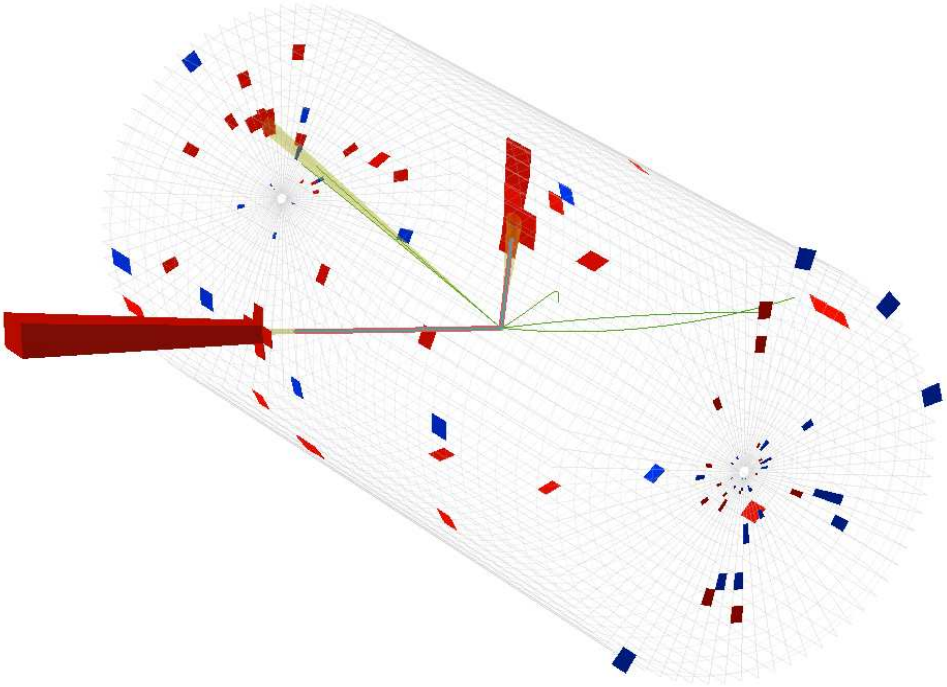
UNIVERSITE LIBRE DE BRUXELLES
FACULTE DES SCIENCES



Study of Drell-Yan production in the di-electron decay channel and search for new physics at the LHC

Dissertation présentée
en vue de l'obtention
du titre de Docteur en Sciences

CHARAF Otman

UNIVERSITE LIBRE DE BRUXELLES

FACULTE DES SCIENCES

Study of Drell-Yan production in the di-electron decay channel and search for new physics at the LHC

Dissertation présentée
en vue de l'obtention
du titre de Docteur en Sciences

CHARAF Otman

October 2010

Doctoral examination

Chair: Prof. J.M. Frère
Supervisor: Prof. B. Clerbaux
Secretary: Prof. C. Vandervelde
Prof. P. Marage
Prof. P. Miné
Dr. M. Mozer

This thesis was performed wit the financial support of IISN

To my parents.
In the heart of human beings,
there is an innate knowledge
that transcends the acquired one.

Contents

Introduction

The aim at understanding the Universe and what it is made of has always driven mankind curiosity. The concept of fundamental constituents or particles has opened a new way to the answer to this question. Today a theory called the Standard Model describes the particles and their interactions. These interactions can be understood as the manifestations of four fundamental forces: gravitation, electromagnetism, strong and weak interactions. Two of them have been unified (electroweak interactions) and three are described in the Standard Model (electromagnetism, weak and strong interactions). The aim of describing all four fundamental forces in a single framework has led to new theoretical approaches called theories beyond the Standard Model (BSM). Several of them predict new heavy bosons which can decay into an electron-positron pair. The study in the present thesis focuses on Drell-Yan production in the di-electron channel and the search for new physics.

Chapter 1 describes the fundamental principles of the Standard Model as well as the spontaneous symmetry breaking mechanism supposed to be at the origin of particle masses. There are strong indications, however, that the Standard Model is only a low energy scale effective theory as it does not provide answers to several fundamental questions for which new theoretical approaches have been proposed. The extra-dimension scenario introduces additional spatial dimensions with finite size. Grand unification theories propose larger gauge groups and introduce new multiplets. These two theoretical frameworks predict the existence of new heavy bosons (at the TeV scale) which can decay into an electron pair.

The CMS experiment at the Large Hadron Collider (LHC) will provide a tool to probe new physics at the TeV scale in the di-electron channel. The LHC, located near the French-Swiss border, produces proton-proton collisions at a centre of mass energy of 7 TeV and aims at covering a broad panel of studies. Due to the composite nature of protons, collisions involve complex physics processes. A presentation of the LHC and a brief introduction to proton-proton collisions with a focus on the parton density functions is provided in chapter 2. The CMS experiment uses a generic detector for numerous physics studies, which is described in chapter 3 with a focus on the two main components essential to this study. The electromagnetic calorimeter measures the energy of photons and electrons. It is characterised by an excellent energy resolution at high energy ($\sigma_E \sim 0.6\%$ at E $\gtrsim$ 100 GeV). The tracker reconstructs the trajectory followed by charged particles in the 3.8 T magnetic field, based on a minimum of 9 measurement points.

The Drell-Yan process, characterised by the presence of an electron-positron pair in the final state, $q\bar{q} \rightarrow \gamma/Z \rightarrow e^+e^-$, is a key process to the search for new physics as no such new physics is expected in the low mass region. It is described in detail in chapter 4 with a focus on the kinematics specific to this process. The Monte Carlo tools used to simulate this process are also presented. Other processes, called background processes, can mimic the signature, in the detector, of the Drell-Yan process. They are also described and their cross sections are given.

In order to look for possible deviations from the Standard Model, a specific analysis

strategy, described in chapter 5, has been developed by the HEEP (High Energy Electron Pairs) group. It relies primarily on the electron/positron selection, to discriminate as much as possible the Drell-Yan events from background contributions. Three regions in the invariant mass spectrum are exploited. The Z peak region ($60 < M_{ee} < 120$ GeV/c^2), with low expected background contributions, is used to measure the electron selection efficiency, from data, using the tag and probe method. In this method, events are selected that contain two objects, with some of the electron characteristics, the 'tag' and the 'probe', where the 'tag' is required to pass stringent selection criteria while the 'probe' is used to measure the efficiency. To ensure a high purity di-electron sample, the invariant mass of the two objects is required to be in the mass range $80 < M_{ee} < 100$ GeV/c^2. The high mass region ($120 < M_{ee} < 600$ GeV/c^2), where no new physics is expected in view of the recent results from Tevatron, is used as a control region where the Drell-Yan cross section will be computed and compared to previous measurements and to theoretical prediction. Finally the discovery region, with $M_{ee} > 600$ GeV/c^2 is dedicated to the direct search for new heavy bosons decaying into an electron pair. In addition, three methods have been designed to estimate the background contributions from data. This analysis strategy was tested, based on pseudo-experiments performed on Monte Carlo samples, considering a 10 TeV centre of mass energy and an integrated luminosity of 100 pb^{-1}. The three data-driven methods to estimate the background are compared. The efficiencies are determined using the tag and probe method. The discovery potential is derived in terms of 5 σ discovery reach and exclusion limits in case no evidence of new physics is observed.

To discriminate as much as possible the Drell-Yan events from the background contributions, the HEEP selection, described in chapter 6, selects events with two high p_t electrons (HPTE). The first step of the selection is the reconstruction of an energy deposit in the electromagnetic calorimeter. The second step of the selection is the electron reconstruction which demands the presence of a track in the tracker with information compatible with the energy deposit in the electromagnetic calorimeter. The third step relies on two sets of criteria: identification criteria and isolation criteria. The former requires more stringent compatibility between the track information and the information from the energy deposit in the ECAL, compared to the electron reconstruction. The latter demands that limited activity, in terms of tracks and calorimeter energy deposits, is present around the electron. The different steps are presented in detail and the efficiencies are determined from Monte Carlo and discussed.

First collisions in the Large Hadron Collider happened on November 23rd, 2009, at 900 GeV centre of mass energy, closely followed by 2.36 TeV collisions on November 30th and finally 7 TeV collisions since March 30th, 2010. A Data Quality Monitoring (DQM) tool, specific to high p_t electrons, was developed by the HEEP group to detect detector problems (noisy channels, miscalibration issues, ...), to perform a fast comparison of data with Monte Carlo predictions and to search for possible deviations from the Standard Model (chapter 7). For each centre of mass energy, distributions of relevant variables from data and Monte Carlo are compared and discussed.

Chapter 8 is dedicated to the analysis of the first LHC data from proton-proton collisions at 7 TeV centre of mass energy. The data collected by the CMS detector from 30/03/2010 till 30/08/2010 (Runs $132440 \rightarrow 144114$) are used, corresponding to a total integrated luminosity of 2.77 pb^{-1}. The study of the Drell-Yan invariant mass spectrum in the di-electron channel is performed. A first comparison between data and Monte Carlo simulations for the variables of the HPTE selection is presented. The invariant mass spectrum is extracted and high mass events are scrutinised using the CMS event display. The analysis focuses then on the Z peak. The high p_t electron selection efficiency is measured from data using the 'Tag and Probe' method and the background contributions are determined from Monte Carlo simulations.

The Drell-Yan cross section at the Z peak is computed and compared to the theoretical predictions at leading and next-to-leading orders. In view of the limited statistics available, no search for new physics was performed.

Chapter 1

The Standard Model and Beyond

> Again, of some bodies, some are composite, others the elements of which these composite bodies are made. These elements are indivisible and unchangeable, and necessarily so, if things are not all to be destroyed and pass into non-existence, but are to be strong enough to endure when the composite bodies are broken up, because they possess, a solid nature and are incapable or being anywhere of anyhow dissolved. It follows that the first beginnings must be indivisible, corporeal entities.
>
> Epicurus, letter to Herodotus, approximately 300 B.C. [1]

The idea that matter is composed of a set of elementary constituents called elementary particles has been proposed long ago (Epicurus letter to Herodotus). Only in the 20th century, however was this idea finalized in a theory called the Standard Model [2] which describes the fundamental constituents of matter and the interactions between them. Its consistency with experiments has been tested extensively and has always shown success [3]. This chapter aims to review the fundamental principles of the Standard Model (section 1.1). The justification for introducing theories beyond the Standard Model as well as the description of some of them constitute the core of section 1.2 and latest results on exclusion limits for new particles predicted by these BSM theories are illustrated in section 1.3.

1.1 The Standard Model

Section 1.1.1 introduces the fundamental constituents of matter and the fundamental interactions between them while section 1.1.2 explains how these fundamental interactions can be interpreted as symmetries of nature. Origins for masses of elementary particles are explained through a symmetry breaking mechanism described in section 1.1.3.

1.1.1 Fundamental constituents of matter and interactions

The fundamental constituents of matter, or elementary particles, are fermions and can be classified into two categories: the quarks and the leptons. Such a classification is sketched in figure 1.1.

Three families exist for the lepton category, the electron and electronic neutrino (e,ν_e), the muon and muonic neutrino (μ,ν_μ), the tau and the tau neutrino (τ,ν_τ). The leptons are characterized by a quantum number called the leptonic number. The electron, muon and tau

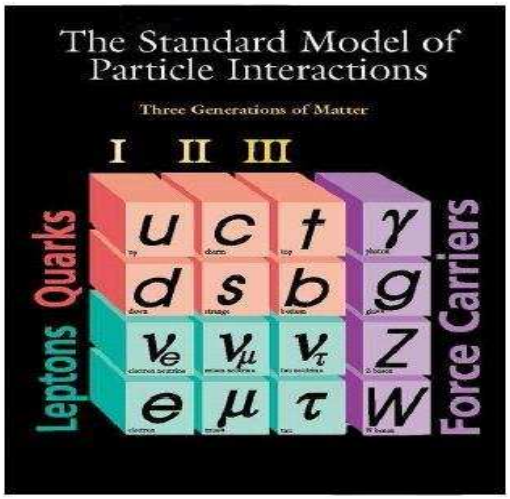


Figure 1.1: The fundamental particles of matter

all have an electric charge equal to -1 but different masses. The neutrinos carry no electric charge and have very small mass.

Similarly, three families are present for the quarks: the up and down quarks (u,d), the charm and strange quarks (c,s) and the top and bottom quarks (t,b). They are characterized by a quantum number called flavour. Each of the six quark species exists in three different colors symbolically denoted blue, red and green. Opposite to the leptons, the quarks carry a fractional electric charge; the up, charm and top quarks have an electric charge equal to 2/3 while the down, strange and bottom quarks have their electric charge equal to -1/3.

These particles of matter all have their corresponding anti-matter partners called anti-particles with the same mass but opposite quantum numbers. As an example, the anti-particle for the electron is the positron with mass 511 KeV/c^2 and electric charge +1.

All the stable matter present in the universe is made of particles from the first families of quarks and leptons: (e,ν_e) and (u,d). The up and down quarks are bound together to form the protons (uud) and the neutrons (udd) present in the atomic nuclei while the electrons around the atomic nuclei bind the atoms together to form the chemical molecules.

All physical processes in the universe can be viewed as the manifestation of a set of fundamental interactions. Up to now, four fundamental interactions have been observed, three of which are described in the Standard Model. They can be seen as the exchange of particle mediators which are bosons. As an example, the electric interaction between two electrons can be modeled as the exchange of a photon between these two electrons, the photon being the particle mediator of electromagnetism. Each fundamental interaction is characterized by an interaction range.

- the electromagnetic interaction acts on all charged particles and is mediated by the photon. It has infinite range. It is described in the Standard Model by a quantum gauge theory called quantum electrodynamics (QED) [4].

- the weak force is responsible for the β decay of radioactive nuclei. It also plays a role in neutrino production in thermonuclear reactions inside the sun. It acts on all particles (quarks and leptons).

- the strong interaction ensures the nucleon cohesion by binding together the quarks inside the nucleon. It acts on any colored objects (quarks and gluons); leptons are not sensitive to the strong interaction. It is mediated by eight bosons called gluons which, in contrast to the photons, carry the corresponding charge and are colored. It is described by a quantum gauge theory called quantum chromodynamics (QCD) [5].
- in classical mechanics, gravitation acts on all massive objects. In general relativity [6], gravitational effects are described by the geometry of local space-time curvature created by the presence of massive objects. A major challenge in modern physics is the edification of a quantum theory of gravitation. The only existing proposal is string theory [7] and has never been tested experimentally. The hypothetic mediator for gravitation, the so-called graviton, is a spin 2 particle, in contrast to the other mediators. Gravitation is not described in the Standard Model.

1.1.2 Interactions as gauge symmetries

Gauge symmetries always played a significant role in physics. A simple example is the freedom of choice for the electromagnetic potential $A^\mu = (\phi, \vec{A})$. The electric and magnetic fields, expressed in terms of ϕ and $\vec{A}$

$$\vec{B} = \vec{\nabla} \wedge \vec{A} \tag{1.1}$$

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}, \tag{1.2}$$

remain unchanged by the following replacements:

$$\vec{A}' = \vec{A} - \vec{\nabla}\chi \tag{1.3}$$

$$\phi' = \phi + \frac{\partial \chi}{\partial t} \tag{1.4}$$

which can be expressed in a single covariant way:

$$A'^\mu = A^\mu + \partial^\mu \chi \tag{1.5}$$

where χ is any arbitrary function.

The invariance of physics under such transformations is often called gauge symmetry and has an associated gauge group. In the following, an overview on the theoretical basis for the description of interactions as gauge symmetries is provided.

In the Standard Model, a free fermion with mass m is described as a spinor[1] ψ by the following Lagrangian:

$$L = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi. \tag{1.6}$$

[1]The following representation for a spinor as a four component column vector is chosen:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$$

The least action principle leads to the Euler-Lagrange equations which translate, for this specific Lagrangian, to the Dirac equation derived by Dirac in 1928:

$$i(\gamma^\mu \partial_\mu - m)\psi = 0. \tag{1.7}$$

Such a Lagrangian for free moving fermions describes, however, a static universe as no dynamics is included. Let us suppose a modification of the wavefunction based on its local phase transformation with rotation parameters $\vec{\epsilon}(x)$ in an internal space with generators $\vec{\tau}$:

$$\psi' = U\psi = e^{i\vec{\epsilon}(x)\frac{\vec{\tau}}{2}}\psi \tag{1.8}$$

such that quantum mechanical observables, based on $|\psi|^2$, remain constant. The Lagrangian 1.6 is generally not invariant and one needs to introduce a so-called covariant derivative D_μ in place of ∂_μ in the Lagrangian:

$$D_\mu = \partial_\mu - ig\frac{\vec{\tau}}{2}\vec{A}_\mu \tag{1.9}$$

where $\vec{A}_\mu$ is a new vector field and g represents the strength of the interaction between the fermion and the field $\vec{A}_\mu$. The Lagrangian is then rewritten as:

$$L = i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi \tag{1.10}$$

$$= i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi + g\bar{\psi}\gamma^\mu \frac{\vec{\tau}}{2}\vec{A}_\mu \psi \tag{1.11}$$

where the last term expresses the interaction between the fermion and the new vector field $\vec{A}_\mu$. Demanding the invariance of the Lagrangian requires:

$$D'_\mu \psi' = U(D_\mu \psi) \tag{1.12}$$

which translates into the following equation for $\vec{A}_\mu$:

$$\frac{\vec{\tau}}{2}\vec{A}'_\mu = -\frac{i}{g}(\partial_\mu U)U^{-1} + U(\frac{\vec{\tau}}{2}\vec{A}_\mu)U^{-1}. \tag{1.13}$$

Let us now take a look at how these symmetries are related to the fundamental interactions mentioned previously.

- The case where $U = e^{-i\chi(x)}$ represents a U(1) phase transformation based on an abelian group, as there is only one parameter[2]. Relation 1.13 translates into:

  $$A'_\mu = A_\mu - \frac{1}{g}\partial_\mu \chi \tag{1.14}$$

  similar to 1.5. This U(1) symmetry is directly linked to electromagnetism. There is only one vector field A_μ related to the electromagnetic field.

- A second symmetry of nature is related to the symmetry of mirror nuclei. This symmetry was first spotted by Heisenberg in 1932 who noticed that energy levels are identical for atoms in which a proton is replaced by a neutron in the nucleus. In other words, the weak interaction makes no difference between protons and neutrons. This lead to

[2]An abelian group is a commutative group, i.e. the product of any two elements of such a group, A, B commutes: [A,B] = 0. Equivalently, the generators of such a group commute.

consider the proton and the neutron as two states of a same object called the nucleon and forming a doublet. This symmetry is thus associated to a SU(2) gauge group, corresponding to a non-abelian theory and is called the isospin symmetry. The latter holds also at the level of the u and d quarks as the valence quarks are uud and udd for the proton and neutron respectively. Three new vector fields W_μ^i are introduced and the three corresponding matrices τ^i are the Pauli matrices.

For a spinor ψ, a classification expressed in terms of chirality is defined as follows:

$$\psi_L = \left(\frac{1-\gamma^5}{2}\right)\psi \tag{1.15}$$

$$\psi_R = \left(\frac{1+\gamma^5}{2}\right)\psi \tag{1.16}$$

with:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{1.17}$$

$$\gamma^{5\dagger} = \gamma^5. \tag{1.18}$$

The spinors ψ_L and ψ_R are called left-handed and right-handed, respectively. The operators $(\frac{1-\gamma^5}{2})$ and $(\frac{1+\gamma^5}{2})$ are called the left and right projectors, respectively. The specificity of the weak interaction is that it interacts only with left-handed particles. The right-handed particles are sterile with respect to this force.

- A third symmetry of nature concerns the quark colors. Indeed, as mentioned in section 1.1.1, each quark exists in three different colors. The associated group is SU(3) and describes the strong interaction. Eight new field vectors are introduced, the gluons, and the corresponding eight matrices τ^i are called the Gell-Mann matrices, usually denoted as λ^i.

Thus the Standard Model is based on the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$.

1.1.3 Introducing mass: the spontaneous symmetry breaking mechanism

The Standard Model as described in the previous section does not include mass terms for the gauge bosons. Introducing mass terms $m\bar{\psi}\psi$ in the Lagrangian breaks the chiral invariance:

$$\bar{\psi}\psi = \bar{\psi_L}\psi_R + \bar{\psi_R}\psi_L \tag{1.19}$$

To introduce mass terms and yet retain the symmetry of the Lagrangian, a mechanism was proposed [8] and [9]. It introduces a new doublet of scalar fields ϕ:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.20}$$

for which the following Lagrangian is introduced:

$$L_\phi = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi) \tag{1.21}$$

$$= (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2 - \lambda(\phi^\dagger\phi) \tag{1.22}$$

The potential $V(\phi)$ has a minimum at 0 if we choose $\mu^2 > 0$. However, if we set μ^2 to be negative, the potential $V(\phi)$ exhibits a shape sketched on figure 1.2. One sees that the minimum of such a potential does not correspond to the vacuum state anymore. The symmetry of the Lagrangian does not correspond to the symmetry of the vacuum state, it is spontaneously broken. This minimum is located at v = μ^2/λ.

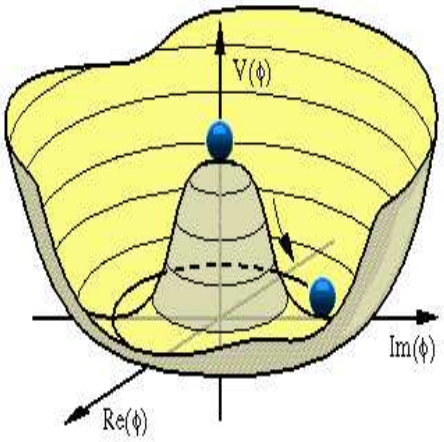


Figure 1.2: The "sombrero" shape of the Higgs potential $V(\phi)$ [10].

Expanding the field ϕ around its minimum in a specific direction and eliminating the Goldstone modes, the field ϕ becomes:

$$\phi = \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.23}$$

where $H(x)$ is a new scalar field called the Higgs field. Inserting 1.23 into 1.22 and using the following derivative, in the case of the $SU(2)_L \times U(1)_Y$ gauge group,

$$D_\mu = \partial_\mu - ig_1 \frac{Y}{2} B_\mu - ig_2 \frac{\vec{\tau}}{2} \vec{A}_\mu \tag{1.24}$$

two mass terms appear $(vg_2/2)^2 W^{\mu+} W^-_\mu = 80.4$ GeV and $(v\sqrt{g_1^2 + g_2^2}/2)^2 Z^\mu Z_\mu/2 = 91.2$ GeV corresponding, respectively, to the masses of the W^+, W^- and the Z while the photon remains massless.

1.2 Beyond the Standard Model

Although the Standard Model has gone with success through a wide variety of experimental tests, it is conceptually incomplete as it does not provide consistent answers to several questions. In the following, a non exhaustive list of the main open questions and shortcomings of the Standard Model are presented and an overview of two theories beyond the Standard Model, relevant for this thesis, is provided: the Grand Unification Theory (GUT) and the Extra Dimension models (ED).

1.2.1 Arguments towards theories beyond the Standard Model

Shortcomings of the Standard Model

- Why are there at least 19 free parameters in the Lagrangian of the Standard Model ? Can a theory constrain these parameters ?

- Quarks and leptons families: Why are there three families of quarks and leptons ?

- Quarks fractional charge: Why do quarks carry fractional charge ? Why is the proton charge exactly the opposite of the electron one ?

- Left-right asymmetry: In the Standard Model, only left-handed particles couple to the weak bosons while right-handed particles are sterile. Such an asymmetry is described in the Standard Model but no reason is provided as to the origin of this asymmetry.

Unification of fundamental interactions

- The present understanding of particles and their interactions includes a gauge group $SU(2)_L \times U(1)_Y \times SU(3)_C$. Can a simpler theory, i.e. a single group, describe all particles and their interactions ?

Theory of gravitational interactions — Hierarchy mass problem

- Gravitation: No description of gravitation is present in the Standard Model. The latter needs to be extended to include a theory of gravitation. There is yet no quantum theory of gravitation which has been tested experimentally.

- The hierarchy mass problem raises the question of the difference in orders of magnitude between the electroweak scale and the Planck scale. While the former is found to be around 10^2 GeV, the latter represents the scale at which quantum gravitational effects become important and is expressed by:

$$M_{Pl} = \sqrt{\frac{\hbar c}{G}} = 1.12 \times 10^{19} \text{ GeV}, \tag{1.25}$$

  where G is the Newton constant. In the framework of the Standard Model, no new physics is expected between these two scales, as the three other forces have been accounted for.

- Naturalness problem: in the Standard Model, the Higgs mass is 'naturally' very large, unless there is a fine-tuning cancellation between the quadratic radiative corrections and the bare mass.

Cosmological issues

- The asymmetry in the universe between matter and antimatter is puzzling. The mechanism at the origin of this asymmetry is intensively studied in cosmology.

- Another cosmological issue comes from the observational evidences for what is called dark matter. Astrophysical best current measurements indicate that around 96% of matter present in the universe is not known to us. No indication in the Standard Model is provided as to the possible constituents of dark matter.

Such questions are indications that the Standard Model is to be viewed only as an effective low energy theory and this encouraged physicists to seek a more global theory that embeds the results of the Standard Model. A cut-off scale, often called Λ, denotes the energy at which the Standard Model is to be replaced by this more fundamental theory. Two extreme cases exist as to the value of Λ. A first approach places it not much below the Planck scale, but it suffers however from the hierarchy problem. A second approach places the scale Λ close to the electroweak scale, and new physics is expected at the TeV scale. An example of the latter case is the supersymmetry (SUSY) theory which introduces a new symmetry between bosons and fermions. It proposes a solution to the naturalness problem, allows unification of the three coupling constants and alleviates the hierarchy mass problem.

Theories at the TeV scale receive great interest as they are directly linked to energies accessible at the LHC.

Many theories have thus been proposed, each focusing on specific points. They constitute the core of what is called theories beyond the Standard Model (BSM theories). In the following, the focus is on two specific kinds of BSM theories and their justifications: the grand unification theory and the extra dimensions scenario, both predicting the existence of new heavy neutral gauge bosons.

1.2.2 The Grand Unification Theories

Theories aiming to unify the electroweak and strong forces have received great interest. Indeed, former examples of unification (e.g. the electroweak theory) have proven to work well.

Grand unification theory (GUT) models refer to models in which the three gauge interactions (electromagnetic, strong and weak) are unified at high energy into a single interaction characterised by a larger gauge symmetry group and one unified coupling constant (rather than three independent ones). Such a unification is possible as the coupling constant is scale dependent in quantum field theory (Renormalization group equations). Observations as the proton decay, for example, could help to test such a theory.

A large panel of gauge groups has been studied. As an example, $SU(5)$, the group defined by 5×5 unitary matrices of determinant 1, is the simplest gauge group that contains the Standard Model gauge groups:

$$SU(5) \supset SU(3) \times SU(2) \times U(1). \tag{1.26}$$

The next simplest one $SO(10)$, the group of 10×10 orthogonal matrices, contains SU(5) and the Standard Model group:

$$SO(10) \supset SU(5) \supset SU(3) \times SU(2) \times U(1). \tag{1.27}$$

Another widely used group is $E(6)$.

A new gauge group involves new gauge bosons and a new organization of particles in multiplets. At low energies, the description of the Standard Model should be recovered and thus one needs to study the symmetry breaking of the GUT group.

$$G_{GUT} \rightarrow G_{SM} = SU(3)_C \times SU(2)_L \times U(1)_Y \tag{1.28}$$

As an example, the $E(6)$ symmetry breaking may proceed through the $SO(10)$ group:

$$E(6) \rightarrow SO(10) \times U(1)_\psi \rightarrow SU(5) \times U(1)_\chi \times U(1)_\psi \rightarrow SU(3)_C \times SU(2)_L \times U(1)_Y \times U(1)'. \tag{1.29}$$

In this particular case, a new heavy neutral boson called Z' is thus predicted, with charge described by a linear combination:

$$U(1)' = U(1)_\chi \cos\theta + U(1)_\psi \sin\theta \tag{1.30}$$

where θ is a free mixing parameter describing the couplings between the new Z' boson and the different fermions. Different models exist, each with a specific value of θ:

- Z'_χ model: $\theta = 0$. This corresponds to the extra Z' introduced in the SO(10) group framework.
- Z'_ψ model: $\theta = \pi/2$. This corresponds to the extra Z' introduced in the E(6) group framework.
- Z'_η model: $\theta = atan(\frac{-\sqrt{5}}{3})$.

In general, GUT models predict the existence of at least one heavy neutral gauge boson, called Z'. There is however no reliable prediction of its mass scale.

In addition to Z'_ψ, Z'_η and Z'_χ, arising from the $E(6)$ and $SO(10)$ groups, the use of the SSM (Sequential Standard Model) Z' is extensively used in the literature as a benchmark model. It supposes the existence of an extra neutral gauge boson Z' with couplings to the other particles identical to the Z boson. The mass of the Z' is a parameter of the model.

Additional models called respectively 'left-right' models and 'alternative left-right' models predict the existence of bosons Z'_{LRM} and Z'_{ALRM}. Their couplings are calculated according to the formalism of [11, 12, 13] assuming couplings to left-handed and right-handed fermions are equal ($g_R = g_L$).

1.2.3 Models with extra spatial dimensions

The extra dimension approach relies on the work originally initiated by Kaluza [14] (1921) and Klein [15] (1926) which attempted at a unification of electromagnetism and gravitation based on the introduction of an additional spatial dimension. They introduced many useful concepts

- the presence of the gravitational field in the higher dimensional space called the bulk reflects the existence of a unified theory in 4+1 dimensions.
- the bulk can be factorized as the product of the usual 4D space-time structure and a compact variety of dimension 1.
- the compactification of the extra dimensions allows the reinterpretation of the five-dimensional field in terms of so-called Kaluza-Klein massive states in four dimensions. The compactification can be applied on any geometry but for simplicity, the torus geometry is adopted with a compactification radius.

In 1998, Arkani-Hamed, Dvani and Dimopoulos proposed the idea of introducing large extra dimensions to address the hierarchy mass problem. The geometry of the extra dimensions is supposed to be responsible for the hierarchy. The gravitational field line is spread throughout the full higher dimensional space, which modifies the behaviour of gravity. This assumption relies on the fact that newtonian behaviour of gravitation has not been tested at distances smaller than a fraction of a millimeter [16].

In these models, the Planck mass appears as an effective scale. This can be expressed using Gauss' theorem:

$$M_{Pl}^2 = M_*^{2+d} R^d, \tag{1.31}$$

where d stands for the number of additional spatial dimensions, M_* is the true fundamental gravitational scale and R is the compactification radius. If we choose M_* to be of the order of the electroweak scale ($\sim$ 1 TeV), one can put a limit on the compactification radius R. $d = 1$ is forbidden as it contradicts the newtonian behaviour of gravitation at large distances (astrophysical scale).

These models predict the existence of massive excitations of the graviton, called Kaluza-Klein excitations or Kaluza-Klein tower. In other models, particles other than the graviton are also allowed to propagate to the bulk, predicting thus the existence of a KK tower for these bosons (KK Z for example).

Another way to solve the hierarchy problem was proposed by Randall and Sundrum [17]. Instead of introducing n large extra dimensions, it supposes the existence of a 'small' additional dimension in a warped geometry, compactified on a circle of radius R_c. The metric of such a geometry is given by:

$$ds^2 = e^{-2kR_c\phi}\eta_{\mu\nu}dx_\mu dx_\nu + R_c^2 d\phi^2, \tag{1.32}$$

where k stands for the curvature scale of the order of TeV. This geometry is non factorisable[3]. Two branes[4] are present at the coordinates $\phi = 0$ and $\phi = \pi$ (infrared brane). Energies of the order of TeV will be localized on the infrared brane. This brane will contain the gauge bosons of the three other interactions while the graviton will be able to propagate to the bulk. The energy scales are thus reduced by a factor $e^{-kR_c\phi}$. The gravity scale Λ_π becomes:

$$\Lambda_\pi = M_{Pl}\ e^{-kR_c\phi} \tag{1.33}$$

and the masses of the massive Kaluza-Klein excitations are given by:

$$M_n = kx_n e^{-kR_c\phi}, \tag{1.34}$$

where x_n is the n^{th} root of the Bessel function J_1. The Randall-Sundrum model is characterized by two parameters: the first mass M_0 and the coupling parameter $c = k/M_{Pl}$.

1.3 Current exclusion limits

Searches for new physics beyond the Standard Model have been extensively performed during the last 30 years, resulting in a strong set of exclusion limits. This section gives an overview of the present exclusion limits on heavy neutral resonances predicted by BSM theories for GUT Z' and Randall-Sundrum heavy gravitons. The most stringent limits have been set by CDF [18] and D0 [19], two experiments operating at the Tevatron which performs $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. The e^+e^- decay channel of the heavy resonance will, more specifically, be considered.

[3] In a factorisable geometry, the metric can be expressed as the sum of the two independent terms, one describing the geometry in the bulk and one describing the geometry in the extra dimensions.

[4] In the language of string theory, a p-brane is a dynamic extended object where p represents its number of spatial dimensions.

1.3.1 Z' exclusion limits

The analysis to put exclusion limits on Z' is performed by D0 [20], using 3.6 fb^{-1}of data from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. For the D0 analysis, the e^+e^- invariant mass spectrum is extracted from data and compared to the expected total background taking into account the acceptance and selection efficiencies. The most important contribution to the total background, called instrumental background, is coming from QCD multijet events in which both jets have been misidentified as isolated electrons. Other sources of background (labeled "Other SM") come from $Z/\gamma^* \to \tau^+\tau^-$, $W + X \to e\nu + X$ where X is a jet/photon misidentified as an electron, $W^+W^- \to e^+e^-\nu_e\bar{\nu}_e$, $W^{\pm}Z$ with $Z \to e^+e^-$, $t\bar{t} \to W^+b + W^-\bar{b} \to e^+\nu_e b + e^-\bar{\nu}_e\bar{b}$ events. Figure 1.3 shows the comparison between the e^+e^- invariant mass spectrum extracted from data and the expected total background in the range 70 GeV/c^2< M_{ee} < 1000 GeV/c^2. Agreement is observed.

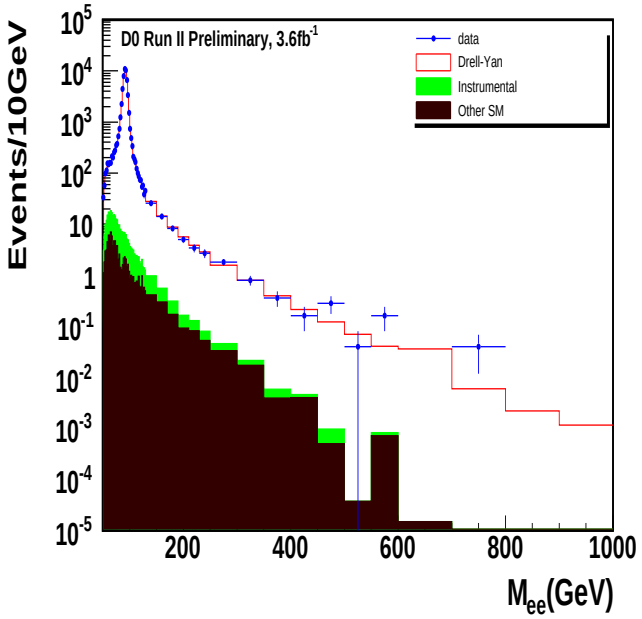


Figure 1.3: Di-electron invariant mass spectrum for data (blue points), with expected total background and the contributions from instrumental and other SM background superimposed for the full range studied [20].

As no significant excess is observed, upper limits on Z' production cross section times the e^+e^- branching ratio, $\sigma \cdot BR(p\bar{p} \to Z' \to ee)$ are derived, based on a Bayesian tool, considering a flat prior probability and accounting for the various uncertainties related to the different parameters involved in the Bayesian tool. Figure 1.4 gives the upper limit at 95% C.L. for the Z' production cross section times branching ratio, together with the theoretical signal production cross section from different models. Lower limits on the Z' mass are summarized in table 1.1. A Z'_{SSM} is thus excluded at $M < 950$ GeV/c^2.

A similar analysis is performed by CDF, using 2.5 fb^{-1} of data from $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV. The Z'_{SSM} in the e^+e^- channel is excluded at 966 GeV. The analysis in the dimuon channel excludes a Z'_{SSM} at 1071 GeV using 4.6 fb^{-1} of data [21].

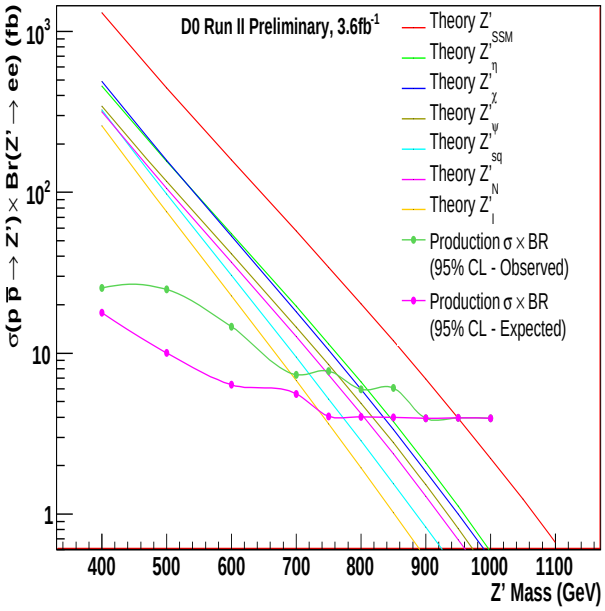


Figure 1.4: The upper limit on the observed and expected cross section at 95% C.L. with superimposed various Z' models [20].

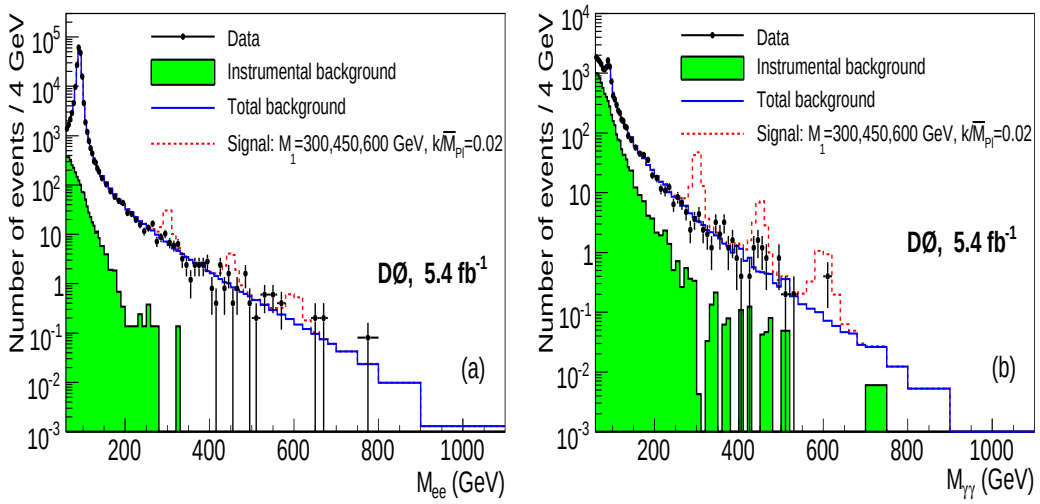
1.3.2 Randall-Sundrum heavy graviton exclusion limits

The Randall-Sundrum model for heavy gravitons is governed by two parameters: the first KK excitation mass M and the coupling $c = k/M_{Pl}$. The most stringent exclusion limit on the heavy graviton mass is put by D0 using 5.4 fb^{-1} of data from $p\bar{p}$ collisions at $\sqrt{s} =$ 1.96 TeV [22]. The analysis is based on the determination of the e^+e^- and $\gamma\gamma$ invariant mass spectra, corresponding to the heavy graviton decay channels $G \to e^+e^-$ and $G \to \gamma\gamma$, respectively. The acceptance and selection efficiencies are taken into account as well as the different sources of background, of which the main source, called instrumental background, comes from QCD jets that mimic the final state.

Figure 1.5 shows the e^+e^- (a) and $\gamma\gamma$ (b) invariant mass spectra from data superimposed on the expected background. The data is in good agreement with the predicted background.

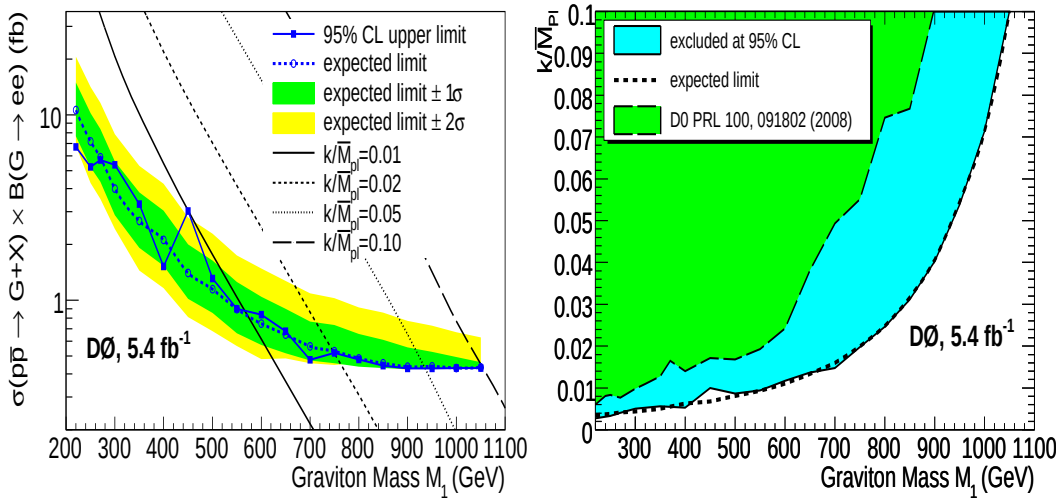
In the absence of a significant excess of data over background, upper limits on the KK heavy graviton production cross section times the corresponding branching fraction (e^+e^-) are set using a log-likelihood ratio (LLR). The LLR tool takes into account systematic uncertainties on background predictions and signal efficiencies. The resulting exclusion limits for the KK heavy graviton production cross section times the e^+e^- branching fraction are given in table 1.3.2 and figure 1.6 (a). These limits can be translated into limits on the coupling $c = k/M_{Pl}$ as a function of mass M using the Randall-Sundrum model cross section predictions, shown in figure 1.6 (b). The lower limit on the KK heavy graviton mass M with coupling c = 0.1 is 1050 GeV/c^2, at 95% C.L. [22].

The corresponding limits for CDF, for a heavy graviton with coupling c = 0.1 are of 850 and 976 GeV for the ee and $\gamma\gamma$ channels, respectively, using 2.5 and 5.4 fb^{-1}, respectively.



a b

Figure 1.5: Invariant mass spectrum from (a) ee and (b) $\gamma\gamma$ data (points). Superimposed are the fitted total background shape from SM processes including instrumental background (open histogram) and the fitted contribution from events with misidentified clusters alone (shaded histogram). The open histogram with dashed line shows the signal expected from KK heavy gravitons with $M_1 = 300$ GeV, 450 GeV, 600 GeV (from left to right) and $k/\overline{M}_{\rm Pl} = 0.02$ on top of the total background. Invariant masses below 200 GeV are taken as the control region [22].



a b

Figure 1.6: (a) 95% C.L. upper limit on $\sigma(p\bar{p} \to G + X) \times B(G \to ee)$ from 5.4 fb^{-1} of integrated luminosity compared with the expected limit and the theoretical predictions for different couplings $k/\overline{M}_{\rm Pl}$. (b) 95% C.L. upper limit on $k/\overline{M}_{\rm Pl}$ versus the heavy graviton mass M_1 from 5.4 fb^{-1} of integrated luminosity compared with the expected limit and the previously published exclusion contour [22].

Model	Nominal		Conservative	
Lower Mass Limit GeV/c^2	Expected	Observed	Expected	Observed
Z'_{SSM}	949	950	942	944
Z'_{η}	844	810	837	800
Z'_{χ}	834	800	827	787
Z'_{ψ}	817	763	809	751
RS G (c=0.1)	826	786	819	767
RS G (c=0.07)	767	708	758	700

Table 1.1: Expected and observed lower mass limits for SSM Z', E6 Z' models and RS heavy gravitons [20], from the D0 experiment (3.6 fb^{-1}).

Heavy Graviton Mass	$\sigma \times B(G \to ee)$ (fb)		Coupling $k/\overline{M}_{\rm Pl}$	
GeV	Expected	Observed	Expected	Observed
220	10.62	6.71	0.0034	0.0027
250	7.18	5.23	0.0038	0.0033
270	5.91	5.69	0.0042	0.0041
300	4.00	5.37	0.0044	0.0050
350	2.67	3.30	0.0051	0.0056
400	2.12	1.52	0.0062	0.0053
450	1.40	3.03	0.0068	0.0099
500	1.15	1.31	0.0081	0.0087
550	0.89	0.90	0.0093	0.0094
600	0.75	0.84	0.0111	0.0117
650	0.65	0.68	0.0133	0.0136
700	0.56	0.48	0.0160	0.0147
750	0.53	0.52	0.0199	0.0197
800	0.48	0.48	0.0248	0.0247
850	0.46	0.44	0.0316	0.0312
900	0.44	0.43	0.0406	0.0403
950	0.44	0.43	0.0545	0.0539
1000	0.43	0.43	0.0713	0.0713
1050	0.43	0.43	0.0969	0.0964

Table 1.2: 95% C.L. upper limit on $\sigma(p\bar{p} \to G + X) \times B(G \to ee)$ and coupling $k/\overline{M}_{\rm Pl}$ from 5.4 fb^{-1} of integrated luminosity from D0 experiment [22].

Chapter 2

Physics at the Large Hadron Collider

Colliding apparatus, to probe the fundamental structure of matter, have been exploited since long ago. The need to probe smaller constituents has led to the design of always higher energy colliding setups, over several decades, since the energy of the probe is inversely related to its wavelength, following the de Broglie relation $\lambda = h/|\vec{p}|$ where $\vec{p}$, the three-momentum of the probe, is related to its energy.

Three main types of colliding machines can be highlighted: the e^+e^- lepton colliders, the hadron colliders which can be proton-proton or proton-antiproton colliders, and the lepton-hadron colliders (electron-proton) which probe the structure of the proton through deep inelastic scattering. While for e^+e^- colliders, the whole energy from the electrons is available in the centre of mass frame, the same is not true for pp or $p\bar{p}$ colliders since protons are composite particles. The e^+e^- colliders are thus able to perform precision measurements, as the centre of mass energy can be tuned to select very specific processes[1].

The main drawback of e^+e^- colliders, however, is their considerable energy loss by synchrotron radiation, which is inversely proportional to the fourth power of the mass m^{-4}. A way to overcome this problem is to accelerate protons instead of electrons, which is one of the main motivations for hadron colliders.

The Large Hadron Collider (LHC) is designed to probe physics at the high energy frontier. It is thus a discovery machine at a very high energy working regime.

The work presented in this thesis is performed in the experimental environment of the (LHC). The LHC is a proton-proton accelerating and colliding system located on the French-Swiss border near Geneva. It started collecting data in November 2009 at a centre of mass energy of 900 GeV and became the highest energy collider a few weeks later with 2.36 TeV in the centre of mass. Since March 2010, the running centre of mass energy has increased to 7 TeV.

This chapter is dedicated to the presentation of the LHC and is organised as follows: the motivations for the LHC are summarized in section 2.1 and the LHC design performance in section 2.2. Section 2.3 presents the LHC data delivery up to June 2010. Section 2.4 of the chapter presents the characteristics of proton proton interactions and section 2.5 introduces the cross sections at LHC.

[1]During the period from 1989 to 1995 the LEP at CERN, an accelerating electron-positron collider performed precise measurements of the Z boson parameters by scanning the centre of mass energy around the Z boson mass of 91 GeV/c^2.

2.1 Motivations for the LHC

The main motivations for the LHC are the search for the Brout-Englert-Higgs boson, assumed to be responsible for the mass of particles (see section 1.1.3), and the search for new physics beyond the Standard Model (see section 1.2). Other important aspects will also be studied, like a deeper understanding of the Standard Model:

- the search for the Brout-Englert-Higgs boson: The origin of mass of particles is still an open question for physicists. The most promising theory to address this question was developed independently by Brout, Englert [8] and Higgs [9]. It invokes a spontaneous symmetry breaking of the Standard Model lagrangian and predicts the existence of a new scalar boson called the Higgs boson (see section 1.1.3). Its mass is however unpredicted by the theory and the cross sections for the various processes involving a Higgs boson depend on its mass. The Higgs boson was searched for at LEP, which put a lower limit for its mass at 115 GeV/c^2 [23]. It is also investigated at the Tevatron which has excluded at 95% C.L. the SM Higgs in the mass range between 158 and 175 GeV/c^2 [24]. Moreover, precision electroweak measurements constrain its mass to be lower than 186 GeV/c^{2}at 95% C.L. If the Higgs boson from the Standard Model does exist, it will certainly be discovered at the LHC.

- the search for new physics as from Supersymmetry: Supersymmetry is proposed as a symmetry between fermions and bosons. For each particle in the Standard Model, a superpartner is predicted. This theory has the advantage of solving at once several open issues such as the naturalness problem and the unification of gauge couplings, and it provides a good dark matter candidate. Possible processes involving superpartners will be investigated at the LHC.

- the search for new physics as from a Grand Unified Theory (GUT): different gauge groups are possible candidates for Grand Unification, such as $SO(10)$ or $E(6)$ (see section 1.2.2). The group G_{GUT} is characterized by a single coupling constant. It leads to a new organization of particles in new multiplets and to new gauge bosons. The G_{GUT} symmetry may be broken at the TeV scale, possibly leading to a new $U(1)$ group. GUT theories predict new massive Z bosons (called Z') which can decay into a lepton or quark pair.

- the search for new physics as from extra dimensions: The Standard Model does not include gravitation. The difficulty in the unification of the four forces is the large scale difference between the Planck scale and the electroweak scale. To overcome this problem, a possibility is the modification of the space-time structure, supposing the existence of a '4+d' space-time structure, the d additional dimensions being compactified (see section 1.2.3). These models predict the existence of new massive particles as for example massive excitations of the graviton or massive gauge bosons ($M \sim$ TeV).

- Deeper understanding of the Standard Model: Known Standard Model processes will also be intensively studied, such as QCD, electroweak processes and top quark physics.

- Specific programmes have also been designed at the LHC to study heavy ion collisions and investigate CP violation through B meson study.

A wide range of physics studies will thus be covered at the LHC. A set of four detectors was designed to detect particles produced during proton-proton or ion-ion interactions. These

are located at four points along the ring, called the interaction points where the collisions between the two proton beams take place. A scheme of the LHC ring together with the detector locations is given in figure 2.1(a).

- CMS (Compact Muon Solenoid) and ATLAS (A Toroidal LHC ApparatuS): These are generic detectors aimed at covering a wide range of physics studies such as Higgs search, investigation of Supersymmetry, theories beyond the Standard Model and standard model physics (top quark physics, QCD and EWK). The CMS detector is described in some detail in chapter 3.
- LHCb (Large Hadron Collider Beauty): This detector will specifically study CP violation through B meson decay. CP violation is a key issue to understand the asymmetry between matter and antimatter in the early universe.
- ALICE (A Large Ion Collider Experiment): This will study processes involved in heavy ion collisions. The aim is to study the quark gluon plasma present in the early universe phase, in order to understand better this period of universe formation.

2.2 The LHC machine: design performance

The design performance of the LHC is governed by three main requirements:

- A high energy in the centre of mass, to explore the high energy range of physics covered by BSM theories, SUSY, ...
- A large luminosity, to account for the small cross sections predicted for Higgs production and BSM or SUSY processes.
- A high bunch crossing rate, to increase the interaction rate and thus the integrated luminosity.

In the following, the LHC design performance expected for 2013 is presented. Proton beams will be accelerated to an energy of 7 TeV, resulting in a centre of mass energy of 14 TeV. The luminosity is expected to be $2 \cdot 10^{33}$ cm^{-2}s^{-1} with bunch crossing every 25 ns. A scheme of the accelerating setup is given in figure 2.1(b).

The main component of the LHC accelerating complex is the 27 km long tunnel, situated underground between 80 and 100 m, where protons will be accelerated from their injection energy of 450 GeV up to the design energy of 7 TeV. This acceleration process is provided by 8 radio-frequency cavities (RF cavities) which boost the beams in total by 16 MeV per turn, by use of a 5.5 MV/m electric field oscillating at 400 MHz. The stability of the beam trajectory is ensured by magnets, of which 1232 superconducting dipole magnets keep the beams on a circular trajectory all along the ring. These magnets have to sustain a strong bending power related to the beam high energy, as expressed by the following relation:

$$E_{beam} = B_{dipole} \times 0.84 \text{ TeV}/T \tag{2.1}$$

which sets a magnetic field of 8.33 T for the magnets, for the nominal LHC beam energy. Such strong magnets have been specially designed for the LHC. They are complemented by 7000 additional magnets to clean and focus the beams.

To reach the design luminosity of the LHC, protons are gathered in packets, called bunches, of 10^{11} protons each, colliding every 25 ns. The beams are collimated to a transverse size of around 16 μm to enhance the collision probability.

Prior to their injection inside the LHC ring at an energy of 450 GeV, protons have already passed through a chain of beam accelerations and operations. They are first accelerated by a linear accelerator, then the 'Booster', and the Proton Synchrotron (PS), reaching the energy of 26 GeV. The PS ensures, in addition, the correct 25 ns spacing between bunches prior to their delivery to the SPS. The latter accelerates protons to the energy of 450 GeV and injects them to the main LHC ring.

2.3 LHC design parameters and plans

The first LHC collisions were performed on November 23rd, 2009 at 900 GeV centre of mass energy, closely followed (December 8th, 2009) by collisions at 2.36 TeV centre of mass energy. First collisions at 7 TeV were performed on March 30th, 2010. In each case, the first events were visible in the detectors and recorded on tape a few minutes later. The data processing chain from initial online triggering to recording on disk proceeded smoothly.

2.3.1 Luminosity measurement

Data taking is usually subdivided into stable periods of continuous data taking, called runs, which can last for several hours, depending on the stability of the beams. A usual variable to quantify the amount of data collected is the integrated luminosity:

$$L_{int} = \int L dt \tag{2.2}$$

where L is the instantaneous luminosity and dt is the duration of data taking. The instantaneous luminosity is defined precisely in [25] and depends on beam currents times the overlap section of the two beams. It is expected to slowly decrease as collisions go on since protons are lost from the beams. The measurement of the instantaneous luminosity is crucial as it provides normalization for all physics cross section measurements. In order to provide a stable measurement of luminosity over time, luminosity sections are defined. They correspond to stable data taking periods of relatively small duration (93 s) during which stable luminosity is expected and luminosity measurements can be averaged. Equation 2.2 becomes then:

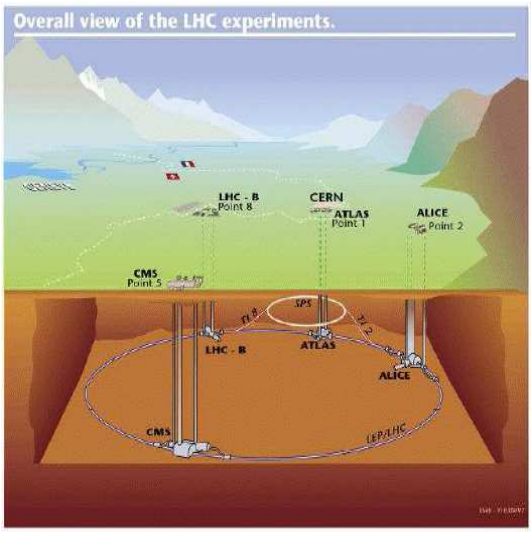
$$L_{int} = \sum_i L_{LS} \times 93s \tag{2.3}$$

where the index i runs over all luminosity sections and L_{LS} is the average instantaneous luminosity per luminosity section. A distinction needs however to be made between the luminosity delivered by the LHC and the luminosity recorded by CMS.

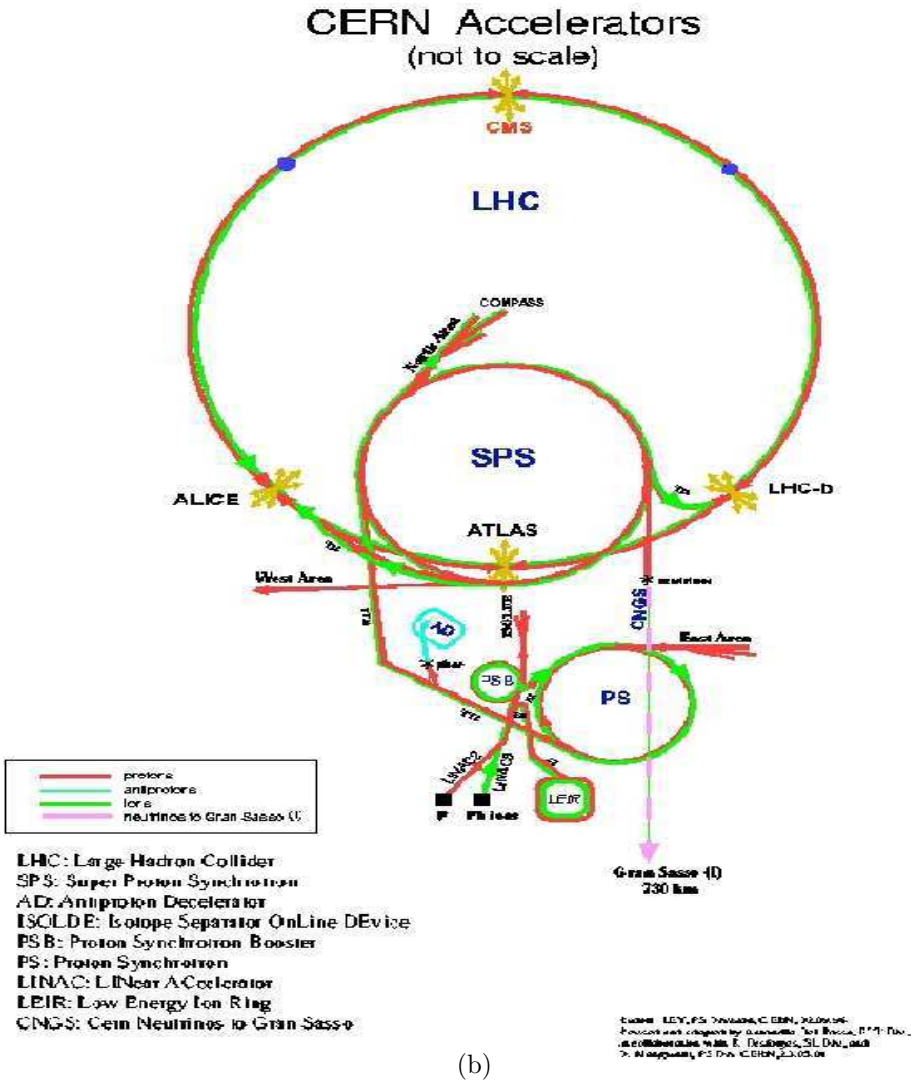
Many methods have been proposed and investigated [26] to provide a real-time luminosity measurement for CMS. The most widely used up to now is the so-called "zero counting" method [25] which uses the fraction of towers with no signal above a given threshold in the "Hadron Forward Calorimeters" (section 3.4). The mean number of interactions μ per bunch crossing is linked to the luminosity via the following relation:

$$\mu = \frac{\sigma L}{f_{BX}} \tag{2.4}$$

where L is the luminosity, f_{BX} is the bunch crossing rate and σ is the total inelastic and diffractive cross section, estimated to be $\sim$ 80 mb. A second method exploits the linear relationship between the total transverse energy in the "Hadron Forward Calorimeters" and the number of interactions and thus the luminosity.



(a)



(b)

Figure 2.1: (a) Overall view of the LHC and the four detectors.(b) The LHC accelerating setup.

Figure 2.2 shows the evolution of the delivered and recorded integrated luminosities up to August 30th 2010 together with the final total numbers. A total of 2.6 pb^{-1} has been delivered of which 2.3 pb^{-1} have been recorded.

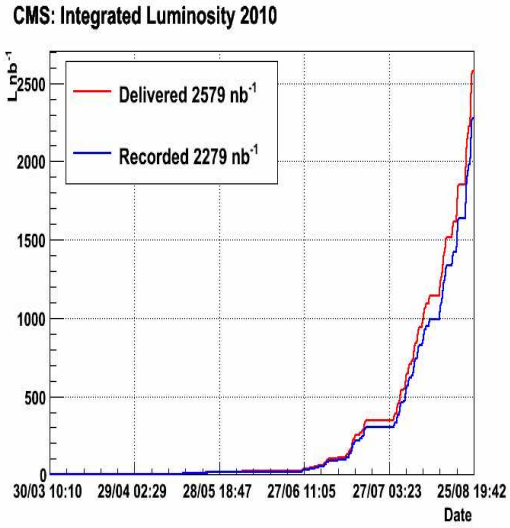


Figure 2.2: Time evolution of delivered and CMS recorded integrated luminosities up to July 19th 2010.

2.3.2 Plans for data taking up to 2020

A specific plan for data-taking has been discussed by the CERN Council up to 2020. This plan assumes an operation mode of the accelerator based on blocks of 2 (3) years interleaved by major shutdown periods. Three main periods are thus identified:

- A two-year data-taking period at 7 TeV in the centre of mass in 2010 and 2011. A total integrated luminosity of 1 fb^{-1} and an instantaneous luminosity of 10^{32} $\text{cm}^{-2}\text{s}^{-1}$ are expected for the end of this run period.
- A 16 months shutdown in the year 2012 to raise the energy to 13-14 TeV in the centre of mass; the luminosity should be above 10^{33} $\text{cm}^{-2}\text{s}^{-1}$ during years 2013, 2014 and possibly 2015.
- A second 16 months shutdown in 2015 (possibly 2016) to raise the luminosity with the help of the new Linac4. The luminosity should be above 10^{34} $\text{cm}^{-2}\text{s}^{-1}$.
- A 19 months shutdown in the year 2020 to allow for the LHC high luminosity phase ($L \sim 4 - 5$ 10^{34} $\text{cm}^{-2}\text{s}^{-1}$).

2.4 Proton-proton interactions

Due to the compositeness of protons, proton-proton interactions involve complex physics processes. The description of such processes is performed using the parton model introduced by Feynman in 1969 to analyze high energy hadron collisions. The very high energy involved in proton interactions at LHC allows the modelling of protons as sets of partons (quarks and

gluons) each carrying, at leading order, a fraction x of the proton momentum, as measured in a fast moving frame ('Breit' frame). These proton constituents are described by the so-called parton density functions (PDF), $f(x, Q^2)$, which quantify the density of partons of type i inside the proton, with momentum fraction x, at a scale Q^2. The scattering of two partons, with momentum fractions x_1 and x_2, forms when a large scale is present (large transverse momentum) what is called a hard event or hard scattering. The cross section for any process in pp collisions ($\sigma(pp \to c + d + X)$) is then derived from the partonic cross section ($\hat{\sigma}(a + b \to c + d)$) with use of the parton density functions:

$$\sigma(pp \to c + d + X) = \sum_{partons\ i,j} \int dx_a\ f_i^a(x_a) \int dx_b\ f_j^b(x_b)\ \hat{\sigma}(a + b \to c + d) \tag{2.5}$$

where i, j are the two partons participating in the hard event, $f_i^a(x_a)$, $f_j^b(x_b)$ the parton density functions for a parton of type i in the first proton with momentum fraction x_a and a parton of type j in the second proton with momentum fraction x_b, respectively. The total partonic centre of mass energy is given by $\hat{s} = x_1 x_2 s$.

The evolution of the parton density functions with respect to the scale Q^2 is provided by the DGLAP equations. The parton density function at a scale Q_1^2 can then be predicted from the corresponding density measured experimentally at a scale Q_2^2 provided that Q_1^2 and Q_2^2 are in the perturbative domain. It allows to perform a fit of the measurements at different scales.

The universality of parton density functions is essential as it allows using deep inelastic scattering measurements of the proton structure to different processes with different energies and hard scales, as achievable at the LHC.

PDF measurements are the subject of intensive research activities as it is an essential tool for the interpretation of hadron collisions. Two main research groups are deeply involved in such activities: the CTEQ and MRST groups. They both use a large sample of measurements from deep inelastic scattering at HERA, neutrino-nucleus target experiments and $p\bar{p}$ collider at Tevatron. They differ in several details in the parametrization of the fits.

At next to leading order (NLO), the accelerated partons (coloured/charged objects) can undergo gluon (photon) radiation. The radiation of extra partons (photons) from initial and final state partons, is called initial state radiation (ISR) and final state radiation (FSR), respectively.

Besides the hard process between two partons, other partons, not involved in the hard scattering, can also interact in the same proton-proton collision. This is called multiple parton interactions (MPI). In addition, beam-beam remnants (BBR) resulting from the hadronization of the partonic constituents that did not participate in other scatterings are also possible. Products of MPI and BBR form what is called the underlying event. The description of such phenomena makes use of phenomenological models. The underlying activity is extensively studied at the Tevatron. The extrapolation from Tevatron energies to LHC energies is however not fully mastered [27].

2.5 LHC cross sections

At the LHC, the total cross section (for the initial two year data taking period) at 7 TeV centre of mass energy is:

$$\sigma_{tot} \sim 100 \text{ mb}. \tag{2.6}$$

For comparison, at 7 TeV centre of mass energy, the cross section for the Drell-Yan events ($pp \rightarrow \gamma/Z \rightarrow e^+e^-X$) with mass above 40 GeV/c^2 is of the order of 10^{-7} mb, the cross section for a 1 TeV/c^2 Z'_{SSM} production ($pp \rightarrow Z' \rightarrow e^+e^-X$) is of the order of 10^{-11} mb and the cross section for QCD dijet production with transverse momentum of the initial parton ($\hat{p_t}$) higher than 15 GeV is of the order of 0.1 mb[2].

[2]These cross-sections were estimated from Monte-Carlo simulation [28].

Chapter 3

The CMS experiment

This chapter describes the 'Compact Muon Solenoid' (CMS) particle detector. After a short presentation of the general detector layout in section 3.1, the four main subdetectors of CMS are presented in sections 3.2 to 3.6 with an emphasis on the systems particularly relevant to this analysis: the tracker system and the electromagnetic calorimeter (ECAL). Section 3.7 presents the trigger system. A detailed description of CMS is given in [29, 30].

3.1 Layout of the experiment

The CMS experiment is located at the activity point 5 on the LHC ring (cf. figure 2.1(a)). The generic detector has a common detector structure where the different parts of the detector (subdetectors) are inserted into each other, starting from the interaction point. It consists of two main parts: a first cylindrical part called the barrel and two circular parts called the endcaps which locate on both sides of the barrel along the beam axis to ensure a maximum detection coverage. A layout of the CMS detector is presented in figure 3.1.

The different subdetectors are clearly visible (from the beam axis towards the outer of the detector):

- The tracker allows to determine trajectories followed by charged particles in the detector based on impact points left by the particle when crossing the layers of the tracker. The curvature and angles of the reconstructed trajectory in the magnetic field allow a measurement of the particle momentum and the sign of its charge.

- The electromagnetic calorimeter (ECAL) measures the position and energy of electrons and photons.

- The hadronic calorimeter (HCAL) measures the position and energy of hadrons involved inside jets of particles.

- The solenoid provides a uniform magnetic field of 3.8 T in the whole detector region (tracker, ECAL and HCAL). The flux is returned through a 10000 ton yoke made of 5 wheels and 2 endcaps interleaved with the muon chambers.

- The muon chambers reconstruct muon trajectories based on their impact points in the muon chambers. The curvature in the magnetic field allows the muon momentum measurement. Together with trajectories reconstructed in the tracker, they can also constrain the momentum measurement.

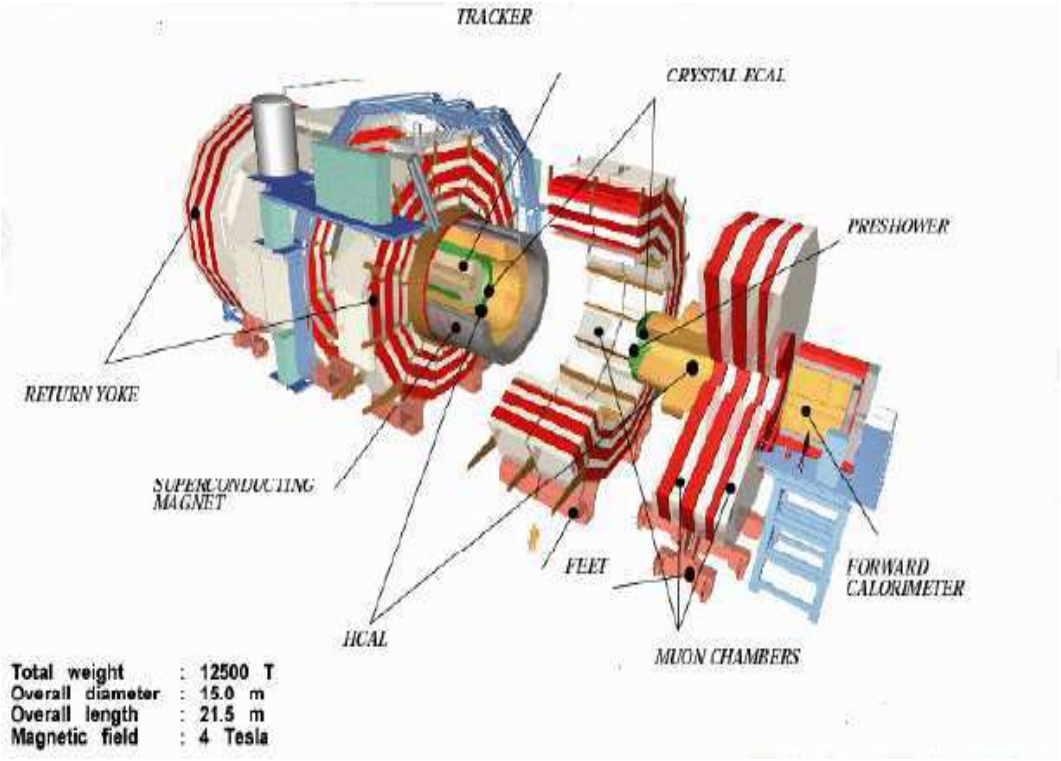


Figure 3.1: (a) A layout of the CMS detector.

Coordinate conventions

Coordinate conventions are necessary to locate the energy deposits in the ECAL and HCAL or the track parameters in the tracker and/or muon chambers. The cylindrical coordinate system is naturally chosen. The z axis is set at the beam axis where beams circulate, the x axis points towards the center of the ring and the y axis, perpendicular to the x and z axis, points towards the surface. The x and y axis form the transverse plane.

The θ coordinate measures the angle with respect to the z axis. The ϕ coordinate measures the angle, in the transverse plane, with respect to the x axis. A widely used coordinate variable is the pseudorapidity η, which is correlated to the θ angle through the formula:

$$\eta = -\ln(\tan(\frac{\theta}{2})). \tag{3.1}$$

As explained in section 2.2, the interactions at LHC constitute a very dense environment in terms of collision rate. Such elements were taken into account in the design of the CMS detector since the early research and development phase. Three main items were highlighted dealing with the constraints imposed by the dense LHC environment on CMS design:

- Radiation hardness: The intense number of particles produced at the LHC during the design luminosity phase is likely to produce serious radiation effects on the different parts of the detector, especially the inner tracker part. A high radiation resistance constraint was brought on the construction of the different detection elements of CMS.
- Fast readout electronics: The large amount of data produced at LHC is processed by readout electronics before being transmitted for storage. Thus, fast readout electronics is necessary in order to sustain the 25 ns bunch crossing rate.

- High granularity: A high detector granularity is needed in order to reduce the occupancy.

The following describes the different subdetectors of CMS.

3.2 The tracker

The tracker reconstructs trajectories of particles (tracks) from hits left by the particle on its way and measures their momentum and charge in the magnetic field based on the track curvature. The performance of the tracker is thus constrained by two main parameters: the hit reconstruction efficiency and the track reconstruction efficiency and parameter estimation from the hits. As the closest subdetector to the beam axis, it also has to sustain a hard radiation environment.

The tracker divides into two main parts: the silicon pixel system and the silicon strip system. The geometry of the silicon modules is different in these two subsystems. Each subsystem can divide into the barrel part and the endcap part.

The most basic detection element in the tracker is called a module which principle is mainly based on a semiconducting device. Semiconductors can be advantageous in the sense that they classify between conductors and insulating in terms of conductivity. Particles crossing the module surface thus ionise the semiconducting medium leading to the creation of electrons caused by the electric field from the high voltage imposed on the semiconducting device. These charges are then collected in an integrated circuit called the APV25 chip which will sample the signal at the LHC frequency of 40 MHz and buffer the data for about 3.2 μs. On receipt of a trigger signal, the signal data are then sent to readout electronics for further processing.

3.2.1 The silicon pixel system

The silicon pixel system represents the inner core of the tracker. It will allow for track segment reconstruction, providing three impact points. Its close proximity to the interaction point makes it also ideal to reconstruct the position of the primary vertex from which the particles originate from and the position of the particle decay secondary vertices.

The modules are made of 66 million 100×150 μm^2 almost square pixels, totaling an active area of around 1 m^2, thus providing this subsystem a high granularity. A layout of the pixel detector is provided in figure 3.2.

It consists of two parts. The barrel part is made of three 53 cm long cylindrical layers placed at radii of 4.4, 7.3 and 10.2 cm respectively. They gather together 48 million pixel modules and will allow coverage for central pseudorapidity tracks. The barrel is complemented by the pixel disks, two on each side of the detector placed at $|z| = 34.5$ and 46.5 cm respectively with radii extending from approximately 6 cm to 15 cm, covering the forward pseudorapidity region. The modules on the pixel disks are tilted by approximately 20°, resulting in a turbine-like geometry.

3.2.2 The silicon strip system

The silicon strip system is divided into four parts. The central region is covered by the Tracker Inner Barrel (TIB) and the Tracker Outer Barrel (TOB). The TIB is composed of four cylindrical layers of modules covering the region 25 cm $<$ r $<$ 50 cm. The TOB is composed of six cylindrical regions covering the region 50 cm $<$ r $<$ 116 cm and $|z| < 118$ cm. The forward region is covered by the Tracker Inner Disks (TID), three disks on each side of

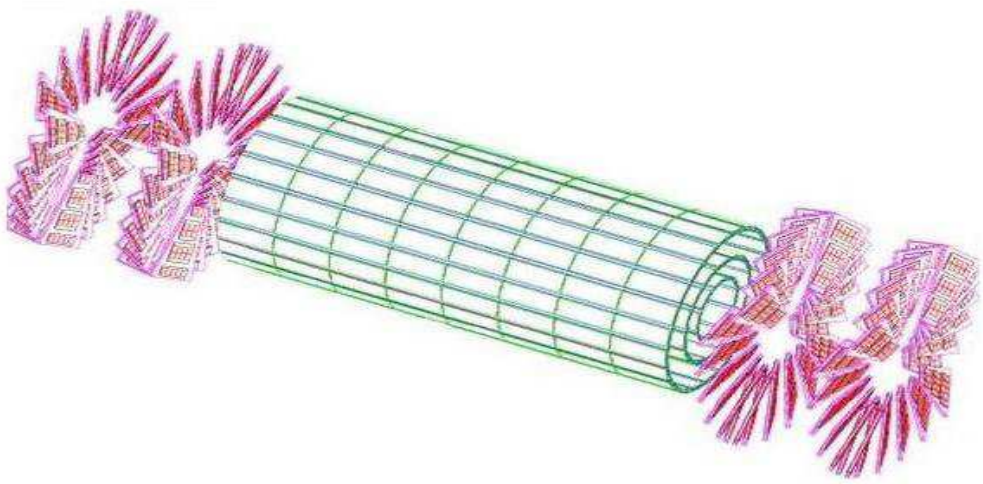


Figure 3.2: (a) A layout of the CMS pixel subsystem.

TIB with each disk made of three rings, and the Tracker Endcaps (TEC) composed of 9 circular so-called wheels covering the region 130 cm $<$ |z| $<$ 270 cm. The first three wheels are composed of 7 rings, the wheels 4 to 6 are composed of 6 rings, the wheels 7-8 are composed of 5 rings and the last wheel is made of 4 rings. Each of this wheel is also divided in ϕ into 8 "petal"-shaped structures also called petals.

The TIB/TID part delivers up to 4 r—ϕ measurements for trajectory reconstruction. The TOB provides another set of 6 r—ϕ measurements for trajectory reconstruction. The TEC provide up to 9 ϕ measurements for trajectory reconstruction.

In addition, the first two layers and rings, respectively, of TIB, TID and TOB as well as the rings 1, 2 and 5 of TEC contain, for each module, a second module mounted back to back to the first one, with an angle of 100 mrad, thus providing a measurement of the second coordinate (z in the barrel and r on the disks). Such layers/rings are qualified as double-sided.

A schematic r—z view of the whole CMS tracker (pixel + strip) is presented in figure 3.3.

The CMS strip tracker, totaling a 198 m^2 silicon detection active area, with its 15148 modules and 9.3 million strips, ensures a minimum of 9 hits, in the full range of the acceptance ($\eta < 2.5$), of which at least 4 are two-dimensional measurements.

3.2.3 Performance

The performance of the tracker relies essentially on the track reconstruction efficiency and parameter estimation based on hits provided by the tracker. Such parameters are the pseudo-rapidity η, the azimuthal angle, the three coordinates (x,y,z) of the point of closest approach to the primary vertex (also called impact parameter) and finally the transverse momentum (p_t).

Track Reconstruction efficiency

The track reconstruction efficiency performance divides into two points:

- The hit reconstruction efficiency, namely the efficiency with which the tracker identifies hits left by the charged particle on its passage. Figure 3.4 presents the number of

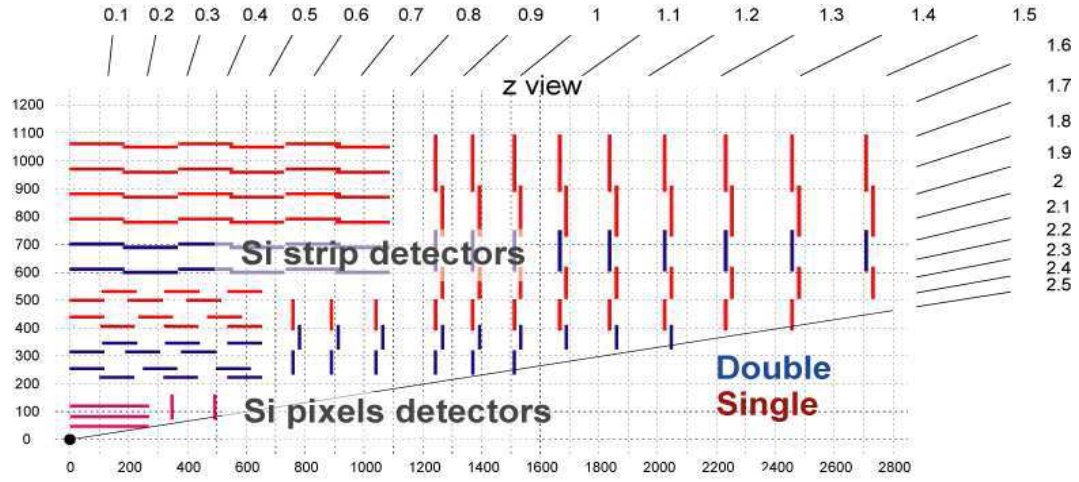


Figure 3.3: (a) A schematic r—z view of the CMS tracker.

measurement points as a a function of the pseudorapidity. As one can see, the minimum requirement of 9 measurements with at least 4 two-dimensional measurements is often fulfilled.

- The efficiency for track reconstruction based on hits provided by the tracker was estimated using samples of muons and pions with transverse momenta of 1, 10 and 100 GeV/c. This efficiency is plotted as a function of $|\eta|$ in figure 3.5 for muons (a) and pions (b). As one can see, the track reconstruction is excellent for pseudorapidity values up to 2.4 up to which the coverage reduces significantly. The performance is however reduced in the case of pions which are much more affected by interactions in the tracker material.

Track Parameter Estimation

The resolution of the transverse momentum, the pseudorapidity, the azimuthal angle and the impact parameter coordinates is provided in figure 3.6 for muons with transverse momenta of 1, 10 and 100 GeV/c.

For muons at high momentum (100 GeV/c), the transverse momentum resolution(3.6(a)) is around 1-2% up to a pseudorapidity of 1.6 after which the lever arm starts to decrease. At lower momentum, the resolution is dominated by multiple scattering effects and the distribution reflects the amount of material in the tracker as a function of η. The resolution of the longitudinal and transverse impact parameters(3.6 (d) and (e)) are fairly constant at high momentum since they are dominated by the resolution of the first hit in the pixel detector. At lower momenta, the resolution is dominated by multiple scattering. One sees, however, an improvement in the longitudinal impact parameter resolution up to pseudorapidites $\eta \sim 0.5$. The resolutions on θ and ϕ show fairly the same shape as the longitudinal and transverse impact parameter resolutions respectively, as expected.

Tracker Material Budget

Particles crossing the tracker can interact with the material, especially electrons which can radiate photons through Bremsstrahlung process and thus lose significant part of their

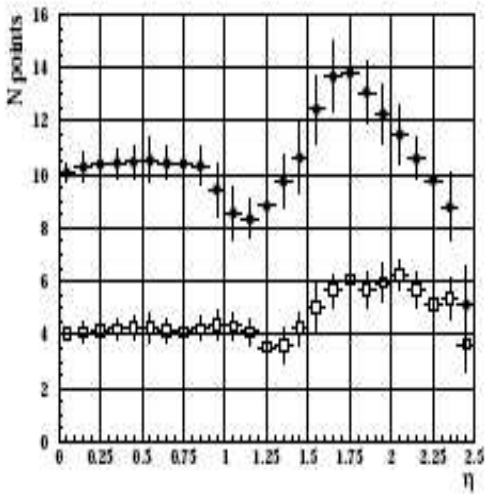
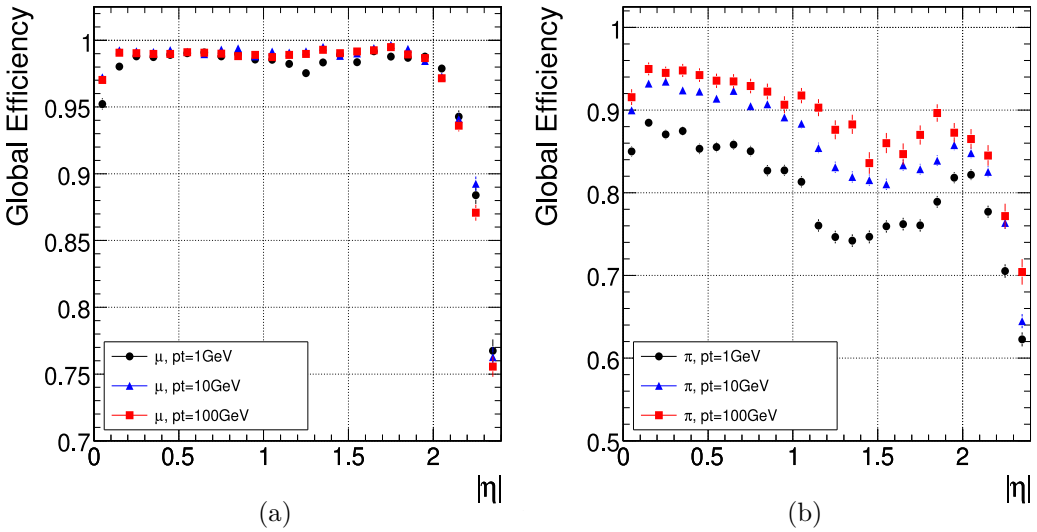


Figure 3.4: (a) Number of measurements of the strip tracker as a function of the pseudorapidity η. Filled circles show the total number (stereo module hits count as one) while open squares show the number of stereo hits [30].



(a) (b)

Figure 3.5: (a) Track reconstruction efficiency as a function of the pseudorapidity $|\eta|$ for muons (a) and pions (b) with transverse momenta of 1, 10 and 100 GeV/c [30].

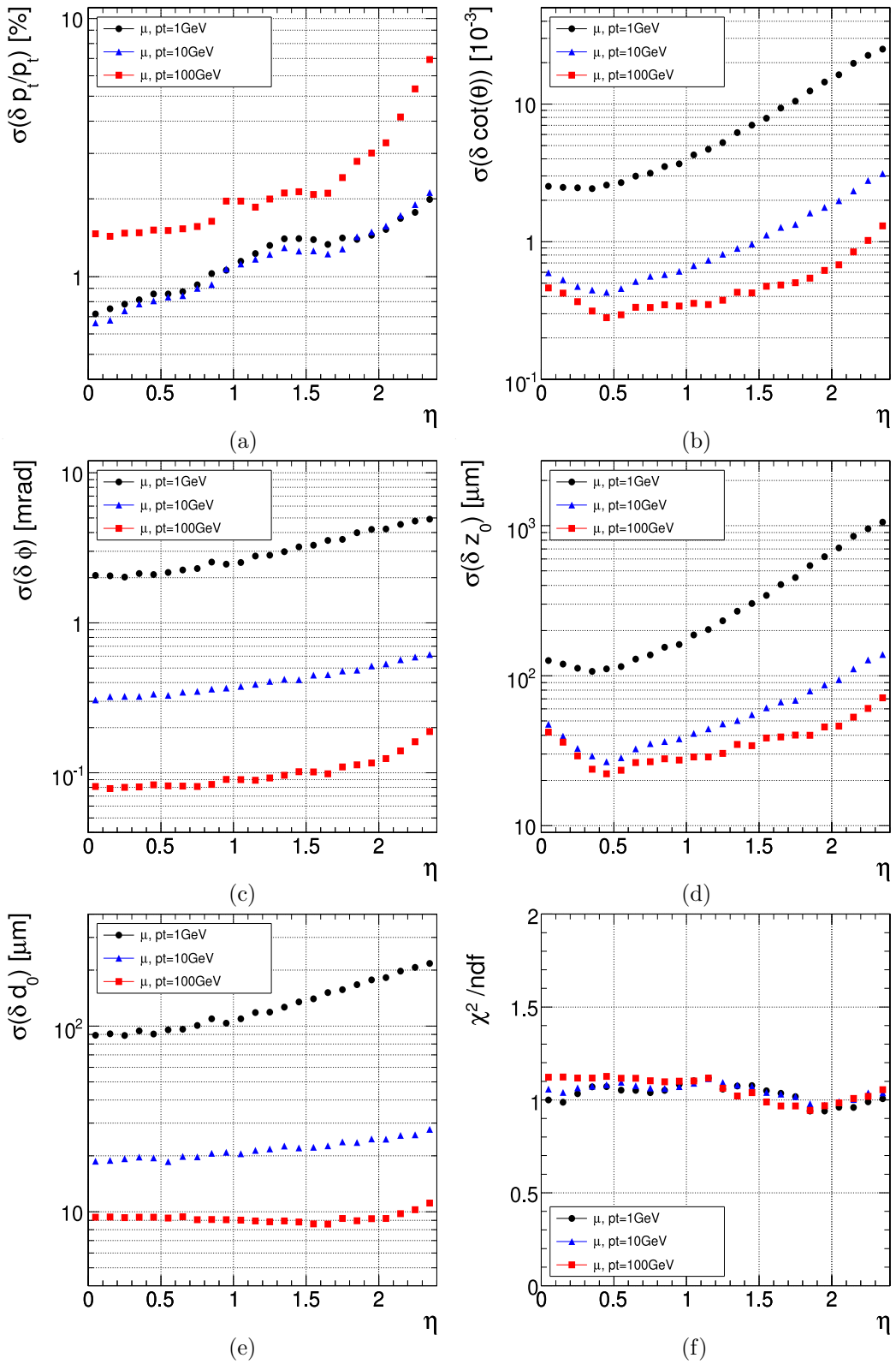


Figure 3.6: Track parameters estimate resolution as a function of pseudorapidity for muons with transverse momenta of 1, 10 and 100 GeV/c: transverse momentum (a), $\cot\theta$ (b), ϕ (c), longitudinal (d), transverse impact parameter (e) and reduced χ^2(f) [30].

energy before having it measured in the electromagnetic calorimeter. Strong pressure was put to minimise the amount of tracker material budget but still, after final construction and installation of the tracker, this amount was found to be a bit higher than expected. It is therefore important to be able to estimate the latter. Figure 3.7 presents the amount of tracker material budget expressed in units of radiation lengths, for the different subsystems, as a function of pseudorapidity $|\eta|$. It increases from 0.4 X0 at $|\eta| \sim 0$ to about 1.8 X_0 at $|\eta| \sim 1.4$, beyond which it falls to about 1 X0 at $|\eta| \sim 2.5$.

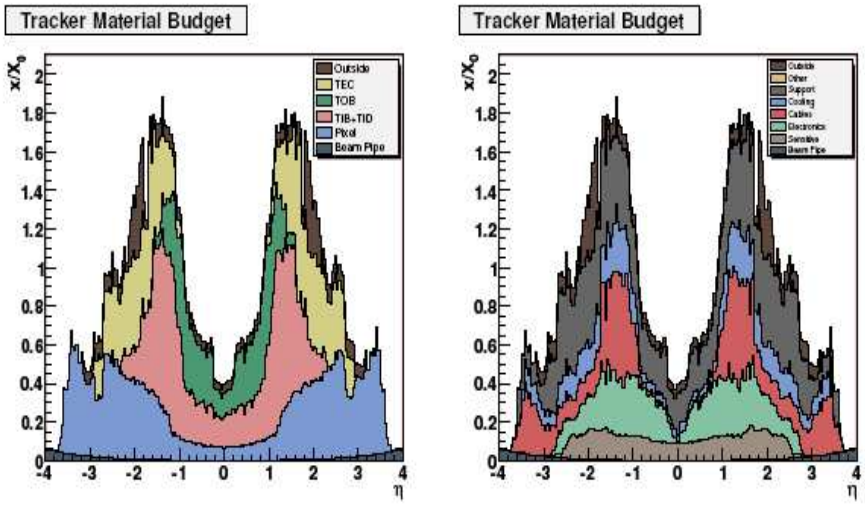


Figure 3.7: (a) Tracker material budget, in units of radiation lengths as a function of pseudorapidity η, for the different subsystems (a) and for the functional contributions (b) [30].

3.3 The electromagnetic calorimeter

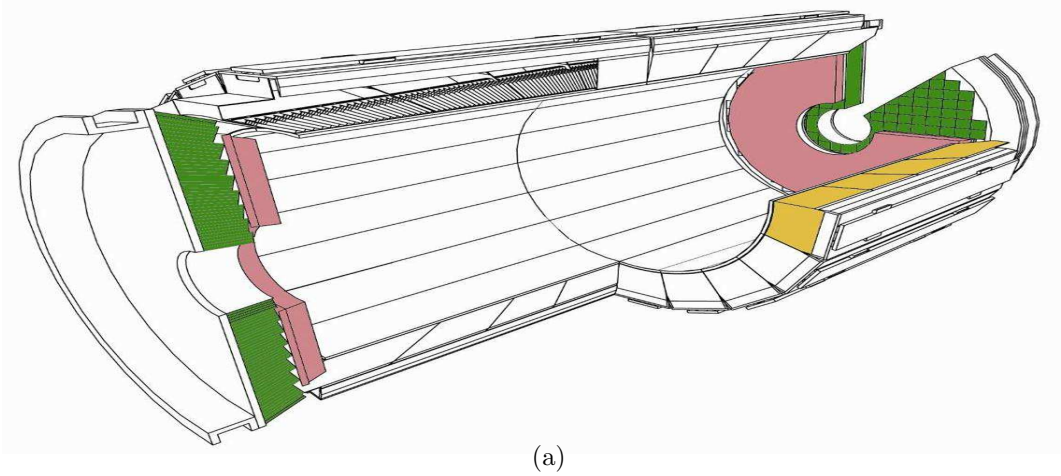
A strong pressure was put in CMS to place the calorimetry (ECAL+HCAL) inside the solenoid to limit the energy loss interactions and thus improve the energy resolution. This of course brought strong constraints on the size of the electromagnetic calorimeter. In addition, a fast response calorimeter and radiation resistant electronics were needed in order to sustain the dense environment of the LHC in terms of collision rate and radiation hardness.

The principle of the CMS ECAL relies on the development of a so called electromagnetic shower initiated by an electron or a photon penetrating a scintillating crystal. Electrons interact with scintillators to radiate photons through Bremsstrahlung processes. These photons also interact to give electron pairs through pair creation, the latter electrons interacting again to give photons, and so on. This process goes on until whole or significant part of the initial electron energy is converted into light. The same way photons penetrate the electromagnetic calorimeter and interact to give electron pairs through pair creation which can then interact to give photons through Bremsstrahlung process. Scintilation light is then created as a result of the interactions of the electrons with the medium. Two quantities are often used to quantify the development of an electromagnetic shower: the radiation length X_0, which quantifies the longitudinal development, is, in average, the distance after which the electron has lost a fraction of its energy equal to $1/e$ and the Moliere radius which quantifies the lateral development. By definition, 95% of the shower energy is contained in a Moliere radius. The length of the crystals is a key issue to achieve a total containment of the shower inside

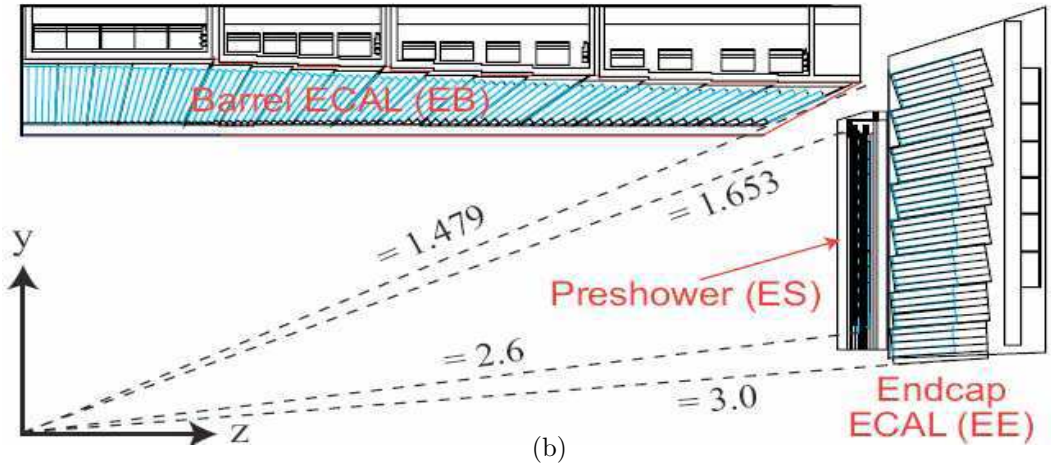
the scintilating medium and thus convert fully the initial electron energy into light in order to measure it efficiently. A typical electromagnetic shower for a few GeV electron/photon often extends laterally in several crystals. The energy is then measured from the individual crystal energy fractions and the position is computed as the barycenter of the hit crystals with weights correlated to the energy fractions.

The choice on the design of the electromagnetic calorimeter was put on a compact and homogeneous lead tungstate scintillating crystal calorimeter. This material has high density ($8.3\ \mathrm{g \cdot cm^{-2}}$), a small Moliere radius (21 mm) and radiation length ($X_0 = 0.89$ cm) and is reasonably resistant to high radiation. In addition, it has fast response (80% of the light is emitted in 25 ns, i.e. one bunch-crossing) but its light yield is relatively limited; approximately 100 photons are emitted per MeV. Signal from the crystals needs therefore to be amplified.

The general layout of the ECAL divides into two parts: a cylindrical shaped structure around the beam axis called the barrel and two circular shaped structures called the endcaps located on each side of the barrel. The crystal geometry and arrangement is different in these two substructures. Figure 3.8(a) gives a layout of the ECAL and figure 3.8(b) gives a schematic r—z view.



(a)



(b)

Figure 3.8: (a) Layout of the ECAL (b) Schematic r—z view.

3.3.1 The barrel

The barrel part has an inner radius of 129 cm and covers the central pseudorapidity region $|\eta| < 1.48$. Crystals have trapezoidal shape with front and rear face dimensions of 22×22 mm^2 and 26×26 mm^2 respectively, equivalent to $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$ and corresponding approximately to one Moliere radius. The crystals length is equal to 230 mm, or 25.8 radiation lengths. Crystals are arranged in submodules (2 rows of five modules) which gather in ϕ to form modules. Modules are then gathered together in η to form supermodules (20 crystals in $\phi \times 85$ crystals in η). A total of 36 supermodules (18 in each half barrel along the beam axis) totaling 61200 crystals constitutes the barrel. Scintillation light provided by the crystals is then directed to so called avalanche photodiodes (APD) which convert light into analog signal, pre-amplify it and send it to outer electronics for further processing and trigger decision.

3.3.2 The endcaps

The endcaps cover the forward region $1.56 < |\eta| < 3.0$ and are placed at $|z| = 314$ cm. Each circular shaped endcap is formed from two structures with a "D" shape called Dees. The crystal front and rear face dimensions are slightly different compared to the barrel part (28.6×28.6 mm^2 and 30×30 mm^2 respectively) and their length is slightly reduced (220 mm) corresponding to 24.7 X_0. Crystals gather in 5×5 matrices called supercrystals. A total of 14648 crystals is present in both endcaps. Due to the higher expected hadron fluency in the forward region, more radiation resistant devices, the vacuum phototriodes (VPT) have been selected for light to analog signal conversion and preamplification.

3.3.3 The preshower

In addition, in both endcaps, a preshower calorimeter was specifically designed to identify neutral pions which decay into two photons almost immediately due to their very short lifetime (8.4×10^{-7} s). It consists of a 3 X_0 sampling calorimeter made of lead to initiate the shower and silicon strips to measure the position of the first charged particle (electrons/positrons) of the electromagnetic shower. This device allows to separate angularly the two photons from the pion and thus distinguish it from a single photon energy deposit.

3.3.4 ECAL Performance

The performance of the electromagnetic calorimeter relies essentially on the energy measurement resolution since the position estimate is very well determined due to its high granularity. This resolution can be expressed using the following formula:

$$(\frac{\sigma}{E})^2 = (\frac{a}{\sqrt{E}})^2 + (\frac{b}{E})^2 + c^2, \tag{3.2}$$

where a stands for the stochastic term (0.05 GeV$^{1/2}$), b describes the contribution from electronics noise (150-250 MeV depending on η) and $c = 0.6\%$ is the constant term; the energy is expressed in GeV. For high energy (hundreds of GeV) electrons, the first two contributions become clearly negligible and the resolution reduces to the constant term contribution of

0.6%, which is an excellent resolution for electromagnetic calorimeters.

The gap and cracks

The region between the barrel and the endcaps ($1.48 < |\eta| < 1.56$) holds the cabling service that provides voltage to the electronics and constitutes a region where the energy estimate resolution is significantly reduced. A lower energy resolution may happen when the electrons cross the detector in spaces between crystals when gathered into submodules, between submodules when gathered into modules, between modules when gathered into supermodules and finally between supermodules themselves. Such regions are called, in a generic way, cracks. Note that the crystals in the barrel and the endcaps are tilted to avoid electrons to be emitted in the cracks. A study on energy resolution as a function of the pseudorapidity is documented here [31].

Energy Leakage

The electromagnetic calorimeter of CMS was mainly designed to measure the energy and position of electrons and photons of energy typically of $\sim 50 - 100$ GeV (for example coming from the decay of a 120 GeV Higgs decaying into two photons). The ability of the ECAL to fully contain the shower, also called the containment, becomes limited when dealing with very high energy (TeV) electrons or photons. In such cases, the showering process finishes in the HCAL cells located directly behind the crystals hits by the electron/photon. Such phenomenon is called energy leakage. A thesis work documented here [32] has been carried out to recover the initial electron energy in such regions, based on the energy deposit in the HCAL cell located directly behind the hit crystals.

Saturation effects

For very high energy (TeV) electrons, when the energy of the hottest crystal in the ECAL energy deposit reaches 1.4 TeV in the barrel (3.0 TeV in the endcaps), the preamplification process in the APD/VPT starts to saturate due to the very high intensity of the analog signal, putting a threshold on the energy measurement. A thesis work described here [33] has been performed to recover the energy of the hottest crystal in such cases, based on the energy fractions deposited in the neighbouring crystals.

3.4 The hadronic calorimeter

The CMS HCAL is a sampling calorimeter formed of layers of plastic scintillating tiles with embedded optical fibers interleaved with layers formed of brass. The latter was chosen for its high density ($8.96\ \mathrm{g \cdot dm^{-3}}$). While the processes involved in the development of an electromagnetic shower are essentially due to electromagnetic interactions, the hadronic shower is mainly caused by the strong interaction between incident hadrons and nuclei from the medium. A significant fraction of the initial hadron energy is transfered to secondary hadrons (essentially pions) produced with high relative transverse momentum, causing a large lateral extent of the shower. Similarly, the cross section for secondary hadrons creation is much reduced compared to electromagnetic processes such as Bremsstrahlung or pair creation. Hadronic shower development thus takes, in average, longer and the longitudinal extension is consequently increased. A radiation length equivalent definition quantity, called the interaction length, is defined for hadronic showers to quantify their longitudinal extension.

Hadrons, involved in jets, penetrating the HCAL cause the scintillating tiles to emit light with maximum emission wavelength between 410 nm and 425 nm. This light is then directed to wavelength shifter optical fibers to shift this maximum to 490 nm corresponding to the maximum absorption of the photomultipliers. This wavelength shifted light is then directed to these photomultipliers through optical fiber waveguides for light/analog signal conversion and transmission to readout electronics for further processing.

The HCAL layout is characterized by a cylindrical part called the barrel and two circular structures, the endcaps, located on each side of the barrel and joined hermetically to it (no gap exists between the HCAL barrel and endcaps as it is the case for the ECAL). In addition, the hadron outer (HO), placed after the solenoid, complements the barrel in the central region and the hadron forward (HF) extends the acceptance coverage in the forward region up to $|\eta| = 5.19$. Figure 3.9 gives a layout of the HCAL.

Figure 3.9: Layout of the HCAL.

The Barrel

The central region ($|\eta| < 1.78$) is covered by the barrel part with inner and outer radii of 1777 mm and 2876 mm respectively. It is made of two halves on the positive and negative side of the z axis, each half dividing into 18 identical wedges covering a ϕ region of 20 degrees. A total of 17 active scintillator tile layers interleaved with stainless steel and brass plates, representing 70000 tiles constitutes the HCAL barrel. Individual scintillator tiles have dimensions of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$, which approximately correspond to a 5×5 crystal matrix in the ECAL. The total depth of the HCAL barrel depends on η and is roughly 5.82 interaction lengths for a central pseudorapidity ($\eta = 0$). As a comparison, the ECAL depth is equivalent to 1.1 interaction length.

The endcaps

Two circular structures (Endcaps) complement the barrel in the forward pseudorapidity region ($1.3 < |\eta| < 3.0$). They also divide in 18 wedges covering a 20 degree ϕ region but contain however 19 scintillator tile layers, slightly more compared to the barrel. The tile

dimensions are similar to the barrel up to pseudorapidities of 1.74, after which the dimensions increse to values of $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$.

A crucial specificity of the endcaps relies on the longitudinal segmentation of the scintillator layers which makes it possible to have two energy estimates of the longitudinal profile of the hadronic shower in the endcaps.

3.5 The solenoid

The CMS solenoid is a supraconducting magnet which provides a 3.8 T magnetic field directed along the beam line (z axis). Charged particle trajectories are thus deviated in the transverse plane.

3.6 The muon system

The muon system, designed to identify muons, relies on three independent systems, referred to as muon chambers which will provide hits for the muon track reconstruction and the measurement of its momentum in the magnetic field.

The drift tube chambers (DT)

The barrel part of the muon system ($|\eta| < 1.2$), where a low neutron-induced background rate and a quite uniform magnetic field are expected, is made of drift tube chambers. They divide in four stations of 250 chambers each and also in five wheels of 12 sectors each, each sector covering a 30 ϕ region. Chambers gather in sets of four chambers called superlayers. The first three stations contain 2 superlayers which provide r—ϕ measurements of the muon coordinate and 1 superlayer in between that provides, in addition, a z measurement along the beam line. The last station provides only r—ϕ measurements.

The cathod strip chambers (CSC)

The forward region ($0.9 < |\eta| < 2.4$) is occupied by cathod strip chambers due to the high expected background rate. There are four cathod strip stations in each endcap made of chambers perpendicular to the beam axis, thus providing r—ϕ measurements of the muon coordinate. Chambers have trapezoidal shape and are made of 7 cathod strip planes leaving 6 gas gaps, the whole being enveloped in the anode wire plane; the wires are roughly perpendicular to the cathode strips. A muon traversing a plane causes gas ionisation and subsequent electron avalanche that creates a charge on the anode wire and an image charge on the strips.

The resistive plate chambers (RPC)

The resistive plate chambers are double gap chambers operating in avalanche mode and are used in muon signal triggering. They have fast response but coarser time resolution compared to DT and CSC chambers. The resistive plate chambers are present in the barrel and in the endcap regions, interleaved between the layers.

3.7 The trigger

The design performance of the LHC is to perform proton-proton collisions with a production rate of 40 MHz (a bunch crossing is expected every 25 ns). In addition, at the design instantaneous luminosity of 10^{34} $cm^{-2}s^{-1}$, approximately 20 proton-proton collisions are expected per bunch crossing. The amount of data produced every second is extremely high and needs drastic reduction. This reduction will be performed through selection criteria based on main kinematic variables such as transverse energy or transverse momentum. Other requirements such as the presence of a track segment in the pixel detector or in the muon chambers can also be present. Such a set of criteria, called a trigger, was designed to reduce the 40 MHz rate to the more affordable value of 100 Hz, selecting the most interesting events.

Moreover, the trigger system can also be used to separate between the different physics data streams, each stream being associated to one or more so-called trigger paths. As an example, the jet trigger requiring the presence of a jet with minimum transverse energy, will be used to preselect with high efficiency events possibly involving jets. Many trigger paths have thus been designed (jet trigger, muon trigger, electron trigger ...) for each type of object or type of event.

The trigger can be divided into two main categories: the level 1 trigger (L1) and the high level trigger (HLT).

3.7.1 The L1 level

The L1 trigger reduces the amount of data to an output rate of 100 KHz and mainly consists in programmable hardware electronics. It uses coarsely segmented data from the calorimeters and the muon system (the expected high track multiplicity represents too big information to be processed rapidly), while holding the data in pipelined memories in front-end electronics. The latency affected to the level 1 trigger decision is 3.2 μs during which a trigger decision has to be taken. Three kind of components can be identified for the level 1 trigger:

- the local trigger, uses information from energy deposits in the calorimeters (ECAL/HCAL) and track segments in the muon chambers. This information is transfered to the regional trigger.
- the regional trigger combines the information from each subdetector local trigger, applies a rank, based on the object parameter estimates (jet/electron energy, track quality), which reflects the level of confidence of these estimates and sorts the objects (electron, muon, track) accordingly. This information is transfered to the global trigger for decision.
- the global calorimeter and muon triggers determine the highest rank object in the whole detector and transfer it to the global trigger which takes the decision to accept or reject the event based on predefined selection requirements. Upon acceptation, the L1 decision is communicated to each subdetector and the data is transfered to the HLT level.

3.7.2 The HLT level

The goal of the high level trigger is to reduce the amount of data to a rate of 100 Hz and consists in a computer farm of roughly one thousand commercial processors. It has access to the complete read-out data and can perform more complex calculations compared to level 1 processing and similar to the ones made offline. Upon positive decision of the level 1 trigger, the data is transfered to the high level trigger which applies the predefined set of

HLT requirements (minimum transverse energy/transverse momentum, presence of a track segment in the pixel detector) to accept or reject the event. Upon acceptation, the data from the event is transfered for definitive storage and further processing/analysis by the physicists. The HLT algorithms depend strongly on the expected instantaneous luminosity.

3.7.3 The electron trigger

A dedicated trigger to select events with at least one or two electrons has been designed for analyses related to electrons. Such a trigger is naturally chosen for this study. Originally, it was tuned for low p_t electrons ($p_t < 40$ GeV/c), namely for the Higgs boson search in the channel $H \rightarrow ZZ^* \rightarrow 4e$, for a centre of mass energy of 14 TeV. Due to the lowering of the latter to the value of 7 TeV, new trigger tables have been defined for the electron which take into account the lower expected rate at 7 TeV.

The level 1 electron trigger

The present definition of the level 1 electron trigger contains essentially two trigger paths. The first one, "L1_SingleEG5" L1 trigger requires the presence of an energy deposit (cluster) in the ECAL with a minimum transverse energy of 5 GeV. This minimal value will be extended to 8 GeV ("L1_SingleEG8" L1 trigger) if the trigger rate is found to be too high during data-taking. Both of these L1 triggers will be used in the case of an instantaneous luminosity of 8×10^{29} cm^{-2}s^{-1}. For an instantaneous luminosity of 10^{31} cm^{-2}s^{-1}, the "L1_SingleEG8" L1 trigger will also be used.

The high level electron trigger

For the high level electron trigger, two trigger paths will be used for the present study, corresponding respectively to 8×10^{29} cm^{-2}s^{-1} and 10^{31} cm^{-2}s^{-1} instantaneous luminosities. The first one, the "HLT_Ele10_LW_LIR" trigger, requires in the ECAL an energy deposit with transverse energy higher than 10 GeV, linked to a pair of pixel hits which is compatible with the cluster transverse energy. In case the trigger rate is too high, the "HLT_Photon15_LIR" photon trigger will be used which requires the presence of a single cluster in the ECAL with transverse energy higher than 15 GeV and no other requirements.

For an instantaneous luminosity of 10^{31} cm^{-2}s^{-1}, the "HLT_Ele20_SW_LIR" electron trigger will be used similar to the "HLT_Ele10_LW_LIR" trigger mentioned above but with a minimum transverse energy of 20 GeV instead of 15 GeV. Several other triggers (electron and photon triggers) can be used for physics analyses or backup. Table 3.1 summarizes the triggers relevant for this study and gives their expected rate for a centre of mass energy of 7 TeV.

Trigger object	HLT path	E_t threshold	expected rate (Hz)	trigger type
Luminosity = 8 $\times 10^{29}$ cm^{-2} s^{-1}				
Electron	HLT_Ele10_LW_L1R	10 GeV	17.30 ± 0.32	physics
Electron	HLT_Ele15_LW_L1R	15 GeV	5.09 ± 0.18	backup
Photon	HLT_Photon15_L1R	15 GeV	10.60 ± 0.25	physics
Photon	HLT_Photon20_L1R	20 GeV	3.65 ± 0.15	backup
Luminosity = 10^{31} cm^{-2} s^{-1}				
Electron	HLT_Ele15_SW_LooseTrackIso_L1R	15 GeV	14.56 ± 1.10	physics
Electron	HLT_Ele20_SW_L1R	20 GeV	14.98 ± 1.12	physics
Photon	HLT_Photon25_L1R	25 GeV	20.50 ± 1.31	physics
Photon	HLT_Photon25_LooseEcalIso_TrackIso_L1R	25 GeV	4.27 ± 0.60	backup
Photon	HLT_DoublePhoton15_L1R	15 GeV	4.69 ± 0.63	physics

Table 3.1: Electron and photon HLT triggers designed for luminosities of 8 $\times 10^{29}$ cm^{-2} s^{-1} and 10^{31} cm^{-2} s^{-1}, respectively. The indication "physics" means that the trigger is intended to be used as a main physics trigger, whereas "backup" means that the trigger could be used in case of excessive rates.

Chapter 4

Drell-Yan production and backgrounds

The subject of this thesis is devoted to the study of the di-electron channel at the LHC with CMS data, with the focus on high e^+e^- invariant masses in order to search for possible deviations from the Standard Model, which could provide hints for new physics. It is thus essential to control well the Drell-Yan spectrum also at lower mass, where no new physics is expected.

This chapter presents the Drell-Yan production and simulation at the LHC (sections 4.1 and 4.2) and the kinematics related to Drell-Yan events (section 4.3). Finally section 4.4 gives an overview of the processes, called generically backgrounds, that can mimic Drell-Yan production, thus contributing to the final di-electron sample.

4.1 Drell-Yan production

At the LHC, Drell-Yan production proceeds through the annihilation of a quark coming from a proton from one beam and an antiquark coming from a proton from the other beam, that create a virtual photon or Z boson which can then decay into a lepton pair ($q\bar{q} \rightarrow \gamma/Z \rightarrow l^+l^-$). Figure 4.1 shows the two Feynman graphs that contribute, at leading order, to this process. The matrix elements for the partonic process are calculable in the Standard Model, applying Feynman rules to the previous graphs. In this chapter, the high virtuality region only will be considered ($M(\gamma^*/Z) > 40$ GeV/c^2).

As collisions at the LHC involve protons, the Drell-Yan cross section needs to be computed. In the reference frame where protons are moving fast ('Breit frame'), the quark and antiquark involved in the hard process carry fractions x_1 and x_2 of the proton momenta, respectively. The parton density functions (PDF) quantify the density of partons of type i in the proton, carrying a fraction x of the proton momentum, at a scale Q^2: $f_i(x, Q^2)$ (see section 2.4). The total cross section can be written as the convolution of the partonic cross section $\sigma(q\bar{q} \rightarrow \gamma/Z \rightarrow e^+e^-)$ and the corresponding parton density functions.

$$\frac{d^2\sigma}{dx_1 dx_2} = \frac{1}{3}\sigma(q\bar{q} \rightarrow \gamma/Z \rightarrow e^+e^-) \sum_{i=1}^{N_f} q_i^2 (f_i(x_1)\bar{f}_i(x_2) + f_i(x_2)\bar{f}_i(x_1)), \tag{4.1}$$

where N_f is the number of flavours and is here equal to 5, q_i is the charge for the quark of type i and $\bar{f}_i(x)$ represents the parton density function for the antiquark of type i.

In addition, next to leading order (NLO) diagrams need to be taken into account. In particular initial and final state radiation are possible. The initial state radiation is characterized

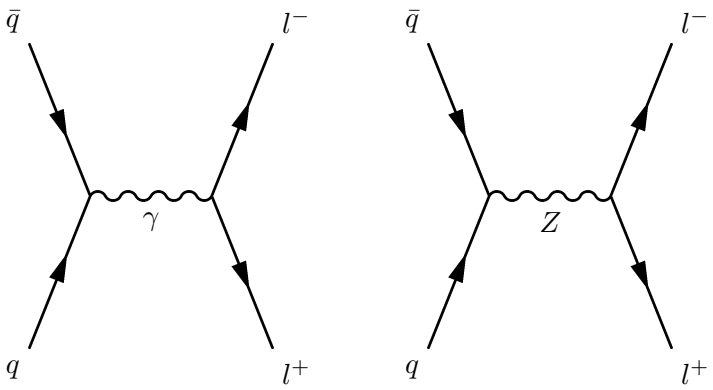


Figure 4.1: The Feynman graphs which contribute, at leading order, to the Drell-Yan process.

by the emission of gluons or quarks (QCD Bremsstrahlung). The emission of a hard (high transverse momentum) gluon or quark gives transverse momentum to the quark which participates in the hard interaction and which transfers it to the final state electrons. The final state radiation for final state electrons is due to the emission of photons (QED Bremsstrahlung) and is inversely proportional to the square of the photon transverse momentum relative to the electron.

4.2 Drell-Yan simulation

Monte Carlo simulations of physics processes have been extensively used during the detector construction and commissioning phase when no data is present, and also during data taking as a basis for the understanding of detector effects and the comparison of theory with data. The simulation tools rely on the statistical nature of physical processes, taking advantage of the computer random number generators. The simulation of Drell-Yan events in CMS is performed in the CMSSW (CMS SoftWare) software framework and can be divided into three main steps:

- the generation: This step simulates the fundamental physical processes as they are produced by Nature and is thus independent of any detector. On basis of the incoming particles four-momentum vectors, the cross section is computed based on matrix elements, and four momentum vectors of all particles produced during the collision are simulated as well as their production vertices.

  The fraction momenta x_1 and x_2 of the two partons participating in the hard event are chosen on the basis of parton density functions provided by the CTEQ or the MRST groups[1]. Initial and final state radiations are also simulated. The generated information is recorded, which is essential to give access to the momentum, pseudorapidity and azimuthal angle of the outgoing electrons before and after FSR. The radiation of extra gluons or quarks from partons is described in the parton showering framework. The showering process stops when the virtuality of the partons gets down to the QCD scale where perturbative QCD loses its predictive power and is followed by a non-perturbative fragmentation of partons, leading to colour-singlet clusters, which are eventually decayed, to form the final state hadrons. Different generators exist and

[1]In many simulation programs, the choice is left to the user as to the parametrization used. The default one used throughout this thesis for Drell-Yan production, unless otherwise stated, is CTEQ6L1

in all this chapter, the Pythia simulation program [34] will be used unless otherwise stated.

- the simulation: The outgoing particles are propagated in all layers of the CMS detector and their interaction with material is simulated using GEANT programme [35]. A very precise simulation of the detector geometry (subdetector position and composition, mechanical support, cables, cooling structures, etc...) is essential. For electrons from Drell-Yan events, hits in the tracker (pixel+strip) and energy deposits in the calorimeters are simulated.

- the reconstruction: Based on the simulation of the particle-matter interaction performed in the last step, the signals in front-end electronics are simulated. Tracks are built from tracker hits and superclusters are reconstructed from energy deposits in the ECAL (cf. section 6.1). The same reconstruction software as for simulation is used for data.

The analysis in this chapter is performed using a sample of 97000 Drell-Yan events with mass above 40 GeV/c^2, simulated with the Pythia simulation programme (version 6), using the CTEQ6L1 parametrization for the PDFs. Two additional samples at masses above 120 and 200 GeV/c^2, of 94000 and 99000 events, respectively, are used to study the high mass part of the spectrum. The total cross sections computed by Pythia at leading order are σ = 803 pb (M > 40 GeV/c^2), σ = 7.80 pb (M > 120 GeV/c^2) and σ = 0.96 pb (M > 200 GeV/c^2). A K factor of 1.28 is applied to account for QCD next-to-leading order effects. Only the generated information is used in this chapter unless otherwise stated.

Figure 4.2 shows the Drell-Yan invariant mass spectrum for mass above 40 GeV/c^2, at generator level and using the three Drell-Yan samples described above. The number of events is normalized to an integrated luminosity of 1 pb^{-1} and the K factos is applied. As one can see the cross section decreases steeply with the high invariant mass. The Z pole region is clearly visible.

4.3 Drell-Yan kinematics

The kinematics of the Drell-Yan process is important for this thesis. It is constrained by many parameters, such as the γ/Z invariant mass and the parton density functions. This section aims at reviewing some aspects of the Drell-Yan event kinematics. Section 4.3.1 reviews the parton density function contribution to the Drell-Yan process while section 4.3.2 describes some features of the Z boson momentum. The momenta of the electrons from Z decay are discussed in section 4.3.3 and the concept of acceptance is introduced and discussed in section 4.3.4.

4.3.1 Parton density functions

Drell-Yan production is due to the annihilation of a quark from a proton by an antiquark (necessarily from the sea) from the other proton. Through the relation:

$$M = \sqrt{x_1 x_2 s}, \tag{4.2}$$

where M is the mass of the produced resonance, $\sqrt{s}$ = 7 TeV the centre of mass energy and $x_{1,2}$ the momentum fractions of the two partons, one sees that for Drell-Yan production around the Z peak ($M \sim 91$ GeV), in the case where $x_1 \sim x_2$, the typical values for $x_{1,2}$ are of the order of 10^{-2}.

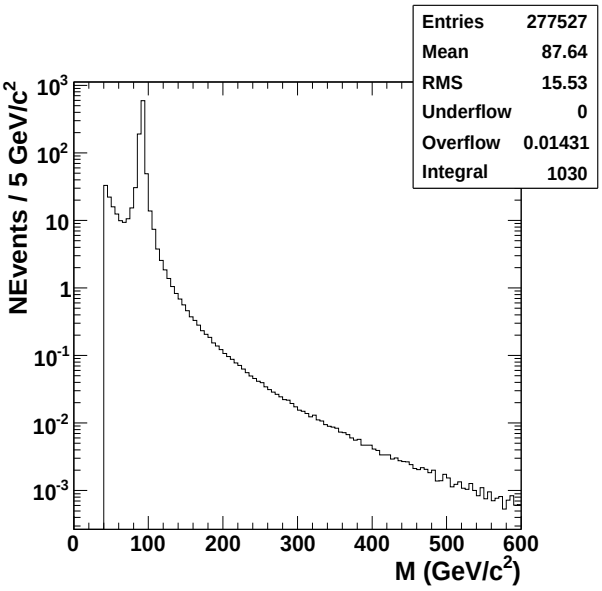


Figure 4.2: Drell-Yan invariant mass spectrum for $M > 40$ GeV/c^2 (generator level), normalised to an integrated luminosity of 1 pb^{-1}, simulated with Pythia, using parametrization CTEQ6L1 for the PDF.

Figure 4.3 presents the distribution of $xf(x)$ for different partons in the proton, at a scale $Q^2 = (91\text{ GeV})^2$, using the CTEQ6L1 parametrization. The gluon density dominates at low values of x ($x \lesssim 10^{-1}$). For values of $x_{1,2}$ around 10^{-2}, the second biggest contributions to parton density come from u and d quarks. A non negligible difference is to be noticed between quarks and antiquarks, the latter being necessarily sea quarks. This difference affects the kinematics of the resonance as it translates into a difference between the proton momentum fractions x_1 and x_2 of the two partons participating to the hard event. The quarks will have, in general, a larger momentum fraction than the antiquark, in particular in the region $x \sim 10^{-1}$ (high x region).

Figure 4.4 shows the particle identifiers of the two partons participating in the hard event for the quarks (black) and the antiquarks (red). The nomenclature of the particle identifiers can be found in [34] and is given in table 4.1 for the five quark types, together with their individual relative contribution. One sees that u and d quarks constitute the core of the partons in Drell-Yan events from pp collisions at 7 TeV at the LHC. The particle identifier for antiquarks is just the opposite of the one for quarks and their relative contribution is just the same, as expected.

4.3.2 Z boson momentum

The Z boson momentum[2] is the sum of the four-momenta of the quark and the antiquark participating in the hard interaction, and depends on the parton density functions.

The fraction momenta for both partons, x_1 and x_2 are often expressed in terms of the Z mass M and its rapidity:

[2] In the following, the γ^*/Z bosons will simply be called Z.

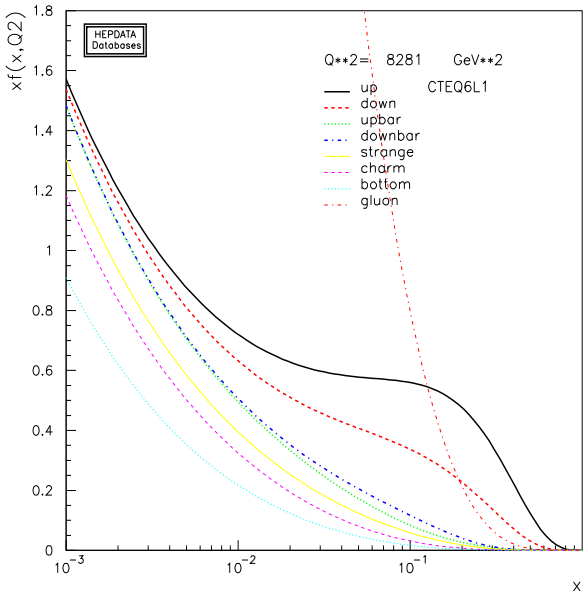


Figure 4.3: Parton density functions for partons in the proton, at a scale $Q^2 = (91\ \text{GeV})^2$ (CTEQ6L1 parametrization [36]).

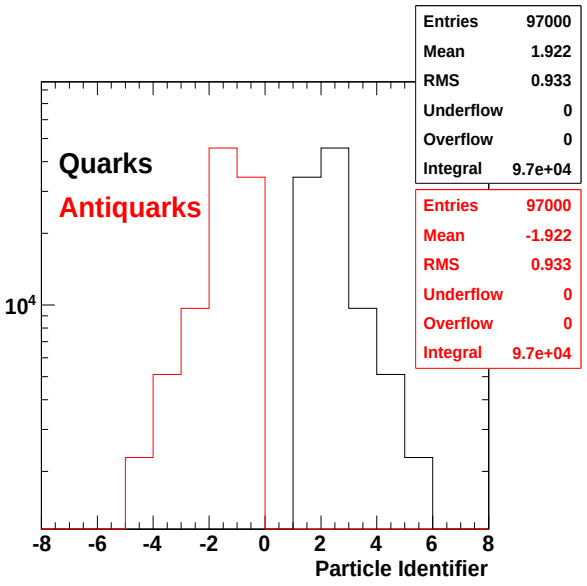


Figure 4.4: Particle identifiers of the quarks and antiquarks participating to the hard event for Drell-Yan events with masses above 40 GeV/c^2.

Quark	Particle identifier	Relative contribution
d	1	35.42 %
u	2	46.98 %
s	3	9.98 %
c	4	5.26 %
b	5	2.36 %

Table 4.1: Quark particle identifiers and their relative contributions in Drell-Yan events with masses above 40 GeV/c^2.

$$x_{1,2} = \frac{M}{\sqrt{s}} e^{\pm y}, \tag{4.3}$$

where the rapidity y is defined according to the Z energy and longitudinal momentum as

$$y = \frac{1}{2} \ln(\frac{E + p_z}{E - p_z}). \tag{4.4}$$

One sees that $M^2 = x_1 x_2 s$ as expected. Using the previous formulae, one can derive:

$$\frac{p_t^2 + M^2}{p_z^2} = \frac{4M^2}{s(x_1 - x_2)^2}. \tag{4.5}$$

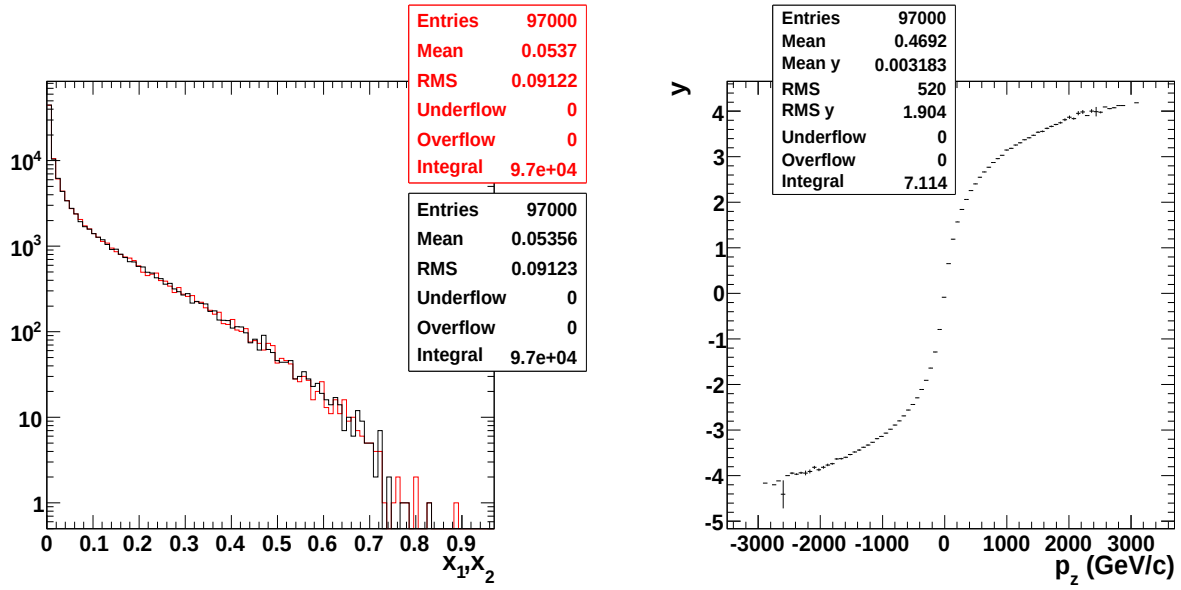
One sees that in case of no ISR, where no transverse momentum is expected, we get that

$$p_z = (x_1 - x_2)\sqrt{s}/2. \tag{4.6}$$

The momentum of the Z boson has thus two origins: the parton distributions x_1 and x_2 influence mainly the longitudinal momentum, and the ISR effects are at the origin of the transverse momentum. Both effects depend on the scale $Q^2 = M^2$. Figure 4.5(a) shows the distribution of x_1 and x_2, as defined in 4.3, for Drell-Yan samples generated at masses above 40 GeV/c^2. The two distributions are identical as expected. Figure 4.5(b) shows the rapidity y of the Z bosons as a function of the longitudinal momentum p_z. For $p_z = 0$, the rapidity is equal to 0 while for high longitudinal momenta, the rapidity increases.

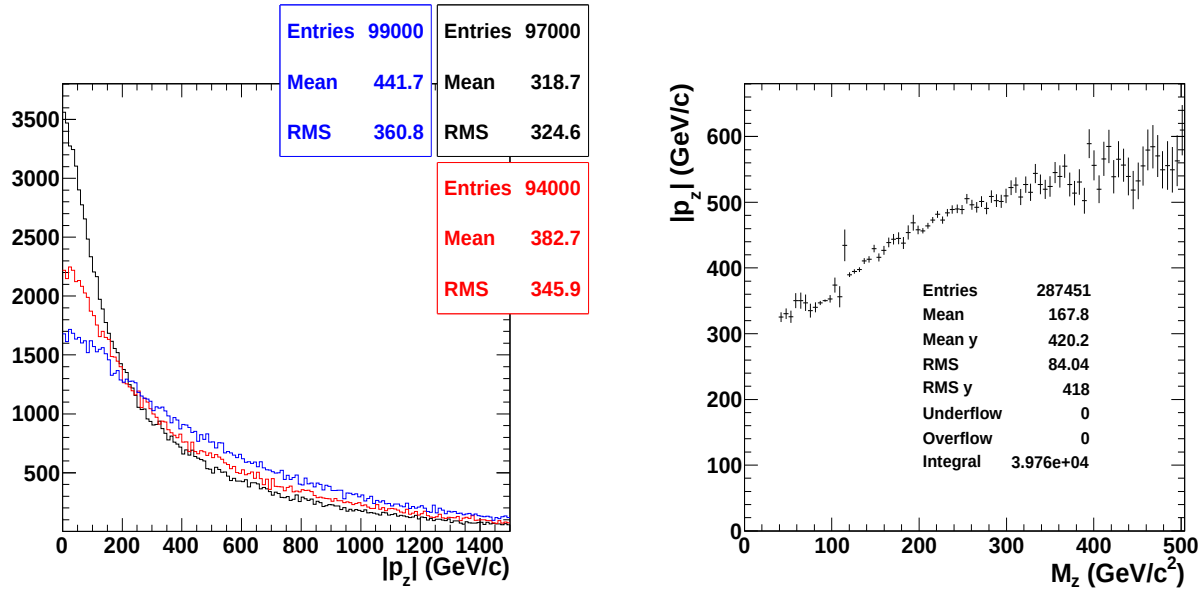
Figure 4.6(a) presents the distribution of the absolute value of the longitudinal momentum of the Z for the three Drell-Yan samples (M > 40, 120 and 200 GeV/c^2). The distributions peak at 0 and the mean value of the absolute longitudinal momentum increases with the Z mass. Indeed, if we consider $x_1 = x_2$, to produce a 50 GeV/c^2 invariant mass Z in a proton-proton collision at 7 TeV centre of mass energy, one needs $x_{1,2}$ around 7×10^{-3} while for a 200 GeV/c^2 invariant mass Z produced in the same conditions, these values are around 2×10^{-2}. The difference $x_1 - x_2$, proportional to the longitudinal momentum of the Z, is expected to be larger in the case of a 200 GeV/c^2 Z than for a 50 GeV/c^2 Z. Figure 4.6(b) presents the absolute value of the Z longitudinal momentum, p_z, as a function of its mass M_Z. The longitudinal momentum increases with mass.

Figure 4.7(a) presents the transverse momentum distribution of the Z for the three Drell-Yan samples generated at masses above 40, 120 and 200 GeV/c^2. The transverse momentum of the Z is expected to come only from initial state radiation effects and thus, in this case also, an increase as a function of the mass is expected as the ISR effects are more important at higher Q^2 values. The evolution of the transverse momentum of the Z as a function of its mass is presented in figure 4.7(b) for a combination of Drell-Yan samples generated at masses above 40, 120 and 200 GeV/c^2.



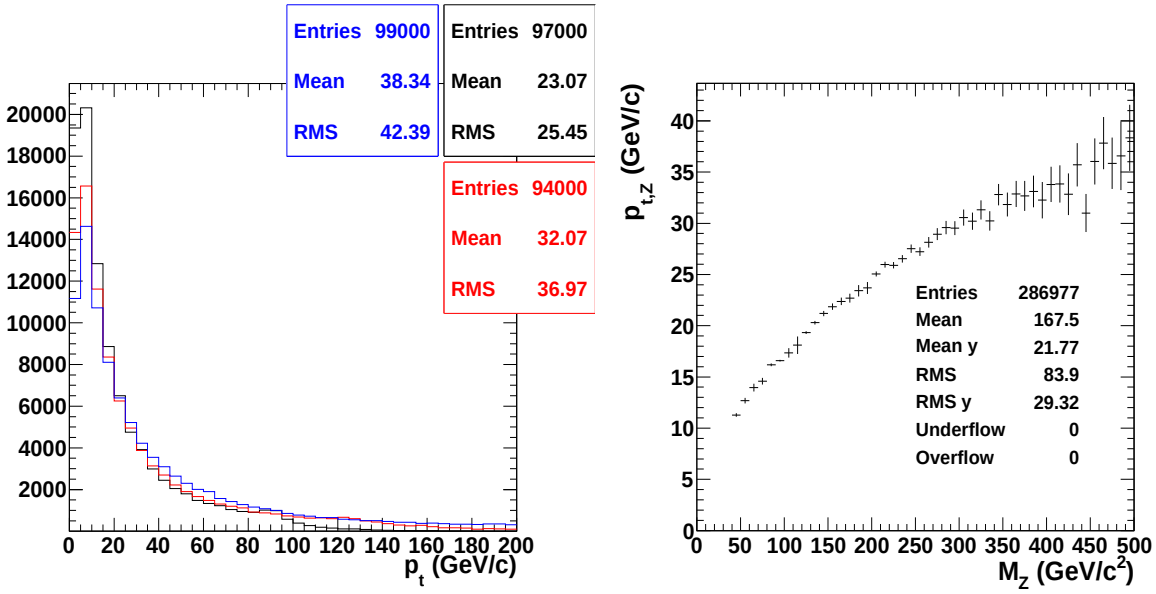
(a) (b)

Figure 4.5: (a) Distribution of the x_1 and x_2 variables (see text for definition) for Drell-Yan samples with masses above 40 GeV/c^2. (b) Rapidity y (see text for definition) as a function of the Z longitudinal momentum, p_z.



(a) (b)

Figure 4.6: (a) Distribution of the absolute value of the longitudinal momentum of the Z for three Drell-Yan samples generated at masses above 40 (black), 120 (red) and 200 (blue) GeV/c^2 and (b) average absolute value of the longitudinal momentum of the Z as a function of its mass M_Z.



(a) (b)

Figure 4.7: (a) Distribution of the transverse momentum of the Z for the three Drell-Yan samples generated at masses above 40 (black), 120 (red) and 200 (blue) GeV/c^2 and (b) transverse momentum of the Z as a function of its mass M_Z.

4.3.3 Momenta of electrons from Z decay

The results derived in section 4.3.2 are helpful to understand the distributions of the electron transverse and longitudinal momenta. Figures 4.8(a), (c) and (e) present the transverse momentum for electrons (black) and positrons (red) from the Z decay in Drell-Yan samples generated at masses above 40 GeV/c^2, 120 GeV/c^2 and 200 GeV/c^2, respectively. No significant difference is observed between electrons and positrons. The transverse momentum distribution for electrons and positrons in Drell-Yan events shows a typical shape. A peak at 45 GeV/c (60 GeV/c, 100 GeV/c), a long tail at low transverse momentum and a short tail at high transverse momentum are visible. These effects will be described in the following.

When the Z is produced at rest ($p_{t,Z} = p_{z,Z} = 0$), the centre of mass frame coincides with the laboratory frame. When the electron pair direction is in the transverse plane of the laboratory frame, the transverse momentum of the electrons is maximal and equal to half the Z mass. Such electrons populate the peaks at 45 GeV/c, 60 GeV/c and 100 GeV/c and also the tail at high transverse momenta.

If the electron pair direction has a component along the z axis, the transverse momentum of the electrons is smaller than half the mass. Such electrons populate the low p_t tail.

The case where the Z is boosted along the z axis ($p_z \neq 0$ and $p_t = 0$) does not change the previous conclusions as the boost along the z axis will conserve the transverse momenta of the two electrons from the decay.

The case where the transverse momentum from the Z is not equal to 0 modifies the transverse momentum of the electrons. If the electron (resp. positron) is emitted in the direction of the Z transverse momentum, its transverse momentum will be higher than half the Z mass. These electrons populate the high transverse momentum tail of the distribution.

Figures 4.8(a), (c) and (e) present the longitudinal momentum for electrons (black) and positrons (red) from the Z decay in Drell-Yan samples generated at masses above 40 GeV/c^2,

120 GeV/c^2 and 200 GeV/c^2, respectively. While no significant difference is observed between electrons and positrons for masses above 40 GeV/c^2, a difference is present for masses above 120 GeV/c^2 and increases for masses above 200 GeV/c^2. This difference between electrons and positrons comes from the (V-A) nature of the coupling of the Z to the fermions. At leading order and at high mass (M $\gtrsim$ 100 GeV/c^2), in the $q\bar{q}$ centre of mass, the e^+e^- pairs are produced back to back, with the electron emitted preferentially in the forward direction (defined as the one from the quark). To go in the laboratory frame, the electron and positron are boosted along the z axis, the z direction of the boost is given by the longitudinal momentum of the Z boson, which is usually the direction of the quark at high mass (M $\gtrsim$ 100 GeV/c^2), and the difference between electrons and positrons is enhanced after the boost. The mean value of the absolute longitudinal momentum at high mass (M $\gtrsim$ 100 GeV/c^2) is thus larger for electrons than for positrons.

As the pseudorapidity of the electrons is related to their transverse and longitudinal momenta ($\eta = -\ln(\tan(\theta/2)) \sim \ln(p_z/p_t)$), a difference in pseudorapidity distributions is also expected between electrons and positrons at high mass. The pseudorapidity distribution for electrons and positrons in Drell-Yan events at masses above 40, 120 and 200 GeV/c^2 is presented in figures 4.9(a), (b) and (c), respectively. No significant difference is observed at masses above 40 GeV/c^2 while a difference is present at masses above 120 GeV/c^2 which increases at masses above 200 GeV/c^2.

At leading order, no transverse momentum is expected for the Z and the two electrons from the Z decay are back to back in the transverse plane, i.e. the difference between their azimuthal angles, $\Delta\phi = \phi_{e^+} - \phi_{e^-}$, expressed in radians, is expected to be around π. This is checked in figures 4.10(a)-(c) which show the quantity $|\Delta\phi|$ for the three Drell-Yan samples with masses above 40, 120 and 200 GeV/c^2. The distributions clearly peak at π. In case of initial state radiation, the two electrons are not expected to be back to back in the transverse plane anymore as the initial state radiation gives transverse momentum to the Z boson which transmits it, when decaying, to the two electrons. Figure 4.10(d) presents the quantity $|\Delta\phi|$ as a function of the transverse momentum of the Z boson. The quantity $|\Delta\phi|$ decreases as the transverse momentum of the Z increases. A further illustration is provided in figure 4.11 which presents an event display for a simulated Drell-Yan event (M = 200.85 GeV/c^2) with initial state radiation. The central window gives a transverse view of the detector where one can see, highlighted in red, the two generated electrons. The two electrons are not back to back and two additional objects displayed as yellow triangles on the upper part of the central view are visible. They correspond to jets initiated from ISR radiation.

4.3.4 Acceptance

The two main subsystems for electron detection, the electromagnetic calorimeter and the tracker, have coverage up to pseudorapidities of 3.0 and 2.5, respectively. The proportion of events where both electrons lie inside the tracker detection coverage region ($|\eta| < 2.5$), is called the acceptance. It is an important correction factor for the estimation of the Drell-Yan cross section, and this factor is computed using simulation.

The acceptance factor is observed to increase with the Z boson mass. The ratio of the mean transverse momentum over the mean absolute longitudinal momentum is presented as a function of the Z mass in figure 4.12(a) for electrons (black) and positrons (red). As $\eta = -\ln(\tan(\theta/2)) \sim \ln(p_z/p_t)$, the fraction of Drell-Yan events with both electrons in the acceptance ($|\eta| < 2.5$) increases when the mass of the Z boson increases. This effect has already been shown in figures 4.9.

Note that the results of section 4.3.2 show that the Z boson longitudinal and transverse

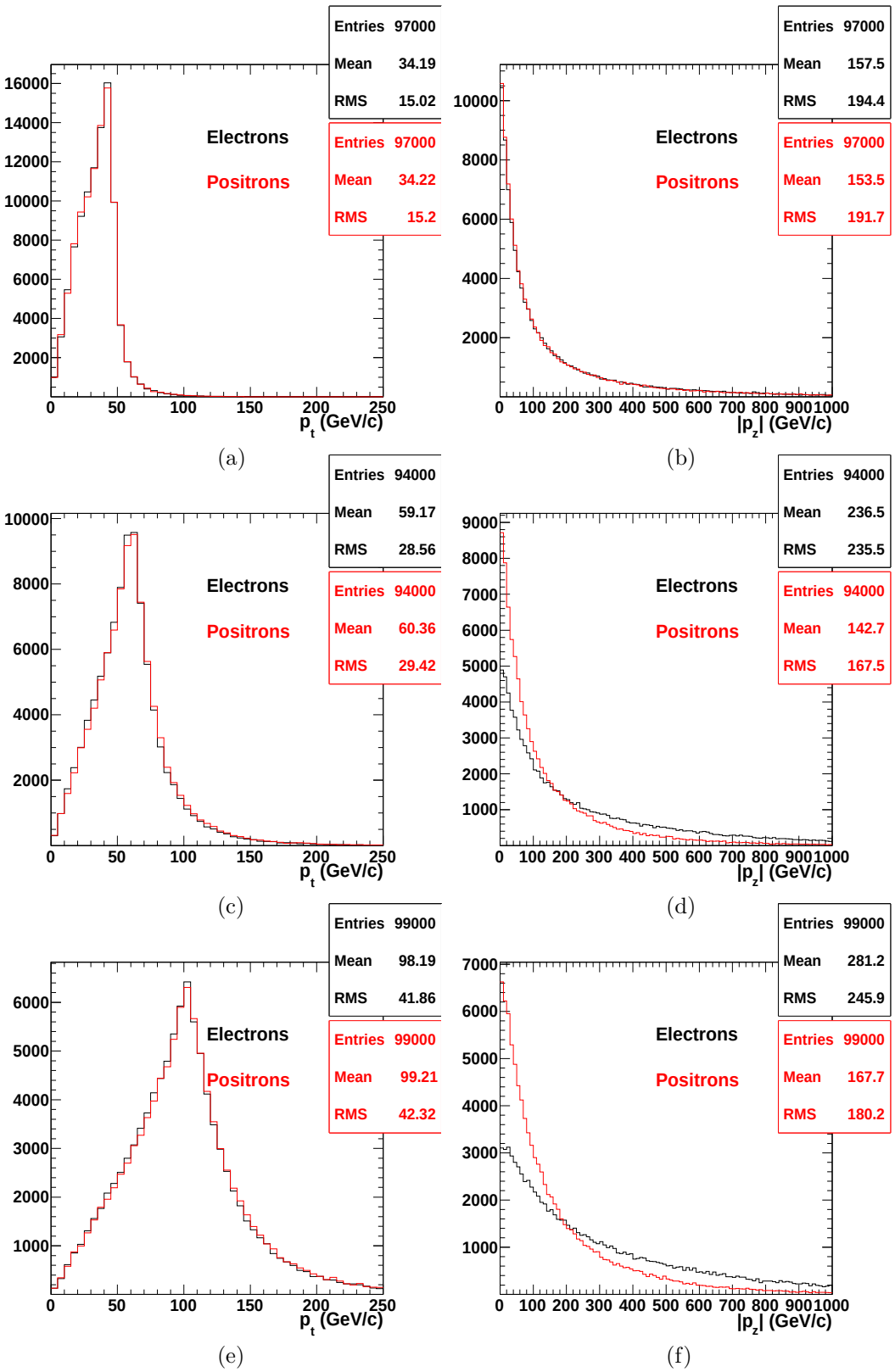
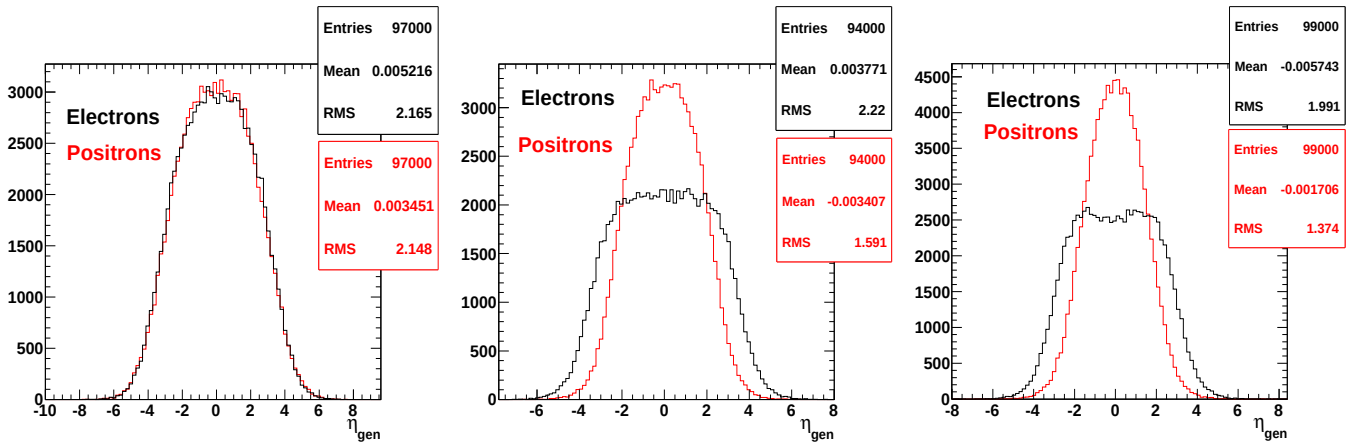


Figure 4.8: Electrons transverse momentum distribution (left) and longitudinal momentum distribution (right) for electrons (black) and positrons (red) from the Z decay in a Drell-Yan sample generated at masses above 40 GeV/c^2(a-b), 120 GeV/c^2 (c-d) and 200 GeV/c^2(e-f).



(a) (b) (c)

Figure 4.9: Pseudorapidity η of electrons (black) and positrons (red) for Drell-Yan samples generated at masses above 40 GeV/c^2(a), 120 GeV/c^2(b) and 200 GeV/c^2(c).

momenta, both increase with respect to the mass of the Z. Figure 4.12(b) presents, for each bin in invariant mass, the ratio of the mean transverse momentum over the mean absolute longitudinal momentum for the Z boson. It is observed that, on average, the Z boson transverse momentum increases more rapidly with mass than its longitudinal momentum.

In this section, the acceptance is defined as the number of events where both the electron and the positron have transverse energy higher than 25 GeV($E_t > 25$ GeV), are emitted outside the ECAL barrel and endcap gap region ($|\eta| \notin [1.442,1.560]$) and in the tracker detection region ($|\eta| < 2.5$), over the total number of events. Figure 4.13 presents the acceptance as a function ofs mass for events where both electrons are emitted in the barrel (blue open circles), in the endcaps (pink downward triangles), one in the endcap and the other in the barrel (red upward triangles) and finally for all events (black full circles). One sees that the total acceptance increases as a function of mass. Table 4.2 gives the total acceptance for Drell-Yan samples generated at masses above 40, 120 and 200 GeV/c^2; the errors, from the Monte Carlo statistics, are of the order of 10^{-3}.

Drell-Yan sample	$M > 40$ GeV/c^2	$M > 120$ GeV/c^2	$M > 200$ GeV/c^2
	34.8%	51.9%	63.7%

Table 4.2: Total acceptance for Drell-Yan events with masses above 40, 120 and 200 GeV/c^2, with both electrons with transverse energy above 25 GeV, emitted outside the ECAL barrel-endcap gap region and in the tracker detection region. Errors from Monte Carlo statistics are of the order of 10^{-3}.

4.4 Background processes

Several Standard Model processes can mimic the final state signal and are then selected in the final di-electron mass spectrum. Their contribution can be minimised by requiring some appropriate conditions (reducible backgrounds). These reducible background processes can be classified into two main categories:

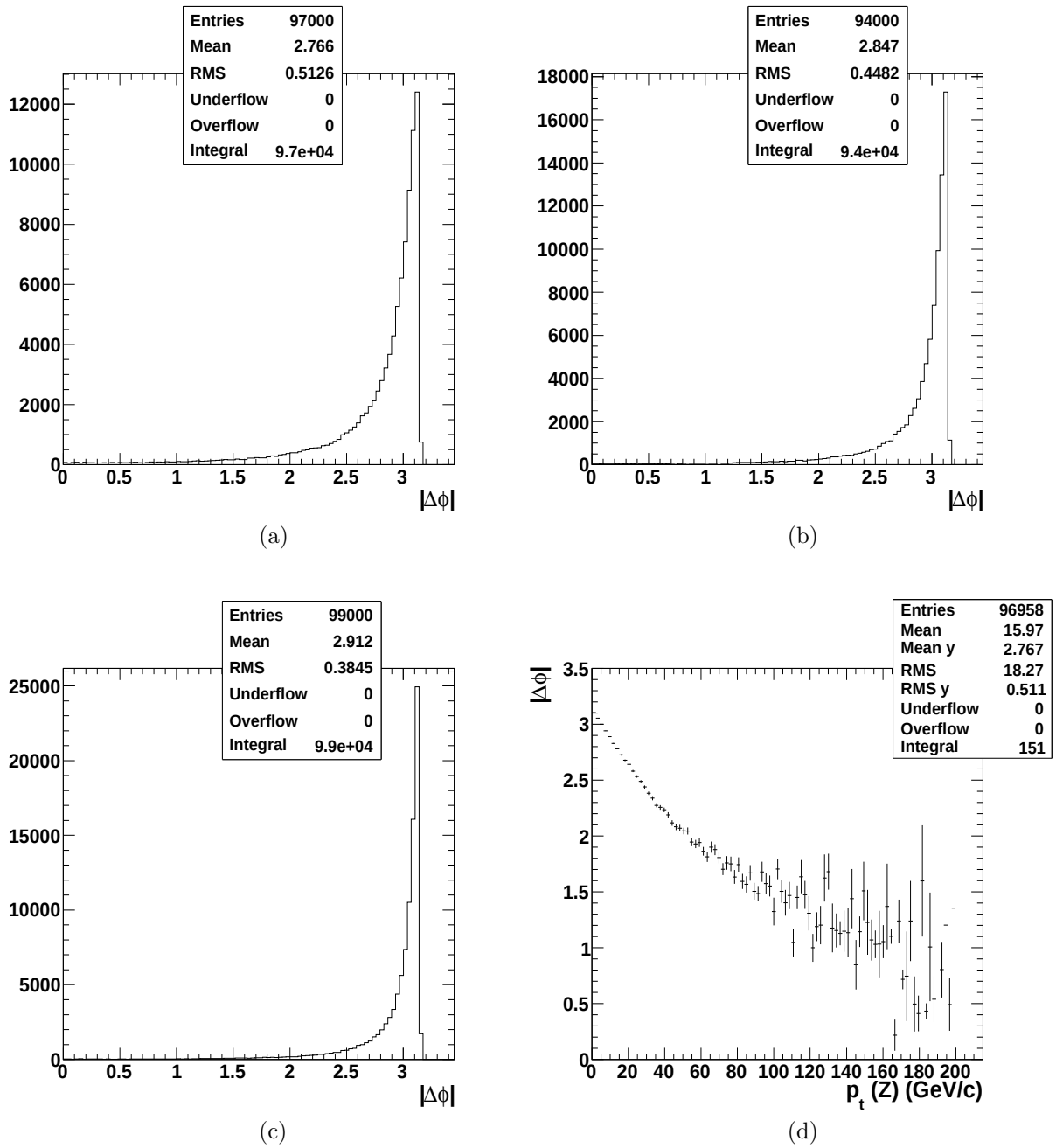


Figure 4.10: Absolute azimuthal angle difference $|\Delta\phi|$ (see text for definition) of the two electrons from the Z decay for Drell-Yan samples generated at masses above 40 GeV/c^2(a), 120 GeV/c^2(b) and 200 GeV/c^2(c). (d) Same difference $|\Delta\phi|$ as a function of the transverse momentum of the Z boson for M $>$ 40 GeV/c^2.

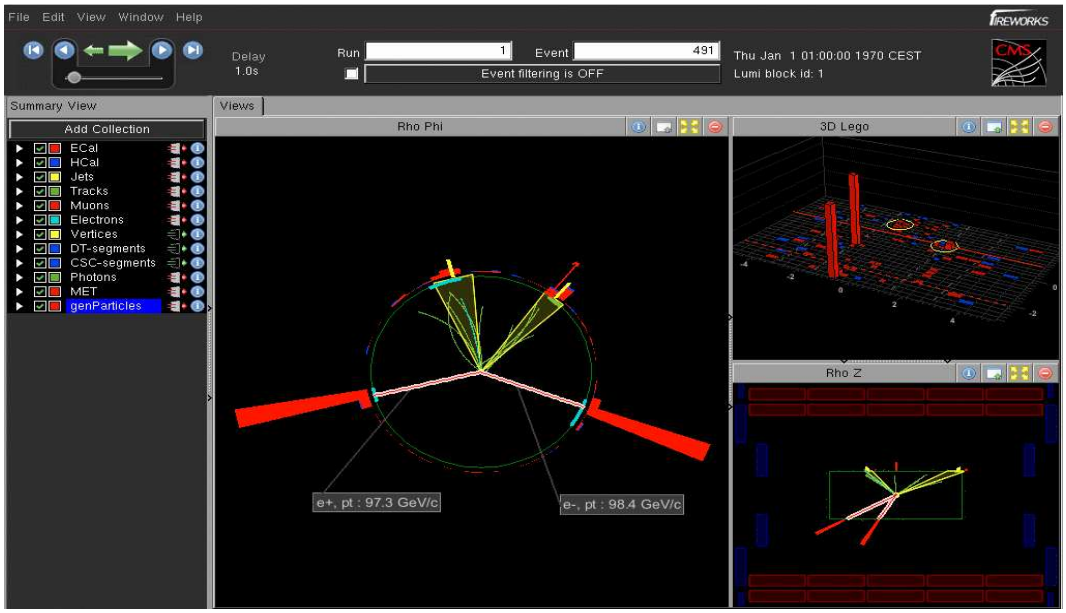
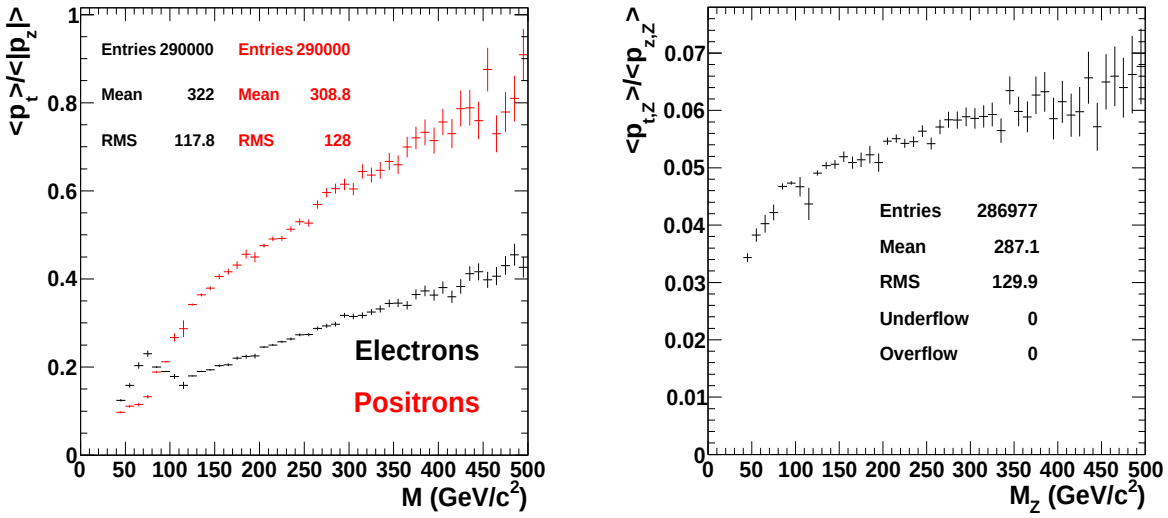


Figure 4.11: Event display of a simulated Drell-Yan event with strong initial state radiation from a Drell-Yan simulation sample with mass above 200 GeV/c^2.



(a) (b)

Figure 4.12: (a) Ratio of mean transverse momentum of the electrons (black) and positrons (red) over their mean absolute longitudinal momentum as a function of the Z mass for Drell-Yan samples generated at masses above 40, 120 and 200 GeV/c^2. (b) Mean value of the ratio of the transverse momentum of the Z over its absolute longitudinal momentum as a function of the Z mass for Drell-Yan samples generated at masses above 40, 120 and 200 GeV/c^2..

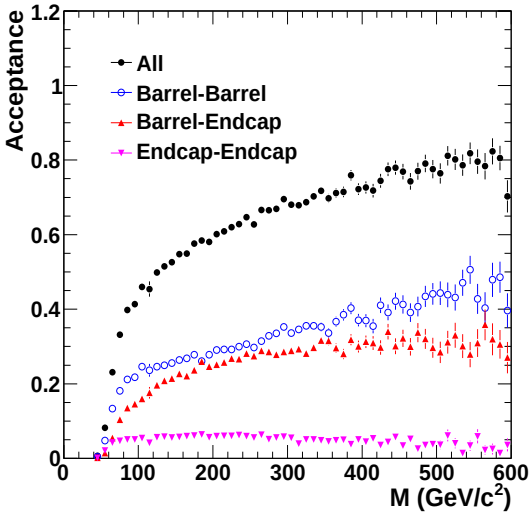


Figure 4.13: Acceptance as a function of the Z mass for Drell-Yan events with M $>$ 40 GeV/c^2, with both electrons with transverse energy above 25 GeV, emitted outside the ECAL barrel-endcap gap region and in the tracker detection region.

- the 'jet background' where at least one jet fake an electron. This background category contains the QCD multijet, W+jet and γ+jet background processes.
- the 'di-electron background' where two electrons are found in the final state. This category contains the $t\bar{t}$, tW, WW, $Z \rightarrow \tau\tau$ and $Z \rightarrow \mu\mu$ background processes.

In addition, the $\gamma\gamma$ background arises when the two photons fake electrons.

In the following, these processes are introduced and their cross sections are provided at leading order (LO) or next to leading order (NLO), considering a 7 TeV centre of mass energy. The ratio between the cross section at next to leading order and the cross section at leading order is called the K factor:

$$K = \frac{\sigma_{NLO}}{\sigma_{LO}} \tag{4.7}$$

4.4.1 Jet background processes

The QCD multijet background

The main background comes from QCD multijet production where two jets each fake an electron. This process was simulated by Pythia in several $\hat{p}_t$ bins where $\hat{p}_t$ is the transverse momentum of the outgoing partons in the centre of mass frame. The cross sections for the various $\hat{p}_t$ bins generated by Pythia are summarized in table 4.3 where no K factor is applied.

The W+jet background process

The W+jet process is another source of background where the W decays into a $e\nu_e$ pair and the jet fakes an electron. It proceeds through quark-gluon fusion or quark-quark fusion.

QCD $\hat{p}_t$ bin	Cross section (pb)
$\hat{p}_t > 15$ GeV/c	8.762×10^8
$\hat{p}_t > 30$ GeV/c	6.041×10^7
$\hat{p}_t > 80$ GeV/c	9.238×10^5
$\hat{p}_t > 170$ GeV/c	2.547×10^4

Table 4.3: Cross sections for the various QCD multijet $\hat{p}_t$ bins as simulated by Pythia6.

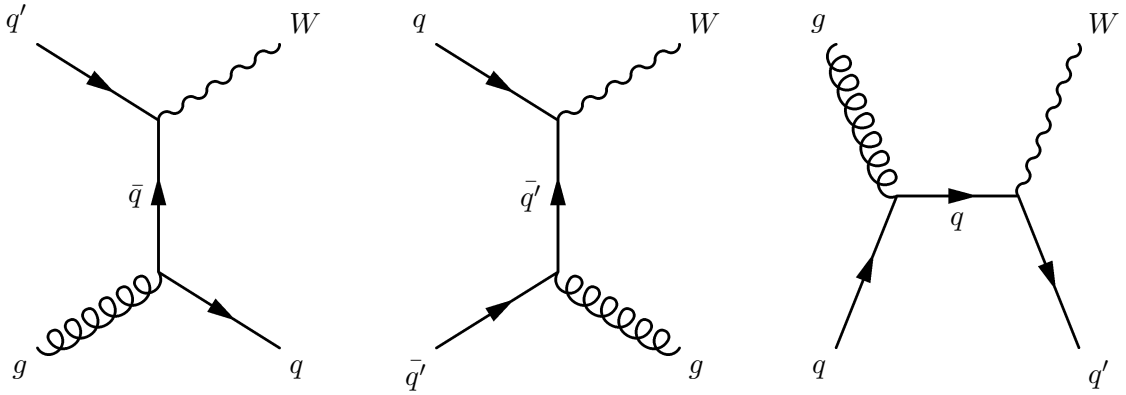


Figure 4.14: The Feynman graphs which contribute, at leading order, to the W+jet process.

Figure 4.14 shows the graphs which contribute, at tree level, to this process. The W+jet process was generated with MadGraph, with a K factor of 1.30, giving a cross section of 3.13×10^4 pb.

The γ+jet background process

The γ+jet process where the photon is misreconstructed as an electron and the jet fakes an electron constitutes another source of background. It proceeds through similar diagrams as for the W+jet process. The γ+jet process was generated with MadGraph.

4.4.2 The di-electron background processes

The $t\bar{t}$ background process

The $t\bar{t}$ pair production process is an important background for the channel under study. Figure 4.15 shows the Feynman graphs which contribute, at leading order, to the $t\bar{t}$ pair production. The process has been generated at leading order with the MadGraph simulation program [37]. The cross section has been computed at next-to-leading order (NNL0) in [38] and the found value is 157.5 pb. A K factor of 1.66 has been applied to the MadGraph cross section in order to account for higher order contributions.

The single top tW background process

Single top in association with W boson, tW is produced at the LHC through gluon-b quark fusion where the b quark originates from the proton sea, as shown in figure 4.16. The W can decay into a $e\nu_e$ pair giving one real electron in the final state. The top quark decays

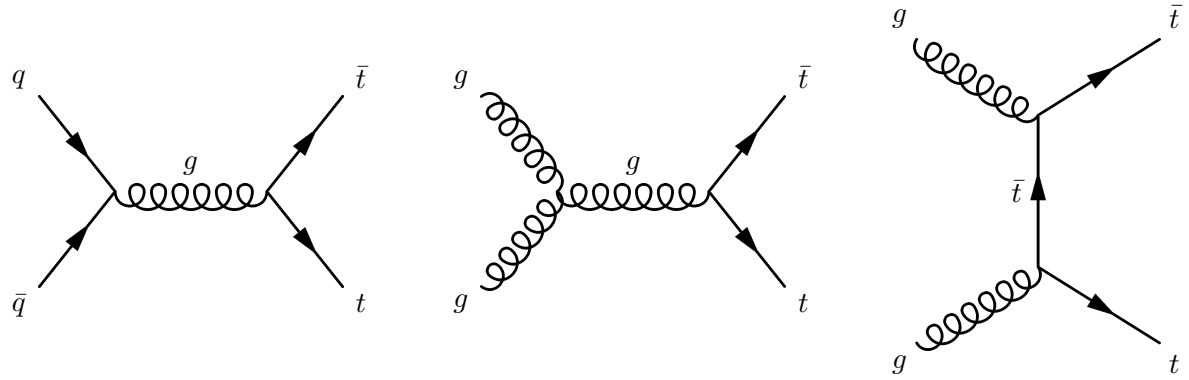


Figure 4.15: The Feynman graphs which contribute, at leading order, to the $t\bar{t}$ process.

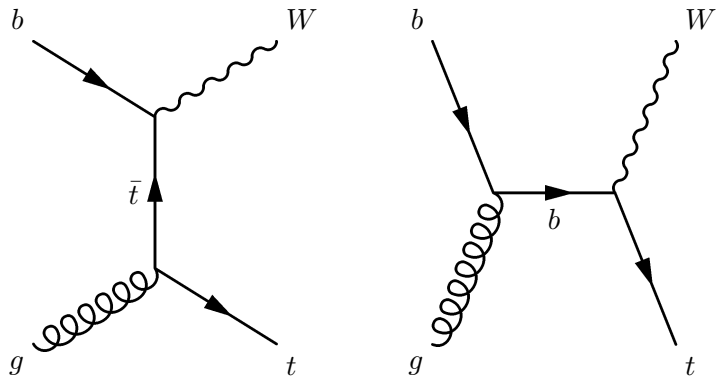


Figure 4.16: The Feynman graphs which contribute, at leading order, to the single top tW process.

into a bW system ($\mathrm{BR}(t \to bW) \sim 100\%$) where the W can also decay into a $e\nu_e$ pair. This background process is thus characterised by the presence of two real electrons in the final state. The tW process was generated with MC@NLO [39]; where no K factor is applied, the cross section being 0.1853 pb.

The WW background process

The WW process with the two W decaying into a $e\nu_e$ pair is a possible background. The W pair system can be produced from quarks or gluons in the initial state through $q\bar{q}$ annihilation or gluon fusion, as shown in figure 4.17. This process gives a contribution to the total background since it gives two real electrons in the final state.

The WW process was generated with Pythia. The NLO cross section is 2.93 pb, where no K factor was applied.

4.4.3 The $\gamma\gamma$ background process

The $\gamma\gamma$ process, where the two photons are misreconstructed as electrons, is another possible background source. The corresponding diagrams are similar to those for WW production. The $\gamma\gamma$ process was generated with Pythia and no K factor was applied.

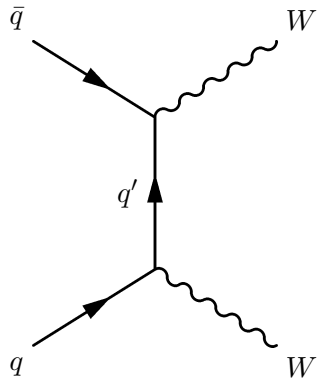


Figure 4.17: The Feynman graph which contributes, at leading order, to the WW pair production process.

The list of Monte Carlo background samples is summarised in table 4.4 together with the corresponding cross sections and the K factors, considering a 7 TeV centre of mass energy.

Sample	Generator	Kinematic cuts (in GeV, c=1)	σ (pb)	K factor
QCD	Pythia6	$\hat{p_t} > 15$	8.762×10^8	/
QCD	Pythia6	$\hat{p_t} > 30$	6.041×10^7	/
QCD	Pythia6	$\hat{p_t} > 80$	9.238×10^5	/
QCD	Pythia6	$\hat{p_t} > 170$	2.547×10^4	/
W+Jets	MadGraph	inclusive	24170	1.30
$t\bar{t}$	MadGraph	inclusive	95	1.66
tW	MC@NLO		0.1853	/
WW	Pythia6		2.93	/

Table 4.4: List of Monte Carlo samples for the background processes. The cross sections and the applied K factors are provided, considering a 7 TeV centre of mass energy.

Chapter 5

Analysis strategy for new physics searches

The Standard Model processes contributing to the di-electron samples, mainly the Drell-Yan process and several backgrounds, have been presented in chapter 4. In this chapter, the strategy for the search for new physics on top of the Standard Model processes is presented.

This strategy has been developed by the HEEP (High Energy Electron Pair) group at CERN. During the first years of data taking, characterized by a limited luminosity and a detector not perfectly understood, the focus will be brought on a clean signal,s as for example heavy resonances clearly identified as a peak in the e^+e^- invariant mass spectrum.

The main information and results presented in this chapter are derived from a study performed using Monte Carlo simulation samples and assuming a centre of mass energy of 10 TeV.

First full study by the HEEP group has led to a CMS publication [40] using simulation samples assuming $\sqrt{s} = 14$ TeV and an integrated luminosity of 100 pb^{-1}.

The analysis has been updated for the 10 TeV case with an improved statistical analysis of the signal significance and exclusion limits and a better understanding of the background contributions. This work led to a CMS publication [41] and its corresponding internal note [42].

This chapter presents the HEEP analysis strategy and the first results that came out of the study at 10 TeV and is organised as follows. The event selection is described in section 5.1, the discovery and control regions are explained in section 5.2. Data driven methods are introduced, to estimate selection efficiencies (section 5.3) and background contributions to the e^+e^- invariant mass spectrum (section 5.4). The final e^+e^- invariant mass spectrum is derived in section 5.5 together with a measurement of the Drell-Yan cross section, and the discovery potential is given in section 5.6. Finally, a scaling of these results to a centre of mass energy of 7 TeV is introduced in section 5.7.

It is stressed again that all results presented in this chapter are derived from Monte-Carlo simulation (for data analysis, see chapters starting from chapter 7).

5.1 Strategy: event selection

5.1.1 The baseline event selection

The main interest is to test the Standard Model Drell-Yan process at high mass and search for possible new resonant structures, in the di-electron channel, through the following process $pp \to (\gamma/Z/Z')XY \to e^+e^-XY$ where XY are particles involved in the underlying event (see section 2.4 for a definition of the underlying event).

These processes are therefore characterized by the presence of two high energy electrons and of the particles involved in the underlying event. In the CMS detector, the two electrons will leave significant energy deposits in the ECAL. Thus, the final topology we are interested in requires the presence of two high energy deposits in the ECAL.

In addition, background processes can contribute and alter the final di-electron mass spectrum as detailed in section 4.4.

The consideration of all those arguments has lead to the design of a specific set of criteria (high p_t electron selection) which aims at separating as much as possible the process involving two high energy, isolated electrons in the final state from the reducible background processes. The detail of the high p_t electron selection is given in chapter 6. The criteria can be divided into three main categories.

- the track criteria: to distinguish between electron energy deposits and photon energy deposits in the ECAL, the presence of a track in the tracker, associated in η and ϕ to the energy deposit, is required; the photon being a neutral particle leaves no signal in the tracker.

- the shower shape criteria: showers initiated in the ECAL by jets involved in the different background processes are different from showers initiated by electromagnetic objects (electrons or photons). These criteria mainly ensure that the showers in the ECAL behave as expected for electrons.

- the isolation criteria: jets are objects that contain an important number of particles, thus giving much activity in the tracker and/or in the ECAL/HCAL. These criteria require that no activity in terms of tracks or energy deposits in the ECAL/HCAL should be found around the electron.

The selection performance for events with a high energy electron pair is described by the efficiency. Quantitatively, it is defined as the ratio of the number of selected signal events over the number of produced signal events. Similarly, the rejection power is defined as the ratio of the number of rejected background events over the total number of produced background events. Another important figure is the purity, defined as the ratio of the number of signal events over the total number of events in the final selected sample. The goal of the HEEP criteria is thus to ensure a very high efficiency while minimising the background contribution as much as possible.

5.2 Discovery and control regions

The HEEP strategy consists in exploiting the different regions of the di-electron mass spectrum to: search for a new heavy resonance, extract essential information such as selection efficiency and background rates, and finally check the detector response and the reconstruction procedures. Consequently, three regions are defined in the di-electron invariant mass spectrum:

- The Z peak region: this region provides a large amount of data with very little background. It will be used to extract efficiency information from data using the "tag-and-probe" method described in section 5.3 and in [43]. This region will also be used for normalization, the computed differential cross sections will be normalized to the cross section at the Z peak. This region contains however essentially low E_t electrons.

- The high mass region (120 GeV/c^2< M < 600 GeV/c^2): this region will be used as a control region. The Drell-Yan differential cross section will be computed and compared to the Standard Model theoretical expectation, the recent results from the Tevatron having shown no sign of new physics in this region. The consistency will prove that the detector response and the reconstruction procedures are under control.

- The discovery region (M > 600 GeV/c^2): this region will be devoted to the direct search for new heavy resonances decaying into an electron pair.

To summarize, the key aspects of the proposed strategy are the search for new heavy resonant structures in the di-electron channel and the determination of the Drell-Yan differential cross section in the mass region 120 GeV/c^2< M < 600 GeV/c^2, normalized to the Z peak cross section.

It is worth noting, however, that the background contributions are expected to be different, depending on the considered invariant mass region. Indeed, in the discovery region, the main expected background is the Drell-Yan irreducible background. In the high mass region, however, where the Drell-Yan differential cross section will be computed, the latter is not considered anymore as a background process but as a "signal" process. The different background contributions to this "signal" in the high mass region are all the reducible backgrounds mentioned in chapter 4.

In a first part, the methods to extract, directly from data, the selection efficiency of di-electron events, will be presented. In a second part, the methods to estimate from data the different backgrounds will be described. The third part will present the final di-electron mass spectrum and the procedure to compute the Drell-Yan cross section. Eventually, the fourth part introduces the procedure devoted to the search of resonant structures in the di-electron channel. The significance estimation will be presented as well as the procedure to put upper limits in case no signal is observed.

5.3 Efficiency extraction from data

The efficiencies are measured directly from data using the "tag-and-probe" method described in [43]. This procedure selects events with a pair of electron candidates. Electron candidates are objects which, in addition to satisfying the electron trigger (section 3.7.3), have fulfilled a preliminary set of additional loose criteria (section 6.2). Stringent selection criteria are applied on one of the electron candidates, the "tag"; the other candidate, the "probe", has to pass looser selection criteria. Since electrons are produced in pairs in the samples relevant for this study, this procedure ensures a high purity electron sample. The efficiency ϵ for an electron to pass a given set of cuts is then computed as the fraction of probes which satisfy this set of cuts. In this analysis, the "tag" is required to pass the full selection criteria while the "probe" is required to be an electron candidate.

The electron reconstruction and identification efficiencies are factorized as the product of two contributions, corresponding to the main steps of the selection: $\epsilon = \epsilon_{cand} \times \epsilon_{sel}$, where ϵ_{cand} is the efficiency for an ECAL cluster to become an electron candidate while ϵ_{sel} is the efficiency for an electron candidate to pass the full selection criteria.

Two regions in the electron pair mass spectrum have been defined in order to extract the efficiencies: the Z peak and the high mass region (200 GeV/c^2< M < 600 GeV/c^2).

5.3.1 Efficiency measurement at the Z peak

The Z peak provides a region with abundant data and very little background. The tag-and-probe pair invariant mass is required to be in the range 81 GeV/c^2< M_{ee} < 101 GeV/c^2. This ensures a high purity electron sample. Note that if both electron candidates pass the "tag" cuts, the pair is considered twice.

Table 5.1 gives the efficiencies of the HEEP selection criteria applied on electron candidates from Drell-Yan events and the corresponding rejection powers for jet background events containing an electron candidate, separately in the barrel and in the endcaps.

	Barrel		Endcaps	
	$Z \to e^+e^-$	jet bg	$Z \to e^+e^-$	jet bg
efficiencies	$93.89 \pm 0.13\%$	–	$94.29 \pm 0.23\%$	–
rejection powers	–	$99.82 \pm 0.13\%$	–	$98.38 \pm 0.44\%$

Table 5.1: Efficiencies of the high p_t electron selection (ϵ_{sel}) criteria, applied to electron candidates, for Drell-Yan events, and corresponding rejection powers for jet background events containing an electron candidate, from simulated samples at $\sqrt{s}$ = 10 TeV, supposing an integrated luminosity of 100 pb^{-1} [42].

The statistical error comes from the efficiency measurement using data for 100 pb^{-1}, as given in table 5.1, it is small compared to the systematic error. For the systematic error, a conservative 50% uncertainty is applied to the estimated number of background events. The uncertainties on the scale factor are ±0.1% (stat.) ±1.7% (syst.) for the barrel and ±0.2% (stat.) ±3.8% (syst.) for the endcap regions. The reliability of the Monte Carlo simulation is estimated quantitatively by comparing the efficiencies obtained by the "tag-and-probe" method, when applied both to data and to Monte Carlo samples. The ratio of these efficiencies is defined as the "scaling factor".

5.3.2 Efficiency measurement in the high mass region

The efficiency will also be extracted using the "tag-and-probe" method in the high mass region (120 GeV/c^2< M < 600 GeV/c^2). In this region, however, the background is expected to be more important than in the Z peak region. More stringent conditions have therefore been imposed on the events in order to reduce the background contaminations. Both electrons are asked to have a transverse energy higher than 50 GeV/c (E_t > 50 GeV/c) and be produced back to back ($|\Delta\phi| > 2.6$) and the ratio of their transverse momenta be comprised between 0.75 and 1.33. In addition, the tag electron is required to have an HCAL energy deposit smaller than 2 GeV.

It was checked [40] that the efficiency ϵ_{cand} can be measured from the data with a systematic uncertainty of about 5%, by varying by a factor 2 the contribution of the background, consisting mainly in W+jet events. The results, taken from Monte Carlo simulation, are of 96 ± 1% in the barrel and 94 ± 1% in the endcaps.

The efficiency ϵ_{sel} for the HEEP selection criteria is measured to be 0.95 ± 0.01 (stat.) ±0.05 (syst.). The background, dominated by the W+jet channel, is measured from the events where the two electron candidates have the same charge; the small charge mismeasurement in the Drell-Yan signal (∼5%) and the charge correlation between the quark and the lepton in the W+jet events are taken into account. The systematic error is obtained by varying by factors 2 and 0.5 the background contribution from the W+jet process (charge

correlation of 1.4) and the QCD contribution (charge correlation of 1.0).

However, during the first phase of data taking, the amount of data available might not be sufficient to provide an efficiency measurement using the "tag-and-probe" method with reasonable uncertainties. In this case, the idea would be to use the efficiency measurement extracted at the Z peak — where high statistics are available — using the "tag-and-probe" method and to extrapolate this measurement to the high mass region using the efficiency dependency in E_t determined from Monte Carlo simulation studies. The Z peak and high mass regions offer the possibility to test this extrapolation method. Indeed, the efficiency estimation from the tag-and-probe method in the high mass region can be cross-checked with the estimation provided by the extrapolation from the Z peak to higher masses.

5.4 Background estimates using data

The HEEP selection aims at rejecting as much background as possible while keeping as much signal as possible. Some sources of background may remain after this selection. They mainly come from the "jet background" where one jet (W+jet) or two jets (QCD jets) fake electrons and also from the $t\bar{t}$ processes giving two real isolated electrons in the final state. It is important to estimate these backgrounds directly from data in startup conditions. Several methods have been developed to estimate these remaining sources of background.

5.4.1 Measurement of the di-electron background with the $e\mu$ method

The $e\mu$ method [44] has been designed to estimate the di-electron background coming from the $t\bar{t}$, $Z \rightarrow \tau\tau$, $WW \rightarrow ll$ and $tW \rightarrow ll$ processes. In all these background processes, two particles decay independently into a final state lepton, which leads to a pair of leptons with either the same flavour or with different flavours. The contribution of the di-electron backgrounds to the e^+e^- spectrum is thus obtained in the data by counting the number of events where the two decays lead to different lepton flavours, e and μ. The number of selected $e\mu$ events $N_{e\mu}$ should be twice that of the ee background.

An estimation of the di-electron background is thus possible through a measurement of $N_{e\mu}$. The muon selection is performed using the standard muon selection criteria with an additional requirement on the tracker isolation.

One has however to take into account the acceptances which are different for electrons and muons as well as the selection efficiencies. In addition, there is a possible contamination coming from the W+jets events. Eventually, the following formula, summarising all these corrections, is used:

$$N_{ee}^{est} = \frac{N_{t\bar{t}\rightarrow ee}}{N_{t\bar{t}\rightarrow e\mu}} \cdot \left(\frac{1}{1+\frac{N_{Wj\rightarrow e\mu}}{N_{t\bar{t}\rightarrow e\mu}}}\right) \cdot N_{e\mu}^{obs} \tag{5.1}$$

$$N_{ee}^{est} = \frac{1}{2} N_{e\mu}^{obs} \cdot \frac{A}{1+R'} \tag{5.2}$$

where A denotes the ratio of the acceptances times selection efficiencies for the ee and $e\mu$ channels, and R denotes the fractional contamination of the $e\mu$ sample from W+jets events.

Table 5.2 gives the number of selected genuine $e\mu$ events from the $t\bar{t}$, $Z \rightarrow \tau\tau$, $WW \rightarrow ll$ and $tW \rightarrow ll$ processes. In addition a possible contamination comes from the W+jets process where the muon comes from the W decay and the jet fakes an electron or vice-versa. Also

Process	number of $e\mu$ events
$t\bar{t}$	486.6
$Z \to \tau\tau$	113.1
WW	60.3
tW	36.0
total di-electron background	696.0 ± 26.4
W+jets	237.1
$Z \to \mu\mu$	30.6
total contamination	267.7 ± 16.3

Table 5.2: Numbers of events with one electron passing the HEEP criteria and one muon passing the muon selection and isolation criteria, with $M_{e\mu} > 40$ GeV/c^2, both leptons having $E_t >$ 25 GeV and the electron being outside a cone of radius $\Delta R = 0.1$ around the muon direction, for an integrated luminosity of 100 pb^{-1}, at $\sqrt{s} = 10$ TeV [42].

another possible contamination comes from the $Z \to \mu\mu$ where one of the muons emits a Bremsstrahlung photon in the ECAL which is reconstructed as an electron. The total contamination, which is of the order of 30%, is corrected using Monte Carlo simulations.

Figure 5.1 presents an estimation of the sum of the di-electron backgrounds to the Drell-Yan spectrum computed with the $e\mu$ method, for events selected using the HEEP selection with $40 < M < 800$ GeV/c^2. Each electron is required to have a transverse energy higher than 25 GeV. The statistics are scaled to an integrated luminosity of 100 pb^{-1}. The different di-electron backgrounds determined from Monte Carlo simulations are also shown for comparison. The agreement is good. The main uncertainty in the background extraction from the $e\mu$ method is in the correction for the contamination of dilepton events due to the W+jet and $Z \to \mu\mu$ processes. A systematic error corresponding to ± 100% of this correction is applied to the $e\mu$ method, to account for the imperfect knowledge of the contamination.

A method, the "b-tag" method, has also been designed specifically to measure the $t\bar{t}$ background as explained in ref. [44], thus allowing for cross-check.

5.4.2 Measurement of the jet background with the "fake rate" method

The "fake rate" method allows to estimate the number of jet background events in which at least one jet fakes an electron. This background contains the QCD multijet process where two jets pass the HEEP selection criteria (events with two fake electrons), the W+jet process (events with a real electron from the W decay and a fake electron) and the γ+jet process (events with a photon reconstructed as an electron and a fake electron). The principle of the method is to determine the probability for a jet to fake an electron. This probability is determined in several bins in E_t, given the E_t dependency, and separately in the barrel and in the endcaps. However, the energy of a jet reconstructed as an electron is different than the energy of this jet reconstructed as a jet. Energy correction factors have thus to be applied in order to account for these differences.

First, a sample is constructed which contains, for each event, an object passing the full HEEP selection. In addition, it is required that no other object in the event fulfills the electron candidate requirements. This criterion gets rid of the contamination coming from the process $Z \to e^+e^-$. Similarly, the contamination coming from the process $W \to e\nu$ is suppressed asking the object that passes the HEEP selection to be back to back in the transverse plane with another jet in the event ($|\Delta\phi - \pi| \leq 30^{\circ}$). The probability is then

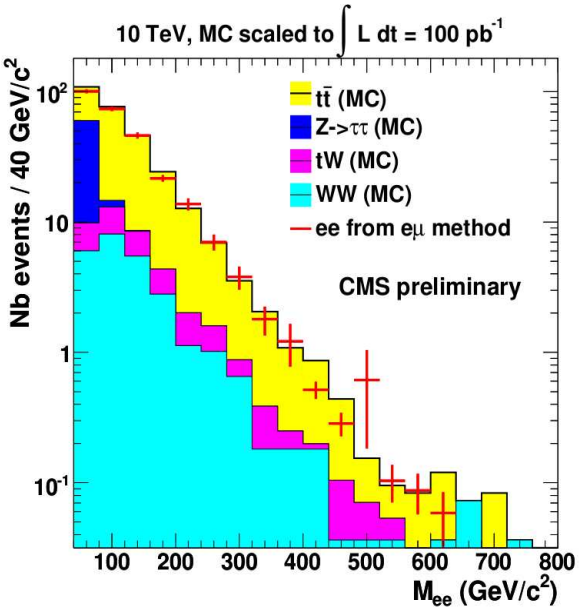


Figure 5.1: (a) Estimation of the sum of the di-electron backgrounds to the e^+e^- Drell-Yan spectrum, computed from the $e\mu$ method, for events selected using the HEEP criteria, with $40 < M < 800$ GeV/c^{2}and each electron with $E_t > 25$ GeV [42].

obtained dividing the E_t spectrum of jets faking electrons by the E_t spectrum of all jets.

Once this probability is determined, a second jet sample is constructed asking, for each event, an object that satisfies the HEEP selection criteria, an additional jet and no other object passing the loose electron criteria. The electron-jet invariant mass spectrum of such events is then weighted by the "fake rate" for the second jet and energy correction factors are applied. This finally provides an estimate of the jet background spectrum.

Figures 5.2 (a) and (b) present the number of di-jet events estimated by the fake-rate method, in the barrel (a) and in the endcaps (b), as a function of the mass. Comparisons with the expectation extracted directly from the Monte Carlo simulation show good agreement. For this comparison, only the H/E, $|\Delta\eta_{in}|$ and $|\Delta\phi_{in}|$ criteria were used to select the fake electrons in order to have enough events for the comparison. The statistics are scaled to 100 pb^{-1}.

5.5 Final di-electron mass spectrum

In order to show that the detector performance and particle reconstruction are understood, the di-electron mass spectrum is studied, from the Z boson peak up to high masses.

Figure 5.3 presents an example of the mass distribution of di-electron events passing the HEEP selection cuts, with electron $E_t > 25$ GeV, to be expected in the data for an integrated luminosity of 100 pb^{-1}. The spectrum is obtained from a randomly chosen pseudo-experiment which includes the contributions of the Drell-Yan process and of a Z' boson of mass 1 TeV/c^2 and with a natural width of 30 GeV. It also includes contributions from the various backgrounds discussed in the introduction: the $\gamma\gamma$ background, the jet backgrounds (QCD multi-jet, W+jet and photon+jets), and the di-electron backgrounds ($t\bar{t}$, tW, WW and $Z \to \tau\tau$). The expected spectrum of these backgrounds, corresponding to large statistics samples scaled to 100 pb^{-1}, are shown on the plots.

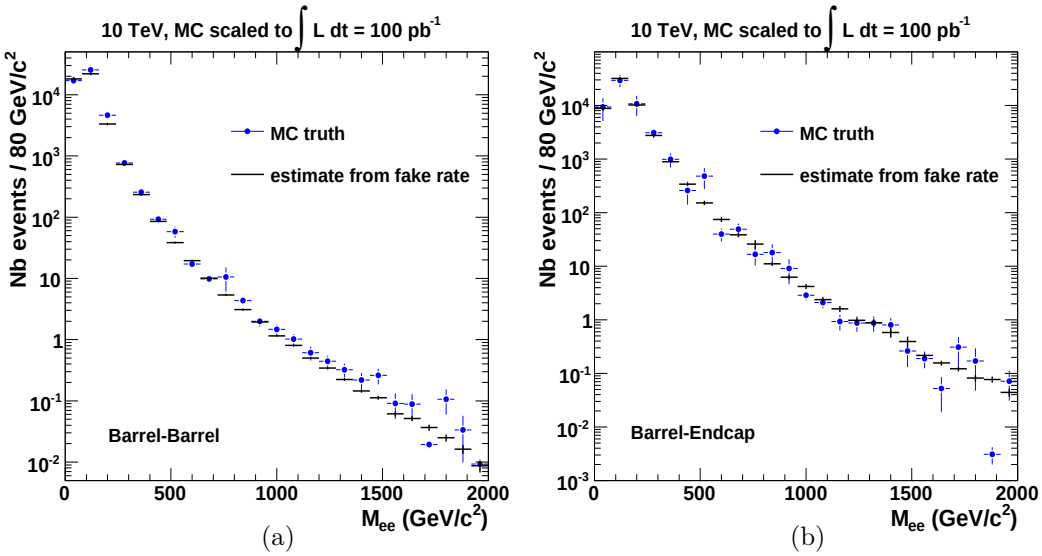


Figure 5.2: Number of di-jet background events estimated by the fake rate method in the barrel (a) and the endcaps (b), compared to the expectation extracted directly from the Monte Carlo simulation, for a loose electron selection (see text), for an integrated luminosity of 100 pb^{-1} [42].

The "data" of the pseudo-experiment and the sum of the Monte Carlo simulations are normalised using the number of events in the Z peak region ($60 \leq M_{ee} \leq 120$ GeV/c^2). As expected, the data in the pseudo-experiment follow well the "true" total spectrum. In the high mass range, $160 \leq M_{ee} \leq 560$ GeV/c^2, the agreement in shape between measurement and expectation is quantified by a χ^2 fit in 5 bins in mass, giving a probability of 63.3%. The few events in excess at 1 TeV/c^2 over negligible background indicate the presence of the Z' boson. As said in section 3.3.4, the ECAL energy resolution for high energy electrons is excellent. Alignment and calibration issues can however alter this resolution. In the scenario of perfect alignment and calibration, the invariant mass resolution is dominated by the energy resolution on the electron pair and is thus excellent: 0.6%.

5.5.1 Drell-Yan cross section measurement

As mentioned in section 5.2, the measurement of the Drell-Yan cross section is a key aspect of the HEEP strategy. In the high mass tail ($120 < M < 600$ GeV/c^2), it will allow to prove that the detector and reconstruction procedures are well under control, comparing the results from the cross section measurement with the expectations from the Standard Model.

In a practical manner, the cross section is extracted from the final selected sample, using the following formula:

$$\sigma(DY) = \frac{N - N_{bg}}{A \cdot \epsilon \cdot L} \tag{5.3}$$

where N is the number of selected events, N_{bg} the number of background events in the final sample, A the acceptance, ϵ the high p_t electron selection efficiency and L the integrated luminosity.

As presented in sections 5.3 and 5.4 respectively, methods to estimate the high p_t electron selection efficiency ϵ and the number of background events N_{bg} are well defined and have

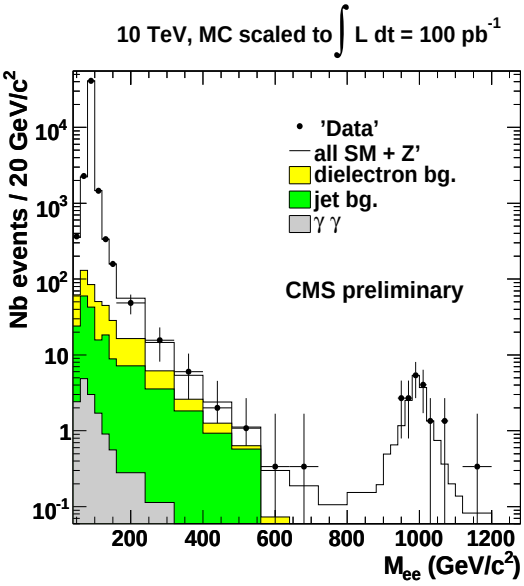


Figure 5.3: Di-Electron invariant mass spectrum for a 100 pb^{-1} pseudo-experiment including a 1 TeV/c^2 Z' signal, compared to the expected signal and SM spectrum for the Drell-Yan process and the jet, di-electron and $\gamma\gamma$ backgrounds, for which the samples are scaled to correspond to an integrated luminosity of 100 pb^{-1} [42].

proven to work well. Based on these methods, it is possible to establish the Drell-Yan invariant mass spectrum together with the contributions from the different backgrounds. Figure 5.4(a) shows the di-electron mass spectrum, in the high mass region ($120 < M < 600$ GeV/c^2), containing the contributions from the Drell-Yan process and the di-electron, jet and $\gamma\gamma$ background processes. The electrons are required to pass the full set of high p_t electron criteria and have a transverse energy above 50 GeV. The samples are scaled to correspond to an integrated luminosity of 100 pb^{-1}.

A pseudo experiment is performed in this mass region, containing the Drell-Yan process and the various backgrounds already mentioned. The total number of events in this pseudo experiment is 501. The contribution from the di-electron background is estimated to be 54.1$\pm$4.3(stat)$\pm$8.5(syst), using the $e\mu$ method. The systematic error comes from the subtraction of the non-genuine dilepton contamination in the $e\mu$ sample (section 5.4.1) which is estimated to be 16%. The contribution from the jet background, using the "fake-rate" method (section 5.4.2), is estimated to be 4.5 events with a systematic error of 50%.

Figure 5.4(b) shows the background subtracted di-electron mass spectrum obtained from a pseudo-experiment containing the Drell-Yan process together with the various background contributions mentioned in the previous section. The Monte Carlo prediction for the Drell-Yan process, for a pseudo-experiment corresponding to an integrated luminosity of 100 pb^{-1}, is shown for comparison. The agreement in shape and normalisation is good. The probability of a χ^2 test between measurement and Monte Carlo prediction, when the pseudo-data are distributed in 7 bins in invariant mass, is 67%.

The measured cross section is $15.92\pm0.89(\text{stat})^{+1.21}_{-0.90} \times 10^3$ fb, using the high p_t electron selection efficiency determined in section 5.3 and the estimated number of background events determined above. This measurement is in agreement with the Standard Model NLO cross section which is 12.20 $\times 10^3 \times$ 1.35 (K factor) = 16.47 $\times 10^3$ fb at 10 TeV.

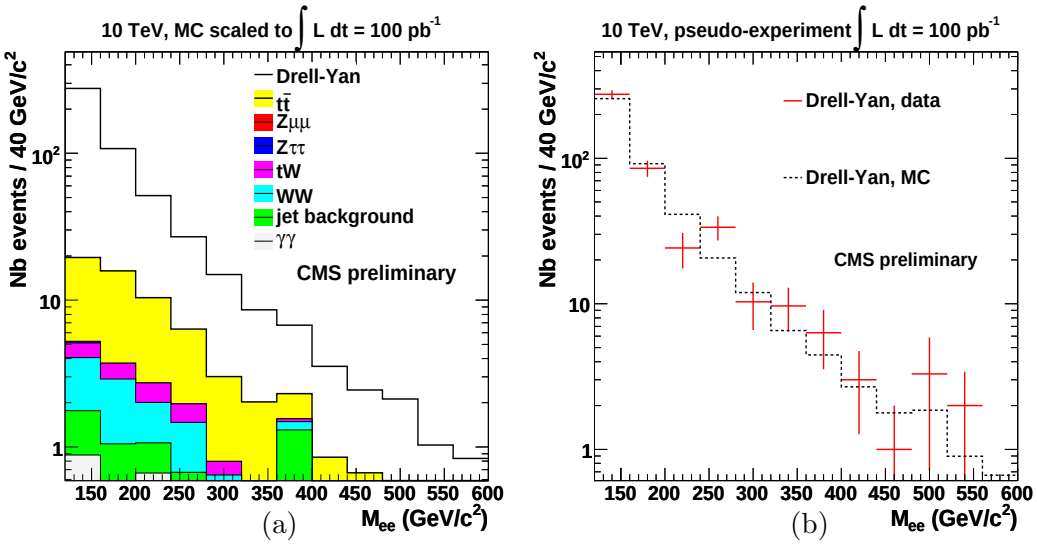


Figure 5.4: (a) Expected di-electron mass spectrum for events passing the high p_t electron selection cuts with electron $E_t > 50$ GeV, containing the contributions of the Drell Yan process and the various di-electron, jet and $\gamma\gamma$ backgrounds. The samples are scaled to correspond to an integrated luminosity of 100 pb^{-1}; (b) measured Drell-Yan spectrum obtained from a pseudo-experiment (for a luminosity of 100 pb^{-1}) containing the Drell-Yan signal and the di-electron and jet backgrounds, after background subtraction; the superimposed histogram is the prediction for a pure SM pseudo-experiment [42].

5.6 The search for resonant structures in the di-electron channel

As explained in section 5.2, the search for resonant structures [42] in the di-electron channel will be carried out in the region with invariant mass above 600 GeV/c^2. The same selection (high p_t electron selection) will be applied in the Z peak, high mass and discovery regions. The selection efficiency for the discovery region will be determined from a Monte Carlo value scaled and extrapolated as explained in section 5.3.2. The reducible backgrounds are neglected in this region, the only background considered comes from the Drell-Yan process. Thus, the final di-electron sample is assumed to contain the contribution from the Drell-Yan process together with a possible signal. In all this section, the background refers to the Drell-Yan background process whereas the signal refers to any possible heavy resonant structure in the di-electron channel (Z' or graviton).

5.6.1 The 5σ discovery reach

The identification of new resonant structures goes through an unbinned likelihood fit to the e^+e^- invariant mass spectrum over a range which includes the Drell-Yan spectrum as well as a possible signal. To show that this excess of events is not due to a fluctuation of the Drell-Yan background, a significance estimator is used which is defined as the likelihood ratio:

$$S = \sqrt{2\ln(L_{S+B}/L_B)}, \tag{5.4}$$

where L_{S+B} is the likelihood obtained from the fit to the background and signal and L_B the likelihood obtained from the fit to the background only. To claim a discovery, the $S > 5$

condition is required, corresponding to a 5 σ probability with a gaussian distribution.

The functions used to describe the shape of the signal and the background are the following. The signal is modeled as a "Breit-Wigner" function with mass and width fixed to the model parameters convoluted with a gaussian of width corresponding to the mass resolution. The background (Drell-Yan process) is described by an exponential function multiplied by a power law, as obtained from a simulation of large Drell-Yan event samples.

Fig. 5.5 shows the integrated luminosity needed to reach the necessary significance as a function of the invariant mass. One sees that, at 10 TeV and with an integrated luminosity of 100 pb^{-1}, CMS has a 5σ discovery potential for Z'_{SSM} bosons with mass up to 1370 GeV/c^2 and for Randall-Sundrum gravitons with mass up to 1310 GeV/c^2 for a coupling parameter c =0.1. These numbers supersede the recent Tevatron results. .

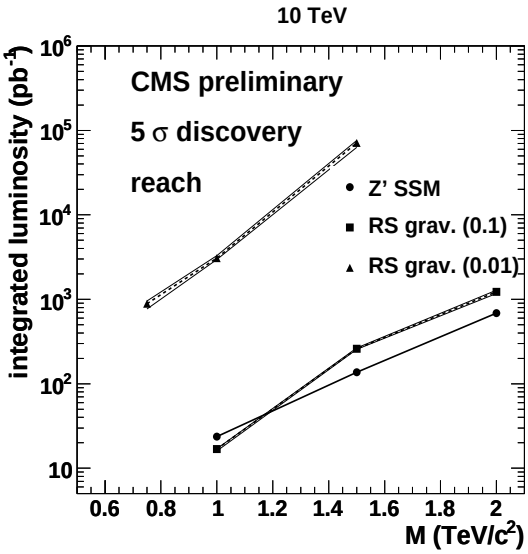


Figure 5.5: Integrated luminosity needed to reach 5σ significance, as a function of mass, for SSM Z' bosons and for RS gravitons with coupling $c = 0.1$ and $c = 0.01$. The three nearly indistinguishable lines for the RS gravitons with $c = 0.1$ and $c = 0.01$ are for an additional variation by $\pm$35% of the prediction of the Drell-Yan background [42].

5.6.2 The exclusion limits in the absence of signal

In case no signal is observed, exclusion limits have to be put. The method consists in determining, given the presence of a known background, how many events have to be added to reach a certain confidence level. A given number of pseudo-experiments is generated containing a mixture of the Drell-Yan background expected at 100 pb^{-1} and a number of λ_S signal events where λ_S follows a Poisson distribution. For each value of λ_S, the significance is computed as well as the significance for the background-only pseudo-experiments. A given number of λ_S signal events is excluded at 95 % if less than 5 % of the pseudo-experiments containing these λ_S signal events give a significance larger than the significance extracted from the background-only hypothesis. Pseudo-experiments are repeated with various values of λ_S, giving a set of points which is then fitted with an exponential. From the fit, one can determine the value of λ_S which corresponds to 5 % of the pseudo-experiments having a larger

significance when fitted with signal than when fitted with background only. For an integrated luminosity of 100 pb^{-1} and a centre-of-mass energy of 10 TeV the expected exclusion limit is found to be 2.4 for a Z'_{SSM} of mass 1 TeV/c^2. This procedure is repeated for different assumptions on the resonance mass, to extract the 95 % exclusion limit as a function of the mass.

The limit on the number of observed events can be turned into a limit on a production cross section using the following equation:

$$N_{Z'}^{limit} = \sigma_{Z'}^{limit} \cdot \epsilon_{Z'}^{tot} \cdot L, \tag{5.5}$$

where $N_{Z'}^{limit}$ is the limit on the observed decays as described above, $\sigma_{Z'}^{limit}$ is the corresponding resonance production cross section, including the branching fraction to electrons, $\epsilon_{Z'}^{tot}$ is the total selection efficiency, including acceptance effects, and L is the integrated luminosity.

This limit on the production cross section can also be normalized to the cross section of Z bosons production defined in the mass window $60 \leq M_{ee} \leq 120$ GeV/c^2, which eliminates the uncertainty on the luminosity measurement, using the following formula:

$$\frac{\sigma_{Z'}^{limit}}{\sigma_Z} = \frac{N_{Z'}^{limit}}{N_Z^{observed}} \cdot \frac{\epsilon_Z^{tot}}{\epsilon_{Z'}^{tot}}. \tag{5.6}$$

The normalized limit on the production cross section as a function of mass is shown in figure 5.6 for a Z'_{SSM} boson and for Randall-Sundrum gravitons with coupling parameters c =0.1 and 0.05. A model can be excluded in the mass range where the limit line lies below the cross section.

In the absence of signal, with an integrated luminosity of 100 pb^{-1} and with a centre-of-mass energy of 10 TeV, CMS can put lower limits at 1.7 TeV/c^2 for a Z'_{SSM} boson, 1.25 TeV/c^2 and 1.55 TeV/c^2 for Randall-Sundrum gravitons with a coupling parameter c =0.05 and c =0.1 respectively. Once again, these results are well above the results reached at the Tevatron.

5.7 Scaling to $\sqrt{s} = 7$ TeV

The eventually chosen operating centre of mass energy for proton-proton interactions at the LHC is 7 TeV, for at least a two year data-taking period. Since the cross section decreases with the centre of mass energy, the discovery potential needed to be re-evaluated. The procedure was based on the evaluation of the cross section ratios at 10 TeV and 7 TeV [45]. The cross sections times the e^+e^- branching ratio were computed, from Monte-Carlo simulation, for the Drell-Yan process with several mass thresholds, for Z'_{SSM} and Z'_{ψ} at several masses and finally for Randall-Sundrum gravitons at several masses with coupling $c = k/M_{Pl} = 0.1$. The results are listed in table 5.3 which shows the cross section times the e^+e^- branching ratio in mb for all processes mentioned above, at 14 TeV, 10 TeV and 7 TeV. The Drell-Yan cross sections are computed at leading order. No interference ($\gamma/Z/Z'$) is included in the simulation for the Z'_{SSM} and Z'_{ψ} processes.

The discovery potential, i.e. the 5σ discovery reach and the exclusion limits in case no signal is observed, are then re-evaluated based on the scaling of the discovery potential presented previously at 10 TeV, taking into account the cross section ratios presented in table 5.3.

The 5σ discovery reach is presented in figure 5.7(a) which shows the integrated luminosity needed to reach a 5σ discovery as a function of the mass. At 7 TeV and with an integrated

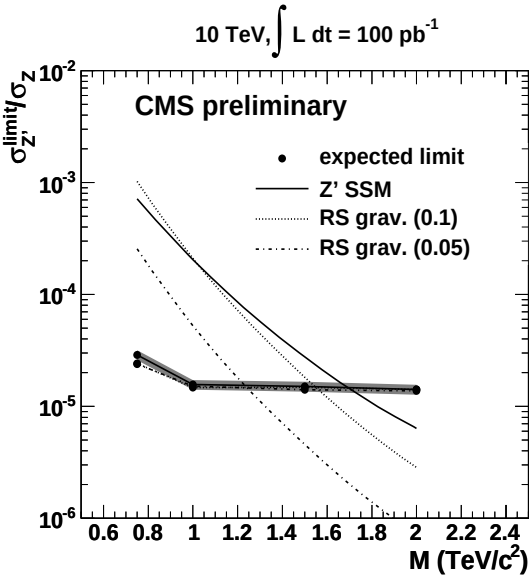
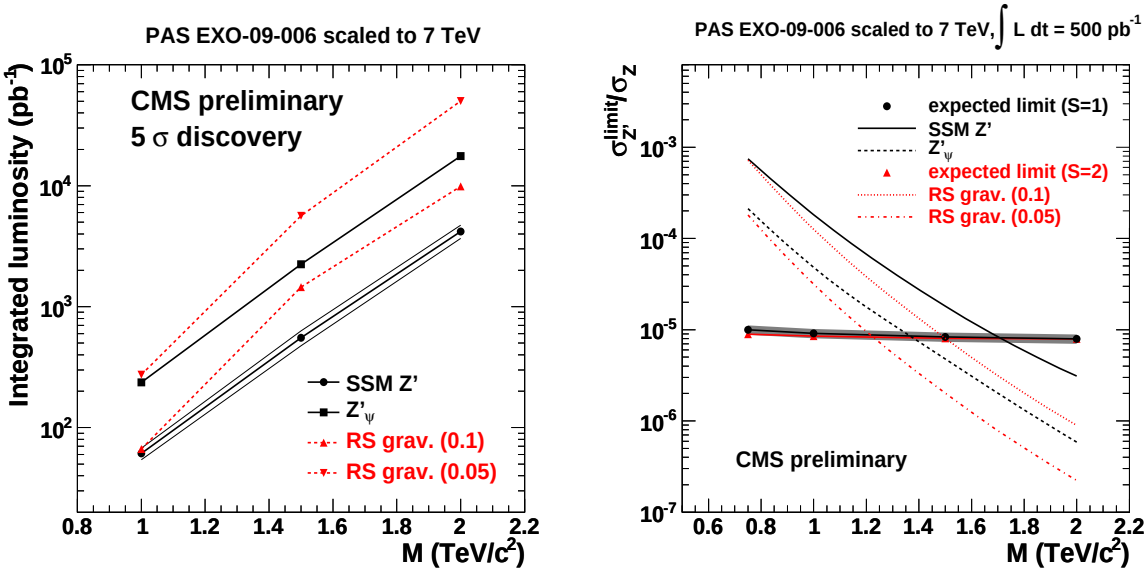


Figure 5.6: Expected 95% CL exclusion limits on the cross section of resonance production, normalised to the Z cross section, as a function of the resonance mass for the same three models; the grey band shows the effect of the variation by $\pm 10\%$ of the high mass selection efficiency [42].

luminosity of 500 pb^{-1}, CMS has a 5σ discovery potential for Z'_{SSM} bosons with mass up to 1.5 TeV/c^2 and for Randall-Sundrum gravitons with mass up to 1.3 TeV/c^2 for a coupling parameter $c = 0.1$.

The exclusion limits are summarized in figure 5.7(b) which presents the normalized limit on the production cross section as a function of mass. In this case an integrated luminosity of 500 pb^{-1} is considered. One sees that with 500 pb^{-1}, at 7 TeV centre of mass energy, CMS can put a lower mass limit at 1.7 TeV/c^2 for a Z'_{SSM}, 1.2 and 1.5 TeV/c^2 for Randall-Sundrum gravitons with a coupling parameter $c = 0.05$ and $c = 0.1$ respectively.



(a) (b)

Figure 5.7: (a) Integrated luminosity needed to reach a 5σ significance, as a function of the mass for SSM Z' bosons and for RS gravitons with coupling $c = 0.1$ and $c = 0.01$. The three nearly indistinguishable lines for the RS gravitons with $c = 0.1$ and $c = 0.01$ are for an additional variation by $\pm 35\%$ of the prediction of the Drell-Yan background. (b) Expected 95% CL exclusion limits on the cross section of resonance production, normalised to the Z cross section, as a function of the resonance mass for the same three models; the grey band shows the effect of the variation by $\pm 10\%$ of the high mass selection efficiency.

Model and resonance mass (GeV)	$\sigma \cdot$ BR (in mb)		
	$\sqrt{s} = 14$ TeV	$\sqrt{s} = 10$ TeV	$\sqrt{s} = 7$ TeV
SSM Z' M=750	$1.397{\cdot}10^{-9}$	$7.609{\cdot}10^{-10}$	$3.587{\cdot}10^{-10}$
SSM Z' M=1000	$4.513{\cdot}10^{-10}$	$2.200{\cdot}10^{-10}$	$8.823{\cdot}10^{-11}$
SSM Z' M=1500	$8.028{\cdot}10^{-11}$	$3.244{\cdot}10^{-11}$	$9.593{\cdot}10^{-12}$
SSM Z' M=2000	$2.169{\cdot}10^{-11}$	$7.326{\cdot}10^{-12}$	$1.627{\cdot}10^{-12}$
Psi Z' M=750	$4.172{\cdot}10^{-10}$	$2.245{\cdot}10^{-10}$	$1.078{\cdot}10^{-10}$
Psi Z' M=1000	$1.349{\cdot}10^{-10}$	$6.526{\cdot}10^{-11}$	$2.591{\cdot}10^{-11}$
Psi Z' M=1500	$2.378{\cdot}10^{-11}$	$9.464{\cdot}10^{-12}$	$2.623{\cdot}10^{-12}$
Psi Z' M=2000	$6.253{\cdot}10^{-12}$	$2.024{\cdot}10^{-12}$	$3.889{\cdot}10^{-13}$
RS Grav. M=750 c=0.1	$2.752{\cdot}10^{-9}$	$1.060{\cdot}10^{-9}$	$3.384{\cdot}10^{-10}$
RS Grav. M=1000 c=0.1	$6.422{\cdot}10^{-10}$	$2.139{\cdot}10^{-10}$	$5.787{\cdot}10^{-11}$
RS Grav. M=1500 c=0.1	$7.038{\cdot}10^{-11}$	$2.011{\cdot}10^{-11}$	$4.018{\cdot}10^{-12}$
RS Grav. M=2000 c=0.1	$1.285{\cdot}10^{-11}$	$2.982{\cdot}10^{-12}$	$4.221{\cdot}10^{-13}$
DY M>40	$1.826{\cdot}10^{-6}$	$1.231{\cdot}10^{-6}$	$8.043{\cdot}10^{-7}$
DY M>120	$1.917{\cdot}10^{-8}$	$1.256{\cdot}10^{-8}$	$7.872{\cdot}10^{-9}$
DY M>200	$2.590{\cdot}10^{-9}$	$1.613{\cdot}10^{-9}$	$9.581{\cdot}10^{-10}$
DY M>500	$1.002{\cdot}10^{-10}$	$5.511{\cdot}10^{-11}$	$2.660{\cdot}10^{-11}$
DY M>1000	$6.434{\cdot}10^{-12}$	$2.839{\cdot}10^{-12}$	$9.617{\cdot}10^{-13}$

Table 5.3: Cross sections (at LO) times e^+e^- decay branching ratio for several SSM Z' bosons and RS extra-dimension gravitons G with coupling $c = 0.1$, and for Drell-Yan production with $M > 40$, 120, 200, 500 and 1000 GeV/c^2, at the LHC centre of mass energy of 14, 10 and 7 TeV.

Chapter 6

The electron reconstruction and identification

In the study of the di-electron invariant mass spectrum at high energy, the HPTE (high p_t electron) selection is a key aspect of the analysis.

As a short reminder, the signal process is the following $q\bar{q} \rightarrow (\gamma/Z/Z') \rightarrow e^+e^-$. It is thus characterised by the presence of two high energy electrons in the final state, leaving therefore a clear signature in the detector (electromagnetic calorimeter + tracking detector). The background processes can alter the final di-electron invariant mass spectrum. Their contribution can be minimised by requiring specific criteria.

This chapter describes the HPTE selection and is organised as follows: the reconstruction of the electron energy deposit in the ECAL, through the clustering algorithm, is described in section 6.1. The electron reconstruction, through the GSF algorithm, is then described in section 6.2 and the efficiencies are given. The electron identification and isolation are discussed in section 6.3.

Throughout this chapter, the Monte Carlo samples used for the analysis correspond to a set of three Drell-Yan samples with masses generated above 40 GeV/c^2, 120 GeV/c^2 and 200 GeV/c^2 (as used in chapter 4, see section 4.2), assuming a centre of mass energy of 7 TeV, with a total statistics of 290 000 events. The simulation was performed with misalignment and miscalibration conditions similar to those expected during the first phase of data-taking.

6.1 Electron energy and position estimates from the ECAL: the supercluster

The subdetector designed to measure the energy and position of high energy electrons is the electromagnetic calorimeter, with excellent resolution at high energy ($\sigma_E/E \sim 0.6\%$). The principle of electron and photon detection in the ECAL has been described in section 3.3. Electron/photon showers deposit their energy in several crystals. A dedicated algorithm [46] has been implemented to measure the direction and the energy of electrons, based on the collected energies and positions of the individual crystals. This algorithm is different in the barrel and the endcaps due to the difference in geometric arrangement of the crystals.

6.1.1 The superclustering algorithm: description

The "hybrid" algorithm in the barrel

In the barrel part with an η—ϕ geometry, the hybrid algorithm is used. It starts looking for "seed" crystals which are local energy maxima with transverse energy above a given threshold $E_{t,hybseed}$. Crystal dominoes aligned in η with the seed crystal are then formed, containing 3 or 5 crystals depending on the seed crystal energy (E_{wing} is the threshold for 5 crystals). This domino-making procedure is repeated, in each direction, along the ϕ direction, for all crystals aligned in η with the original seed crystal, with a maximum of N_{crys} crystals. Only dominoes with energy above a given threshold E_{thresh} are kept. The dominoes are then clustered in ϕ and each distinct cluster of dominoes is required to have a seed domino with energy above a given threshold E_{seed}.

The clusters of dominoes, also called clusters, are gathered together to form what is called a supercluster. The total energy of the supercluster is computed from the sum of the energies of the individual crystals accounted in the supercluster and the position is computed as the barycenter of the individual crystals where the weights are the relative energy contributions of each crystal to the total energy.

The domino construction step is illustrated in figure 6.1 and the parameters introduced for the hybrid algorithm are listed in table 6.1.

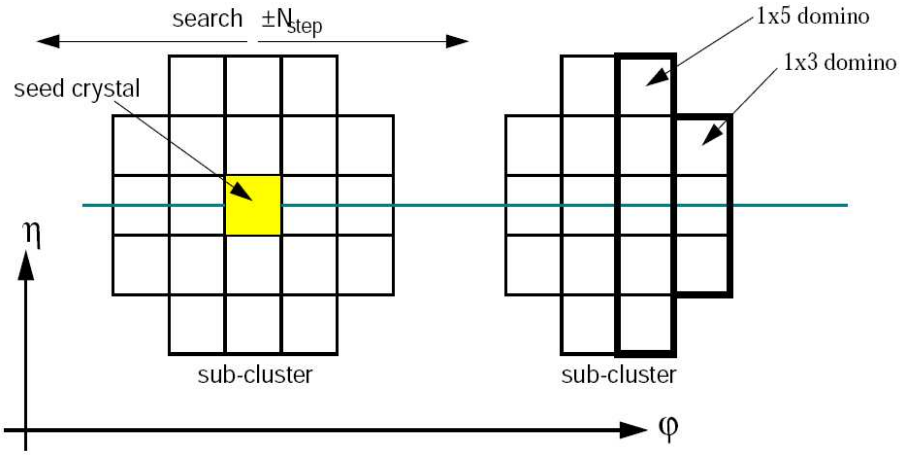


Figure 6.1: Domino construction step of "hybrid" algorithm.

Parameter	Value
$E_{t,hybseed}$	1 GeV
E_{wing}	0 GeV
N_{crys}	17 crystals
E_{thresh}	0.1 GeV
E_{seed}	0.35 GeV

Table 6.1: Parameter values used in the clustering steps of the "hybrid" algorithm.

The "Multi5x5" algorithm in the endcaps

The endcap part makes use of the "Multi5x5" algorithm. Based on the individual crystal energies, the algorithm aims to build fixed size clusters of 5x5 crystal arrays. Note that the crystals in the endcaps have a larger size compared to the barrel ones. The basic clusters are then clustered into higher level objects: the superclusters.

The algorithm starts with local energy maxima crystals with transverse energy above a given threshold $E_{t,seed}$ called seeds. Starting with the highest energy crystal seed, a 5x5 crystal cluster is made around the seed.

The algorithm then combines those basic clusters into superclusters. Energy deposits from Bremsstrahlung photons are most likely to be spread in ϕ direction. The procedure starts with a list of basic cluster seeds, i.e. clusters with transverse energy above a given threshold $E_{t,seed,BC}$. These basic clusters are then added to the supercluster if they are aligned in η and ϕ with the seed cluster within a $\Delta\eta$—$\Delta\phi$ window.

The parameters for the "Multi5x5" algorithm are listed in table 6.2.

Parameter	Value
$E_{t,seed}$	0.18 GeV
$E_{t,seed,BC}$	1 GeV
$\Delta\eta$	0.14 rad
$\Delta\phi$	0.6

Table 6.2: Parameter values used in the clustering steps of the "Multi5x5" algorithm.

The final object built by the clustering algorithm, in the barrel or in the endcaps, a supercluster, provides energy and position estimates of the electron. Note, however, that the supercluster position is computed with respect to the nominal point with coordinates (0,0,0), assuming that the electrons were emitted from this nominal point. The primary vertex is however not always located at the nominal point and its position can thus bias the pseudorapidity estimate of the supercluster[1].

To study the superclustering algorithm performance (resolution and efficiency), Drell-Yan simulation samples are used and a matching in ΔR is made between the generated electron direction (η_{ele},ϕ_{ele}) and the supercluster position (η_{SC},ϕ_{SC}) as determined by the barycenter (a cut of $\Delta R < 0.2$ is used):

$$(\Delta R)^2 = (\Delta\eta)^2 + (\Delta\phi)^2 = (\eta_{ele} - \eta_{SC})^2 + (\phi_{ele} - \phi_{SC})^2 \tag{6.1}$$

6.1.2 Superclustering algorithm in pseudorapidity, azimuthal angle and energy

Figures 6.2 present the difference between the generated pseudorapidity and the pseudorapidity estimate from the supercluster (a) and the difference between the generated azimuthal angle and the azimuthal angle estimate from the supercluster (b), separately for superclusters associated to electrons and positrons. The RMS values of these different distributions are of the order of 0.019 and 0.02 rad respectively, both for superclusters associated to electrons and positrons. The distributions of the azimuthal angle difference $\phi_{gen} - \phi_{SC}$ for superclusters

[1]The supercluster pseudorapidity can be corrected with respect to the primary vertex position if the latter is known. This new estimate, when used, will be referred as the corrected supercluster pseudorapidity.

associated to electrons and positrons peak at opposite values. This is due to the bending of the electron trajectory in the magnetic field.

Figures 6.2 present the generated pseudorapidity as a function of the supercluster pseudorapidity (c) and the generated azimuthal angle as a function of the supercluster azimuthal angle (d) for superclusters associated to electrons and positrons. The difference between the supercluster azimuthal angle and the generated azimuthal angle does not depend on the generated azimuthal angle. The difference between the supercluster pseudorapidity and the generated pseudorapidity is a bit larger for central pseudorapidities.

The resolution is better in pseudorapidity compared to azimuthal angle. This is due to photons created by Bremsstrahlung processes initiated by electrons propagating through the tracker. The energy deposits from these photons in the ECAL can sometimes not be gathered in the main supercluster from the electron but form their own single supercluster. Such effects have an impact on the energy resolution since energy of such Bremsstrahlung photon superclusters is not considered in the electron energy estimate. This is visible in Figure 6.2(e) which presents the ratio of the supercluster energy estimate over the generated corresponding electron energy for electrons (black) and positrons (red). A clear tail at low ratios is visible, corresponding to Bremsstrahlung photon energies being uncollected during the clustering process. Both distributions peak closely around 1 and in average, 94% of the initial electron energy is comprised in the supercluster energy. No significant difference is observed between electrons and positrons.

The electrons are, however, not always reconstructed as superclusters in the electromagnetic calorimeter, as the latter acceptance is limited to pseudorapidities smaller than 3. Another effect is present at pseudorapidities around ± 1.5 corresponding to the ECAL gap regions between the barrel and the endcaps. These regions are specific as the clustering algorithm is problematic here. Electrons emitted in these regions are excluded from the analysis. The gap region is defined in a conservative way as the pseudorapidity interval ($1.442 < |\eta| < 1.56$).

As the HEEP group is mainly interested in high energy electrons, a preselection cut is imposed on the transverse energy at 25 GeV in addition to the cuts on acceptance and gaps. This transverse energy cut can be lowered in the first low luminosity phase of data-taking where limited statistics are available.

6.1.3 Superclustering algorithm efficiency

The superclustering efficiency is defined as N_{rec}/N_{gen} where N_{gen} is the number of generated electrons with transverse energy higher than 25 GeV, emitted in the tracker acceptance ($|\eta| < 2.5$) and out of the gap and N_{rec} is the number of such generated electrons which pass the reconstruction step of the superclustering algorithm. This efficiency is plotted in figure 6.3 as a function of the electron transverse energy (a), pseudorapidity (b) and azimuthal angle (c). The efficiency behaviour is very uniform with respect to the three kinematic variables and the overall efficiency number is very high, around 99.8 %.

For information, the ratio N'_{rec}/N_{gen} is introduced, where N'_{rec} is the number of generated electrons with transverse energy higher than 25 GeV, emitted in the tracker acceptance ($|\eta| < 2.5$) and out of the gap which are reconstructed as superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap. This ratio, called the kinematic superclustering efficiency, is plotted in figure 6.4 as a function of the electron transverse energy (a), pseudorapidity (b) and azimuthal angle (c). One sees a drop in the low transverse energy region and around the ECAL barrel-endcap gap region. Some electrons emitted in the tracker

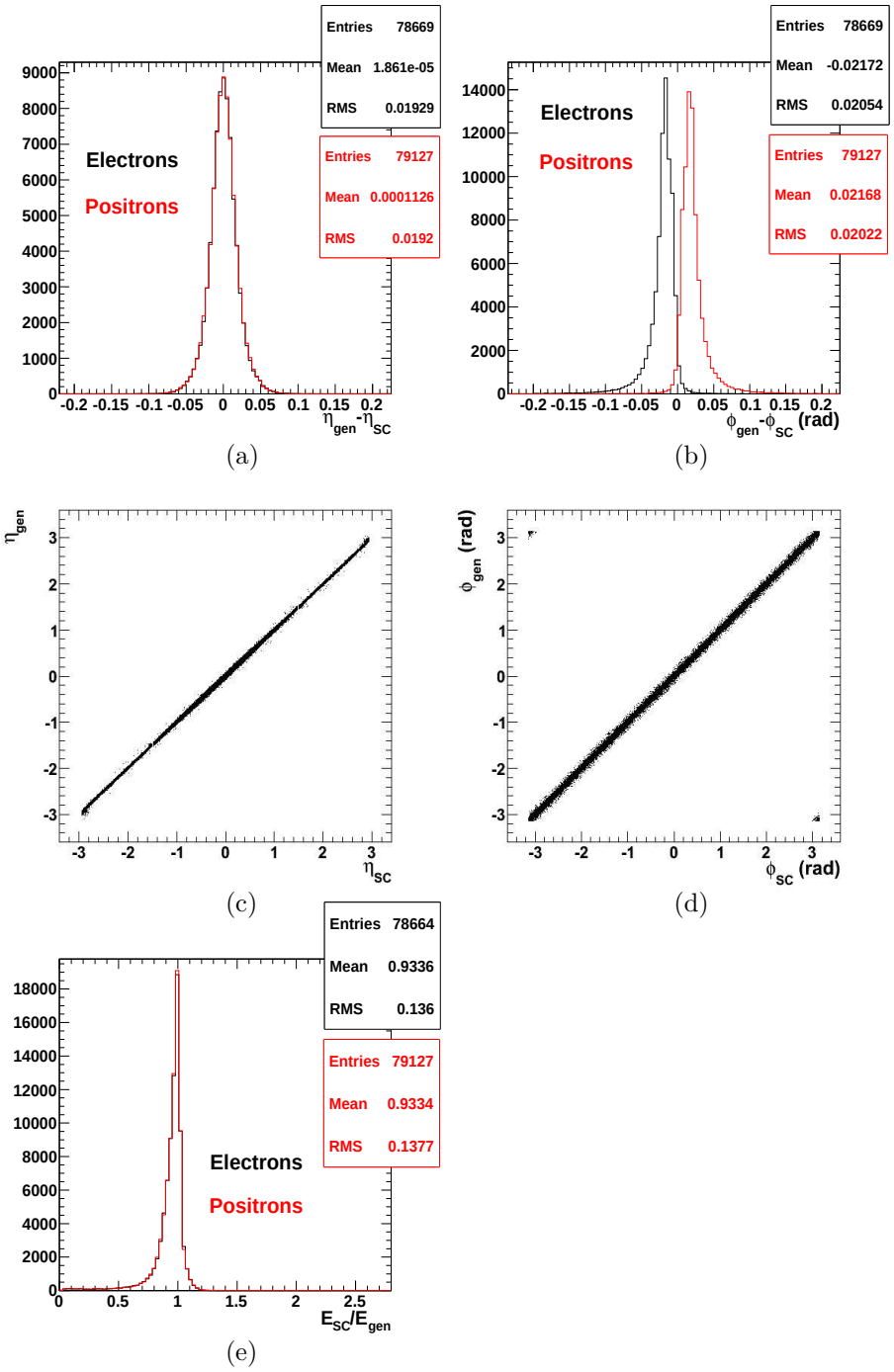
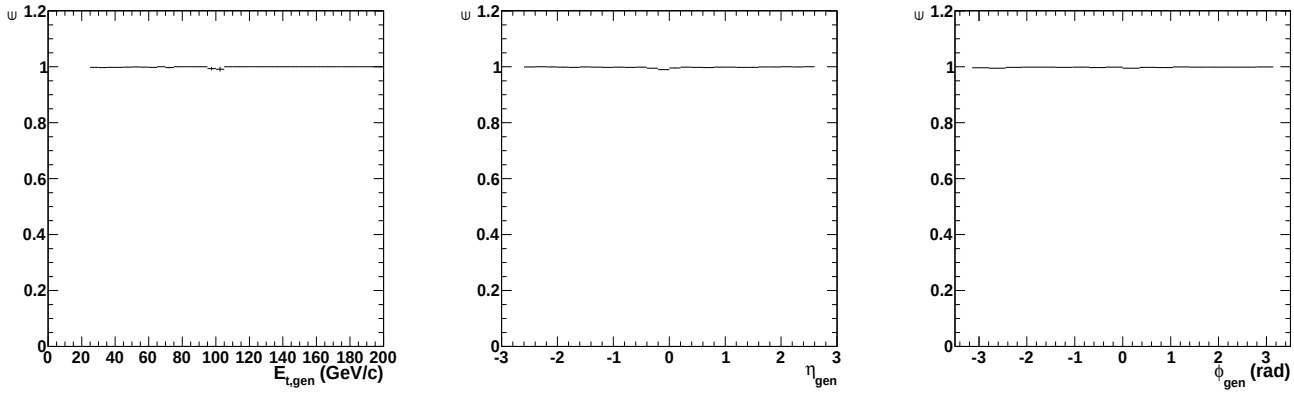


Figure 6.2: Resolution on supercluster information. Difference between the generated pseudorapidity and the pseudorapidity estimate from the supercluster (a) and difference between the generated azimuthal angle and the azimuthal angle estimate from the supercluster (b), separately for superclusters associated to electrons and positrons. Generated pseudorapidity as a function of the supercluster pseudorapidity (c) and generated azimuthal angle as a function of the supercluster azimuthal angle estimate (d) for superclusters associated to electrons/positrons. Ratio of the supercluster energy over the corresponding electron energy (e), separately for superclusters associated to electrons and positrons. (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b) (c)

Figure 6.3: Superclustering efficiency as a function of the electron transverse energy (a), pseudorapidity (b) and azimuthal angle (c) for electrons emitted in the acceptance, out of the gap region, with transverse energy above 25 GeV. (Drell-Yan production with mass above 40 GeV/c^2).

acceptance ($|\eta| < 2.5$), out of the gap and with transverse energy higher than 25 GeV are not reconstructed as superclusters in the same kinematic region by the superclustering algorithm.

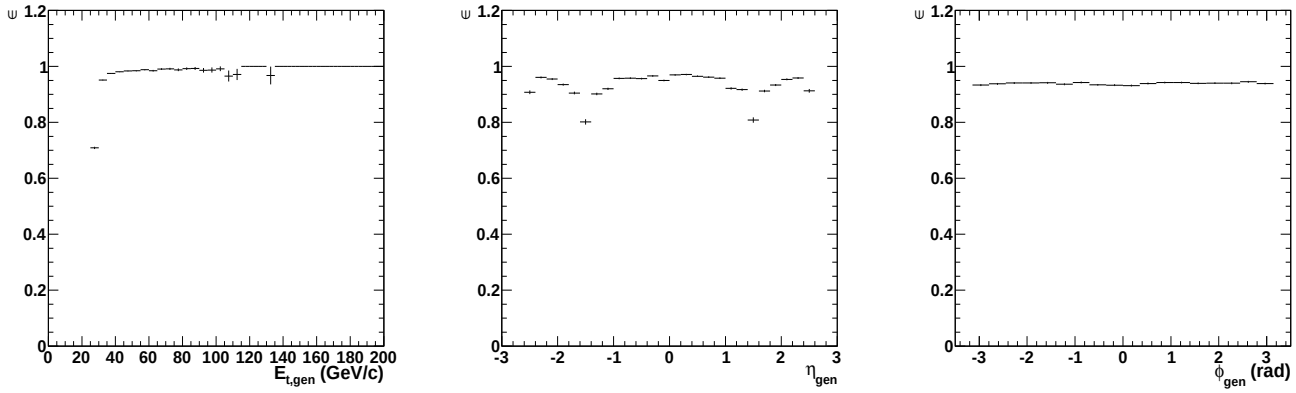
Figure 6.5 presents the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c) distributions for electrons with transverse energy higher than 25 GeV, in the tracker acceptance, out of the gap, reconstructed as superclusters but not in the same kinematic region. The majority have transverse energy lower than 25 GeV and are located in the cracks. Such electrons are generated with a transverse energy slightly higher than 25 GeV and are reconstructed as superclusters with a smaller transverse energy (< 25 GeV) as they are emitted in the gaps and inter-module cracks in η and ϕ. Also, a structure around η equal to zero is visible. This is the region between the two halves of the barrel along the z axis which holds service material.

6.2 Electron reconstruction: the GSF electron candidate

The first step of the electron selection is the 'reconstruction' part, which demands the presence of a track in the tracker with parameters compatible with the ECAL energy deposit information. This has also the advantage to discriminate between electrons and photons. The "pixelmatchgsf" algorithm [47], also here refered to as the GSF algorithm, has been designed to reconstruct a track compatible with the supercluster information. It is organised in three steps which will be described in detail: (i) the seeding, (ii) the trajectory building and the track fitting which constitute the track reconstruction and (iii) the compatibility test, also called preselection. In the following, the three steps of the "pixelmatchgsf" algorithm are detailed (sections 6.2.1 to 6.2.3) using a Monte Carlo simulation sample containing 97000 events with mass above 40 GeV/c^2.

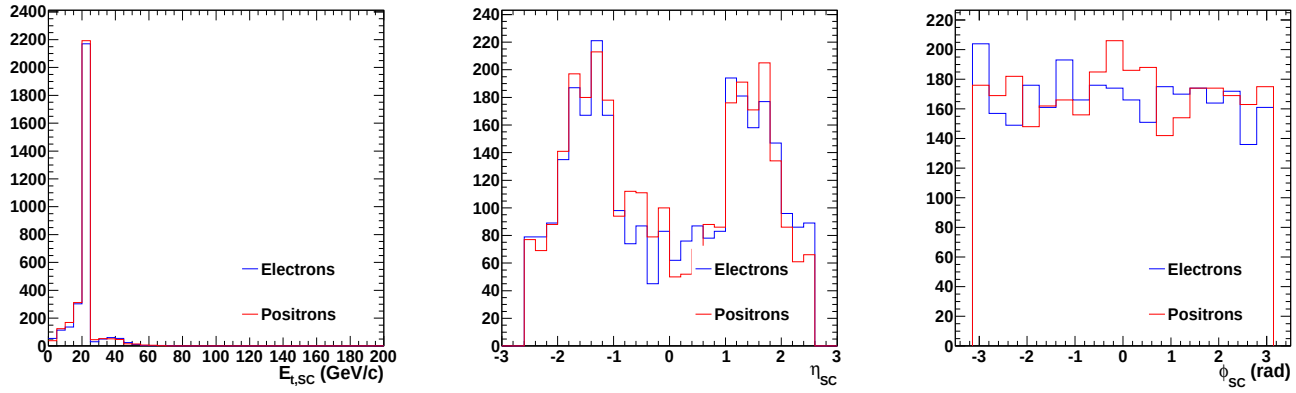
6.2.1 Trajectory seeding

The trajectory seeding consists of finding a pair of hits or triplet (set of 3 hits) in the pixel detector, called a seed, compatible with the supercluster information (energy and position).



(a) (b) (c)

Figure 6.4: Efficiency for reconstructing a supercluster with transverse energy above 25 GeV, in the acceptance, out of the gap region for electrons generated with transverse energy above 25 GeV, in the tracker acceptance, out of the gap region as a function of the electron transverse energy (a), pseudorapidity (b) and azimuthal angle (c). (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b) (c)

Figure 6.5: Supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c) for electrons emitted in the acceptance, out of the gap region, reconstructed as superclusters, but not with transverse energy above 25 GeV, in the acceptance, out of the gap region. (Drell-Yan production with mass above 40 GeV/c^2).

Starting with the supercluster position and energy, the algorithm extrapolates this position to the inner layers and forward disks of the pixel detector, based on the trajectory followed by a charged particle in the magnetic field, considering both charge hypotheses and the beam spot position. Pixel hits are then searched for in the found region using a ($\Delta\phi$ — Δz) window, in the first and second layers or the forward disks of the pixel detector successively. In addition to the pixel detector, hits can also be looked for in the silicon part of the tracker (TID and TEC). For commodity, the $\Delta\phi$ window is made dependent on the supercluster transverse energy ($E_{t,SC}$) for the first hit; the values were tuned using samples of electrons with transverse momentum between 5 and 50 GeV/c. The Δz window, for the first hit, is also variable and is computed from the beam spot position along the z axis (z_{BS}), the z-width of the beam interaction point ($\sigma_{BS,z}$) and its error ($\epsilon_{\sigma_{BS,z}}$). The corresponding parameters and their definitions are listed in table 6.3 and in equations 6.2 to 6.5. All $\Delta\phi$ values are expressed in radians and Δz values are in centimeters.

$$\Delta\phi_1 = \begin{cases} 0.23 & \text{if } E_t < 5 \text{ GeV} \\ 0.875/E_t + 0.055 & \text{if } 5 \text{ GeV} < E_t < 35 \text{ GeV} \\ 0.08 & \text{if } E_t > 35 \text{ GeV} \end{cases} \tag{6.2}$$

$$\Delta\phi_2 = 0.008 \tag{6.3}$$

$$z_{min,1} = z_{BS} - 5 \times \sqrt{(\sigma_{BS,z})^2 + (\epsilon_{\sigma_{BS,z}})^2} \tag{6.4}$$

$$z_{max,1} = z_{BS} + 5 \times \sqrt{(\sigma_{BS,z})^2 + (\epsilon_{\sigma_{BS,z}})^2} \tag{6.5}$$

Variables	Cutoff values
$\phi_{1,min,e-}, \phi_{1,max,e-}$	$-0.675 \times \Delta\phi_1$, $0.325 \times \Delta\phi_1$ (see equations)
$\phi_{1,min,e+}, \phi_{1,max,e+}$	$-0.325 \times \Delta\phi_1$, $0.675 \times \Delta\phi_1$ (see equations)
Δz_1	$z_{min,1}$, $z_{max,1}$ (see equations)
$\phi_{2,min}, \phi_{2,max}$	$-\Delta\phi_2/2$, $\Delta\phi_2/2$ (see equations)
Δz_2	$\pm$ 0.07 cm

Table 6.3: ($\Delta\phi$ — Δz) window parameter values to search for hits in the first ($\Delta\phi_1$ — Δz_1) and second layer ($\Delta\phi_2$ — Δz_2) of the pixel detector during the trajectory seeding step of the pixelmatchgsf algorithm.

Figure 6.6 presents the value of $\Delta\phi_1$ as a function of the supercluster transverse energy ($E_{t,SC}$).

The first pixel hit is looked for in the first layer of the pixel detector. A first pixel hit is also looked for in the second layer in case it would not be found in the first layer. Once a first pixel hit is found, the primary vertex position is estimated from a straight line, using the supercluster and the found first pixel hit positions. The estimate of the primary vertex position together with the first pixel hit position allows to extrapolate forward the trajectory to the next layer (second or third layer, depending on the layer where the first hit is found), assuming again the trajectory of a charged particle in a magnetic field and considering both charge hypotheses. The second pixel hit is then searched for in the extrapolated region using cutoff values listed in table 6.3. A seed is any pair of first and second compatible hits and a triplet is any set of three compatible hits. Seeds built through the algorithm described here are often referred to as 'ecaldriven' seeds.

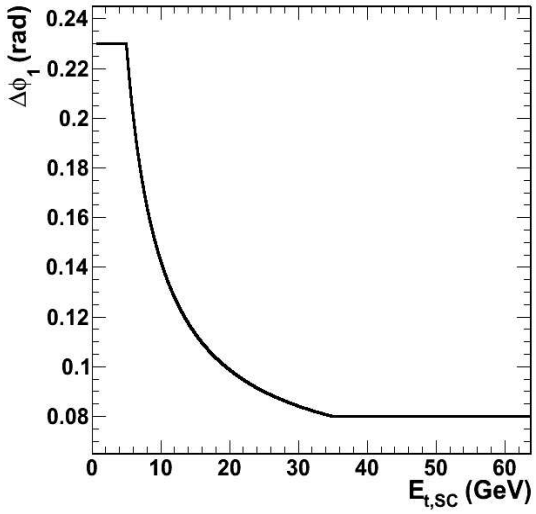


Figure 6.6: Value of $\Delta\phi_1$ window (radians) for the first hit, as a function of the supercluster transverse energy.

For very low transverse momentum electrons ($p_t < 1$ GeV), the above seeding method starts to be inefficient as the seeding window becomes too small. For this purpose, ecaldriven seeds are complemented by so-called 'trackerdriven' seeds which are built from general tracks followed as precisely as possible to the ECAL surface and matched to clusters.

In the case of high transverse momentum electrons, trajectories are expected to be rather straight[2]. As a result, it is possible, for a given supercluster, that first and second pixel hits are found for both charge hypotheses. In this case, two first hits will be found in the first layer, one for each charge hypothesis (+/-). Extrapolated to the second layer from these two first hits, two second hits will be found (+/-). Also, two first hits will be found in the second layer (+/-). Again, extrapolating to the third layer, two second hits will be found (+/-). This results in a total of eight hits, four first hits and four second hits for which the following first-second hits combinations are possible and are accounted for as seeds.

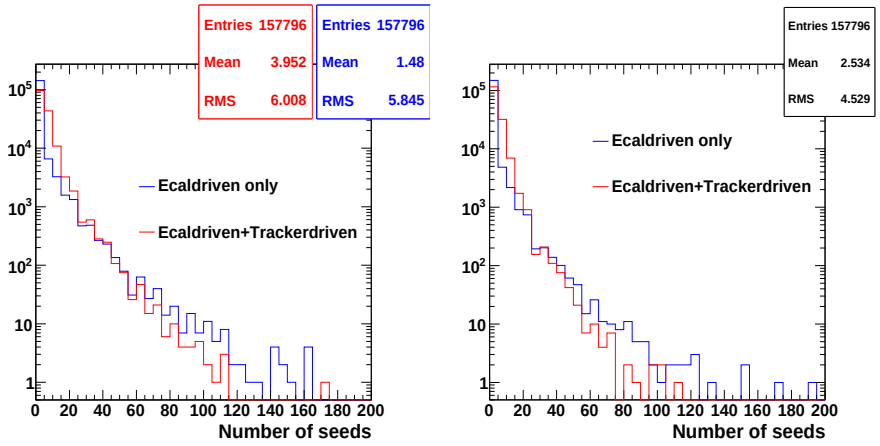
- the combination of a first + hit in the first layer and a second + hit in the second layer.
- the combination of a first - hit in the first layer and a second - hit in the second layer.
- the combination of a first + hit in the second layer and a second + hit in the third layer.
- the combination of a first - hit in the second layer and a second - hit in the third layer.

In Monte-Carlo simulation, for each supercluster, it is possible to check if the seeds found have a charge hypothesis identical or opposite to the generated electron charge.

[2]The 'straightness' of a track can be estimated through the formula $p_t = 0.3\ B\ R$ where B is the magnetic field in tesla, R is the curvature radius in meters and p_t the transverse momentum in GeV/c. Thus a 50 GeV/c transverse momentum electron has a track with curvature radius of approximately 44 m. In comparison, the distance from the interaction point to the ECAL barrel surface is about 1.3 m.

Figure 6.7 presents the number of seeds found during the seeding phase of the GSF algorithm for superclusters associated to either an electron or a positron, with (a) charge corresponding to the generated electron charge and (b) charge opposite to the generated electron charge. Ecaldriven only seeds are presented in black and seeds which are both ecaldriven and trackerdriven are presented in red.

In general, one sees that a high number of seeds is found for each supercluster associated to an electron/positron. The majority of seeds are both ecaldriven and trackerdriven. A relatively large number of seeds of charge opposite to the generated electron charge is found, in average only slightly smaller than the number of seeds with the same charge as the electron one. This is a direct effect of the relatively large $\Delta\phi$ window in which many hits and thus seeds are found for an almost straight track from a high energy electron.

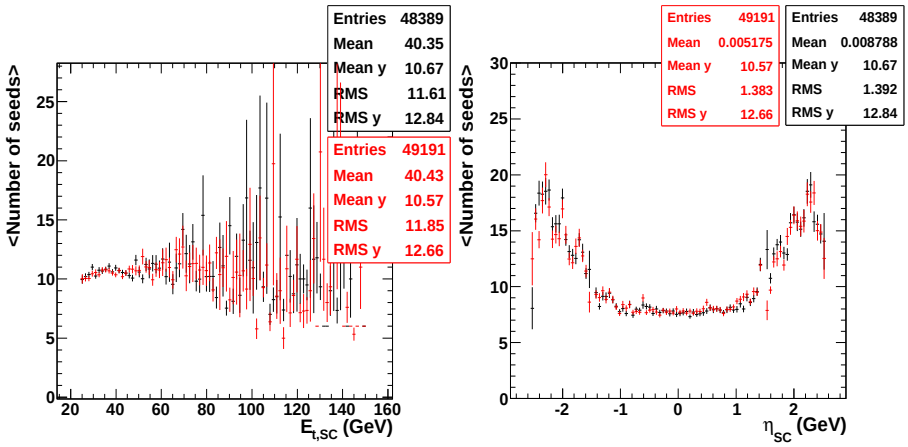


(a) (b)

Figure 6.7: Number of ecaldriven only seeds (black) and both trackerdriven and ecaldriven seeds (red) with same charge (a) and opposite charge (b) for superclusters associated to the electron/positron. (Drell-Yan production with mass above 40 GeV/c^2).

Figure 6.8(a) presents the total number of seeds (ecaldriven only or ecaldriven and trackerdriven and same and opposite charge as the generated electron charge) found as a function of the supercluster transverse energy for superclusters associated to the electrons, with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap. The total number of found seeds is fairly stable for high supercluster transverse energy ($\gtrsim$ 100 GeV). Figure 6.8(b) presents the same total number of seeds as a function of the supercluster pseudorapidity for superclusters associated to electrons, with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap. A rather constant behaviour is observed in a pseudorapidity interval up to 1, which corresponds to the beginning of the acceptance of the forward region. At pseudorapidities higher than 1, the number of seeds increases since hits in the forward disks are combined with hits in central region layers, therefore increasing the number of possible combinations. Also, the tracker material has an impact on the number of seeds as electrons from Bremsstrahlung photon conversions will induce more seeds. The total number of seeds eventually starts decreasing at pseudorapidities around 2.5 which marks the end of the tracker acceptance. No difference is observed between electrons and positrons.

Electrons with an associated supercluster for which no seed is found are not reconstructed. The seeding efficiency is thus a key information for the electron reconstruction. Many features contribute to the seeding inefficiency, the main ones being summarized hereafter:



(a) (b)

Figure 6.8: Mean value of the total number of seeds (ecaldriven only, ecaldriven and tracker-driven and same and opposite charge as the generated electron charge) found as a function of the supercluster transverse energy (a) and pseudorapidity (b) for superclusters associated to electrons (black) and positrons (red), with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap. (Drell-Yan production with mass above 40 GeV/c^2).

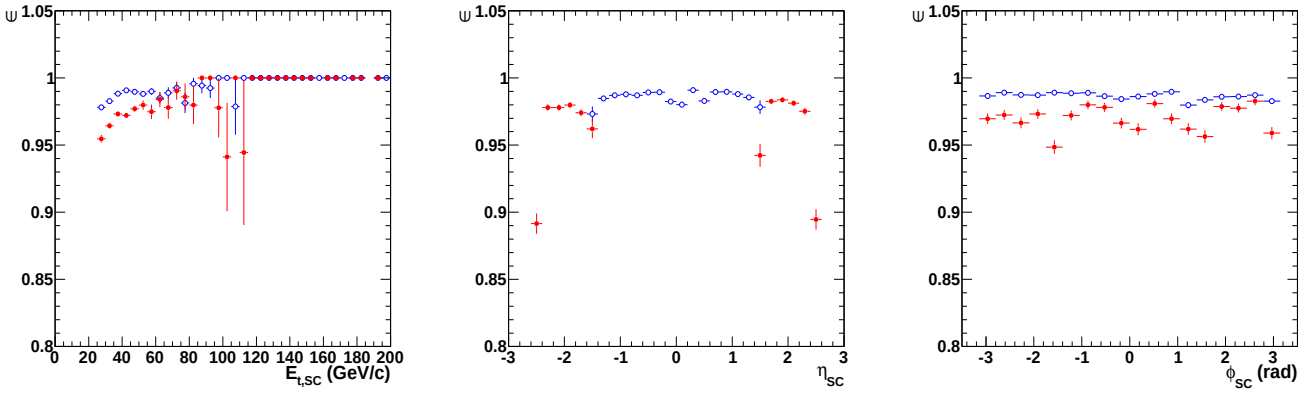
- The electron can be emitted out of the tracker acceptance. The tracker acceptance ($|\eta| < 2.5$) is smaller than the ECAL acceptance ($|\eta| < 3.0$); an electron emitted in the intermediate region ($2.5 < |\eta| < 3.0$) leaves an energy deposit in the ECAL but leaves no hits in the tracker and thus no seed is reconstructed.

- The initial electron energy is poorly reconstructed. The seeding procedure is driven by the electron energy estimate from the supercluster. Thus, a wrong estimate of the initial electron energy leads to a wrong pixel hit search region during the backward propagation from the supercluster position to the pixel detector. Such biased energy reconstruction often happens when the electron is emitted in the problematic regions of the gap or cracks in ECAL.

- The hit inefficiency. A seed is any pair of hits compatible with the supercluster position and energy estimates. Cases where hits are not reconstructed in the pixel detector are possible and are general to any kind of particle.

Starting with a number of superclusters associated to electrons/positrons, with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, the number of such superclusters for which at least one seed is found is deduced. The seeding efficiency is defined as the ratio of these two numbers and is plotted in figure 6.9 as a function of the supercluster transverse energy (a), the pseudorapidity (b) and the azimuthal angle (c), separately in the barrel and in the endcaps. The ϕ distribution shows a uniform behaviour in the barrel; the behaviour in the endcaps is due to inter-boundary regions between supercrystals where a x-y geometry is present. The η distribution shows a relatively uniform distribution, with however efficiency decrease observed in the gap, around $\eta = 0$ and close to the end of the acceptance $\eta \sim 2.5$. The E_t distribution shows a decrease in the low E_t region. The overall numbers are 98.6 % in the barrel and 97.0 % in the endcaps.

Note the decrease of efficiency at high generated pseudorapidities due to the distribution of the primary vertex position (interaction point) along the z axis which influences the acceptance. Indeed, an electron emitted with a pseudorapidity slightly smaller than 2.5 can be out



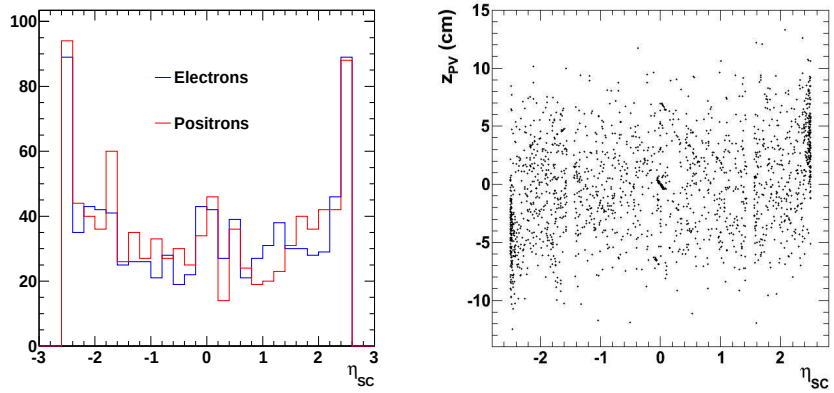
(a) (b) (c)

Figure 6.9: Seeding efficiency (see definition) for superclusters associated to electrons/positrons, with transverse energy above 25 GeV, in the tracker acceptance and out of the gap, as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel (blue open circles) and in the endcaps (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).

of the tracker acceptance if its primary vertex position is smeared to relatively large positive values (say $\sim$ 5 cm). Such an effect is presented in figure 6.10(b) which shows the primary vertex position along the z axis (z_{PV}) as a function of the supercluster pseudorapidity for superclusters, associated to electrons/positrons, for which no seed is found. One sees that superclusters with positive pseudorapidity slightly lower than 2.5 are out of the acceptance when the primary vertex is smeared to positive values. The same way, superclusters with negative pseudorapidity slightly higher than -2.5 are out of the acceptance when the primary vertex is smeared to negative values.

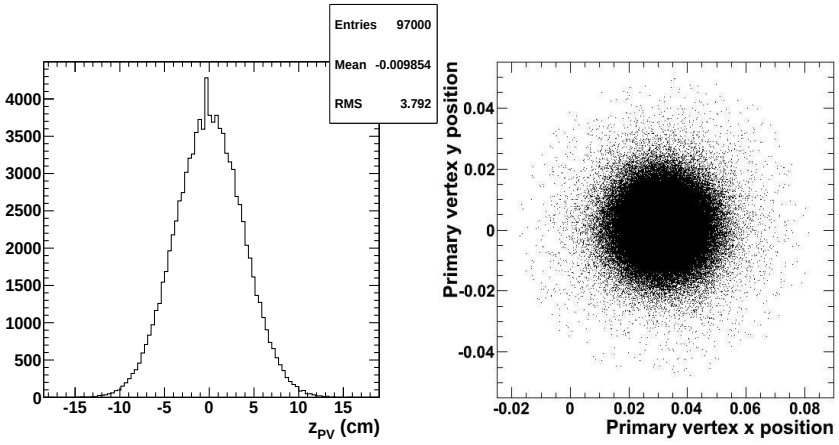
Figure 6.10(a) presents the supercluster pseudorapidity for superclusters, associated to electrons (blue) and positrons (red) from Drell-Yan production, with transverse energy above 25 GeV, in the tracker acceptance and out of the gap, for which no seed is found. The majority are located outside the tracker acceptance ($|\eta| < 2.5$) (note the difference for superclusters associated to electrons and positrons, coming from kinematics, electrons are more forward). Figure 6.10(b) the primary vertex position along the z axis as a function of the supercluster pseudorapidity for superclusters, associated to electrons and positrons, with transverse energy above 25 GeV, in the tracker acceptance and out of the gap, for which no seed is found. Structures are visible around pseudorapidities of 0.45, 0.8, 1.15 corresponding to the cracks between modules and around pseudorapidities equal to 0 corresponding to the crack between the two halves of the ECAL barrel along the z axis. In such regions, the initial energy reconstruction is problematic.

Figure 6.11(a) gives the distribution of the z position in centimeters for the primary vertex. A gaussian distribution with standard deviation of around 3.8 cm is observed. For comparison, the primary vertex position (centimeters) in the transverse plane is plotted in figure 6.11(b). The smearing is much smaller (tenths of millimeters) in the transverse plane.



(a) (b)

Figure 6.10: Supercluster pseudorapidity (a) and primary vertex z position as a function of supercluster pseudorapidity (b) for superclusters, associated to electrons and positrons, with transverse energy above 25 GeV, in the tracker acceptance and out of the gap, for which no seed is found. (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b)

Figure 6.11: Primary vertex position (centimeters) along the z axis (a) and in the transverse plane (b) (Drell-Yan production with mass above 40 GeV/c^2.).

6.2.2 Trajectory building and track fitting

Each seed is used as a starting point to build a candidate trajectory followed by the particle in the tracking detector, based on a Kalman filter [48]. The candidate trajectory is described by a five-parameter vector containing the information on momentum and direction, together with the associated covariant matrix. Starting with the two pixel hits or triplets, the trajectory is extrapolated to the next layers of the tracker, taking into account the magnetic field bending, energy loss in the material and multiple scattering. Hits are then looked for in the extrapolated region and trajectory parameters are updated, combining the information from the extrapolated parameters and the found hits; the compatibility between the former and the latter is quantified by a χ^2 value. The possibility of, at most, one missing (lost) hit is also allowed, in case of detector inefficiencies. A new track candidate is created for each compatible hit and, at each stage, only the five best ones are kept for further propagation. The procedure is repeated for each layer of the tracking detector until the last one or until a stopping condition is satisfied.

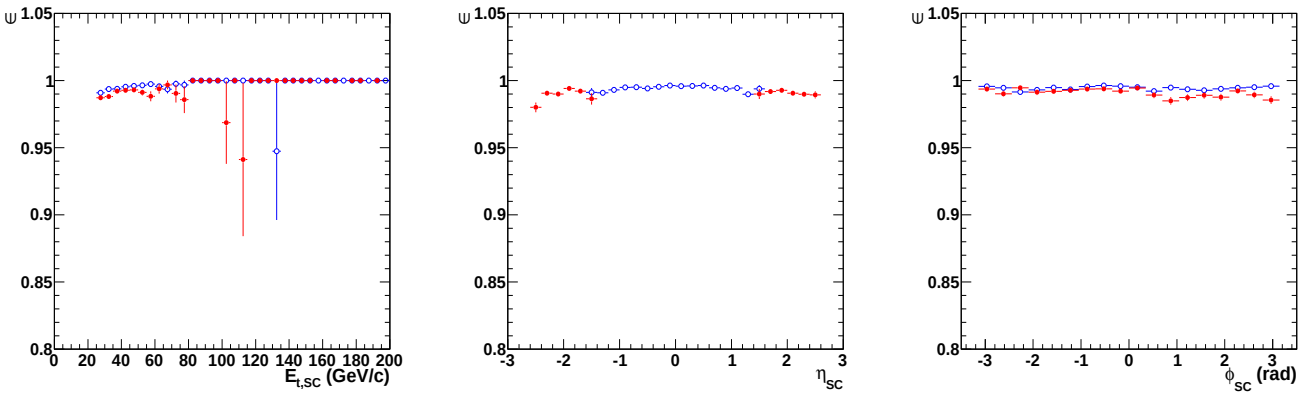
In a second step, the track fitting and the smoothing of the final track parameters are performed on all layers, using a 'Gaussian Sum Filter' (GSF) algorithm [49, 50]. This algorithm is similar, in principle, to the Kalman filter but better integrates Bremsstrahlung effects associated to electrons crossing the tracker, as it approximates the Bremsstrahlung energy loss of electrons with a mixture of gaussians rather than a single gaussian. The GSF filter is applied to all hits of the track candidate in both directions, outside-in and inside-out. The predicted results from both filters are combined to give optimal estimates of the trajectory parameters at the module surface associated with each hit. The track parameters are also updated using the information on the track incident angle on the layer to which the hit belongs. At the end of this phase, only the best three tracks (i.e. the three tracks with the smallest χ^2 value) are kept. Tracks are required to have a minimum of five hits.

We consider the number of superclusters associated to electrons/positrons, with transverse energy higher than 25 GeV, in the tracker acceptance, out of the gap which have at least one seed associated N_{before}. Among such superclusters, we consider the number of superclusters which have an associated GSF track, N_{after}. The GSF track building efficiency is defined as the ratio of these two numbers $\epsilon = N_{after}/N_{before}$. Figure 6.12 shows this efficiency as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel and in the endcaps. All distributions show relatively uniform behaviour and high efficiency. The overall numbers are 99.4% in the barrel and 99.1% in the endcaps.

Figure 6.13(a) presents the distribution of the number of GSF tracks built during the trajectory building and track fitting phase of the GSF algorithm, for superclusters associated to the electrons (black) and positrons (red). One sees that, most often and in average, one gsf track is found for each supercluster. No significant difference between electrons and positrons is observed. The pseudorapidity distribution of superclusters associated to electrons/positrons which have at least one seed but no gsf track associated is shown in figure 6.13(b).

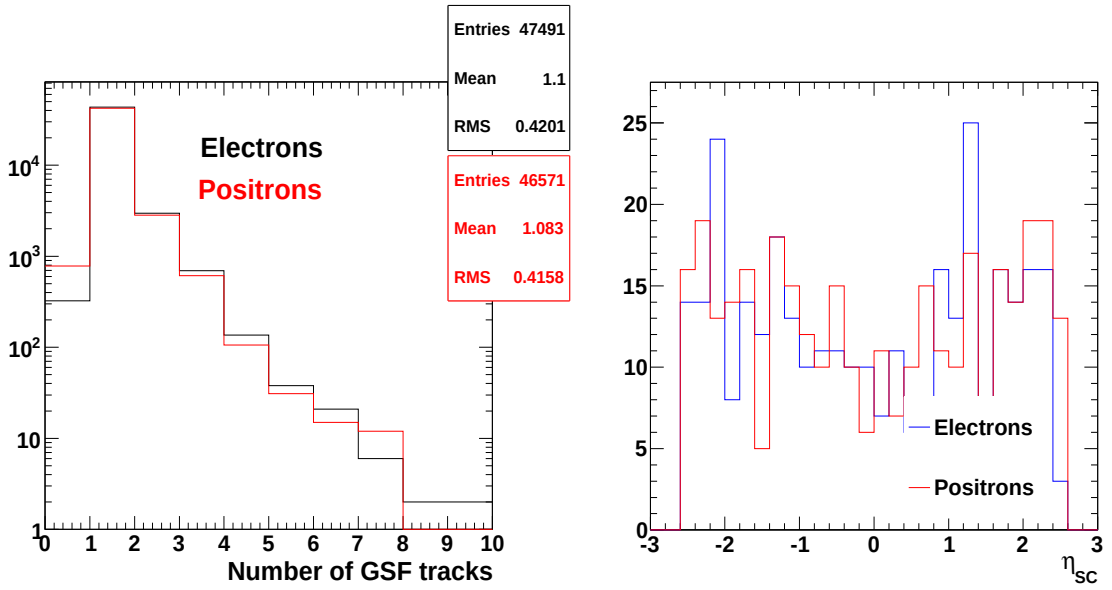
6.2.3 Preselection

The last step of the GSF algorithm requires the compatibility between the information from the reconstructed track (momentum and direction) and the information from the supercluster (energy and position). This compatibility is based on conditions applied on some predefined variables described hereafter. For each condition, the efficiency is defined as the fraction of the number of superclusters with an associated GSF track that fulfill the condition.



(a) (b) (c)

Figure 6.12: GSF track efficiency for superclusters associated to electrons/positrons, with transverse energy above 25 GeV, in the tracker acceptance and out of the gap, as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel (blue open circles) and in the endcaps (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b)

Figure 6.13: (a) Number of GSF tracks for each supercluster associated to an electron (black) or positron (red), with transverse energy above 25 GeV, in the tracker acceptance and out of the gap. (b) Pseudorapidity for superclusters associated to electrons (black) or positrons (red), with transverse energy above 25 GeV, in the tracker acceptance and out of the gap, with at least one seed but no GSF track associated. (Drell-Yan production with mass above 40 GeV/c^2).

Transverse momentum condition

The transverse energy determined from the supercluster is required to be higher than 4 GeV/c.

$$E_{t,SC} > 4 \text{ GeV} \tag{6.6}$$

This condition has no effect as we consider only superclusters with transverse energy higher than 25 GeV.

HCAL energy deposit condition

The design of the electromagnetic calorimeter has been studied to ensure the containment of electromagnetic showers initiated by electrons and photons. Thus, no energy deposit is expected for electrons in the hadronic calorimeter. Criteria are therefore imposed on the amount of energy present in the HCAL cells located directly behind the supercluster position. The position of such cells is deduced from a cone requirement with $\Delta R < 0.15$ where $\Delta R = \sqrt{(\eta_{SC} - \eta_{cell})^2 + (\phi_{SC} - \phi_{cell})^2}$. The ratio of the energy measured in the HCAL cells inside this cone, H, to the energy of the supercluster, E, has to be below a given threshold[3]. Advantage is taken of the two segmentations of the hadronic calorimeter in this case, defining an H/E value for each segmentation. Note however that two segmentations are present only from pseudorapidities starting from around $|\eta| > 1.21$.

$$H/E \ (1^{st} HCAL \ segmentation) < 0.1 \tag{6.7}$$
$$H/E \ (2^{nd} HCAL \ segmentation) < 0.1 \tag{6.8}$$

A cut on the H/E variable is already put at the seeding level. The H/E value is required to be less than 0.1. The effect of such a cut is not studied here.

Track-supercluster direction matching condition

The supercluster position (η_{SC},ϕ_{SC}) is compared to the track direction estimated at the vertex, extrapolated to the ECAL surface, assuming a helix (η_{in}^{extrap},ϕ_{in}^{extrap}). The differences $\Delta\eta$ and $\Delta\phi$ are required to be under a given threshold:

$$|\Delta\eta_{in}| = |\eta_{SC} - \eta_{in}^{extrap}| < 0.02 \tag{6.9}$$
$$|\Delta\phi_{in}| = |\phi_{SC} - \phi_{in}^{extrap}| < 0.15 \ rad \tag{6.10}$$

The condition on $|\Delta\phi|$ is looser due to the magnetic field bending which acts in the ϕ direction. The inefficiency related to the track-supercluster direction matching condition, for electron candidates with transverse energy above 25 GeV, in the acceptance, out of the gap region, is 2.28 %. Figure 6.14 presents the supercluster pseudorapidity distribution for electron candidates with transverse energy higher than 25 GeV, in the tracker acceptance, out of the gap, which have fulfilled the transverse momentum and HCAL energy deposit conditions but not the track-supercluster direction matching one. The distribution approximately follows the

[3]Considerations for the design of the electromagnetic calorimeter were based on low energy electron showers (up to $\sim$ 100 GeV/c). In the case of high energy electrons present in this study, a fraction of the electromagnetic shower is expected to take place in the hadronic calorimeter. This "leakage" phenomenon is characteristic of high energy electrons and justifies the loose threshold value.

one for the tracker material budget as a function of pseudorapidity, presented in section 3.2.3 (figure 3.7). The probability for Bremsstrahlung emission and photon conversion in the tracker is strongly affected by its material budget, and as it increases, the track measurement from the inner layer is expected to drift when extrapolated to the electromagnetic calorimeter surface.

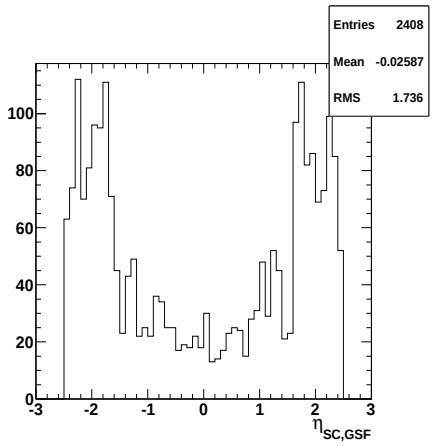
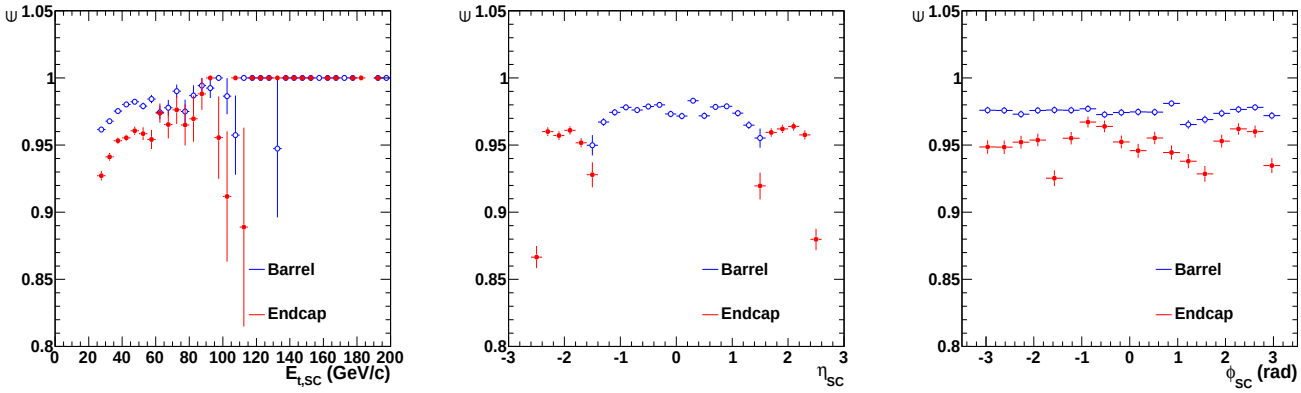


Figure 6.14: Supercluster pseudorapidity of electron candidates with transverse energy higher than 25 GeV, in the tracker acceptance, out of the gap, which have fulfilled the transverse momentum and HCAL energy deposit conditions but not the track-supercluster direction matching condition. (Drell-Yan production with mass above 40 GeV/c^2).

6.2.4 Final electron candidates

The final electron candidate mixes the information from the built track and the reconstructed supercluster and fulfills the preselection cuts mentioned in section 6.2.3. For what concerns the electron candidate energy, the information from the supercluster is used while for the angles (pseudorapidity and azimuthal angle), the information from the track is used. Figure 6.15 presents the GSF efficiency for superclusters associated to electrons/positrons with transverse energy higher than 25 GeV, emitted in the tracker acceptance and out of the gap, as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel and in the endcaps. This efficiency combines all efficiencies described in the previous sections (seeding, trajectory building, track fitting and preselection cuts). The overall efficiencies for superclusters with transverse energy above 25 GeV, in the acceptance, out of the gap region and associated to electrons from Drell-Yan production with mass above 40 GeV/c^2, are: 97.5% for the barrel and 94.9% for the endcaps.

The electron pseudorapidity and azimuthal angle can be determined either from the track direction ($\eta_{track}, \phi_{track}$) or from the supercluster position (η_{SC}, ϕ_{SC}) and the primary vertex position. Figure 6.16 shows the pseudorapidity difference $\Delta\eta$ (a) and azimuthal angle difference $\Delta\phi$ (b) for GSF superclusters associated to electrons, in the tracker acceptance, out of the gap and with transverse energy higher than 25 GeV. The difference between the GSF supercluster pseudorapidity and the generated pseudorapidity $\eta_{SC} - \eta_{gen}$ (difference between the GSF supercluster azimuthal angle and the generated azimuthal angle $\phi_{SC} - \phi_{gen}$ resp.) is shown in black, the difference between the GSF supercluster pseudorapidity corrected with respect to the primary vertex position and the generated pseudorapidity is shown in red and



(a) (b) (c)

Figure 6.15: Total GSF efficiency (see definition) for superclusters associated to electrons/positrons, with transverse energy above 25 GeV, in the tracker acceptance and out of the gap, as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel (blue open circles) and in the endcaps (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).

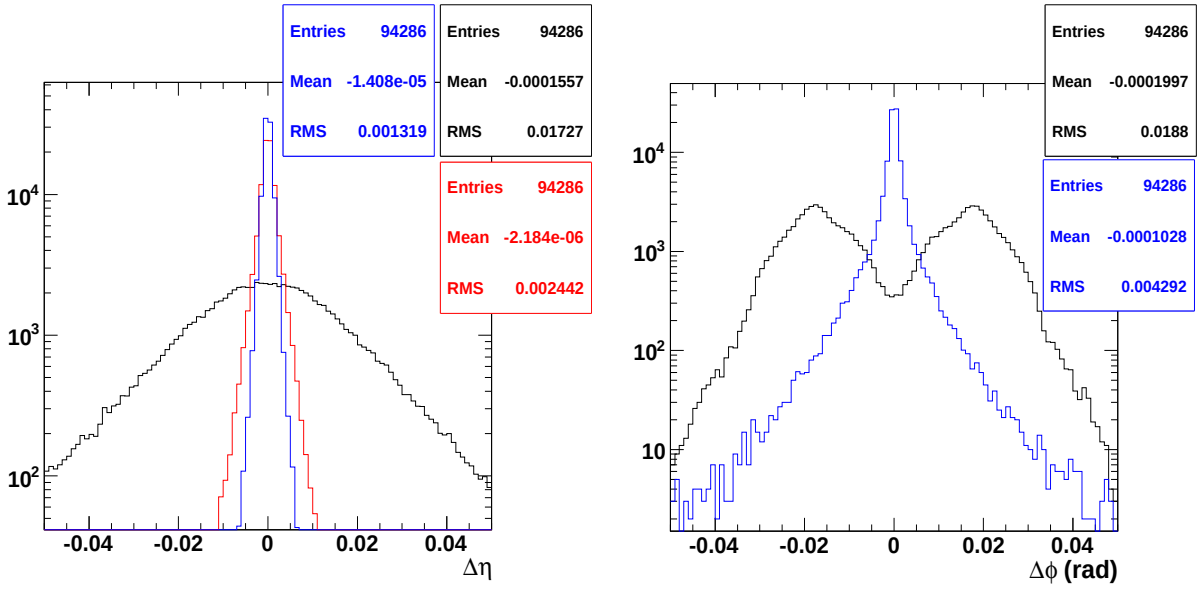
the difference between the track pseudorapidity as measured in the inner layers of the tracker and the generated pseudorapidity $\eta_{track} - \eta_{gen}$ (difference between the track azimuthal angle as measured in the inner layers of the tracker and the generated azimuthal angle $\phi_{track} - \phi_{gen}$) is shown in blue.

The best electron pseudorapidity estimate remains the track pseudorapidity as measured in the inner layers of the tracker. The same is true for the azimuthal angle. Less events are present around $\Delta\phi = 0$ for the difference $\phi_{SC} - \phi_{gen}$ due to the bending of the electron trajectory in the magnetic field.

The electron candidate charge is determined by a method which combines three different charge estimates and takes the value from the two out of three estimates that are in agreement. The first estimate is the charge as determined from the electron candidate GSF track. The second estimate takes the charge as determined from the general track matched to the GSF track, asking for at least one hit shared in the innermost tracker part (pixels). The third estimate, the 'supercluster charge', is obtained by computing the sign of the azimuthal angle difference between the vector joining the beam spot and the supercluster position and the one joining the beam spot and the first hit of the electron candidate GSF track.

The electron candidate charge is presented in figure 6.17 (a) for electron candidates associated to electrons (red) and positrons (blue). GSF superclusters are required to be in the tracker acceptance, out of the gap and with transverse energy higher than 25 GeV. One sees that a small fraction of electron candidates have a charge incorrectly assigned: 1.9% for GSF associated to electrons and 2.22% for GSF associated to positrons.

The charge misidentification rate, defined by the proportion of GSF which have their charge incorrectly assigned, is presented in figure 6.17 as a function of the GSF supercluster pseudorapidity (b) and azimuthal angle (c). GSF superclusters are required to be in the tracker acceptance, out of the gap and with transverse energy higher than 25 GeV. The charge misidentification rate increases with respect to the pseudorapidity and follows a shape related to the tracker material shape (cf. fig. 3.7). As the tracker material budget increases, the Bremsstrahlung photon radiation followed by photon conversion in the tracker, becomes more



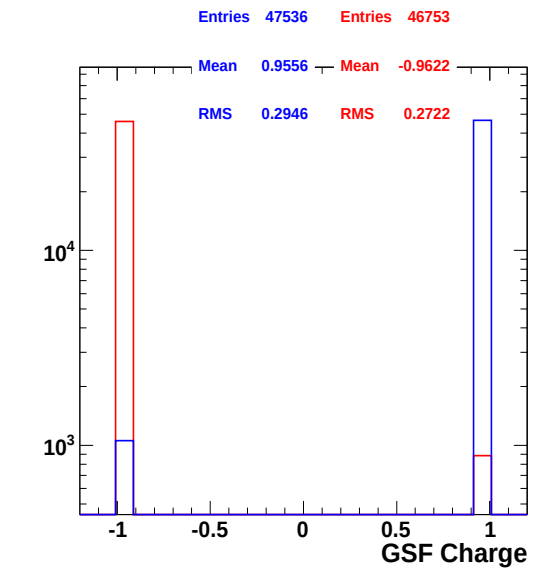
(a) (b)

Figure 6.16: Pseudorapidity difference $\Delta\eta$ (a) and azimuthal angle difference $\Delta\phi$ (b) for GSF superclusters associated to electrons, in the tracker acceptance, out of the gap and with transverse energy higher than 25 GeV. The differences $\eta_{SC} - \eta_{gen}$ and $\phi_{SC} - \phi_{gen}$ are shown in black, the difference $\eta_{SC,corr} - \eta_{gen}$ in red and the differences $\eta_{track} - \eta_{gen}$ and $\phi_{track} - \phi_{gen}$ are shown in blue. (Drell-Yan production with mass above 40 GeV/c^2).

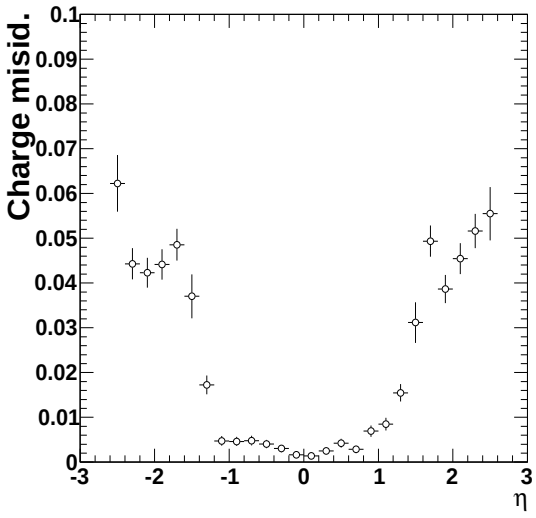
important. During the trajectory building phase, one of the two photon conversion tracks can then be taken into account and the charge determination precision is thus affected. Figure 6.17 (d) shows the charge misidentification rate as a function of the GSF supercluster transverse energy only for superclusters with absolute pseudorapidity less than 1. GSF superclusters are required to be in the tracker acceptance, out of the gap and with transverse energy higher than 25 GeV. The total charge misidentification rate for GSF superclusters with transverse energy higher than 25 GeV, in the acceptance, out of the gap region and associated to electrons from Drell-Yan events with mass above 40 GeV/c^2 is $(2.06 \pm 0.05)\%$.

6.3 Electron identification and isolation

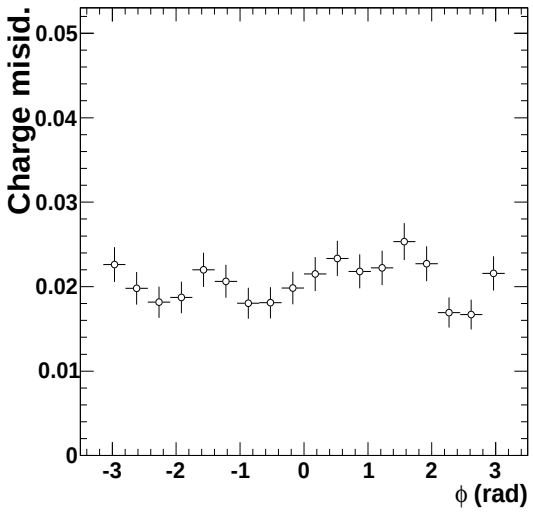
The GSF electron candidate constitutes an object that already fulfills a set of preselection criteria reducing the background rate. However, the electron candidate is a standard object with loose cuts to be used in a wide variety of analysis and preselection criteria are kept as loose as possible. The GSF electron can not thus be considered as the final selected object for the HEEP analysis and tighter cuts have to be imposed. Two categories of cuts, the identification cuts and the isolation cuts, have been defined to keep the signal efficiency as high as possible while rejecting a considerable fraction of the background. Some variables are used at the electron reconstruction level but with tighter cuts. In the following, the electron identification and isolation variable definitions are given and variables are plotted as a function of the GSF supercluster transverse energy, pseudorapidity and azimuthal angle for GSF superclusters matched to the generated electrons. For some variables, the GSF supercluster energy can be used instead of the GSF supercluster transverse energy when it



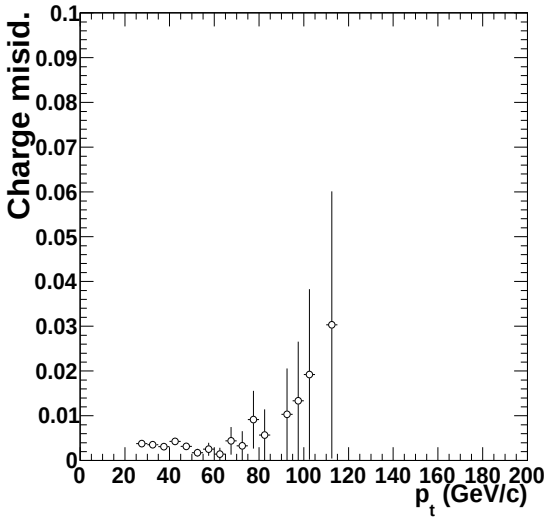
(a)



(b)



(c)



(d)

Figure 6.17: Electron candidate charge for superclusters associated to electrons (red) and positrons (blue) (a). Charge misidentification rate (see text for definition) as a function of the supercluster pseudorapidity (b) and azimuthal angle (c). Charge misidentification rate as a function of the supercluster transverse energy for superclusters with absolute pseudorapidity smaller than 1 (d). The GSF superclusters are required to be in the tracker acceptance, out of the gap region and with transverse energy higher than 25 GeV. (Drell-Yan production with mass above 40 GeV/c^2).

is the most relevant variable.

6.3.1 Electron identification

The electron identification is based mainly on shower shape variables and track-cluster matching variables. The shower shape variables select electrons based on the shape of the showers in the ECAL as it is different for electrons than for hadrons involved in most background processes. Moreover, the shower initiated in the ECAL by an electron is expected to be contained longitudinally and only a small fraction of the electron energy is expected to be collected in the HCAL cells located behind (leakage, see section 3.3.4, [32]).

The track-cluster matching variables require consistency of the information from the energy deposit in the ECAL (supercluster) and the track information (GSF track). These variables are already used at the GSF preselection level.

The definition of the electron identification variables is listed hereafter. For each variable, the corresponding distribution is given separately in the barrel and in the endcaps, as a function of the supercluster transverse energy, pseudorapidity and azimuthal angle. The GSF superclusters are matched to generated electrons/positrons and required to be in the tracker acceptance, out of the gap, with a transverse energy higher than 25 GeV.

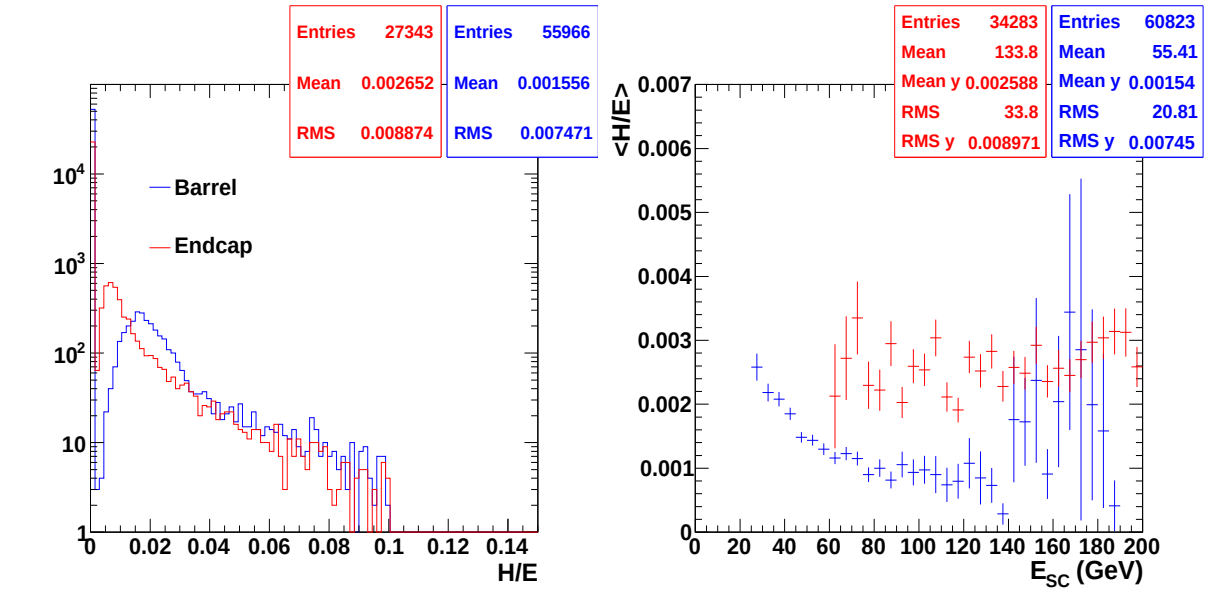
HCAL over ECAL energy deposits: H/E variable

This variable is defined as the ratio of the HCAL energy deposited in "caloTowers" in a cone of radius of 0.15 centered on the supercluster position to the supercluster energy. To reduce the noise contribution, a cut at 0.7 GeV (0.8 GeV, respectively) is put on the HCAL energy deposit in the barrel (in the endcaps, respectively). This variable is expected to be lower in the case of electrons compared to hadrons involved in jets for which the energy is mainly located in the HCAL.

Two main effects contribute to the HCAL energy H: the noise in the HCAL and the energy deposits from the underlying event. The first contribution is expected to be independent of the supercluster energy, pseudorapidity and azimuthal angle. The second contribution is expected to be more important in the forward regions. Small additional contributions from leakage are expected around the intermodule crack regions (in η and ϕ) where a fraction of the initial electron energy is deposited in the HCAL cells placed behind the energy deposit in the ECAL.

Figures 6.18 present the distribution of the variable H/E (a), and the average values of H/E as a function of the supercluster energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps. The H/E value decreases with respect to η in the barrel since the supercluster energy increases at higher pseudorapidities. In the endcaps, the H/E value becomes more important as the underlying event contribution becomes more important. The H/E variable distribution as a function of the supercluster azimuthal angle shows a shape typical of the endcaps where an $x-y$ geometry is present.

The H/E value is required to be smaller than 5% both for the barrel and the endcaps. Figures 6.19 present the efficiency of the H/E cut as a function of the supercluster transverse energy (a), the supercluster pseudorapidity (b) and azimuthal angle (c) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps. The efficiency is high and is constant with respect to the supercluster transverse energy, pseudorapidity and azimuthal angle. Note



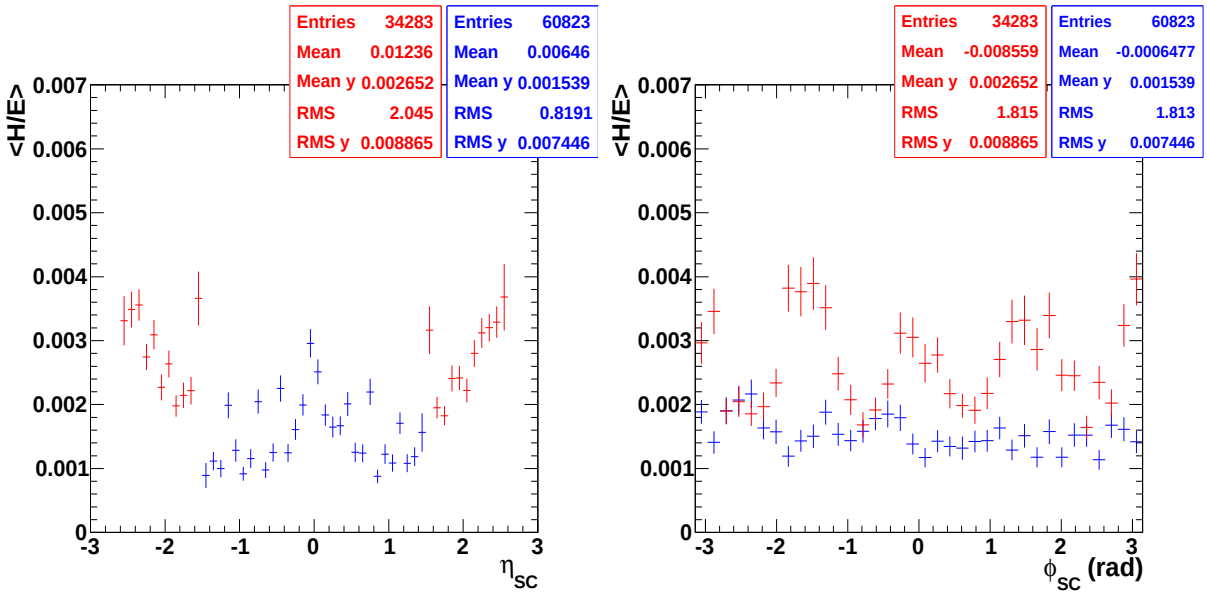
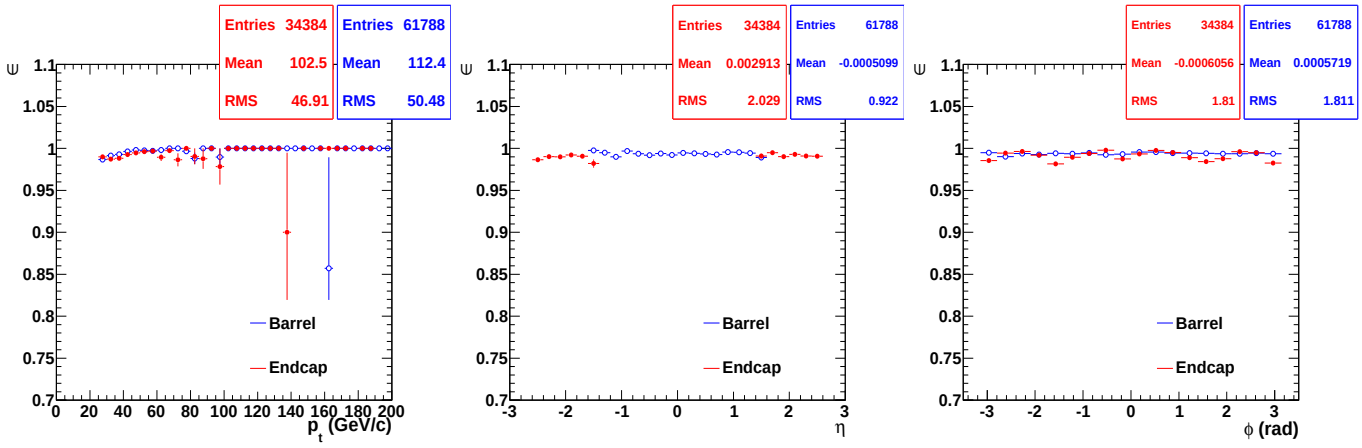


Figure 6.18: Distribution of the H/E variable (a), and average value of H/E as a function of the supercluster energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel (blue) and in the endcaps (red). (Drell-Yan production with mass above 40 GeV/c^2).

that a cut at 10 % on the H/E value is already applied from the preselection criteria of the electron reconstruction (cf. section 6.2.3, eqns 6.2.3).



(a) (b) (c)

Figure 6.19: Efficiency of the H/E cut for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel (blue open circles) and in the endcaps (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).

Track-supercluster pseudorapidity matching: $\Delta\eta_{in}$ variable

This variable is defined as the difference between the supercluster pseudorapidity and the pseudorapidity derived from the track, as measured in the inner tracker layers, extrapolated to the calorimeter surface.

$$\Delta\eta_{in} = \eta_{SC} - \eta_{in} \tag{6.11}$$

The matching between the supercluster position and the track direction is better for electrons than for fake electrons from QCD multijet events.

Figures 6.20 present the distribution of the variable $|\Delta\eta_{in}|$ (a), and the average values of $|\Delta\eta_{in}|$ as a function of the supercluster transverse energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps.

The $\Delta\eta_{in}$ variable is affected by the supercluster transverse energy as shown in figure 6.20(b). The distribution as a function of the pseudorapidity is symmetric with respect to $\eta = 0$ and follows roughly the distribution of the tracker material budget presented in figure 3.7, showing notably larger differences in the endcap. One also sees structures around $\eta_{SC} = 0.45$, 0.8, 1.2 corresponding to the inter module cracks. In these regions, the supercluster reconstruction performance is limited and the determination of the supercluster pseudorapidity comes with limited precision. The behaviour is quite uniform with respect to the azimuthal angle.

The absolute value of $\Delta\eta_{in}$ is required to be smaller than 0.005 in the barrel and 0.007 in the endcaps. Figures 6.21 present the efficiency of the $\Delta\eta_{in}$ cut as a function of the supercluster transverse energy (a), the supercluster pseudorapidity (b) and azimuthal angle (c) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker

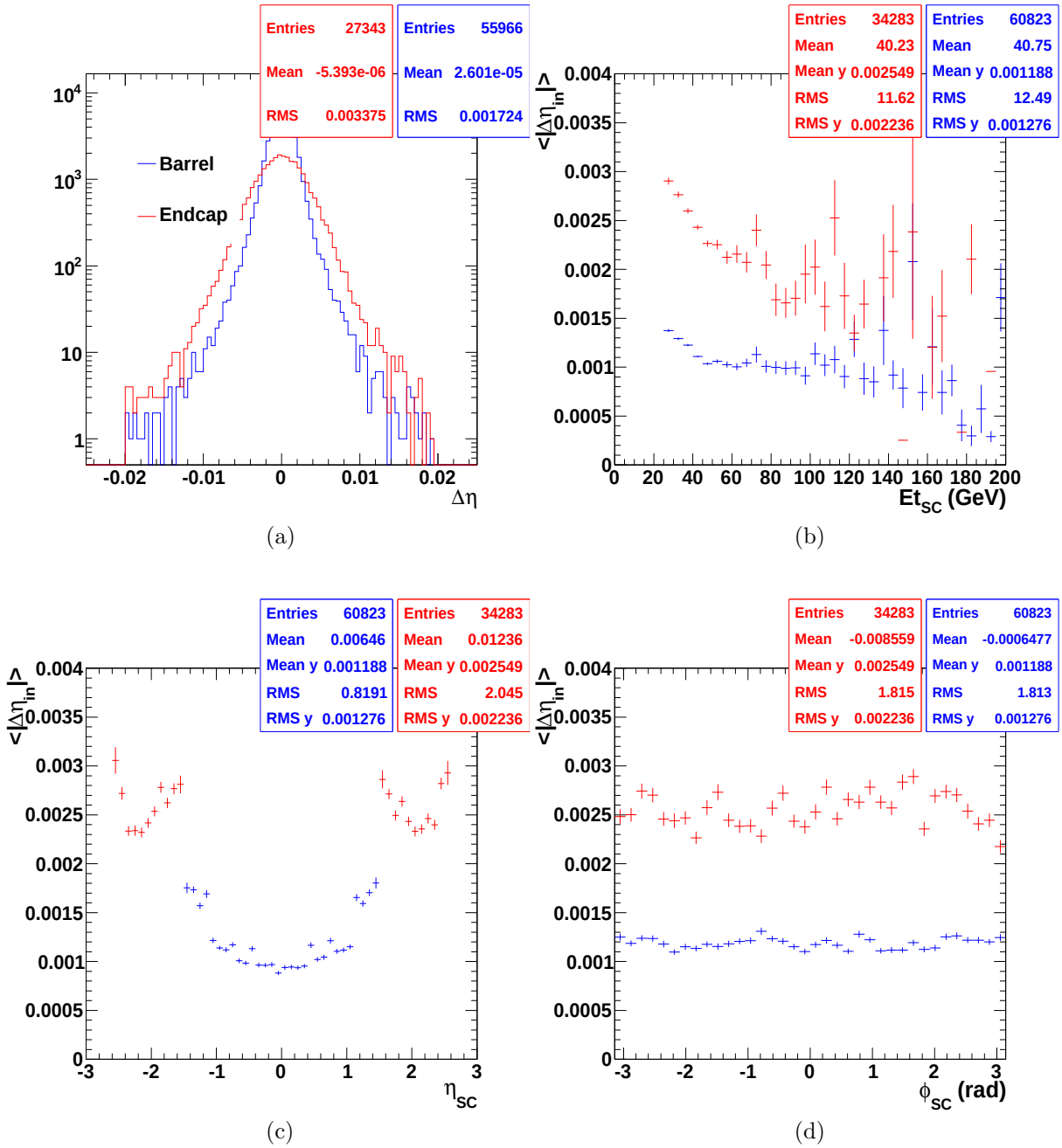
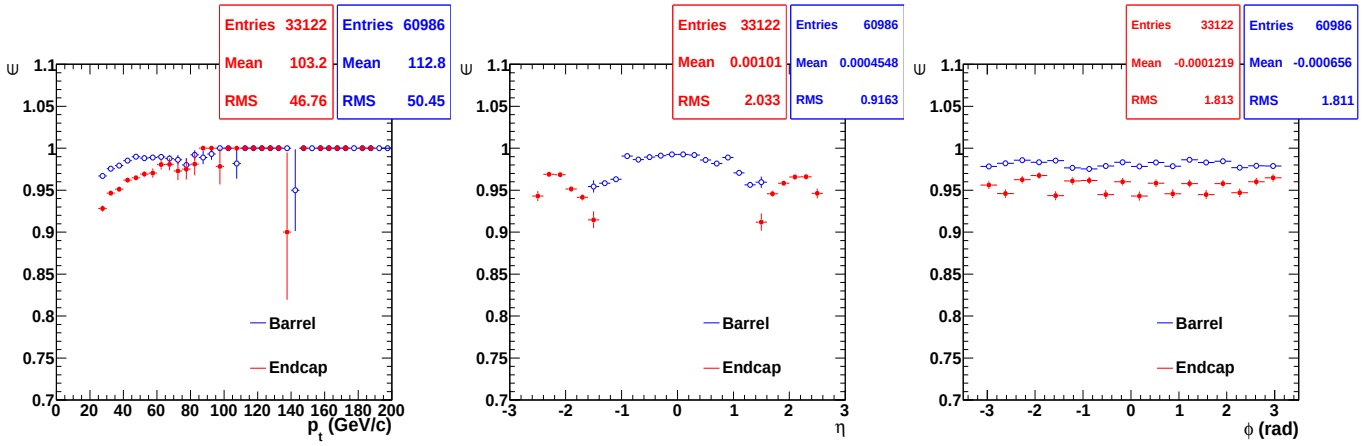


Figure 6.20: Distribution of the $\Delta\eta_{in}$ variable (a), and average value of $\Delta\eta_{in}$ as a function of the supercluster transverse energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel (blue) and in the endcaps (red). (Drell-Yan production with mass above 40 GeV/c^2).

acceptance and out of the gap, separately in the barrel and in the endcaps. The efficiency is uniform at high transverse energy ($E_{t,SC} > 60$ GeV/c) and decreases at low transverse energy (see fig. 6.20(b)). The dependency with respect to the azimuthal angle is due to the tracker material (see fig. 6.20(c)).



(a) (b) (c)

Figure 6.21: Efficiency of the $\Delta\eta_{in}$ cut for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel (blue open circles) and in the endcaps (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).

Track-supercluster azimuthal angle matching: $\Delta\phi_{in}$ variable

This variable is defined as the difference between the supercluster azimuthal angle and the azimuthal angle derived from the track, as measured in the inner tracker layers, extrapolated to the calorimeter.

$$\Delta\phi_{in} = \phi_{SC} - \phi_{in} \tag{6.12}$$

Figures 6.22 present the distribution of the variable $|\Delta\phi_{in}|$ (a), and the average values of $|\Delta\phi_{in}|$ as a function of the supercluster transverse energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps.

As with $\Delta\eta_{in}$, the tracker material budget shows its impact on the $\Delta\phi_{in}$ variable. The dependency of $\Delta\phi_{in}$ with respect to the supercluster transverse energy follows qualitatively the same evolution as with $\Delta\eta_{in}$.

The absolute value of $\Delta\phi_{in}$ is required to be smaller than 0.09 radians, both for the barrel and the endcaps. Figures 6.23 present the efficiency of the $\Delta\phi_{in}$ cut as a function of the supercluster transverse energy (a), the supercluster pseudorapidity (b) and azimuthal angle (c) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps. The behaviour is rather uniform with respect to the supercluster transverse energy, pseudorapidity and azimuthal angle. A slight dependency on the pseudorapidity is observed.

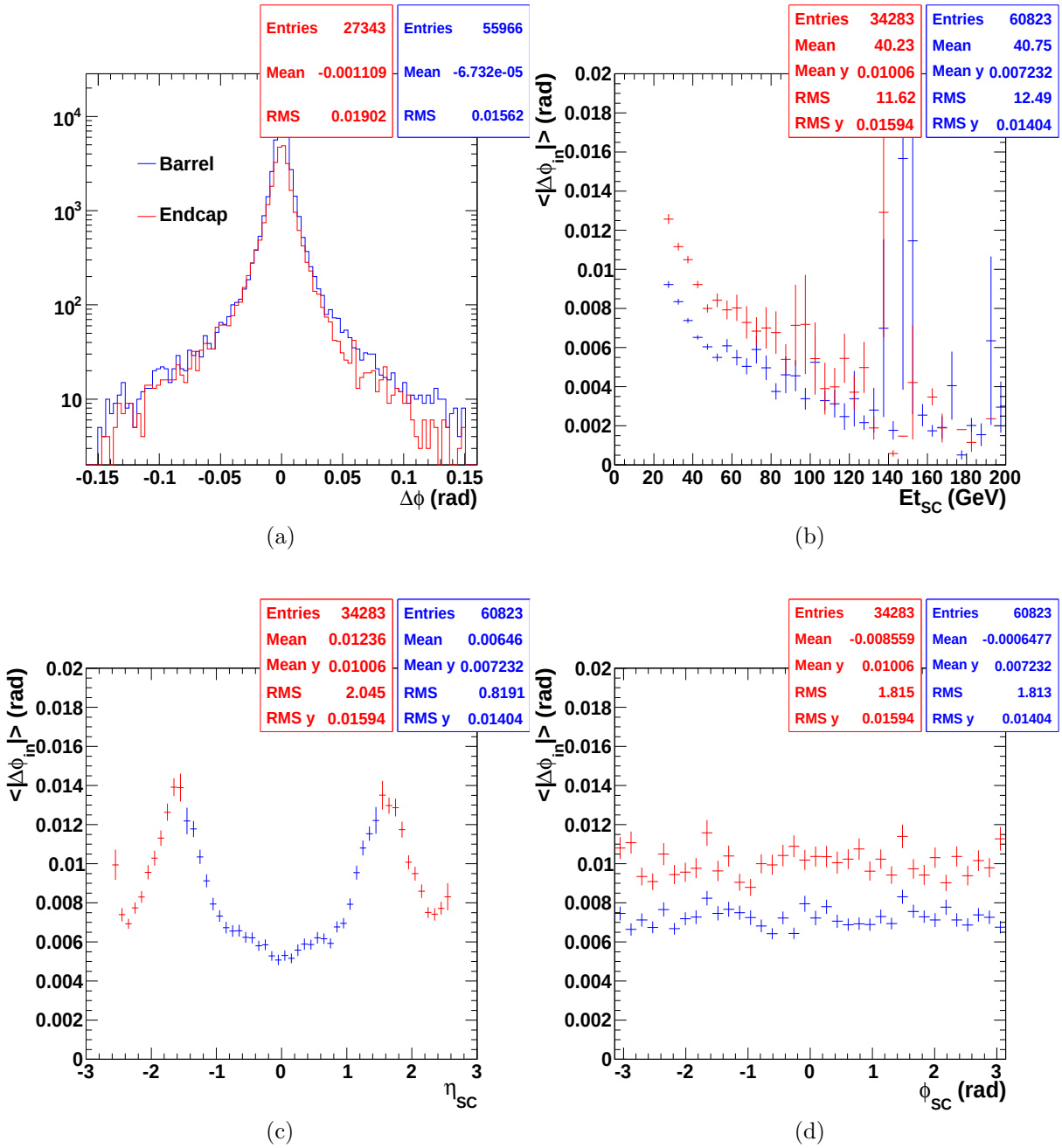
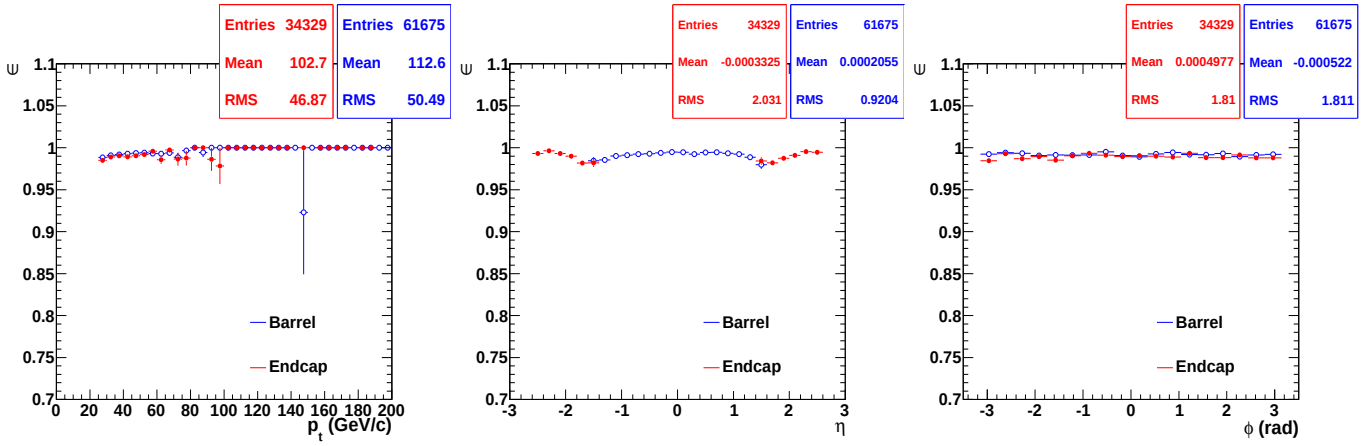


Figure 6.22: Distribution of the $\Delta\phi_{in}$ variable (a), and average value of $\Delta\phi_{in}$ as a function of the supercluster transverse energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel (blue) and in the endcaps (red). (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b) (c)

Figure 6.23: Efficiency of the $\Delta\phi_{in}$ cut for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel (blue open circles) and in the endcaps (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).

Shower shape: $E_{1\times5}/E_{5\times5}$ variable

This shower shape variable relies on the fact that the shape of the shower initiated in the ECAL is different for electrons than for hadrons involved in jets. It is defined as the ratio of the energy inside a 1x5 ($\eta \times \phi$) ECAL crystal matrix centered around the highest energy crystal of the supercluster to the energy inside a 5x5 ($\eta \times \phi$) ECAL crystal matrix centered around the highest energy crystal of the supercluster. The cut on this variable is defined only for the barrel.

The lateral profile of the electromagnetic shower initiated in the ECAL is approximately independent of the electron energy and direction. This variable should therefore not depend on the supercluster energy, pseudorapidity or azimuthal angle. Additional effects such as intermodule cracks, edge effects or different crystal geometric arrangements in the barrel and in the endcaps can however bring dependencies.

Figures 6.24 present the distribution of the variable $E_{1\times5}/E_{5\times5}$ (a), and the average values of $E_{1\times5}/E_{5\times5}$ as a function of the supercluster energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, in the barrel .

This variable does not show any dependency with respect to the pseudorapidity in the barrel. No dependency is observed with respect to the azimuthal angle in the barrel.

Shower shape: $E_{2\times5}/E_{5\times5}$ variable

This shower shape variable is defined as the ratio of the energy inside a 2x5 ($\eta \times \phi$) ECAL crystal matrix centered around the highest energy crystal of the supercluster to the energy inside a 5x5 ($\eta \times \phi$) ECAL crystal matrix centered around the highest energy crystal of the supercluster. The cut on this variable is defined only for the barrel.

Figures 6.25 present the distribution of the variable $E_{2\times5}/E_{5\times5}$ (a), and the average values of $E_{2\times5}/E_{5\times5}$ as a function of the supercluster energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than

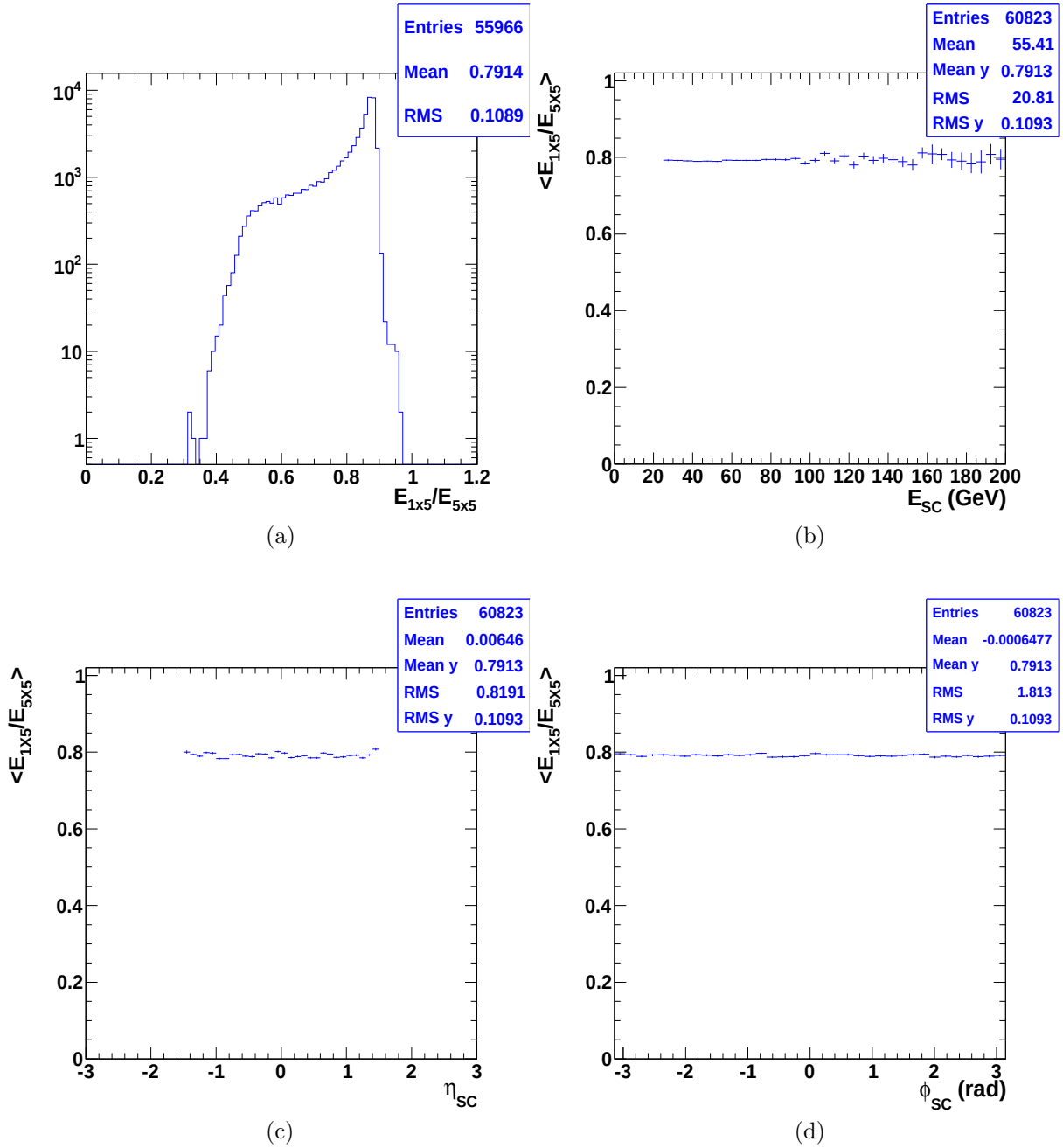


Figure 6.24: Distribution of the $E_{1\times5}/E_{5\times5}$ variable (a), and average value of $E_{1\times5}/E_{5\times5}$ as a function of the supercluster energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, in the barrel only (blue). (Drell-Yan production with mass above 40 GeV/c^2).

25 GeV, in the tracker acceptance and out of the gap, in the barrel. The same conclusions as with the $E_{1\times5}/E_{5\times5}$ variable can be driven.

In the high p_t electron selection, the variables $E_{1\times5}/E_{5\times5}$ and $E_{2\times5}/E_{5\times5}$ are used together (see table 6.4). The value $E_{1\times5}/E_{5\times5}$ is required to be higher than 0.83 or the value $E_{2\times5}/E_{5\times5}$ is required to be higher than 0.94, in the barrel only. No criterion is applied for these two variables, in the endcaps. Figures 6.26 present the efficiency of the combined $E_{1\times5}/E_{5\times5}$-$E_{2\times5}/E_{5\times5}$ cut as a function of the supercluster transverse energy (a), the supercluster pseudorapidity (b) and azimuthal angle (c) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, in the barrel. The behaviour is constant with respect to the supercluster transverse energy, pseudorapidity and azimuthal angle.

Shower shape: $\sigma_{i\eta i\eta}$ variable

This shower shape variable measures the spread in η in units of crystals of the energy deposited inside a 5x5 crystal matrix centered on the highest energy crystal of the supercluster. It is defined by the following formula:

$$\sigma_{i\eta i\eta} = \frac{\sum_i^{5\times5} w_i^2(\eta_i - \eta_{max} - \bar{\eta}_{5x5})^2(\Delta\eta)^2}{\sum_i^{5\times5} w_i}, \tag{6.13}$$

with

$$\bar{\eta_{5x5}} = \frac{\sum_i^{5\times5}(\eta_i - \eta_{max})E_i}{\sum_i^{5\times5} E_i} \tag{6.14}$$

and

$$w_i = 4.7 + \ln(\frac{E_i}{E_{5\times5}}) \tag{6.15}$$

The values w_i, η_i and E_i are the weights, the pseudorapidity and the energy of each crystal in the 5x5 crystal matrix, respectively. The value $\Delta\eta$ is the size of the crystals in units of pseudorapidity and is equal to 0.01745 in the barrel and 0.0447 in the endcaps. The value η_{max} is the pseudorapidity of the highest energy crystal of the supercluster. The cut on this variable is used only in the endcaps.

Figures 6.27 present the distribution of the variable $\sigma_{i\eta i\eta}$ (a), and the average values of $\sigma_{i\eta i\eta}$ as a function of the supercluster energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, in the endcaps. A dependency as a function of the supercluster pseudorapidity is observed in the endcaps. The behaviour with respect to the supercluster azimuthal angle is uniform. A slight decrease is observed in the low energy region (fig. 6.27(b)).

The $\sigma_{i\eta i\eta}$ value is required to be smaller than 0.03, in the endcaps only. No criterion on this variable is applied in the barrel. Figures 6.28 present the efficiency of the $\sigma_{i\eta i\eta}$ cut as a function of the supercluster transverse energy (a), the supercluster pseudorapidity (b) and azimuthal angle (c) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, in the endcaps. The behaviour is uniform with an efficiency decrease in the low transverse energy region.

6.3.2 Electron isolation

In addition to identification variables, variables to quantify the electron isolation are defined. They rely on the assumption that little activity in the calorimeters and the tracker is expected

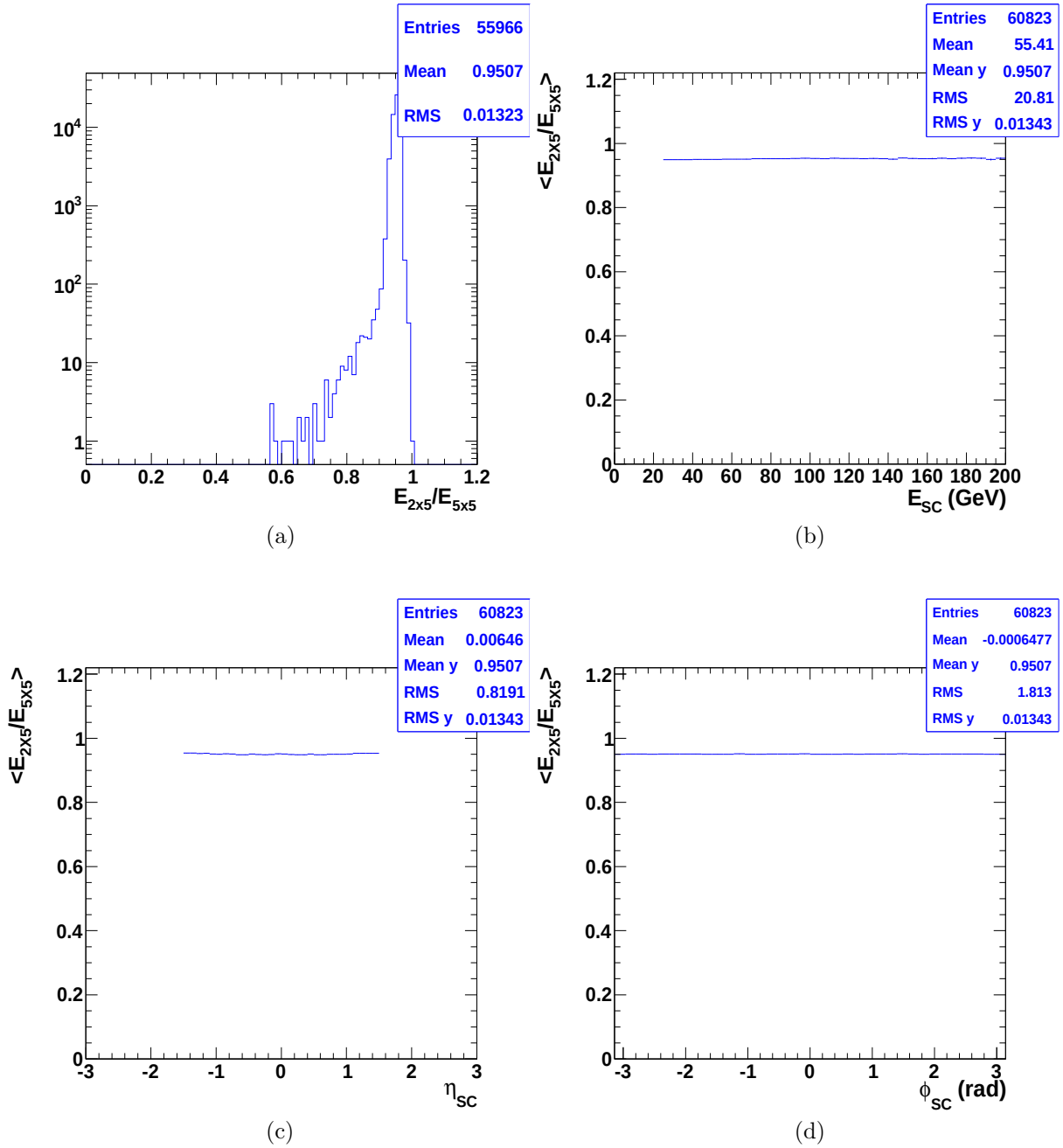
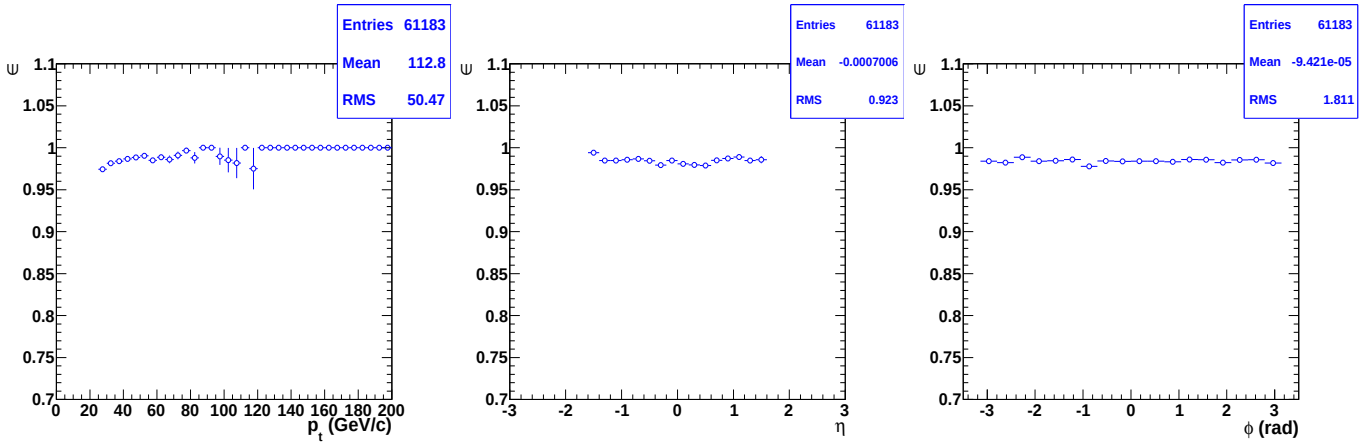


Figure 6.25: Distribution of the $E_{2\times5}/E_{5\times5}$ variable (a), and average value of $E_{2\times5}/E_{5\times5}$ as a function of the supercluster energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, in the barrel only (blue). (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b) (c)

Figure 6.26: Efficiency of the combined $E_{1\times5}/E_{5\times5}$-$E_{2\times5}/E_{5\times5}$ cut for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), in the barrel only (blue open circles). (Drell-Yan production with mass above 40 GeV/c^2).

around the electron direction, contrary to jets, which contain many particles. The tracker activity can be described by the number of tracks around the electron direction or by the sum of the transverse momenta of such tracks. The calorimetry activity can be derived from the amount of energy collected in the ECAL or the HCAL cells around the electron direction. The definitions for the isolation variables are listed hereafter:

ECAL isolation

This isolation definition is based on the transverse energy in all ECAL cells in a cone of radius 0.3 centered on the electron direction in the calorimeter. The cells in a cone of size corresponding to 3 crystals and the cells in an η strip of 3 crystals are not considered. In addition a cut on the absolute energy of each cell to be considered is applied; 0.08 GeV in the barrel and 0.1 GeV in the endcaps.

There are three main contributions to the ECAL energy deposits around the electron direction: the energy deposits due to the underlying event, the fake energy deposits resulting from the noise in the ECAL and finally the energy deposits coming from the electron itself, that would lie outside the cone size of 3 crystals (see [51]).

Figures 6.29 present the distribution of the ECAL isolation variable (a), and the average values of the ECAL isolation variable as a function of the supercluster transverse energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps.

In the dependency as a function of the supercluster pseudorapidity, an effect, related to the tracker material budget, is visible. Topological local effects also appear in the distribution as outliers around $\eta \sim 1.5$ corresponding the barrel-endcaps gap region and a small bump around $\eta \sim 0$ corresponding to the crack region between the two halves of the barrel.

The dependency with respect to the supercluster azimuthal angle is presented in figure 6.29(c) separately for the barrel and the endcaps. The variation is small and the variable is approximately constant as a function of ϕ_{SC}.

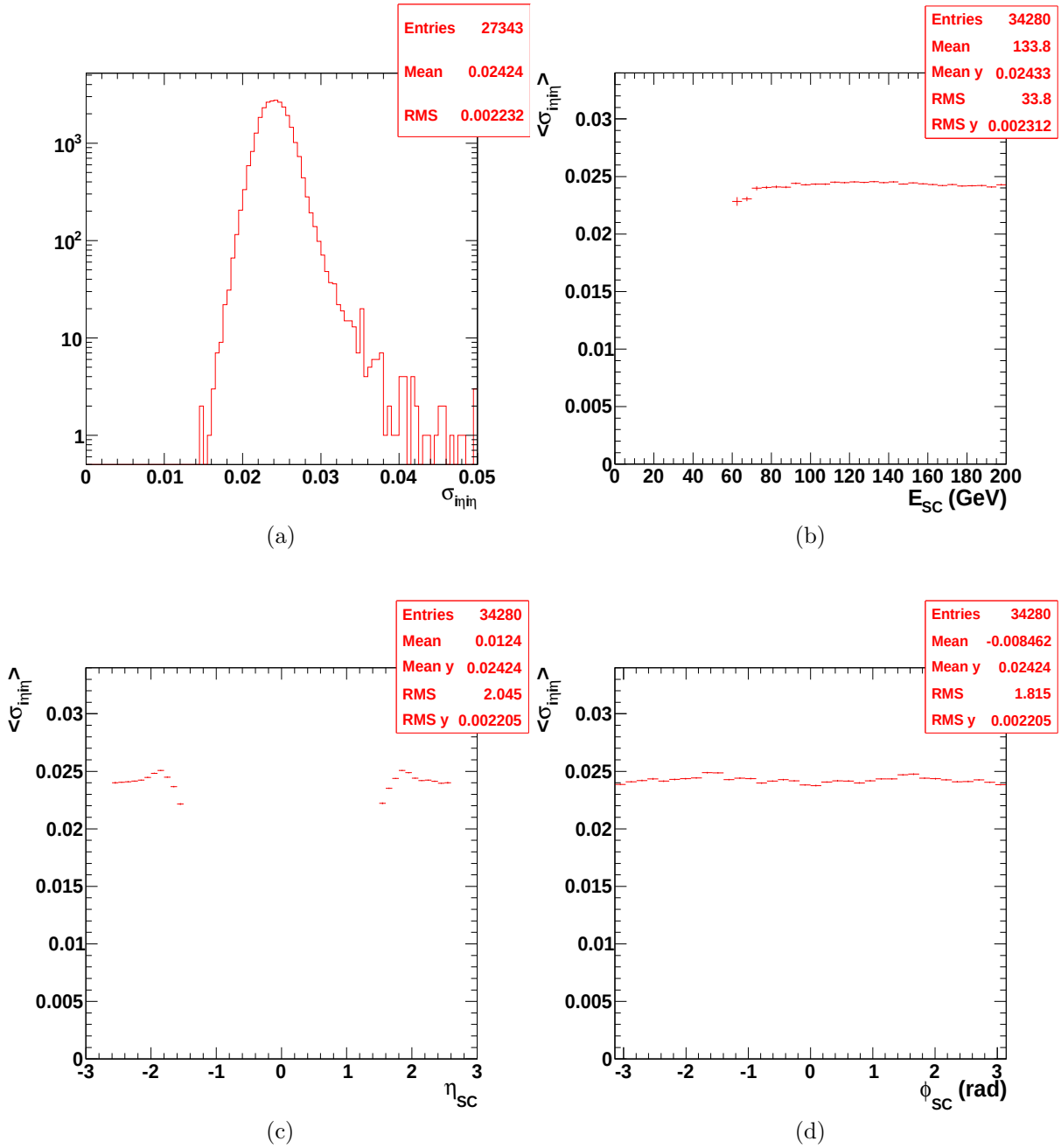
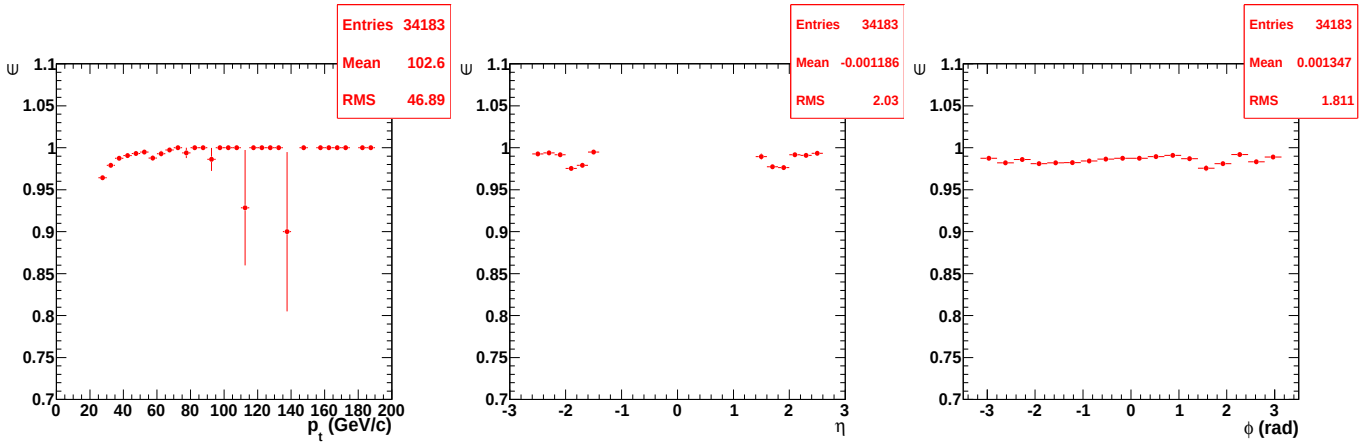


Figure 6.27: Distribution of the $\sigma_{i\eta i\eta}$ variable (a), and average value of $\sigma_{i\eta i\eta}$ as a function of the supercluster energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, in the endcaps only (red). (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b) (c)

Figure 6.28: Efficiency of the $\sigma_{i\eta i\eta}$ cut for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), in the endcaps only (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).

HCAL1 isolation

This isolation definition is based on the hadronic energy of all caloTowers in the first segmentation of the HCAL, within a cone of radius 0.3 centered on the electron position in the calorimeter, excluding caloTowers in a cone of radius 0.15. The relevance of this isolation variable is to reduce the QCD dijet background. In pure signal case, the only contribution expected comes from the noise in the HCAL, energy deposits from particles of the underlying event or electron energy leakage to the HCAL for very high energy electrons or in problematic regions of the ECAL (gap, cracks). Thus, no overall dependency on the supercluster energy is expected for this variable.

Figures 6.30 present the distribution of the HCAL1 isolation variable (a), and the average values of the HCAL1 isolation variable as a function of the supercluster transverse energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps.

In the final HPTE selection, the energy from the HCAL1 isolation is combined with the energy from the ECAL isolation. In the barrel, the ECAL+HCAL1 isolation value is required to be smaller than 2 GeV $+3\%E_t$ where E_t is the transverse energy of electron candidates, expressed in GeV. In the endcaps, two cases are considered. If the electron transverse energy E_t is smaller than 50 GeV, the ECAL+HCAL1 isolation value is required to be smaller than 2.5 GeV. In the other case ($E_t > 50$ GeV), the ECAL+HCAL1 isolation value is required to be smaller than 2.5 GeV + 3% (E_t - 50) where the transvere energy, E_t, is expressed in GeV. Figures 6.31 present the efficiency of the ECAL+HCAL1 isolation cut as a function of the supercluster transverse energy (a), the supercluster pseudorapidity (b) and azimuthal angle (c), respectively, for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately for the barrel and the endcaps.

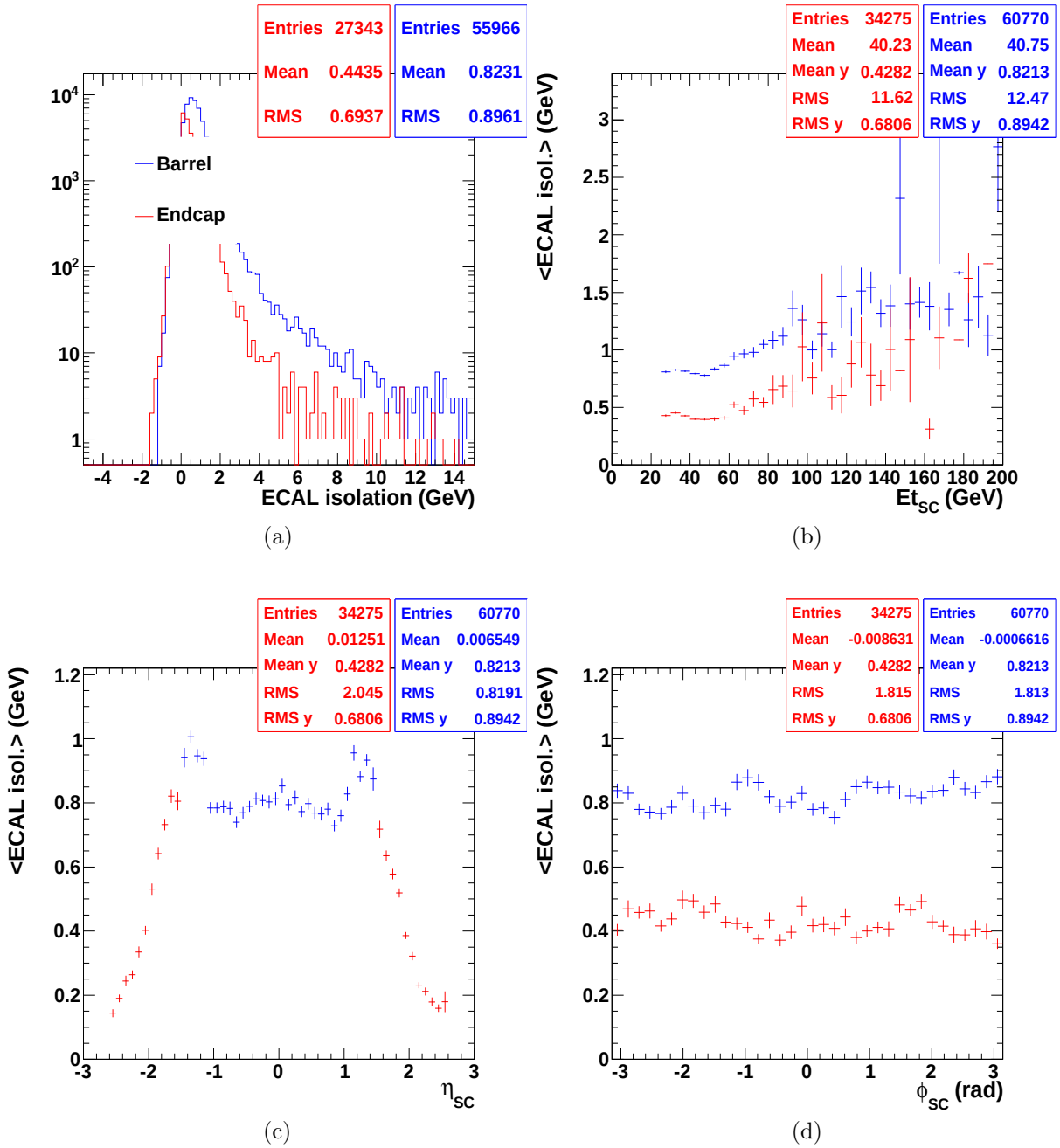


Figure 6.29: Distribution of the ECAL isolation variable (a), and average value of the ECAL isolation variable as a function of the supercluster transverse energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel (blue) and in the endcaps (red). (Drell-Yan production with mass above 40 GeV/c^2).

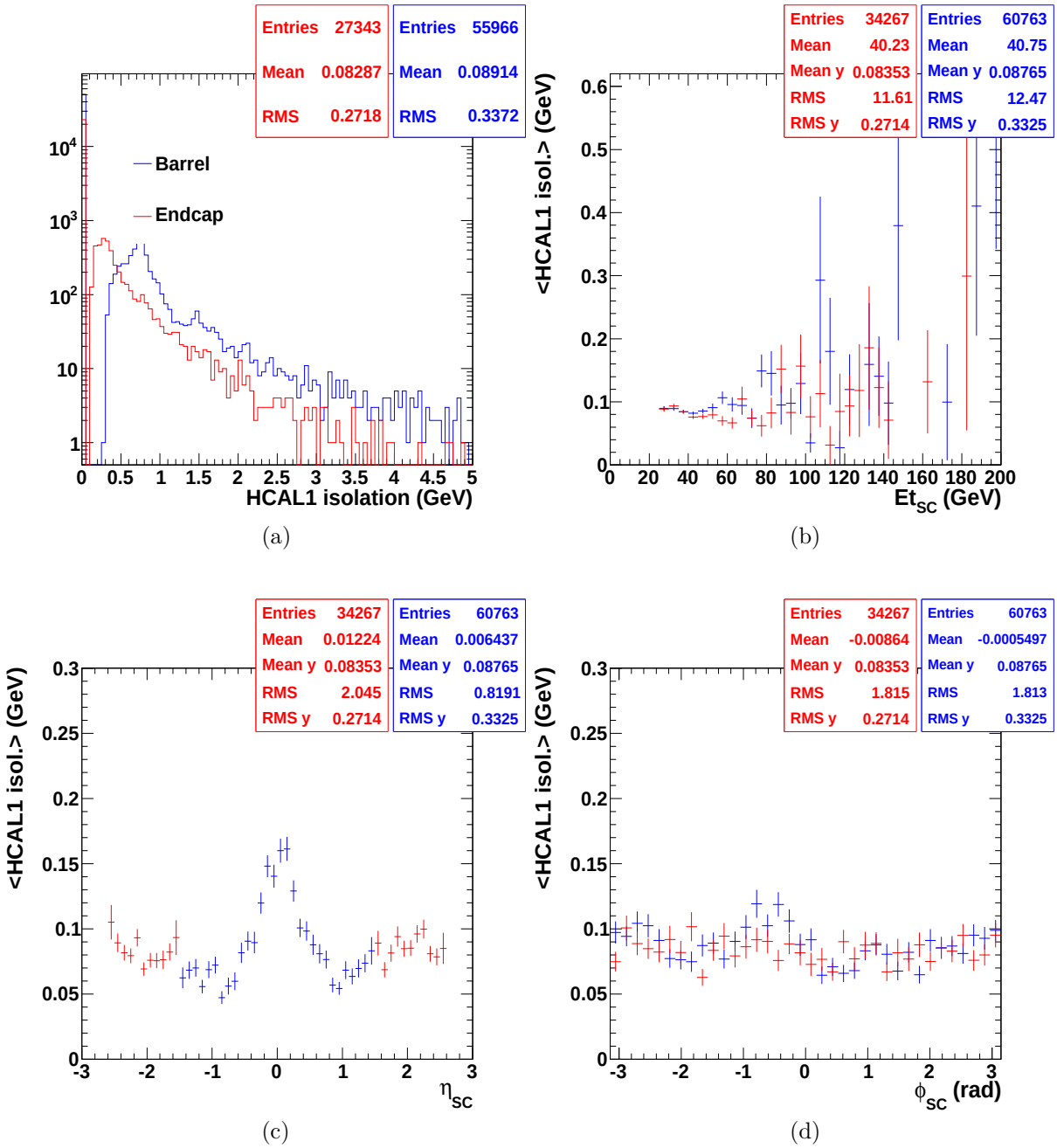
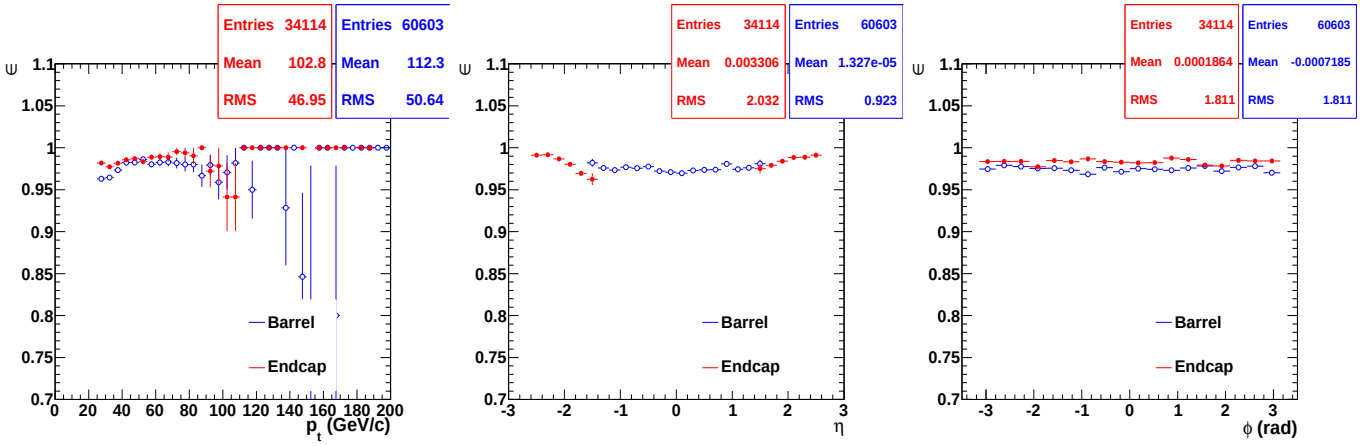


Figure 6.30: Distribution of the HCAL isolation (first segmentation) variable (a), and average value of the HCAL isolation (first segmentation) as a function of the supercluster transverse energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel (blue) and in the endcaps (red). (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b) (c)

Figure 6.31: Efficiency of the ECAL+HCAL1 isolation cut for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel (blue open circles) and in the endcaps (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).

HCAL2 isolation

This isolation definition is based on the hadronic energy of all caloTowers in the second segmentation of the HCAL, in a cone of radius 0.3, centered on the electron position in the calorimeter, excluding caloTowers in a cone of radius 0.15.

Figures 6.32 present the distribution of the HCAL2 isolation variable (a), and the average values of the HCAL2 isolation variable as a function of the supercluster transverse energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, in the endcaps.

The arguments concerning the contribution to this isolation variable are identical to those of the HCAL1 variable. However, one needs to remind the HCAL is segmented only starting from pseudorapidities of $\eta \sim 1.2$ which corresponds to tower number 18. The HCAL2 isolation is rather constant with respect to the transverse energy and the azimuthal angle.

The H/E value is required to be smaller than 0.5 GeV, in the endcaps only. No criterion is applied for this variable in the barrel. Figures 6.33 present the efficiency of the HCAL2 isolation cut as a function of the supercluster transverse energy (a), the supercluster pseudorapidity (b) and azimuthal angle (c) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps. A uniform behaviour is observed with respect to the three kinematic variables.

Track isolation

This isolation is computed as the scalar sum of transverse momenta of all general tracks (not GSF tracks but tracks from the standard tracking algorithm) between two cones of radii 0.04 and 0.3 centered around the electron track direction. Only tracks with transverse momentum above 0.7 GeV/c and with a z_0 distance between -0.2 and 0.2 cm are considered; the z_0 distance being the closest approach distance of the track with respect to the primary vertex.

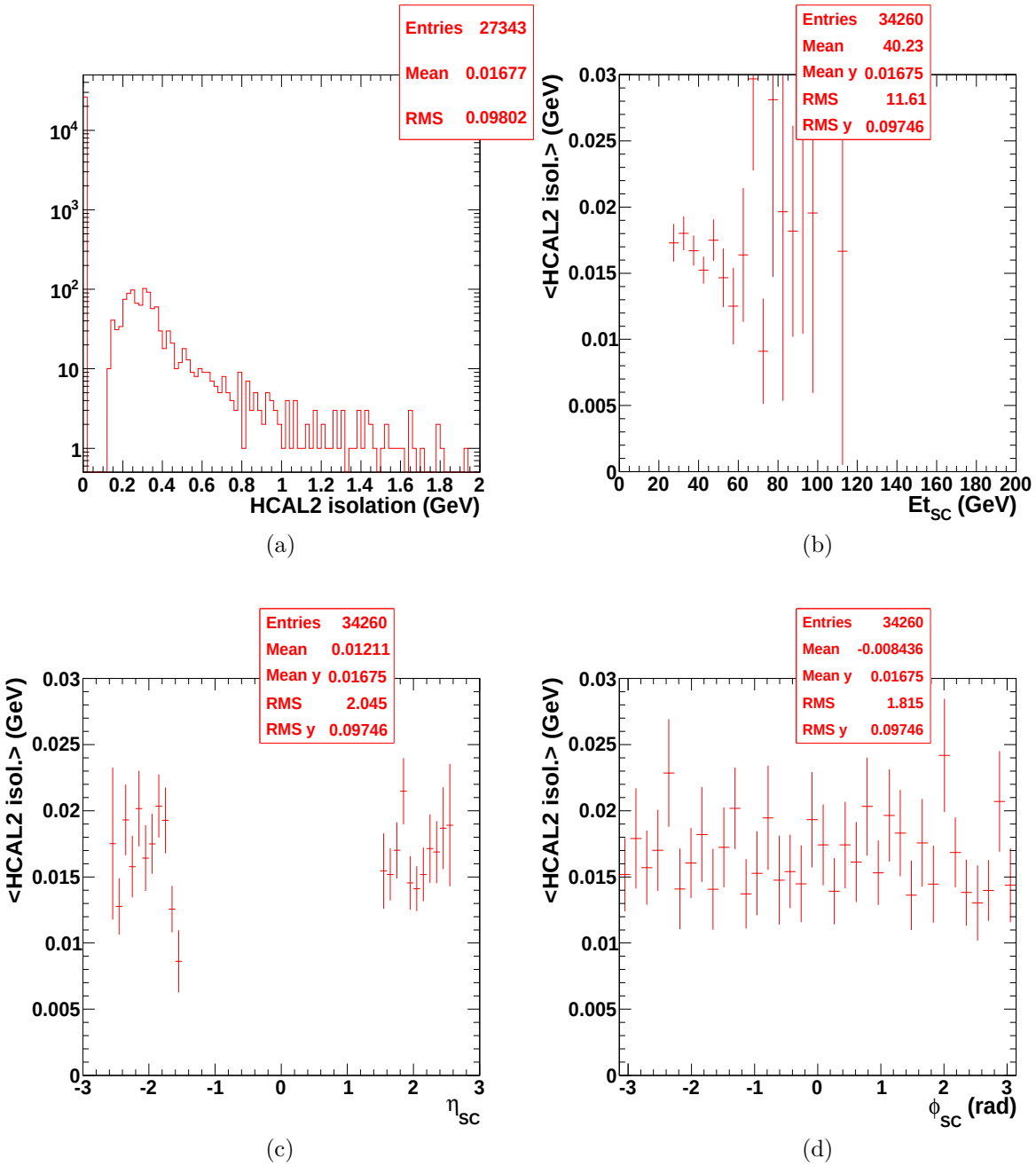
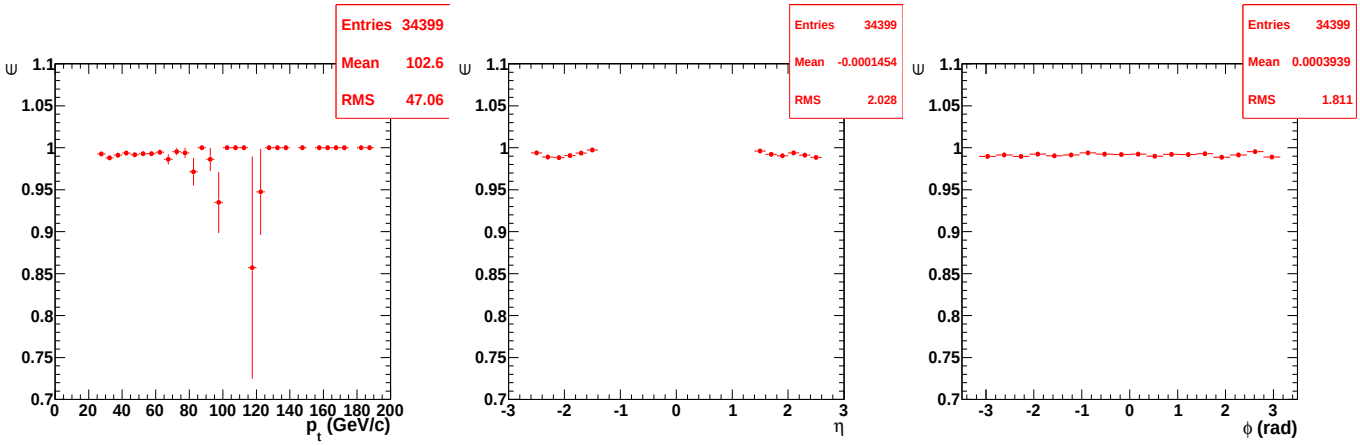


Figure 6.32: Distribution of the HCAL isolation (second segmentation) variable (a), and average value of the HCAL isolation (second segmentation) as a function of the supercluster transverse energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, in the endcaps only (red). (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b) (c)

Figure 6.33: Efficiency of the HCAL isolation (second segmentation) cut for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), in the endcaps only (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).

Figures 6.34 present the distribution of the track isolation variable (a), and the average values of the track isolation variable as a function of the supercluster transverse energy (b), the supercluster pseudorapidity (c) and azimuthal angle (d) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps.

The contribution to this variable is expected from QCD background dijets. In pure Drell-Yan events, it receives contributions from the particles of the underlying event and also from Bremsstrahlung photon conversions collimated to the electron direction. While the underlying event activity in terms of tracks is nearly constant as a function of the supercluster pseudorapidity, it is not the case for photon conversions as the tracker material budget increases from $\eta \sim 0$ to $\eta \sim 1.5$. The tracker isolation is thus expected to increase, which is what is observed in figure 6.34(c). The isolation decreases for $\eta > 1.5$, and at very high pseudorapidities ($|\eta| \gtrsim 2.4$) the acceptance starts to be limited and edge effects start to be significant.

No dependency on the supercluster azimuthal angle or transverse energy is observed. The increase of the track isolation activity at high pseudorapidities has been the subject of a specific study documented in [52] and available in appendix A.

The tracker isolation value is required to be smaller than 7.5 GeV in the barrel and 15 GeV in the endcaps. Figures 6.35 present the efficiency of the track isolation cut as a function of the supercluster transverse energy (a), the supercluster pseudorapidity (b) and azimuthal angle (c) respectively for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel and in the endcaps. A uniform behaviour is observed with respect to supercluster transverse energy, pseudorapidity and azimuthal angle.

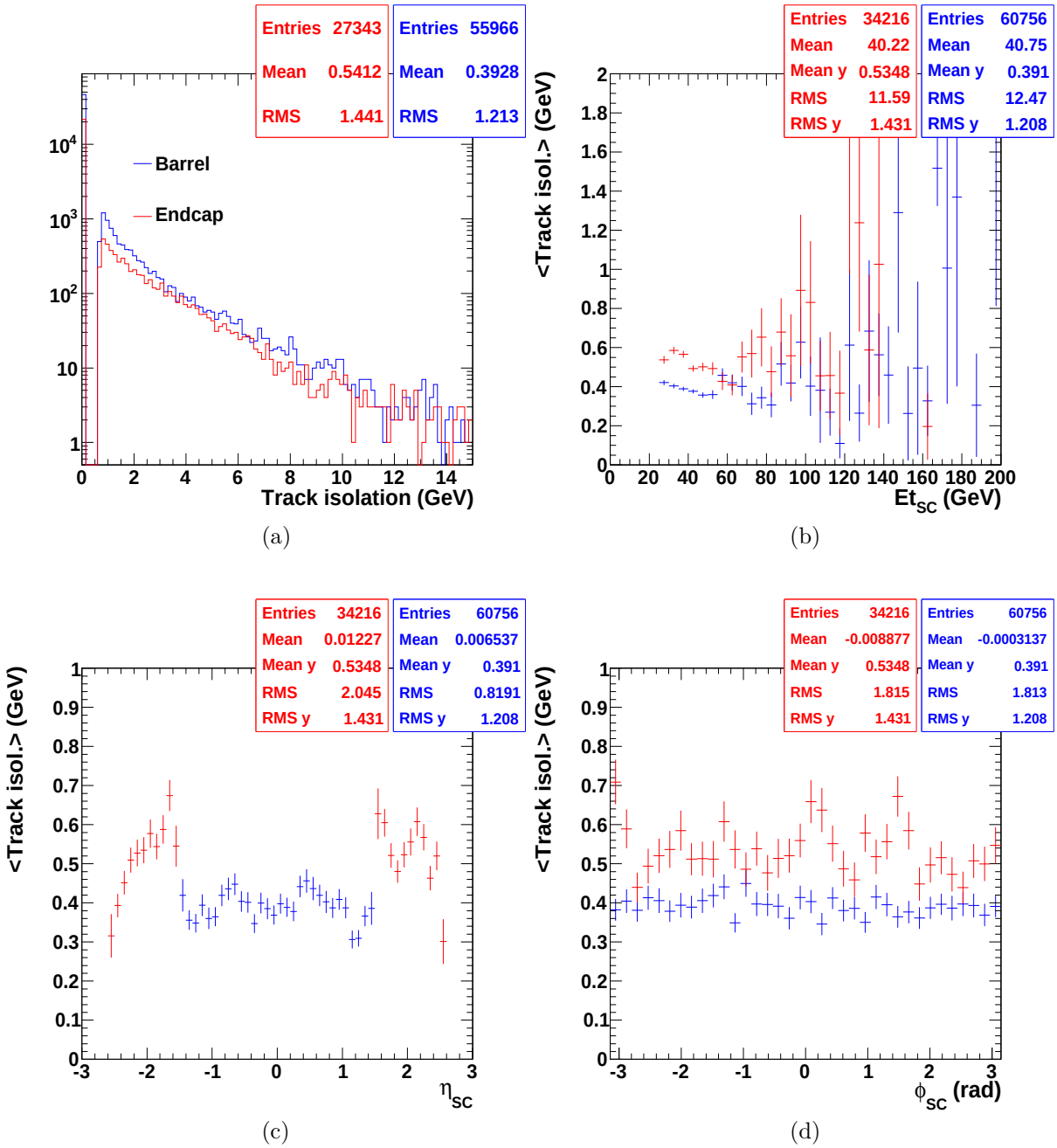
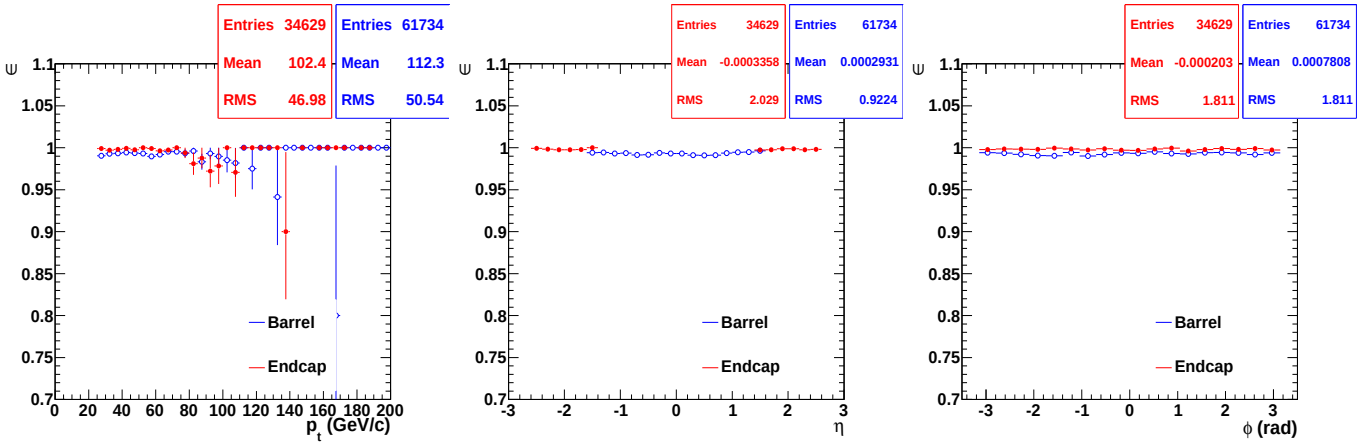


Figure 6.34: Distribution of the Track isolation variable (a), and average value of the Track isolation variable as a function of the supercluster transverse energy (b), pseudorapidity (c) and azimuthal angle (d), for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap, separately in the barrel (blue) and in the endcaps (red). (Drell-Yan production with mass above 40 GeV/c^2).



(a) (b) (c)

Figure 6.35: Efficiency of the Track isolation cut for superclusters with transverse energy higher than 25 GeV, in the tracker acceptance and out of the gap as a function of the supercluster transverse energy (a), pseudorapidity (b) and azimuthal angle (c), separately in the barrel (blue open circles) and in the endcaps (red full circles). (Drell-Yan production with mass above 40 GeV/c^2).

6.3.3 Additional criteria

In addition to these two categories of criteria, a cut is imposed on the electron transverse energy to reduce the huge amount of background expected at low regime. This quantity is defined as the product of the energy from the supercluster and the sinus of the track angle:

$$E_t = E_{SC} \times \sin(\theta_{track}) \tag{6.16}$$

The supercluster pseudorapidity is used to separate between the barrel and the endcaps. Eventually, the GSF electron candidates are required to be "ecaldriven".

6.4 Summary of HPTE criteria

Table 6.4 presents the list of HPTE criteria used for the HEEP analysis, separately for the barrel and the endcaps. The two criteria categories are separated as two blocks.

The efficiencies are provided individually for each variable of the high p_t electron selection for three Drell-Yan samples at masses above 40, 120 and 200 GeV/c^2, separately in the barrel (table 6.5) and in the endcaps (table 6.6). The efficiencies are, in general, higher than 98%, the lowest efficiencies come from the $\Delta\eta$ variable in the endcaps. In addition, the total HPTE efficiency and the global efficiency which accounts for the GSF reconstruction efficiency are given. They are observed to increase with respect to the mass.

6.5 HEEP event selection

6.6 Efficiencies from Monte Carlo

The HEEP event selection consists in asking at least two GSF electrons, in the event, that both fulfill the high p_t electron selection. The HEEP efficiency is defined as the ratio N_{HEEP}/N_{kine} where N_{HEEP} is the number of events where the GSF electrons, associated

Variable	Barrel	Endcaps
E_t	> 25 GeV	> 25 GeV
$\|\eta_{SC}\|$	< 1.442	$1.560 < \|\eta_{SC}\| < 2.5$
isEcalDriven	=1	=1
H/E	< 0.05	< 0.05
$\Delta\eta_{in}$	< 0.005	< 0.007
$\Delta\phi_{in}$	< 0.09	< 0.09
$E_{2\times5}/E_{5\times5}$	> 0.94 or $E_{1\times5}/E_{5\times5} > 0.83$	not applied
$\sigma_{i\eta i\eta}$	not applied	< 0.03
isol. ECAL+HCAL1	< 2+0.03 Et (GeV)	<2.5 (Et < 50 GeV) < 2.5+0.03 (Et-50) (Et > 50 GeV)
isol. HCAL2	not applied	< 0.5 GeV
isol. Tracker	< 7.5 GeV/c	< 15 GeV/c

Table 6.4: HPTE selection criteria for the barrel and the endcaps.

HPTE selection cuts (single efficiencies)	M > 40 GeV	M > 120 GeV	M > 200 GeV
H/E	0.994	0.993	0.992
$E_{1\times5}/E_{5\times5}$-$E_{2\times5}/E_{5\times5}$	0.984	0.986	0.987
$\sigma_{i\eta i\eta}$	not applied	not applied	not applied
$\Delta\eta_{in}$	0.981	0.984	0.986
$\Delta\phi_{in}$	0.992	0.994	0.996
isol. ECAL+HCAL1	0.975	0.974	0.974
isol. HCAL2	not applied	not applied	not applied
isol. Tracker	0.993	0.991	0.988
HPTE	0.930	0.935	0.939

Table 6.5: HPTE selection single efficiencies for electrons from Drell-Yan process with various masses, emitted in the barrel.

HPTE selection cuts (single efficiencies)	M > 40 GeV	M > 120 GeV	M > 200 GeV
H/E	0.991	0.990	0.990
$E_{1\times5}/E_{5\times5}$-$E_{2\times5}/E_{5\times5}$	not applied	not applied	not applied
$\sigma_{i\eta i\eta}$	0.985	0.991	0.994
$\Delta\eta_{in}$	0.955	0.971	0.979
$\Delta\phi_{in}$	0.989	0.991	0.993
isol. ECAL+HCAL1	0.983	0.980	0.981
isol. HCAL2	0.991	0.990	0.989
isol. Tracker	0.998	0.997	0.996
HPTE	0.907	0.924	0.935

Table 6.6: HPTE selection single efficiencies for electrons from Drell-Yan process with various masses, emitted in the endcaps.

to the generated electrons, pass the high p_t electron selection given in table 6.4 and N_{kine} is the number of events where the two GSF electrons associated to the generated electrons are emitted in the acceptance, out of the gap and with transverse energy higher than 25 GeV.

The HEEP efficiency is presented in figure 6.36 as a function of the invariant mass of the two electron candidates for all events (black full circles), for events where both electron candidates are in the barrel (blue open circles), events where both electron candidates lie in the endcaps (pink downward triangles) and for events where one electron candidate is in the barrel and the other in the endcaps (red upward triangles). The invariant mass is computed using the supercluster energies, track pseudorapidities and azimuthal angless.

The HEEP efficiency is observed to increase with respect to the invariant mass of the GSF pair. The overall number is 87.08%.

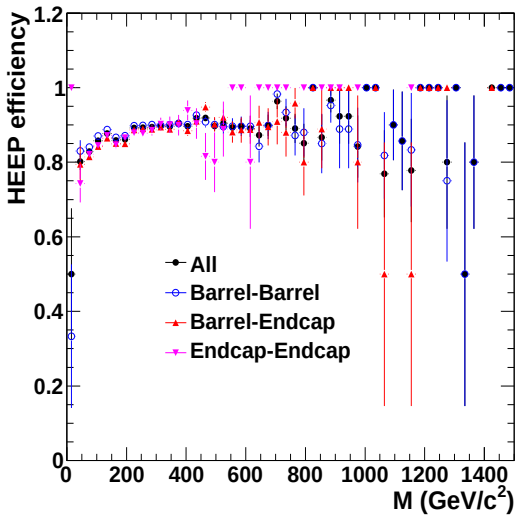


Figure 6.36: HEEP efficiency (see text for definition) as a function of the GSF pair invariant mass for all events (black full circles), events where both electron candidates lie in the barrel (blue open circles), events where both electron candidates lie in the endcaps (pink downward triangles) and for events where one electron candidate is in the barrel and the other in the endcaps (red upward triangles). (Drell-Yan production with mass above 40 GeV/c^2).

Chapter 7

Data Quality Monitoring at $\sqrt{s} =$ 900 and 2360 GeV

First collisions in the Large Hadron Collider happened on November 23rd 2009. They were performed with beams of energy of 450 GeV from the injection chain, giving a centre of mass energy of 900 GeV. They were closely followed by collisions with 2.36 TeV energy in the centre of mass on November 30th, breaking the world record. Even with limited integrated luminosity and energy, data from these collisions were essential as they serve as a basis to check the trigger, the detector response and calibration as well as the energy corrections and the alignment of the different parts of the detector using real collision data information [1]. In that view, a data quality monitoring tool was developed by CMS [53] in order to perform a fast check of first online and offline data as they are provided.

7.1 Data Quality Monitoring for high energy electrons

A similar DQM tool, specific to the HEEP analysis requirements was developed by the HEEP group well before first collision data [54]. It focuses on electron reconstruction as defined in section 6.2(supercluster—track reconstruction and linking) as well as Drell-Yan events kinematics with a specific emphasis on high energy electron events. It emphasizes on three key points:

- Spot detector problems (noisy channels, miscalibration issues, ...) and react quickly
- Compare data and Monte Carlo expectations for Monte Carlo tuning (beam spot position, ...)
- Search for new physics behaviour compared to Standard Model

The DQM consists of a set of plots showing distributions of predefined variables, where the distributions from data are superimposed to the distributions from Monte Carlo for comparison. The Monte Carlo simulates events with conditions similar to what we expect to see in coming data. Any difference between Monte Carlo and data spots either new unexpected physics phenomena or problems related to the detector/reconstruction software. The list of predefined variables divides into three main categories, in ascending order of complexity:

[1] A first check of the detector response and in particular the tracker and the detector alignment were performed earlier using cosmic muon data.

- The supercluster level: This level is intended to check behaviour of lowest possible level objects, to distinguish for instance between pions and electrons/photons. It gathers a list of variables associated to the supercluster:
  - Supercluster energy variables: Uncorrected and corrected ("raw") energy and transverse energy, pseudorapidity η and azimuthal angle ϕ.
  - Supercluster shower shape variables: superclusters widths in η and ϕ, $\sigma_{i\eta i\eta}$ and $\sigma_{i\phi i\phi}$. Also additional variables such as $E_{1\times5}$ and $E_{2\times5}$ which correspond to the energies collected in one (resp. two) domino(es) of five crystals aligned in η and centered around the maximal energy crystal. The H/E variable is the ratio of the energy collected in the HCAL cells located directly behind the supercluster position over the supercluster energy.
- The electron candidate (GSF electron) level: This level is intended to check higher level objects such as GSF electrons which combine information from energy deposits in ECAL (superclusters) and tracks in the tracker.
  - GSF basic variables: Energy and transverse energy, pseudorapidity (η) and azimuthal angle ϕ.
  - Supercluster-track information consistency: $\Delta\eta$, $\Delta\phi$ and E/p.
  - GSF shower shape variables: $\sigma_{i\eta i\eta}$, $\sigma_{i\phi i\phi}$.
  - GSF isolation variables: ECAL, HCAL, tracker isolation and H/E
- The HEEP event level: The latter is intended for events where, at least, two electron candidates with minimal transverse energy are present. Such events are obviously to direct interest to the HEEP group and a list of pertinent variables was predefined:
  - General event variables: Primary vertex position (x,y,z), number of muons, tracks, electron candidates, vertices and missing transverse energy in the event.
  - Variables for the highest energy superclusters in the event: energy E, transverse energy (E_t) and its components (E_x,E_y,E_z). Pseudorapidity (η), azimuthal angle ϕ. Invariant mass (M_{inv}), angular separation $\Delta\phi$ of two highest energy superclusters.
  - Variables for the highest energy GSF in the event: energy E, transverse energy (E_t) and its components (E_x,E_y,E_z). Pseudorapidity (η), azimuthal angle ϕ.
  - HEEP event variables: Invariant mass (M_{inv}) and angular separation $\Delta\phi$ for two highest energy superclusters.

7.2 Data from 900 GeV and 2.36 TeV collisions

The HEEP DQM tool was first applied on data from 900 GeV and 2.36 TeV centre of mass energy collisions. For each centre of mass energy, Monte Carlo samples were produced with conditions similar to those expected in collisions. The DQM tool was also applied on these Monte Carlo samples. Distributions from real collision data and from Monte Carlo samples were superimposed and compared. For commodity, the number of events in the Monte Carlo was scaled to the number of events from real data. Any difference shows either new unexpected phenomena or problems related to the detector/reconstruction software and was scrutinized for further understanding.

In the following, the DQM tool applied on first data is presented for the three levels considered (supercluster level, GSF electron candidate level and HEEP event level). However, due to limited statistics, some p_t thresholds were lowered, especially at the HEEP event level (see section 7.5).

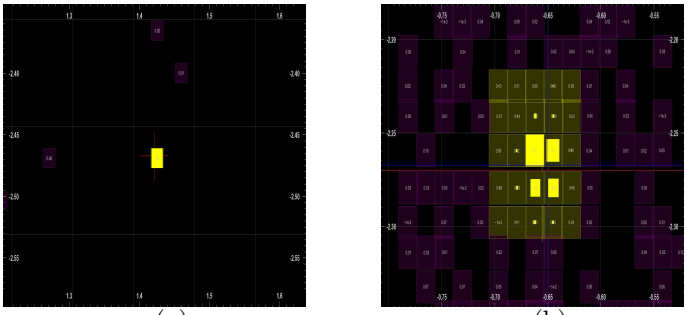
7.3 DQM for supercluster variables

The energy variables distributions at the supercluster level are presented in figures 7.2 separately in the barrel (a) and in the endcaps (b) for 900 GeV collision data. Dotted and solid histograms correspond to collision data and Monte Carlo simulation respectively. Error bars are given for collision data. While, for the endcaps, the distributions from data have very good agreement with the Monte Carlo distributions, in the barrel, however, some discrepancies are clearly visible in the energy distributions (energy, rawenergy, energy1x5, energy2x5, energy5x5) where the tails are more important in the data distributions.

Similar discrepancies are visible in figures 7.3(a) and (b) where distributions for shower shape at the supercluster level are presented, separately for the barrel and the endcaps, for 900 GeV collision data. Spikes clearly appear at values equal to zero for the data distributions of $\sigma_{\eta\eta}$, $\sigma_{\phi\phi}$, $\sigma_{i\eta i\eta}$ and $\sigma_{i\phi i\phi}$. Once again, this effect appears only in the barrel, the endcap distributions agree reasonably well.

These discrepancies are due to relatively high energy deposits in the barrel contained in a single crystal which are not expected to be compatible with energy deposits from electrons or photons that deposit their energy in several crystals (usually a 5x5 crystal matrix) or even from pions.

Figure 7.1(a) presents an η—ϕ view of the fake energy deposit in the electromagnetic calorimeter barrel with individual crystal energies in the case of so-called "spiky" events. The horizontal and vertical axis correspond to η and ϕ variables respectively. As a comparison, the energy deposit for a true electron coming from a simulated Z event is shown in figure 7.1(b).



(a) (b)

Figure 7.1: An η—ϕ view of the energy deposit with individual crystal energies for "hot channels" in 900 GeV collision data (a) and for a true electron from a simulated Z event (b).

A specific task force [55] in the CMS community has been formed to investigate on the origins of such "spikes" effects. The present understanding includes the possibility of neutral and charged hadrons (primary or secondary) depositing energy directly in the avalanche photo diodes active layers of the ECAL readout and thus creating a fake signal. These effects are characterised by a different pulse shape and timing response compared to real electrons as there is no scintillation process in the crystals in these cases.

Such effects were taken into account in the HEEP DQM tool and a specific criterion was imposed to veto on such fake energy deposits, using the quantity $E1/E9$ where $E1$

represents the energy deposited in the highest energy crystal of the supercluster and $E9$ is the energy deposited in a 3x3 crystal matrix centered around the highest energy crystal in the supercluster. For a normal electron deposit, this quantity is not expected to be more than 90%; which defines the threshold to identify spike fake energy deposits. Figure 7.4 shows the energy variables (a) and shower shape variables (b) distributions, at the supercluster level, in the barrel, after removing the spikes. One sees that the distributions are now well in agreement.

7.4 DQM for GSF variables

Similar plots were performed for the electron candidate variables, running on data from 900 GeV and 2.36 TeV collisions. For information, only the plots for the barrel are shown, in figure 7.5. The electron candidates with fake energy deposits (spike events) are removed from data and shown separately in red. The distributions from data and Monte Carlo simulation agree well.

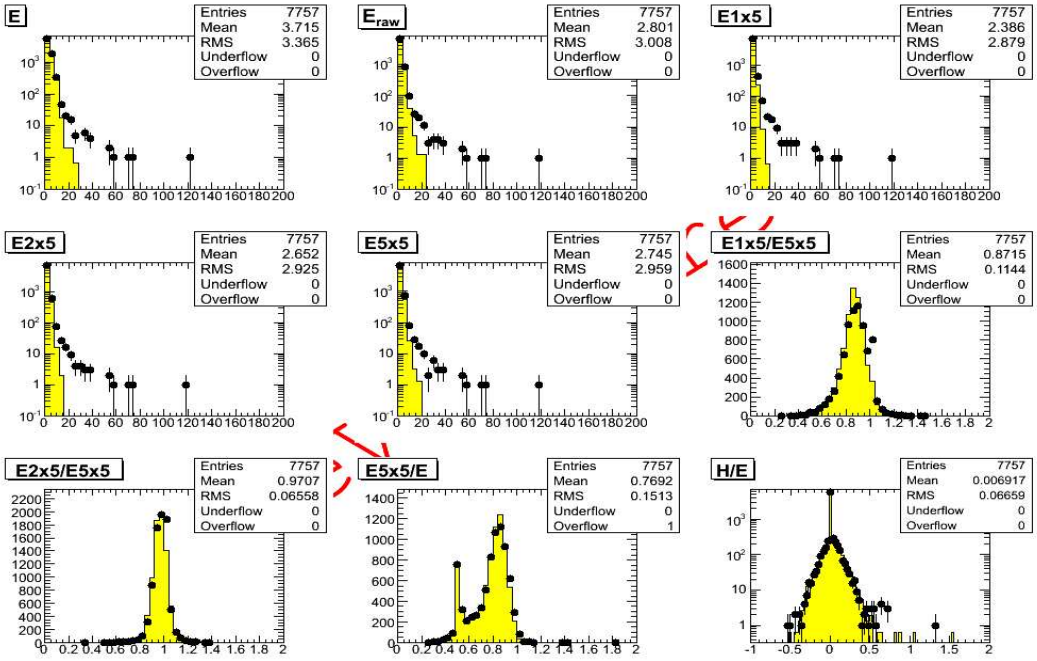
7.5 DQM for HEEP event variables

Monitoring the data quality at the level of the high energy electron pair events is of primary motivation for the HEEP group, especially at startup. Such events are characterized, in the DQM tool, by the presence of at least two electron candidates with minimum transverse momentum. However, due to limited integrated luminosity, very small statistics was available after selection of such events for the 900 GeV and 2.36 TeV data-taking. The selection criteria were thus loosened to the requirement of events with at least one or two superclusters (two separate streams) with transverse energy higher than 4 GeV. The latter requirement will be gradually tightened, as integrated luminosity increases, to finally get to the original selection of two high p_t electron candidates.

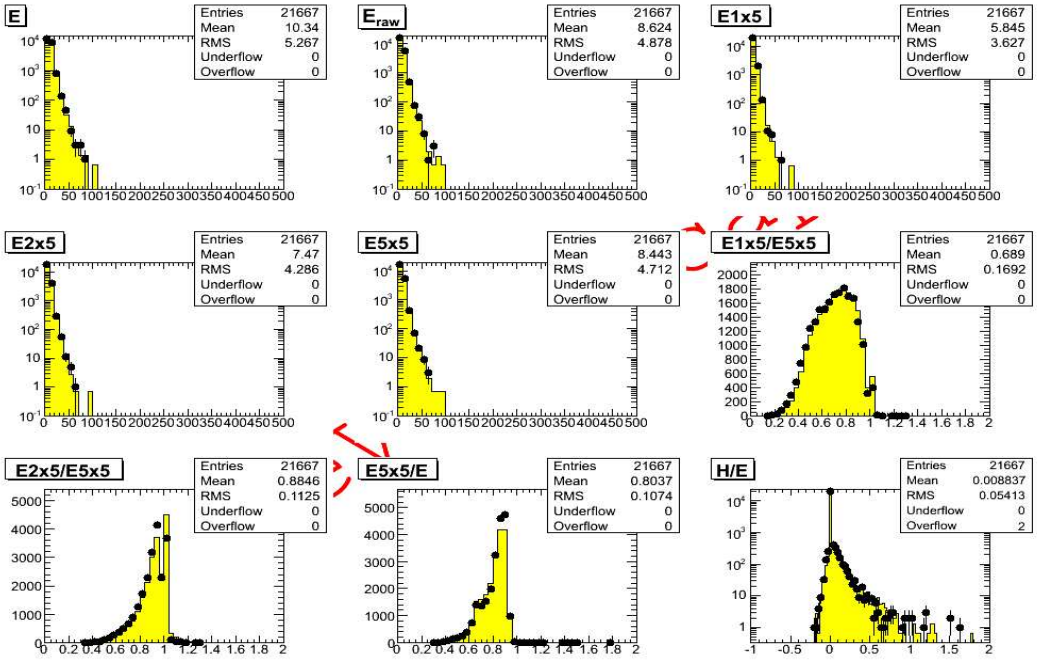
Figure 7.6 presents the supercluster variables for the highest E_t supercluster (a) and the second highest E_t supercluster (b), for events with at least two superclusters with transverse energy higher than 4 GeV from 900 GeV collision data. In addition, the invariant mass and the angular separation($\Delta\phi$) are shown for the two highest E_t superclusters in the event in figures 7.6 (b). Blue triangles correspond to "spiky" fake energy deposit events (i.e. events with at least one supercluster with $E1/E9$ higher than 90%) and are clearly present in the barrel only. These fake energy deposit events were removed from the data (dotted points). Even with a loosened event selection, the statistics is still very low for exact comparison. However, the distributions from Monte Carlo and data agree reasonably well within the error bars.

7.6 The CMS Event visualization tool

In addition to the HEEP DQM tool, a visualization tool [56] was developed by the CMS community to display events recorded inside CMS and its geometry. It is composed of several window applications which present each a specific view of the detector (ρ—ϕ, ρ—z and 3D view). The energy deposits in the calorimetry (ECAL/HCAL) are made visible through lego bars with height proportional to the energy deposit and the tracker and muon tracks are symbolized by curved trajectories as reconstructed by the tracking algorithm. The high level objects (electrons, muons, photons, jets) are also individually represented through their energy deposit and/or track. In addition, the event display provides information about high

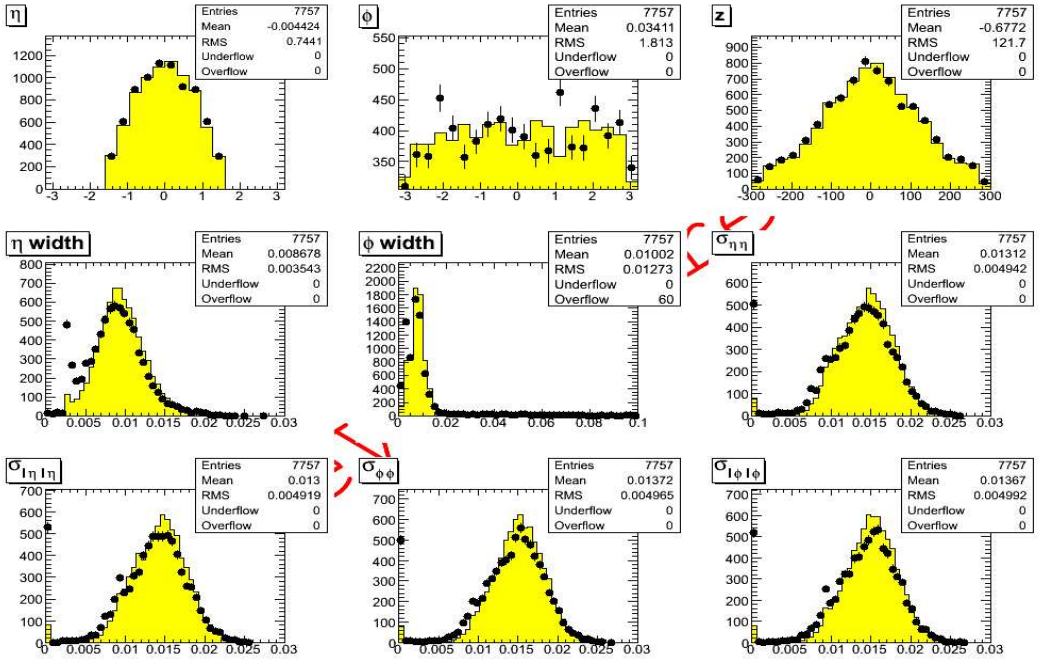


(a)

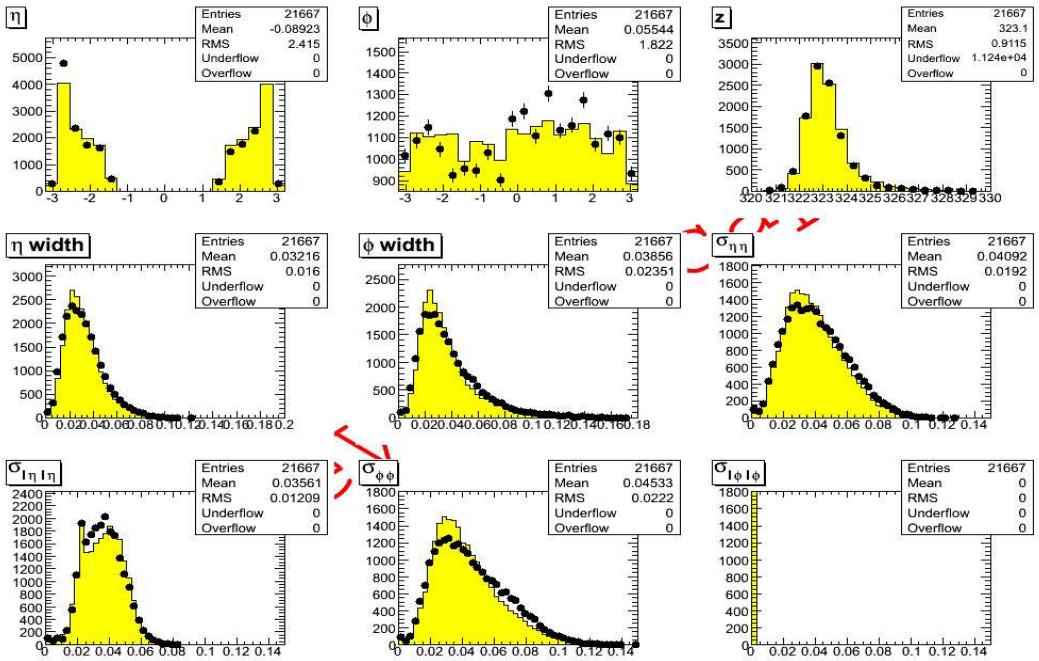


(b)

Figure 7.2: HEEP data quality monitoring distributions for energy variables at the supercluster level for the barrel (a) and the endcaps (b), for 900 GeV collision data.

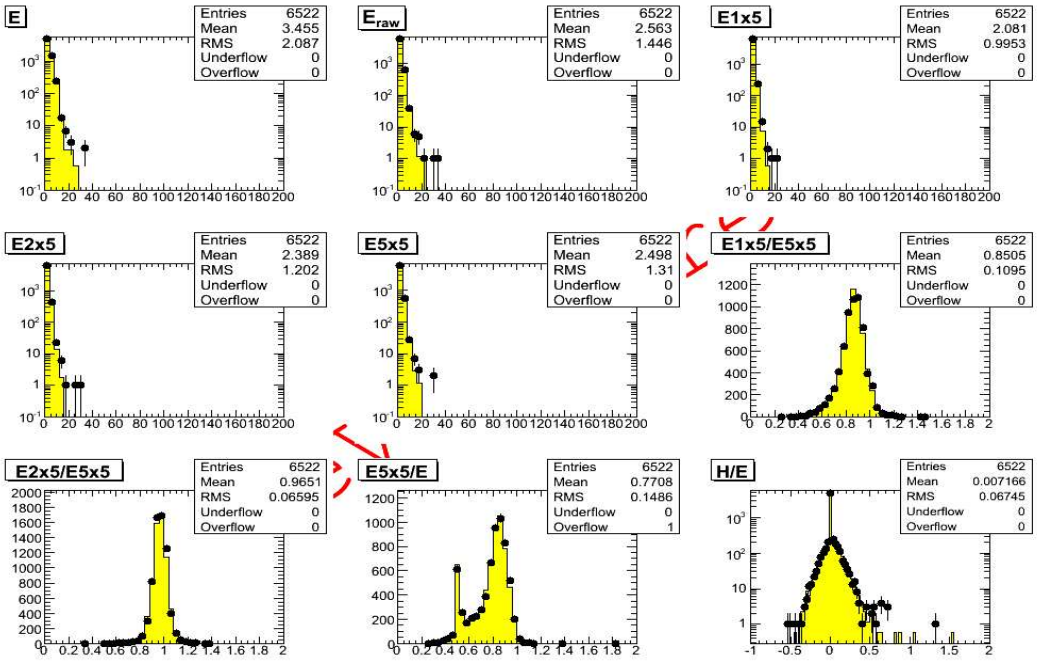


(a)

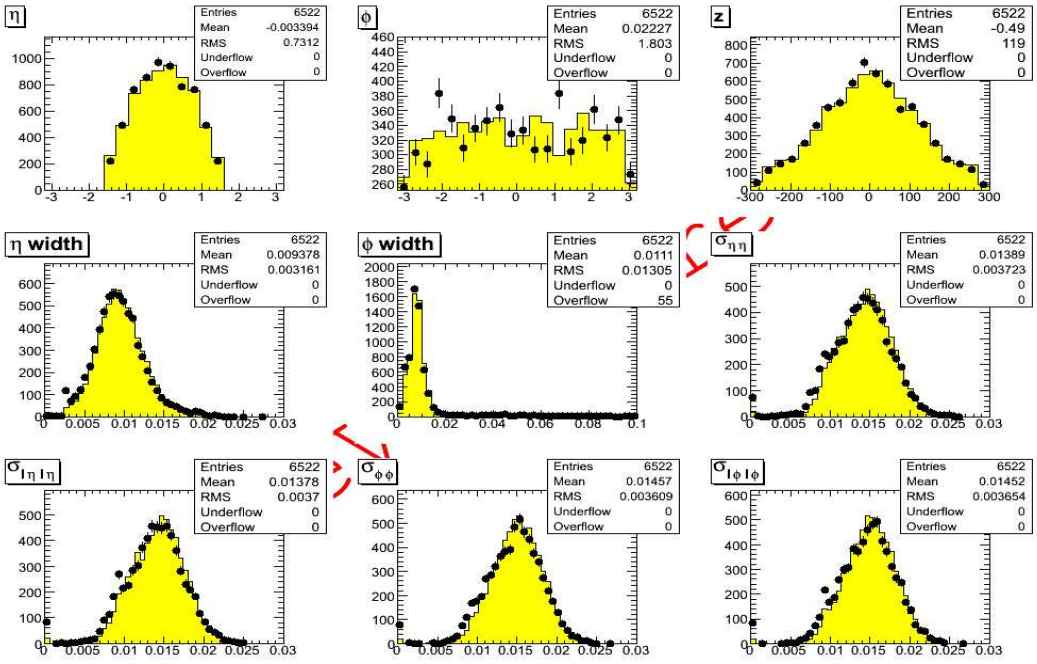


(b)

Figure 7.3: HEEP data quality monitoring distributions for shower shape variables at the supercluster level for the barrel (a) and the endcaps (b), for 900 GeV collision data.



(a)



(b)

Figure 7.4: HEEP data quality monitoring distributions for energy (a) and shower shape (b) variables at the supercluster level for the barrel, after spike removal, for 900 GeV collision data.

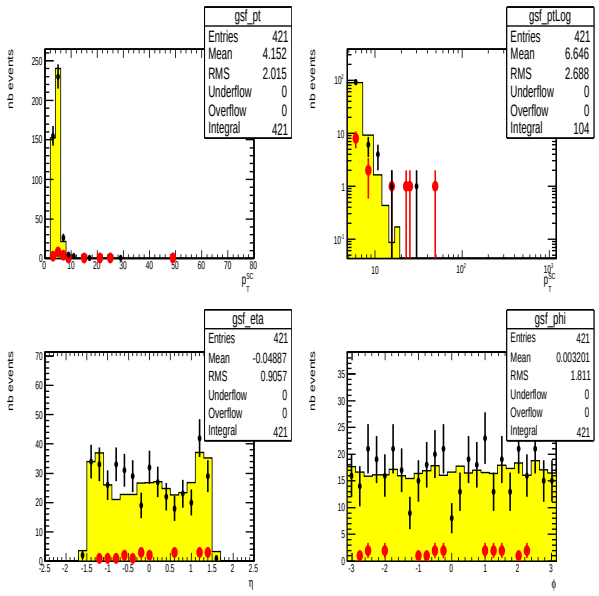
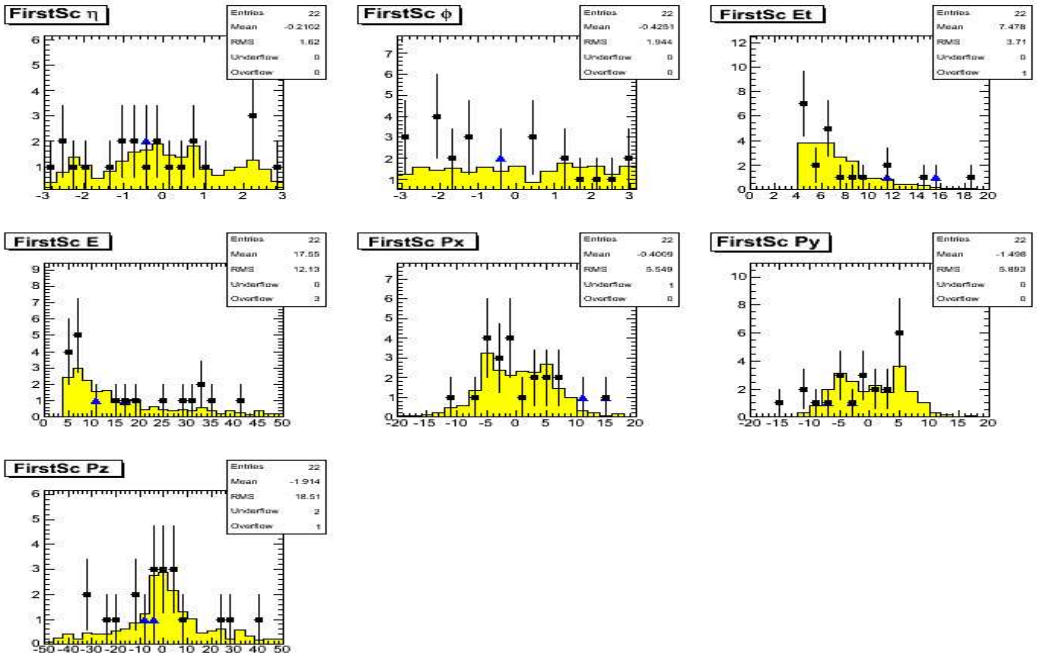


Figure 7.5: HEEP data quality monitoring distributions for variables at the GSF level for the barrel, for 900 GeV collision data. The "spiky" electron candidates are removed from data and shown separately as red points.

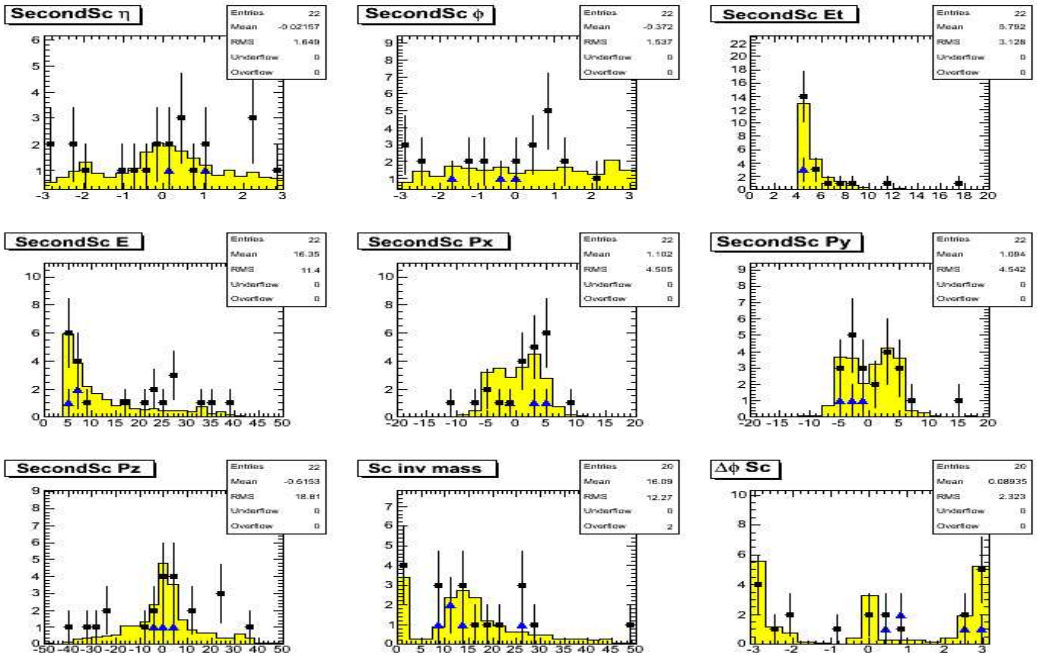
level objects such as transverse momentum, pseudorapidity and azimuthal angle as well as information about triggers that were fired by the event. It also offers the possibility to perform basic object/event selection.

Figure 7.7(a) presents an overview of the event display for the event with the highest invariant mass for the two superclusters for 900 GeV collision data. The central window shows a transverse view of the detector (ρ—ϕ view). The muon chamber stations are represented as red blocks while the energy deposits in the electromagnetic calorimeter (resp. hadronic calorimeter) are symbolized by red (resp. blue) lego bars with height proportional to the energy deposit. The size of the cell for the lego bar base corresponds to a 5x5 crystal matrix. Two other views (3D Lego view and ρ—z view) are available in the right part. The two superclusters contributing to the invariant mass are highlighted and one can see that the energy deposits are both located in the endcaps and spread among a large number of crystals, much larger than the expected number from an electron energy deposit. The left part provides information about high level objects. No GSF electron is present in the event. Figure 7.7(b) presents a complete three-dimensional view of the event.

Figure 7.8 presents a similar highest invariant mass event from 2.36 TeV collision data. The two superclusters contributing to the invariant mass are highlighted. An electron candidate is present in the event.

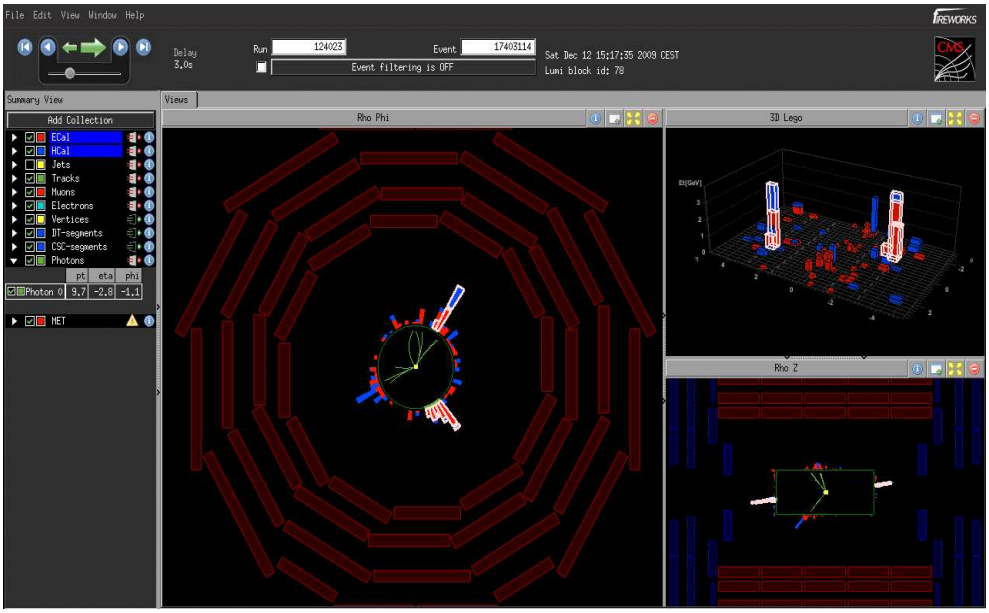


(a)

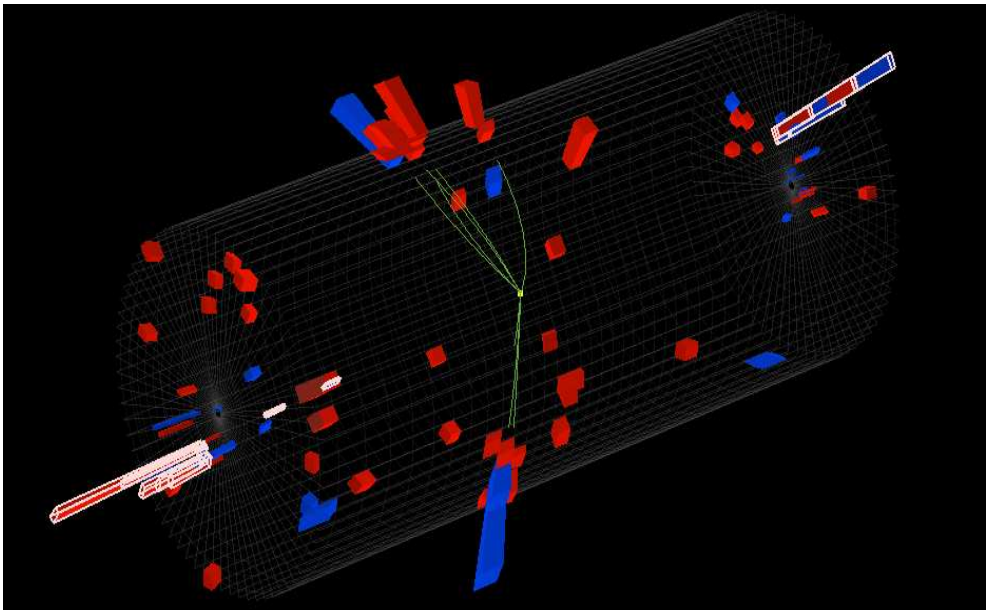


(b)

Figure 7.6: HEEP data quality monitoring distributions for supercluster variables at the HEEP event level for the highest E_t supercluster (a) and the second highest E_t supercluster (b), for 900 GeV collision data.

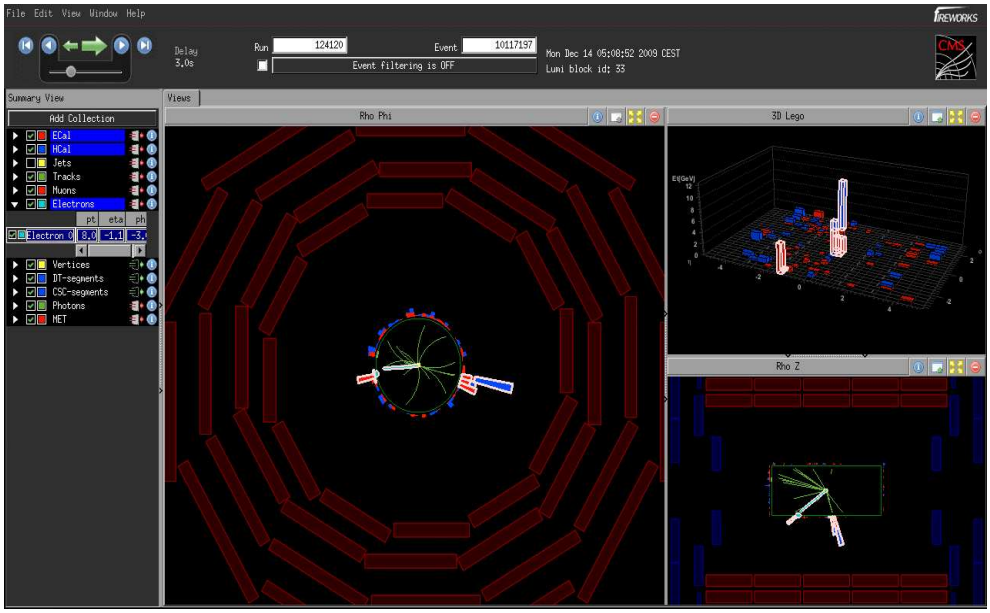


(a)

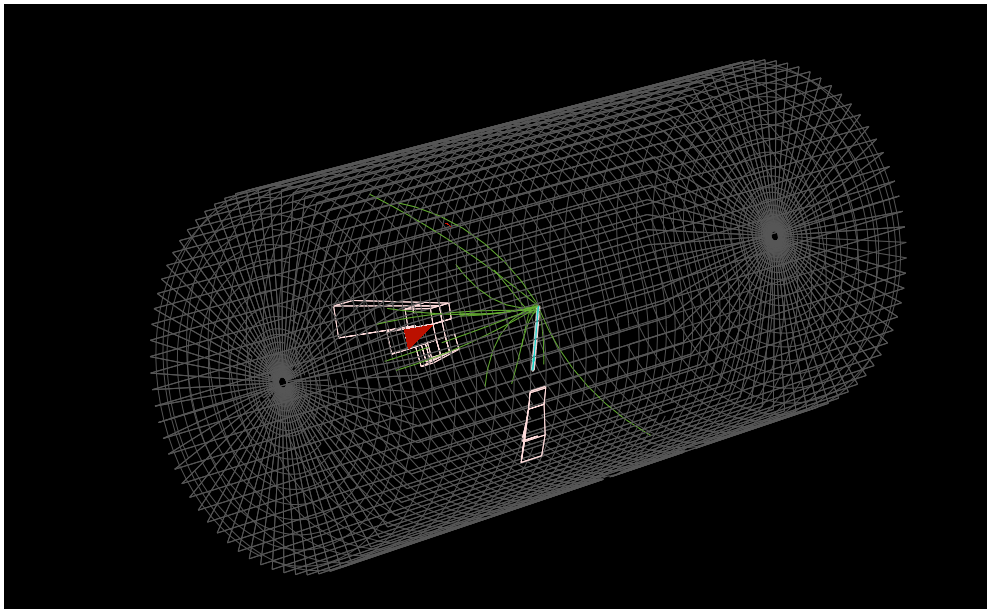


(b)

Figure 7.7: An event display for the event with highest supercluster pair invariant mass from 900 GeV collision data. Main window application (a) and global three-dimensional view (b).



(a)



(b)

Figure 7.8: An event display for the event with highest supercluster pair invariant mass from 2360 GeV collision data.Main window application (a) and global three-dimensional view (b).

Chapter 8

Data analysis at $\sqrt{s} = 7$ TeV

This chapter presents the study of the di-electron invariant mass spectrum performed on the very first LHC data at 7 TeV centre of mass energy. The samples used (data + Monte Carlo) are described in section 8.1. Section 8.2 presents the HEEP event candidates: the HEEP selection is described, the comparison between data and Monte Carlo expectations is shown and the di-electron invariant mass spectrum is presented with a focus on the high mass events. Section 8.3 presents the analysis performed at the Z peak ($60 < M < 120$ GeV/c^2): the selection efficiency is determined from data using the 'Tag and Probe' method, the background contributions to the mass spectrum are estimated from Monte Carlo and the cross section is computed with statistical, systematic and luminosity errors.

8.1 Samples

All the analysis presented in this chapter is based on data from pp collisions at 7 TeV centre of mass energy and Monte Carlo samples to simulate the Drell-Yan and background processes.

8.1.1 Data samples

The total integrated luminosity for the data is 2.77 pb^{-1} corresponding to a data-taking period from 30/03/2010 to 30/08/2010 (Runs 132440 $\rightarrow$ 144114). For each run, the detectors were required to be in a good status with respect to the CMS DQM requirements [53, 57]. The L1 trigger requires the presence in the ECAL of an energy deposit with transverse energy above 5 GeV while the HLT trigger requires the presence in the ECAL of an energy deposit with transverse energy above 15 GeV.

8.1.2 Monte Carlo samples

The Monte Carlo samples were all generated assuming a 7 TeV centre of mass energy, with detector conditions similar to those present during data-taking[1]. Several simulation generators were used (Pythia [34], MadGraph [58], PowHeg [59, 60, 61], MC@NLO [39]).

The Drell-Yan process ($Z \rightarrow e^+e^-$, $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow \tau^+\tau^-$) was generated with Pythia simulation program with matrix elements computed at leading order effects. In addition, a second Drell-Yan ($Z \rightarrow e^+e^-$) sample was generated with PowHeg simulation program with

[1] During 2009, prior to first collisions, first steps of alignment and calibration of the different parts of the detector were performed using cosmic muons. The detector conditions considered in the Monte Carlo samples correspond to the level of alignment and calibration achieved at the end of this period.

matrix elements computed at next-to-leading order effects. A cut on the generated invariant mass of the two electrons is put at 20 GeV/c^2.

The QCD multijet process was generated with Pythia simulation program. Four bins in $\hat{p_t}$ were simulated ($\hat{p_t}$ > 15 GeV/c, 30 GeV/c, 80 GeV/c and 170 GeV/c). However, large statistics is needed to simulate fake electrons from QCD jets as the fake rate is low. For this reason, QCD samples with enhanced electromagnetic content (QCD Em Enriched) were simulated with Pythia in 3 bins in $\hat{p_t}$ ([20-30], [30-80], [80-170]). Only light quarks (u,d,s) were considered in these processes. In addition, QCD samples with an electron from a b/c quark decay chain (QCD $b/c \rightarrow e$) were simulated with Pythia in 3 bins in $\hat{p_t}$ ([20-30], [30-80], [80-170]). The QCD Em Enriched and QCD $b/c \rightarrow e$ samples are often referred to as biased QCD samples while the QCD samples are referred to as unbiased QCD samples.

The $t\bar{t}$ process was generated, at leading order, with Pythia simulation program. A $t\bar{t}$+jets sample was generated at leading order using Madgraph simulation program.

The W+jet, WW and tW processes were generated using MadGraph, Pythia and MC@NLO simulation programs, respectively.

Table 8.1 lists all the samples presented above and provides the cross sections as well as the K factors to account for higher order effects. For the QCD biased samples, an enrichment factor is taken into account.

Sample	Generator	Kinematic cuts (in GeV, c=1)	σ (pb)	K factor (Enrichment factor)
Z→ee	powheg (NLO)	M(ee) > 20	1522	1.09
Z→ee	Pythia6 (LO)	M(ee) > 20	1300	1.28
Z→$\mu\mu$	Pythia6	M($\mu\mu$) > 20	1300	1.28
Z→$\tau\tau$	Pythia6	M($\tau\tau$) > 20	1300	1.28
W+Jets	MadGraph	inclusive	24170	1.30
$t\bar{t}$+jets	MadGraph	inclusive	95	1.66
$t\bar{t}$	Pythia6		95	1.66
QCD	Pythia6	pthat > 15	8.762 10^8	/
QCD	Pythia6	pthat > 30	6.041 10^7	/
QCD	Pythia6	pthat > 80	9.238 10^5	/
QCD	Pythia6	pthat > 170	2.547 10^4	/
QCD Em Enriched	Pythia6	$20 < \hat{p_t} < 30$	2.355 10^8	(0.0073)
QCD Em Enriched	Pythia6	$30 < \hat{p_t} < 80$	5.93 10^7	(0.059)
QCD Em Enriched	Pythia6	$80 < \hat{p_t} < 170$	9.06 10^5	(0.148)
QCD b/c → e	Pythia6	$20 < \hat{p_t} < 30$	2.355 10^8	(0.00046)
QCD b/c → e	Pythia6	$30 < \hat{p_t} < 80$	5.93 10^7	(0.00234)
QCD b/c → e	Pythia6	$80 < \hat{p_t} < 170$	9.06 10^5	(0.0104)
WW	Pythia6		2.93	/
tW	MC@NLO	diagram removal	0.1853	/

Table 8.1: List of Monte Carlo samples for the signal and background processes. The cross sections and the K factors are provided. For QCD biased samples, an enrichment factor is provided.

In the following, the Drell-Yan sample from PowHeg is used as the matrix element computation is considered at next-to-leading order. For the $t\bar{t}$ process, the $t\bar{t}$+jets sample from MadGraph is used. For the QCD multijet, the QCD biased (QCD Em Enriched + b/c → e) samples are used as high statistics is needed in the case of QCD unbiased samples to simulate

the fake electrons from jets. These samples are the sames used by the CMS 'electroweak' group [62].

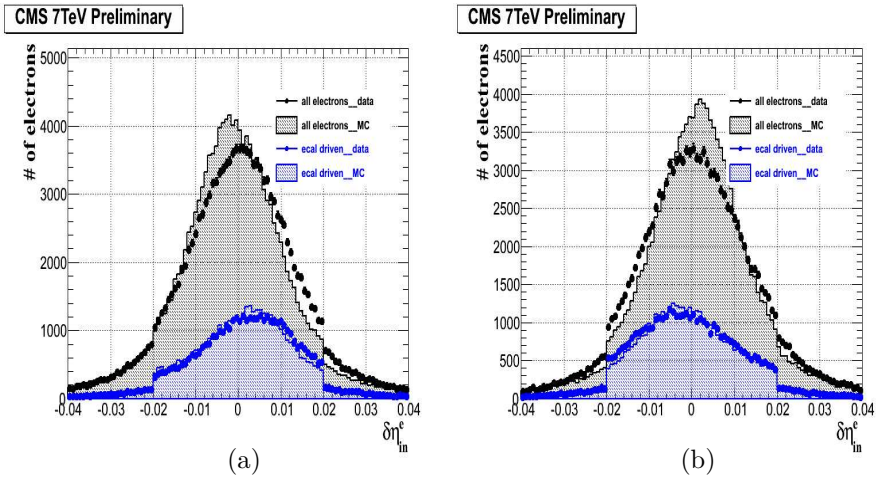
8.2 HEEP event candidates

This section focuses on the events with two electron candidates that fulfill the HPTE requirements described in section 6.5 and summarized in table 6.4. Section 8.2.1 presents the selection and section 8.2.2 makes the comparison between data and Monte Carlo for the different variables of the HPTE selection. The di-electron invariant mass spectrum is presented in section 8.2.3 and the event displays for selected data events with mass above 120 GeV/c^2 are presented in section 8.2.4.

8.2.1 Selection

Selected events are required to have at least two electron candidates that pass the high p_t electron requirements described in table 6.4. However, a large discrepancy appears in the endcaps when comparing distributions from data and Monte Carlo for the $\Delta\eta$ variable in a sample of electron candidates. This discrepancy is visible in figure 8.1 which shows the distribution of $\Delta\eta$ in both endcaps (- (a), + (b)) for electron candidates in data and in Monte Carlo samples. This discrepancy is due to a problem of alignment between the tracker and the ECAL endcaps and is corrected for in later releases of the CMS software. For the moment, the $\Delta\eta$ variable is chosen not to be applied in the endcaps. The selection criteria used in this analysis are summarized in table 8.2.

In case more than two electron candidates are present in the event, the two candidates with highest transverse energy are considered.



(a) (b)

Figure 8.1: Distribution of $\Delta\eta$ for electron candidates in EE- (a) and EE+ (b), in data (black points) and Monte Carlo sample (solid lines), taken from [63].

In addition to the high p_t electron requirements, a criterion is applied on electrons to reject possible spikes. Such problematic effects have been introduced in section 7.3 and a criterion on the quantity $E1/E9$ was applied. In the present selection the latter criteria is replaced by a new one called 'SwissCross'. This variable is computed from the energy contained in the highest energy crystal of the supercluster (E_1) and the sum of the energies of the four neighbouring crystals forming a 'Swiss Cross' shape (E_4). To avoid negative energy

Variable	Barrel	Endcaps
E_t	> 25 GeV	> 25 GeV
$\lvert\eta_{SC}\rvert$	< 1.442	$1.560 < \lvert\eta_{SC}\rvert < 2.5$
isEcalDriven	=1	=1
Identification variables		
H/E	< 0.05	< 0.05
$\Delta\eta_{in}$	< 0.005	$<$ not applied
$\Delta\phi_{in}$	< 0.09	< 0.09
$E_{2\times5}/E_{5\times5}$	> 0.94 or $E_{1\times5}/E_{5\times5} > 0.83$	not applied
$\sigma_{i\eta i\eta}$	not applied	< 0.03
Isolation variables		
isol. ECAL+HCAL1	< 2+0.03 E_t (GeV)	$<$2.5 ($E_t < 50$ GeV)
		< 2.5+0.03 (E_t-50) ($E_t > 50$ GeV)
isol. HCAL2	not applied	< 0.5 GeV
isol. Tracker	< 7.5 GeV/c	< 15 GeV/c

Table 8.2: HPTE selection criteria for the barrel and the endcaps.

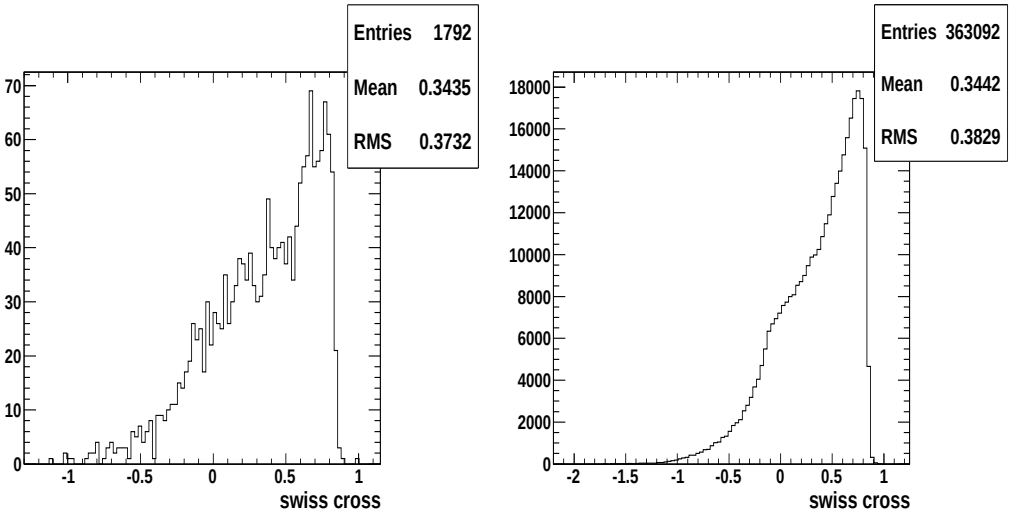
from noise contributions, a minimal energy threshold of 0 GeV is applied on the considered crystals. The 'SwissCross' variable is defined as the quantity 1 - E_4/E_1. If the crystals energy threshold is not fulfilled, the variable is set to 0. It is presently also set to 0 in the endcaps as the spike effects appear only in the barrel. The distribution of the 'SwissCross' variable is given in figure 8.2 for 'ecaldriven' electron candidates (GSF) in the barrel, out of the gap, with transverse energy higher than 25 GeV which pass the HPTE selection in data (a) and in a Monte Carlo Drell-Yan sample (Powheg) with mass above 20 GeV/c^2 (b). In the Monte Carlo Drell-Yan sample, the electron candidates are in addition required to be matched to the generated electrons.

A cut at 95% is applied on electron candidates in data and in Monte Carlo for consistency. The total number of selected electron candidates in data is 1792 among which 1 has a swiss cross value higher than 95%. In the Monte Carlo Drell-Yan sample, the contribution of selected electron candidates with a swiss cross value higher than 95% to the total number of selected electron candidates is of the order of 0.001%.

8.2.2 Comparison data — Monte Carlo

To check the agreement between data and Monte Carlo expectations, a comparison is performed in the distributions of the kinematic variables used in the HPTE requirements (GSF transverse energy, GSF supercluster pseudorapidity and azimuthal angle, see table 8.2). Figures 8.3 present for each of the three kinematic variables the distributions from data and Monte Carlo for electron candidates which fulfill the HPTE requirements. The distributions from Monte Carlo are stacked, normalised to the integrated luminosity and superimposed to distributions from data. Good agreement is observed between data and Monte Carlo expectations for the distributions of the GSF supercluster pseudorapidity and azimuthal angle. The GSF transverse energy from data presents a distribution slightly shifted to lower transverse energy values compared to the Monte Carlo expectations. This indicated that the ECAL energy measurement is slightly underestimated (from the order of 1%). This has been corrected in the new versions of the reconstruction software.

The agreement between data and Monte Carlo expectations is also checked in the distri-



(a) (b)

Figure 8.2: Distribution of 'SwissCross' variable for electron candidates (GSF) matched to generated electrons, in the tracker acceptance, out of the gap and with transverse energy higher than 25 GeV in the barrel only, in data (a) and in Drell-Yan sample (Powheg) with mass above 20 GeV/c^2 (b).

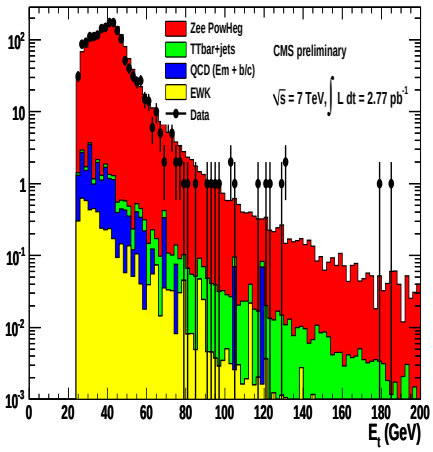
butions of the variables used in the HPTE requirements (table 8.2). The comparison is based on the so-called $N-1$ distributions; the distributions of each variable from data and Monte Carlo are superimposed for GSF electrons which fulfill the HPTE requirements except the requirement for the variable under study.

Figures 8.4 and 8.5 present, for the barrel and the endcaps respectively, the $N-1$ distributions of the identification variables ($\Delta\eta$ (a), $\Delta\phi$ (b), H/E (c) and $\sigma_{i\eta i\eta}$ (d)) for candidate electrons (black points) and the Monte Carlo samples (solid lines), see legend. The Monte Carlo histograms are stacked, normalised to the integrated luminosity and superimposed to distributions from data. The same distributions for the isolation variables (ECAL+HCAL1 isolation (a), HCAL2 isolation (b) and Tracker isolation (c)) are shown in figures 8.6 and 8.7 for the barrel and the endcaps respectively.

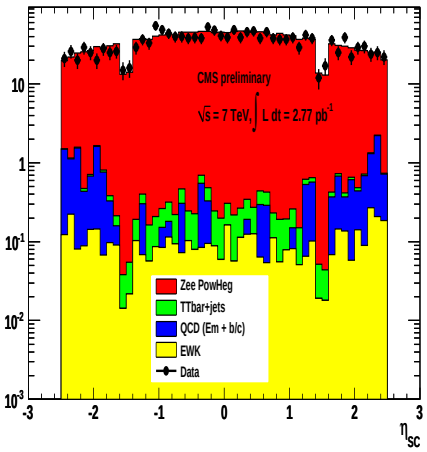
The $\Delta\eta$ barrel distribution in data is shifted to the right with respect to the Monte Carlo. In the endcaps, a clear discrepancy appears between data and Monte Carlo. This discrepancy was already mentioned in section 8.2.1. The distribution of the ECAL+HCAL1 isolation in data is also shifted to the right with respect to the Monte Carlo. This is possibly due to different noise threshold values in data and in Monte Carlo. Other than that, the distributions from data are in relatively good agreement with the Monte Carlo expectations.

8.2.3 Di-electron invariant mass spectrum

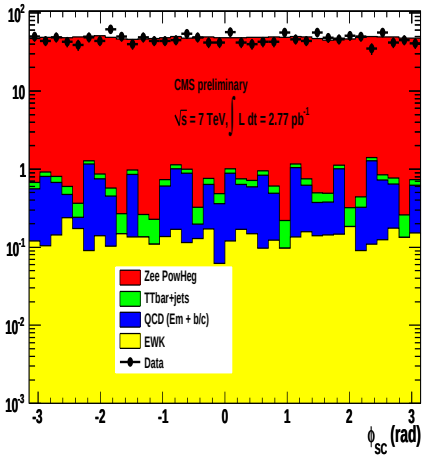
The invariant mass spectrum for the selected events is derived from the two electrons energy-momentum four-vector. The supercluster information is used for the energy estimates (E_{SC}) as it has better resolution than the momentum derived from the track. The track information as determined from measurements at inner part of tracker is used for the angle estimates (θ_{tr},ϕ_{tr}) as it has better resolution than the supercluster position:



(a)

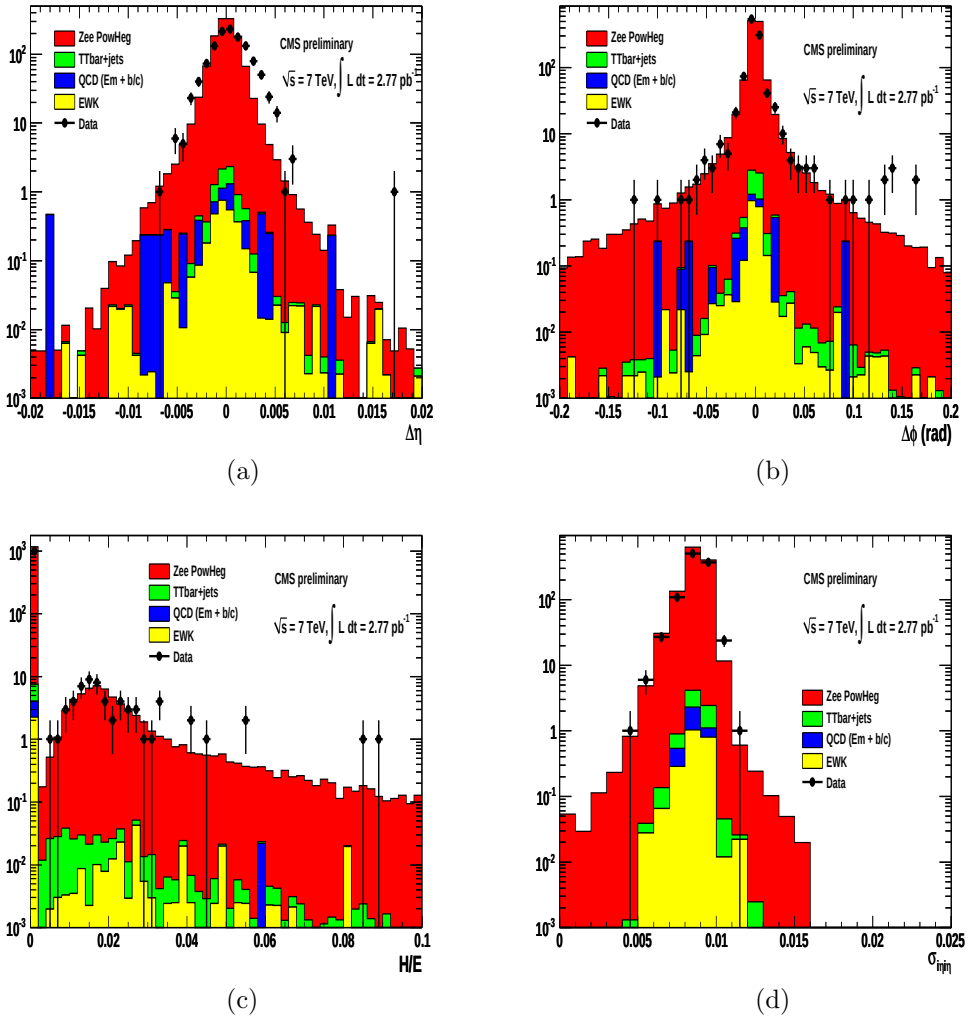


(b)



(c)

Figure 8.3: Distributions of the kinematic variables (GSF transverse energy (a), GSF supercluster pseudorapidity (b) and GSF supercluster azimuthal angle (c)) for HPTE electron candidates from data (black points) and from the Monte Carlo samples (solid lines), see legend. The Monte Carlo histograms are stacked, normalised to the integrated luminosity and superimposed to distributions from data.



(a) (b)

(c) (d)

Figure 8.4: $N-1$ distributions of the identification variables ($\Delta\eta$ (a), $\Delta\phi$ (b), H/E (c) and $\sigma_{i\eta i\eta}$ (d)) for electron candidates in the barrel, from data (black points) and from the Monte Carlo samples (solid lines), see legend. The Monte Carlo histograms are stacked, normalised to the integrated luminosity and superimposed to distributions from data.

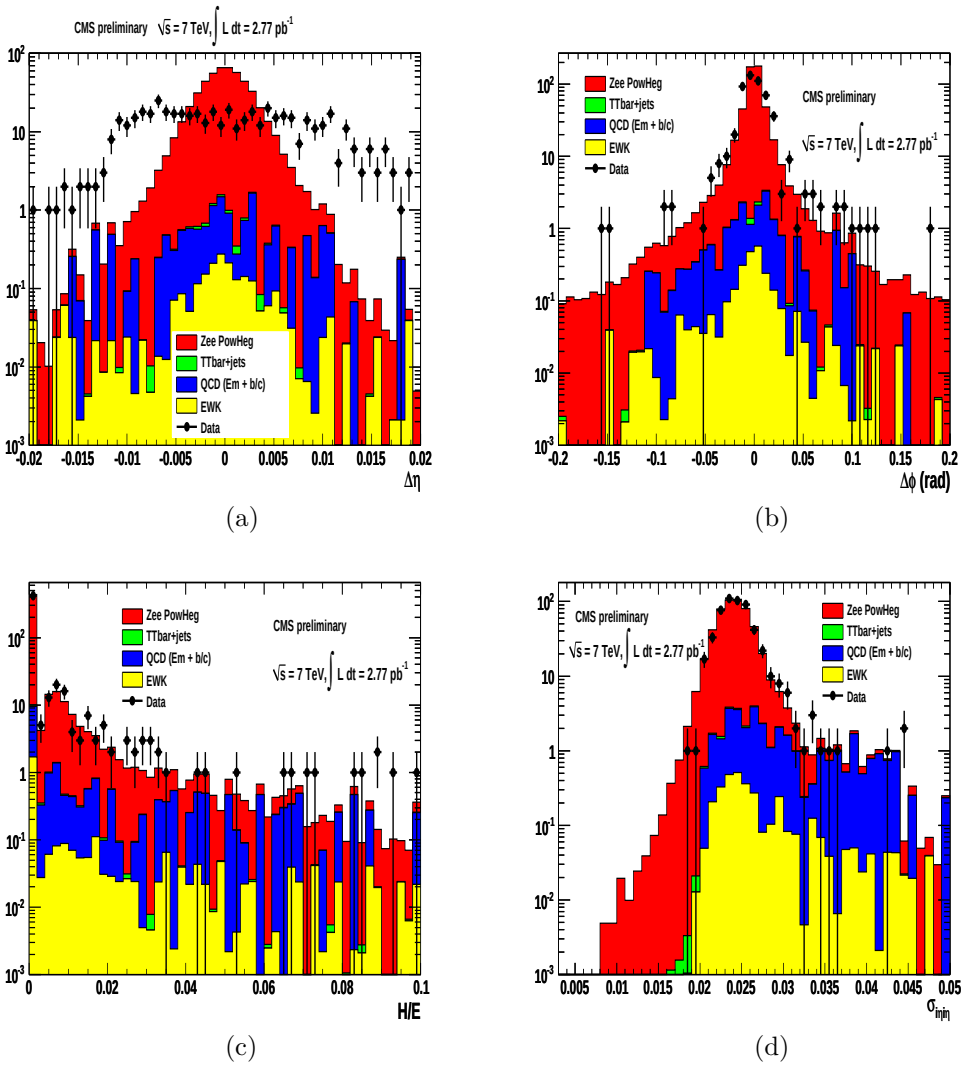
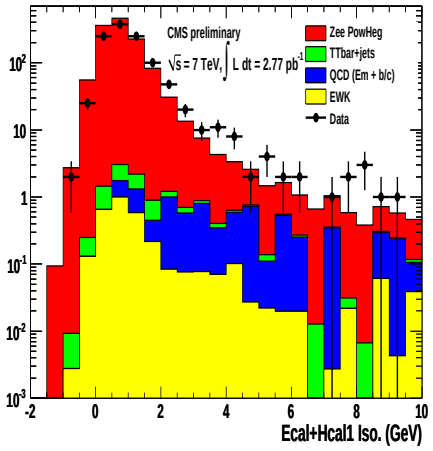
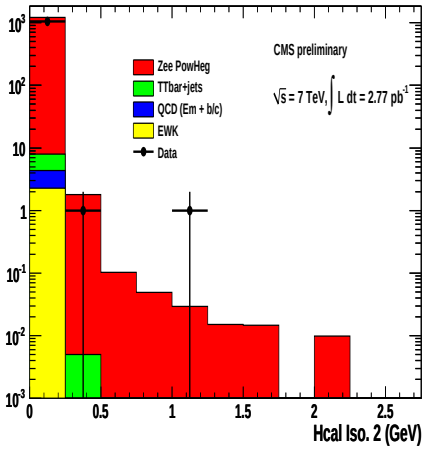


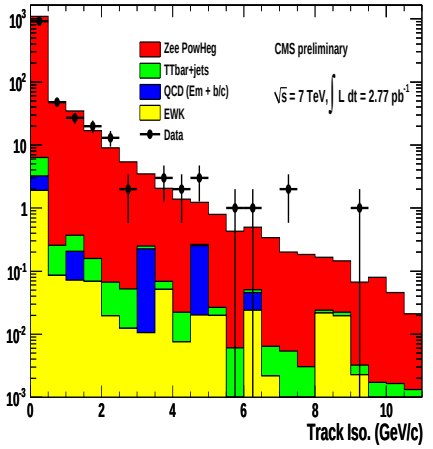
Figure 8.5: $N-1$ distributions of the identification variables ($\Delta\eta$ (a), $\Delta\phi$ (b), H/E (c) and $\sigma_{i\eta i\eta}$ (d)) for electron candidates in the endcaps, from data (black points) and from the Monte Carlo samples (solid lines), see legend. The Monte Carlo histograms are stacked, normalised to the integrated luminosity and superimposed to distributions from data.



(a)

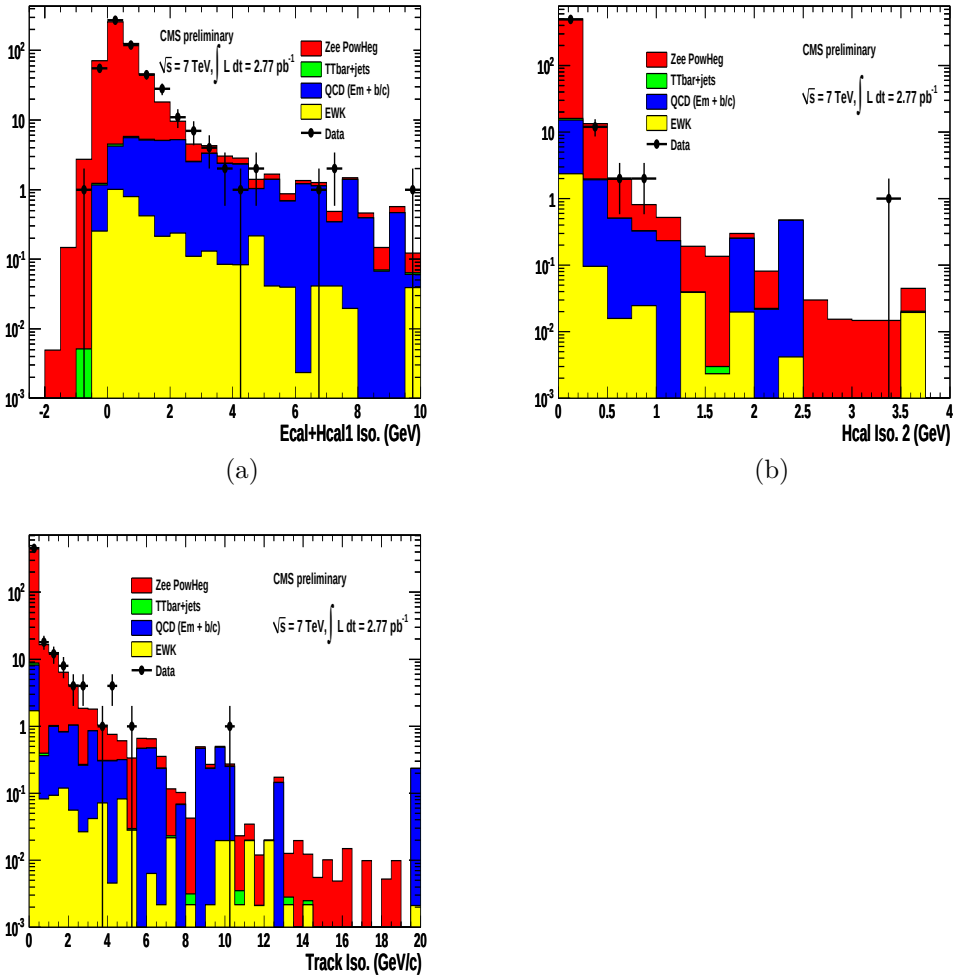


(b)



(c)

Figure 8.6: $N-1$ distributions of the isolation variables (ECAL+HCAL1 isolation (a), HCAL2 isolation (b) and Tracker isolation (c)) for electron candidates in the barrel, from data (black points) and from the Monte Carlo samples (solid lines), see legend. The Monte Carlo histograms are stacked, normalised to the integrated luminosity and superimposed to distributions from data.



(a) (b) (c)

Figure 8.7: $N-1$ distributions of the isolation variables (ECAL+HCAL1 isolation (a), HCAL2 isolation (b) and Tracker isolation (c)) for electron candidates in the endcaps, from data (black points) and from the Monte Carlo samples (solid lines), see legend. The Monte Carlo histograms are stacked, normalised to the integrated luminosity and superimposed to distributions from data.

$$p^{\mu} = \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} E_{SC} \\ E_{SC} \sin\theta_{tr} \cos\phi_{tr} \\ E_{SC} \sin\theta_{tr} \sin\phi_{tr} \\ E_{SC} \cos\theta_{tr} \end{pmatrix} \tag{8.1}$$

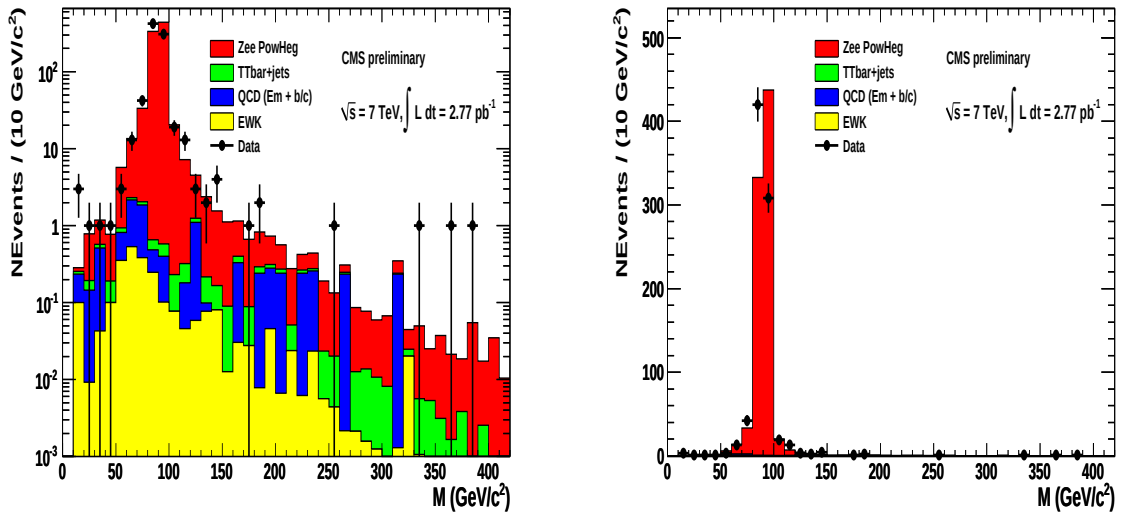
The HEEP di-electron invariant mass spectrum is shown in figure 8.8 for the data (black points) and the Monte Carlo samples (solid lines). The Monte Carlo histograms are stacked, superimposed and normalised to the integrated luminosity. The plots are shown in logarithmic (a) and linear scale (b).

For an integrated luminosity of 2.77 pb^{-1} and using the HEEP selection, a total of 837 HEEP events with mass above 20 GeV/c^2 is present in the data.

The mass spectrum for the data is in agreement with the Monte Carlo expectation except the peak in the data distribution which is slightly shifted with respect to the Monte Carlo. This effect is a consequence of the peak shift observed in the transverse momentum distribution of the HPTE electrons (cf. section 8.2.2). The contributions from the different Monte Carlo backgrounds are very small. The two biggest background contributions come from the QCD dijet samples (mainly from the 30—80 $\hat{p}_t$ bin) and the $t\bar{t}$ samples.

The number of HEEP candidate events with mass above 20 GeV/c^2 are presented in table 8.3 for data and for the different Monte Carlo samples. The contributions from the Monte Carlo are normalised to the integrated luminosity. The contributions are added for the QCD (QCD Em Enriched + QCD b/c, 3 $\hat{p}_t$ bins together) and the $Z \to \mu\mu$, $Z \to \tau\tau$, $W + Jets$, WW and tW contributions are added to form the electroweak contribution. The $t\bar{t}$ and QCD unbiased processes represent 63.6% and 18.9% of the total background contribution, respectively.

According to the Monte Carlo, the total background contribution represents 1.43% of the di-electron invariant mass spectrum in the whole mass range.



(a) (b)

Figure 8.8: HEEP di-electron invariant mass spectrum for the Monte Carlo samples (solid lines) and the data (black points) in logarithmic scale (a) and linear scale (b). The Monte Carlo histograms are stacked, superimposed and normalised to the integrated luminosity.

Data	837
Z $\rightarrow$ ee (Powheg)	857.2
$t\bar{t}$ (Madgraph)	2.2
QCD Em Enriched $\hat{p_t}$ [20—30]	0.1
QCD Em Enriched $\hat{p_t}$ [30—80]	6.8
QCD Em Enriched $\hat{p_t}$ [80—170]	0.7
QCD b/c $\rightarrow$ e $\hat{p_t}$ [20—30]	0
QCD b/c $\rightarrow$ e $\hat{p_t}$ [30—80]	0.2
QCD b/c $\rightarrow$ e $\hat{p_t}$ [80—170]	0.0
QCD	7.9
Z $\rightarrow \mu\mu$	0.0
Z $\rightarrow \tau\tau$	1.0
$W + Jets$	1.0
WW	0.3
tW	0.0
Electroweak	2.3
Total background contribution	12.4

Table 8.3: Number of HEEP candidate events with mass above 20 GeV/c^2, for an integrated luminosity of 2.77 pb^{-1}, for data (black points) and for Monte Carlo (solid lines). The contributions from the Monte Carlo are normalised to the integrated luminosity.

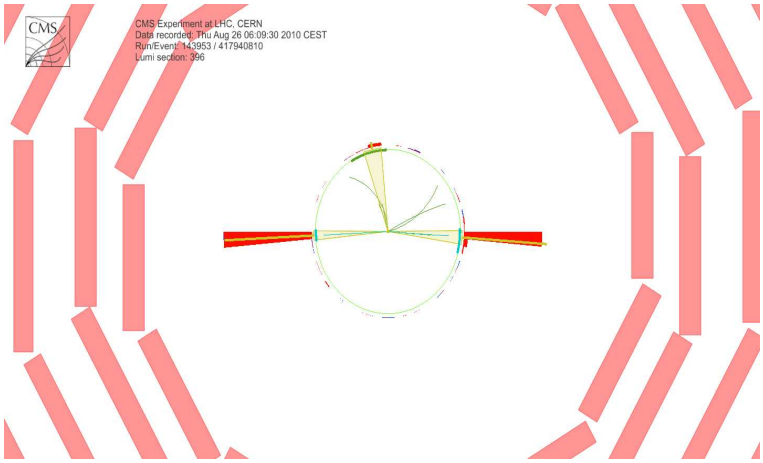
8.2.4 High mass event display

The HEEP events at high mass (M > 120 GeV/c^2) are scrutinised with help of the CMS event display [56] already introduced in section 7.6. 16 events are present in the high mass region as can be seen from figures 8.8 with a total of 16.4 expected from Monte Carlo predictions. The event display for the highest mass event is presented in figures 8.9 (a) (ρ—ϕ view), (b) (ρ—z view) and (c) (3D lego view). The two HPTE electron candidates are back to back and isolated. Their energy deposits are mainly located in the ECAL (red bars) and very little energy is present in the HCAL cells located behind (blue bars). A third jet is present in the event (yellow triangle) and one of the two electrons is emitted in the endcaps.

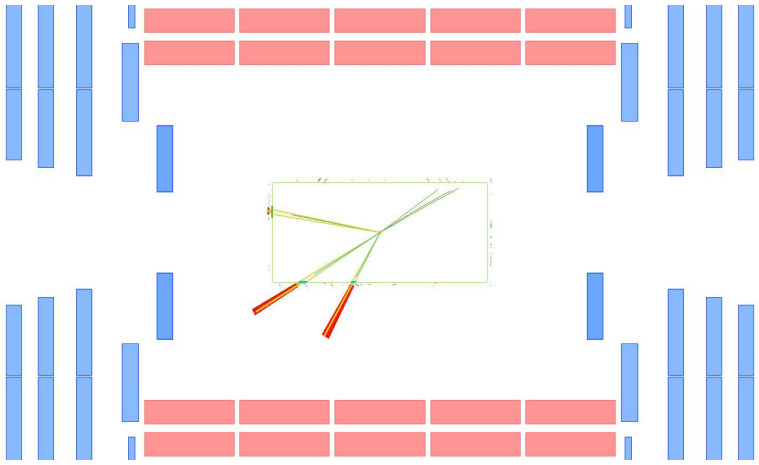
Some information about this event (Run/Event number and di-electron invariant mass) as well as for the two electron candidates (E_t, η, ϕ, identification and isolation variables) is summarized in table 8.4.

8.3 Analysis at the Z peak

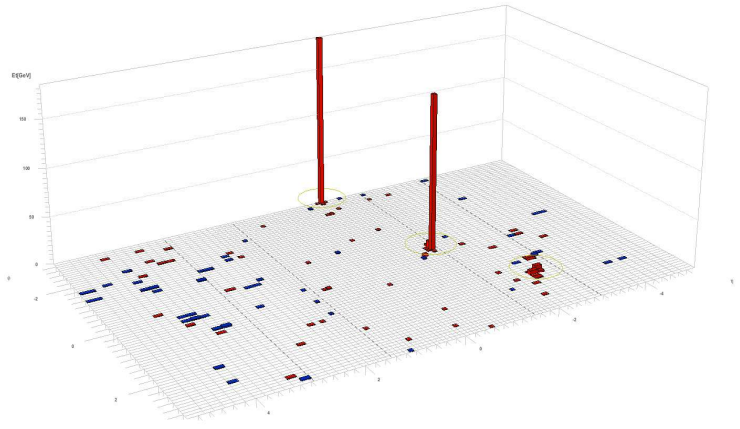
This section focuses specifically on the Z peak region defined by the mass range $60 < M < 120$ GeV/c^2. The high p_t electron selection efficiency is derived from data in section 8.3.1 using the "Tag and Probe" method described in section 5.3 and documented in [43]. The background contributions to the HEEP mass spectrum at the Z peak are discussed and estimated from Monte Carlo in section 8.3.2. Finally, the estimation of the cross section at the Z peak is performed in section 8.3.3.



(a)



(b)



(c)

Figure 8.9: High mass HEEP event display. ρ—ϕ view (a), ρ—z view (b) and 3D lego view (c).

Run 143953, Event 417940810, Mass = 388 GeV		
Variable	Electron 1	Electron 2
E_t(GeV)	184.33	178.09
η_{SC}	-0.57	-1.34
ϕ_{SC}(rad)	-3.09	-0.07
SwissCross	0.33	0.52
H/E	0	0
$\Delta\eta_{in}$	0.00	0.00
$\Delta\phi_{in}$(rad)	-0.00	-0.00
$E_{1\times5}/E_{5\times5}$	0.88	0.87
$E_{2\times5}/E_{5\times5}$	0.94	0.96
$\sigma_{i\eta i\eta}$	0.01	0.01
ECAL isol. (GeV)	1.78	1.38
HCAL1 isol. (GeV)	0	0
HCAL2 isol. (GeV)	0	0
Tracker isol. (GeV/c)	0	0
Charge	1	-1

Table 8.4: High mass event information. Run number, event number, di-electron invariant mass and values of identification and isolation variables for both electrons.

8.3.1 Selection efficiency from the "Tag and Probe" method

In the Z peak region, where little background contamination is expected, the HPTE selection efficiency is determined directly from data using the "Tag and Probe" method. Such a method was studied using simulation at centre of mass energies of 14 and 10 TeV in section 5.3 and is here applied to the data.

This method relies on the presence in the event of two objects, the 'tag' and the 'probe' where the 'tag' is required to pass stringent selection criteria while the 'probe' is used to measure the efficiency. To ensure a high purity di-electron sample, the invariant mass of the two objects is required to be in the mass range $80 < M < 100$ GeV/c^2.

In our case, the 'probe' is required to be an 'ecaldriven' GSF electron candidate[2] with transverse energy higher than 25 GeV, emitted in the tracker acceptance, out of the gap and with a 'SwissCross' value smaller than 95% (N_{probe}). The 'tag' is required to fulfill the 'probe' requirements and the high p_t electron selection ($N_{probe,HPTE}$). In addition, a maximal value of 1.5 is required on the tag on the quantity E/p defined as the ratio of the supercluster energy over the track momentum.

The efficiency is determined separately for the barrel and the endcaps, from the ratio:

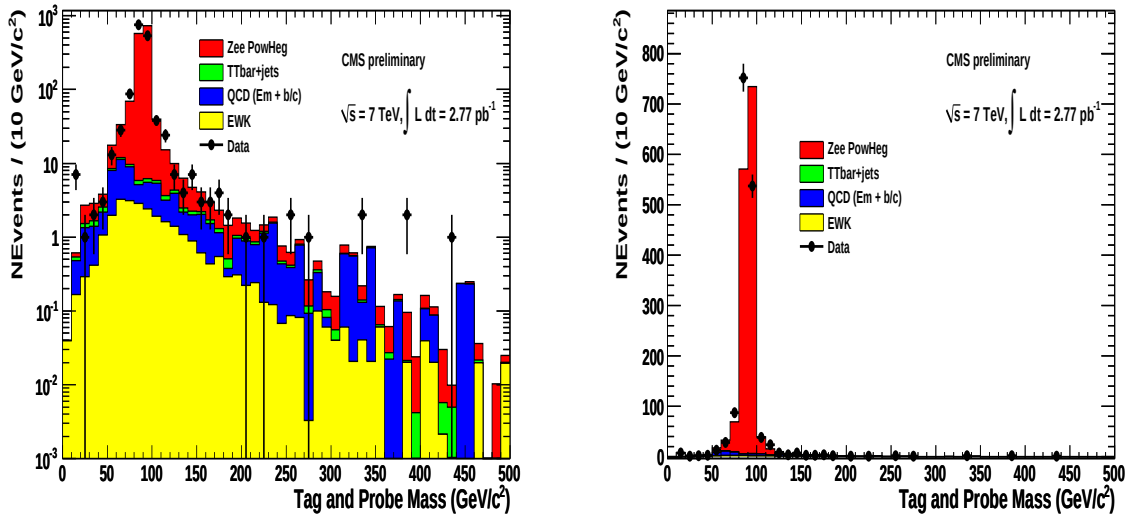
$$\epsilon_{HPTE,TP} = \frac{N_{probe,HPTE}}{N_{probe}} \tag{8.2}$$

Note that if both electrons satisfy the requirements of the 'tag', the pair is considered twice. Also, for consistency with section 8.2, the $\Delta\eta$ criterion is not applied in the endcaps.

The invariant mass spectrum of all tag and probe pairs in the whole mass range is given in figure 8.10 for data (black points) and Monte Carlo (solid histograms). The Monte Carlo distributions are stacked, superimposed and normalised to the integrated luminosity. Plots are shown in logarithmic scale (a) and linear scale (b). A total of 1533 tag-probe pairs is

[2]Note that the GSF preselection (section 6.2.3 is accounted for).

found in data in the whole mass range[3]. As for figure 8.8, the Z peak is shifted in data with respect to Monte Carlo expectations. The contamination from the backgrounds is expected to be more important as the selection is less stringent in the tag-probe pairs.



(a) (b)

Figure 8.10: Tag and Probe pair invariant mass spectrum for the data (black points) and the Monte Carlo samples (solid lines) in logarithmic scale (a) and linear scale (b). The Monte Carlo histograms are stacked, superimposed and normalised to the integrated luminosity.

The numbers of tag-probe pairs in the mass range $80 < M < 100$ GeV/c^2 are presented in table 8.5 for data and for the different Monte Carlo samples. The contributions from the Monte Carlo are normalised to the integrated luminosity.

A total of 1289 tag-probe pairs is found in data in the mass range $80 < M < 100$ GeV/c^2 (of which 893 have their probes in the barrel and 396 in the endcaps). According to the Monte Carlo, the total background contamination represents 0.69% of the total number of tag-probe pairs in the mass range $80 < M < 100$ GeV/c^2.

The high p_t electron selection efficiency is computed from the number of probes that pass the high p_t electron selection requirements. Table 8.6 gives the number of probes and the number of probes which pass the high p_t electron selection requirements, for data, Monte Carlo Drell-Yan and background samples (all added together), separately for the barrel and the endcaps. The Monte Carlo contributions are normalised to the integrated luminosity.

The background contaminations are very small in the sample of probes which fulfill the high p_t electron selection criteria (0.08 % and 0.14 % in the barrel and the endcaps respectively) and are neglected. The relative background contaminations in the sample of probes are of 0.7% and 1.3% for the barrel and the endcaps respectively.

The high p_t electron selection efficiency numbers are deduced, for the barrel and the endcaps, based on the numbers from table 8.6. The statistical errors are computed using binomial statistics:

[3]Note that the number of tag-probe pairs in the whole mass range is not higher than twice the number of HEEP events since there is an additional E/p requirement on the tag.

Data	1289
Z→ee	1293.4
$t\bar{t}$	1.3
QCD Em Enriched $\hat{p}_t$ [20-30]	0
QCD Em Enriched $\hat{p}_t$ [30-80]	4.0
QCD Em Enriched $\hat{p}_t$ [80-170]	0.9
QCD b/c → e $\hat{p}_t$ [20-30]	0
QCD b/c → e $\hat{p}_t$ [30-80]	0.4
QCD b/c → e $\hat{p}_t$ [80-170]	0.2
QCD	5.4
Z→$\mu\mu$	0.0
Z→$\tau\tau$	0.2
$W + Jets$	5.0
WW	0.1
tW	0.0
Electroweak	5.3
Total background contribution	9.0

Table 8.5: Numbers of tag-probe pairs in the mass range $80 < M < 100$ GeV/c^2 for data (black points) and for different Monte Carlo contributions (solid lines). The contributions from the Monte Carlo are normalised to the integrated luminosity.

	N_{probe}	$N_{probe,HPTE}$	N_{probe}	$N_{probe,HPTE}$
	Barrel		Endcap	
Data	893	816	396	367
Z→ee	907.2	865.0	386.2	367.7
$t\bar{t}$	1.1	0.5	0.2	0.1
QCD	2.2	0	3.3	0.2
Electroweak	3.6	0.2	1.8	0.2
Total background contamination (MC)	6.8	0.7	5.2	0.5
Relative background (%) contamination (MC)	0.7	0.1	1.3	0.1

Table 8.6: Number of probes and number of HPTE probes in data, Monte Carlo Drell-Yan and background samples, in the mass range $80 < M < 100$ GeV/c^2. The total and relative background contamination from Monte Carlo are given.

$$\epsilon_{EB,Data} = \frac{816}{893} = (91.4 \pm 0.9)\% \tag{8.3}$$

$$\epsilon_{EE,Data} = \frac{367}{396} = (92.7 \pm 1.3)\% \tag{8.4}$$

As a comparison, the efficiencies determined from Monte Carlo, considering the Drell-Yan and the background contamination, using the 'Tag and Probe' method are:

$$\epsilon_{EB,MC} = (94.71 \pm 0.05\ (stat.) \pm 0.31\ (syst.))\% \tag{8.5}$$

$$\epsilon_{EE,MC} = (94.08 \pm 0.08\ (stat.) \pm 0.56\ (syst.))\%, \tag{8.6}$$

where the statistical errors are computed using binomial estimators and the total number of entries before normalization in order to account for the large statistics available in the Monte Carlo samples. The systematic errors are computed considering a 50% systematic error on the background contamination. The efficiency numbers from data and Monte Carlo agree reasonably well within the uncertainties. The efficiency in the barrel is a bit low compared to the Monte Carlo prediction.

The background contamination (low HPTE efficiency) introduces a bias in the determination of the efficiency from data, using the 'Tag and Probe' method. This bias is expected to be limited as the background contamination is small due to the E/p requirement on the 'tag'. To quantify this bias, the efficiencies are determined from the Monte Carlo Drell-Yan sample, using the 'Tag and Probe' method:

$$\epsilon_{EB,DY} = \frac{865.0}{907.2} = (95.35 \pm 0.05)\% \tag{8.7}$$

$$\epsilon_{EE,DY} = \frac{367.7}{386.2} = (95.22 \pm 0.08)\%, \tag{8.8}$$

the statistical errors are calculated using binomial estimators. The bias is defined as the difference between the efficiencies in equations 8.8 and 8.6. It is equal to 0.6% in the barrel and 1.1% in the endcaps. The efficiencies determined from data, using the 'Tag and Probe' method are corrected for the bias induced by the background contamination:

$$\epsilon_{EB,data} = (92.0 \pm 0.9\ (stat.) \pm 0.31\ (syst.))\% \tag{8.9}$$

$$\epsilon_{EE,data} = (93.8 \pm 1.3\ (stat.) \pm 0.56\ (syst.))\%, \tag{8.10}$$

where the systematic errors are computed considering a 50% systematic error on the background contamination (cf. eqns. 8.6).

The overall number is:

$$\epsilon_{Data} = (92.6 \pm 0.8\ (stat.) \pm 0.4\ (syst.))\% \tag{8.11}$$

8.3.2 Background estimation from Monte Carlo

The background contributions to the HEEP mass spectrum in the mass range $60 < M < 120$ GeV/c^2 are estimated from Monte Carlo expectations. In this section, in addition to the QCD (Em Enriched + b/c $\rightarrow$ e) samples and the $t\bar{t}$+jets samples from Madgraph, a set of

four QCD samples with no bias (4 $\hat{p_t}$ bins) and a $t\bar{t}$ sample from Pythia (cf. section 8.1) are used[4].

The different contributions (Monte Carlo + data) to the HEEP mass spectrum in the mass range $60 < M < 120$ GeV/c^2 are presented in table 8.7. The contributions from the Monte Carlo are normalised to the integrated luminosity. The contributions are added for the QCD (QCD Em Enriched + QCD b/c, 3 bins together), the QCD unbiased (4 bins together) and the $Z \rightarrow \mu\mu$, $Z \rightarrow \tau\tau$, $W + Jets$, WW and tW contributions (electroweak contribution).

The two predictions for the $t\bar{t}$ process are in good agreement while the predictions from the two Drell-Yan samples (Pythia, PowHeg) differ by around 3%. No event in the four QCD unbiased bins passes the HEEP selection.

If we consider only the QCD biased (Em + b/c $\rightarrow$ e), $t\bar{t}$+jets (Madgraph), electroweak and Drell-Yan (PowHeg) samples, according to the Monte Carlo, for an integrated luminosity of 2.77 pb^{-1} and using the HEEP selection described above (section 8.2.1), the total background contribution represents 0.73% of the di-electron invariant mass spectrum in the mass range $60 < M < 120$ GeV/c^2. This is smaller than the contribution obtained considering the whole mass range (section 8.2.3) as expected since we are at the Z peak.

Based on the numbers from table 8.7, the background contribution to the HEEP mass spectrum from data can be deduced. The statistical error is computed using a gaussian estimator and considering the weights of the different samples[5]:

$$N_b = 6.1 \pm 0.8 \ (stat.) \tag{8.12}$$

A 50% systematic error is considered in the determination of the background contribution.

8.3.3 Cross section estimation

The cross section is estimated at the Z peak ($60 < M < 120$ GeV/c^2), using the following formula:

$$\sigma = \frac{N - N_b}{A \ \epsilon \ L}, \tag{8.13}$$

where N is the total number of events in the final selected sample in data, N_b is the number of background events in data estimated from Monte Carlo (section 8.3.2). A is the acceptance, L the integrated luminosity and ϵ the total HEEP event selection efficiency. These numbers will be individually treated in separate sections.

Number of events

The total number of events in data that pass the HEEP selection in the mass range $60 < M <$ 120 GeV/c has been already evaluated (table 8.7). It is recalled here:

$$N = 815 \pm 29 \ (stat.), \tag{8.14}$$

where the statistical error is computed on basis of a gaussian estimator.

[4]The four QCD unbiased bins are generated with $\hat{p_t} > 15$, 30, 80, 170 GeV/c. In order not to count events twice when merging the four bins together, the events with $\hat{p_t} > 30$ GeV/c ($\hat{p_t} > 80$ GeV/c, $\hat{p_t} > 170$ GeV/crespectively) in the bin $\hat{p_t} > 15$ GeV/c ($\hat{p_t} > 30$ GeV/c, $\hat{p_t} > 80$ GeV/c respectively) are discarded.

[5]In the QCD Monte Carlo, a small number of pairs with large weights is present in the final sample, which explains the relatively large statistical error.

Sample	N
Data	815
Z→ee (PowHeg,NLO)	838.5
Z→ee (Pythia,LO)	815.0
$t\bar{t}$+jets (Madgraph)	1.0
$t\bar{t}$ (Pythia)	1.0
QCD $\hat{p}_t$ [20-30]	0
QCD $\hat{p}_t$ [30-80]	0
QCD $\hat{p}_t$ [80-170]	0
QCD $\hat{p}_t$ [170-∞]	0
QCD unbiased	0
QCD Em Enriched $\hat{p}_t$ [20-30]	0.1
QCD Em Enriched $\hat{p}_t$ [30-80]	3.3
QCD Em Enriched $\hat{p}_t$ [80-170]	0.3
QCD b/c → e $\hat{p}_t$ [20-30]	0
QCD b/c → e $\hat{p}_t$ [30-80]	0
QCD b/c → e $\hat{p}_t$ [80-170]	0
QCD biased	3.7
Z→$\mu\mu$	0
Z→$\tau\tau$	0.7
W+Jets	0.5
WW	0.2
tW	0.0
Electroweak	1.4
Total background contribution	6.1
Total relative background contribution (%)	0.7

Table 8.7: Contributions of different processes to the HEEP mass spectrum at the Z peak ($60 < M < 120$ GeV/c^2) from data (black points) and Monte Carlo (solid lines). The contributions from the Monte Carlo are normalised to the integrated luminosity.

Number of background events

The background contribution to the HEEP mass spectrum from data in the mass range 60 $< M <$ 120 GeV/c^2 has been evaluated from Monte Carlo (section 8.3.2). The result is recalled here:

$$N_b = 6.1 \pm 0.8\ (stat.) \pm 3.1\ (syst.), \tag{8.15}$$

where the statistical error is computed on basis of a gaussian estimator and a 50% systematic error is considered.

HEEP event selection efficiency

The HEEP event selection efficiency is defined as the ratio $N_{HEEP}/N_{DY,acc}$ where N_{HEEP} is the number of HEEP events and $N_{DY,acc}$ is the number of Drell-Yan events where both electrons are emitted in the tracker acceptance, out of the gap and have a transverse energy higher than 25 GeV. The HEEP event selection is equal to the square of the total electron selection efficiency:

$$\epsilon_{HEEP} = (\epsilon_{elec})^2 \tag{8.16}$$

Accordingly, the total electron selection efficiency is defined as the ratio N_{ele}/N_{acc} where N_{acc} is the number of electrons from Drell-Yan events emitted in the tracker acceptance, out of the gap, with transverse energy higher than 25 GeV and N_{ele} is the number of such electrons which pass the total electron selection. The clustering efficiency has been determined from Monte Carlo in section 6.1.3 and is considered here to be 100%. The 'trigger' efficiency is the efficiency for an electron from a Drell-Yan event to fire the triggers mentioned in section 8.1. For an electron with p_t higher than 15 GeV/c, it has been shown [62] that this efficiency is 100%. The total electron selection efficiency is factorised as follows:

$$\epsilon_{elec} = \epsilon_{GSF} \times \epsilon_{HPTE}. \tag{8.17}$$

The electron reconstruction (GSF) efficiency is the fraction of electrons, from Drell-Yan events, in the kinematic regime specific to this study, which are reconstructed as electron candidates (GSF electrons). This efficiency has been determined from Monte Carlo in section 6.2.4, using Pythia Drell-Yan samples. It is reevaluated here using the PowHeg Drell-Yan sample:

$$\epsilon_{GSF} = (97.4 \pm 0.0)\%, \tag{8.18}$$

the statistical errors are computed using binomial estimators.

The high p_t electron (HPTE) selection efficiency is the fraction of electron candidates that pass the HPTE criteria described in table 8.2. This efficiency has been determined from data, using the 'Tag and Probe' method, in section 8.3.1:

$$\epsilon_{HPTE,TP} = (92.6 \pm 0.8\ (stat.) \pm 0.4\ (syst.))\% \tag{8.19}$$

The total electron selection efficiency for electrons from Drell-Yan events in the kinematic regime specific to this study is evaluated:

$$\epsilon_{electron} = (90.2 \pm 0.8\ (stat.) \pm 0.4\ (syst.))\%, \tag{8.20}$$

the total statistical and systematic errors are obtained independently by adding quadratically the statistical and systematic errors of the individual contributions.

Acceptance

The acceptance is defined as the ratio $N_{DY,acc}/N_{DY}$ where $N_{DY,acc}$ is the number of Drell-Yan events in the mass range $60 < \mathrm{M} < 120$ GeV/c^2, for which both electrons are reconstructed as superclusters in the tracker acceptance, out of the gaps, with transverse energy higher than 25 GeV and N_{DY} is the total number of Drell-Yan events in the same mass range. This quantity is determined from Monte Carlo, using samples from both generators (Pythia and PowHeg):

$$A_{PowHeg} = (37.02 \pm 0.07\ (stat.))\% \tag{8.21}$$

$$A_{Pythia} = (36.77 \pm 0.04\ (stat.))\% \tag{8.22}$$

the statistical errors are computed using binomial estimators. As the simulation from PowHeg accounts for next-to-leading order effects, the acceptance from PowHeg is considered as the final result.

Luminosity

The total integrated luminosity considered in this analysis is 2.77 pb^{-1}. The measurement of the luminosity is the task of a specific group [64] in the CMS collaboration and is common to all physics analysis groups in CMS. The systematic errors on the measured luminosity have been estimated here [65] and a total systematic error of 11% has been found. This systematic error is quoted as a separate error (lumi.).

$$L = (2.77 \pm 0.31)\ \mathrm{pb}^{-1} \tag{8.23}$$

Final result for cross section

The final result for the measured Drell-Yan cross section at the Z peak ($60 < M < 120$ GeV/c^2) in the di-electron channel, from 2.77 pb^{-1}, using the HEEP selection is:

$$\sigma_{\gamma^*/Z} \times BR(\gamma^*/Z \rightarrow e^+e^-) = (970 \pm 34\ (stat.) \pm 37\ (syst.) \pm 107\ (lumi.))\ \mathrm{pb}. \tag{8.24}$$

The statistical errors are added quadratically and the error on the luminosity is quoted as a separate error. The major source of uncertainty comes from the limited amount of statistics used in this study.

As a comparison, the theoretical predictions to the cross section $\sigma_{\gamma^*/Z} \times BR(\gamma^*/Z \rightarrow e^+e^-)$ at the Z peak ($60 < M < 120$ GeV/c^2), at leading order and next-to-leading order are:

$$\sigma_{TH,\gamma^*/Z} \times BR(\gamma^*/Z \rightarrow e^+e^-) = 740\ \mathrm{pb}\ (LO) \tag{8.25}$$

$$\sigma_{TH,\gamma^*/Z} \times BR(\gamma^*/Z \rightarrow e^+e^-) = 911\ \mathrm{pb}\ (NLO) \tag{8.26}$$

The Drell-Yan cross section measurement of eqn. 8.24 is in agreement with the theoretical prediction within uncertainties.

Chapter 9

Conclusions

In this thesis, the study of the Drell-Yan invariant mass spectrum was performed with a focus on possible deviations from the Standard Model. Such deviations are implied by certain theories beyond the Standard Model which predict the existence of new heavy bosons (Z' or heavy gravitons) at the TeV scale which can decay into an electron-positron pair. Such theories have been presented in chapter 1 with a focus on the extra-dimension scenario and the grand unification theory models. Exclusion limits, set by the Tevatron, have been put for the graviton and the Z', using 5.4 fb^{-1} and 3.6 fb^{-1} of data respectively. The Z'_{SSM} is excluded at $M < 1071$ GeV/c^2 and the graviton with coupling c = 0.1 is excluded at 1050 GeV/c^2, at 95% C.L.

This study was performed in the framework of the CMS experiment at the LHC. Its centre of mass energy of 7 TeV and its instantaneous luminosity of 10^{31} cm^{-2}s^{-1} provide a good sensitivity to the study of the Drell-Yan invariant mass spectrum as well as for the search for new heavy bosons decaying into an electron pair (chapter 2). In addition, the excellent energy resolution of the CMS electromagnetic calorimeter, 0.6% at $E \gtrsim 100$ GeV and the performance of the tracker to reconstruct the trajectory followed by the electron in the intense magnetic field of 3.8 T bring much precision in the detection of high energy electrons from heavy boson decays (chapter 3).

The Monte Carlo simulation tools have been extensively used in this study. They simulate the physics of Drell-Yan events and give an estimate of the cross section. They have been introduced in chapter 4 together with the kinematics of the Drell-Yan events. The background processes that can contribute to the final di-electron sample were also presented and their cross sections were given. The most important ones come from QCD multijet processes where at least two jets fake each an electron and from the $t\bar{t}$ process where two real electrons are present in the final state. The cross section for Drell-Yan events from pp collisions at 7 TeV, with mass above 40 GeV/c^2 is 803 pb, at leading order. As a comparison the cross section for QCD dijets with transverse momentum of the initial partons in their centre of mass frame ($\hat{p_t}$) higher than 15 GeV is of the order of 0.1 mb.

To search for possible deviations from the Standard Model, a specific analysis strategy has been designed (chapter 5) by the HEEP group. This strategy has been tested on pseudo-experiments performed on Monte Carlo simulation data, considering a 10 TeV centre of mass energy and an integrated luminosity of 100 pb^{-1}. It relies primarily on a baseline selection (high p_t electron selection) to discriminate as much as possible the Drell-Yan events from the background contributions. Three different regions in the invariant mass spectrum are exploited. The Z peak region ($60 < \mathrm{M} < 120$ GeV/c^2) is used to extract efficiencies using a data-driven approach. The high mass region ($120 < \mathrm{M} < 600$ GeV/c^2), where no new physics is expected, is used to measure the Drell-Yan cross section and compare to the theoretical

prediction. The discovery region (M > 600 GeV/c^2) is devoted to the direct search for new resonances. To measure the background contributions using a data-driven approach, three different methods have been tested and compared and have shown good agreement. The Drell-Yan invariant mass spectrum has been presented, including the Drell-Yan and background processes as well as a Z'_{SSM} at 1 TeV/c^2. The agreement in shape, in the high mass range, between measurement and expectation, is quantified by a fit in 5 bins in mass, giving a probability of 63.3%. Finally, the discovery potential and exclusion limits in the absence of signal are estimated at 10 TeV and 7 TeV. For a 7 TeV centre of mass energy and considering an integrated luminosity of 500 pb^{-1} expected after the whole data-taking period at 7 TeV, in 2010 and 2011, the 5σ discovery reach is possible for Z'_{SSM} bosons with mass up to 1.5 TeV/c^2 and for Randall-Sundrum gravitons with mass up to 1.1 TeV/c^2 and 1.3 TeV/c^2, for a coupling parameter $c = 0.05$ and $c = 0.1$ respectively. In case no signal is observed, lower mass limits can be put at 1.7 TeV/c^2 for a Z'_{SSM} boson, 1.2 TeV/c^2 and 1.5 TeV/c^2 for Randall-Sundrum gravitons with a coupling parameter $c = 0.05$ and $c = 0.1$ respectively.

The high p_t electron selection has been described in chapter 6. The first step, the clustering, estimates the electron energy and direction based on its energy deposit in the ECAL. The energy resolution is affected by the energy loss in the tracker material budget through photon radiation. The second step, the electron reconstruction, requires the presence of a track in the tracker with information compatible with the information from the energy deposit in the ECAL. The different parts of the electron reconstruction are explained and the efficiencies are given. The third step requires the electron to fulfill predefined identification and isolation criteria. The former criteria require more stringent compatibility, compared to the second step, between the information from the track and the information of the energy deposit from the ECAL. The latter criteria require that limited activity in terms of transverse momenta of tracks and ECAL/HCAL energy deposits be found around the electron. In addition a cut at 25 GeV on the electron transverse energy is applied.

As first collisions at LHC took place, the necessity came to perform a fast check of online and offline data. A data quality monitoring tool (DQM), specific to high p_t electrons, was developed by the HEEP group to detect detector problems, compare data and Monte Carlo expectations for Monte Carlo tuning and search for possible deviations from the Standard Model (chapter 7). This DQM tool was developed in the framework of the CMS software version 3_5_X. Discrepancies in the barrel were observed, due to the presence of fake energy deposits contained in a single crystal.

Chapter 8 was dedicated to the analysis of the first LHC data from proton-proton collisions at 7 TeV centre of mass energy. The data collected by the CMS detector from 30/03/2010 till 30/08/2010 (Runs 132440 $\rightarrow$ 144114) were used, corresponding to a total integrated luminosity of 2.77 pb^{-1}. The study of the Drell-Yan invariant mass spectrum in the di-electron channel has been performed. The distribution of the kinematic, identification and isolation variables have been compared to Monte Carlo and in general good agreement was found. A discrepancy between data and Monte Carlo predictions was found in the endcaps for the case of the $\Delta\eta$ variable which quantifies the consistency between the track pseudorapidity and the energy deposit pseudorapidity of the electron candidate. This is related to the alignment, in the endcaps, between the tracker and the ECAL. The criterion on this variable was removed in the endcaps in the final selection. The Drell-Yan invariant mass spectrum from data was compared to the Monte Carlo and a shift was spotted in the data spectrum with respect to the Monte Carlo, due to calibration issues. These two effects have been corrected in the reconstruction part of the new versions of the CMS software (CMSSW 3_8_X). The background contributions were estimated from Monte Carlo expectations and the efficiencies

were extracted using the 'Tag and Probe' method. The cross section was finally estimated at the Z peak:

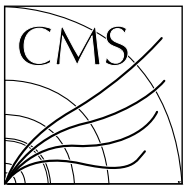
$$\sigma_{\gamma^*/Z} \times BR(\gamma^*/Z \rightarrow e^+e^-) = (970 \pm 34\ (stat.) \pm 37\ (syst.) \pm 107\ (lumi.))\ \text{pb}. \qquad (9.1)$$

As a comparison, the theoretical predictions to the cross section $\sigma_{\gamma^*/Z} \times BR(\gamma^*/Z \rightarrow e^+e^-)$ at the Z peak ($60 < M < 120$ GeV/c^2), at leading order and next-to-leading order are 740 pb (leading order) and 911 pb (next-to-leading order). The Drell-Yan cross section measurement (eqn. 9.1) is in agreement with the theoretical prediction within uncertainties.

Appendix A

Tracker isolation studies: rejection of Bremsstrahlung photon conversion tracks around high energy electrons

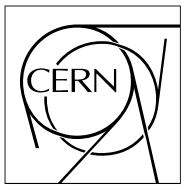
Available on CMS information server **CMS AN 2009/072**



The Compact Muon Solenoid Experiment

Analysis Note

The content of this note is intended for CMS internal use and distribution only



07/05/2009

abstract This note presents a study of the tracker activity around high energy electrons. Using Drell-Yan events with high mass ($M > 200$ GeV/c^2), this activity is observed to be higher than in the underlying event, which is attributed to tracks from Bremsstrahlung photon conversion. An improved tracker isolation procedure for the identification of electrons with $p_t > 50$ GeV/c is proposed, using only tracks associated to the primary vertex. The performance of the new tracker isolation criteria is studied for electrons from Drell-Yan events and for fake electrons from QCD jets. This work is a contribution to the CMS HEEP (High Energy Electron Pairs) group effort.

A.1 Introduction

The study presented in this note is performed in the framework of the HEEP (High Energy Electron Pairs) group [66] which has defined a complete strategy to search for the electron-positron decay channel of new heavy neutral bosons [40, 67]. The electron reconstruction and identification at high energy is a crucial part of the HEEP analysis. In this selection, the isolation cuts are powerful tools to suppress the QCD jet backgrounds, where a jet with a high electromagnetic deposit in the electromagnetic (ECAL) calorimeter with an additional track associated to the cluster fakes an electron. Isolation criteria are designed

using activities around the electron candidate direction in the hadronic calorimeter (HCAL), the electromagnetic calorimeter (ECAL) and the tracker detectors, with a strong correlation between these three detectors.

This note presents a study of the activity in the tracker detector around high energy electrons ($p_t > 50$ GeV/c), and in particular the effect of the conversion in the tracker material of Bremsstrahlung photons emitted by the electron. It complements the calorimeter isolation studies performed by the HEEP group in [51].

The analysis is performed within the CMSSW_1_6_11 framework using the two following datasets : (i) Drell-Yan events with generated mass above 200 GeV/c^2 from CSA07 production CMSSW_1_6_7, and (ii) QCD events from the CMSSW_1_6_7 CSA07 gumbo soup production, gathering $\hat{p}_t$ bins up to 3500 GeV/c.

Before studying the tracker activity around high energy electrons, it is worth noting that the tracker underlying event activity is uniform in η and ϕ. This is observed in Figure A.1, which presents the $\eta - \phi$ density of all reconstructed tracks (from the "ctfWithMaterialTracks" collection) with $p_t > 1.5$ GeV/c for all generated Drell-Yan events with mass above 200 GeV/c^2. The decrease of the track density at high values of $|\eta|$ ($|\eta| \gtrsim 2.5$) is due to the drop of the track reconstruction efficiency caused by the decrease of the acceptance.

The note is organised as follows. Section A.2 presents a detailed analysis of the characteristics of the tracker activity around high energy electrons, using Drell-Yan events (decaying into an electron-positron pair) with mass above 200 GeV/c^2 : the high energy electron selection, the definitions of the tracker isolation variables and the study of the tracker activity as a function of various quantities are reported in subsections A.2.1, A.2.2 and A.2.3, respectively; subsection A.2.4 complements the study of the tracker activity using information concerning the simulation details. In section A.3, new tracker isolation criteria are presented, using only tracks associated to the primary vertex. The performance of the proposed criteria for the identification of high energy electrons from Drell-Yan events and the rejection of fake electrons from QCD jets is studied in section A.4. Finally, conclusions are given in section A.5, followed by appendices.

A.2 Tracker activity around high energy electron direction

In this section, the high energy electron selection is detailed, followed by the definition of the tracker isolation variables. The tracker activity around the electron direction is then studied using Drell-Yan events generated with mass above 200 GeV/c^2 and is compared to the activity expected from the underlying event. Finally the last subsection complements the tracker isolation studies using the full simulated information.

The activity in the underlying event is estimated in the direction at 90 $^\circ$ in ϕ from the electron direction. Note that NLO QCD effects, such as gluon radiation from quarks, may lead in Drell-Yan event to the presence of a recoil jet back to back in p_t with the Z boson, which may produce strong activity in the tracker. However the Z decay electron and positron directions are not correlated to that of the Z boson, nor of this recoil jet.

A.2.1 Electron selection

Throughout the note, the standard object used for reconstructed electrons is the "PixelMatchGsfElectron" (GSF electron) built by the "pixelmatchgsf algorithm". The latter gathers information from energy deposits in the ECAL (energy and position) with information provided by the track (momentum and direction), seeded with pixel tracker hits. Cuts are

applied on the cluster-track matching in η and in ϕ, and a loose cut $H/E < 0.2$ is applied, where H/E represents the ratio of the energy deposit in the HCAL calorimeter behind the ECAL cluster attributed to the electron candidate, to the energy of the latter. Details of the pixelmatchgsf algorithm, definitions of the cuts and cut values can be found in [47].

The following additional cuts are applied to the electron candidate everywhere in this paper : the GSF electron should be reconstructed with a transverse momentum $p_{t,GSF} > 50$ GeV/c, in the tracker acceptance ($|\eta_{\rm GSF}| < 2.5$) and outside the gap between the ECAL barrel and endcaps ($1.442 < |\eta| < 1.560$).

A.2.2 Definition of tracker isolation variables

The tracker activity is studied in a cone with radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ around the electron candidate direction. In order not to count the electron track itself, an inner veto cone is used. The tracker activity is quantified by the number of tracks between the two cones and by the scalar sum of their transverse momenta. Tracks are taken into account if they satisfy in addition two quality criteria : a minimum transverse momentum to suppress fake tracks, and a maximum emission distance $|\Delta z|$ along the beam direction between the vertex position and the emission point of the track, computed as corresponding to that of the closest approach to the beam. It will be seen in the following that the tracker activity is very sensitive to the cut values chosen, in particular for the inner veto cone radius and for the maximum emission Δz distance.

The tracker activity is computed for an outer cone around the electron direction with radius $\Delta R_{\rm out} = 0.2$ and an inner veto cone with radius $\Delta R_{\rm in} = 0.02$. All reconstructed tracks from the "ctfWithMaterialTracks" collection with transverse momenta higher than 1.5 GeV/c and maximum emission distances $|\Delta z| < 0.1$ cm are considered.

The activity between the two cones is described by the following three main variables [1] : (i) the number of tracks, N_{tr}; (ii) the average value of the transverse momentum of the tracks, $\langle p_t \rangle$; (iii) the scalar sum of the transverse momenta of the tracks, Σp_t.

The same selections and variable definitions apply for the study of the tracker activity in the underlying event.

A.2.3 Dependence of the tracker isolation variables on η, Δz and cone sizes

In this section, the dependences on η, Δz and cone sizes of the isolation variables are presented for high energy electrons from Drell-Yan events generated with mass above 200 GeV/c^2. After applying the electron reconstruction as defined in section A.2.1, a total of 33457 electrons is used for the study.

Figure A.2 shows the average number of tracks between two cones around the electron direction (a) and at 90 $^\circ$ in ϕ (underlying event) (b), as a function of the pseudorapidity of the electron, $\eta_{\rm GSF}$. The activity around the electron direction is significantly larger at high values of $|\eta_{\rm GSF}|$ than in the underlying event. This behaviour is not consistent with the uniform density shown in Figure A.1. It will be shown that this difference is explained in terms of Bremsstrahlung effects. Tracks due to the conversion in the tracker material of Bremsstrahlung photons emitted by the electron are indeed expected to be close to the electron direction, and absent from the underlying event. To the contrary, the activity observed

[1] It should be noted that another tracker isolation variable, mentioned in the past, the ratio of Σp_t to the transverse momentum of the GSF electron, is not used here, since it was shown not to be performant [51].

at low values of $|\eta_{\rm GSF}|$ is low and similar in both regions (an average of 0.05 track per $\Delta\eta$ bin of width 0.2), as expected for isolated electrons.

The dependency of this activity with respect to the inner cone radius is presented in Figure A.3, which shows the average number of tracks around the electron direction as a function of $\eta_{\rm GSF}$ for three different veto cone radii $\Delta R_{\rm in}$: (a) 0.02, (b) 0.04 and (c) 0.07. The extra activity observed at high $|\eta|$ decreases significantly to become uniform in η for a veto cone radius of 0.07. This shows that the extra activity is strongly concentrated along the electron direction. The average value of the transverse momentum of tracks between the two cones is shown in Figure A.3 (d−f) as a function of $\eta_{\rm GSF}$ for the three veto cone sizes. For low values of $|\eta|$, the transverse momenta are roughly the same for the three cases (around 3 GeV/c), whereas it is larger for higher values of $|\eta|$. All these observations are consistent with the Bremsstrahlung hypothesis.

The estimation of the tracker activity also relies on the $|\Delta z|$ cut value. Figure A.4 shows the average number of tracks around the electron direction as a function of $\eta_{\rm GSF}$ for $|\Delta z| <$ 0.1 cm and $|\Delta z| < 0.3$ cm. The extra activity at high $|\eta|$ increases even more when relaxing the $|\Delta z|$ cut whereas the activity remains the same in both cases for low values of $|\eta_{\rm GSF}|$. This is understood by the fact that, for similar precision of the angle measurement, high $|\eta|$ tracks have a larger uncertainty on the position of their emission point than low $|\eta|$ ones. Therefore, relaxing the $|\Delta z|$ cut allows more tracks to be accounted for.

The increase of activity at high $\eta_{\rm GSF}$ thus results from two effects: the increase of tracker material with η (with a maximum value at $|\eta| \gtrsim 1.5$) and the pointing effect described in the previous paragraph.

A.2.4 Check of the Bremsstrahlung hypothesis using simulated track information

In order to check the hypothesis that the extra activity observed along the electron direction at high values of $|\eta_{\rm GSF}|$ comes from tracks due to the conversion of Bremsstrahlung photons emitted by the electron candidate, information from the GEANT simulation [35] is used. For each reconstructed track, the associated simulated track is determined, using the package "TrackAssociatorByHits" within the CMSSW_1_6_11 framework. The simulated track information gives access to the identification of the simulated particle and of its mother. More details on this package can be found in [68].

Figure A.5 presents an r_t — z view of the vertex positions of the simulated particles associated to the reconstructed tracks around the electron direction, where r_t is the distance between the vertex and the beam direction in the transverse plane. A large number of tracks is observed not to originate from the primary vertex (vertices represented by red points), with only a fraction of the emission vertices corresponding to the primary vertex with small r_t and $|z|$ values (black points). In practice, these two categories are established using the information on the mother particle : if the number of mother particles for the particle giving the simulated track is equal to zero, the simulated track is considered to originate from the primary vertex, otherwise it is considered to originate from a secondary vertex. The vertex positions indicate that the extra activity at high $|\eta|$ in the previous plots is due to tracks which do not originate from the primary vertex emission, with a small contribution from tracks which originate from the primary vertex. The tracker structure is indeed visible in Figure A.5, where the majority of (red) secondary vertices are located along a direction corresponding to $|\eta| \simeq 1.5$.

Figures A.6(a) and (b) show the identifier code of the particles corresponding to the tracks

reconstructed around the electron direction, for tracks which originate (a) or do not originate (b) from the primary vertex. Most particles which do not originate from the primary vertex are electrons, as expected from photon conversion, whereas pions are mostly emitted from the primary vertex, as expected for the underlying event. Figure A.6(c) shows the identifier code of the mother particles for tracks not coming from the primary vertex, of which the major part consists in electrons and photons.

Therefore, the extra activity at high $|\eta|$ is mainly due to the conversion of Bremsstrahlung photons emitted by the primary electron.

A.3 New tracker isolation criteria using only tracks associated to the primary vertex

In the previous sections, the tracker activity between two cones around the electron direction was studied as a function of the η_{GSF}. An increase of the activity at high $|\eta|$ was observed, implying a dependence of the tracker activity on the $|\Delta z|$ cut value and the size of the inner cone. It was showed that this extra activity mainly consists in electron tracks not associated to the primary vertex, but coming from the conversion of Bremsstrahlung photons emitted by the primary electron. Considering only reconstructed tracks associated to the primary vertex when estimating the tracker activity thus allows us to suppress this extra activity at high $|\eta|$.

The association of the reconstructed tracks to the primary vertex is performed using the primary vertex finder algorithm [69]. Once tracks are reconstructed, they are clustered based on the z_0 coordinate of their point of closest approach to the beam. Vertices are then formed, based on this clustering of tracks, and each track is associated to the corresponding vertex. The vertices are then sorted by decreasing sum of the transverse momenta squared Σp_t^2 of the associated tracks. The first of these vertices, i.e. the vertex with the largest Σp_t^2, is tagged as the primary vertex.

A new tracker isolation is proposed in this paper, where only tracks associated to the primary vertex are taken into account to compute the tracker isolation variables defined in section A.2.2. For this, the radii of the inner veto cone and of the outer cone are $\Delta R_{\mathrm{in}} = 0.02$ and $\Delta R_{\mathrm{out}} = 0.2$, and the minimum transverse momentum $p_t > 1.5$ GeV/c cut is required, as above. By definition, no $|\Delta z|$ cut is imposed.

Figures A.7(a), (b) and (c) present, as a function of η_{GSF}, the average number of tracks, their average transverse momentum and the average scalar sum of the transverse momenta of all tracks between the two cones around the GSF electron direction, according to the new tracker isolation definition. The corresponding distributions for the underlying event, estimated between two similar cones at 90 $^\circ$ in ϕ off the electron candidate direction, are shown in Figures A.7(d), (e) and (f), respectively. The extra activity at high $|\eta|$ observed in previous sections is no longer visible, and the variable distributions are similar around the electron direction and for the underlying event.

The activity is studied with respect to different inner cone radii ΔR_{in} in Figure A.8, which presents, as a function of η_{GSF}, the average number of tracks around the electron direction for $\Delta R_{\mathrm{in}} = 0.02$ (a), 0.04 (b) and 0.07 (c). The shapes are similar in the three cases. The small decrease of the average number of tracks observed when increasing the inner cone size is due to the fact that the area between the two cones is slightly reduced.

In summary, the new tracker isolation definition, using only tracks associated to the primary vertex, has the advantage to present a uniform η dependence, similar to the one for

the underlying event. The tracker isolation variables are independent of the inner cone radius $\Delta R_{\rm in}$ and, by definition, they are independent of the choice of the $|\Delta z|$ cut.

A visualisation of the tracker activity is given in Appendix C for a high mass Drell-Yan event. The positions of the tracker hits both at the simulated and at the reconstructed levels are displayed.

A.4 Performance of the tracker isolation algorithm

In this section, the performance of the new tracker isolation algorithm is checked and compared to the old HEEP tracker isolation algorithm (version 1.2) used in the analysis of [51].

The old HEEP tracker isolation algorithm requires that the scalar sum, Σp_t, of the transverse momenta of all tracks with $p_t > 1.5$ GeV/c, between two cones of radii 0.02 and 0.2 around the electron direction, is less than 7.5 GeV/c in the ECAL barrel and 15.0 GeV/c in the ECAL endcaps. The higher cut value in the endcaps is the consequence of the higher activity at high $|\eta_{\rm GSF}|$ observed in section A.2.3.

The new tracker isolation algorithm proposed in this note uses the same cone sizes and minimum track p_t, but considers only the tracks associated to the primary vertex in the estimation of the tracker activity, and thus does not require any Δz cut.

The tracker activity performance is studied both in terms of the identification efficiency for high energy electrons from Drell-Yan production with mass above 200 GeV/c^2, and of the rejection power for fake electrons from jets, using QCD jet samples. For the new tracker isolation algorithm two variables are presented, the number of tracks, N_{tr}, and the scalar sum of transverse momenta of all tracks, Σp_t, between two cones of radii 0.02 and 0.2 along the electron direction.

In the case of QCD events, more activity is expected around fake GSF electrons from jets. In order to check this hypothesis, the QCD samples defined in section A.1 are used. Note that the QCD events used here are preselected with the $e\gamma$ trigger [70], which have different requirements depending on the electron candidate transverse momentum : the Very High E_t trigger only requires an ECAL energy deposit with $p_t > 200$ GeV/c; the High E_t trigger requires an ECAL deposit with $p_t > 80$ GeV/c and loose isolation requirements in the ECAL, the HCAL and the tracker, no track matching with the ECAL energy deposit being required; the Relaxed Single Electron trigger requires an energy deposit with $p_t > 18$ GeV/c in the ECAL calorimeter, direction and energy matching between the ECAL deposit and a track in the tracking detector, and isolation conditions in the HCAL and the tracker.

The electron reconstruction and selection procedure is applied as defined in section A.2.1, which requires in particular for the fake GSF electron candidate $p_{t,GSF} > 50$ GeV/c and $H/E < 0.2$. This leads to a total of 409390 fake electron candidates in the QCD jet samples considered in this study. These fake electrons are characterised by an atypically high electromagnetic fraction, a low H/E ratio and the presence of a track associated to the electromagnetic cluster.

Figures A.9(a), (b) and (c) present, for QCD events, the average number of tracks, the average transverse momentum of tracks and the average scalar sum of the transverse momenta of tracks around the fake electron, as a function of $\eta_{\rm GSF}$, according to the new tracker isolation criteria. The corresponding distributions for the underlying event, taken at 90 $^\circ$ in ϕ from the electron candidate direction, are shown in Figures A.9(d), (e) and (f) respectively. Along the fake electron direction, much more activity is observed than in the Drell-Yan case, with an average around 2 tracks instead of 0.05 and an average sum of transverse momenta, Σp_t, around 15 GeV/c instead of 0.15 GeV/c, per bin in $\Delta\eta = 0.2$. The underlying event distribution shows much less activity compared to that around the fake electron, (the errors

	Drell-Yan efficiency ϵ		QCD background rejection $1-\epsilon$			Drell-Yan efficiency ϵ		QCD background rejection $1-\epsilon$	
	barrel	endcap	barrel	endcap		barrel	endcap	barrel	endcap
$N_{tr} = 0$	0.96	0.97	0.71	0.60	$\Sigma p_t < 2$ GeV/c	0.98	0.98	0.66	0.53
$N_{tr} \leq 1$	0.99	1.00	0.48	0.33	$\Sigma p_t < 3$ GeV/c	0.99	0.99	0.59	0.47
$N_{tr} \leq 2$	1.00	1.00	0.32	0.18	$\Sigma p_t < 4$ GeV/c	0.99	0.99	0.56	0.44
$N_{tr} \leq 3$	1.00	1.00	0.23	0.11	$\Sigma p_t < 5$ GeV/c	1.00	0.99	0.54	0.40
					$\Sigma p_t < 6$ GeV/c	1.00	1.00	0.51	0.37

Table A.1: Efficiencies of the new tracker isolation criteria for high energy electrons from Drell-Yan events with $M > 200$ GeV/c^2, and rejection power for fake electrons from QCD jet events, for several cut values of the isolation variables, N_{tr} and Σp_t, separately for the GSF electron cluster situated in the ECAL barrel and endcaps.

on the average values are computed following the procedure defined in Appendix B, to take into account the weighting of the event samples).

The tracker isolation efficiency (ϵ) for high energy electrons from Drell-Yan events and the rejection power (1-ϵ) of fake electrons from QCD jet events are summarized in Table A.1, for the new tracker isolation criteria. The results are presented separately for GSF electron clusters situated in the ECAL barrel and endcaps. The efficiency is defined as the number of GSF electrons passing the tracker isolation cut divided by the total number of GSF electrons, as defined in section A.2.1. The corresponding plots, as a function of $p_{t,GSF}$ for several cut values on N_{tr} and Σp_t, can be found in section A.7. One observes that for $N_{tr} < 2$ and $\Sigma p_t <$ 3 GeV/c efficiencies are close to 99% for Drell-Yan events while typically between 30 and 60 % of the fake electrons from QCD jet are rejected; note that $\Sigma p_t < 3$ GeV/c implies $N_{tr} <$ 2, due to the minimum track p_t of 1.5 GeV/c in the estimation of isolation. The efficiencies are similar to those for the old tracker isolation criteria [51].

In the rest of this section, the efficiencies of the new isolation criteria are given, after having applied the criteria of the HEEP electron selection defined in [67, 51] and summarised in section A.6. In particular, the ECAL isolation criterion implies a cut on the sum of the transverse energy deposits measured in all ECAL cells contained within a cone of radius $\Delta R = 0.3$ around the electron cluster barycentre, except for the cells belonging to the electron candidate cluster. The HCAL isolation criterion includes a cut on the sum of the transverse energy deposits in all HCAL cells contained between two cones of radii $\Delta R = 0.1$ and 0.3 around the electron candidate direction; for $|\eta| > 1.560$ (endcap region), the longitudinal segmentation of the HCAL is used, considering separately energy deposits in the HCAL section close to (away from) the interaction point, called HCAL Depth 1 (HCAL Depth 2).

The efficiencies for high energy electrons from Drell-Yan events and the rejection powers for fake electrons from QCD jet events, after having applied the other criteria of the HEEP selection, are given in Table A.2, both for the new and old isolation criteria, for $p_{t,GSF} > 50$ GeV/c and $p_{t,GSF} > 200$ GeV/c. The efficiencies for the new isolation criteria are observed to be uniform as a function of $p_{t,GSF}$, as shown in Figures A.12 and A.13 in section A.7, and are similar to those for the old tracker isolation criteria [51].

In view of the results discussed here, it is proposed to use tracker isolation criteria defined as the following : the scalar sum of the transverse momenta of tracks with $p_t > 1.5$ GeV/c, associated to the primary vertex and contained between two cones of radii $\Delta R_{in} = 0.02$ and

$\Delta R_{out} = 0.2$, must be smaller than 3 GeV/c :

$$\Sigma\ p_t < 3\ \text{GeV/c} \tag{A.1}$$

The correlation between the isolation criteria designed using activities around the electron candidate direction in the hadronic calorimeter (HCAL), the electromagnetic calorimeter (ECAL) and the tracker detector are studied. For this, the HEEP electron selection criteria defined in section A.6 are applied, except for the isolation criteria (four last rows in Table A.5).

The cumulative efficiencies for various cuts are defined as the efficiencies when the cuts are applied successively. The correlations between the calorimeter isolation criteria, ECAL + HCAL depth 1 (HCAL1) and HCAL depth 2 (HCAL2, only for the endcap region), and the new tracker isolation criterion are presented in Tables A.3 and A.4, which gives the cumulative efficiencies of the various detector isolation cuts for Drell-Yan events and the cumulative rejections for fake electrons in QCD jet samples, after having applied the other HEEP selection criteria.

In the barrel (endcaps), the new tracker isolation criterion alone rejects 1 (1) % of the signal electrons from Drell-Yan events and around 60 (47) % of the fake electrons from QCD jet events. The ECAL + HCAL depth 1 calorimeter isolation criterion is most efficient in rejecting QCD fake electrons and it rejects about 6-8% of the signal electrons from Drell-Yan events; it is also highly correlated with the two other criteria. Note that, even if they are correlated, the three detector isolation criteria are more powerful when they are all applied, with a total signal efficiency around 91 (92) % for electrons from Drell-Yan events, and a background rejection power around 89 (77) % for fake electrons from QCD jet events.

A.5 Summary and Conclusions

Tracker isolation criteria for the validation of high energy electrons have been studied in the framework of the High Energy Electron pairs (HEEP) analysis.

The tracker activities between two cones around the electron direction and in the underlying event (defined in a direction shifted in ϕ by 90 °) have been compared, showing an extra activity around the electron direction at high $|\eta|$. This activity has been studied with respect to the size of the inner isolation cone and the distance $|\Delta z|$ between the primary vertex and the emission point of the particule giving the track. A clear dependency on these parameters was found.

Using the simulated information from GEANT, it was shown that the extra activity around the electron direction at high $|\eta|$ comes from additional tracks due to the conversion in the tracker material of Bremsstrahlung photons emitted by the electron.

A new tracker isolation criterion is thus used : the scalar sum of the transverse momenta of tracks with $p_t > 1.5$ GeV/c, associated to the primary vertex and contained between two cones of radii $\Delta R_{in} = 0.02$ and $\Delta R_{out} = 0.2$, must be smaller than 3 GeV/c. This procedure suppresses the activity due to tracks from Bremsstrahlung photon conversion and restores behaviours similar to those in the underlying event. The new isolation criterion has a uniform dependence on the inner isolation cone size and on the kinematical variables as $p_{t,GSF}$, $|\eta|$ and ϕ. By definition it is independent of the emission distance $|\Delta z|$.

The identification efficiency for high energy isolated electrons from Drell-Yan events with mass above 200 GeV/c^2, and the rejection power for fake electrons from QCD jet events have been studied. The performance of the new tracker isolation criterion is similar to that of the old criteria. Efficiencies were studied with respect to the other criteria of the HEEP selection.

A.6 HEEP electron selection

The HEEP electron selection criteria and related efficiencies are defined and discussed in detail in [67, 51].

The criteria applied on reconstructed electron candidates are designed to ensure high efficiencies for high energy electrons, whilst efficiently rejecting the QCD background. They are required to be essentially insensitive to the quality of the detector calibration and therefore useful in startup conditions.

High energy electrons are selected within the tracker acceptance ($|\eta| < 2.5$), using the criteria summarised in Table A.5. The values of the cuts on the shower shape, track matching and isolation variables to be different for electrons emitted in the barrel ($|\eta| < 1.442$) and in the endcap ($1.560 < |\eta| < 2.5$) regions, because of the different detector geometries and activities in the isolation cones.

A.7 Tracker isolation efficiency

Figures A.10 and A.11 present, for the new tracker isolation criteria, efficiencies for high energy electrons from Drell-Yan events and rejection powers for fake electrons from QCD events, as a function of $p_{t,GSF}$, for several cut values on N_{tr} and Σp_t, the corresponding data for the HEEP tracker isolation (version 1.2) can be found in [51]. For Drell-Yan events, one observes that efficiencies with the new selection criteria are uniform with respect to p_t. Taking as an example the values $N_{tr} < 2$ and $\Sigma p_t < 3$ GeV/c (red points), they are close to 99% for high energy electrons. At high $p_{t,GSF}$, the tracker isolation criteria reject between 70 and 90 % of fake electrons in QCD jet events. The change in shape visible at 200 GeV/c^2 in Figure A.11 is due to the $e\gamma$ trigger preselection criteria [70] summarised in section A.4 : above 200 GeV/c^2, all QCD jets faking electrons satisfy the Very High E_t trigger.

The tracker isolation efficiencies are also estimated with respect to the other criteria of the HEEP selection [67], which includes in particular calorimeter isolation, as shown for $\Sigma p_t <$ 3 GeV/c and $N_{tr} < 2$ in Figure A.12 for high energy electrons from Drell-Yan events and in Figure A.13 for fake electrons from QCD jet events. These efficiencies are similar to those for the old tracker isolation criteria [51].

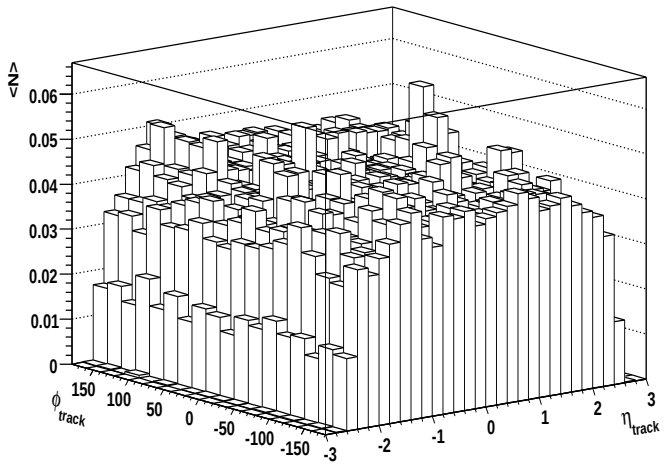


Figure A.1: η — ϕ density distribution of all tracks with $p_t > 1.5$ GeV/c, for Drell-Yan events with mass above 200 GeV/c^2.

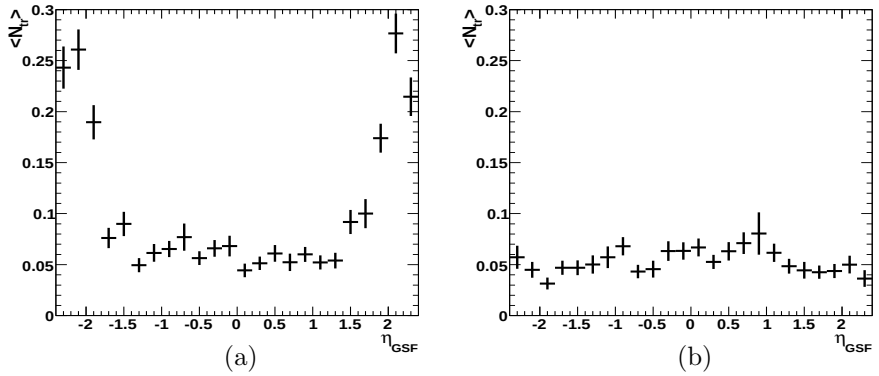


Figure A.2: (a) Distribution of the average number of tracks with $p_t > 1.5$ GeV/c and $|\Delta z| < 0.1$ cm between two cones of radii 0.02 and 0.2 around the electron direction, as a function of $\eta_{\rm GSF}$ for Drell-Yan events with mass above 200 GeV/c^2; (b) same distribution for the underlying event (90° in ϕ off the electron direction).

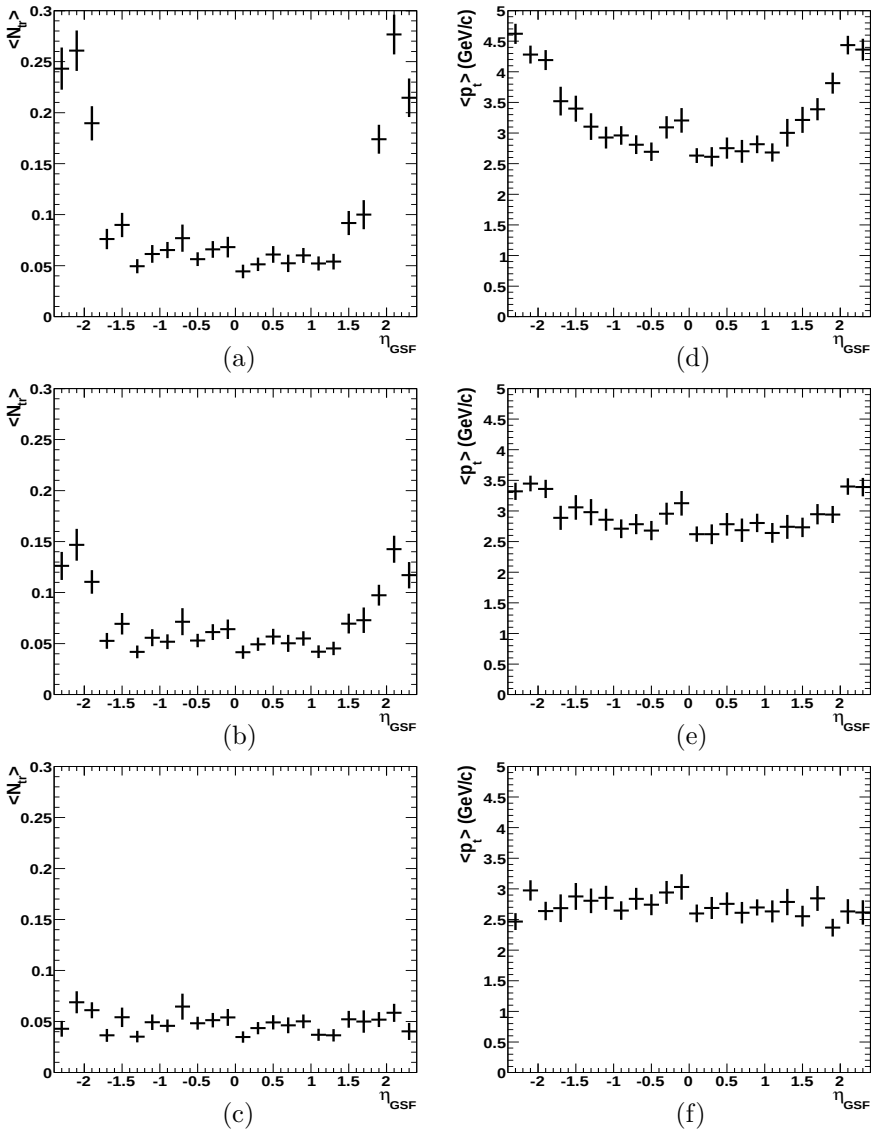


Figure A.3: Distribution of the average number of tracks with $p_t > 1.5$ GeV/c and $|\Delta z| < 0.1$ cm, comprised between two cones around the electron direction, as a function of $\eta_{\rm GSF}$ for three inner veto cone radii : (a) 0.02, (b) 0.04, (c) 0.07, the outer cone radius being 0.2; same distributions for the average transverse momentum of each track around the electron direction for the three inner veto cone radii : (d) 0.02, (e) 0.04, (f) 0.07.

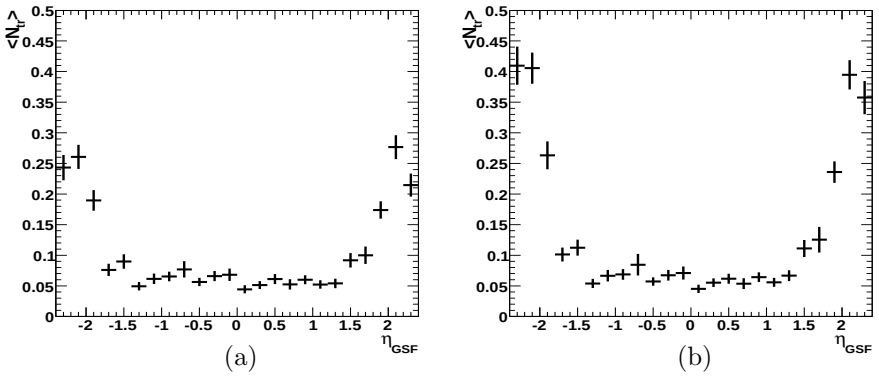


Figure A.4: Distribution of the average number of tracks with $p_t > 1.5$ GeV/c and (a) $|\Delta z| < 0.1$ cm and (b) $|\Delta z| < 0.3$ cm, between two cones of radii 0.02 and 0.2 around the electron direction, as a function of η_{GSF}.

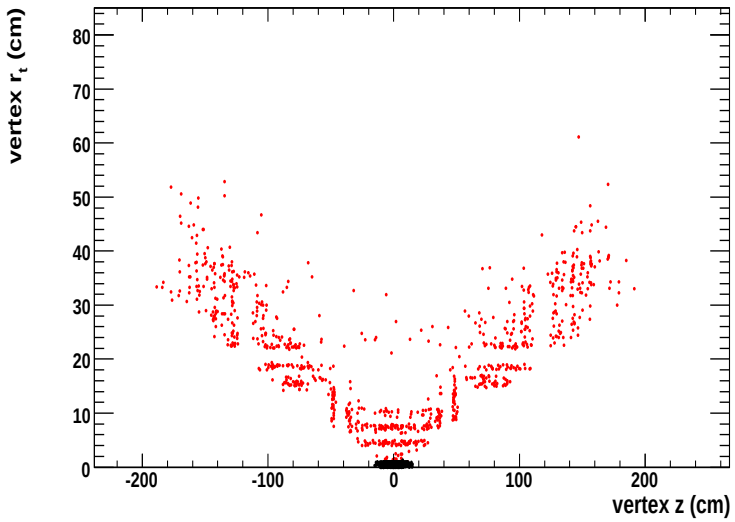
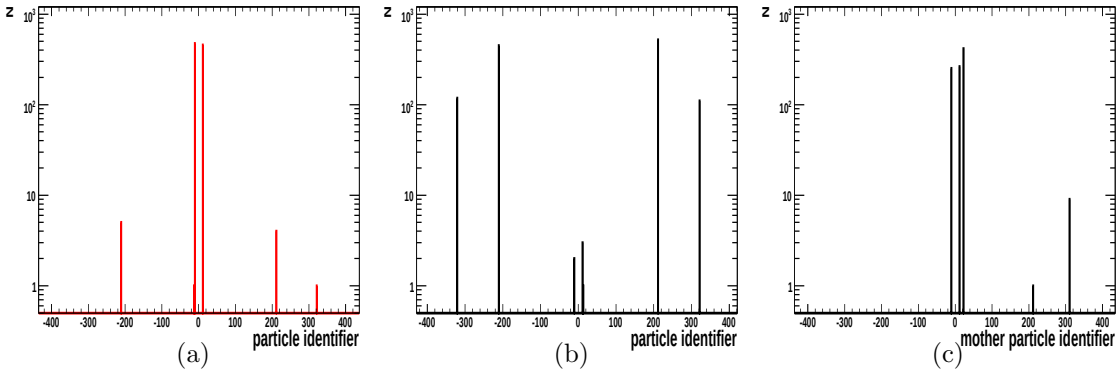


Figure A.5: r_t — z position of the emission vertices of the simulated particles associated to the reconstructed tracks around the electron direction. Black (red) points correspond to vertices of the simulated tracks which coincide (do not coincide) with the primary vertex.



(a) (b) (c)

Figure A.6: Identifier code of the particles corresponding to simulated tracks associated to reconstructed tracks around the electron direction, for simulated tracks associated (a) and not associated (b) to the primary vertex; (c) identifier code of the mother particles, for simulated tracks not associated to the primary vertex. The following code is used for the identifier : photon=22, electron=11, positron=-11, positive pion=211, negative pion=-211, positive kaon=321 and negative kaon=-321.

	Drell-Yan M > 200 GeV/c^2 : efficiency ϵ					
	$\Sigma p_t < 3$ GeV/c		$N_{tr} < 2$		Old tracker isolation	
	barrel	endcap	barrel	endcap	barrel	endcap
$p_{t,GSF} > 50$ GeV/c	0.99	0.99	1.00	1.00	1.00	0.99
$p_{t,GSF} > 200$ GeV/c	0.99	1.00	1.00	1.00	1.00	0.99
	QCD jet background : rejection $1 - \epsilon$					
	$\Sigma p_t < 3$ GeV/c		$N_{tr} < 2$		Old tracker isolation	
	barrel	endcap	barrel	endcap	barrel	endcap
$p_{t,GSF} > 50$ GeV/c	0.18	0.09	0.03	0.04	0.06	0.07
$p_{t,GSF} > 200$ GeV/c	0.39	0.54	0.13	0.33	0.23	0.29

Table A.2: Efficiencies for high energy electrons from Drell-Yan events and rejection power for fake electrons from QCD jet events, after application of the other HEEP cuts given in [67], for the new tracker isolation cuts : $\Sigma p_t < 3$ GeV/c and $N_{tr} < 2$, and for the old tracker isolation criteria.

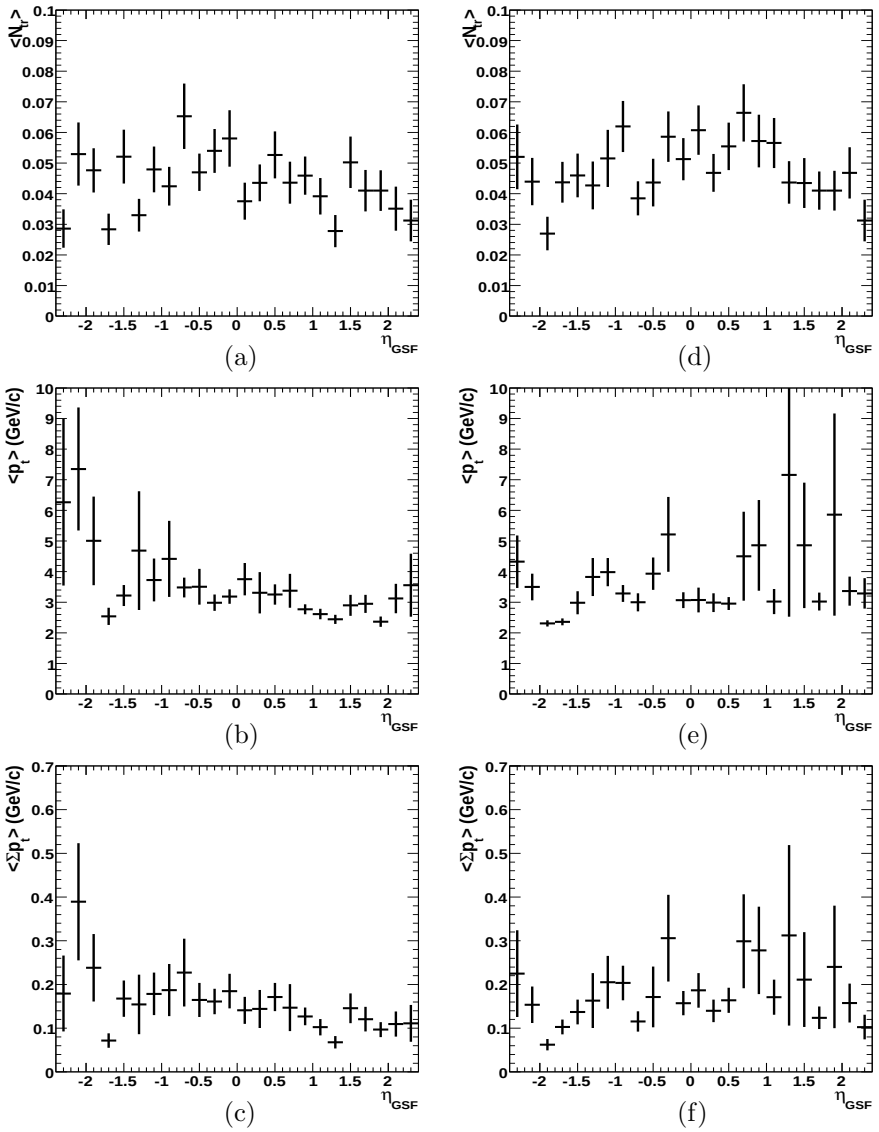
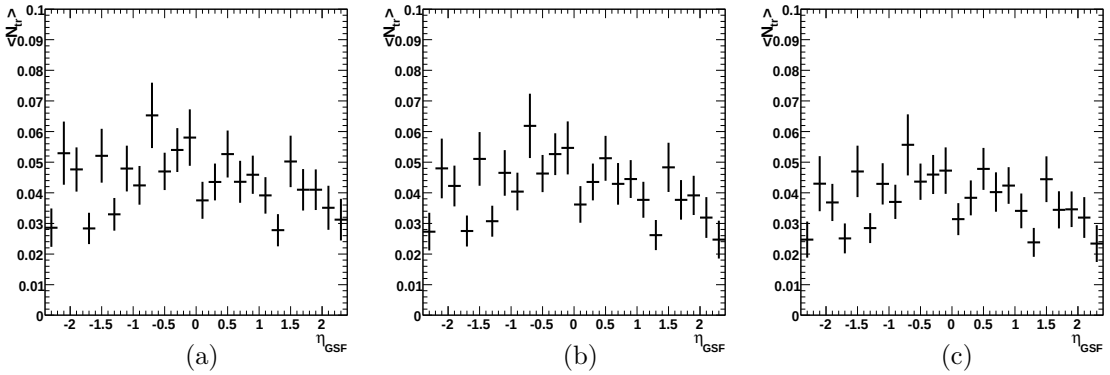


Figure A.7: Distribution, as a function of η_{GSF}, of the average number of tracks (a), the average p_t (b) and the average Σp_t (c) for tracks with $p_t > 1.5$ GeV/c between two cones of radii 0.02 and 0.2 around the electron direction, using high mass Drell-Yan events, according to the new definition of tracker isolation, i.e. considering only tracks associated to the primary vertex, same distributions for the underlying event (90° in ϕ off the electron direction) in (d), (e) and (f), respectively.



(a) (b) (c)

Figure A.8: Distribution, as a function of η_{GSF}, of the average number of tracks with $p_t > 1.5$ GeV/c between two cones around the electron direction for an inner cone radius of 0.02 (a), 0.04 (b) and 0.07 (c), the outer cone radius being 0.2, according to the new definition of tracker isolation, i.e. considering only tracks associated to the primary vertex.

ECAL barrel					
ECAL+HCAL isol.	91.9			tracker isol.	98.8
tracker isol.	91.2			ECAL+HCAL isol.	91.2
ECAL endcaps					
ECAL+HCAL1 isol.	93.9	HCAL2 isol.	98.5	tracker isol.	98.8
HCAL2 isol.	92.9	ECAL+HCAL1 isol.	92.9	ECAL+HCAL1 isol.	93.3
tracker isol.	92.4	tracker isol.	92.4	HCAL2 isol.	92.4

Table A.3: Cumulative efficiencies of the ECAL and HCAL HEEP isolation criteria defined in [67] and of the new tracker isolation criterion ($\Sigma p_t < 3$ GeV/c), with respect to the other HEEP selection criteria for Drell-Yan events with mass above 200 GeV/c^2 in the barrel (upper rows) and the endcaps (lower rows).

ECAL barrel					
ECAL+HCAL isol.	87.1			tracker isol.	57.7
tracker isol.	89.4			ECAL+HCAL isol.	89.4
ECAL endcaps					
ECAL+HCAL1 isol.	72.8	HCAL2 isol.	35.5	tracker isol.	47.2
HCAL2 isol.	73.5	ECAL+HCAL1 isol.	73.5	ECAL+HCAL1 isol.	75.9
tracker isol.	77.1	tracker isol.	77.1	HCAL2 isol.	77.1

Table A.4: Cumulative rejections of the ECAL and HCAL HEEP isolation criteria defined in [67] and of the new tracker isolation criterion ($\Sigma p_t < 3$ GeV/c), with respect to the other HEEP selection criteria for QCD background events in the barrel (upper rows) and the endcaps (lower rows).

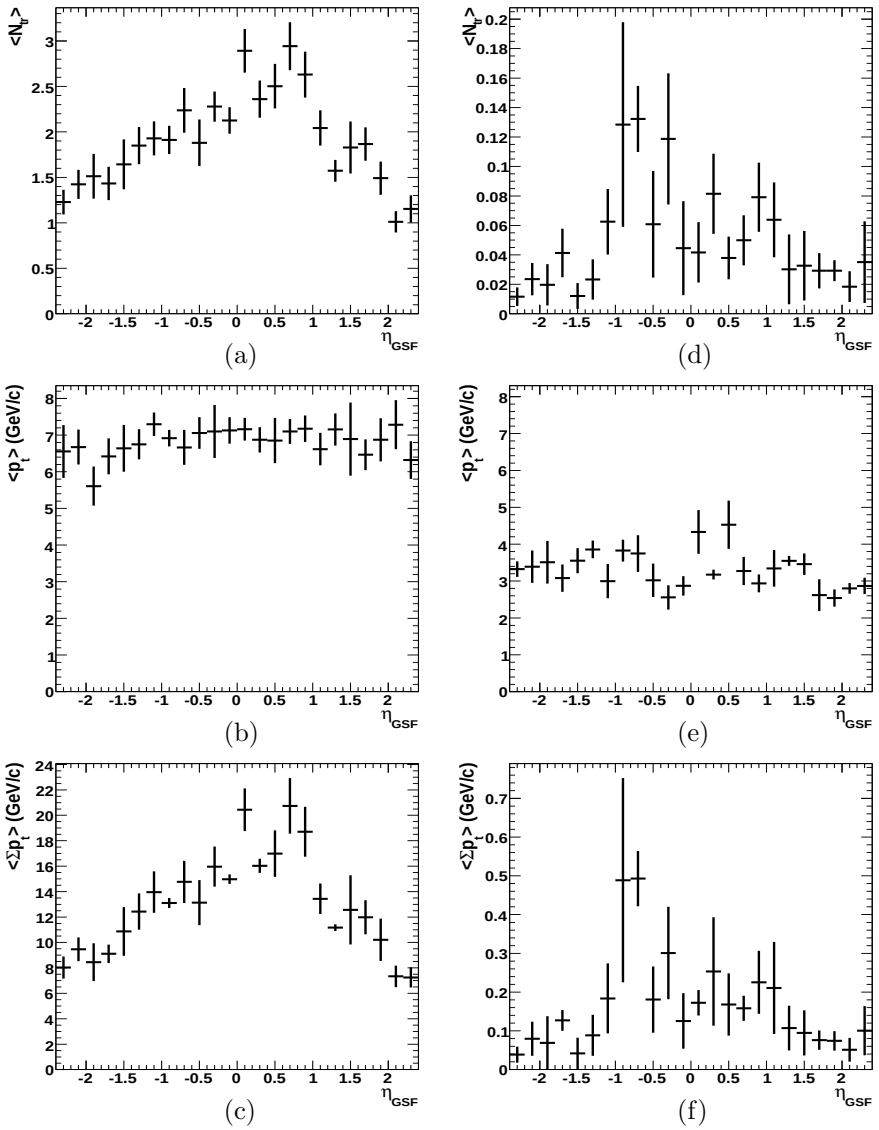


Figure A.9: Same distributions as in Figure A.7, but for fake electron candidates in the QCD jet samples. Note the different scales in (a)-(d) and in (e)-(f).

Variable	Barrel	Endcap
E_t	> 30 GeV	> 30 GeV
η_{SC}	$\lvert\eta_{SC}\rvert < 1.442$	$1.560 < \lvert\eta_{SC}\rvert < 2.5$
classification	< 40	≥ 100
$\Delta\eta_{in}$	$\lvert\Delta\eta_{in}\rvert < 0.005$	$\lvert\Delta\eta_{in}\rvert < 0.007$
$\Delta\phi_{in}$	$\lvert\Delta\phi_{in}\rvert < 0.09$	$\lvert\Delta\phi_{in}\rvert < 0.09$
H/E	< 0.05	< 0.1
$\sigma_{\eta\eta}$	< 0.011	< 0.0275
ECAL + HCAL1 isol.	< 5 GeV for $E_t < 65$ GeV else < 5 GeV $+ 0.02 * (E_t - 65)$	< 4 GeV for $E_t < 65$ GeV else < 4 GeV $+ 0.04 * (E_t - 65)$
HCAL2 isol.	not applied	< 1 GeV $+ 0.005 * E_t$[GeV]
Tracker isol.: Σp_t	< 7.5 GeV	< 15 GeV
Duplicate cleaning	E/p closest to 1	E/p closest to 1

Table A.5: HEEP electron selection criteria (version 1.2) from Ref. [67].

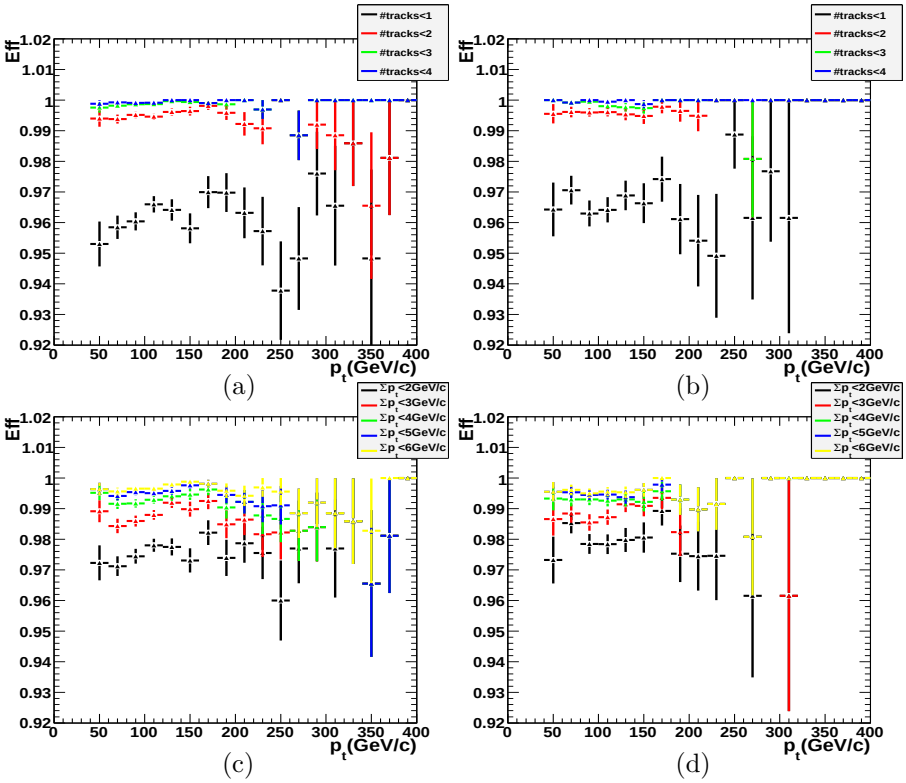


Figure A.10: Efficiencies of the new tracker isolation criteria for high energy electrons from Drell-Yan events with mass above 200 GeV/c^2 , as a function of $p_{t,GSF}$, for several cut values of N_{tr} in the barrel (a) and the endcaps (b), and for several cut values of Σp_t in the barrel (c) and the endcaps (d).

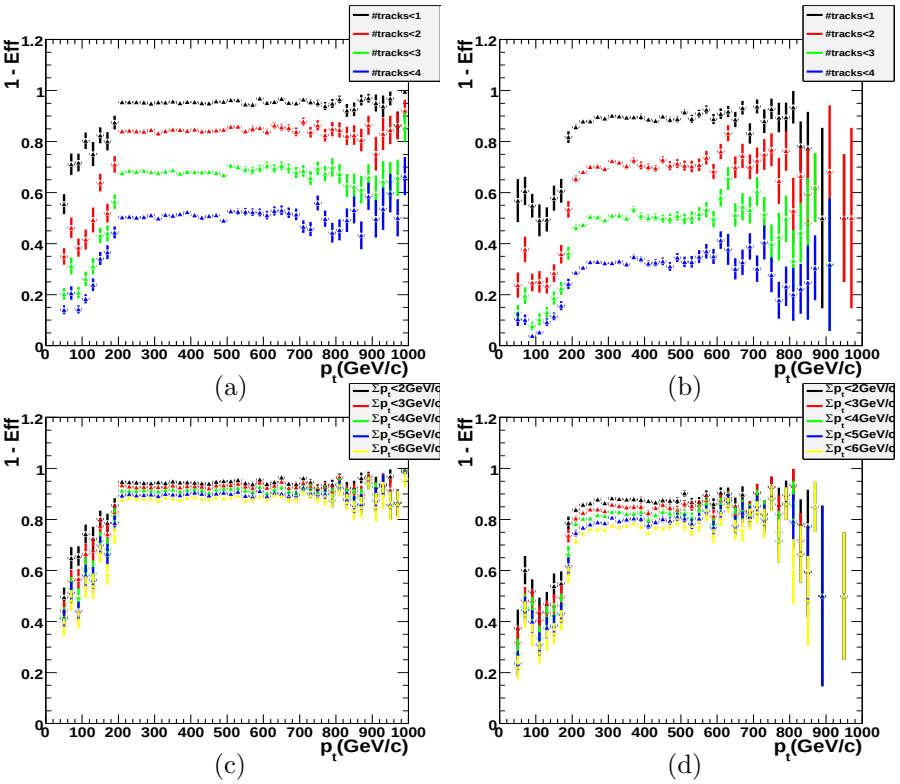


Figure A.11: Rejection power of the new tracker isolation criteria for fake electrons from QCD jet events, as a function of $p_{t,GSF}$, for several cut values of N_{tr}, in the barrel (a) and the endcaps (b), and for several cut values of Σp_t in the barrel (c) and the endcaps (d).

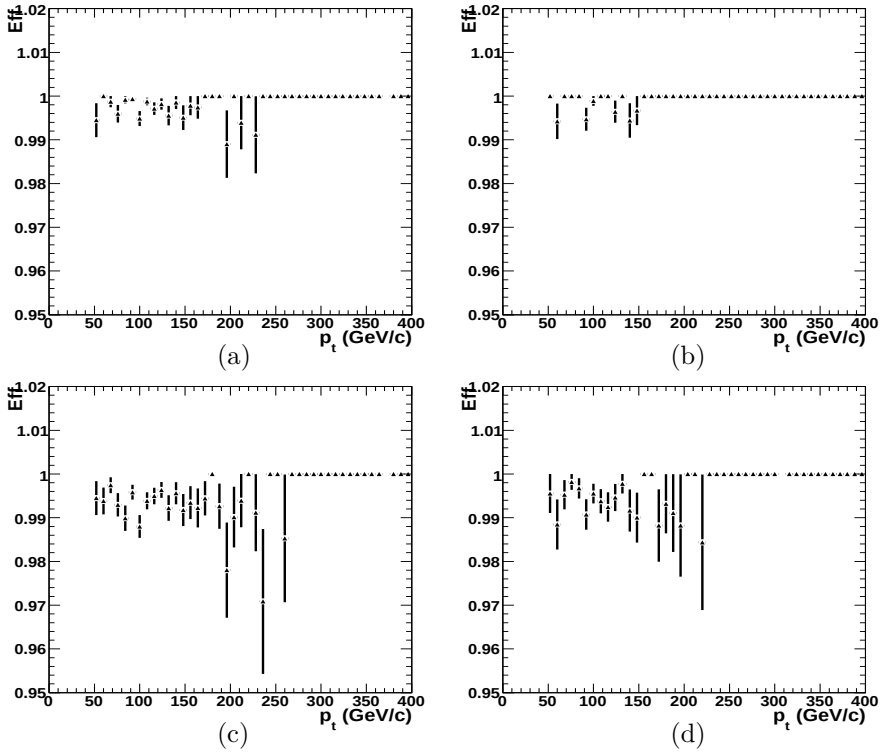


Figure A.12: Efficiencies of the new tracker isolation criteria for high energy electrons from Drell-Yan events with mass above 200 GeV/c^2, with respect to the other criteria of the HEEP selection, as a function of $p_{t,GSF}$, for $N_{tr} < 2$ in the barrel (a) and the endcaps (b), and for $\Sigma p_t < 3$ GeV/c in the barrel (c) and the endcaps (d).

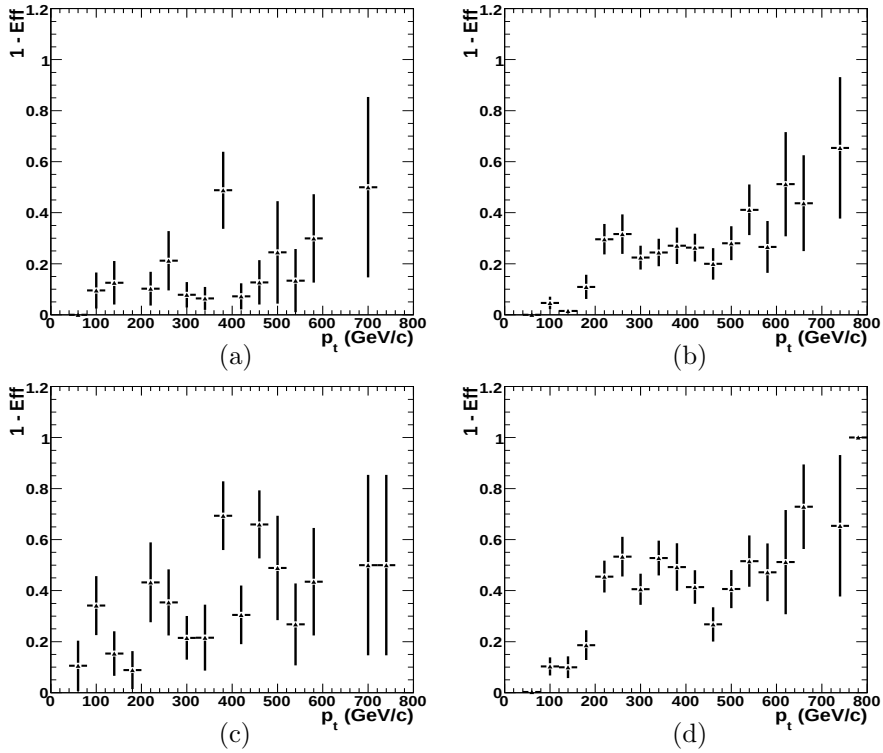


Figure A.13: Rejection power of the new tracker isolation criteria for fake electrons from QCD jet events with respect to the other criteria of the HEEP selection, as a function of $p_{t,GSF}$, for $N_{tr} < 2$ in the barrel (a) and the endcaps (b), and for $\Sigma p_t < 3$ GeV/c in the barrel (c) and the endcaps (d).

Appendix B

Error estimation on average values with weighted events

As said in the introduction, QCD samples are composed of several $\hat{p_t}$ bins. Each bin is specified by a number of simulated entries and a corresponding weight. The number of entries in each bin multiplied by its weight gives the number of events expected in that bin for an integrated luminosity of 100 pb^{-1}.

Therefore, in each bin in $\Delta\eta$, several entries from several $\hat{p_t}$ bins combine to give the average value that we see on figures A.9. Since each $\hat{p_t}$ bin has its own distribution and its own number of simulated events, we will classify the entries in each $\Delta\eta$ bin per weight. Thus, in each $\Delta\eta$ bin, we will create classes i specified by its weight w_i and the corresponding number of entries in that $\Delta\eta$ bin N_i.

Thus, for each class i, we have the mean value :

$$\bar{x_i} = \frac{1}{N_i}\sum_{n=1}^{N_i} x_{in} \tag{B.1}$$

the variance is given by :

$$\sigma_i^2 = \frac{1}{N_i}\sum_{n=1}^{N_i}(x_{in} - \bar{x_i})^2 \tag{B.2}$$

and the error is then given by :

$$\sigma_{\bar{x_i}} = \frac{\sigma_i}{\sqrt{N_i}} \tag{B.3}$$

Each individual average value $\bar{x_i}$ contributes to the general average value of the $\Delta\eta$ bin $\bar{x}$ with a weight $w_i/\sigma^2_{\bar{x_i}}$ and the general average value is then given by :

$$\bar{x} = \frac{\sum_{i=1}^{N} w_i\bar{x_i}/\sigma^2_{\bar{x_i}}}{\sum_{i=1}^{N} w_i/\sigma^2_{\bar{x_i}}} \tag{B.4}$$

It is now straightforward to compute the error on $\bar{x}$. Let's define :

$$X = \sum_{i=1}^{N}\frac{w_i}{\sigma^2_{\bar{x_i}}}\bar{x_i} = \sum_{i=1}^{N} a_i\bar{x_i} \tag{B.5}$$

The error on X is given by :

$$\sigma_X^2 = \sum_{i=1}^{N} a_i^2 \sigma_{\bar{x}_i}^2 = \sum_{i=1}^{N} \frac{w_i^2}{\sigma_{\bar{x}_i}^2} \tag{B.6}$$

The general value $\bar{x}$ can then be expressed as a function of X via :

$$\bar{x} = \frac{X}{\sum_{i=1}^{N} \frac{w_i}{\sigma_{\bar{x}_i}^2}} \tag{B.7}$$

and the error on $\bar{x}$ can now be expressed as :

$$\sigma_{\bar{x}}^2 = \frac{\sigma_X^2}{(\sum_{i=1}^{N} \frac{w_i}{\sigma_{\bar{x}_i}^2})^2} = \frac{\sum_{i=1}^{N} \frac{w_i^2}{\sigma_{\bar{x}_i}^2}}{(\sum_{i=1}^{N} \frac{w_i}{\sigma_{\bar{x}_i}^2})^2} \tag{B.8}$$

Appendix C

Example of a Drell-Yan event display

This appendix presents the visualisation of the tracker activity, given in different views, for a high mass Drell-Yan event.

Figure C.1 gives the position of the different type of tracker hits in the event both at the simulated and the reconstructed level. Hits from reconstructed tracks in the event are displayed as red crosses. Hits from simulated tracks associated to reconstructed tracks in the event are displayed as blue triangles. Finally, vertices of simulated tracks associated to reconstructed tracks are displayed as black circles.

The r_t—ϕ view is shown in Figures C.1 (a) and (b), the latter giving a zoomed view on the GSF track of the electron, r_t being the distance between the vertex and the beam direction in the transverse plane. The r_t—z and x—y views are given in Figures C.1 (c) and (d) respectively.

Figure C.1 (b) focuses on the region where the GSF electron track is visible (characterised by $p_{t,GSF} = 83.8$ GeV/c, $|\eta| = 1.91$ and $\phi = 0.484$). One can also see a reconstructed track ($p_t = 14.1$ GeV/c, $|\eta| = 1.91$ and $\phi = 0.474$) in the double cone around this GSF electron direction. The simulated information from GEANT indicates that this track is an electron track. From the "mother" particle information, it is found that this electron comes from a photon through pair creation ($\gamma \rightarrow e^+e^-$), the photon itself is coming from the GSF electron through a Bremsstrahlung process : $e \rightarrow e\gamma$.

This reconstructed track in the double cone around the electron direction is found : (i) to have an emission distance $|\Delta z| < 0.1$ cm and would then contribute to the old tracker isolation definition, (ii) not to be associated to the primary vertex. This example shows that using only tracks associated to the primary vertex gets rid of tracks due to the conversion of Bremsstrahlung photons emitted by the primary electrons when estimating the activity around the electron direction.

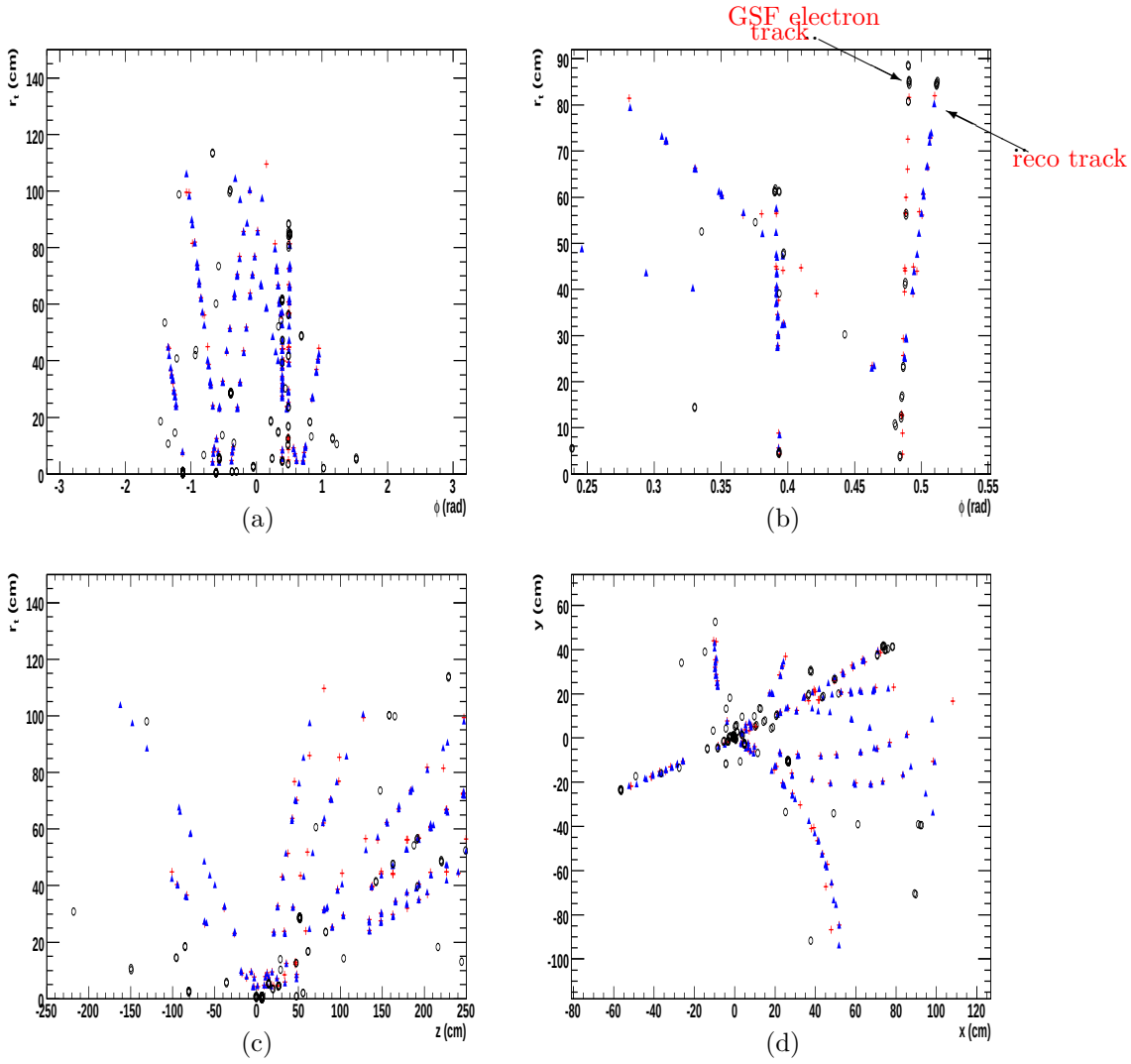


Figure C.1: Different views of a high mass Drell-Yan event showing the positions of reconstructed hits (red crosses), hits of simulated tracks associated to reconstructed tracks in the event (blue triangles) and vertices of the latter simulated tracks (black circles); (a) r_t—ϕ view, where r_t is the distance between the vertex and the beam direction in the transverse plane; (b) zoom of the latter view showing the GSF electron track; (c) r_t—z view; (d) x—y view.

Bibliography

[1] *Lives and Opinions of the eminent philosophers* by Diogene Laertius. Translation from C.D. Yonge

[2] C. Burgess, G. Moore *The Standard Model A Primer* Cambridge University Press ISBN 0-521-86036-9

[3] LEP Collaborations *A combination of preliminary electroweak measurements and constraints on the Standard Model* (2005) arXiv:hep-ex/0412015v2

[4] R.P. Feynman *Quantum electrodynamics* Westview Press, New Ed. Edition (1998). ISBN 978-0201360752

[5] Greiner, Walter, Schafer, Andreas *Quantum Chromodynamics* Springer. ISBN 0-387-57103-5

[6] S.M. Carroll *Spacetime and Geometry : An introduction to General Relativity* San Francisco : Addison Wesley, ISBN 0-8053-8732-3

[7] T. Mohaupt *Introduction to String Theory* (2002) arXiv:hep-th/0207249-v1

[8] F. Englert and R.Brout, Phys. Rev. Letter, **13** (1964) 321-322.

[9] P.W. Higgs, Phys. Rev. Letter, **13** (1964) 508-509.

[10] `http://www.hep.ph.ic.ac.uk/cms/physics/higgs.html`

[11] R.N. Mohapatra, S.A. Goldstein, D. Money *Unification and Supersymmetry : The Frontiers of quark-lepton physics* Springer-Verlag, 2002. Third Edition.

[12] M. Cvetic and S. Godfrey *Discovery and identification of extra gauge bosons* in 'Electroweak symmetry breaking and new physics at the TeV scale', T. Barlow and S. Dawson, pp. 383-415. World Scientific, 1996. arXiv:hep-ph/9504216

[13] S. Godfrey *Update of discovery limits for extra neutral gauge bosons at hadron colliders* in 'Proceedings of Snowmass 2001', p. P344. Snowmass, Colorado, June-July, 2001. arXiv:hep-ph/0201093.

[14] T. Kaluza *On the problem of unity in physics*, Sitzungsber. Preuss. Akad. Wiss. Berlin. (Math. Phys.)(1921) 966-972.

[15] O. Klein *Quantum theory and five dimensional theory of relativity*, Z. Phys. **37** (1926) 895-906

[16] E.G. Adelberger et al. *Torsion Balance Experiments : A low energy frontier of Particle Physics* http://www.npl.washington.edu/eotwash/publications/pdf/lowfrontier2.pdf

[17] L. Randall, R. Sundrum *A Large Mass Hierarchy from a small Extra Dimension* (1999) Phys. Rev. Letters **83** (17) 3370-3373

[18] http://www-cdf.fnal.gov/physics/exotic/exotic.html

[19] http://www-d0.fnal.gov/Run2Physics/WWW/results/np.htm

[20] D0 Collaboration *Search for high-mass narrow resonances in the di-electron channel at D0* D0 Note 5923-CONF

[21] The CDF Collaboration *Search for high mass resonances decaying to muon pairs* http://www-cdf.fnal.gov/physics/exotic/r2a/20100527.zprime_mumu/conference_note.pdf

[22] The D0 Collaboration *Search for Randall-Sundrum gravitons in the dielectron and diphoton final states with 5.4* fb^{-1}*of data from* $p\bar{p}$ *collisions at* $\sqrt{s}$ = 1.96 TeV arXiv:hep-ex/10041826v1

[23] The LEP Higgs working group *Higgs Boson searches at LEP* arXiv:hep-ex/08044146v1

[24] The D0 and CDF collaborations *Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 6.7* fb^{-1} *of Data* arXiv:hep-ex/10074587v1

[25] CMS Collaboration, Detector Performance and Software, Technical Design Report, volume I, CERN/LHCC 2006-001.

[26] N. Adam & al. *The CMS Luminosity Readout System* CMS IN 2007/052

[27] CMS Collaboration *Measurement of the underlying event activity at the LHC with* $\sqrt{s}$ = *7* TeV *and comparison with* $\sqrt{s}$ = *900* GeV, CMS Physics Analysis Summary PAS QCD-10-010.

[28] http://twiki.cern.ch/twiki/bin/view/CMS/HEEP7TeVscaled

[29] http://cmsdoc.cern.ch

[30] CMS Collaboration, *The CMS experiment at the LHC* Journal of Instrumentation 2008 JINST 3 S08004

[31] Sh. Elgammal & al., *Saturation and energy corrections for* TeV *electrons and photons* CMS AN-2006/149

[32] Sh. Elgammal & al., *Recovery of high* p_t *electrons from large mass pairs, lost in CMS ECAL cracks and in case of a missing endcap* CMS AN-2008/029

[33] Sh. Elgammal & al., *Validation of ECAL calibration at high energy and correction for electronics saturation* CMS AN-2009/049

[34] T. Sjostrand, P. Mrenna, S. Skands *Pythia 6.4 Physics Manual* arXiv:hep-ph/0603175v2

[35] http://cern.ch/geant4

[36] `http://www.phys.psu.edu/\~{}cteq/`

[37] `http://madgraph.hep.uiuc.edu/`

[38] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, *Updated predictions for the total production cross sections of top and of heavier quark pairs at the Tevatron and at the LHC*, JHEP 09 (2008) 127, arXiv:0804.2800. doi:10.1088/1126-6708/2008/09/127.

[39] `http://www.hep.phy.cam.ac.uk/theory/webber/MCatNLO`

[40] CMS Collaboration, *Search for high mass resonance production decaying into an electron pair in the CMS experiment*, CMS Physics Analysis Summary PAS EXO-2008-001.

[41] CMS Collaboration, *Search for high mass resonances decaying into an electron pair in CMS at 10* TeV *with 100* pb^{-1}, CMS Physics Analysis Summary PAS EXO-2009-006.

[42] CMS Collaboration, *Search for high mass resonances decaying into an electron pair in CMS at 10* TeV *with 100* pb^{-1}, CMS AN 2009/097.

[43] Measuring electron efficiencies at CMS with early data. CMS PAS, EGM-07-001, 2007. 2

[44] D. Bandurin et al., "Study of backgrounds to high mass dielectron (Drell-Yan) final states" CMS AN, 2008-044 (2008).

[45] The CMS Collaboration *The CMS physics reach for searches at 7* TeV CMS Note 2010-008

[46] M. Anderson & al. *A review of clustering algorithms and energy corrections in the electromagnetic calorimeter* CMS IN 2010/008

[47] S. Baffioni & al., *Electron Reconstruction in CMS*, CMS NOTE 2006/040.

[48] W. Adam and al. *Track Reconstruction in the CMS Tracker* CMS Note 2006/041 (2006)

[49] C. Charlot et al. *Reconstruction of electron tracks using Gaussian Sum Filter in CMS* CMS AN 2005/011.

[50] R. Fruhwirth Comput. Phys. Commun., 100 (1997) 1.

[51] O. Charaf, B. Clerbaux, S. Harper, D. Newbold, C. Shepherd-Themistocleous, R. Walton *Electron Identification at high energies*, CMS AN 2008/045.

[52] O. Charaf & al. *Tracker isolation studies: rejection of Bremsstrahlung photon conversion tracks around high energy electrons* CMS AN-2009/071

[53] L. Tuura, A. Meyer, I. Segoni, G. Della Ricca *CMS Data Quality Monitoring : Systems and Experiences* Journal of Physics Conference Series 219 (2010) 072020

[54] `http://twiki.cern.ch/twiki/bin/view/CMS/HPTEgroupDQM`

[55] `http://twiki.cern.ch/twiki/bin/view/CMS/EcalFirstBeam09Anomalous`

[56] `http://twiki.cern.ch/twiki/bin/viewauth/CMS/WorkBookFireworks`

[57] `https://twiki.cern.ch/twiki/bin/view/CMS/DQM`

[58] http://madgraph.roma2.infn.it

[59] S. Alioli, P. Nason, C. Oleari & al. *NLO vector-boson production matched with shower in POWHEG*, JHEP 07 (2008) 060, arXiv:hep-ph/0805:4802

[60] P. Nason *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP 11 (2004) 040, arXiv:hep-ph/0409146

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

[62] The CMS Collaboration *Measurement of inclusive W and Z productions in pp collisions at* $\sqrt{s} = 7$ TeV CMS AN-10-116

[63] http://llr.in2p3.fr/\~baffioni/HNelectron/ElePlots_deltaeta.html

[64] https://twiki.cern.ch/twiki/bin/view/CMS/LumiWikiHome

[65] The CMS Collaboration *Measurement of CMS Luminosity*, CMS PAS EWK-10-004

[66] https://twiki.cern.ch/twiki/bin/view/CMS/SUSYBSMHeep

[67] D. Evans et al., *Search for high mass resonance production decaying into an electron pair in the CMS experiment*, CMS NOTE AN 2008/048

[68] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideTrackingTools\#Track_Association

[69] S. Cucciarelli & al., *Track reconstruction, primary vertex finding and seed generation with the Pixel Detector*, CMS NOTE 2006/026.

[70] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideEgammaHLT

Acknowledgements

The carry out of a thesis is not a task to be claimed individually. Many persons have contributed to the realisation of this work, to which I especially think about and would like to show my gratitude.

To start with, I would like to thank my family who has always been there to support me and who believed in me. I especially think about my parents from whom I inherit important qualities and who, in a very significant part, made the person I am. Mother, Father, even if you do not necessarily get the topic of this thesis and its content, it is dedicated to you. I, of course, also would like to mention my brothers and sisters who always believed in me, even in hard times, and who partly own this work. I will not forget your support during all those long and labourous years.

I would like to thank my supervisor who has been more than patient with me and has often supported my bad mood. It is inevitable, also, to think about Pierre Marage whose knowledge and experience were extremely valuable to carry out my work. Thanks, Pierre, for all the explanations and particularly the corrections you brought to this manuscript. Its quality would not have been so without your patient reading and numerous but ô so precious comments.

It would be shameful not to mention my friends who 'hit the road' with me since long ago. Julien Riton, doctor in Mathematics, which has been my friend since I started my studies at university. Long days and nights studying together nourished our interest for rigour. Samuel Challeat, the geographer, who has shared with me long nights of sky observation. Open and globular clusters, galaxies, nebulae, planets and sometimes comets were our 'nightly' share. I will not forget what this experience brought to me. Angelique Lagoutte, the physicist, that I have known now for almost 12 years. I do not forget you. Last, but not least, Jennyfer Jordan, the smart and beautiful english teacher with whom I shared so many great moments (learning, laughing, having fun...).

Special thanks go to my housemates which all became also very good friends. When I arrived in Belgium four years ago, I hardly knew anyone. Together, in this mangificent house that has been ours, we built the Marcqstraat community based on international exchange, moments of laugh, big feasts and delicious barbecues in the garden during summer and also sometimes less happy moments. Roser Gasol, the spanish biologist, studying oiled wildlife response country profiles; Ricard Aymar, the spanish top artist of next generation; Liesbeth Goossens, the flamish girl, (my ultimate theory still holds); Leontien Krul, the deutsch girl; Andrea De Marchis, the italian cooker ('pasta al dente'). These valuable housemates left the house and were replaced by others: John Moodie, the biritsh dancer (John, so far, you have been my most serious rival). Koen and Liesbeth Willaert, the two unseparable brother and sister and last but not least, the sweet Yoko Theeuws, best illustrator for next generation (I wish you all the best but know it's going to happen).

I cannot end this section without thanking my colleagues, especially the ones who have endured a lot, working in the same office. Sharing the same working place builds connections.

Vincent Dero, Sherif Elgammal, Laurent Thomas, I owe you a big thank. Also, TOP students, Gregory Hammad, Petra Van Mulders, Joris Maes, Ilaria Vilella and Jan Heyninck. Having you as close colleagues, during those four years, has been a memorable part of my life. You have all, in a way, contributed significantly to this work.