Dissertation ETH Nr. 15788

Electron reconstruction and electroweak processes as tools to achieve precision measurements at a hadron collider: from CDF to CMS

A dissertation submitted to the
Swiss Federal Institute of Technology Zürich

for the degree of
Dr. Sc. ETH Zürich

presented by
Anne-Sylvie Giolo-Nicollerat
Dipl.-Phys. Lausanne University
born June 5th. 1978 in Lausanne, Switzerland
Citizen of Switzerland

Prof. Felicitas Pauss (examiner)
Prof. Zoltan Kunszt (co-examiner)
Dr. Michael Dittmar (co-examiner)

- 2004 -

Contents

Abstract

Precision measurements are an important aspect of hadron colliders physics program. This thesis describes a method, together with a first application, of how to achieve and use precision measurements at the LHC. The idea is to use reference processes to control the detector systematics and to constrain the theoretical predictions. We chose as reference processes single W and Z production, selected through their decays into electrons. The study presents results for both the CDF experiment at the Tevatron (USA), for which a data analysis was performed and the CMS experiment at the future CERN LHC (Switzerland) for which Monte Carlo simulations were done.

In the first part, we determine the luminosity of the data collected by the CDF detector between February 2002 and May 2003, using the $p\bar{p} \rightarrow W \rightarrow e\nu$ and $p\bar{p} \rightarrow Z \rightarrow ee$ processes. The obtained total luminosity of 125.5 ± 0.6 (stat.) $\pm$ 7.1 (syst.) pb^{-1} is in good agreement with the 'traditional' luminosity measurement of 125.5 ± 7.3 pb^{-1}. Due to the high statistics of W and Z events at the LHC, a more precise relative luminosity measurement can be performed by CMS, reducing the error on the luminosity determination perhaps to about 1%, thus allowing accurate cross sections determinations.

In a second step we show, using a full detector simulation, how leptonic W and Z events can be identified with CMS, concentrating on the electron reconstruction and identification. Different selection variables for electrons are defined and systematic errors issues are discussed.

Finally, the potential of CMS to perform precision measurements is illustrated using the leptonic decays of a hypothetical Z′ boson. For the first time we demonstrate how well the Z′ properties can be constrained at the LHC combining well-known variables like the cross section times branching ratio, the forward backward charge asymmetry on and off peak. We also show that the Z′ rapidity distribution can be used to constrain the Z′ couplings to the quarks. Z′ bosons from different theoretical models can be discriminated for Z′ masses up to 2-2.5 TeV, assuming an integrated luminosity of 100 fb^{-1}.

Résumé

Les mesures de précision représentent un aspect important du programme de physique des collisionneurs hadroniques. Ce travail de thèse présente une méthode avec une première application permettant de faire de telles mesures avec le LHC. L'idée est d'utiliser des réactions de référence pour contrôler aussi bien les erreurs systématiques liées au détecteur que les prédictions théoriques. Dans ce cas précis, les réactions de référence choisies sont la production de W et de Z, ces derniers étant sélectionnés à l'aide de leur désintégration en électrons. Cette étude présente des résultats pour deux expériences : CDF, situé sur l'anneau du Tevatron aux USA, où nous avons procédé à une analyse de données et CMS, sis sur le futur LHC du CERN en Suisse, pour lequel des simulations Monte Carlo ont été effectuées.

Premièrement, nous avons mesuré la luminosité des données collectées par CDF entre février 2002 et mai 2003 en utilisant les réactions suivantes: $p\bar{p} \rightarrow W \rightarrow e\nu$ et $p\bar{p} \rightarrow Z \rightarrow ee$. Nous avons trouvé une luminosité totale de 125.5 ± 0.6 (stat.) $\pm$ 7.1 (syst.) pb^{-1} en accord avec le résultat obtenu de manière traditionelle, soit 125.5 ± 7.3 pb^{-1}. Grâce au nombre important de W et de Z produits au LHC, il pourrait être possible de réduire l'incertitude sur la luminosité relative à moins de 1%, permettant ainsi d'augmenter la précision des mesures de sections efficaces.

Dans un second temps, nous avons étudié la manière d'identifier les désintégrations leptoniques des événements W, Z avec CMS en se concentrant plus spécifiquement sur la reconstruction des électrons et leur identification. Diverses variables ont été définies pour sélectionner les électrons et différentes sources potentielles d'incertitudes systématiques sont discutées.

Finalement pour illustrer le potentiel de CMS pour effectuer des mesures de précision, une étude de la détermination des propriétés d'un hypothétique Z′ a été conduite. Après la découverte du Z′ en utilisant sa désintégration en leptons, ses propriétés peuvent êtres restreintes à l'aide des observables suivantes bien connues: la section efficace multipliée par la largeur du Z′, l'asymétrie de charge avant-arrière au sommet et à côté du pic de masse. Ces observables ont été combinées pour la première fois pour étudier la sensitivité du LHC. Nous avons également montré que la distribution de rapidité du Z′ peut aussi être utilisée pour determiner le couplage du Z′ aux quarks. Il devrait être possible avec CMS de distinguer les Z′ de différents modèles pour des masses inférieures à 2-2.5 TeV avec une luminosité de 100 fb^{-1}.

Acknoledgements

"It seems to me that if we want to improve scientific research in this country, it's not more money or equipment that needs to be infused in the system but the accumulated experience of the best minds who've ever done science"

Robert Scott Root-Bernstein

I would like to thank Felicitas Pauss for giving me the possibility to make my PhD within the ETH group and for her unconditional support during these three years.

I am also grateful to Michael Dittmar to have supervised me throughout my PhD and to have shown such patience in transmitting to me his experience in high energy physics research. Thanks also for his friendship and making me almost part of his family!

Thanks to Jaco Konigsberg and the members of the CLC group to have welcomed me in their group and to Pasha Murat for his help with the CDF Standard Ntuple. I am also grateful to Dee Hahn for her judicious ACE choice three years ago... Thanks to Steffen Kappler and Alexander Schmidt for introducing me to the world of PAX and to Abdelhak Djouadi for the nice collaboration on the topic of the Z'.

I would like also to thank my colleagues of the Institute of Particle Physics of ETHZ for their help and support and also for their friendship. My thanks go specially to Alison Lister, André Holzner, Gabriele Kogler, Giovanna Davatz, Günther Dissertori, Jan Ehlers and Radek Oficrzynski.

Finally I would like to thank my family and my friends for supporting me during these years.

Introduction

"I myself, being fairly ignorant of scientific literature, could find more knowledge new to me in an hour's time spent at the library than I could find at my workbench in a month or a year. It is not truth I am searching for, it is new truth. A scientific researcher has to be attracted by those [blank] spots on the map of human knowledge, and if need be, be willing to give his life for filling them in." A. Szent-Györgyi

It might not be a contradiction to say that Particle Physics, the science studying the basic components of matter and their interactions, is a young science based on an old idea. Already three centuries before Jesus Christ the Greek philosopher, Democritus, brought forward the idea that matter was composed of fundamental unbreakable entities that he called "atoms". During the 20^{th} century, the development of particle colliders of ever increasing energy and detectors of ever improving performances, allowed the exploration of matter at smaller and smaller scales.

The development of Quantum Mechanics at the beginning of the twentieth century, Quantum Field Theory some decades later and finally the Standard Model of particle physics, formulated in the sixties, represented major breakthroughs in our way to understand matter's building blocks and their interactions. Using these theories, physicists managed to explain with an unprecedented success the more and more accurate experimental observations.

The Standard Model of particle physics makes the following basic statements, as summarized in Figure 1: all the known matter can be described by 12 fundamental entities, named fermions (spin 1/2 particles), together with interaction carriers, named vector bosons (spin 1 particles). These particles can interact via three types of forces[1]: the electromagnetic interaction (carried by photons), the strong interaction (carried by gluons) and the weak interaction (carried by W and Z bosons). The twelve fundamental fermions are sorted using their different behaviors under these interactions. For example, quarks are defined as particles which feel the strong interaction and leptons as particles which do not. These twelve particles can also be sorted into three doublet families, each family having the same general behavior under the weak interaction, the so-called universality. Particles from the first family form the ordinary matter: up and down quarks are combined to form protons and neutrons. Atoms are made out of the combination of protons, neutrons and electrons. The last building block of the first family is the electron neutrino (ν_e) which was postulated to get a correct description of the radioactive decays. The two other families (μ, ν_μ, c and s quarks, τ, ν_τ, t and b quarks) are used to describe high energy states like the ones which were present at the beginning of the Universe and can now be created and studied in detail with particle accelerators.

However, at least two fundamental questions remain to be clarified. The first is to discover the last building block of the Standard Model, the so-called Higgs boson. Its existence was postulated in the sixties by the Scottish physicist Peter Higgs. His idea consists in assuming the existence of a scalar field, the Higgs field, which, by coupling to the different particles, will give them a mass after a spontaneous symmetry breaking [2]. Although the Higgs field was originally postulated to explain the exceptionally high mass of the W and Z vector bosons, the same mechanism can be extended to account for the fermion masses. The interactions of the Higgs

[1] We exclude here gravity, which is about 10^{41} times weaker than the electromagnetic force (for two u quarks at 10^{-17}m) and is up to now not included in that model.

[2] The Standard Model alone cannot explain the origin of the different masses of the particles as the inclusion of mass terms in the Standard Model Lagrangian breaks its gauge symmetry.

FERMIONS — matter constituents, spin = 1/2, 3/2, 5/2, ...

Leptons spin = 1/2

Flavor	Mass GeV/c^2	Electric charge
ν_e electron neutrino	$<1\times10^{-8}$	0
e electron	0.000511	−1
ν_μ muon neutrino	<0.0002	0
μ muon	0.106	−1
ν_τ tau neutrino	<0.02	0
τ tau	1.7771	−1

Quarks spin = 1/2

Flavor	Approx. Mass GeV/c^2	Electric charge
u up	0.003	2/3
d down	0.006	−1/3
c charm	1.3	2/3
s strange	0.1	−1/3
t top	175	2/3
b bottom	4.3	−1/3

BOSONS — force carriers, spin = 0, 1, 2, ...

Unified Electroweak spin = 1

Name	Mass GeV/c^2	Electric charge
γ photon	0	0
W^-	80.4	−1
W^+	80.4	+1
Z^0	91.187	0

Strong (color) spin = 1

Name	Mass GeV/c^2	Electric charge
g gluon	0	0

PROPERTIES OF THE INTERACTIONS

Property \ Interaction	Gravitational	Weak	Electromagnetic	Strong: Fundamental	Strong: Residual
Acts on:	Mass – Energy	Flavor	Electric Charge	Color Charge	See Residual Strong Interaction Note
Particles experiencing:	All	Quarks, Leptons	Electrically charged	Quarks, Gluons	Hadrons
Particles mediating:	Graviton (not yet observed)	W^+ W^- Z^0	γ	Gluons	Mesons
Strength relative to electromag for two u quarks at: 10^{-18} m	10^{-41}	0.8	1	25	Not applicable to quarks
Strength relative to electromag for two u quarks at: 3×10^{-17} m	10^{-41}	10^{-4}	1	60	
for two protons in nucleus	10^{-36}	10^{-7}	1	Not applicable to hadrons	20

FIGURE 1: *The Standard Model of particle physics in a nutshell ! The first table gives a description of the matter components, the fermions. The second describes the force carriers, the vector bosons. The last one sums up the different interactions [1].*
NOTE "residual strong interaction": the strong interaction of color-neutral protons and neutrons to form nuclei is due to a residual strong interaction between their color-charged constituents, like e.g. the Van der Waals forces in water which is an electric residual force.

with the other particles, its so-called couplings, are calculable within the Standard Model, but its mass is not directly predicted. However it is possible to constrain the Higgs mass by assuming the overall validity of the Standard Model and combining experimental observables that are sensitive to the Higgs mass through electroweak corrections [3]. Putting all the electroweak observables together, the best-fit value of the expected Higgs mass given by the Standard Model is 117 GeV, the upper limit, with 95% confidence level, being 251 GeV [2]. Moreover, direct search at LEP excluded a Higgs with a mass smaller than 114.4 GeV with a 95% confidence level [3].

There are many reasons to think that the Standard Model is not the final theory and that new phenomena should appear at higher energies. Taking into account that the Standard Model is thought to be an effective theory, valid up to a given energy scale (Λ) and which suffers from theoretical problems that could be solved by replacing or extending it, the next question would be to determine up to what energy the Standard Model is valid or alternatively to try and discover new physics. Some alternative models have been developed and are waiting to be confirmed or refuted by observations. The most popular extension of the Standard Model is Supersymmetry which postulates a symmetry between fermions (half integer spin) and bosons (integer spin). It solves some theoretical problems of the Standard Model and manages to unify the strong and the electroweak interactions. Within this theory, the number of particles is doubled, since for each fermion/boson a corresponding supersymmetric boson/fermion should exist. However, up to now, no supersymmetric particles have been seen. Another challenge is to develop a theory which brings gravity and the other interactions into a coherent picture and which makes measurable predictions.

To search for the Higgs or phenomena beyond the Standard Model, a new proton-proton collider, the **LHC**, is being built at CERN (Geneva, Switzerland) and is currently expected to start running in Summer 2007 at energies never reached before. At Fermilab, close to Chicago (USA), there is a proton-antiproton collider, the **Tevatron**, which started its second phase of data taking in 2001. The Tevatron and the LHC have a very similar working environment as they both collide hadrons. Thus many experimental physics aspects relevant for the LHC can be studied now at the Tevatron.

For this work, data from **CDF**, one of the detectors at Tevatron, were analyzed with the purpose of understanding better physics problems that will be encountered at the LHC. Then, specific questions related to **CMS**, one of the two general purpose detectors at LHC, will be studied, using Monte Carlo simulations.

In the following sections some general statements will be made on the physics at hadron colliders. Two of its aspects will be developed: the discovery of new particles and precision measurements made to test the predictive power of a theory.

Discovery at hadron colliders

Throughout the history of Particle Physics, hadron colliders have been attributed the role of "discovery machines" par excellence. Due to the high mass of the proton, compared for instance to the electron, it is easier to accelerate and collide protons to reach energies never investigated before. The W and the Z bosons for instance were discovered at the CERN $Sp\bar{p}S$ and the top quark at the Tevatron, both proton-antiproton colliders. However the price to pay for using hadrons is a harder event analyze and reconstruction since protons have substructure and are strong interacting objects.

In order to be detected, a process of interest should have a high enough cross section. The whole 'art' of the physicist will consist in trying to reduce the backgrounds, keeping at the same time a large enough signal. Traditionally a discovery is claimed when a deviation of more than

[3] One should be careful with the interpretation of this fit results since the Higgs mass comes in the electroweak corrections like $\ln m_H$, whereas the top mass, known to an accuracy of 3%, comes in like m_t^2. A 2% variation in the top mass results in a 13% variation in the Higgs mass limit.

five standard deviations (σ) away from the background is observed, where $\sigma = S/\sqrt{B}$ [4], with S being the number of signal events and B the number of background events. Exploring new energies and trying to find new physics will very often resemble to looking for a needle in a haystack as small interesting signals have to be separated from huge backgrounds. Since no one knows what to expect at new energy scales, the detector design has to be as general and flexible as possible. A special interest should be paid to high transverse momentum physics, as the presence of high transverse energy particles is often the sign of the decay of a heavy particle.

In general, new heavy particles have a very short lifetime and decay before they reach the detector elements. A new phenomenon will be mostly identified through a deviation in a kinematic distribution of specific well-known "stable particles" and jets. For instance, a new particle is often discovered by studying invariant mass distributions. The signatures of a detectable process can be sorted into two categories. The 'easy' ones, for which a narrow mass peak can be reconstructed and the ones for which a mass peak cannot be reconstructed. The signal has then to be found in distributions like e.g. the transverse momentum of a particular particle, the missing energy of the event or specific event topologies. The first type of signatures will need a very good energy and momentum determination as the width of the mass peak will depend mainly on the energy resolution, whereas for the second type of signatures, a very good angular determination and an extended detector coverage will be needed. Table 1 shows which types of measurements could reveal which types of new physics.

To determine whether a new particle can be discovered, it is not sufficient to consider $S/\sqrt{B}$ it is also important to take S/B into account. For signatures where a narrow mass peak can be reconstructed, a discovery can be made already with a low signal to background ratio, since the backgrounds are well under control as they can be estimated from the regions away from the peak. On the contrary for the other type of signatures, a high signal to background ratio has to be required, since the backgrounds have sometimes to be estimated from a Monte Carlo simulation and fluctuations due to the uncertainty on this estimation can easily hide a signal.

The Higgs search at the LHC provides a good illustration of these ideas. If the Standard Model Higgs has a mass of 130 GeV, its decay into two photons would be a channel to look for. It is expected to be detected as a peak above a large irreducible background. Figure 2 (Top) shows how this peak might look like in CMS. If the Higgs has a mass around 140 GeV, it could be detected in the following decay chain: $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$. For this channel, no Higgs mass peak can be reconstructed, a signal could thus be seen in the lepton transverse momentum spectrum, as illustrated in Figure 2 (Bottom). It is evident from this plot that in this case, a precise background normalization is more difficult to achieve.

Precision at hadron colliders

Whereas trying to discover a new particle may look more appealing, precision measurements are another important aspect of the physics program of hadron colliders.

Precision measurements are fundamental to test the predictive power of the Standard Model or any competing theory at the new energy scale probed. This will be done for instance by comparing the cross section of different processes with the theoretical predictions. Moreover, after having discovered a new particle, it is important to measure its properties in order to constrain the theoretical frame.

Precision measurements are also needed to enable discoveries, in keeping the background under control. More precisely, it is fundamental to get an accurate measurement of the kinematics and the cross sections of 'well-known' Standard Model processes which are backgrounds for discovery channels. This is especially important in cases where the Standard Model background

[4] This formula holds for S and B higher than $\approx$20. Otherwise, σ has to be estimated using the Poisson statistics. Note that other estimators of the statistical significance can also be used, like for instance the ones based on likelihood ratios.

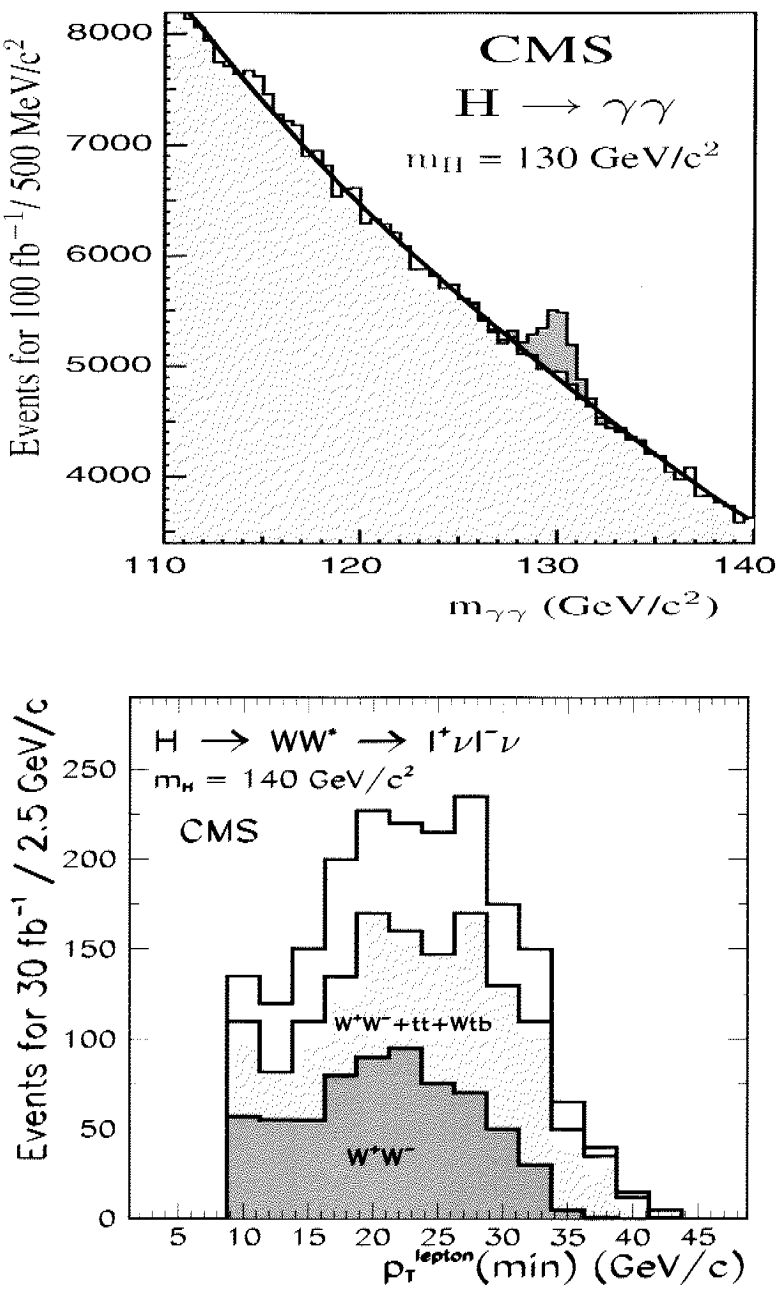

FIGURE 2: *Two different ways to look for the Standard Model Higgs at the LHC: its decay into $\gamma\,\gamma$ where the signal is reconstructed as a mass peak (top plot) and its decay into $WW \to \ell\nu\ell\nu$ where the signal is seen in the lepton transverse energy spectrum (bottom plot). The top plot shows the reconstructed dipholon mass for a luminosity of 100 fb^{-1} and a Higgs mass of 130 GeV. The bottom plot shows the transverse momentum of the lepton with the lowest momentum, for a 140 GeV Higgs decaying into $WW^* \to \ell\nu\ell\nu$ and a luminosity of 30 fb^{-1} (in white) together with the main backgrounds (in yellow and blue) [4].*

Type of measurement	indicates	required for
isolated high p_t $e^{\pm}$, $\mu^{\pm}$	$W^{(*)}$, $Z^{(*)}$ decays	Higgs search top physics, 'all' searches (e.g. Z')
isolated high p_t γ's	electromagnetic process	Higgs search
τ and b-quark tagging	'rare' processes	special Higgs-like searches, Supersymmetry
large missing p_t, E_t	event with ν-like particles; W, Z decays	Higgs, Supersymmetry, exotic 'exotica'
jets	quarks and gluons	QCD, understanding of backgrounds/efficiencies

TABLE 1: *Examples of what a detector should be able to measure in order to explore a new energy scale [5].*

for a given discovery channel has to be estimated from a Monte Carlo simulation.

The event reconstruction and identification at hadron colliders is quite complicated. In fact, the detector will have to isolate from the whole proton-(anti)proton interaction, the products of the parton-parton interaction to be studied. A way to solve that problem is to get a high number of collisions, increasing the statistics of the processes. Like that, harder requirements can be put on the event selection, allowing to use only the well-reconstructed events. A high collision rate will however require a fast detector response, as well as a good triggering efficiency. Thus, to reach a good precision, stringent requirements have to be put on the design of the detectors in order to obtain an excellent energy and momentum resolution, as well as the largest acceptance possible.

Precision measurements is a general concept. In the following, we shall discuss the case of cross section measurements and how to constrain the parameters of the Standard Model, giving examples from the LHC and the Tevatron.

Measuring cross sections

The measurement of the cross sections for different processes allows to test the predictions of the Standard Model at the energy scale probed. This can also lead to a discovery since an excess in the cross section could be the sign of a new phenomenon. The cross section depends on the following variables: the number of signal events after background subtraction, $S = N - B$, the cut efficiency, ε_{eff} and the luminosity, $\mathcal{L}_{pp}$:

$$\sigma = \frac{S}{\varepsilon_{eff} \cdot \mathcal{L}_{pp}}$$

Therefore, three main factors can limit the precision of a cross section measurement:

1. *The number of signal events (S)*

   The statistics will depend on the energy of the colliding particles and on the luminosity delivered by the accelerator. The error on the number of signal events scales like $\sqrt{\mathcal{L}}$.

   For instance, the top cross section measurement at the Tevatron Run I was mainly limited by the statistics available [6]. The error was about 25% and all the other uncertainties could be reduced to about half. In the case of the LHC, running at much higher energy with a high design luminosity should produce about 10^7 $t\bar{t}$ which is enormous compared to the total of 10^4 events at the Tevatron, thus resulting in a very small statistical error on the cross section. Figure 3 shows the proton proton cross section energy dependence for

different processes together with some experimental measurements. Many processes will be measured with a good statistical accuracy at the LHC.

2. *The luminosity ($\mathcal{L}_{pp}$)*

   Next is the uncertainty related to the luminosity determination. The luminosity is a variable which depends on the machine parameters. It is usually obtained at hadrons colliders by measuring the proton-(anti)proton elastic or inelastic cross section using counters close to the beam line. The error on the luminosity at the Tevatron is currently 5%. The main reason is the uncertainty on the proton-antiproton total inelastic cross section and the difficulty in estimating the acceptance of the luminosity detectors. The prospects for the CMS experiment were assumed to be similar, as stated in the 1994 Technical Proposal.

   In the following we discuss a proposal to reduce significantly the error on the luminosity determination by using single W and single Z production, following a method described in [8], the so-called parton luminosity. The goal is to bring this error down to 1%. A first application of this method on real data could already be performed using the Tevatron Run II data and will be presented in Chapter 2.

   An example of a cross section measurement where the limiting factor was the error on the luminosity, is the W and Z cross section at the Tevatron Run I [9]. There, the signatures are very clean leading to small systematic uncertainties. The statistical errors are also low resulting from the relatively high cross section of these processes.

3. *The efficiency (ε_{eff}) and the background estimate (B)*

   The last source of error is linked to the cut efficiency and background estimates. Such uncertainties become usually important when the signature is difficult to reconstruct. They are also quite dependent on the detector type, depending, for instance, on the energy and momentum resolution. In this case, it is hard to make general statements. An increase in the number of events should have only a limited effect on these systematic errors. However, what will be possible at LHC is to use reference processes like single W and single Z production to control the efficiencies and the background. In Chapter 3, some systematic effects will be studied for CMS using a full simulation.

   An example of a cross section measurement which was limited by systematics is the dijet cross section at the Tevatron Run I [10], as it is experimentally difficult to reconstruct jets and to interpret the signature.

At the LHC the uncertainty on the cross section of many processes should be significantly reduced for the following reasons: high statistics for many interesting processes will be reached, allowing to almost neglect the errors due to the statistics and to make sufficiently hard cuts to get the background systematics under control. However this increases the errors on the cut efficiency. Moreover, the use of reference processes like single W and single Z production to determine the parton luminosity but also to study some systematics should allow the errors on the measured cross sections to be significantly reduced.

Using precision measurements to constrain model parameters

Precision measurements should allow to test the consistency of a model at a given energy scale by constraining the different parameters of the model. For instance, there are about 19 free parameters in the Standard Model (plus another 10 assuming massive neutrini). Apart from that, alternative theories predicting new physics have free parameters which have to be extracted from the data. Precision measurements can also be used to discriminate between the different models describing a newly discovered particle. For these different questions, precision measurements of cross sections are well suited. Other potential good observables are particle masses and

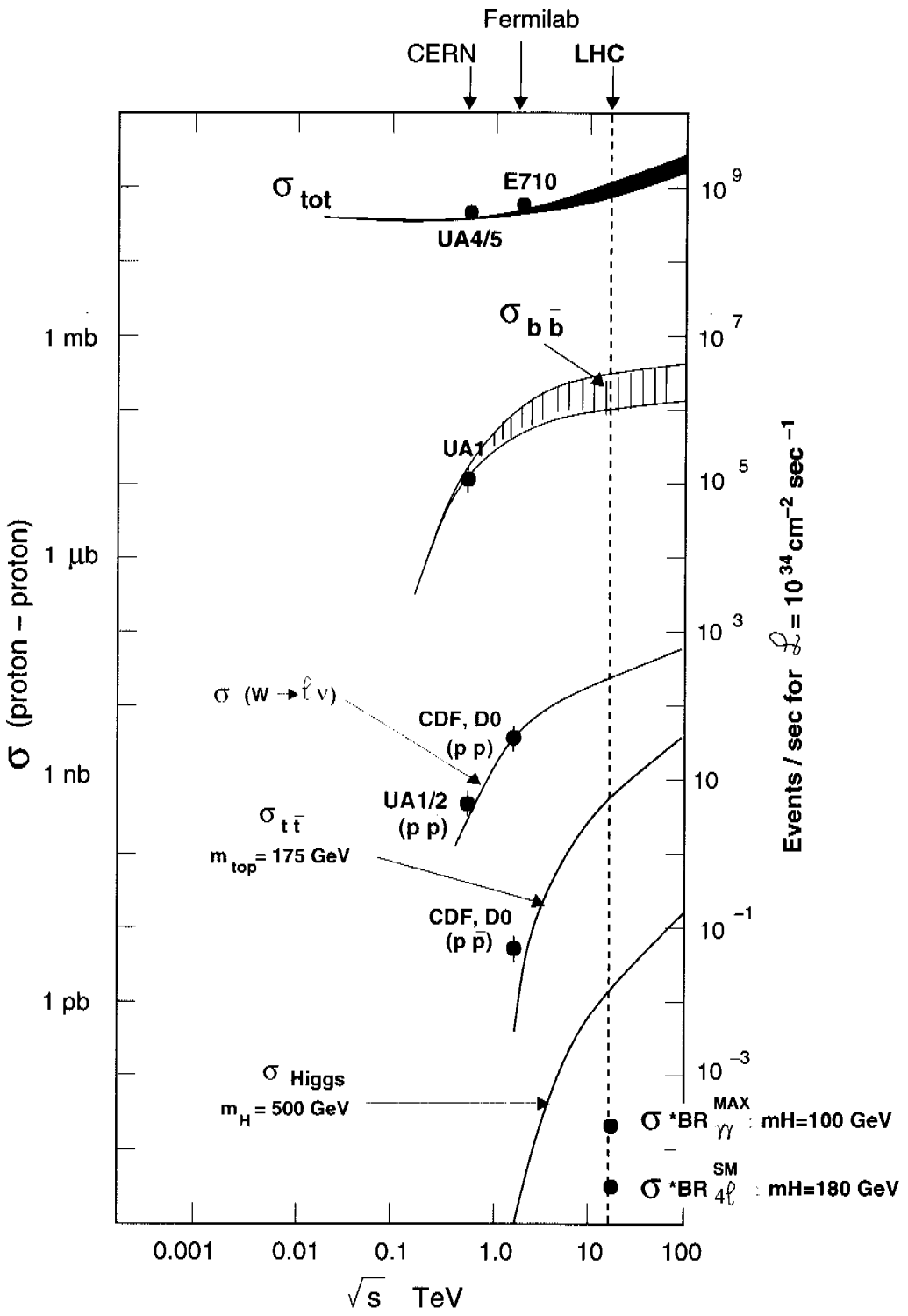

FIGURE 3: *The proton-proton cross section energy dependence for different processes together with some experimental results. The points corresponding to the Tevatron results (CDF, DØ) are not perfectly on the curves, since they measured proton-antiproton cross sections [7].*

couplings. We give in the following some example on how precisely masses and couplings of Standard Model particles are expected be measured at the LHC.

- Masses

  It will be possible to measure with a certain accuracy the masses of the "heavy particles" like the Higgs, the top and the W at the LHC. This will allow the Standard Model to be over-constrained, as these masses are dependent from each other. This type of measurement will need a very good determination of the lepton energy and momentum scale as well as of the jet scale, since the mass is determined from a mass peak (or a transverse mass peak). For instance, the Higgs mass could be determined at the LHC with an accuracy of $\Delta M_H/M_H = 0.1\%$ [11] for almost the whole possible Higgs mass range. The top mass could be measured with an accuracy of 2 GeV (the present uncertainty lies around 5 GeV). The uncertainty on the W mass is expected to be reduced down to 25 MeV [11]. The last results from LEP and Tevatron claim an error of 40 MeV, respectively 60 MeV on the W mass [12],[13].

- Couplings

  The couplings of a given particle can usually be measured by comparing different processes. For instance, if the Higgs has a mass between 300 and 600 GeV, the LHC experiments should be able to constrain the relative Higgs couplings to fermions and bosons. This can be done by measuring the ratio of the cross section where the Higgs is produced through weak boson fusion, where it couples to W and Z and the Higgs produced though gluon fusion, where it couples to the top quark. This measurement could be achieved with a statistical precision of around 10% [14]. In that case, as ratios are measured, the hope is to get systematic errors much smaller than the statistical ones.

  Another example is the measurement of the couplings of the vector bosons (W, Z and γ) with each other, the so-called triple gauge couplings, which are quite sensitive to new physics. Although the Standard Model is tested up to a precision of 0.1% or better, the parameters characterizing the triple gauge couplings are known with a precision of 10% from LEP and the Tevatron. One goal of the LHC would be to improve this measurement by about a factor of 10. A way of determining these couplings is to fit the measured transverse momentum spectrum of one of the vector bosons, for instance by measuring the p_l spectrum of the Z in $pp \to ZW \to \ell\ell\ell\nu$ process.

From a precise determination of the different particle couplings, particles masses and cross sections of typical processes, it is possible to test the predictions of the Standard Model over a range of energies. Moreover, if a new particle is discovered, the measurement of its couplings, its mass and other parameters that are characteristic of this particle like its spin, should allow constraining the different models. An example with an exotic particle, the Z', will be presented in Chapter 4.

In the next chapter, the CDF and CMS detectors will be described and compared. Then, the method using single W and single Z production to measure the luminosity, the so-called parton luminosity, will be presented and a first application will be carried out on the CDF data. It will be shown that, in the case of the Tevatron, this method is competitive with the 'traditional' luminosity measurement employed up to know by CDF. For the LHC the parton luminosity looks promising due to the high expected statistics of single W and Z production. This represents a first step toward precision measurement at the LHC.

Then we will show how electroweak processes such as single W production could be reconstructed with the CMS detector and what selection procedure could be used for the electron

reconstruction. Moreover, some systematics issues using a full detector simulation will be analyzed in order to discuss the CMS detector capabilities.

Finally it will be shown that – thanks to precision measurements of specific variables – it should be possible to identify a hypothetical new gauge boson, the Z', at the LHC.

Chapter 1

The detectors: CDF and CMS

The results described in this thesis were obtained by analyzing the data of the CDF detector and from simulations of the CMS detector. CDF and CMS are multi-purpose detectors built to analyze hadron collisions from respectively the Tevatron accelerator, colliding protons with anti-protons at an energy of 1.96 TeV and the LHC accelerator, which will collide protons with protons at an energy of 14 TeV. First we give a description of those two detectors and then analyze their common points and differences.

1.1 The CDF detector

The Collider Detector at Fermilab (CDF) is a general purpose detector located in the Tevatron accelerator at Fermilab. The first collisions were produced and detected in October 1985. From 1992 to 1996, the detector recorded about 110 pb^{-1} of data during the so-called Run I. The analysis of this data resulted in about 200 publications, the highlight being the discovery of the top quark in 1994 [15]. This was followed by a major accelerator and detector upgrade completed in 2001. CDF started a second Run of data taking in April 2001.

1.1.1 The Tevatron accelerator

The Tevatron is a synchrotron ring of 1 km radius designed to accelerate protons and anti-protons up to a collision center of mass energy of $\sqrt{s} = 1.96$ TeV, using super-conductive magnets with a field of 4.2 T. Figure 1.1 shows a simplified view of the acceleration system.

In this $p\bar{p}$ collider, protons are produced by ionizing hydrogen. The H^- ions are accelerated up to an energy of 750 keV by a Cockroft-Walton accelerator followed by a Linac which brings the particles up to an energy of 400 MeV. The ions are then passed through a thin layer of graphite to strip the electrons. The resulting protons are accelerated in a small 150 m radius synchrotron (Booster) up to an energy of 8 GeV, divided into bunches and then passed to the Main Injector.

The anti-protons are produced by the interaction of 120 GeV protons from the Main Injector and a target of nickel. After being collimated at an energy of 8 GeV they are debunched into a continuous beam and stochastically cooled down to reduce the energy dispersion. The anti-protons are then accumulated in the accumulation ring. When the number of anti-protons is sufficiently high they are injected into the recycler ring. The recycler ring is one of the accelerator improvements in Run II. It allows to store the anti-protons from the accumulator and it also recycles the anti-protons that did not interact in the Tevatron.

After the first acceleration step, protons and anti-protons are divided into 36 bunches and accelerated to 980 GeV. The instantaneous luminosity is given by the following formula:

$$\mathcal{L}_{int} \propto \frac{fBN_pN_{\bar{p}}}{2\pi(\sigma_p^2\sigma_{\bar{p}}^2)}$$

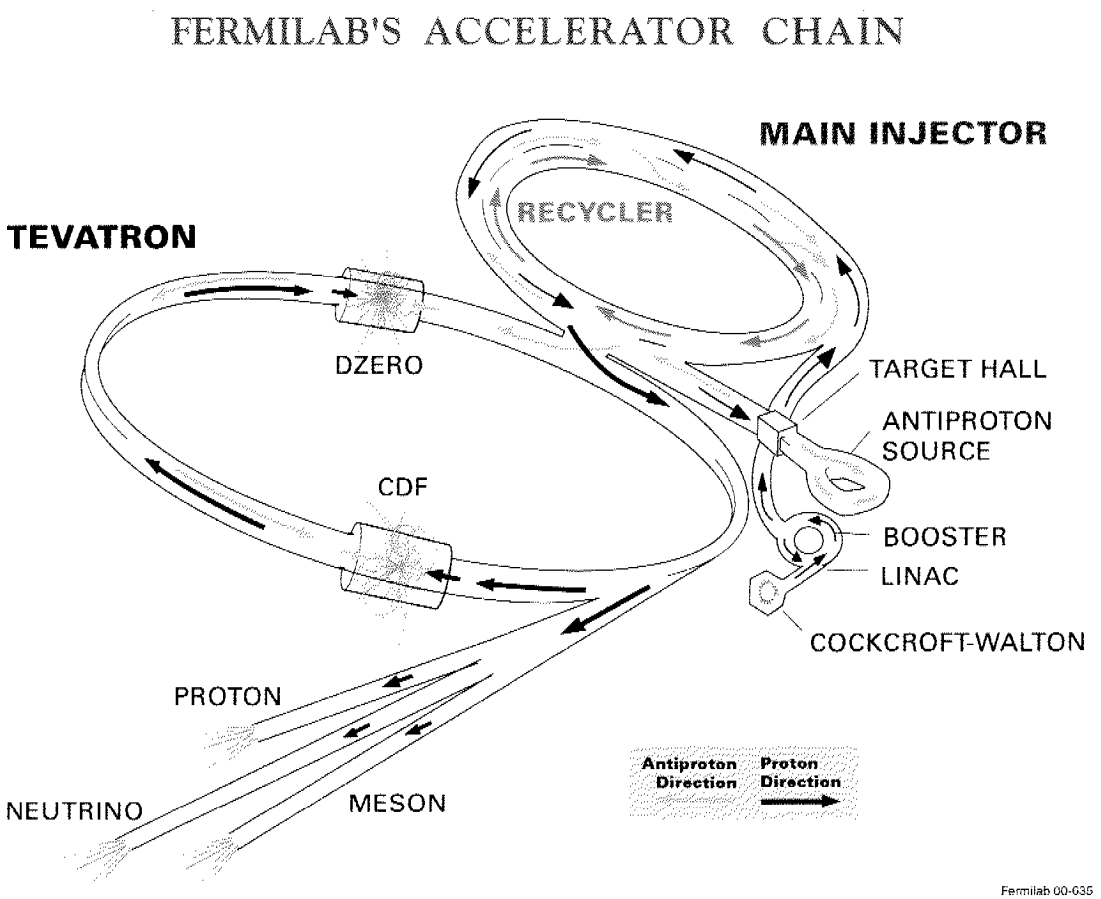

FIGURE 1.1: *The Fermilab accelerator chain.*

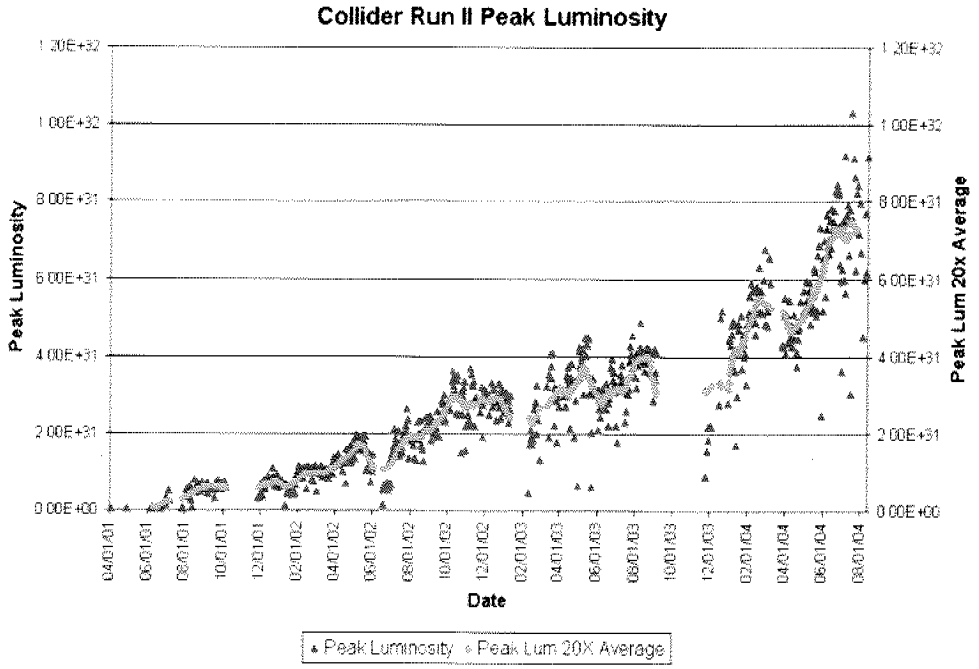

FIGURE 1.2: *The Tevatron collider peak luminosity for Run II [16].*

where f is the revolution frequency, B is the number of bunches, N_p and N_p the number of protons and anti-protons per bunch and σ_p^2 and σ_p^2 are the transverse dimensions of the beams at the collision point. The proportionality factor depends on the longitudinal resolution of the beam and on his dispersion in the phase space. In Run II the number of bunches is increased, but the number of particles per bunch is roughly the same as in Run I. The luminosity at the end of 2003 (when the study described here was done) was about $\mathcal{L} \sim 3 \times 10^{31}\text{cm}^{-2}\text{s}^{-1}$, as Figure 1.2 shows. In June 2004, a record luminosity of $1.02 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$ has been reached.

Two detectors are build along the Tevatron, CDF and DØ. Their physics program is similar. In the following the CDF detector will be described.

1.1.2 The CDF detector

CDF consists of the following elements: a tracking system to detect tracks of charged particles, an electromagnetic and hadronic calorimeter to identify and measure the energy of the electrons, photons and jets; and finally the muon chambers. Figure 1.3 shows a view of the CDF detector. A detailed description of the detector can be found in [17]. In the following, we give a brief description of the tracking system, the calorimeters, the trigger system and the detectors used for the luminosity determination as these are the sub-detectors relevant for our analysis.

The tracking system

The tracking system was fully replaced for Run II and is shown on Figure 1.4. Charged particle tracks are reconstructed using the combined informations from silicon-based detectors in the most inner part and cell-drift chambers in the outer part.

The silicon tracking system is composed of 3 different detectors subsystems: Layer00 is a layer of silicon detectors installed directly on the beam pipe to improve the impact parameter resolution. Then comes the silicon vertex detector (SVXII), which is composed of five layers of double-sided silicon sensors with a combination of both 90-degree and small angle stereo layers to give 3-dimensional information. The expected vertex resolution is $\sigma_\phi < 30\,\mu\text{m}$ and $\sigma_z < 60\,\mu\text{m}$. Finally, the Intermediate Silicon Layer (ISL) is a large radius silicon tracker with a total active

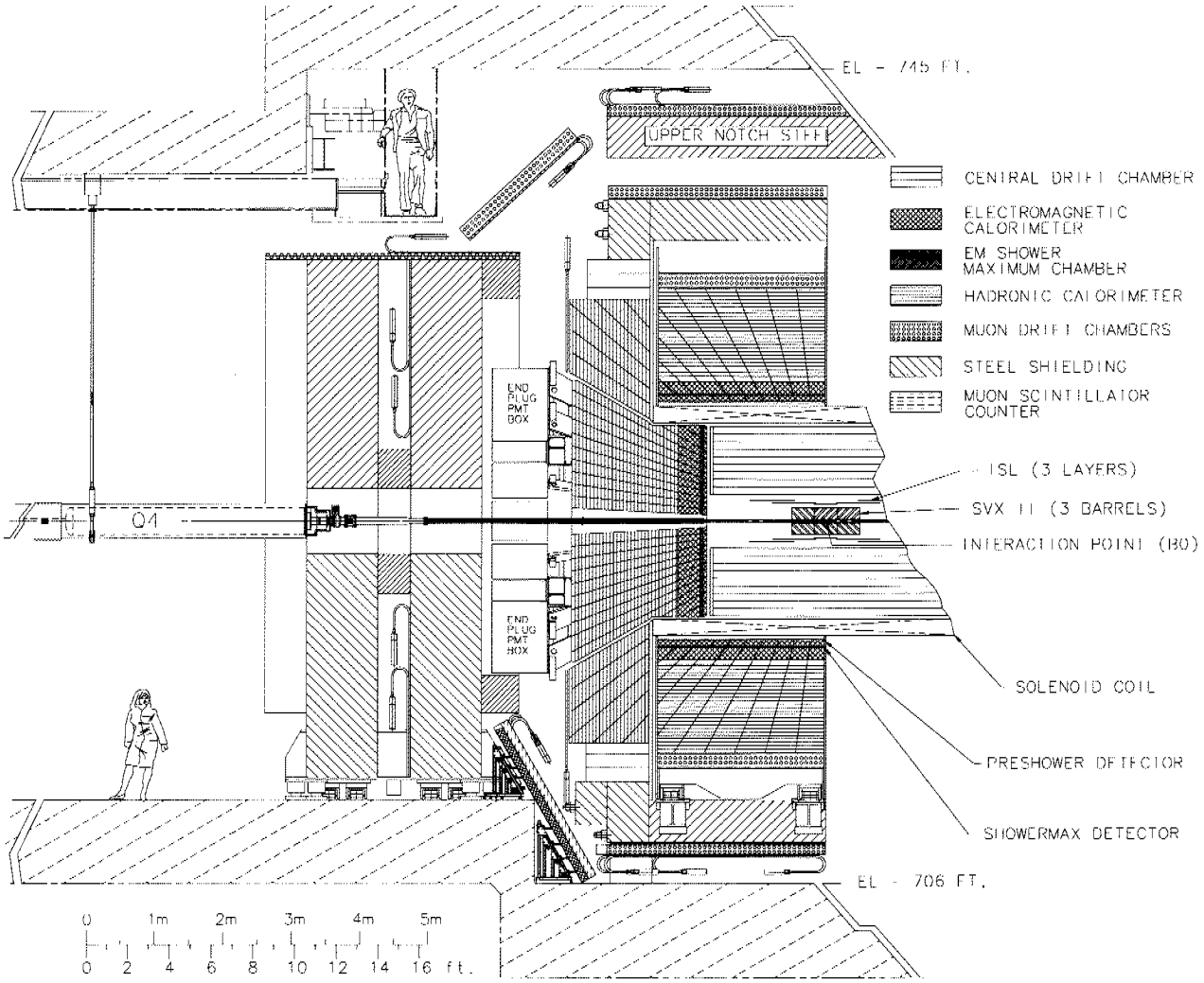

FIGURE 1.3: *Elevation view of one half of the CDF II detector [17].*

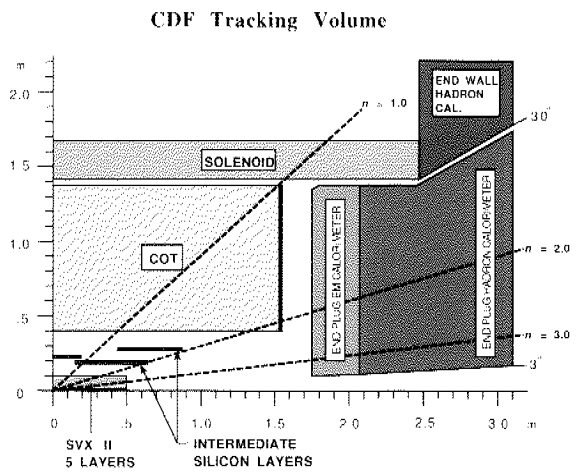

FIGURE 1.4: *Longitudinal view of the CDF II Tracking system [17].*

area of about 3.5 m^2. It is composed of 296 basic units, called ladders, made of three silicon sensors bonded together in order to form one electrical unit. It is located between the silicon vertex detector and the central outer tracker. It covers a pseudorapidity [1] region up to $|\eta| = 2$.

The Central Outer Tracker (COT) is the new CDF II central tracking chamber. It is an open cell drift chamber with a maximum drift time of about 100 ns and thus able to operate at a beam crossing time of about 132 ns. The COT consists of 96 layers arranged in four axial and four stereo superlayers. Combining the hits to reconstruct the tracks, the COT reaches an efficiency close to 100% for high p_t isolated tracks. The expected hit resolution, based on the Run I experience is $\sigma < 180\,\mu$m. It also provides some dE/dx information for particle identification.

Between the COT and the solenoid is the new time of flight detector (TOF). This scintillator based detector with a 110 ps resolution has the capability to tag charged kaons in the p_t range from 0.6 to 1.6 GeV.

Around the tracking system stands the superconducting coil. With a current density of 1200 A/m, it provides an axial magnetic field of 1.5 T over a useful volume of 2.8 m in diameter and a length of 3.5 m. The design momentum resolution of the tracking system is $\delta p_t/p_t^2 = 0.1\%$ for $|\eta| < 1.0$ and of $\delta p_t/p_t^2 \sim 0.4\%$ $1.0 < |\eta| < 2.0$.

The calorimeter system

Outside the solenoid, a scintillator-based calorimeter covers the region $|\eta| < 3.6$. As shown in Fig.1.3 the calorimeter consists of an inner electromagnetic section followed by an outer hadronic section. Both sections consist of alternate layers of scintillators and passive material. Lead is used for the electromagnetic calorimeter and iron for the hadronic one. The calorimeter can be divided in a central region (detecting particles up to $|\eta| = 1.1$ for an event vertex in the detector center) and a forward region (detecting particles in the range $1.1 < |\eta| < 3.6$). Good identification of isolated electron is possible up to a pseudorapidity of $|\eta| < 2$. The central region is instrumented by the central electro-magnetic calorimeter (CEM) and the central hadronic calorimeter (CHA). The CHA is then further extended in η by the endwall hadron calorimeter (WHA). The calorimeter system is segmented in towers along the η and ϕ directions pointing toward the interaction point. The towers are common to the electromagnetic and hadronic calorimeter. In the central (forward) region, each tower covers a $\Delta\phi$ of 15° (7.5°-15°) and a $\Delta\eta$ of 0.11 (0.09-0.64).

The resolutions obtained in a test beam (single electrons and pions) are $\frac{\Delta E}{E} = \frac{16\%}{\sqrt{E}} \oplus 1\%$ for the central electromagnetic calorimeter and $\frac{\Delta E}{E} = \frac{80\%}{\sqrt{E}} \oplus 5\%$ for hadrons [17]. In the central region there are also two position detectors, the central pre-radiator gas chamber (CPR) just before the electromagnetic calorimeter and the central E-M strip/wire gas chamber (CES) inside it. The shower maximum chambers contribute to the identification of electrons and photons using the position measurement to match the "clusters" with the tracks, the transverse shower profile to separate photons from π^0s, and the pulse height.

[1] A commonly used variable in hadron colliders is the *rapidity*, Y, defined like:

$$Y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}$$

It is defined with respect to a direction, z, usually taken parallel to the beam axis. If $p >> m$, the rapidity can be approximated by the *pseudorapidity*, defined as:

$$\eta = -\ln(\tan(\theta/2))$$

where θ is the angle between the particle momentum and the beam axis.

The trigger system

Due to the very high total interaction rate compared to the relatively small rate of interesting interactions (as shown in Figure 3, on page 8), it is fundamental to have a very sophisticated trigger able to extract the most interesting physics events from the large number of collisions. In Run II, the maximum data to disk recording rate for CDF is 75 Hz, while the collision rate is 2.5 MHz. The CDF II trigger is organized in three different levels. The Level 1 trigger (L1) provides fast drift chamber tracks, muon and electron triggers based on signatures in the muon chambers and calorimeter in combination with drift chamber tracks, and calorimeter based triggers based on missing transverse energy, photons and jets. This trigger has almost no dead-time. It possesses a 42 stage pipeline and can make a trigger decision every 132 ns with a total latency time of 5544 ns. A new on-line processor reconstructs COT tracks (eXtremly Fast Tracker). The L1 accept rate achieved in 2003 was 20 kHz. The L2 trigger adds information within about 22 μs, to the objects found by the L1 trigger. It does a limited event reconstruction which is processed in programmable processors and reduces the event rate to about 300 Hz. The events that pass L2 are processed by a network of parallel processors called farms, which do a final filtering, reducing the event rate down to 50 Hz.

1.1.3 The proton–anti-proton luminosity measurement at CDF

The method used by CDF to measure the luminosity is based on the rate of inelastic $p\bar{p}$ interactions[2], using the formula: $\mu \cdot f_{BC} = \sigma_{in} \cdot \mathcal{L}$ where μ is the average number of interactions per bunch crossing and f_{BC} is the rate of bunch crossings in the Tevatron.

The luminosity monitors, the so-called Čerenkov Luminosity Counters (CLC), shown on Figure 1.5, are designed to measure μ within a few percent, since the average number of interactions per crossing is proportional to the average number of hits in the CLC per bunch crossing. To detect efficiently inelastic $p\bar{p}$ events, the luminosity detectors are put at small angles where the cross section for that process is higher. Two Čerenkov modules were installed in the proton (east) and anti-proton (west) directions with a rapidity coverage between 3.75 and 4.75. Each module consists of 48 thin, long, gas-filled, Čerenkov counters. The counters are arranged around the beam-pipe in three concentric layers, with 16 counters each, and pointing to the center of the interaction region. They are built with reflective aluminum mylar sheets of 0.1 mm thickness and have a conical shape. The Čerenkov light is detected with fast photomultiplier tubes. The counters are mounted inside a thin pressure vessel made of aluminum and filled with isobutane. A primary particle from $p\bar{p}$ interaction will cross the full length of the counter and generate a large PMT signal, while the secondary particle produced in the beam pipe and materials surrounding the CLC cross the counter at different angles and yield much smaller signals.

The proton–anti-proton luminosity is then determined using the rate of the inelastic $p\bar{p}$ events measured with the CLC luminosity monitor $R_{p\bar{p}}$, the CLC geometric acceptance ε_{clc} and the inelastic cross-section σ_{in}:

$$\mathcal{L} = \frac{R_{p\bar{p}}}{\varepsilon_{clc} \cdot \sigma_{in}}$$

with the systematic errors of 1.8% on $R_{p\bar{p}}$, 4.0% on ε_{clc} and 3.8% on σ_{in} respectively. The total systematic error on the luminosity was estimated to be 5.8% [18].

[2]The CLC has zero acceptance for elastic $p\bar{p}$ events.

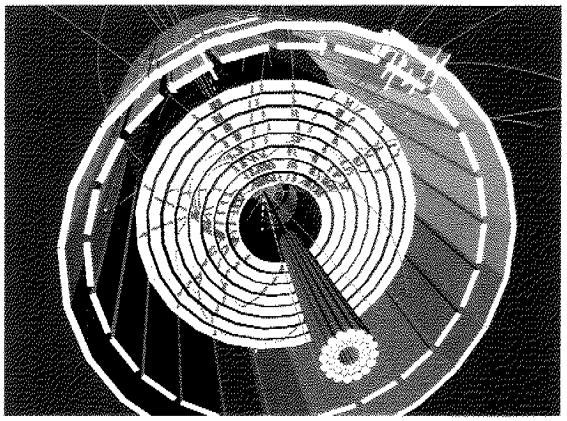

FIGURE 1.5: *A schematic view of the CLC.*

1.2 The CMS detector

The Compact Muons Solenoid (CMS) experiment is a general purpose detector which will be located at the Large Hadron Collider (LHC) accelerator at CERN (Geneva). The present schedule foresees the first proton proton collisions in Summer 2007.

1.2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a proton-proton collider with a center of mass energy of 14 TeV and a design luminosity of $\mathcal{L} = 10^{34}$ cm^{-1} s^{-1}. The LHC, which is currently being constructed in the existing LEP tunnel, has a circumference of 27 km. Figure 1.6 shows a schematic view of the LHC ring and Table 1.1 presents some of the LHC parameters.

The LHC will be connected to the CERN accelerating facilities. The SPS accelerator, bringing protons from an energy of 26 GeV up to an energy of 450 GeV will inject protons into the LHC ring. Inside the LHC, the protons will then be accelerated to an energy of 7 TeV with the help of sixteen 400 MHz radio-frequency cavities, assembled in 4 modules.

The design of the LHC is made in such a way that protons will circulate in both directions in two parallel vacuum tubes embedded in one magnet system (the so-called '2-in-1' magnet). The magnetic field to bend the trajectory of the protons is thus quite complicated, since it has to be in the opposite direction for each of the two tubes, as shown in Figure 1.7. There will be 1232 main dipoles and about 400 quadrupoles along the ring to guide the particles and correct their trajectories. Given the radius of the tunnel and the energy to be reached one can calculate that the magnetic field to keep the protons inside the ring has to be of about 8.3 T. The only way to obtain such a high magnetic field is to use superconductivity. In order to make the magnet superconductive, the dipoles will be cooled down to 1.9 K with super-fluid Helium, implying a large cryogenic system inside the LHC tunnel.

Currently the LHC is supposed to start with an instantaneous luminosity of 10^{33} cm^{-2} s^{-1}, followed by a second phase where the accelerator will operate with an instantaneous luminosity of 10^{34} cm^{-2} s^{-1}. The LHC should also run a couple of months every year colliding lead nuclei. The center-of-mass energy per nucleon pair should be of 5.4 TeV. In such high energy heavy nuclei collisions, a quark-gluon plasma might be formed and its properties will be studied.

Figure 1.6 shows the location of the four detectors which will study collisions at the LHC: ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muons Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb. The first two are multi-purpose detectors. Their main physics goal will be the Higgs search and the exploration of possible new phenomena at high masses. ALICE is a dedicated detector, which will analyze the collisions of heavy ions. Finally LHCb is

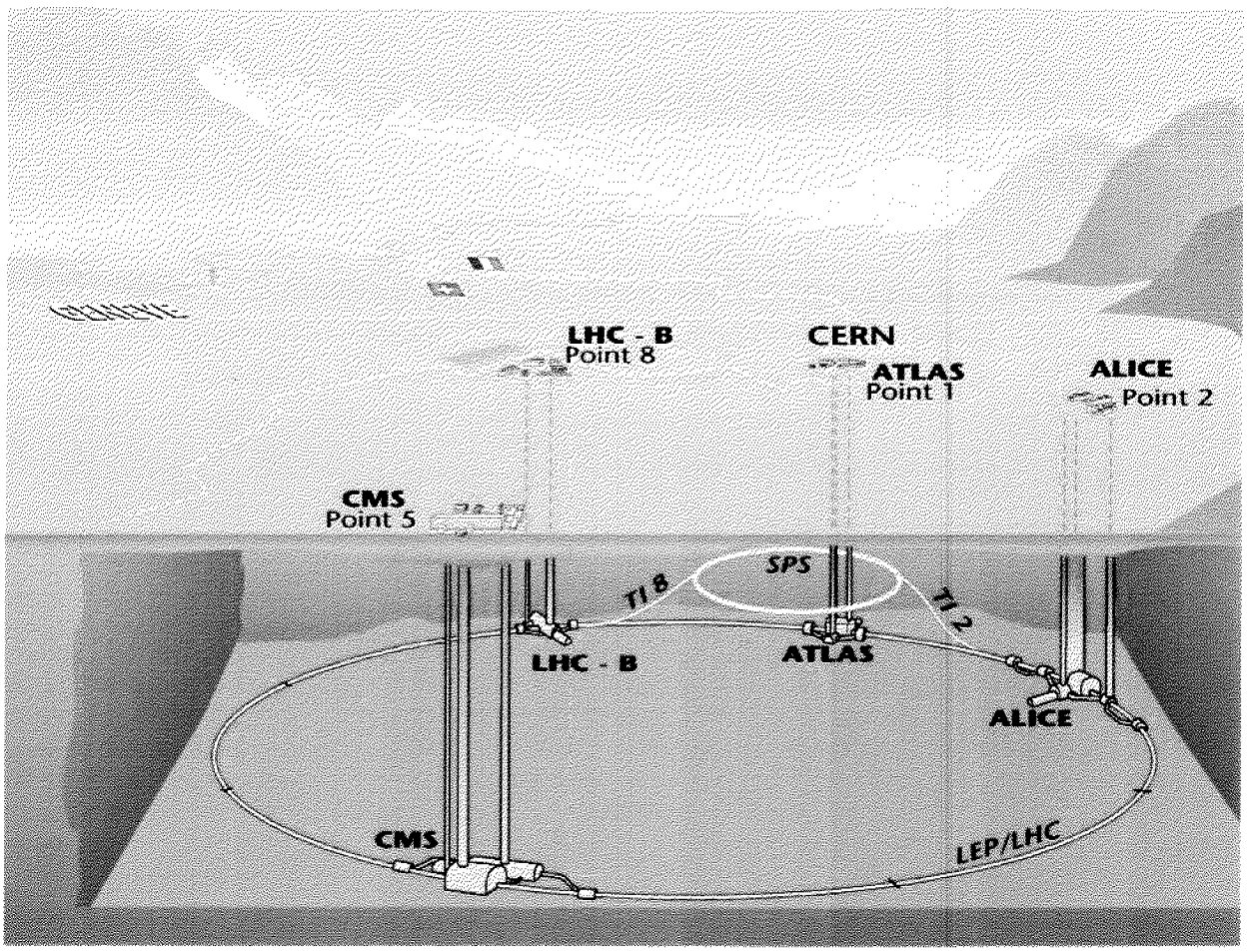

FIGURE 1.6: *Schematic view of the four LHC experiments [19].*

Momentum at collision	7 TeV
Momentum at injection	450 GeV
Dipole field at 7 TeV	8.33 Tesla
Circumference	26658 m
Design Luminosity	10^{34} cm^{-2} s^{-1}
Number of bunches	2808
Particles per bunch	$1.1{\cdot}10^{11}$
Collision rate	45 MHz
Luminosity per crossing	2 barn
DC beam current	0.56 A
Stored energy per beam	350 MJ

TABLE 1.1: *Some of the LHC design parameters [20].*

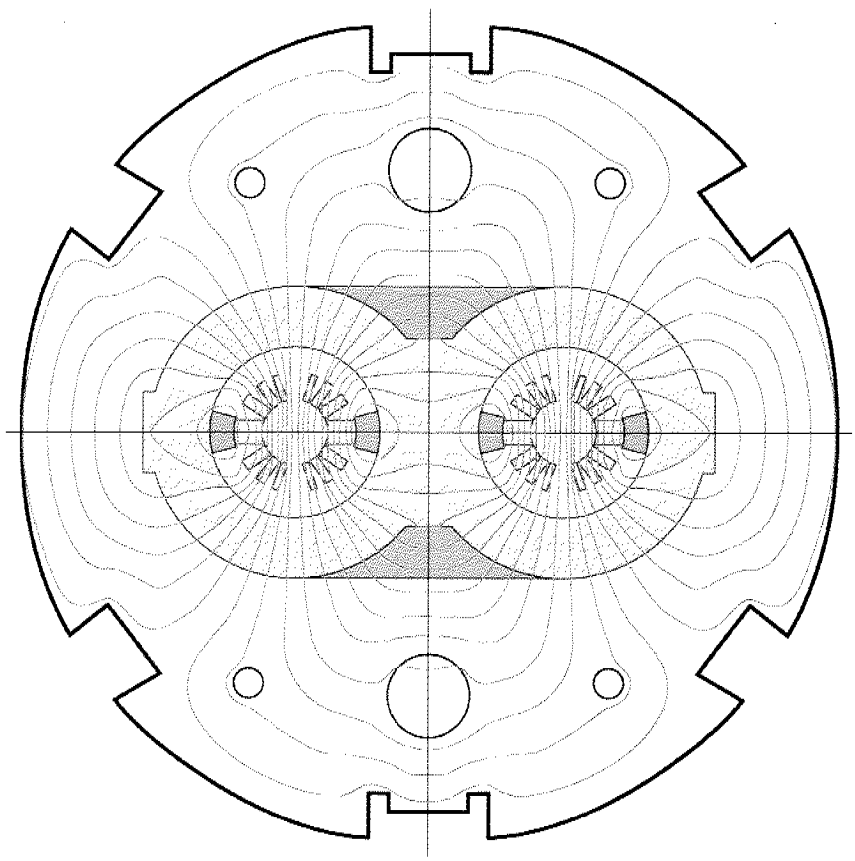

FIGURE 1.7: *A LHC dipole ('2-in-1' magnet) with its calculated magnetic field for a 10 T field [19].*

also a dedicated detector, which will study CP violation and other rare phenomena in the decay of the beauty quark.

1.2.2 The CMS detector

About 2000 scientists from more than 160 institutions and 36 countries are involved in the CMS experiment. A detailed description of this detector can be found in [21]. Figure 1.8 shows a schematic view of CMS.

Like CDF, CMS is a multi-purpose detector with a tracking system to detect charged particles, electromagnetic and hadronic calorimeters to measure the energy of electrons, photons and jets; and muon chambers to detect muons. In the following, a short detector description will be given with particular emphasize on the sub-detectors that will be used in this study, i.e. the electromagnetic calorimeter.

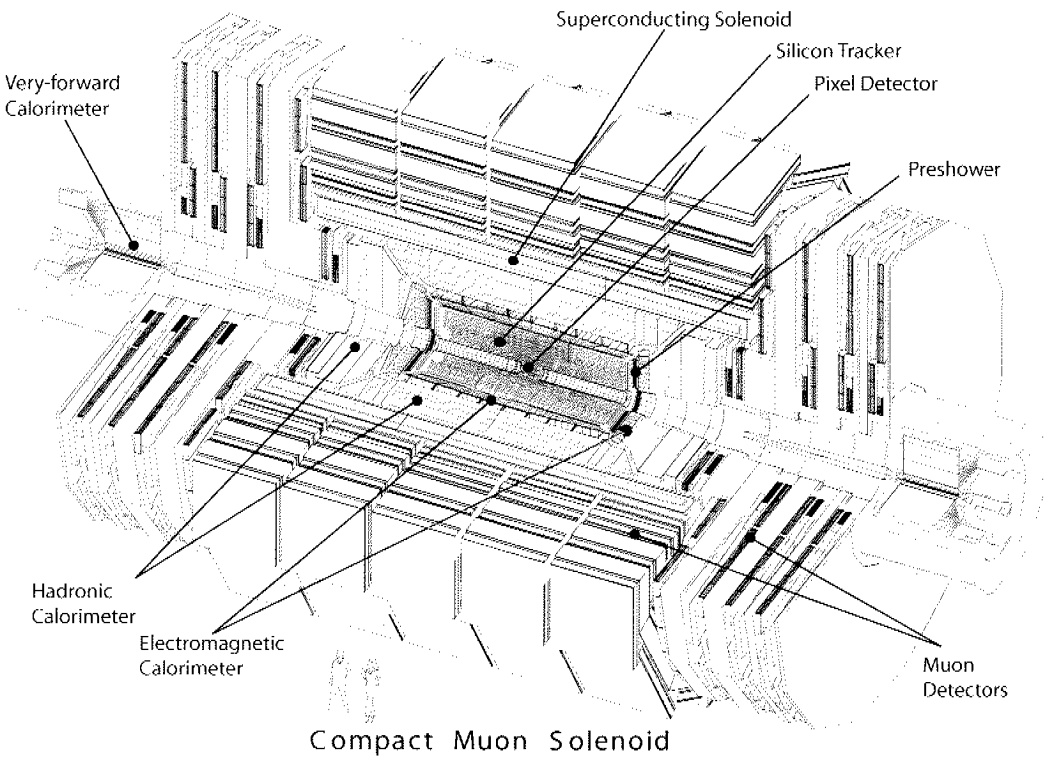

FIGURE 1.8: *Schematic view of the CMS detector at LHC [21].*

CMS stands for "Compact Muon Solenoid", but has nevertheless a diameter of 15 m and a length of 21.6 m. Despite its size, the design of CMS is driven by the idea of being as compact as possible without compromising its physics performances.

The "S" of CMS stands for solenoid. The magnet return yoke is indeed the main component of CMS in terms of weight, size and plays a structural role in supporting the other components of the detector. CMS will have a very high magnetic field together with a large magnet volume, as the coil will be placed after the calorimeters. The required magnetic field is about 4 T with a diameter of roughly 6 m and a length of 13 m. Thus, 2 GJ of energy will be stored in the magnet. Such a high magnetic field will be created by a very high current (20 kA) circulating in superconducting strands. The magnet will thus be cooled down to 4.2°K with liquid helium to make it superconducting[3]. The magnetic flux is returned through a 1.5 m thick saturated iron

[3]The LHC superconducting magnets have about the same amount of electrical current circulating but will

yoke, which is split into five barrel rings and two end-cap disks housing the muon chambers. Its solenoid is placed after the calorimeters, thus not degrading the energy resolution.

In the following a description of the different CMS sub-detectors is given.

The tracking system

Placed around the interaction point, CMS will have tracking chambers consisting in silicon pixels layers and silicon micro-strip layers. They will occupy a cylindrical volume with a length of about 5.4 m and a diameter of about 2.4 m. The large volume of the tracker together with the 4 T magnetic field allows a significant bending of the track and therefore an accurate momentum measurement of high energy charged particles.

Three layers of pixel detectors, placed right around the beam line will allow a precise vertex reconstruction and will provide also the first step in the track reconstruction. The expected pixel hit resolution is $\sigma_{r\phi} \sim 10\mu$m and $\sigma_{rz} \sim 17\mu$m. The silicon micro-strips placed after the pixel detectors will allow a precise track reconstruction. It consists of four inner barrel layers, two double-sided outer barrel layers and four single-sided outer barrel layers. The expected hit resolution for the silicon strip is $\sigma_{r\phi} = 10 - 60\mu$m and $\sigma_{rz} = 500\mu$m. Combining these numbers, the expected CMS tracking resolution ranges from $\delta p_t/p_t^2 = 0.015\%$ for $|\eta| < 1.6$ up to $\delta p_t/p_t^2 = 0.06\%$ for $|\eta| = 2.5$.

The calorimeter system

The electromagnetic calorimeter (ECAL) is installed after the tracking chambers. It is a crystal calorimeter, consisting in 75'848 lead-tungstate ($PbWO_4$) crystals. The barrel crystal dimension is roughly 2.2 cm × 2.2 cm × 23 cm. In the endcaps, the crystal dimension is 2.5 cm × 2.5 cm × 22 cm. The choice of lead tungstate was motivated by its fast light decay time (100 ns are enough to collect all the light emitted by the crystal, about 60% of the light should be emitted already after 15 ns against 300 ns for BGO[4] crystals) and its good radiation resistance, as the crystals have to endure very high radiation doses. Lead tungstate has also a short radiation length (the crystal length represents about 26 radiation lengths) and a small Molière radius, which means that the shower should be well contained laterally inside a few crystals. For 35 GeV electrons, about 50% of the electron energy is contained in one crystal and 80% in a 3 × 3 crystal array. A drawback of the lead tungstate is its relatively low light yield, which is about 14 times less than BGO crystals. To collect the scintillation light emitted by the crystals, two avalanches photo-diodes (APD) per crystal in the barrel region and one vacuum photo-triodes (VPD) per crystal in the end-cap region, will be glued at the end of each crystal[5].

The 76'000 crystals of the ECAL are assembled together in a modular structure. The design of the structure which holds the crystals together, is done in a way to minimize the cracks. This is the reason why the crystals are tilted in the transverse plane by 3 degrees, leading to a quite complicated geometry, as shown in Figure 1.9. In η, the crystals are tilted from 0 degrees up to 3 degrees in order to allow the electron trajectories to be in most cases parallel to the crystal axis. The barrel ECAL is divided into sub-modules, modules and super-modules, as Figure 1.10 illustrates: a sub-module is composed of 2×5 crystals assembled into a fiberglass alveolar structure. Forty or fifty sub-modules are then assembled into a module, consisting

need to be cooled down to an even lower temperature, of 1.9°K. This is because the magnetic field created by the LHC magnets is higher than the CMS one, requiring thus the critical temperature to be lower to keep the magnet superconducting.

[4]The crystals used e.g. in the L3 detector.

[5]The photodetectors have to operate in a rather hostile environment, in a strong magnetic field of 4 T and under unprecedented radiation levels. No single mass-produced photodetector exists that can handle both these challenges and therefore two types have to be used: avalanche photodiodes can operate in strong transverse magnetic fields and will be used in the barrel part of the calorimeter. In the endcaps, the vacuum phototriodes will be used in order to cope with the higher levels of radiation. [21]

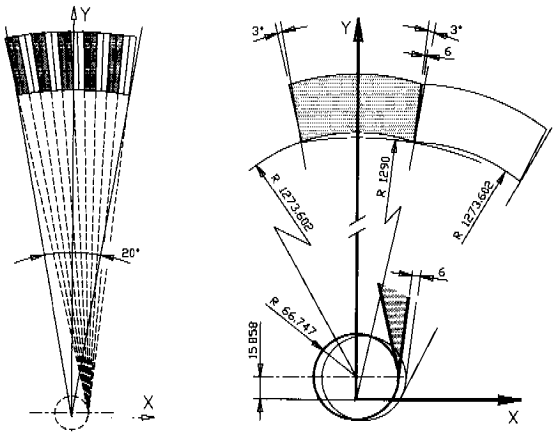

FIGURE 1.9: *Construction of the crystal ϕ-tilt. The plot on the right shows how the super-modules are put together in the transverse plane. [21]*

thus of 400-500 crystals. Finally, four modules are put together to form a super-module. The design guarantees a maximum distance between crystals faces of 0.4 mm within a sub-module and 0.6 mm across two sub-modules. A crack of about 6 mm is expected between two super-modules in ϕ and 6.8 and 7.8 mm between two modules at different η. There are in total 36 super-modules, subtending an angle of 20 degree. Figure 1.11 shows an artistic view of the ECAL.

The endcap ECAL is built up of identical 5×5 crystals. To ensure a hermetic design, the crystals will be oriented toward a point located 1.3 cm away from the interaction point. Thus the crystals are off-pointing to a similar extent as the barrel crystals.

A preshower build in the front of the end-cap calorimeter should allow to reject π^0 by measuring the transverse profile of the electromagnetic shower after roughly three interaction lengths. The preshower is built like a sampling calorimeter with lead as absorber and a layer of silicon strip sensors for the measurement of the charged particles created in the shower. The expected resolution is around 300 μm for a 50 GeV π^0, corresponding to an angle of 0.1 mrad with respect to the detector center. For comparison, the opening angle is about 2 mrad, for 50 GeV photons coming for a π^0 decay at the interaction point.

The hadronic calorimeter is placed right after the ECAL. It is a sampling calorimeter, made of copper absorber plates, interleaved with 4 mm thick plastic scintillator tiles. The produced blue scintillation light is captured and shifted toward green in wavelength shifting fibers and then transported to photo-diodes. Its purpose is, together with the electromagnetic calorimeter, to reconstruct jets and missing energy. A calorimeter granularity of $\Delta\eta \times \Delta\phi$ has been chosen so that highly boosted dijets from W and Z decays can still be distinguished. In order to get a good measurement of the missing energy, jets are expected to be reconstructed up to a rapidity of 5

The muon system

The muons chambers are located in the outermost part of CMS, inside the return yoke of the CMS magnet. The muon system will have basically three tasks: the muon identification, a redundant muon momentum measurement and the triggering of physics events. Muons, unlike the electrons, make essentially no Bremsstrahlung and are expected to give very clean signals. In the barrel, the muon system will consist of drift tubes and resistive plate chambers (RPC). In the endcaps, cathode strip chambers will be installed. Together with the information from the tracking chamber a resolution of about 1%−1.5% for 10 GeV muons and 6%−17% for 1 TeV

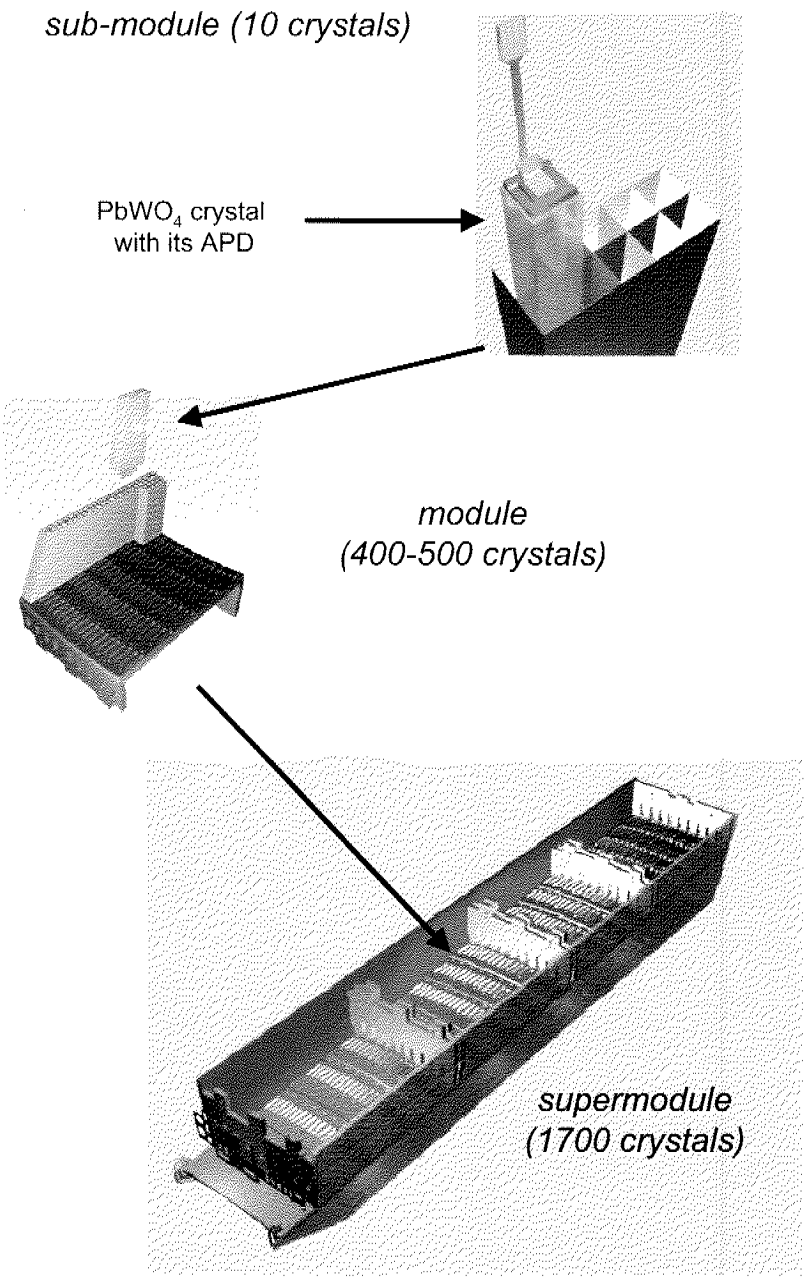

FIGURE 1.10: *Structure of the barrel ECAL, showing the different assembly steps starting from a sub-module (10 crystals), to form a module (400-500 crystals) and finally a super-module (1700 crystals).*

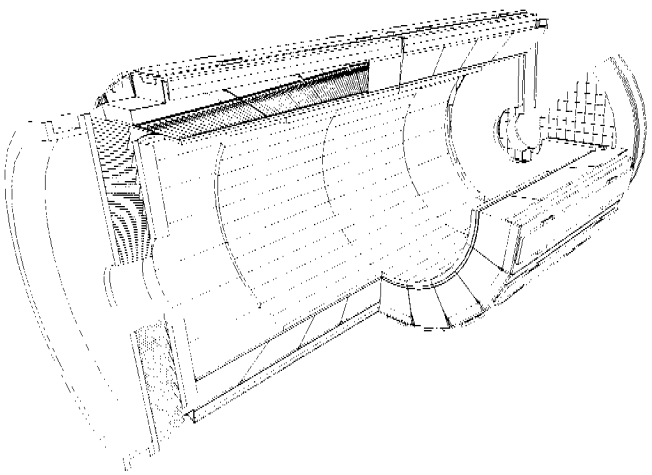

FIGURE 1.11: *An artistic view of the ECAL, showing its super-module structure and the crystal tilt in η [21].*

muons is predicted.

The trigger system

Another important aspect of the CMS detector is the trigger system. There will be a bunch crossing every 25 ns together with about 25 multiple interactions at $\mathcal{L} \sim \times 10^{34}$ cm^{-1} s^{-1}. The data flow has to be reduced from about 45 MHz down to about 100 Hz (output rate on disk).

CMS has chosen to reduce this rate in two steps: at the first level, all data are stored for 3.8 μs (the equivalent of 192 bunch crossings). A maximum event rate of 100 kHz is then forwarded to the High Level Trigger (HLT). The Level 1 trigger is a very fast hardware trigger, using coarse information from the calorimeters and the muon RPC's to determine the basic event properties in order to decide if the event should be discarded. The HLT is a software trigger and is based on about 1000 processors grouped in a so-called farm. The data from the detector front-end electronics are passed to the processor farm using a high bandwidth switching network. The data flow through the switch is about one Terabit per second. The functionality of the CMS HLT is three-fold. First to perform the readout of the front-end electronics after a Level 1 trigger accept. To execute physics selection algorithms on the events read-out, in order to accept the ones with the most interesting physics content. Finally to forward these events, as well as a small sample of rejected events, to the online services that monitor the performance of CMS. The accepted events are finally archived in mass storage, on tapes.

1.3 LHC versus Tevatron and a comparison between CDF and CMS.

In this section some differences between CDF (and the Tevatron accelerator) and CMS (and the LHC accelerator) will be discussed. A summary of typical design variables for the two experiments are shown in Table 1.2. It is important to keep in mind that CMS will start its first data-taking about 6 years after the first collision took place in the upgraded CDF detector. In the mean time, different technologies were developed allowing to achieve performances at CMS that were not possible 6 years earlier.

- *The cross sections*

	Tevatron	LHC
Center of mass energy	1.96 TeV	14 TeV
Particles colliding	protons-anti-protons $(p-\bar{p})$	protons-protons $(p-p)$
Design luminosity	$(1-2)\cdot 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$	$10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
Multiple interactions per crossing (at design luminosity)	~ 5	~ 25
	CDF	CMS
Detector magnetic field	1.5 T	4 T
Tracking resolution	$\delta p_t/p_t^2 = 0.1\%$, $\lvert\eta\rvert < 1.0$ $\delta p_t/p_t^2 \sim 0.4\%$ $1.0 < \lvert\eta\rvert < 2.0$	from $\delta p_t/p_t^2 = 0.015\%$, $\lvert\eta\rvert < 1.6$ up to $\delta p_t/p_t^2 = 0.06\%$, $\lvert\eta\rvert = 2.5$
Tracking chamber radius	1.4 m	1.3 m
Tracking chamber total length	3 m	6 m
Number of channels in the tracker	$40.5\cdot 10^4$	$11\cdot 10^6$
Silicone area	1.9 m^2	223 m^2
ECAL type	sampling: lead/scintillator	$\sim 76'000$ $PbWO_4$ crystals
ECAL resolution	$\frac{16\%}{\sqrt{E}} \oplus 1\%$	$\frac{2.9\%}{\sqrt{E}} \oplus \frac{0.13}{E} \oplus 0.4\%$ *
ECAL granularity (barrel) $(\Delta\phi \times \Delta\eta)$	0.26×0.1	0.0175×0.0175
Radiation lengths in front of the ECAL (for $\theta = 90°$)	$1.8\% X_0$ (tracker) + $1X_0$(solenoid)	0.57 X_0 (tracker)
Radiation lengths in the barrel ECAL	19-21 X_0	26 X_0
HCAL resolution	$\frac{80\%}{\sqrt{E}} \oplus 5\%$	$\frac{120\%}{\sqrt{E}} \oplus 5\%$
Barrel HCAL granularity $(\Delta\phi \times \Delta\eta)$	0.26×0.1	$\sim 0.09 \times 0.09$
Trigger reduction rate	2.5 MHz to 75Hz	45MHz to 100Hz

TABLE 1.2: *A comparison between CDF, Tevatron and CMS, LHC (for the resolutions, E and p_t are given in GeV). The numbers are taken from the technical design reports of the two experiments: [17] and [21].*

* *as given from the 2003 test beam results [22]*

Compared to the Tevatron, the LHC is going to collide protons with a factor seven higher energy, increasing the cross sections for almost all interesting physics processes, as Figure 1.12 illustrates. For instance, the cross section for single W and single Z production will increase roughly by a factor ten.

Moreover the LHC luminosity is expected to be a factor 10 to 100 times higher than the Tevatron.

- *Jets and backgrounds*

  Because of the parton momentum distribution in the proton and the different center of mass energies (1.96 TeV at the Tevatron and 14 TeV at the LHC), the type of partons interacting at an energy of $Q^2 = 10^4$ GeV, will be most probably quarks at the Tevatron and most probably gluons at the LHC. For instance, at the Tevatron 85% of the top antitop pairs produced will come from quarks and 15% from gluons, whereas, at the LHC, 10% of the top antitop pairs produced will come from quarks and 90% from gluons. The jet production at the LHC is also expected to be higher than at the Tevatron. As an example, the ratio (electron/jet) between the inclusive rate of electrons coming from single W and the inclusive rate of jets with $p_t > 20$ GeV, is 10^{-3} at the Tevatron and 10^{-5} at the LHC. To achieve comparable results at a fixed energy, the particle identification of the LHC detectors should thus be two orders of magnitude better than those of the Tevatron experiments [24].

- *Proton-proton versus proton-antiproton collisions*

  The Tevatron is colliding protons with anti-protons, whereas the LHC will collide protons with protons. For single W/Z production, the two quarks will be most of the time valence quarks at the Tevatron, whereas, for the LHC, one quark will have to come from the sea. Because of that, the W^+ rapidity distribution will be asymmetric in rapidity, whereas at the LHC it will be symmetric, as it will be discussed in detail in Chapter 2.

  One physics goal of the Tevatron is to use these proton-antiproton collisions to measure the forward-backward asymmetry of the Z resonance. At the LHC, due to the symmetric collisions, such a measurement is more problematic. Nevertheless, as it will be shown in Chapter 4, even here, a forward-backward asymmetry can be measured. This can be done if a cut on the rapidity of the system is applied.

- *The detector response and the tracking*

  A crossing takes place inside CDF every 396 ns. For CMS this time is decreased to 25 ns. This will require a much faster detector response for CMS.

  For instance, unlike CDF having a drift chamber, the CMS tracking volume will consist only in silicon detectors. Drift chambers, like the CDF one, have a dead time of about 100 ns, which is too slow for the LHC planned crossing time of 25 ns. For comparison, silicon detectors have a dead time of about 15 ns, which is limited by readout electronics.

- *The track reconstruction*

  The momentum, p, of a particle with charge ze, moving in a magnetic field is given by the following formula:

  $$p\cos(\lambda) = 0.3zBR$$

  where B is the magnetic field in Tesla, R the radius of curvature in meters and λ the angle between the track momentum and the magnetic fixed direction. For a fixed momentum and a uniform magnetic field, the momentum resolution depends mainly on the projected length of the track onto the bending plane, L', the measurement error at each point, ε,

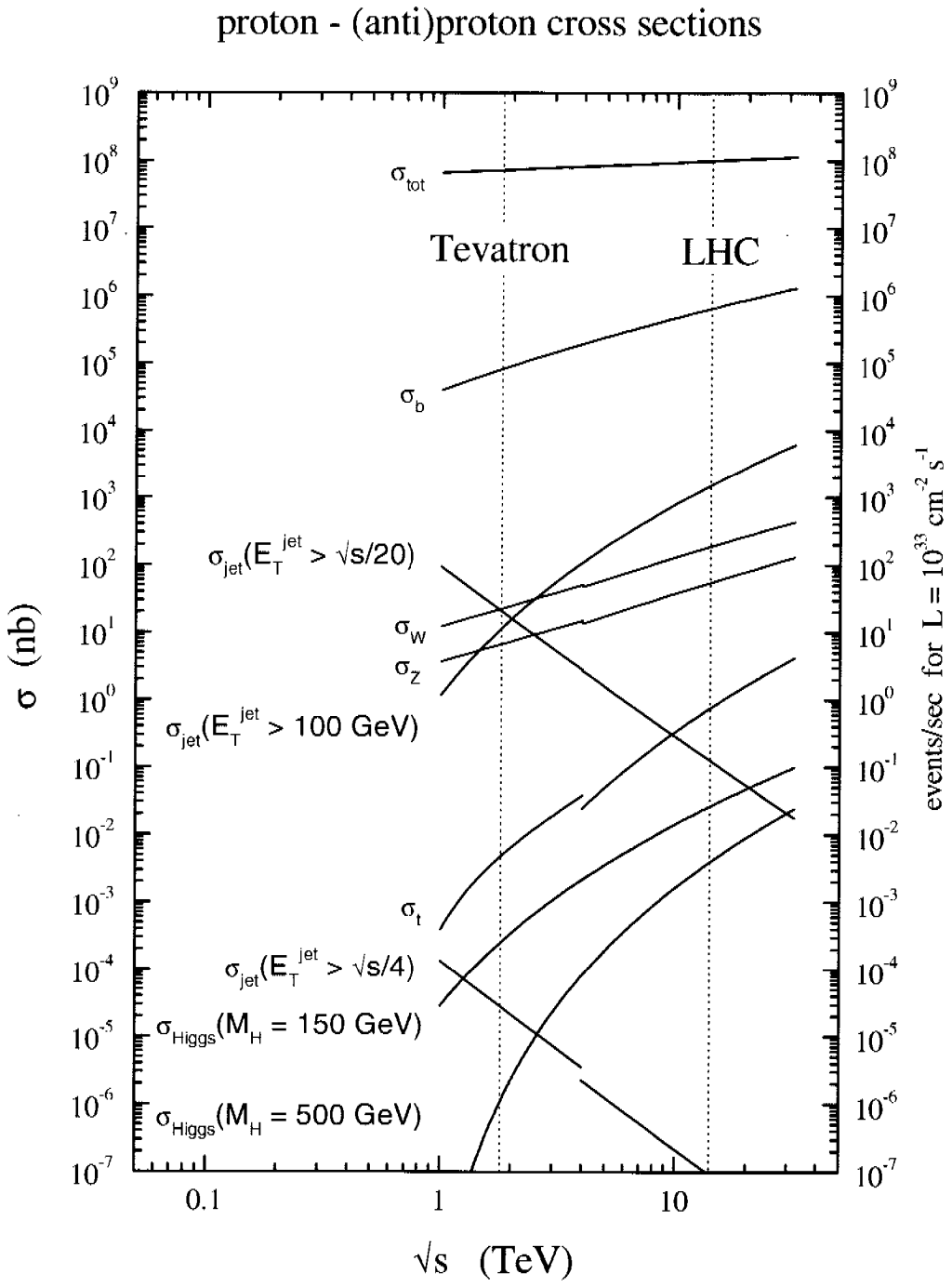

FIGURE 1.12: *Calculated cross sections for hard scattering versus the center of mass energy, $\sqrt{s}$, all cross sections except σ_{tot} are calculated using the latest MRST PDF [23]. The discontinuities in the curves arise from the difference between proton-anti-proton (shown for $\sqrt{s}$ lower than 4 TeV) and proton-proton cross sections (shown for $\sqrt{s}$ higher than 4 TeV). $\sigma_{jet}(E_T^{jet} > \sqrt{s}/20)$ stands for the jet production cross section requiring jets to have a minimum energy scaled with the center of mass energy of the collision.*

and the number of points measured along the track, N. If the number of points exceeds 10, a good approximation of the error on the momentum resolution is [25]:

$$\frac{\delta p_t}{p_t}(res) = \frac{\varepsilon}{B \cdot L'^2}\sqrt{\frac{702}{N+4}}$$

The resolution is also affected by multiple scattering effects, which can be approximated with:

$$\frac{\delta p_t}{p_t}(ms) \approx \frac{0.05}{B\sqrt{L' X_0}}$$

where X_0, the radiation length, is for instance 9.4 cm for silicon and 11000 cm for argon (used in the CDF COT). From this formula, one sees that the errors are very sensitive to L'. To improve the resolution, everything should be done (staying within reasonable costs) to maximize the length of the track. CMS and CDF have almost the same chamber outer radius but as CMS has a higher magnetic field, it should have a better track resolution. For a 10 GeV charged particle in the barrel, CDF expects a track momentum resolution of 1% while in CMS the resolution is 0.15%.

CMS has a tracker made only out of silicon (its silicon surface is about a factor 100 higher than the one of CDF), which will require a strong cooling and an important support. Thus, there will be a lot of material before the calorimeter, as illustrated on Figure 1.13. A consequence of this is that the track reconstruction for electrons will be more problematic in CMS, due to the electron interacting with the tracker material and loosing energy through Bremsstrahlung. The number of interaction lengths for electrons in the tracking volume for central rapidities goes indeed from $1.8\% X_0$ for CDF to $53\% X_0$ for CMS. The effect of the Bremsstrahlung and its influence on the electron reconstruction will be discussed in more detail in Chapter 3.

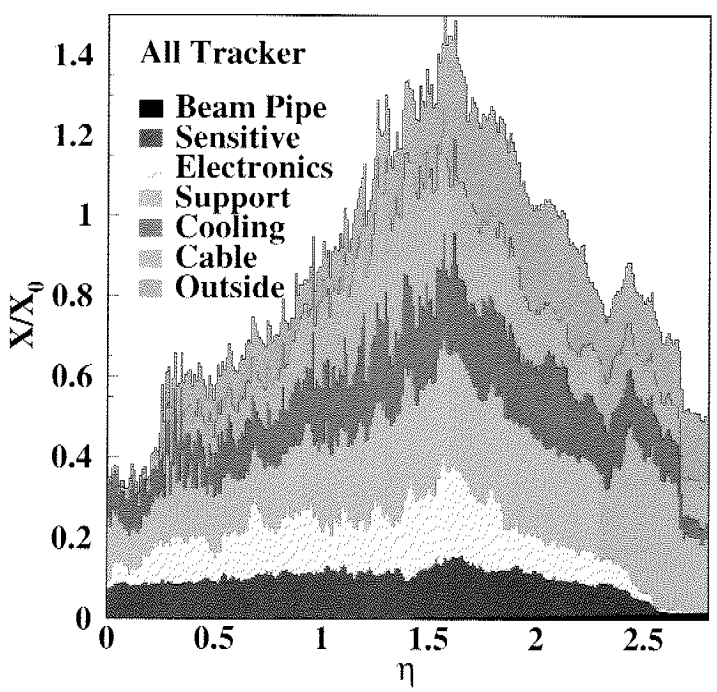

FIGURE 1.13: *The tracker material budget as a function of the pseudorapidity. [26]*

- *The calorimeter*

  CDF has a sampling calorimeter while CMS has a crystal calorimeter. The advantage of a crystal calorimeter over a sampling calorimeter is a better energy resolution, as the crystals are at the same time absorbers and scintillators. With a sampling calorimeter, an optimization of the thickness of sampling and absorbing layers has to be done in order to maximize the resolution. For 50 GeV electrons, CDF has a resolution of 2.5% (from test beam data) against 0.65% (from test beam data [27]) for CMS.

  The longitudinal shower shape can be better estimated with a sampling calorimeter, as the shower development can be measured along the different scintillator layers. This allows to discriminate between the showers created by electrons and the ones generated by pions. In a crystal calorimeter, only the sum of emitted photons will be measured. However, the fine granularity of the CMS ECAL alone should allow an excellent electron-pion discrimination by measuring the lateral size of the shower. Also the presence of the preshower in front of the calorimeter endcaps should provide a good pion rejection.

- *The trigger*

  CMS has to reduce its data from 45 MHz down to 100 Hz. Taking also into account the multiple interactions (about 25 for CMS and 5 for CDF), this represents roughly a factor 100 more incoming data than for CDF for an almost similar outgoing rate! The architecture of the trigger is different in the two experiments as CDF has three trigger levels while CMS will have a hardware Level 1 and a 'high level trigger', based on CPU farms.

  It is a common feature to hadron machines to have high trigger requirements, as opposed to lepton colliders where massive trigger reduction is not really needed. In proton collisions the total interaction cross section is about 100 mb, while the cross section of electroweak processes are 6 order of magnitude smaller. In lepton collisions – the lepton do not interact through the strong interaction – the total interaction cross section is much smaller. For instance, the $e^+e^- \rightarrow Z$ cross section is about 40 nb. LEP I was running with a mean luminosity of about $10^{31} cm^{-2} s^{-1}$ resulting in an event rate of about 0.4 Hz !

Since both CMS and CDF are multi-purpose detectors, they possess a very similar general structure, consisting of a tracking chamber, calorimeters and muons chambers. Compared to CDF, CMS is expected to have a much better energy and momentum resolution. This will allow to fulfill the broad CMS physics program, given the fact that the event reconstruction and analyze at the LHC is expected to be much harder, with an enhanced jet production and a very high collision rate, requiring a fast detector response.

Seite Leer /
Blank leaf

Chapter 2

Towards a precise luminosity determination at hadron colliders

2.1 The parton luminosity method

An important type of precision measurements at hadron colliders is the cross section determination. A good knowledge of different processes cross sections allows to test the predictions of a given model. Cross sections are also good observables to study the properties of 'known' processes that are background for new phenomena searches.

At hadron colliders, one of the limitations on the accuracy of cross section measurement is the uncertainty on the luminosity determination (expected to be less than 5%) and on the Parton Distribution Functions (PDF) (around 5% depending on the parton type). In the following, we will describe and give a first application of a method that should allow to reduce these uncertainties down to 1%, allowing a more precise cross section determination.

Essentially all measurable cross sections ($\sigma_{pp(p\bar{p})} \rightarrow X$) and the corresponding signal event rates $S_X = N - B$ at a hadron collider are related with the $pp(p\bar{p})$ luminosity and the cut efficiency (ε):

$$S_X = N - B = \sigma_X \times \mathcal{L}_{pp(p\bar{p})} \times \varepsilon \tag{2.1}$$

However, most "interesting" processes do not involve directly the protons and anti-protons but their basic constituents, the partons. Theoretical cross sections are calculated assuming two partons interacting and can be expressed for the corresponding proton-(anti)proton interaction using the parton distribution functions.

The parton distribution functions

The parton distribution functions (PDF) provide the link between the colliding protons and their interacting constituents, parameterizing the quark and gluon structure of the proton. The cross section for a given process $pp \rightarrow X$ for two hadrons with momenta P_1, P_2 can be written as [28]:

$$\sigma_X(P_1, P_2) = \sum_{a,b} \int_0^1 dx_1 dx_2 f_a(x_1, \mu^2) f_b(x_2, \mu^2) \times \hat{\sigma}_{ab \rightarrow X}\left(p_1, p_2, \alpha_S(\mu^2), Q^2/\mu^2\right) \tag{2.2}$$

where the sum runs on the involved partons a and b, x_1 and x_2 are the parton momentum fractions ($x_i = p_{\text{parton } i}/P_{\text{hadron } i}$) integrated over the whole kinematic range, f_a and f_b the corresponding PDF, defined at a factorization scale (μ^2) and α_s is the strong coupling constant.

It has been shown that the total cross section ($pp \rightarrow X$) can be divided into a high momentum and low momentum part (for a review, see [29]). The high momentum cross section ($\hat{\sigma}$) is

calculated considering the interaction of two partons and is characterized by an energy scale Q^2 (which could be for instance the mass of a weak boson). The factorization scale, which is an arbitrary parameter, can be thought as the scale which separates the high and low momentum processes. Since the strong coupling constant is small at high energy, the high momentum process can be calculated using perturbation QCD. The low momentum part is factorized into the PDF. This can be done since the high momentum cross section is insensitive to the physics of low-momentum. In particular, $\hat{\sigma}$, involving only interacting partons, is independent of the type of the incoming hadron. Moreover, the PDF are common to all processes involving the same type of hadrons.

Since the PDF include the low momentum part of the interaction, they are not a perturbatively calculable quantity and must be extracted from experimental data. However the measurable part of the PDF should be independent of the factorization scale, μ. Requiring this and setting $\mu^2 = Q^2$, one gets the DGLAP equations [30]:

$$\begin{aligned} \frac{\partial q_i(x,Q^2)}{\partial \log(Q^2)} &= \frac{\alpha_S}{2\pi}\int_x^1 \frac{dy}{y}\{P_{q_iq_j}(y,\alpha_S)q_j(\frac{x}{y},Q^2) + P_{q_ig}(y,\alpha_S)g(\frac{x}{y},Q^2)\} \\ \frac{\partial g(x,Q^2)}{\partial \log(Q^2)} &= \frac{\alpha_S}{2\pi}\int_x^1 \frac{dy}{y}\{P_{gq_j}(y,\alpha_S)q_j(\frac{x}{y},Q^2) + P_{gg}(y,\alpha_S)g(\frac{x}{y},Q^2)\} \end{aligned} \tag{2.3}$$

where the $P_{q_iq_j}$ are the so-called splitting functions giving the probability that a parton of type q_i is evolving in a parton of type q_j. The splitting functions have been calculated up to three loops (NNLO[1]) [31].

The Q^2 dependence of the PDF can be calculated using the DGLAP equations, but the PDF are also dependent on x. This dependence is parameterized at some low scale and the DGLAP equations are used to specify the distributions at the higher scales where data exist. The PDF parameters can then be determined for instance from deep inelastic scattering data, taking into account that the PDF are process-independent. Figure 2.1 illustrates the range in x and Q^2 which is reached at HERA (electron-proton deep inelastic scattering) and the range which will be probed with the LHC.

Todays knowledge on the PDF comes from different measurements: from experiments based on deep inelastic scattering (like SLAC [33], BCDMS [34], NMC [35], E665 [36], H1 [37] or ZEUS [38]), on the Drell-Yan processes (like E605 [39] or E866 [40]) or measurements of high E_t jets (CDF [41], DØ [42]), W rapidity asymmetry (CDF [43]), and νN dimuon (CCFR, NuTeV [44]).

Different groups used these data to extract the PDF (like MRST [45], CTEQ [46] or Alekhin [47]). There is some freedom in the way to fit these data, which leads to variations in the different PDF sets obtained. This depends on the different data put in the fit (the selection of the data, the treatment of experimental systematic errors) and the difference in the theoretical treatment (like the choice of the fit tolerance: CTEQ allowed $\Delta\chi^2 = 100$ and Alekhin $\Delta\chi^2 = 1$, the factorization and renormalization scheme and scale, the PDF parametric form, the treatment of heavy flavors, etc.).

This leads to uncertainties in the PDF. For instance, the predicted value of the NLO W cross section at the LHC is 204 ± 4 nb with the MRST2002 set, 205 ± 8 nb with the CTEQ6 set and 215 ± 6 nb with the Alekhin02 set [32].

The parton luminosities: constraining the PDF at Hadron Colliders

A way to reduce the uncertainties from the PDF and on the luminosity is to measure directly the different *parton luminosities* using the single W and Z production identified through their leptonic decays. The *parton luminosities*, $\mathcal{L}_{\text{parton}_1,\text{parton}_2}(Q^2, x_1, x_2)$, predict the frequency with

[1]The perturbation order at which a process is calculated is usually expressed in the following way: LO means leading order, NLO, next-to-leading order and NNLO, next-to-next-to-leading order.

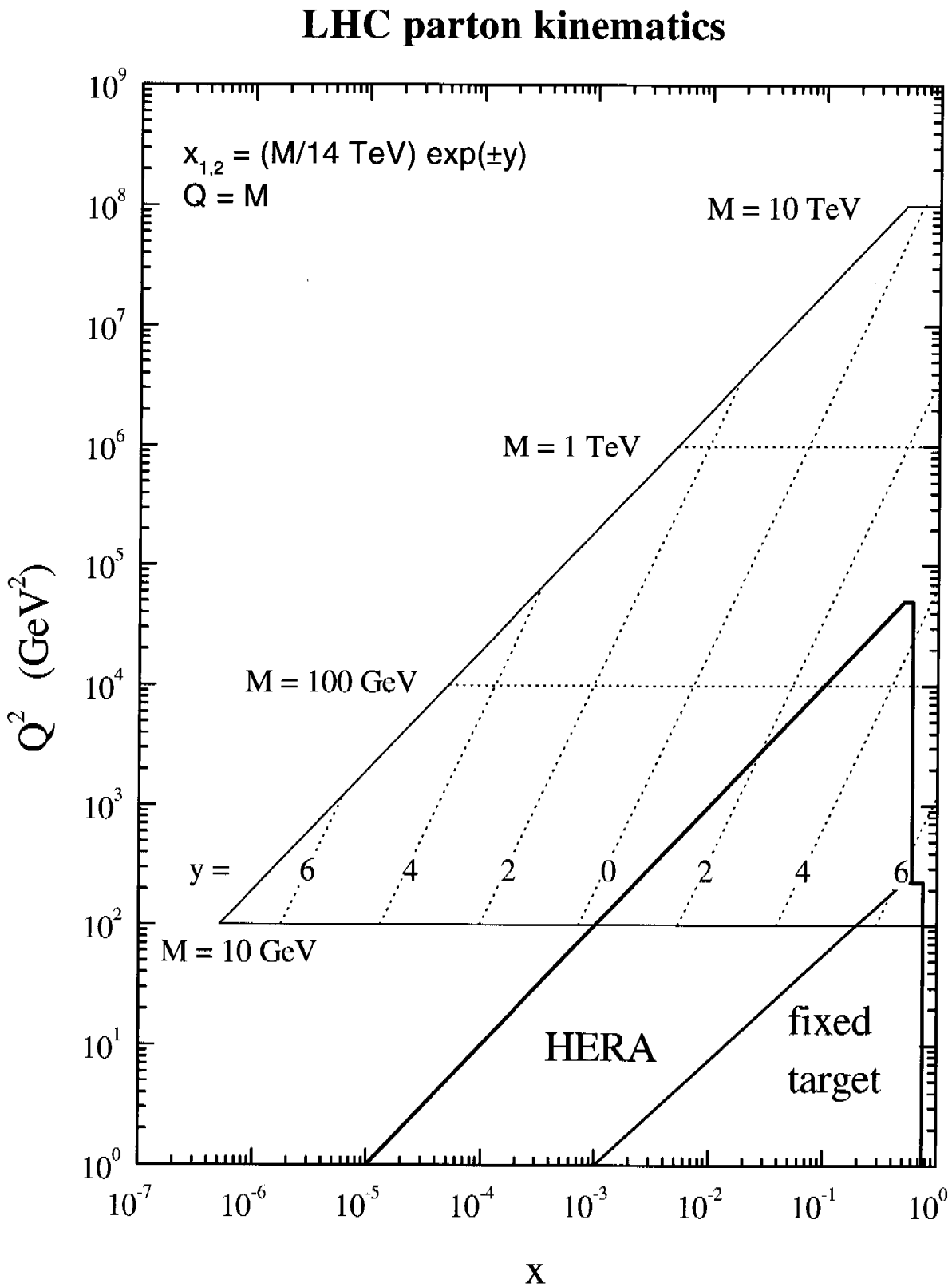

FIGURE 2.1: *The range in x and Q^2 which is reached at HERA and will be probed with the LHC [32].*

which two partons of given types and given momenta will interact with each other at a definite energy scale. They depend obviously on the following physics parameters: the parton types (q,$\bar{q}$ or g), the parton momentum fractions (x_1 and x_2) and also the energy scale of the process (Q^2). The number of signal events can then be expressed in the following way:

$$S = N - B = \hat{\sigma}_{\mathrm{parton}_1,\mathrm{parton}_2 \to X} \times \mathcal{L}_{\mathrm{parton}_1,\mathrm{parton}_2}(x_1, x_2, Q^2) \times \varepsilon \tag{2.4}$$

Taking into account that theoretical estimates for high Q^2 processes are based on the interactions of quarks and gluons, parton luminosities appears to be a natural quantity [8]. While this approach to the luminosity question was described in detail for the LHC, it applies as well to any hadron collider. The basic ideas are the following:

- The $u(\bar{u})$ and $d(\bar{d})$ parton distribution functions can be studied using the following processes: $u\bar{u}(d\bar{d}) \to Z \to \ell\ell$, $u\bar{d} \to W^+ \to \ell^+\nu$ and $d\bar{u} \to W^- \to \ell^-\bar{\nu}$, taking into account that the W and Z branching fraction to lepton have been measured with high accuracy at LEP. The quark/antiquark – gluon interaction could be studied using the $qg \to \gamma, Z, W$+jet processes.

  The heavy quark flavor component has to be also taken into account. About 8% (10%) of the W/Z produced at the Tevatron (LHC) come from heavy quarks. Figure 2.2 shows the parton decomposition of the cross section for single W and single Z production.

- In the case of single W/Z production, the product of the quarks momenta fraction is directly constrained as this process occurs at a determined mass:

$$\hat{s} = s \cdot x_1 \cdot x_2 \tag{2.5}$$

  where $\sqrt{\hat{s}} = M_W, M_Z$ is the mass of the vector boson and s the center of mass energy (for LHC: 14 TeV and for Tevatron: 1.96 TeV).

- The crucial observable is the W and Z rapidity (Y). The rapidity distribution depends directly on the ratio between x_1 and x_2:

$$Y = 1/2 \ \ln(x_1/x_2) \tag{2.6}$$

  Since the Q^2 dependence can be extrapolated to a given scale, $\hat{s}$, the parton luminosity depends then on the vector boson rapidity distribution.

  For e.g., single W^+: $\mathcal{L}_{u_1,\bar{d}_2}(Q^2, x_1, x_2) \equiv \mathcal{L}_{u_1,\bar{d}_2}(Y_W)$ [2]. Combining the relations (2.5) and (2.6) one finds that the parton x_1, x_2 range for central W production corresponds to $x_1 \approx 0.002$ and $x_2 \approx 0.016$ for Y(W)=1 and $x_1 \approx 0.042$ and $x_2 \approx 0.0008$ for Y(W)=2 for a center of mass energy of 14 TeV (at the LHC). The sensitivity of the rapidity distribution to the PDF is illustrated in Figure 2.3 which shows the resulting W rapidity distribution for two choices of PDF sets: the MRST "default" and the so-called MRST "conservative". The PDF at NLO are sensitive to the value of x_{min}, which is the minimal x value down for which the data are included in the PDF fit. The conservative PDF set requires $x_{min} > 0.005$, while no cut is applied for the default set. The comparison of the rapidities predicted by the two PDF sets give a rough indication of the theoretical uncertainty due to the parton distributions. A drop of almost 20% in the W cross section calculated with the two PDF sets is observed, but the central part of the rapidity distribution remains basically unchanged. The PDF are expected to be less sensitive to the cut on x_{min} once NNLO DGLAP is used.

[2] Note that in hadronic collisions it is not possible to reconstruct the W rapidity distribution. However one can use the pseudorapidity ($\eta = -\ln(\tan\theta/2)$) distribution of the lepton coming from the W decay.

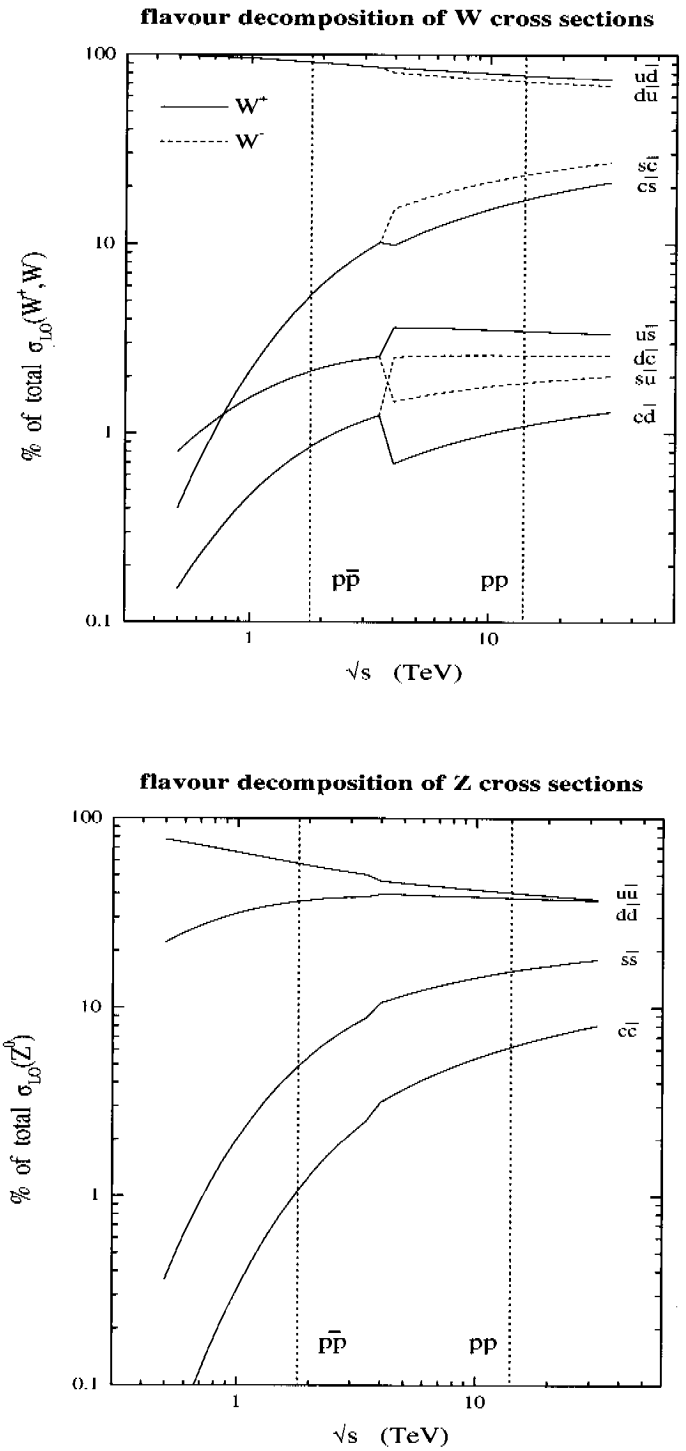

FIGURE 2.2: *Parton decomposition of the W^+ (solid line, top), W^- (dashed line, top) and Z (bottom) total cross section in $p\bar{p}$ and pp collisions. Individual contributions are shown as a percentage of the total leading-order cross section. In $p\bar{p}$ collisions, the decomposition is the same for W^+ and W^-. The two vertical dashed lines shows the values corresponding to the Tevatron (1.96 TeV) and the LHC (14 TeV). The discontinuities in the curves arise from the differences between proton-anti-proton and proton-proton interactions [48]*

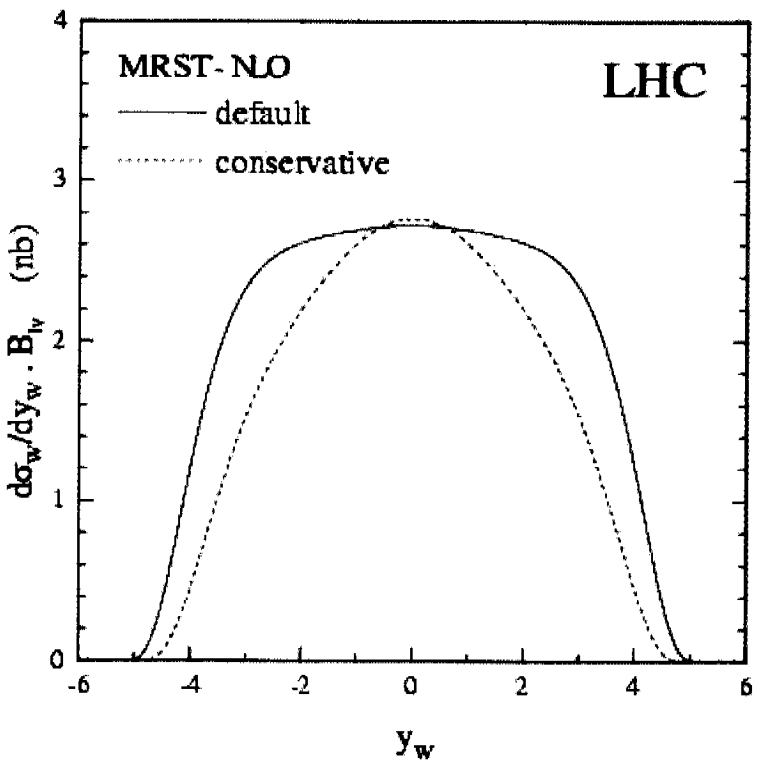

FIGURE 2.3: *The prediction for the rapidity distribution of the W cross section at the LHC for two different PDF sets. [49]*

- The single W and Z processes provide thus a tool to measure parton-parton (quark and antiquark) luminosities, which can be used as an input for the theoretical predictions of related physics processes. Interesting are the pair production of vector bosons and the Drell-Yan dileptons which depend on $q\bar{q}$ luminosities at different values of Q^2. Using the W, Z luminosity normalization procedure, the above examples might eventually be calculable and measurable with statistical and systematic uncertainties reaching $\pm 1\%$ [8].

- In addition, systematic uncertainties for processes which are identified with isolated leptons will become much smaller since the results are normalized to the production of W and Z decays. It is thus interesting to develop a general selection that can be also applied to different processes. By taking ratios, it will allow to reduce the systematic errors.

It is not the first time that single W and single Z production are used as "standard candle" processes. Rapidity distributions, cross sections and lepton asymmetries of Drell–Yan lepton pairs have been used since many years as a tool to constrain the PDF. The parton luminosity goes however one step further allowing to determine at the same time the luminosity and the PDF ! Moreover it is well known that the accuracy of many measurements and their theoretical interpretations can be considerably improved if appropriate ratios can be measured. Many errors, including the luminosity uncertainty, are irrelevant if one measures the ratio between the inclusive W and Z cross section.

The full parton luminosity method is adapted to the LHC where the very high W/Z rate will allow to reduce drastically statistical uncertainties. However, already with the existing CDF data, the first part of the method can be studied: the signal reconstruction and the counting of the total number of W and Z events. Knowing the cross section for such processes and the detector efficiency, one can get the proton-antiproton luminosity and compare it with the results from the traditional CDF luminosity measurement[3].

[3]To obtain the parton luminosity, a measurement of the W and Z rapidity distributions is needed. However given the statistics of CDF data, such a measurement was not accurate enough to be really interesting.

2.2 W and Z counting using one year and a half of CDF data

2.2.1 Single W and single Z production at the Tevatron

The theoretical calculation for single W/Z production is known to a good accuracy. The next-to-next-to-leading order (NNLO) cross sections for single W/Z production at $\sqrt{s}$ =1.96 TeV, in the narrow width approximation and a "given" Parton Distribution Function, have theoretical uncertainties of less than 3% [49] :

$$\begin{aligned} \sigma_{NNLO}(p\bar{p} \rightarrow W \rightarrow e\nu) &= 2.687 \pm 0.054 \text{ (PDF)} \pm 0.054 \text{ (hard process) nb} \\ \sigma_{NNLO}(p\bar{p} \rightarrow Z \rightarrow ee) &= 0.251 \pm 0.005 \text{ (PDF)} \pm 0.005 \text{ (hard process) nb} \\ \frac{\sigma_{NNLO}(p\bar{p} \rightarrow Z \rightarrow ee)}{\sigma_{NNLO}(p\bar{p} \rightarrow W \rightarrow e\nu)} &= 0.094 \pm 0.001 \text{ (PDF)} \end{aligned} \quad (2.7)$$

The first uncertainty, of about 2%, reflects the error on the PDF determination. As explained in [49], the error on the PDF can be split into a theoretical and an experimental contribution. The experimental contribution comprises the systematical and statistical errors of the data used in the global fit performed to determine the PDF. The theoretical contribution includes mainly the uncertainties coming from the corrections to the standard DGLAP equations. Other sources of theoretical errors like the choice of input parametrization, heavy target corrections necessary to fit to neutrino data, assumptions about the strange quark sea and isospin violation can be neglected.

The second uncertainty reflects the error on the cross section of the hard process involving the partons, $\hat{\sigma}_{q_1,q_2 \rightarrow W/Z}$. As the main issue consists in knowing whether the perturbation serie converges or not, the uncertainty on $\hat{\sigma}_{q_1,q_2 \rightarrow W/Z}$ is estimated to be 2% which is half of the variation between the cross section at NLO and NNLO.

It is interesting to note that if the cross section changes by about 30% when going from LO to NLO and NNLO, the ratio of the Z to W cross sections remains essentially constant (within 1%). One can expect that, for similar processes, many theoretical errors cancel and a more accurate prediction for their ratio is possible. The theoretical error on the W to Z ratio can then be estimated to 1% [49].

Combining the calculated cross section, the number of signal events and the selection efficiency, one can measure the luminosity. The W and Z will be selected using their decays into electrons. First, the samples used for this study are presented and a description of the event selection together with the method to count the number of Z and W is explained. Then the systematics sources are discussed, considering particularly the detector stability over time. The luminosity is then measured using three different types of samples: Z events where the two electrons are in the central calorimeter, Z events where one electron is in the central calorimeter and the other is in the plug calorimeter[4] and W events with a central electron. Finally some distributions for the W and Z bosons are studied and differences between PYTHIA and HERWIG generators are discussed.

2.2.2 Data and Monte Carlo Event samples

The data collected by the CDF detector from February 2002 until May 2003 (run number 138425 to run number 163527) were used for this analysis. The runs are required to satisfy minimal quality requirement [50]. This corresponds to an integrated luminosity of 125.5 pb^{-1}, as estimated from the CLC luminosity monitor[5]. The data were processed using version 4.8.4 of the CDF reconstruction program. Specific root trees [51], the so-called Standard ntuples, were written and analyzed [52]. The data were divided in 7 subsets as shown in Table 2.1.

The following trigger conditions were analyzed: ELECTRON_CENTRAL_18, used for the selection of W and Z events, and W_NOTRACK, used to estimate the trigger efficiencies.

[4]In CDF the forward and backward parts "closing" the detector are called the "plugs" whereas in CMS they

Dataset	Runs number	CLC luminosity (good runs)
1	138425 - 147866	11.6 pb^{-1}
2	147869 - 152616	19.2 pb^{-1}
3	152630 - 155116	24.4 pb^{-1}
4	155121 - 156487	18.9 pb^{-1}
5	159603 - 161409	18.4 pb^{-1}
6	161410 - 162631	18.5 pb^{-1}
7	162663 - 163527	14.5 pb^{-1}

TABLE 2.1: *CLC luminosity and run numbers for the different datasets*

The ELECTRON_CENTRAL_18 trigger requires a central electromagnetic cluster with $E_t > 18$ GeV and the ratio of the energy in the hadronic calorimeter to the energy in the electromagnetic one to be lower than 0.125. A track from the central outer tracker with $p_t > 9$ GeV must also match the electromagnetic cluster. The W_NOTRACK trigger requires a central electromagnetic cluster with $E_t > 25$ GeV and missing energy higher than 25 GeV. Since the trigger efficiency to get a cluster in the electromagnetic calorimeter is about 1 for high energy electrons, the two triggers are assumed to be independent and can be combined to measure the trigger efficiency. During the studied period, the requirements for an event to pass the Level 1 trigger for ELECTRON_CENTRAL_18 changed two times, leading to small variations in the trigger efficiency.

In order to have a more workable data sets, "obvious" dijet QCD backgrounds were removed. This preselection was chosen such that it allowed to keep essentially all detectable $Z \rightarrow ee$ decays and essentially all $W \rightarrow e\nu$ decays which did not show additional jet activity. This reduction was achieved by demanding that each event contained at least one reconstructed track, at least one electromagnetic cluster with a transverse energy of at least 20 GeV and satisfying the so-called *preselection cuts*: requiring the cluster to be isolated[6] with $\frac{\text{Iso}}{\text{Iso}+E_t} < 0.25$ and a small ratio of the electromagnetic over hadronic energy: $\frac{E_{had}}{E_{em}} < 0.1$.

Then, an event was kept if:

- It contains at least two clusters fulfilling the preselection cuts (for $Z \rightarrow ee$ events).
- It has exactly one electromagnetic cluster fulfilling the preselection cuts and no additional jet with a minimum p_t(jet) of 10 GeV (for $W \rightarrow e\nu$ events). Jets are reconstructed using a clustering algorithm with a cone size of 0.4.
- Harder requirements were then put on events with one electromagnetic cluster fulfilling the preselection cuts and one or more jets, to reduce the QCD dijets background. Such events were kept for further analysis if either the angle ϕ in the plane transverse to the beam between the highest p_t jet and the electromagnetic cluster is smaller than 160^0, or if the electron candidate fulfills the following stricter criteria:
  - the electromagnetic cluster is isolated ($\frac{\text{Iso}}{\text{Iso}+E_t} < 0.05$)
  - its energy fraction in the electromagnetic calorimeter is large ($\frac{E_{had}}{E_{em}} < 0.025$)
  - it has a good E/p ratio ($0.6 < E/p < 1.6$)

are called "endcaps".

[5]To measure the luminosity, CDF uses Čerenkov counters placed close to the beam line, as explained on page 16. This sub-detector is called the CLC.

[6]The variable Iso represents the sum of the transverse energy found in a cone with a $\Delta R = (\Delta\eta^2+\Delta\phi^2)^{1/2} = 0.4$ around the electromagnetic cluster

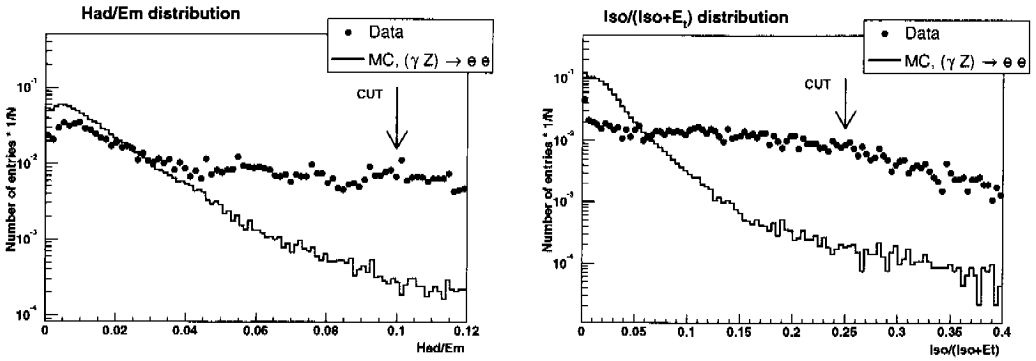

FIGURE 2.4: *The distributions for the isolation and electromagnetic energy fraction of electron candidates for a small sample of events which passed the ELECTRON_CENTRAL_18 trigger and for simulated electrons in the Z Monte Carlo sample. Only a very small fraction of signal events is removed by the preselection cuts used for the filtering.*

From the Monte Carlo simulation for the signal, one finds that only about 20% of all W events with jets have an angle ϕ between the jet and the electron larger than 160^0. About 2/3 of these events are kept with the applied electron criteria. Taking into account that additionally to that filtering, only the good runs were written to file, the applied criteria reduced the data by roughly a factor of 10 and 394,756 events were kept.

The Monte Carlo event samples used to simulate the signal and the background are described in Table 2.2. They were generated using the version 4.9.1 of the CDF reconstruction software. Signals simulated with PYTHIA were used for the analysis, but HERWIG signal events were also studied in order to compare the two generator outputs.

MC set	Generator	Number of events
$W \rightarrow e\nu X$	PYTHIA	200,000
	HERWIG	100,000
$\gamma, Z \rightarrow e^+e^-X$	PYTHIA	100,000
	HERWIG	50,000
$W \rightarrow \tau\nu X$	PYTHIA	200,000
QCD dijet	PYTHIA	$2.7 \cdot 10^6$

TABLE 2.2: *Monte Carlo sets used for the analysis. The sample of QCD dijets was already processed in standard ntuples, see [53].*

2.2.3 Selection of $W \rightarrow e\nu$ and $Z \rightarrow ee$ events

The obvious goal for the selection criteria of W and Z events is to count signal events above a small background, keeping the unvertainties on the cut efficiency small. In contrast to the $Z \rightarrow ee$ event selection, with a narrow peak in the two electron mass distribution, the signal for on-shell $W \rightarrow e\nu$ events has to be extracted from a broad peak, either in the transverse mass distribution of the electron neutrino system or from the electron p_t spectrum. In addition, the presence of jets increases potential backgrounds, influences the missing transverse energy and also the transverse mass distribution.

The Z events do not only provide a clean signal, but can also be used to calibrate the energy scale of the electromagnetic calorimeter and are a useful tool to measure electron selection efficiencies with respect to the Monte Carlo prediction.

The following electron selection strategy for Z and W events was used:

- All potential W, Z candidate events had to be triggered by the central electron trigger (ELECTRON_CENTRAL_18). This trigger efficiency could be measured from the number of signal events which were accepted by trigger conditions independent from the central outer tracker (W_NOTRACK).

- Only a few electron identification criteria were required. These are based on the ratio of electromagnetic over hadronic energy deposit, the electron isolation and the ratio between the tracking and the calorimeter response (E/p).

- Z signals are obtained from combinations of different electron reconstruction qualities in order to estimate potential efficiency differences between the data and the Monte Carlo simulation.

- W events are selected using events with one electron having a good reconstruction quality. No jets with a $p_t > 10$ GeV should be in the event. The model dependence of this jet veto can be controlled using the data and the Z signal in events with or without additional jets and applying a correction with respect to the expected ratio using the Monte Carlo simulation. Potential differences between the jet definition in the data and the W Monte Carlo simulation can be controlled using the Z data. The similarity of their production mechanism allows to determine a correction factor with respect to the Monte Carlo for the fraction of events without jets. More details about this correction will be discussed in section 2.2.7, on page 59.

Selection of good electron candidates

As explained before, the electron identification combines three requirements: the electron has to be isolated, it has to deposit almost all its energy in the electromagnetic calorimeter and the ratio between the energy deposited in the ECAL and the track momentum has to be close to one, since the track and the cluster come from the same particle. To account for the quality of the reconstruction, different exclusive classes of electrons are defined, based on combinations of the following requirements:

1. $\frac{\text{Iso}}{\text{Iso}+E_t} < 0.05$
2. $\frac{E_{had}}{E_{em}} < 0.025$
3. $0.6 < \frac{E_{cluster}}{p_{track}} < 1.6$

In detail, "gold" electrons satisfy all three conditions, "silver" electrons have to fulfill the third condition and either the first or the second one and "bronze" electrons satisfy only one of the three conditions.

In addition, the potential electron candidates should fulfill the following requirements of the kinematic of the event: a transverse energy of at least 25 GeV, a z vertex position of $|z| < 60$ cm and a maximum pseudorapidity $|\eta| < 3.0$.

As can be seen in Figure 2.5, data and Monte Carlo are not in perfect agreement and resolutions are somewhat worse in the data. For instance as shown on Figure 2.6, for the Z sample with two central electrons, the sample combining a bronze and a golden electron shows an excess in the data with respect to the Monte Carlo, which is compensated by a deficit in the sample combining two silver electrons. The differences between the data and the Monte Carlo seen here should not result in large discrepancies as the cuts are chosen such that most of the signal events are still accepted.

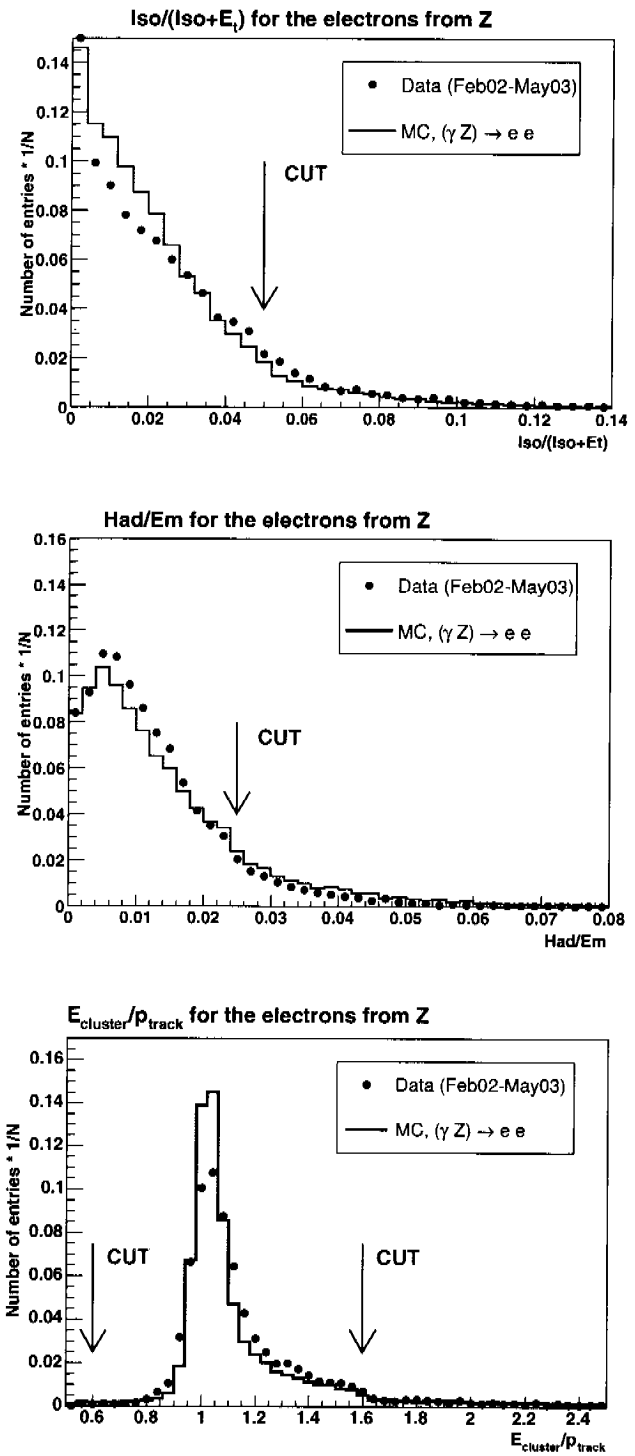

FIGURE 2.5: *Distributions for good electron candidates for events with at least two electron candidates in the data and in the $Z \rightarrow ee$ MC. All other cuts, besides the one for the variable shown in the plot, are applied.*

Selection of the $Z \rightarrow ee$ sample

As the electrons can be either in the central or in the plug calorimeter, we will divide the Z sample in two datasets: one where the two electrons are central (Z central–central) and one where one electron is in the central calorimeter and one is in the plug (Z central–plug). Z bosons are identified using the following classification: for the central–central Z selection, electron pairs were accepted if at least one electron is identified as a golden electron while the second one has to be at least a bronze electron. For the central–plug electron pairs, clean Z signals are obtained if the central electron candidate is identified as a golden or silver electron while for the plug electron, we demand that it fulfills at least the conditions of a bronze electron. Figures 2.6 and 2.7 show the resulting invariant mass distributions for the different combinations of electron reconstruction qualities.

Combinations involving bronze electrons show an excess of events in the data for an invariant mass between 60 and 70 GeV for the central–central Z sample. This is probably due to a background contamination, as the cuts for the bronze electrons are rather loose and a smaller excess can also be seen in the bronze-gold combination. The main background is dijet events, where the jets are reconstructed as electron. This process is characterized by a steeply falling jet transverse energy spectrum, with a cutoff around 50 GeV originating from the cut put on the electron transverse energy, which is what the plots show.

For the selected candidate events, the mass and width (σ) are obtained from a Gaussian fit. The energy of the central electrons in the data as well as in the Monte Carlo was already well calibrated. The energy of the plug electrons in the data needed an additional calibration which was done using the Z peak. The energy of the electrons going in the east plug (positive rapidity) has to be increased by 8% and for the electrons going in the west plug (negative rapidity) the energy has to be increased by 6% so that the position of the Z peak matches in the data and the Monte Carlo.

However, the Z peak in the Monte Carlo is narrower than in the data. For the central (resp. plug) electrons, an additional random Gaussian smearing, with a sigma of 2.7% (resp. 5%) of the electron energy is thus applied to the Monte Carlo so that the width of the simulated Z peak matches the one in the data.

The fitted mass and σ (obtained from the Gaussian fit) after the energy corrections are given in Table 2.3 for the data and for the Monte Carlo.

	central–central		central–plug	
	Mass [GeV]	σ [GeV]	Mass [GeV]	σ [GeV]
Data	91.1	4.6	91.3	4.9
Monte Carlo	91.2	4.4	91.2	4.7

TABLE 2.3: *Results from a Gaussian fit to the Z peak. The corresponding numbers for the Monte Carlo, including the additional energy smearing are also given. The statistical errors for the mass and the σ from the fit to the data are about 80 MeV.*

Figure 2.8 shows the resulting mass distribution for the data and the Monte Carlo and for central central and central–plug electron pairs after the calibration described above is applied.

Selection of the $W \rightarrow e\nu$ sample

Events with exactly one golden electron candidate found in the central calorimeter and missing transverse energy larger than 25 GeV are kept as W candidate events. For this analysis, W signal events are counted only in the case of zero jet events[7] and if the electron transverse mass fulfills 60 GeV $< M_T <$ 90 GeV. The transverse mass distribution is shown in Figure 2.9.

[7]Jets are defined using a cone algorithm of 0.4 and requiring $p_t^{\text{jet}} > 10$ GeV.

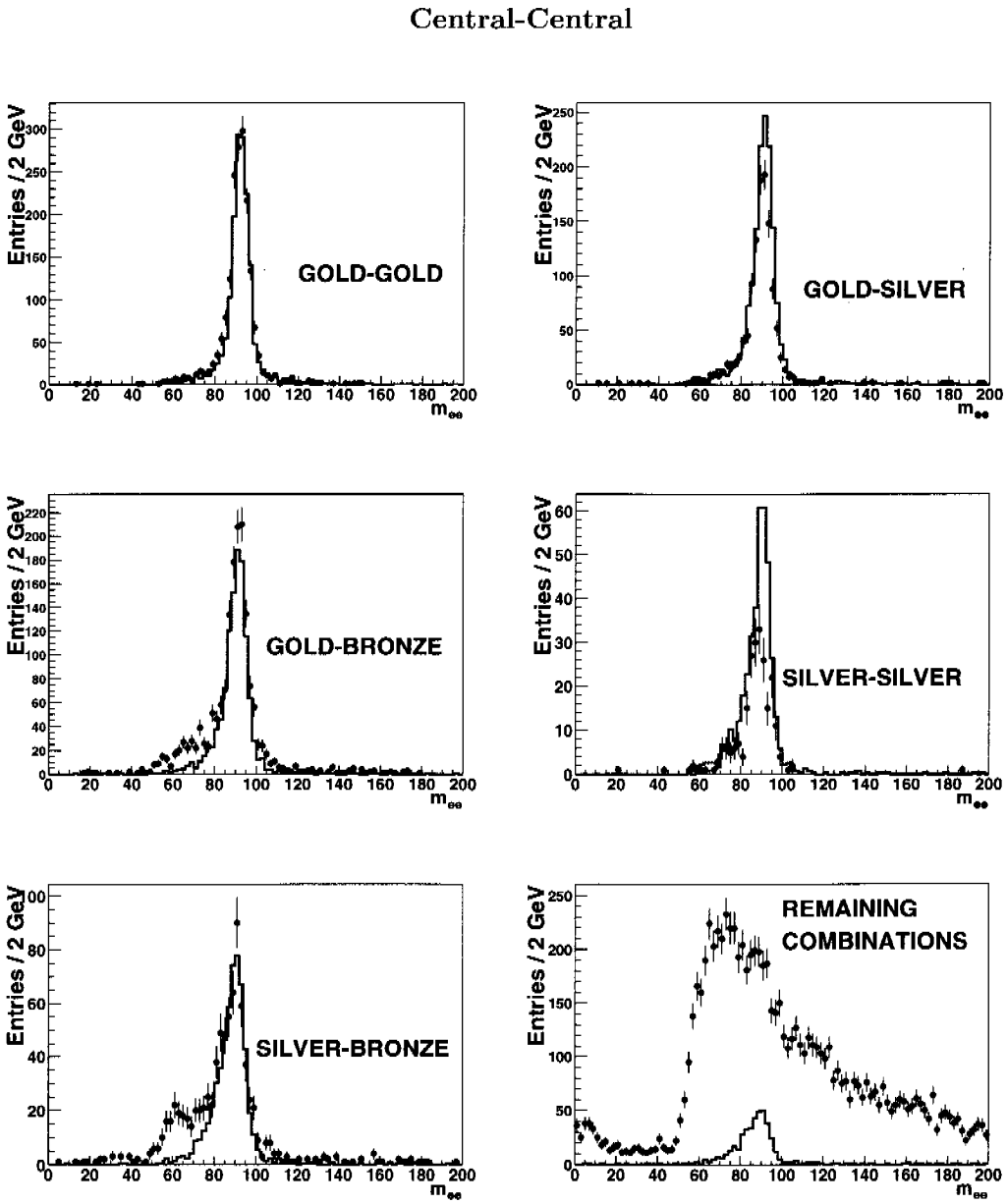

FIGURE 2.6: *Mass distribution for central–central electron pairs in the data ($\mathcal{L}$ =125.5 pb^{-1}) and the different quality combination of electrons. The mass distributions are shown for gold–gold, gold–silver, gold–bronze, silver–silver, silver–bronze and the remaining combinations of electron pair candidates. The first three combinations are used to define the Z signal and to determine the normalization factor between data and Monte Carlo which is then applied to all the other combinations.*

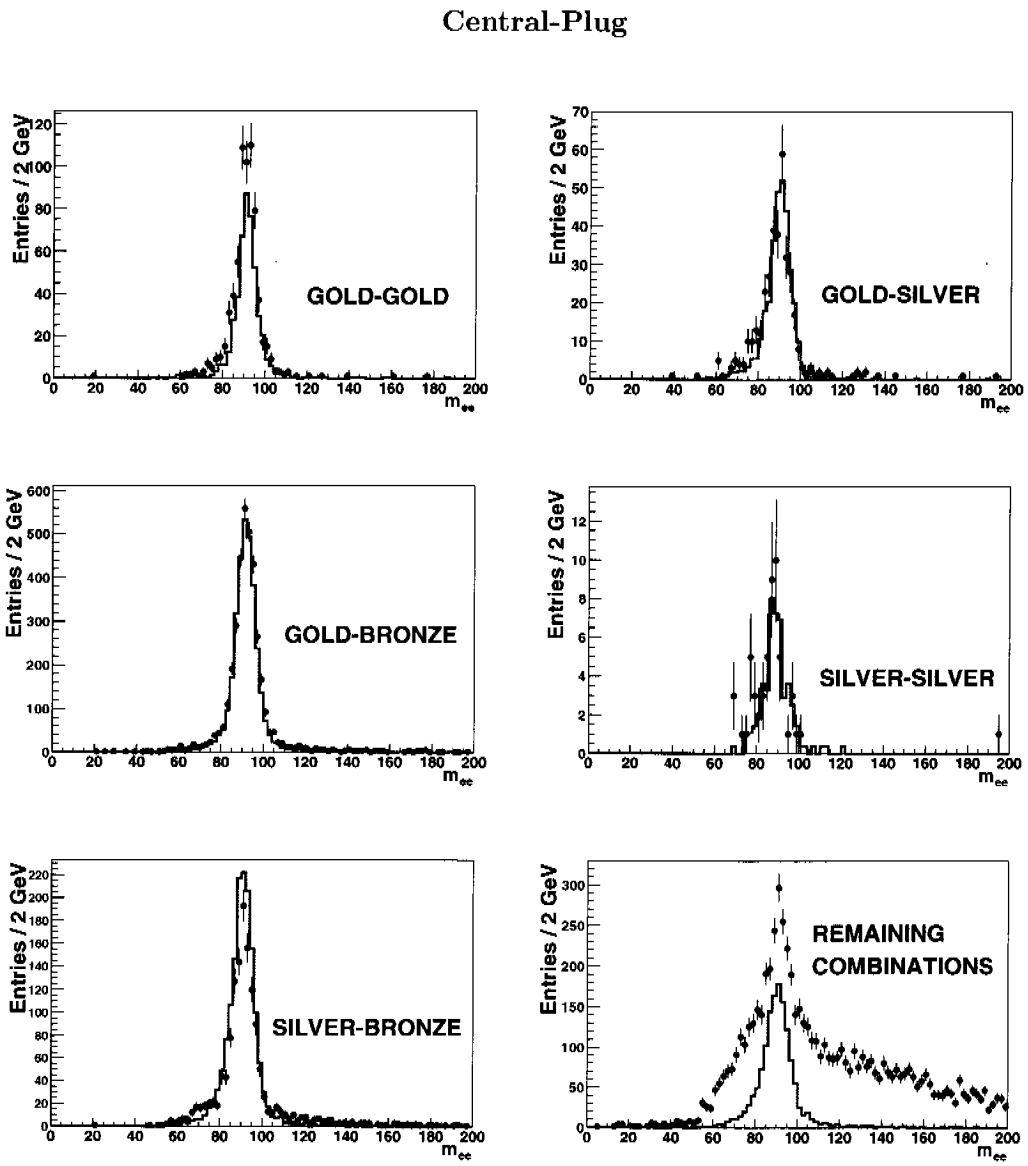

FIGURE 2.7: *Mass distribution for central-plug electron pairs in the data ($\mathcal{L}$ =125.5 pb^{-1}) and the different quality combination of electrons. The mass distributions are shown for gold-gold, gold-silver, gold-bronze, silver-silver, silver-bronze and the remaining possible combinations of electron pair candidates. The Z signal is defined as a combination between a central golden or silver electron with a bronze, silver or golden plug electron. The normalization factor between data and Monte Carlo is determined with the sum of those combinations and is then applied to all the other combinations.*

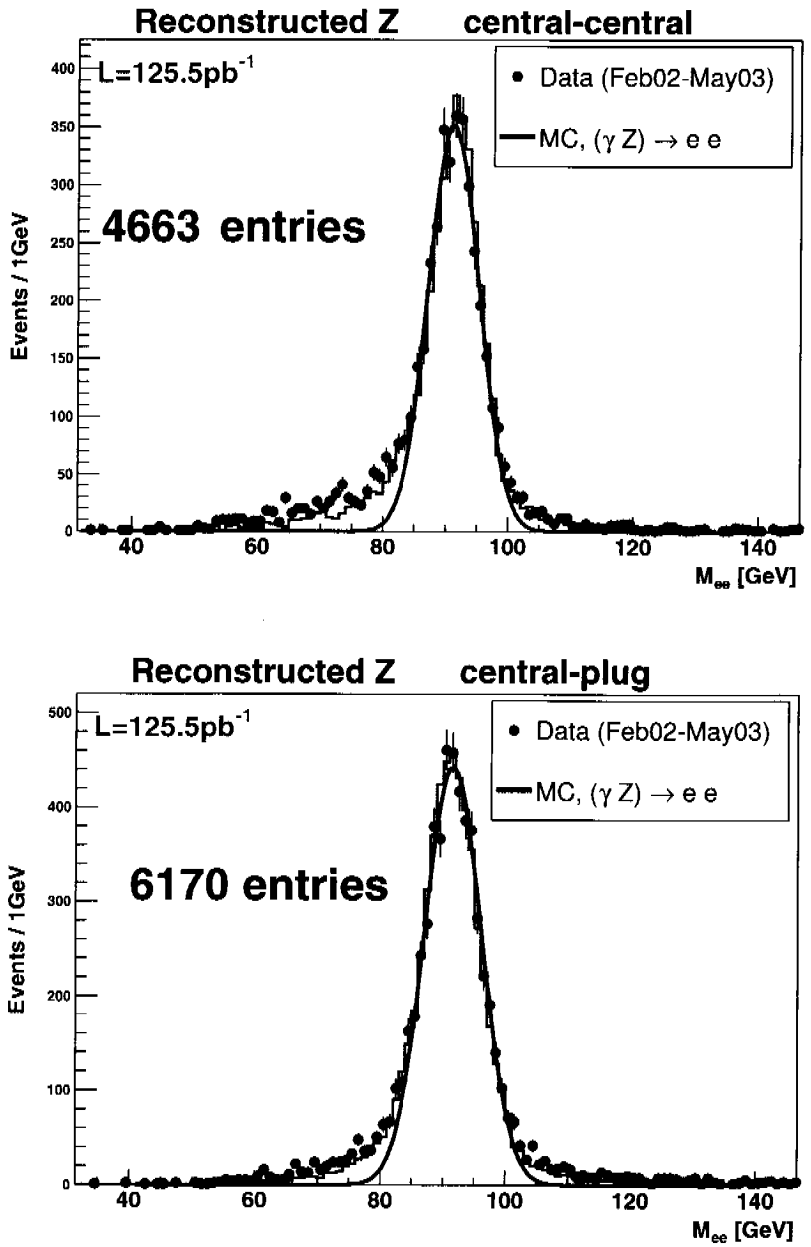

FIGURE 2.8: *Reconstructed Z signal for data (dots with error bars) and Monte Carlo (histogram). The curve is the result of a Gaussian fit to the data. The Monte Carlo is normalized to the number of events found within $\pm 2\sigma$ around the fitted Z peak.*

After the electron energy calibration, which was determined with the Z sample, a shift of about 250-500 MeV between data and Monte Carlo remains. The uncertainty arising from this miss-calibration is found to give a 1% systematic uncertainty on the final luminosity.

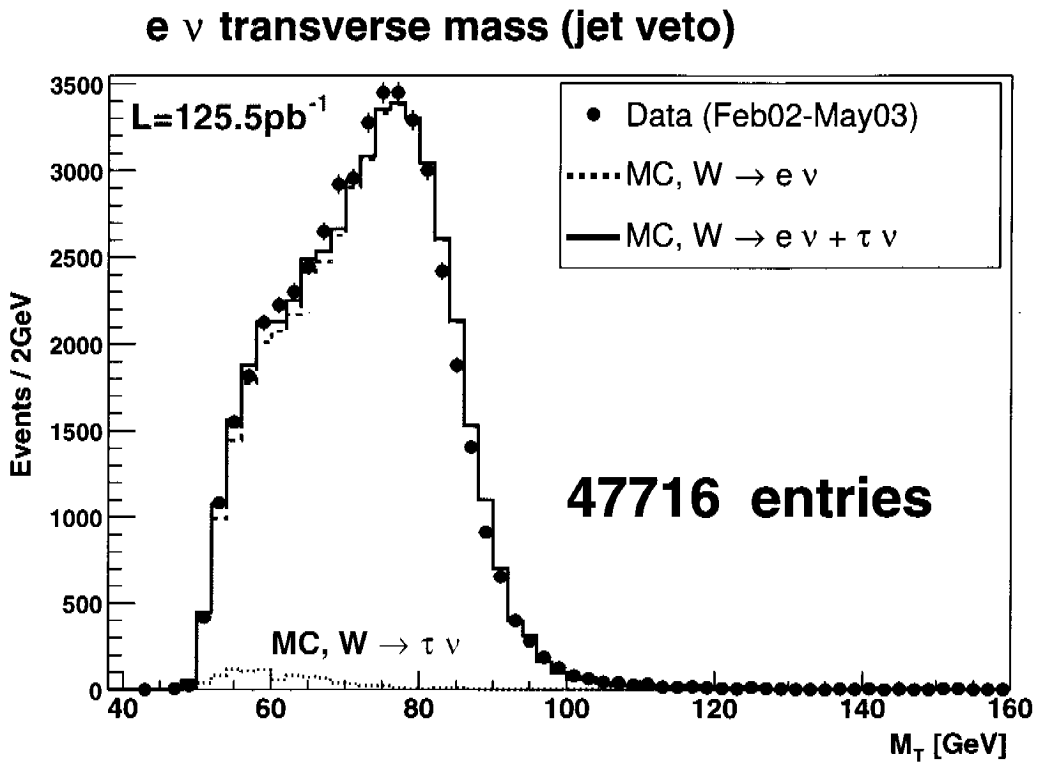

FIGURE 2.9: *Reconstructed W transverse mass distribution for events with golden electrons and events without jets, in the data (dots with error bars), in the $W \rightarrow e\nu$ Monte Carlo (dashed red line) and in the $W \rightarrow e\nu$ Monte Carlo together with the simulated $W \rightarrow \tau\nu$ background process (solid blue line). The pink dotted line shows the $W \rightarrow \tau\nu$ alone*

As determined from the Z sample, the electron energy in the Monte Carlo was smeared using a random Gaussian function, having a σ of 2.7% of the electron energy.

2.2.4 The effective luminous z–vertex region

In contrast to the luminosity measurement with the CLC, which "sees" interactions essentially over the entire collision region along the beam line (±100 cm), most measurements are limited to a smaller z–region. Traditionally, it is required that the event vertex for high p_t physics should be found with a z–vertex coordinate, z_{vertex}, ±60 cm around the center of CDF. The efficiency loss due to this requirement has thus to be known[8].

The fraction of events, which will not be accepted due to the z_{vertex} vertex position condition, has been determined from the data using reasonably well reconstructed tracks from the central outer tracker. However the efficiencies, especially for the trigger and the tracking, are smaller for large z_{vertex} positions. To minimize that problem, only events that are "boosted" into the detector acceptance are used to measure the luminous region as those events have a better acceptance. Such events are selected requiring the p_z of the decay particles and the measured z_{vertex} position to have opposite signs. Their z_{vertex} distribution is shown in Figure 2.10. The

[8]No additional correction needs to be applied if the rates for other studied high p_t processes are normalized to the W and Z counting method as inefficiencies from long tails in the vertex distributions will be identical for the studied reactions.

resulting "Gaussian" distribution are 31.4 cm and 27.2 cm in the data and the Monte Carlo respectively. Note also that the mean z-vertex position is not in $z_{vertex} = 0$ but in $z_{vertex} = 2.1 \pm 0.1$ cm in the data and 1.8 ± 0.2 cm (resp. 2.2 ± 0.1 cm) in the W (resp. Z) Monte Carlo. This shift has however little effect on the results since the z-vertex distribution has a large spread.

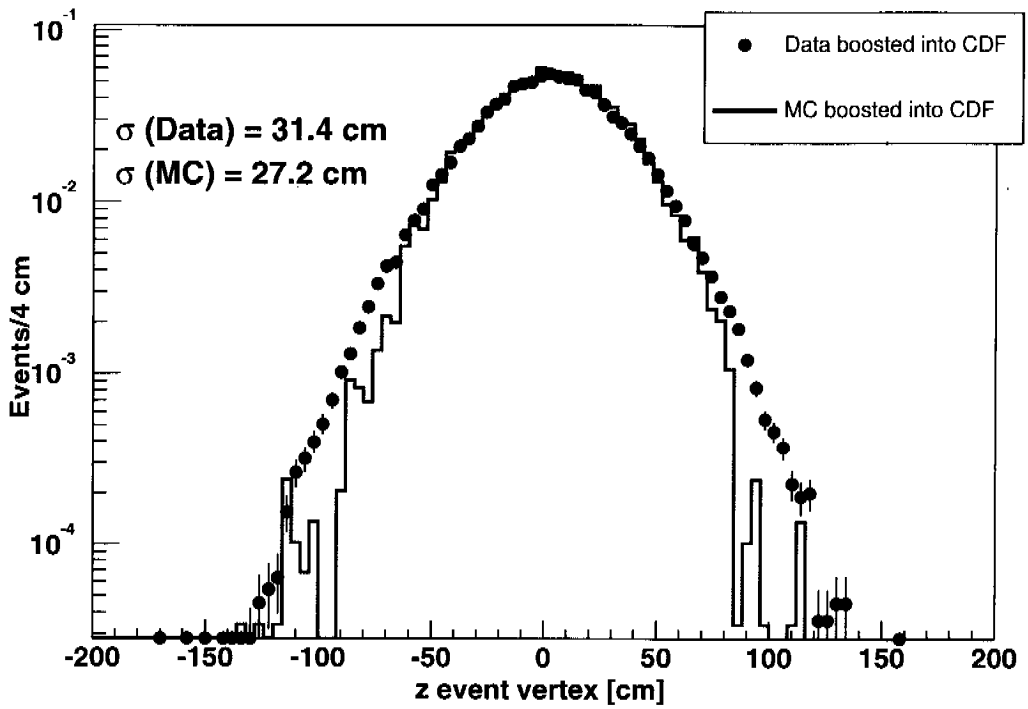

FIGURE 2.10: *Distribution of the reconstructed z-vertex in the data and in the Monte Carlo, requiring that the events are boosted inside the detector (i.e. the sign of the p_z component and the sign of the z-vertex have to be opposite).*

Using events which are boosted inside the detector, one finds that 2.9% of the W and Z Monte Carlo events have a vertex larger than $\pm$ 60 cm, compared to 5.3% in the data. Given the fact that the CLC counters are counting essentially the complete luminous region an additional efficiency correction of 0.98 has to be applied on top of the actual W and Z Monte Carlo vertex inefficiency.

2.2.5 Counting resonance W and Z decays

While theoretical cross section estimates are given for on-shell "resonance" production of W and Z bosons, using the narrow width approximation, the data include also off-shell production and backgrounds. In comparing data with theory it is thus important that efficiency and background corrections match as closely as possible the one used for the NNLO calculations. The following W and Z event counting scheme is used:

- For the Z signal, a straight forward counting method is used. Candidate events are counted within $\pm 2\sigma$ around the Z peak, as determined by a Gaussian fit.

- The backgrounds are determined from a side band method, using the number of events found between $3 - 5\sigma$ on both sides of the Z peak.

- The Z cut efficiency is calculated for events where the generated mass is on-shell, requiring $91.2 \pm 7.5 (= 3\Gamma_Z)$ GeV and the generated rapidity Y of the Z satisfies $|Y| < 2$. The number of accepted events in the Monte Carlo are calculated like in the data, including the subtraction of "signal" events found in the side band regions. A correction of 1.5% (resp. 0.6%) for central–central (resp. central–plug) Z is applied to the cut efficiency to take into account the contamination of the off-shell (γZ^*) events, populating more the side bands than the central region. This correction factor is estimated using the Monte Carlo.

- The counting of the W signal relies on the transverse mass, M_T, distribution calculated from the electron neutrino system and requiring 60 GeV $< M_T <$ 90 GeV.

- Backgrounds are separated into off-shell production of $W^* \rightarrow e\nu$, which are reconstructed within the signal region, as well as background from the $W \rightarrow \tau\nu$, $Z \rightarrow ee$ and "QCD" dijet events.

- The W cut efficiency is calculated, like in the case of the Z, for events where the generated mass is on-shell, requiring $80.4 \pm 6.4 (= 3\Gamma_W)$ GeV and the generated rapidity Y of the W satisfies $|Y| < 2$. The number of accepted events in the Monte Carlo are calculated as in the data.

Estimating the background for the W sample

The sources of background for the W signal were etimated using Monte Carlo simulations. The major source of background for W events comes from off-shell W production and is estimated to be $6.5 \pm 0.2\%$ of the selected events, using the PYTHIA Monte Carlo sample and keeping only events where the generated mass is either smaller than 74 GeV or higher than 86.8 GeV (i.e. 3Γ around the mean generated mass).

The background coming from the $W \rightarrow \tau\nu$ was estimated using a sample of 102,000 events generated with PYTHIA and found to be $1.5 \pm 0.1\%$ of the selected W signal events.

For the $Z \rightarrow ee$ background a sample of 100,000 Monte Carlo PYTHIA events was used and the contamination is found to be $0.47 \pm 0.02\%$ of the selected W signal events.

The majority of the QCD dijet events are removed with the application of the jet veto. To study this background, a sample of Monte Carlo QCD dijet events containing $2.7 \cdot 10^6$ events was used. As the cross section for these type of events is 50 μb, we expect $6 \cdot 10^9$ events for a luminosity of 125.5 pb^{-1}. The sample we had contains then less than 0.04% of the events corresponding to a luminosity of 125.5 pb^{-1}. Thus to get around this problem, two different factorization methods were used.

The first way to approximate the QCD background relies on the idea that a jet–jet event fakes a W event if one jet is detected as an electron and the other jet gives missing E_t. One can then measure the probability that a jet is detected as an electron and the probability that a jet is mis-measured, resulting in missing energy. The combined probability is then found to be 10^{-10} which represents 0.6 ± 0.6 events for a luminosity of 125.5 pb^{-1}.

The second way of estimating the QCD background relies on the observation that the p_t spectrum of the jets from QCD dijet events is steeply falling, leading to more background in the lower p_t region. Furthermore, the region of the low missing transverse energy is poorly described by the Monte Carlo simulation if no QCD dijet background is assumed. A sample in the Monte Carlo QCD dijet was thus selected containing events where a jet is back to back to the missing energy vector in the transverse plane (arising from a "not–reconstructed" jet). One can then determine the shape of the missing transverse energy spectrum of this sample for different jet energies. This background gets dangerous when a jet is misidentified as an electron. Assuming that this happens from time to time and that it enhances the low missing transverse energy region, one can scale the dijet contribution, so that, when added to the other Monte Carlo

components, the simulation describes better the lower part of the missing energy spectrum in the data.

Figure 2.11 shows the data (triangles) together with the Monte Carlo prediction for $W \rightarrow e\nu$, $W \rightarrow \tau\nu$, $Z \rightarrow ee$ (dashed black line) and the Monte Carlo including also the QCD dijet sample (solid blue line). The missing transverse energy distribution is plotted for different electron transverse energies (20-25 GeV, 25-30 GeV, 30-35 GeV and 35-45 GeV). For the four different regions the scaling factors to be applied to the QCD jet sample is 0.18, 0.14, 0.08 and 0.0018 respectively. Multiplying this scaling factor with the number of QCD jet events found in the regions where the counting of W bosons is done ($E_t^{miss} > 30$ GeV and 30 GeV $< E_t^{electron} < 45$ GeV), the QCD dijet background is estimated to be 1.3 ± 0.9 events per 125.5pb^{-1}.

Both methods show that the QCD background can be neglected with the selection used.

The total background for W events is estimated to be 8.5 ± 0.7 % with the following composition: 6.5% of the background events are from off-shell W bosons, 1.5% are from W decays into τ's and 0.5% are from Z bosons.

2.2.6 Stability of the detector

Trigger efficiency

The efficiency for ELECTRON_CENTRAL_18 trigger, used to select the W/Z events, has been estimated from the coincidences between the W_NOTRACK trigger and the ELECTRON_CENTRAL_18 trigger. It is found to vary from 0.92 to 0.97 over the different periods (see Table 2.4, page 58). Moreover the results for the trigger efficiencies for the Z sample have been cross-checked using the Z_NO_TRACK trigger.

The trigger efficiency depends on the electron pseudorapidity. Figure 2.12a shows the trigger efficiency as function of the electron rapidity, for golden and silver electrons, for ELECTRON_CENTRAL_18 trigger (black dots) and W_NO_TRACK trigger (green triangles).

The space bars at the center of the tracking chamber causes inefficiencies for the track finding around $\eta = 0$, leading to an overall lower ELECTRON_CENTRAL_18 trigger efficiency for small rapidities. However, as the event vertex is quite spread in z, one should consider the η dependence of the trigger efficiency for different z_{vertex} position. If the event vertex is shifted in z, the position of the central detector space bars will correspond to a shifted η. Figure 2.12b shows the trigger efficiency for events for which $z_{vertex} < -10$ cm. A clear drop at the value of pseudorapidity corresponding to the position of the bars for such event vertices can be seen. The presence of events with different z_{vertex} positions explains then the "up and down" behavior of the overall trigger efficiency around $\eta = 0$ in Figure 2.12a.

For the W_NOTRACK trigger efficiency (green triangles), no significant η dependence is observed as this trigger does not have any tracking requirements and is thus only sensitive to the calorimeter efficiency. In Figure 2.12b one can see a small reduction of the W_NO_TRACK trigger efficiency for $\eta = -0.5$ which come from the gap between the central and plug calorimeter.

Counting stability over time

To study the stability of the detector the number of Z and W bosons have been counted for each of the different periods as defined in Table 2.1 on page 39.

The number of W and Z events per period divided by the CLC luminosity estimate of that period, including corrections for the variations in the trigger efficiency are shown in Figure 2.13. The data are found to be roughly constant within $\pm$ 5% (a χ^2 test leads to 10.9, 13.1, 14.0 with 6 degrees of freedom for Z events with two central electrons, with one central and one plug electron and W events respectively).

The same can be done with the ratios between the number of Z events having both electrons in the central calorimeter, the number of Z events where one electron is in the plug calorimeter

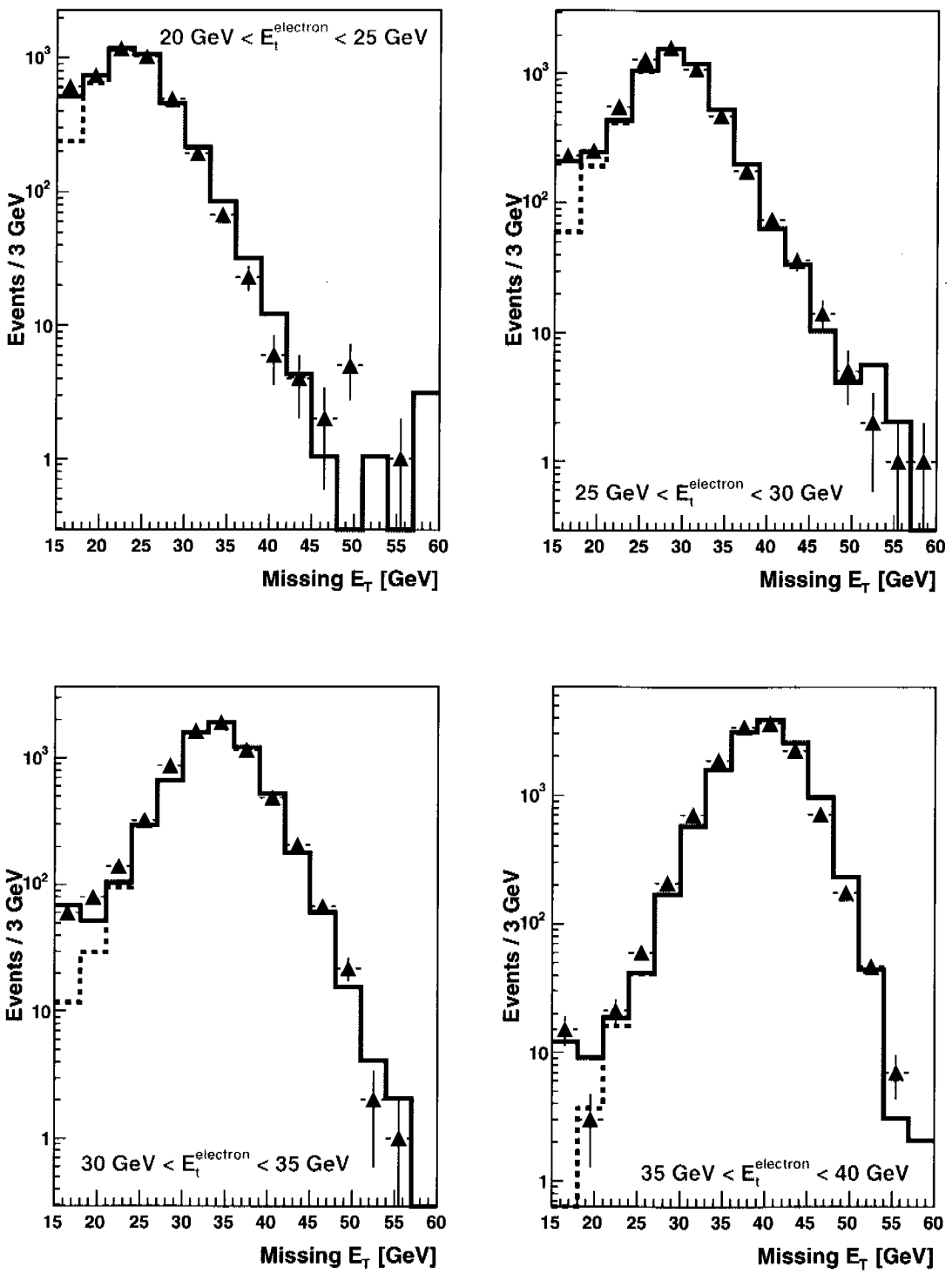

FIGURE 2.11: *Missing transverse energy spectrum for different electron energy intervals: for data (triangles), Monte Carlo prediction including the $W \rightarrow e\nu$, $W \rightarrow \tau\nu$ and $Z \rightarrow ee$ contributions (dashed black line) and Monte Carlo prediction with the additional QCD dijets contribution with appropriate scaling (solid blue line).*

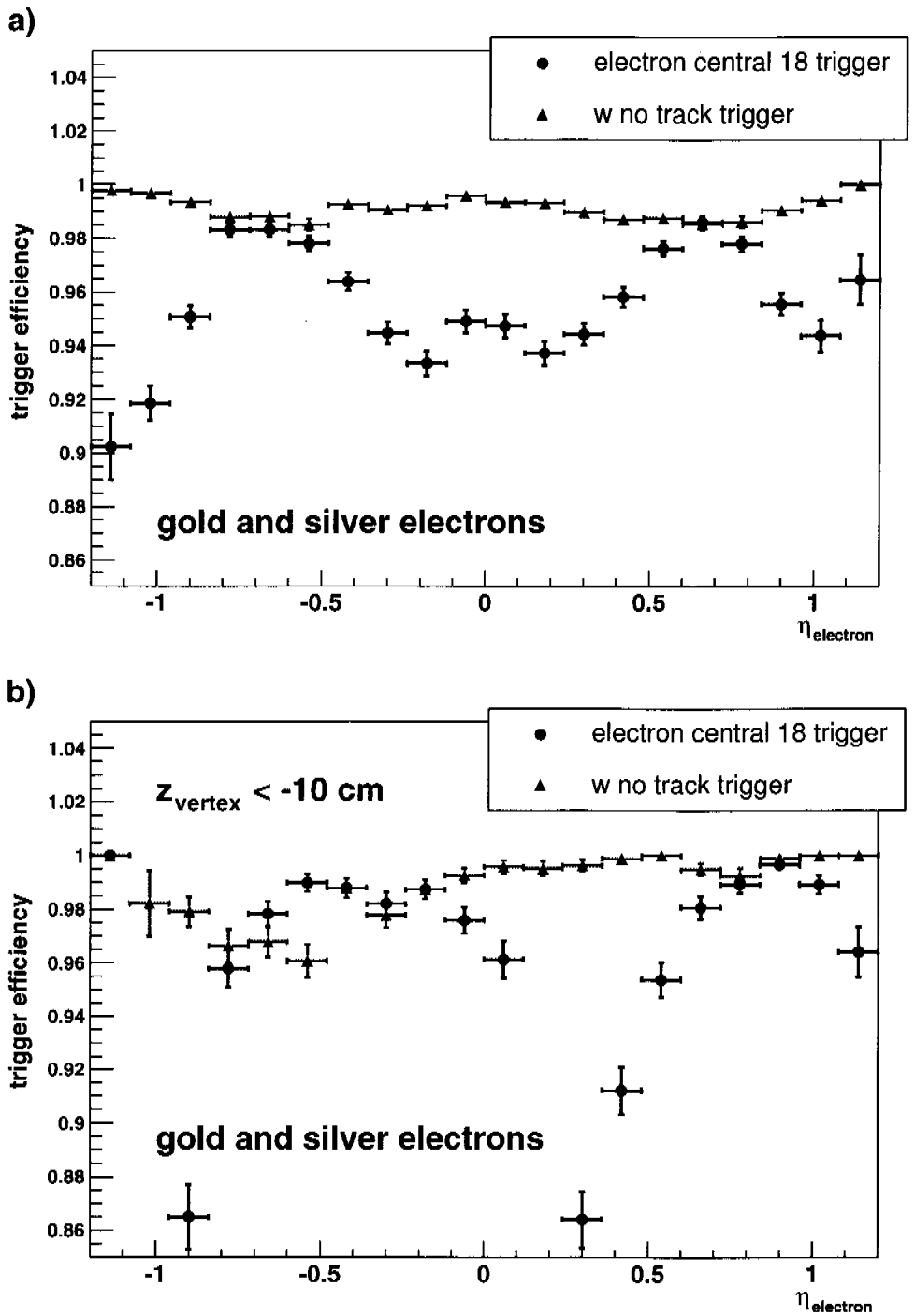

FIGURE 2.12: *The trigger efficiencies measured from the W sample as a function of the electron pseudorapidity for the ELECTRON_CENTRAL_18 trigger and W_NOTRACK trigger (a). The same but for events where the z position of the vertex is smaller than -10 cm (b). Statistical errors are shown. Electrons are required to be gold or silver candidates and pass the selection cuts for a W event.*

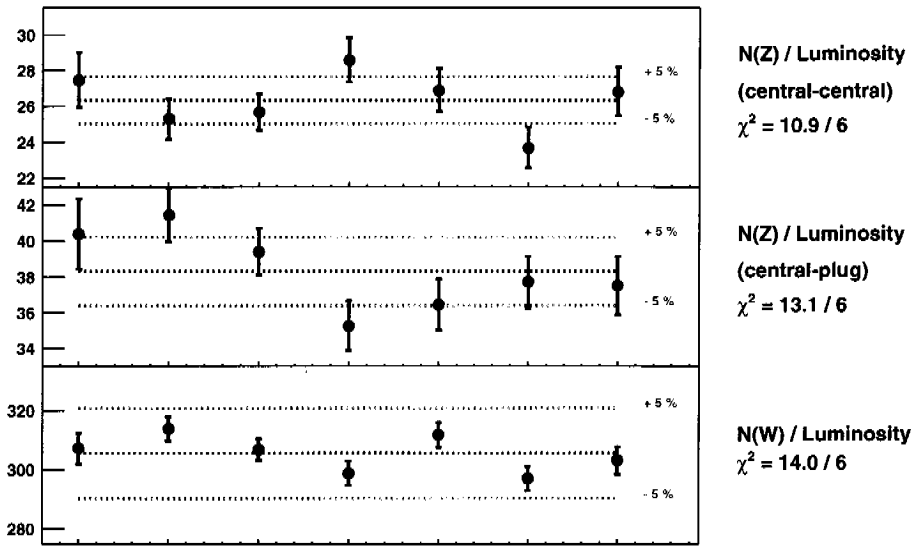

FIGURE 2.13: *Ratios between the number of W and Z events corrected with the trigger efficiency and the corresponding CLC luminosity estimates for the seven different data periods as defined in Table 2.1. The central line shows the average for all the data. The data are found to be constant within ± 5%. Only statistical errors are shown.*

and the number of W events. These ratios are found to be also constant within 5%, as it can be seen in Figure 2.14 (a χ^2 test leads to 16.3, 8.5, 10.6 with 6 degrees of freedom for Z central–central events, Z central–plug events and W events respectively) and are consistent with the statistical errors.

From these results, no significant time-dependent effect can be seen. Even if some points show a discrepancy (the χ^2 test is not very good), a global trend for the same period is not observed in the other channels. More statistics would be needed to be able to analyze smaller periods (for instance each time the trigger table is modified or the running conditions are changed) keeping at the same time reasonably small statistical errors.

Homogeneity of the calorimeter

To study the homogeneity of the calorimeter, the calorimeter towers where the electron coming from the W is detected are studied as a function of η and ϕ. No hole or significant inefficiency was found in any of the towers.

The calorimeter towers, grouped in ϕ intervals, are found to be constant for the data and the Monte Carlo, as it can be seen on Figure 2.15. Fitting a straight line through those points results in a χ^2 of 26.8/23 for the data and 46.6/23 for the Monte Carlo.

Figure 2.16a shows the calorimeter towers where the electron coming from a W is detected, grouped in η intervals, for data and Monte Carlo. The Monte Carlo is normalized to the data. The first and the last tower are not used to get the normalization factor as they are not well simulated in the Monte Carlo. The data are corrected with the trigger efficiency, which depends on η (as shown on Figure 2.12, page 52). Quite a big discrepancy is seen between the central and non-central region description from data and Monte Carlo, as shown in Figure 2.16a. The origin of this effect was found to originate mainly from a bad reconstruction of the variable

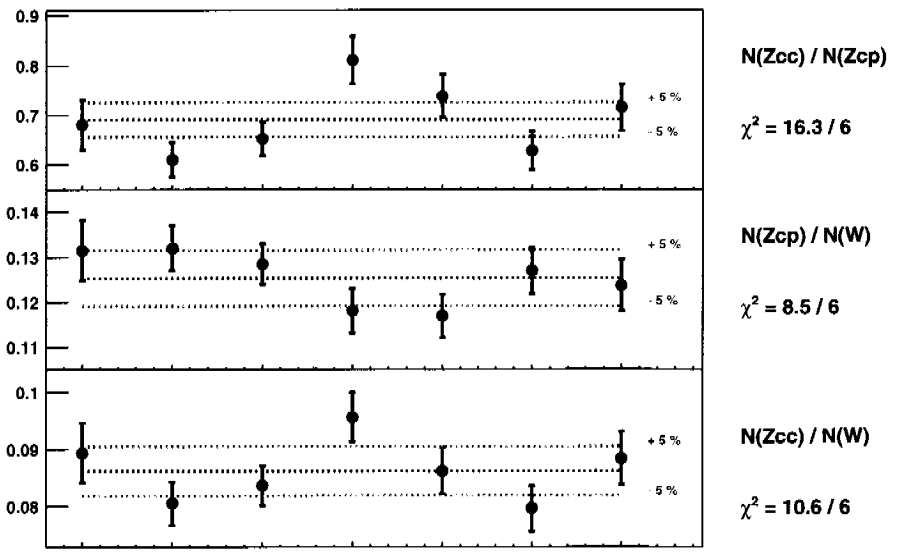

FIGURE 2.14: *Ratios between the number of Z central–central events to Z central–plug events (upper plot), of the number of Z central–plug events to the number of W events (middle plot) and of the number of Z central central events to the number of W events corrected with the trigger efficiency (bottom plot), for the different data periods (Table 2.1). The data are found to be constant within ± 5%. Only statistical errors are shown.*

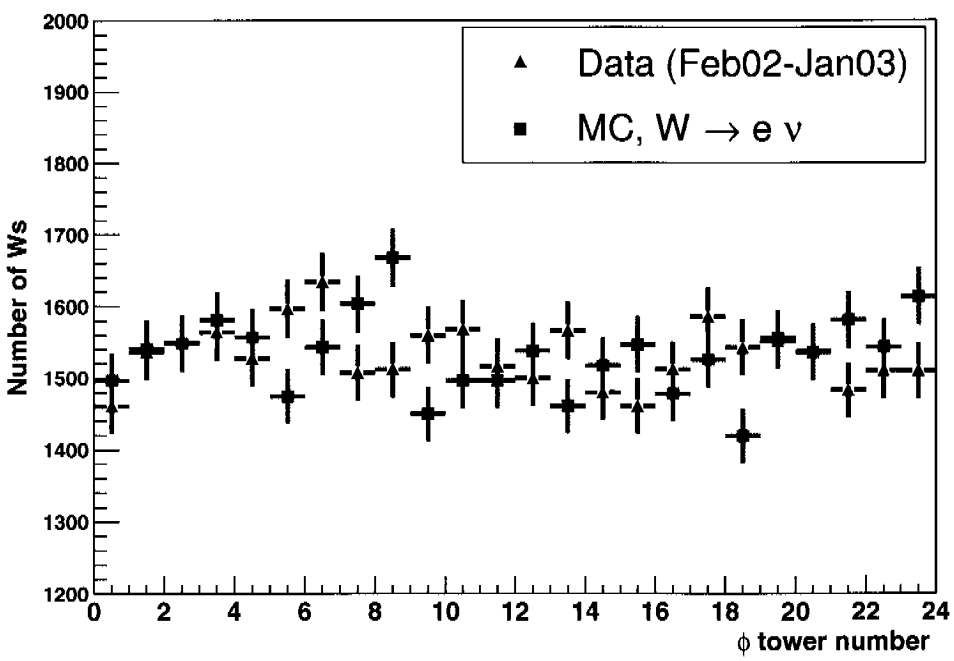

FIGURE 2.15: *Calorimeter towers where the electron coming from a W is detected, grouped in φ intervals for data and Monte Carlo.*

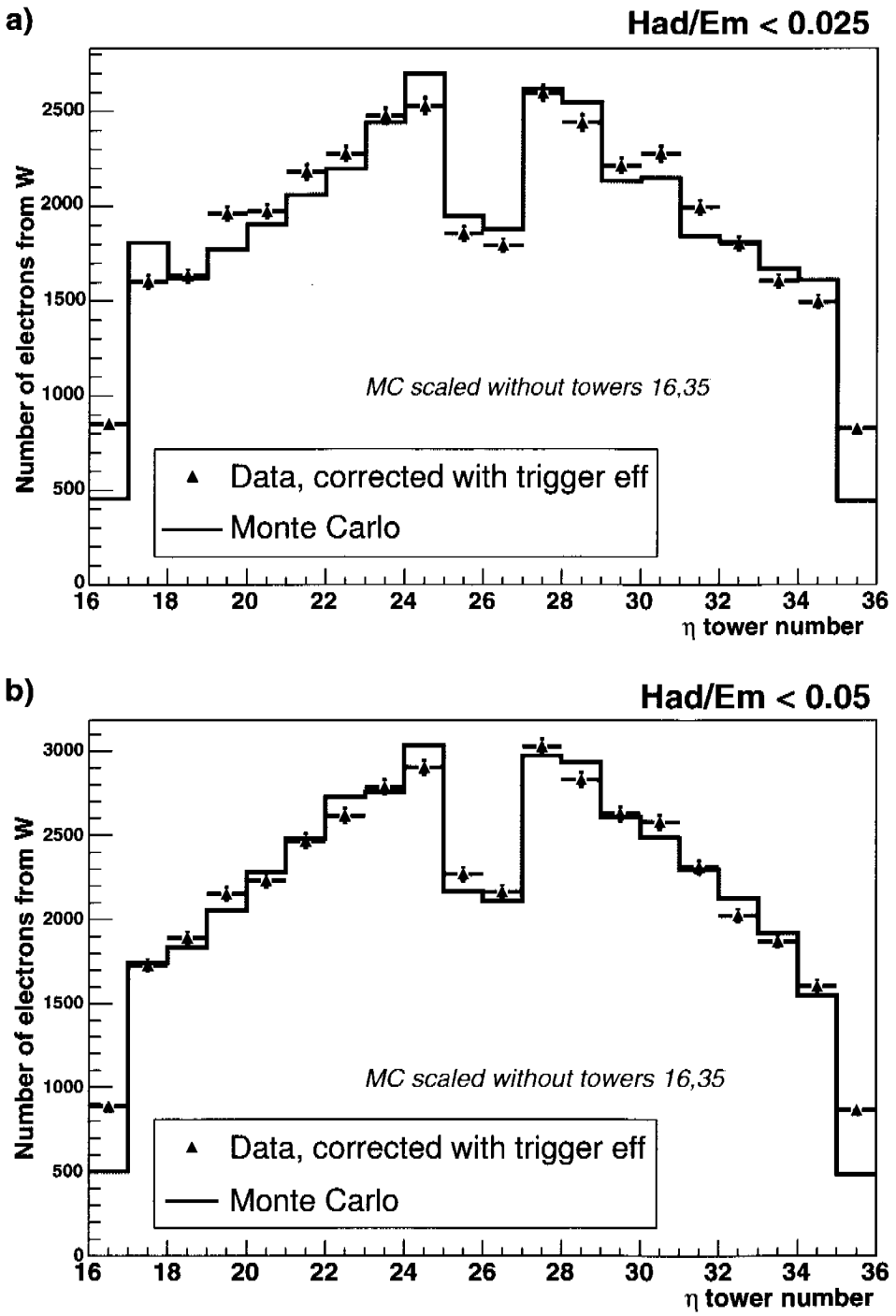

FIGURE 2.16: *Calorimeter towers where the electron coming from a W is detected, grouped in η intervals, for data and Monte Carlo (a). The same but when the cut on $\frac{E_{had}}{E_{em}}$ for the electron selection is relaxed from 0.025 to 0.05 (b). The Monte Carlo is normalized to the data without taking into account the two outer edge towers.*

$\frac{E_{had}}{E_{em}}$ in the central towers. The mean $\frac{E_{had}}{E_{em}}$ is found to be higher for central towers than for the remaining ones in the data, whereas such an effect is not seen in the Monte Carlo, as illustrated in Figure 2.17.

This could be explained by the fact that due to the space bars in the center of the calorimeter, more electromagnetic energy gets lost, resulting in a higher $\frac{E_{had}}{E_{em}}$ value. If the W selection is done with the cut on this variable relaxed from 0.025 to 0.05, as shown on Figure 2.16b, the distribution is in agreement with the Monte Carlo. However, in this analysis, the value of $\frac{E_{had}}{E_{em}}$ was left to 0.025 as it allows a cleaner selection. The discrepancy between data and Monte Carlo is taken into account in the evaluation of the systematic uncertainty.

Finally, the position of the Z peak is found to vary over the different periods from about 1.5% of the mean reconstructed mass for the events where both electrons are central and about 2.5% of the mean reconstructed mass for the events with one electron in the plug.

2.2.7 Luminosity estimation with W and Z bosons

Luminosity determination

The relation between the luminosity $\mathcal{L}$ and the number of signal events N_{signal} is given by:

$$\mathcal{L} = \frac{N_{signal}}{\sigma_{NNLO} \cdot \varepsilon_{z-event} \cdot \varepsilon_{|Y|<2} \cdot \varepsilon_{cuts} \cdot \varepsilon_{trigger}} \tag{2.8}$$

The NNLO theoretical cross section σ_{NNLO}, including the branching ratios to electrons at a center of mass energy $\sqrt{s} = 1.96$ TeV is 2.687 nb for inclusive on shell W production and 0.2513 nb for the Z production, as given in Formula (2.7), page 38.

As explained in section 2.2.4, $\varepsilon_{z-event}$ is a factor that corrects the efficiency of the cut requiring that $z_{event\,vertex} = \pm 60$ cm is lower in the data than in the Monte Carlo. This factor is found to be 0.98.

$\varepsilon_{|Y|<2}$ is the efficiency of the cut requiring that the generated vector bosons have a rapidity smaller than 2 and is found to be 0.92 for Z events and 0.89 for W events.

ε_{cuts} is defined from the ratio of accepted signal events divided by the number of generated events in the Monte Carlo. The generated events must have a generated Z mass fulfilling: $|91.2\,\text{GeV} - M_{gen}| < 3\Gamma$ and a Z rapidity of $|Y_Z| < 2$. The efficiency is 0.115 for the Z central–central, 0.198 for the Z central–plug and 0.123 for the W events.

The results for this luminosity estimate with the different signals, central–central and central–plug Z and for central W are given in Table 2.4.

The luminosity values obtained using the Z events with two central electrons agree with the ones obtained using W events and are also in good agreement with the CLC luminosity estimation.

The luminosity obtained with the Z sample where one electron is in the plug calorimeter and the other is central are systematically too low by about 15 to 20%. It is interesting to note, as seen on Figures 2.13, page 53 and 2.14, page 54, that the ratio between the Z sample for a central and a plug electron and the other samples and the CLC luminosity is constant within 5%. No obvious experimental explanation was found so far and the Z sample where electrons have a larger rapidity needs to be investigated in more details. It is also possible that the discrepancy for the high rapidity region comes from inaccuracies in the PDF. Figure 2.3 on page 36 shows that for $Y > 1.5$ the errors on the PDF start to be rather important.

Systematics

Systematic errors can be put on each of the factors appearing in Formula (2.8):

- $\mathbf{N_{signal}}$: The systematic error on the number of signal events was approximated by studying how the luminosity varies if one changes the definition of the signal and background

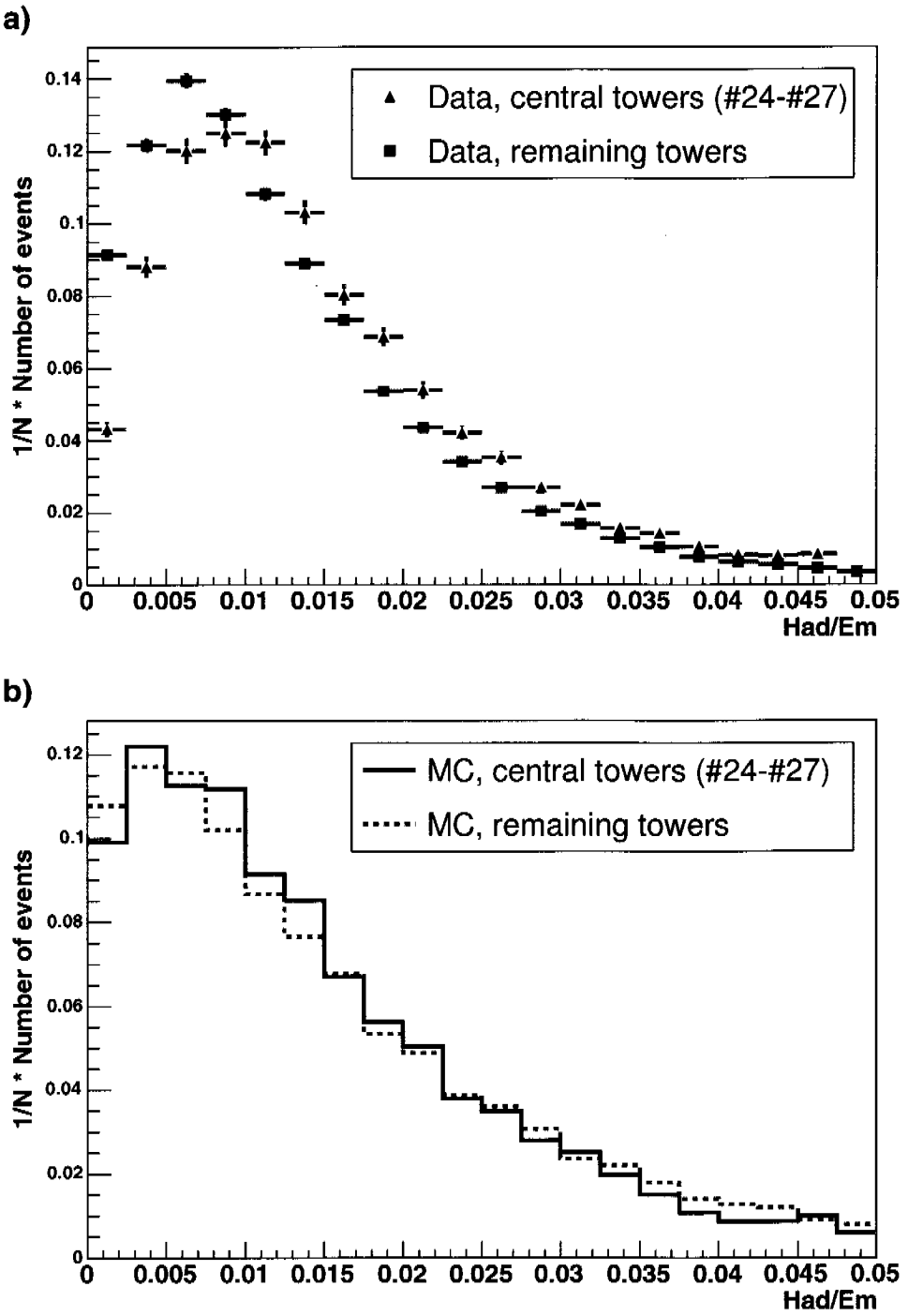

FIGURE 2.17: $\frac{E_{had}}{E_{em}}$ *for electrons which are detected in the four central towers of the electromagnetic calorimeter (red triangles) and in the remaining ones (brown squares), for data (a) and Monte Carlo (b).*

Data period	Sample type	Signal events	$\varepsilon_{trigger}$	$L_{p\bar{p}}$(W,Z) [pb^{-1}]	$L_{p\bar{p}}$(CLC) [pb^{-1}]	Ratio
1	Z cc	317	0.995	12.3 ± 0.7	11.6	1.06 ± 0.06
	Z cp	433	0.924	10.4 ± 0.5	11.6	0.90 ± 0.04
	W 0 jet	3293	0.924	11.7 ± 0.2	11.6	1.01 ± 0.02
2	Z cc	485	0.999	18.7 ± 0.8	19.2	0.97 ± 0.04
	Z cp	775	0.974	17.7 ± 0.6	19.2	0.92 ± 0.03
	W–0 jet	5870	0.974	19.7 ± 0.3	19.2	1.03 ± 0.01
3	Z cc	625	0.998	24.1 ± 0.9	24.4	0.99 ± 0.04
	Z cp	920	0.957	21.3 ± 0.7	24.4	0.87 ± 0.03
	W–0 jet	7164	0.957	24.5 ± 0.3	24.4	1.00 ± 0.01
4	Z cc	539	0.998	20.8 ± 0.9	18.9	1.10 ± 0.05
	Z cp	635	0.953	14.8 ± 0.6	18.9	0.78 ± 0.03
	W–0 jet	5382	0.953	18.4 ± 0.3	18.9	0.97 ± 0.01
5	Z cc	494	0.999	19.1 ± 0.9	18.4	1.04 ± 0.05
	Z cp	647	0.965	14.9 ± 0.6	18.4	0.81 ± 0.03
	W–0 jet	5537	0.965	18.7 ± 0.3	18.4	1.02 ± 0.01
6	Z cc	437	0.999	16.8 ± 0.8	18.5	0.91 ± 0.04
	Z cp	671	0.962	15.5 ± 0.6	18.5	0.84 ± 0.03
	W–0 jet	5286	0.962	18.0 ± 0.2	18.5	0.97 ± 0.01
7	Z cc	388	0.998	14.9 ± 0.8	14.5	1.03 ± 0.05
	Z cp	522	0.960	12.1 ± 0.5	14.5	0.83 ± 0.04
	W–0 jet	4220	0.960	14.4 ± 0.2	14.5	0.99 ± 0.02
TOTAL	**Z cc**	**3285**	**0.998**	**126.3 ± 2.2**	**125.5**	**1.01 ± 0.02**
	Z cp	**4603**	**0.957**	**106.4 ± 1.6**	**125.5**	**0.85 ± 0.01**
	W–0 jet	**36752**	**0.964**	**125.4 ± 0.7**	**125.5**	**1.00 ± 0.01**

TABLE 2.4: *The observed number of Z events for central–central (cc) central plug (cp) and W events (W: without background subtraction) in the data and for the different periods, the estimated trigger efficiencies and the resulting luminosity with statistical errors. The last column shows the ratio between the W/Z luminosity ($L_{p\bar{p}}$(W,Z)) and the CLC luminosity ($L_{p\bar{p}}$(CLC)), assuming a negligible statistical error on the CLC estimate.*

zones used for the counting. For the W signal, one finds a variation of 1% in the obtained luminosity if one varies the cut on the W transverse mass used for the signal definition. Similar tests were also performed on the Z sample and a variation of 0.5% of the luminosity was found.

- $\varepsilon_{|\mathbf{Y}|<\mathbf{2}}$: Depending on the available statistics, the signal can be split into several rapidity bins. For this measurement, the efficiency corrections of W and Z events are calculated such that the generated rapidity of W and Z bosons satisfies $|Y| < 2$. The fraction of events with larger rapidities is found to be 7.7% for the Z's and 10.7% for the W's. This fraction depends on the assumed PDF knowledge at small and large x.

  Using the same PDF, a slightly different rapidity distribution of the W and the Z bosons is found for the simulation made with PYTHIA and the one made with HERWIG. Figure 2.18 shows the output of PYTHIA and HERWIG for the generated rapidity of the W for all events and for the selected signal events. After having talked to the authors of PYTHIA and of HERWIG, no obvious explanation was found.

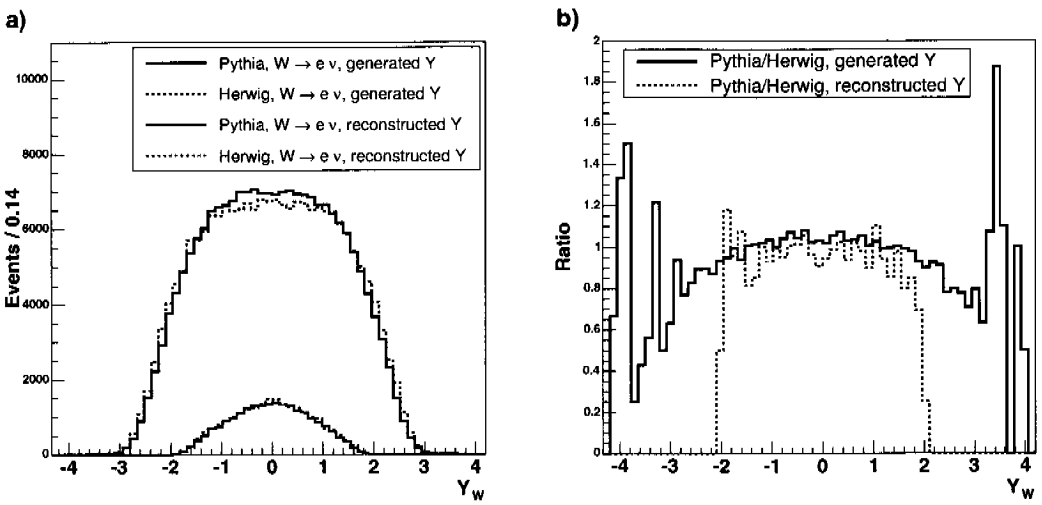

FIGURE 2.18: *(a) Distribution of the generated (without any basic kinematic selection) and reconstructed rapidity distribution for W events in the PYTHIA (solid black and red lines) and HERWIG simulation (dashed blue and green lines). (b) The ratio between the W rapidity prediction of PYTHIA and HERWIG simulations for all generated events (solid black line) and for the reconstructed ones (dashed red line).*

A systematic error on the final result of $\pm 1.5\%$ is attributed to this rapidity extrapolation.

- $\sigma_{\mathbf{NNLO}}$: The theoretical error on both cross sections is assumed to be 2.8% as explained on page 38.

- $\varepsilon_{\mathbf{cuts}}$:

  The systematic error on the cut efficiency is estimated by varying the values of the cuts recording the change in the luminosity. We found a 2.2% variation in the luminosity for the Z sample where both electrons are central, 2.2% variation for the Z sample when one electron is in the plug and 4.7% for the W sample. The main source of error comes from the variation of the cut on $\frac{E_{had}}{E_{em}}$. This distribution shows discrepancies between the Monte Carlo prediction and the data, as illustrated in Figure 2.17, on page 57.

  The efficiency of the jet veto applied in the W event selection was also studied. The accuracy of the Monte Carlo prediction can be checked using the Z sample with two

central electrons. The fraction of Z events without jets to the total number of Z events is found to be 0.799±0.007 in the data and 0.827±0.004 for the PYTHIA Monte Carlo. A correction factor of 1.035±0.0125 is then applied in the Monte Carlo efficiency to correct for the jet veto cut.

- $\varepsilon_{\mathbf{trigger}}$: The error on the trigger efficiency is estimated using the the statistical error on the number of events firing both ELECTRON_CENTRAL_18 trigger and W_NO_TRACK trigger and is found to be negligible (between 0.05 and 0.2%).

Combining all these errors, we find a total systematic error of 3.9% for both Z samples and 5.8% for the W sample.

Table 2.5 gives a summary of the list of all studied systematics.

Systematic errors for:	Zcc	Zcp	W
N_{signal}	0.5%	0.5%	1.0%
σ_{NNLO}	2.8%	2.8%	2.8%
$\varepsilon_{\|Y\|<2}$	1.5%	1.5%	1.5%
ε_{cuts}	2.2%	2.2%	4.7%
$\varepsilon_{trigger}$	0.05%	0.2%	0.2%
Total	3.9%	3.9%	5.8%

TABLE 2.5: *List of studied systematic errors.*

This study of the systematic errors should be carried on in more details in a further analysis. An important point left aside is the detailed study of the plug calorimeter where currently a lower luminosity is systematically found in each data period. The reason could be caused by inaccuracies in the PDF, but more likely to an inaccurate simulation of the plug calorimeter and to unknown detector inefficiencies.

2.2.8 Results and interpretation of W and Z production

Kinematics of the W and Z bosons events

In this section, different kinematic distributions are studied. For the signal simulation, a comparison between the predictions of PYTHIA and HERWIG is performed. The transverse momentum spectrum and rapidity distribution are discussed.

Figures 2.19 and 2.20 show the p_t for Z bosons reconstructed with central–central and central–plug electron pairs. The energy correction based on the position of the Z peak, as explained on page 43 is applied on the electrons. For Z bosons with large p_t and two central electrons, one observes 491 events with a transverse momentum higher than 20 GeV against 352 events expected by the PYTHIA Monte Carlo, which represents an excess of more than 5σ. From the HERWIG Monte Carlo, however one expects 461 events with a p_t higher than 20 GeV, which is in good agreement.

It is also interesting to consider the low p_t part of the spectrum. Looking first at the Z sample with two central electrons, one sees that HERWIG describes quite well that part of the spectrum. However, for the Z bosons where one electron is central and the other is in the plug, both HERWIG and PYTHIA do not describe the data.

The shape of the p_t spectrum in Leading Order Monte Carlo generators will be determined by the initial state radiation (ISR). For soft or collinear partons, the Altarelli-Parisi approach can be used to simulate the ISR. However, this probabilistic procedure does not cover the whole parameter space and for the high p_t part of the spectrum, the exact matrix elements of the given process have to be used.

In the collinear/soft approximation, the following parameters control the evolution of the ISR: the value of α_s, the Q^2 scale at which the partons stop showering and the intrinsic transverse momentum (k_T) of the partons in the proton. These parameters were set differently in PYTHIA and in HERWIG. The hard part of the p_t spectrum will depend on how the matrix element corrections are implemented, which is different in PYTHIA and in HERWIG [54]. This could explain the discrepancies observed between the two generators.

For low Z momenta, less events are seen in the data as expected from the simulation. This might explain the systematically too low measured luminosity using the central–plug sample. However this 10% effect could not explain alone the 15-20% discrepency observed.

Figure 2.21 shows the rapidity for Z bosons reconstructed with central–central and central–plug electron pairs. No significant difference is observed between data and the HERWIG and PYTHIA predictions. As discussed in [8], the measurement of the Z rapidity distribution constrains the PDF and the x_1, x_2 range, which leads directly to the "parton–parton" luminosity.

Figure 2.22 shows the shape of the electron pseudorapidity distribution for the $W^{\pm}$ events with zero jets in the data and the PYTHIA and HERWIG Monte Carlo. The data are corrected with the trigger efficiency. As explained in Section 2.2.6, page 53, in order to get a better shape description of the data by the Monte Carlo, the cut on $\frac{E_{had}}{E_{em}}$ is relaxed from 0.025 to 0.05. Due to the different momenta distributions in the proton for the u, d, $\bar{u}$ and $\bar{d}$ quarks, to be able to compare the W^+ with the W^- rapidity distribution, the sign of the rapidity for the W^- distribution needs to be inverted with respect to the W^- rapidity distribution. All four curves are in reasonable agreement.

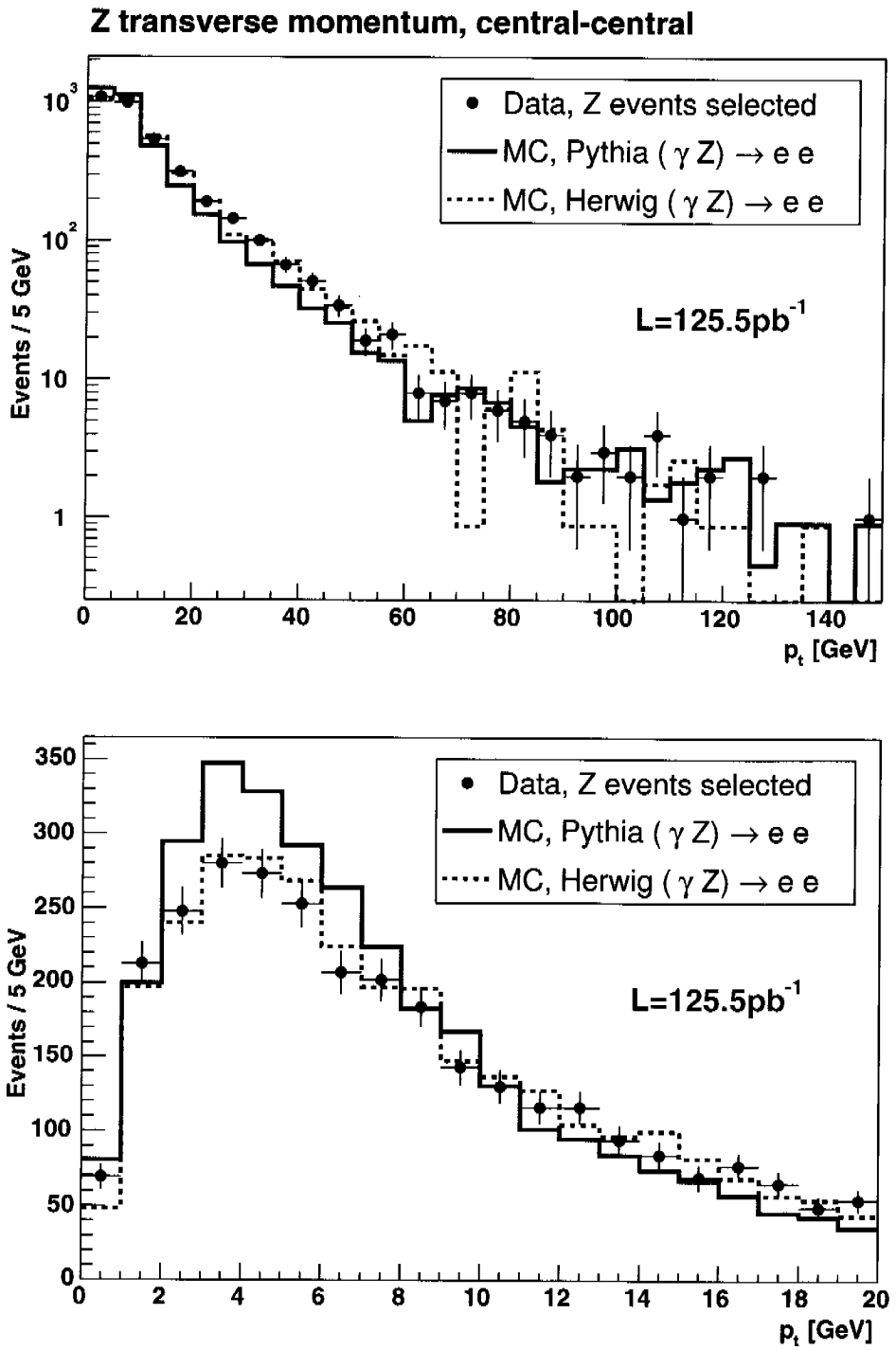

FIGURE 2.19: *Reconstructed Z transverse momentum for the PYTHIA Monte Carlo (histogram) and data (dots with error bars) for central–central events on two different scales. The dotted line is the expected distribution obtained with the HERWIG Monte Carlo program.*

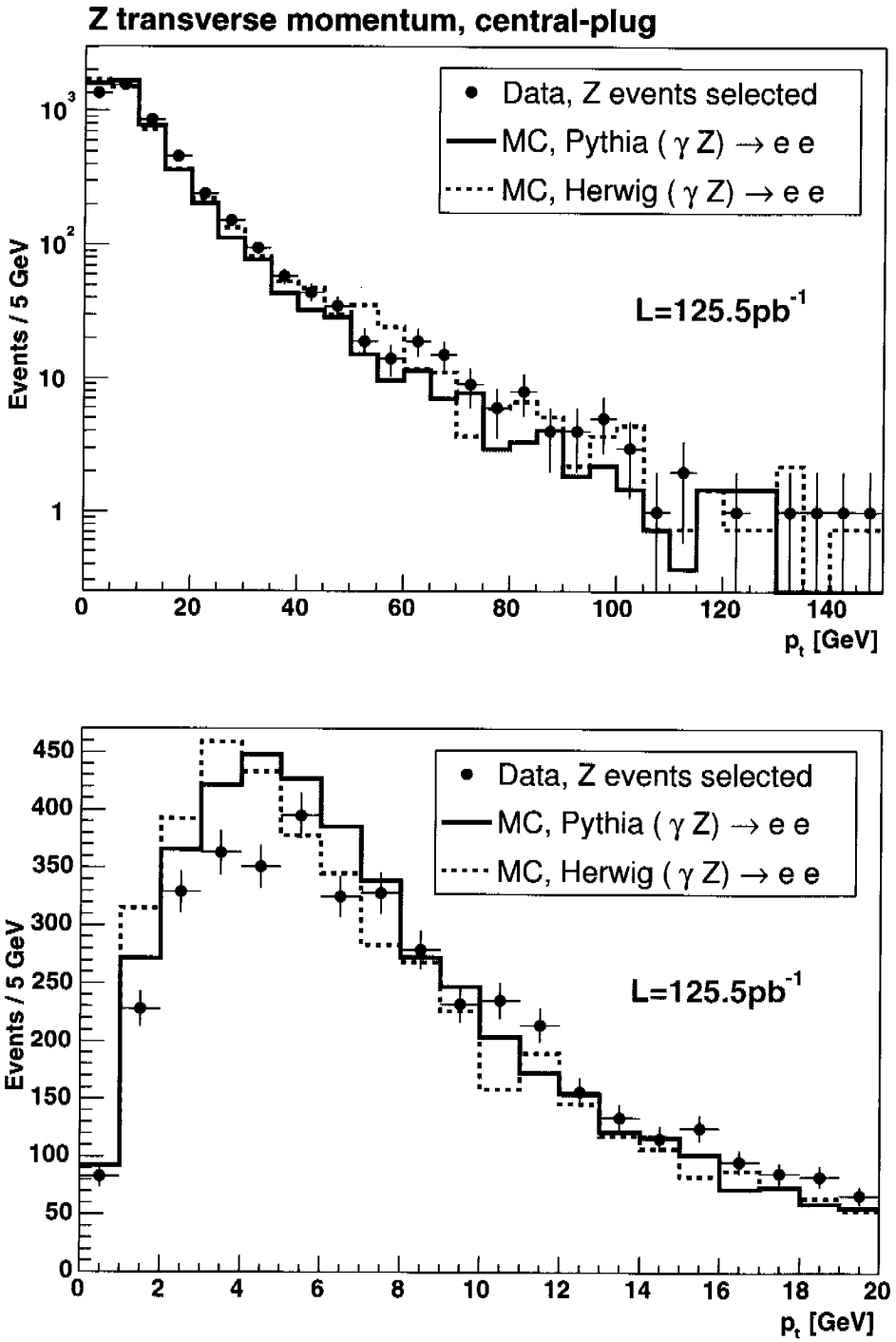

FIGURE 2.20: *Reconstructed Z transverse momentum for the PYTHIA Monte Carlo (histogram) and data (dots with error bars) for central–plug events on two different scales. The dotted line is the expected distribution obtained with the HERWIG Monte Carlo program.*

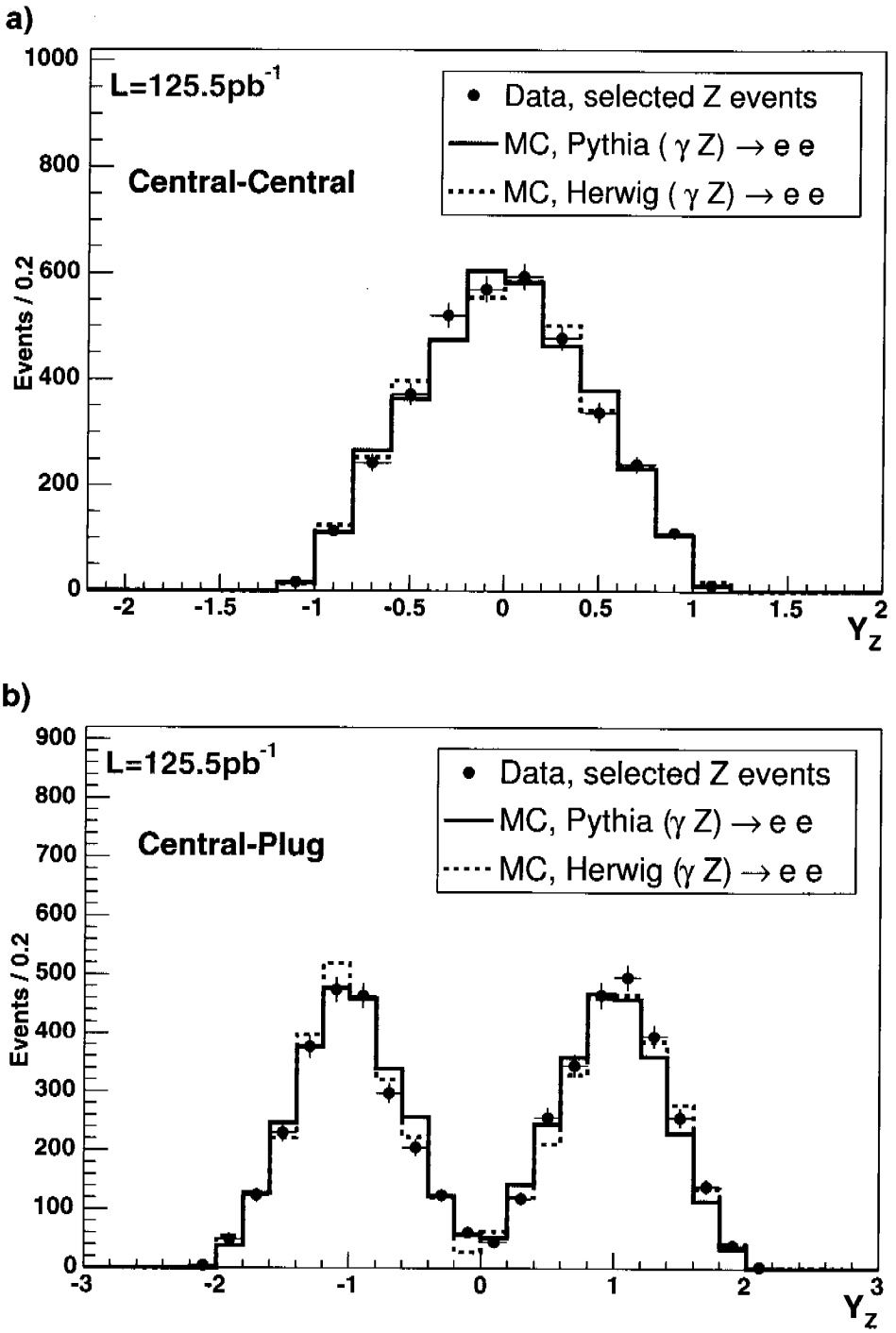

FIGURE 2.21: *Rapidity distribution for the observed Z signal in the data (dots with error bars) and the Monte Carlo (histogram) for central–central (a) and central plug (b) events. The dotted line is the expected distribution obtained with the HERWIG Monte Carlo program.*

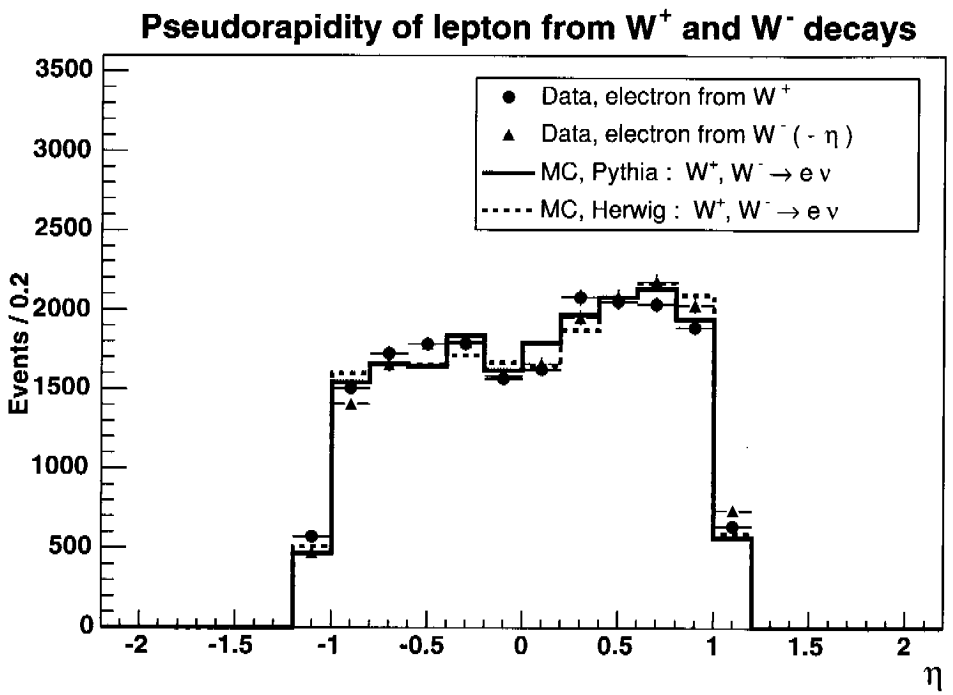

FIGURE 2.22: *Lepton pseudorapidity distribution of the $W \rightarrow e\nu$ events for data (dots with error bars) and PYTHIA simulated events (histogram). For data both W^- and W^+ contributions are shown. The dotted line is the expected distribution obtained with the HERWIG Monte Carlo program. As explained before, the cut on $\frac{E_{had}}{E_{em}}$ is relaxed in order to allow the Monte Carlo to fit better the data. The data are also corrected with the trigger efficiency.*

Luminosity with W and Z and the cross section ratio

The final results for the luminosity are then:

$$\begin{array}{rcl} \mathcal{L}_{p\bar{p}}(\mathrm{Zcc}) & = & 126.3 \pm 2.2(\mathrm{stat.}) \pm 4.9(\mathrm{syst.})\ \mathrm{pb}^{-1} \\ \mathcal{L}_{p\bar{p}}(\mathrm{Zcp}) & = & 106.4 \pm 1.6(\mathrm{stat.}) \pm 4.1(\mathrm{syst.})\ \mathrm{pb}^{-1} \\ \mathcal{L}_{p\bar{p}}(\mathrm{W}) & = & 125.4 \pm 0.7(\mathrm{stat.}) \pm 7.3(\mathrm{syst.})\ \mathrm{pb}^{-1} \end{array}$$

The ratio between the Z and W cross section can be also measured. This ratio is determined using only the samples involving central electrons and it is found to be:

$$\begin{array}{rcl} \text{Measured}: \dfrac{pp \to Z \to ee}{pp \to W \to e\nu} & = & 0.094 \pm 0.002(\mathrm{stat.}) \pm 0.006(\mathrm{syst.}) \\ \text{Calculated[49]}: \dfrac{pp \to Z \to ee}{pp \to W \to e\nu} & = & 0.093 \pm 0.001 \end{array}$$

which is in good agreement with the theoretical prediction. The systematic error on the ratio could be even reduced in a further study, as more sources of error could cancel in the ratio, like for instance the error on the cut efficiency or the error on the rapidity extrapolation.

2.2.9 Summary: one year and a half of luminosity determination at CDF

No major obstacle was found for the W and Z counting method which was for the first time applied to the CDF data. This method can be summarized as follows:

- We defined first a way to select electrons being as simple as possible to allow a good control on the efficiencies, but sufficiently effective in background suppression.
- Then a W and a Z sample was selected using the kinematic properties of the processes and the electron identification cuts.
- The Z sample was used to test the quality of the Monte Carlo simulation and to control the cut efficiencies.
- Counting regions for the Z and the W samples were defined and the number of Z and W were counted. Knowing the cross section and the efficiency, the luminosity can be calculated. The next step will be to count the W and the Z in rapidity bins in order to obtain the parton luminosity.
- The systematic errors were finally determined, like for instance the detector stability over time, the detector homogeneity and the uncertainty in the electron selection.

In summary, the data collected by CDF between February 2002 and May 2003 were analyzed in order to get a luminosity estimate. 3285 Z events with two central electrons, 4603 Z events with one central and one plug electron, and 36752 W events were selected. Combining the W and Z with central electrons only a luminosity of 125.5 $\pm$ 0.6 (stat.) $\pm$ 7.1 (syst.) pb^{-1} is obtained which is in good agreement with the CLC measurement of 125.5 $\pm$ 7.3 pb^{-1}, where a 5.8% total uncertainty is assumed [18].

In contrast, the Z sample with a central and a plug electron gives a luminosity estimation systematically 15 to 20% lower than the one predicted with the other samples. This effect could be due to an inaccurate description of the plug calorimeter in the simulation but also to a PDF theoretical error.

The ratio of the Z to W cross section for central electrons only has been measured and is found to be 0.094 $\pm$ 0.002(stat.) $\pm$ 0.006(syst.), in agreement with the theoretical prediction of 0.093 $\pm$ 0.002.

Before applying this method at the LHC, some questions still need to be answered.

$\mathcal{L} = 2 \cdot 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$		
Signal	Background	Total
$W \rightarrow e\nu$: 10Hz	$\pi^{\pm}/\pi^0$ overlap: 5Hz π^0 conversions: 10Hz $b/c \rightarrow e$: 8Hz	33Hz
$Z \rightarrow ee$: 1 Hz	~ 0	1Hz

TABLE 2.6: *The expected rates of single W and single Z into electrons at the LHC with the CMS High Level Trigger during the low-luminosity period [21].*

- *The narrow width approximation*

  To take into account the narrow width approximation used for the theoretical predictions, a cut on the generated W mass was set requiring the generated W events to lie within 3Γ around the generated mass. The theoretical systematic errors linked to this way of implementing the narrow width approximation should be studied in more detail in the future.

- *Determining the scale of the process*

  The precise measurement of the rapidity distribution of the vector boson is fundamental to obtain the parton distribution, as explained on page 34. This rapidity distribution has to be measured at a given scale, which is, in the case of single W/Z, the mass of the vector boson. However, when there are jets in the event, the scale is modified. The influence of the jets on the scale of the process should be studied, also to make sure that the uncertainties on the jet reconstruction do not limit the accuracy of the method. Moreover, a Z+jets sample could be used to confirm the Q^2 evolution of the PDF.

- *Combining LO, NLO and NNLO in a correct way*

  PYTHIA (or HERWIG) are leading order generators. They simulate a transverse momentum of the generated system using initial and final state radiation. The process simulated is thus half way between a LO and NLO process. It is then not clear whether it is better to use the generators with NLO PDF or LO PDF. For the simulation of a specific process, it would be useful to define when NLO calculations are mandatory and when they are less important and could be replaced for instance by a reweighting method [55].

2.3 From CDF to CMS

The techniques applied here to count the W and the Z to measure the luminosity can be applied in a similar way at CMS. For instance, the idea to use the Z sample in order to control the cut efficiencies for the W sample is also relevant for CMS.

A very high rate of W and Z should be produced in the proton proton collisions at the LHC: for a 14 TeV center-of-mass energy, the cross sections for single W and single Z production is about a factor ten compared to the ones at the Tevatron and moreover the LHC is expected to deliver 10 to 100 times more luminosity than the Tevatron. Table 2.6 shows the expected rates of vector bosons and potential background accepted by the CMS high-level trigger. The number of W and Z decaying into electrons expected to be triggered each day during the LHC low-luminosity phase is about 10^6 and 10^5. It will therefore be possible to obtain the parton luminosities by measuring very precisely the rapidity distribution of the W and the Z bosons.

As shown before, the first step to select W and Z is to learn how to select efficiently electrons[9]. The general principles to identify electrons are more or less common to CDF and CMS. However,

[9]Event if this study concentrated on electrons, it is clear that the signal with muons should also be used.

differences in the design of the two detectors will have an influence on the variables used by these two experiments to select electrons: CDF and CMS have a different type of electromagnetic calorimeter. CMS, with its crystal calorimeter, has a much finer granularity than CDF. This high granularity can be used to make cuts on the transversal shower shape. Another important difference is the amount of material in the tracker: about $0.57X_0$ for CMS against $0.02X_0$ for CDF. Electrons in CMS will loose quite a lot of energy through Bremsstrahlung in the tracker, leading to a more difficult momentum determination and a particular shower shape in the calorimeter.

In the following, the electron selection in CMS will be discussed in detail using a full detector simulation. Several systematics problems that are specific to that detector will be analyzed.

Chapter 3

Identifying electrons in the CMS detector

Nowadays a growing effort is put into the simulation of the detector physics capabilities during the design and constructing phase of an experiment. Such Monte Carlo simulations should allow to study the detector reconstruction performance in order to determine which type of measurement might be possible. It should also help to define search strategies and analysis methods to isolate different types of processes.

For instance, in CMS, Monte Carlo simulations are fundamental to design an efficient trigger, able to reduce the event rate from 45 MHz down to 100 Hz. More specifically, for electrons, a basic selection has to be applied already on the High Level Trigger in order to remove part of the high jet background. At low luminosity the expected ratio of electrons (mainly from W decays), to background (like pions or leptons coming from b and c decays), passing the single electron trigger, is about 0.5 for electrons having a transverse energy higher than 29 GeV, as given in the trigger TDR [21]. However, for most measurements, clean electron signals with much improved signal to background ratios are needed. The 'offline' electron reconstruction and selection is thus a second aspect to be studied using a simulation.

In the following the capability of the CMS detector to identify electrons will be discussed using a full detector Monte Carlo simulation. After a description of the CMS reconstruction software chain, selection variables will be presented for electrons at different energies and for potential backgrounds. We will discuss how these variables could be used to provide an efficient electron selection which is also as general as possible. If the same electron selection can be applied in different channels, the errors related to the cut efficiency will cancel, if proper ratios are taken, reducing the systematic uncertainties. Ratios are also the base of the parton luminosity method described before. This electron selection developed here will be used then to select single W events. Finally the homogeneity and the electron energy resolution of the CMS detector will be analyzed.

3.1 The CMS reconstruction chain

The simulation of the detector response to a given physics process can be divided into different steps. First, the kinematics of the physics process to be studied is simulated using a Monte Carlo generator. Out of all the particles produced in the interaction, the detector can only detect the stable ones[1]: electromagnetic particles (electrons, photons and muons) and hadrons (charged: $\pi^{\pm}$, $K^{\pm}$, protons and neutral: K_L^0, neutrons). The interactions of these stable particles with the detector are simulated, determining the different sub-detectors responses. Then, for Monte

[1] "Stable particles" for us are the ones having a life time longer than a nanosecond, implying that they do not have enough time to decay before being detected.

Carlo as well as for data, these signals need to be combined to reconstruct the event. With such a simulation, the capability of the detector to reconstruct a given type of signature can be estimated and the reconstruction and analysis procedure discussed and optimized.

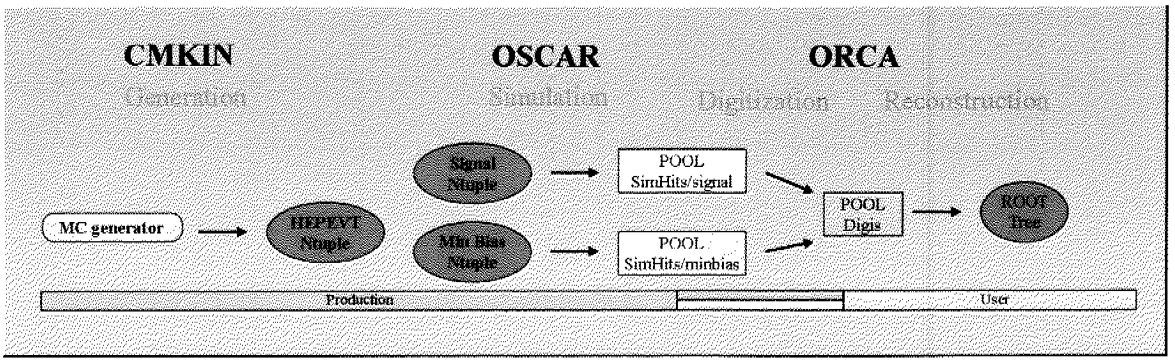

FIGURE 3.1: *The CMS simulation chain from the production to the user. The upper part shows the name of the programs used and the lower part, the type of data storage.*

The program chain used to simulate the CMS detector is shown in Figure 3.1 and will be explained in the following in some detail.

First the CMKIN program [56] simulates the kinematics of the physics process to be studied, by running an event generator like PYTHIA, Isajet or HERWIG and writes a file as output containing a list of the generated particles types together with their 4-vectors. This file format has a structure similar to the HEPEVT Fortran common block used in most event generators.

These stable particles will be propagated through the detector and the interactions with the different detector elements are simulated with the GEANT program [57]. The study presented in the following was done during the transition phase between Fortran, GEANT-3, and C++, GEANT-4. The CMS specific version of GEANT, which uses the detector geometry and materials as well as the magnetic field configuration, is called OSCAR (and CMSIM for the Fortran version). The output of GEANT is a collection of so-called hits[2]. A hit is defined every time a particle crosses a sensitive element of the detector and contains information about the position and time of the hit, the energy or charge deposit.

Next, the hits have to be transformed to signals, usually simulating the electronic outputs, getting the so-called *digis*. This step is done with the ORCA (Object Oriented Reconstruction for CMS Analysis) program [58]. Using these *digis*, ORCA reconstructs the so-called *RecHits*, which are a translation of the electronic signal into a usable quantity like for instance an energy. For example in the electromagnetic calorimeter, this means translating a group of ADC counts in the amount of energy deposited in a crystal. One can add here effects of pile-up and of minimum bias events.

In a second step the sub-detectors responses are combined within ORCA to reconstruct the different elements of the physics process signature, i.e. the electrons, photons, muons, taus and jets. This step is common to the Monte Carlo simulation and to the data. Combining the *RecHits*, ORCA will write the so-called *RecObjects* as an output.

ORCA and OSCAR are written in C++, where the program building blocks are the so-called objects. It is thus convenient to store the data as objects (for example an electromagnetic cluster 'object' should be stored together with the information related to it, e.g. which crystals are in the cluster, its total energy, etc.). Then the programmer can get these objects back together with the informations associated to them from the data stored in the file. In CMS, POOL [59]

[2]For the experts: the hits written by OSCAR are called *SimHits*. Indeed only OSCAR produces directly *SimHits*, CMSIM writes zebra files that have to be formatted in *SimHits* using ORCA.

is used for this.

For the analysis described in the following, *digis* were processed running ORCA to reconstruct the different elements (tracks, clusters, etc.) of a given process. PAX was run together with ORCA to write root files [51]. PAX [60] is a set of C++ classes, that provides convenient tools to analyze the event to be reconstructed, its base components being 4-vectors with additional information. The final step of the analysis could be performed on those root files, allowing a faster processing: with a 2.4 GHz Pentium processor, running on 1000 events with ORCA and requiring the tracks to be reconstructed takes about 4 hours, whereas running on 1000 events takes a few seconds on a root file written with PAX. In the future, running on events with ORCA should be faster due to the introduction of DSTs containing the high level objects, like e.g. tracks, already reconstructed.

3.2 Electromagnetic versus hadronic showers

The identification of electrons rely on the differences between the calorimeter showers initiated by the electrons and the showers initiated by hadrons. The different selection variables will be discussed in the following. Some general characteristics of electromagnetic and hadronic showers will be discussed first. Table 3.1 shows a summary of these characteristics and a detailed description can be found in [61].

High-energy electrons and photons crossing a dense material, like for instance a crystal of the CMS calorimeter, will interact mainly through Bremsstrahlung and pair production. Secondary electrons and photons produced in those processes will then start to interact again through Bremsstrahlung and pair production. This cascade of interactions will develop as a shower in the crystal. Figure 3.2 shows a simulation of an electromagnetic shower in a crystal of the CMS electromagnetic calorimeter. The electromagnetic shower shape scales, to a good approximation, longitudinally with the radiation length, X_0, and laterally with the Molière radius. These variables are material dependent[3]. The shower maximum is reached at a depth t_{max}, which depends logarithmically on the initial energy. After this maximum, the number and energy of the secondaries start to decrease, until the average energy per particle becomes low enough to stop further multiplication. This critical energy, ε, defined as the energy at which the rate of energy loss per radiation lengths equals the total energy of the electron, is about 550 MeV/Z. A 35 GeV electron in the CMS calorimeter has about 80% of its energy contained in a 3×3 crystal array.

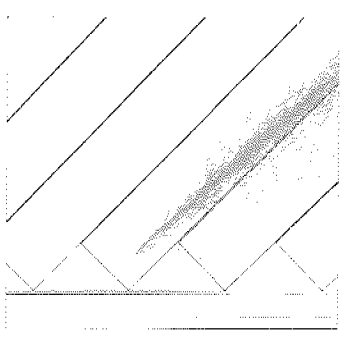

FIGURE 3.2: *Simulation of an electromagnetic shower in a crystal of the CMS ECAL (CMS-ECAL, TDR[21]).*

A shower initiated by a hadron is dominated by a succession of inelastic interactions and is thus more complex to describe. The characteristic dimensions of hadronic showers are determined by the nuclear absorption length, λ [4]. At high energy, these showers are characterized by multi-particle production and nuclear physics effects associated with the excitation of the absorber nuclei. A fraction of the available energy is converted into excitation or breakup of the nuclei, part of which will be undetectable. Moreover π^0s are relatively frequent inside showers initiated by hadrons, bringing an electromagnetic component to the shower. The average fraction of a hadronic shower converted into π^0s mainly depends on the nature of the first inelastic interaction. Hadronic showers have a large shower-to-shower fluctuation and are expected to be

[3]For the $PbWO_4$ crystals of the CMS calorimeter, the Molière radius is 2.19 cm and the radiation length, 0.89 cm.

[4]For the $PbWO_4$ crystals of the CMS calorimeter, the nuclear absorption length is about 23.6 cm.

Quantity	Electromagnetic showers	Hadronic showers
Mean free path	$9X_0/7$ for γ's, $X_0 \approx 180A/Z^2$ [g·cm^{-2}]	$\lambda \approx A/(N_{Avogadro} \cdot \sigma) \propto A^{1/4}$
Secondary particles	$e^{\pm}$, γ; below critical energy $\varepsilon \approx 550$ MeV/Z ionization loss only; inelasticity $\kappa = 1$ (all energy used in particle production)	Fast nucleons, pions; medium energy ($\sim$ 100 MeV): p, n; low-energy ($\sim$ 10 MeV): p, n, γ; nuclear fragments; inelasticity $\kappa \sim 0.5$
Shower maximum	$t_{max}[X_0] \approx \ln(E/\varepsilon) - \alpha$, ($\alpha \approx 1$ for e, ≈ 0.5 for γ)	$t_{max}[\lambda] \approx 0.6 \ln(E[GeV]) - 0.2$
Depth for $\approx 95\%$ longitudinal containment	$L_{0.95}[X_0] \approx t_{max} + 0.08Z + 9.6$	$L_{0.95}[\lambda] \approx t_{max} + 4E^{0.15}[GeV]$
Radius for $\approx 95\%$ radial containment	$R \approx 2\rho_M \approx 14A/Z$ [g·cm^{-2}]	$R \approx \lambda$

TABLE 3.1: *Average properties of electromagnetic and hadronic showers [61]. For a 100 GeV electron in lead, the shower maximum is at about 8.6 X_0 and $L_{0.95} = 24.8\ X_0$ and for a 100 GeV π^+, the shower maximum is at about 2.6 X_0 and $L_{0.95} = 10.6\ X_0$.*

more spread than electromagnetic showers. In fact, electromagnetic particles loose their energy mainly through Bremsstrahlung and pair production for which the angle between the primary and secondary particles is small. On the contrary, for hadronic showers, when the hadrons initiate nuclear interactions, the secondary particles coming from the decay of a high mass system will receive transverse energy and thus have a different direction than the primary particle. This will cause the shower to be broader.

3.3 ORCA reconstruction of electrons, the EGAMMA software

Within ORCA, EGAMMA is providing tools currently developed to reconstruct electrons and photons. Combining the calorimeter and tracking information, it aims at reconstructing the electron energy and position. Details about the EGAMMA package are given in [62].

To reconstruct electrons one first has to look for high energy deposits in a single crystal of the calorimeter, which is called a "seed". The energy contained in the neighboring crystals will be added to the seed in order to form "basic-clusters". Two algorithms are implemented in EGAMMA to reconstruct these clusters, the *island* and the *hybrid algorithm*.

For the *island algorithm*, the reconstruction of basic-clusters goes as follows: First of all, "seed" crystals are defined and ordered in ascending energy. The transverse energy threshold for a crystal to be used as a seed is per default at 500 MeV for the barrel and 180 MeV for the endcaps[5]. Then, starting with the most energetic seed crystal, the energy of the neighboring crystals in both directions in ϕ and η from the seed position, will be collected until a crystal without energy or a crystal with more energy than its neighboring one is found. Double counting is prevented since all the collected crystals are marked as belonging to a given cluster and cannot be used any more.

The *hybrid algorithm* has been developed for electrons in the barrel, with $E_t > 10$ GeV [62]. Starting from a seed crystal (the cut for a crystal to be a seed is set here at $E_t > 1$ GeV), *dominoes* of 1×3 crystals in $\eta \times \phi$ (or 1×5 if the energy in the seed crystal is higher than $E > 1$ GeV) are created, each with their central crystal aligned in η with the seed crystal. As Figure 3.3 shows, dominoes are created for each crystal closer than 10 crystals from the seed in

[5]Note that a transverse energy of 180 MeV at a rapidity of 1.7 corresponds to a total energy of about 500 MeV.

ϕ. An adjacent domino is added to the cluster if its energy is higher than E >100 MeV and a non-adjacent domino if its energy is higher than E >350 MeV.

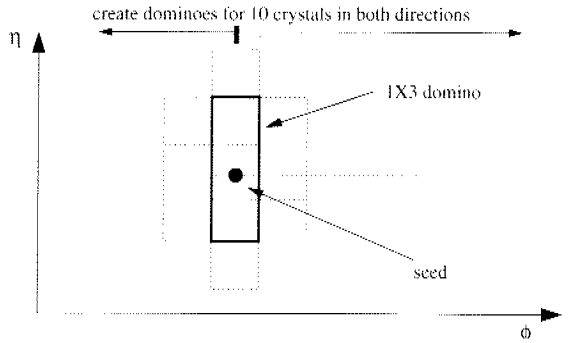

FIGURE 3.3: *The principle of the island algorithm.*

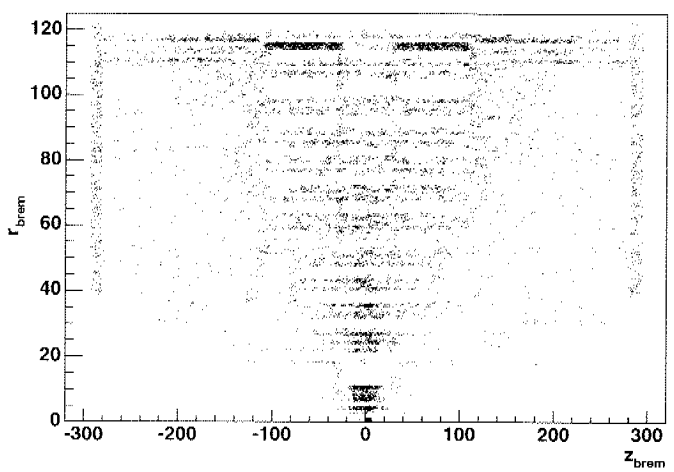

FIGURE 3.4: *The position, given in cylindric coordinates (r, z), where the electron emitted a Bremsstrahlung photon in the tracker. It is a good indicator of the material location in the tracker.*

The *hybrid algorithm* exploits the fact that in an axial magnetic field, the electromagnetic shower will be mainly extended in ϕ and remain narrow in η. Moreover due to the amount of material before the calorimeter, the electrons can radiate part of their energy through Bremsstrahlung. Figure 3.4 shows the different positions given in cylindric coordinates[6] (r, z) at which the electrons emitted a Bremsstrahlung photon. An electron loosing its energy through Bremsstrahlung will have a shower extended in ϕ due to the photons emitted.

To recover the energy lost in the tracker through Bremsstrahlung, *basic-clusters* are associated together to create *super-clusters*. Two basic-clusters are merged if they lie in a narrow window in η and a wider window in ϕ, more precisely if $\Delta\phi < 0.8$ for clusters in the barrel (0.4 for clusters in the endcaps) and $\Delta\eta < 0.06$ for clusters in the barrel (0.14 for clusters in the endcaps). The energy threshold for a super-cluster to be kept is set at 4 GeV.

The simulated energy loss of electrons in the tracker material is given by GEANT. The ratio between the total energy radiated by the electron through Bremsstrahlung and the reconstructed

[6] z is taken parallel to the beam pipe direction.

energy of the super-cluster, E_{brem}/E_{SC} will be used in the following to study the Bremsstrahlung impact on the electron reconstruction and identification.

Finally, the shower position is corrected to take into account the non-pointing geometry of the crystals. The super-cluster energy is corrected in order to get the ratio E_{meas}/E_{true} peaking at 1, using currently a function which depends on the number of crystals in the cluster.

Tracks which will be combined with the super-clusters to distinguish between electrons and photons, are reconstructed using a standard algorithm based on a Kalman filter (named "Combinatorial track finder") [7].

3.4 Datasets and software versions

The list of samples used for this study is given in Table 3.2. They are divided in two main groups. The first group is composed of isolated mono energetic particles, called "study samples". These samples are useful to concentrate on a particular aspect of the reconstruction without worrying about effects linked to the kinematics of a given process. The second group of samples consists of physics processes. These samples give a more realistic simulation of the CMS detector in the LHC environment. In contrast to the other samples, the $H \rightarrow ZZ \rightarrow 2e2\mu$ and the $WZ \rightarrow 3\ell$ samples are generated including pile-up.

Except for the diphoton sample based on OSCAR/GEANT-4, the simulation of the particle interaction with the detector elements was based on CMSIM/GEANT-3. It has been shown previously that CMSIM and OSCAR give similar results [63].

Digitized information (*Digis*) were processed running ORCA together with PAX. Due to the rapid version changes of ORCA, different versions of the software had to be used. However, the results of the study presented here do not depend on the version used.

Sample	Number of events	ORCA version	Detector simulation
Study samples (mono-energetic, isolated particles)			
Dielectrons, $E_t = 35$ GeV	99000	7_5_2	CMSIM
Diphotons, $E_t = 35$ GeV	97500	7_5_2	OSCAR, no tracker
Single charged pions, $E_t = 100$ GeV	90000	7_5_2	CMSIM
Physics samples			
$H \rightarrow ZZ^* \rightarrow 4e$, $m_H = 150$ GeV	32000	7_5_2	CMSIM
$H \rightarrow ZZ \rightarrow 2e2\mu$, $m_H = 500$ GeV	32000	7_6_1	CMSIM
$WZ \rightarrow 3\ell$ (ℓ : e,μ,τ)	33000	7_6_1	CMSIM
$pp \rightarrow \gamma$ jet	34000	7_6_1	CMSIM
Drell-Yan $m_{e^+e^-} > 1300$ GeV	6000	7_2_2	CMSIM

TABLE 3.2: *Technical details of the samples used. CMSIM means that the detector simulation is based on GEANT-3, and OSCAR means that the detector simulation is based on GEANT-4.*

The following samples were used to study the detector response for various particles and with different energies. In particular:

[7] The 'default' tracks proposed by the EGAMMA group and optimized for trigger studies uses the same reconstruction algorithm than the standard one, but with different parameters. Namely, requiring a χ^2 for the track to be smaller than 5, whereas the standard value is 30 and requiring the tracks to have at least 3 hits whereas the standard value is 5 hits.

- Electrons with a transverse energy E_t of 35 GeV (Dielectron sample), with an average transverse energy, $< E_t >$, of 40 GeV ($H \rightarrow ZZ^* \rightarrow 4e$ sample), with $< E_t >$ of 100 GeV ($H \rightarrow ZZ \rightarrow 2e2\mu$ sample) and with $< E_t >$ of 650 GeV (Drell-Yan $m_{e^+e^-} > 1300$ GeV sample).
- Photons with a transverse energy E_t of 35 GeV (Diphotons sample) and an average transverse energy, $< E_t >$, of 65 GeV ($pp \rightarrow \gamma jet$ sample). Photons are reconstructed vetoing a matching track.
- Pions with a transverse energy E_t of 100 GeV, having a mean super-cluster transverse energy of 80 GeV (Single charged pions sample).
- Taus decaying to hadrons having a mean super-cluster transverse energy of 25 GeV ($WZ \rightarrow 3\ell\,(\ell : e, \mu, \tau)$ sample).
- Jets having a mean super-cluster transverse energy of 30 GeV ($pp \rightarrow \gamma jet$ sample).

The generated transverse energy distribution of the studied particles is shown in Figure 3.5. For the jets and the taus, the reconstructed mean energy of the super-clusters is smaller than the generated one as they do not deposit all their energy in the ECAL.

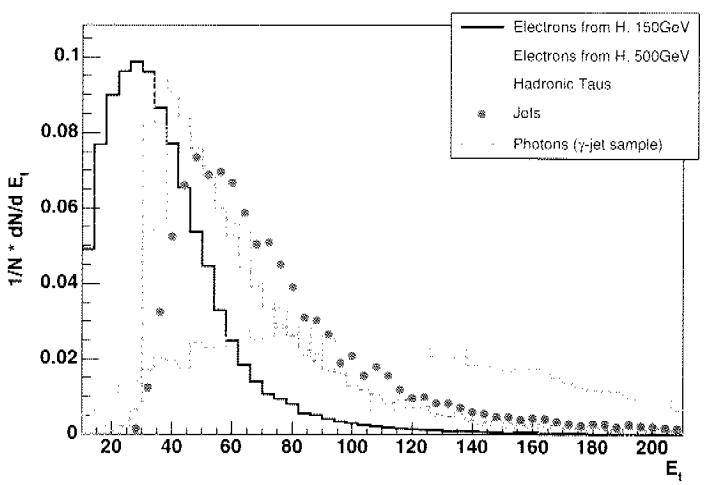

FIGURE 3.5: *The generated transverse energy distribution of the different particles types and for the different samples studied, requiring* $|\eta| < 1.4$.

3.5 Electron identification strategy

An electron is characterized by a narrow cluster in the electromagnetic calorimeter together with a track matching it. A set of selection variables needs to be defined in order to separate electrons from the "electron-looking background". For this purpose, the properties of the electromagnetic showers described before are used to define an efficient electron selection. Guided by the techniques already developed in other experiments like CDF or L3 (which also had a crystal calorimeter), general selection variables for an electron candidate with a transverse energy higher than 10 GeV can be defined.

- *The cluster isolation:*

  Electrons coming from the decay of heavy particles such as the W ($W \rightarrow \ell\nu$), the top ($t \rightarrow Wb \rightarrow \ell\nu b$) or Higgs (e.g. $H \rightarrow ZZ \rightarrow \ell\ell\ell\ell$) decays should be identified in the

detector as isolated high transverse energy clusters. Requiring the electrons candidates to be isolated is particularly efficient to reject the background crated by neutral pions coming from jets.

- *The ratio between the energy deposit in the electromagnetic and hadronic calorimeter:*

  The energy deposited by an electron is almost fully contained in the electromagnetic calorimeter. In contrast, the hadrons will tend to leave most of their energy in the hadronic calorimeter.

- *The shower shape:*

  As explained before, showers initiated by electromagnetic and hadronic particles have different shapes. Electromagnetic clusters are expected to be contained to a large extend in a 3×3 crystals array of the CMS ECAL. In contrast, hadronic clusters are more spread in all directions. Due to the magnetic field and the electrons emitting Bremsstrahlung, the electromagnetic clusters will remain narrow in η but will be somewhat extended in ϕ.

- *The ratio between the energy in the electromagnetic calorimeter and the track momentum:*

  For electrons this ratio is expected to be around 1 within the measurement errors.

- *The precise matching in ϕ between tracks and clusters:*

  Clusters initiated by electrons, have their associated track pointing directly to the energy weighted cluster center. It will be shown, however, that some care is needed to use this variable since it is quite sensitive to electron Bremsstrahlung.

3.6 Selection variables

After this qualitative description, a quantitative analysis of the electron selection is presented. The goal is to define a selection, which is as much as possible independent on the event properties. However, the cuts proposed for a given physics channel should be further tuned depending on the background. For instance, to select a Higgs decaying into four electrons, one would obviously not apply the same cuts on all four electrons, loosening the requirements once one or two "good quality" electrons are found. The cuts have also to be optimized in order to maximize the signal to background ratio. However, general selection variables can already be defined, allowing to get a basis for an efficient electron selection.

3.6.1 Preselection

First a preselection is applied for the different samples. Only the super-clusters in the barrel (with $|\eta| < 1.4$) and with E_t larger than 10 GeV are kept for the study. To compare the reconstructed electrons with the generated ones, it is required that these super-clusters originate from the interaction of an electron (or the other particles studied) with the calorimeter. A super-cluster is said to be matched, if it lies inside a "large" cone in $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ of 0.15 around the generated particle. The kinematic properties of the generated particles are directly taken from the Monte Carlo generator information.

Sometimes, more than one super-cluster is found within this cone (which happened for instance in about 2% of the cases for the $H \rightarrow ZZ^* \rightarrow 4e$ sample). In this case, the one with the energy closest to that of the generated electron was selected. Figure 3.6 shows the ΔR distribution between the generated electron and all super-clusters in the event for the $H \rightarrow ZZ^* \rightarrow 4e$ sample (left) and the single electrons sample (right). In contrast to the single electron sample, the ΔR distribution for the $H \rightarrow ZZ^* \rightarrow 4e$ sample shows a tail. These high values of ΔR come from 'wrong combinations' with the other super-clusters present in the interaction.

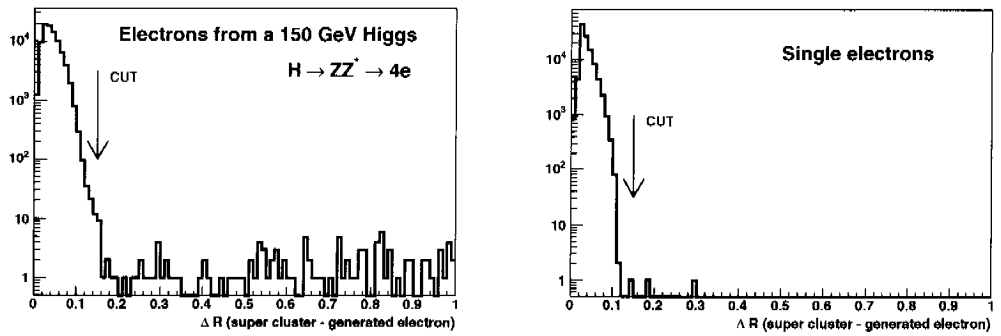

FIGURE 3.6: *The smallest ΔR between a generated electron and all reconstructed super-clusters in the $H \rightarrow ZZ^* \rightarrow 4e$ sample (left) and in the single 35 GeV electron sample (right).*

For 0.4% of the generated electrons, no matching super-clusters were found. Figure 3.7 shows the pseudorapidity distribution of the generated electrons were no super-clusters were found for the $H \rightarrow ZZ^* \rightarrow 4e$ sample. Most of these electrons were 'lost' in a crack of the calorimeter. In that case, either no basic-cluster could be reconstructed or a basic-cluster was reconstructed with too little energy[8]. The effects of the gaps and a study of the homogeneity of the electromagnetic calorimeter is discussed in more detail in Section 3.7.

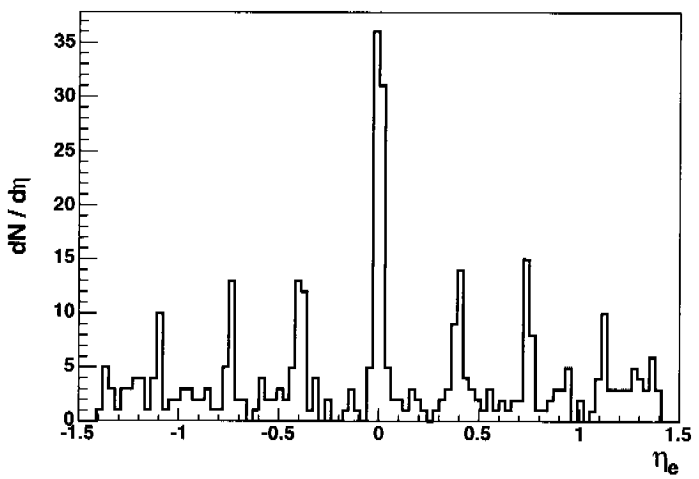

FIGURE 3.7: *Pseudorapidity distribution of the generated electrons for which no matching super-cluster was found (0.4% of the generated electrons). Electrons from the $H \rightarrow ZZ^* \rightarrow 4e$ sample ($m_H = 150$ GeV) were used.*

The second part of this preselection consists in requiring that the super-cluster is in spatial coincidence with a reconstructed track. The tracks should fulfill loose quality requirements, i.e. having a transverse momentum of at least 5 GeV and at least 5 hits (using the silicon strip and pixel tracker information). To associate a super-cluster with a track, the same method as before is applied. For 94% of the electrons, a matching track was found. Figure 3.8 shows the rapidity distribution of the super-clusters with and without matching track. As expected, track finding is less efficient for higher pseudo-rapidities. The matching between tracks and super-clusters will be investigated in detail in the following and used as a selection variable.

[8]To be kept, a super-cluster is required to have an energy above 4 GeV.

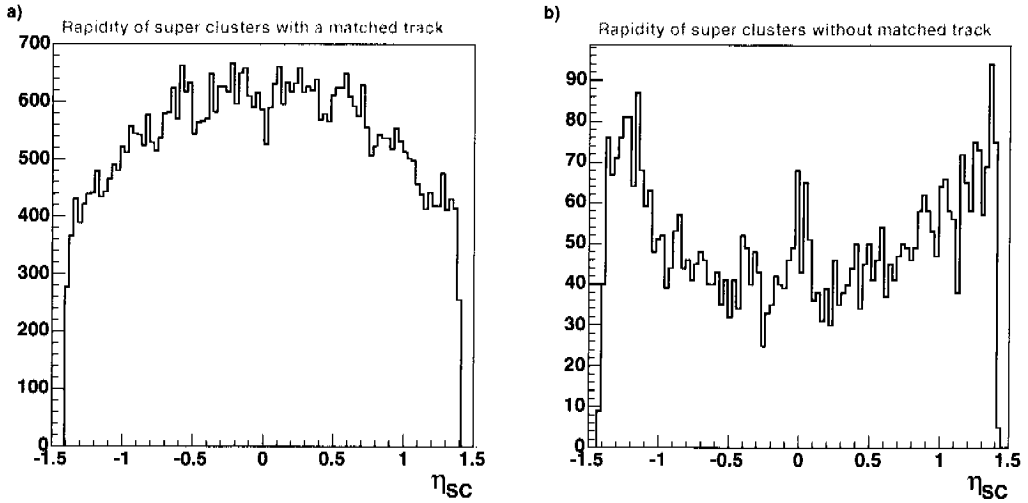

FIGURE 3.8: *Pseudorapidity distribution of the super-clusters with (a) and without (b) matching tracks (94% of the electrons have a matching track).*

The different electron selection variables can now be studied starting from these preselected samples.

3.6.2 The cluster isolation

The isolation can be defined using the tracking or/and the calorimeter. A possible definition of a calorimeter-based isolation could be to consider the energy in the ECAL, lying in a cone of defined size around the super-cluster seed. However, as electrons loose energy through Bremsstrahlung they will not be perfectly isolated. It is then impossible to determine whether an energy deposit in the electron vicinity was caused by photons emitted through Bremsstrahlung or whether this energy deposit was caused by a particle belonging to a hadronic shower.

For this purpose, we chose to characterize the isolation using the tracker only and define isolation by taking the sum of the transverse momentum of all the tracks which lie inside a cone of 0.35 in Δr around the super-cluster position divided by the super-cluster transverse energy measured in the calorimeter:

$$ISO_{trk} = \frac{1}{E_t^{SC}} \left(\sum_{\Delta r_{tr-SC}<0.35} p_t^{\text{track}} - p_t^{\text{matched trk}} \right)$$

The track matching the super-cluster is not taken in the sum, as it is supposed to come from the electron. In the following this variable will be used.

The efficiency of a cut on the isolation is almost independent on the event type. Figure 3.9a shows ISO_{trk} for Higgs events with four electrons, from $H \rightarrow ZZ^* \rightarrow 4e$, m_H =150 GeV (solid line), with two electrons, from $H \rightarrow ZZ \rightarrow 2e2\mu$, m_H =500 GeV (dashed line) and for photons, from $pp \rightarrow \gamma$jet sample (dashed dotted line). As expected, the electrons are less isolated if the final state topology is more complicated.

Figure 3.9b shows the same variable for jets from $pp \rightarrow \gamma$jet sample (dash-dotted line) and taus decaying into hadrons, from $WZ \rightarrow 3\ell$ sample (solid line). As the hadronization process results normally in multi-particle production, most of the time jets are not isolated. The super-clusters initiated by jets are then less isolated than the ones from electrons and photons. On the contrary, the tau decay products will be isolated, since 50% of the taus decay into one charged hadron (and several π^0s) and about 15% decay in three charged particles (and several π^0s). Therefore, taus decaying into hadrons are often called "isolated jets". For tau jets with

$E_t > 50$ GeV, about 90% of their energy is contained in a very small region in (η, ϕ) space of radius 0.15 to 0.2, and about 98% in a radius of 0.4, a property used for instance for the tau selection by the High Level Trigger. The algorithm looks for isolated high E_t hadrons in the calorimeter ($\Delta R = 0.13$) surrounded by an isolation region ($\Delta R = 0.4$) [64].

The isolation cut is found to be almost independent on the electron transverse energy. A cut requiring ISO_{trk} to be less than 0.2 could be applied to select isolated electrons.

3.6.3 The ratio between the energy in the hadronic and electromagnetic calorimeter

The ratio between the energy reconstructed in the hadronic calorimeter and in the electromagnetic calorimeter, E_{had}/E_{em} is quite different for electrons and hadrons. Figure 3.10 shows the E_{had}/E_{em} distribution for: (a) electrons with a mean E_t of 40 GeV (solid line), with a mean E_t of 100 GeV (dashed line) and for photons (dash-dotted line); and (b) for jets (dash-dotted line), taus decaying into hadrons (solid line) and single charged pions (dotted line). As expected, the E_{had}/E_{em} ratio is in general close to zero for electromagnetic particles. It is higher for hadrons, which tend to deposit a large fraction of their energy in the HCAL.

This variable is almost independent on the electron energy since the distributions for the electrons with a mean E_t of 40 GeV and 100 GeV look similar. Jets and taus decaying into hadrons show also a similar E_{had}/E_{em} distribution, unlike the distribution for 100 GeV charged pions which looks much flatter. The jet hadronization and the tau decay will produce π^0's which will increase the energy fraction left in the ECAL. For the 100 GeV charged pions, the energy fraction left in ECAL will depend on where the first interaction starts, leading to a flat distribution of E_{had}/E_{em}.

A possible cut value would be to require E_{had}/E_{em} smaller than 0.05.

3.6.4 The shower shape

Selection variables based on the shower shape exploit the fact that electromagnetic showers are more concentric and dense than hadronic ones. The electron energy is almost all contained in a 3×3 crystal array. To characterize the density of the shower, one can compare the energy in crystal arrays of different size: taking the seed only, E_{SEED}, the 2-by-2 more energetic crystal array containing the seed crystal, $E_{2\times2}$, the 3×3 crystal array centered on the seed, $E_{3\times3}$, and the 5×5 crystal array centered on the seed, $E_{5\times5}$. About 90% of the super-cluster energy is contained in $E_{5\times5}$, 80% in $E_{3\times3}$ and 75% in $E_{2\times2}$ for 35 GeV electrons.

The ratio $E_{3\times3}/E_{5\times5}$ will be studied first. Figure 3.11 shows the $E_{3\times3}/E_{5\times5}$ distribution for: (a) electrons with a mean E_t of 40 GeV (solid line), with a mean E_t of 100 GeV (dashed line) and for photons (dash-dotted line); and (b) for jets (dash-dotted line), taus decaying into hadrons (solid line) and single charged pions (dotted line). The distribution is characterized by a peak and a long tail.

Figure 3.12a shows that the $E_{3\times3}/E_{5\times5}$ distribution gets narrower when the energy of the electron is increased: the solid line shows the distribution for all electrons, the dash-dotted line for electron with $E_t > 50$ GeV, the dashed line for electron with $E_t > 100$ GeV and finally the dotted line for electrons with $E_t > 200$ GeV. These distributions show that showers is denser at higher electron energies. This observation can be explained considering the Bremsstrahlung photons emitted by the electrons in the tracker. For lower energies, the electrons are separated from these Bremsstrahlung photons by the magnetic field making the subsequent showers more spread. This leads to a lower $E_{3\times3}/E_{5\times5}$ ratio. At higher energies, the electrons have a smaller curvature in the magnetic field and the shower remains dense.

Figure 3.12b shows the efficiency of the cut requiring $E_{3\times3}/E_{5\times5} > 0.9$, as a function of the generated electron transverse energy. The efficiency starts to be flat for E_t above 150 GeV. It

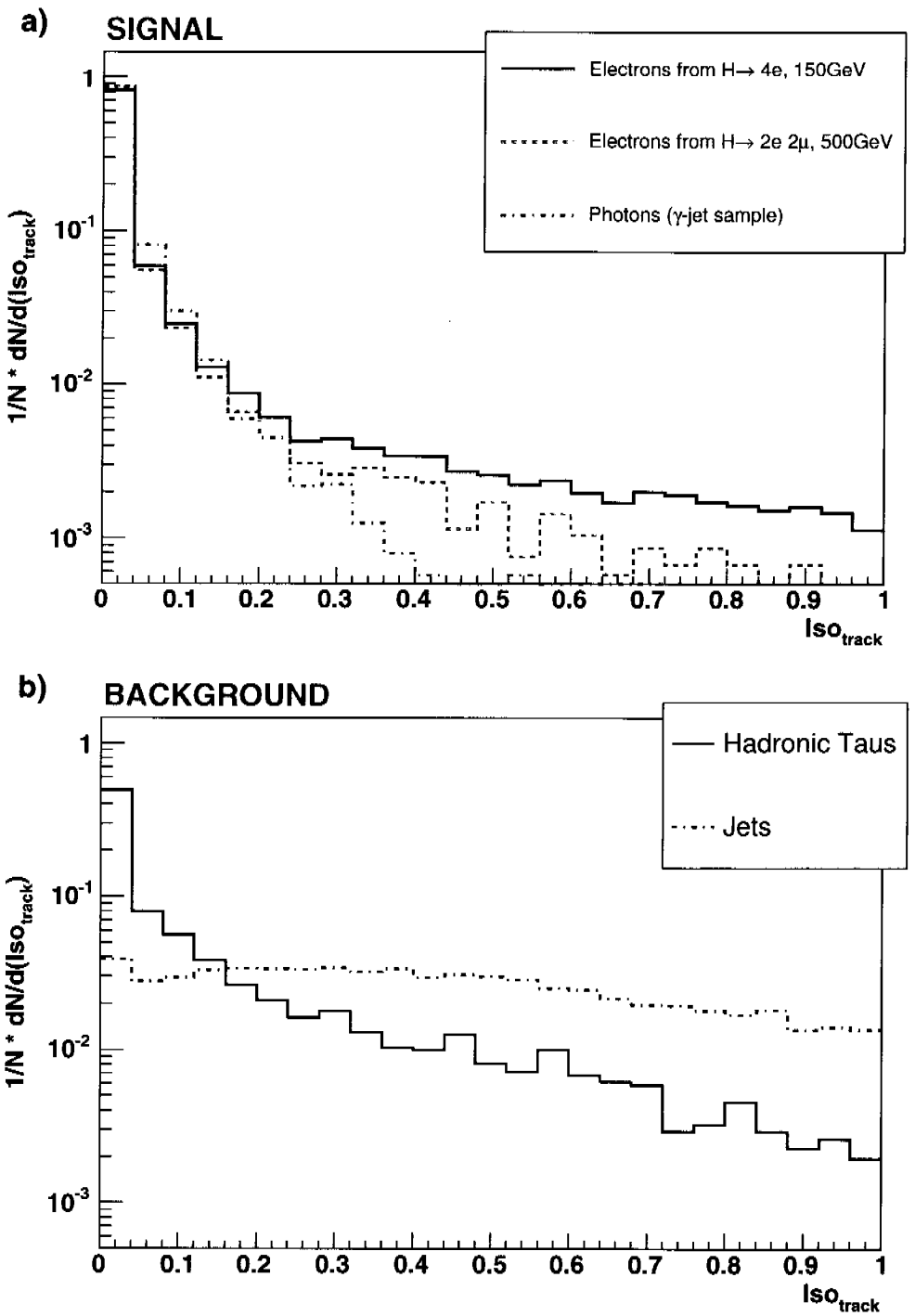

FIGURE 3.9: *The isolation defined with the tracker: (a) for events with four electrons (solid black line, from the $H \rightarrow ZZ^* \rightarrow 4e$ sample), with two electrons (dashed red line, from the $H \rightarrow ZZ \rightarrow 2e2\mu$ sample) and for photons (green dash-dotted line, from the $pp \rightarrow \gamma$jet sample). (b) for hadronic taus (pink solid line, from the $WZ \rightarrow 3\ell$ sample), jets (blue dash-dotted line, from the $pp \rightarrow \gamma$jet sample).*

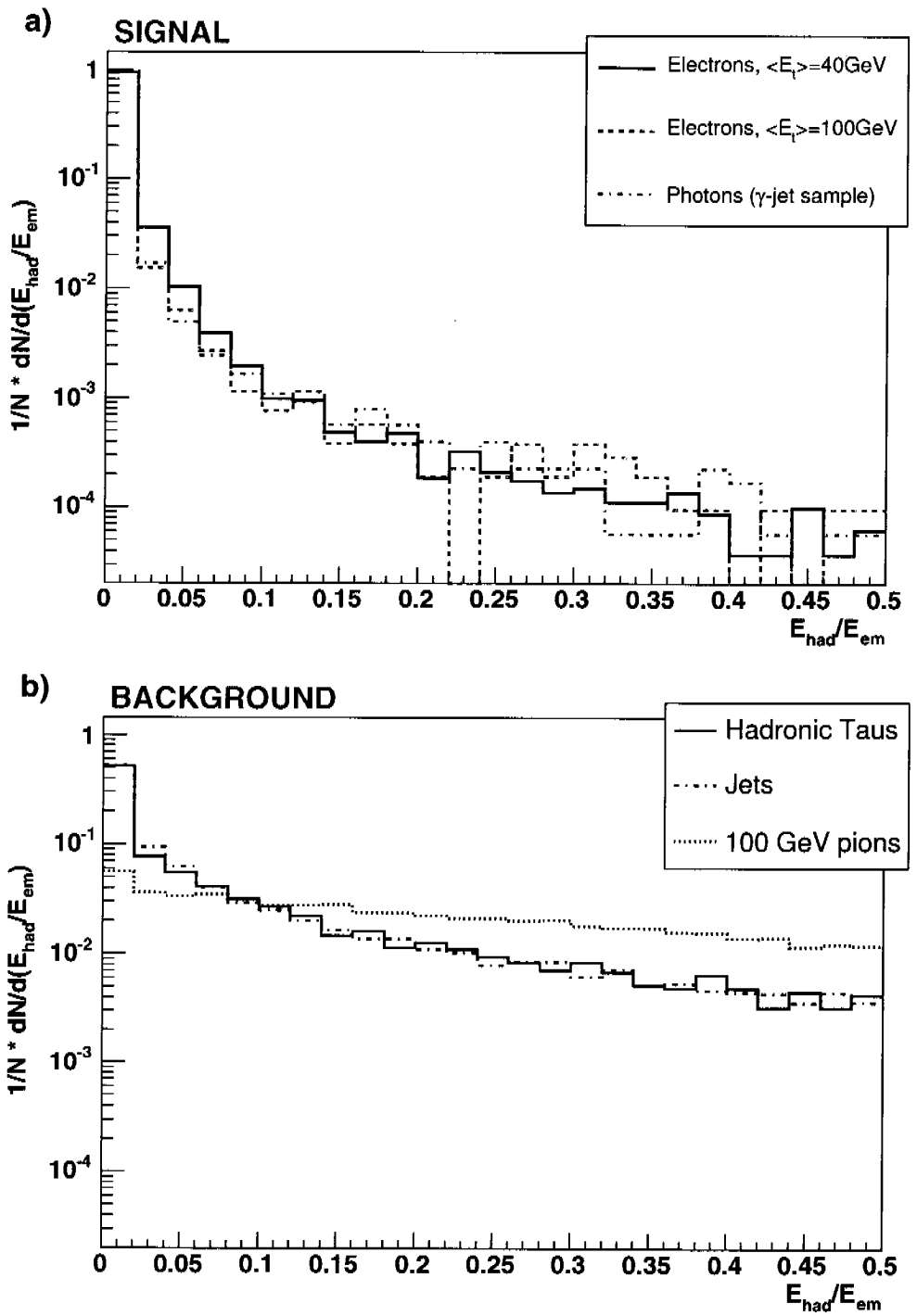

FIGURE 3.10: *(a) E_{had}/E_{em} for electrons with a mean E_t of 40 GeV (solid black line, from the $H \rightarrow ZZ^* \rightarrow 4e$ sample), for electrons with a mean E_t of 100 GeV (dashed red line, from the $H \rightarrow ZZ \rightarrow 2e2\mu$ sample) and photons (green dash-dotted line from the $pp \rightarrow \gamma jet$). (b) E_{had}/E_{em} for hadronic taus (pink solid line), jets (blue dash-dotted line) and single charged pions (black dotted line).*

is thus preferable to use this selection variable only for "high-energy" electrons or alternatively make the cut on this variable energy dependent. This variable is well-suited to select photons.

Other variables exploiting the shower shape were also studied like $E_{2\times2}/E_{SC}$ and $(E_{5\times5} - E_{3\times3})/E_{\text{SEED}}$, a variable sensitive to the energy spread in the crystals with 'less energy'. Figure 3.13 show how these variables look like for the electrons, photons and hadrons. These variables vary also as a function of the electron energy.

An alternative variable, based on the shower shape and independent of the particle transverse energy (at least for E_t ranging between 10 and 300 GeV) is the shower spread in η, $\sigma_{\eta\eta}$. It is defined by making the energy weighted sum of the difference squared between a particular crystal pseudorapidity and the seed pseudorapidity. The sum is taken over all the crystals in a five by five array centered on the seed crystal:

$$\sigma_{\eta\eta} = \sum_{5\times5\,\text{crystals}} (\eta_{\text{crystal}} - \eta_{\text{seed}})^2 \cdot \frac{E_{\text{crystal}}}{E_{5\times5}}$$

Figure 3.14 shows the $\sigma_{\eta\eta}$ distribution (a) for electrons with a mean E_t of 40 GeV (solid line), with a mean E_t of 100 GeV (dashed line) and for photons (dash-dotted line); and (b) for jets (dash-dotted line), taus decaying into hadrons (solid line) and single charged pions (dotted line). The $\sigma_{\eta\eta}$ distributions look very similar for the electron having a mean E_t of 40 and 100 GeV and for the photons. Also the distributions for the different hadronic samples look similar.

The efficiency of the cut requiring $\sigma_{\eta\eta} < 0.0001$ as a function of the generated electron transverse energy is shown in Figure 3.15. The efficiency is found to be almost independent of the electron transverse energy up to a E_t of 300 GeV.

For future studies a similar variable could be defined with the shower spread in ϕ. This is more complicated as the shower spread in ϕ comes not only from the intrinsic shower but also from Bremsstrahlung. To get a 'usable' variable, $\sigma_{\phi\phi}$ could be split in two parts: one where the sum is made on crystals having a ϕ higher than the ϕ of the seed and the other where the sum is made on crystals having a ϕ lower than the ϕ of the seed. Knowing the charge of the track, one can select the one which is on the opposite side of the Bremsstrahlung tail.

3.6.5 The ratio between the energy in the electromagnetic calorimeter and the track momentum

The next two electron identification variables combine the track and calorimeter information. A first variable is the ratio between the energy in the electromagnetic calorimeter and the track momentum, E_{SC}/p_{track}. For electrons that ratio is close to unity. On the contrary, e.g. for taus decaying into hadrons, this ratio is uniformly distributed as the track comes from a $\pi^{\pm}$ and the cluster from a π^0 decaying into two photons.

Figure 3.16 shows the E_{SC}/p_{track} ratio (a) for electrons with a mean E_t of 40 GeV (solid line) and with a mean E_t of 100 GeV (dashed line); (b) for jets (dash-dotted line), taus decaying into hadrons (dotted line) and single charged pions (dotted line).

This variable depends on the electron energy as shown in Figure 3.16a. The peak for 100 GeV electrons is a bit broader for the following reasons. When the electron emits Bremsstrahlung photons, their measured momentum in the tracker tends to be underestimated and the track p_t resolution gets worse at higher energies. On the contrary the super-cluster will tend to contain also the energy from the Bremsstrahlung photons. This will make the E_{SC}/p_{track} ratio higher.

The dotted line on Figure 3.16a shows the electrons with a mean E_t of 100 GeV and an additional cut requiring the electron not to loose too much energy through Bremsstrahlung, $E_{brem}/E_{SC} < 0.5$[9]. In this case, the distribution for higher energies electrons gets closer to the one obtained with 40 GeV electrons (solid line).

[9]The simulated electron energy loss in the tracker is given by GEANT. The ratio between the total energy radiated by the electron through Bremsstrahlung and the reconstructed energy of the super-cluster, E_{brem}/E_{SC} was used here to study the Bremsstrahlung impact on the electron reconstruction and identification.

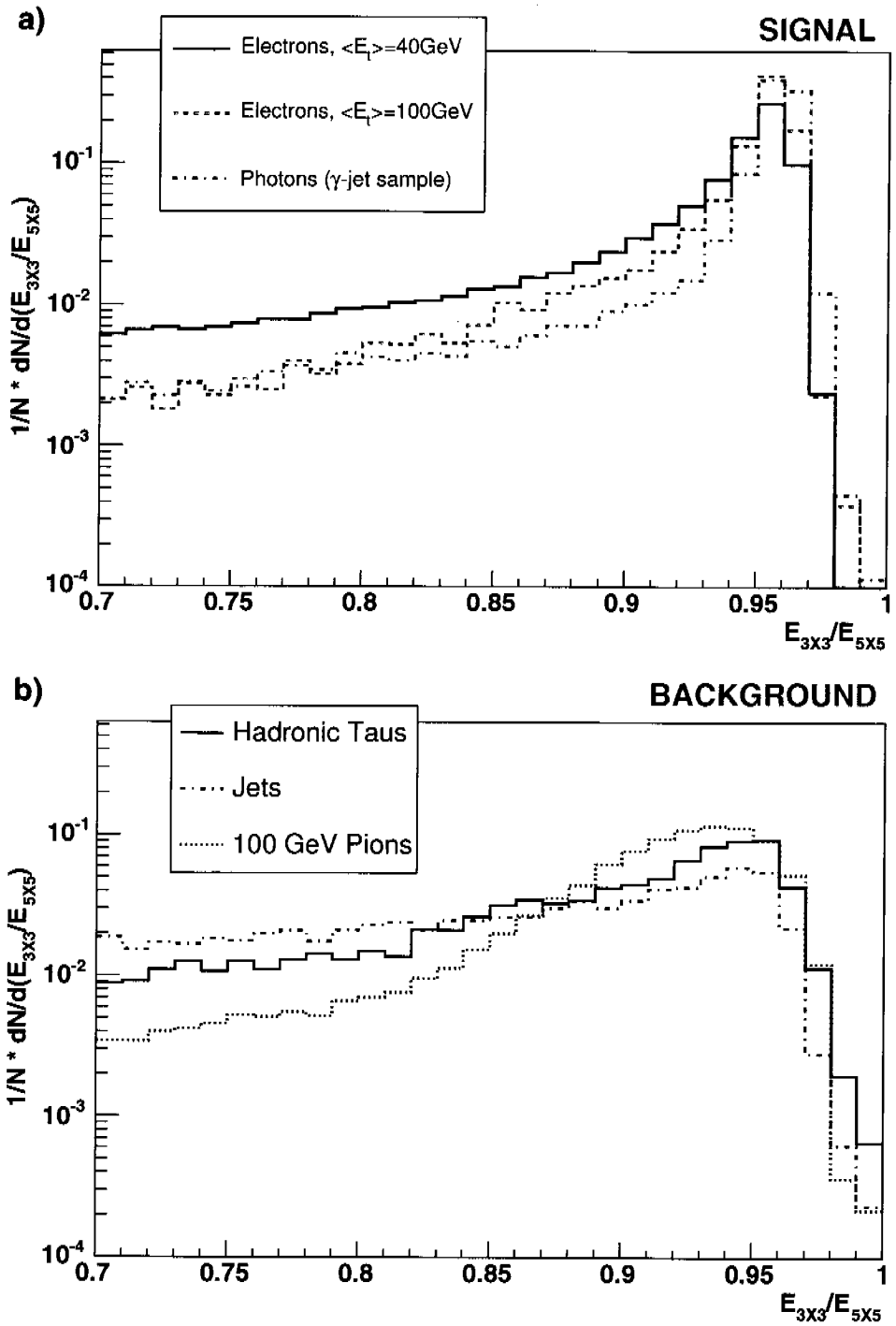

FIGURE 3.11: *(a) $E_{3\times3}/E_{5\times5}$ for electrons with a mean E_t of 40 GeV (solid black line, from the $H \rightarrow ZZ^* \rightarrow 4e$ sample), for electrons with a mean E_t of 100 GeV (dashed red line from the $H \rightarrow ZZ \rightarrow 2e2\mu$ sample) and photons (green dash-dotted line from the $pp \rightarrow \gamma jet$). (b) $E_{3\times3}/E_{5\times5}$ for hadronic taus (pink solid line), jets (blue dash-dotted line) and single charged pions (black dotted line).*

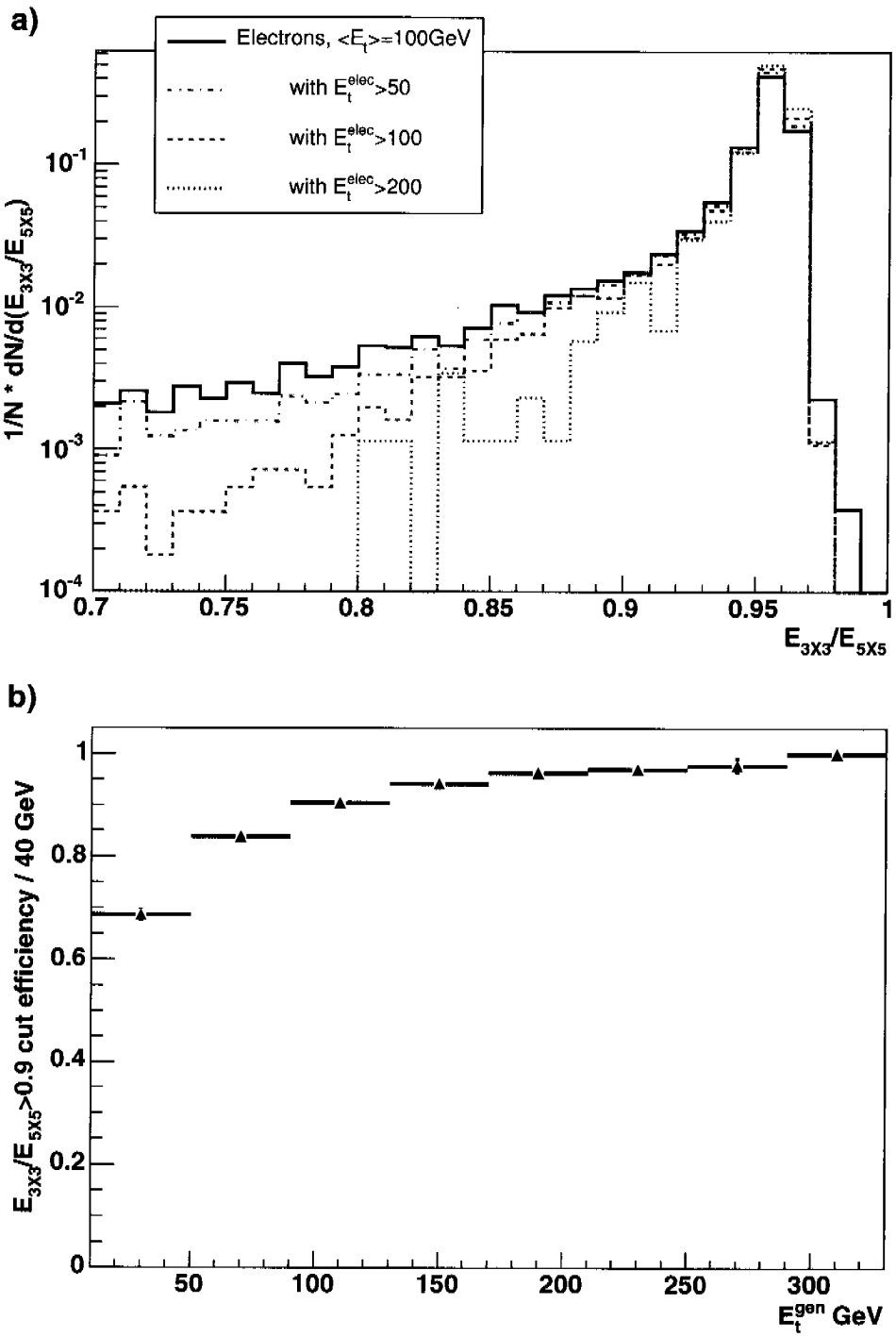

FIGURE 3.12: *(a) $E_{3\times3}/E_{5\times5}$, for the electrons with a mean E_t of 100 GeV. All electrons from this sample (solid black line), electrons with $E_t^{el} > 50$ GeV (red dash-dotted line), $E_t^{el} > 100$ GeV (brown dashed line) and $E_t^{el} > 200$ GeV (brown dotted line). (b) The efficiency of the cut requiring $E_{3\times3}/E_{5\times5} > 0.9$ as a function of the generated electron transverse energy.*

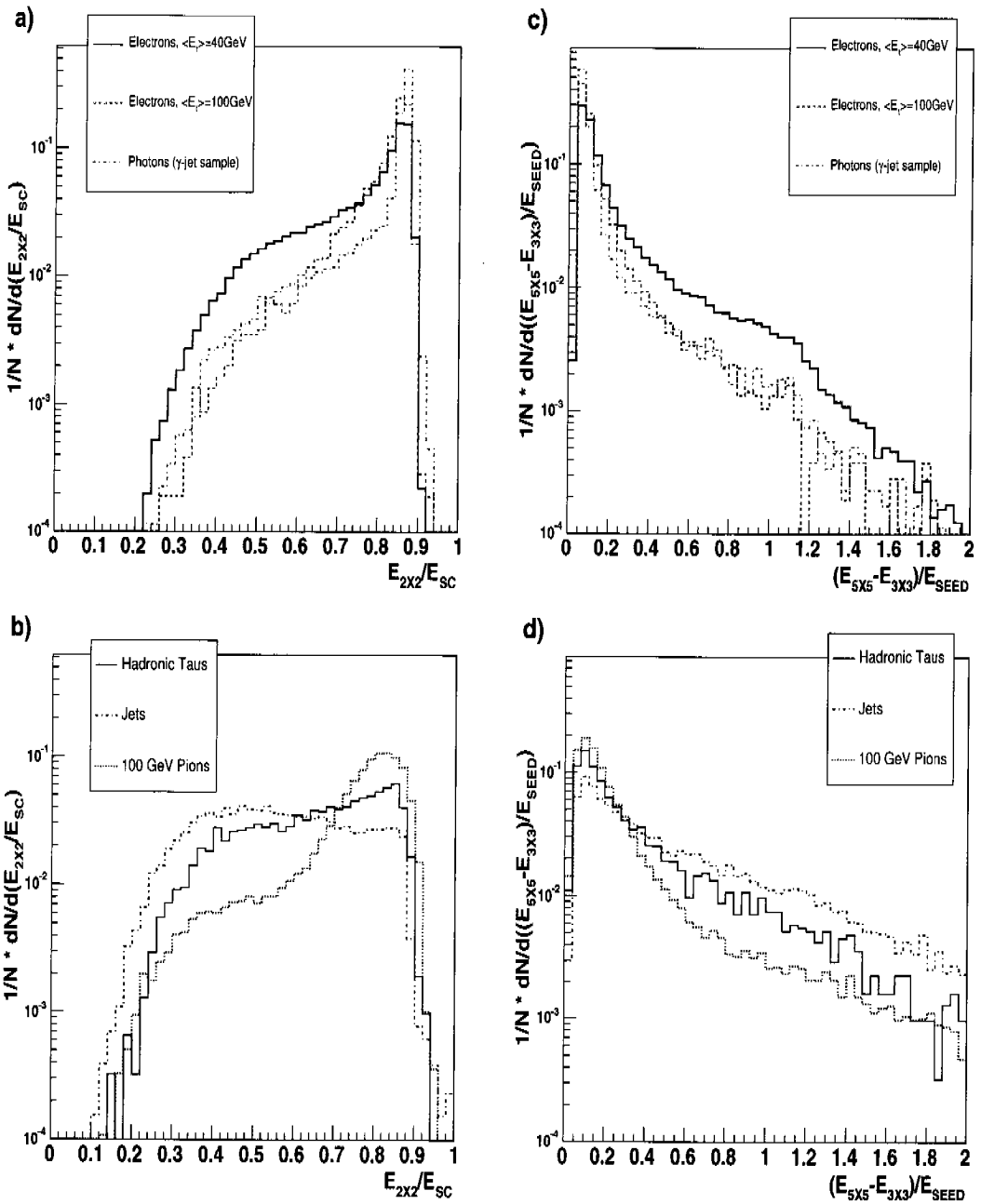

FIGURE 3.13: $E_{2\times2}/E_{SC}$ *(a,b) and* $(E_{5\times5} - E_{3\times3})/E_{\text{SEED}}$ *(c,d). (a and c) for electrons with a mean* E_t *of 40 GeV (solid black line, from the* $H \rightarrow ZZ^* \rightarrow 4e$ *sample), for electrons with a mean* E_t *of 100 GeV (dashed red line from the* $H \rightarrow ZZ \rightarrow 2e2\mu$ *sample) and photons (green dash-dotted line from the* $pp \rightarrow \gamma jet$*). (b and d) for hadronic taus (pink solid line), jets (blue dash-dotted line) and single charged pions (black dotted line).*

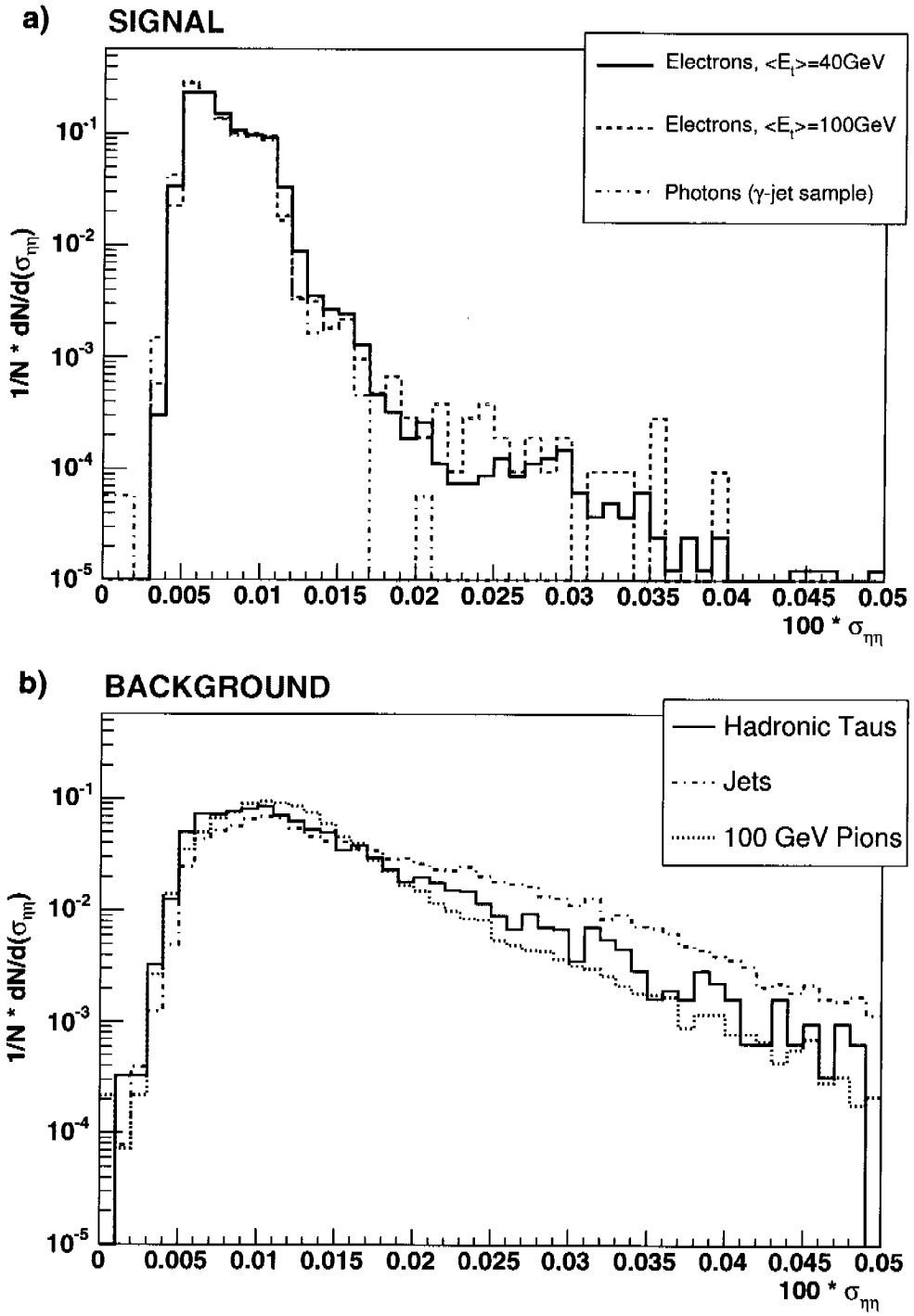

FIGURE 3.14: *The shower extend in η, $\sigma_{\eta\eta}$. (a) for electrons with a mean E_t of 40 GeV (solid black line, from the $H \rightarrow ZZ^* \rightarrow 4e$ sample), for electrons with a mean E_t of 100 GeV (dashed red line from the $H \rightarrow ZZ \rightarrow 2e2\mu$ sample) and photons (green dash-dotted line from the $pp \rightarrow \gamma jet$). (b) for hadronic taus (pink solid line), jets (blue dash-dotted line) and single charged pions (black dotted line).*

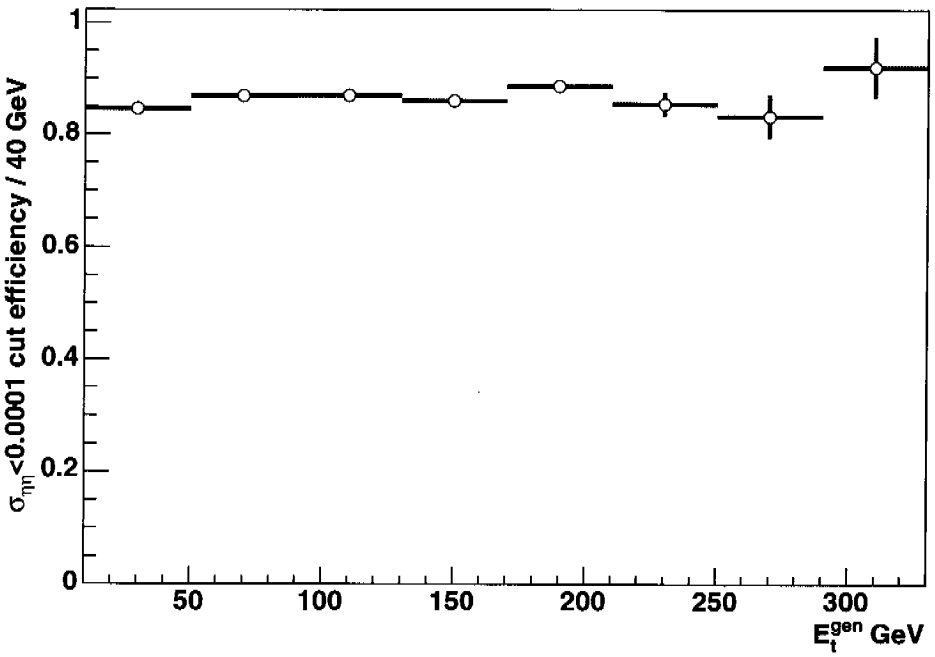

FIGURE 3.15: *The efficiency of the cut requiring $\sigma_{\eta\eta} < 0.0001$ as a function of the generated electron transverse energy on the preselected sample. (No other cuts are applied)*

As Figure 3.16 shows, a cut on the lower part of the E_{SC}/p_{track} distribution (e.g. $E_{SC}/p_{track} > 0.8$) would be efficient against single pions, as they tend to leave only part of their energy in the ECAL while their full momentum is measured by the tracker. To get rid of the jets (mainly their $\pi^{\pm}\pi^0$ component), it is also good to cut on the large values, requiring for instance $E_{SC}/p_{track} < 2$. However, a cut on the higher part of the spectrum will be quite energy dependent. A possibility is then to redefine the variable in the following way: $|1/E_{SC} - 1/p_{track}|$, in order to be less sensitive to high energy electrons. This also makes sense as the uncertainty on p_{track} is Gaussian in $1/p_{track}$.

Figure 3.17 shows $|1/E_{SC} - 1/p_{track}|$ (a) for electrons with a mean E_t of 40 GeV (solid line) and with a mean E_t of 100 GeV (dashed line); (b) for jets (dash-dotted line), taus decaying into hadrons (dotted line) and single charged pions (dotted line). The distribution looks now similar for electrons with a mean E_t of 40 and of 100 GeV. Figure 3.17b shows that a cut requiring $|1/E_{SC} - 1/p_{track}|$ smaller than 0.02 will be efficient against jets and hadronic tau decays but less efficient against single pions.

Figure 3.18a shows the evolution of the cut efficiency for $0.9 < E_{SC}/p_{track} < 2$ as a function of the generated electron transverse energy. Figure 3.18b shows the same but for $|1/E_{SC} - 1/p_{track}| < 0.02$. The cut on $|1/E_{SC} - 1/p_{track}|$ shows a variation of 10% in the efficiency as compared to the cut on E_{SC}/p_{track} which varies 40% between 10 and 1000 GeV. As said before, this efficiency decrease is mainly due to the requirement $E_{SC}/p_{track} < 2$.

A good solution is to cut simultaneously on both variables, requiring $|1/E_{SC} - 1/p_{track}| > 0.02$ and $E_{SC}/p_{track} > 0.9$. That way, the cut efficiency is varying less than 10% between 10 and 50 GeV and is then flat up to $E_t = 300$ GeV. Moreover, such a cut efficiently removes all the three backgrounds. However, some care should be taken since these two variables are not independent.

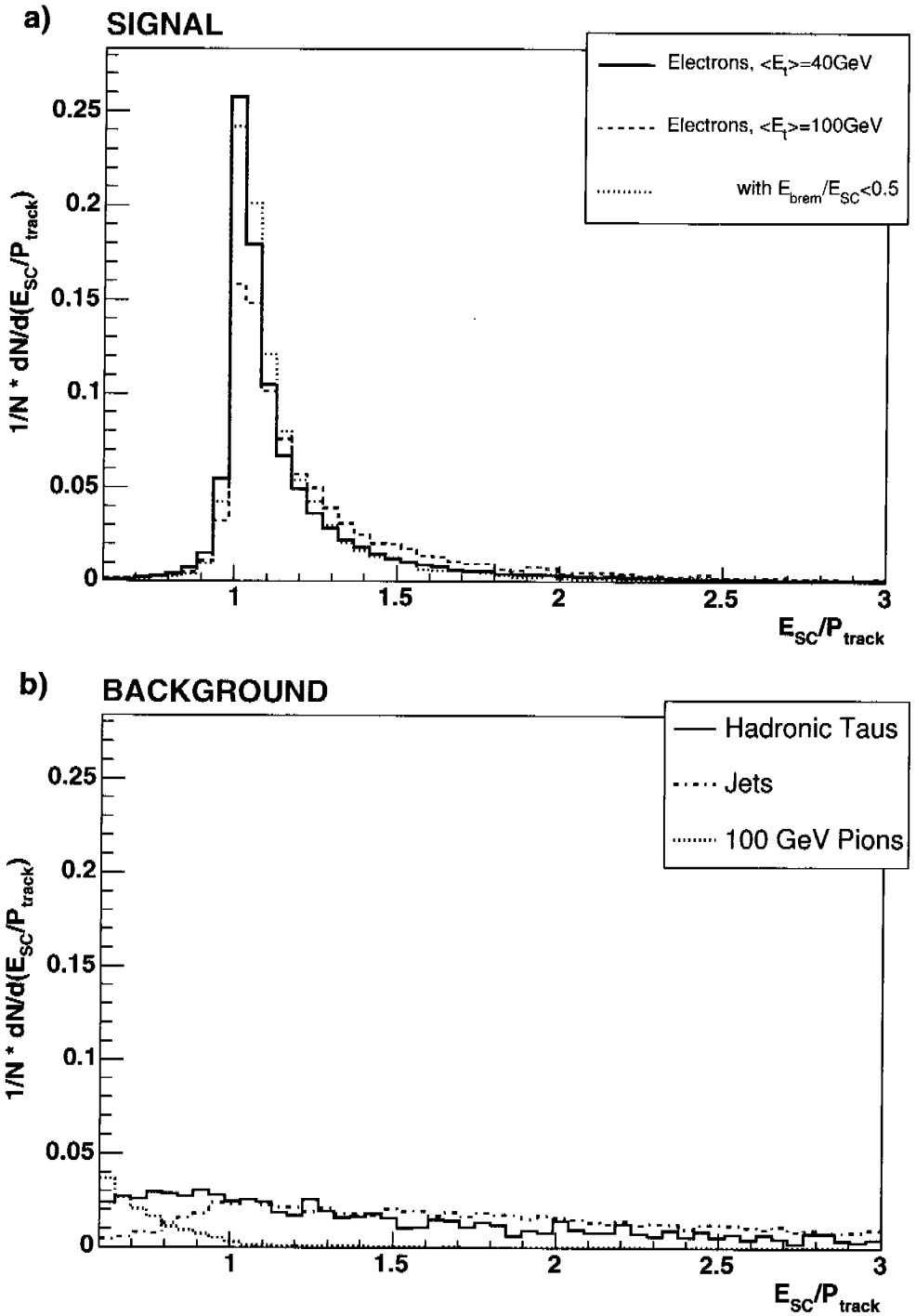

FIGURE 3.16: *The E_{SC}/p_{track} distribution (a) for electrons with a mean E_t of 40 GeV (solid black line, from the $H \rightarrow ZZ^* \rightarrow 4e$ sample), for electrons with a mean E_t of 100 GeV (dashed red line from the $H \rightarrow ZZ \rightarrow 2e2\mu$ sample) and for the electrons with $E_{brem}/E_{SC} < 0.5$ and a mean E_t of 100 GeV (dotted brown line). (b) for hadronic taus (pink solid line), jets (blue dash-dotted line) and single charged pions (black dotted line).*

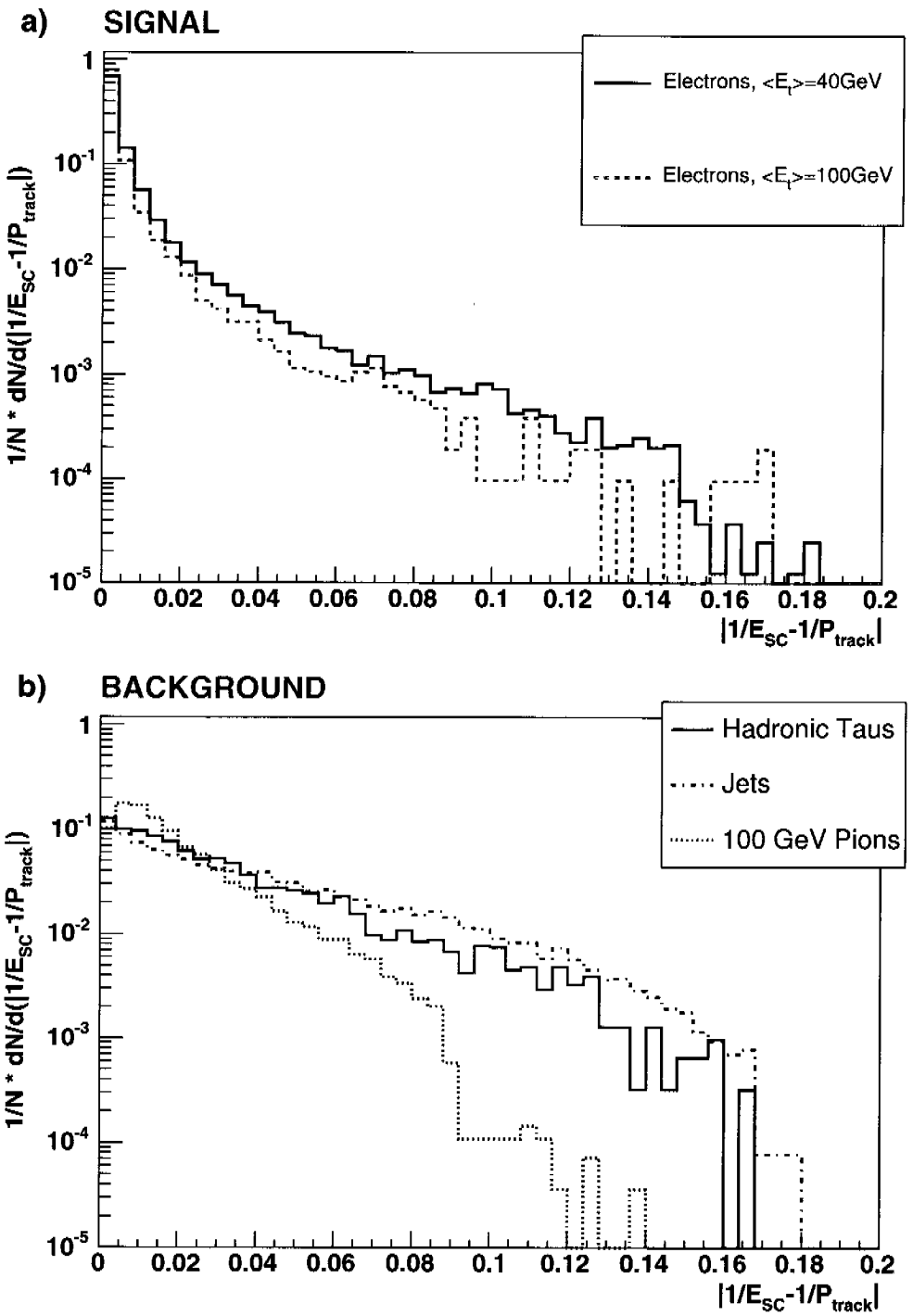

FIGURE 3.17: *The $|1/E_{SC} - 1/p_{track}|$ distribution (a) for electrons with a mean E_t of 40 GeV (solid black line, from the $H \rightarrow ZZ^* \rightarrow 4e$ sample) and for electrons with a mean E_t of 100 GeV (dashed red line from the $H \rightarrow ZZ \rightarrow 2e2\mu$ sample); (b) for hadronic taus (pink solid line), jets (blue dash-dotted line) and single charged pions (black dotted line)*

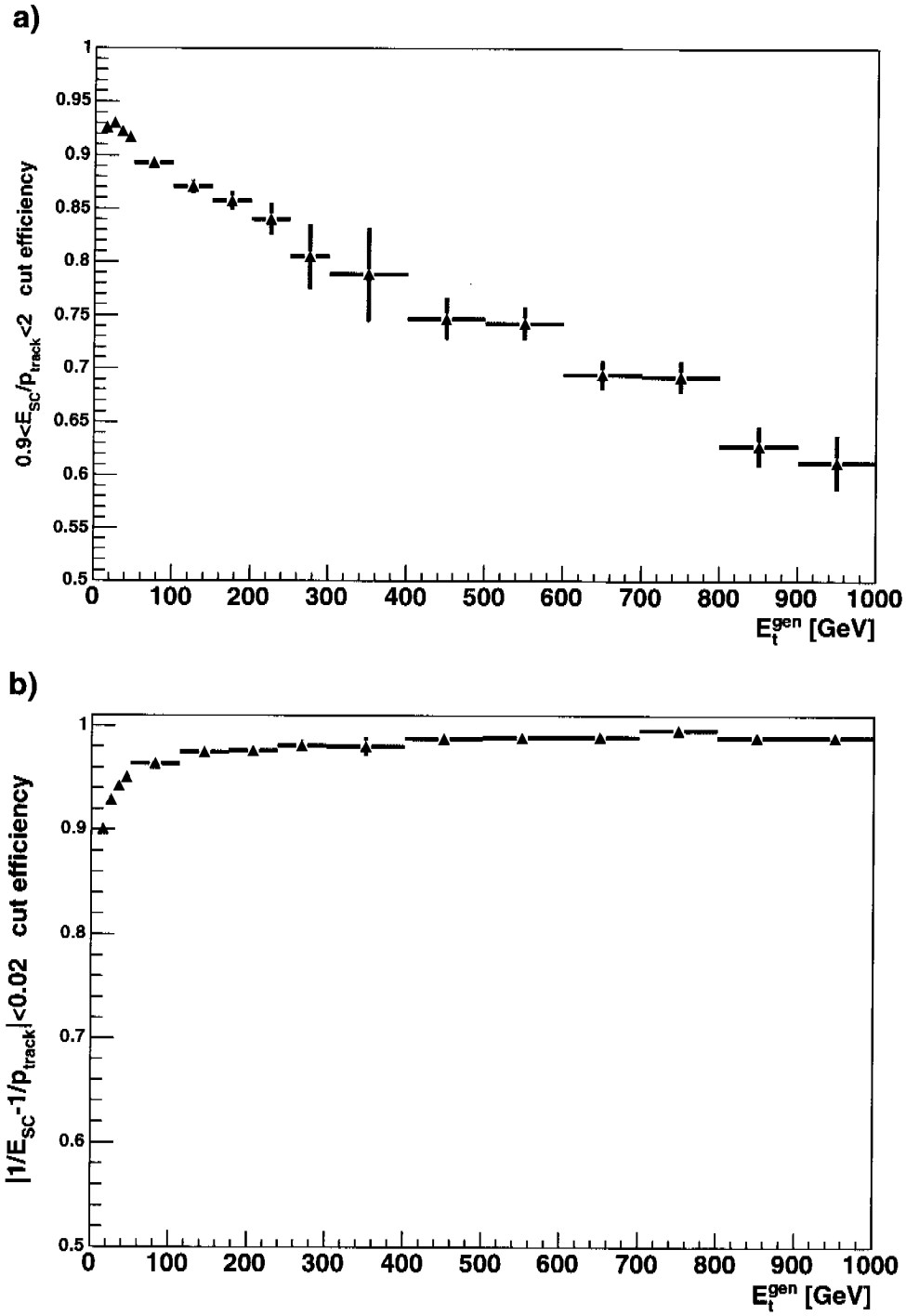

FIGURE 3.18: *(a) The efficiency for the cut:* $0.9 < E_{SC}/p_{track} < 2$ *as a function of the generated electron transverse energy. (b) The efficiency for the cut:* $|1/E_{SC} - 1/p_{track}| < 0.02$ *as a function of the generated electron transverse energy.*

3.6.6 The matching between the track and the cluster

A second variable combining the tracking and the calorimeter is the matching in ϕ between the super-cluster and the propagated track, $|\phi_{SC} - \phi_{tr}^{prop}|$. ϕ_{SC} is the angle of the super-cluster position (which is obtained from the shower shape). ϕ_{tr}^{prop} is the track angle extrapolated from the vertex to the cluster position, taking only the effect of the magnetic field into account. Quite tight cuts can be made on this variable since the tracking angular resolution is very accurate in ϕ, as Figure 3.19 shows.

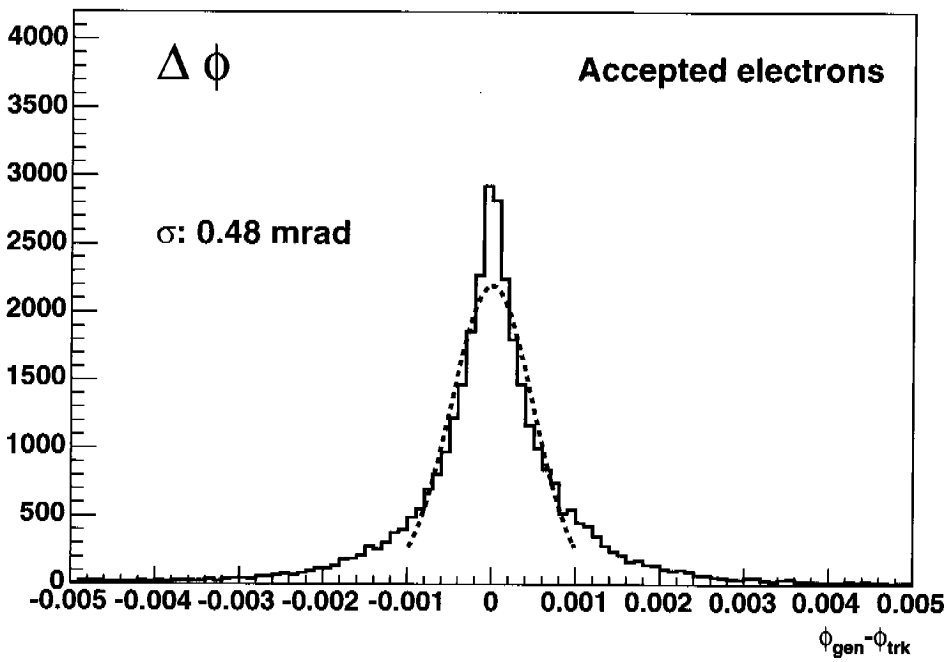

FIGURE 3.19: *Difference between the generated ϕ and the ϕ of the track (taken at the vertex), giving the track resolution in ϕ for electrons having a mean E_t of 40 GeV.*

Figure 3.20a shows $|\phi_{SC} - \phi_{tr}^{prop}|$ for electrons with a mean E_t of 40 GeV (solid line) and with a mean E_t of 100 GeV (dashed line). This variable is sensitive to Bremsstrahlung: when electrons emit Bremsstrahlung photons, their track will be more curved, leading to a larger value of $|\phi_{SC} - \phi_{tr}^{prop}|$, as illustrated in Figure 3.21. This explains the different distributions obtained for the two electron samples in the $|\phi_{SC} - \phi_{tr}^{prop}|$ range between 0.005 and 0.015. If we require $E_{brem}/E_{SC} < 0.1$ for the 40 GeV electrons (dotted line on Figure 3.20a), the distribution obtained gets close to the one obtained for the 100 GeV electrons for this $|\phi_{SC} - \phi_{tr}^{prop}|$ interval. For the higher values of $|\phi_{SC} - \phi_{tr}^{prop}|$, a difference remains. It was checked that this comes from the track momentum resolution. If the track transverse momentum is not well-measured, ϕ_{tr}^{prop} will also have a big uncertainty. This effect is more important for low energy electrons.

Figure 3.20b shows $|\phi_{SC} - \phi_{tr}^{prop}|$ for jets (dash-dotted line), taus decaying into hadrons (solid line) and single charged pions (dotted line). The single pions have a "bump" around $|\phi_{SC} - \phi_{tr}^{prop}| = 0.01$. This comes from the pions that leave only a small part of their energy in the ECAL, creating a mismatch between the track and the cluster. This excess disappears if for instance, a cut requiring $E_{SC}/p_{track} > 0.9$ is applied.

Comparing Figures 3.20a and b, it turns out that cutting on this variable is very efficient for background reduction. Figure 3.22 shows how the efficiency for a cut requiring $|\phi_{SC} - \phi_{tr}^{prop}| < 0.006$ evolves with the electron transverse energy. The efficiency becomes flat for transverse energies above 150 GeV. As expected, high energy electrons will have a higher efficiency than

Variable	Cut value	E_t dependence	Bremsstrahlung dependence	Process dependence
ISO_{TRK}	< 0.2	Small	No	Yes
E_{HAD}/E_{EM}	< 0.05	No	No	No
$E_{3\times3}/E_{5\times5}$	>0.9 for E_t >150GeV	Yes	Yes	No
$\sigma_{\eta\eta}$	< 0.0001	No	No	No
$\lvert\phi^{prop}_{track} - \phi_{SC}\rvert$	<0.006 for E_t >150GeV	Yes	Yes	No
E_{SC}/p_{track}	> 0.9	Yes	Yes	No
$\lvert 1/E_{SC} - 1/p_{track}\rvert$	< 0.02	small	No	No

TABLE 3.3: *Summary of the proposed electron identification criteria. It is also specified whether a particular cut depends on the electron E_t, on the Bremsstrahlung and on the process type.*

lower energy ones.

3.6.7 Electron identification summary: a selection of W events

Table 3.3 presents a possible set of variables to identify electrons in the barrel. The Table summarizes, together with the proposed cuts values, whether the variable is dependent on the event configuration, on the particle transverse energy or on the Bremsstrahlung.

The efficiency of a cut on ISO_{TRK}, E_{HAD}/E_{EM} and $\sigma_{\eta\eta}$ are found to have little variation: less than 10% for electrons with a transverse energy between 10 and 1000 GeV. The energy dependence of the other variables ($E_{3\times3}/E_{5\times5}$, $|\phi^{prop}_{track} - \phi_{SC}|$ and E_{SC}/p_{track}) should be taken into account if used in a selection.

Three variables are sensitive to the Bremsstrahlung: $E_{3\times3}/E_{5\times5}$, $|\phi^{prop}_{track} - \phi_{SC}|$ and E_{SC}/p_{track}. They can be used to discriminate between the electrons which emitted Bremsstrahlung or not.

The cut on ISO_{TRK} will depend on the process type, since the electrons are less isolated if the event has more leptons and jets in the final state.

This electron selection can be applied to select $W \rightarrow e\nu$ events, using the "W calibration sample". Only the events having a central cluster ($|\eta| < 1.4$) with E_t higher than 30 GeV will be studied and the efficiencies will be calculated with respect to the number of these events. Out of these events, 94% were found to have a matching track. The list of the cuts applied, together with their corresponding efficiencies, are summarized in Table 3.4. The electron selection has an overall efficiency of 74%, which is acceptable but could be raised by tunning the electron selection cuts on particular channels.

Apart from these electron identification cuts, some kinematic selection has also to be done. A possible selection consists in requiring the missing transverse energy ($\not\!\!E_T$) to be higher than 30 GeV[10]. To remove the background from the QCD dijets events a jet veto has to be applied, requiring for instance, no jets in the event with an E_t higher then 30 GeV. Finally the W's are required to have a transverse mass (M_T) between 60 and 90 GeV. This selection has a total efficiency of 48%.

The events have also to pass the single electron trigger. In this case, the overall efficiency goes down from 48% to 41%, as shown in the two last columns of Table 3.4. A large fraction, about 35% of the untriggered W events are rejected by the cut requiring $E/p < 1.5$, then 13% of the rejected events fail the High Level Trigger because no matching 'trigger'-track is found and finally the rest of events are mainly rejected due to a bad reconstructed electron, which emitted Bremsstrahlung. Figure 3.23 shows the E_{SC}/p_{track} distribution for the selected electrons (solid

[10]The missing transverse energy is built from the sum of all "jets" in the event. A jet is reconstructed using a cone algorithm with a cone size of 0.5 (called *RecJet-Itercone0.5* in ORCA) and is kept if its E_t is above 5 GeV.

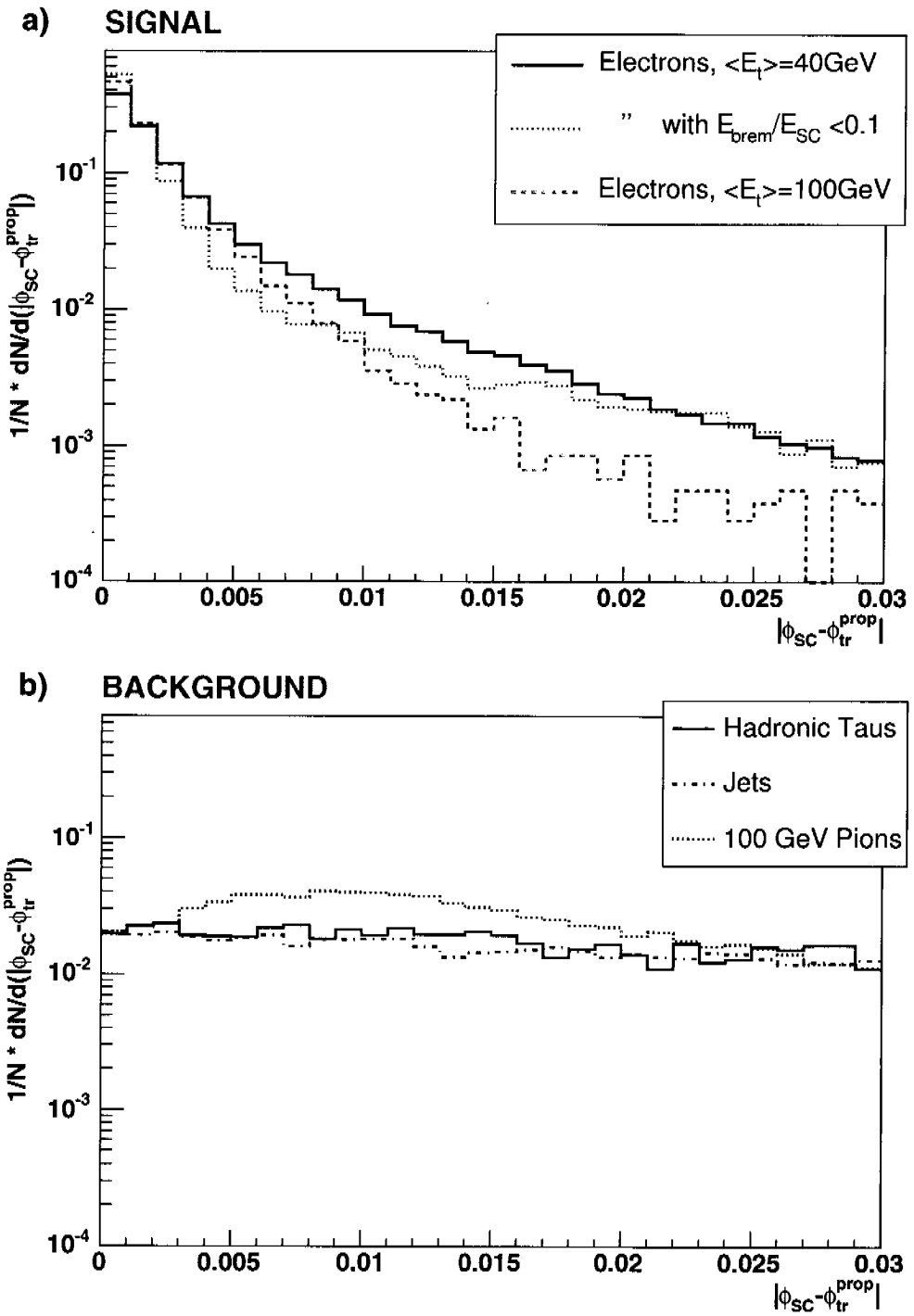

FIGURE 3.20: *(a)* $|\phi_{SC} - \phi_{tr}^{prop}|$ *for electrons with a mean* E_t *of 40 GeV (solid black line, from the* $H \rightarrow ZZ^* \rightarrow 4e$ *sample) and for electrons with a mean* E_t *of 100 GeV (dashed red line from the* $H \rightarrow ZZ \rightarrow 2e2\mu$ *sample) (b)* $|\phi_{SC} - \phi_{tr}^{prop}|$ *for hadronic taus (pink solid line), jets (blue dash-dotted line) and single charged pions (black dotted line).*

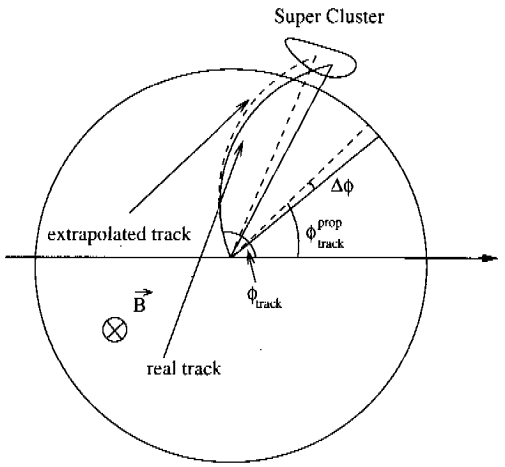

FIGURE 3.21: *To define a matching in ϕ between the track and the cluster, the track angle has to be extrapolated to the cluster position, taking into account the magnetic field. The track angle at the super-cluster position in the ECAL is called ϕ_{tr}^{prop}. $|\phi_{SC} - \phi_{tr}^{prop}|$ is a good variable to determine whether the electron emitted Bremsstrahlung, since in that case, the cluster tend to be more extended in ϕ and the track more curved (marked as the "real track" on the Figure), increasing the difference between the ϕ coordinate of the super-cluster and the track, marked as $\Delta\phi$ on the plot.*

Cut	ε_{tot}	ε_{cut}	ε_{tot} Triggered	ε_{cut} Triggered
Matching track	0.94	0.94	0.94	0.94
Single electron trigger	-	-	0.75	0.80
$ISO_{TRK} < 0.2$	0.91	0.97	0.73	0.97
$E_{HAD}/E_{EM} < 0.05$	0.90	0.99	0.72	0.99
$\sigma_{\eta\eta} < 0.0001$	0.78	0.87	0.64	0.89
$E_{SC}/p_{track} > 0.9$	0.77	0.99	0.63	0.98
$\lvert 1/E_{SC} - 1/p_{track}\rvert < 0.02$	0.74	0.96	0.63	1.00
Kinematic cuts: $\not\!\!E_T$	0.53	0.72	0.45	0.71
jet veto ($E_T^{jet} < 30$GeV)	0.50	0.94	0.43	0.96
$60 > M_T > 90$ GeV	0.48	0.96	0.41	0.95

TABLE 3.4: *The selection of W events. The efficiencies are given with respect to the number of events with a central super-cluster having a transverse energy higher than 30 GeV, for each cut (ε_{cut}) and together with the previous cuts (ε_{tot}). In the two last columns the efficiencies are given for electrons that pass the single electron trigger.*

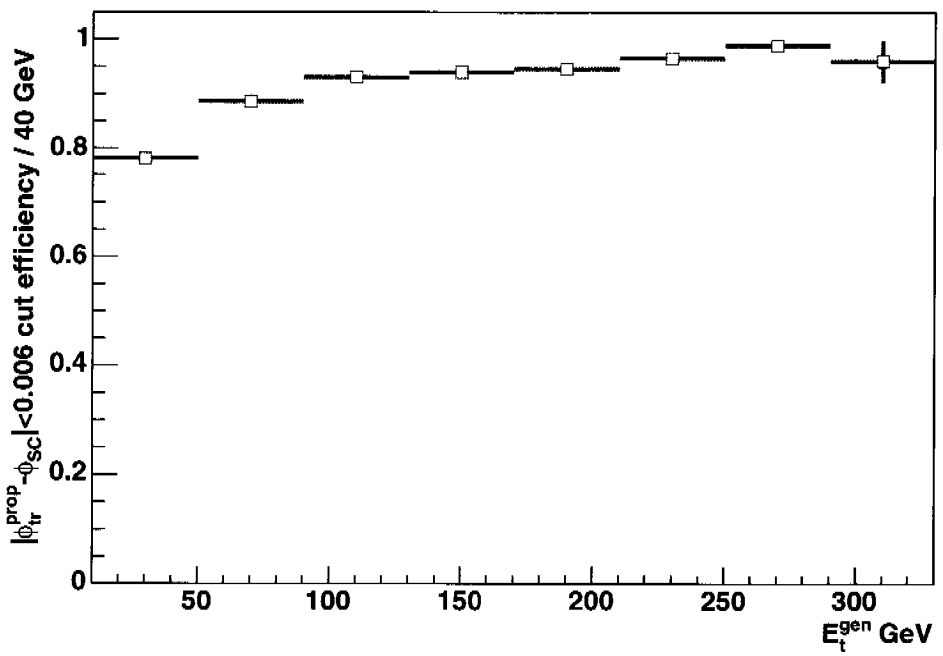

FIGURE 3.22: *The efficiency of the cut requiring* $|\phi_{SC} - \phi_{tr}^{prop}| < 0.006$ *as a function of the electron transverse energy.*

	Single electron (35GeV)		Jets		Single pion (100GeV)		Tau to hadrons	
Cut	ε tot	ε_{cut}	ε tot	ε_{cut}	ε tot	ε_{cut}	ε tot	ε_{cut}
Matching track	0.94	0.94	0.74	0.74	0.99	0.99	0.79	0.79
$ISO_{TRK} < 0.2$	0.94	1.00	0.30	0.40	0.98	0.99	0.65	0.83
$E_{HAD}/E_{EM} < 0.05$	0.93	0.99	0.20	0.67	0.10	0.11	0.40	0.62
$\sigma_{\eta\eta} < 0.0001$	0.82	0.88	0.065	0.32	0.039	0.37	0.18	0.44
$E_{SC}/p_{track} > 0.9$	0.80	0.98	0.063	0.98	0.01	0.26	0.16	0.88
$\lvert 1/E_{SC} - 1/p_{track}\rvert < 0.02$	0.76	0.95	0.012	0.19	0.009	0.87	0.044	0.28
$\lvert\phi^{prop}_{track} - \phi_{SC}\rvert < 0.006$	0.68	0.88	0.002	0.32	0.008	0.37	0.010	0.44

TABLE 3.5: *Electrons selection applied on different samples. The efficiency is calculated with respect to the number of central super-clusters found in the barrel with a transverse energy higher than 10 GeV.*

black line) and the ones removed by the trigger (red dashed line). Some more detailed study and optimization might be required for the trigger.

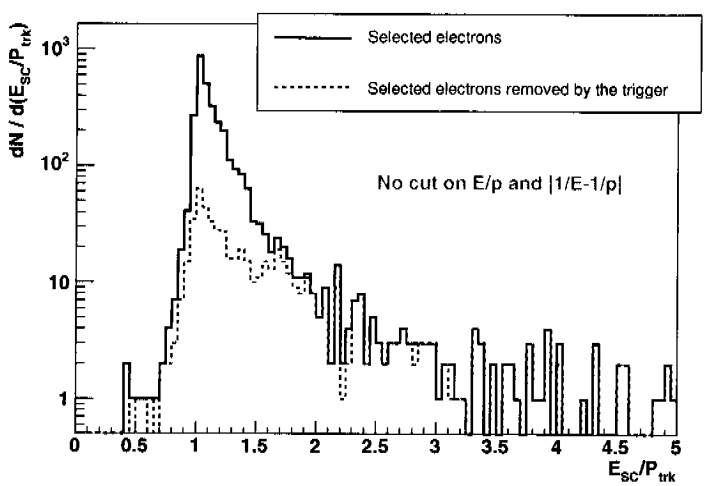

FIGURE 3.23: *The E_{SC}/p_{track} distribution for the selected electrons (without the cuts on E_{SC}/p_{track} and $|1/E_{SC} - 1/p_{track}|$): (black solid line) all selected electrons (red dashed line) selected electrons rejected by the trigger.*

These electron selection cuts were also applied on the following samples: single electrons, jets, single charged pions and taus decaying into hadrons. The efficiencies are summarized in Table 3.5. The efficiencies were calculated with respect to the number of central super-clusters with a transverse energy higher than 10 GeV. Note that for instance only half of the single pions have such an associated super-cluster.

The electron selection cuts presented for the W sample, reduces significantly the three different backgrounds (QCD jets, hadronic taus and single charged pions), keeping a reasonable efficiency for the electrons. An additional cut on $|\phi^{prop}_{track} - \phi_{SC}|$ can be added if a high purity electron sample is required.

3.7 Homogeneity

The CMS electromagnetic calorimeter is designed to minimize the cracks. However, separations between the modules and the super-modules still have an influence on the electron reconstruction. In the following we will try to determine regions in the central barrel where the reconstruction is optimal, the so-called "fiducial regions".

To determine the regions where the electron reconstruction is more difficult, we will select the electrons having only 90% of their energy reconstructed, requiring $E_{SC}/E_{gen} < 0.9$. Figure 3.24 shows the η distribution of such electrons for the 40 GeV (a) and the 100 GeV (b) electron samples. These poorly reconstructed electrons represent 4% of the 40 GeV electrons and 1.2% of the 100 GeV electrons.

The 'peaks' around $\eta \approx 0$, 0.45, 0.8 and 1.15 arise from the energy lost in the cracks present at the separations between the ECAL modules. The peaks are at the same position than the ones observed on Figure 3.7, page 77 (showing the generated electron where no matching super-cluster was found). These 'small' inefficiencies are more important for the 40 GeV electrons than for the 100 GeV electrons. This comes from the fact that for high energy electrons the amount of energy lost in a crack represents a smaller part of the whole electron energy (assuming that the energy lost in the crack does not depend too much on the initial electron energy).

Around the peaks and for the lower energy electrons, there is also a small 'continuum' of poorly reconstructed electrons which increases with η. For the high energy electrons, this 'continuum' is almost absent. The cause of this effect is the Bremsstrahlung. For the low energy electrons the probability that some Bremsstrahlung photons are not clustered together with the electron is higher. It will then be more probable to find electrons with less than 90% of their energy reconstructed. More of these badly reconstructed electrons are found for higher rapidities as the electron is crossing more material, increasing its probability to emit Bremsstrahlung.

In Figure 3.24, regions in η are defined (vertical dashed line) where the electron is inside a crack (if $|\eta| < 0.04$ or $0.41 < |\eta| < 0.47$ or $0.76 < |\eta| < 0.83$ or $1.1 < |\eta| < 1.18$).

Having identified the cracks in η, one can look for the cracks in ϕ. In the design of the CMS ECAL, there is a separation between two super-modules each 20 degrees in the transverse plane. Since the same structure is repeated every 20 degrees and to get more statistics, the super-modules can be folded on each other, by dividing ϕ (in degrees) by 20, keeping the rest of this division. This variable is called $\phi_{\text{mod } 20}$. Figure 3.25 shows $\phi_{\text{mod } 20}$ against the mean ratio between the super-cluster energy and the generated electron energy for the selected electrons. The ratio is lower where there is the super-module separation since some energy gets lost in the crack, but this effect is small. The separation between two super-modules corresponds to a value of $\phi_{\text{mod } 20}$ between 9 and 11 degrees. In the following two different domains were defined: a domain inside the crack, where $9 < \phi_{\text{mod } 20} < 11$ and a domain away from the crack where either $\phi_{\text{mod } 20} < 8$ or $\phi_{\text{mod } 20} > 12$.

A fiducial volume in ϕ and η can be defined in the following way:

- Fiducial volume:
  $0.04 < |\eta| < 0.41$ or $0.47 < |\eta| < 0.76$ or $0.83 < |\eta| < 1.1$ or $1.18 < |\eta| < 1.4$.
  $\phi \mod 20 < 8$ or $\phi \mod 20 > 12$

- Non-fiducial volume:
  $|\eta| < 0.04$ or $0.41 < |\eta| < 0.47$ or $0.76 < |\eta| < 0.83$ or $1.1 < |\eta| < 1.18$.
  $9 < \phi \mod 20 < 11$

For the 40 GeV electrons, 2.2% of the electrons are found to be outside the fiducial volume and 2.5% for the 100 GeV electrons.

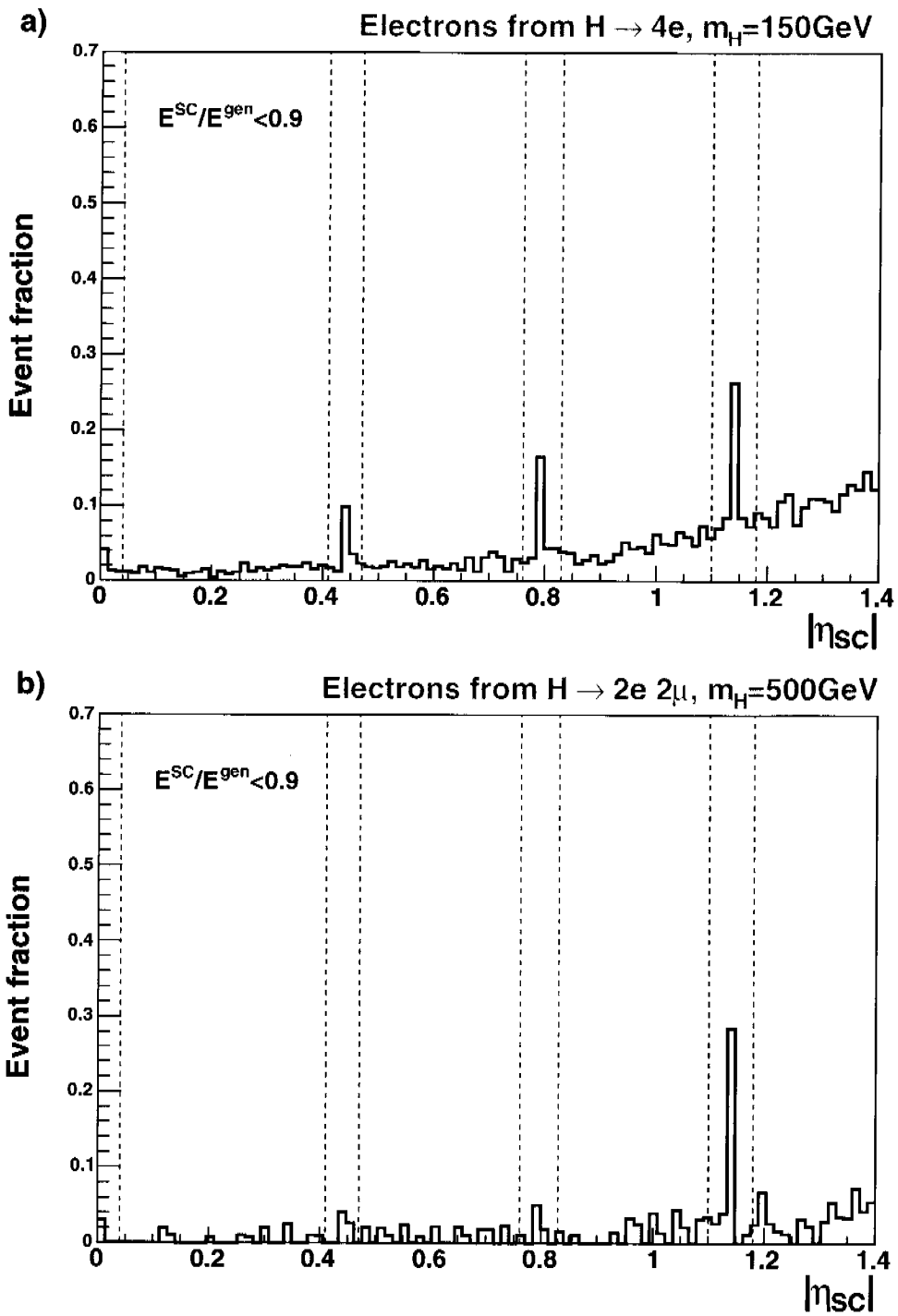

FIGURE 3.24: *The fraction of the electrons having less than 90% of their energy reconstructed as a function of the super-cluster pseudorapidity. (a) for electrons with a mean E_t of 40 GeV, from the $H \rightarrow ZZ^* \rightarrow 4e$ sample, (b) for electrons with a mean E_t of 100 GeV, from the $H \rightarrow ZZ \rightarrow 2e2\mu$ sample. The electrons are selected as explained on page 92.*

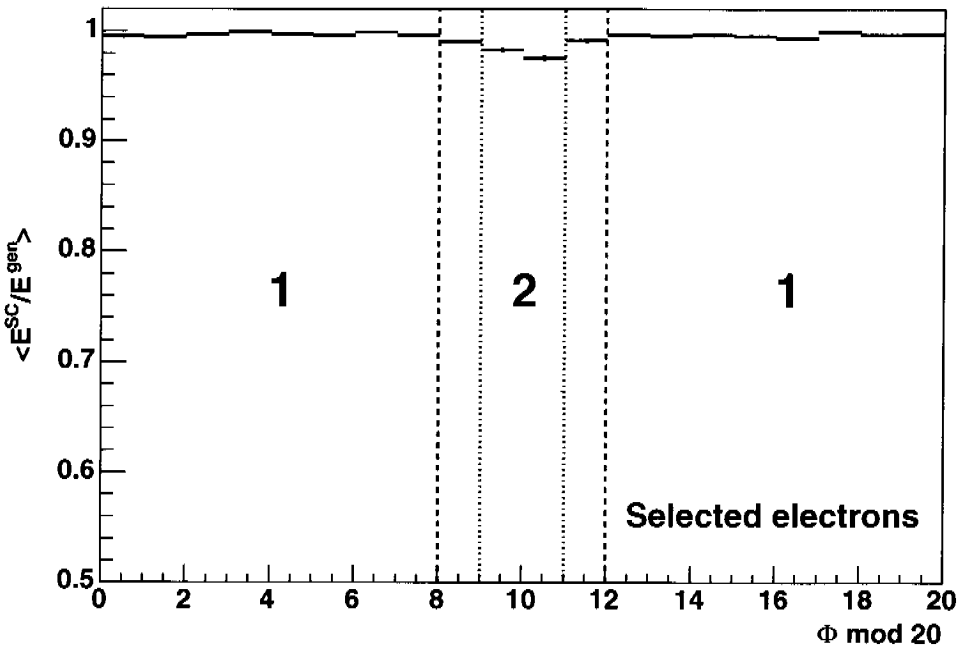

FIGURE 3.25: *The mean ratio between the super-cluster energy and the electron energy as a function of $\phi_{mod\,20}$. The drop comes from the super-module separation every 20 degrees in ϕ. Electrons from the W sample were used.*

3.8 Energy resolution

The energy resolution is determined from the distribution of the ratio between the super-cluster energy and the generated energy. Figure 3.26 shows the E^{SC}/E^{gen} distribution for electrons coming from the $W \rightarrow e\nu$ decays (without restriction on the fiducial volume). The distribution is not really Gaussian and has large tails due to electrons emitting Bremsstrahlung. Nevertheless one would like to define a resolution. Three quantities to quantify the quality of the electron reconstruction can be defined:

- The peak resolution, σ_{peak}, which is the variance of a Gaussian fit of the E^{SC}/E^{gen} distribution, fitted between 0.98 and 1.05.
- The overall resolution, σ_{all}, which is the variance of a Gaussian fit of the E^{SC}/E^{gen} distribution, making a blind Gaussian fit between 0.7 and 1.15. As the fit interval is broader, it does not fit the distribution, it is however sensitive to tails.
- The RMS of the E^{SC}/E^{gen} distribution.

The values for the two different resolutions and the RMS are given in Figure 3.26, together with the fit curves used to obtain the peak (solid line) and the overall (dashed line) resolution.

Figure 3.27 shows a comparison of the electron resolution inside and outside of the fiducial region in the ECAL: Figure 3.27a shows the E^{SC}/E^{gen} when the electron is inside (black solid line) and outside of the fiducial volume in η (red dashed line). Figure 3.27b shows the same but for the electrons inside (black solid line) and outside (red dashed line) of the fiducial volume in ϕ. Figure 3.27c shows the E^{SC}/E^{gen} distribution together with the two resolutions and the RMS for fiducial electrons and Figure 3.27d shows it for electrons outside of the fiducial volume. As the number of electrons outside of the fiducial volume represent only about 2% of the total

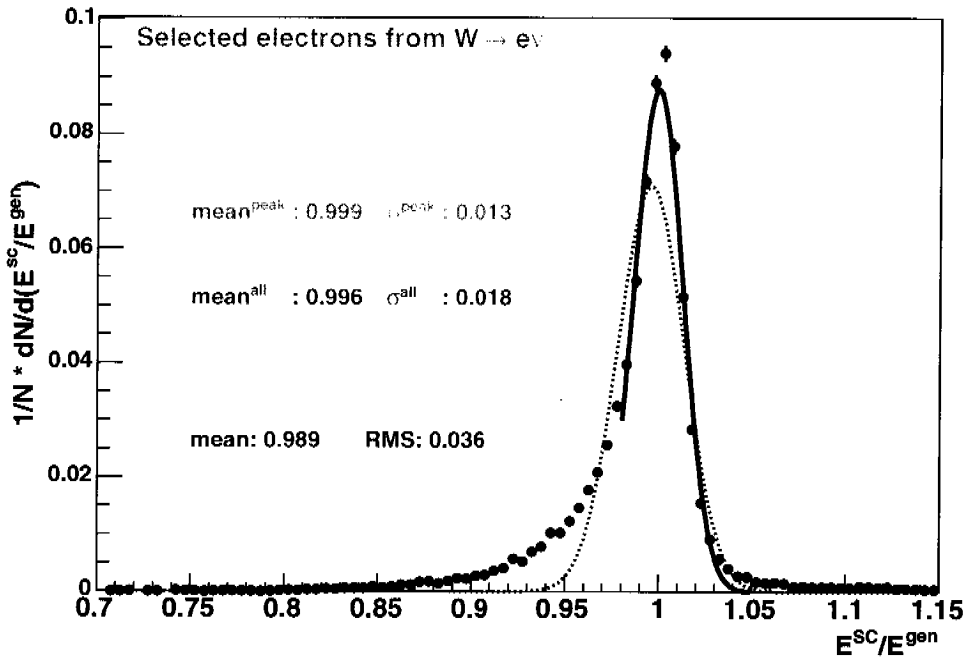

FIGURE 3.26: *E^{SC}/E^{gen} for the selected electrons from the W without restriction on the fiducial volume. The three different resolutions are given: the peak resolution, the overall resolution and the RMS. The solid line shows the fit result for the peak resolution and the dashed line, the fit result for the overall resolution.*

	Peak Resolution		Overall Resolution		RMS	fraction of
	Mean	σ_{peak}	Mean	σ_{all}		electrons
Fiducial	1.001	0.013	0.998	0.017	0.034	98.1%
No fiducial	0.983	0.017	0.977	0.025	0.048	1.9%

TABLE 3.6: *Comparison of the resolutions for electrons from W decays inside and outside of the fiducial volume. The last column gives the fraction of electrons inside and outside the fiducial volume.*

number of electrons, the effect of the cracks on the electron resolution is important but only for a small number of electrons.

The two different resolutions and the RMS inside and outside of the fiducial volume are given in Table 3.6. As expected, σ_{all} is about 30% broader when the electrons are outside of the fiducial volume, σ_{peak} is 25% broader. The mean of the fit for the electron outside of the fiducial volume is also shifted to a smaller value of E^{SC}/E^{gen}, revealing that some energy gets lost in the cracks.

The electron energy resolution depends also on the energy of the electron. The energy resolution of a crystal is usually parameterized in the following way:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$

where the first term accounts for the statistical fluctuations in the development of the shower, the second term accounts for the electronic noise and the noise entering the crystals, coming for instance from the pile-up, and the last term is the constant term which mainly originates from

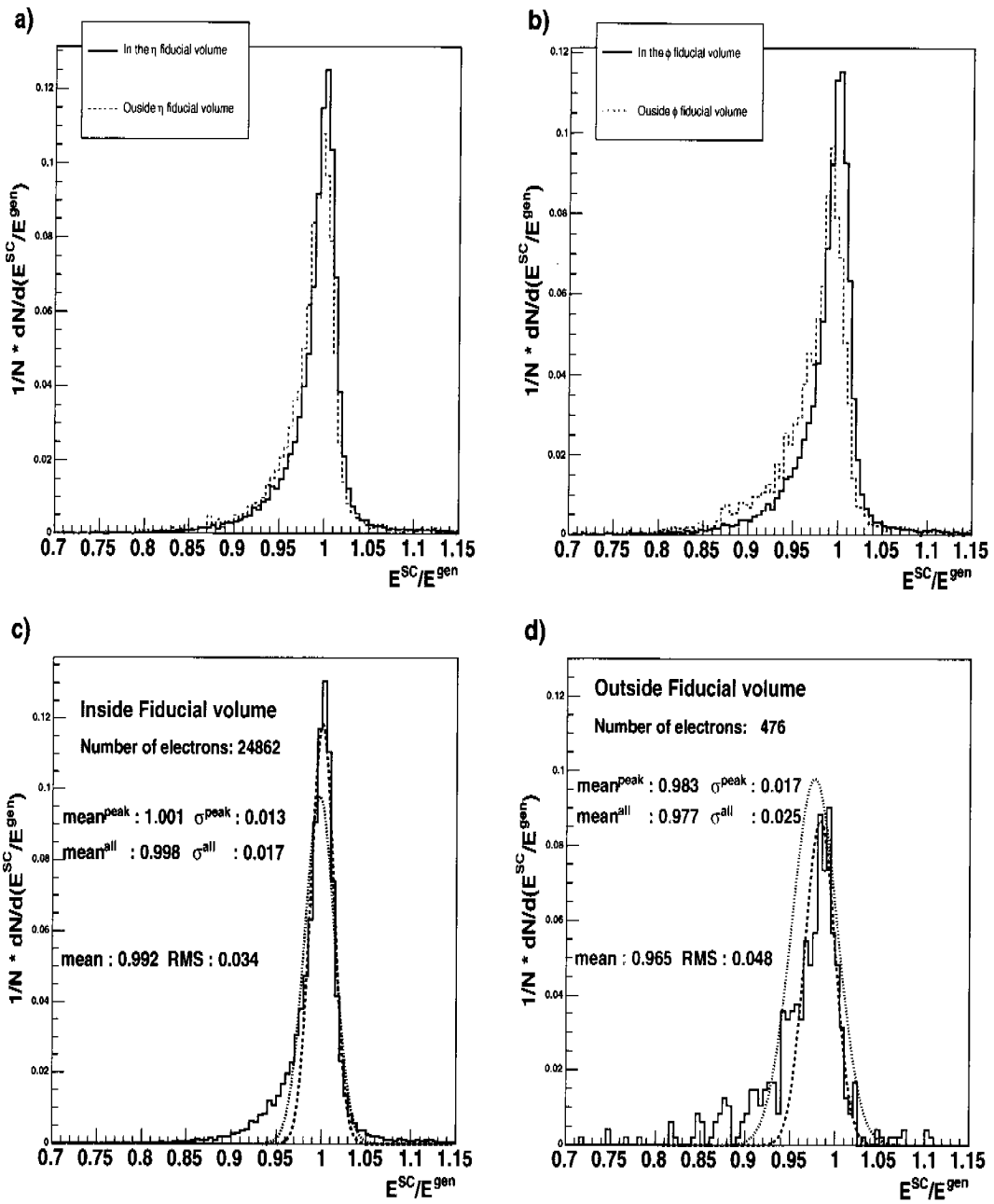

FIGURE 3.27: *The electron energy resolution. (a) in the fiducial volume in η (black solid line) and outside of the fiducial volume in η (red dashed line). (b) in the fiducial volume in ϕ (black solid line) and outside of the fiducial volume in ϕ (red dashed line). (c) The resolution in the overall fiducial volume (d) outside of the overall fiducial volume. All plots are normalized to the total number of events.*

the calibration uncertainty. The events were generated with the GEANT simulation assuming $a = 2.25\%$, $b = 0.04$ GeV and $c = 0.5\%$. In the following we would like to get the effective overall electron resolution including geometrical effects and electron reconstruction.

The Figure 3.28 shows the energy dependence of $\frac{\sigma}{E}$ for the electrons having a mean E_t of 40 GeV (a) and 100 GeV (b). For comparison, the resolution for 35 GeV photons was added on the first plot. The electrons were selected using the selection criteria described on page 92. Several functions and fit-regions were tried. Here, a fit using the following function: $\sqrt{\frac{a^2}{E} + c}$ was applied on the two curves. The idea is to get a good description of the high energy region and there the b/E term can be neglected. The results are very sensitive on the fit region chosen. A fit region between 30 and 100 GeV for the 40 GeV electrons and between 30 and 200 GeV for the 100 GeV electrons was chosen such that the "high energy" region was well described. An "effective" constant term of about 1% is found for both electron energy ranges. For the photons a constant term of 0.8% is found. This term takes into account the effects of the crystal geometry and the clustering method used to reconstruct the electrons. This factor is dominant for electrons with an energy higher than 100 GeV.

The Bremsstrahlung is deteriorating the resolution. Figure 3.29 shows $\frac{\sigma}{E}$ as a function of the pseudorapidity for the 35 GeV dielectrons (black squares) and the 35 GeV diphotons (red circles). The resolution for the electrons is becoming worse as the pseudorapidity increases, which is mainly due to the fact that at high η, electrons cross more material and thus the effect of the Bremsstrahlung becomes more important. If specific cuts against Bremsstrahlung are added to the electron selection, the energy resolution gets closer to the photon one (blue triangles). The cuts applied to minimize the Bremsstrahlung were the following: $E/p < 1.4$, $|\phi_{SC} - \phi_{tr}^{prop}| < 0.003$ and $E_{3\times3}/E_{5\times5} > 0.9$. The efficiency for these cuts is about 60%.

3.9 Summary

Clean electrons are expected to be efficiently selected in the CMS detector using the following variables: the track isolation, E_{had}/E_{EM}, $E_{3\times3}/E_{5\times5}$, $\sigma_{\eta\eta}$, E_{SC}/p_{track} and $|\phi_{SC} - \phi_{tr}^{prop}|$. The energy and Bremsstrahlung dependence of these variables were discussed.

Furthermore a fiducial volume for the ECAL barrel was defined, taking into account the cracks between the modules and super-modules. There are only about 2-3% electrons outside of the fiducial volume. These electrons are found to have a slightly worse energy resolution.

Finally the electron energy resolution and its energy dependence were studied. An "effective" constant term of about 1% for electrons and 0.8% for photons was found. This difference comes mainly from Bremsstrahlung.

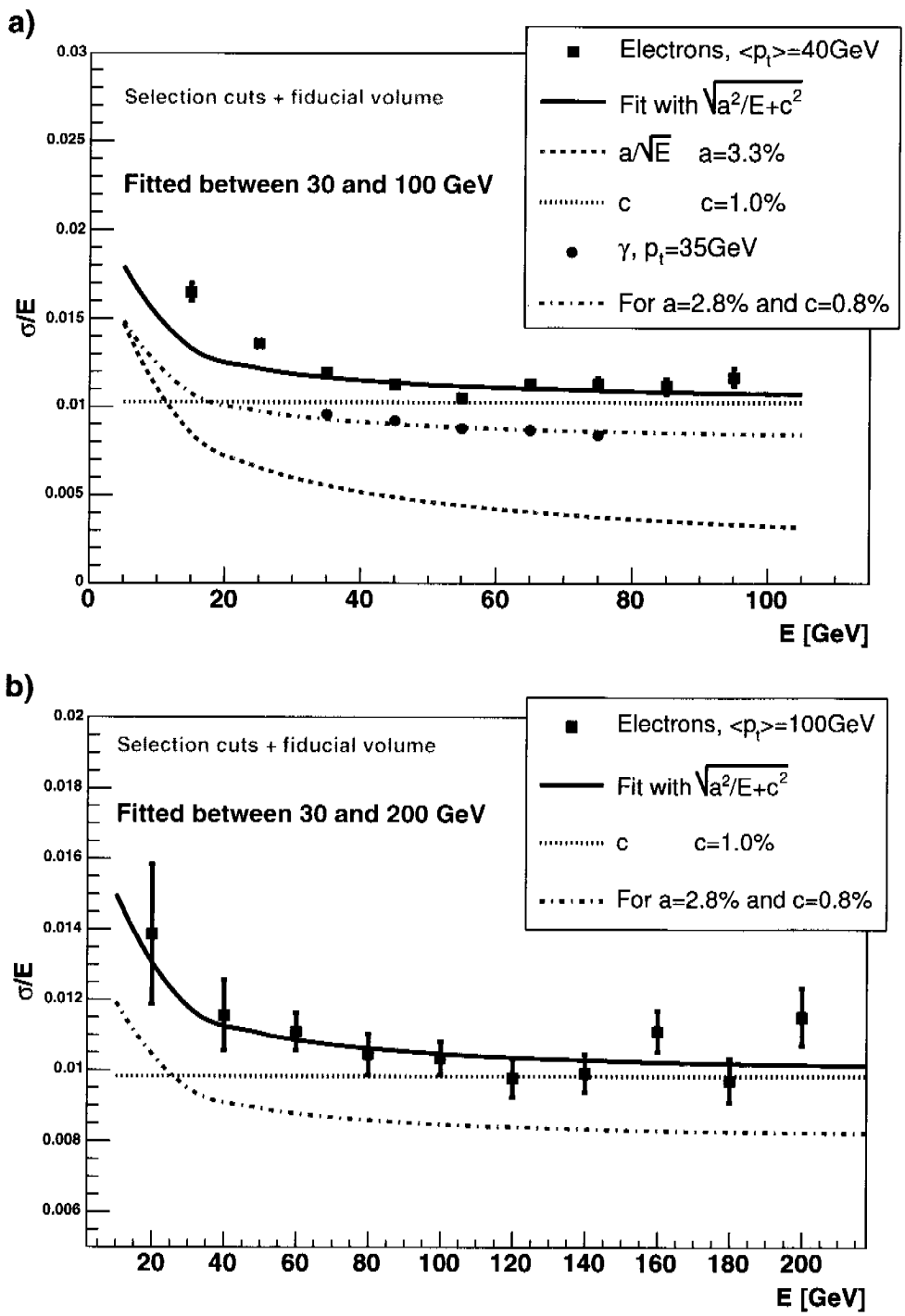

FIGURE 3.28: *The energy resolution as a function of the electron transverse energy shown together with different fits. (a) for the electrons with* $< E_t >= 40$ *GeV (from a 150 GeV Higgs) and 35 GeV single photons. (b) for the electrons with* $< E_t >= 100$ *GeV (from a 500 GeV Higgs)*

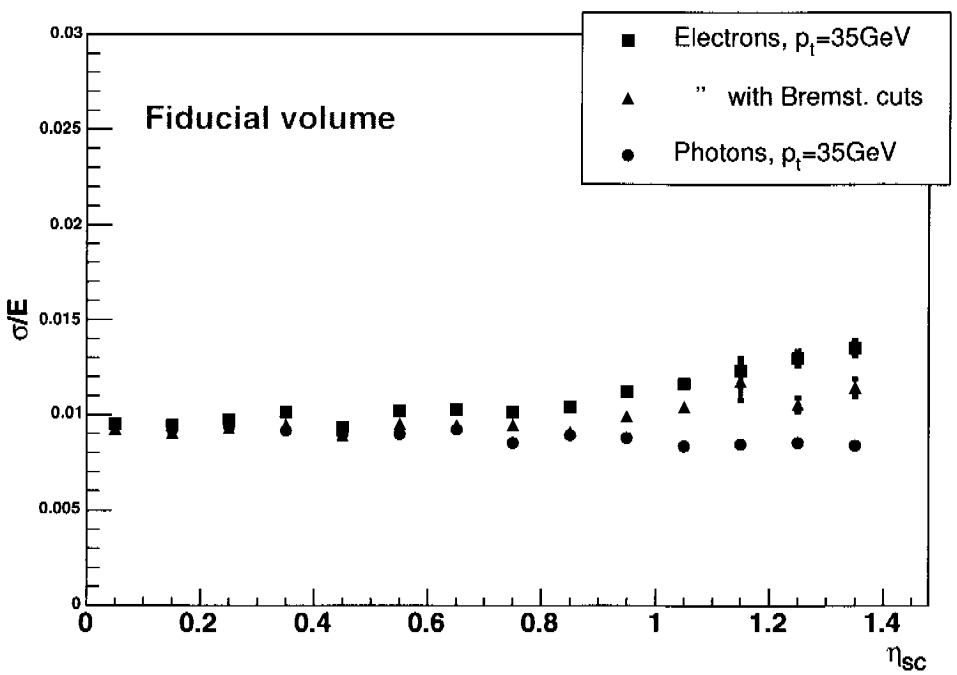

FIGURE 3.29: *The energy resolution, σ/E as a function of the electron pseudorapidity. The electrons and photons are selected in the way described on page 92. The red dots show the resolution for photons, the black squares for electrons and the blue triangles for electrons where three cuts against electron emitting Bremsstrahlung were added: $E/p < 1.4$, $|\phi_{SC} - \phi_{tr}^{prop}| < 0.003$ and $E_{3\times3}/E_{5\times5} > 0.9$.*

Chapter 4

Identifying a Z' at the LHC

Although the Standard Model of the electroweak and strong interactions describes nearly all experimental data available today, it is widely believed that it is not the ultimate theory. Many theories have been developed to extend or replace to Standard Model at higher energies. Assuming that the LHC discovers new phenomena, one would like to constrain the parameters associated to this new physics as much as possible.

In the following we will concentrate on the Z', a new gauge boson predicted by different theories. We will see what kind of observables could be used at the LHC in order to discriminate between the different models.

A possibility to extend the Standard Model is achieved through Grand Unified Theories (GUT) which unify the different interaction couplings at high energies, by embedding the Standard Model gauge group $SU(3) \times SU(2) \times U(1)$ into a higher order group like SU(5), SO(10) or E_6 [65]. Many of these GUT models predict the existence of new neutral gauge bosons, which might be light enough to be accessible in future collider experiments; for reviews see [66]. New vector bosons appear also in models of dynamical symmetry breaking [67]. Recently, "little Higgs" models have been proposed to solve the hierarchy problem of the Standard Model [68]: they also have large gauge group structures and therefore predict many new gauge bosons with masses in the TeV range.

The search for these Z' particles in the Drell-Yan process $p\bar{p} \rightarrow Z' \rightarrow \ell^+\ell^-$, with $\ell = e, \mu$ [69] is an important aspect of the experimental physics program of future high-energy colliders. Present limits from direct production at the Tevatron and virtual effects at LEP, through interference or mixing with the Z boson, imply that new Z' bosons are rather heavy and mix very little with the Z boson. Depending on the considered theoretical models, Z' masses of the order of 500 to 800 GeV and Z–Z' mixing angles at the level of a few per-mile are excluded[1] [72]. A Z' boson, if lighter than about 1 TeV, could be discovered at Run II of the Tevatron [73]. Detailed theoretical [73] and experimental [74, 75, 11] analyzes have shown that the discovery potential of the LHC experiments is about 5 TeV.

After the discovery of a Z' boson, some information about its couplings needs to be obtained in order to constrain the theoretical frame. For this purpose, the forward-backward charge asymmetry for leptons A_{FB}^{ℓ} has been advocated as being a powerful tool [76]; the most direct method to actually measure A_{FB}^{ℓ} at the LHC has been described in [77]. In addition to the information from the total Z' cross section, it has been argued that the measurement of ratios of Z' cross sections in different rapidity bins might provide some information about the Z' couplings to up and down quarks [78].

[1] In contrast, some experimental data on atomic parity violation and deep inelastic neutrino-nucleon scattering, although controversial and of small statistical significance [see Ref. [70] for instance], can be explained by the presence of a Z' boson [71].

While numerous theoretical and experimentally motivated Z′ studies had already been performed, the combination of all sensitive LHC variables had not been done before the study presented here. We will perform the studies using the PYTHIA program [79] and a fast LHC detector simulation[2]. First, following the method proposed in [77], the forward-backward charge asymmetries, on and off the Z′ resonance peak, are analyzed together with the cross section in order to differentiate between the different models[3]. Then, we show that a direct fit of the rapidity distribution provides additional information which could be used to disentangle between Z′ bosons from various models through their different couplings to up-type and down-type quarks.

The following section defines the theoretical framework in which the analysis is performed. Then, the relevant observables that can be measured at the LHC are defined, namely the dilepton cross section times the Z′ total width, the on-peak and off-peak forward-backward charge asymmetries and the rapidity distribution. In section 4.3, we analyze the resolving power of these observables and finally discuss some detector effects.

4.1 The Z′ models considered

To simplify the discussion, we will focus here on two effective theories of well motivated models that lead to an extra gauge boson:

1) *An effective* $SU(2)_L \times U(1)_Y \times U(1)_{Y'}$ *model*, which originates from the breaking of the exceptional group E_6, general enough to include many interesting possibilities. Indeed, in the breaking of this group down to the Standard Model symmetry, two additional neutral gauge bosons could appear. For simplicity we assume that only the lightest Z′ can be produced at the LHC. It is defined as

$$Z' = Z'_\chi \cos\beta + Z'_\psi \sin\beta \tag{4.1}$$

and can be parametrized in terms of the hypercharges of the two groups $U(1)_\psi$ and $U(1)_\chi$ which are involved in the breaking chain:

$$E_6 \to SO(10) \times U(1)_\psi \to SU(5) \times U(1)_\chi \times U(1)_\psi \to SU(3)_c \times SU(2)_L \times U(1)_Y \times U(1)_{Y'}$$

The particular models that will be studied in the following correspond to the values $\beta = 0$ and $\beta = \pi/2$ and are, respectively, pure Z'_χ and Z'_ψ bosons, and the value $\beta = \arctan(-\sqrt{5/3})$ that is the Z'_η boson originating from the direct breaking of E_6 to a rank-5 group in Superstring inspired models.

2) *Left-right (LR) models*, based on the symmetry group $SU(2)_R \times SU(2)_L \times U(1)_{B-L}$, where B and L are the baryon and lepton numbers. Even though we investigate only the Z′ in this paper, it should be recalled that new charged vector bosons, potentially observable at the LHC, also appear in these models. The most general neutral boson will couple to a linear combination of the right-handed and B–L currents:

$$J^\mu_{LR} = \alpha_{LR} J^\mu_{3R} - (1/2\alpha_{LR}) J^\mu_{B-L} \quad \text{with } \alpha_{LR} = \sqrt{(c_W^2 g_R^2 / s_W^2 g_L^2) - 1} \tag{4.2}$$

where $g_L{=}e/s_W$ and g_R are the $SU(2)_L$ and $SU(2)_R$ coupling constants with $s_W^2 = 1 - c_W^2 \equiv \sin^2\theta_W$. The parameter α_{LR} is restricted to lie in the range $\sqrt{2/3} \lesssim \alpha_{LR} \lesssim \sqrt{2}$: the upper bound corresponds to a LR-symmetric model with $g_R = g_L$, and which will be studied in the following as Z'_{LR}, while the lower bound corresponds to the Z'_χ model discussed in the first model, since SO(10) can lead to both $SU(5) \times U(1)$ and $SU(2)_R \times SU(2)_L \times U(1)$ breaking patterns.

[2]Fast simulation means that we just simulated the acceptance of the detector, no resolution effects were simulated.

[3]Recently, the off-peak forward-backward asymmetry has also been used in [80] to study Kaluza-Klein excitations of gauge bosons.

For a complete comparison, we will also discuss the case of a sequential boson Z'_{SM}, which has the same fermion couplings as the Standard Model Z boson, although it is not a theoretically viable theory. We will also take the case of a Z′ boson, denoted by Z'_d, with vanishing axial and vectorial couplings to u quarks and which, in E_6 models, corresponds to the choice $\cos\beta = \sqrt{5/8}$.

The left- and right-handed couplings of the Z′ boson to fermions, defined as:

$$g_{Z'} J^{\mu}_{Z'} Z'_{\mu} = \frac{e}{c_W} \sum_f \gamma^{\mu} \left[\frac{1-\gamma_5}{2} g_L^{fZ'} + \frac{1+\gamma_5}{2} g_R^{fZ'} \right] \tag{4.3}$$

are given in Table 4.1 for the first-generation fermions in the two scenarios. Given the experimental results, one knows that the mixing between the Z and Z′ bosons is very small [72] and will be neglected in our discussion.

f	$g_L^{fZ'}\vert_{E_6}$	$g_R^{fZ'}\vert_{E_6}$	$g_L^{fZ'}\vert_{LR}$	$g_R^{fZ'}\vert_{LR}$
ν_e	$\frac{3\cos\beta}{2\sqrt{6}} + \frac{\sqrt{10}\sin\beta}{12}$	0	$\frac{1}{2\alpha_{LR}}$	0
e	$\frac{3\cos\beta}{2\sqrt{6}} + \frac{\sqrt{10}\sin\beta}{12}$	$\frac{\cos\beta}{2\sqrt{6}} - \frac{\sqrt{10}\sin\beta}{12}$	$\frac{1}{2\alpha_{LR}}$	$\frac{1}{2\alpha_{LR}} - \frac{\alpha_{LR}}{2}$
u	$-\frac{\cos\beta}{2\sqrt{6}} + \frac{\sqrt{10}\sin\beta}{12}$	$\frac{\cos\beta}{2\sqrt{6}} - \frac{\sqrt{10}\sin\beta}{12}$	$-\frac{1}{6\alpha_{LR}}$	$-\frac{1}{6\alpha_{LR}} + \frac{\alpha_{LR}}{2}$
d	$-\frac{\cos\beta}{2\sqrt{6}} + \frac{\sqrt{10}\sin\beta}{12}$	$-\frac{3\cos\beta}{2\sqrt{6}} - \frac{\sqrt{10}\sin\beta}{12}$	$-\frac{1}{6\alpha_{LR}}$	$-\frac{1}{6\alpha_{LR}} - \frac{\alpha_{LR}}{2}$

TABLE 4.1: *Left- and right-handed couplings of the Z′ boson to the Standard Model fermions with the notation of the first generation in the E_6 (left panels) and LR (right panels) models.*

The Z′ partial decay width into a massless fermion-antifermion pair is:

$$\Gamma^f_{Z'} = N_c \frac{\alpha M_{Z'}}{6c_W^2} \left[(g_L^{fZ'})^2 + (g_R^{fZ'})^2 \right] \tag{4.4}$$

with N_c the color factor and the electromagnetic coupling constant to be evaluated at the scale $M_{Z'}$. In the absence of any exotic decay channel, the branching fractions for decays into the first-generation leptons and quarks are shown in Figure 4.1 for E_6 and LR models as functions of $\cos\beta$ and α_{LR}, respectively. As can be seen, the decay fractions into $\ell^+\ell^-$ pairs are rather small, varying between 6.6% and 3.4% for E_6 models and 6.6% and 2.3% for LR models; in the latter case the decay branching fraction is largest for the symmetric case $g_L = g_R$ and smallest for $\alpha_{LR} \simeq \sqrt{2}$. The Z′ total decay width, normalized to $M_{Z'}$, is also shown in Figure 4.1: it is largest when $\cos\beta = \pm 1$ in E_6 models and $\alpha_{LR} \simeq \sqrt{2}$ in LR ones. The Z′ bosons studied here are thus narrow resonances, as their total decay width does not exceed 2% of their masses[4].

In the limit of negligible fermion masses, the differential cross section for the subprocess $q\bar{q} \to \ell^+\ell^-$, with respect to θ^* defined as the angle between the initial quark q and the final lepton ℓ^- in the Z′ rest frame, is given by

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\cos\theta^*}(q\bar{q} \to \gamma, \mathrm{Z}, \mathrm{Z}' \to \ell^+\ell^-) = \frac{1}{9}\frac{\pi\alpha^2}{2\hat{s}} \left[(1+\cos^2\theta^*)Q_1 + 2\cos\theta^* Q_3 \right] \tag{4.5}$$

where $\hat{s} = M^2_{\ell\ell}$ is the center of mass energy of the subprocess and the charges Q_1 and Q_3 are given by [81]

$$Q_{1/3} = \left[|Q_{LL}|^2 + |Q_{RR}|^2 \pm |Q_{RL}|^2 \pm |Q_{LR}|^2 \right]/4 \tag{4.6}$$

In terms of the left- and right-handed couplings of the Z′ boson defined previously, and of those of the Z boson [$g_L^{fZ} = I_{3L}^f - Q^f s_W^2$, $g_R^{fZ} = -Q_f s_W^2$] and the photon [$g_L^{f\gamma} = g_R^{f\gamma} = Q_f$] with

[4]Note however that non-standard decays, such as decays into supersymmetric particles and/or decays into exotic fermions, are possible; if kinetically allowed, they can increase the total decay width and hence decrease the $Z' \to \ell^+\ell^-$ branching ratios.

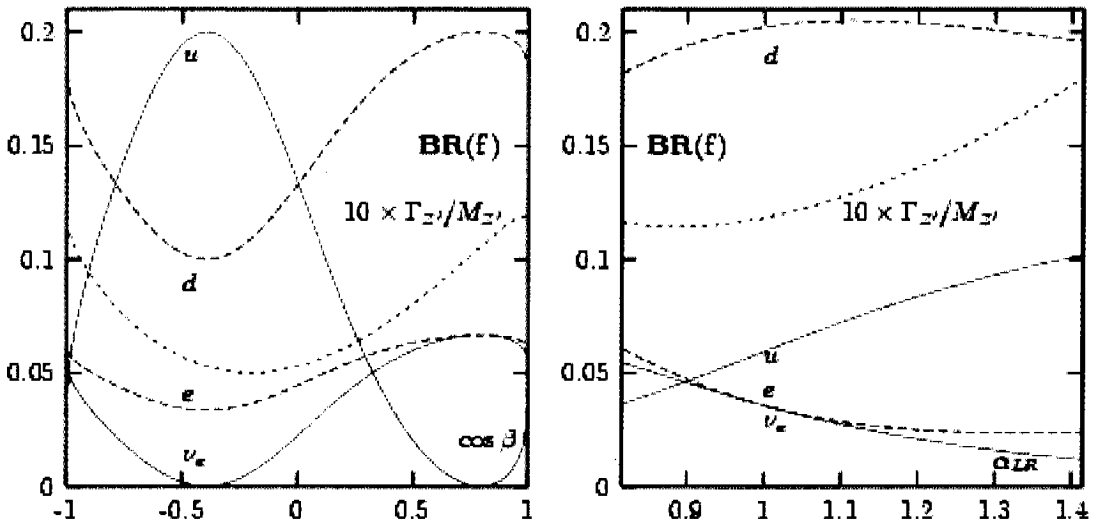

FIGURE 4.1: *The branching ratios of the decays* $Z' \to f\bar{f}$ *in* E_6 *models as a function of* $\cos\beta$ *(left) and in LR models as a function of* α_{LR} *(right). The total* Z' *decay widths, normalized to* $10/M_{Z'}$*, are also shown.*

Q^f the electric charge and I_{3L}^f the left-handed weak isospin, the helicity amplitudes Q_{ij} with $i,j = L,R$ for a given initial $q\bar{q}$ state read

$$Q_{ij}^q = g_i^{q\gamma} g_j^{\ell\gamma} + \frac{g_i^{qZ} g_j^{\ell Z}}{s_W^2 c_W^2} \frac{\hat{s}}{\hat{s} - M_Z^2 + i\Gamma_Z M_Z} + \frac{g_i^{qZ'} g_j^{\ell Z'}}{c_W^2} \frac{\hat{s}}{\hat{s} - M_{Z'}^2 + i\Gamma_{Z'} M_{Z'}} \tag{4.7}$$

To obtain the total hadronic cross section[5] and forward-backward asymmetries, we must sum over the contributing quarks and fold with the parton luminositics.

A few points are worth recalling concerning the forward-backward asymmetry in E_6 models [76]:

1. Since the up-type quarks have no axial couplings to the Z' boson, $Q_3^q = 0$, they do not contribute to $A_{\rm FB}^\ell$ on the Z' peak.

2. The asymmetry completely vanishes for three β values: $\beta = \arctan(-\sqrt{3/5})$ and $\beta = \pm\pi/2$ (corresponding to a Z'_ψ), where the left- and right-handed Z' couplings of both d-quarks and charged leptons are equal.

3. There is always an off-peak asymmetry that is generated by the Z boson and γ couplings and its interferences with the Z'.

4.2 Observables sensitive to Z' properties

The LHC discovery potential for a Z' detected as a mass peak above a small background in the reaction $pp \to Z' \to \ell^+\ell^-$, with $\ell = e, \mu$, is well known. The required luminosity to discover a Z' basically depends only on its cross section times branching ratio, and therefore on its mass and couplings to quarks and leptons.

Once a Z' boson is observed at the LHC, it is straightforward to measure its mass, its total width and cross section. Furthermore, forward-backward charge asymmetries on and off the Z' resonance provide additional information about its couplings and interference effects with the Z boson and the photon. In addition one can include the analysis of the Z' rapidity distribution,

[5] A K-factor of the order of $K_{\rm DY} \sim 1.4$ [82] for the production cross section can be also included.

which is sensitive to the Z' couplings to $u\bar{u}$ and $d\bar{d}$ quarks. Such future measurements can be performed as follows at the LHC:

- *The total decay width of the Z'*

  It is obtained from a fit to the invariant mass distribution of the reconstructed dilepton system using a non-relativistic Breit-Wigner function:

  $$\frac{a_0}{[(M_{\ell\ell}^2 - M_{Z'}^2)^2 + a_1]}$$

  with $a_1 = \Gamma_{Z'}^2 M_{Z'}^2$.

- *The Z' cross section times leptonic branching ratio*

  It is calculated from the number of reconstructed dilepton events lying for example within $\pm 3\Gamma$ around the observed peak[6].

- *The leptonic forward-backward charge asymmetry*

  A_{FB}^{ℓ} is defined from the lepton angular distribution with respect to the quark direction in the Z' center-of-mass frame, as:

  $$\frac{d\sigma}{d\cos\theta^*} \propto \frac{3}{8}(1 + \cos^2\theta^*) + A_{FB}^{\ell} \cos\theta^* \tag{4.8}$$

  A_{FB}^{ℓ} can then be determined with an unbinned maximum likelihood fit to the $\cos\theta^*$ distribution. A_{FB}^{ℓ} cannot be measured directly in a proton-proton collider, as the original quark direction is not known. However, it can be extracted from the kinematics of the dilepton system, as it was shown in detail in [77]. The method is based on the different momentum fraction (x_i) spectra of the quarks and antiquarks in the proton, which allows to approximate the quark direction with the boost direction of the $\ell\ell$ system with respect to the beam axis (the z axis). Consequently, the probability to assign the correct quark direction increases for larger rapidities of the dilepton system.

  Figure 4.2a shows the $\cos\theta^*$ distribution for the Standard Drell-Yan pairs, when the angle is calculated between the lepton and (1) the quark direction (black solid line), (2) the direction of the boost of the dilepton system (blue dotted line) and (3) with a random direction (red dashed line). Compared to the angle between the lepton and a random direction, the $\cos\theta^*$ distribution where the angle is taken between the lepton and the boost of the dilepton system shows an asymmetry. It is however smaller than the real one where $\cos\theta^*$ is taken between the quark and the lepton. Figure 4.2b shows the rapidity of the dilepton system for the Standard Model Drell-Yan for all accepted events (red dashed curve) and for the events where the initial quark direction is in the same direction than the boost of the dilepton system (blue dotted line). At high rapidities almost all Z' are produced in such a way that the quark direction is parallel to the boost direction. A purer, though smaller, signal sample can thus be obtained by introducing a rapidity cut. For the following studies we will require $|Y_{\ell\ell}| > 0.8$.

  Figure 4.3a shows A_{FB}^{ℓ}, assuming that the quark direction is known, as a function of the dilepton mass for a Z'_{χ} and a Z'_{SM} boson, assuming a mass $M_{Z'} = 1.5$ TeV. A_{FB}^{ℓ} varies strongly with the dilepton mass and is very different in the two models. Figure 4.3b shows the measurable asymmetry in different dilepton rapidity intervals for a 1.5 TeV Z'_{χ} together with the theoretical prediction.

[6] As noted previously, both the total width and the cross section times the leptonic branching ratio can be altered if exotic decays of the Z' boson are present. However, this dependence disappears in the product, and it is this quantity that should be used in discriminating models independently of the decays.

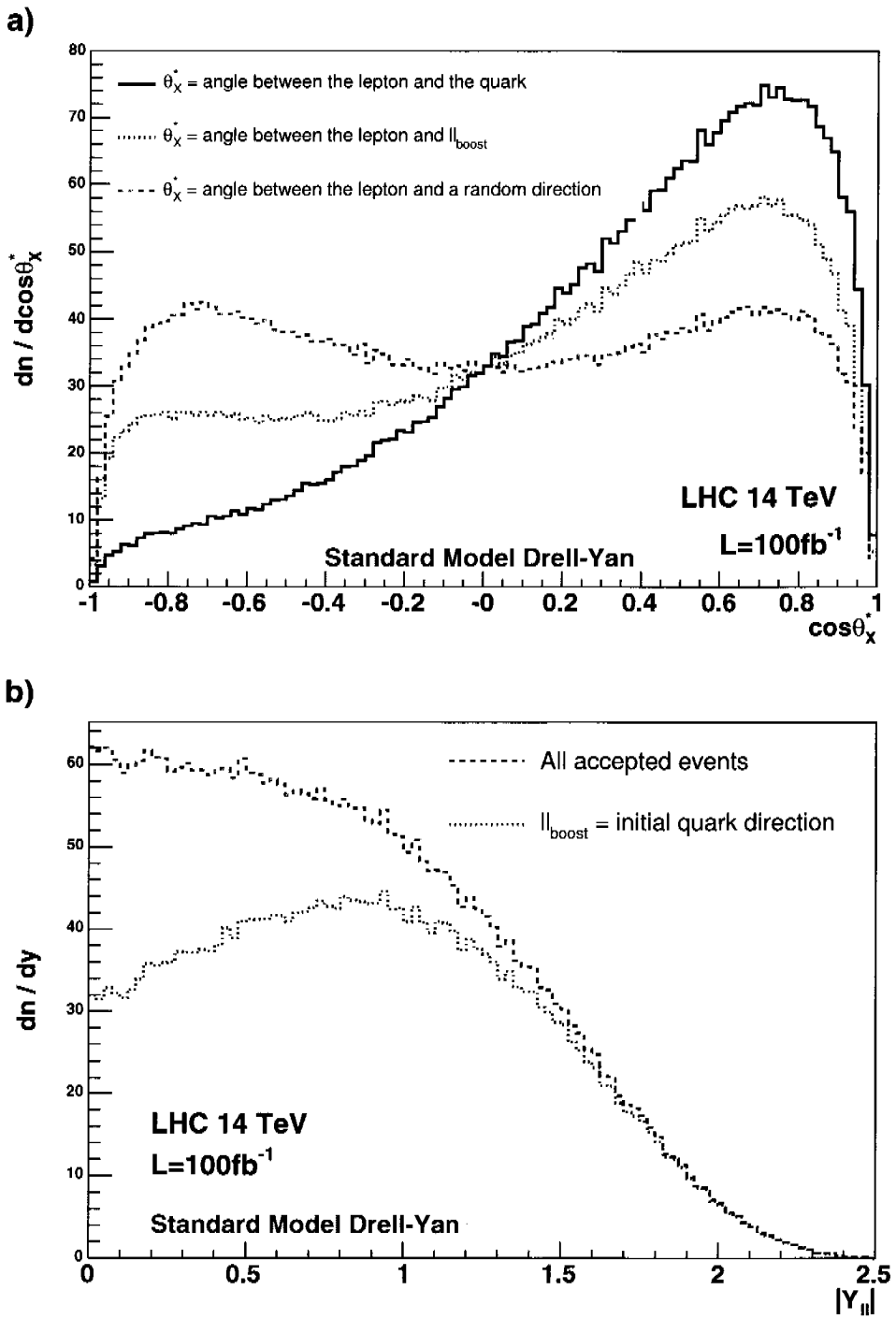

FIGURE 4.2: *(a) The* $\cos\theta^*$ *distribution for the Standard Drell-Yan pairs, when the angle is calculated between the lepton and: the quark direction (black solid line), the direction of the boost of the dilepton system (blue dotted line) and with a random direction (red dashed line). (b) The rapidity of the dilepton system for the Standard Model Drell-Yan for all accepted events (red dashed line) and for the events where the initial quark direction is in the same direction than the boost of the dilepton system (blue dotted line).*

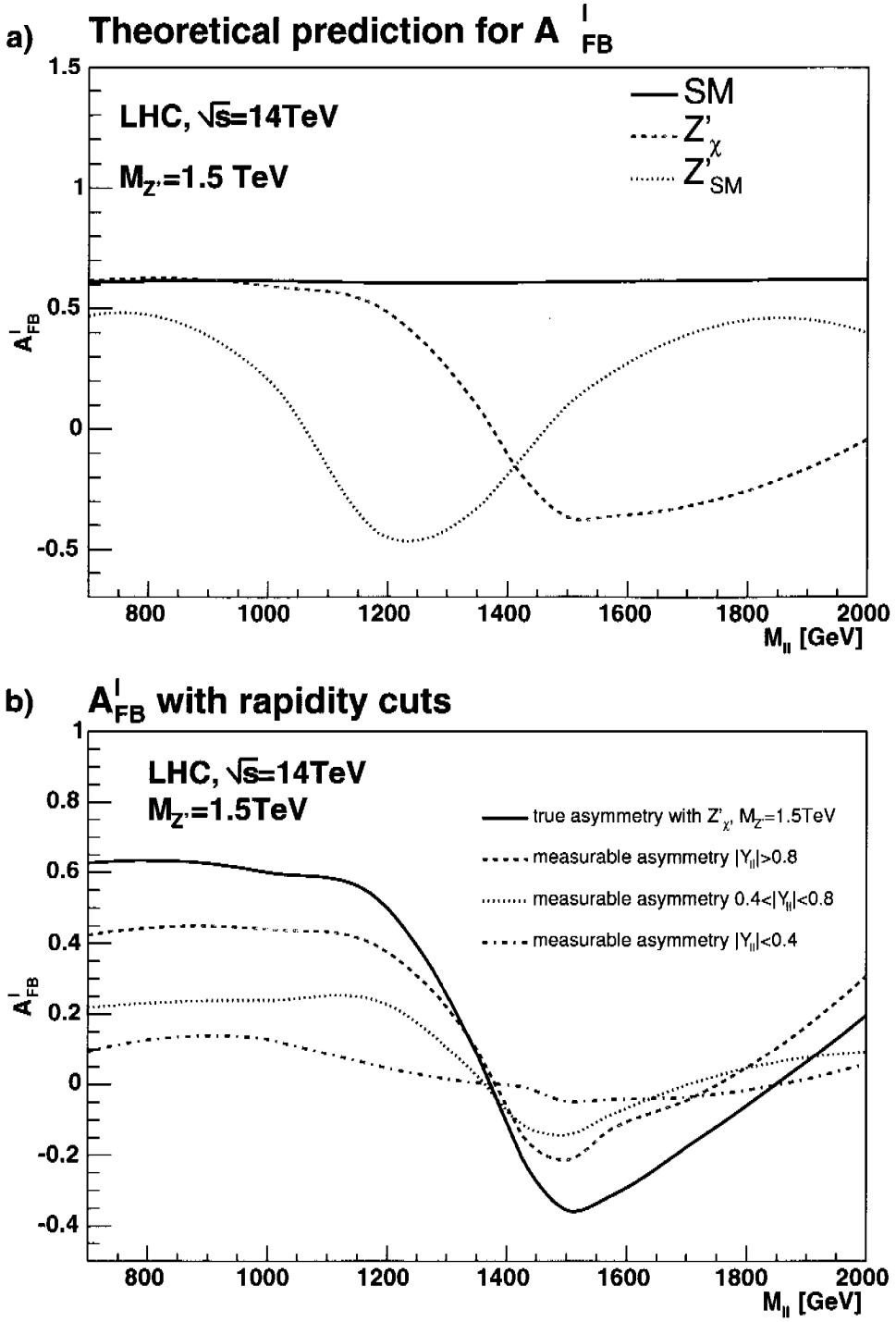

FIGURE 4.3: *A_{FB}^{ℓ} as a function of the dilepton mass: (a) the predicted asymmetry for a Z'_{χ} (red dashed line) and a Z'_{SM} (green dotted) together with the Standard Model prediction (black solid line) and (b) the measurable asymmetry in different dilepton rapidity intervals for a 1.5 TeV Z'_{χ} (broken lines) and for the theoretical prediction (black line).*

- *The Z' rapidity distribution.*

  This rapidity distribution allows us to obtain the fraction of Z' bosons produced from $u\bar{u}$ and $d\bar{d}$ initial states. The shape of the rapidity spectrum for a given flavor of initial quarks depends only on the parton distribution functions and the Z' mass. Assuming that the $W^{\pm}$ and Z boson rapidity distributions have been measured in detail, as discussed in Chapter 2, relative parton distribution functions for u and d quarks, as well as for the corresponding sea quarks and antiquarks are known well enough. Thus, the shape of the rapidity spectra can be calculated separately for $u\bar{u}$ and $d\bar{d}$, as well as for sea quark antiquark annihilation and extrapolated to the mass region of interest. Using the shape of these distributions, a fit can be performed to the Z' rapidity distribution, which allows to obtain the corresponding fractions of the Z' boson produced from $u\bar{u}$, $d\bar{d}$ as well as for sea quark-antiquark annihilation[7], which is directly dependent on the Z' coupling strenght to u and d quarks.

In the present analysis, PYTHIA events of the type $pp \rightarrow \gamma^*, Z, Z' \rightarrow ee/\mu\mu$ were simulated at a center-of-mass energy of 14 TeV, and for the Z' models discussed in section 4.1. The CTEQ5L parton distribution function was used [83]. The Z' masses were varied from 1 TeV up to 5 TeV. These events were analyzed, using simple acceptance cuts following the design criteria of ATLAS and CMS. Following the results from previous studies and the expected excellent detector resolutions, the obtained values are known to be rather insensitive to measurement errors, especially for the e^+e^- final states. We therefore do not include any resolution for the current study. In detail, the following basic event selection criteria were used:

- The transverse momenta of the leptons, p_T^{ℓ}, should be at least 20 GeV.
- The pseudorapidity $|\eta|$ of each lepton should be smaller than 2.5.
- The leptons should be isolated, requiring that the lepton carries at least 95% of the total transverse energy found in a cone of size of 0.5 around the lepton.
- There should be exactly two isolated leptons with opposite charge in each event.
- The two leptons should be back to back in the plane transverse to the beam direction, so that the opening angle between them is larger than 160°.

Figure 4.4 shows the expected number of events for masses intervals of 500 GeV for a luminosity of 100 fb^{-1}. The Standard Model background (Drell-Yan lepton pair production) relative to the signal cross section is found to be essentially negligible for the considered Z' models. We thus reconfirm the known Z' boson LHC discovery potential, to reach masses up to about 5 TeV for a luminosity of 100 fb^{-1} [73].

Figure 4.5a shows the invariant mass distribution for the dilepton system, as expected for different models with $M_{Z'}$ fixed to 1.5 TeV and for the Standard Model using a luminosity of 100 fb^{-1}. For all Z' models, huge peaks, corresponding to 3000–6000 signal events, are found above a small background. The cross sections for Z' bosons in the various models are also strongly varying.

The forward-backward charge asymmetries expected as a function of the dilepton mass and for the different Z' models, are shown in Figure 4.5b. In order to get an impression of how an experimental signal with statistical fluctuations would look like, the measurable forward-backward asymmetry in the Z'_{η} case has been generated with the number of events corresponding

[7]Following this procedure, it would be imaginable even to measure also the forward-backward charge asymmetries separately for u and d quarks. Charge asymmetries for different Z' rapidity intervals would have to be measured and, with the knowledge of the corresponding $u\bar{u}$ and $d\bar{d}$ fractions from the entire rapidity distribution, the corresponding u and d asymmetries could eventually be disentangled. However, an estimate of the potential sensitivity indicates that an interesting statistical sensitivity would require a luminosity of at least 1000 fb^{-1}.

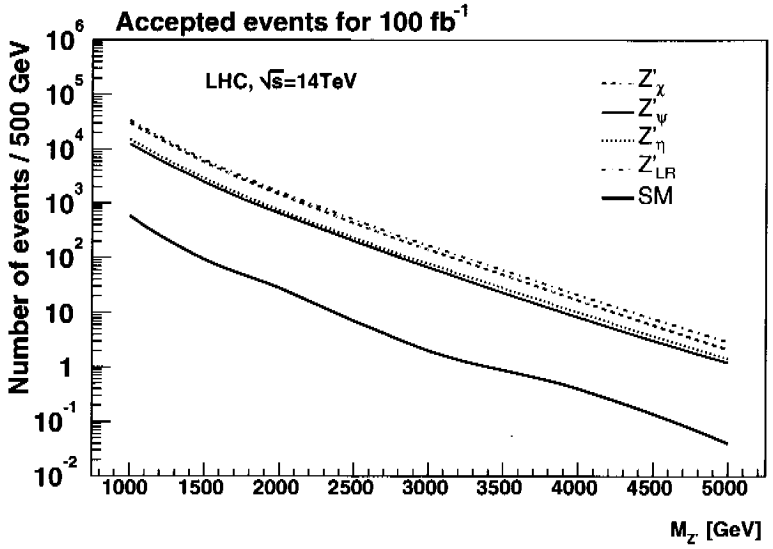

FIGURE 4.4: *The LHC discovery reach: the number of events fulfilling the selection cuts for a luminosity of 100 fb^{-1}, for the different Z' models and in the Standard Model case.*

to 100 fb^{-1}, as shown in Figure 4.5b. We find that additional and complementary information is also obtained from $A_{\rm FB}^{\ell}$ measured in the interference region (off-peak). To quantify the study for a Z' mass of 1.5 TeV, "on-peak events" are counted if the dilepton mass is found in the interval 1.45 TeV $\leq M_{\ell\ell} \leq$ 1.55 TeV. The "interference region" is defined accordingly and satisfy 1 TeV $\leq M_{\ell\ell} \leq$ 1.45 TeV.

Finally, the rapidity distribution is analyzed. Figure 4.6a shows the normalized distributions for a Z' with a mass of 1.5 TeV produced from $u\bar{u}$ (green solid line), $d\bar{d}$ (red dashed line) and sea-antisea quark annihilation (blue dashed-dotted line), which depends only on the parton distribution functions. Especially the Z' rapidity distribution from $u\bar{u}$ annihilation appears to be significantly different from the other two distributions. Figure 4.6b shows the expected rapidity distribution for the Z'_η model. A particular Z' rapidity distribution is fitted using a linear combination of the three pure quark-antiquark rapidity distributions shown in Figure 4.6b. The fit output gives the $u\bar{u}$, $d\bar{d}$ and *sea* quarks fraction in the sample. This will thus reveal how the Z' couples to different quark flavors in a particular model.

In order to demonstrate the analysis power of this method, we also show the rapidity distribution in the case of the Z'_ψ boson, which has equal couplings to up-type and down-type quarks. As can be qualitatively expected from the distributions shown in Figure 4.6a, the used fitting procedure provides very accurate results for the known generated fraction $R_{u\bar{u}}$ of $u\bar{u}$/all. Some correlations between $d\bar{d}$ and the sea-antisea Z' production limits the accuracy of the measurement for the $d\bar{d}$ fractions. For example, for the Z'_η model, the generated event fractions from $u\bar{u}$, $d\bar{d}$ and sea-antisea quarks are 0.71, 0.26 and 0.03 respectively. The corresponding numbers from the fit and 100 fb^{-1} are 0.71±0.07, 0.29±0.08 and 0.01±0.02.

Table 4.2 shows the value of the cross section times the total decay width, the forward-backward charge asymmetry for the on-peak and interference regions as defined above, and the ratio of Z' events produced from $u\bar{u}$ annihilation as obtained from the fit to the Z' rapidity distribution.

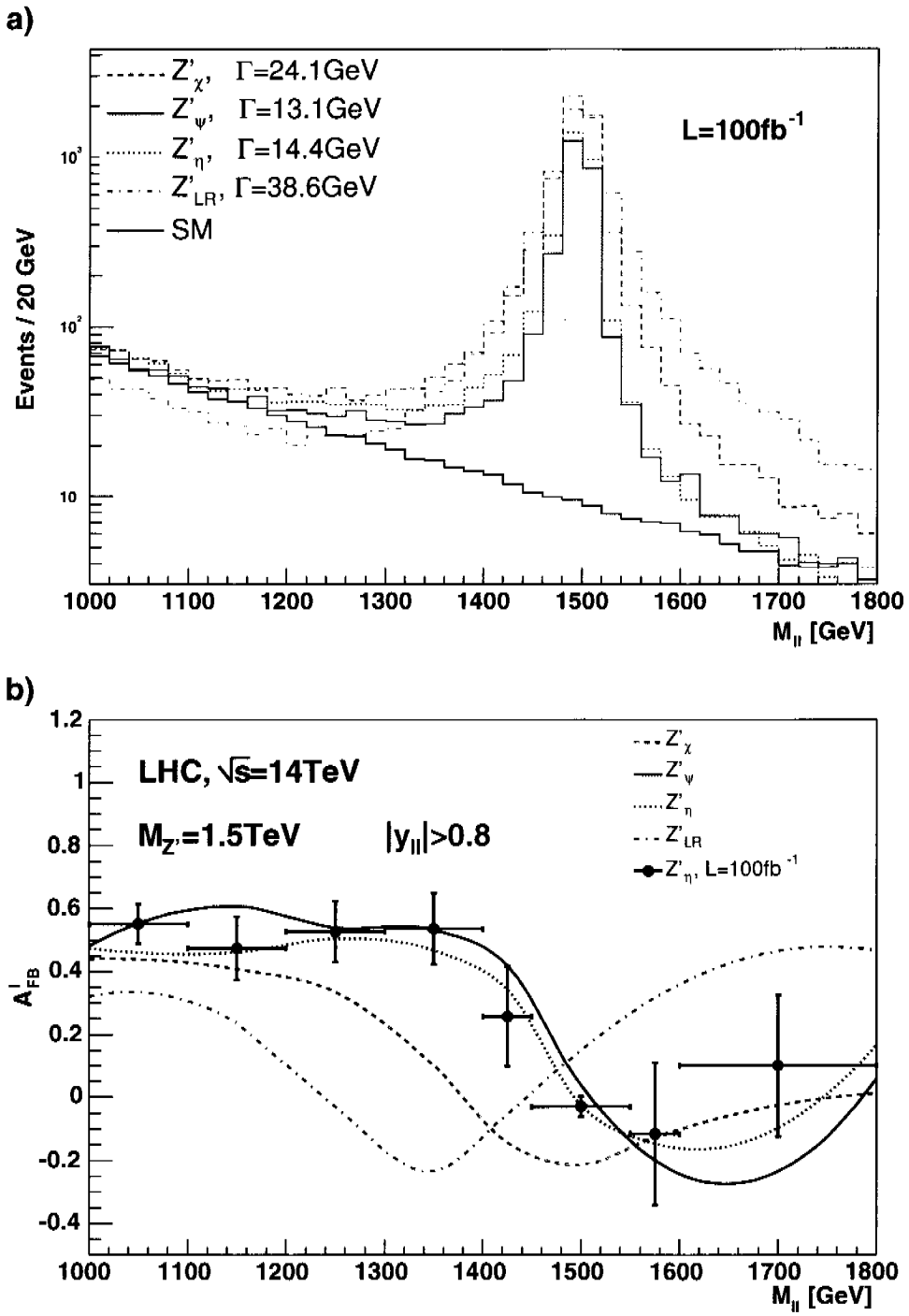

FIGURE 4.5: *The dilepton invariant mass spectrum (a) and A_{FB}^{ℓ} (b) as a function of $M_{\ell\ell}$ for four Z' models. For the forward-backward charge asymmetry, the rapidity of the dilepton system is required to be larger than 0.8. A simulation of the statistical errors, including random fluctuations of the Z'_η model and with errors corresponding to a luminosity of 100 fb^{-1} has been included in (b).*

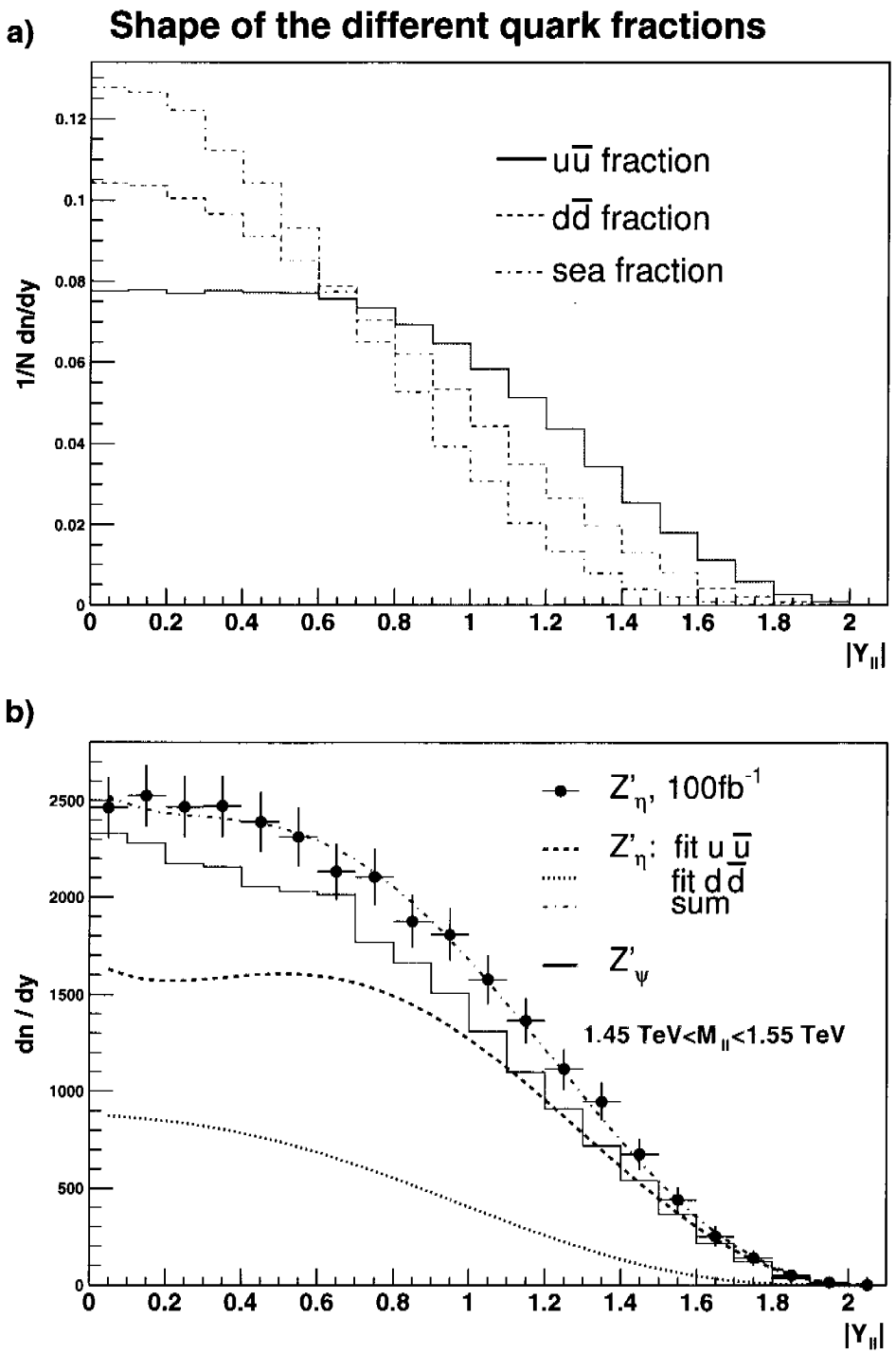

FIGURE 4.6: *The normalized rapidity distribution of* Z′ *with a mass of* 1.5 ± 0.05 *TeV produced from the different types of quarks (a). The observable rapidity distribution for two different* Z′ *models is shown in (b), including the fit results that determine the different types of* $q\bar{q}$ *fractions.*

Model	$\sigma_{ll}^{3\Gamma} \times \Gamma$ [fb·GeV]	$A_{\mathrm{FB}}^{\mathrm{on-peak}}$	$A_{\mathrm{FB}}^{\mathrm{off-peak}}$	$R_{u\bar{u}}$
Z'_{ψ}	487 ± 5	0.04 ± 0.03	0.53 ± 0.04	0.60 ± 0.07
Z'_{η}	630 ± 20	−0.03 ± 0.03	0.45 ± 0.04	0.71 ± 0.07
Z'_{χ}	2050 ± 40	−0.23 ± 0.02	0.26 ± 0.05	0.22 ± 0.05
Z'_{LR}	3630 ± 80	0.15 ± 0.02	0.06 ± 0.06	0.45 ± 0.05
Z'_{SM}	8000 ± 140	0.07 ± 0.02	0.18 ± 0.03	0.05 ± 0.04
Z'_{d}	1520 ± 40	−0.50 ± 0.02	0.26 ± 0.05	0.00 ± 0.01

TABLE 4.2: *The values of the four basic observables, the signal cross section, multiplied by the total width, the forward-backward charge asymmetry on- and off-peak, and the ratio $R_{u\bar{u}}$ for various Z' models and with a Z' mass of 1.5 TeV. The quoted statistical errors are those that can be expected for a luminosity of 100 fb^{-1}.*

4.3 Distinction between models and parameter determination

Let us now discuss how well the different Z' models can be distinguished experimentally using the observables defined before: $\sigma_{\ell\ell}^{3\Gamma} \cdot \Gamma$, A_{FB}^{ℓ} on- and off-peak, as well as $R_{u\bar{u}}$ as obtained from the rapidity distribution. As a working hypothesis, a luminosity of 100 fb^{-1} and a Z' mass of 1.5 TeV will be assumed in the following.

A precise knowledge of the cross section times the total width allows a first good distinction to be made between some models, as shown in the upper two plots of Figure 4.7. It is not obvious how accurately absolute cross sections can be measured and interpreted at the LHC. However, following the procedure outlined in [8], comparable reactions, in this case Z' and Z boson production, should be counted with respect to each other. The use of such ratio measurements should allow us to minimize systematic uncertainties, and an accuracy of ±1% might be achievable [8]. As can be seen from the other plots in Figure 4.7, the additional variables show a different sensitivity for the different couplings.

For example, very similar cross sections are expected for the E_6 Z' models with $\cos\beta \sim \pm 1$ and for left-right models with $\alpha_{LR} \lesssim 1.3$. However, these two different models show a very different behavior for on- and especially off-peak asymmetries and for the couplings to up-type and down-type quarks. Obviously, the maximum sensitivity can be obtained by using all observables together. Having said this, one also needs to point out that some ambiguities between the different models remain, even after a complete analysis of 100 fb^{-1} of data.

Assuming that a particular model has been selected, one would like to know how well the parameter(s), such as $\cos\beta$ or α_{LR}, can be constrained. In the case of the E_6 model for instance, one finds that $\cos\beta$ cannot always be determined unambiguously. Very similar results can be expected for different observables but using very different values for $\cos\beta$. Again, the combination of the various measurements helps to reduce some ambiguities.

If the Z' mass is increased, the number of events decreases drastically and the differences between the models start to become covered within the statistical fluctuations. For the assumed luminosity of 100 fb^{-1}, we could still distinguish a Z'_{χ} from a Z'_{LR} over a large parameter range; the A_{FB}^{ℓ} measurements provide some statistical significance up to $M_{Z'} = 2$ 2.5 TeV. On the contrary, a Z'_{η} could be differentiated from a Z'_{ψ} only up to a Z' mass of at most 2 TeV as, in that case, the dependence of A_{FB}^{ℓ} is almost identical in the two models.

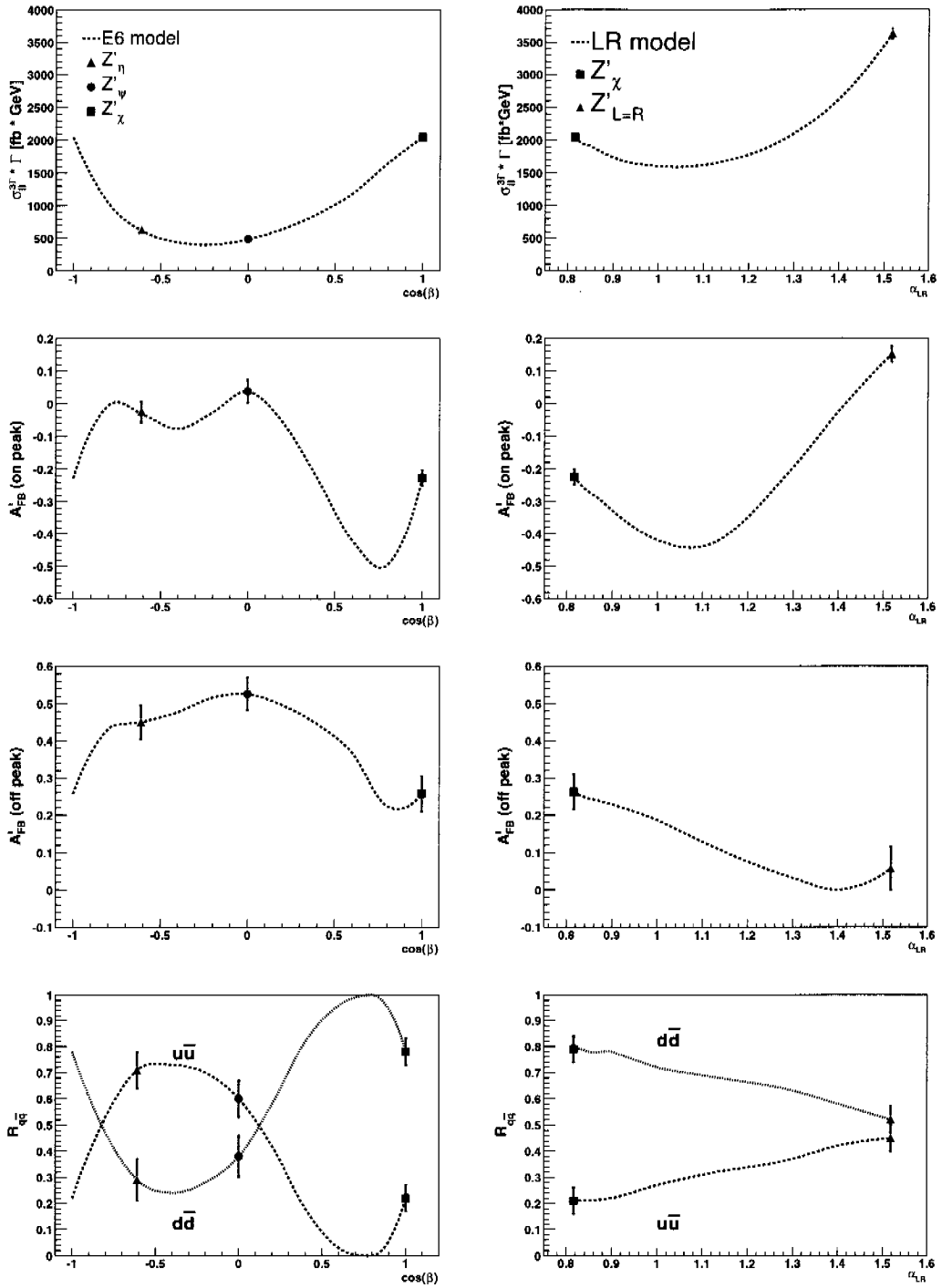

FIGURE 4.7: *Variation of* $\sigma_{\ell\ell}^{3\Gamma} \cdot \Gamma$, $A_{FB}^{\text{on-peak}}$, $A_{FB}^{\text{off-peak}}$ *and the ratio* $R_{q\bar{q}}$ *as a function of the* E_6 *model parameter* $\cos\beta$ *(left) and the LR-model parameter* α_{LR}. *The points corresponding to the particular* Z′ *models are also shown together with the statistical errors corresponding to the number of events for a luminosity of 100* fb^{-1}.

4.4 Where a full detector simulation might be important

For the study described above, PYTHIA was used together with kinematics cuts to reproduce the detector acceptance. One could wonder whether there might be additional detector effects that could alter these results or if only some minor changes are expected. As the experimental signature is very clean with a low background, experimental effects are expected to result in only minor modification of the sensitivity.

- The lepton selection

  The electron selection described in Chapter 3 has to be adapted to very high energy electrons. For instance, the energy leakage in the HCAL will force us to release the cut on E_{had}/E_{em}. However the electron selection can be done in a looser way as only small background is expected.

  The identification of high energy muons should also be studied.

- The mass peak reconstruction

  The accuracy of the mass peak reconstruction will depend mainly on the energy resolution (for electrons it is dominated by the constant term). Since the muon energy resolution is low at such high energies it might be required to measure the Z' width using only the electrons. For electrons, problems related to saturation effects should be correctly taken into account. In a study done for Randall-Sundrum excitations of gravitons at an energy if 1.5 TeV (an experimental signature that looks like a Z' except for the spin 2 nature of the graviton), the variable $E_{3\times3} - E_{2\times2}$ (the difference between the energy in a 3 times 3 crystal array centered on the seed crystal and the energy in the most energetic 2 times 2 crystal array containing the seed crystal) was advocated to correct the energy saturation [84].

- The detector acceptance and $A_{\rm FB}^{\ell}$

  As the detector acceptance is expected to be charge symmetric, it should not have any influence on the measurement of the forward-backward asymmetry. Figure 4.8 shows the effect of the detector acceptance with the fast simulation done with PYTHIA on the $\cos\theta^*$ distribution for a 1.5 TeV Z'_χ.

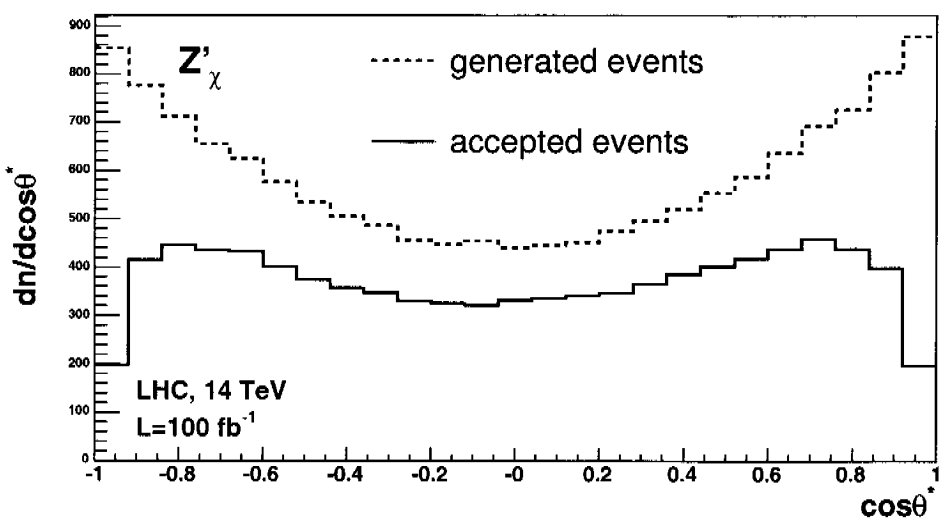

FIGURE 4.8: *The $\cos\theta^*$ distribution for a 1.5 TeV Z'_χ for all generated events (black dotted line) and for accepted events within PYTHIA (red solid line).*

A potential problem for the determination of $A_{\rm FB}^{\ell}$ could be the charge measurement, as at those high momenta the track curvature is very small. However, following the CMS tracker TDR [21], the charge mis-assignment probability for tracks of $p_t = 1$ TeV is about 10^{-4} in the barrel region and increases up to 0.5% at the largest pseudorapidity, as Figure 4.9 shows.

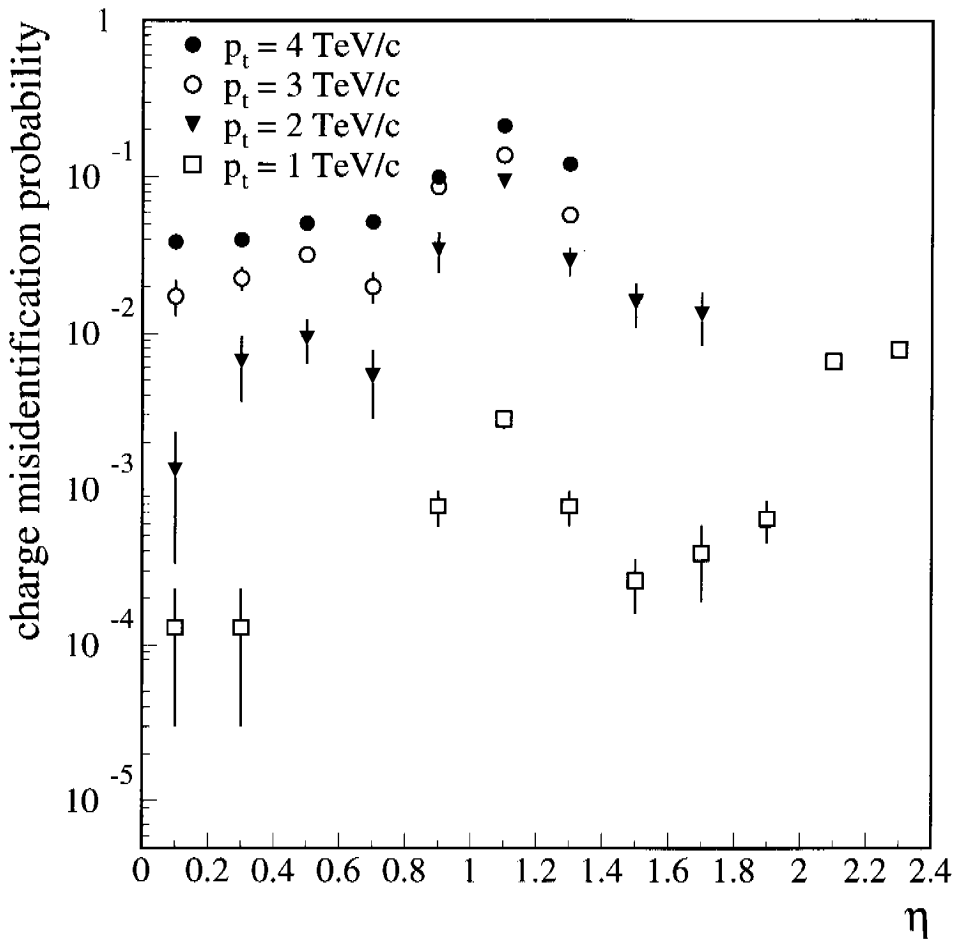

FIGURE 4.9: *Charge misidentification probability for high p_t tracks as a function of pseudorapidity [21].*

- The rapidity distribution and the cross section

  A variation in the PDF can influence quite a lot the Z′ rapidity distribution. However as the quark PDF are known to a sufficient accuracy and could be constrained using a Drell-Yan or single Z samples, this should not have a significant effect on the determination of the Z′ rapidity distribution.

  Figure 4.10a shows the expected rapidity distribution for a 1.5 TeV Z'_{ψ} using two different sets of PDF, MRST99(c-g) and CTEQ5L. The difference between the two distributions are less than 10%. A way to reduce even further this uncertainty is to use Drell-Yan processes to constrain the PDF. Figure 4.10b shows how the rapidity distribution of Drell-Yan lepton pairs would look like at high mass for the two PDF sets. The rapidity distribution of Drell-Yan pairs can be accurately measured at a smaller energy scale and extrapolated at an energy of 1.5 TeV using the DGLAP evolution. One can see on Figure 4.10c that when the ratio between the Z'_{ψ} and Drell-Yan rapidity distribution is taken, the dependence on the PDF practically disappears.

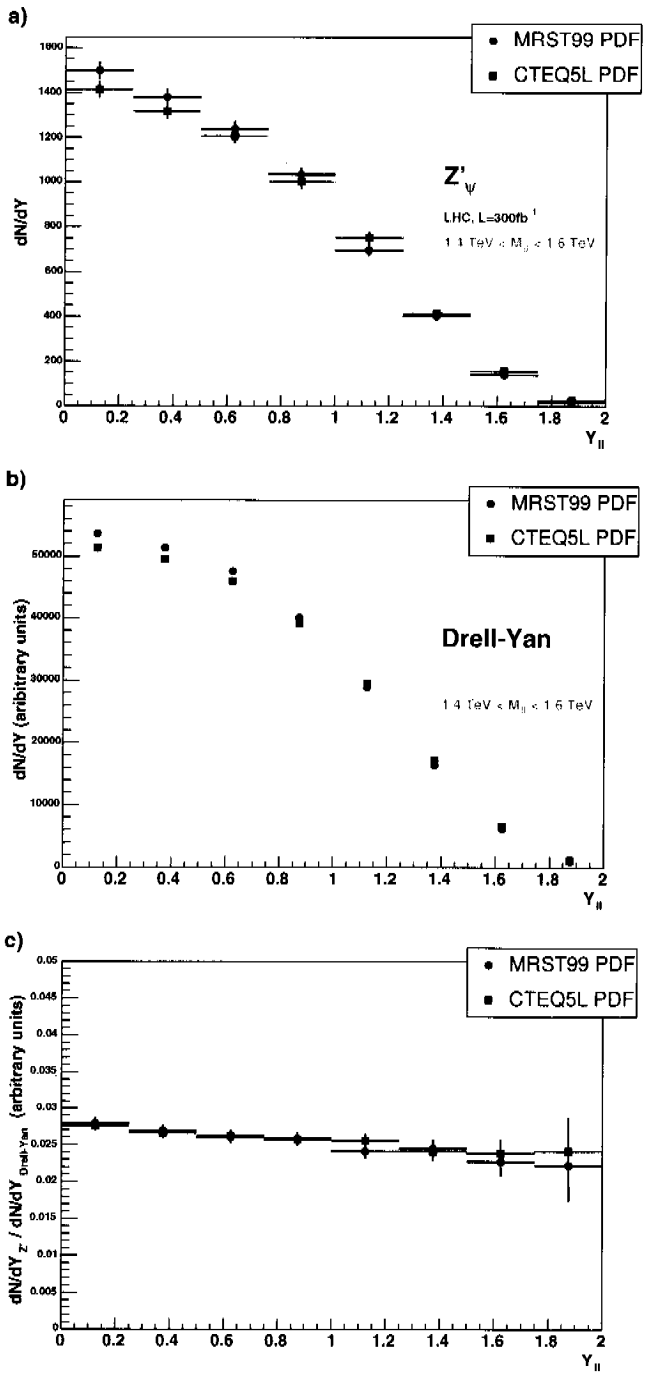

FIGURE 4.10: *a) The expected rapidity for a Z'_ψ for two different PDF sets: MRST99(c-h) [85] (black dots) and CTEQ5L (soft valence) [86] (red squares). The numbers correspond to a luminosity of 300fb^{-1}. b) The same but for the Drell-Yan process with $1.4 < m_{\ell\ell} < 1.6$ TeV. c) The ratio between the Z'_ψ and Drell-Yan for the two different PDF.*

The same procedure can be used to measure the cross section for the Z′ to a good accuracy. This could be done by normalizing it using a single Z sample.

This work concentrated mainly on electrons. CMS discovery capacities for a Z′ decaying into two muons have been studied with a full detector simulation [75]. Figure 4.11 shows the CMS discovery potential for this channel. So far no study on the forward-backward asymmetry using a full detector simulation has been carried out.

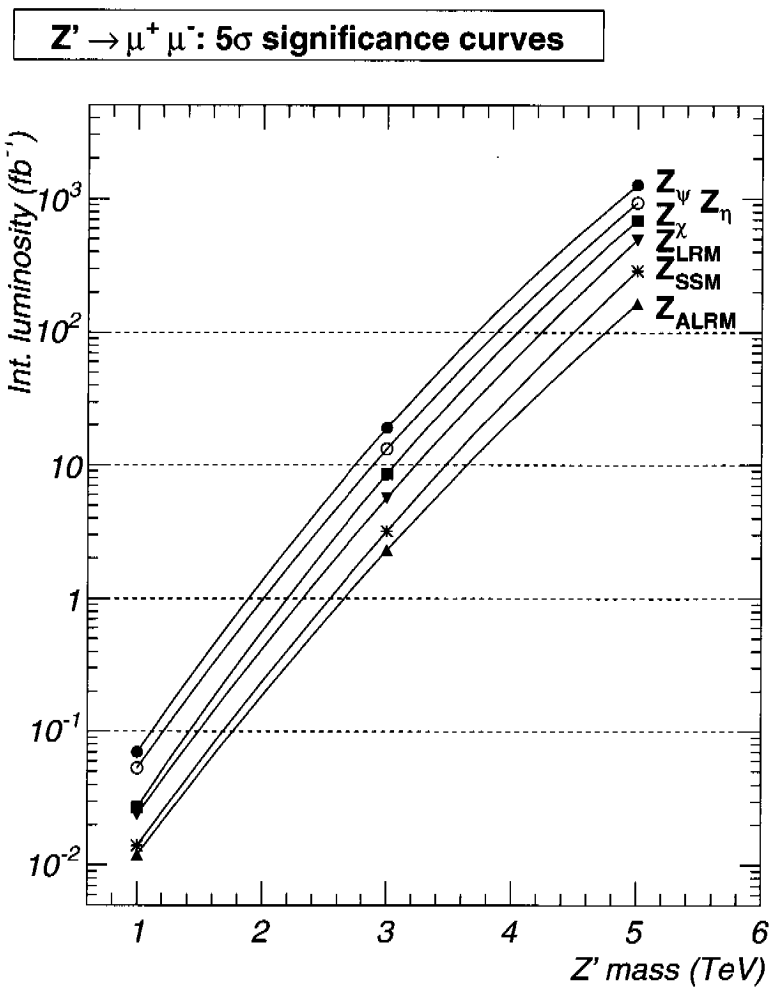

FIGURE 4.11: *CMS discovery potential for Z′ decaying into two muons [75].*

In summary, a realistic simulation of the study of the properties of Z′ bosons originating from various theoretical models has been performed for the LHC. We have shown that, in addition to the Z′ production cross section times total decay width, the measurement of the forward-backward lepton charge asymmetry, both on the Z′ peak and in the interference region, provide complementary information. We have also shown that a fit of the rapidity distribution can provide a sensitivity to the Z′ couplings to up-type and down-type quarks. The combination of all these observables would allow us to discriminate between Z′ bosons of different models or classes of models for masses up to 2–2.5 TeV, if a luminosity of 100 fb^{-1} is collected.

Conclusion

The CMS experiment is currently expected to start in 2007 with goals as different as searching for the Higgs boson, supersymmetric particles or try to measure accurately the mass of the top quark. To reach the goals of such a broad physics program it will be fundamental to obtain precision measurements. A way to improve the measurement accuracy is to use reference processes that are theoretically well-known and experimentally well-measured in order to understand detector systematics and to tune theoretical models. For this purpose, single W and Z production processes are very well suited, since their leptonic decays provide very clean signatures. These processes can be used to measure the luminosity with high precision, constraining at the same time the parton distribution functions, a procedure called the parton luminosity. This will allow, for instance, to obtain more precise cross sections measurements.

This thesis gives examples on how to detect and measure precisely single W/Z production, concentrating on their decays into electrons, for two high energy experiments, CDF, a running experiment located at the Tevatron proton-antiproton collider (running at a center-of-mass energy of 1.96 TeV) and CMS, an experiment located on the LHC proton-proton accelerator (running at a center-of-mass energy of 14 TeV) which should start data taking in 2007.

First the $p\bar{p} \rightarrow W \rightarrow e\nu$ and $p\bar{p} \rightarrow Z \rightarrow ee$ processes were selected from the CDF data collected between February 2002 and May 2003. The total luminosity was measured combining the calculated cross section, the number of signal events and the selection efficiency. The accuracy obtained was comparable to the measurement done with the 'traditional' luminosity counters. With higher statistics it would be possible to determine the parton luminosity by measuring the number of single W/Z events in different rapidity intervals. This should be possible with CMS, where a high rate of W and Z will be produced. Applying this method at the LHC, should reduce the uncertainty on the luminosity from 5% down to 1-2%.

In a second step, the reconstruction of such processes was studied for CMS, concentrating on the electron selection and using a full detector simulation. One particularity of CMS, compared to CDF, is its crystal electromagnetic calorimeter with a very fine granularity providing an excellent energy resolution. A CMS specific problem is the large amount of material in the front of the calorimeter, causing electrons to emit Bremsstrahlung. We showed that electrons are expected to be efficiently selected in the CMS detector using the following variables: the track isolation, E_{had}/E_{EM}, $E_{3\times3}/E_{5\times5}$, $\sigma_{\eta\eta}$, E_{SC}/p_{track} and $|\phi_{SC} - \phi_{tr}^{prop}|$. This will allow to measure processes like single W/Z production with a very good accuracy. Some systematics effects like the energy and Bremsstrahlung dependence of these variables were discussed. The detector homogeneity and the expected energy resolution from the simulation was also studied.

Finally, to illustrate the capability of CMS to do precision measurements, the ability to determine the properties of Z′ bosons with CMS were studied using a fast simulation. Various models inspired from grand unified theories or Superstring theories predict the existence of a new high mass neutral vector boson, the Z′. After the discovery of such a particle, the next step would be to determine which model could describe the measured properties. We have shown that, in addition to the Z′ production cross section (which is expected to be precisely measured by comparing it to the single Z cross section) and its total decay width, the measurement of the forward-backward lepton charge asymmetry, both on the Z′ peak and in the interference region, provide complementary information. We have also shown that a fit of the rapidity distribution

can provide a sensitivity to the Z′ couplings to up-type and down-type quarks. The combination of all these observables would allow to discriminate between Z′ bosons of different models or classes of models for masses up to 2–2.5 TeV, if a luminosity of 100 fb^{-1} is collected by CMS.

Bibliography

[1] The plots are taken from http://particleadventure.org.

[2] V. M. Abazov *et al.* [D0 Collaboration], Nature **429** (2004) 638 [arXiv:hep-ex/0406031].

[3] R. Barate *et al.*, Phys. Lett. B **565** (2003) 61 [arXiv:hep-ex/0306033].

[4] PH-HIGGS, Summary of the CMS Potential for the Higgs Boson Discovery, CERN-CMS-NOTE-2003-033.

[5] M. Dittmar, [arXiv:hep-ex/9901009].

[6] *see Tevatron Run I measurement of the top cross section* ($t\bar{t} \to \ell\nu + 4jets$)*:*
DØ: V. M. Abazov et al., Phys. Rev. D 67 (2003) arXiv:hep-ex/0205019.
CDF: T. Affolder et al., Phys. Rev. D 64 (2001) arXiv:hep-ex/0101036.

[7] F. Pauss, private communication.

[8] M. Dittmar, F. Pauss and D. Zürcher, Phys. Rev. **D56** (1997) 7284, [arXiv:hep-ex/9705004].

[9] *see Tevatron Run I measurement of the* $Z \to ee$ *cross section:*
DØ: B. Abbott et al., Phys. Rev. D 60 (1999) arXiv:hep-ex/9901040.
CDF: F. Abe et al., Phys. Rev. Lett. 76 (1996) arXiv:hep-ex/9509010.

[10] *see Tevatron Run IB measurement of the dijet cross section:*
DØ: B. Abbott et al., Phys. Rev. D 64 (2001) arXiv:hep-ex/0012046.
CDF: F. Abe et al., Phys. Rev. Lett. 77 (1996) arXiv:hep-ex/9601008.
and T. Affolder et al., Phys. Rev. D 64 (2001) arXiv:hep-ph/0102074.

[11] ATLAS Collaboration, Technical Design Report, Detector and Physics Performance, CERN Report No. CERN/LHCC99-15, ATLAS TDR 15, May 1999.

[12] A. Straessner, arXiv:hep-ex/0405005.

[13] [CDF Collaboration], arXiv:hep-ex/0311039.

[14] M. Dittmar, A.-S. Nicollerat, High mass Higgs studies using gluon fusion and weak boson fusion at the LHC, CERN-CMS-NOTE-2001-036.

[15] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **73** (1994) 225 [arXiv:hep-ex/9405005].

[16] See the following web site: http://www-bdnew.fnal.gov/operations/lum/lum.html.

[17] CDF Run II Collaboration, The CDF Run II technical design report, Fermilab-Pub-96-390-E, www-cdf.fnal.gov/internal/upgrades/tdr/tdr.html.

[18] S. Klimenko, J. Konigsberg, T. Liss, CDF note 6314.

[19] See the following web site: http://acwebimages.cern.ch/.

[20] From a talk from R. Schmidt given at the SUSSP summer school, http://www.ippp.dur.ac.uk/sussp57/.

[21] *The general report but a bit outdated:*
CMS Collaboration, G.L. Bayatian et al. , Technical Proposal, CERN Report No. CERN/LHCC 94-38, LHCC/P1, 15. Dec. 1994.
The specific technical design reports:
The CMS Magnet Project, Technical Design Report, CERN/LHCC 97-10, May 1997.
The CMS HCAL Technical Design Report, CERN/LHCC 97-31, June 1997.
The CMS ECAL Technical Design Report, CERN/LHCC 97-33, December 1997.
The CMS MUON Technical Design Report, CERN/LHCC 97-32, December 1997.
The CMS Tracker Project, Technical Design Report, CERN/LHCC 98-6, April 1998
Addendum: CERN/LHCC 2000-016, February 2000.
The CMS TriDAS Project, Technical Design Report, Volume 1: The Trigger Systems, CERN/LHCC 2000 - 38, December 2000.
The CMS TriDAS Project, Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger, CERN/LHCC 2002 - 26, December 2002.

[22] G. Dewhirst, R. Brunelière, Energy resolution of the CMS ECAL barrel super-module using MGPA electronics, CERN-CMS-RN-2004/004.

[23] Plot from J. Strinling taken from: S. Catani *et al.*, [arXiv:hep-ph/0005025].

[24] LHC / LC Study Group Report: http://www.ippp.dur.ac.uk/ georg/lhclc/.

[25] R. L. Gluckstern, Nucl. Instrum. Meth. **24** (1963) 381.

[26] Ariane Frey private communication.

[27] CMS Note RN-2004/004, G. Dewhirst, R. Bruemière, Energy resolution of the CMS ECAL barrel super-module using MGPA electronics.

[28] See for instance: K. Ellis, J. Stirling, B. Webber, QCD and particle physics, Cambridge University Press, 1996.

[29] J. C. Collins, D.E. Soper, Ann. Rev. Nucl. Part. Sci **37** (1987) 383.

[30] V. N. Gribov and L. N. Lipatov, Yad. Fiz. **15** (1972) 781 [Sov. J. Nucl. Phys. **15** 438]
V. N. Gribov and L. N. Lipatov, Yad. Fiz. **15** (1972) 1218 [Sov. J. Nucl. Phys. **15** 675]
L. N. Lipatov, Sov. J. Nucl. Phys. **20** (1975) 94
G. Altarelli and G. Parisi, Nucl. Phys. **B126** (1977) 298
Yu. L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641.

[31] S. Moch, J. A. M. Vermaseren and A. Vogt, [arXiv:hep-ph/0403192].

[32] All these numbers are taken from the talk of J. Stirling given at the HERA-LHC workshop, CERN, March 26th 2004.

[33] L. W. Whitlow, E. M. Riordan, S. Dasu, S. Rock and A. Bodek, Phys. Lett. B **282** (1992) 475.

[34] A. C. Benvenuti *et al.* [BCDMS Collaboration], Phys. Lett. B **223** (1989) 485.
A. C. Benvenuti *et al.* [BCDMS Collaboration], Phys. Lett. B **237** (1990) 599.

[35] M. Arneodo *et al.* [New Muon Collaboration], Nucl. Phys. B **483** (1997) 3 [arXiv:hep-ph/9610231].

[36] M. R. Adams *et al.* [E665 Collaboration], Phys. Rev. D **54** (1996) 3006.

[37] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **13** (2000) 609 [arXiv:hep-ex/9908059]
C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **19** (2001) 269 [arXiv:hep-ex/0012052]
C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **21** (2001) 33 [arXiv:hep-ex/0012053].

[38] J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **12** (2000) 35 [arXiv:hep-ex/9908012].
S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **21** (2001) 443 [arXiv:hep-ex/0105090].

[39] G. Moreno *et al.*, Phys. Rev. D **43** (1991) 2815.

[40] R. S. Towell *et al.* [FNAL E866/NuSea Collaboration], Phys. Rev. D **64** (2001) 052002 [arXiv:hep-ex/0103030].

[41] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **81** (1998) 5754 [arXiv:hep-ex/9809001].

[42] B. Abbott *et al.* [D0 Collaboration], Phys. Rev. Lett. **86** (2001) 1707 [arXiv:hep-ex/0011036].

[43] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. D **64** (2001) 032001 [Erratum-ibid. D **65** (2002) 039903] [arXiv:hep-ph/0102074].

[44] W. G. Seligman *et al.* [CCFR Collaboration], Phys. Rev. Lett. **79** (1997) 1213.
U. K. Yang *et al.* [CCFR/NuTeV Collaboration], Phys. Rev. Lett. **86** (2001) 2742 [arXiv:hep-ex/0009041].
M. Goncharov *et al.* [NuTeV Collaboration], [arXiv:hep-ex/0102049].

[45] MRST PDF: http://durpdg.dur.ac.uk/hepdata/mrs.html/.

[46] CTEQ PDF: http://durpdg.dur.ac.uk/hepdata/cteq.html/.

[47] Alekhin PDF: http://sirius.ihep.su/ alekhin/pdfa02/.

[48] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. C **14** (2000) 133 [arXiv:hep-ph/9907231].

[49] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. C **28** (2003) 455, [arXiv:hep-ph/0211080].
A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, [arXiv:hep-ph/0308087] and private communication.

[50] To be accepted, a run has to fulfill quality criteria based on the tracking chamber and on the calorimeter. The runs which are selected are the same than the ones used for the W/Z cross section measurement:
www-cdf.fnal.gov/internal/physics/ewk/tools_and_datasets/good_run_list.html.

[51] Everything about ROOT: http://root.cern.ch.

[52] The Stntuple analysis frame is written and maintained by Pasha Murat. The version used here was dev_240.

[53] see http://ncdf41.fnal.gov/murat/Stntuple/datasets.html.

[54] G. Corcella and M. H. Seymour, Nucl. Phys. B **565** (2000) 227 [arXiv:hep-ph/9908388].
G. Miu and T. Sjöstrand, Phys. Lett. B **449** (1999) 313 [arXiv:hep-ph/9812455].

[55] G. Davatz, G. Dissertori, M. Dittmar, M. Grazzini and F. Pauss, JHEP **0405** (2004) 009 [arXiv:hep-ph/0402218].

[56] see Monte Carlo GENERATORS for CMS: http://cmsdoc.cern.ch/cms/generators/.

[57] Everything about GEANT4: http://wwwasd.web.cern.ch/wwwasd/geant4/geant4.html.

[58] Everything about ORCA: http://cmsdoc.cern.ch/orca.

[59] Everything about POOL: http://lcgapp.cern.ch/project/persist/.

[60] PAX homepage: http://www-ekp.physik.uni-karlsruhe.de/ erdmann/paxguide/.
See also: M. Erdmann *et al.*, eConf **C0303241** (2003) THLT008 [arXiv:physics/0306085].

[61] C. W. Fabjan and T. Ludlam, Ann. Rev. Nucl. Part. Sci. **32** (1982) 335.

[62] E. Meschi et al, Electron reconstruction in the CMS Electromagnetic Calorimeter, CERN-CMS-NOTE-2001-034
and see also the ORCA web page: http://cmsdoc.cern.ch/orca.

[63] G. Karapostoli, Validation of OSCAR Simulation Performance of CMS, CERN-CMS-IN-2004/004.

[64] C. Seez, The CMS trigger system, CERN-CMS-CR-2003-008.

[65] *Unification using SU(5):* H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32** (1974) 438.
Unification using SO(10): H. Fritzsch and P. Minkowski, Annals Phys. **93** (1975) 193.
Unification using E_6 (reviews and references): D. London and J. L. Rosner, Phys. Rev. D **34** (1986) 1530.

[66] J. L. Hewett and T. G. Rizzo, Phys. Rept. **183** (1989) 193;
A. Leike, Phys. Rept. **317** (1999) 143 [arXiv:hep-ph/9805494];
M. Cvetič and P. Langacker, [arXiv:hep-ph/9707451].

[67] See for instance, C. T. Hill and E. H. Simmons, Phys. Rept. **381** (2003) 235 [Erratum-ibid. **390** (2004) 553] [arXiv:hep-ph/0203079].

[68] N. Arkani-Hamed et al., JHEP 0208 (2002) 021, [arXiv:hep-ph/0206021].
For the phenomenological aspects, see: T. Han et al., Phys. Rev. D67 (2003) 095004, [arXiv:hep-ph/0301040];
J. Hewett, F. Petriello and T. Rizzo, JHEP **0310** (2003) 062, [arXiv:hep-ph/0211218];
C. Csaki et al., Phys. Rev. D **68** (2003) 035009, [arXiv:hep-ph/0303236].

[69] R. W. Robinett and J. L. Rosner, Phys. Rev. D **25** (1982) 3036 [Erratum-ibid. D **27** (1983) 679].

[70] The LEP Electroweak Working Group and the SLD Heavy Flavour Group, Note LEPEWWG/2003-01; http://lepewwg.web.cern.ch/LEPEWWG.

[71] R. Casalbuoni et al., Phys. Lett. B460 (1999) 135, [arXiv:hep-ph/9905568];
J. Rosner, Phys. Rev. D **61** (2000) 016006, [arXiv:hep-ph/9907524];
J. Erler and P. Langacker, Phys. Rev. Lett. **84** (2000) 212, [arXiv:hep-ph/9910315].

[72] For an account, see J. Erler and P. Langacker, Phys. Lett. B **456** (1999) 68 [arXiv:hep-ph/9903476];
see also for an updated analysis of LEP1 limits, F. Richard, [arXiv:hep-ph/0303107].

[73] M. Cvetič, S. Godfrey et al., DPF study in *Electroweak Symmetry Breaking and Beyond the Standard Model*, [arXiv:hep-ph/9504216];
T.G. Rizzo, [arXiv:hep-ph/9612440];
S. Godfrey, in *Proc. of the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001)* ed. N. Graf, eConf **C010630** (2001) E3065, [arXiv:hep-ph/0201092] and [arXiv:hep-ph/0201093].

[74] Proc. of the ECFA Large Hadron Collider Worshop, Aachen (Germany) 1990, Reports CERN 90–10 and ECFA 90–133.

[75] PRS-Analysis Note, R. Cousins et al., Detection of Z' Gauge Bosons in the Di-muon Decay Mode, CERN-CMS-AN-2004-002.

[76] P. Langacker, R. W. Robinett and J. L. Rosner, Phys. Rev. D **30** (1984) 1470;
V. D. Barger, N. Deshpande, R. J. N. Phillips and K. Whisnant, Phys. Rev. D **33** (1986) 1912 [Erratum-ibid. D **35** (1987) 1741].

[77] M. Dittmar, Phys. Rev. D **55** (1997) 161 [arXiv:hep-ex/9606002].

[78] F. del Aguila, M. Cvetic and P. Langacker, Phys. Rev. D **48** (1993) 969 [arXiv:hep-ph/9303299].

[79] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, Comput. Phys. Commun. **135** (2001) 238 [arXiv:hep-ph/0010017].

[80] G. Azuelos and G. Polesello, [arXiv:hep-ph/0204031];
see also, T. Rizzo, JHEP **0306** (2003) 021, [arXiv:hep-ph/0305077].

[81] A. Djouadi, A. Leike, T. Riemann, D. Schaile and C. Verzegnassi, Z. Phys. C **56** (1992) 289.

[82] G. Altarelli, R. K. Ellis and G. Martinelli, Nucl. Phys. B **143** (1978) 521 [Erratum-ibid. B **146** (1978) 544];
J. Kubar-Andre and F. E. Paige, Phys. Rev. D **19** (1979) 221.

[83] H. L. Lai *et al.* [CTEQ Collaboration], Eur. Phys. J. C **12** (2000) 375, [arXiv:hep-ph/9903282].

[84] CMS Note 2004/xxx, M.C. Lemaire, C. Collard, Search with the CMS Detector for Randall-Sundrum excitations of gravitons decaying into an electron pair.

[85] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. C **4** (1998) 463 [arXiv:hep-ph/9803445].

[86] H. L. Lai *et al.* [CTEQ Collaboration], Eur. Phys. J. C **12** (2000) 375 [arXiv:hep-ph/9903282].

[87] M. Dittmar, A.-S. Nicollerat, CDF note 6110 (September 6, 2002).

[88] M. Dittmar, A.-S. Nicollerat, CDF note 6118 (September 18, 2002).

[89] M. Dittmar, A.-S. Nicollerat, CDF note 6411 (April 9, 2003).

[90] M. Dittmar, A-S. Giolo-Nicollerat, Identifying central electrons in the CMS detector, CERN-CMS-IN-2004-036.

[91] M. Dittmar, A. S. Nicollerat and A. Djouadi Phys. Lett. B **583** (2004) 111 [arXiv:hep-ph/0307020].

[92] B. C. Allanach *et al.* [Beyond the Standard Model Working Group Collaboration], [arXiv:hep-ph/0402295].

[93] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **91** (2003) 241804 [arXiv:hep-ex/0307080].
D. Acosta *et al.* [CDF II Collaboration], Phys. Rev. D **68** (2003) 072004 [arXiv:hep-ex/0310043].
D. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **68** (2003) 091101 [arXiv:hep-ex/0308059].
D. Acosta *et al.* [CDF II Collaboration], Phys. Rev. Lett. **93** (2004) 072001 [arXiv:hep-ex/0312021].
D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **93** (2004) 032001 [arXiv:hep-ex/0403032].
D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **93** (2004) 142001 [arXiv:hep-ex/0404036].
D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **93** (2004) 221802 [arXiv:hep-ex/0406073].
D. Acosta *et al.* [CDF II Collaboration], [arXiv:hep-ex/0406078].
D. Acosta *et al.* [CDF-II Collaboration], "Measurement of the t anti-t production cross section in p anti-p collisions [arXiv:hep-ex/0409029].
D. Acosta *et al.* [CDF II Collaboration], [arXiv:hep-ex/0410008].
D. Acosta *et al.*, [arXiv:hep-ex/0410013].
D. Acosta *et al.* [CDF Collaboration], FERMILAB-PUB-04-148-E
D. Acosta *et al.* [CDF Collaboration], "Measurement of the t anti-t production cross section in p anti-p collisions [arXiv:hep-ex/0410041].
D. Acosta *et al.* [CDF Collaboration], [arXiv:hep-ex/0410053].
D. Acosta *et al.* [CDF Collaboration], [arXiv:hep-ex/0410058].
D. Acosta *et al.* [CDF Collaboration], [arXiv:hep-ex/0410076].
D. Acosta *et al.* [CDF Collaboration], FERMILAB-PUB-04-113-E
D. Acosta *et al.* [CDF Collaboration], arXiv:hep-ex/0411059].