first CMS data, this time relying on the fact that electrons coming from the Z decay are always oppositely-charged.

Measuring this charge misidentification rate not only allows to perform a real-time check of the reconstruction quality during data taking, but also has an important role in the study of some physics channels. One of the studies where the charge misidentification has an important influence is the W^+/W^- cross section ratio, that represent a test of the Standard Model which does not need a precise knowledge of the machine luminosity, that will be difficult to achieve with the first data.

Chapter 1

The Large Hadron Collider

The Large Hadron Collider (LHC) [1] is the proton-proton ($p-p$) collider under construction at CERN. It will collide protons with a center of mass energy $\sqrt{s}$= 14 TeV with a design luminosity $\mathcal{L}$=10^{34} cm^{-2}s^{-1}. The first LHC single beams successfully circulated through the whole collider on the 10th September 2008; this successful breakthrough was spoiled by the accident occurred in sector 3-4 on the 19th September, when a large amount of Helium leaked into the tunnel caused the damage of a number of magnets. The repairs of the damaged region has started during the winter shutdown; this will cause a delay on the schedule and the next circulating beam in the LHC are foreseen for late 2009. During the first 3 years of data taking, the luminosity is expected to be 2×10^{33} cm^{-2}s^{-1} (the so called "low luminosity" phase), while at the start up LHC is forseen to run in the luminosity range $\mathcal{L}$= 10^{30} - 10^{32} cm^{-2}s^{-1}.
LHC is installed in the already existent LEP tunnel and the available CERN accelerators are employed in the injection chain: the proton beam exiting a small linear accelerator at 50 MeV are be injected in the PS at 1.4 GeV, then in the SPS at 25 GeV and finally in the LHC ring at 450 GeV (Fig. 1.1). One of the critical aspects in accelerating the protons up to an energy of 7 TeV is the required bending magnetic field which, for the LHC bending radius ($R \sim 2780$ m) is about 8.4 T. 1232 LHC superconducting 14.2 m long dipole magnets will create this magnetic field; they are placed in the eight curved sections which connect the straight sections of the LHC ring. The super-conducting magnets use a Ni-Ti conductor, cooled down to 1.9 K by means of super-fluid Helium. The choice of a $p-p$ collider obliges to install two separate magnetic chambers which lays in the same mechanical structure and cryostat for economical reasons.
The high luminosity of the LHC is obtained by a high frequency bunch crossing and by a high number of protons per bunch: two beams of protons with an energy of 7 TeV (3-5 TeV in the initial physics runs), circulating in two different vacuum chambers, will contain each 2808 bunches filled with about 1.15×10^{11} protons. The beams will cross at the interaction point at the rate of 40 MHz, with a spread of 7.5 cm along the beam axis and 15 μm in the transverse directions. The main machine parameters are summarized in Table 1.
The operating conditions at the LHC are extremely challenging for the experiments. The $p-p$ total inelastic cross section at $\sqrt{s}$= 14 TeV is about 80 mb, several orders of magnitude larger than the typical cross section for events with large momentum transfer. Most of the

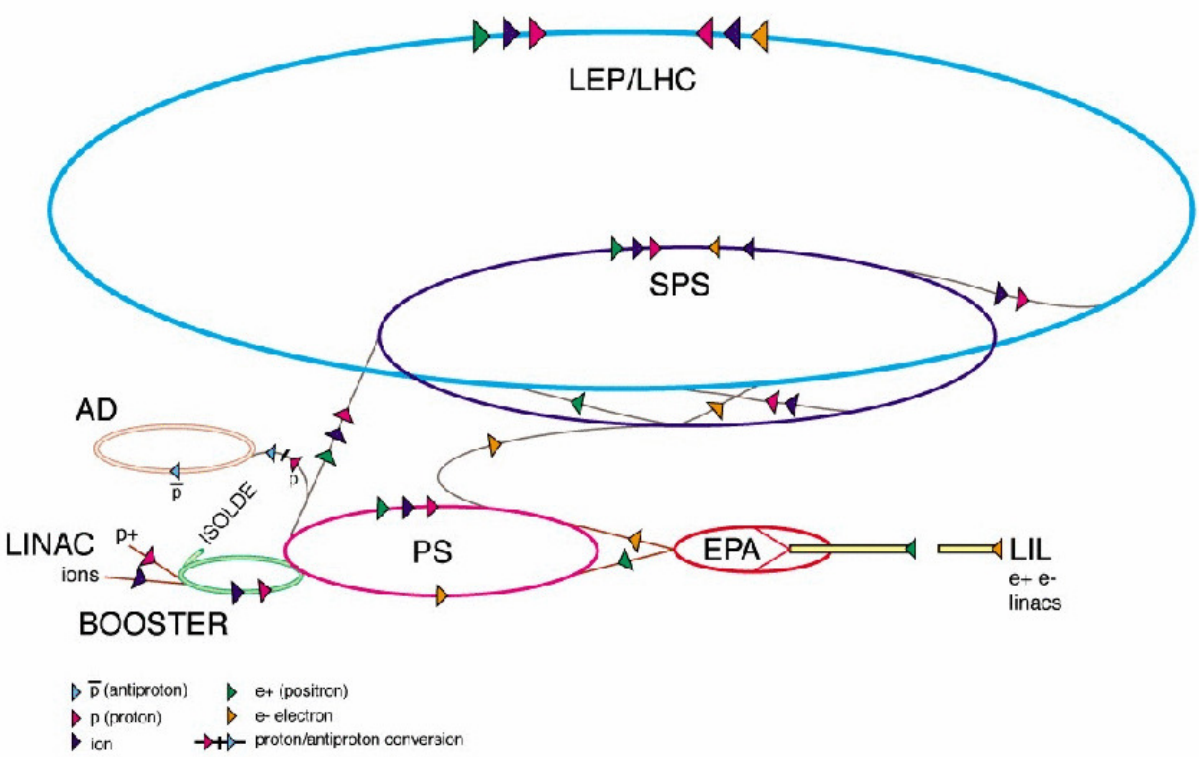


Figure 1.1: Scheme of the LHC injection chain.

inelastic events consist of soft $p-p$ interactions characterized by outgoing particles with a low transverse momentum. These events are referred to as *minimum bias*. It is expected that each bunch crossing will produce about 20 minimum bias events in the high luminosity phase and 5 minimum bias events in the low luminosity phase. Hence, each interesting event will be readout entangled with a large number of minimum bias events, which constitute the *pile-up*. The high interaction rate and the high bunch crossing frequency ($\sim 10^9$ interactions/s) impose stringent requirements on the data acquisition and trigger systems and on the detectors. The trigger has to provide an high rejection factor, maintaining at the same time an high efficiency in selecting the interesting events. The detectors must have a fast response time (25-50 ns) and a fine granularity (and therefore a large number of readout channels) in order to minimize the effect of the pile-up. Furthermore the high flux of particles coming from the $p-p$ interactions implies that each component of the detector, including the read-out electronics, has to be radiation resistant.

Beam parameters	
Beam energy	7 TeV
Maximum luminosity	$10^{34}\text{cm}^{-2}\text{s}^{-1}$
Time between collisions	25 ns
Bunch length	7.7 cm
RMS beam radius at the interaction point	16.7 μm
Technical parameters	
Ring length	26658.9 m
Radiofrequency	400.8 MHz
Number of bunches	2808
Number of dipoles	1232
Dipole magnetic field	8.33 T

Table 1.1: The relevant LHC parameters for $p-p$ collisions.

1.1 The CMS detector

The CMS (*Compact Muon Solenoid*) detector is one of the two general purpose experiment that will take data at the LHC. The CMS structure is a typical one for experiments at colliders: a cylindrical central section (the *barrel*) closed at its end by two caps (the *endcaps*), as sketched in Fig. 1.2. The coordinates system in CMS are chosen with the z axis along the beam direction, the x axis directed toward the center of the LHC ring and the y axis directed upward, orthogonally to the z and x axes. Given the cylindrical structure of CMS, a convenient and commonly used coordinate system is r, ϕ, η, where r is the distance from the z axis, ϕ is the azimuthal angle in the xy plane and η is the pseudorapidity defined as:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{1.1}$$

where θ is the angle with respect to the beam axis. The use of pseudorapidity instead of the polar angle is motivated by the fact that the difference in pseudorapidity between two particles is invariant under Lorentz boosts along the beam axis.
CMS is characterized by high hermeticity with a full coverage in ϕ and up to $|\eta| = 5$ in pseudorapidity. The detector is constituted by different subdetectors with different tasks. Starting from the innermost region to the outermost they are: the inner tracking system; the electromagnetic calorimeter (ECAL); the hadron calorimeter (HCAL) and the muon chambers. A characteristic feature of CMS is that it is embedded in a 3.8 T solenoidal magnetic field, which lead to a compact design for the detector.

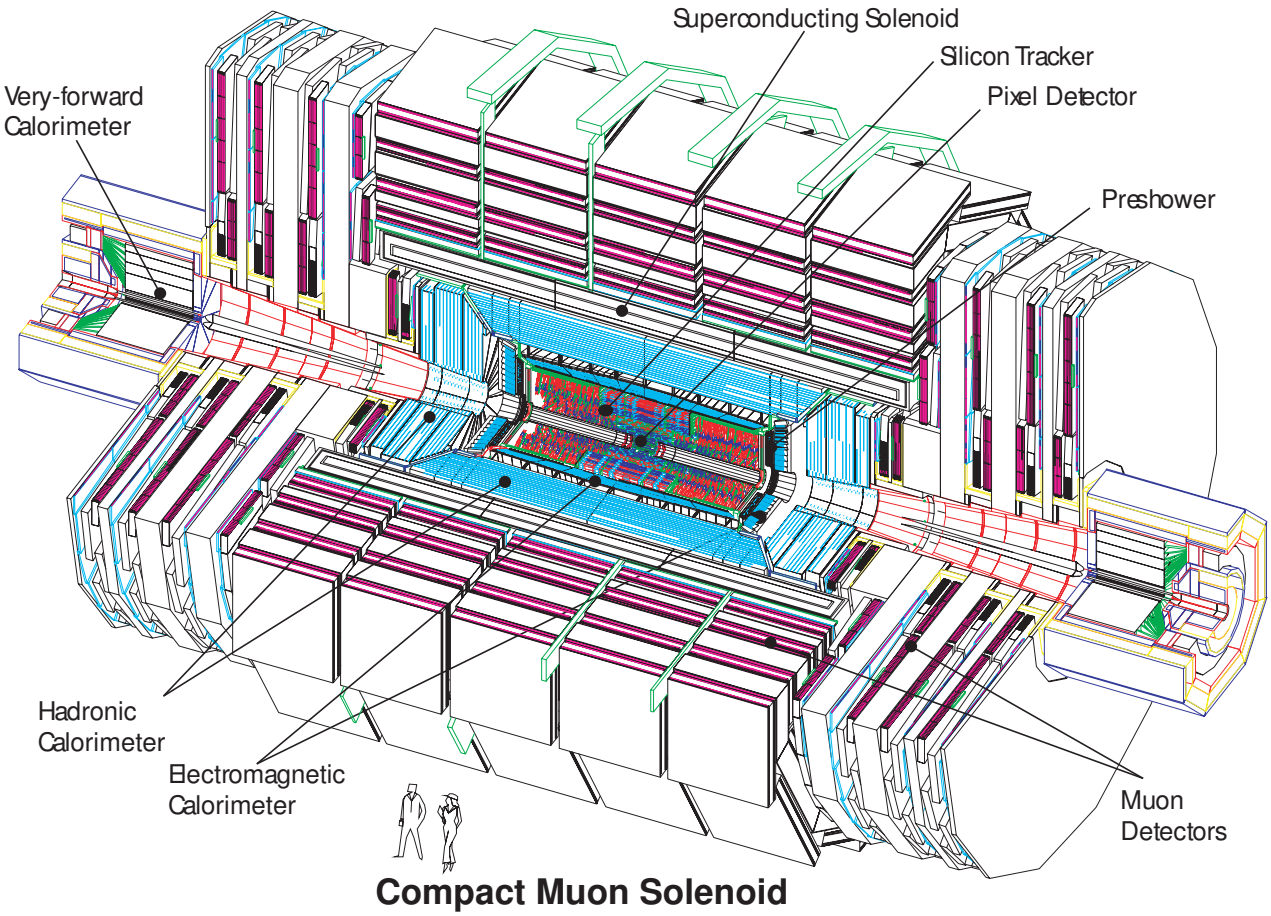


Figure 1.2: CMS overview.

1.1.1 The tracking system

CMS tracker [2] is the subdetector closer to the interaction point, placed in the superconductive solenoid; it is designed to determine the interaction vertex, measure with good accuracy the momentum of the charged particles and identify the presence of secondary vertexes.
The tracking system must be able to operate without degrading its performances in the hard radiation environment of LHC and it has to comply with severe material budget (see Fig. 1.3) constraints, in order not to degrade the excellent energy resolution of the electromagnetic calorimeter. The CMS collaboration has adopted silicon technology for the whole tracker. Three regions can be delineated, considering the charged particle flux at different radii at high luminosity:

- Closest to the interaction vertex where the particle flux is highest ($\sim 10^7 s^{-1}$ at $r \sim 10$ cm) pixel detectors are placed. The size of the pixels is $\sim 100 \times 150$ μm^2, leading to an occupancy of 10^{-4} per pixel per LHC crossing.

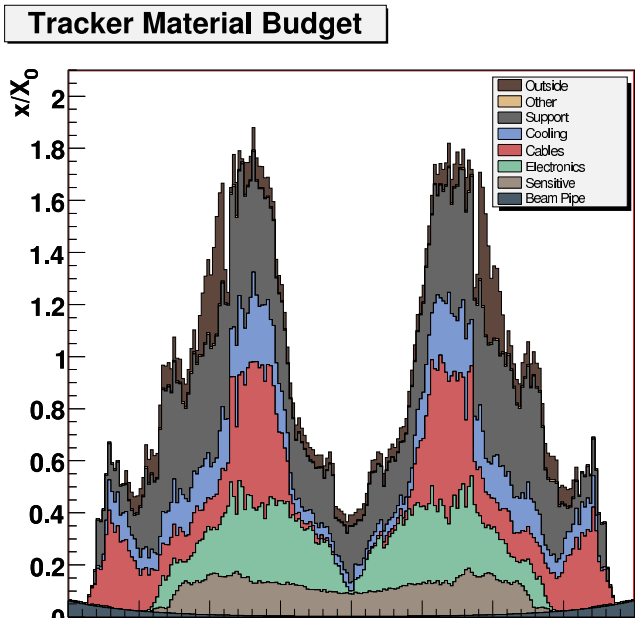


Figure 1.3: Material budget as a function of η. The thickness is expressed in terms of radiation length (X_0). The peak around η=1.5 corresponds to the cables and services of the tracker.

- In the intermediate region with 20 cm $< r <$ 55 cm the particle flux becomes low enough to allow the use of silicon microstrip detectors, with a minimum cell size of $\sim$ 10 cm $\times$ 80 μm, giving an occupancy of 2-3% per LHC crossing.
- The outermost region is characterized by sufficiently low fluxes that enable to adopt larger-pitch silicon microstrips with a maximum cell size of $\sim$ 25 cm $\times$ 80 μm, keeping the occupancy to $\sim$1%.

The pixel detector consists of three barrel layers and two endcap disks at each side (Fig. 1.4). The barrel layers are located at 4.4 cm, 7.3 cm and 10.2 cm and are 53 cm long. The two end disks, extending from 6 to 15 cm in radius, are placed on each side at $|z|$= 34.5 cm and 46.5 cm. This design allows to obtain at least two points per track in the $|\eta|$ <2.2 region for tracks originating within $2\sigma_z$ from the central interaction point. The total number of channels is about 44 millions, organized in about 16000 modules of 52 columns and 80 rows. The total active area is close to 0.92 m^2. The presence of high magnetic field causes a noticeable drift of the electrons (and a smaller drift for the holes) from the ionization point along the track with a Lorentz angle of about 32°. This leads to a charge sharing between pixels which, using an analog readout, can be exploited to considerably improve the resolution down to about 10 μm. In the endcap the modules of the detector are arranged in a turbine-like shape with a 20° tilt, again in order to enhance the charge sharing.
The inner and outer tracker detector (see Fig. 1.5) are based on silicon strips. They are p^+ strips on a n-type bulk whose thickness is close to 300 and 500 μm respectively in the inner

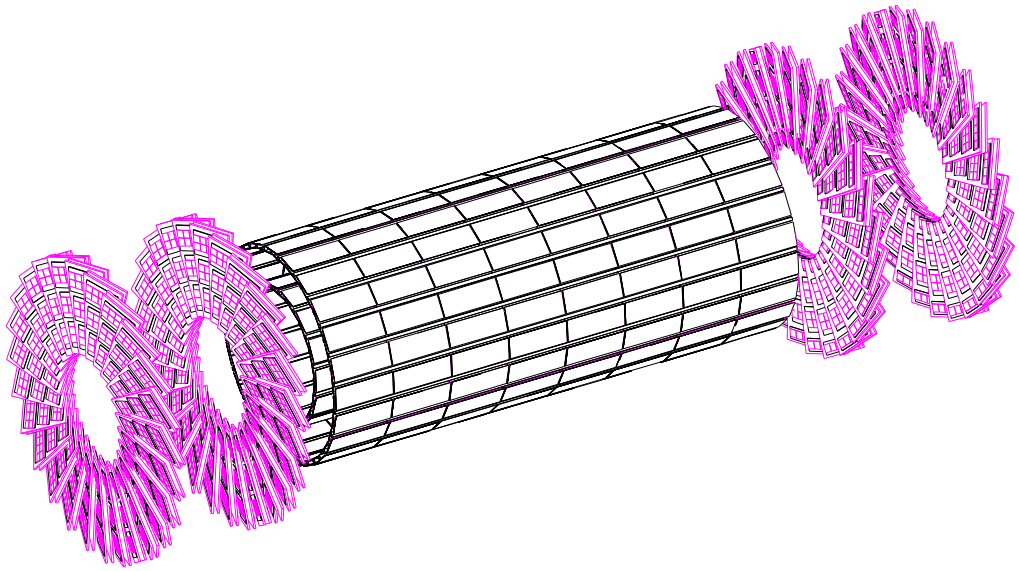


Figure 1.4: The inner pixel detector. The three barrel section and the two disks of the endcap with blades disposed in a turbine-like shape are visible.

and outer tracker. In the barrel the strips are parallel to the beam axis while for the endcaps they have a radial orientation. The inner tracker is made of 4 barrel layers, the two innermost are double sided and the endcaps count 3 disks each. The outer tracker consists of 6 layers in the barrel (the two innermost are double sided) while the endcaps are made of 9 layers (the first, the second and the fifth are double sided). On the whole the silicon trackers is made of about 10 millions of channels for an active area close to 198 m^2.

1.1.2 The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) measures the energy of the electrons and photons. The design of the CMS ECAL [3] was driven by the requirements imposed by the search of the Higgs boson in the channel $H \to \gamma\gamma$, where a peak in the di-photon invariant mass placed at the Higgs mass has to be distinguished from a continuous background. A good resolution and a fine granularity are therefore required: both of them improve the invariant mass resolution on the di-photon system by improving respectively the energy and the angle measurement of the two photons. The fine granularity also helps to obtain a good π^0/γ separation.
In order not to deteriorate the energy resolution ECAL is placed inside the solenoid, hence compact calorimeter is required. ECAL is a hermetic, homogeneous calorimeter made of lead tungstate ($PbWO_4$) crystals, 61200 of them mounted in the central barrel part, and 7324 crystals in each endcap (Fig. 1.6). The choice of lead tungstate scintillating crystals [4] was driven by their characteristics: they have a short radiation length ($X_0 = 0.89$ cm) and a small Moliere radius ($R_M = 2.2$ cm); they are fast, as the 80% of the scintillation light is emitted within 25 ns and radiation hard. The use of $PbWO_4$ crystals has thus allowed to

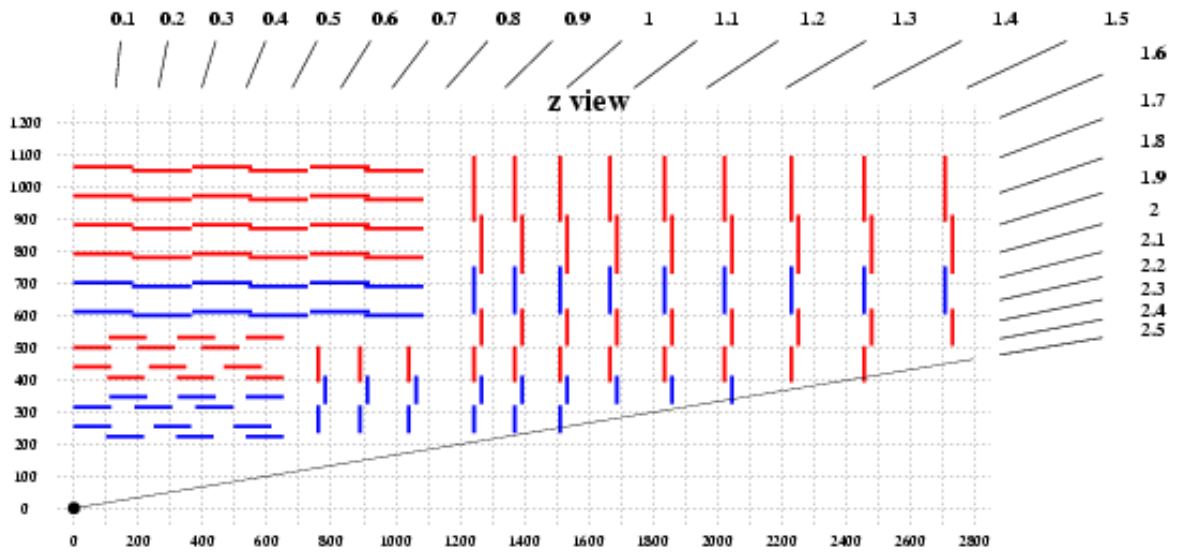


Figure 1.5: CMS tracker: layout of the silicon strip detectors (1/4 of the z view). Red (blue) indicates single-sided (double-sided) layers.

design of a compact calorimeter placed inside the solenoid, fast and with fine granularity and radiation resistant. However the relative low light yield (80-100 γ/MeV) requires the use of photodetectors with intrinsic gain that can operate in a magnetic field. In the barrel, silicon avalanche photodiodes (APDs) are used as photodetectors, while vacuum phototriodes (VPTs) have been chosen for the endcaps. In addition, the sensitivity of both the crystals and the APDs response to temperature changes requires temperature stability; details on the temperature stability of the High Voltage system can be found in [5]. A water cooling system guarantees a long term stability at the 0.05°C in order to preserve the ECAL energy resolution performances.
The barrel region has a pseudorapidity coverage up to $|\eta| < 1.479$. It has an inner radius of 129 cm and is structured in 36 supermodules, each containing 1700 crystals having the shape of a truncated pyramid, covering half the barrel length and subtending a 20° angle in ϕ. Each supermodule is divided along η into four modules which in turn are made of submodules, the basic assembling alveolar units, containing 5×2 crystals each. The barrel crystals have a front face cross-section of $\sim 22\times$ 22 mm^2 and have a length of 230 mm, corresponding to $25.8X_0$. The crystal axes are oriented with a 3° tilt with respect to the pointing geometry to avoid that particles can directly escape into the dead regions between the crystals, as can be seen in Fig. 1.7. The granularity of the barrel is $\Delta\phi \times \Delta\eta = 0.0175 \times 0.0175$ and the crystals are grouped, from the readout point of view, into 5×5 arrays corresponding to the *trigger towers*. The endcaps cover the pseudorapidity region $1.48 < |\eta| < 3.0$, ensuring precision measurements up to $|\eta| < 2.5$. The endcap crystals have dimensions of 28.6×28.6×220 mm^2. Each

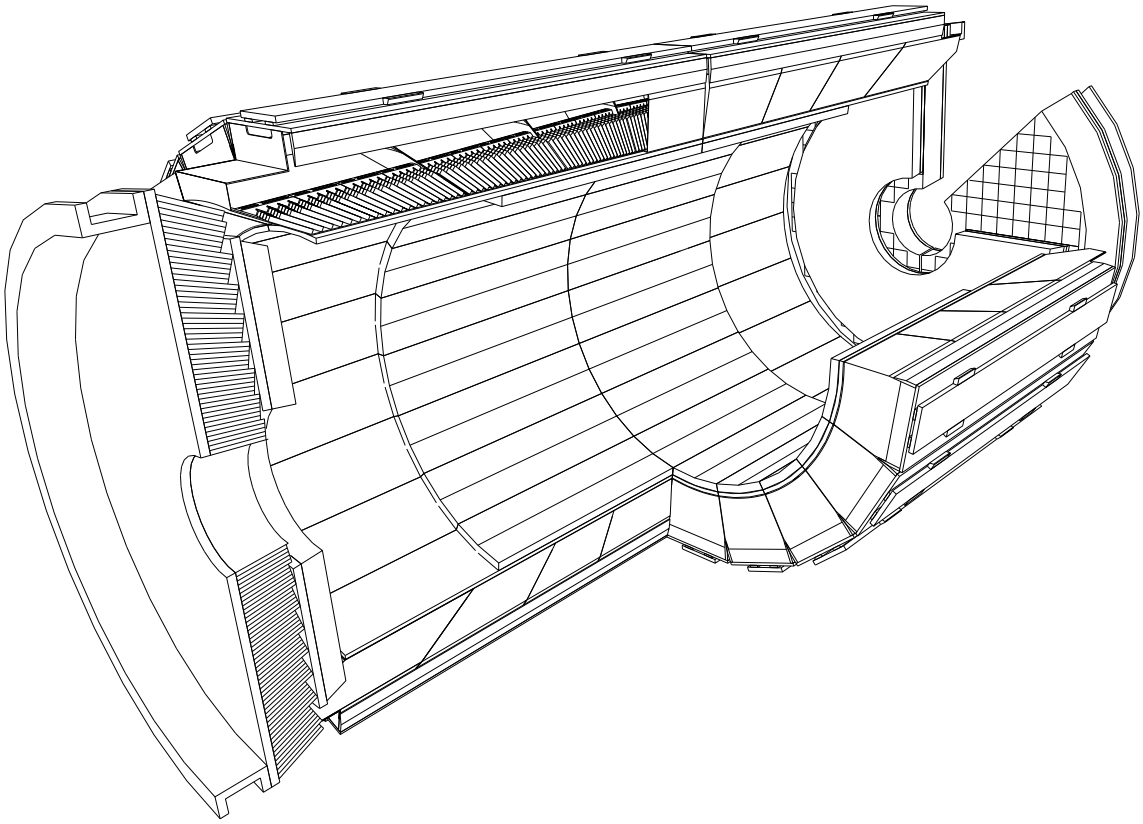


Figure 1.6: Scheme of the barrel and of the endcaps of the CMS ECAL.

endcap is structured in two "Dees" consisting of semi-circular aluminum plates from which are cantilevered structural units of 5×5 crystals, known as "supercrystals".
A preshower device, whose principal aim is to identify neutral pions in the endcaps within $1.653 < |\eta| < 2.6$, is placed in front of the crystal calorimeter. The active elements are two planes of silicon strip detectors which lie behind disks of lead absorber at depths of $2X_0$ and $3X_0$.

ECAL goals and performances

One of the relevant issue in evaluating the performances of the electromagnetic calorimeter is its energy resolution. In the relevant energy range between 25 GeV and 500 GeV, the energy resolution is usually parametrized as the sum in quadrature of three different terms:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad , \tag{1.2}$$

where a, b and c are named respectively stochastic, noise and constant term, and E is the energy expressed in GeV.
The target values for CMS are about 2.7% for a, (120) 200 MeV when adding the signal of

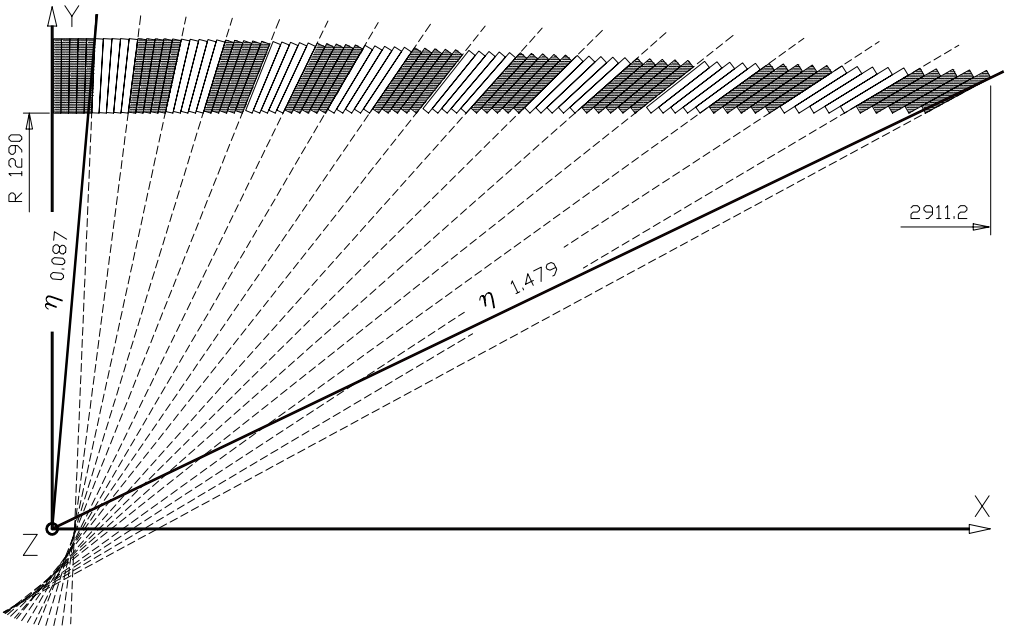


Figure 1.7: Cross section view of the crystals in the calorimeter. The groups of five crystals are highlightened by the changing of color. The shape of the crystals is different between the groups. The 3° tilt with respect to the pointing geometry is clearly visible.

(3x3) 5x5 crystals for b, and 0.5% for c; their individual contribution to the energy resolution, as a function of energy, is shown in Fig. 1.8.
Different effects contribute to the different terms in eq. (1.2):

- the stochastic term a receives a contribution from the fluctuations in the number of electrons which reaches the preamplifier (n_e). These fluctuations are proportional to $\sqrt{n_e}$ and therefore proportional to the square root of the deposited energy. Contributions to this term come from the light yield of the crystals, from the efficiency in collecting light onto the photodetector surface and from the quantum efficiency of the photodetector.

- The noise term b accounts for all the effects that can alter the measurements of the energy deposit independently of the energy itself. This term receives contributions from the electronic noise and from the pile-up events, whose contribution are different in the barrel and in the endcaps and can vary with the luminosity of LHC. The target values for the barrel (at $\eta = 0$) and the endcaps (at $\eta = 2$) in the low luminosity running are respectively 155 MeV and 205 MeV while in the high luminosity running they are 210 MeV and 245 MeV [6].

- The constant term determines the energy resolution at high energy. Many different effects contribute to this term: the stability of the operating conditions such as the temperature and the high voltage of the photodetectors; the electromagnetic shower

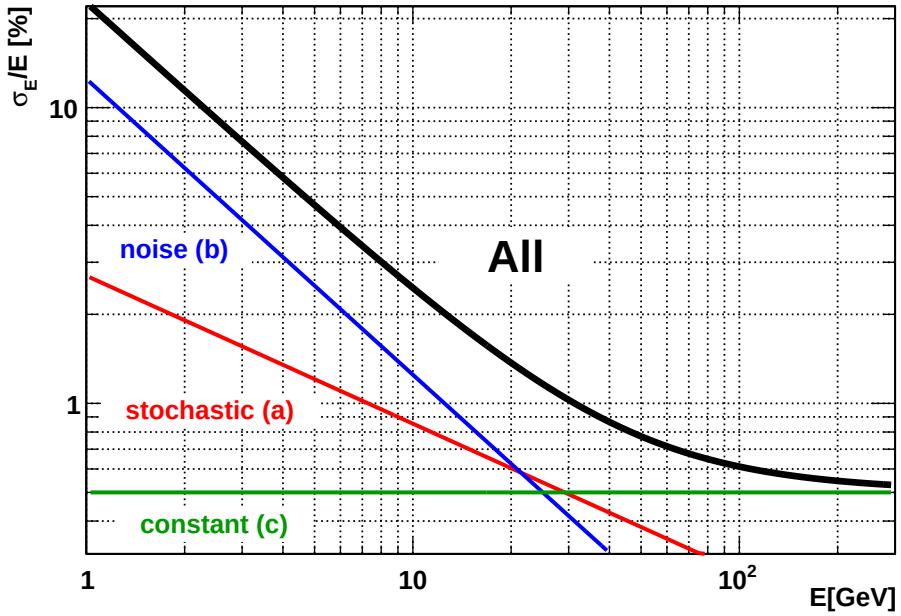


Figure 1.8: Contribution of the different terms to the ECAL energy resolution as a function of energy; design values for the Barrel are used for the terms a, b and c (see text for details). The noise term (b) refers to matrices of 3x3 crystals.

containment and the presence of the dead material of the supporting structure between the crystals; the light collection uniformity along the crystal axis; the intercalibration between the channels which contributes almost directly to the overall energy resolution since the most of the energy is contained into few crystals; the radiation damage of the $PbWO_4$ crystals [1].

As visible in Fig. 1.9 [7], measurements conducted on the ECAL Barrel with electron test beam at CERN [7, 8, 9] showed that:

- The average measured noise of the ECAL read-out electronics is 129 MeV equivalent for 3x3 crystal matrices, meeting the design goal of the electronics, and there is no significant correlated noise in arrays of 3x3 and 5x5 crystals used to reconstruct the energy of incident electrons.

- The energy resolution was measured for incident electrons of several energy values from 20 to 250 GeV; the measured resolutions are consistent with expectations, with mean

[1] Changes in the crystal transparency caused by irradiation and subsequent annealing lead to a variation in the crystal response to a given deposited energy; this effect develops on a short term scale ($\sim$ hours) and need to be tracked and corrected properly using a laser monitoring system.

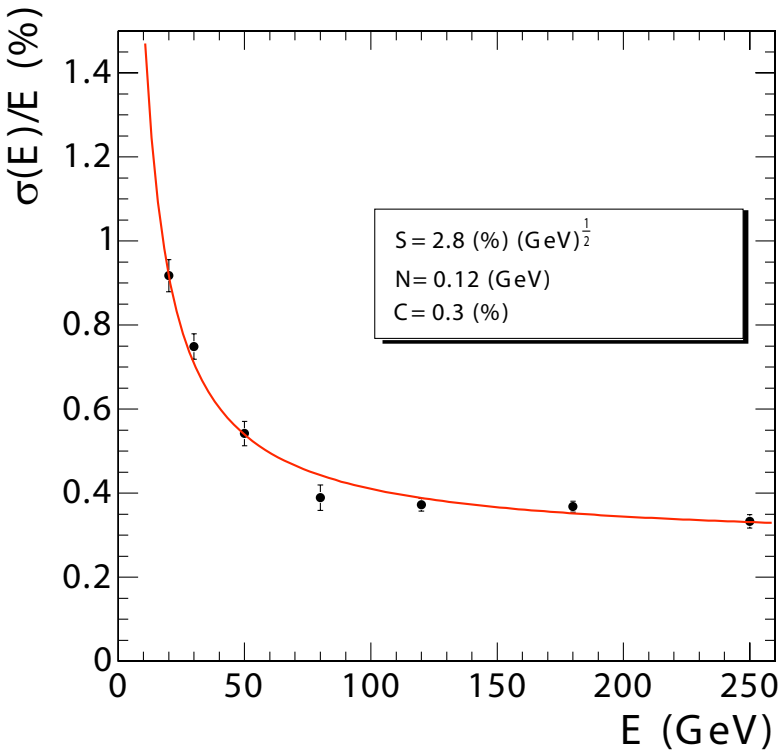


Figure 1.9: The energy resolution of the ECAL as measured at test beam facility; see text for notation and details.

values of the stochastic and constant terms of around 2.8% and around 0.3%, respectively.

These results therefore show that the CMS ECAL should perform consistently with the design goals of the experiment.

1.1.3 The hadron calorimeter

The hadron calorimeter (HCAL) [10], placed just outside the electromagnetic calorimeter, plays a major role in the reconstruction of jets and missing energy. Its resolution must guarantee a good reconstruction of the di-jets invariant mass and an efficient measurement of the missing energy which represents an effective signature in many channels of physics beyond the Standard Model. Similarly to the other subdetectors, HCAL has to provide a good hermeticity, which is critical for determining the missing energy, and a quite fine granularity to allow a clear separation of di-jets from resonance decays to be made and improve the resolution in the invariant mass of the di-jets. Moreover it has to provide a number of interaction lengths sufficient to contain the energetic particles from high transverse momentum jets. The dynamic

range has to be large enough to detect signals ranging from the signal of a single minimum ionizing muon up to an energy of 3 TeV.
The pseudorapidity region $|\eta| < 3$ is covered by the Hadron Barrel (HB, up to $|\eta| < 1.74$) and the two Hadron Endcaps (HE) calorimeters . HCAL is composed of brass layers as absorbers interleaved with 4 mm thick plastic scintillator layers used as active medium. The thickness of the absorber layers is between 60 mm in the barrel and 80 mm in the endcaps. In terms of interaction lengths λ, the barrel ranges from 5.46λ at $|\eta|$ =0 up to 10.82λ at $|\eta|$ =1.3; the endcap corresponds on average to 11λ. The scintillator in each layer is divided into tiles with a granularity matching the granularity of the ECAL trigger towers ($\Delta\eta \times \Delta\phi = 0.0875 \times 0.0875$) and the light is collected by wavelength shifters.
The hadron calorimeter barrel is radially restricted between the outer extent of the electromagnetic calorimeter (R = 1.77 m) and the inner extent of the magnet coil (R = 2.95 m); this constrains the total amount of material which can be put in to absorb the hadronic shower. Therefore, an outer hadron calorimeter or tail catcher (HO) is placed outside the solenoid complementing the barrel calorimeter. The HO will utilize the solenoid coil as an additional absorber equal to $1.4/\sin\theta$ interaction lengths and will be used to identify late starting showers and to measure the shower energy deposited after HB. The HO is constrained by the geometry of the muon system. The central ring (ring 0) has two layers of HO scintillators on either side of a 19.5 cm thick piece of iron (the tail catcher iron) at radial distances of 3.82 m and 4.07 m, respectively; all other rings have a single HO layer at a radial distance of 4.07 m. The effect of shower leakage has a direct consequence on the measurement of missing transverse energy in an event ($\not\!\!E_T$), leading to the production of the so-called *fake* $\not\!\!E_T$.
The two hadronic forward (HF) calorimeters improve the HCAL hermeticity, covering the pseudorapidity region 3< $|\eta|$ <5. They are placed at 11.15 m from the interaction point outside the magnetic field. Due to the extremely harsh radiation environment a different detection technique is used: a grid of quartz (radiation hard) fibers is embedded in a iron absorber.

1.1.4 The magnetic field

An important aspect of the CMS experiment is its solenoidal high magnetic field. The magnet system of CMS [11] is composed of three main parts: the superconducting solenoid, the barrel return yoke and the endcap return yoke. The 3.8 T magnetic field allows the muon momentum to be measured efficiently up to a pseudorapidity of 2.4. The return yoke is made of iron and contains the muon detectors. It is a 12-sided cylindrical structure, with a total length is about 11 m and it is divided into five rings of about 2.5 m each. The barrel part has an outer diameter of 14 m and a total weight of about 7000 tons. Each ring is divided into three iron layers where the muon detectors are inserted. The thickness of the outer layers is 630 mm and the middle layer is 295 mm thick. Each endcap is composed by three independent disks, the outermost is 300 mm thick and the others are 600 mm thick.
The superconductive coil is housed into a vacuum tank and kept at the temperature of the liquid helium. The vacuum tank is supported only by the central barrel ring of the yoke and in turn supports the calorimeter system (ECAL and HCAL) and the tracker.

1.1.5 The muon system

The muon system [12] has three functions: muon identification, momentum measurement, and triggering. Good muon momentum resolution and trigger capability are enabled by the high-field solenoidal magnet and its flux-return yoke. The latter also serves as a hadron absorber for the identification of muons. In CMS the muon detectors are placed beyond the calorimeters and the solenoid. Because the muon system consists of about 25 000 m^2 of detection planes, the muon chambers have to be inexpensive, reliable, and robust.

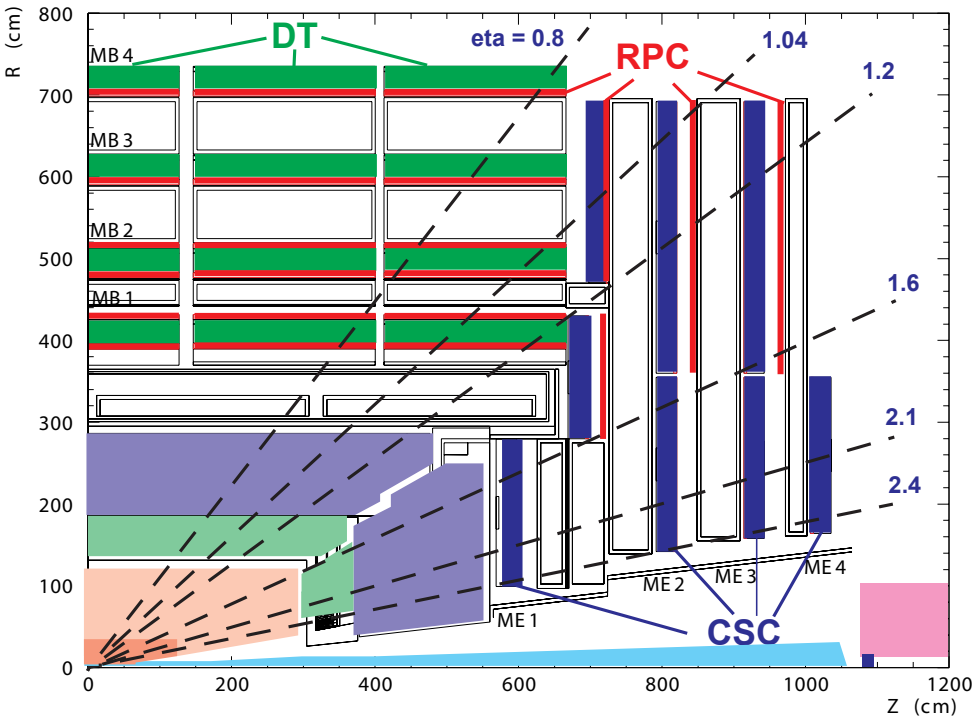


Figure 1.10: Layout of one quarter of the CMS muon system

The muon system layout is schematically depicted in Fig. 1.10. The barrel drift tube (DT) chambers cover the pseudorapidity region $|\eta| < 1.2$ and are organized into 4 stations interspersed with the layers of the flux return plates. The number of chambers in each station and their orientation were chosen to provide good efficiency for linking together muon hits from different stations into a single muon track and for rejecting background hits. In the two endcap regions of CMS, where the muon rates and background levels are high and the magnetic field is large and non-uniform, the muon system uses cathode strip chambers (CSC). The CSCs identify muons between $0.9 < |\eta| < 2.4$. There are 4 stations of CSCs in each endcap, with chambers positioned perpendicular to the beam line and interspersed between the flux return plates. A crucial characteristic of the DT and CSC subsystems is that they can each trigger on the p_T of muons with good efficiency and high background rejection, independent of the rest of the detector.

Because of the uncertainty in the ability of the muon system to measure the correct beam-crossing time when the LHC reaches full luminosity, a complementary, dedicated trigger system consisting of resistive plate chambers (RPC) was added in both the barrel and endcap regions. The RPCs provide a fast, independent, and highly-segmented trigger with a sharp p_T threshold over a large portion of the rapidity range ($|\eta| < 1.6$ for the first period of data taking) of the muon system. The RPCs are double-gap chambers, operated in avalanche mode to ensure good operation at high rates.

1.1.6 The trigger system

At the nominal LHC luminosity the expected event rate is about 10^9 Hz; given the typical size of a raw event ($\sim$ 1 MB) it is not possible to record all the information for all the events. Indeed, the event rate is largely dominated by soft $p - p$ interactions with particles of low transverse momentum. The triggering system must have a large reduction factor and maintain at the same time a high efficiency on the potentially interesting events, reducing the rate down to about 200 Hz, which is the maximum sustainable rate for storing events. The trigger system consists of two main steps: a Level 1 Trigger (L1) and a High Level Trigger. L1 consists of custom-designed, largely programmable electronics, whereas the HLT is a software system implemented in a filter farm of about one thousand commercial processors. The rate reduction capability is designed to be a factor of 10^7 for the combined L1 and HLT.

The Level 1 trigger

The Level 1 trigger (L1) [13] reduces the rate of selected events down to 50 (100) kHz for the low (high) luminosity running. The full data are stored in pipelines of processing elements, while waiting for the trigger decision. The L1 decision about taking or discarding data from a particular bunch crossing has to be taken in 3.2 μs; if the L1 accepts the event, the data are moved to be processed by the High Level Trigger. To deal with the high bunch crossing rate, the L1 trigger has to take a decision in a time too short to read data from the whole detector, therefore it employs the calorimetric and muons information only, since the tracker algorithms are too slow for this purpose. The Level-1 trigger is organized into a Calorimeter Trigger and a Muon trigger whose information are transferred to the Global Trigger which takes the accept-reject decision.
The Calorimeter Trigger is based on trigger towers, arrays of 5 crystals in ECAL , which match the granularity of the HCAL towers. The trigger towers are grouped in calorimetric region of 4×4 trigger towers. The Calorimeter Trigger identifies, from the calorimetric region information, the best four candidates of each of the following classes: electrons and photons, central jets, forward jets and τ-jets identified from the shape of the deposited energy. The information of these objects is passed to the Global Trigger, together with the measured missing E_T. The Muon trigger is performed separately for each muon detector (see Sec. 1.1.5). The information is then merged and the best four muon candidates are transferred to the Global Trigger.
The Global Triggers takes the accept-reject decision exploiting both the characteristic of the

single objects and of combination of them.

The High Level Trigger

The High Level Trigger (HLT) [14] reduces the output rate down to around 200 Hz. The idea of the HLT trigger software is the regional reconstruction on demand, that is only those objects in the useful regions are reconstructed and the uninteresting events are rejected as soon as possible. This leads to the development of three "virtual trigger" levels: at the first level only the full information of the muon system and of the calorimeters is used, in the second level the information of the tracker pixels is added and in the third and final level the full event information is available.

1.2 CMS software components

The high-level goals of the CMS software are to process and select events inside the High Level Trigger Farm [14], to deliver the processed results to experimenters within the CMS Collaboration, and to provide tools for them to analyze the processed information in order to produce physics results. The overall collection of software, now referred to as CMSSW [15], is built around a Framework, an Event Data Model, Services needed by the simulation, calibration and alignment and reconstruction modules that process event data so that physicists can perform analysis. The physics and utility modules are written by detector groups. The modules can be plugged into the application framework at run time, independently of the computing environment. The software should be developed keeping in mind not only performance but also modularity, exibility, maintainability, quality assurance and documentation. CMS has adopted an object-oriented development methodology, based primarily on the C++ programming language.
The primary goal of the CMS Framework and Event Data Model (EDM) is to facilitate the development and deployment of reconstruction and analysis software. The event data model is centered around the Event. The Event holds all data that was taken during a triggered physics event as well as all data derived from the data taking (e.g. calibration and alignment constants).
The detailed CMS detector and physics simulation is currently based on the GEANT 4 [16] simulation toolkit and the CMS object-oriented framework and event model. The simulation is implemented for all CMS detectors in both the central and forward regions (Tracker, Calorimeters and Muon Systems) and in the very forward region (CASTOR calorimeter, TOTEM telescopes, Roman Pot detectors and the Zero Degree Calorimeter, ZDC), including the field map from the 3.8 T solenoid. In addition, several test-beam prototypes and layouts have been simulated. The full simulation program implements the sensitive detector behavior, track selection mechanisms, hit collection and digitization (i.e. detector response). The detailed simulation workflow is as follows:

- A physics group configures an appropriate Monte Carlo event generator (several are used) to produce the data samples of interest.

- The production team/system runs the generator software to produce generator event data files.
- The physics group validates the generator data samples and selects a configuration for the GEANT 4 simulation (detector configuration, physics cuts, etc.).
- The production team/system runs the GEANT 4-based simulation of CMS, with generator events as input, to produce (using the standard CMS framework) persistent hits in the detectors.
- The physics group validates these hit data which are then used as input to the subsequent digitization step, allowing for pile-up to be included. This step converts hits into digitizations (also known as "digis") which correspond to the output of the CMS electronics.

As mentioned above, the full simulation relies on the GEANT 4 toolkit. GEANT 4 provides a rich set of physics processes describing electromagnetic and hadronic interactions in detail. It also provides tools for modelling the full CMS detector geometry and the interfaces required for retrieving information from particle tracking through these detectors and the magnetic field. The validation of GEANT 4 in the context of CMS is described in detail in [17]. The CMS GEANT 4-based simulation program uses the standard CMS software framework and utilities, as used by the reconstruction programs.
The digitization step, following the hit creation step, constitutes the simulation of the electronic readout used to acquire data by the detector and DAQ systems. It starts from the hit positions and simulated energy losses in the detectors, and produces an output that needs to be as close as possible to real data coming from CMS. Information from the generation stage (e.g. particle type and momentum) is preserved in the digitization step.
The CMS Fast Simulation framework [18] has been built in view of doing physics analyses, developing and tuning reconstruction algorithms, designing detector upgrades, without being penalized by CPU time considerations while still benefitting from an accurate simulation of the detector effects. It is an object-oriented system for which C++ has been chosen as programming language. The Fast Simulation does not depend on the GEANT software at all, and this allows an execution time at least two orders of magnitude faster than the full simulation to be obtained. To achieve this performance, a number of simplifying assumptions were made, a number of dedicated parametrizations were used, and some optimized reconstruction algorithms were developed (e.g., for tracking). The Fast Simulation is validated and tuned to the full simulation.

1.2.1 Event reconstruction

Physics at hadron colliders is characterized by the presence of a great variety of particles, each of them with its own properties: the understanding of these drives the development of reconstruction and identification algorithms, which give as results the *Physical Objects* [19].

Jets

Due to to the color symmetry in QCD, only colorless final states can be observed: a single quark or gluon state cannot be observed directly. What is actually seen in the detector is the result of the *hadronization* process, in which the parton is fragmented into colorless hadrons: this results in a grossly collimated hadron flow detected by the tracking and the calorimetric systems, called *hadronic jet* or simply jet.
A jet algorithm is a procedure which assembles a collection of objects (calorimetric towers, particles). Besides giving precise estimates of the initial parton energy and direction, a good algorithm should also be *collinear safe*, so that the result obtained must be unchanged if e.g. the energy carried by one particle is distributed among two collinear particles, and *infrared safe*, which means that the algorithm is stable against the introduction of soft particles. Three algorithms were considered:

- *Iterative Cone* (IC): this is a simple cone-based algorithm also used online in the CMS High Level Trigger. Input objects with $E_T > 1$ GeV sorted by descending order are used as *seeds* for the iterative search for stable cones associating all the inputs with $R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < R_C$ from the cone axis, with R_C the cone amplitude. This algorithm is one of the most used in high energy physics, it has a short and well predictable computation time, but it is neither collinear nor infrared safe.

- *Midpoint Cone* (MP): this algorithm also is based on fixed angular magnitude R_C cones. An improvement compared to the IC algorithm is given by considering as seeds also each pair of proto-jets (a proto-jet is a not yet definitive jet) closer than $2R_C$. Furthermore each input can be associated with several proto-jets, and splitting/merging algorithms are applied afterwards to ensure each input belongs to a single jet. This algorithm has proved to be infrared safe for Leading and Next-to-Leading order perturbative QCD, but not beyond.

- *Fast* k_T [20]: this is an implementation of the k_T algorithm [21] that dramatically reduced the computational time. This algorithm merges two inputs if they are near in the metric:

$$d_{ij} = 2\frac{min(p_{T,i}^2, p_{T,j}^2)\Delta R_{ij}^2}{D} \tag{1.3}$$

$$d_i = p_{T,i}^2 \tag{1.4}$$

  The parameter D plays role similar to the cone amplitude R_C. If $d_{ij} < d_i$ for all j then the two objects are merged in a proto-jet, otherwise the i input is removed from the list. The k_T algorithm is infrared and collinear safe.

The inputs for the calorimetric jets created by these algorithms are the calorimetric towers built using the Scheme B prescriptions described in Table 1.2. The resulting jets are called also

HB [GeV]	HO [GeV]	HE [GeV]	$\sum$ EB [GeV]	$\sum$ EE [GeV]
0.90	1.10	1.40	0.20	0.45

Table 1.2: Energy thresholds for "Scheme B" calorimeter noise suppression. $\sum$EB ($\sum$EE) refers to the sum of ECAL energy deposits associated with the same tower in the barrel (endcap)

raw jets, as no correction has been applied to them. Grossly speaking, the major corrections to be applied are: the dependence on the calorimeter response, in order to regain a flat η response, and the absolute p_T correction, which rescales the jet energy by a calibration factor retrieved by MonteCarlo (the current default) or by data as γ+jet imbalance.

Missing Transverse Energy ($\not\!\!E_T$)

Many particles cannot be measured directly in a detector, for example neutrinos: they can be nevertheless inferred them from the total momentum imbalance in the transverse plane. The Missing Transverse Momentum $\vec{E_T}^{miss}$ is defined as the sum over energy deposits in uncorrected, projective calorimeter towers:

$$\vec{E_T}^{miss} = -\sum(E_n \sin\theta_n \cos\phi_n \hat{\mathbf{i}} + E_n \sin\theta_n \sin\phi_n \hat{\mathbf{j}}) \quad (1.5)$$

$$= E_x^{miss}\hat{\mathbf{i}} + E_j^{miss}\hat{\mathbf{j}} \quad (1.6)$$

where the calorimeter towers are objects composed at least by one HCAL cell and the corresponding 5x5 ECAL crystal matrices. The threshold values used for the calorimeter cells in the different subdetectors are given by the standard "Scheme B" prescription; an overall tower threshold $E_T > 0.5$ GeV is also applied. The Missing Transverse Energy $\not\!\!E_T$ is defined as $|\vec{E_T}^{miss}|$. Despite the fact that this quantity has a very simple definition, it is extremely sensitive to detector malfunctions (e.g. hot or dead channels) or to particles hitting poorly instrumented regions of the detector. All these issues can give rise to artificial $\not\!\!E_T$ which can mimic signal for new physics. Furthermore, the low- and the high-$\not\!\!E_T$ regions have different experimental challenges: the former is dependent on the environmental backgrounds and needs a good low energy resolution, while the latter needs the understanding of the tails, that for example can be given by QCD jets in a detector environment with hot/dead cells.
To the measured $\not\!\!E_T$, called also *raw* $\not\!\!E_T$, corrections from jet energy response and energy scale, muons (which have negligible energy deposits in the calorimeters) and also particle-based algorithms can be applied to improve the $\not\!\!E_T$ resolution and central value. The current implementation of $\not\!\!E_T$ corrections uses MonteCarlo corrected jets, taking into account the detector response and energy scale:

$$\vec{E_T}^{miss\,corr} = \vec{E_T}^{miss} - \sum_{i=1}^{N_{jets}} [\vec{p}_{T,i}^{corr} - \vec{p}_{T,i}^{raw}]$$

Tracks

Tracking has reached a level of great complexity at CMS. The final fit is based on the Kalman Filter technique [22] and performs an iterative search for the hits in the tracker detector, starting from the innermost layer till the outmost one and then refitting back, with a chi-square check between the last hit measurement and the predicted track.

Muons

Muons are reconstructed in CMS using two different algorithms, one using only information from the muon chambers and the other with also the silicon tracker data. In the former case the segments or the tridimensional points are used as inputs for a Kalman Filter fit: the result is then propagated till the nominal interaction point, with a χ^2 check on track goodness: this object is called *Stand Alone Muon* (STA). Starting from a STA muon, the track in the silicon tracker is extrapolated from the innermost muon chamber hit to the outermost tracker hit: then a Kalman Filter fit is applied till a compatible track has been found. The final result is called Global Muon.

Electrons

The sequence used to reconstruct the electrons [23] is called `pixelMatchGsfElectrons`. Starting from an ECAL SuperCluster, defined as a *cluster of clusters* [17], a pixel track seed matching the SuperCluster is searched for. If the seed is found, the pattern recognition is performed with the Combinatorial Track Finder algorithm in a loose cut configuration while the final fit with the Gaussian Sum Filter [24] (GSF). GSF is a fitting algorithm dedicated to electrons that accounts for the electron bremsstrahlung energy loss. As reconstructed pions are often misidentified as electrons, the electrons have to pass the tight version of the category based electron-id algorithm [25]. Electron isolation is also required: no tracks with $p_T > 1.5$ GeV have to lie in a cone with $0.02 < \Delta R < 0.2$ around the electron; the ECAL deposit within a cone with $\Delta R = 0.3$ is required to be $< 0.05 \times E_{SC}$, while the HCAL deposit in a cone with $0.15 < \Delta R < 0.3$ has to be $< 0.2 \times E_{SC}$.

Chapter 2

Calibration of the CMS electromagnetic calorimeter with Z events

The present chapter develops some experimental methods that are useful in order to calibrate the electromagnetic calorimeter of the CMS detector.
After a brief review of the CMS ECAL calibration strategy, it is shown how it is possible to evaluate the level of ECAL calibration from the experimental Z width in the dielectron channel; it is then analyzed with what accuracy it is expected to measure the ECAL absolute scale using $Z \to e^+e^-$ events, and an iterative method for calibrating different calorimeter regions is introduced and described in detail.
As in all parts of the present work, a centre-of-mass energy of 10 TeV will be assumed in the calibration studies.
The main focus of the present chapter is on *methods*; a detailed, realistic study of the background to $Z \to e^+e^-$ events is therefore performed in Chapters 3 and 4.

2.1 ECAL calibration strategy

Calibration is a severe technical challenge for the operation of the CMS electromagnetic calorimeter. It can be seen as composed of a global component, giving the *absolute* energy scale, and a uniform single channel response, which is referred to as intercalibration.
The final goal of the calibration strategy is to achive the most accurate energy measurement for electrons and photons. Schematically, the reconstructed energy can be factorized into three terms:

$$E_{e,\gamma} = \mathcal{F} \times \sum_i G \times c_i \times A_i \tag{2.1}$$

where G is the global absolute scale; the function $\mathcal{F}$ is a correction function depending on the type of particle, its position, its momentum and on the clustering algorithm used. The c_i are the intercalibration coefficients and A_i are the signal amplitudes, which are summed over all

the clustered crystals.
The intercalibration at the start-up relies on laboratory measurements of the crystal light yield, on test beam precalibration of some supermodules and on the commissioning of the supermodules with cosmic rays [26].
After the assembly of the detector, the *in situ* calibration with physics events will be performed, exploiting different tools:

- The ϕ-symmetry of the deposited energy allows to rapidly improve the intercalibration within rings at constant η: with few tens of millions of minimum bias events, equivalent to about 10 hours of data taking under the assumption of 1 kHz of Level-1 trigger bandwidth, it is possible to approach the limit precision of 0.5% in the barrel region and 1-3% in the endcap region. The main limit on precision comes from non-uniformities and inhomogeneities in the tracker material.

- The measurement of the momentum of isolated electrons from $W \to e\nu$ will provide a useful tool to intercalibrate crystals, once the tracker is fully operational and well aligned [27, 28]. The calibration precision dependence as a function of η follows the tracker material budget distribution and it depends strongly on the available statistics; a precision better than 1.5% can be achieved in the barrel region with 5 fb^{-1}.

- The Z mass constrain in $Z \to ee$ decays is a further powerful mean for ECAL calibration. Many uses are envisaged, from the tuning of algorithmic corrections, to the intercalibration of ECAL regions (as a complement to the ϕ-symmetry method at the start-up), to the absolute scale calibration. The use of $Z \to e^+e^-$ for calibration purposes at startup are extensively described in the present work.

- The calibration using mass reconstruction of π^0, $\eta \to \gamma\gamma$ and $Z \to \mu\mu\gamma$ has also been studied [29].

  As far as $\pi^0 \to \gamma\gamma$ is concerned, several independent calibration algorithms have been developed and proved to give consistent results in CMS [30]; it has been shown that, assuming a low-luminosity scenario for the LHC (10^{33}cm^{-2}s^{-1} at $\sqrt{s} = 14$ TeV), this method allows to calibrate the majority of the ECAL Barrel calorimeter, up to $|\eta| < 1.4$, to at least a 1% (0.5%) precision after about 30 (130) hours of data-taking. The typical energies of photons from π^0 decay are around 5 - 10 GeV, but since the ECAL response has shown a linearity better than 1% (see Appendix A) the calibration constants from the π^0 method are also useful for the reconstruction of higher energy electrons.

  Though the speed of intercalibration with $\eta \to \gamma\gamma$ decays is expected to be much slower than intercalibration with $\pi^0 \to \gamma\gamma$ decays, it is going to be very valuable to have η calibration as a cross check [31]. In addition, the average energy of the photons from the η decays is about a factor of two higher than that of the photons from the selected π^0 decays, therefore the photon reconstruction from η decays will be less affected by the experimental systematics, for example, the containment variations versus different pseudorapidity regions of the ECAL barrel [32].

2.1.1 Barrel calibration status at startup

Between 2004 and 2006, all the crystals belonging to 11 of the 36 ECAL Barrel SuperModules have been calibrated using electron beams at CERN test facilities:

- 10 SuperModules (SM) at H4 test beam facility in 2004 and 2006 (calibration precision around 0.3%)
- 1 SM at H2 test beam facility in 2006 (calibration precision 1-2%)

In addition, all the 36 Barrel SuperModules have been intercalibrated using Light Yield (LY) measurements (precision 4.5 - 6%) and cosmic rays (precision 1.5 - 2.2%).

Combined calibration constants for Barrel crystals

For most of the SuperModules more than one calibration constant per crystal was available (i.e. Test Beam, cosmic rays and Light Yield). They have been combined producing one "optimal", combined intercalibration constant for each of the Barrel crystals.
The basic algorithm that defined "optimal" calibration constants has been:

1. If the calibration constant coming from the test beam is present, then:
   - if the relative uncertainty on this constant is less than 1%, then the test beam constant is chosen as "optimal"
   - if the relative uncertainty on this constant is greater than 1%, then a combination of the test beam constant and the cosmic-ray constant is calculated and chosen as "optimal"
2. If no constant from the test beam is present, then:
   - if the relative uncertainty on the constant from the cosmic rays calibration is less than 3%, then the constant coming from the cosmic rays calibration is chosen as "optimal"
   - if the relative uncertainty on the constant from the cosmic rays calibration is greater than 3%, then a combination of the Light Yield constant and the cosmic-ray constant is calculated and chosen as "optimal"

The outcome of this method is a set of "optimized", combined calibration coefficients whose precision throughout the ECAL Barrel is shown in Fig. 2.1 .

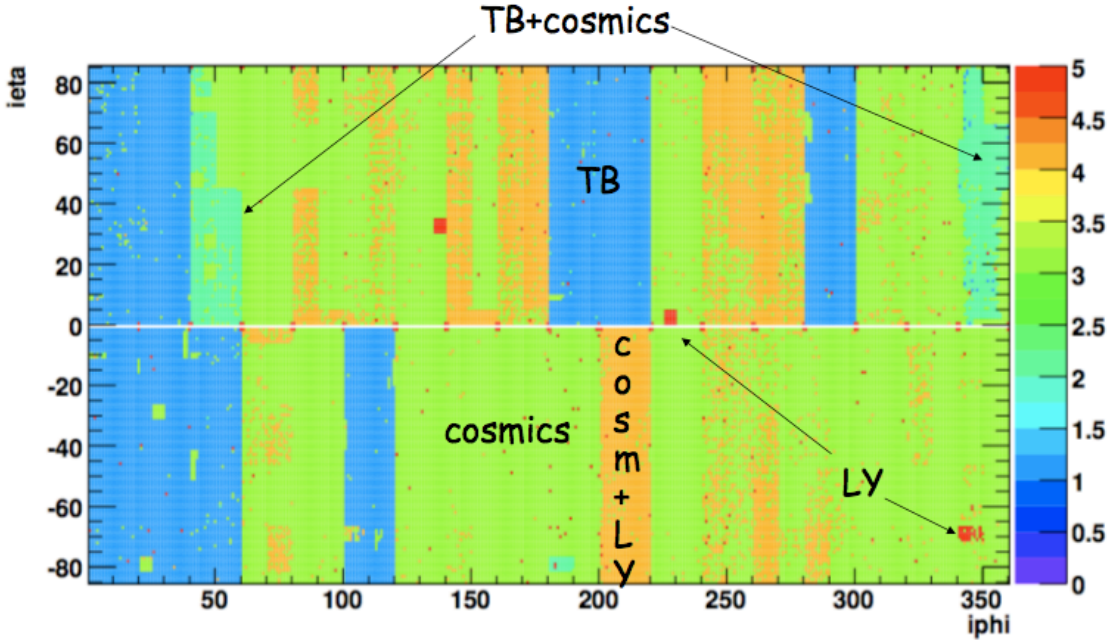


Figure 2.1: Calibration map of the ECAL Barrel: different colours indicate different precisions on the calibration constants.

2.2 $Z \rightarrow e^+e^-$ events at the LHC

The Z boson is primarily produced at the LHC via quark-antiquark fusion (see Fig. 2.2), where the quark can be a valence or a sea quark and the antiquark comes from the sea: the products of the Z decay are two oppositely-charge electrons revealed by the tracking system and the electromagnetic calorimeter.
Fig. 2.3(a) and 2.3(b) display a fully-simulated and reconstructed $Z \rightarrow e^+e^-$ event in the CMS detector: the two electrons (tracks and calorimetric deposits) are clearly visible.

The Z boson

The experimental strategy to select $Z \rightarrow e^+e^-$ events consists in choosing events with two electrons (the so-called "Drell-Yan" events with electrons); the two electrons can come from the competitive processes of γ^* and Z exchange and their interference. A "Z event" is defined as a dielectron event where the dielectron invariant mass is close to the Z mass value (around 91 GeV/c^2,), where there is a peak of the cross section due to the Z boson exchange.
The Z events shown in the present work have been generated using the PYTHIA event generator [33], and fully reconstructed through the official reconstruction framework of the CMS experiment.
Fig. 2.4 shows the Z boson mass reconstructed at the MonteCarlo level as well as the mass resolution, i.e. the difference between MonteCarlo mass and mass generated using dielectron invariant mass in fully simulated $Z \rightarrow e^+e^-$ events in the CMS detector; it shows that, in a

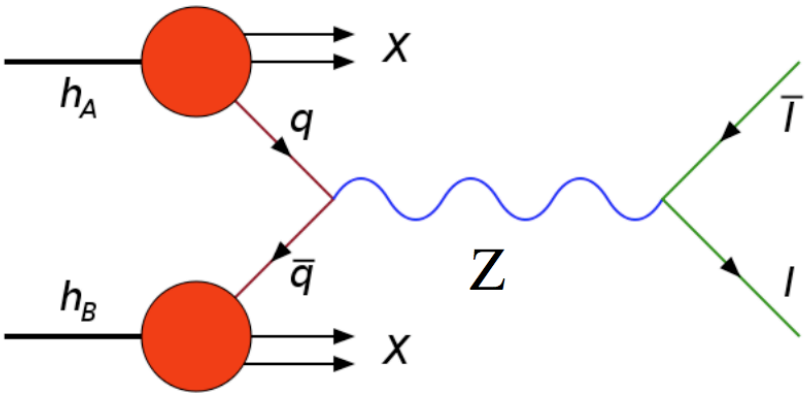


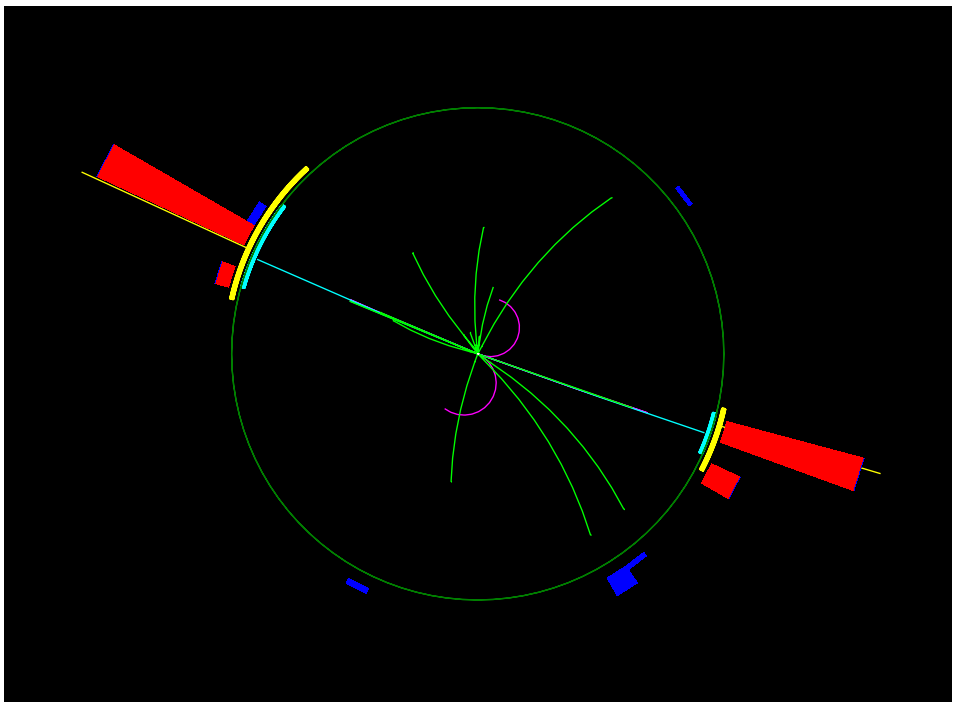
Figure 2.2: Feynman diagram of Z production and subsequent decay to leptons (Drell-Yan process)

perfectly calibrated CMS detector, the Z mass is reconstructed with a resolution around 1.3 GeV/c^2, and the peak value is very well reconstructed with a bias of around 55 MeV/c^2.
Fig. 2.4(a) shows the Z mass peak around 91 GeV/c^2 as well as lower and higher-invariant mass tails, where the contribution from γ^*-exchange becomes dominant.

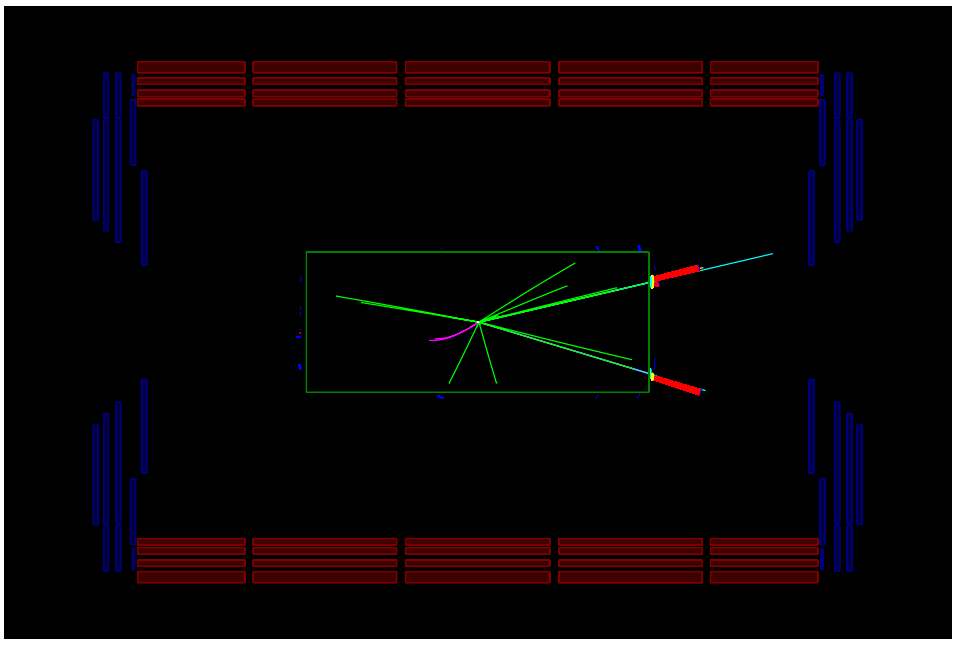
2.2.1 Electrons from Z decay

Fig. 2.7, 2.8 and 2.9 show some reconstructed quantities of interest for this work, relative to electrons in $Z \to e^+e^-$ events triggered using Single Electron HLT paths; electrons from the Z decay have a p_T spectrum with a characteristic Jacobian distribution around 45 GeV/c and are well distributed in pseudorapidity along the CMS detector.
In particular, Fig. 2.5 shows - with arbitrary normalization - the occupancy distribution of reconstructed electrons from Z decay throughout the ECAL surface: it is evident that the electron reconstruction efficiency drops off in correspondence of module-module cracks in the Barrel, and in correspondence of Barrel-Endcap transition region.
Around 50% of all $Z \to e^+e^-$ events have both Z electrons impinging onto the ECAL Barrel; around 40% have one electron in the Barrel and one in the Endcap, while in 10% of the events both Z electrons impinge onto the ECAL Endcaps.
It is interesting to observe (see Fig. 2.6) how the average separation angle between the two electrons from Z decay gets smaller as the Z boson is more boosted along the beam axis: when the Z boson is produced at rest, the average separation angle is 180°, i.e. the two electrons are emitted mainly back-to-back, while for a Z momentum of 100 GeV/c the average angle separation falls to around 80°.
The average energy resolution of Z electrons varies with $|\eta|$ as shown in Fig. 2.10 [1]: it has a

[1]Fig. 2.10 is produced as follows: at fixed η the distribution of E_{reco}/E_{MC} is fitted with a Gaussian, and the resulting Gaussian σ is reported. The same method has been followed for Fig. 2.11

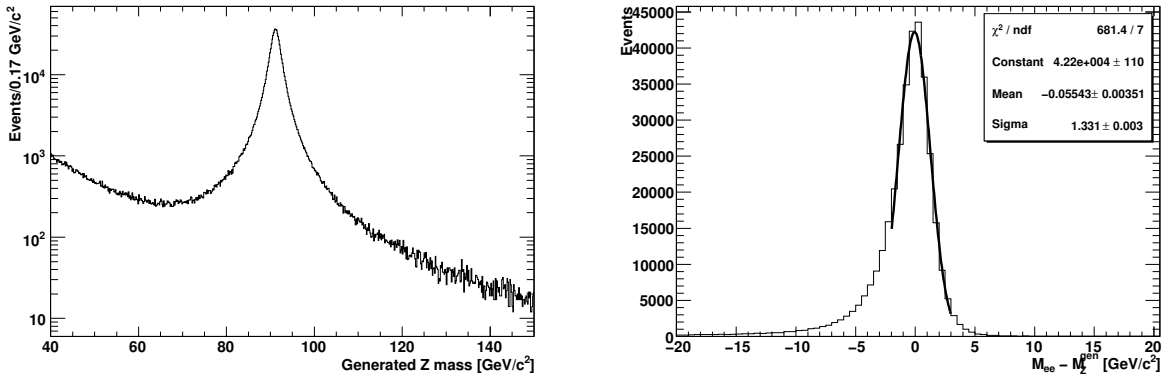


(a) $R - \phi$ view



(b) $R - z$ view

Figure 2.3: Views of a fully-simulated $Z \rightarrow e^+e^-$ event in the CMS detector. Red (blue) segments represent energy deposits in the ECAL (HCAL), their length being proportional to the deposited energy. Green lines represent tracks. Cyan segments indicate reconstructed electrons, composed by a track and an energy deposit.



(a) Generated Z mass (b) Mass resolution

Figure 2.4: The Z boson generated mass and Z mass resolution in CMS simulated events

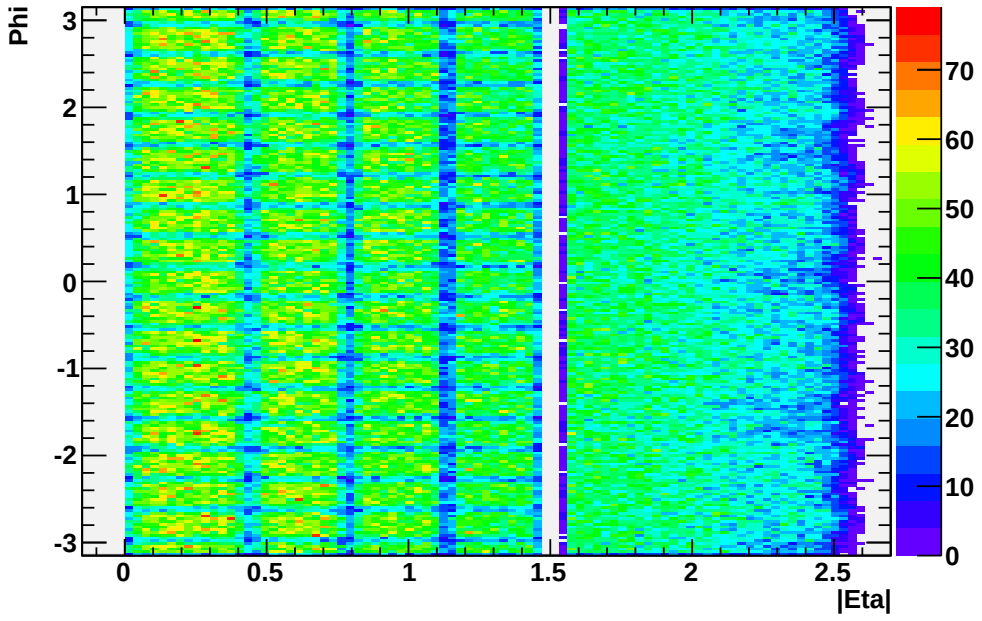


Figure 2.5: Occupancy distribution of reconstructed electrons from Z decay throughout the ECAL surface

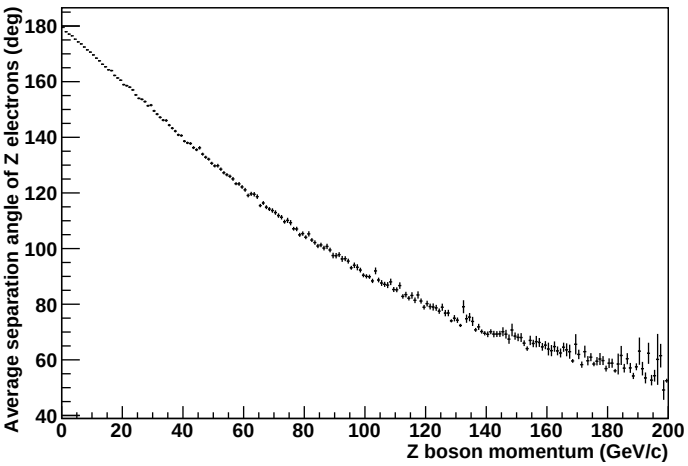
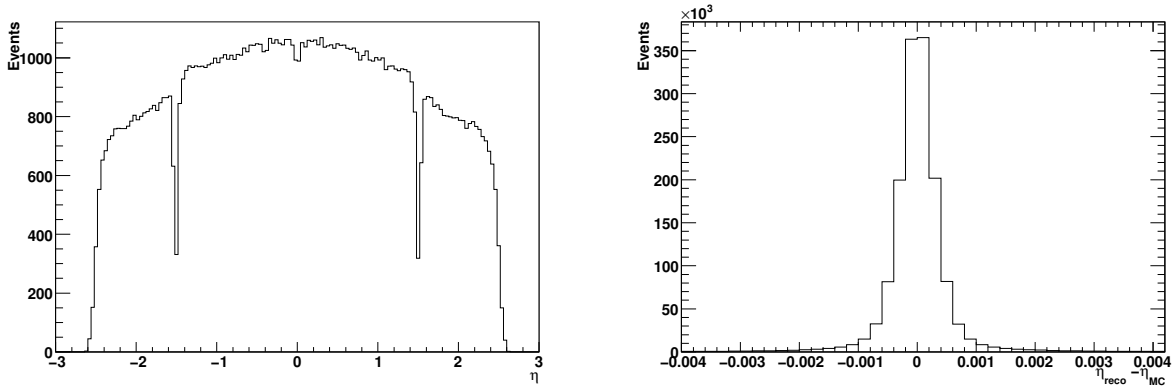


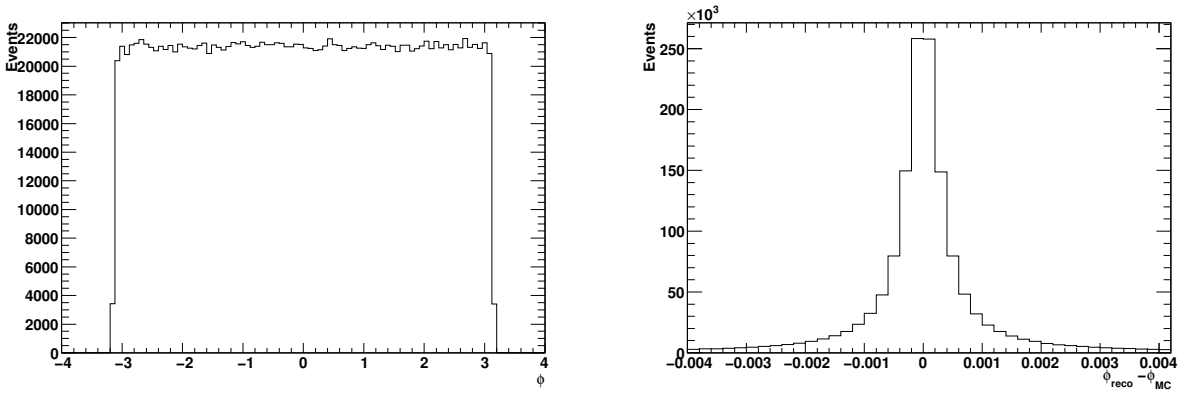
Figure 2.6: Average value of the separation angle between Z electrons as a function of the Z boson momentum (quantities at generator level)

minimum (around 1.5%) at $\eta \simeq 0$, and reaches a maximum value ($\simeq 5\%$) around the Barrel-Endcap transition crack, clearly reflecting the variation of the material budget in front of calorimeter.
The average p_T resolution (Fig. 2.11) rises with increasing $|\eta|$, with a minimum value of 4% and a maximum of around 30%; the Barrel-Endcap transition crack is clearly visible at $|\eta| \simeq 1.5$.



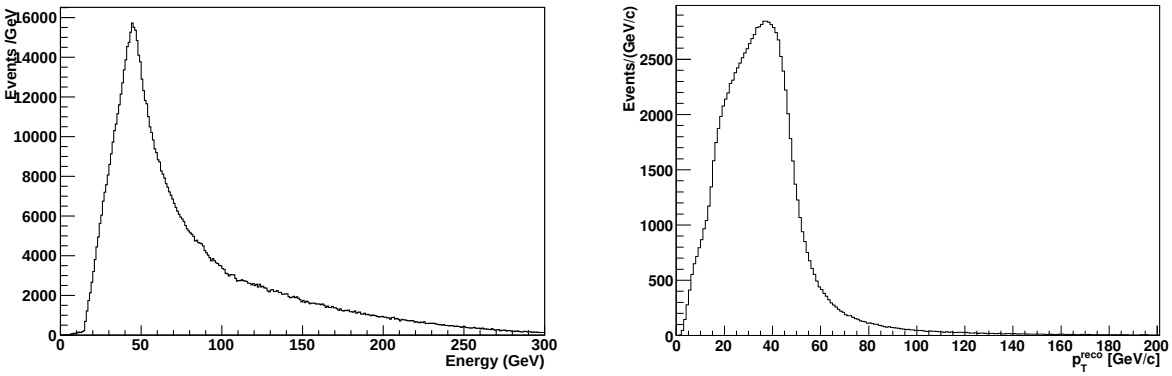
(a) Pseudorapidity (b) Pseudorapidity resolution

Figure 2.7: Pseudorapidity distribution of Z electrons and its resolution in CMS simulated events



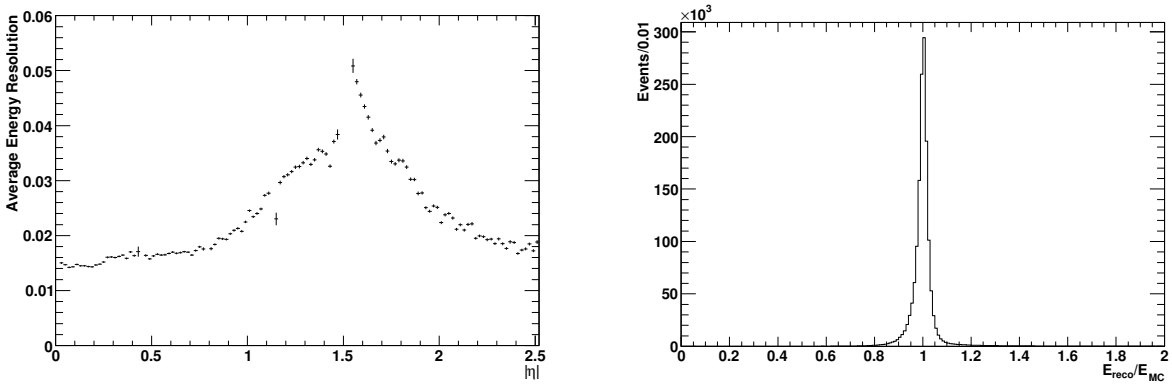
(a) Azimuthal angle (b) Azimuthal angle resolution

Figure 2.8: Azimuthal angle (ϕ) distribution of Z electrons and its resolution in CMS simulated events



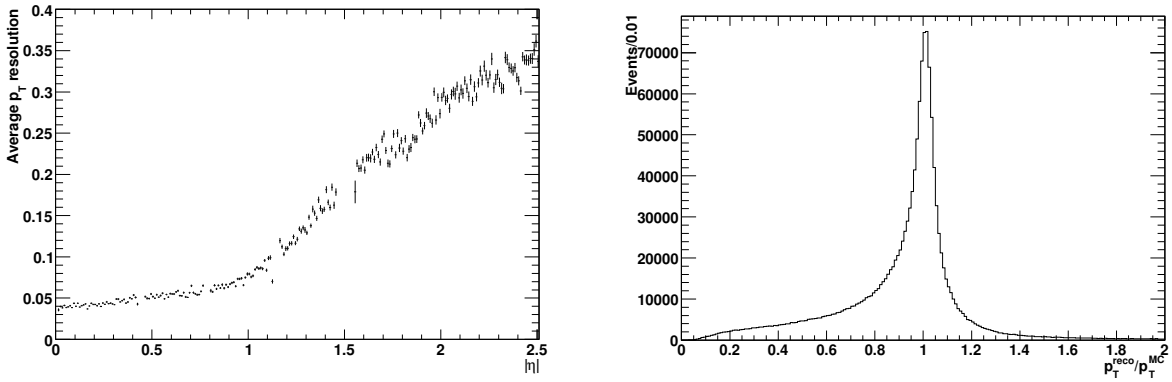
(a) Energy (b) Transverse momentum

Figure 2.9: Energy and transverse momentum of electrons in $Z \to e^+e^-$ HLT-selected events.



(a) Energy resolution vs $|\eta|$ (b) Energy resolution

Figure 2.10: Electron energy resolution



(a) p_T resolution vs $|\eta|$ (b) p_T resolution

Figure 2.11: Electron transverse momentum resolution

2.3 Evaluation of the level of detector calibration from the Z width

During the data taking of the CMS experiment, looking at the experimental width of the Z peak allows to extract the calibration status of the electromagnetic calorimeter. This procedure requires the use of MonteCarlo simulation: $Z \rightarrow e^+e^-$ simulated events are fully reconstructed using CMS algorithms, after insertion of an artificial miscalibration of the ECAL. Starting from real data it is possible to detect and correct miscalibration patterns occurring on a crystal-to-crystal basis, but also on rings of crystals having the same value of η ("η-rings"); in general it is possible to simulate any possible miscalibration pattern (by Modules, Supermodules etc.).
A simple choice for simulating a crystal-to-crystal miscalibration is to assign to the crystals a miscalibration constant extracted from a Gaussian distribution with mean 1 and rms σ; this type of miscalibration is reasonable if one thinks of the startup conditions of the ECAL Endcaps (while for the Barrel, no detectable Gaussian miscalibration is expected at the startup). Then the method proceeds as follows: at each value of miscalibration - i.e. at each value of Gaussian σ - the miscalibration constants are applied to the crystals before the reconstruction, then the full reconstruction is performed and the experimental width of the Z peak is extracted using a fitting procedure.
This provides a MonteCarlo-based curve of Γ_Z^{reco} as a function of σ; during data taking, this curve can be used so that, when a value of Γ_Z^{reco} is measured, it can be put in correspondence with a value of $\sigma = \sigma(\Gamma_Z^{reco})$ that is a "snapshot" of the calibration status of the electromagnetic calorimeter.
It is useful to analyze separately events with both Z electrons in the ECAL Barrel and events with both electrons in the ECAL Endcaps, because this allows to monitor the Barrel and Endcaps miscalibration status separately. The ECAL Barrel will start the data taking with a crystal-to-crystal resolution of 1-2%, which is too low for this method to have a sensitivity; on the other hand, ECAL Endcaps will be calibrated at 7-10% at startup so this method gets more promising for this part of the detector.
Fig. 2.12 shows how the ratio σ_Z/M_Z^{reco} changes with increasing Endcap miscalibration, when only $Z \rightarrow e^+e^-$ events with both electrons in the ECAL Endcaps have been used to build the Z peak. The experimental width of the Z boson is iteratively fitted with a Gaussian function in the range $[-1.5, +2]\sigma$ around the peak.
The points in Fig. 2.12 are fitted with the following function:

$$\frac{\sigma_M}{M} = \sqrt{\sigma_0^2 + (\kappa \cdot \mathrm{misc})^2} \tag{2.2}$$

The first parameter σ_0 contains, in quadrature, the intrinsic width of the Z boson plus a term due to the intrinsic energy resolution and the resolution of the angle between the two electrons:

$$\sigma_0 = \left(\frac{\Gamma_Z}{M_Z}\right) \oplus \sigma(E_1, E_2, \theta_{12}, \sigma_{E_1}, \sigma_{E_2}, \sigma_{\theta_{12}}) \tag{2.3}$$

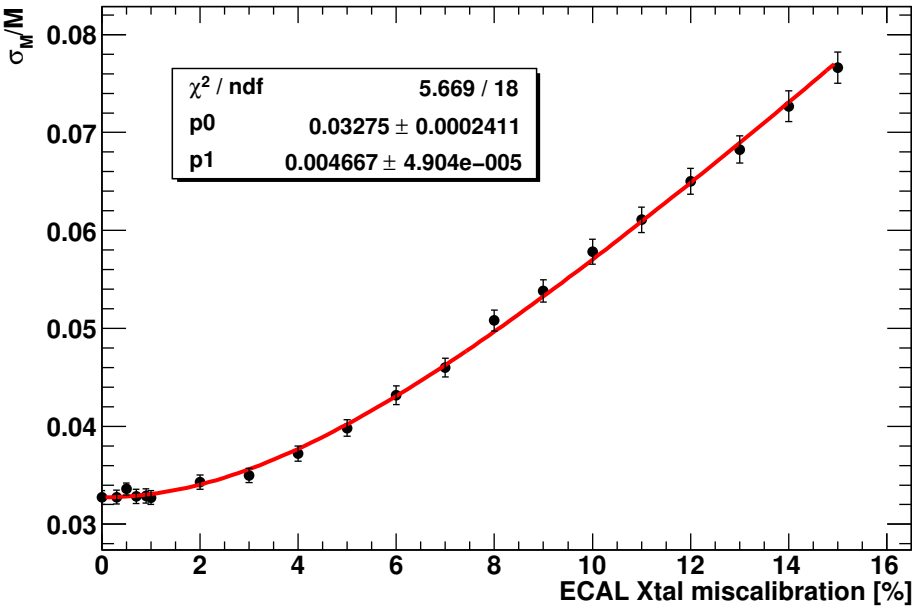


Figure 2.12: Experimental resolution of the Z boson mass vs crystal-to-crystal miscalibration; plot refers to events with both Z electrons in ECAL Endcaps.

from the fit in Fig. 2.12, $\sigma_0 = (3.27 \pm 0.02)\%$.
The parameter κ - once multiplied by 100 because of the format of the x-axis in Fig. 2.12 - is approximately equal to $1/\sqrt{2}$ times the average fraction of the electron energy carried by its most energetic crystal:

$$\kappa \simeq \frac{1}{\sqrt{2}} < \frac{E_1}{E_{SC}} > \tag{2.4}$$

where the average is taken over all the selected electron in the ECAL Endcaps; the fit gives $\kappa = (0.467 \pm 0.005)$ which corresponds to an average fraction of (0.660 ± 0.007).
The reason for eq. 2.4 is the following: each electron deposits its energy in a large number of (miscalibrated) crystals, but most of the energy is deposited in only one crystal. The electron energy can be then written as:

$$E_{ele} = (1 + \epsilon_1)E_1 + \sum_i (1 + \epsilon_i)E_i \tag{2.5}$$

where E_i are the energy deposited in each of the crystals. If the cluster is made up of many crystal, it is likely to find, in the sum of eq. 2.6, pairs of terms of the kind $\epsilon_k E_k$ that cancel each other. So eq. 2.6 becomes:

$$E_{ele} \simeq (1 + \epsilon_1)E_1 + \sum_i E_i = E + \epsilon_1 E_1 = (1 + f_1\epsilon_1)E \tag{2.6}$$

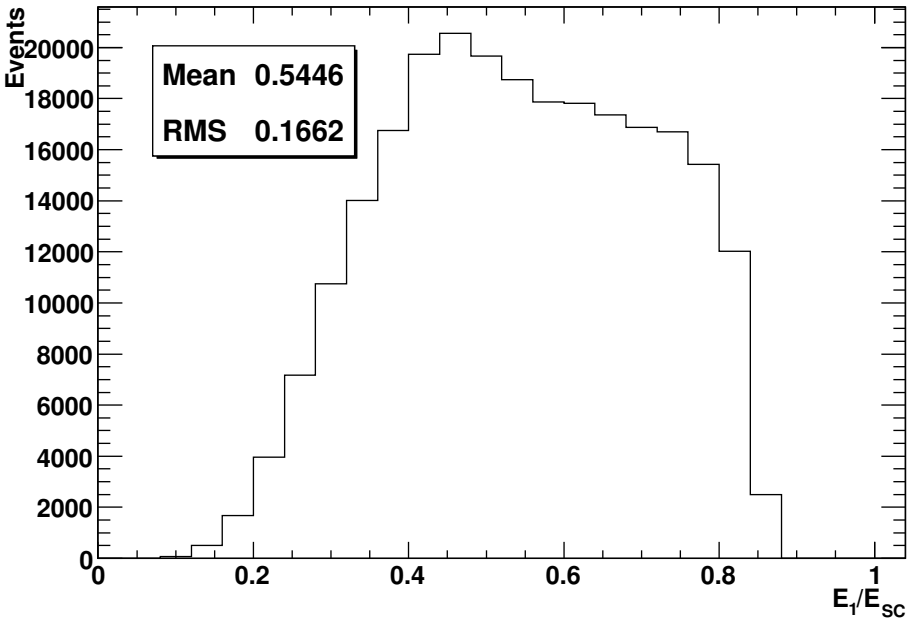


Figure 2.13: E_1/E_{SC} distribution for Z electrons in ECAL Endcaps.

where E is the electron energy in the absence of crystal miscalibration and $f_1 = E_1/E_{SC}$. Eq. 2.6 shows that the crystal miscalibration, when propagated to the electrons, shall be approximately diluted by a factor f_1. The parameter κ represents this diluition factor; the prefactor $1/\sqrt{2}$ comes in because there are two (miscalibrated) electrons in each event.
The distribution of E_1/E_{SC} of electrons in the ECAL Endcaps is plotted in Fig. 2.13; the average f_1 from this plot is around 0.55 which is a bit lower than the fit value, indicating that on average not only the hottest crystal of the cluster contributes to the electron energy miscalibration.

Influence of the Tracker miscalibration

It can be argued that what determines the experimental width of the Z peak is not only electromagnetic calorimeter calibration, but also the tracker calibration, since the electron-positron invariant mass is built using at least the θ and ϕ angle coming from the tracker information. More specifically, the resolution on the dielectron invariant mass is:

$$\frac{\sigma_M}{M} = \frac{1}{2}\left[\frac{\sigma_{E_1}}{E_1} \oplus \frac{\sigma_{E_2}}{E_2} \oplus \frac{\sigma_{\theta_{12}}}{\tan(\theta_{12}/2)}\right] \tag{2.7}$$

Fig. 2.14 shows, however, that the angular term in Eq. 2.8 is negligible with respect to the energy terms: for each reconstructed $Z \to e^+e^-$ event, the differences $\Delta E_{1,2}$ between the reconstructed and the generated ($E_{1,2}^{true}$) energy of the two electrons, as well as $\Delta\theta_{12}$, i.e. the

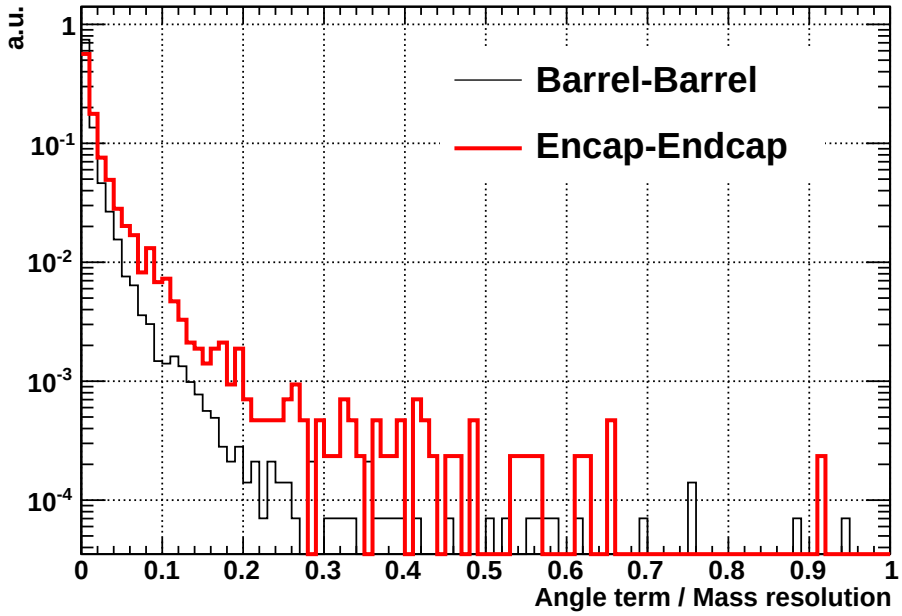


Figure 2.14: Percentage of Z mass resolution given by the angular term in Eq. 2.8 (see text for details)

difference between the reconstructed and the generator-level (θ_{12}^{true}) angle between the two electrons are calculated. Then, the histogram is filled with the ratio:

$$\frac{\frac{\Delta\theta_{12}}{\tan(\theta_{12}^{true}/2)}}{\left[\frac{\Delta E_1}{E_1^{true}} \oplus \frac{\Delta E_2}{E_2^{true}} \oplus \frac{\Delta\theta_{12}}{\tan(\theta_{12}^{true}/2)}\right]} \tag{2.8}$$

2.4 Measurement of the ECAL absolute scale

The need for accurate experimental measurements of particle masses does not require much justification, given the progress made in particle physics over the past decades through such measurements. Nevertheless, a few examples of what might be achieved with LHC and CMS are listed below and described in some detail.
If supersymmetry (SUSY) were to be discovered at the LHC, the more accurate the measurement of the new particles, the tighter the constraints on the fundamental parameters of the underlying SUSY model will be.
Very precise measurements of the masses of the W boson and of the top quark, beyond those which will have been (or will be) achieved at the Tevatron and at LEP2, will provide further constraints on the Standard Model.

In case a Higgs boson were to be discovered, an accurate measurement of the Higgs boson mass will provide strong constraints on the underlying model; the power of a constraint at this level has been demonstrated in the context of the global fits to the parameters of minimal supergravity models.
This section examines at what level of accuracy it is possible to set the absolute scale of the CMS electromagnetic calorimeter using $Z \to e^+e^-$ events as a a function of the available data statistics.
At the LHC startup, ECAL absolute scale will be known with an accuracy of around 1% for the Barrel, as a result of the calibration at test beam facilities and with cosmic muons [34].
In $Z \to e^+e^-$ events, the absolute scale is defined as the ratio between the value of the Z peak, as reconstructed in dielectron events inside the CMS detector, and the peak value which is expected in a perfectly-calibrated detector, i.e. the LEP value of M_Z in first approximation. Neglecting biases due to the tracking system, if the above ratio equals one, it means that the CMS electromagnetic calorimeter scale is properly set.
The reconstructed Z lineshape can be fitted using an asymmetric Gaussian function:

$$f(x) = A \exp\left[-\frac{(x-\mu)^2}{2\sigma^2 + \alpha(x-\mu)^2}\right] \tag{2.9}$$

that has 6 free parameters because σ and α are fitted separately for $x < \mu$ and $x > \mu$. This function allows to extract the signal yield and the peak position with great accuracy, even if it does not give a direct and intuitive access to the width of the Z boson.
Fig. 2.15 shows the goodness of such a fit to the invariant mass spectrum from a sample of fully-reconstructed $Z \to e^+e^-$ events in the CMS detector.
It is desirable to extract two different absolute scale, one for the ECAL Barrel and one for each the ECAL Endcap (the two Endcaps will be trated as one unique detector in the following, just for the sake of simplicity). Therefore, $Z \to e^+e^-$ events are divided into events having both electrons in the Barrel and events having both electrons in the Endcap.
Indeed, if both electrons are in the same subdetector, then the ratio between the fitted peak and M_Z actually gives the absolute scale for that subdetector, i.e. the number for which all the calibration constants of that subdetector must be multiplied in order to obtain the correct reconstructed scale:

$$\frac{M_{ee}^{BB,EE}}{M_Z} = a_{B,E} \tag{2.10}$$

In events with one electron in the Barrel and one electron in Endcap, on the other hand, the above ratio gives a folded effect, i.e. the square root of the product of the two multiplicative factor related to each subdetector:

$$\frac{M_{ee}^{EB}}{M_Z} = \sqrt{a_B a_E} \tag{2.11}$$

In order to evaluate with what *precision* it is possible to extract the calorimeter absolute scale - given a fixed data statistics - the data corresponding to the statistics is divided into separate (and therefore, independent) subsamples, and for each of these subsamples:

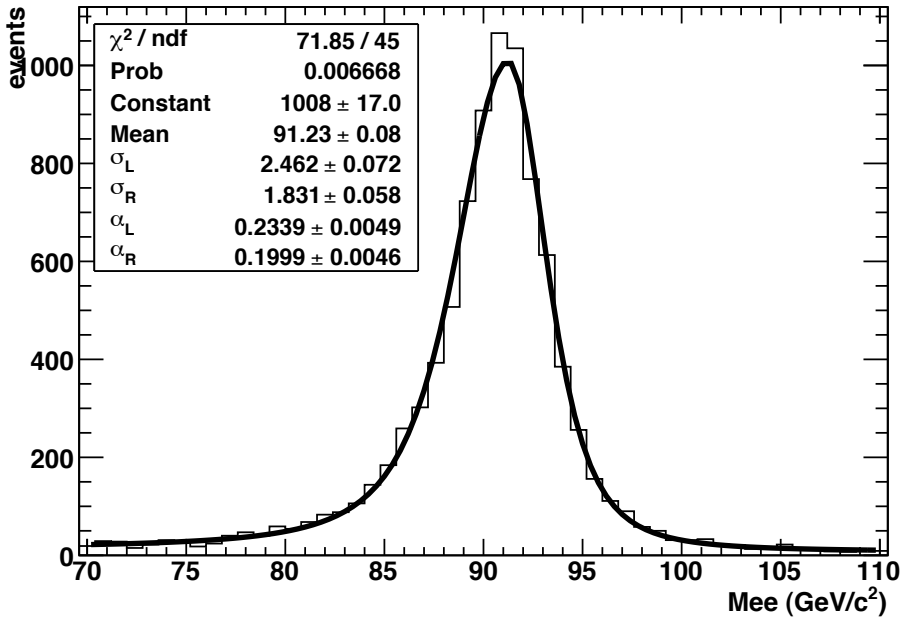


Figure 2.15: Reconstructed dielectron spectrum fitted with an asymmetric Gaussian lineshape

- the dielectron mass spectrum is fitted
- the fit error on the peak is divided by the fitted peak value - so to give the relative precision on the absolute scale and this ratio is put into a histogram
- the histogram mean gives the precision on the absolute scale achievable with that particular statistics, while the RMS of the distribution gives the corresponding uncertainty

It is possible to repeat the above procedure for different data statistics, so to obtain a curve of accuracy in the measurement of absolute scale as a function of integrated luminosity (2.16). This curve is fitted with a function $f(x) = p0 + p1/\sqrt{x}$, that expresses the Poisson-scaling with the number of events, and shows how it is possible to set the absolute scale of the CMS electromagnetic calorimeter at the level of 0.25% with as much as 50 $\text{pb}^{-1}$of available data. In addition to this, it is possible to look at the *bias* in the reconstruction of the calorimeter absolute scale; in this case the procedure is similar, and, for each subsample of a given statistics:

- the dielectron mass spectrum is fitted
- the fit error on the peak is divided by the nominal mass of the Z boson - and this ratio is put into a histogram
- the mean of the histogram gives the precision on the absolute scale achievable with that particular statistics, while the RMS of the distribution gives the corresponding uncertainty

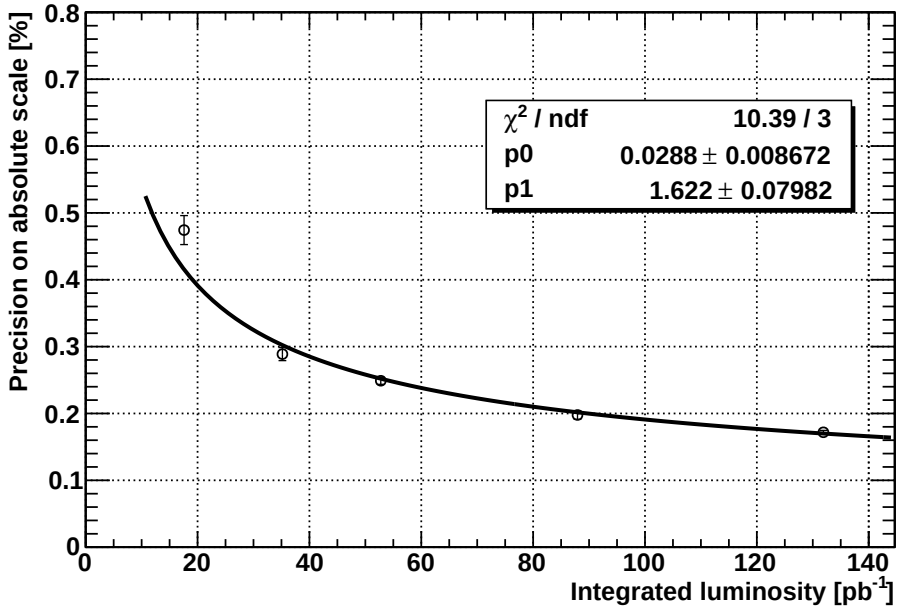


Figure 2.16: Precision on the measurement of absolute scale as a function of integrated luminosity

Fig. 2.17 inspects whether a bias in the reconstruction of the absolute scale is present as a function of the available data statistics. No significant deviations from the input absolute scale is observed.
A number of systematic uncertainties on the Z peak position in a perfectly calibrated CMS detector will be discussed in detail in Section 2.5.2.

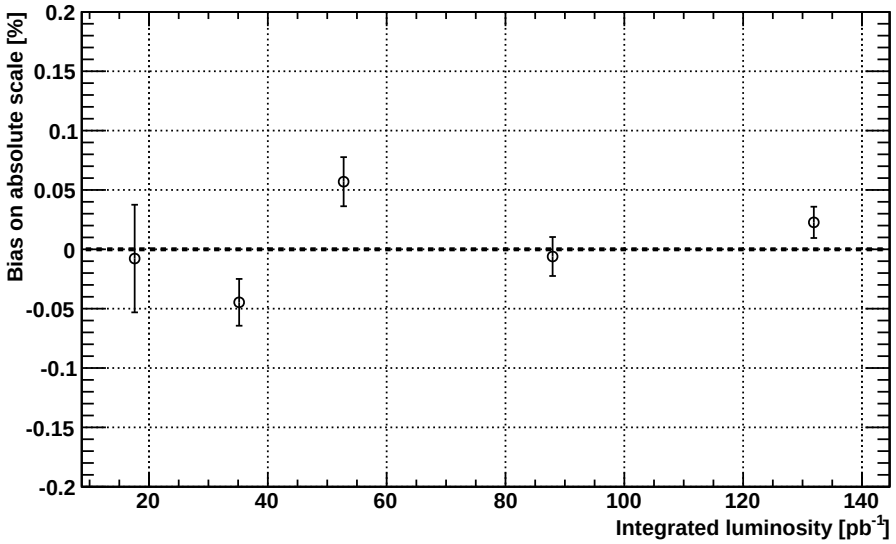


Figure 2.17: Bias in the measurement of absolute scale as a function of integrated luminosity

2.5 An iterative method to calibrate the CMS electromagnetic calorimeter

The present chapter describes an experimental method that uses $Z \to e^+e^-$ events to calibrate the CMS electromagnetic calorimeter, exploiting the very precise knowledge of the Z boson mass coming from LEP experiments. The same method can in principle be applied to any other dielectron resonance, i.e. J/ψ or Υ; $Z \to e^+e^-$ events provide calibration constants using electrons with energy of the order of 100 GeV (see energy distribution in Fig. 2.9), and the use of such lower-mass resonances would validate these constants also at electron energies of 5-10 GeV.

2.5.1 Description of the method

In order to describe the method it is useful to consider at first a single electron that, in an event labelled i, hits a certain number of regions (indexed j) of the calorimeter. These region can be single crystals, η-rings or other kinds of region. If region j is miscalibrated of an amount ϵ_j, the following holds for a single region:

$$E_{meas}^{region\ j} = (1 + \epsilon_j) E_{true}^{j} \tag{2.12}$$

so for the whole electron energy, spreaded over a certain number of regions, can be expressed as:

$$E_{ele,meas}^{event\ i} = \sum_{j \in ele} (1 + \epsilon_j) E_{true}^{j} = E_{ele,true}^{i} + \sum_{j \in ele} \epsilon_j E_{true}^{j} \tag{2.13}$$

The true electron energy is built as the sum over the energies contained in the regions; the single region has a weight given by:

$$w^j_{true} = \frac{E^j_{true}}{E_{ele,true}} = \frac{E^j_{meas}/(1+\epsilon_j)}{E_{ele,meas}/(1+\sum_{j\ \in\ ele}\epsilon_j)} = w^j_{meas} + O(\epsilon) \tag{2.14}$$

where w^j_{true} is obviously unknown but w^j_{meas} can be measured; so at lowest order in ϵ_j it holds:

$$E^{event\ i}_{ele,meas} = E^i_{ele,true}(1 + \sum_{j\ \in\ ele}\epsilon^j w^j_{meas}) \tag{2.15}$$

For events with two electrons (which in the following are always assumed to be massless):

$$M^i_{inv} = \sqrt{2E^i_{ele1}E^i_{ele2}(1-\cos\theta_{12})} = M_Z\sqrt{1 + \sum_{j\ \in\ ele1,ele2}\epsilon^j w^j_{meas} + O(\epsilon^2)} \tag{2.16}$$

where M_Z denotes the true value of mass of the Z boson in event i. Algebrical manipulation yields:

$$\left(\frac{M^i_{inv}}{M_Z}\right)^2 - 1 = \sum_{j\ \in\ ele1,ele2}\epsilon^j w^j_{meas} + O(\epsilon^2) \tag{2.17}$$

and being $\sum_{j\ \in\ ele1,ele2} w^j_{meas} = 2$ the following noticeable equality is obtained:

$$\frac{1}{2}\left[\left(\frac{M^i_{inv}}{M_Z}\right)^2 - 1\right] = \frac{\sum_{j\ \in\ ele1,ele2}\epsilon^j w^j_{meas}}{\sum_{j\ \in\ ele1,ele2} w^j_{meas}} \tag{2.18}$$

The quantity on the right side represents the weighted mean value of the miscalibration of the calorimeter regions around the two electrons, where the weight is given - at first order - by the fraction of the electron energy that is carried by the region itself. This is an information not directly related to any particular region, being the single region information folded in an average effect.
A practical estimator for ϵ_j is obtained by creating a histogram for each region j and filling this histogram, for each event i, with the quantity:

$$\frac{1}{2}\left[\left(\frac{M^i_{inv}}{M_Z}\right)^2 - 1\right] \tag{2.19}$$

(where M_Z is the well-known experimental value of the Z mass peak [35]) weighted with:

$$w^j_{meas} = \frac{E^j_{meas}}{E_{ele,meas}}$$

The mean value of the distribution gives an estimation of ϵ_k; in order to weaken the sensitivity to tails, an asymmetric Gaussian fit (Eq. 2.9) is performed on the distribution, and the resulting peak is considered instead of the mean value. Fig. 2.18 shows an example of fitted distribution of ϵ quantities for one η-ring.

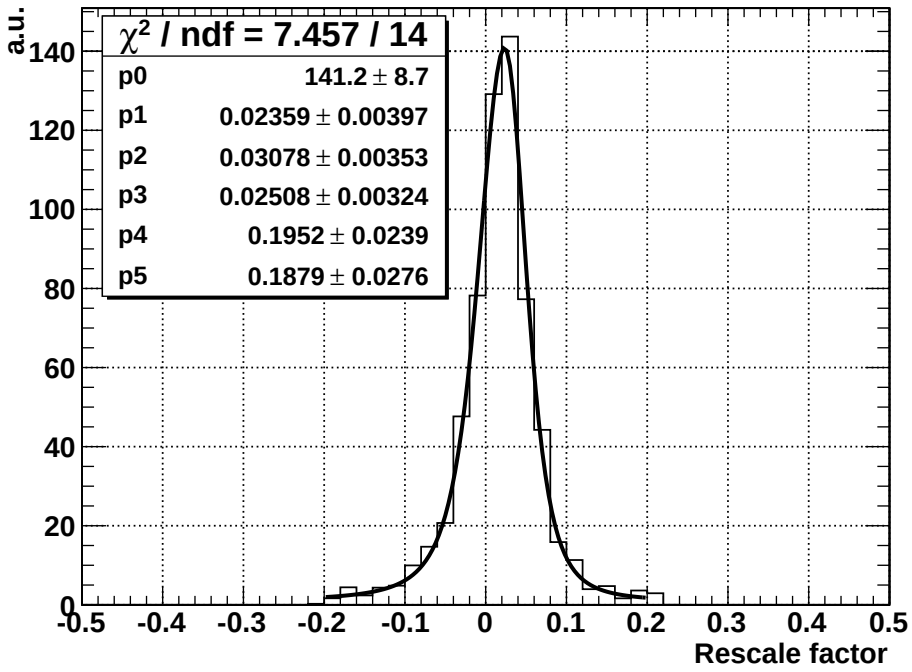
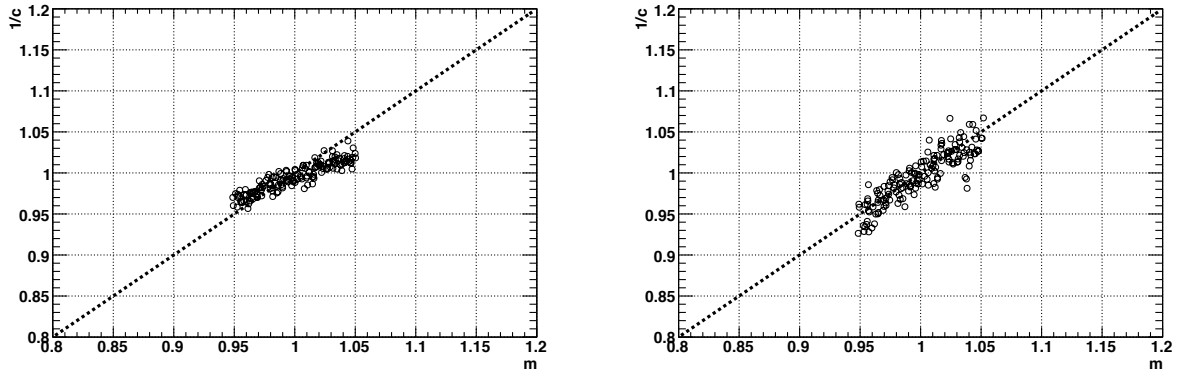


Figure 2.18: Distribution of the quantity ϵ (as defined in eq. 2.19) for one η-ring, with superimposed fit.



(a) After one iteration

(b) At convergence

Figure 2.19: Relationship between miscalibration constants and (inverse of) recalibration constants given out by the algorithm

It can be interesting to focus on the reason why is it necessary to apply the method iteratively: the origin of this iterative nature is the fact that only terms linear in the miscalibrations ϵ_j are retained in the calculations.
As mentioned above, the histogram related to a region j brings the effect of the miscalibration of the region j, convoluted with all the miscalibration constants of all the other regions: as a first approximation, the peak value of each histogram is:

$$\text{peak}_j = \sqrt{c_j < c_{i \neq j} >} \tag{2.20}$$

where $< c_{i \neq j} >$ is the weighted average of all the miscalibration constants of the other regions. Therefore, after one iteration the correction is diluted and the recalibration constants is approximately $c_j^{-1/2}$ instead of c_j^{-1}. Iteration after iteration, c_j^{-1} approaches 1 (smaller and smaller recalibrations are needed) and so the difference between $c_j^{-1/2}$ and c_j^{-1} becomes negligible.
This effect is clearly displayed by Fig. 2.19: after a sufficient number of iterations, r_j get equal to m_j, while after only one iteration of the algorithm the miscalibration constants are only partially recovered by the recalibration constants given by the algorithm.
Fig. 2.20 provides both a clear explanation of the iterative nature of the method, and a MonteCarlo validation of the method itself: an artificial miscalibration coefficient of 1.03 is applied to all Barrel crystals, then the algorithm is applied. After one iteration, part of the recalibration coefficient is applied to the Endcap crystals too, so that recalibration constants less than 1 are found for all the ECAL crystals; after 10 iterations, Endcap crystals are correctly recalibrated with coefficients close to 1 while Barrel crystals are recalibrated with coefficients around $1/1.03 \simeq 0.97$.
This happens because of the presence of $Z \rightarrow e^+e^-$ events with one electron in the Barrel and one electron in an Endcap: for these events, the recalibration coefficient is around $\sqrt{0.97} \simeq 0.985$, and is attributed both to the involved Barrel region (whose miscalibration is only

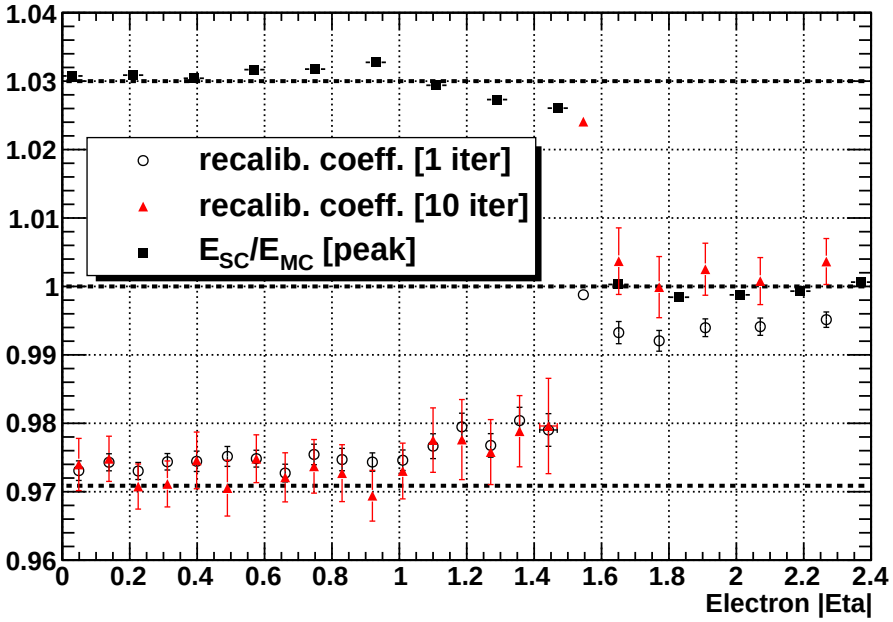


Figure 2.20: Progressive convergence of the algorithm: Barrel crystals are scaled down by a factor 0.97 while Endcap crystals have not been miscalibrated.

partially recovered) and to the involved Endcap region (whose energy response is artificially lowered).
Fig. 2.20 provides a MonteCarlo validation of the algorithm, by showing the peak of the distribution of E_{SC}/E_{MC} in slices of electron $|\eta|$; the E_{SC}/E_{MC} distribution are nicely fitted using an asymmetric Gaussian function.
Fig. 2.21 shows the recalibration coefficients, for 1 iteration and at convergence, obtained when using the so-called "raw" SuperCluster energy of the electron, i.e. just the sum of the energy deposits from the crystals of the calorimetric cluster associated with the electron object.
At convergence, the energy resolution points and the recalibration coefficients are symmetric with respect to the horizontal line at 1; after one iteration, this effect is clearly diluted.

2.5.2 Considerations on the Z peak position

The mass of the Z boson has been measured with great accuracy by the four LEP collaborations; its latest estimate is:

$$M_Z = (91.1876 \pm 0.0021)\ \mathrm{GeV/c^2} \tag{2.21}$$

Is it now worth questioning whether this is the value of M_Z that must be inserted into eq. 2.18 in order to obtain the correct calibration constant for the ECAL regions, i.e. what would

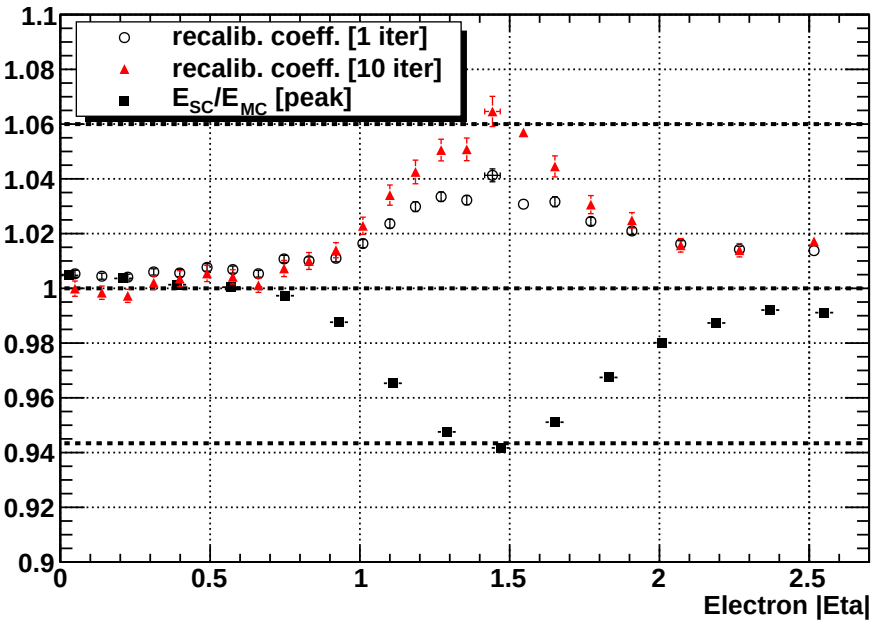


Figure 2.21: Progressive convergence of the algorithm: in this case, no artificial miscalibration constants are applied to the crystals, but all clustering corrections are neglected, and the electron energy is just computed as the sum of the energy deposits of the SuperCluster crystals ("raw energy").

the peak value of the Z mass be in a perfectly-calibrated CMS detector.
When reconstructing the Z lineshape using dielectron events, the peak value in eq. 2.21 can be altered by at least two phenomena:

- *outer* Bremsstrahlung: electrons from Z decay lose energy through emission of photon induced by interaction with the detector material
- *inner* Bremsstrahlung: electrons from Z decay lose energy through emission of photon before traversing the detector material

Both processes involve the emission of a photon; there is however a major difference.
Events with relevant *outer* Bremsstrahlung can be discarded by requiring a small value of f_{brem}; events with large *inner* Bremsstrahlung cannot be easily identified since the photon is radiated at the interaction vertex, so that:

1. the photon impinges onto the ECAL far away from the Z-electrons, which are bent along ϕ by the magnetic field;
2. f_{brem} is not useful because also with large *inner* Bremsstrahlung it is possible to have small f_{brem}.

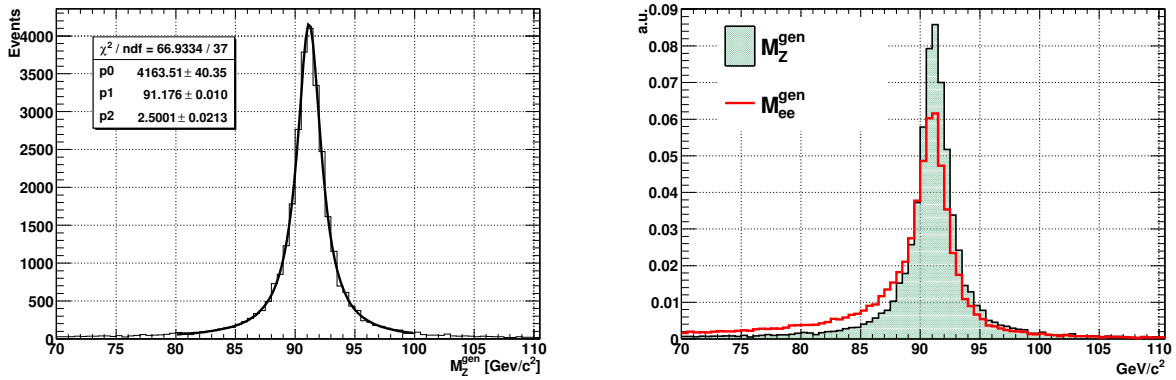
In order to reconstruct properly the events with *inner* Bremsstrahlung, it would be necessary to include in the analysis events of the kind $Z \to ee\gamma$ (or the probably cleaner $Z \to \mu\mu\gamma$); this addition can be ignored in the present analysis, only if it is shown that the *inner* Bremsstrahlung phenomenon does not significantly affect the position of the reconstructed Z peak.
The following paragraph analyze these issues: it is important to notice that it is necessary to rely on the way the MonteCarlo simulates both FSR/inner Bremsstrahlung and interaction of electrons with the detector material in order to draw conclusions on the role of inner/outer Bremsstrahlung.

Inner Bremsstrahlung

In order to show the effect of *inner* Bremsstrahlung it is not necessary to examine reconstructed events since the phenomenon is evident at the generator level: Fig. 2.22 shows the mass distribution of the generated Z bosons compared to the invariant mass distribution of the two MonteCarlo electrons from the Z decay.
The mass distribution of the generated Z bosons is very well fitted by a relativistic Breit-Wigner function:

$$f(M_{ee}; M_Z, \Gamma_Z) \propto \frac{1}{(M_{ee}^2 - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \tag{2.22}$$

while the dielectron mass shows a distorted shape with some low-mass tail enrichment, because of the inner Bremsstrahlung (the photon four-momentum is not included in the invariant mass calculation).



(a) Relativistic Breit-Wigner fit

(b) Generator level: Z mass and dielectron invariant mass

Figure 2.22: Mass distribution of the generated Z bosons compared to the invariant mass distribution of the two MonteCarlo electrons from the Z decay.

Fig. 2.23 shows the event-by-event difference between the generated Z mass and the generated-electron invariant mass; it can be noticed that:

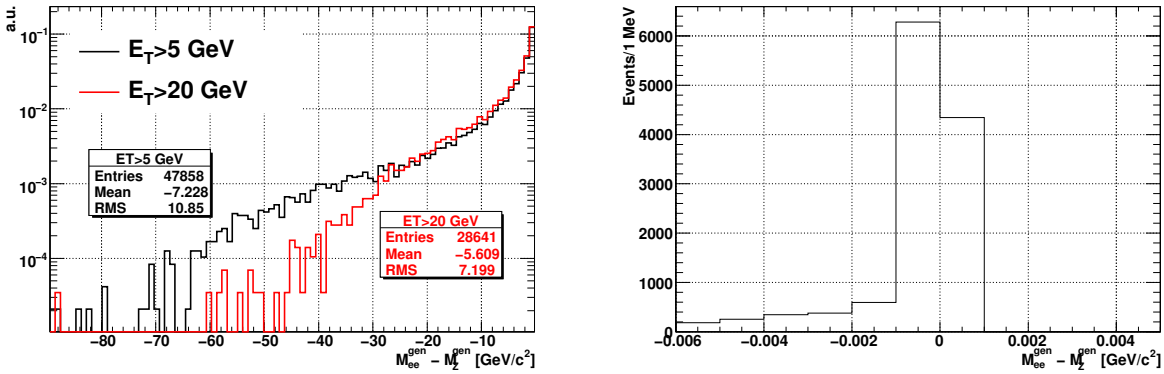
- the peak does not show a significant shift;
- there appears a significant tail on the left of the $M_{ee}^{gen} - M_Z^{gen}$ distribution; this tail gets larger as the electron-E_T cut is lowered. Lowering the electron-E_T cut allows to select events where electrons are soft because they radiated a hard inner-Brem photon: the cut-dependent behaviour appears therefore as expected.

Outer Bremsstrahlung

The effect of the outer Bremsstrahlung on the Z peak position can be investigated using reconstructed events, by choosing whether/how the Z peak moves as the outer Bremsstrahlung cut on the reconstructed Z-electrons is varied.
The amount of outer Bremsstrahlung can be experimentally estimated by comparing the electron momentum reconstructed at the inner tracker layer (p_{in}) and the electron momentum at the outermost tracker layer, just before the ECAL (p_{out}): the quantity

$$f_{brem} \equiv \frac{p_{in} - p_{out}}{p_{in}}$$

can then be used, so that $f_{brem} = 0$ corresponds to no outer Bremsstrahlung, and $f_{brem} = 1$ corresponds to maximum outer Bremsstrahlung. The quantity f_{brem} represents the fraction of the initial electron momentum subtracted by outer Bremsstrahlung photons as the electron moves across the tracker material.



(a) Overall spectrum

(b) Zoom on peak

Figure 2.23: Difference between generated Z mass and Z-electron invariant mass (generator level)

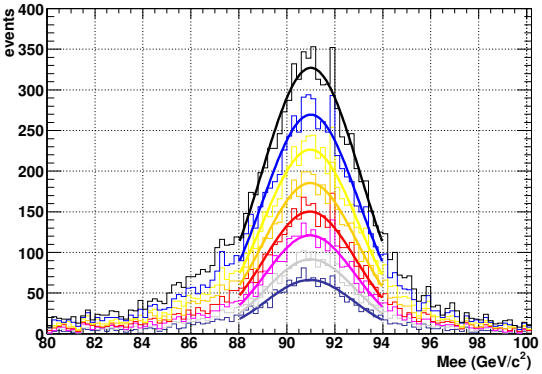
Fig. 2.24 shows how the Z peak position varies as the f_{brem} of the electrons is varied from 0.2 to 1 in steps of 0.1; the selection is the same for both electrons. The peak value tends to increase slightly with reduced outer Bremsstrahlung - as expected - but the fit uncertainty is rather significant and an extrapolation to $f_{brem} = 0$ is not straightforward. This suggest to introduce a moderate f_{brem} cut in the selection of Z events to use for ECAL calibration.

2.5.3 Systematics related to the Parton Distribution Functions

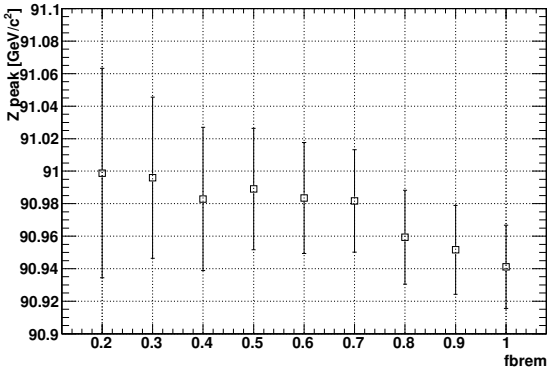
It has been shown by studies inside the CMS collaboration [36] that using different Parton Distribution Functions in the simulation of Z decay to leptons has a very small impact on the position of the Z (i.e. dilepton) mass peak: therefore, the related systematic uncertainty will not be considered in the present work.

2.5.4 Systematics related to event kinematics

In addition to the "global" value of M_Z to be used in the calibration algorithm, there is another systematic effect that is, in principle, even more important: the Z-peak value to use could, in principle, be a function of the event kinematics. If this is the case then it would be necessary, in the calibration algorithm, to use a denominator value M_Z that changes on an event-to-event basis. Using a unique value for M_Z brings to a systematic uncertainty that should then be taking into account.
In the first place, changing the minimum E_T threshold of the Z electrons brings to a bias in the Z lineshape: a higher threshold brings a preference towards higher Z masses. In order to estimate the importance of this effect, the reconstructed Z peak position has been plotted against the electron E_T threshold (applied symmetrically on both Z electrons). The plot in Fig. 2.25 shows that, in a range of reasonable CMS-like cuts, the systematics can be estimated to be around 1 permille: cutting both electrons above 40 GeV, which is undoubtedly a tight



(a) Peak with superimposed Gaussian fit: stricter f_{brem} cuts corresponds to lower (smaller-statistics) curves



(b) Fitted peak vs f_{brem} cut

Figure 2.24: Reconstructed Z peak as f_{brem} for both electrons (symmetric selection) varies from 0.2 to 1

cut, in fact brings a raise of around 2.5 permille in the Z mass peak, so reasonable intermediate $Z \to e^+e^-$ selections will bring, on average, a 1 permille effect.
On the other hand, it is possible that FSR and internal Bremsstrahlung yields show a correlation with respect to the event kinematics: for example, events with high-η (i.e. tipically, very high energy) electrons could have more collinear Bremsstrahlung photons and, therefore, better Bremsstrahlung recovery and a less-distorted invariant mass peak. In order to isolate this effect, the electron-positron invariant mass peak has been studied at the generator level (before any reconstruction): the η-region between -3 and 3 has been subdivided into 11 equally-spaced regions, and to each region an invariant mass histogram has been associated. One $Z \to e^+e^-$ event fills two histograms, one corresponding to the η-bin of the electron and the other corresponding to the η-bin of the positron; each of the 11 histograms is then fitted with an asymmetric Gaussian, and the peak position is plotted against the bin central η-value.
The result of this analysis is shown in Fig. 2.26: no significant bias related to such a mechanism.

2.5.5 Systematics related to choice of the recalibration quantities

The specific choice of fitting the recalibration factors as defined in Eq. 2.19, even if motivated, can introduce some systematic uncertainty. For this reason, the same Z events were used to extract the recalibration coefficients twice, using:

- the quadratic quantities ϵ as defined by Eq. 2.19
- the simple, linear mass ratios $M_{ee}/M_Z - 1$

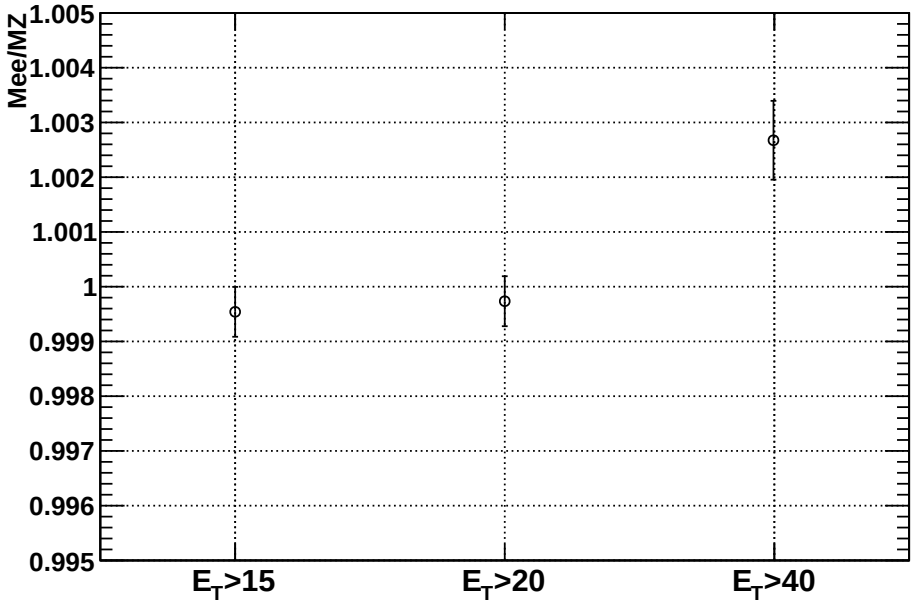


Figure 2.25: Effect of an E_T cut on the reconstructed Z peak

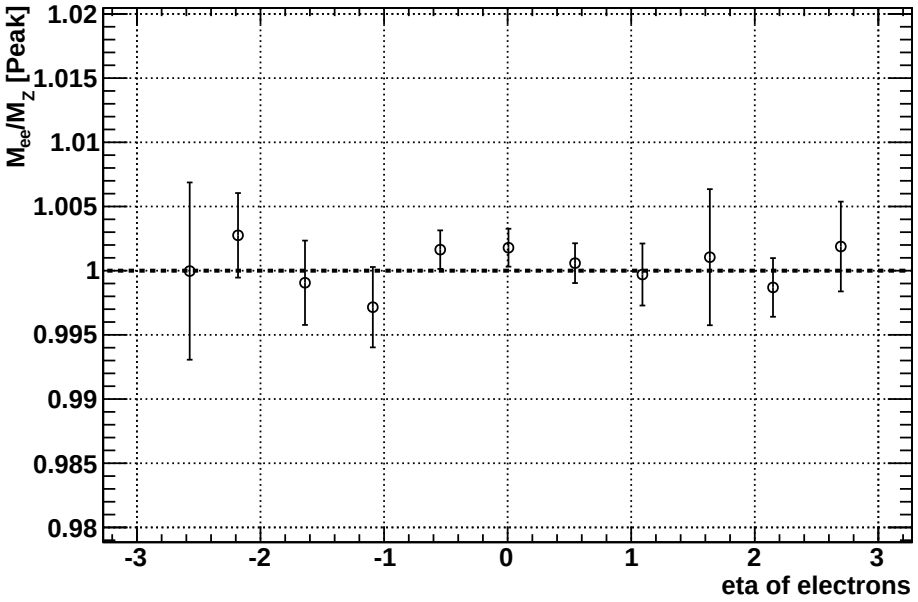


Figure 2.26: Z peak position in η bins of each of the two electrons (see text for details)

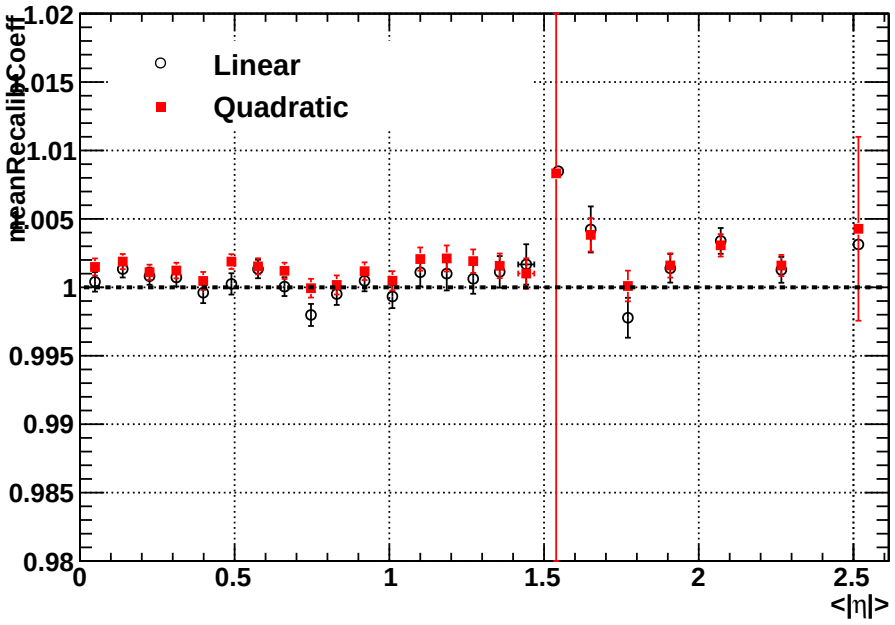


Figure 2.27: Systematics related to the choice of the quantities used for recalibration

The results are shown in Fig. 2.27, and show that the related systematic uncertainty is under control, being around 0.1%.

2.5.6 Event selection

2.5.7 MonteCarlo validation of the method

The iterative method described in the previous sections can be tested by applying known artificial miscalibration constants c_i to the perfectly-calibrated ECAL, and check what recalibration constants r_i are given in output by the iterative algorithm. Here the subscript i denotes the ECAL regions.
At convergence, the algorithm should give a value of r_i such that

$$r_i \approx c_i^{-1} \tag{2.23}$$

for each of the regions; the residual detector miscalibration can therefore be estimated by the dispersion of the quantity:

$$r_i c_i^{-1} - 1 \tag{2.24}$$

which should be peaked at zero with the smallest possible spread. After convergence the spread depends only on the available data statistics, and drops off approximately as $\sqrt{N}$, where N is the number of events.

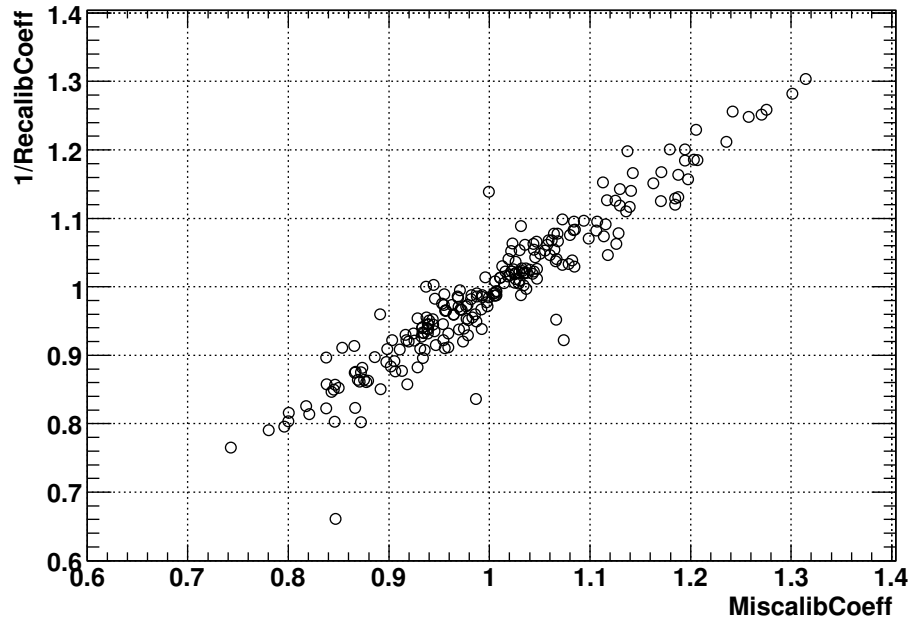
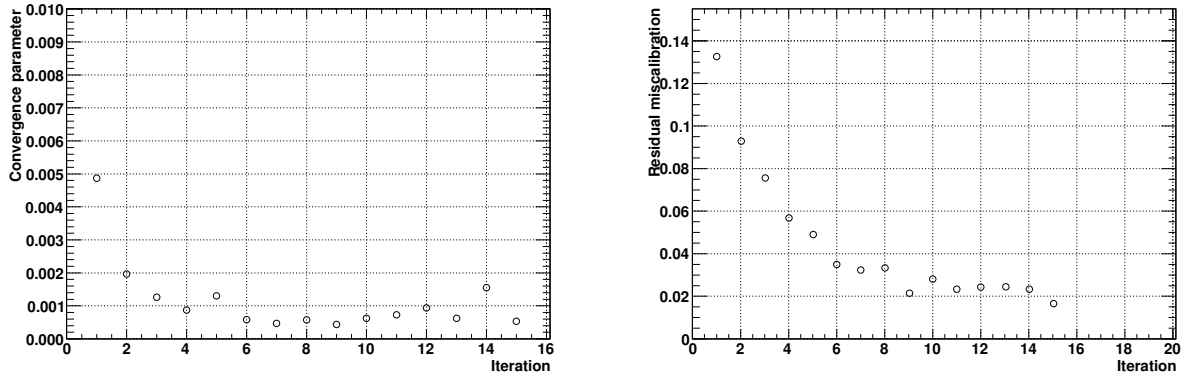


Figure 2.28: r_i^{-1} vs c_i after convergence of the iterative algorithm

A simple test can consist in simulating a large (10%), Gaussian η-ring miscalibration. The CMS ECAL is subdivided into 248 η-rings (i.e. rings made by one crystal along η): 170 rings for the Barrel and 39 for each of the two Endcaps. Each Barrel ring is formed by 360 crystals in ϕ; for the Endcap rings, the number of crystals for each ring varies from around 40 to around 150.
The 12 rings located next to module-module cracks of the Barrel have not been considered in the calibration process: because of the lacks in the containment of the electromagnetic shower, their recalibration constants would have been artificially high. Six rings next to Barrel-Endcap and Barrel-Barrel junction (at $\eta = 0$) - and the last ring of each Endcap - have not been calibrated as well.
Consequently, the total number of calibrated rings is 228.
In addition, the event selection discards all the electrons whose calorimetric seed (i.e. the crystal with the highest energy deposit) lies at the boundary of a module or region of the ECAL: this cut brings an efficiency of around 80%.
Fig. 2.30 shows the relationship between the various c_i and the corresponding r_i at convergence.
As a criterion for the convergence of the algorithm, the convergence parameter:

$$d_k = \frac{1}{n} \sum_{i=1}^{n} \sqrt{(r_{k-1} - r_k)^2} \tag{2.25}$$

where i runs over the n regions to intercalibrate, has been inspected as a function of the



(a) Convergence parameter d vs iteration

(b) Residual miscalibration vs iteration

Figure 2.29: Algorithm behaviour vs number of iterations

iteration number k: the algorithm typically needs 10-15 iterations to converge. The expression in 2.25 basically represents the average "distance" between the recalibration coefficients at iteration k and the ones obtained after iteration $k-1$; when this distance gets small and constant, further iterations bring no improvement and the algorithm can be stopped.
Fig. 2.29 shows how the convergence parameter in eq. 2.25 and the residual miscalibration - computed as the gaussian σ of $r_i c_i - 1$ - vary along the iterations.

Intercalibration potential with early CMS data

Fig. 2.31 shows the best calibration achievable by using the $Z \to e^+e^-$ iterative algorithm as a function of the integrated luminosity collected by the CMS experiment. If a startup intercalibration of around 2% is assumed for the ECAL Barrel, while 7-8% is assumed for the Endcaps, it is clear that application of the $Z \to e^+e^-$ iterative method can bring some improvement to the Endcap calibration even in the presence of a perfectly-Gaussian miscalibration of the detector.

Impact of the Z natural width

The $Z \to e^+e^-$ calibration method described in the present Chapter aims to achieve a calibration between regions better than the ratio between the natural width and the mass peak of the Z boson, i.e. $\Gamma_Z/M_Z \simeq 2.7\%$.
The fact that the Z boson is a resonance having a natural width of 2.7% brings an interesting effect that is shown in Fig. 2.31: going from the initial, large (10% in the specific case of Fig. 2.31) artificial miscalibration to a calibration of 2.7% requires a very small data statistics - around 10 pb^{-1}- while improving from 2.7% to 1% requires a statistics 10 times larger.
In general:

- obtaining a calibration of order Γ_Z/M_Z requires a very small number of Z events: apart

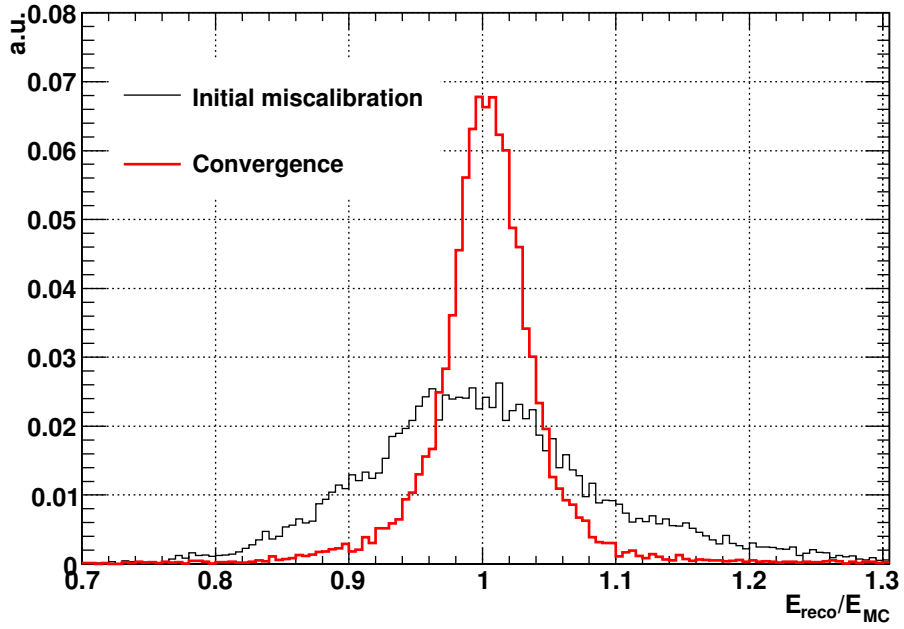


Figure 2.30: Electron energy resolution, before and after convergence of the iterative algorithm

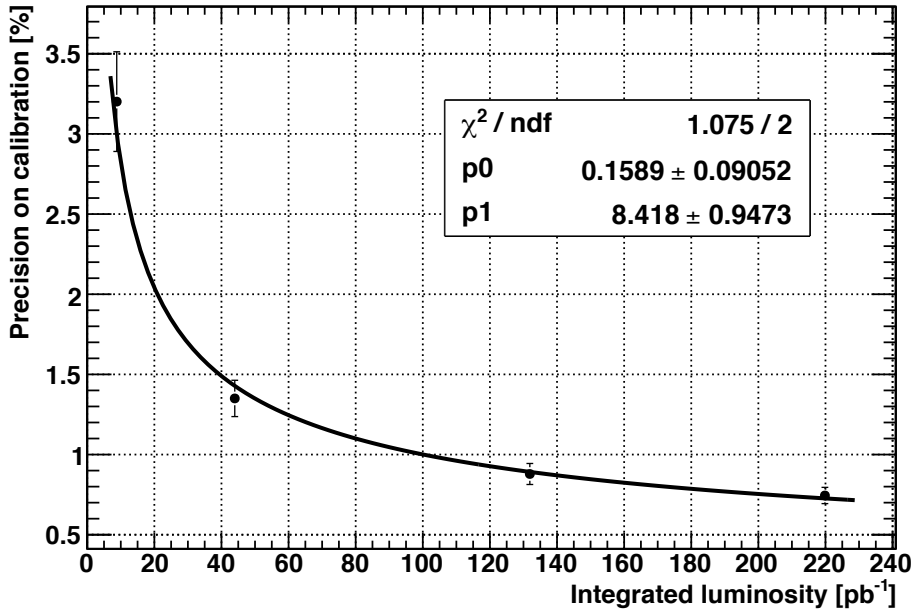
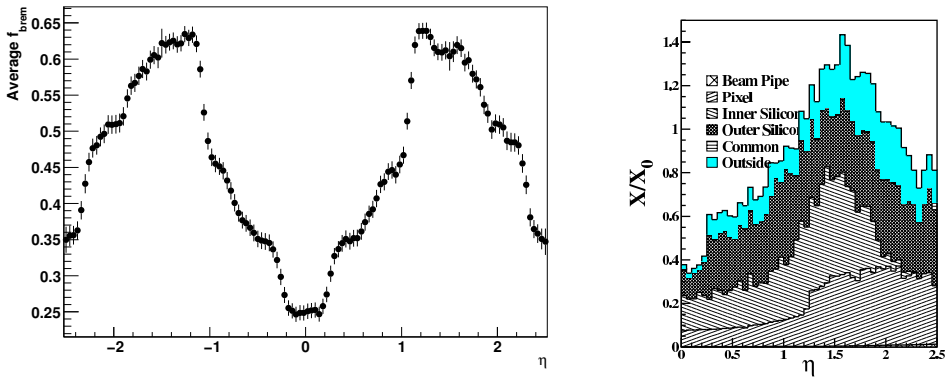


Figure 2.31: Calibration achievable by using the $Z \to e^+e^-$ iterative algorithm as a function of the integrated luminosity



(a) Average f_{brem} vs η

(b) Material budget in front of the ECAL

Figure 2.32: Average value of f_{brem} for electrons from Z decay, after HLT selection, as a function of electron η; this is compared to the budget of material in front of the ECAL

from the cross-talk between different ECAL regions (that motivates the iterative approach), one single event per region can calibrate the ECAL regions to a level of 2.7%;

- if a calibration better than Γ_Z/M_Z is to be obtained, then it becomes important to accurately determine the value of the fitted peak in the regional calibration histograms (like the one shown in Fig. 2.18), and this requires a larger number of events.

2.5.8 Applications of the iterative method

The example in the previous section clearly shows the power and effectiveness of the calibration method; it is not very realistic though. The ECAL is expected to start data taking with a calibration status which has been achieved through various strategies, the residual crystal-to-crystal miscalibration being expected to be around 1-2% in the Barrel and around 7-10% in the Endcaps. For example, the startup miscalibration between η-rings is equal to the crystal miscalibration divided by $/\sqrt{N_{xtal}}$, where N_{xtal} is the number of crystals inside the ring; this means around 0.1% in the Barrel and 0.7% in the Endcaps. With a data statistics of around 100 pb^{-1}it is possible to detect and correct such a small miscalibration in the ECAL Endcaps using the iterative algorithm (see Fig. 2.31); in addition to this, there is a number of situations that the algorithm can effectively investigate and correct.

Tuning of clustering algorithms

Even with a perfectly-calibrated detector, electrons emit Bremsstrahlung photons; the amount of Bremsstrahlung radiation varies with η as depicted in Fig. 2.32.
The visible effect of this phenomenon is that, since some of the radiated energy is not recovered by the clustering algorithms, the reconstructed electron energy is lower than the true electron energy by an amount that is η-dependent.

The electron reconstruction algorithms in CMS incorporate an η-dependent correction factor; running the iterative method over a perfectly-calibrated detector checks and tunes these correction factors (that, in this case, correspond to the recalibration coefficients r_i).
It is worth noticing that:

- these η-dependent correction factors are particle-specific, i.e. refer only to electrons because electrons radiate. These correction factors do not refer to the calibration status of the detector, but to the quality of the electron reconstruction;
- since electrons specifically need η-tuned corrections, if the detector is not perfectly calibrated, then for each region i the iterative method gives a recalibration constant r_i that is the product of two terms:

$$r_i = r_i^{detector}\ r_i^{ele} \tag{2.26}$$

  where $r_i^{detector}$ is the factor that corrects for the region-averaged detector miscalibration and r_i^{ele} is the average electron-specific correction for that region.

If the iterative method is applied to an almost perfectly-calibrated detector, then $r_i^{detector} \approx 1$ and the method measures r_i^{ele}; if the detector miscalibration is consistent, then a disentanglement of the two terms in eq. 2.26 can be desirable.
In this latter case, there are two possible calibration strategies:

- the detector could be calibrated using physics channels with photons in the final state (i.e. $\pi^0/\eta \to \gamma\gamma$), and then the present method could deliver electron corrections;
- the present method can be implemented twice, with different electron selections. The first time it can be implemented only on good-quality electrons (i.e. with very low f_{brem}), so that $r_i^{ele} \approx 1$ and one obtains $r_i^{detector}$; then, it can be applied to a looser selection in order to obtain r_i^{ele} separately for the desired electron selection.

Fig. 2.33 shows the electron correction factors as a function of electron η, for two E_T - regions; no particular cuts have been requested for the electron, except for $E_T > 15$ GeV. No miscalibration factors have been inserted, so what is plotted is actually r_{ele}, but the electron energy is calculated just as the sum of the energy deposits in the cluster crystals (without any further correction factor). In this condition, electron-specific corrections up to 6% are needed.
With increasing E_T , electrons generally need less Bremsstrahlung correction because the Bremsstrahlung photons are, on average, emitted more collinear with respect to the electron direction and more energetic electrons are bent less by the solenoidal magnetic field, so that the recovery is more efficient.
Fig. 2.34 shows the corrections coming from the iterative Z algorithm, as calculated on top (i.e. after) all the corrections incorporated as of today into the CMS reconstruction framework.

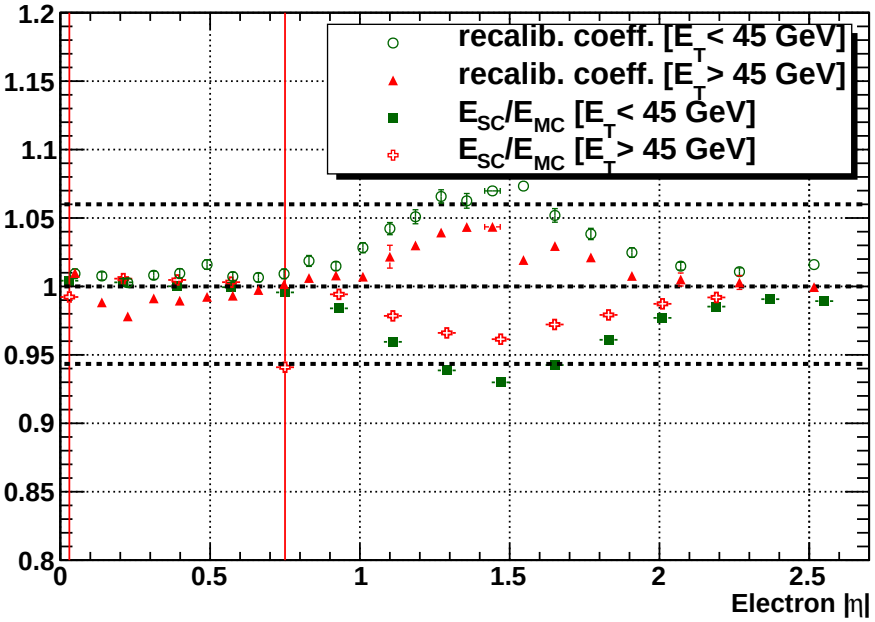


Figure 2.33: Electron-specific correction factors along η, as calculated over *raw* electron energy

In this case, the electron energy looks well reconstructed, except for the need of some residual (few-permille) correction for $\eta \simeq 1.3$.
During the CMS data taking, it will be possible to perform a "fine-tuning" of the electron corrections already implemented into the reconstruction algorithms.
Of course, changing the electron selection has effects on Brem radiation - and, more generally, on the electron reconstruction quality - and changes the necessary correction factors.
It is therefore possible to classify the electrons according to a certain number of reconstructed variables (i.e. f_{brem}, number of clusters...) and compute class-specific correction factors.

Intercalibration of ECAL Barrel modules

The Z iterative method can be also used in order to intercalibrate the ECAL Barrel modules. Each of the 36 Barrel SuperModules is divided into four modules along η, so the Barrel is made up of 144 modules; the 36 modules symmetrically closest to $\eta = 0$ are made up of 25 crystals in η and 20 along ϕ (500 crystals) while the others are made up of 25 crystals in η and 20 along ϕ (400 crystals).
Some modules are entirely composed of crystals from a batch of Chinese production, and these have a different light yield with respect to all the other crystals: this explains how non-Gaussian miscalibration between modules is somehow expected, and so the Z algorithm can be particularly useful in calibrating them.

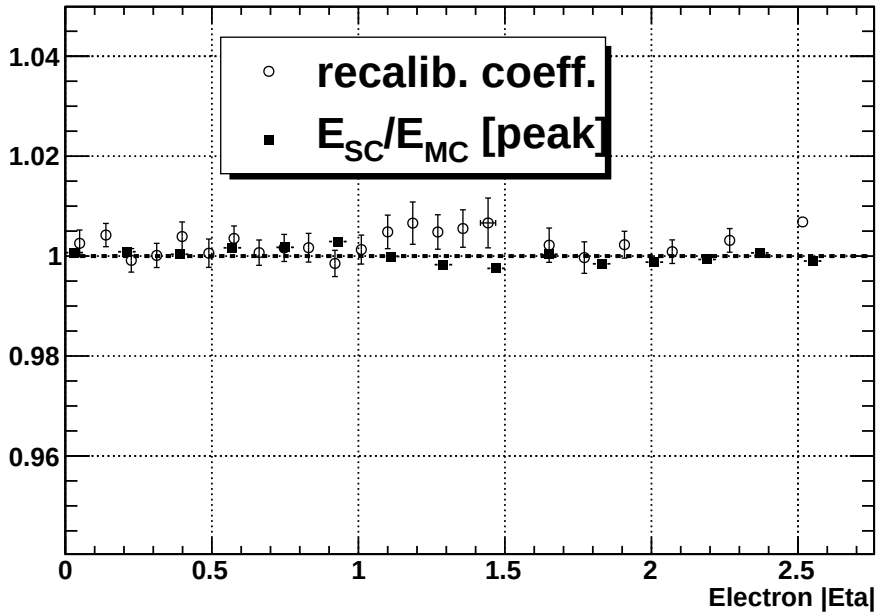


Figure 2.34: Electron-specific correction factors along η, as calculated over corrected electron energy

2.5.9 Extension to lower-energy electrons

Since the Z algorithm uses electron with energies in the 50-100 GeV (see Fig. 2.9), it is not trivial to extend the recalibration coefficients to electron belonging to different energy scales. These extensions are feasible only if the ECAL is verified to be linear in the response to electrons in that particular energy range.
This work shows (see Appendix A) that a linearity test of the ECAL has been performed at CERN in 2006-2007, and the outcoming results indicate an ECAL linearity better than 1% in the electron energy range 9-100 GeV; this means that the recalibration coefficients obtained using the Z iterative algorithm can be safely applied to lower-energy (i.e. 10 GeV) electrons, with an uncertainty lower than 1%.

Chapter 3

Measurement of efficiency of electron charge identification from data

For some physics processes it seems not always essential to unambiguously determine the charge of leptons to uniquely identify a certain decay signature since the knowledge of the existence of a lepton per se suffices. For other processes however a reliable charge identification can be a powerful instrument to discriminate the signal from the background events. A current example would be a decay channel for supersymmetric particles into a final state with two same-sign leptons. A major background for this signal are top-anti-top ($t\bar{t}$) events, with a decay signature resembling the jet-rich supersymmetric decay except for the two leptons carrying an opposite charge with respect to each other.
Electric charge is assigned to reconstructed electrons on the basis of the sign of the curvature of the track that, associated to a calorimetric SuperCluster, forms the electron object. The electron charge has a finite probability of being wrongly-assigned, this probability depending on electron E_T and η as well as on the quality selection made on the electron itself.
The probability of wrong charge assignment to electrons can be taken directly from Monte Carlo; anyway, it is possible to extract it from real data, thus avoiding the dependence on the Monte Carlo model of the detector.

Applications of the measurement

The measurement of wrong charge assignment to electrons has several interesting applications:

- It enters directly all measurements where the charge of an electron is used to define the signal, i.e. charge asymmetries
- In the measurement of the $Z \rightarrow e^+e^-$ cross section, the event selection strategy can include choosing only events with two opposite-sign electrons; the mischarge probability is then important in order to correct the estimate of the signal events.
- A number of new physics channels have same-sign signatures with electrons; precise knowledge of the mischarge probability of electrons is needed in order to accurately determine both the signal and the SM background yields.

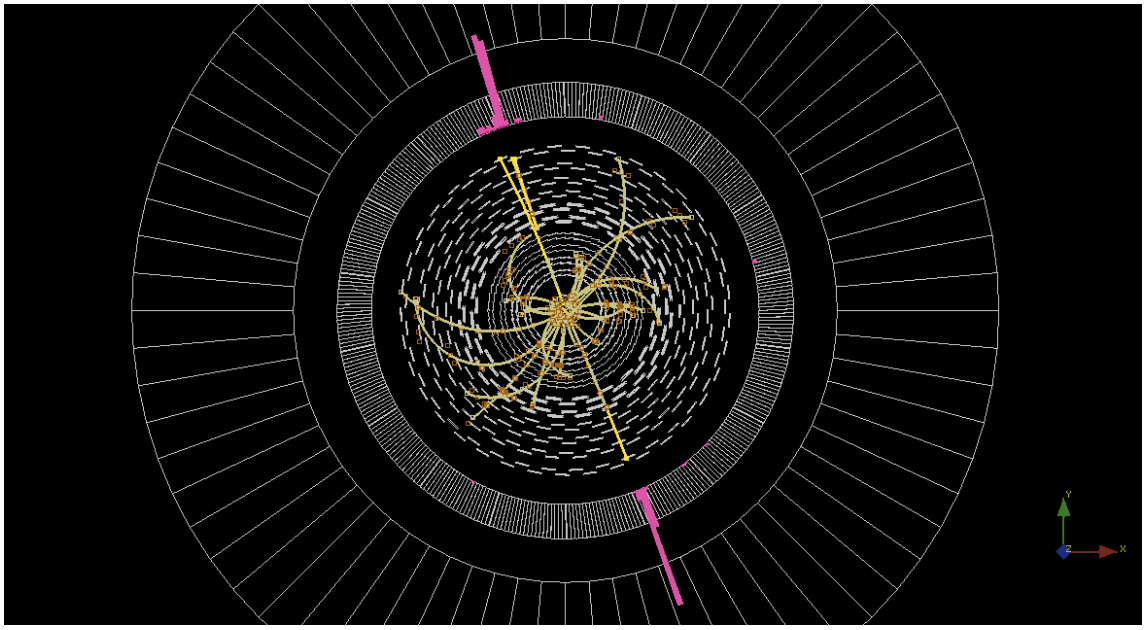


Figure 3.1: Visualization of a $Z \rightarrow e^+e^-$ event with a charge-misidentified electron (at the top left of the event)

Causes of electron charge misidentification

Fig. 3.1 shows the view [37] of a simulated $Z \rightarrow e^+e^-$ event in the CMS detector, where the charge of one two electrons is wrongly reconstructed. Other examples of such events can be found in Appendix B.
Electron charge misidentification is mainly caused by an interaction of the electron with the detector material. For electron energies above several hundred MeV almost all energy loss is due to Bremsstrahlung and scattering or ionization effects play only a minor role.
Bremsstrahlung in matter leads to deflections and kinks in the track of the electron and such result in a wrong measurement of the sign of the curvature of the track. Also, there is the possibility of subsequent conversion of the Bremsstrahlung photon also having the ability to confuse the reconstruction algorithms.
This mechanism is schematically depicted in fig. 3.2.
In order to validate this "Bremsstrahlung + conversion" hypothesis, the ratio E/p is plotted in fig. 3.3(a) for MonteCarlo-matched electrons coming from Z decay - separately for electrons that have correct charge ID and electrons whose charge is wrongly reconstructed[1].
For electrons whose charge is correctly reconstructed, E/p has a sharp peak at 1 while for wrongly reconstructed electrons, E/p has a smooth shape with long tail at values well greater than 1. This smooth shape is due to underestimation of the electron momentum, that points towards the presence of Bremsstrahlung and photon conversion phenomena. The confidence

[1] A reconstructed electron here is defined as "matched" to a MonteCarlo electron if the two object have a distance of less than 0.05 in $\eta - \phi$ space

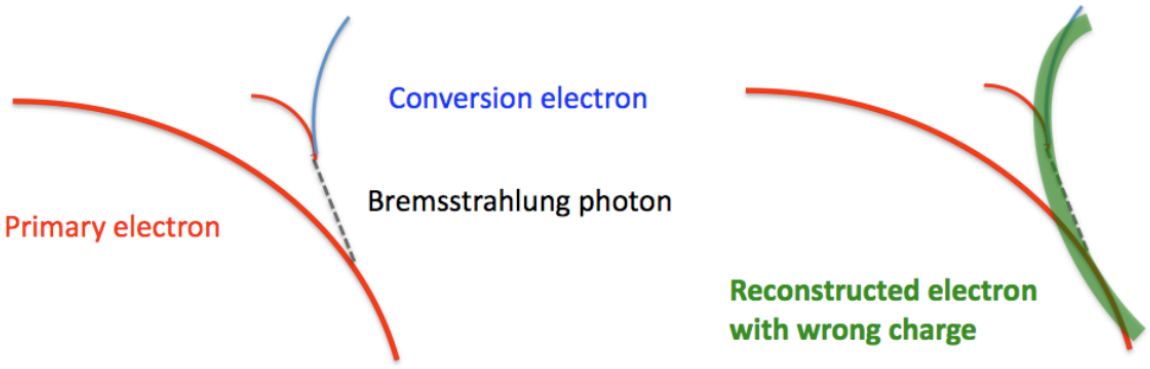
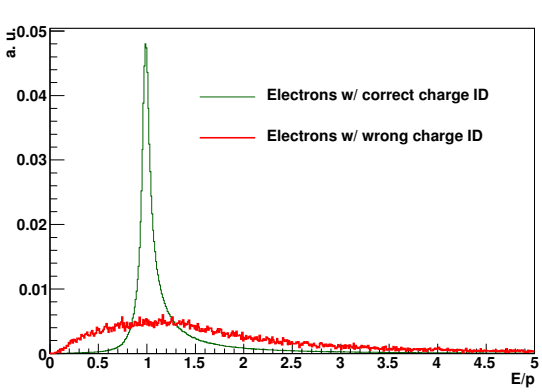
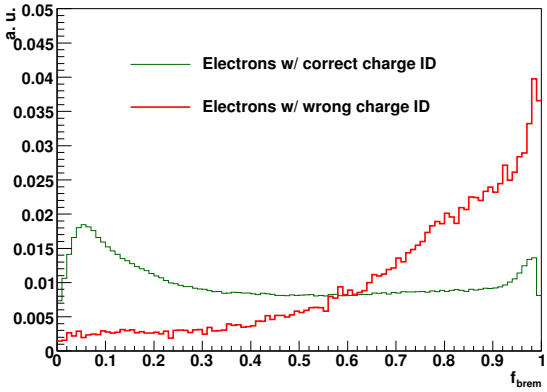


Figure 3.2: Possible cause of electron charge misidentification mechanism: emission of a Bremsstrahlung photon and subsequent conversion (left) may confuse the reconstruction algorithms and make it build a track with wrong electric charge (right).

in the calorimeter cluster energy measurement is relatively strong due to the following reasons: the cluster collects energies within a certain spatial range and thus Bremsstrahlung photons are likely to contribute to the energy of the electron. Also, the situation for conversion electrons is not too different. Either the conversion electrons carry only a tiny energy fraction compared to the original energy, then their paths are probably too curved to contribute to the cluster, but since their energies were low to begin with the cluster energy is affected only slightly. High energy conversion electrons will however contribute to the same cluster. The suspicion then is that the discrepancy between the cluster energy measurement E and the track momentum measurement p comes from an erroneous track momentum measurement.
Fig. 3.3(b) shows the distribution of the Bremsstrahlung yield (i.e. the reconstructed quantity f_{brem} as defined in Chapter 2) for reconstructed electrons in $Z \rightarrow e^+e^-$ events: it is evident that electrons with misidentified charge have, on average, large f_{brem} and this validates the simple model described above.
Fig. 3.4 shows the energy and transverse momentum resolution for electrons with correct and wrong charge reconstruction: for misidentified electrons the energy resolution doesn't worsen much, while the p_T resolution is extremely worsened and the p_T^{reco}/p_T^{true} distribution is broadened flat.
Fig. 3.5 shows the pseudorapidity and azimuthal angle resolution for electrons with correct and wrong charge reconstruction: in this case, the term "resolution" does not refer to the ratio of reconstructed and MonteCarlo quantities, but to their difference. The pseudorapidity resolution is approximately the same for the two classes of electrons while the ϕ-resolution shows a peculiar structure for misidentified electrons: it is not centered at zero, and is symmetric around zero for negative and positive electrons.
This is expected because reconstructing the wrong charge is linked to reconstructing the wrong curvature and, therefore, the ϕ position at vertex and the Transverse Impact Parameter (TIP) of the track are expected to be affected simultaneously, as depicted in Fig. 3.6, where it is

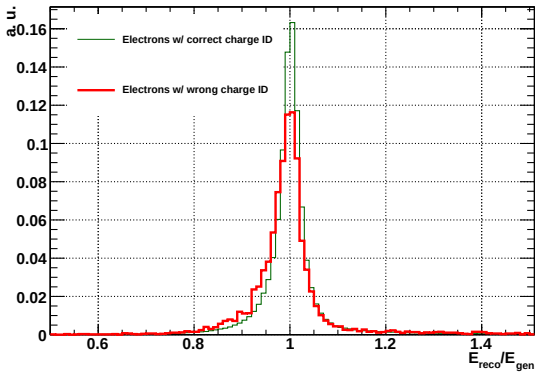


(a) E/p

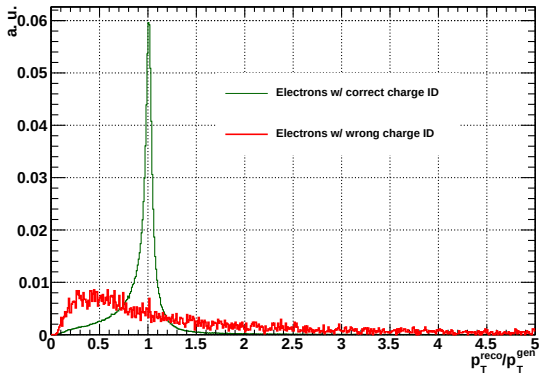


(b) f_{brem}

Figure 3.3: E/p and f_{brem} distribution for electrons with correct and wrong charge ID in CMS

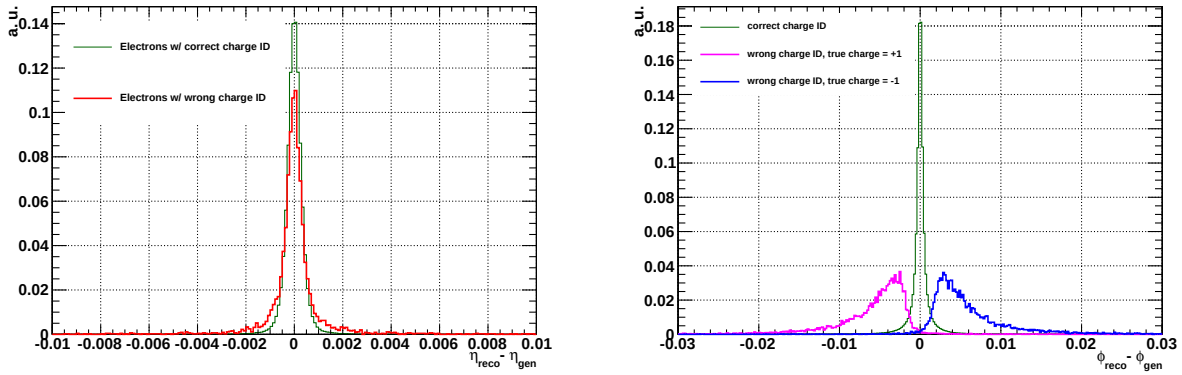


(a) Energy resolution



(b) p_T resolution

Figure 3.4: Energy and transverse momentum resolution for electrons with correct and wrong charge ID in CMS



(a) η resolution

(b) ϕ resolution

Figure 3.5: Pseudorapidity and azimuthal angle resolution for electrons with correct and wrong charge ID in CMS

evident that electrons with large TIP usually have also large values of $|\phi_{reco} - \phi_{true}|$. Fig. 3.7 clearly shows how the TIP is on average larger for electrons with wrong reconstructed charge. In particular, it is simple to observe that the shift in ϕ has always the same sign for misidentified negative electrons, and it is the opposite for misidentified positive electrons.
Fig. 3.8 shows how, for electrons with wrongly reconstructed charge, there is some amount of correlation between high TIP value and large deviations from zero of $\phi_{reco} - \phi_{true}$.
In this section we describe two different methods that employ $Z \rightarrow e^+e^-$ events to extract the probability of wrong charge assignment to electrons using real data only. These methods are the "Symmetric Method" and the "Tag&Probe" method.

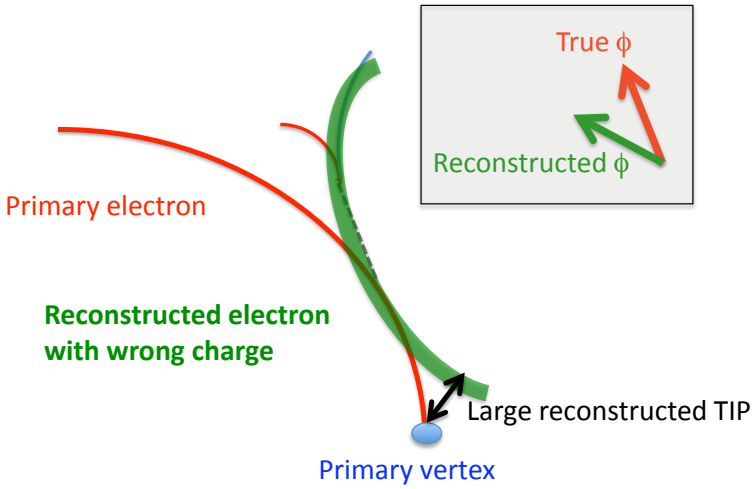


Figure 3.6: Assignment of the wrong track charge to the electron simultaneously brings to an increased TIP and a shift in the azimuthal angle ϕ. The sign of the ϕ-shift depends on the true charge of the electron.

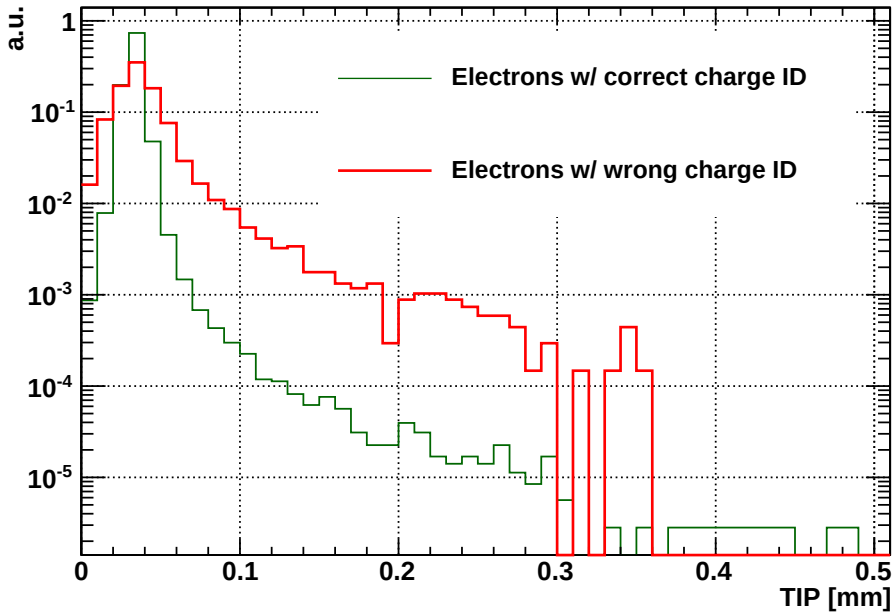


Figure 3.7: Transverse Impact Parameter (TIP) distribution for electrons in $Z \rightarrow e^+e^-$ events

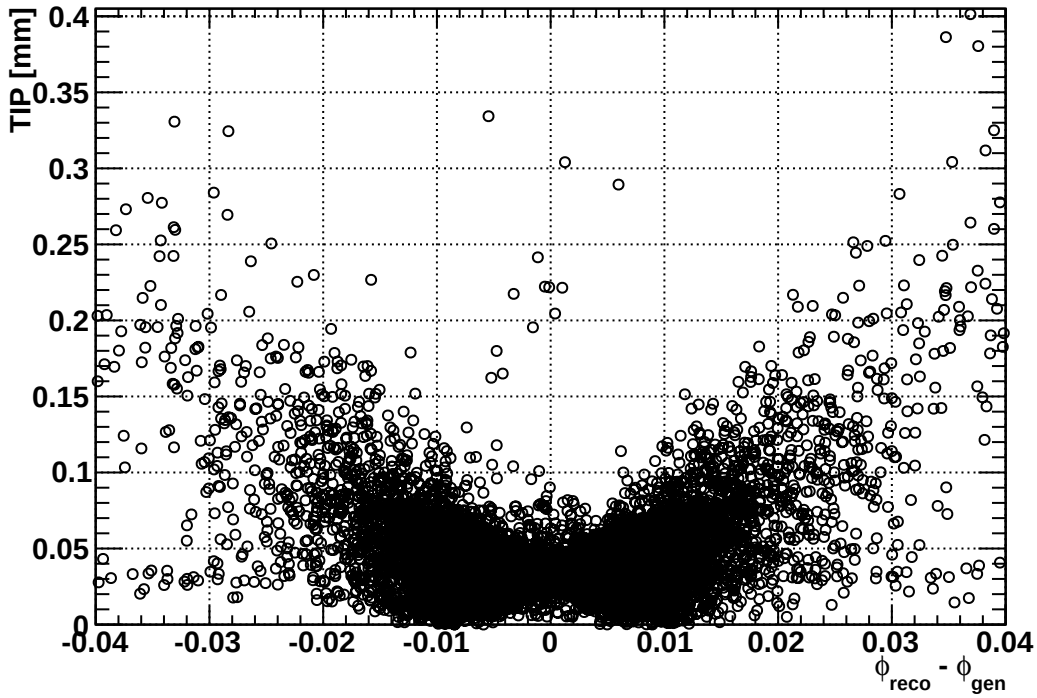


Figure 3.8: Transverse Impact Parameter and azimuthal angle resolution for electrons with wrongly reconstructed charge

3.1 The Symmetric Method

The basic principle of the "Symmetric Method" is to use the same selection for both the electrons coming from the Z decay, so that the mischarge probability P is the same for both of them. In the absence of background, the number of $Z \rightarrow e^+e^-$ events having reconstructed electrons with the same charge is:

$$N_{SS,signal} = 2P(1-P)N_{signal} \tag{3.1}$$

where $P(1-P)$ is the probability of the first electron being reconstructed with the correct charge (probability $1-P$) and the other electron's charge has been mismeasured (probability P), and the factor 2 keeps into account the symmetrical case. Thus by just counting the number of same-sign events passing certain electron selection, it is possible to extract the mischarge probability P related to electrons passing the selection criteria:

$$P = \frac{1}{2}\left[1 - \sqrt{1 - 2\frac{N_{SS}}{N_{tot}}}\right] \simeq \frac{1}{2}\frac{N_{SS}}{N_{tot}} \tag{3.2}$$

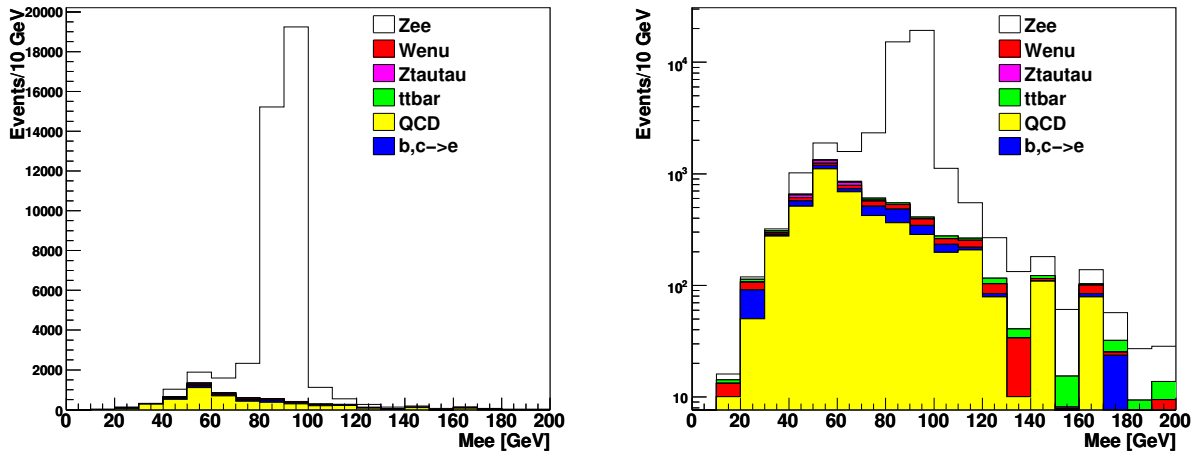
where the last expression comes from the first term of a Taylor expansion and holds if $P << 1$. It is worth noticing that the mischarge probability, as it comes from the Symmetric method, is an integrated probability, because a binning in electron pseudorapidity or transverse energy is not possible. As far as other analyses are concerned, it would be useful to get a value of P for electrons that pass a selection that is of general interest, i.e. not really severe and specific; for example, high-energy and track-isolated electrons - with an isolation cut that is not too tight - can be an interesting sample for a general result useful to the physics programme of CMS.
In particular, the following criteria have been requested for both Z electrons:

- $E_T > 20$ GeV;
- $|\eta| < 1.44$ and $1.56 < |\eta| < 2.5$;
- track isolation, i.e. the request that

$$\sum_{\text{tracks}} \frac{p_T^{\text{track}}}{p_T^{\text{ele}}} < 0.1$$

  for all reconstructed tracks having $p_T > 5$ GeV/c in a cone of radius $\Delta R = 0.6$ centered onto the electron direction.

The pseudorapidity cut is performed in order to reject events where electrons impinge onto the module-module cracks in the Barrel or onto the Barrel-Endcap junction.
The track isolation is important to reject QCD events having one real electron inside a jet, or with a jet misidentified as an electron ("fake electron").



(a) linear scale

(b) logarithmic scale

Figure 3.9: Invariant mass spectrum of events passing selection of the Symmetric Method. Histograms are superimposed.

When loosening the selection on both Z legs it is important to control the amount of background passing the selection, and subtract it effectively if possible.
Fig. 3.9 shows, both in linear and log scale, the mass spectrum of signal and different backgrounds passing the selection described above; Fig. 3.10 shows these selected events separated into same-sign (SS) and opposite-sign (OS) events.
The signal and background samples used throughout this work have been generated using PYTHIA [33] and fully reconstructed in CMSSW, the official software framework of the CMS experiment. All these data samples are listed in Appendix C; in particular, in Fig. 3.9:

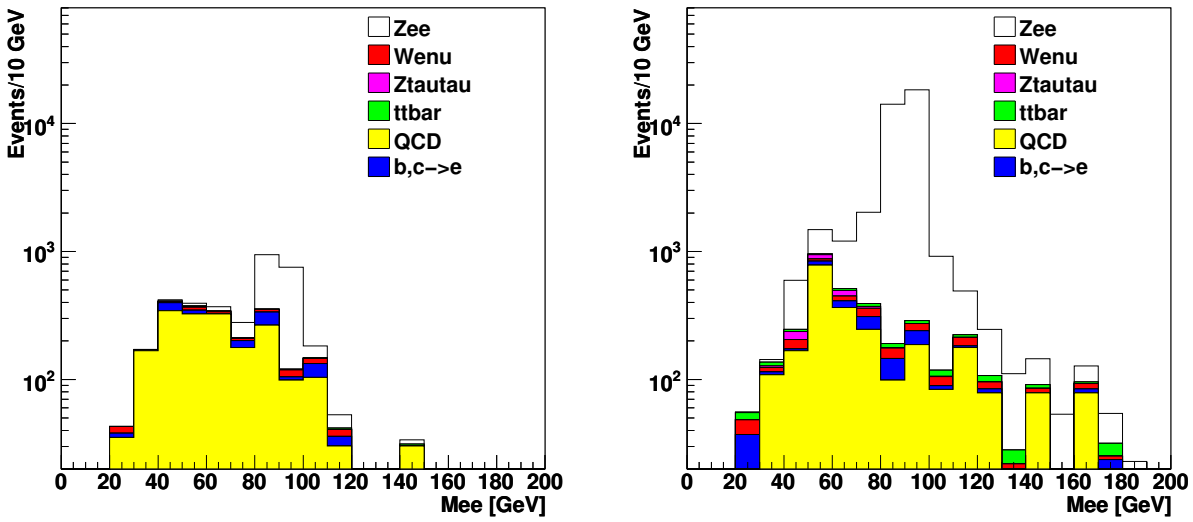
- "QCD" refers to a sample of QCD events where jets are originated from light quarks (u, d, s);
- "$b, c \to e$" refers to a sample of QCD events where jets are originated from heavy quark flavours (b and c), that also undergo semileptonic decay involving a real electron.

Two different methods have been thought of for the purpose of background subtraction.

3.1.1 Background subtraction from S/B ratio

The first strategy of background subtraction for the Symmetric Method uses the MonteCarlo to extract the fraction f_S of signal events in the dataset, i.e. the signal purity.
Putting $Z \to e^+e^-$ signal and all the background together, the number of same-sign events is:

$$N_{SS} = 2P(1-P)N_s + aN_{bg} \tag{3.3}$$



(a) Same-sign (SS) events

(b) Opposite-sign (OS) events

Figure 3.10: Invariant mass spectra of events passing selection of the Symmetric Method: same-sign and opposite-sign events are shown separately.

where $a \in [0,1]$ is the fraction of background events that is reconstructed as same-sign dielectron events; if the quantity $f_s = N_s/(N_s + N_{bg})$ is estimated from MonteCarlo simulation, the above formula gives the following expression for P :

$$P = \frac{1}{2}\left[1 - \sqrt{(1-2a) + \frac{2a - 2\frac{N_{SS}}{N_{tot}}}{f_s}}\right] \tag{3.4}$$

For the chosen electron selection, the following values are given as input:

- $f_s = (98.91 \pm 0.06)\%$ from MonteCarlo (error is only statistical);
- $a = 0.5$ from SS/OS studies (see above), physically in agreement with the lack of charge correlations in the predominant QCD background.

and thus eq. 3.4 gives the following value for the mischarge probability:

$$P_{S/B} = (1.83 \pm 0.06)\% \tag{3.5}$$

which is very good when compared to the MonteCarlo value:

$$P_{MC} = (1.84 \pm 0.06)\% \tag{3.6}$$

F	P [%]
0.5	1.71 ± 0.06
0.75	1.77 ± 0.06
1	1.83 ± 0.06
1.25	1.90 ± 0.06
1.5	1.96 ± 0.06
2	2.08 ± 0.06

Table 3.1: Sensitivity of the S/B method to the background yield. F is the multiplicative factor of the background yield.

Systematics due to choice of counting region

Performing the same analysis using an invariant mass window of [75,105] GeV/c^2 (which doubles the background fraction from 1.8% to 3.5%) yields a value of $P_{sidebands} = (1.74 \pm 0.06)\%$, so, by difference with respect to the value in Eq. 3.5, this systematic uncertainty can be conservatively estimated to be around 0.09%.

Systematics due to estimation of signal purity

The MonteCarlo estimation of the signal and background yields is affected by a theoretical uncertainty; in order to quantify the impact of this uncertainty on the measurement of P, the analysis has been performed fixing f_s to the value indicated by the MonteCarlo, while both the SS and OS backgrounds have been rescaled by a global factor F.
Table 3.1 shows the sensitivity of the method to f_s; if, for example, the real background is twice the MonteCarlo-predicted value (F=2), the method yields a value for P which is 14% greater than the true value.
If a (conservative) uncertainty of a factor 2 in the background yield is taken into account, the resulting systematic uncertainty on P is around 0.25%.

Systematics due to estimation of SS/OS ratio in the background

This systematic uncertainty can be obtained by preserving the overall background yield while changing the "composition" of the background in terms of SS and OS components.
As stated above, it is reasonable to assume that SS/OS = 1 (within statistical error) for the background, i.e. $a = 1/2$ in Eq. 3.3; varying a is equivalent to changing the mean SS/OS ratio for the background, as illustrated in Table 3.2.
If an uncertainty of a factor 1.5 in the SS/OS ratio of background is taken into account, the resulting systematic uncertainty on P is around 0.06%, as shown in Table 3.3.

Overall systematic uncertainty

If all the above systematics are taken into account and added in quadrature, the resulting measured value of P with the S/B method becomes:

SS/OS	a	OS-bg rescale factor	SS-bg rescale factor
0.5	0.33	1.33	0.67
0.75	0.43	1.14	0.86
1	0.50	1.00	1.00
1.5	0.60	0.80	1.20
2	0.67	0.67	1.33

Table 3.2: Correspondence between the a factor, as defined in Eq. 3.3, and the SS/OS ratio of the background.

Background SS/OS	P [%]	Pull wrt SS/OS=1
0.5	1.74 ± 0.06	-1.06
0.75	1.79 ± 0.06	-0.47
1	1.83 ± 0.06	0.00
1.5	1.89 ± 0.06	0.71
2	1.92 ± 0.06	1.06

Table 3.3: Sensitivity of the S/B method to the SS/OS ratio of the background.

$$P_{S/B} = (1.83 \pm 0.06\ [\text{stat.}] \pm 0.36\ [\text{syst.}])\% \tag{3.7}$$

which shows that the uncertainty is clearly dominated by systematics, even at very low integrated luminosities (i.e. around 10 pb^{-1}).

3.1.2 Background subtraction from sidebands

The basic assumption of this subtraction method is that, in the same-sign spectrum, the invariant mass region above and below the Z peak is signal-free, so it is possible to infer the number of SS-background events under the peak just performing a combined fit of the sidebands and interpolate.
The peak region can be defined as the range [85,95] GeV/c^2; the sidebands can be defined as the [60,80] GeV/c^2 and the [110,140] GeV/c^2 regions; the specific choice of the "signal" and "sidebands" regions brings in a systematic uncertainty that will be investigated later. The functional form chosen to fit the sidebands is an exponential:

$$f(M_{ee}) = Ae^{-BM_{ee}} \tag{3.8}$$

As a second step it can be assumed that the SS and OS spectra of the background are compatible within statistical errors; if this equality holds, then it is possible to get the expected number of background events in the SS-peak region, and use it in a straightforward manner to estimate the number of OS-background under the Z peak.
This "SS=OS" equality for the background will be taken as an assumption, since it cannot be effectively validated using the available MonteCarlo samples having large weights (see

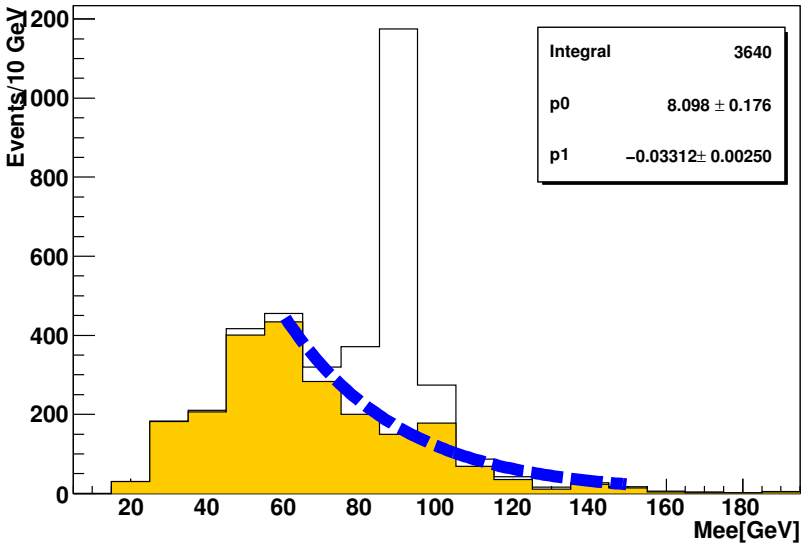


Figure 3.11: Fit to SS sidebands: open histogram is reconstructed data, filled yellow histogram represents the "true" background (from generator-level information).

Appendix C).
The following table shows the number of Signal and Background events coming from the background subtraction method in the SS sample(computed in the [85,95] GeV/c^2 mass window):

	MonteCarlo	Method
S	1025	1007
B	150	168

While the following table shows the same numbers referred to the OS sample:

	MonteCarlo	Method
S	27284	27278
B	162	168

Once the number of background-subtracted SS and OS events is known, the resulting value of P is:

$$P_{sidebands} = (1.81 \pm 0.06)\% \tag{3.9}$$

that has to be compared with the MonteCarlo value coming from the tables above, that is:

$$P_{MC} = (1.84 \pm 0.06)\% \tag{3.10}$$

Systematics due to choice of counting region

The counting region has been defined as the [85,95] GeV/c^2 mass window in the analysis above. The systematics due to the specific definition of the region in which the event counting is performed can be estimated by performing the analysis with different counting regions.
Performing the same analysis using a [75,105] GeV/c^2 region (which doubles the background fraction from 1.8% to 3.5%) yields a value of $P_{sidebands} = (1.88 \pm 0.06)\%$, so this systematic uncertainty can be estimated around 0.07%.

Systematics due to SS/OS ratio in the background

The method is sensitive to the yield of the SS background, but it assumes the OS background to be the same as the SS background, so different yields of SS and OS backgrounds are not taken into account by the method and this can bring in systematic uncertainty.
This systematic can be estimated by multiplying the OS background using a scale factor at the MonteCarlo level; performing this analysis actually shows that, even if the OS background is doubled (i.e. OS/SS = 2 for the whole background), the relative systematic uncertainty on P is below 1% and can be therefore considered negligible.
This is due to the fact that, using these selection criteria, the background contribute only a very small percentage to the number of OS under the Z peak, S/B ratio being around 150 in that region; so changes in the OS-background yield do not affect the measurement significantly.

Systematics due to choice of fit function

The systematic uncertainty on P due to the particular choice of the fitting function can be estimated by repeating the above analysis using another function; for example, a linear function of the kind

$$f(M_{ee}) = A + BM_{ee} \tag{3.11}$$

is considered a reasonable choice. Choosing this linear function brings a systematic of around 0.03%.

Overall systematic uncertainty

If all the above systematics are taken into account and added in quadrature, the resulting measured value of P with the S/B method becomes:

$$P_{sidebands} = (1.81 \pm 0.06\ [\text{stat.}] \pm 0.07\ [\text{syst.}])\% \tag{3.12}$$

which shows that, with 100 pb^{-1}of data, the statistical uncertainty already reaches the systematic term, which is mainly due to the characteristics of the fit procedure.

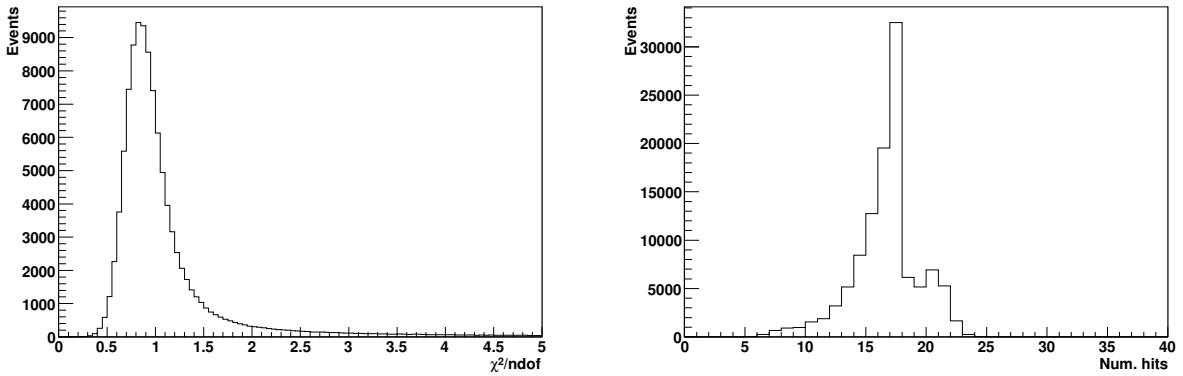
3.2 The Tag and Probe method

Another method that allows to measure the electron mischarge probability P is based on the so-called Tag&Probe (TP) approach.
The TP method, which has been successfully used in some form or another by both CDF and D0 experiments at the Tevatron collider, relies upon $Z \rightarrow e^+e^-$ decays to provide an unbiased, high-purity, electron sample with which to measure the rate of a particular selection or trigger. In this method, a single electron trigger sample is used, from which a subset of di-electron events are selected. One of the electrons, the Tag, is required to pass stringent quality criteria whilst the other electron, the Probe, is required to pass a set of identification criteria depending on the rate under study. The invariant mass of the Tag&Probe electron candidates are required to be within a window around M_Z=91 GeV.
The tight criteria imposed on the Tag, coupled with the invariant mass requirement, is sufficient to ensure high electron purity, i.e. to leave a very low probability that the Probe is not actually a real electron coming from the Z decay.
Note that the tight criteria imposed on the Tag can (and perhaps should) be the same for all rate measurements, but that the Probe criteria necessarily will vary depending on the specifics of the rate that is being measured and how the overall rate has been factorized.
We discuss a particular implementation of this method in the following Sections.
We then present the validation of these tools by comparing the rate obtained by them with that expected from MonteCarlo information. Even though the Tag+invariant mass requirement generally provides a high purity di-electron sample, there will inevitably be some residual background contamination due to various physics channels (i.e. W+jets, γ+jets, $Z \rightarrow \tau\tau$, QCD) having two real or fake electrons. The contribution of this background has also been considered and discussed.

3.2.1 Tag and Probe definition

In the particular case of the charge misidentification rate, the "Tag" electron must have characteristics that ensure that its charge is very likely to be correctly reconstructed.
In particular, since the electron charge in CMS is extracted from the curvature of the electron track, a high-quality track is the basic requirement; so for a Tag electron the following features are required:

- $E_T > 20$ GeV
- $|\eta| < 1.44$ (ECAL Barrel only)
- number of valid reconstructed hits in the track > 10
- normalized $\chi^2$of the reconstructed track < 1.2
- track isolation: $\sum_{tracks} \frac{p_T^{track}}{p_T^{ele}} < 0.2$ for all tracks in a volume around the electron track between a cone with inner radius 0.02 and a cone with outer radius 0.6



(a) χ^2/ndof of electron track

(b) Number of hits of electron track

Figure 3.12: Distribution of track quality parameters of electrons coming from Z decay, after HLT selection

- low Bremsstrahlung radiation: $f_{brem} < 0.1$, where $f_{brem} = 1 - p_{out}/p_{in}$ and $p_{in(out)}$ is the electron momentum measured at the innermost (outermost) tracker layer.

On the other hand, "Probe" electrons are just high-E_T , track-isolated electrons in ECAL fiducial region, defined by the cuts:

- $E_T > 20$ GeV;
- $|\eta| < 1.44$ and $1.56 < |\eta| < 2.5$;
- track isolation (defined the same as for Tag electrons).

This - somehow loose - definition ensures that this method gives information on a very general class of electrons, with cuts that are typical of physics analysis with electrons in the final state. It is always possible to repeat the present analysis using a different Probe definition, thus extracting P values that refer to the new electron class. For a "Tag-Probe" pair to be selected in this analysis, it is additionally requested that the Tag-Probe invariant mass be around M_Z, namely between 85 and 95 GeV.
All the electron four-momenta in this analysis are built using mainly ECAL energy (E); polar (θ) and azimuthal (ϕ) angles are defined using the tracker information:

$$p^{\mu}_{ele} = (E, p_x, p_y, p_z) = (E, E\ \sin\theta\cos\phi, E\ \sin\theta\sin\phi, E\ \cos\theta) \tag{3.13}$$

this choice makes this analysis more detector-oriented with respect to the ECAL, and loosens the analysis dependence on the reconstruction performance of the tracker.

3.2.2 Tag and Probe combinatory

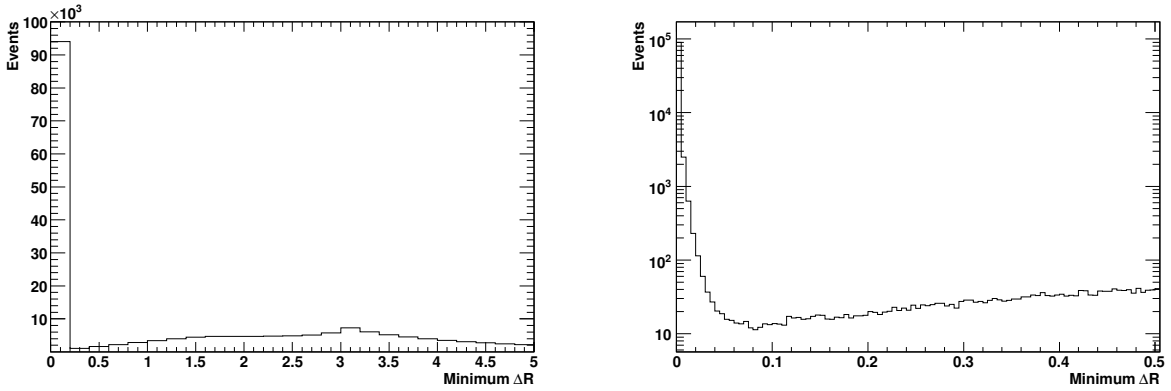
For each event, the method finds Tag electrons (if any); for each Tag electron, a loop is performed on the remaining electrons in the events and those passing the Probe requirements (including the invariant mass with respect to the considered Tag) are coupled to that Tag. Each Tag can thus be linked to multiple Probes, and more than one Tag-Probe pair can be created; in order to avoid biases through unjustified choices, if a Tag is linked to more than one Probe, then these Tag-Probe pairs are discarded and the method moves on to the next Tag electron and repeats the procedure. Similarly, an event is discarded if it contains two or more separate Tag-Probe pairs.
On the other hand, if only one Probe electron is found for a Tag, then the resulting Tag-Probe pair is kept for the measurement.
In the TP method, the probability of wrong charge assignment is always relative to the Probe electrons; it is measured as the number of TP events where the Probe has the same reconstructed charge as the Tag (which is assumed to be correct), divided by the total number of TP events:

$$P_{Probe} = N_{SS}/N_{tot} \tag{3.14}$$

Formula 3.14 assumes that $P_{Tag} = 0$, i.e. the Tag charge is always correctly reconstructed; the more realistic case $P_{Tag} > 0$ will be considered later in this chapter.
For a di-electron event, it is possible that zero, one, or two of electrons may pass the Tag criteria. Likewise, it is also possible that zero, one, or two of the electrons may pass the Probe criteria. There are thus three possible types of events with a Tag&Probe pair: TT, TP and TF where T =passing Tag criteria, P = passing Probe selection, but not Tag criteria, and F = passing neither the Tag nor the Probe selection. The procedure described here loops twice over TT events: at first, one electron plays the role of the Tag while the other is the Probe (the "Tag" selection includes the "Probe" one) and then it turns the other way round. So each reconstructed dielectron event can correspond to zero (PP, TF, PF, FF), one (TP) or even two (TT) Tag-Probe events.

3.2.3 MonteCarlo validation of the Tag & Probe method

In order to verify that the tools described in the previous section can be used to effectively measure wrong charge assignment to electrons from data, the results obtained applying the Tag & Probe method have been checked against those coming from the MonteCarlo truth. In addition to the requirements described above, the Probe electron is therefore asked to match a MonteCarlo electron inside a cone of radius ($\Delta R = \sqrt{(\eta_{ele} - \eta_{MC})^2 + (\phi_{ele} - \phi_{MC})^2}$) equals to 0.05.
This cut value comes from observing the distribution of the minimum distance between a MonteCarlo electron and all the reconstructed electrons in the same event (fig. 3.13); this distribution shows a sharp peak at 0 and ending around 0.1, that corresponds to well-matched electrons. The shallow peak around 3 corresponds to events in which one of the two Z electrons has not been reconstructed. If MonteCarlo electrons are named "mc1" and "mc2" and only



(a) ΔR (b) ΔR - zoomed

Figure 3.13: Distribution of the minimum distance between a MonteCarlo electron and all the reconstructed electrons in the same event

one electron (say "e1", the one corresponding to "mc1") is reconstructed in the event, the reconstructed electron nearest to "mc2" is "e1" itself, and this gives a high value of minimum ΔR.
No background has still been inserted for the purpose of this validation.
The MonteCarlo mischarge probability (P_{MC}) is defined as the percentage of events where the reconstructed Probe has a charge opposite to its matched MonteCarlo electron.
The comparison of P_{MC} with P obtained through the Tag&Probe (P_{TP}) gives an estimation of the reliability and quality of the Tag&Probe method; Fig. 3.14 and 3.16 compare the Tag&Probe results and the MonteCarlo value of P as a function of the pseudorapidity and transverse momentum distribution of the Probe electron, showing an extremely good agreement.
Fig. 3.19 shows a very good agreement in P evaluated as a function of Probe f_{brem}.

Dependence of Charge Misidentification Probability on Pseudorapidity

The shape of the η profile, shown in Fig. 3.14 [2], can be explained as follows: with increasing $|\eta|$, the tracker material increases, the electrons tend to produce on average more Bremsstrahlung radiation and the track reconstruction gets worser. At $\eta = 0$, the mischarge probability is around 0.5%, while at $|\eta| = 2.25$ it gets as high as 5%.
This explanation is supported by the plot of f_{brem} vs η relative to all the electrons coming from $Z \to e^+e^-$ events (Fig. 3.15).
Note that, even if the detector design is intended to be symmetric around $\eta = 0$, it has been chosen to examine the behaviour of P as a function of η, and not $|\eta|$, because during detector

[2]The plots in this Chapter are obtained by using exactly the same set of simulated events for the Tag & Probe method and the generator-level MonteCarlo: this explains why the agreement is better than the statistical fluctuations.

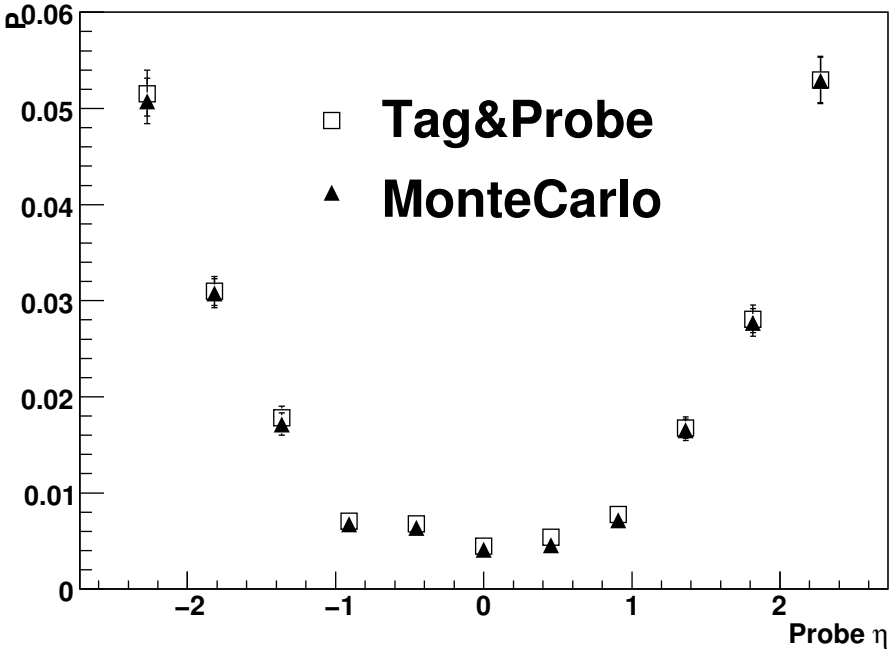
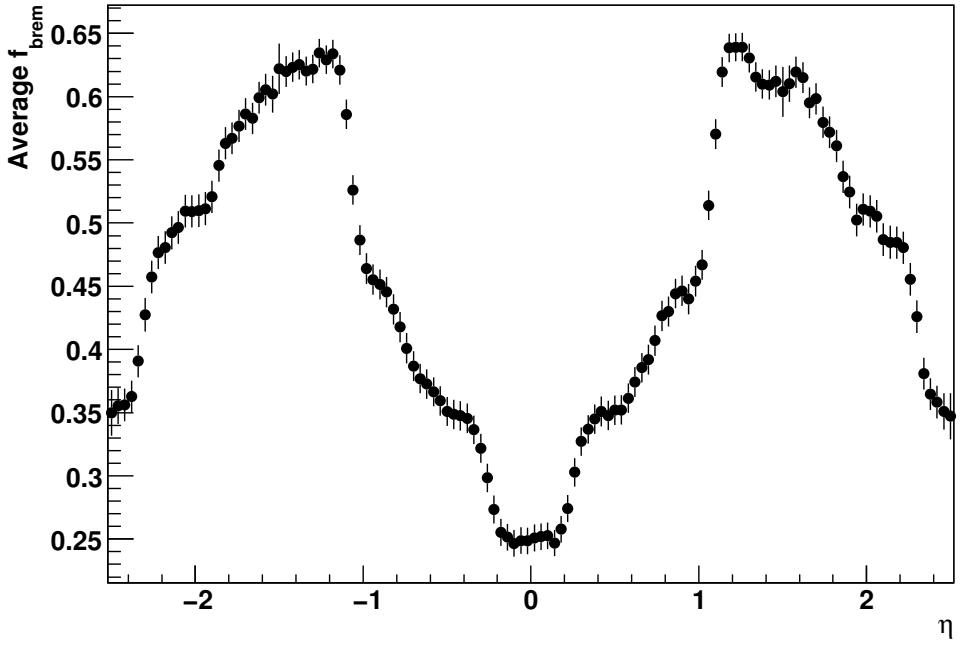
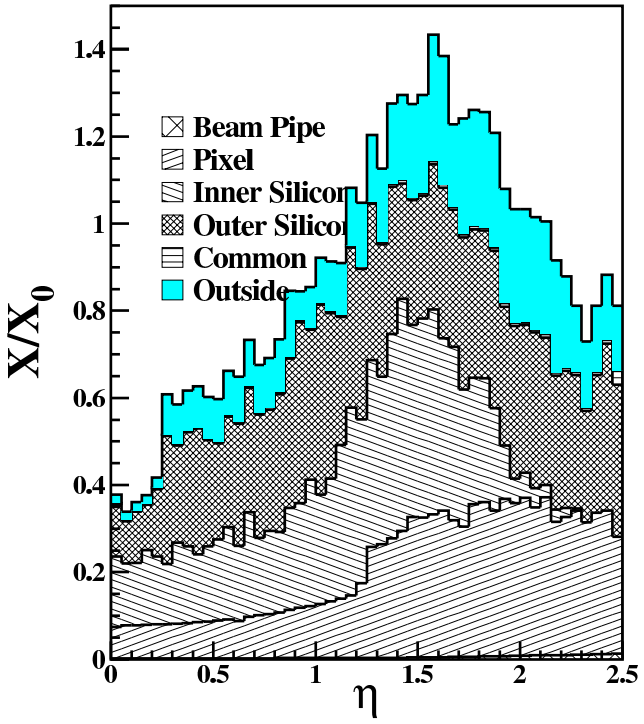


Figure 3.14: Comparison between Probe mischarge probability obtained from Tag&Probe method (white squares) and obtained from MonteCarlo (black triangles): η-profile



(a) Average f_{brem} vs η



(b) Material budget in front of the ECAL

Figure 3.15: Average value of f_{brem} for electrons coming from Z decay, after HLT selection, as a function of electron η; this is compared to the budget of material in front of the ECAL

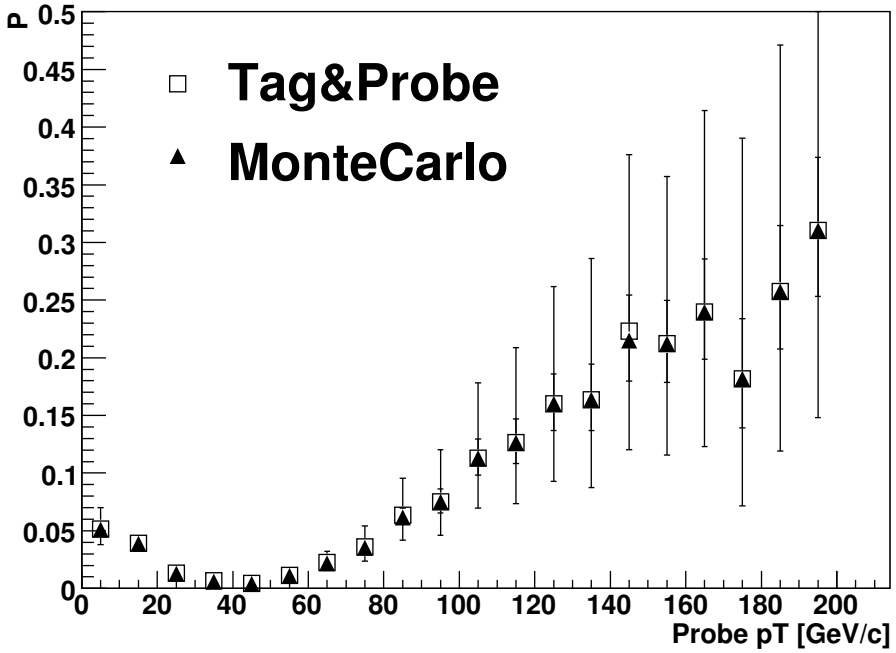
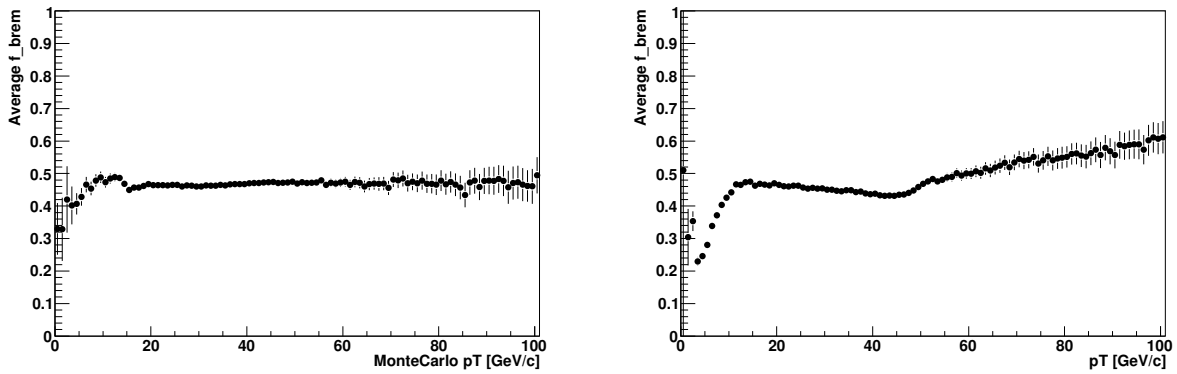


Figure 3.16: Probe mischarge probability obtained from Tag&Probe method: p_T-profile

operation this plot can be a real cross-check of this symmetry, and point to biases and detector asymmetries whatsoever.

Dependence of Charge Misidentification Probability on Transverse Momentum

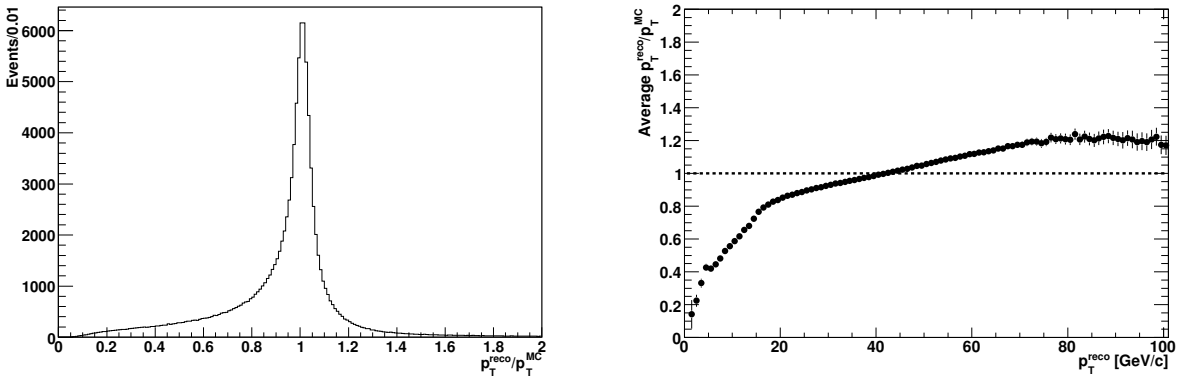
The behaviour of P as a function of the p_T of the reconstructed electron is shown in fig. 3.16; the electron momentum is reconstructed using the tracker information.
In this case a peculiar shape is observed, with mischarge probability reaching a minimum value around 40 GeV/c and then raising for lower and upper momenta; the explanation is that low-momenta bins - i.e. bins with p_T around 20 GeV/c - are populated by electrons that actually have higher transverse momentum - p_T around 40 GeV/c, which is typical of Z electrons - but that emitted high amount of Bremsstrahlung radiation (see Fig. 3.17). This emission worsens the track reconstruction, and has the twofold effect of both reducing the measured p_T with respect to the true one (see fig. 3.18) and producing worse charge reconstruction.
On the other side, P in high-p_T electrons becomes bigger because of the poorer performance of the track reconstruction algorithm as the track sagitta gets shorter and shorter.



(a) f_{brem} vs MonteCarlo p_T

(b) f_{brem} vs reconstructed p_T

Figure 3.17: f_{brem} as a function of p_T : electrons from Z, MC-matched, after HLT selection



(a) Integrated value of the ratio p_T^{reco}/p_T^{true}

(b) Average value of the ratio p_T^{reco}/p_T^{true} vs reconstructed p_T

Figure 3.18: p_T resolution of MC-matched electrons from Z, after HLT selection

Dependence of Charge Misidentification Probability on Bremsstrahlung yield

The charge misidentification probability clearly rises with increasing f_{brem} , as expected since Bremsstrahlung is probably one of the main reasons for the wrong charge assignment phenomenon; this is shown in Fig. 3.19.

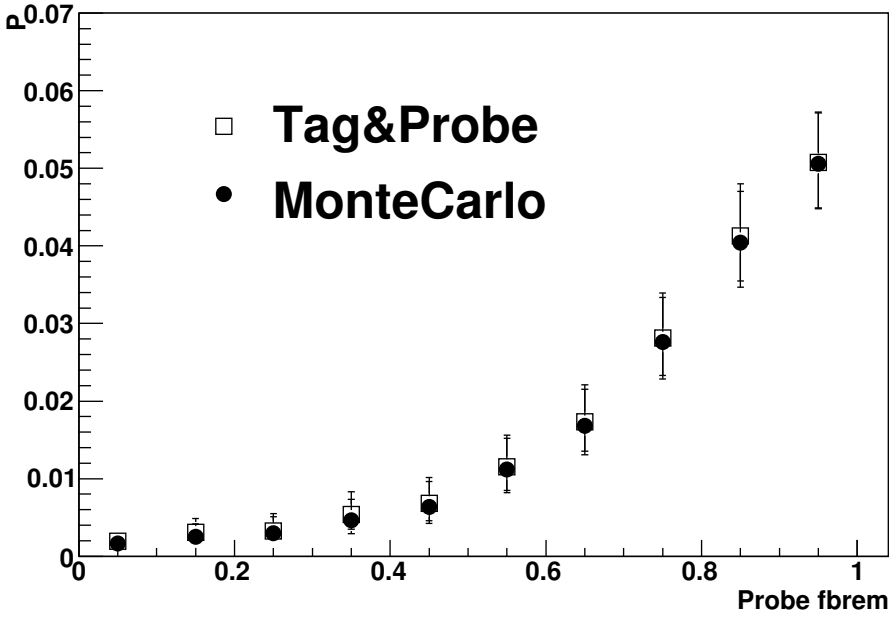
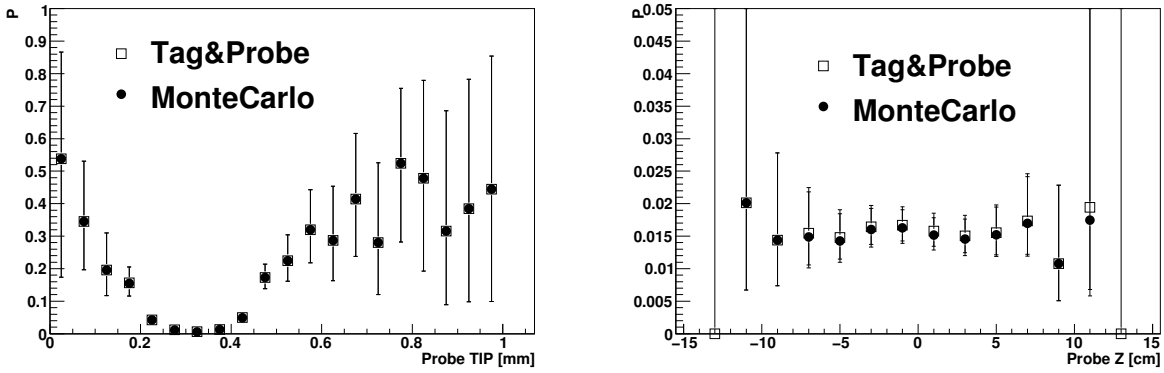


Figure 3.19: Comparison between Probe mischarge probability obtained from Tag&Probe method (white squares) and obtained from MonteCarlo (black triangles): f_{brem} -profile

Dependence of Charge Misidentification Probability on TIP and z-coordinate of the electron track

The charge misidentification probability shows some dependence on the Transverse Impact Parameter (TIP) of the electron track, i.e. the distance - on the $r - \phi$ plane - between the interaction vertex and the extrapolated point of closest approach of the electron track with respect to the interaction vertex (see Fig. 3.20(a)).
If the simple model for the charge misID is correct, it is expected that mismeasured electrons show a somehow different distribution of the TIP; wrong charge ID can occur if the track reconstruction algorithm builds a track made of hits from the primary electrons plus some hits from a conversion electron of opposite charge. The resulting fitted track is expected not to point accurately to the interaction vertex. This is indeed shown in Fig. 3.6 and 3.7.
On the other hand, no significant dependence of the mischarge probability on the beam axis (i.e. z) coordinate of the electron track has been observed (see Fig. 3.20(b)).



(a) P vs Transverse Impact Parameter

(b) P vs Z coordinate

Figure 3.20: Mischarge probability from Tag&Probe method, as a function of coordinates of the electron track at the interaction vertex

Integrated value the of Charge Misidentification Probability

Comparison of the integrated values of mischarge probability P also gives very good results:

$$P_{MC} = (1.49 \pm 0.04)\% \tag{3.15}$$

$$P_{TP} = (1.52 \pm 0.04)\% \tag{3.16}$$

The uncertainty here refers to 100/pb data and is purely statistical; a small bias (around 0.03%) is observed; as explained in the following section, this is due to the fact that P_{Tag} is not zero, i.e. in some cases the Tag charge is wrongly reconstructed.
The integrated value of P in eq. 3.16 is different from the integrated value obtained using the Symmetric Method in eq. 3.5; in particular, the Tag&Probe result reports a smaller mischarge probability.
This is due to the fact that the Probes used in the Tag&Probe method have a pseudorapidity distribution which is different with respect to the electrons used in the Symmetric Method; in particular, the Probes tend to be more central (their average $|\eta|$ is smaller) because they tend to be back-to-back with respect to the Tags, and the Tags are selected to be in the ECAL Barrel. Since the mischarge probability clearly decreases towards $\eta = 0$ (see Fig. 3.14), the integrated value of P in the Tag&Probe is consequently smaller.

3.2.4 Stability of the method versus Tag and Probe definitions

It is extremely important to show that the Tag&Probe method described here is stable with respect to the Tag and Probe definitions, i.e. that it gives accurate results when applied to different electron selections, so to be flexible and useful for any physics analysis.
Fig. 3.21 shows the good agreement of the method with the MonteCarlo truth, for some different definitions of Probe that imply different misID rates; in particular, it shows that the

method remains accurate even when the mischarge probability gets low, as it could get in a better-understood detector or with an optimized charge-reconstruction algorithm.

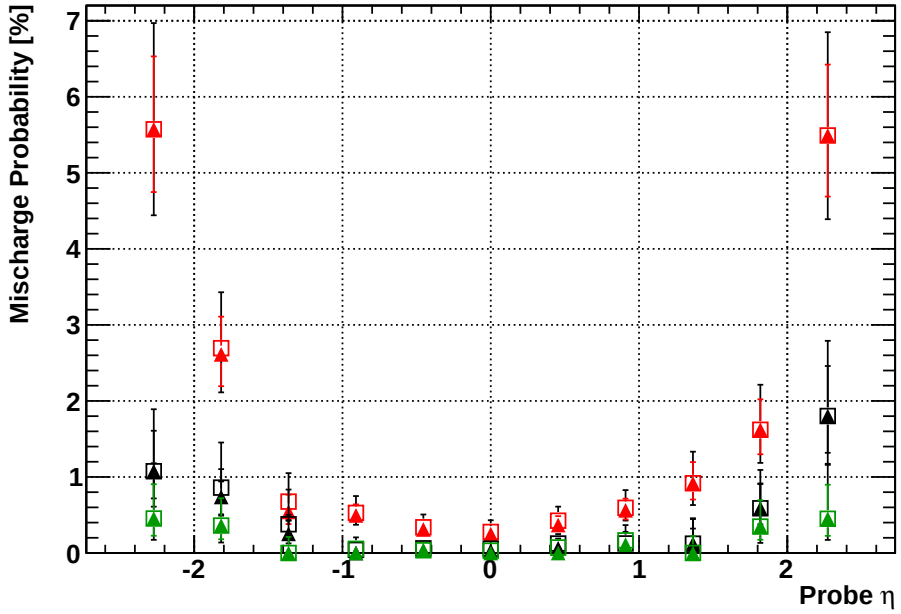


Figure 3.21: MonteCarlo validation of misID versus Probe pseudorapidity, for different Probe definitions: open squares are TagProbe data, full triangles are MonteCarlo

3.2.5 Charge symmetry of the misID rate

It is interesting to investigate whether, given a specific Probe selection, the charge misID rate is charge-symmetric; this is relevant, for example, when applying the charge misID rate to the W^+/W^- ratio, as shown later in the present work.
The charge symmetry can be checked by additionally requesting the Tag electron to have a specific electric charge (i.e. negative Tags) , then repeating the exact same analysis with inverted charge requirement (positive Tags) and compare the two results. Fig. 3.22 and 3.23 show that no significant deviations between the two cases is observable.

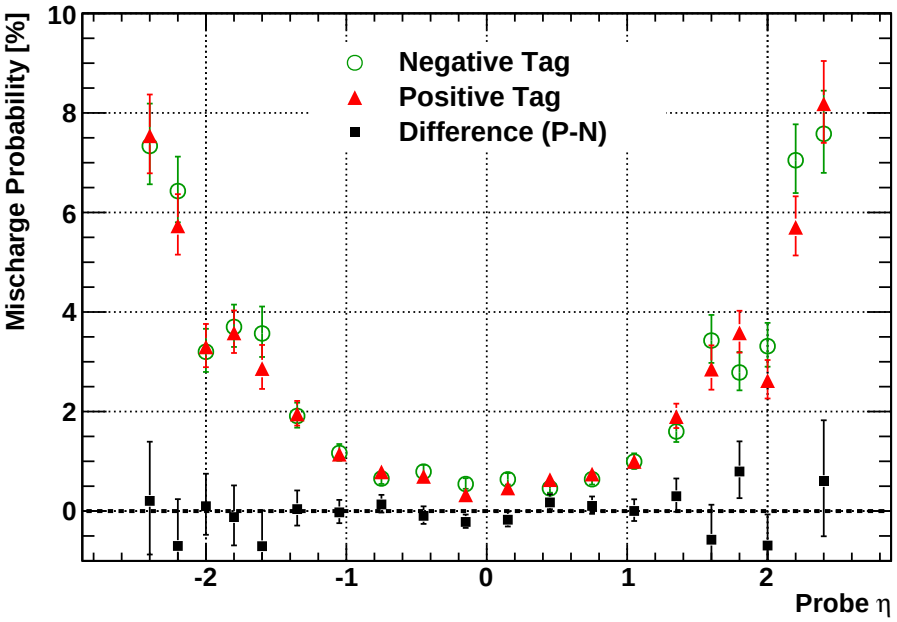


Figure 3.22: Charge symmetry of the charge misID rate.

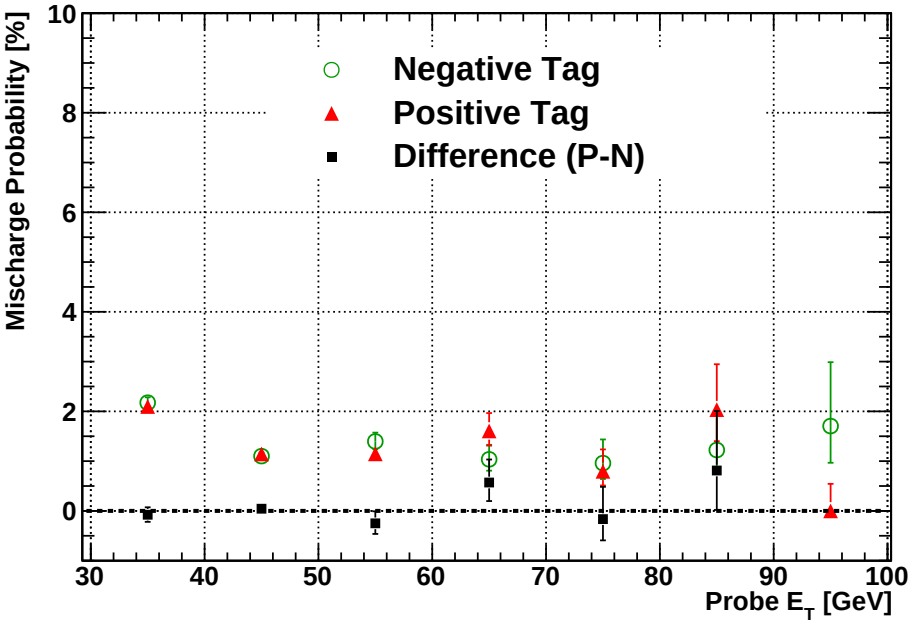


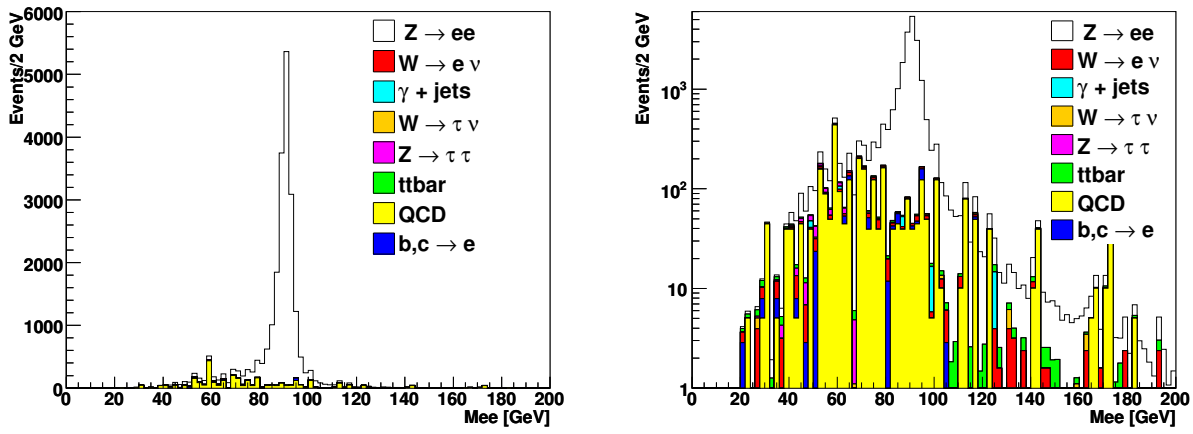
Figure 3.23: Charge symmetry of the charge misID rate.

3.2.6 Backgrounds in Tag&Probe events

Requesting one Tag and one Probe electron in a narrow invariant mass window around the Z peak eliminates most of the background; Table 3.4 shows, for each physics channel, the number of events passing the HLT and the offline (that is, the Tag&Probe) selection: the S/B ratio is around 52, i.e. the background contamination is kept as low as around 2%.
Since the available background dataset only represent a small statistics (due to containment of computational effort), background events passing the selection must be multiplied by large weights in order to represent the desired 100/pb statistics; this brings in some spikes in the invariant mass spectrum (instead of a smooth-shaped spectrum), and concentrates the background in only some E_T and η bins, as will be explained below.
This effect is shown in Fig. 3.25 and 3.27, where the number of weighted and unweighted selected events is shown for each pseudorapidity bin.
The large weight issue can be understood by looking at the pseudorapidity profile of P: adding the background causes some specific bins to largely deviate from the MonteCarlo P. Looking at the background distribution along η it appears that in those bins, only one (unweighted) event of the MonteCarlo sample "QCD-Pt30to80" passes the Tag&Probe selection, but the weighted number of events is around 40; therefore, where 0 events out of 1 do not have opposite-sign electrons, P is pushed higher because the counting method assumes that 0 events out of 40 are opposite-sign.

Physics Channel	Events passing HLT	Events passing Offline Selection
$Z \rightarrow e^+e^-$	71435	14662
$W \rightarrow e\nu$	630535	17
$Z \rightarrow \tau^+\tau^-$	12875	0
$W \rightarrow \tau\nu$	55260	0
QCD-Pt20to30	9682070	0
QCD-Pt30to80	27543900	236
QCD-Pt80to170	535310	10
BCtoE-Pt20to30	673446	0
BCtoE-Pt30to80	1097970	0
BCtoE-Pt80to170	54808	0
ttbar	3428	6
PhotonJet-Pt15to20	272518	0
PhotonJet-Pt20to25	188006	0
PhotonJet-Pt25to30	103202	0
PhotonJet-Pt30to35	56035	0
PhotonJet-Pt35	95971	11
All background	41005333	280

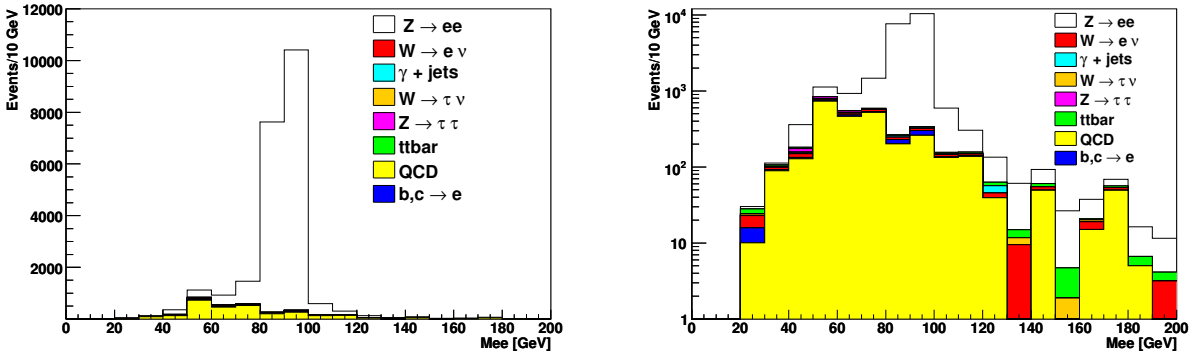
Table 3.4: Number of weighted events passing the HLT and offline selection



(a) Linear scale, 2 GeV bins

(b) Log scale, 2 GeV bins

Figure 3.24: Invariant mass spectra of events passing selection of the Tag&Probe Method



(a) Linear scale, 10 GeV bins

(b) Log scale, 10 GeV bins

Figure 3.25: Invariant mass spectra of events passing selection of the Tag&Probe Method

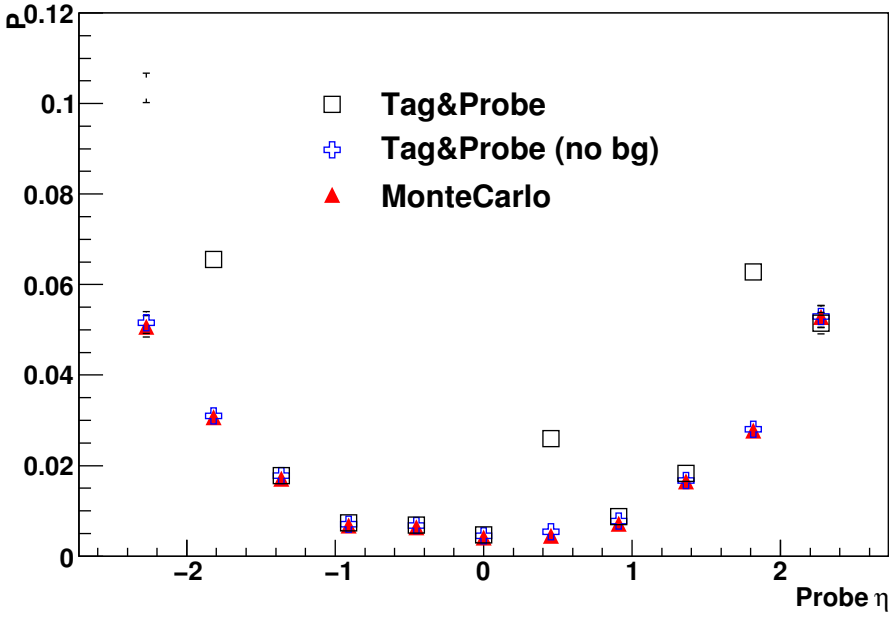
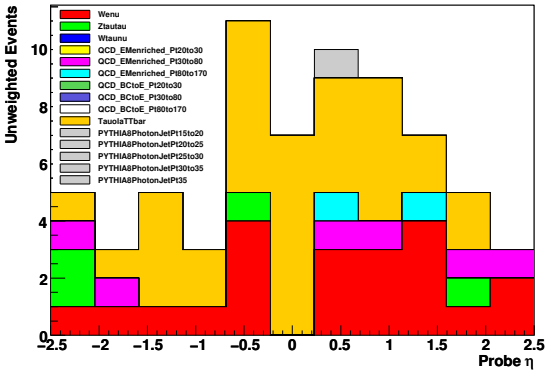
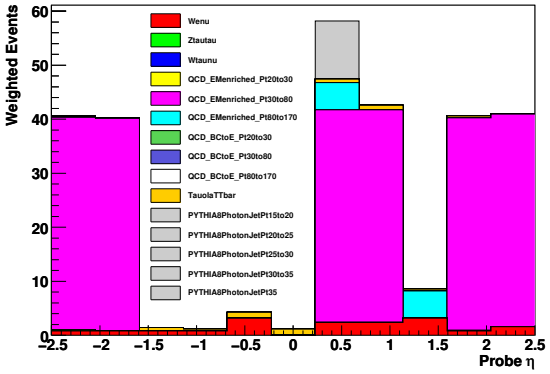


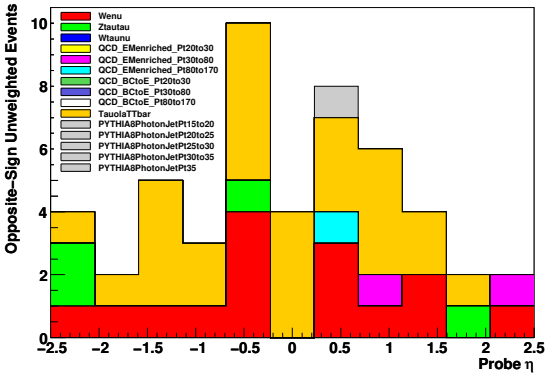
Figure 3.26: P vs pseudorapidity, showing effects of background insertion



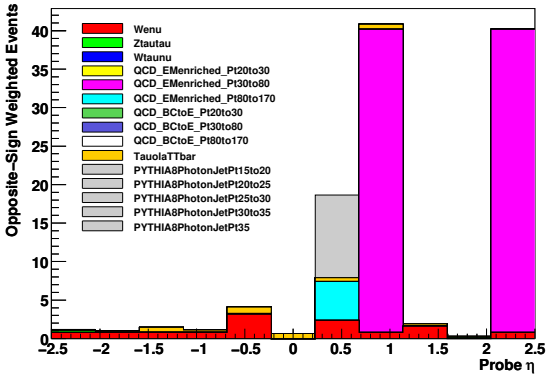
(a) All, unweighted



(b) All, weighted



(c) Opposite sign, unweighted



(d) Opposite sign, weighted

Figure 3.27: Distribution of background events as a function of Probe η

3.2.7 Systematics of the method

Wrong reconstruction of the Tag charge

Up to this point (see Eq. 3.14) it has been assumed that $P_{Tag} = 0$, i.e. Tag charge is always correctly reconstructed. In the most general case where $P_{Tag} \neq 0$, the following holds:

$$N_{SS} = [P_{Tag}(1 - P_{Probe}) + P_{Probe}(1 - P_{Tag})]N_{tot} \tag{3.17}$$

$$N_{SS}/N_{tot} = P_{Tag} + P_{Probe} - 2P_{Tag}P_{Probe} \simeq P_{Tag} + P_{Probe} \tag{3.18}$$

This can be checked by performing different selections on the Tag, thus obtaining different values of P_{Tag} that can be checked using MonteCarlo or data (through the Symmetric Method), and checking how the difference $P_{TP} - P_{MC}$ varies as a function of P_{Tag}. In particular, the results in table 3.5 - where different values of P_{Tag} were practically obtained by varying the f_{brem} cut - and in fig. 3.28 clearly validate eq. 3.18.

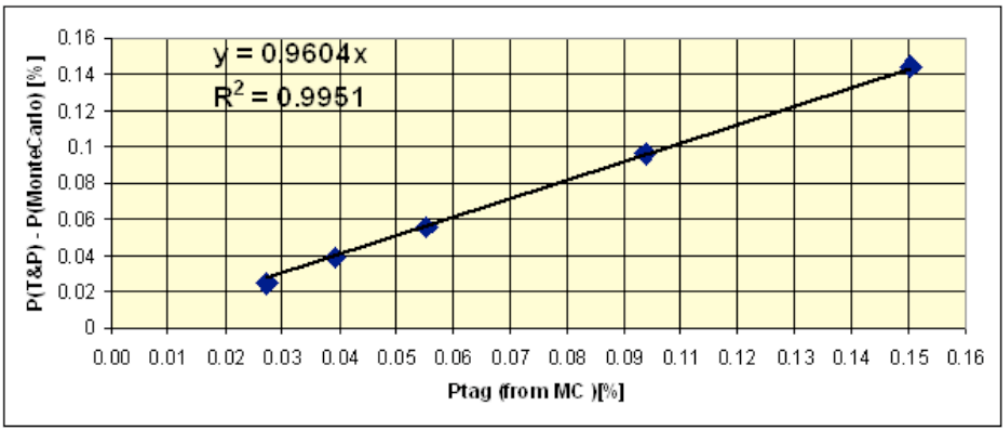


Figure 3.28: P_{Tag} systematics: $P_{T\&P}$ - P_{MC} fitted linearly as a function of P_{Tag}

The value of P_{Tag} can be measured using the Symmetric Method described above, by requesting a Tag selection on both legs of the Z decay. The resulting value can be subtracted to the P values obtained by the Tag&Probe in order to correct them and to almost cancel this systematic bias.

P_{tag}[%]	$P_{T\&P}$ [%]	P_{MC}[%]	$P_{T\&P}$ - P_{MC} [%]
0.029	1.522	1.498	0.024
0.041	1.535	1.496	0.039
0.058	1.600	1.5441	0.056
0.099	1.574	1.478	0.096
0.147	1.735	1.5903	0.144

Table 3.5: Systematics related to $P_{tag} > 0$

Invariant mass window

In the Tag&Probe method, Tag-Probe invariant mass is required to stay inside a region just around the Z peak; changing this mass window can affect the measurement, and in order to estimate this systematic uncertainty, the Tag&Probe analysis was repeated using different reasonable definitions for the dielectron invariant mass window. This sensitivity study has been performed without inserting the background; this somehow brings to an underestimation of the systematic uncertainty, but the presence of a spiky background spectrum would have spoiled the results even further.
The uncertainty related to the particular choice of the mass window is around 0.03%.

Background level

The relative uncertainty of the background yield is, in a conservative approach, as high as 300%: this is predominantly stemmed in the uncertainty about the "fake electrons" yield.
In order to estimate the corresponding systematic uncertainty in the measure of P, background event weight has been multiplied by an additional "enhancement factor"; the results on P are shown in fig. 3.7, and show an upper limit on this uncertainty of around (1.43 - 1.38) % = 0.05 %.

3.3 Summary of the Results

The two method described here to measure the electron charge misidentification probability show good effectiveness.
The Symmetric Method provides a global, integrated value of mischarge probability; it suffers from systematics related to the procedure of background subtraction.
The Tag&Probe method provides values of mischarge probability binned with respect to any other electron quantity (i.e. η, p_T , f_{brem}) and shows well-contained systematic uncertainties.
The Symmetric Method is useful to provide a correction to the Tag & Probe method that takes into account the charge misidentification of the Tag electron.

3.4 Applications to physics analyses

Measuring the electron charge misID at the CMS detector is important not only per se - since it gives a figure of merit of the reconstruction algorithm - but also because charge misID

Mee window $[GeV/c^2]$	P_{TP} [%]	P_{MC} [%]
85-95	1.37 ± 0.03	1.33 ± 0.03
80-100	1.39 ± 0.03	1.34 ± 0.03
70-110	1.40 ± 0.03	1.35 ± 0.03

Table 3.6: Systematic uncertainty related to the choice of the dielectron invariant mass window

Background enhancement factor	P_{TP} [%]	P_{MC} [%]
1 (no enhance)	1.38 ± 0.09	1.35 ± 0.06
2	1.41 ± 0.09	1.35 ± 0.06
3	1.43 ± 0.09	1.35 ± 0.06

Table 3.7: Systematic uncertainty related to the background yield: background event weight has been multiplied by an "enhancement factor"

plays a crucial role in some physics analyses with electron in the final states, both inside and outside the Standard Model:

- One of the first Standard Model measurement at the CMS detector will be the W^+/W^- cross section ratio (or the W boson charge asymmetry), in which electron charge misID dilutes the measurement
- A number of theories beyond the Standard Model predicts events with a peculiar signature given by two same-sign leptons.

  For example, in models where the Higgs is a pseudo-Goldstone boson [38] heavy partners of the top quark may exist with electric charge +5/3 ($T_{5/3}$) that decay to two same-sign W bosons and a *b*-jet. If the new-fermion masses are below 1 TeV, it can be discovered with 100 pb^{-1}at the LHC. The main background to this channel is given by $t\bar{t}$+jets, that can be subtracted away using the same-sign requirement; the potential of this discovery is therefore highly affected by the charge misID rate.
- Majorana neutrinos can be produced at the LHC mainly through the decay from a W boson ($W \rightarrow Ne$); since the Majorana neutrino itself then decays into a W boson and a lepton ($N \rightarrow$W e), this channel gives two same-sign electrons. Even in this case, knowing and reducing the charge misID rate is extremely important in order to state the significance of a discovery [39].

3.4.1 Improvement of the electron charge reconstruction

The charge of electrons in CMS is currently reconstructed as the sign of the curvature radius of the electron track; several different methods are under investigation that aim to show a better performance, i.e. a lower misidentification probability.
It is possible, for example, to define the charge of the electron (q) as the sign of the difference between the ϕ-coordinates of the electron track and the ϕ-coordinates of the electron calorimetric cluster:

$$q = \text{sign}(\phi_{track} - \phi_{cluster}) \tag{3.19}$$

where ϕ_{track} is measured at the innermost layer of the tracking detector. Preliminary results on these studies already show good performances. Fig. 3.29 shows the average value of $|\phi_{track} - \phi_{cluster}|$ for electrons coming from Z decay.

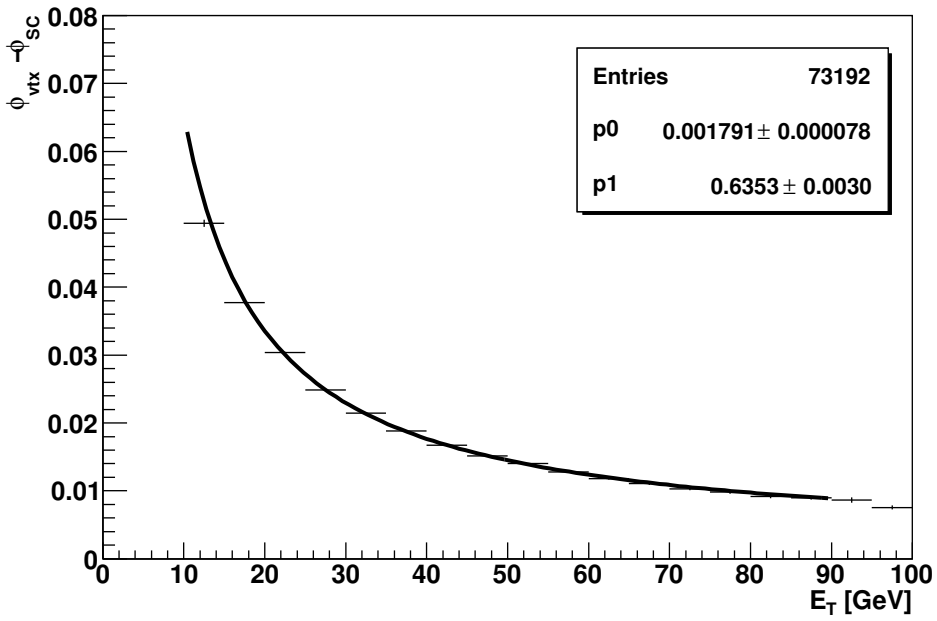


Figure 3.29: Average value of $|\phi_{track} - \phi_{cluster}|$ for electrons coming from Z decay, after HLT selection, as a function of electron transverse energy. The superimpose fit is $y = p_0 + p_1/x$.

The following, more articulate, definition of electron charge is currently under investigation in CMS, and will probably be applied to the first real data.
The charge stored in the electron object is chosen among three possible charges:

1. the charge value from the curvature of the GSF track associated to the electron ("GSF charge");
2. the charge value of the Kalman-filter track closest to the electron (if any) ("Kalman charge");
3. the charge value obtained using information from the Supercluster position and the first hits in the pixel detector ("pixel track")

If the GSF charge and the Kalman charge agree, their charge is blessed as the electron charge. If they disagree and there is a closest Kalman track, its charge is used instead. In they disagree and there is no Kalman track, the pixel charge is used.
It is important to stress that the methods described in this work, aimed to measure the electron charge misID, can be applied in a straightforward way, whatever charge definition is used by CMS.

Chapter 4

Charge misidentification correction to the W^+/W^- cross section ratio

The study of the production of W and Z events at the LHC is fundamental in several respects. First, the calculation of higher order corrections to these simple, colour singlet final states is very advanced, with a residual theoretical uncertainty smaller than 1%. Such precision makes W and Z production a stringent test of QCD.
Finally, a number of fundamental electroweak parameters can be accessed through W and Z final states (M_W , through the W boson decay distributions; $\sin^2\theta_W$, via the Z forward-backward asymmetry; lepton universality, by comparing electron and muon cross-sections). These measurements are long term applications where the understanding of the hadronic environment at the LHC is crucial, and to which the above-mentioned measurements are necessary inputs.
First physics measurement at the LHC will undoubtedly suffer from the large (order 10% at LHC startup) systematic uncertainty in luminosity measurement.
It is therefore resonable to be interested into cross section *ratios* like:

$$R \equiv \frac{\sigma(W^+ \to e^+\nu_e)}{\sigma(W^- \to e^-\overline{\nu}_e)} \equiv \sigma^+/\sigma^- \tag{4.1}$$

or *asymmetries* like:

$$A \equiv \frac{\sigma^+ - \sigma^-}{\sigma^+ + \sigma^-} \tag{4.2}$$

These measurements, indeed, do not suffer from the luminosity systematic uncertainty as well as some other systematics related to reconstruction efficiencies, that are factored out in the ratios.
This means that such cross-section ratios eventually become ratios between *number of events*.
The ratio in Eq. 4.1 is of great interest for the first LHC times, because it can provide useful information about Parton Distribution Functions (PDF) of quarks inside the proton at an unprecedent energy scale, and therefore at unprecedent values of (x, Q^2).

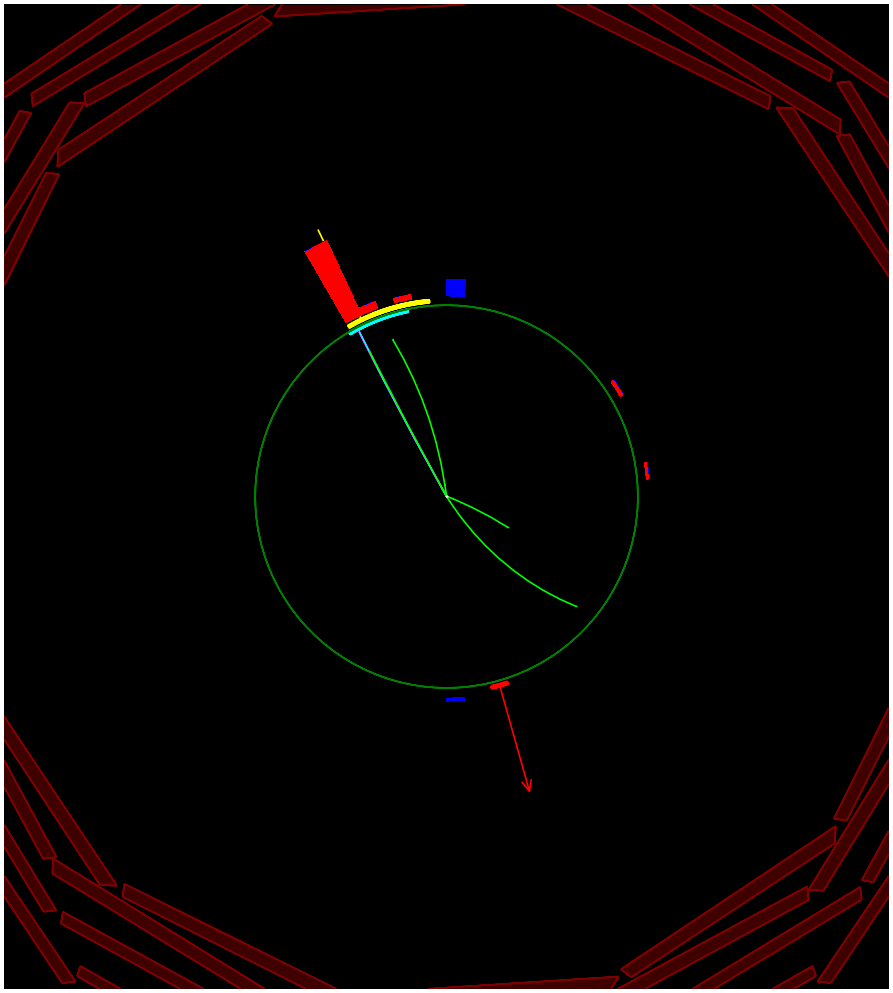


Figure 4.1: $R-\phi$ view of a fully simulated $W \to e\nu$ event in the CMS detector

4.1 $W \to e\nu$ event selection

4.1.1 Trigger and online reconstruction

A single electron trigger with a threshold of $E_T > 15$ GeV is used for the $W \to e\nu$ cross section measurements.
The trigger is seeded by a Level-1 ECAL trigger with a L1 E_T-threshold of 10 GeV and no explicit isolation requirements. Using the L1 candidates as a starting point, the HLT algorithm regionally reconstructs an ECAL supercluster.
The HLT supercluster seeds an electron tracking algorithm with a large window for searching for pixel hits consistent with a track. In the low instantaneous luminosities anticipated for the first 10 pb^{-1}, no further electron selection is expected to be required at this E_T threshold. Higher background rates will require further background rejection, either via tighter pixel match criteria or via loose application of isolation or identication criteria used offline.

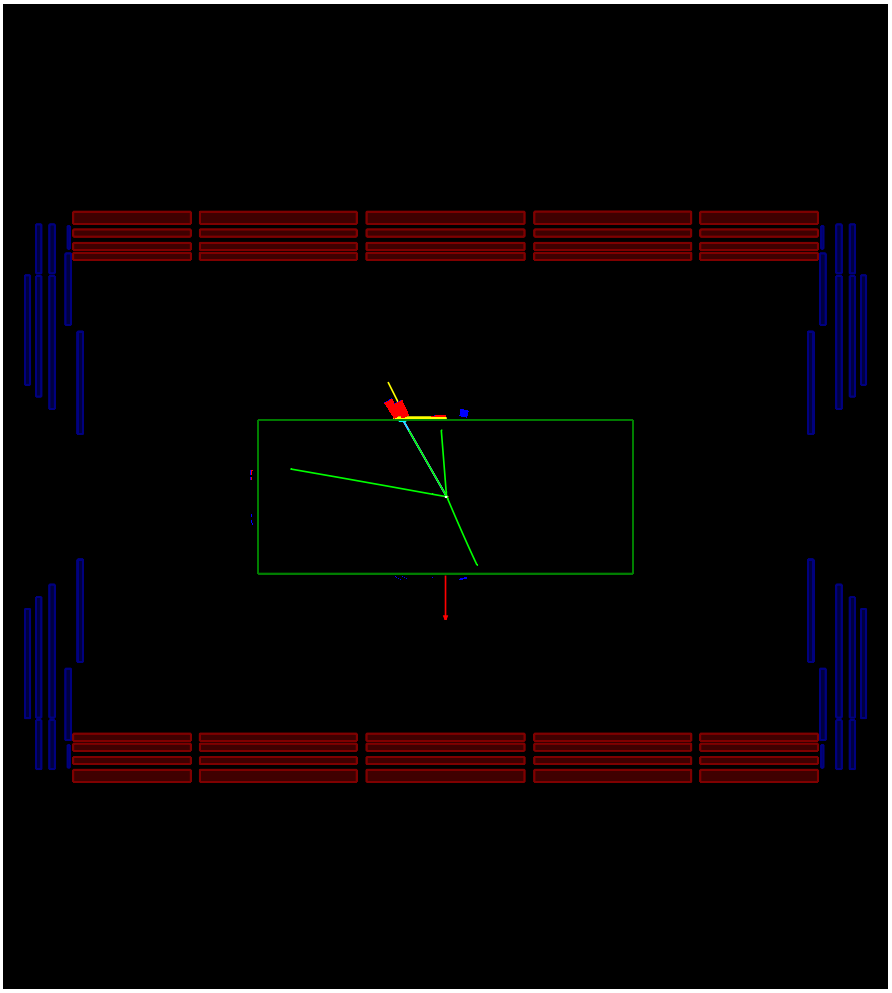


Figure 4.2: $R - z$ view of a fully simulated $W \to e\nu$ event in the CMS detector

	$\sigma_{i\eta i\eta}$	$\Delta\eta_{in}$	$\Delta\phi_{in}$
Barrel	0.0099	0.025	0.0040
Endcap	0.028	0.020	0.0066

Table 4.1: Electron identification criteria for $W \to e\nu$ candidates

	Track p_T [GeV/c]	ECAL E_T [GeV]	HCAL E_T [GeV]
Barrel	2.2	4.2	2.0
Endcap	1.1	3.4	1.3

Table 4.2: Isolation criteria for $W \to e\nu$ candidates

4.1.2 Electron identification and isolation

For the offline selection, the intent is to use only simple criteria expected to be well understood in the early phases of data-taking. For the $Z \to e^+e^-$ selection, cut thresholds are loose since the background is small, while, for instance, for the $W \to e\nu$ selection tighter cut thresholds are going to be used in order to reduce the large QCD background.
The shape variable $\sigma_{i\eta i\eta}$ measures the RMS shower width in the η direction. The $\Delta\eta_{in}$ and $\Delta\phi_{in}$ variables quantify the geometric match between the GSF track trajectory and the ECAL supercluster.
Electrons are also required to pass isolation criteria. The isolation variables consist of sums of ECAL RecHits, HCAL sections of the CaloTowers, and track pT within regions (cones) of $\Delta R < 0.4$ made for ECAL, HCAL and track isolation, respectively. The regions were centered on the supercluster position for the calorimetric isolation variables, and on the track direction at the vertex for the track isolation. In the case of the ECAL, a threshold is placed on the RecHits to be included in the sum. In all three cases, the possible track or energy footprint of the electron is removed by excluding an inner cone, and in the case of the ECAL sum, an additional narrow strip in the ϕ direction. More details on the isolation definitions can be found in [40].

Cut optimization

The values of the selection cuts are chosen using an algorithm that maximizes the background rejection power for a given signal efficiency [40]. The algorithm maps out a trajectory in efficiency versus rejection space, and each point on this trajectory corresponds to a set of cuts, where the rejection power is shared between the cuts so as to maximize the total background rejection for a given signal efficiency. A discrete set of cut values is considered in this optimization, but the optimum is found to be independent of the spacing.
The algorithm is supplied with candidate electrons: a signal sample and a background sample. A set of variables on which to cut is defined. At each step in an iteration it is required to improve the signal/background ratio by a small increment by tightening the cut on a

single selection variable, that is the one that decreases the signal efficiency by the smallest amount. The algorithm rapidly converges on the trajectory in efficiency versus rejection space that consists of the sets of cuts that maximize background rejection for any given signal efficiency (and vice versa). It has been demonstrated that the algorithm is rather tolerant of contaminated signal and background samples. Nearly optimal sharing of rejection power between the selection variables is still achieved even using rather impure input samples. In tests described in [41], the background sample was taken as the single electron sample with $\not\!\!E_T < 20$ GeV (and also some cuts to reject $Z \to e^+e^-$). A rather pure input signal sample can be derived from $Z \to e^+e^-$, but the algorithm works well with a signal sample defined as the single electron sample with $\not\!\!E_T > 30$ GeV . With 10 pb^{-1}data samples input, selections were obtained from the algorithm that were at most 1 or 2% less efficient for signal than the optimum for a given background rejection.

4.1.3 $W \to e\nu$ selection

The $W \to e\nu$ events are selected from events that pass the single electron High Level Trigger. A high-p_T electron is required, which is formed from the association of a high-E_T ECAL supercluster and a high-p_T GSF track in the Tracker; the offline candidate electron must also match the direction of the HLT candidate electron.
The absence of a second reconstructed GSF electron with supercluster $E_T > 20$ GeV is required. This cut reduces the $Z \to e^+e^-$ background by a factor of 3, reducing slightly the $W \to e\nu$ signal efficiency by 0.9%.
An additional criterion that could help in background rejection would be to apply a vertex compatibility requirement on the tracks. It would be particularly helpful in rejecting jets from secondary vertices like b-jets and electrons from converted photons. Especially for the first measurements such a selection criterion is expected to have reduced rejection power due to the misalignment of the Tracker. It can be useful in higher integrated luminosities when the Tracker alignment will be improved. No impact parameter or lifetime cuts are applied in the present analysis.
Since the leptonic W decay gives an undetectable neutrino, the W candidate events should show an imbalance of the measured momentum. Since the colliding partons have an overall $p_T \simeq 0$, the missing transverse energy in the event is identified with the neutrino p_T.
The following online and offline selection has been used for the $W \to e\nu$ cross section analysis:

- event passes the single electron HLT
- one electron in ECAL fiducial volume ($|\eta| < 2.5$ with $1.4442 < |\eta| < 1.560$ excluded)
- electron supercluster having $E_T > 30$ GeV
- electron passes Isolation criteria as defined in Table 4.2
- electron passes Electron ID criteria as defined in Table 4.1
- no second electron in the event with $E_T > 20$ GeV

Selection Criterion	Efficiency for $W \to e\nu$ [%]
single electron HLT	94.1
GsfElectron, $E_T > 30$ GeV, in fiducial	98.2
track isolation	93.0
ECAL isolation	94.7
HCAL isolation	97.4
passes ID criteria as defined in Table 4.1	93.0
no 2nd electron with $E_T > 20$ GeV	99.1

Table 4.3: Signal selection efficiencies relative to $W \to e\nu$ events in which a supercluster with $E_T > 30$ GeV was falling within the ECAL fiducial volume.

The efficiencies of the selection criteria on a signal sample can be found in Table 4.3. For the correct interpretation of the numbers it is necessary to take into account that the selection was applied on a fully simulated $W \to e\nu$ sample without pile-up in which the electrons were falling within the kinematic and geometrical acceptance (a supercluster within the ECAL fiducial with $E_T > 30$ GeV).

Cut optimization

The number of events passing the above selection is shown in Table 4.4; the overall S/B ratio is around 2.3.
Fig. 4.3, 4.4 and 4.5 show some reconstructed quantities for signal and background events. In particular, Fig. 4.3 underlines how distinctive a variable is the missing transverse energy for $W \to e\nu$ with respect to the background channels, while Fig. 4.5 shows how the signal-to-background ratio varies along $|\eta|$, with a minimum value of around 1 at $|\eta| > 2$ where the contribution of QCD events becomes larger.

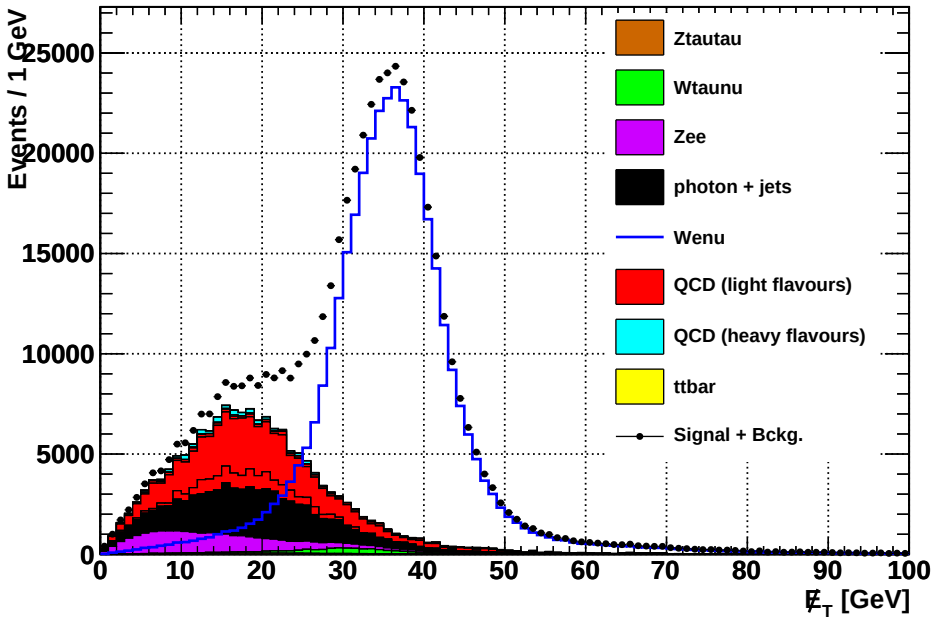


Figure 4.3: Missing transverse energy distribution for signal and background events passing both the online and offline selection

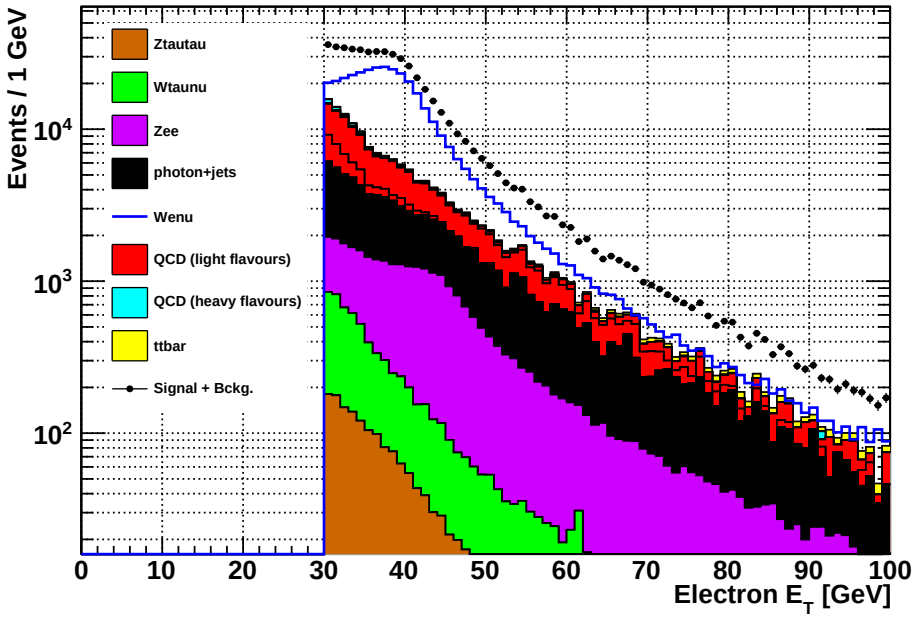


Figure 4.4: Electron transverse energy distribution for signal and background events passing both the online and offline selection

Physics channel	Weighted selected events
$W \to e\nu$	371937
$Z \to \tau\tau$	1589
$W \to \tau\nu$	5018
$Z \to ee$	23650
$\gamma + \text{jets}(15 < \hat{p_T} < 20)$	837
$\gamma + \text{jets}(20 < \hat{p_T} < 25)$	2573
$\gamma + \text{jets}(25 < \hat{p_T} < 30)$	7722
$\gamma + \text{jets}(30 < \hat{p_T} < 35)$	11357
$\gamma + \text{jets}(\hat{p_T} > 35)$	30649
QCD light flavours $(20 < \hat{p_T} < 30)$	12825
QCD light flavours $(30 < \hat{p_T} < 80)$	55384
QCD light flavours $(80 < \hat{p_T} < 170)$	4444
QCD heavy flavours $(20 < \hat{p_T} < 30)$	480
QCD heavy flavours $(30 < \hat{p_T} < 80)$	3499
QCD heavy flavours $(80 < \hat{p_T} < 170)$	245
$t\bar{t}$	1875

Table 4.4: Number of signal and background events passing the selection for a statistics of 100 pb^{-1}

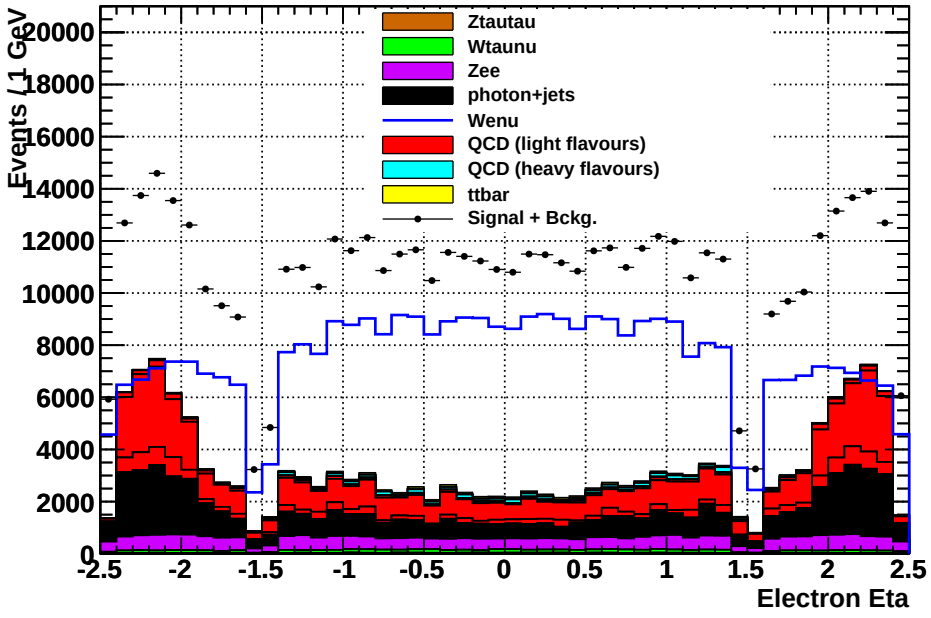


Figure 4.5: Electron pseudorapidity distribution for signal and background events passing both the online and offline selection

4.2 Sign asymmetry

In the proton-proton collisions at the LHC, the initial state is charge-asymmetric; therefore, the production of positively-charged W bosons is expected to be enhanced with respect to negatively-charged W bosons (the overall W^+/W^- ratio is greater than 1).
W^+ are produced by u valence (or sea) quarks while W^- are produced by d quarks.
Since in the proton u-type quarks carry, on average, greater momentum fraction than d-type quarks (see Fig. 4.6), W bosons heavily boosted along the beam direction are more likely to be W^+ than W^-.
This has the consequence that the W^+/W^- ratio tends to increase with increasing W rapidity; since the W rapidity and the pseudorapidity of the electron coming from the $W \to e\nu$ decay are correlated (high-pseudorapidity electrons come from high-rapidity W bosons, as shown in fig. 4.10), the electron sign-ratio e^+/e^- is expected to increase with increasing electron pseudorapidities.
This is confirmed by the signal plot in Fig. 4.7.

In addition to the difference in the PDF of up/down quarks, there is another concurrent effect that contributes to determine the e^+/e^- pattern for the signal in Fig. 4.7: if a W boson is produced by one valence quark and by a sea antiquark, since W are produced by left-handed quarks the W produced by valence quarks of a given beam are 100% polarized along the direction of that beam. The overall W polarization, at a given rapidity of the W boson, is a function of the fraction of W bosons produced by each of the two beams, so it is zero at $y = 0$ but becomes significant at (even moderately) large W-rapidities.
Polarized W^+ and W^- behave differently in their decay: the positron (antimatter) will be more likely emitted in the direction opposite to the W^+ rapidity and electron (matter) will be more likely emitted in the direction of W^- rapidity.
This polarization effect somehow dilutes the PDF effect described above: anyway both these mechanisms are included in the MonteCarlo simulation and Fig. 4.7 shows their overall, combined effect.

As far as the background is concerned, it is interesting to observe how Fig. 4.7, 4.8 and 4.9 show that:

- the background as a whole does not show a significant charge asymmetry. This property of the background can be useful when designing a background-subtraction strategy, as outlined later in Section 4.4.

- among the electroweak background channels, selected $W \to \tau\nu$ events show a e^+/e^- behaviour very similar to $W \to e\nu$: this is expected since τ's are produced charge-asymmetrically, and the final state electron has the same charge of the τ because it comes from the $\tau \to e\nu_e\nu_\tau$ decay.

Fig. 4.11 and 4.12 show the pseudorapidity and transverse energy distribution of the reconstructed positive/negative electrons coming from the decay of positive/negative W bosons

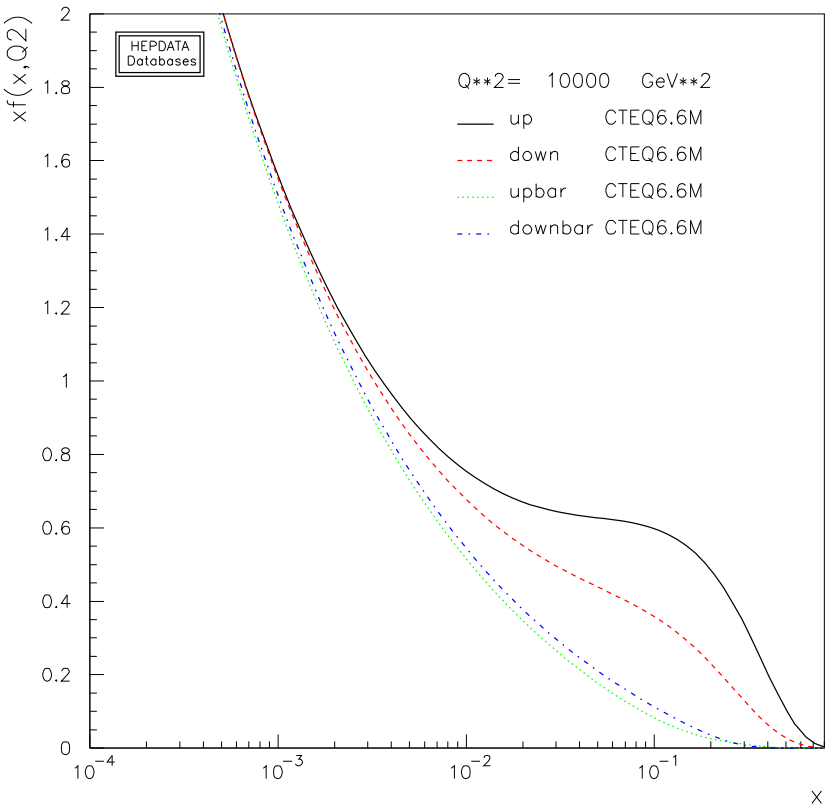


Figure 4.6: Parton Distribution Functions for u- and d-type quarks and antiquarks in the proton at $Q^2 = (10 \text{ TeV})^2$: the plot shows how, on average, larger momentum fraction is carried by u-type than d-type quarks in the proton.

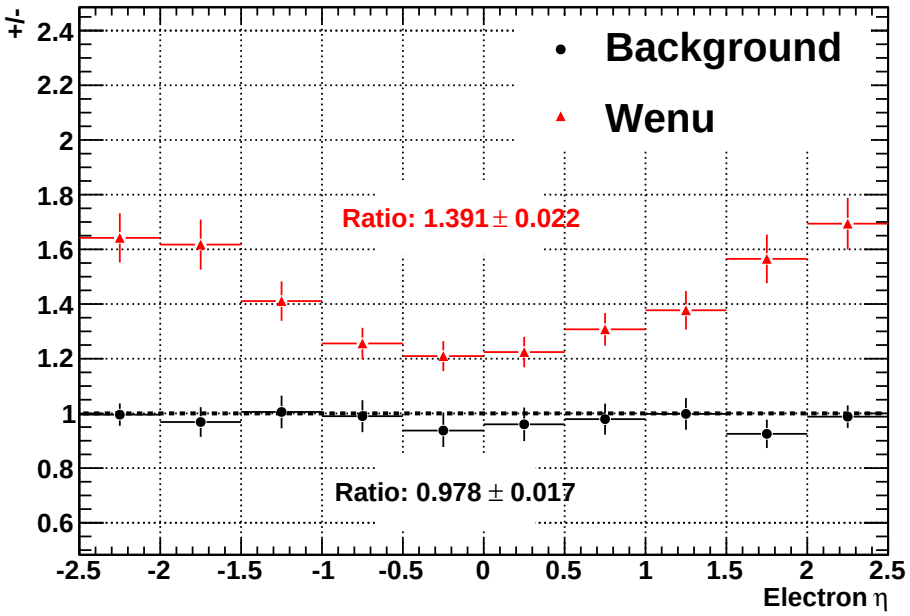


Figure 4.7: e^+/e^- ratio as a function of electron pseudorapidity for $W \to e\nu$ and background reconstructed events

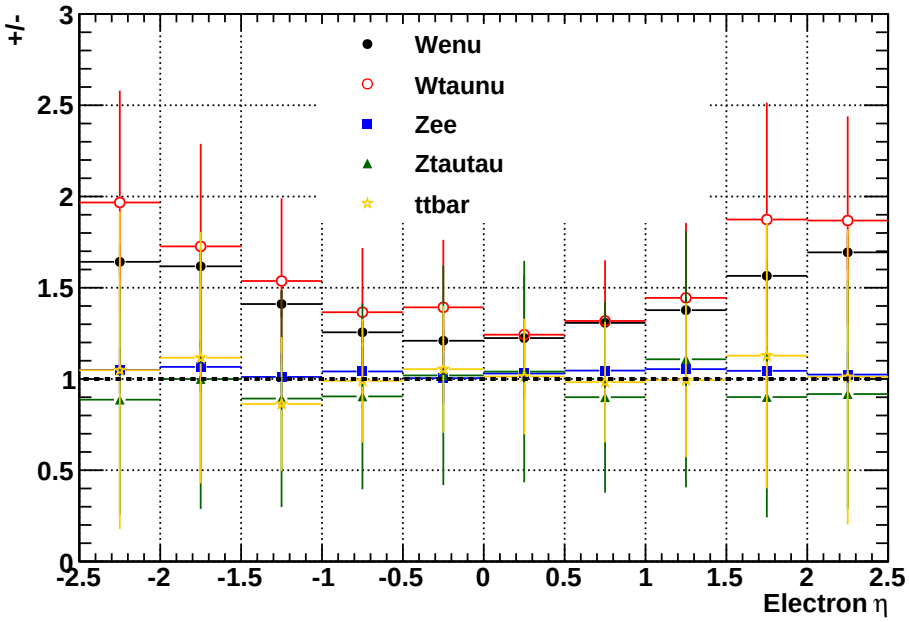


Figure 4.8: e^+/e^- ratio as a function of electron pseudorapidity for reconstructed events belonging to different electroweak backgrounds

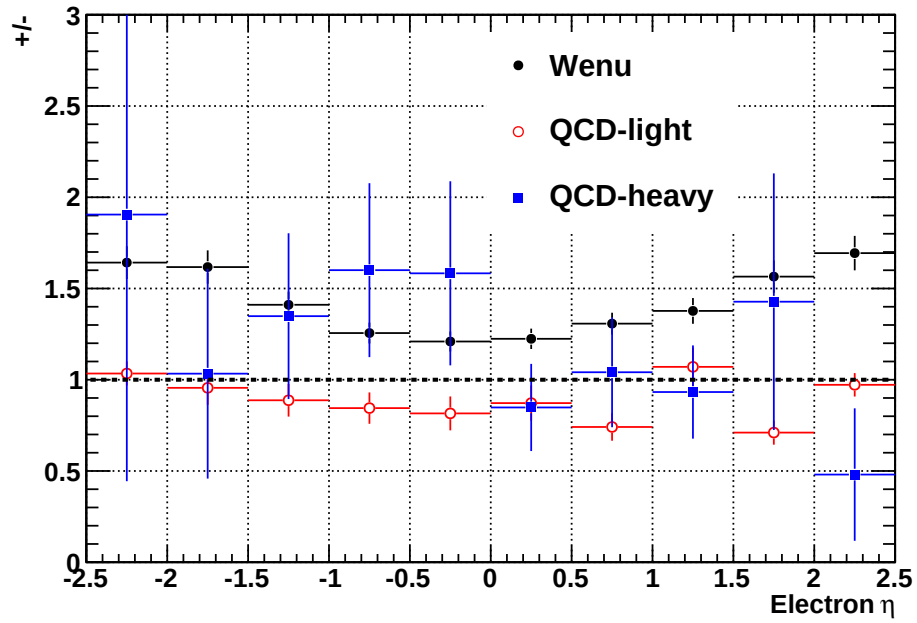
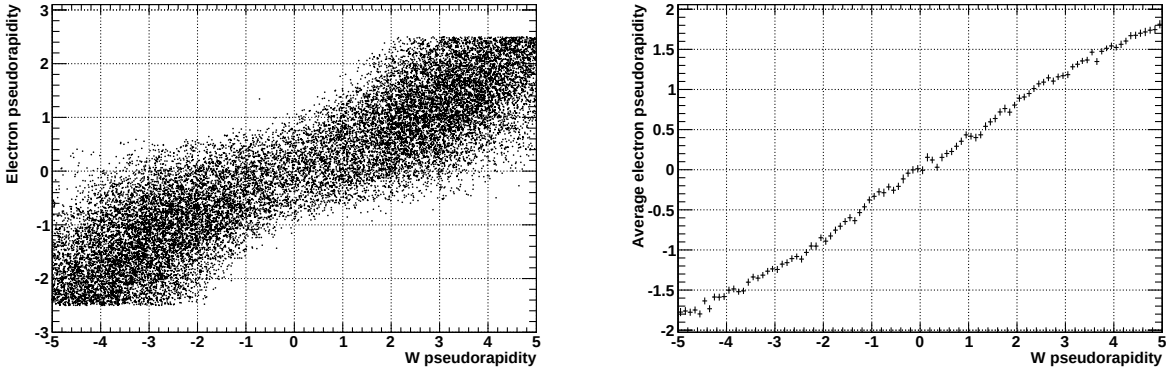


Figure 4.9: e^+/e^- ratio as a function of electron pseudorapidity for reconstructed events belonging to different QCD backgrounds



(a) 2D plot

(b) Average value of W-electron pseudorapidity vs W boson pseudorapidity

Figure 4.10: W-electron pseudorapidity vs W boson pseudorapidity (generator level). A generator-level filter excludes events where $|\eta|$ of the electron is greater than 2.5.

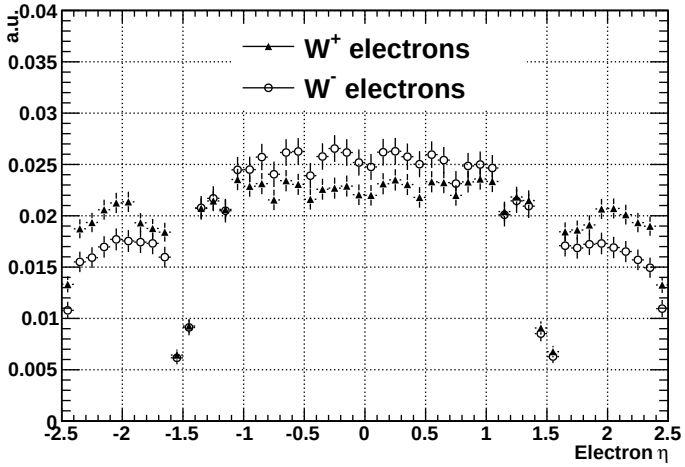


Figure 4.11: Reconstructed pseudorapidity distribution of positive/negative electrons coming from positive/negative W bosons

respectively (except for the misID phenomenon that will be covered later on in this chapter): the transverse energy distributions are very similar, while the pseudorapidity plot shows how negative electrons populate less than positrons the high-pseudorapidity regions of the detector.
The experimental method to measure the integrated W^+/W^- ratio in the electron channel consists of measuring the integrated e^+/e^- ratio: once the background has been properly subtracted, in the absence of electron charge misID the e^+/e^- ratio indeed corresponds to the W^+/W^- ratio.
In a physics analysis perspective, both the integrated and the differential e^+/e^- ratio (as a function of the electron pseudorapidity or transverse energy) provide a useful tool to constraint the Parton Distribution Functions (PDF), as highlighted later in Section 4.5.
As the cross section is directly (or inversely) proportional to each of these factors, the square of the total relative uncertainty of the cross section is simply the sum in quadrature of the relative uncertainties of the above factors.

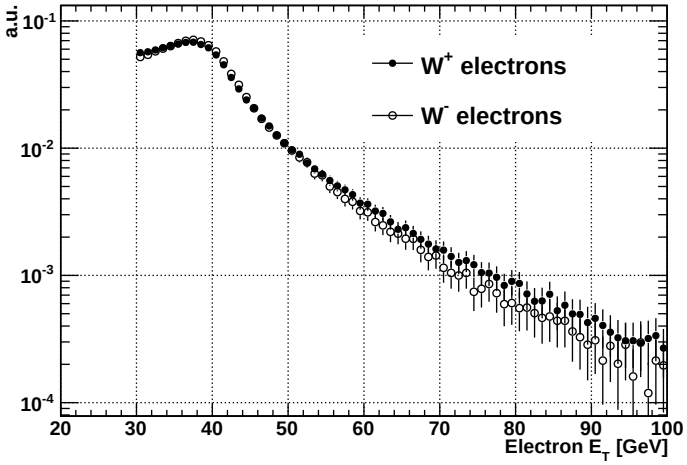


Figure 4.12: Reconstructed transverse energy distribution of positive/negative electrons coming from positive/negative W boson: transverse energy is defined of the transverse energy of the electron superCluster in the ECAL

4.3 Insertion of a charge misID correction

4.3.1 Effect of the charge misID on the W^+/W^- ratio

In each bin, if $N^{\pm}$ ($T^{\pm}$) denotes the observed (true) number of electrons of positive/negative charge, and P denotes the charge misID rate, then:

$$\frac{N^+}{N^-} = \frac{T^+(1-P) + PT^-}{T^-(1-P) + PT^+} \tag{4.3}$$

After some algebrical manipulations, if $R_{meas} = N^+/N^-$ and $R^{corr}_{meas} = T^+/T^-$, it yields:

$$R^{corr}_{meas} = \frac{(1-P)R_{meas} - P}{(1-P) - PR_{meas}} \tag{4.4}$$

Eq. 4.3 shows how, if the real ratio R_{true} were equals to 1, the measured ratio would not be affected by charge misID; on the contrary, when the measured ratio is greater than 1 (as it is the case in the W^+/W^- ratio), the existence of the charge misID phenomenon brings to a dilution of the measured ratio ($R_{meas} < R_{true}$).
Eq. 4.4 also shows that the misID corrections become particularly important at high values of $|\eta|$, where both R_{meas} and P become larger: Fig. 4.13 shows the values of the correction R^{corr}_{meas}/R_{meas} as a function of R_{meas} and charge misID rate.

4.3.2 Measurement and insertion of the charge misID correction

In order to compute charge misID rates specifically applicable to the W^+/W^- ratio analysis, it is necessary to follow the Tag-Probe procedure described in Chapter 3 of the present work,

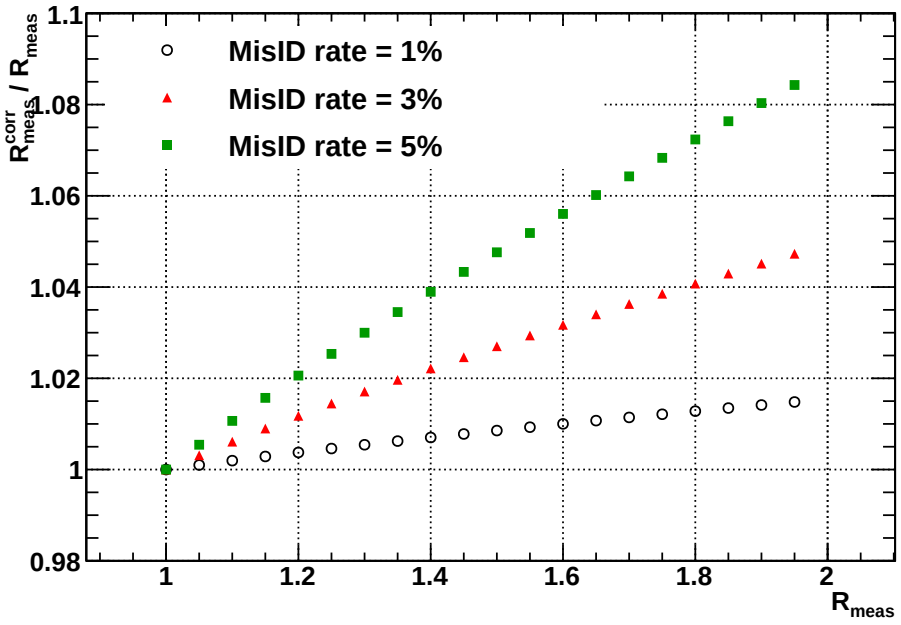


Figure 4.13: Values of the correction R^{corr}_{meas}/R_{meas} as a function of R_{meas} and charge misID rate.

selecting events where the Probe electrons are required to pass the exact same selection criteria that are requested in the $W \to e\nu$ analysis.
Fig. 4.14 shows the charge misID rate as a function of the electron pseudorapidity, for electrons passing the $W \to e\nu$ selection criteria and an integrated luminosity of 100 pb^{-1}.
The resulting integrated misID rate (obtained from signal only, since background is assumed to be safely negligible with such a tight selection) is:

$$P = (1.27 \pm 0.08)\% \tag{4.5}$$

that is to be compared with a MonteCarlo value of:

$$P_{MC} = (1.22 \pm 0.08)\% \tag{4.6}$$

Using negative and positive Probes indifferently in order to determine the Probe charge misID is obviously desirable, since the Probe selection is quite stringent and the available event statistics is limited.
On the other hand, using negative and positive Probes indifferently is possible only if the reconstruction algorithms in CMS have the same quality for negative and positive electrons: if a bias were present, it would be then necessary to separately compute charge misID rates for negative and positive electrons.
Fig. 4.15 shows that charge symmetry of the misID rate is verified along electron pseudorapidity: this legitimates the merged use of negative and positive Probes to define a common

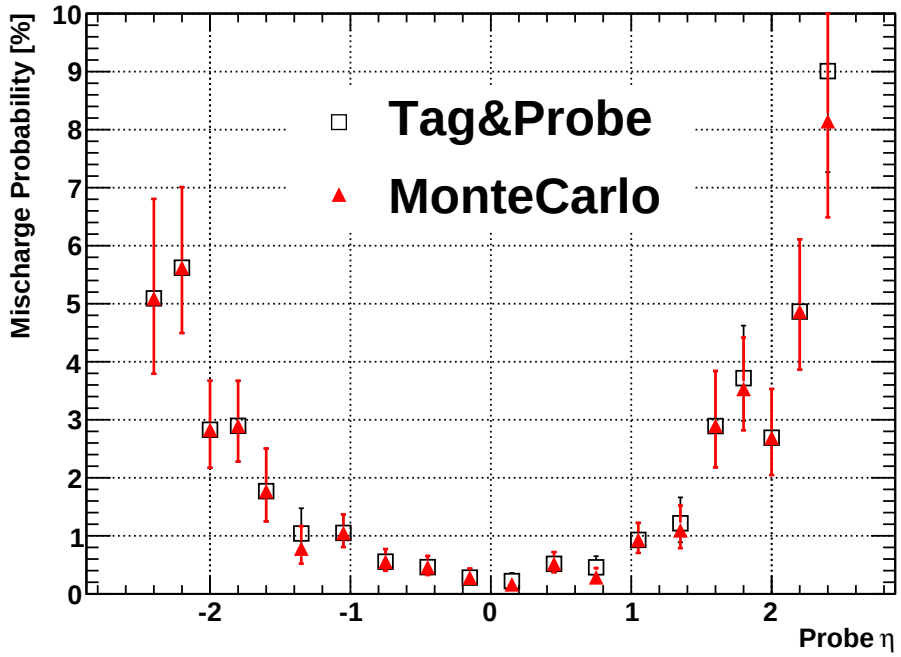


Figure 4.14: MisID rate from Tag and Probe method, as a function of electron pseudorapidity: Probe electrons are selected using W-electron criteria

misID rate.
Fig. 4.17 shows the effect of inserting such a misID correction on the $W \rightarrow e\nu$ signal (no background considered here): the integrated value of R is brought closer to the MonteCarlo value, with appreciable corrections at high electron pseudorapidity.

$$R_{MC} = 1.435 \pm 0.004 \tag{4.7}$$

$$R_{reco} = 1.397 \pm 0.004 \tag{4.8}$$

$$R_{reco}^{corr} = 1.425 \pm 0.005 \tag{4.9}$$

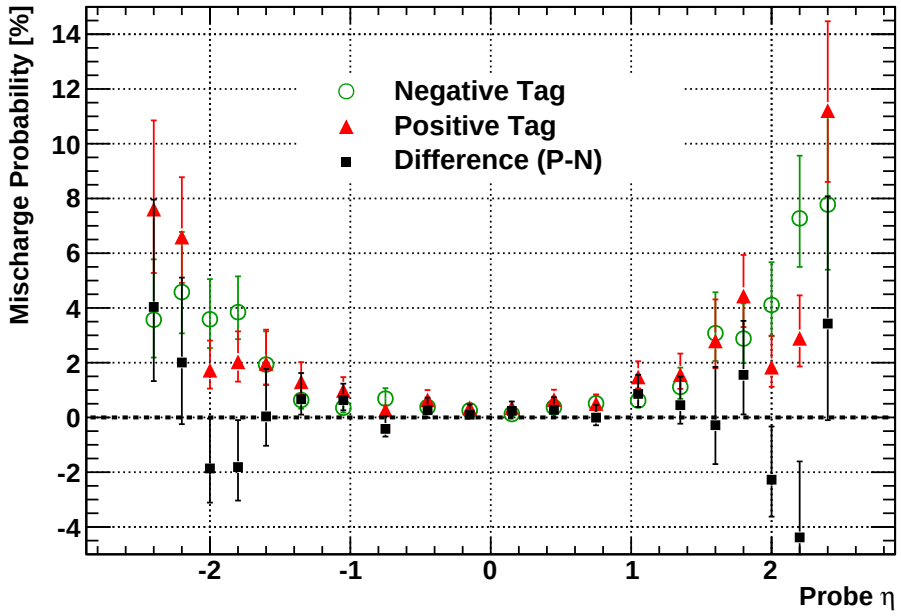


Figure 4.15: Charge symmetry of the charge misID rate: Probe electrons are selected using W-electron criteria

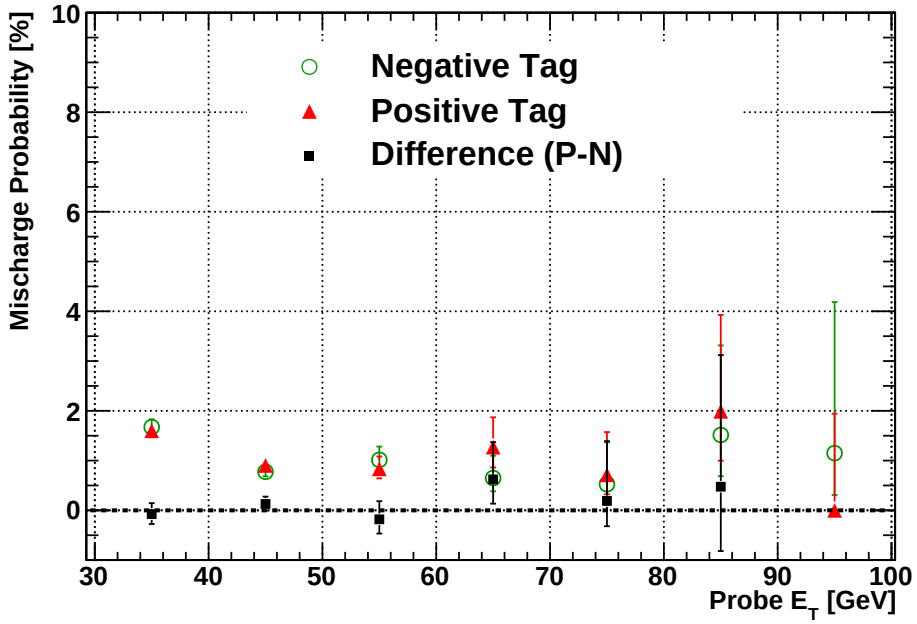


Figure 4.16: Charge symmetry of the charge misID rate: Probe electrons are selected using W-electron criteria

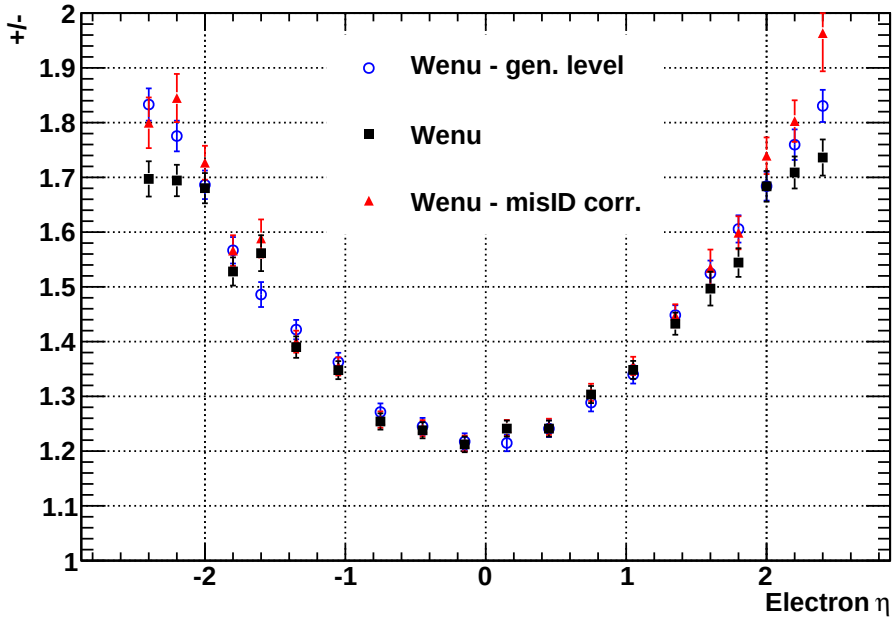


Figure 4.17: W^+/W^- ratio as a function of the W-electron pseudorapidity

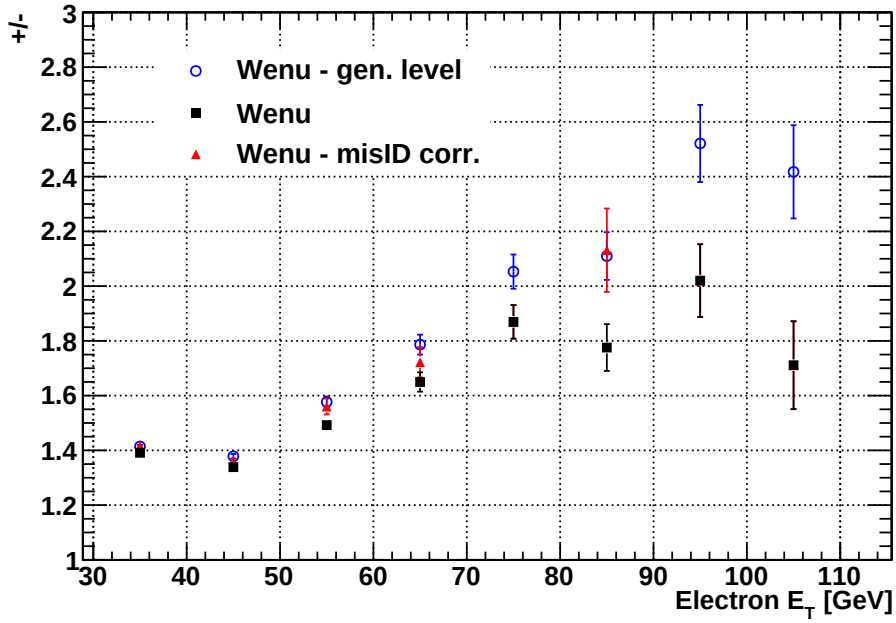


Figure 4.18: W^+/W^- ratio as a function of the W-electron transverse energy

4.4 Background subtraction: strategy proposals

For the first LHC data, two methods can be possibly applied to this analysis for background subtraction, depending on whether one wants an integrated or binned value of the W^+/W^- ratio.
The basic idea is that measurements and backgroound-subtraction strategies should be kept as robust and simple as possible at the LHC startup.

4.4.1 Integrated W^+/W^- ratio with integrated charge misID

The simplest measurement is a single, integrated value of R, without any binning; in this case, one can proceed as follows:

1. Apply the $W \to e\nu$ selection to the data
2. Estimate the number of background events B, i.e. making use of the strategy implemented by CMS in [40].
3. Use the MonteCarlo information (shown in Fig. 4.7) that the background is equally divided into positive- and negative- sign events, i.e. $B^+ = B^- = B/2$: separate the data sample into positive- and negative-charge events, and subtract $B/2$ from each
4. Compute the background-subtracted ratio R
5. Correct the ratio using Eq. 4.4, where P comes to be the integrated value of misID in Eq. 4.5

4.4.2 Binned W^+/W^- ratio

A slightly more articulated strategy could look at the quite strong η-dependence of both the background contamination (see Fig. 4.5) and the charge misID rate (see Fig. 4.14), neglecting the residual, weaker E_T-dependence in order to avoid complications arising from the poor bin statistics.
In this case, the *shape* of the η-distribution of the background can be extracted from MonteCarlo (as in Fig. 4.5), while the normalization can be extracted from data.
The difference with respect to the integrated results gives a hint on the size of the approximation that one makes when assuming a flat misID.

4.5 Discrimination of Parton Distribution Functions

The LHC will be able to explore the Parton Distribution Functions of the proton by getting information from regions of the (x, Q^2) [1] plane not covered by precedent experiments (see Fig. 4.19).

[1] x is the fraction of proton momentum carried by the quark and Q^2 is the squared momentum transferred to the quark

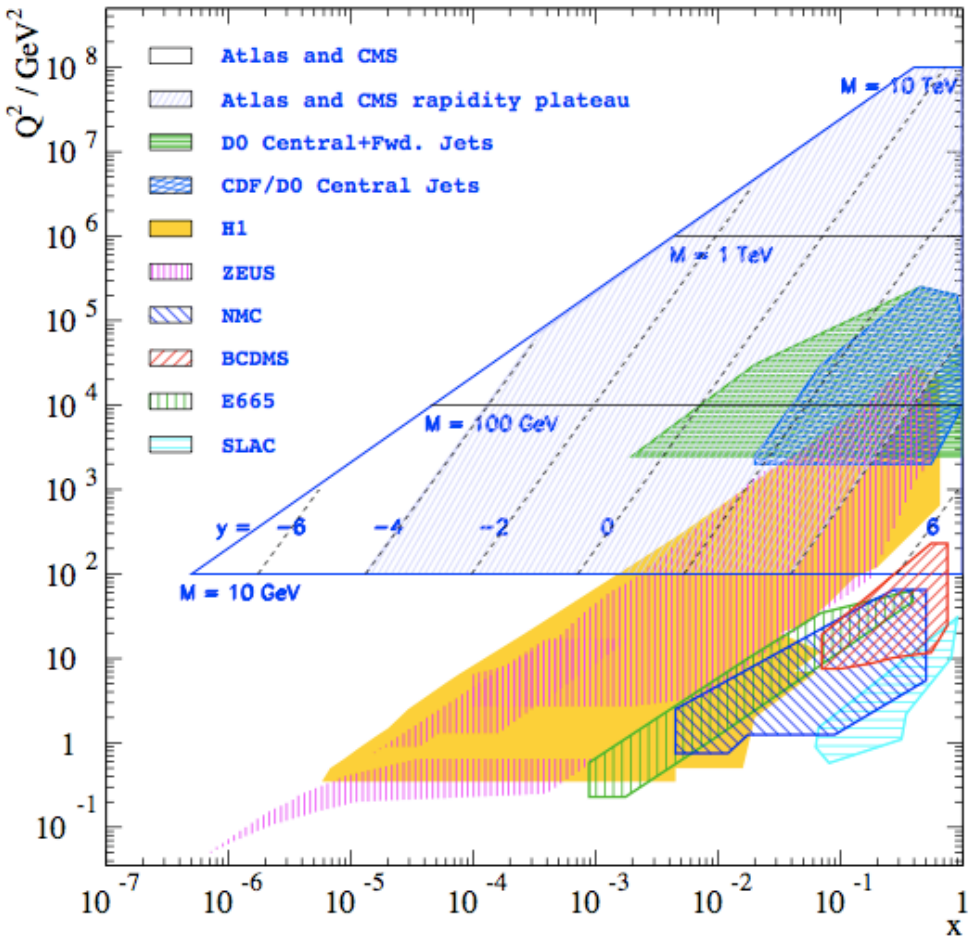
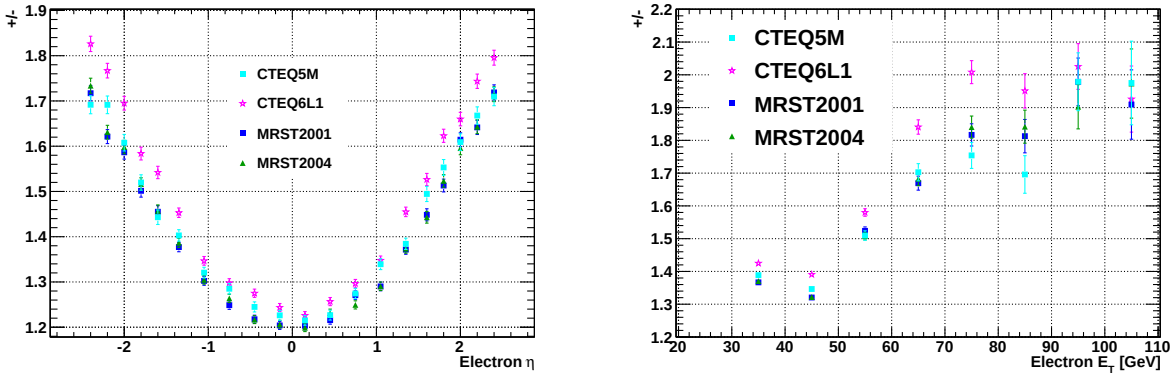


Figure 4.19: Coverage of the (x, Q^2) plane provided by high-energy experiments.



(a) Pseudorapidity

(b) Transverse energy

Figure 4.20: W^+/W^- asymmetry as a function of the electron pseudorapidity, as obtained from different PDF models

Different PDFs can give rise to different W^+/W^- distributions (see Fig. 4.20); measuring W^+/W^- ratio, and adequately correcting it for electron charge misID, can provide a useful tool to constrain PDF parameters and/or models from data.
In order to give a feeling of such discrimination power with 100 pb^{-1}of data, Fig. 4.22 shows the W^+/W^- measured (and misID-corrected) ratio superimposed to the MonteCarlo predictions obtained with the CTEQ 6L1 - that is the PDF model that has been used by the CMS collaboration to generate the $W \to e\nu$ events analyzed in the present work - and other PDF models at generator level.
The predictions of the various PDFs have been simulated by making use of the LHAPDF libraries [42] within the PYTHIA generator [33].
In particular, the following PDFs were simulated in PYTHIA:

- CTEQ 6L1 (LO) [PYTHIA parameter MSTP(51) = 10042]
- CTEQ 5M (NLO, $\overline{\text{MS}}$ scheme), [PYTHIA parameter MSTP(51) = 19050]
- MRST 2001 (NLO, $\overline{\text{MS}}$ scheme) [PYTHIA parameter MSTP(51) = 20050]
- MRST 2004 (NLO, $\overline{\text{MS}}$ scheme) [PYTHIA parameter MSTP(51) = 20450]

Fig. 4.22 shows that, with a statistics of 100 pb^{-1}, it is possible to exclude some PDF models, like the CTEQ 5M; the discrepancies between data and CTEQ 5M become particularly evident for $|\eta| > 1.5$.
By comparing Fig. 4.22 and Fig. 4.21, where charge misID has *not* been inserted, it appears that the charge misID correction is important: with no misID correction, the data-CTEQ 5M discrepancy is not very evident.
The same effect can be observed when the e^+/e^- ratio is plotted as a function of the electron transverse energy, as in Fig. 4.24, 4.23 (both zoomed in Fig. 4.26, 4.25).

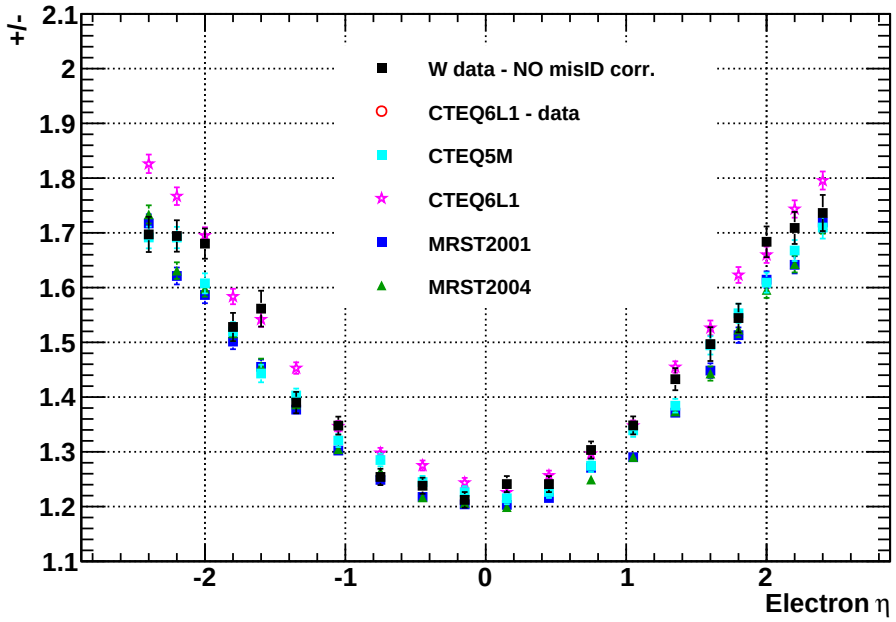


Figure 4.21: W^+/W^- asymmetry as a function of the electron pseudorapidity, as obtained from data without misID-correction (100 pb^{-1}), superimposed to the different PDF models

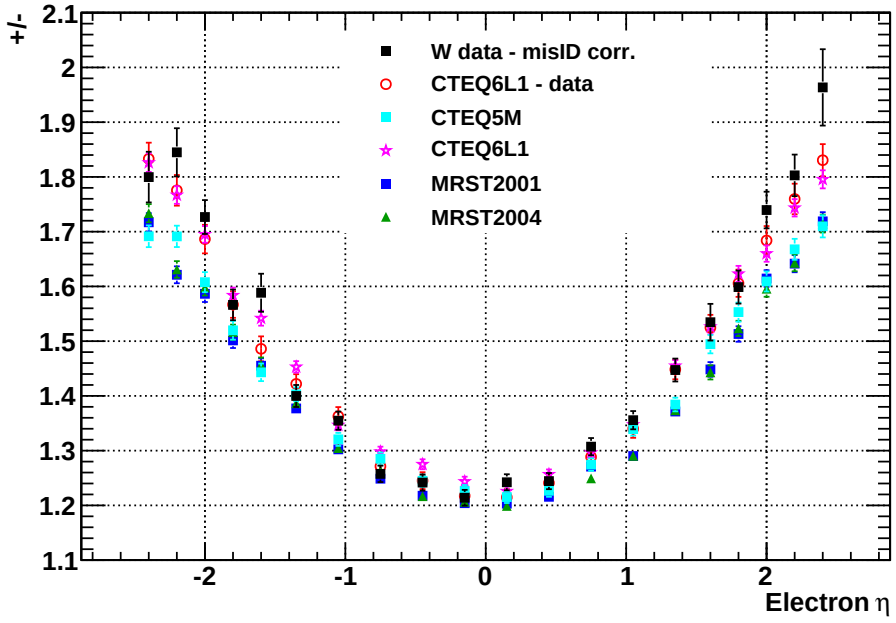


Figure 4.22: W^+/W^- asymmetry as a function of the electron pseudorapidity, as obtained from misID-corrected data (100 pb^{-1}), superimposed to the different PDF models

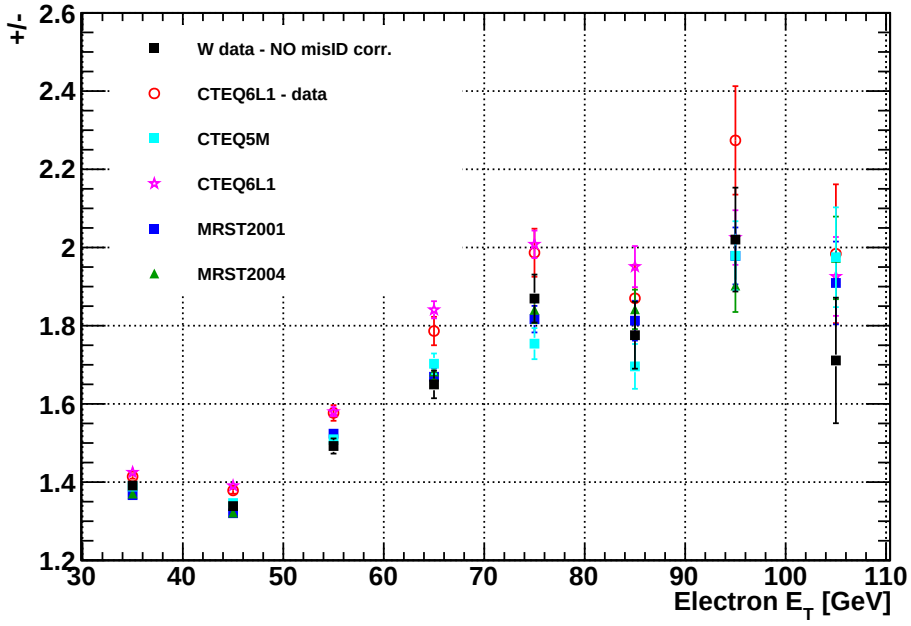


Figure 4.23: W^+/W^- asymmetry as a function of the electron transverse energy, as obtained from data without misID-correction (100 pb^{-1}), superimposed to the different PDF models

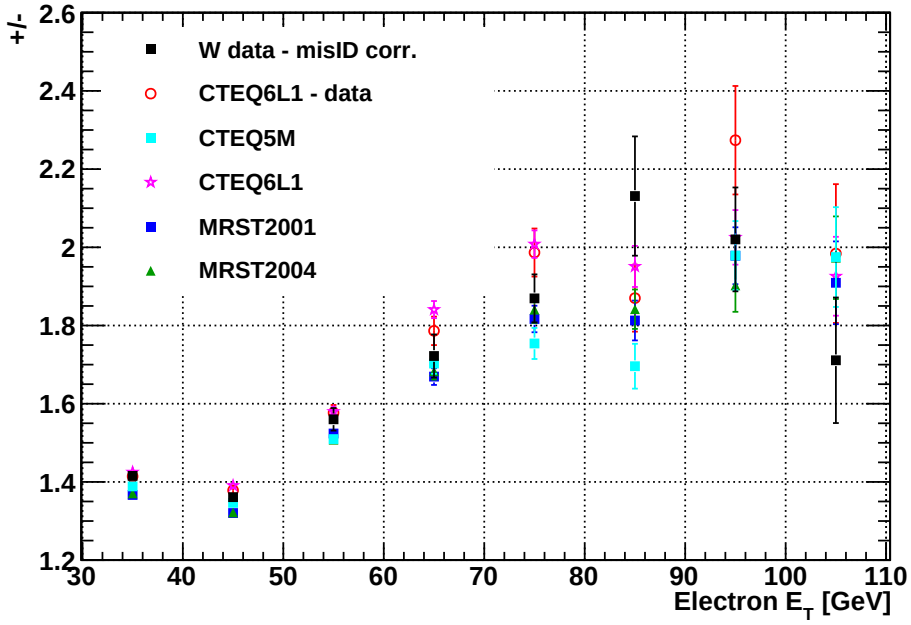


Figure 4.24: W^+/W^- asymmetry as a function of the electron transverse energy, as obtained from misID-corrected data (100 pb^{-1}), superimposed to the different PDF models

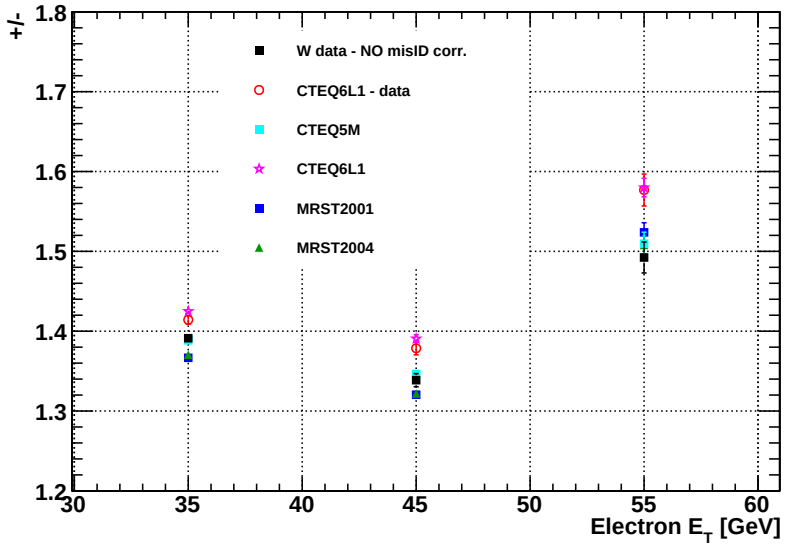


Figure 4.25: W^+/W^- asymmetry as a function of the electron transverse energy, as obtained from data without misID-correction (100 pb^{-1}), superimposed to the different PDF models: zoom on high-statistics bins

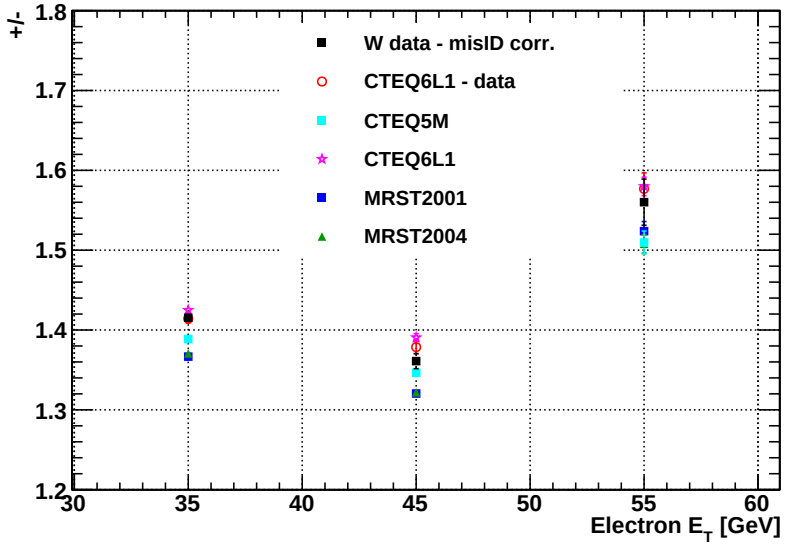


Figure 4.26: W^+/W^- asymmetry as a function of the electron transverse energy, as obtained from misID-corrected data (100 pb^{-1}), superimposed to the different PDF models: zoom on high-statistics bins

Conclusions

The Large Hadron Collider (LHC) at CERN will provide proton-proton collision at an initial centre-of-mass energy of approximately 10 TeV. The Compact Muon Solenoid (CMS) is one of the four high-energy experiments on the LHC ring; through its multiple detecting systems, it will allow to measure the energy and momenta of the particles coming out of the primary interactions and explore the TeV region, where many theories beyond the Standard Model predict a rich phenomenology. In particular, electron and photon energy will be revealed through a homogeneous electromagnetic calorimeter (ECAL) made of 75,848 scintillating crystals.

At the beginning of the data taking, the complex subdetectors of the CMS experiments will be precisely calibrated using physics events.
The first part of this work shows how it is possible to adopt an iterative method that exploits $Z \rightarrow e^+e^-$ events and the very precise knowledge of the Z mass to calibrate regions of the ECAL; the method basically uses the ratio between reconstructed and nominal Z mass in each region and extract regional calibration constants. An extensive MonteCarlo validation has been carried out in order to ensure the reliability of the method. Among the possible applications are pseudorapidity- and energy-dependent correction factors for electrons, measurement of the ECAL absolute scale and calibration of ECAL rings.
Test beam analyses - carried out at the H2 facility at CERN - have shown that the ECAL response to electrons has a non-linearity less than 1% in the energy range 9-100 GeV, so the calibration constants coming out of the iterative algorithm - and obtained from electrons with energies around 50-100 GeV - retain their validity even for lower energy electrons.

The second part of this work describes how to measure the electron charge misidentification rate from the first (100 pb^{-1}) CMS data. The method employed is the Tag & Probe method: on $Z \rightarrow e^+e^-$ events, one of the electrons (the "Tag") is required to have a high-quality track so that all the cases where the second electron (the "Probe") has the same charge as the Tag can be stated as charge-misidentified events. The reliability and stability of the method has been thoroughly investigated through MonteCarlo.
Measuring the electron charge misID rate not only is a commissioning task with respect to the reconstruction algorithms of the CMS experiment, but also allows to measure an ingredient which is important to both Standard Model and "exotic" phsyics studies.

A physical observable that is useful to test the Standard Model at the LHC, and on which the electron charge misID has a relevant role, is the W^+/W^- cross section ratio. At the LHC the initial state of the collision is charge-asymmetric and favours the production of positive W bosons, especially in the forward direction, so a W^+/W^- ratio greater than 1 is expected. The measurement of the W^+/W^- ratio will not suffer from the large (around 10%) luminosity uncertainty which is expected with the first LHC data.
In the third part of this work, the basic strategy for the measurement of the W^+/W^- ratio is described, and the effect of a charge misID correction is presented. Different PDFs predict different values for the W^+/W^- ratio; the possibility to put a constraint on the various PDF models with 100 $pb^{-1}$of collected data is also investigated. As an example, it has been shown that with a statistics of 100 pb^{-1}it is possible to exclude some PDF models, but the correction for charge misidentification is mandatory for such a conclusion.

Appendix A

Linearity of the CMS electromagnetic calorimeter

In 2006 the CMS collaboration conducted a test of the combined electromagnetic and hadronic calorimeter (ECAL and HCAL) at CERN; it allowed to study the calorimetry of the CMS experiment as a whole.
This section reports the basic elements of an investigation of the linearity of the ECAL supermodule set up on the CMS ECAL-HCAL combined test beam facility "H2", using electrons and positrons in the H2 beam line. The event selection is described including how the beam contamination by other types of particles are strongly reduced by using the particle ID detectors.
Measuring the linearity of the ECAL in a wide range of energy is one of the most interesting points of the test beam analysis because it leads to an understand of the ECAL response to low energy particles.
This report is relevant to the topic of the $Z \rightarrow e^+e^-$ calibration because it points out whether the ECAL response to electrons and positrons is linear in the range 9 - 100 GeV: if this is the case, then the calibration constants derived using electrons in $Z \rightarrow e^+e^-$ events - that have energies of order 100 GeV - can, on principle, be extrapolated down to lower energies, keeping uncertainty under control.
On the other hand, if the ECAL shows good linearity, this linearity can be considered as a cross-check piece of information in calibration studies using low-mass dielectron resonances, like J/ψ or Υ, whose electrons have energies of order of 10 GeV.

A.1 H2 experimental setup

Fig. A.1 schematically depicts the CERN H2 beam line. The beam line is designed to operate in two different modes. In the high energy mode, 400 GeV/c protons, extracted from the Super Proton Synchrotron (SPS), strike a target (T2) positioned 590.9 m upstream of the calorimeters and produce particles with momenta between 10 and 350 GeV/c. In the Very Low Energy (VLE) mode, an additional target (T22) is inserted 97.0 m upstream of the

calorimeters; tertiary particles coming out of T22 have momenta between 1 and 9 GeV/c, although at 1 GeV/c the beam quality is very poor.

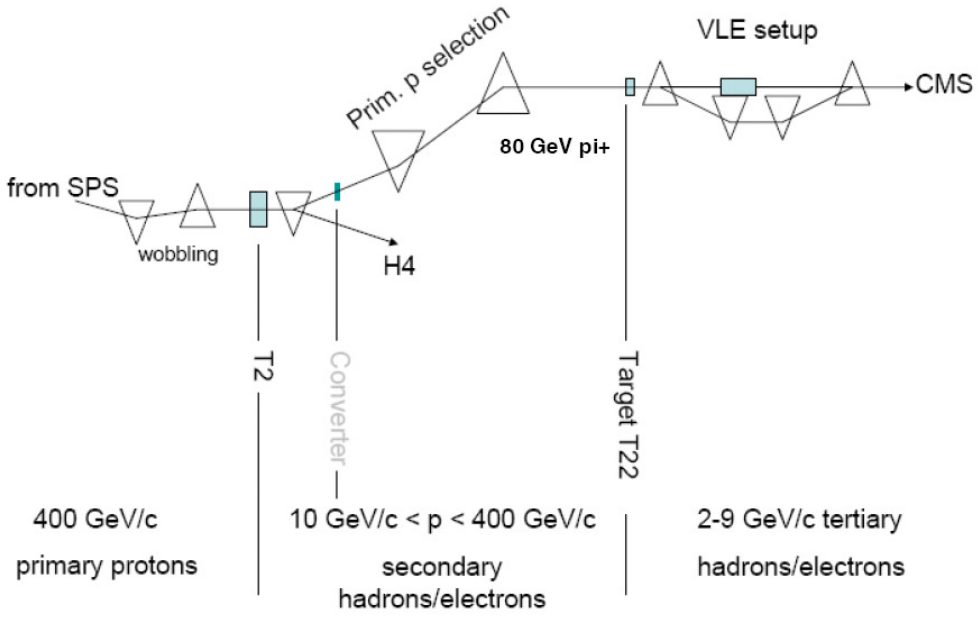


Figure A.1: A schematic view of the H2 beam line.

A number of detectors, shown in Fig. A.2, are positioned along the beamline in order to allow particle identification (PID) and event selection:

- eight Wire Chamber planes (WC 1, 2, 3 and WC A to WC E);
- three Cherenkov counters (CK 1, 2, 3);
- a Time of Flight (TOF) system made up of two stations (TOF1 and TOF2);
- four scintillation Beam Halo counters (BH1 through BH4);
- polystyrene Scintillators (S1 through S4);
- Muon Veto counters (Muon Veto Front, Muon Veto Back and a Wall of eight Muon Veto: VM1 to VM8)

In this analysis, all these detectors except the Cherenkov and the Time of Flight are used.

A.2 Data selection

In order to select events where only one electron or positron strikes the calorimeter, it is necessary to apply a selection based on Particle ID detectors and some ECAL variables. The main contaminants are multi-particle events and events with muons or hadrons.

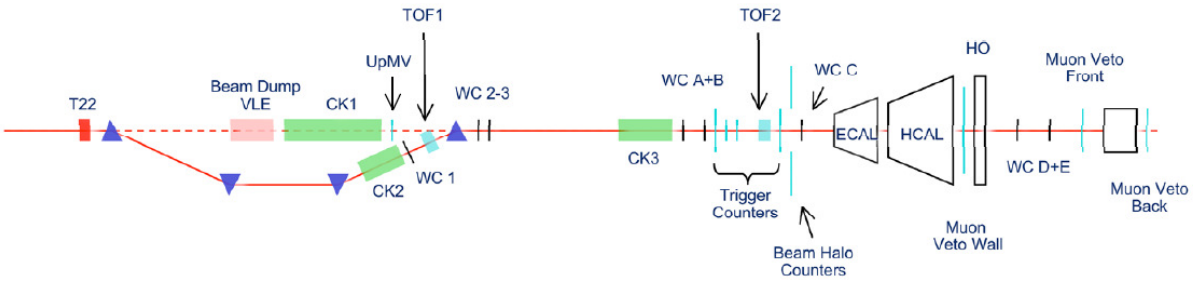
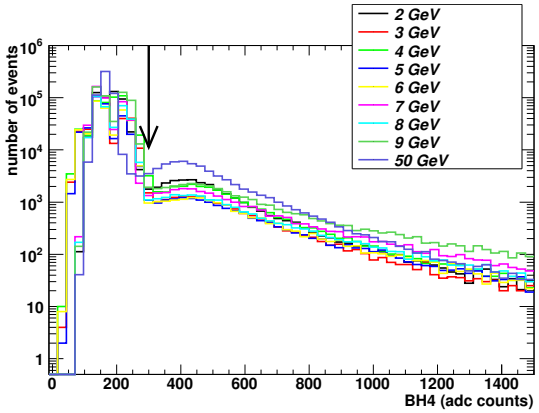


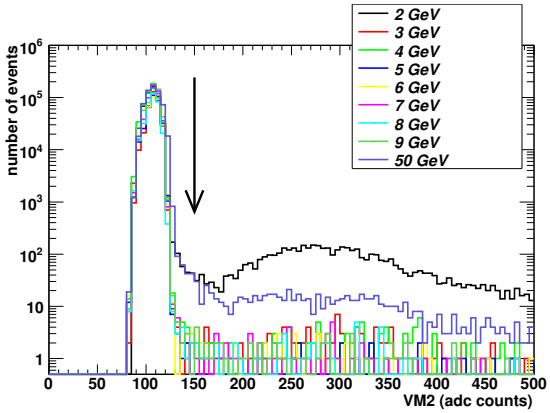
Figure A.2: A schematic view of PID detectors along the beam line.

A.2.1 Selection with Particle ID variables

A number of cuts in the signal of various PID detectors along the beamline were made. In order to exclude events with more than one particle striking the calorimeters, exactly one hit was allowed in WC B (both on the x and the y coordinates), which was demonstrated to be the most efficient among the WC's. We also made cuts on the signal from the scintillation counters. Cuts were also made on the signal from beam halo counters and the muon contamination was strongly reduced using the Muon Veto Wall (MVW) counters.
Some typical distributions for the signals coming from electron events can be seen in Fig. A.3.



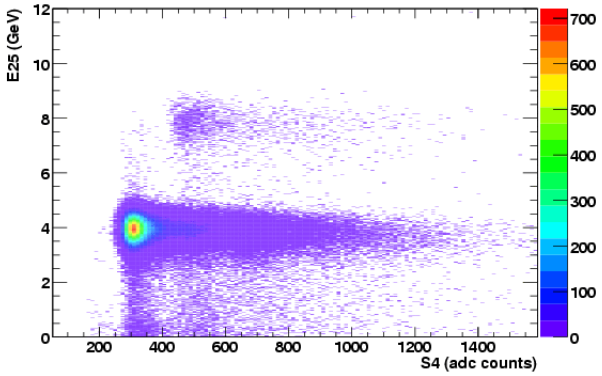
(a) Beam Halo 4 signal



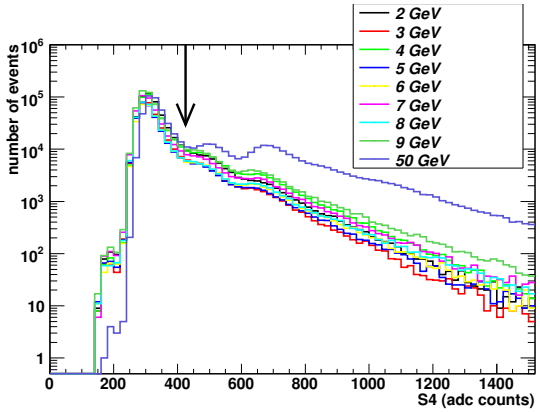
(b) Muon Veto Wall 2 signal

Figure A.3: Distribution of the signal of some PID detectors along the H2 beamline for electron beam at different energies; events to the left of the arrows are selected.

In Fig. A.4 it is clearly shown - for 4 GeV electron beam - that events with more than one particle hitting the ECAL, depositing more energy in the crystals, correspond to events with a high signal (in ADC counts) in the scintillators; in particular, a certain number of events with twice as energy as expected (i.e. with two particles) are visible.



(a) Reconstructed energy (E25) vs Scintillator 4 signal (4 GeV electron beam)



(b) Scintillator 4 signal for electron beam at different energies. Events to the left of the arrow are selected.

Figure A.4: Scintillators are used to reject events with multiple particle hitting the ECAL.

A.2.2 Selection with ECAL variables

Some calorimetric variables turned out to be useful in the selection of electrons. The few pions that pass the previous cuts are rejected using the ratio $E1/E9$, where E1 is the energy deposited in the ECAL crystal with the maximum energy deposit and E9 is the sum of the energies in a matrix of 3 x 3 crystals around the central one. Pions deposit energy in a very narrow region of ECAL, so that $E1/E9$ is about 1 for these particles. On the other hand, electrons have a different distribution of values below 1 for $E1/E9$; this yields the distribution in Fig. A.5, where it is clear how a cut $E1/E9 < 0.95$ excludes the pions from the analysis.

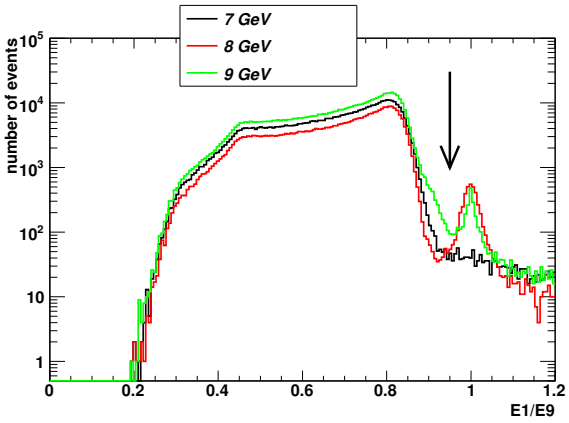


Figure A.5: $E1/E9$ distribution for 7-9 GeV electrons after PID selection; a residual contamination of minimum ionizing pions is visible around $E1/E9 = 1$. Events in the right tails, with $E1/E9 > 1$, correspond to $E1$ of few hundreds MeV and negative values of $(E9 - E1)$, compatible with noise in the eight crystals around the most energetic one.

Another class of events to be rejected are those where the beam particle hits the ECAL far from the center of a crystal, so that a non negligible amount of energy is lost in the intercrystal gaps. The beam is not centered with respect to the crystal and the beam transverse profile is larger than one crystal front face; both these beam features are clearly visible if $E1/E9$ is plotted against the pseudorapidity (η) value of the 5 x 5 cluster as in Fig. A.6; the crystal-by-crystal shape of the ECAL is clearly visible.
Fig. A.7a (b) shows the ratio $E1/E9$ versus the difference between the η (ϕ) coordinate of the most energetic crystal and the η (ϕ) coordinate of ECAL cluster, calculated with the log-weighted mean of the energy in the 5 x 5 crystals matrix [**?**]. Since the latter is computed at the depth of the shower maximum, while the crystal coordinates refer to the front face, the maximum of $E1/E9$ is not centered in zero, due to the 3° tilt of the crystals.
The $E1/E9$ distribution is fitted in both dimensions with a polynomial of 4^{th} degree; fit results are used to select events in the central part of the crystal. Even if the fit functions are very similar, cut values ($\Delta\eta$ and $\Delta\phi$) are computed at each beam energy in order to select the same area (about 60 % of the crystal front face). Selection efficiencies, for positron and electron beams, are given in Table A.1.

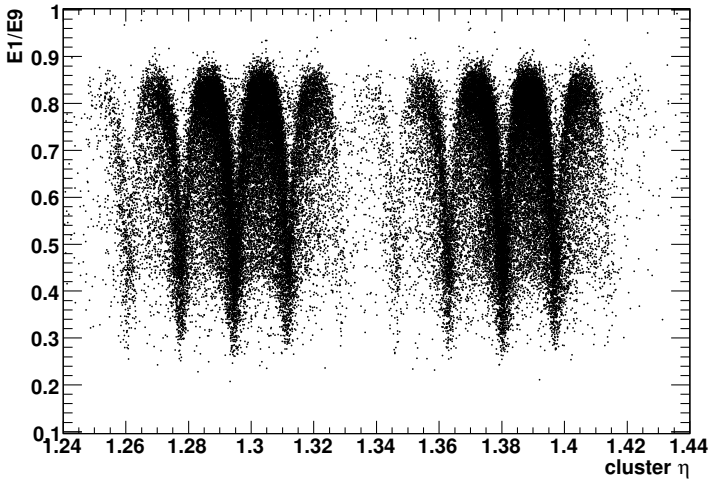
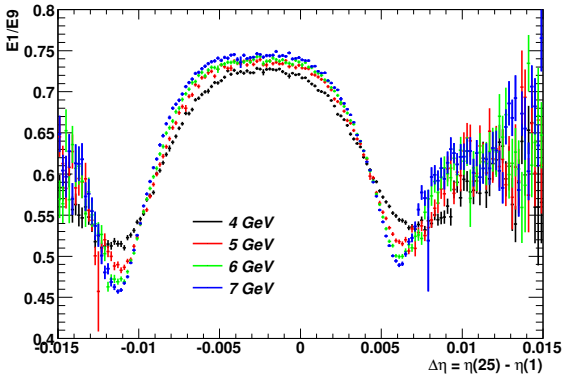


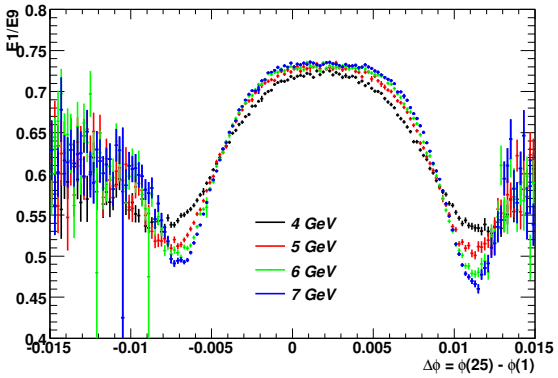
Figure A.6: $E1/E9$ for two 6 GeV electron beam runs as a function of η of 5 x 5 cluster around the most energetic crystal.

Table A.1: Selection efficiency of electron and positron beam data. VLE efficiencies are larger because contamination in VLE beam is lower than in high energy beam.

Beam Particle	Beam Energy (GeV)	Number of Events	PID selection ε	ECAL selection ε	global ε
e^-	2	527 k	65.9 %	59.4 %	39.1 %
e^-	3	359 k	69.9 %	59.7 %	41.7 %
e^-	4	512 k	66.4 %	58.2 %	38.6 %
e^-	5	335 k	69.4 %	57.3 %	39.8 %
e^-	6	334 k	68.2 %	56.7 %	38.6 %
e^-	7	468 k	68.6 %	57.1 %	39.2 %
e^-	8	354 k	68.3 %	57.3 %	39.1 %
e^-	9	590 k	67.6 %	56.9 %	38.5 %
e^+	9	59 k	59.2 %	53.0 %	31.3 %
e^+	15	17 k	49.2 %	47.8 %	23.5 %
e^+	20	18 k	51.6 %	45.3 %	23.3 %
e^+	30	102 k	51.0 %	30.9 %	15.8 %
e^+	50	2911 k	33.2 %	43.8 %	14.5 %
e^+	100	81 k	52.0 %	57.6 %	29.9 %



(a) E1/E9 vs η difference



(b) E1/E9 vs ϕ difference

Figure A.7: E1/E9 for different low energy electrons as a function of the difference in η and ϕ between the cluster and the the most energetic crystal.

.

A.3 Linearity Results

The energy of electrons and positrons is reconstructed using E25, defined as the sum of the reconstructed energy in a 5 x 5 matrix of calibrated crystals around the most energetic crystal in the event; the E25 distribution is then fitted with with a Crystal Ball function[**?**]:

$$CB(x) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x-x_0}{\sigma}\right)^2} & \frac{x-x_0}{\sigma} > -|\alpha| \\ \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{1}{2}|\alpha|^2} \left[\frac{n}{|\alpha|} - |\alpha| - \frac{x-x_0}{\sigma}\right]^{-n} & \frac{x-x_0}{\sigma} < -|\alpha| \end{cases} \quad \text{(A.1)}$$

The ECAL linearity is investigated by computing the ratio $E25_{peak}/E_{beam}$ where $E25_{peak}$ is the Crystal Ball peak (x_0) and E_{beam} is the beam energy. For a perfectly linear ECAL, we expect $E25_{peak}/E_{beam}$ to be independent of E_{beam}. An iterative Gaussian fit of E25 distribution, in

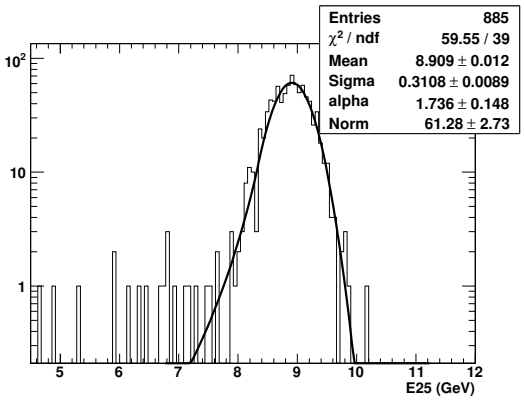
the range $[-1\sigma, +1.5\sigma]$ around the peak, is used as a cross-check. The results from the two fits are always in excellent agreement so that systematic errors, due to the fit procedure, is considered neglibible.
The corrections applied, in this analysis, on ECAL energy measurements are described in details in Ref. [43]. Several effect have been taken into account: the temperature variation during the data taking, the energy loss along the beam line, the energy leakage of the crystals, the beam energy scale variation and beam energy spread.

A.3.1 High Energy Beam

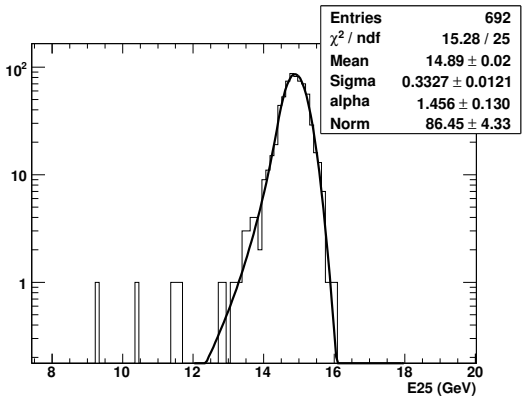
Most of the positron data was taken at 50 GeV to intercalibrate the supermodule crystals. As reported in Table A.1, a few thousand events were collected at different energies, namely from 9 to 100 GeV, with beam impinging on module 1 crystals. After selection we identified two crystals (268 and 288 [**?**]) with at least 600 events for each energy bin. The E25 distribution, fitted with a Crystal Ball, are shown in Fig. A.8 for crystal 288.
The intercalibration of the ECAL crystals in the H2 supermodule has been performed using 50 GeV positrons and a 3 x 3 matrix around the most energetic one. Intercalibration constants obtained with cosmic ray muons are taken for crystals having insufficient H2 data (see Chapter 2). In this analysis the absolute scale is fixed by requiring that, in average, 50 GeV was the value of the calibrated energy in a matrix of 5 x 5 crystals. Clearly, the ECAL absolute scale doesn't affect the linearity results described in this note.
E25 distribution peaks normalized to the beam energy are shown in Fig. A.9 for crystals 268 and 288; errors are the fit statistical errors of the distribution peak. A linearity better than 1% is observed in both crystals in the energy range 9-100 GeV.
Defining R_i as the ratio $E25_{peak}/E_{beam}$ at beam energy i GeV, the weighted average:

$$< R >= \frac{\sum_{j=9}^{100} R_j/\sigma_{R_j}^2}{\sum_{j=9}^{100} 1/\sigma_{R_j}^2} \tag{A.2}$$

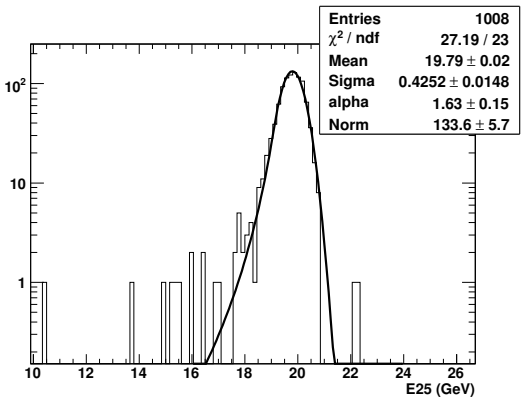
is directly related to the intercalibration constant of the crystal; σ_R is the mean statistical error, normalized to the beam energy, from the fit (Fig. A.8).
In order to remove, in Fig. A.9, the intercalibration contribution, R_i of the two crystals is rescaled, using $< R >$. Rescaled data are then reported in Fig. A.10; the two curves represent the beam energy scale uncertainty due to the error on the current measurements in the magnet power supplies, as discussed in [43].



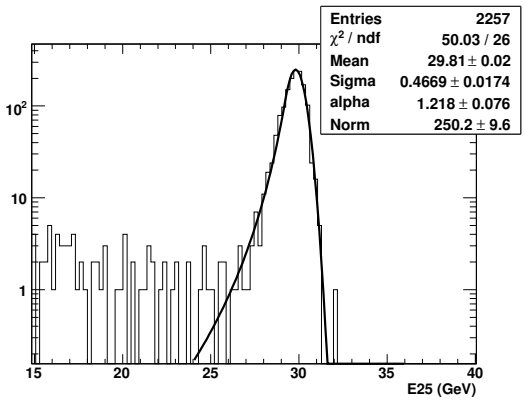
(a) 9 GeV positron beam



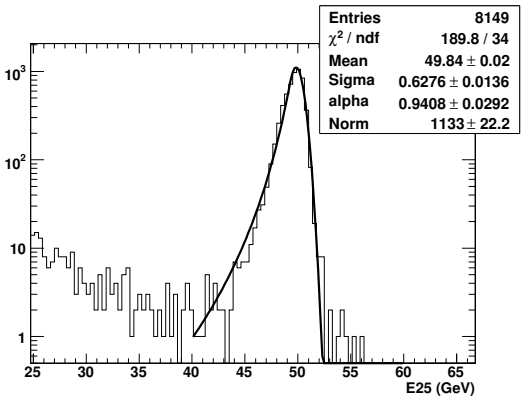
(b) 15 GeV positron beam



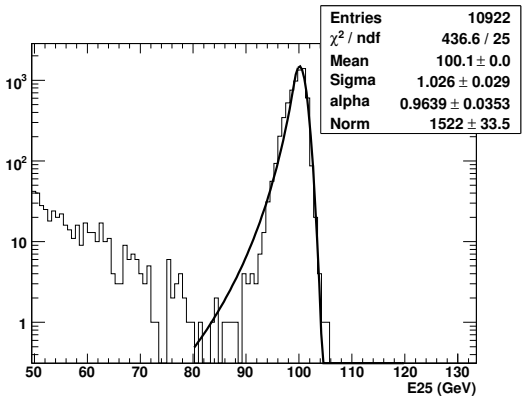
(c) 20 GeV positron beam



(d) 30 GeV positron beam



(e) 50 GeV positron beam



(f) 100 GeV positron beam

Figure A.8: Sum of the reconstructed energies in a 5 x 5 matrix around crystal 288, using events in which this crystal has the largest energy deposit.

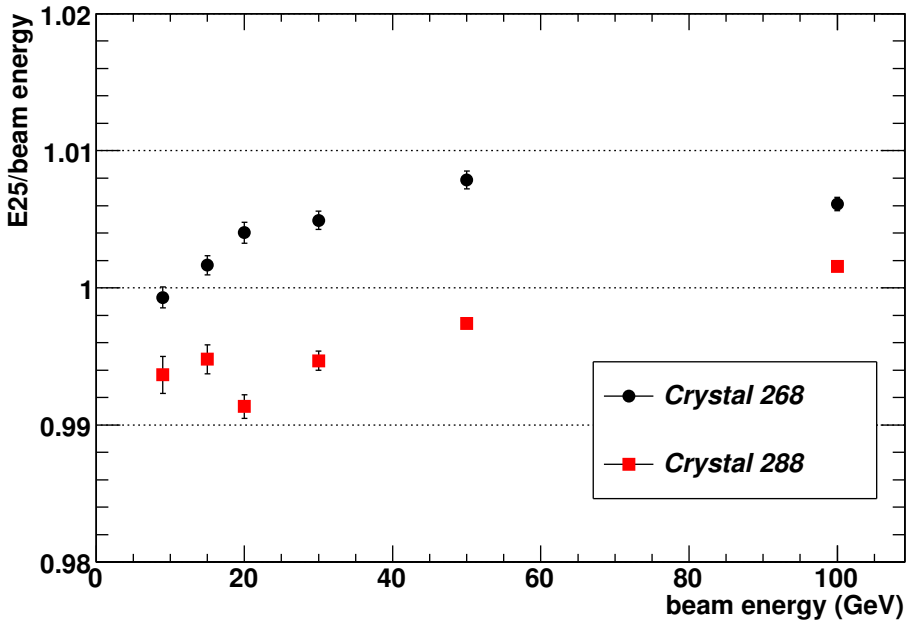


Figure A.9: E25 distribution peaks normalized to the beam energy for positron beam.

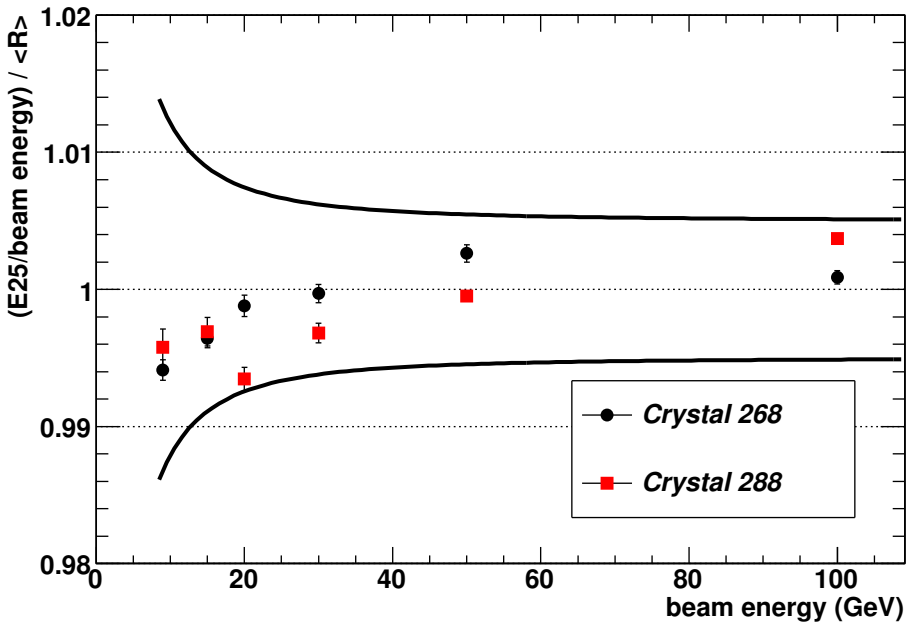


Figure A.10: E25 distribution peaks normalized to the beam energy after $< R >$ rescaling; the curves represent the beam energy scale uncertainty.

Appendix B

$Z \to e^+e^-$ events with charge misidentification in CMS

In the following pages, some fully simulated $Z \to e^+e^-$ events where one of the two Z electrons has wrong reconstructed charge are shown using the CMS visualization package Fireworks [37]. Yellow points represent tracker hits, while magenta towers represent calorimetric deposits. Visual inspections of the plots shows that misidentified electrons are characterized by secondary electron production that comes from the conversion of a Bremsstrahlung photon: in the plot, this appears as a "double-structure" of the calorimetric deposit of the misidentified electron.

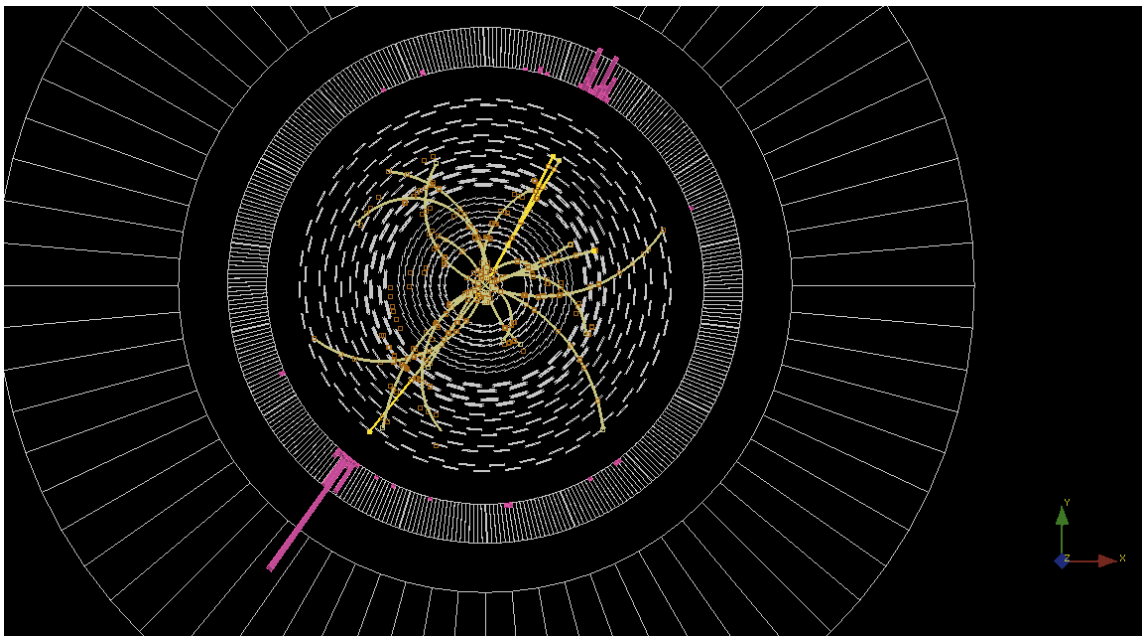


Figure B.1:

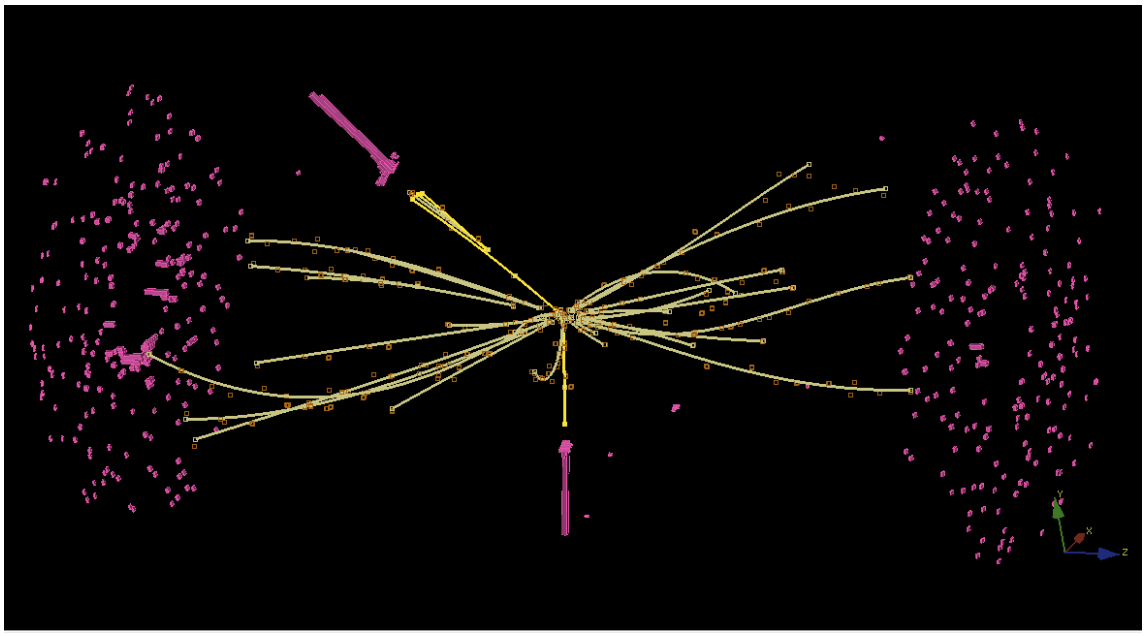


Figure B.2:

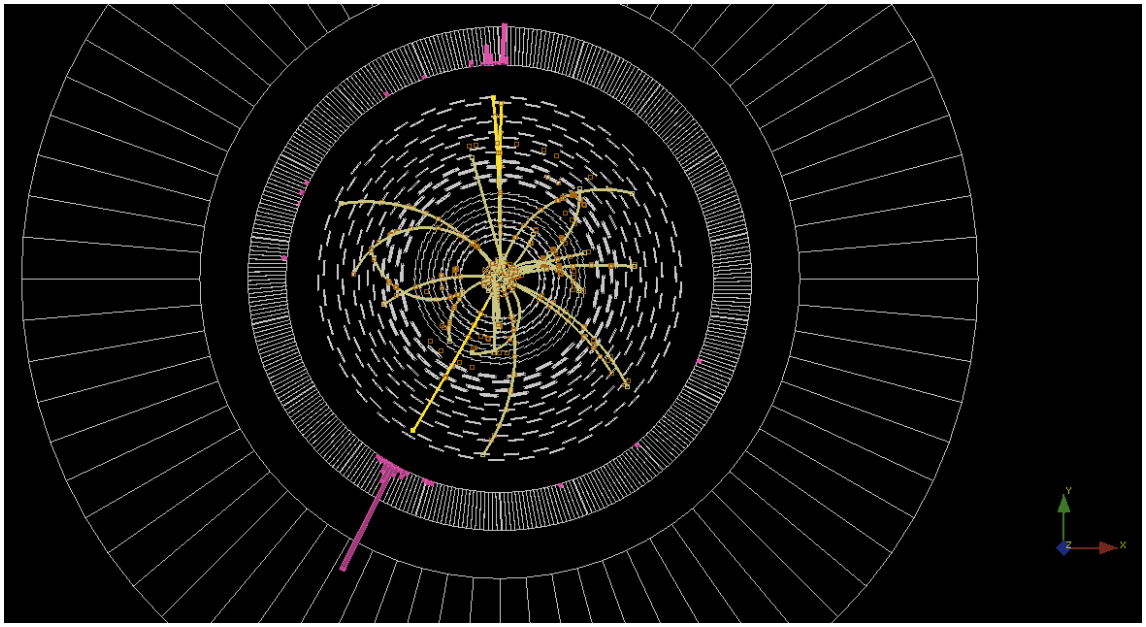


Figure B.3:

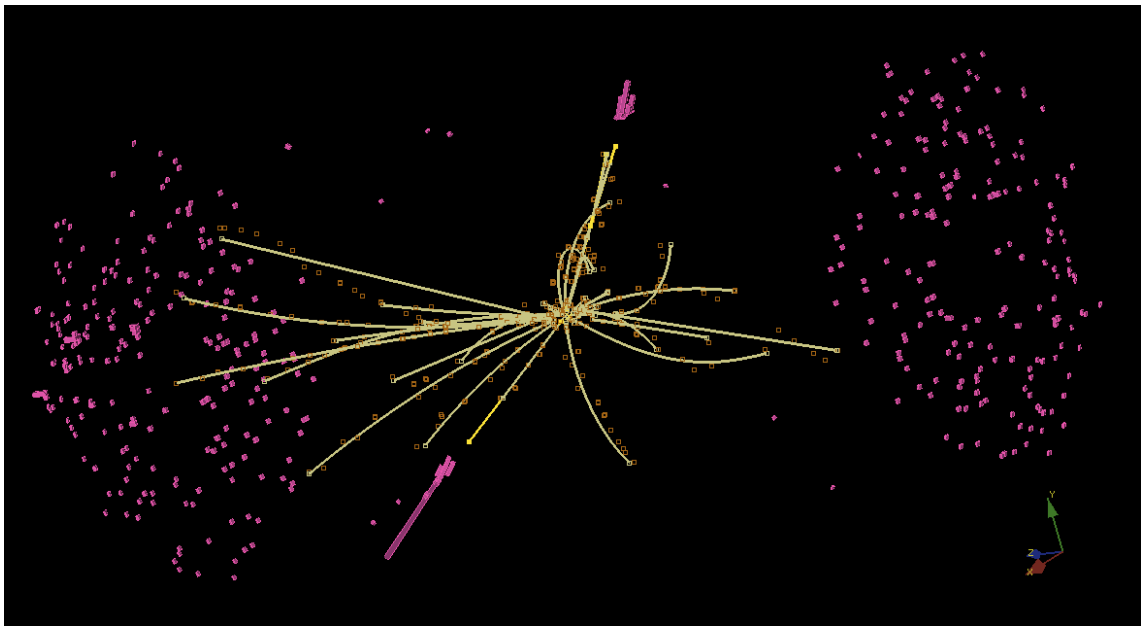


Figure B.4:

Appendix C

Data samples

Physics Channel	N.events	$\int$ L dt [pb^{-1}]	Weight w.r.t. $Z \rightarrow e^+e^-$
$Z \rightarrow e^+e^-$	455787	200.4	1.0
$W \rightarrow e\nu$	1112967	108.6	1.8
$Z \rightarrow \tau^+\tau^-$	1245500	899.1	0.2
$W \rightarrow \tau\nu$	1098500	78.5	2.6
QCD light flavours ($20 < \hat{p_T} < 30$)	20359765	6.4	31.5
QCD light flavours ($30 < \hat{p_T} < 80$)	38298918	8.1	24.6
QCD light flavours ($80 < \hat{p_T} < 170$)	5970425	20.9	9.6
QCD heavy flavours ($20 < \hat{p_T} < 30$)	1997072	10.4	19.3
QCD heavy flavours ($30 < \hat{p_T} < 80$)	2016487	8.4	23.9
QCD heavy flavours ($80 < \hat{p_T} < 170$)	1075822	47.2	4.2
$t\bar{t}$	103253	292.6	0.7
$\gamma + \text{jets}(15 < \hat{p_T} < 20)$	94811	1.1	186.4
$\gamma + \text{jets}(20 < \hat{p_T} < 25)$	115284	3.4	59.3
$\gamma + \text{jets}(25 < \hat{p_T} < 30)$	137708	8.9	22.4
$\gamma + \text{jets}(30 < \hat{p_T} < 35)$	167879	21.9	9.2
$\gamma + \text{jets}(\hat{p_T} > 35)$	162464	13.4	15.0

Bibliography

[1] P. Bryant L. Evans. LHC Machine. *JINST*, S08001:3, 2008.

[2] The CMS Collaboration. The Tracker Project Technical Design Report. *CERN/LHCC 98-006. Addendum CERN/LHCC 2000-016*, 1998.

[3] The CMS Collaboration. The Electromagnetic Calorimeter Technical Design Report. *CERN/LHCC 97-033. Addendum CERN/LHCC 2002-027*, 1997.

[4] L.M. Barone et al. Precise measurements of light yield and transmission of PbWO4 crystals at the INFN-ENEA regional centre. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 562:76–84, 2006.

[5] L.M. Barone et al. High Voltage system for the CMS electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 582:462–468, 2007.

[6] CMS ECAL Technical Design Report. *CERN/LHCC 97-33*, 1997.

[7] S. Chatrchyan et al. The CMS Collaboration. The CMS experiment at the CERN LHC. *JINST 3 S08004*, 2008.

[8] P.Adzic et al. Energy resolution of the barrel of the CMS Electromagnetic Calorimeter. *JINST 2 P04004*, 2007.

[9] P. Adzic et al. Results of the first performance tests of the CMS electromagnetic calorimeter . *Eur. Phys. C*, 44:1–10, 2006.

[10] The CMS Collaboration. The Hadron Calorimeter Technical Design Report. *CERN/LHCC 97-031*, 1997.

[11] The CMS Collaboration. The Magnet Technical Design Report. *CERN/LHCC 97-10*, 1997.

[12] The CMS Collaboration. The Muon Project Technical Design Report. *CERN/LHCC 97-032*, 1997.

[13] The CMS Collaboration. The TriDAS Project Technical Design Report, Vol. 1. *CERN/LHCC 2000-038*, 2000.

[14] The CMS Collaboration. The TriDAS Project Technical Design Report, Vol. 2. *CERN/LHCC 2002-026*, 2002.

[15] The CMS Offline SW Guide. `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuide`.

[16] The GEANT4 Collaboration. GEANT4: A simulation toolkit. *Nuclear Instruments and Methods in Physics Research A*, 250:250–303, 2003.

[17] The CMS Collaboration. CMS Physics Performance, Volume 2. *CERN-LHCC-2006-021*, 2006.

[18] The CMS FastSimulation Homepage. `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideFastSimulation`.

[19] The CMS Collaboration. CMS Physics Performance, Volume 1. *CERN-LHCC-2006-001*, 2006.

[20] Matteo Cacciari and Gavin P. Salam. Dispelling the n^3 myth for the kt jet-finder. *Physics Letters B*, 641:57, 2006.

[21] M. H. Seymour S. Catani, Y. L. Dokshitzer and B. R. Webber. Longitudinally invariant kt clustering algorithms for hadron hadron collisions. *Nuclear Physics B*, 406:187–224, 1993.

[22] R. Fruhwirth. Application of Kalman Filtering to Track and Vertex Fitting. *Nuclear Instruments and Methods in Physics Research A*, 262:444, 1987.

[23] S. Baffioni et al. Electron reconstruction in CMS. *CMS NOTE-2006/040*, 2006.

[24] Y. Sirois C. Charlot, C. Rovelli. Reconstruction of Electron Tracks Using Gaussian Sum Filter in CMS. *CMS AN-2005/011*, 2005.

[25] J. Branson et al. A cut based method for electron identification in CMS. *CMS AN-2008/082*, 2008.

[26] The CMS Electromagnetic Calorimeter Group et al. Intercalibration of the barrel electromagnetic calorimeter of the CMS experiment at start-up. *JINST*, P10007:3, 2008.

[27] The CMS Collaboration. The CMS Silicon Strip Tracker Operation and Performance with Cosmic Rays in 3.8 T Magnetic Field. *CMS PAPER CFT-09-002*, 2009.

[28] The CMS Collaboration. Alignment of the CMS Silicon Tracker During Commissioning with Cosmic Ray Particles. *CMS PAPER CFT-09-003*, 2009.

[29] The CMS Collaboration. From Detector to Analysis: CMS Data Processing Workflows During an Extended Cosmic Ray Run. *CMS PAPER CFT-09-007*, 2009.

[30] M. Gataullin et al. InterCalibration of the CMS Barrel Electromagnetic Calorimeter Using $\pi^0 \to \gamma\gamma$ Decays. *CMS DN 2007/013*, 2007.

[31] M. Gataullin et al. Online Selection of $\eta \to \gamma\gamma$ Decays and Performance of the η InterCalibration Method at the LHC Startup. *CMS DN 2009/007*, 2009.

[32] R.Paramatti et al. Shower containment studies in ECAL Barrel. *CMS DN 2009/004*, 2009.

[33] T. Sjostrand et al. A Brief Introduction to PYTHIA 8.1 . *arXiv:0710.3820*, 2007.

[34] The CMS Collaboration. Measurement of the muon stopping power in Lead Tungstate. *CFT-09-005*, 2009.

[35] C. Amsler et al. (Particle Data Group). Review of particle physics. *Physics Letters B*, 667:1, 2008.

[36] Juan Alcaraz Maestre. private communication.

[37] https://twiki.cern.ch/twiki/bin/view/cms/workbookfireworks.

[38] G. Servant R. Contino. Discovering the top partners at the LHC using same-sign dilepton final states. *arXiv:0801.1679v1*, 2008.

[39] W.Clarida et al. Searching for Majorana Neutrinos by same-sign dilepton final state at $\sqrt{s}$=10 TeV at the LHC. *CMS AN-2009/103*, 2009.

[40] N. Adam et al. Towards a Measurement of the Inclusive $W \to e\nu$ and $Z \to e^+e^-$ Cross Sections in pp Collisions at $\sqrt{s}$=10 TeV . *CMS AN-2009/098*, 2009.

[41] J. Branson et al. A cut based method for electron identification in CMS . *CMS AN-2008/082*, 2008.

[42] M. R. Whalley, D. Bourilkov, and R. C. Group. The Les Houches Accord PDFs (LHAPDF) and Lhaglue. 2005.

[43] R. Paramatti F. Cavallari, A. Palma. Linearity of the CMS Electromagnetic Calorimeter from H2 test beam data . *CMS DN-2007/019*, 2007.