Le calorimetre electromagnetique de CMS et la recherche du boson de Higgs dans le canal de desintegration H $\rightarrow$ WW$^{(*)}$ $\rightarrow 2e2\nu$

Chiara Rovelli

DOCTEUR DE L'ECOLE POLYTECHNIQUE

Spécialité: Physique

par

Chiara Ilaria Rovelli

The CMS electromagnetic calorimeter and the search for the Higgs boson in the decay channel $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$

Soutenue le 10 janvier 2006 devant le jury composé de:

P. Bloch	CERN, Genève	Rapporteur
F. Navarria	Università di Bologna, Bologna	Rapporteur
P. Busson	Ecole Polytechnique, Palaiseau	Directeur de thèse
S. Ragazzi	Università di Milano-Bicocca, Milano	Directeur de thèse

Facoltà di Scienze Matematiche Fisiche e Naturali

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

The CMS electromagnetic calorimeter and the search for the Higgs boson in the decay channel $H \to WW^{(*)} \to 2e2\nu$

Coordinatore: Prof. Claudio Destri
Tutore: Prof. Stefano Ragazzi
Tutore: Dott. Philippe Busson

Tesi di Dottorato di
Chiara Ilaria Rovelli
Matricola R00285

XVIII Ciclo
Anno Accademico 2004–2005

Contents

Introduction

The Standard Model of the electroweak interactions is today one of the best theories of modern physics. Most of its building blocks have been tested up to a very high precision over a large range of energies and an agreement at the per mill level with the theoretical predictions has been found. Despite a so large success, however, some questions are still open. In particular, the origin of the particles masses remains still not explained. The spontaneous symmetry breaking mechanism, which could explain the particles masses, indeed requires the existence of a scalar particle, the Higgs boson, which has not yet been found experimentally.
The search for the Higgs boson and the possibility to investigate the existence of new physics beyond the Standard Model are among the main reasons which led to the project of the Large Hadron Collider, a proton-proton collider which is currently under construction at CERN, Geneva. The first collisions are foreseen for 2007. Four experiments will take data at LHC: two general purpose ones (ATLAS and CMS), an experiment dedicated to heavy ions physics (ALICE) and a fourth one for the study of b-physics (LHCb). The four experiments will exploit the LHC high luminosity and center of mass energy to cover a large physics program.
This thesis presents the work I did within the CMS collaboration in the past three years. Large part of my work was devoted to the development of electron reconstruction tools and aimed at improving the Higgs boson discovery potential of CMS, in particular in the $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$ channel. A major role in the electron reconstruction is played by the electromagnetic calorimeter ECAL, an homogeneous calorimeter made of scintillating $PbWO_4$ crystals. In these years I also joined the efforts for the ECAL construction and development, taking care in particular of some aspects related to its calibration.
The thesis is organized in the following way. The first three chapters give an overview of LHC and CMS. Chapters 4, 5 and 6 summarize my own work and the original contributions I gave.
In chapter 1 the motivations which have brought to the LHC project are presented. The physics of the Higgs boson and the current limits on its mass are discussed, together with some hints on the Higgs boson searches at the LHC.
Chapter 2 gives a general overview of both LHC and the CMS detector. Special

emphasis is given to the electromagnetic calorimeter, which is also the subject of chapter 3 and 4. Chapter 3 describes the techniques which have been foreseen for the ECAL calibration and stability monitoring. In chapter 4 the analysis of the data collected during the 2003 electromagnetic calorimeter test beam is presented. First the problem of the intercalibration at the test beam is addressed. This is a major task, since the precision of the intercalibration directly affects the constant term of the energy resolution, for which the CMS goal is to reach a precision better than 0.5%. The first part of the chapter is dedicated to the study of the statistical precision and of the effects of possible systematics whose contribution is evaluated. The good initial intercalibration, anyway, could be spoiled during the data taking by the effects of the radiation on the crystals, which can change the relative responses of the channels. The problem of monitoring the intercalibration and the check of its stability in time are the subject of the second part of the chapter. A monitoring laser system is foreseen at CMS. The possibility to check the calibration stability and to correct the changes in the response with a precision within the required limits is demonstrated.
Chapter five describes the electron reconstruction and identification in CMS. A crucial problem for the electron reconstruction is represented by the bremsstrahlung emission in the tracker. A tracking procedure dealing with the bremsstrahlung energy loss is discussed. Together with an improvement in the reconstruction efficiency, the procedure allows to identify electrons with a small fraction of radiated energy, which can be usefully exploited for the ECAL calibration. The combination of tracker and calorimetric informations to build electrons and an electron identification strategy are also discussed.
The developed algorithms are applied in chapter 6, which presents the study of the CMS discovery potential of the Higgs boson in the $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$ channel. This is the discovery channel in the range of masses between $2m_W$ and $2m_Z$. Here the possibility to extend the study also to the low mass region up to m_H = 120 GeV/c^2 is investigated. The CMS sensitivity to the $H \rightarrow WW^{(*)}$ channel is presented and the main background sources (WW, $t\bar{t}$ and Wtb production) are discussed, together with a set of selections to reduce them. Some hints to the possibility of reconstructing the Higgs mass despite the absence of a narrow mass peak are given.

Chapter 1

The Higgs boson in the Standard Model

Most of the building blocks of the Standard Model have been successfully tested to a very high precision over a wide range of energies in recent years. However the Higgs particle, which plays a central role in the theory providing the electroweak symmetry breaking which originates the mass of both the gauge bosons and the fermions, has not yet been found experimentally.
Theoretical consistency imposes both upper and lower limits on the Higgs boson mass. A lower limit of about 114 GeV/c^2 has been also set by the direct searches of the experiments which took data at the LEP accelerator. To continue the search beyond this limit a new particle accelerator, the Large Hadron Collider, is currently under construction at CERN and it will cover a wide range of masses up to the TeV scale.
In this chapter the basic concepts of the Standard Model are summarized, with a particular attention devoted to the electroweak symmetry breaking mechanism and to the Higgs sector. The limits on the Higgs mass, set on both theoretical and experimental grounds, are then discussed. Finally an overview of the Higgs production and decay mechanisms at LHC is given and the main search strategies are presented.

1.1 The Standard Model and the Higgs mechanism

The Standard Model is a gauge quantum field theory describing all the fundamental interactions except gravity.
It is based on the SU(3)$\times$SU(2)$\times$U(1) symmetry group. The SU(3) symmetry of Quantum Chromodynamics [1] is related to the rotations in the colour space and it describes the interactions of quarks by means of eight massless coloured gauge

bosons, the gluons. The Electroweak Theory [2] unifies the weak interactions and Quantum Electrodynamics. It has a SU(2)×U(1) symmetry related to rotations in the weak isospin T and in the weak hypercharge Y spaces. The two symmetries are related by the Gell-Mann and Nishijma relation $Q = T_3 + Y/2$, being Q the electric charge and T_3 the third component of the weak isospin.
In the Electroweak Theory quarks and leptons are organized in the following multiplets of the gauge group

$$\begin{array}{rcl} Q_L^i = \binom{u_L^i}{d_L^i} & \longrightarrow & (2, 1/3) \\ u_R^i & \longrightarrow & (1, 4/3) \\ d_R^i & \longrightarrow & (1, -2/3) \\ L_L^i = \binom{\nu_L^i}{l_L^i} & \longrightarrow & (2, -1) \\ l_R^i & \longrightarrow & (1, -2) \end{array} \tag{1.1}$$

The two numbers associated with each multiplet are respectively the dimension of the SU(2) representation and the value of the weak hypercharge, while the index i runs over the three quarks and leptons families.
Four gauge fields correspond to the SU(2)×U(1) group, three related to the SU(2) symmetry (W^i, i = 1, 2, 3) and the last one corresponding to the U(1) symmetry, B. The two charged bosons $W^\pm$ are given by the linear combination of two of them,

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp W_\mu^2 \right) \tag{1.2}$$

The photon A and the neutral boson W^0 are obtained with the combinations

$$\begin{array}{rcl} B_\mu &=& A_\mu \cos\theta_W - Z_\mu \sin\theta_W \\ W_\mu^3 &=& A_\mu \sin\theta_W + Z_\mu \cos\theta_W \end{array} \tag{1.3}$$

where θ_W is the Weinberg angle.

The gauge theories are not compatible with explicit mass terms in the lagrangian because these would break the gauge invariance and lead to non renormalizable theories and to non unitary scattering matrices. An alternative way to give mass both to the fermions and to the vector bosons is the spontaneous symmetry breaking mechanism, first introduced by Higgs [3], which requires the presence of a scalar field with the proper quantum numbers.
In the contest of the Standard Model, the symmetry breaking mechanism has to happen in such a way to keep the photon massless. The simplest way to do that is the use of a SU(2) doublet with hypercharge equal to one,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \qquad Y(\phi) = 1 \tag{1.4}$$

The most general scalar potential consistent with gauge invariance and renormalizability is

$$V(\phi) \;=\; m^2 \mid \phi \mid^2 \;+ \lambda \mid \phi \mid^4 \tag{1.5}$$

Two parameters, λ and m, enter the potential. If $m^2 \geq 0$ the potential has a minimum in ϕ = 0. If $m^2 < 0$ the parameter m can not be interpreted as a squared mass for the field ϕ and the potential has an infinite number of degenerate minima given by all those field configurations for which

$$\mid \phi \mid^2 = \; -\frac{m^2}{2\lambda} \;\equiv\; \frac{1}{2}v^2 \tag{1.6}$$

All these minimum configurations are connected by gauge transformations that change the phase of the complex field without affecting its modulus. The lowest energy state (the vacuum) corresponds to one of these minima and when one of the possible solutions is chosen as vacuum expectation value the symmetry is spontaneously broken. To keep the photon massless, the minimum which stays invariant under the symmetry having the electric charge as generator has to be chosen and this is obtained taking the vacuum expectation value with hypercharge $Y = \pm 1$. Considering the first choice, $Y = 1$, the scalar field ϕ can be parameterized as

$$\phi \;=\; \frac{1}{\sqrt{2}} e^{i\tau^i \theta^i(x)/v} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.7}$$

where $\theta^i(x)$ and $H(x)$ are real and the τ^is are the SU(2) generators. Such parameterization is convenient when working in the unitary gauge θ_i = 0.
The gauge bosons acquire masses through their couplings to the Higgs field,

$$m_W^2 \;=\; \frac{1}{4} g^2 v^2 \tag{1.8}$$

$$m_Z^2 \;=\; \frac{1}{4}\left(g^2 \;+\; g'^2\right) v^2 \tag{1.9}$$

while the photon stays massless. The Higgs boson mass is

$$m_H^2 \;=\; 2\lambda v^2 \tag{1.10}$$

The value of v, the vacuum expectation value of the neutral component of the Higgs doublet, can be obtained from the measured value of the Fermi constant and it is $v \;=\; \sqrt{\frac{1}{G_F\sqrt{2}}} \;\simeq$ 246.22 GeV. The value of the quartic coupling λ is not fixed.
The Higgs mechanism described here to give mass to the vector bosons also works to generate the masses of the fermions without adding symmetry breaking terms

to the lagrangian. This is done introducing Yukawa couplings, which are specific for each charged lepton and quark, between the Higgs doublet and the fermion fields. As a result, the mass terms for the fermions are proportional to the Yukawa couplings,

$$m_f = g_f \frac{v}{\sqrt{2}} \tag{1.11}$$

Since the couplings are arbitrary, the masses are non calculable parameters of the Standard Model.

1.2 Beyond the Standard Model

The Standard Model has undergone intensive experimental tests which have shown an agreement with its predictions at the per mill level up to energies around 100 GeV [4].
Despite such remarkable agreement, anyway, some questions are still open. From an experimental point of view, the Higgs particle necessary for the spontaneous symmetry breaking mechanism has not yet been discovered. A large region of masses allowed for the Higgs boson mass is anyway still unexplored, so the presence of a Standard Model Higgs is not necessarily in contradiction with the present measurements.
From the theoretical point of view, instead, there are some arguments suggesting that the Standard Model is not the ultimate theory but an effective field theory valid up to a certain scale Λ. Among them there are the number of free parameters to be determined experimentally, which is too large for an ultimate theory, and the problem of the unification of the interactions.
One of the main drawback of the Standard Model is then the so called *hierarchy problem* [5], that is the big difference existing between the scale of the Electroweak Theory ($\sim 10^3$ GeV) and the Planck scale ($\sim 10^{19}$ GeV) of Quantum Gravity. The problem manifests itself when computing the quantum corrections to the Higgs mass given by the virtual effects of every particle which couples to the Higgs. The Higgs coupling to a fermion with mass m_f yields a correction

$$\Delta m_H^2 \sim \frac{\mid \lambda_f \mid^2}{16\pi^2} \left(-2\Lambda^2 + 6m_f^2 ln \frac{\Lambda}{m_f} \right) \tag{1.12}$$

where Λ is the momentum cut off used to avoid divergences and can be interpreted as the energy scale up to which the Standard Model is valid. Λ can not be as large as the Planck scale, otherwise $\Delta\ m_H$ will be too large with respect to the Higgs mass. Divergent terms could be canceled by mass counterterms but the latters should be fined tuned with a procedure which is theoretically possible, but which is considered unnatural.

A solution to the hierarchy problem comes from the SUperSYmmetric theories in which a higher symmetry is broken at low energy. For each fermionic particle there is a scalar partner belonging to the same SUSY multiplet and whose contribution to the Higgs mass cancels the divergences from its partner. In case of existence of the supersymmetry, new particles would be observed at the TeV scale.
For what concerns the Higgs sector, at least two Higgs doublets are necessary in a supersymmetric scenario. In the Minimal Supersymmetric Standard Model (MSSM) [6] two complex SU(2) Higgs doublets are foreseen, which correspond to eight degrees of freedom. When the symmetry is spontaneously broken, the Ws and the Z acquire mass as in the Standard Model, absorbing three degrees of freedom. The remaining five are scalar Higgs fields: two charged Higgs bosons $H^{\pm}$, a CP-odd neutral boson A and two CP-even neutral bosons h and H. Two parameters are necessary to describe the Higgs sector at the tree level in the MSSM. The usual choices for them are the mass of the CP-odd neutral boson m_A and the ratio between the vacuum expectation values of the two Higgs doublets, $tan\beta$.
In the following of the chapter we will concentrate only on the Standard Model Higgs boson.

1.3 Limits on the Higgs mass

Some limits can be set on the Higgs mass both on theoretical and experimental grounds.

Theoretical limits

From a theoretical point of view, some constraints on the Higgs mass come from consistency arguments of the Standard Model [7]. The Standard Model is an effective field theory which provides a very good description of the physics of fundamental interactions at an energy scale of O(100 GeV) and below. It existes anyway some energy scale Λ at which the Standard Model is no longer adequate to describe the theory and some degrees of freedom associate with new physics become relevant.
According to equation 1.10, the mass of the Higgs boson depends on the two parameters λ and v. While v is determined by the gauge bosons masses, λ is a free parameter and therefore the Higgs mass is unknown. As a coupling in a gauge theory, λ evolves with the energy scale. As it is shown in figure 1.1 [8], its behaviour depends on the value at the TeV scale and therefore on the Higgs mass according to equation 1.10. A lower limit on the Higgs mass can be set requiring that the Higgs potential has a state of minimum and this corresponds

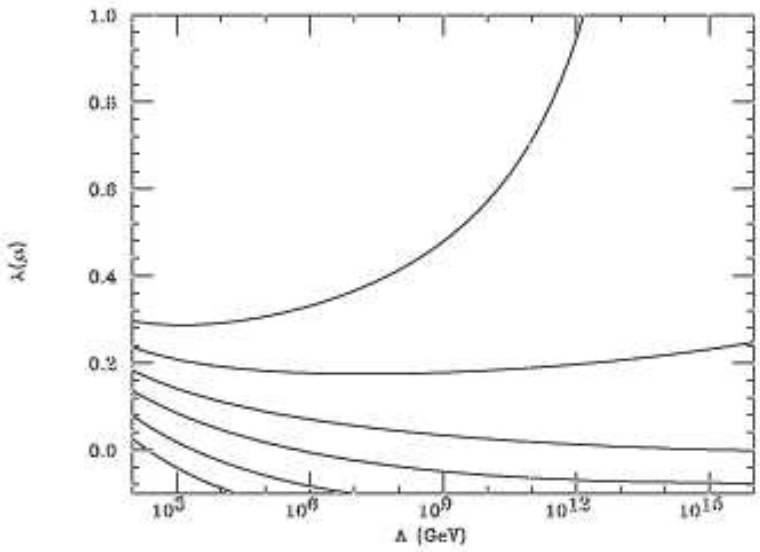

Figure 1.1: *The running coupling constant $\lambda(\mu)$ for different values of $\lambda(m_Z)$ [8].*

to the requirement that λ stays positive, since for negative couplings the potential is unbounded from below. An upper limit on the Higgs mass (the so called $triviality\ bound$) comes instead from the requirement that λ remains finite. For large values of the Higgs mass indeed the coupling constant λ grows when the scale increases and it can leave the perturbative domain since the solution of the renormalization group for λ presents a Landau singularity. It is worth noticing that both the upper and the lower limit depend on the scale Λ up to which the Standard Model is still valid, since λ is required to stay finite and positive up to that scale. The theoretical bounds on the Higgs mass as a function of the scale Λ are given in fig.1.2. Depending on Λ, a different domain is allowed for the Higgs boson mass. As Λ increases, the range which is permitted for the Higgs mass becomes narrower and for the Standard Model to be valid up to the grand unification scale $\sim 10^{16}$ GeV the Higgs mass has to stay within the 100-200 GeV/c^2 range. It is worth noticing that these values are consistent with the indirect experimental limits.

Experimental limits

Experimental limits on the Higgs mass come both from direct and indirect searches. The existence of the Standard Model Higgs particle has an impact on the value of most electroweak parameters via higher order corrections, which present a logarithmic sensitivity to the Higgs mass [9]. The best constraint to the Higgs mass

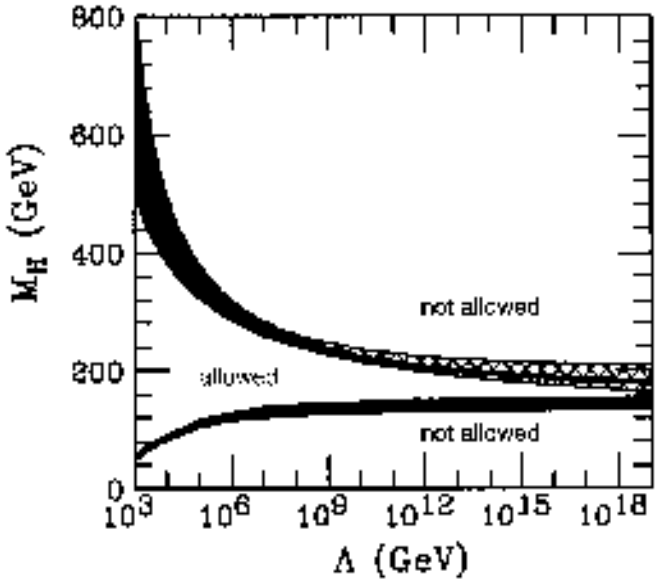

Figure 1.2: *Upper and lower theoretical limits on the Higgs mass as a function of the energy scale Λ up to which the Standard Model is valid. The shaded area indicates the theoretical uncertainties in the calculation of the bounds. Here the values m_{top} = 175 GeV/c^2 and $\alpha_s(M_Z)$ = 0.118 have been used [7].*

is given by a global fit of all the available electroweak data. The data which are used include mainly LEP results but also other measurements like the top quark mass from Tevatron or SLD results at the Z peak. Figure 1.3 on the left shows the $\Delta\chi^2$ curve from precision electroweak measurements as a function of the Higgs mass when assuming the Standard Model to be the correct theory of nature. The 95% confidence level upper limit, including results up to summer 2005, is $m_H < 219$ GeV/c^2, which is consistent with the direct searches at LEP discussed in the following. The theoretical uncertainties originating from the uncomputed higher order corrections are indicated by the shaded band. The result which is obtained when using an alternative estimation of the contribution of light quarks to the photon vacuum polarization is also shown ($\Delta\alpha^{(5)}(m_Z^2)$). Among the experimental inputs for the global fit, the W mass has a large impact. The top quark mass also enters the fit as a parameter. The green band in figure 1.3 on the right shows how the Standard Model constraints between the W and the top masses depend on the Higgs mass, which enter the loop corrections.
The indirect limits discussed here of course only hold in the contest of the Standard Model and a heavier Higgs can enter theories which extend the model.

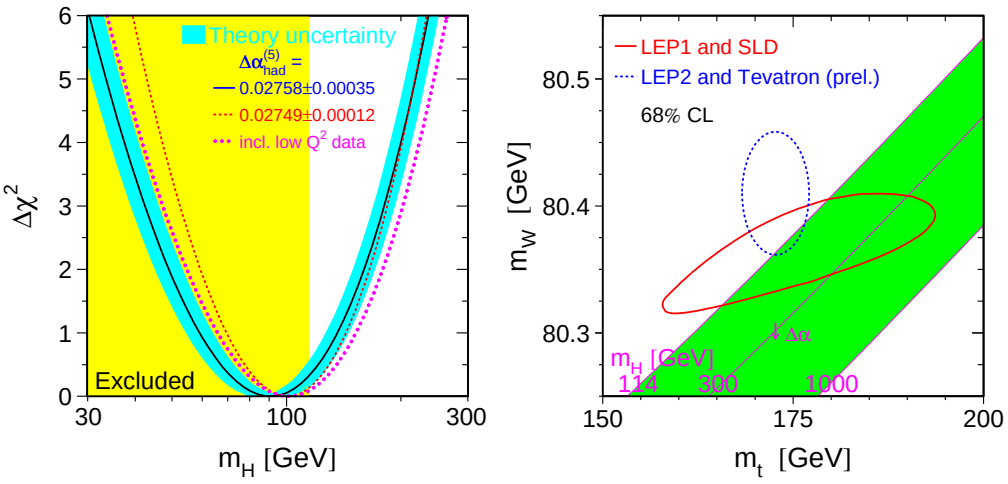

Figure 1.3: *Left: observed $\Delta\chi^2$ curve of a common fit to different electroweak measurements as a function of the Higgs mass. The yellow band corresponds to the mass region excluded by direct searches. The blue band gives the theoretical uncertainty. Right: contour curves of 68% CL in the plane of the W mass versus the top mass. The green band shows the dependence on the Higgs mass of the Standard Model constraints between the two masses. $\Delta\alpha$ shows how the relation between m_W and m_t changes when $\alpha(m_Z^2)$ is changed by one standard deviation [9].*

The direct searches of the Higgs boson did not show any evidence up to now. The strongest limit comes from the combined results of the four LEP experiments [10]. At LEP the Higgs boson was mainly produced through the 'Higgs-strahlung' process $e^+e^- \rightarrow HZ$ in association with the Z boson. The data collected by the four experiments have been examined in a likelihood test for their consistency with the background only hypothesis and the signal plus background hypothesis. The final result, shown in figure 1.4, gives $m_H > 114.4$ GeV/c^2 as lower limit on the Higgs mass at the 95% confidence level.

1.4 Higgs production and decay mechanisms

As it was discussed in the previous sections, the Higgs mass is not predicted by the theory. In contrast, the Higgs couplings to the fermions (bosons) are predicted by the theory to be proportional to the corresponding particle masses (squared

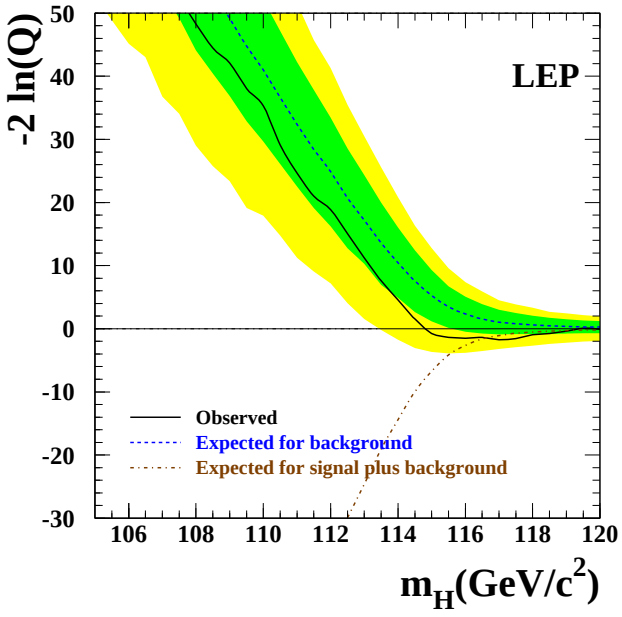

Figure 1.4: *Observed and expected behaviour of the test statistics -2lnQ as a function of the Higgs mass. Q is the ratio between the signal plus background likelihood and the background only likelihood. The result is the combination of the data collected by the four LEP experiments. The solid line is the observed curve, the dashed curve is the median background expectation and the dash-dotted curve is the expectation in the signal plus background hypothesis. The dark and the light shaded bands represent the 68% and 95% probability bounds around the median background expectation. The expected curves and the associated uncertainty bands have been obtained by replacing data by MonteCarlo generated events in a large number of toy experiments [10].*

masses) and they are given by

$$\begin{aligned} g_{hf\overline{f}} &= \frac{m_f}{v} \\ g_{hVV} &= \frac{2m_V^2}{v} \end{aligned} \tag{1.13}$$

For this reason, in the Higgs production and decay processes, the dominant mechanisms involve the coupling of the Higgs boson to heavy particles and therefore mainly to W, Z and the third generation fermions. At leading order the Higgs does not couple to gluons and photons. The Higgs coupling to gluons is induced by an one-loop graph in which the Higgs couples to a virtual $q\overline{q}$ pair. The coupling to photons is generated in an analogous way, with the dominant contribution given

by an one-loop graph in which the Higgs couples to a virtual W^+W^- pair.

1.4.1 Higgs boson production

The cross sections for the various Higgs production mechanisms at a hadron colliders are shown in fig.1.5 [11] for a center of mass energy $\sqrt{s}$ = 14 TeV, corresponding to the design value at the Large Hadron Collider. The interesting pro-

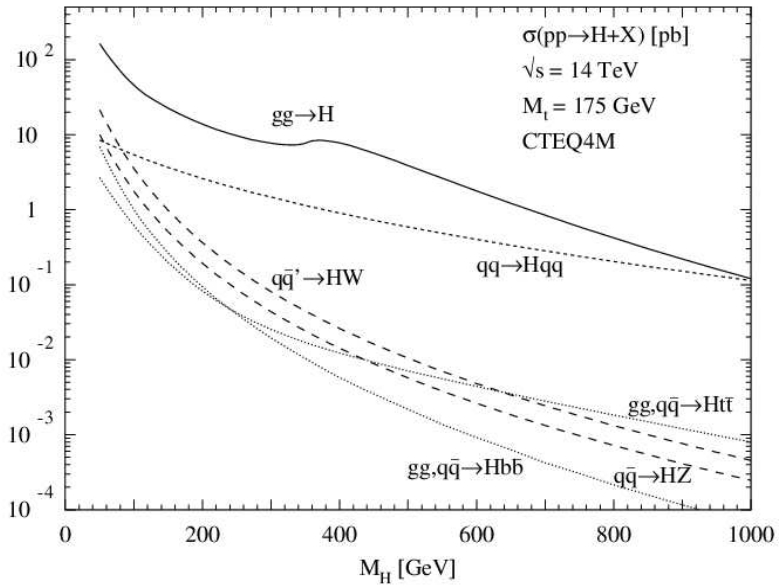

Figure 1.5: *Higgs production cross sections at $\sqrt{s}$ = 14 TeV as a function of the Higgs mass.*

cesses are the gluon fusion $gg \rightarrow H$, the vector boson fusion $qq \rightarrow Hqq$, the Higgs-strahlung $qq \rightarrow VH$ and the associated production with a $q\overline{q}$ pair. The leading order diagrams for these processes are shown in figure 1.6.

The **gluon fusion** is the dominating mechanism for the Higgs production at LHC over the whole Higgs mass range, due to the gluons large luminosity. The process proceeds mainly through the top or bottom quark triangle loop described before as shown in figure 1.6 [12]. The next-to-leading order QCD corrections to this process [13] are relevant and positive and they increase the cross section up to about a factor 2 with respect to the lowest order results. They are usually taken into account with the K-factor, which is defined as the ratio between the NLO cross section and the leading order one and which includes both real and virtual contributions. Recently the next-to-next-to-leading order calculations [14] have

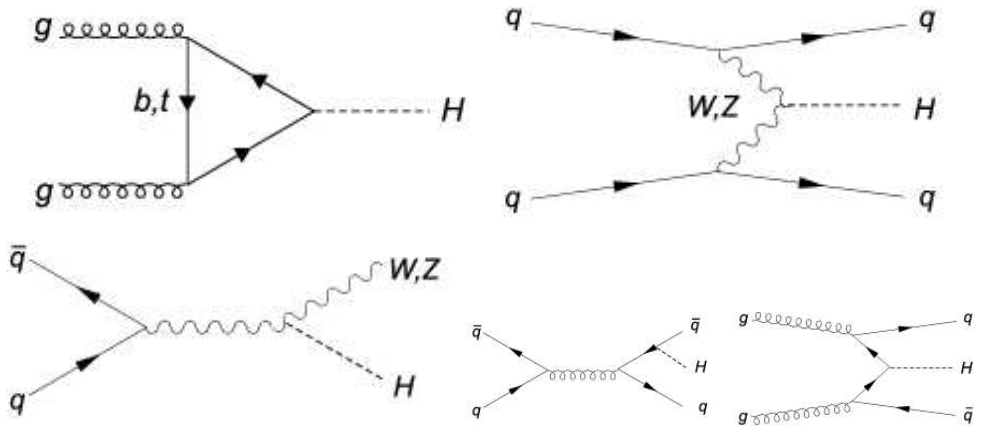

Figure 1.6: *Diagrams for the various Higgs production mechanisms. On the top: left) gluon fusion, right) vector boson fusion. On the bottom: left) Higgs-strahlung, right) associated production with a* $q\overline{q}$ *pair.*

been also computed in the large top mass approximation and they show a further enhancement of about 10% to 30% depending on the Higgs mass and on the center of mass energy of the collider. The scale dependence and the effects of higher order terms not yet computed are estimated to give a theoretical uncertainty of 10%-20%. The dependence of the gluon fusion cross section on different parton densities yields roughly an additional uncertainty of order 10%.

The **W** and **Z fusion** is the second largest contribution to the Higgs production cross section. It is one order of magnitude smaller than the gluon fusion in the intermediate mass range and the two processes become competitive only at high masses around the TeV. Both the quark and the antiquark radiate virtual bosons which annihilate to produce the Higgs, so the final state is characterized by the presence of two quarks. Two forward jets with high invariant mass and no hadron activity in the central region are the main experimental signatures of this channel, which can be exploited to improve the signal over background ratio, despite the low cross section. The cross section for this process is known with small uncertainties both at LO and at NLO. Higher order QCD corrections are quite small and the K-factor is about 1.1 [15].

In the **Higgs-strahlungh** process the Higgs is produced in association with a W or a Z boson, whose decay products can be used to tag the event. The cross section is orders of magnitude smaller with respect to the gluon fusion one [16]. The QCD corrections are large and the K-factor ranges from about 1.2 to 1.4 de-

pending on the Higgs mass [17].

The **associated production** [18] with a $t\bar{t}$ pair is the last production mechanism for the Higgs boson. The cross section is much lower with respect to that for the gluon fusion, but as for the Higgs-strahlungh case the presence of a $t\bar{t}$ pair in the final state can be exploited to tag the event. Higher order corrections give a K-factor around 1.2. The size of the QCD corrections depends sensitively on the choice of the scale and on the parton distributions functions.

1.4.2 Higgs boson decay

Depending on the Higgs mass, different decay channels can be exploited to detect the particle. The Higgs total decay width, the lifetime and the different decay branching ratios all depend on the Higgs couplings to the vector bosons and to the fermions in the Standard Model lagrangian given in 1.13.
Due to the dependence of Higgs couplings on the particle masses, the Higgs tends to decay into the heaviest particles which are kinematically allowed. This behaviour is shown in figure 1.7 [19], which shows the Higgs decay branching ratios

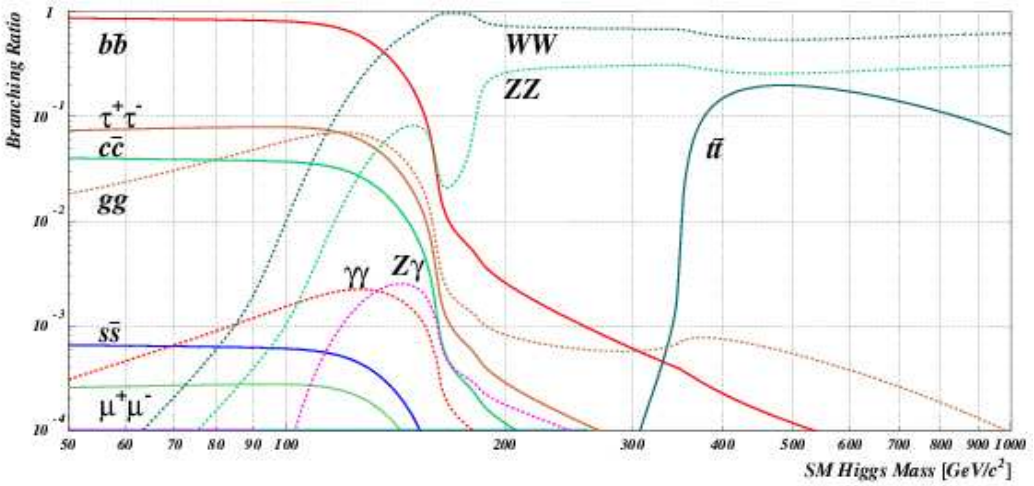

Figure 1.7: *Higgs decay branching ratios in the different channels as a function of the mass.*

.

including also NLO QCD and QED corrections. It can be seen how new chan-

nels open up when increasing the Higgs mass and how they dominate the branching ratio. Fermions decay modes contribute in the low mass region (up to $\sim$150 GeV/c^2), where the branching ratio is dominated by the channel $H \rightarrow b\overline{b}$. Once the decay into pairs of bosons is allowed, it quickly dominates; a peak in the WW production is visible around 160 GeV/c^2, when the production of two on-shell Ws becomes possible and the production of a real Z pair is still not permitted. At high masses ($\simeq$ 350 GeV/c^2) also $t\overline{t}$ pairs can be produced. As it was discussed before, the Higgs boson does not couple to photons and gluons at tree level. Such couplings can arise via fermion or W loops and they give a contribution in the low mass region.
The total width, given by the sum over all the possible decay channels, is shown in fig. 1.8. It quickly increases with the Higgs mass due to the new channels

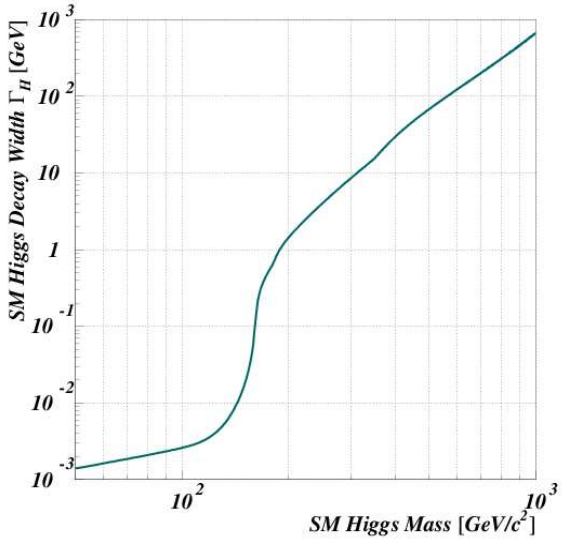

Figure 1.8: *Higgs total decay width as a function of the mass.*

.

which open and it is almost as large as the Higgs mass around 1 TeV, where it becomes almost impossible to separate the Higgs resonance from the VV continuum. The most challenging region is that of the low masses, because since up to 200 GeV/c^2 the Higgs width is lower than 1 GeV and it is therefore dominated by the experimental resolution.

It is worth mentioning that the importance of a given channel depends not only on its total cross section but also on the capability of experimentally detecting the signal rejecting the backgrounds. Fully hadronic final states are the

most abundant at LHC but they are difficult to discriminate with respect to the QCD backgrounds, therefore leptonic final states are preferred. Figure 1.9 gives the statistical significance as a function of the Higgs mass for the CMS detector,

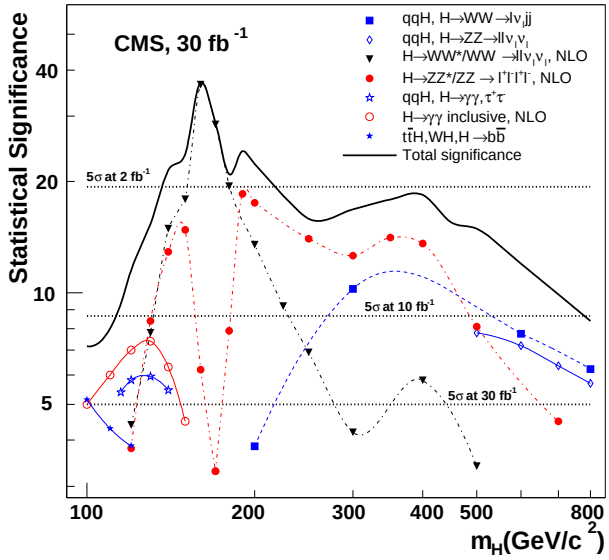

Figure 1.9: *Expected statistical significance with the CMS detector with 30 fb^{-1} for the Standard Model Higgs boson as a function of the mass. The NLO cross-sections are included for the inclusive $H \to \gamma\gamma$ channel, the $H \to \gamma\gamma$ channel in the H+jet production, the $H \to ZZ^* \to 4l$ channel and the $H \to WW^* \to 2l2\nu$ channel.*

which will be presented in chapter 2. Among the most promising channels there are

- $H \to \gamma\gamma$ in the low mass region, $m_H < 130$ GeV/c^2. It requires excellent performances of the electromagnetic calorimeter in term of energy and direction resolution. Despite the higher branching ratio, the channel $H \to b\overline{b}$ alone can hardly be exploited due to the large di-jet background and the difficulties in selecting the events at trigger level. A better situation can be obtained selecting the associated production and the Higgs-strahlungh channels, which have additional leptons in the final state and can be therefore exploited despite the lower cross section.

- $H \to WW^{(*)} \to 2l2\nu$ and $H \to ZZ^{(*)} \to 4l$ in the intermediate and high mass region. The Higgs decay in the WW channel is particularly important

above the W-pair and below the Z-pair threshold, as it will be discussed in the following. In the region $m_H > 2m_Z$, both the Zs in the $H \to ZZ$ final state are real and the signature for this decay channel is well clear.

Chapter 2

The CMS experiment at the Large Hadron Collider

The Large Hadron Collider is a new particle accelerator which will collide protons at the centre of mass energy $\sqrt{s}$ = 14 TeV with a design luminosity of 10^{34} cm^{-2} s^{-1}. The experiments which will take data at LHC will exploit its unique characteristics to cover a wide physics program ranging from the problem of the origin of the mass to the search for new physics beyond the Standard Model.
In the first part of this chapter an overview of the machine and of the requirements for the LHC experiments is given. The second part of the chapter is devoted to the description of the CMS detector.

2.1 The Large Hadron Collider

As it was discussed in the previous chapter, the Standard Model is not able to answer all the many questions which raise in particle physics, the most burning ones being the problem of the origin of the mass and the possible presence of new physics beyond the Standard Model. This is the main reason which led to the project of a new particle accelerator, the Large Hadron Collider, presently under construction at CERN and foreseen to start its operations in 2007.
The Large Hadron Collider (LHC) [20] is a proton-proton collider with a centre of mass energy of 14 TeV. A pp collider has been chosen instead of an e^+e^- one to reduce the synchrotron radiation and to be able to accelerate the particles up to a very large energy; it was preferred to a $p\overline{p}$ collider because it allows to reach higher luminosity.
The two proton beams, which will circulate in two different vacuum chambers, will contain 2808 bunches each, with an average of 1.1×10^{11} protons per bunch. The beams will collide at the rate of 40 MHz at the four interaction points, with

a spatial distribution having a RMS of 7.5 cm along the beam axis and 15 μm in the transverse directions. The high frequency of the bunch crossing and the high number of protons per bunch make it possible to obtain a very high luminosity. During the first three years the LHC will work at a luminosity of $10^{33}\text{cm}^{-2}\text{s}^{-1}$, corresponding to an integrated luminosity of 10 fb^{-1}/year, then it will reach the nominal luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ (100 fb^{-1}/year).
The collider will be installed in the already existing LEP tunnel. The available CERN accelerators will be used in the injection chain, as it is shown in fig 2.1. The proton beams exit a 50 MeV linear accelerator, they enter the PS at 1.4 GeV,

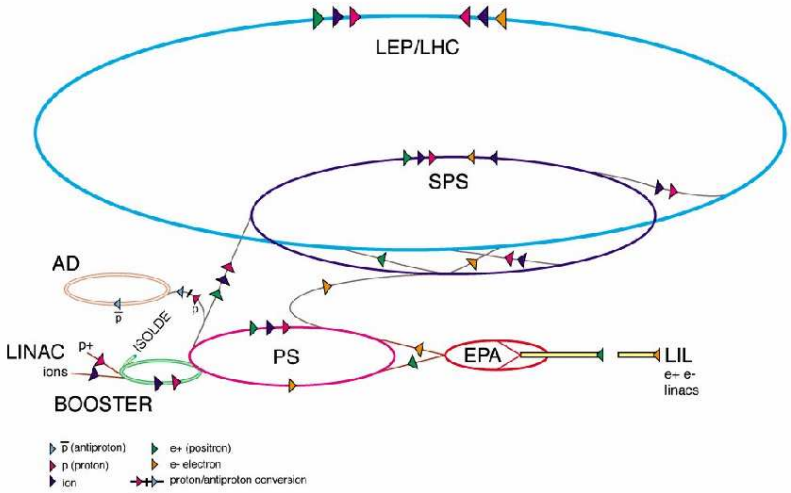

Figure 2.1: *The LHC injection chain.*

then the SPS at 25 GeV and finally the LHC at 450 GeV. In the LHC each beam will reach the 7 TeV energy. Since the proton is not an elementary particle, the available energy in the collisions between its constituents depends on the proton structure functions and on average will be smaller, but nevertheless in the TeV range.
A serious difficulty comes from the high rigidity of a 7 TeV proton beam, which requires a very strong magnetic field for the bending. The LHC will exploit the most advanced superconducting technology, with Ni-Ti conducting dipole magnets which are designed for a maximum field of 8.4 T and which will be cooled down to 1.9 K by means of superfluid Helium.
The LHC will also collide heavy ions beams with a total energy up to 2.76 TeV per nucleon in Pb-Pb collisions, for the study of the quark-gluon plasma.

In figure 2.2 the cross sections for many processes in a proton-proton collision

are shown as a function of the centre of mass energy.

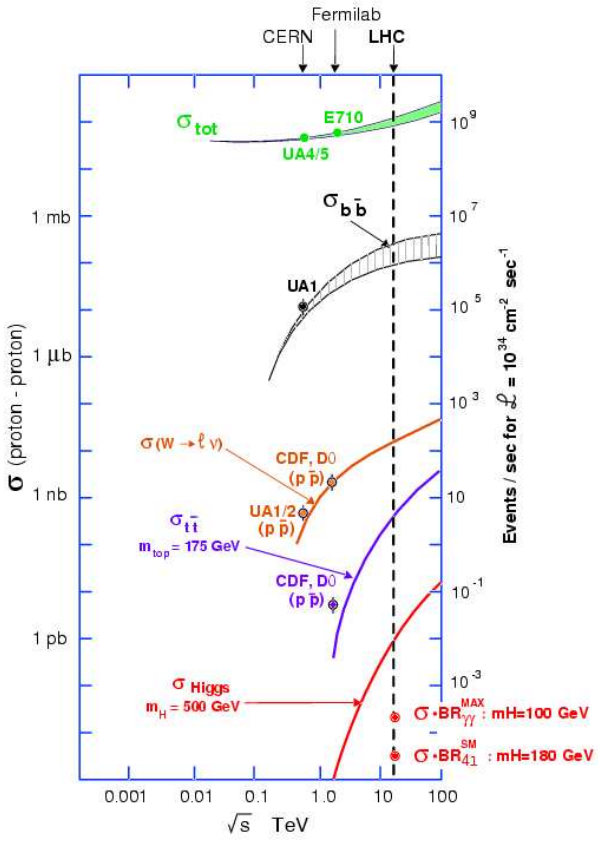

Figure 2.2: *Cross sections for different processes in a proton-proton collision as a function of the centre of mass energy.*

As it is discussed in [21], the expected total pp cross section is σ_{tot} = 100 mb and the inelastic cross section is σ_{in} = 60 mb. The inelastic interactions are divided into two classes. The first one consists of head on collisions at small distance between two partons coming from the two protons. These interactions are characterized by high transferred momentum and they bring to the creation of massive particles at large angle with respect to the beam line. Unfortunately these events are rare. The majority of the interactions consists of *minimum bias events*, large distance collisions between the two incoming partons, with a small transferred momentum and low transverse momentum in the final state.
With the LHC nominal parameters, 20 minimum bias events are expected on average per bunch crossing at high luminosity, giving a significant pile-up to each interaction. Furthermore the detectors have to face the problem of large QCD

backgrounds with cross sections orders of magnitude larger than those for interesting events (fig. 2.2 and table 2.1). Such a difficult environment imposes

process	events/s
$W \rightarrow e\nu$	20
$Z \rightarrow l^+l^-$	2
$t\bar{t}$	4
$b\bar{b}$	10^5
QCD jets ($p_T > 200$ GeV/c)	10^2

Table 2.1: *Expected rates for some important processes at the LHC low luminosity.*

important requirements on the detectors design. A fast response (of the order of 25-50 ns) and a fine granularity are needed to separate the large number of particles and to minimize the pile-up effects. The high rate of events ($\sim 10^{\,9}$ ev/s) makes a fast and efficient trigger and data acquisition system necessary. Finally, due to the high flux of particles, a good radiation resistance for all the detector components is required, especially in the forward regions. All these reasons make the project of a detector for LHC very challenging.

Four detectors will be installed at the LHC. ATLAS [22] and CMS [23], both general purpose detectors, will be installed in the two high luminosity interaction points. LHCb [24], devoted to B-physics studies and the ALICE [25], dedicated to the study of heavy ions collisions, will work instead at a reduced luminosity.
In the following a brief description of the CMS detector is given.

2.2 The CMS experiment

CMS (Compact Muon Solenoid) is one of the two general purpose experiments which will tale data at the LHC. Its physics goals range from the search for the Higgs boson to the searches for new physics beyond the Standard Model, to the precision measurements of already known particles and phenomena. To achieve these goals excellent performances in lepton reconstruction and particle identification are necessary. Besides, an optimal hermeticity is required to detect invisible particles through missing energy measurements.
CMS has a central cylindrical section (the barrel) closed at both ends by disks which are orthogonal to the beam pipe (the endcaps). The central part of the detector is a 13m long superconducting solenoid of 6m diameter, providing a 4 T

axial magnetic field. The longitudinal view of one quarter of the detector is shown in figure 2.3 and the transverse view of the barrel region is given in fig 2.4.
The natural coordinate frame to describe the detector is a right handed cartesian

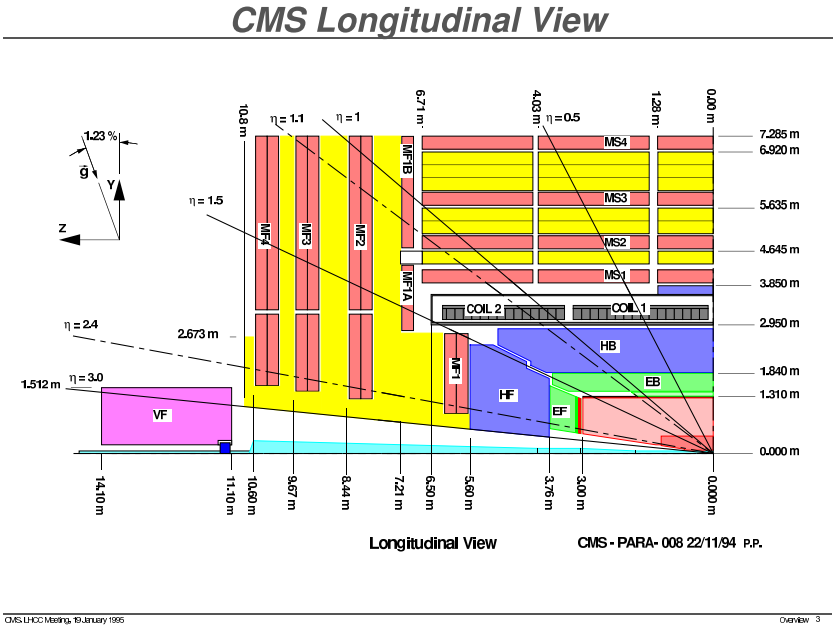

Figure 2.3: *Longitudinal view of one quarter of the CMS detector.*

system with the x axis pointing toward the centre of the LHC ring, the z axis directed along the beam axis and the y axis directed upward. Given the cylindrical symmetry of CMS, a convenient coordinate system is given by the triplet (r,ϕ,η), being r the distance from the z axis, ϕ the azimuthal coordinate with respect to the x axis and η the pseudorapidity, which is defined as $\eta = -ln(tan(\theta/2)$, where θ is the polar angle with respect to the z axis.
CMS consists of different subdetectors. Starting from the beam line there are

- the tracker, to measure the momentum of charged particles in the magnetic field and to identify the interaction vertex and the secondary vertices
- the electromagnetic calorimeter ECAL, for an accurate measurement of the energy and of the position of photons and electrons
- the hadronic calorimeter HCAL, to measure the energy of both neutral and charged hadronic particles
- the muon system, to reconstruct the muonic tracks and to measure their momenta from the bending in the magnetic field

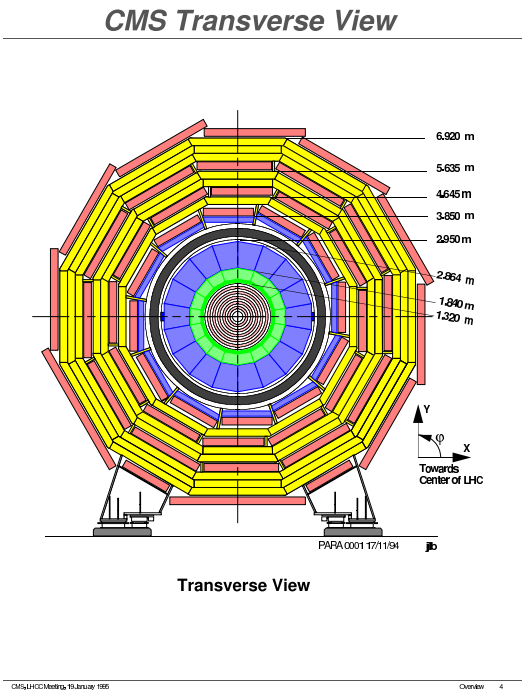

Figure 2.4: *Transverse view of the barrel region of CMS detector.*

The overall length of the detector is approximately 22m, its width is 15m and the total weight is about 12500 tons.
A brief description of each subdetector follows in the next sections.

2.2.1 The tracker

The tracker [26], placed within the magnetic field, is the subdetector which is closer to the interaction point. It is dedicated to track and vertex finding. High p_T charged particles, both isolated and within jets, have to be reconstructed with high efficiency and good momentum resolution; the interaction vertex has to be reconstructed and possible secondary vertices (which can be usefully exploited for jet tagging) have also to be identified.
The silicon technology has been chosen for the whole tracker and different constraints have driven the tracker design. Given the high number of charged tracks which are expected per bunch crossing, the tracker has to provide high granularity and large hit redundancy to perform better the pattern recognition. Severe material budget constraints are imposed by the necessity not to degrade the ECAL performances. Finally, the tracker performances must degrade as least as possible in an environment with high radiation, due to the high number of hadrons and

back scattered neutrons. Indeed, the irradiation can bring both superficial and bulk damages [27] [28] to the silicon structure. In order to survive for a long time, the tracking system will operate at $-10\ ^{\circ}C$.
The tracker covers the region up to 120 cm in the radial direction and up to $\mid z \mid$ = 270 cm along the beam. It consists of an inner silicon pixel detector and an outer silicon microstrip system.

The **pixel detector** consists of three barrel layers and two endcap disks at each side (fig. 2.5). The barrel layers, 53 cm long, are positioned at radii of 4.4, 7.3

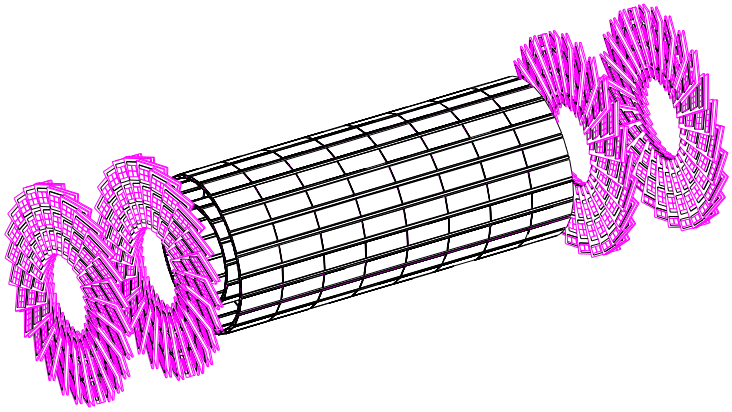

Figure 2.5: *The pixel detector. The barrel section and the two disks of the endcaps are visible.*

and 10.2 cm. The first layer will be replaced by an outer layer at r = 13 cm during the high luminosity phase, to reduce the radiation damage. Each of the endcaps consists of 2 disks with 24 blades arranged in a turbin-like shape, having the inner radius of 6 cm and the outer one of 15 cm. The total area covered with pixels is close to 0.92 m^2.
The inner detector is designed to ensures at least 2 hits for each track having the vertex within $2\sigma_z^2$ from the central interaction point in the whole pseudorapidity region $\mid \eta \mid < 2.2$. Due to the high density of tracks, a fine granularity is required to assure low occupancy for each cell; 100 $\times$ 150 μm^2 pixels have been chosen. The estimated resolution on the single hit is 10 μm for the (r, ϕ) coordinate and 15 μm for z in the barrel, 15 μm and 20 μm respectively in the endcaps.

The **silicon strip detector** consists of an inner system composed of 4 barrel layers and 3 disks at each side and an outer system with 6 barrel layers and 9 disks in the endcaps (fig. 2.6). It covers the area between 20 cm and 120 cm in the transverse radius and the pseudorapidity region up to $\mid \eta \mid < 2.5$. The strips are parallel to the

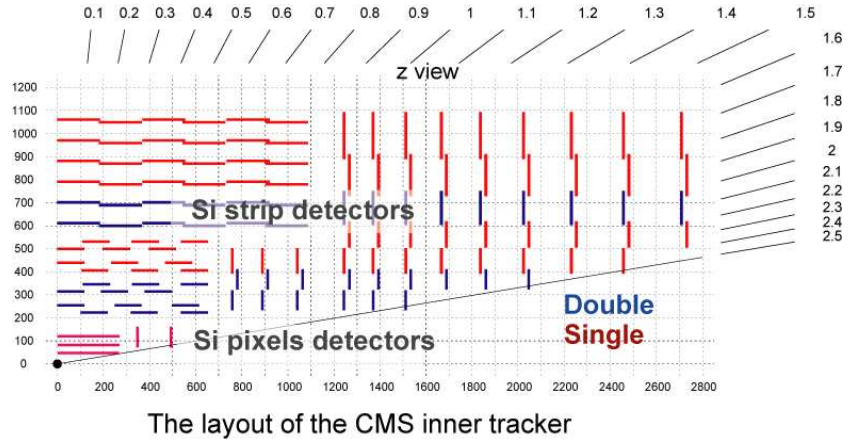

Figure 2.6: *Longitudinal view of one quarter of the CMS silicon strip tracking system. The red lines are modules composed of one detector only, the blue lines are detectors composed of two modules.*

beam in the barrel and have radial orientation in the endcaps and their dimensions have been chosen to assure good spatial resolution and occupancy lower than 1%. The expected spatial resolution is order of 40-60 μm in the (r, ϕ) plane and it is about 500 μm along z [29].

The performances of the CMS tracker are summarized in fig 2.7 and 2.8 [29]. High energy electrons are reconstructed with efficiency higher than 90%. The reconstruction efficiency for isolated muons with $p_T > 1$ GeV/c is close to $100\,\%$ in the $|\,\eta\,| < 2$ region and the transverse momentum resolution is better than 3% for $p_T \lesssim 100$ GeV/c. The vertex reconstruction can be performed using the pixel detector alone with a spatial resolution between 20 μm and 70 μm and efficiency around 95%. Such reconstruction is fast enough to be used in the High Level Trigger. The resolution on the vertex position can be improved down to 15 μm by using the information from whole tracking system. This can be usefully exploited in the accurate off-line analysis.

2.2.2 The electromagnetic calorimeter

The main goal of the electromagnetic calorimeter ECAL is the precise measurement of the energy of electrons and photons.
The design of the ECAL [30] was led by the requests imposed by the $H \rightarrow \gamma\gamma$ channel, which is studied looking for a peak in the di-photon invariant mass distribution. It was discussed in the previous chapter that in the mass region $m_H < 140$

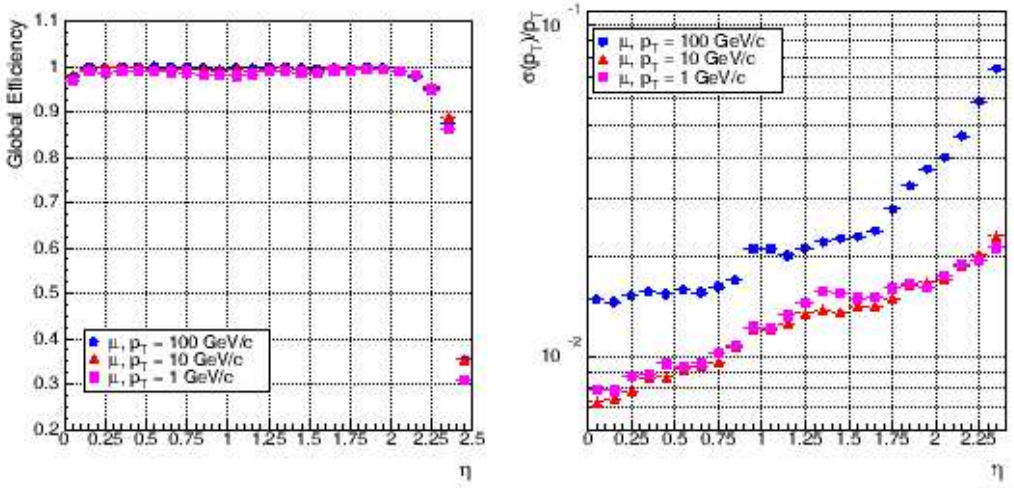

Figure 2.7: *Left: global efficiency of the tracking algorithm for single muons events as a function of the pseudorapidity (p_T = 1, 10, 100 GeV/c). Right: transverse momentum resolution for single muons events as a function of the pseudorapidity (p_T = 1, 10, 100 GeV/c) [29].*

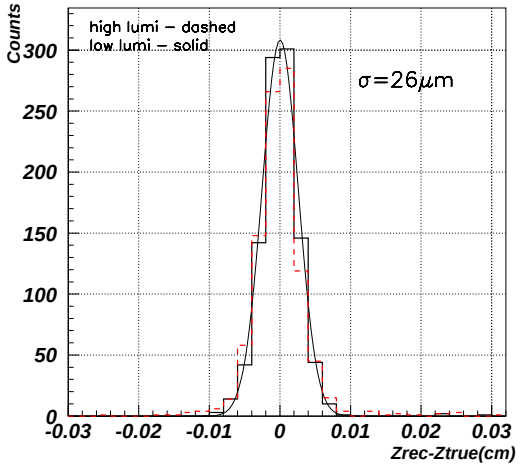

Figure 2.8: *Difference in z between the reconstructed and the generated vertex. Only the information from the pixel detector is used in the reconstruction. A Gaussian fit is superimposed. The use of the full tracker information allows to improve the resolution down to 15 μm [29].*

GeV/c^2 the intrinsic Higgs width is lower than 30 MeV, therefore the invariant mass resolution is dominated by the experimental resolution, which is required to

be order of 1%. Besides, an high granularity is necessary, to improve the measurement of the angle between the two photons and to obtain a good π^0/γ separation.

The general structure

An homogeneous calorimeter has been chosen. ECAL consists of almost 76000 Lead Tungstate $PbWO_4$ scintillating crystals divided into a barrel and two endcaps; a three dimensional view of the calorimeter is given in fig 2.9.
The barrel ($\mid \eta \mid < 1.479$) consists of 36 supermodules, each one containing 20

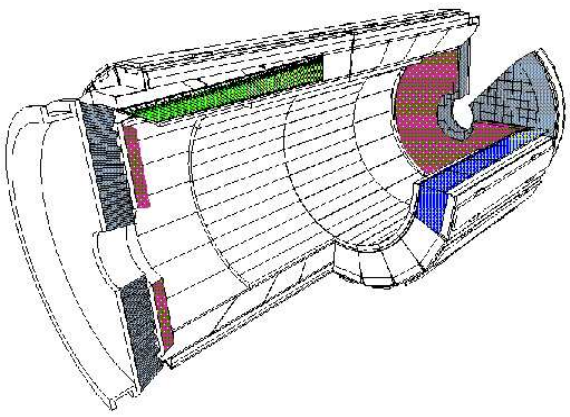

Figure 2.9: *Three dimensional view of the electromagnetic calorimeter.*

crystals in ϕ × 85 crystals in η and covering an azimuthal arc of $20°$. The supermodules are divided along η in 4 modules made of submodules, which are the basic units of ECAL and which consist of 5×2 crystals each. The geometrical shape of the crystals slightly changes along η and there are 17 types of crystals, with length close to 230 mm and front face area of about 22×22 mm^2. The barrel granularity is $\Delta\phi \times \Delta\eta$ = 0.0175×0.0175, the crystals are grouped into 5 × 5 arrays corresponding to the trigger towers. To avoid that cracks might align with the particles trajectories, the crystal axes are tilted by 3 degrees with respect to the direction from the interaction point, both in η and in ϕ.
The endcaps consist of two halves (*Dees*) and cover the pseudorapidity region $1.48 < \mid \eta \mid < 3$. All the crystals have the same shape (220×24.7×24.7 mm^3) and they are grouped in structures of 5 × 5 crystals called super-crystals. The granularity varies from $\Delta\phi \times \Delta\eta$=0.0175×0.0175 to 0.05×0.05. As for the barrel, the crystals have a non pointing geometry. To ensure good hermeticity, an overlap of half crystal between the endcaps and the barrel is obtained by orienting the crystals axis to point 1300 mm beyond the interaction point.

To improve the π^0/γ separation and the vertex identification, a preshower is designed to cover the region between $\mid \eta \mid = 1.6$ and $\mid \eta \mid = 2.6$. It consists of two lead converters ($2X_0$ and X_0 thick) followed by silicon strips with a pitch of less than 2 mm. The strips following the two absorbers are disposed in orthogonal way. The preshower will operate at the temperature of -5°C.

Different reasons brought to the choice of the $PbWO_4$ as active medium for ECAL. Its short radiation length (X_0 = 0.89 cm) and Moliere Radius (R_M = 2.19 cm) allow to build a compact and high granularity calorimeter. An important aspect is the fast response ($\sim$80% of the light is collected within 25ns), which is compatible with the high LHC rate. Finally the $PbWO_4$ has a good intrinsic radiation hardness, which makes it suitable to work in the hard LHC environment. The main drawback of the $PbWO_4$ crystals is the low light yield ($\sim$ 80 γ/MeV), which makes an internal amplification for the photodetectors necessary.
An intensive R&D work has been done in recent years on the $PbWO_4$ crystals. As it will be described in details in the next chapter, it succeeded in improving the radiation resistance, by the improvement of the growing techniques and by the doping with Yttrium Y and Nobium Nb; the crystals show a light yield loss which is on average lower than 3% when irradiated with 0.15 Gy/h photons. The R&D also succeeded in improving the transmission of the light along the crystal axis (as it is shown in fig 2.10) and in reducing the light collection non uniformity under the level of 0.35%/X_0 [31] [32], which is the ECAL goal as will be discussed in par 2.2.2. The main parameters of the $PbWO_4$ crystals used in the ECAL are

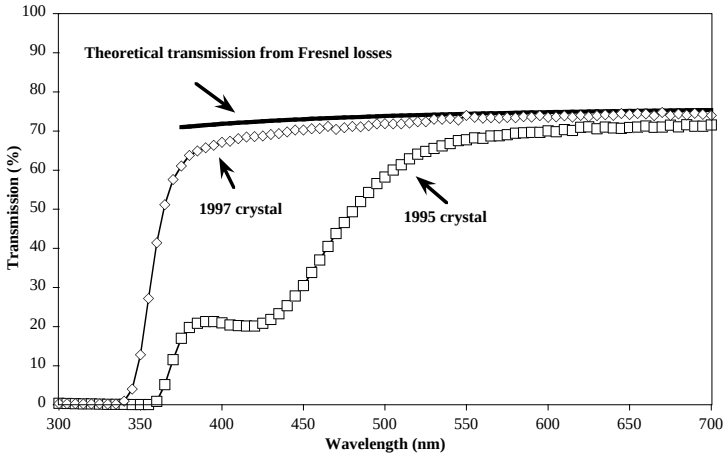

Figure 2.10: *The longitudinal transmission of the $PbW0_4$, at the beginning of the R&D and for the ECAL final crystals.*

summarized in table 2.2. More details about the scintillation mechanism and the

radiation damage are given in chapter 3.

$PbWO_4$ characteristics	
Density (g/cm^3)	8.28
Radiation length X_0(cm)	0.89
Interaction length λ_{int} (cm)	22.4
Molière radius (cm)	2.19
Light decay time (ns)	5 (39%)
	15 (60%)
	100 (1%)
Refractive index at 500 nm	2.30
Maximum of emission (nm)	440
Light yield (γ/MeV)	$\sim$80

Table 2.2: *Main characteristics of the $PbWO_4$ ECAL crystals.*

The photodetectors

The low light yield of the $PbWO_4$ makes the use of photodetectors with an intrinsic gain necessary. At the same time, the photodetectors for ECAL have to be radiation hard, fast and able to operate in the strong CMS magnetic field. The devices which match these characteristics and that have been chosen for the electromagnetic calorimeter are the Avalanche PhotoDiodes (APDs) for the barrel and the Vacuum PhotoTriodes (VPTs) for the endcaps.
The APDs are silicon detectors. The scheme of functioning is shown in figure 2.11: a 5 μm thick p^+ layer acts as photoconverter, the photoelectrons are accelerated and multiplied through the p-n junction and then a n^{++} doped region provides the ohmic contact with the preamplifier. The active area is a $5\times5mm^2$ surface. Even if it is quite small, the APDs have an high quantum efficiency ($\sim$75% at 430 nm) which well matches the emission spectrum of the $PbWO_4$; besides, each crystal is equipped with 2 APDs to increase the acceptance to the scintillation photons.
The APDs affect all the terms of the energy resolution which will be discussed later in the chapter [33]. The statistical fluctuations in the multiplication process influence the stochastic term of the energy resolution. The effect can be quantified using the excess noise factor F which is related to the APD gain M by the relation

$$F^2 = 1 + \left(\frac{\sigma_M^2}{M^2}\right) \tag{2.1}$$

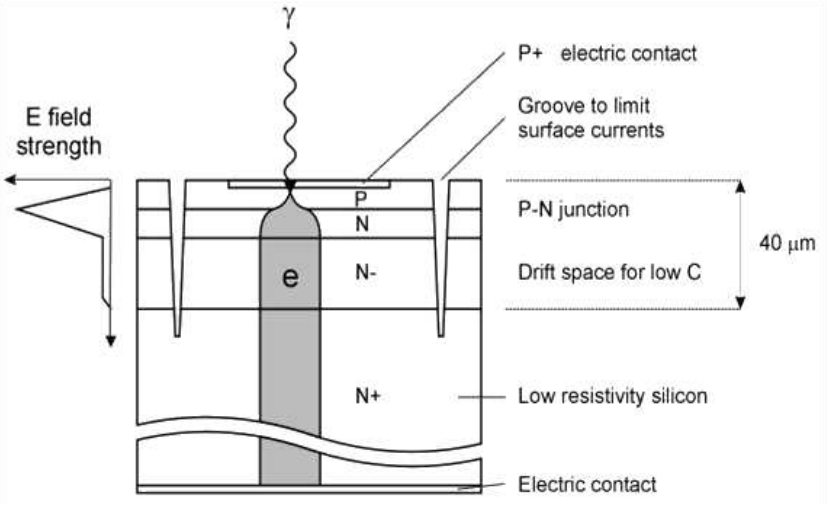

Figure 2.11: *Scheme of an Avalanche Photo Diode.*

F is about 2 for a gain M = 50, which is the value set for the ECAL barrel APDs. The calorimeter noise is affected by the APD capacitance and by the leakage currents flowing on the surface and in the bulk of the APD. Finally, the sensitivity of the APD gain to the biasing voltage and to the temperature directly influences the constant term in the energy resolution.
The APDs are seriously affected by the radiations. The damage is mainly due to the neutrons which create defects in the silicon increasing the leakage currents. Since the APDs can not survive the radiation doses of the endcaps, the technology of the Vacuum PhotoTriodes has been chosen for those regions.
The VPTs are phototubes with a bi-alkali photocathode deposited on a glass window (fig. 2.12). The electrons emitted from the cathode are accelerated towards a 10μm thick anode; the fraction of them passing the anode grid impacts on a reflective dynode with a planar geometry and then emits new electrons ($\sim$ 20 secondary electrons are emitted for each impinging electron), which are accelerated back toward the anode. The VPT quantum efficiency is about 15% at the peak of the $PbWO_4$ emission spectrum, but since the active area is almost 300 mm^2 the total light collection is at the same level as for the APDs. The VPTs operate at gain around 8 - 10 which is much lower with respect the one of the APDs but which presents very little dependence on the temperature fluctuations and on the biasing voltage. The capacitance is low (few pF), the leakage current is $<$ 2nA and the excess noise factor around 2.5-3. The VPTs use a radiation hard glass which has been appositely developed and whose transparency is not strongly affected by irradiating particles. Tests have shown that the loss in response can be kept under

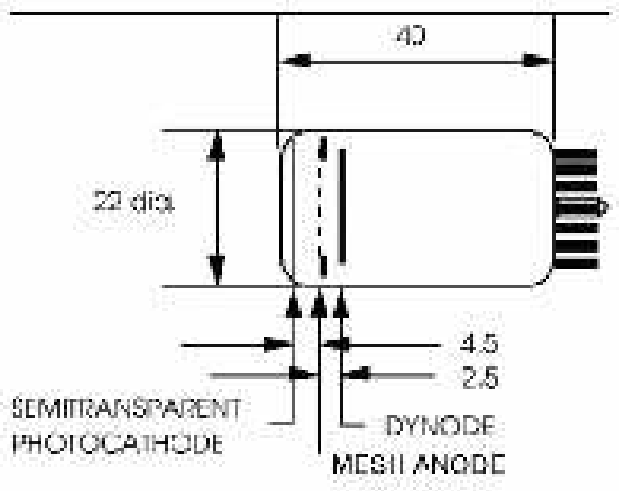

Figure 2.12: *Scheme of a Vacuum PhotoTriode*

10% in 10 years of LHC.

The electronics chain

Important requirements are imposed on the front-end (FE) electronics. It has to be fast to match the 25 ns LHC crossing rate and it must keep the noise level below $\sim$50MeV per crystal over a dynamic range of about 95dB; since it is placed on the detector it has to be radiation hard.
The electronics mirrors the trigger structure and its basic element is a group of 5×5 crystals called trigger tower. The signals coming from each photodetector are sent to a motherboard which hosts 5 Very Front End cards (VFE) and a Low Voltage Regulator card (LVR) providing the power for the VFE cards [34]. The output of the 5 VFEs is fed into a FE card which processes the digitised data of one trigger tower. Each VFE card houses 5 identical channels, where the signal from the photodetector is amplified, shaped and then sampled by a 12-bit sampling ADC working at 40MHz.
The full electronics chain underwent a heavy design review at the beginning of 2002. The first design foresaw that the signals from the photodetectors were preamplified and shaped by a trans-impedance amplifier with internal shaping, followed by a four-range amplification and logic stage to adapt the output to the dynamic range of a 12-bit ADC. Everything was integrated into a unique radiation-hard integrated circuit, the FPPA (Floating Point Pre-Amplifier), which was nevertheless affected by an important noise due to parasitic resistances and capacitances.

The current scheme employs the radiation tolerant 0.25μm CMOS technology. Each channel of the VFE cards now consists of a Multi Gain Pre-Amplifier MGPA, an ADC and a buffer. The MGPA [35] provides three outputs at three different gains, which are digitised in parallel by a new four channel 12-bit ADC working at 40 MHz with an integrated digital gain switching logic. In addition, each VFE card has a Detector Control Unit which measures the crystal temperature and the APD leakage currents.
The new electronics scheme also brought a considerable reduction of the off-detector electronics. The first design indeed foresaw that the samples of each channel were transferred via optical link to the off-detector electronics for triggering purposes. In the new scheme (shown in fig 2.13), one chip for the L1 Trigger primitive generation (the *Fenix Asics*) is put on each trigger tower within the detector to store the data until the positive response of L1 and to sum the pulse shape data from a raw of 5 crystals to produce a strip energy; in the ECAL barrel, all the strip energies are also summed to compute the energy which is deposited in the trigger tower. Trigger data are then transmitted to the off-detector electronics through a serial link which operates at 800 Mb/s, which allows the data transmission every 25 ns. In case of a L1 accept, the data from the triggered events are also transmitted to the off-line electronics through a separate serial data link also working at 800 Mb/s. This allows to reduce both the costs of the optical link system and the quantity of the off-detectors electronics by about a factor 8.

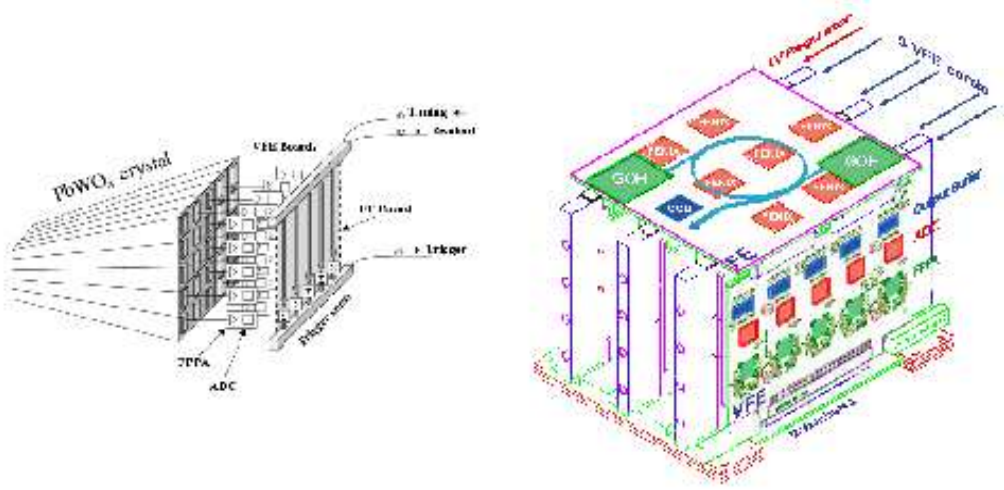

Figure 2.13: *Left: schematic view of the ECAL electronics for 5×5 crystals. Right: the arrangement of the front-end electronics into VFE and FE boards.*

The laser monitoring system

The ECAL monitoring system is based on the injection of laser light into each crystal. During the LHC data taking, the monitoring system will regularly provide light pulses during the 3μs gap the beam will have every 89μs; dedicated runs to follow the crystal recovery are also foreseen during the LHC refills.
The laser distribution system proceeds in three steps. First the light emitted by the laser source is split by means of an optical switch into each of the 80 channels the calorimeter is organized in (72 half supermodules in the barrel and 8 groups of supercrystals in the endcaps), then a 2 steps distribution system at the level of the single unit further splits the pulses to reach each crystal. A *level 2 fan-out* splits the light into 5 or 6 (according to which part of the supermodule is fed) quartz fibers; 4 or 5 fibers go to the *level 1 fan-out* and the last one is sent to a reference PN diode for the monitoring of the pulse amplitude. The *level 1 fan-out* then splits the light into 200 fibers going to the crystals. Two output fibers finally bring the light to two monitoring PNs. The light distribution scheme is shown in fig. 2.14
In the barrel the light is injected from the front face of the crystal and it directly

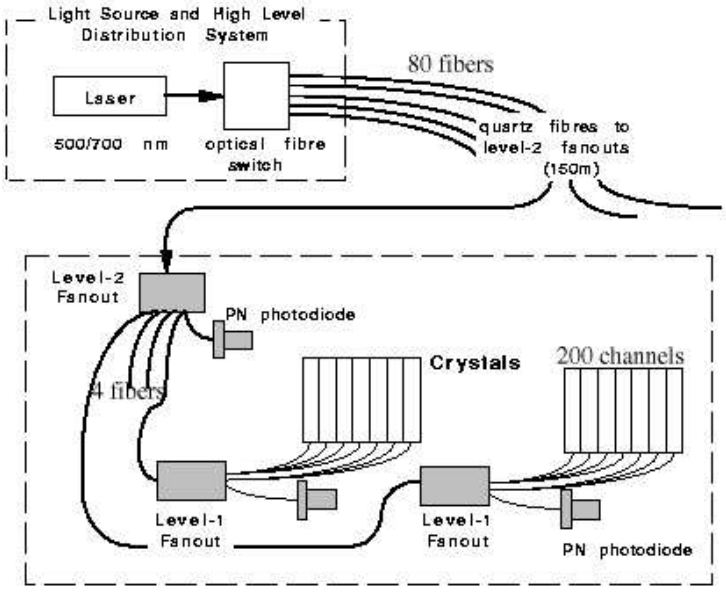

Figure 2.14: *Schematic representation of the three splitting levels in the light distribution system.*

reaches the APD, which is on the rear face; in the endcaps, the light is injected from the rear face and it reaches the photodetector after a reflection on the front face.

The pulse energy is 1mJ/pulse at the monitoring wavelenght, corresponding to 1.3 TeV in the dynamic range.

The energy resolution

The energy resolution of an homogeneous calorimeter is usually written as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{2.2}$$

where a, b and c represent respectively the *stochastic*, *noise* and *constant term* of the energy resolution. Different effects contribute to each term in eq 2.2; the relative contributions are shown in figure 2.15.

The **stochastic term** a includes the contribution of the fluctuations in the number

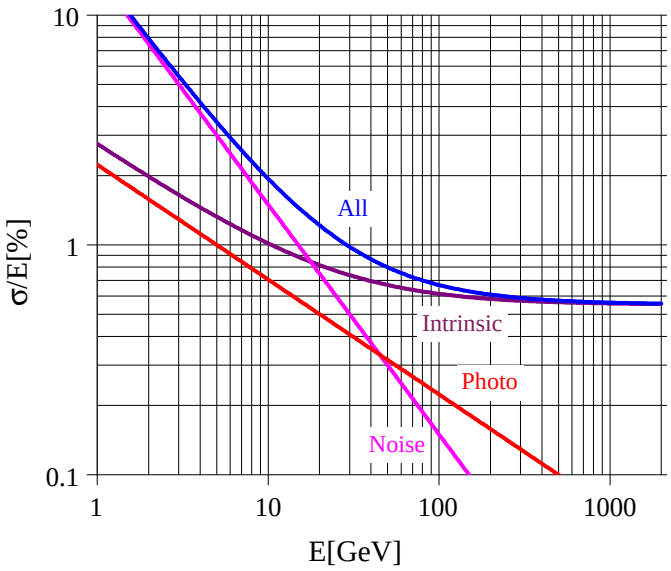

Figure 2.15: *The expected ECAL energy resolution versus the energy of the impacting electron. The different contributions are superimposed separately. The term "Intrinsic" includes the shower containment and a constant term (0.55%).*

of electrons which are produced and collected. Since the fluctuations are poissonian, the stochastic term is $a = \frac{1}{\sqrt{n_{pe}}}$, where n_{pe} is the number of photoelectrons which are emitted per energy unit. Contributions come from the light yield of the crystals, from the efficiency in the light collection and from the quantum efficiency of the photodetectors. Important contributions also come from the fluctuations in the multiplication process inside the photodetectors, which are described

by the excess noise factor introduced in par. 2.2.2. The target value for ECAL is a = 0.027 for the barrel and 0.057 for the endcaps, where the main contribution ($5\%/\sqrt{E}$) comes from the preshower sampling term.
The **noise term** b includes contributions from the electronic noise, both due to the photodetector and to the preamplifier, and from pile-up events. The contributions change at the different pseudorapidities and with the luminosity of the machine. The target values in the low and in the high luminosity phases are respectively 155 MeV and 210 MeV at $\mid \eta \mid$ = 0, 205 MeV and 245 MeV at $\mid \eta \mid$ = 2.
The **constant term** c is the dominating term at high energies and it includes many different contributions. Among them, the most important are:

- the stability of the operating conditions, such as the temperature and the high voltage.
  Both the scintillation mechanism and the APD gain are affected by the temperature and the response for a given energy deposit varies with the temperature of the calorimeter with a slope which is around -4$\%/^{\circ}$C for the barrel. The stability of the temperature within 0.05°C is required to keep the contribution to the constant term below 0.1%.
  For what concerns the photodetectors, the APD gain strongly depends on the bias voltage according to the law $\frac{1}{M}\frac{dM}{dV} \sim 3\%/V$ for a gain around 50. A stability better than 30 mV is required to keep the contribution to the constant term smaller than 0.1%.

- the presence of dead materials between the crystals and the rear and lateral leakage of the electromagnetic shower.

- the longitudinal non uniformity of the crystal light yield. A strong focusing effect of the light takes place due to the tronco-pyramidal shape of the crystals and to the high refractive index, so the light collection is not uniform. It has been shown that such non uniformity has to be kept at the level of 0.35%/X_0 to keep the contribution to the constant term within 0.3%.

- the intercalibration errors.

- the radiation damage of the crystals, which changes their response to a certain amount of deposited energy when exposed to high radiation dose rates. The details of the radiation damage mechanism are presented in chapter 3.

The target value for the constant term of the CMS ECAL is 0.5%.

2.2.3 The hadron calorimeter

The main goal of the hadron calorimeter HCAL is to contribute to the reconstruction of events which involve both hadrons and invisible particles, by means of jet

and missing energy reconstruction. High hermeticity and transverse granularity are necessary, together with a number of hadron interaction lengths sufficient to contain the energetic particles coming from high transverse momentum jets.
The CMS central hadron calorimeter [37] is placed within the magnet. It's a sampling calorimeter with brass layers used as absorbers and plastic scintillators as active medium, while the structural elements are made of stainless steel. The plastic scintillators are read by fibers which also act as wavelength shifters. In each layer they are divided into tiles with a granularity of $\Delta\eta \times \Delta\phi$ = 0.0875×0.0875, which matches the granularity of the ECAL trigger towers and which ensures the required jet separation. Brass has been chosen as absorber instead of iron because it is easier to machine and it has a 10% shorter hadron interaction length.

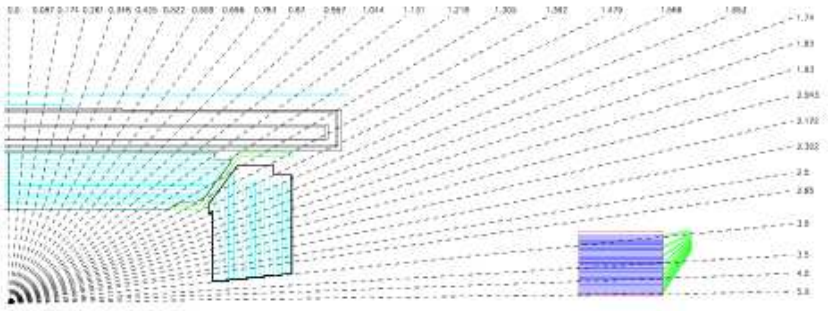

Figure 2.16: *Longitudinal view of one quarter of the CMS hadron calorimeter. The very forward calorimeter is also shown.*

The hadron calorimeter consists of the barrel and two endcap disks. The barrel extends up to $\mid \eta \mid < 1.78$ and the two endcaps cover the pseudorapidity region between 1.305 and 3, partially overlapping the barrel. For low values of pseudorapidity the tickness of the hadron calorimeter can not ensure a satisfactory containment. To assure the necessary coverage in term of interaction lengths over the whole pseudorapidity range, an outer scintillator layer is placed outside the solenoid coil in the region up to $\mid \eta \mid < 1.305$ and two other scintillators are foreseen for the central region $\mid \eta \mid < 0.348$.
To improve the hermeticity, a separate very forward calorimeter is placed outside the magnet yoke, 11m away from the interaction point both in the forward and in the backward direction, extending the coverage up to $\mid \eta \mid = 5.191$. Due to the severe irradiation in the region, radiation hard quarz fibers have been chosen as active elements, interleaved with bulky steel working as absorber; the light is read out by means of photomultipliers.
The CMS calorimetric system is not compensated, being the response to the electromagnetic part of an hadron shower different from the response to the hadron

part. The non compensation effects, which degrade both the linearity and the resolution, can be reduced by an appropriate weighting of the response of the ECAL and of the different layers of the HCAL. Detailed MonteCarlo studies and test beam analysis have shown that the energy resolution is only marginally improved by the usage of energy dependent weights; fixed weights have therefore been chosen. The hadronic energy resolution when combining informations from HCAL and ECAL is:

$$\frac{\sigma_E}{E} = \frac{1.00}{\sqrt{E/GeV}} \oplus 0.045 \tag{2.3}$$

The expected energy resolution for the very forward calorimeter is given for electrons and hadrons by

$$\begin{aligned} \frac{\sigma_E}{E} &= \frac{1.38}{\sqrt{E/GeV}} \oplus 0.05 \qquad (electrons) \\ \frac{\sigma_E}{E} &= \frac{1.82}{\sqrt{E/GeV}} \oplus 0.09 \qquad (hadrons) \end{aligned} \tag{2.4}$$

2.2.4 The magnet

Both the tracker and the two calorimeters are within a 4 T magnetic field which is generated by a solenoidal magnet coaxial to the beam. The presence of the field allows the momentum measurement of the tracker thanks to the use of the curvature radius and at the same time reduces the effect of the pile-up by preventing the low energy particles to reach the ECAL barrel. Finally, the magnetic field in the return yoke is used for the reconstruction of the muon tracks in the muon chambers.

The CMS magnet system [38] consists of a superconducting coil housed in a vacuum tank and of a return yoke. The iron return yoke has a 12-sided cylindrical structure. The central part is divided in 5 coaxial rings, each one consisting of three layers where the muon chambers are hosted; the endcaps yokes instead are made of three disks, divided into 12 sectors. The superconducting coil is cooled down by liquid helium. It is housed in a vacuum tank which also works as supporting structure for the ECAL, the HCAL and the tracker. The main parameters of the magnet are given in table 2.3.

The choice of the field intensity is a compromise between trigger efficiency requirements and reconstruction requirements.

2.2.5 The muon system

The muon system is the outermost of the CMS subdetectors. Its main goals are the identification of muons, thanks to their high penetrating power, and a precise mea-

The magnet parameters	
Magnetic field at the interaction point	4T
Coil length	12.48 m
Stored energy	2.70×10^9 J
Magnetic radial pressure	6.47×10^6 Pa
Axial compressive force at mid plane	148×10^6 N
Circulating current	20 kA

Table 2.3: *Some properties of the CMS solenoid.*

surement of their momentum, with the help of the information coming from the tracker. The muon system also works as trigger for events which involve muons and it provides a precise time measurement of the bunch crossing.
The CMS muon system [39] relies on three kinds of gaseous detectors: drift tubes, cathode strip chambers and resistive plate chambers. The drift tubes and the cathode strip chambers provide an excellent spatial resolution and the resistive plate chambers have a very good timing. The active parts of the muon system are hosted into stations which are interleaved by the iron layers of the return yoke of the magnet. The longitudinal view of a quarter of the muon system is given in figure 2.17. The barrel extends up to $\mid \eta \mid < 1.3$, the endcaps up to $\mid \eta \mid < 2.4$.

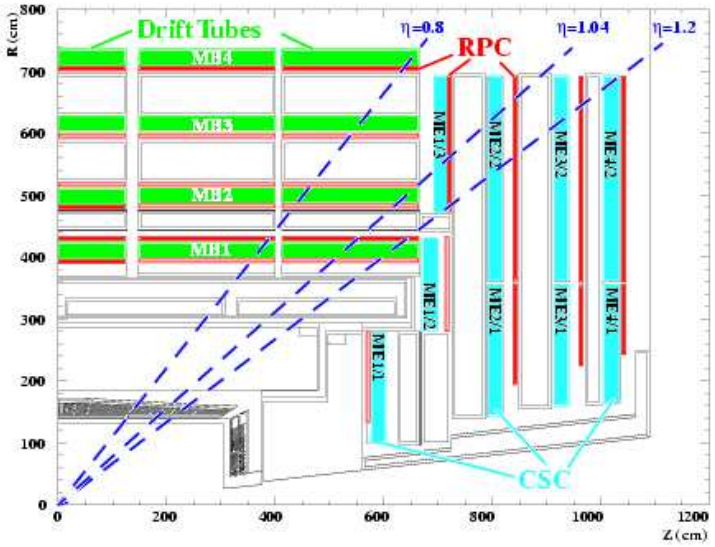

Figure 2.17: *Longitudinal view of one quarter of the CMS muon system.*

The **drif tube chambers** can very well operate in the barrel region, where the track occupancy is low, the neutron presence is negligible and the residual magnetic field is not too high thanks to the presence of the magnet return joke.
The basic element of a drift tube chamber is the drift cell, which is shown in figure 2.18. A stainless steel anode wire is placed between 2 parallel aluminium

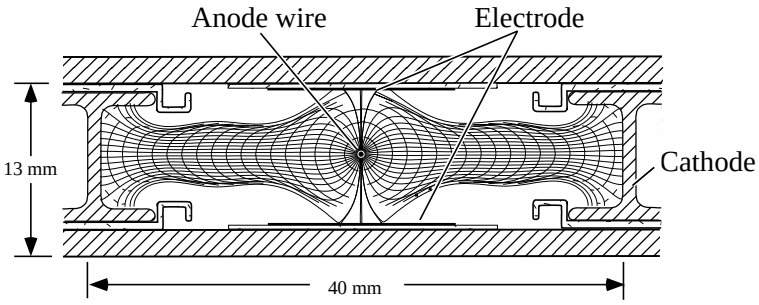

Figure 2.18: *Transverse view of a drift tube cell.*

layers; two 'I'-shaped electrodes, which define the boundaries of the cell, work as cathodes and shape the electric field. The distance of the track from the wire is measured by the drift time of the electrons and the chosen mixture of 80% Ar and 20% CO_2 ensures a good space-time linearity. The single cell, which works in condition of saturated drift velocity, has efficiency around 99.8% and spatial resolution at the level of 180 μm.
Four layers of parallel staggering drift cells are glued together to form a superlayer. This allows to solve the left-right ambiguity of a single layer. Each muon station consists of three superlayers; the central one measures z and the 2 others the azimuthal coordinate ϕ. Despite the long drift time of a single layer, the combination of the responses of the 4 layers ensures a good time resolution.

The **cathode strip chambers** are multi-wire proportional chambers. Being able to work also in a high radiation environment and in presence of inhomogeneus magnetic field, they have been chosen as detectors for the two endcaps. The chambers are composed of six layers, each one consisting of an array of anode wires between two cathode planes, one of which is segmented in the radial direction to provide the ϕ measurements. The region among the cathodes is filled with a mixture of Ar, CO_2 and CF_4 (30%, 50%, 20%). The passage of a particle induces a signal on many wires and strips and the particle position is obtained by an

interpolation, with a resolution between 50 and 100 μm for the ϕ measurements (done by the strips) and about 5 mm for r, measured by the wires. In each endcap station, the presence of 6 layers of cathode strip chambers improves the timing and gives an efficiency better than 99%.

The **resistive plate chambers** are used both in the barrel and in the endcaps to provide a fast answer which is suitable for triggering purposes. The RPCs have four bakelite electrodes forming 2 coupled gaps. They are filled with a gas mixture of freon (C_2 H_2 F_4, 95%) and isobutane (i-C_4 H_{10}, 5%). The outer face of the bakelite planes is covered with graphite to distribute the high voltage over the whole surface. The RPCs operate in the avalanche mode instead of the streamer mode, to better sustain the high flux of particles; the amplitude of the signal is smaller since the gas multiplication is reduced, but this is compensated by an electronic amplification. The readout is made by aluminium strips. The main characteristic of the RPCs is their excellent time resolution, which is better than 2 ns.

2.2.6 The trigger

At the nominal LHC luminosity, a total event rate of 10^9 Hz is expected. Given the typical size of a raw event ($\sim$1MB), it is impossible to record the information corresponding to all the events; the rate has therefore to be reduced to the order of 100 Hz, which is the upper limit for storing events. As it was discussed in chapter one, the rate is dominated by low transverse momentum events. The trigger system therefore must have a huge reduction factor and at the same time it must maintain high efficiency on interesting events. This requires for the trigger a level of complexity which is comparable with the offline reconstruction and at the same time the necessity to work fast. This is done in two main steps.

The **Level 1 trigger (L1)** [40] reduces the rate to about 50 (100) kHz for the low (high) luminosity phase. At a first level, the full data are stored in pipelines of processing elements, each one taking a decision in less than 25 ns. At each bunch crossing, each element passes its results to the following one and it receives new informations. The L1 decision about taking or discarding data from a particular bunch crossing has to be taken in 3.2 μs. If the first level trigger accepts the event, the data are moved to be processed by the HLT.
To deal with the 25 ns bunch crossing rate, the L1 trigger has to take decisions in a time which is too short to read all the raw data from the whole detector, therefore it uses the calorimetric and muons information only. It is organized in a Calorimeter Trigger and a Muon Trigger; they both pass the information to the Global Trigger which takes a global decision. The Calorimetric Trigger consists of trigger towers which match the granularity of ECAL and which are grouped in regions of 4 $\times$

4 trigger towers. Four categories of objects (electrons and photons, central jets, forward jets and τ jets) are analized and the best 4 candidates of each classes are passed to the Global Trigger together with the missing E_T. The Muon Trigger analizes separately the three different detectors of the muon system and then it passes the four best muon candidates to the Global Trigger. The Global Trigger uses a logical combination of the data with the corresponding thresholds and it takes a decision. The estimate output rate is almost a factor 3 lower than the sustainable one.

The **High Level Trigger (HLT)** [41] reduces the output rate down to 100 Hz. The idea behind the HLT software is the regional reconstruction *on demand*, that means that only the objects which are in useful regions are reconstructed and the not interesting events are rejected as soon as possible. The HLT can be splitted into three logical levels. At the first one, only the full information of the muon system and of the calorimeters is used; at the second level the information from the tracker hits is added and finally, at a third level, the full information is available.

Chapter 3

The electromagnetic calorimeter calibration

The calibration defines the ultimate performances of the CMS electromagnetic calorimeter. A careful monitoring and the individual calibration of all the channels are necessary to keep the constant term in the energy resolution below 0.5%, which is the ECAL target as discussed in section 2.2.2.
The ECAL calibration will be done in three steps. First, before the installation in CMS, the channels will be intercalibrated using electron beams, cosmic rays and laboratory measurements, to have an initial set of coefficients at startup. Then, when installed in CMS, all the crystals will be calibrated *in-situ* using physical events. Finally, since the radiation damage affects the transparency of the crystals, the response of each channel will be constantly monitored with a monitoring system based on the injection of laser light.
In the following, the ECAL calibration procedure is described. In the first part of the chapter, some details about the pre-calibration and the *in-situ* calibration are given; in the second part, the radiation damage is discussed and the concept of the laser monitoring is presented. Results on calibration issues at test beam will be given in the next chapter.

3.1 The ECAL pre-calibration

The overall initial spread in the channel response is expected to be about 8%. To improve the pre-calibration precision, all the supermodules were initially foreseen to be tested on dedicated electron beams at two energies before the installation in CMS. The clear situation at test beam indeed allows to have an initial set of coefficients at startup with great precision (of the order of 0.5% - 1%), using the procedure which will be discussed in detail in the next chapter.

Due to time constraints, only few supermodules will be pre-calibrated in this way and alternative strategies have been suggested to provide initial coefficients for the remaining crystals. The estimate of the intercalibration precision which can be reached using such alternative methods is necessary.
The first proposed strategy exploits the laboratory measurements performed during the assembly phase that all the crystals have to undergo before being accepted for ECAL [42] [43]. Such measurements aim at determining with the highest possible precision the optical properties of the crystals and the gain of the electronics for each crystal. The light yield is obtained by measuring the photopeak position in the scintillation energy spectrum of a ^{60}Co source [44].
A great effort has been recently done to improve the stability and the precision of the measurements and an independent way of determining the light yield has been found, which exploits the correlation between the light yield and the crystal longitudinal transmission at the wavelength λ = 360 nm [45]. The precision on the intercalibration which can be reached combining all the available measurements is about 4-5% (fig 3.1), as it was demonstrated by the comparison with test beam data [46].

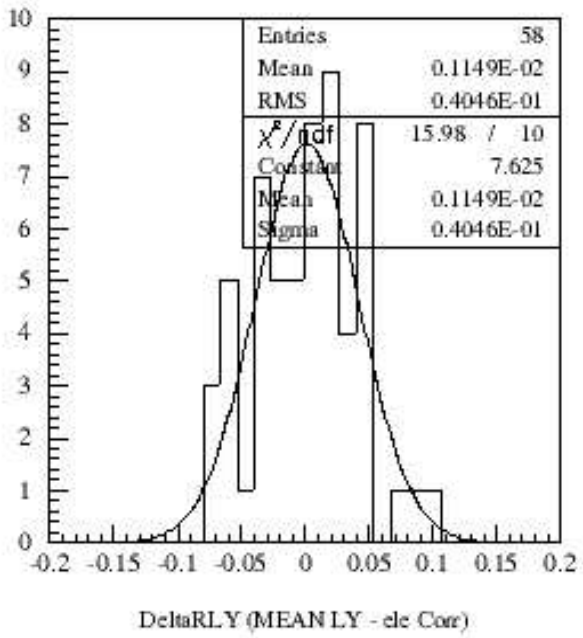

Figure 3.1: *Relative difference between the intercalibration coefficients from laboratory measurements (with the combination of the direct light yield measurement and of the prediction of the light yield from the longitudinal transmission) and the intercalibration obtained with 2003 ECAL test beam data [46].*

The previous procedure is performed on individual components before assembly. A possible approach to pre-calibrate the full detector is the use of cosmic rays [47] [48]. The idea consists in preparing a comics telescope and in selecting only those events in which the muon traverses the crystal along the direction parallel to its axis. This can be done by choosing only the events where one single crystal presents a significative signal by vetoing in the neighboring crystals or by using an external tracking device. A complete simulation of the ECAL response to cosmic rays has been recently produced [49], which has shown the feasibility of a calibration with cosmic rays. The comparison of real cosmic data and test beam data has shown that the overall precision in the determination of the calibration constants is about 3.5% (fig 3.2). Given the large associated statistical uncertainty (cosmic data have been collected only for few days), the ultimate systematic uncertainty can be estimated to be between 2% and 3%. A longer test to calibrate the

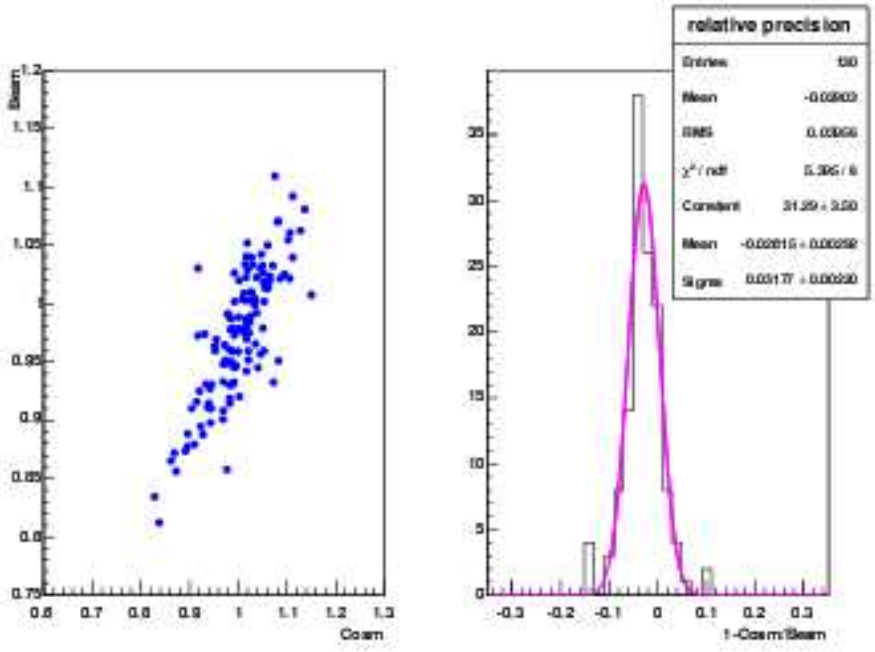

Figure 3.2: *Left: correlation between the intercalibration coefficients from test beam and from the test with cosmic rays. Right: relative precision of the cosmics calibration constants. Only those channels with more than 50 events for cosmic run are considered [49].*

ECAL supermodules with cosmic rays started in summer 2005 and it is currently on-going.

3.2 The *in-situ* calibration

After the installation in CMS, ECAL will undergo a calibration procedure *in-situ* using physical events. Three main different steps are foreseen. First rings of crystals in the same η interval will be intercalibrated exploiting the ϕ-symmetry of the energy deposition. The different η rings will be then intercalibrated using the $Z \rightarrow e^+e^-$ decays, which will be also used to set the absolute energy scale. As soon as the tracker is aligned, the intercalibration of different crystals within a single module will be done using the energy over momentum measurements of isolated electrons coming from the $W \rightarrow \nu e$ decay, which allows higher statistics than the $Z \rightarrow e^+e^-$ decay. The use of the $\pi^0 \rightarrow \gamma\gamma$ and $\eta \rightarrow \gamma\gamma$ processes has been also foreseen at low energy. The target uniformity is at the 0.5% level.

3.2.1 Intercalibration using the ϕ-symmetry

The ϕ-symmetry method exploits the fact that the total transverse energy which is deposited from a large number of events is the same for all the crystals belonging to the same η ring. The intercalibration can therefore be performed by comparing the total energy deposited in each crystal with the mean of the distribution of total energies for all the crystals at that pseudorapidity.
The method has been tested on a sample of 11 millions of jet trigger events [50], which correspond to about 3 hours at startup assuming 1kHz Level 1 Trigger bandwith dedicated to jet triggers. Without using any knowledge about the material distribution in the tracker, the limit on the precision to which crystals can be intercalibrated in ϕ was found to be close to 1.5% throughout the barrel and between 1.0% and 3.0% for the fiducial regions of the endcaps. The precision is mainly limited by the inhomogeneity of the distribution of tracker material. This result is shown in figure 3.3.
Similar results have been also obtained using minimum-bias events [51]. Nevertheless, due to the higher transverse energies which are characteristic of jet trigger events, the use of jet triggers is likely to be a more robust alternative to the use of minimum bias events, which was proposed first.

3.2.2 Intercalibration using $Z \rightarrow e^+e^-$ events

At the LHC energy $\sqrt{s}$ = 14 TeV, the $Z \rightarrow e^+e^-$ cross section is about 1.7 nb.
Together with the clear signature of the channel, this assures enough data both to set the ECAL absolute energy scale and for the intercalibration. Z events can be used to intercalibrate pairs of rings in η exploiting the electron pair invariant mass reconstruction. Since the Z gives energetic correlated electrons in different regions of the ECAL, a large fraction of events can be used to intercalibrate the

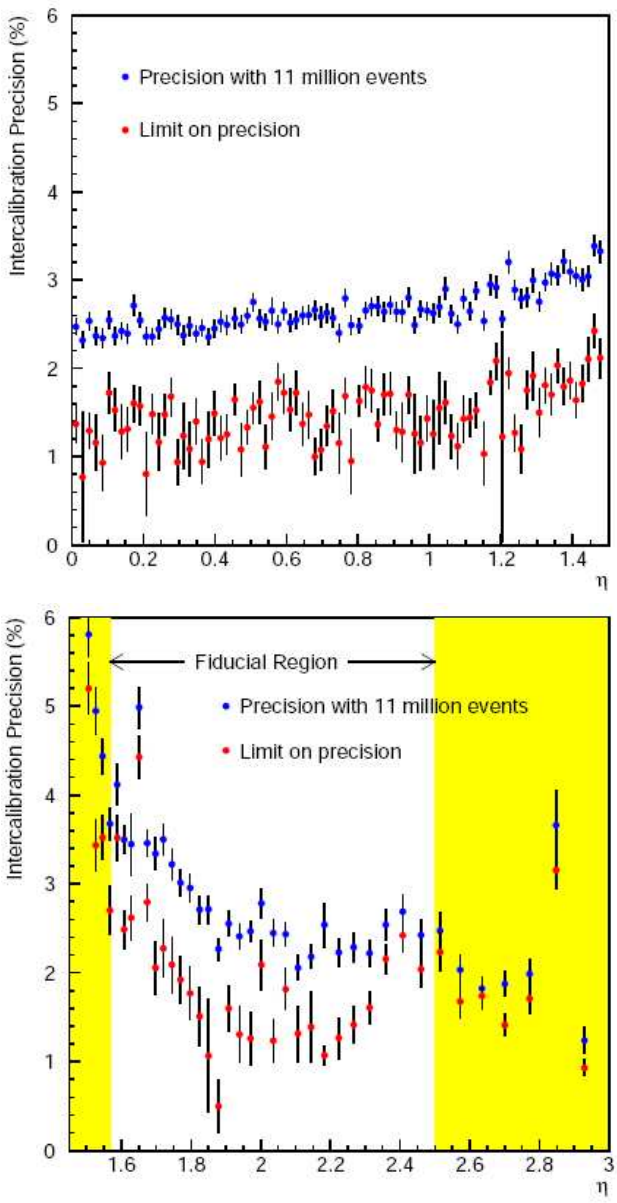

Figure 3.3: *Intercalibration precision obtainable with 11 million Level1 jet trigger events and limit on the intercalibration precision due to the tracker material inhomogeneity as a function of η for (top) the 85 pairs of rings of crystals in the ECAL barrel and (bottom) the 39 pairs of rings of crystals in the ECAL endcaps. The fiducial region is the region of the ECAL used for high-precision calorimetry [50].*

endcaps with respect to the barrel.
A problem related to the method is the fact that the shower from an electron involves about 25 crystals, so the calibration constants have to be deconvoluted. To solve the deconvolution problem, an iterative method developed to solve a similar problem at the L3 experiment [52] has been proposed. Looping over the events, the η ring intercalibration constants are computed as

$$C_N(\eta_i) = C_{N-1}(\eta_i) \cdot \frac{\sum_n \frac{M_Z}{M_n^{inv}} w_{i,n}}{\sum_n w_{i,n}} \tag{3.1}$$

where N is the iteration index, n is the event index and the sum is done over the events with one energy cluster in the ring i. The calibration constants are in this way deconvoluted by the weight factors w_i which are proportional to the deposited energies. The iteration stops when the calibration constants converge.
It has been shown [53] that with some proper electron selections a barrel ring intercalibration precision of 0.6% can be obtained with an integrated luminosity of 1 fb^{-1}. The obtained precision ranges from 0.3% in the type 1 module region up to 1.3% in the type 4 module region, being this difference mainly due to the change with the pseudorapidity in the efficiency in selecting the electrons for the analysis. The method converges after few iterations and the result is not affected by the initial miscalibration. The possibility to set the absolute energy scale using Z events was also tested and it was proved that a variation of the global miscalibration scale from -8% to 6% can be followed with a precision better than 0.05% using events corresponding to an integrated luminosity of 1.0 fb^{-1}. This is shown in figure 3.4.

3.2.3 Intercalibration using $W \to e\nu$ events

The $W \to e\nu$ channel will be exploited to intercalibrate the individual crystals by combining tracking and calorimetric informations.
As in the Z channel, the shower involves an array of crystals so the calibration constants have to be deconvoluted. Two methods are currently under study, the matrix inversion algorithm [54] and the L3 algorithm. In the latter, the quantity $E_{Ecal}/P_{Tracker}$ replaces the M_Z/M_{inv} ratio in eq 3.1. A serious problem is represented by the bremsstrahlung emission, which is responsible for a long tail in the E/P distribution. Tight cuts are applied to reject the events with large bremsstrahlung emission. The hardness of these cuts is a compromise between the purity of the selected sample and the time necessary to acquire the statistics needed for a precise calibration.
It has been shown [55] [56] that the precision which can be reached depends on the pseudorapidity of the crystals and on the quality and the number of selected

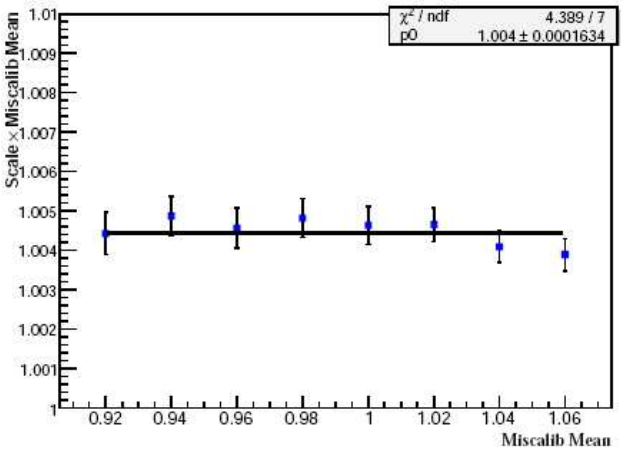

Figure 3.4: *Scale factor multiplied by the mean of the global miscalibration value versus the injected miscalibration mean value. A fit with a constant function is superimposed [53].*

electrons per crystal. With the proposed cuts to select non radiating electrons, the target of 0.5% calibration precision is achieved with an integrated luminosity of 5 fb^{-1} at low luminosity for the low pseudorapidity region up to $\mid \eta \mid < 1$. For higher η, the calibration precision decreases and it is expected to vary between 1.0% and 1.5% in the endcaps with 7.2 fb^{-1} (fig 3.5).
The described procedure requires to have the whole detector commissioned and the tracker already aligned. A smearing of the track momentum at 2% level has been applied to study the effect of the tracker misalignment and miscalibration. The calibration precision for ECAL is reduced by 18%, compared with the case of a perfect momentum measurement. For this reason, the required time scale could be larger than the time strictly needed to collect data.

3.3 The monitoring with the laser

It has already been mentioned in the second chapter that the radiation damage reduces the transparency of the crystals. Since the effect is different from crystal to crystal, this affects the system calibration and it is therefore necessary to correct for it. To cope with this effect, a laser monitoring system has been developed to follow the time evolution of each channel response between two successive calibrations with physical events. The radiation damage and the monitoring strategy are discussed in the following.

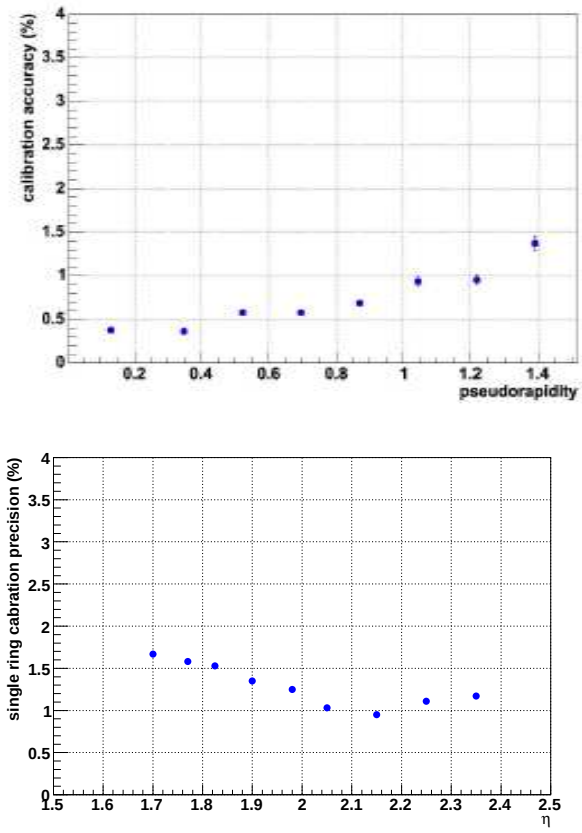

Figure 3.5: *Calibration precision as a function of the pseudorapidity using $W \rightarrow e\nu$ events. On the top: the ECAL barrel, for an integrated luminosity of 5 fb^{-1} at low luminosity. On the bottom: the endcaps, for an integrated luminosity of 7.2 fb^{-1} at low luminosity [55][56].*

3.3.1 The scintillation mechanism and the radiation damage for $PbWO_4$ crystals

The lattice structure of the $PbWO_4$ plays a crucial role in its scintillation mechanism. In the lattice it is possible to identify WO_4^{2-} groups linked to Pb^{2+} ions. Each Pb^{2+} ion is surrounded by eight oxygen atoms which belong to the WO_4^{2-} structure; the latter is a tetrahedron with the W^{6+} ion in the centre, surrounded by four O^{2-} ions at the vertices.
When an highly energetic particle crosses the $PbWO_4$, the electrons of the Pb^{2+} in the valence band can be excited to the conductive band. Such electrons lose part of their energy with a non radiative transition and descend to energy levels which are just below the conductive band and which are related to the WO_4^{2-} structure. The

successive deexcitation, bringing the electron to the ground state, can occur with the emission of a photon (radiative process) or by exchanging phonons with the crystal lattice (non radiative process). The balance between the two mechanisms determine the scintillation of the crystal. Since the non radiative process presents a strong dependence on the temperature and the radiative process does not, both the light yield and the decay time depend on the temperature itself.

An important role in the scintillation mechanism of the $PbWO_4$ is played by possible defects in the lattice [57][58]. A perfect $PbWO_4$ cell would emit blue light ($\lambda \sim 420$ nm) with an exponential behaviour with time constant $\tau_1 \sim 5$ ns. The presence of the defects in the lattice creates allowed energy levels between the conduction and the valence band, so influencing the scintillation light spectrum.
In the $PbWO_4$ the defects mainly originate from a oxygen ions vacancy or from the deficiency of the cation in the W position in the tetrahedron. In the first case, an oxygen atom escapes and a pair composed by a WO_3 molecule and a vacancy $V_{O^{2-}}$ is formed to compensate the local charge imbalance. The WO_3 centers are responsible of light emission in the green band with a time constant $\tau_2 \sim 14$ ns; the number of emitted green photons is almost equal to one third of those emitted by a perfect cell. In the case of the W vacancy, an electron from a lead ion falls into an oxygen vacancy. Eased by the absence in the lattice of a W^{6+} ion, the creation of a lead vacancy may cause one of the electrons of a distinct Pb^{2+} ion to be trapped in $V_{Pb^{2+}}$ when decaying from its excited state. These color centers are responsible for the red band in the scintillation light, with a long time constant $\tau_3 \sim 110$ ns and with an intensity which corresponds to about 5% of the blue photons.
Due to the presence of the three kinds of color centers, the dynamic of the light emission can be parameterized as the sum of exponentials with different amplitudes and time constants.

When a sample of $PbWO_4$ is exposed to intense irradiation, some damages in the lattice structure can arise and the defects play an important role in the process. The electrons and the holes which are created by the radiation separate in the cooling and in the diffusion processes and they can eventually fall into the traps which arise in the lattice due to the defects in the structure. The local charge imbalance which is created in this way is recovered in a new local equilibrium with the creation of new color centers.
According to this model, the radiation affects the light transmittance, but it does not change the scintillation mechanism and does not influence the intrinsic light yield; the loss in the light production is therefore only due to the loss in transparency. This was confirmed by measurements of excitation and luminescence spectra, which give informations on the available energy levels. Figure 3.6 [59] shows

that no change in the spectra occurs with the irradiation.

Due to the role that the defects have in the radiation damage, an intensive R&D

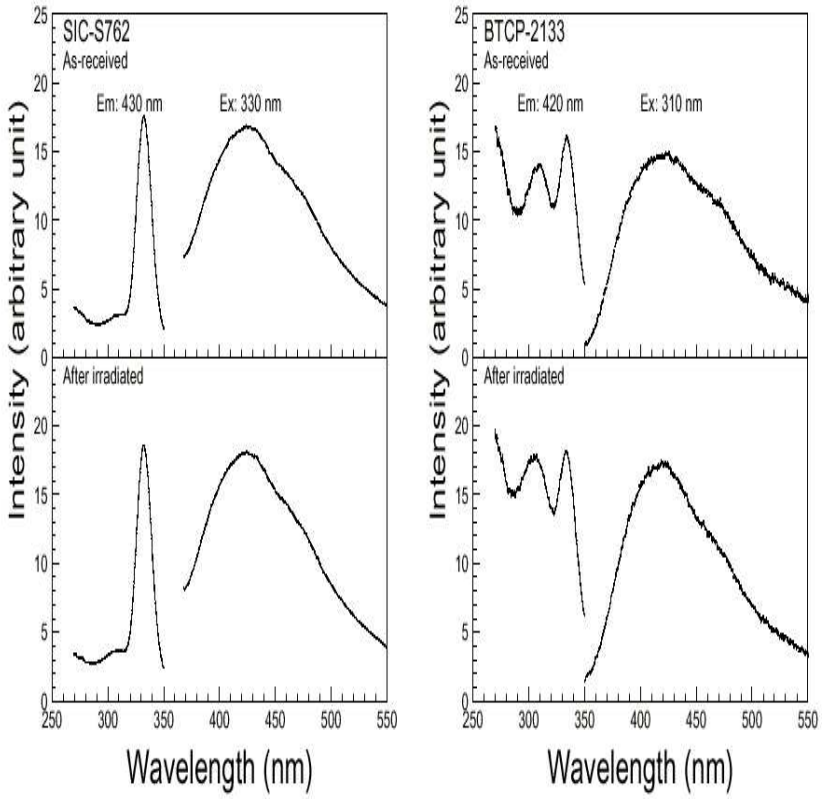

Figure 3.6: *Excitation and photo luminescence spectra before (top) and after (bottom) the irradiation. No evident change can be seen. Left: a crystal from SIC (Shanghai Institute of Ceramics). Right: a crystal from BTCP (Bogoroditsk Techno-Chemical Plant) [59].*

has been devoted to reduce them. Since the defects originate from the vacancy of an oxygen or a tungstate ion, their concentration changes according to the concentration of W and O in the melt, which varies during the growth of the crystal. That causes a dependence of the defects concentration on the position along the crystal axis, so affecting the uniformity of the crystal response. The growth of the crystal in a stoichiometric controlled melt or the introduction of doping elements during the growth can help in reducing the non uniformity of the defects concentration. Nobium and Yttrium have been chosen as dopings for the ECAL crystals, after it was shown that they improve the transmittance, the uniformity and the radiation hardness of the crystals [58] [60].

The radiation induced color centers spontaneously annihilate and the two processes of creation and annihilation coexist in the lattice [59]. The annihilation rate is proportional to the density N of color centers while the production rate is proportional to the dose rate R and to the density of traps still available to create color centers. The variation of N in a time interval dt during the irradiation is therefore

$$dN = \sum_{i=1}^{n} \left\{ -a_i N_i + \left(N_i^{\text{Max}} - N_i \right) b_i R \right\} dt \tag{3.2}$$

where the sum is performed on the n different kinds of color centers. The solution of equation 3.2 is

$$N(t) = \sum_{i=1}^{n} \left\{ \frac{b_i R N_i^{\text{Max}}}{a_i + b_i R} \left[1 - e^{-(a_i + b_i R)t} \right] + N_i^0 e^{-(a_i + b_i R)t} \right\} \tag{3.3}$$

being N_i^0 the initial density of the i-th color center type. After an initial increase, an asymptotic equilibrium is reached for the color centers density, which approaches the value

$$N_{eq} = \sum_{i=1}^{n} \frac{b_i R N_i^{\text{Max}}}{a_i + b_i R} \tag{3.4}$$

The concentration is higher with higher dose rates and this influences the crystal light yield, as it is shown in figure 3.7.

3.3.2 The idea of the monitoring system

The changes in the crystal response to the impinging particles can be ascribed both to the radiation damage and, on a longer term scale, to ageing effects. The monitoring system discussed in section 3.3 was developed to follow the evolution of such response. Here we concentrate on the monitoring of the radiation induced changes.

The crystal responses to the passage of a particle with energy E can be parameterized as [61]

$$S = E \int N(E, z) C(t, z, \lambda) P(T, z, \lambda) \overline{M}(V, T, t, \lambda) d\lambda dz \tag{3.5}$$

where

- $N(E, z)$ is the longitudinal profile of the scintillation light. It depends on the particle type and energy and on the impinging angle with respect to the

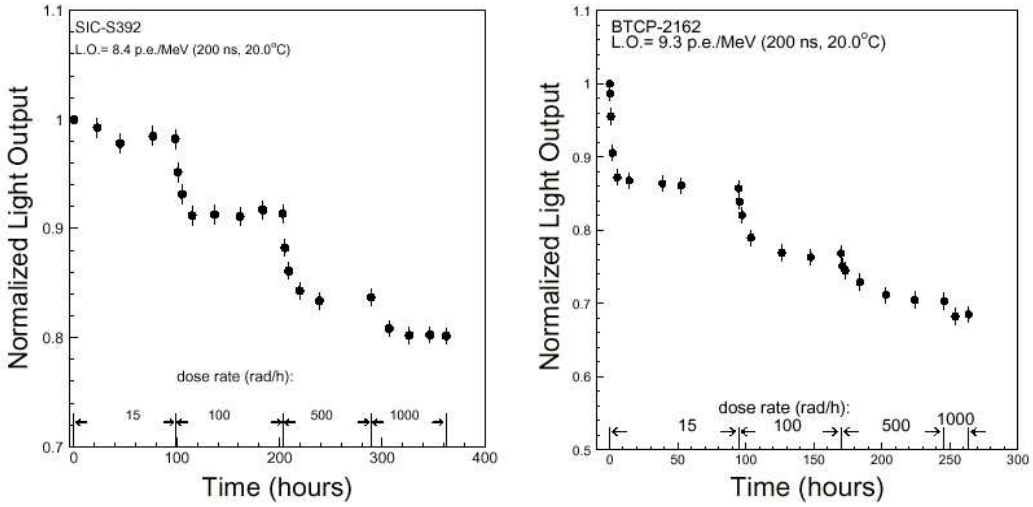

Figure 3.7: *Light output as a function of time. Left: a crystal from SIC. Right: a crystal from BTCP. The different dose rates are shown on the time axis.*

crystal axis. The average longitudinal shower profile for 120 GeV electrons which hit the center of the crystal is shown in figure 3.8

- $C(t, z, \lambda)$ is the transmission term for the scintillation light. It combines the geometrical acceptance and the light attenuation
- $P(T, z, \lambda)$ is the spectrum of the scintillation light in term of number of photons per unit of deposited energy and wavelenght
- $\overline{M}(V, t, T, \lambda)$ is the photodetector response. It combines the quantum efficiency and the intrinsic gain of the APD

In eq 3.5, t is the time, λ is the wavelength of the scintillation light, z is the position along the crystal axis, T is the temperature and V is the bias voltage for the APD.
As it was discussed in par 3.3.1, the radiation damage does not affect the scintillation mechanism, therefore we are here only interested in the transmission term C. The time evolution of the transmittance is regulated by the time evolution of the attenuation length

$$\Lambda(t, z, \lambda) = 1/\rho(t, z) \cdot \sigma_C(\lambda) \tag{3.6}$$

where we indicate with $\rho(t, z)$ the density of radiation induced color centers and with $\sigma_C(\lambda)$ their cross section for interaction with a wavelength λ. The basic idea

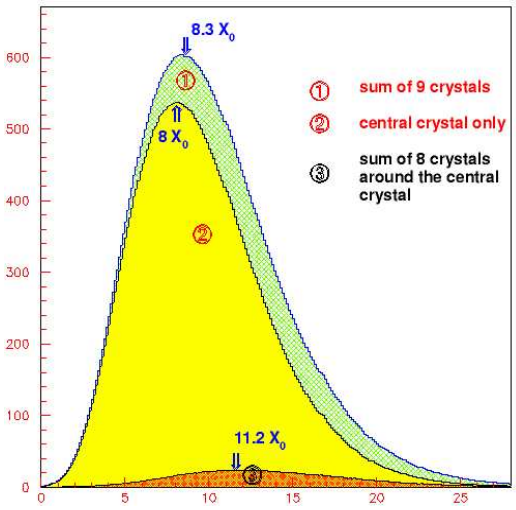

Figure 3.8: *Shower profile in $PbWO_4$ for 120 GeV electrons impinging at the center of the crystal [61].*

of the monitoring system is to control the time evolution of $\Lambda(t, z, \lambda)$ by looking at the evolution of the crystal response to laser light with wavelength λ.
The response to the laser light can be parameterized as

$$R = a(t, \lambda)L(t, \lambda)\beta(t, \lambda)\overline{M}(V, t, T, \lambda) \tag{3.7}$$

Since the laser amplitude can fluctuate over time, as will be discussed in the next chapter, it is monitored via a PN diode. In equation 3.7, $L(t, \lambda)$ gives the number of photons arriving on the PN, $\beta(t, \lambda)$ is the transmission of the laser light and $a(t, \lambda)$ gives the transmission of the fiber feeding a given channel relative to the one feeding the PN diode, which can be considered with good approximation constant.

The accurate control of transmission terms of the injected light β and of the scintillation light C is crucial for the monitoring system.
The different patterns followed by the laser light to propagate from the injection point to the APD can be considered similar. With such approximation, the transmittance of the laser light in a crystal with lenght L is given by

$$\beta(t, \lambda) = \beta_0(\lambda) \exp\left(-r_1 \int_0^L \frac{dz}{\Lambda(t, z, \lambda)}\right) \cdot \frac{1}{1 - k \exp\left(-2r_2 \int_0^L \frac{dz}{\Lambda(t,z,\lambda)}\right)} \tag{3.8}$$

$\beta_0(\lambda)$ is the initial transmittance and it includes the geometrical acceptance and the optical couplings; r_1 and r_2 are factors close to 1 which account for the dif-

ferent patterns followed by the light and the last term takes into account the multiple reflections along the crystal with a reflection coefficient k. Equation 3.8 clearly shows that the laser light gives informations about the average evolution of $\Lambda(t, z, \lambda)$ but it can not be used to investigate its z dependence, since the attenuation length only enters the integration along the whole axis.
The scintillation light can basically follow two different patterns to reach the APD, the first one going from the scintillation point directly to the rear crystal face equipped with the APD and the other moving first towards the front face of the crystal and then back to the APD after a reflection. With the same approximation introduced for the laser light, the transmittance for a scintillation happening at the z position along the crystal axis is given by

$$C(t,z,\lambda) = \left(\Omega_1(z) e^{-s_1 \int_z^L \frac{du}{\Lambda(t,u,\lambda)}} + \Omega_2(z) e^{-s_2 \int_0^L \frac{du}{\Lambda(t,u,\lambda)}} e^{-s_3 \int_0^z \frac{du}{\Lambda(t,u,\lambda)}}\right) \cdot \cdot \frac{1}{1 - k' \exp\left(-2 s_4 \int_0^L \frac{du}{\Lambda(t,u,\lambda)}\right)}, \tag{3.9}$$

where $\Omega_1(z)$ and $\Omega_2(z)$ represent the geometrical acceptance for the scintillation light emitted at z respectively for the path without and with reflection on the crystal front face.
Equations 3.8 and 3.9 suggest that the relation between β and C is not simple. If we call R_0 the response to the laser before irradiation and $R(t)$ the response at time t, the evolution of Λ can be deduced from the ratio $R(t)/R_0$ on the basis of a model for the evolution of the radiation induced color centers density. Then Λ is used as input for the equation 3.9 to calculate the evolution of the transmittance for the scintillation light. In this way the change in the crystal response to an impinging particle can be evaluated and corrected for.
A different possibility is to establish a direct relation between the change in the response to the impinging particles ($S(t)/S_0$) and to the laser ($R(t)/R_0$). Using the equations 3.8 and 3.9 and following the discussion in [61], the evolution of the electron signal compared to the monitoring signal can be approximated with

$$\frac{S(t)}{S_0} \simeq \left(\frac{R(t)}{R_0}\right)^{\alpha} \left[1 + \eta ln\left(\frac{R(t)}{R_0}\right)\right] \tag{3.10}$$

The parameter α depends on the coefficients introduced in equations 3.8 and 3.9 and on the attenuation coefficients for the monitoring and the scintillation light. η is a parameter of the order of ± 0.1 to ± 0.5, therefore an almost linear relation between S and R is expected in the domain $0.95 \leq R/R_O \leq 1$. Relation 3.10 will be used in the analysis discussed in the next chapter.

Chapter 4

The 2003 ECAL test beam

In recent years an extensive test beam programme has been conducted on some modules of ECAL. The campaign aimed at testing the response of the calorimeter in a working condition close to the final one and at the same time in a controlled environment like the one offered by the test beam.
In 2003 two partially equipped supermodules, $SM0$ and $SM1$, underwent a full system test to check the calorimeter performances. Among the main goals of the test beam there was the demonstration of the feasibility of the precalibration strategy and of the reliability of the monitoring system.
In this chapter, the analysis of the data collected at the 2003 test beam is presented. First a quick description of the test beam setup and of the data taking is given. The intercalibration strategy at the test beam is then presented, together with some results about the possible systematics affecting it and the statistical precision which can be reached. In the last part of the chapter, the monitoring of the crystals behaviour with the laser is discussed and a procedure to correct for the radiation damage effects is introduced.

4.1 The H4 beam

The 2003 ECAL test beam took place on the H4 extraction line of the SPS, which is located in the CERN North Area.
The electron beam in H4 is a tertiary beam which can be produced using two main configurations. In the first one, which is show in figure 4.1, two magnets $B1$ and $B2$ bend the primary proton beam coming from the SPS before it impacts on a berillium target $T2$. The secondary particles are bent by the magnet $B3$ according to their charge, so that the positively charged particles reach the H2 extraction line and the negatively charged particles go towards H4. These last ones, which are

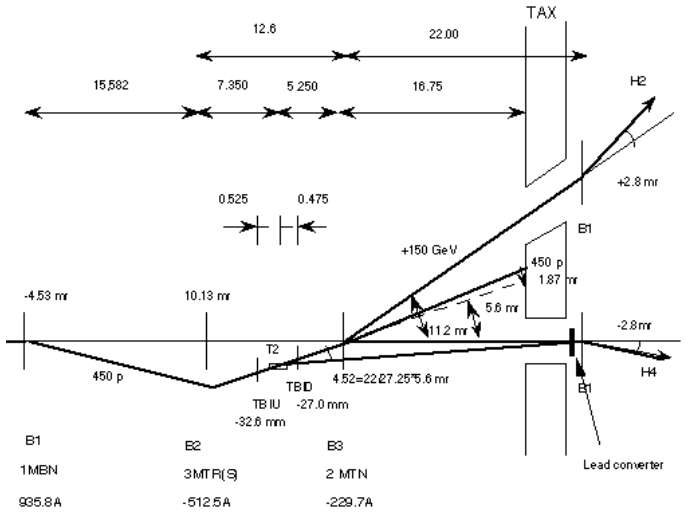

Figure 4.1: *Optics configuration for the production of an electron beam at H4.*

backgrounds for a clean electron beam, are then eliminated by some optics in the following. The π^0 produced by the protons impacting on $T2$ decay into photons. The fraction of photons going to H4 impacts on a lead converter originating e^+e^- pairs; some optics which follow then select only the electrons.
In such configuration, it is difficult to produce a large number of electrons with high momentum, due to the impinging angle of the primary particles on $T2$. A second configuration in which the primary beam is not bent is also foreseen. The beam impacts perpendicularly on $T2$, then the magnet $B3$ provides the high magnetic field which is necessary to avoid the proton beam enter H4. In this way an higher flux of particles can reach H4 also at high energy. Typical values are about thousands electrons per spill at 280 GeV and about 200 electrons per spill at 300 GeV. This configuration, which is shown in figure 4.2, doesn't give an electron beam in H2, but it can used to produce a pion beam. A set of collimators allows to limit the transverse momentum and the size of the beam and the number of electrons per spill.

4.2 The test beam setup

In the test area, the crystals are put on a rotating table which allows to reproduce the almost pointing geometry of CMS. Each crystal is positioned on the beam with a 3 degrees tilt between its axis and the beam direction, such as if the particles were originated in the LHC interaction point. Six scintillators, called $S1$ to $S6$, work as trigger along the beam line. In the standard configuration the trigger

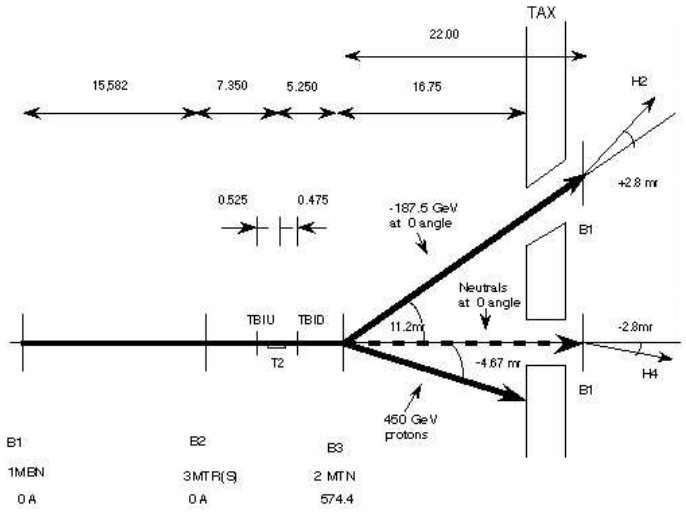

Figure 4.2: *Alternative optics configuration for the production of an electron beam at H4. It allows a more intense electron beam with respect to the configuration in figure 4.1.*

selects events on the whole surface of the crystal.
The impact position of the particles on the crystal surface is given by a hodoscope system [62]. In consists of two stations, each one with two layers equipped with 1 mm scintillating fibers. The two stations are separated by 2.5 m and the fiber planes are placed orthogonally. Each plane provides the particle position with resolution around 200 μm and the combination of the two stations allows to reach a 150 μm resolution on the two coordinates orthogonal to the beam.
A scheme of the setup is shown in figure 4.3. The main difference between the test beam setup and the final experimental one are the absence of material in front of the crystals and the absence of the magnetic field.
During the 2003 test beam some crystals belonging to the supermodules SM0 and SM1 have been tested; the maps of the two supermodules are given in fig 4.4 and 4.5. In both cases the active channels were type 15 and 16 crystals belonging to the fourth module, which corresponds to the preudorapidity region around 1.26 and 1.35.

Both SM0 and SM1 were built according to the final design. In the following a short description of the cooling system, the monitoring system and the readout electronics is given.

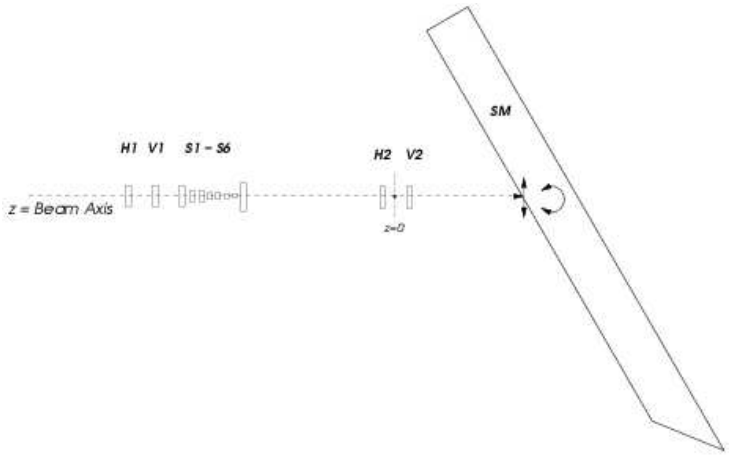

Figure 4.3: *The test beam setup at H4. The four scintillating planes composing the hodoscopy system (H1, H2, V1, V2), the scintillators working as triggers (S1 to S6) and the supermodule position are shown.*

The readout electronics

Both the old and the new electronic chains, described in section 2.2.2, have been tested. The SM0 crystals were equipped with FPPA chips, having four parallel amplifiers with different gains ($\times 1$, $\times 5$, $\times 9$, $\times 33$) to amplify the signal and comparators to select at each clock the highest gain giving a signal within the specified range. The FPPA had already been tested on the beam in 2002. The SM1 crystals were equipped with the MGPA electronics, which had never been used before at the test beam. The MGPA chips have three parallel amplifiers working at three different gains ($\times 12$, $\times 6$, $\times 1$), covering respectively the energy range 1-140, 140-300, 300-1250 GeV.

The laser monitoring system

At the 2003 test beam two different laser systems were available, providing lights of four wavelengths: blue (440 nm), green (495 nm), red (700 nm) and infrared (800 nm). The laser system was made of a Nd:YLF pump and of a tunable Ti:S laser, with quartz fibers splitting the light and bringing it to each crystal according the scheme discussed in the second chapter. To avoid fluctuations in the laser amplitude due to the temperature, the system was placed in a thermalized area. The laser pulse amplitude was monitored both via PN diodes placed within the ECAL module and via an independent system with its own electronics and acquisition. The stability of the laser system will be discussed in the following.

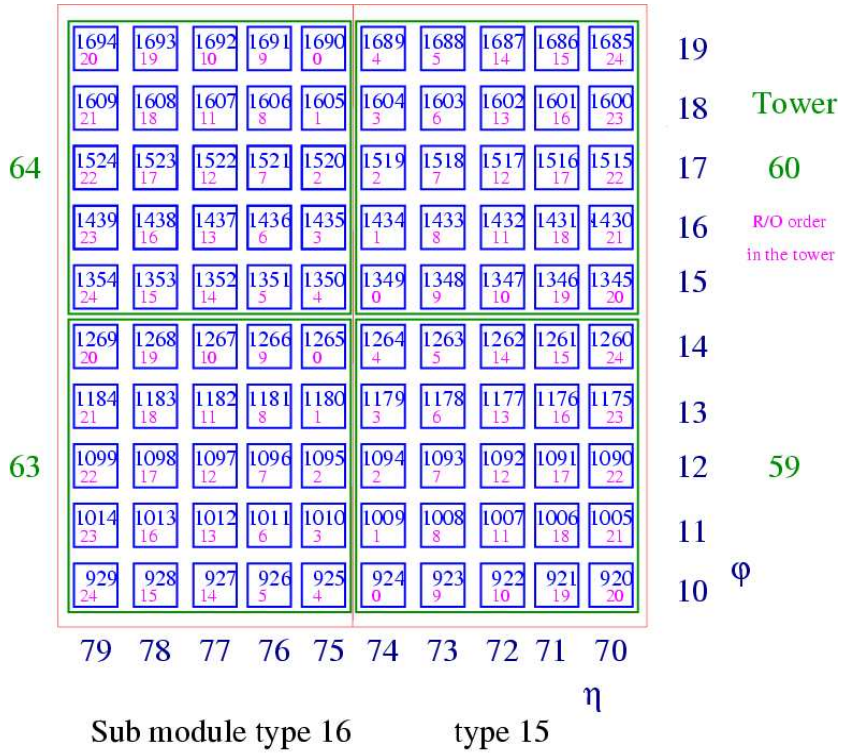

Figure 4.4: *Scheme of the numbering and of the positions of the crystals inside the supermodule SM0. The tower numbering refers to the number of the trigger towers inside the supermodule.*

The cooling system

The ECAL cooling system [63] employs a water flux to thermalize the detector at the level of 0.05°C. An aluminium grid separates the crystals from the mother boards which host the electronic cards. The water reaching the supermodule is first split in parallel streams which run inside the grid, then it is collected by the grid return pipes and it is distributed to a set of aluminium cooling bars. Such devices are connected in parallel and have been designed to absorb the heat dissipated by the electronic cards which are thermally coupled to the bars themselves. The water flowing in the cooling circuit is thermalized at the level of 0.01°C.
The 2003 test beam was the first test for the final cooling system design. Differently with respect to the final detector, the thermalization at the test beam has also to take care of changes in the environment temperature. An additional cooling

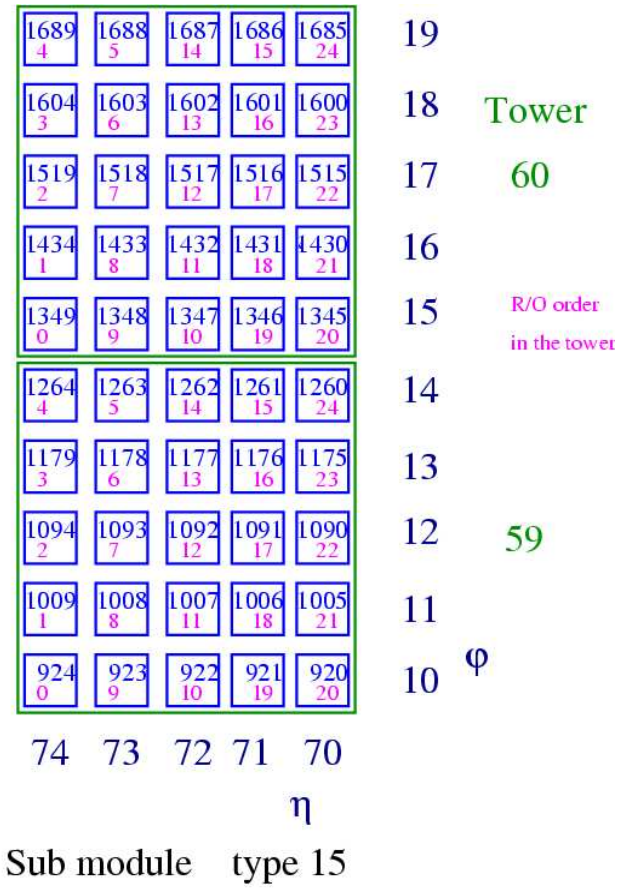

Figure 4.5: *Scheme of the numbering and of the positions of the crystals inside the supermodule SM1. The tower numbering refers to the number of the trigger towers inside the supermodule.*

system therefore was used to insulate the module from the environment.

4.3 Data and signal amplitude reconstruction

The measurements done at the test beam can be organized in four types of runs:

- **beam runs**, in which the events originate from the impact of particles on the crystal. The particles arrive in spills which are set along the H4 line once for each SPS cycle. The number of particles per spill is set through the collimators placed along the line and it is around 2 k in the calibration runs

and between 20 k and 40 k in the irradiation runs. The number of events per run is typically between 20 k and 30 k. The acquisition rate, which is dominated by the acquisition dead time, is about 500 events per spill

- **laser runs**, which are recorded during the inter spill of the SPS cycle in parallel over all the crystals. A typical laser run has 1500 events per wavelength
- **tempdark runs**, in which temperature and dark current measurements are performed at the same time. The information for all the crystals is read in parallel
- **pedestal runs**, in which the pedestal for each channel is registered in order to subtract its contribution when reconstructing the pulse shape. The pedestal runs are recorded during the interspill of the beam at the rate of 500 events per spill. The electronics works with forced gain and data for each available gain are registered

At the 2003 test beam, 14 samples of the pulse sampling were digitised and stored at each clock of the ADC. Two different techniques were then available to reconstruct the pulse amplitude starting from the samples, the *analytic fit method* and the *weights method*.
The analytic fit method consists in fitting the samples with the function

$$f(t) = A\left(\frac{t - t_{\text{off}}}{t_{\text{peak}}}\right)^{\alpha} \exp\left(-\alpha\frac{(t - t_{\text{off}}) - t_{\text{peak}}}{t_{\text{peak}}}\right), \tag{4.1}$$

where t_{peak} is the time corresponding to the maximum of the pulse and t_{off} accounts for the fact that some samples are recorded before the start of the signal and it gives the time offset between the start of the pulse and the first sample. Function 4.1 is determined starting from the transfer function of the readout electronics. At the test beam the clock runs independently of the trigger and a TDC is therefore employed to register the time difference between the start of the pulse and the first sample following the clock. This information allows to average on many events to compute the values for α and t_{peak} for each crystal. The pulse shape is then fitted in each event to obtain the amplitude factor A as an estimate of the energy. To improve the agreement with the data, the fit is restricted to the region from one sample before the maximum to three samples after it. At the 2003 test beam, the values for α and t_{peak} have been found to be uniform between the different crystals and a common value has been therefore set for all the channels.
The second method used to reconstruct the pulse shape is the weights method [64]. The signal amplitude is computed as the sum over the samples

$$A = \sum S_i W_i \tag{4.2}$$

where the individually digitised samples S_i are the sum of the contributions of the signal f_i, of the noise b_i and of a common pedestal baseline P; the signal contributions f_i are computed using eq 4.1. The set of optimal weights is determined by minimizing the expected variance of the reconstructed amplitude. The value of the initial offset, which changes from event to event, brings to different values for f_i; for the analysis of the test beam data, the 25 ns offset range has been divided into 25 bins and a different set of weights has been computed for each of these bins.

4.4 Intercalibration studies

As it was discussed in the previous chapters the error on the intercalibration of the different ECAL channels directly enters the constant term of the energy resolution. For the physics ECAL plans to do, the accuracy on the intercalibration has to be kept at a level better than 0.5%. It was discussed how the ECAL calibration strategy foresees some of the crystals being intercalibrated at test beam before the beginning of the LHC data taking. The coefficients computed at the test beam will provide part of the initial intercalibration at the beginning of the data taking and they will be used to check the *in-situ* intercalibration.
An intercalibration procedure has been developed and first introduced in the 2000 year test beam [65]. To check its robustness, several scans of the whole matrix have been performed with high statistics ($\sim$ 30000 events/crystal) during the 2003 test.
In the following the intercalibration procedure will be described, together with the results of detailed checks aimed to study its reliability.

4.4.1 The intercalibration procedure

The reconstructed amplitude in a single crystal varies as a function of the electron impact position (x,y) on the crystal face because of the variation of the electromagnetic shower containment. Figure 4.6 shows the dependence of the crystal response on the electron impact position as it is measured by the set of hodoscopes. The two coordinates X and Y are shown separately, for a crystal taken as example hit by a 120 GeV beam.
In order to intercalibrate the channels, the difference in response due to the different impact points needs to be corrected for. First, the average dependence of the response on the particle impact position is experimentally determined. The distributions in figure 4.6 suggest that such dependence can be approximately factorized in the two coordinates, h(x,y) = f_X(x) $\cdot$ g_Y(y). Both f_X(x) and g_Y(y) can

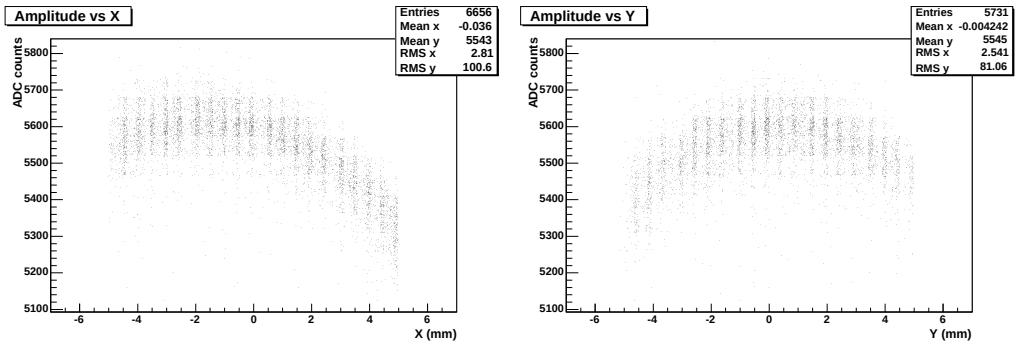

Figure 4.6: *Raw crystal response as a function of the X (left) and Y (right) coordinate of the electron impact point on the crystal surface. An average on a narrow central strip along Y (X) was done.*

be parameterized using a 4-th order polynomial which is determined by fitting the crystal response as a function of the X and the Y coordinates separately (fig.4.7). The small differences between the two polynomials in the two coordinates arise from the differences in the crystals size in the two directions. To better determine

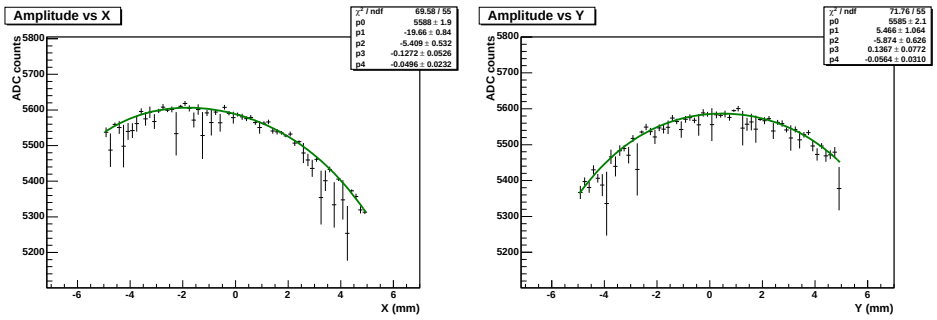

Figure 4.7: *Raw crystal response as a function of the X (left) and Y (right) coordinate of the electron impact point on the crystal surface. A 4-th order polynomial fit is superimposed. An average on a narrow central strip along Y (X) was done.*

the polynomial for the correction in the X (Y) coordinate, only the events falling in a narrow region ($\pm$ 2 mm) around the nominal centre of the crystal along the Y (X) coordinate are selected. The corrected response S^i_{corr} in the i-th event is then obtained, event by event, from the measured amplitude S^i_{meas} using

$$S^i_{corr} = S^i_{meas} \frac{P^{max}_x \cdot P^{max}_y}{P_x(X_i) \cdot P_y(Y_i)} \tag{4.3}$$

Here X_i and Y_i are the coordinates of the electron impact point on the crystal surface and P_j^{max} is the maximum value which is assumed by the polynomial.
The correction works fine in removing the dependence of the crystal response on the impact position, as it is shown in figure 4.8. Since in the central region of the

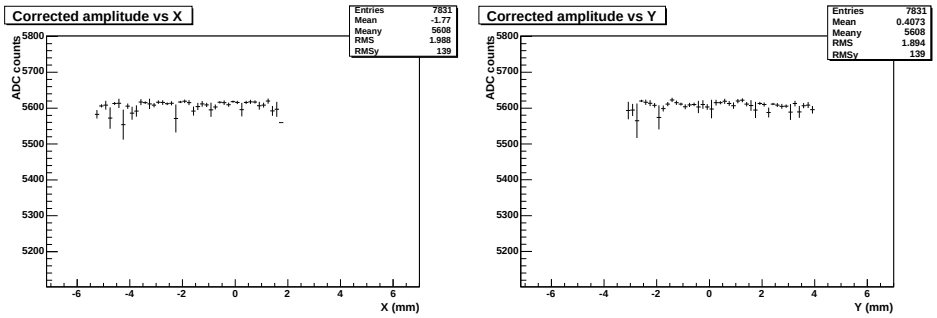

Figure 4.8: *Crystal response as a function of the X (left) and Y (right) coordinate after the correction for the electron impact point on the crystal surface has been applied.*

crystal surface the correction is less affected by fluctuations and it has a smaller size, only the events falling in a 7×7 mm^2 window around the centre have been selected for the analysis described here. This cut selects about 25% of the events.
Two alternative ways to determine the polynomials for the fit have been tested. A different polynomial can be fit on each crystal. It can be noticed anyway that the dependence of the response on the electron impact position has a similar shape for all the crystals and the only differences among the channels are in the amplitude of the response and in the coordinate of the point of maximum response. The polynomial parameters can therefore be obtained fitting two reference polynomials P_X^{ref} and P_Y^{ref} on a reference crystal. The responses of the other crystals are then fitted using the same shapes with a shift x_0 in the coordinate and an overall multiplicative factor A_i as free parameters

$$\begin{aligned} P_x(x) &= A_x \cdot P_X^{ref}(x - x_0) \\ P_y(y) &= A_y \cdot P_Y^{ref}(y - y_0) \end{aligned} \tag{4.4}$$

The two methods are equivalent, as will be discussed in the following. The choice of a reference polynomial anyway makes the whole strategy more stable, expecially when working with low statistics. Local fluctuations can indeed arise and distort the polynomial shape, making a fit with many free parameters difficult.
Figure 4.9 shows the distribution of the corrected energy for an example crystal hit by electrons at 120 GeV. Since a tail on the low energy side is present, the

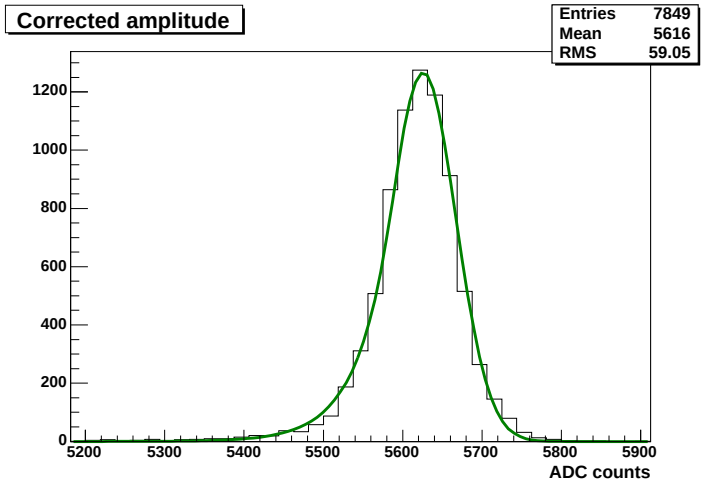

Figure 4.9: *Distribution of the crystal response to a 120 GeV electron beam after the correction for the impact position. A Gaussian fit with an exponential left tail is superimposed.*

distribution can be well fitted using a Gaussian plus an exponential left tail (for the region up to 1σ from the peak position). The position of the peak of the fitted function is chosen as an estimate of the channel response.
The intercalibration coefficient C_i for the $i-th$ channel is then defined as the ratio between its response M_i and the response of a reference crystal M_{ref},

$$C_i = \frac{M_i}{M_{ref}} \tag{4.5}$$

4.4.2 Accuracy of the intercalibration procedure

The stability of the intercalibration procedure and the reproducibility of the results can be affected by many factors, mainly related to the calibration conditions.
In this section some of the factors possibly influencing it are investigated in detail. Unfortunately, due to the strict time constraints, only one scan was performed on SM1 and no check of the reproducibility was possible. All the tests presented in the following therefore refer to the SM0 supermodule.

Influence of the amplitude reconstruction method

It was discussed in section 4.3 that the signal amplitude is reconstructed starting from the 14 samples of the pulse shape which are digitised by the ADC and that for the 2003 test beam data analysis two different techniques were available. The

choice of the *analytic fit method* or of the *weights method* to reconstruct the pulse shape could in principle affect the intercalibration coefficients. Figure 4.10 shows the difference between two sets of coefficients computed using the two reconstruction methods. The agreement is at the 0.05% level, meaning that the impact of the reconstruction method is almost negligible on the intercalibration. All the results presented in the following have been obtained using the weights method.

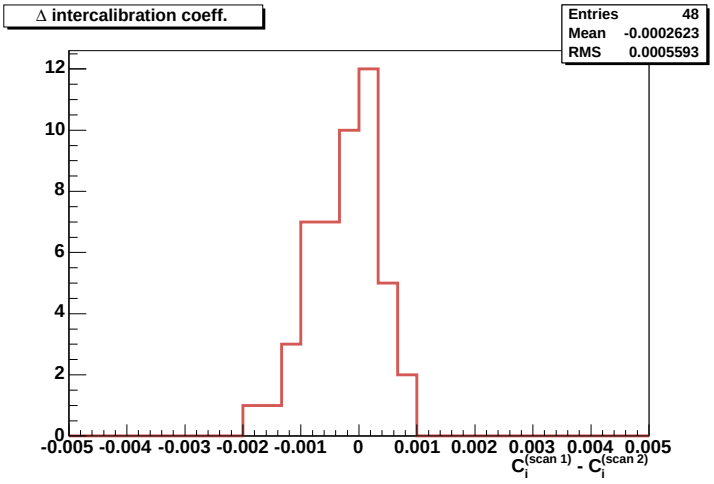

Figure 4.10: *Distribution of the differences among the two sets of intercalibration coefficients obtained with the method of the analytic fit and the weights method as pulse shape reconstruction algorithm. The results refer to the data collected on SM0 crystals with a 120 GeV electron beam. The agreement is at the 0.05% level.*

Influence of the electronics operation mode

During the first part of the test beam, the SM0 supermodule was equipped with a front-end electronics based on the FPPA preamplifier. The electronics gain could be fixed to the same value for all the samples or could be let free to switch among the four available values ($\times 1$, $\times 5$, $\times 9$, $\times 33$) according to the sample amplitude (the so called *autogain* mode). The gain mode could in principle affect the amplitude reconstruction and by consequence the intercalibration coefficients.
Two consecutive scans have been performed with the amplifier working first in the autogain and then in the fixed gain ($\times 9$) modality. Figure 4.11 shows the comparison among the intercalibration coefficients computed in the two scans. The agreement is at the 0.3% level.

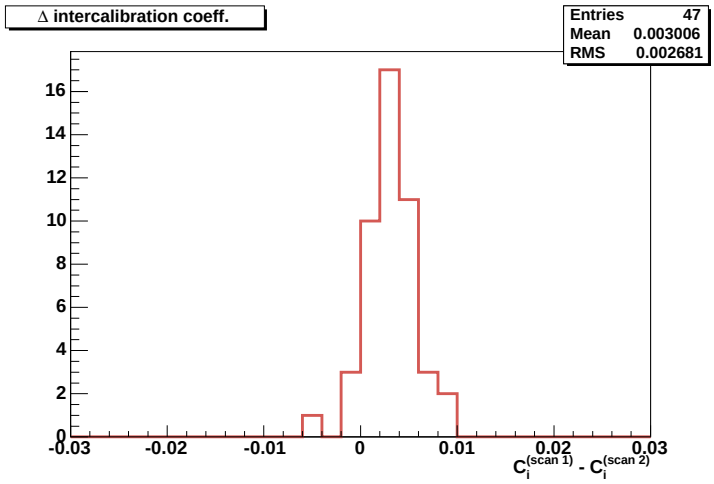

Figure 4.11: *Distribution of the differences among the two sets of intercalibration coefficients obtained with the FPPA electronics working in the fixed gain and in the autogain mode. The results refer to the data collected on SM0 crystals with a 120 GeV electron beam. The agreement is at the 0.3% level.*

Influence of the beam energy

The possible dependence of the intercalibration on the electron beam energy has been studied. A comparison between the coefficients computed using 120 GeV and 50 GeV electrons has been performed and the result is shown in figure 4.12. The agreement is at the 0.3% level.

Influence of the noise

The analysis of the 2003 test beam data has shown the presence of a low frequency noise, giving origin to oscillations of the pedestal line between the different events. Two methods have been developed to remove it. The first one [66] exploits the noise correlation between the different crystals; it uses the informations coming from crystals far away from the one hit by the beam to evaluate the oscillations of the pedestal line. The second method [67] exploits the information of the two samples before the start of the pulse for a dynamical subtraction of the pedestal.
To check the effect of the noise on the intercalibration, the channel responses with and without the pedestal subtraction have been compared. The result is presented in figure 4.13 for crystals belonging to the SM0 and SM1 supermodules. The scan on the SM0 crystals was done with 120 GeV electrons, which correspond to $\sim$ 5000 ADC counts; the scan on SM1 was instead done with a 50 GeV beam, corresponding to 1200 ADC counts on average.

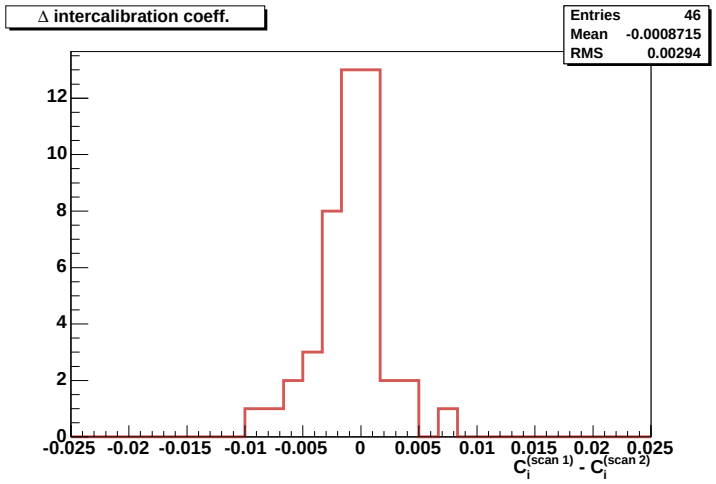

Figure 4.12: *Distribution of the differences among the two sets of intercalibration coefficients obtained with electron beams of two different energies, 120 GeV and 50 GeV. The results refer to the data collected on SM0 crystals. The agreement is at the 0.3% level.*

The noise affects the intercalibration coefficients less than 0.01%. This is reaso-

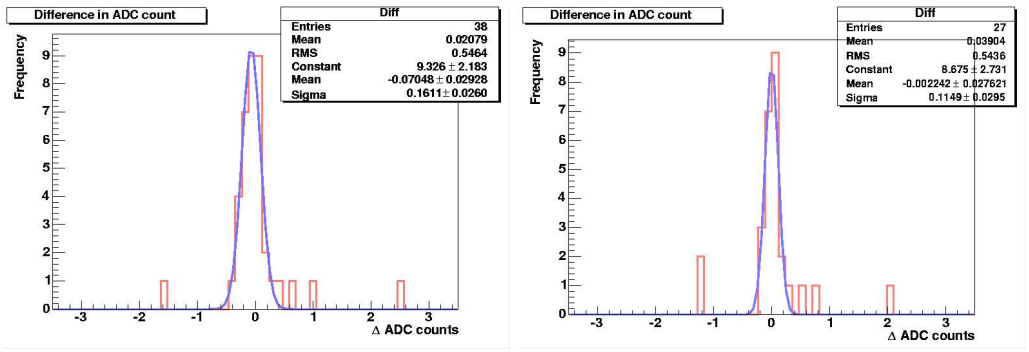

Figure 4.13: *Distribution of the differences among the crystal responses with and without the noise subtraction. Left (right), results obtained on the SM0 (SM1) crystals with 120 (50) GeV electrons.*

nable, since the coefficients are computed using the peak of the corrected energy distribution, which is built with high statistics. Since the mean value of the noise is zero, the noise contribution to the peak position is RMS $_{noise}/\sqrt{N}$, where N is the number of events which are employed. The contribution of the noise was checked to remain negligible even when working with a reduced statistics.

Influence of the polynomial determination

All the factors which have been investigated up to now are related to the calibration conditions. In this section, the effect of the polynomial determination on the precision of the intercalibration is discussed, as a possible source of inaccuracy related to the procedure itself.
As it was already mentioned, the correction of the crystal response for the impact position of the electron on the crystal surface can be done using the same reference polynomial for all the crystals or determining a new polynomial with a free fit for each channel. The first strategy works better when operating with small statistics, due to the fluctuations which can possibly occur. Anyway, when working with high statistics, the coefficients computed with the two methods are in agreement at the level of 0.1%, as it is shown in figure 4.14. That also suggests that the choice of the reference crystal when working with a reference polynomial does not affect the intercalibration in a significative way.

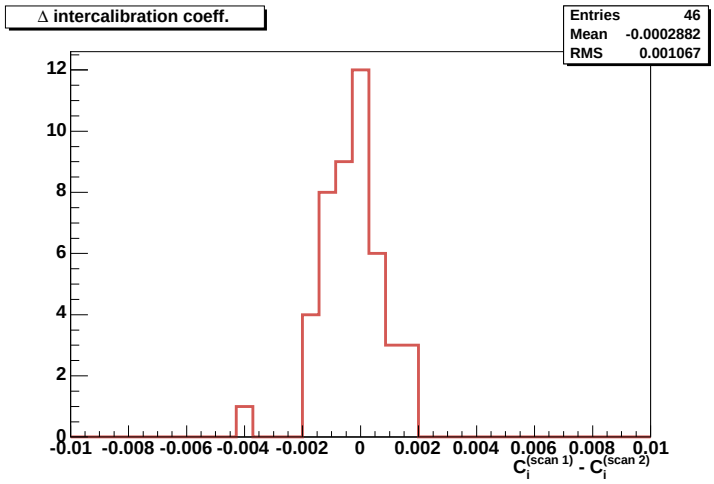

Figure 4.14: *Distribution of the differences among the two sets of intercalibration coefficients computed using different procedures to correct for the electron impact point on the crystal surface. In the first case the same reference polynomial has been used for all the crystals, with a shift and an overall moltiplicative factor as free parameters for the fit. In the second case a free fit on each crystal has been performed. The results refer to the data collected on SM0 crystals with a 120 GeV electron beam.*

The dependence of the crystal response on the particle impact point could be in principle affected by the shape of the crystal itself. All the crystals of the SM0

supermodule under analysis at the 2003 test beam had the same shape, being of type 15 and 16. To check a possible dependence, the polynomials determined during the 2002 test beam on type 8 and 9 crystals have been used for the fit. These crystals, which belong to the different pseudorapidity range around η = 0.7, had indeed a different shape with respect to the ones under analysis in 2003. In figure 4.15 the reconstructed energy versus the impact point of the electron is shown. The fit is good and the precision seems not to be affected by the choice of a reference polynomial in a different pseudorapidity region.
The coefficients obtained using the polynomials from the 2002 test beam and those computed for the analysis presented here have been compared. The agreement is at the 0.1% level. This same result was obtained using different reference polynomials in the same η range [68].

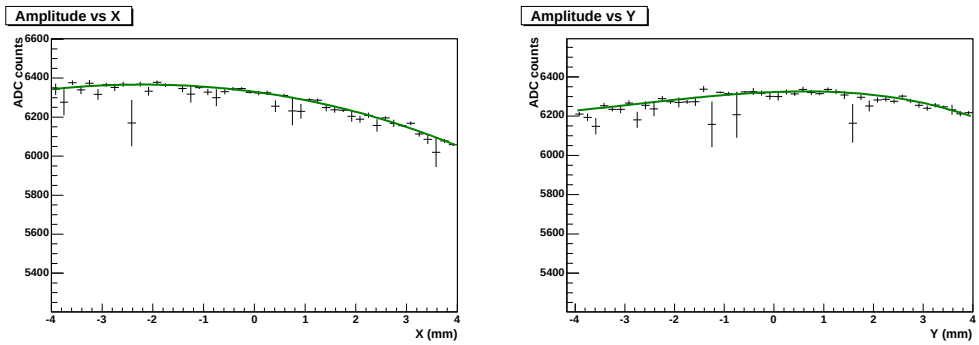

Figure 4.15: *Crystal response versus the electron impact point before any correction. On the left (right), the X (Y) coordinate is shown. A reference polynomial determined during the 2002 test beam has been used for the fit, leaving as free parameters only the global normalization and a shift in the coordinate.*

Even if it is not directly related to the intercalibration procedure, an estimate of the accuracy in the determination of the point of maximum response can be interesting. Due to the tilt between the crystal axis and the beam direction, the point of maximum response differs from the geometrical centre of the crystal and it depends on the shower dept and so on the beam energy. The comparisons of the results obtained from different datasets both at the same 120 GeV energy and at two different energies (120 GeV and 50 GeV) are shown in figure 4.16. A free fit on each crystal has been done to determine the polynomial. In both cases the RMS is about 200 μm, slightly larger in X than in Y; the reproducibility is slightly better when working with a fixed shape polynomial. A systematic shift in the point of maximum response is visible when comparing the results at different

energies, and it is compatible with what is expected due to the tilt of the crystal with respect to a perfect pointing geometry. Again a difference arise in the two coordinates.

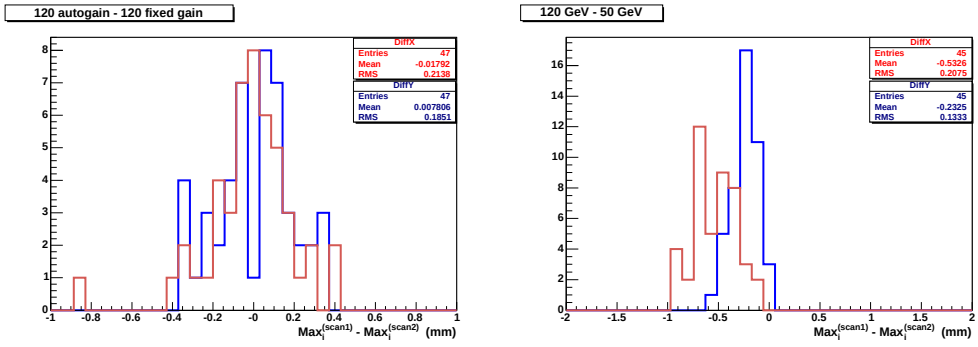

Figure 4.16: *Distribution of the differences among the positions of the points of maximum response from two different scans at the same 120 GeV energy (left) and at 120 and 50 GeV (right). A different polynomial was used to fit each crystal.*

The reproducibility of the results

For the SM0 supermodule, all the scans have been performed in different conditions. All the datasets which have been collected differ for at least one of the calibration parameters, so a check of the reproducibility of the intercalibration coefficients in exactly the same conditions is not possible. Such study has been done during the 2002 test beam on the crystals of the M0$'$ supermodule. The results obtained in two scans taken in exactly the same conditions were compared and the agreement among the coefficients was found to be at the 0.3% level [68]. The discrepancy can not be attributed to the statistical fluctuations only, as will be shown in the following section. The effect of the non reproducibility could in principle therefore affect also the comparisons performed during the 2003 test beam which have been discussed in the previous sections.

4.4.3 Statistical precision of the intercalibration

The intercalibration studies at the 2003 test beam have been performed with a statistics of $\sim$ 30000 events for each crystal, corresponding to about 15 minutes data taking. To make an extensive intercalibration possible, it is necessary to reduce this time. It is therefore interesting to understand how the intercalibration precision depends on the statistics which is used.

The study of the statistical precision of the intercalibration has been done using a scan of the SM0 matrix with 120 GeV electrons. As it was already discussed, the best way to operate when working with low statistics is to use a fix shape polynomial which is determined using a large number of events; the full chain is then applied on the available statistics. In this way, only few runs with high statistics are necessary to determine the polynomial shape.
In principle, the influence of the statistics can be estimated by comparing the results from several subsamples for one given crystal. However, assuming that all the channels behave the same way, this is equivalent to compare the results from two subruns only employing many crystals. This second procedure has been used in the present analysis. The dependence of the intercalibration precision on the statistics is shown in figure 4.17 and a fit with the function $\sqrt{p_0^2 + (\frac{p_1}{\sqrt{N}})^2}$ is superimposed. The intercalibration precision should vary as $1/\sqrt{N}$, being N the number of events which are used; the constant term p_0 represents the limit on the precision. 1000 events per crystal before the fiducial cuts (which correspond to

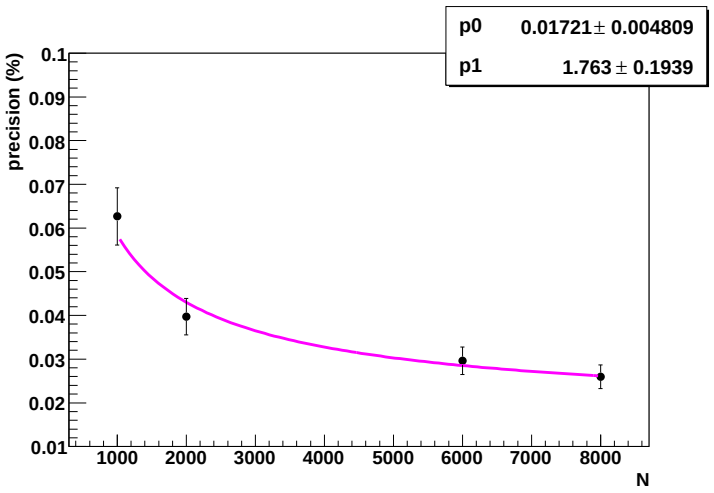

Figure 4.17: *The trend of the statistical precision with respect to the number of events used to intercalibrate the crystals (before the fiducial cuts). The curve represents the fit function* $\sqrt{p_0^2 + (\frac{p_1}{\sqrt{N}})^2}$.

$\sim$ 250 events after the selections) are already enough to keep the statistical accuracy better than 0.1%. The proposed strategy can therefore be employed for an extensive intercalibration on the beam.

4.5 Irradiation studies

It was discussed in the previous chapter how the crystals change their behaviour when esposed to high radiations. Since such an effect is different from crystal to crystal, the intercalibration constants are modified and this can spoil the energy resolution. A correction for the effect is therefore necessary and a laser monitoring system has been developed in CMS to check the crystal response evolution. During the 2003 test beam, some of the SM0 crystals have been irradiated with electrons and pions to test the monitoring laser system and to check the possibility of correcting for the ageing due to the radiation damage. The analysis of the data collected is presented in the following.

4.5.1 The test beam irradiations

During the 2003 test beam, some of the SM0 crystals underwent an irradiation with 120 GeV electrons and 120 GeV pions. The irradiated crystals, the corresponding dose rates and the integrated doses are reported in table 4.1. Figure 4.18 shows the average doses versus the time for the irradiated crystals.

Irradiation with electrons			
Crystal	Dose rate at the shower max (Gy/h)	Mean dose rate (Gy/h)	Integrated mean dose (Gy)
1522	0.366	0.169	1.20
1097	0.346	0.160	1.06
1607	0.211	0.098	0.45
1602	0.315	0.146	0.46
1182	0.402	0.186	1.06
Irradiation with pions			
Crystal	Dose rate at the shower max (Gy/h)	Mean dose rate (Gy/h)	Integrated mean dose (Gy)
1522	-	0.133	1.82
1097	-	0.122	1.79
1607	-	0.08	0.98
1353	-	0.087	1.10
1182	-	0.117	1.37

Table 4.1: *Dose rates and integrated doses for the various irradiated crystals.*

At the test beam, counters were put along the beam line to register the number N of particles impinging on the crystal under study. The dose absorbed by the

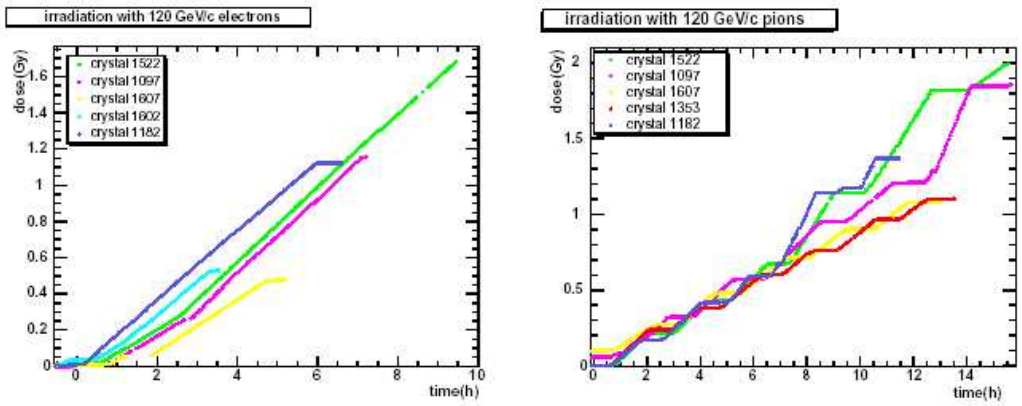

Figure 4.18: *Mean dose versus time for the various crystals under electron (left) and pion (right) irradiation.*

crystal has been computed taking into account the beam transverse dimensions and the shower profile. For an electron beam of 120 GeV with a gaussian profile of 1.5 cm RMS both in the x and y coordinate, the dose at the shower maximum can be computed using the relation

$$Dose_{max}(e, 120\,\text{GeV/c}) = 6.7 \times 10^{-8} \times N \quad [Gy] \tag{4.6}$$

with an extimated accuracy between 10% and 20% [69].
Since the exact value of the absorbed dose is not relevant for the analysis presented here, the average doses absorbed by the crystals have been used. Taking into account the average energy which is deposited in a crystal by a 120 GeV electron or a 120 GeV pion, the mean doses can be computed as

$$Dose(e, 120\,\text{GeV/c}) = 3.1 \times 10^{-8} \times N \quad [Gy] \tag{4.7}$$

$$Dose(\pi, 120\,\text{GeV/c}) = 3.0 \times 10^{-9} \times N \quad [Gy] \tag{4.8}$$

According to these relations, the irradiation rates ranged from about 0.10 Gy/h to 0.19 Gy/h for the irradiation with electrons and from about 0.08 Gy/h to 0.13 Gy/h for pions. These values are at the level of the rates which are expected during the high luminosity phase in the ECAL barrel, which are between 0.15 and 0.30 Gy/h. More details about the expected values at the LHC are given in figure 4.19, where the doses expected in the whole ECAL pseudorapidity range are reported.
During the irradiation period, the electrons runs were interleaved with laser runs for monitoring purposes. When irradiating with pions, the evolution of the response of the crystals to electromagnetic showers was followed through periodic exposures of the crystals to the 120 GeV electron beam (two electron runs and two laser runs were performed after each irradiation run).

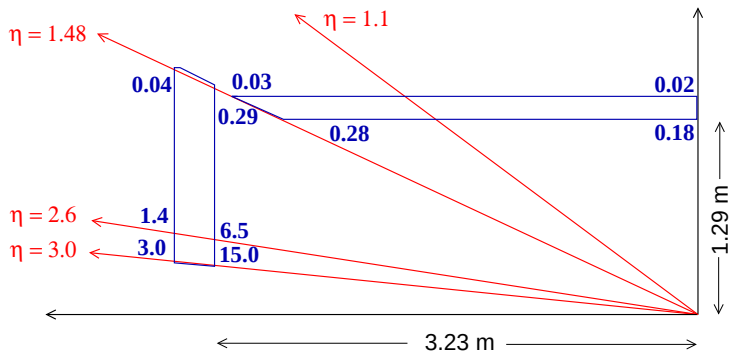

Figure 4.19: *Scheme of the dose rate (in Gy/h) expected in the ECAL barrel and in the endcaps for the LHC luminosity of* 10^{34} $cm^{-2}s^{-1}$. *The values are given for the maximum of the shower energy deposition.*

4.5.2 The laser stability and the data treatment

The response to the particle beam and to the laser light has been reconstructed for the analysis described in the following with two different methods.
The electron energy was reconstructed using the same correction for the impact position described in section 4.4.1 for the intercalibration studies. The electron runs have been divided in a sequence of subruns, each one corresponding to $\sim$ 6500 events, to have a finer sampling of the response loss evolution. This turned out to be quite useful expecially at the beginning of the irradiation for each crystal, when the slope of the response versus the dose is steeper. The precision on the peak position was order of 0.1% on each subrun.
No division in subruns was applied for the study of the laser runs, which had about 1500 events for each colour. Four laser wavelengths were available at the test beam but only the blue and the infrared lasers have been used for the whole irradiation period.
The peak of the pulse amplitude distribution could not be used as an estimate of the response to the laser since it presented both fluctuations from event to event in a single run and from run to run. The pulse amplitude distribution for a typical run is shown in figure 4.20 for the four wavelengths. As a consequence of the amplitude fluctuations from event to event, the spread of the distributions is quite large (some percents) and the determination of the peak position is not precise. Moreover, the laser amplitude presents sometimes sudden important changes during a single run (fig 4.22) that make the raw response to the laser not meaningful. The fluctuations of the laser amplitude from run to run instead can mimick the variation in the crystal response due to the radiation damage and so make the cor-

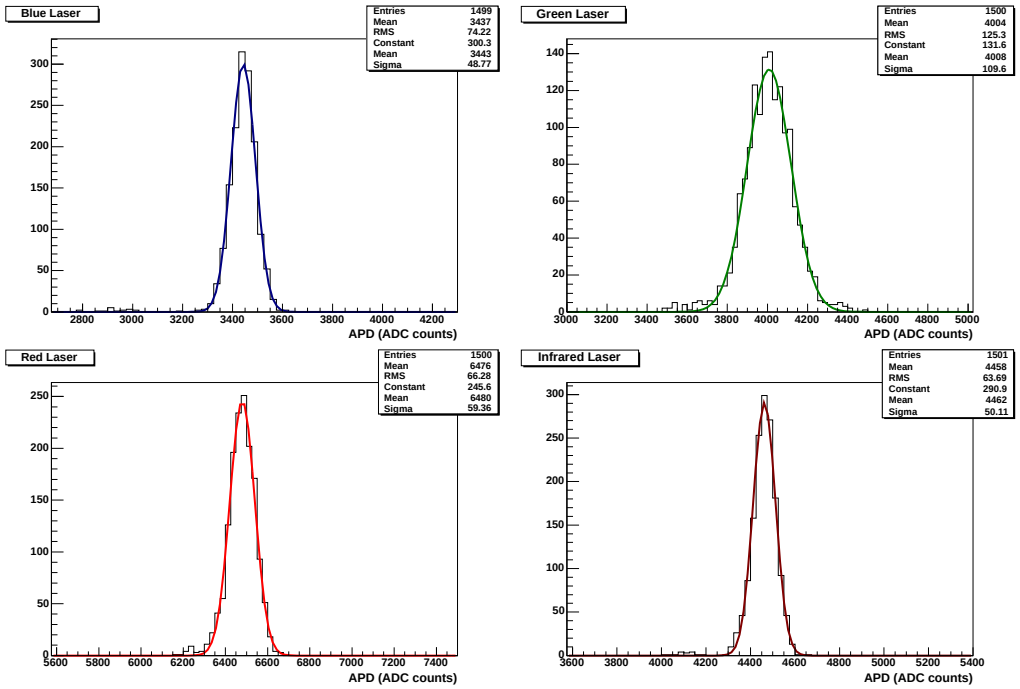

Figure 4.20: *Distributions of the response to the laser light as read by the APDs for the four available wavelengths. The data refer to the crystal 1097 of SM0 before the irradiation.*

rection wrong.
Reference PN diodes have been employed to monitor in an independent way both kinds of fluctuations. The distribution of the ratio of the laser signal measured by the APD normalized, event by event, to the PN reponse is shown in figure 4.21. The RMS of the distribution on the single run scale is about 0.5% and the correction works efficiently also in case of problematic runs like the one which is shown in figure 4.22.
In the following we call laser response of a crystal the APD/PN ratio.

In the long term ($\sim$ 1 month), sizeable drifts up to 15% of the laser amplitude were observed and small non linearities between the APD and the PN could alter the stability of the APD/PN ratio. The APD/PN ratio is shown in figure 4.23 versus the PN absolute value for a large number of laser runs before the irradiation. A flat distribution would be expected in case of full linearity, instead a correlation is clearly visible. We have corrected for this effect by using the quantity $\left.\frac{APD}{PN}\right|_{corr}$

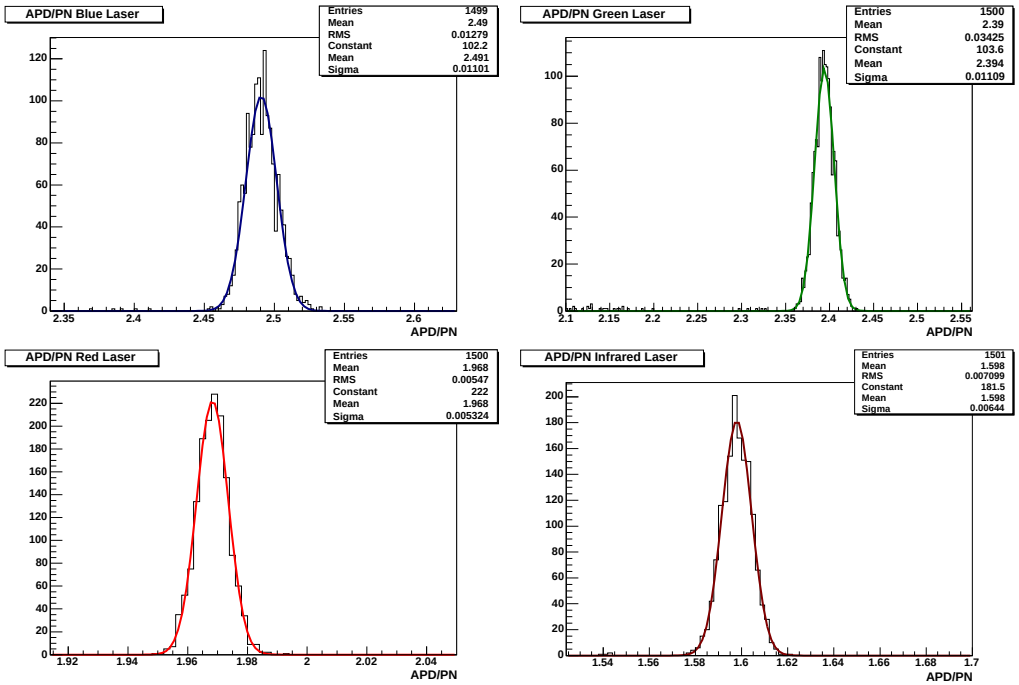

Figure 4.21: *Distribution of the ratio APD/PN in a laser run for the four available wavelengths before the correction for the non linearity. The data refer to the crystal 1097 of SM0 before the irradiation.*

defined as

$$\left.\frac{APD}{PN}\right|_{corr} = \left.\frac{APD}{PN}\right|_{raw} - k \cdot (PN - \overline{PN}) \tag{4.9}$$

$\overline{PN}$ is an arbitrary reference value, which is the same for all the crystals and the runs and we have used the middle point in the range of variation of the PN values. The coefficient k is determined fitting distributions like those shown in figure 4.23 for each crystal before the irradiation. For the green laser, which is the less stable, there is not a clear dependence, therefore an higher degree polynomial has been used for the fit. In principle k could vary during the irradiation since the crystal response to the laser changes and the PN does not; this effect is anyway negligible ($\sim 10^{-5}$). The correction for the non linearity is at the level of 10^{-3} or lower.
After the correction for the non linearity, the stability of the monitoring system on the time scale of a typical irradiation study (about 10 h) was found to be at the level of 0.03% - 0.04% (fig. 4.24), increasing up to 0.1% over a period of about 2 days. This spread is larger than what expected from the dispersion on a single run (which is 0.5% / $\sqrt{N}$), meaning that some fluctuations arise on a time scale

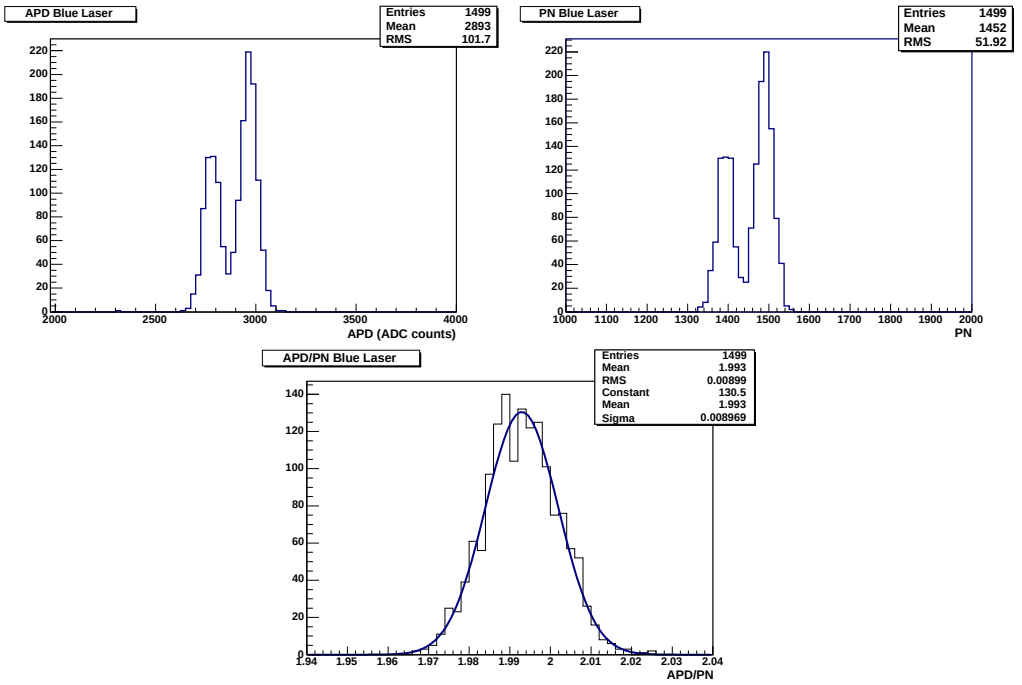

Figure 4.22: *Distribution of the response to the blue laser light as read by the APD and by the PN diode (on the top) in the case of a problematic run; on the bottom the distribution of the APD/PN ratio is shown. The fluctuations are canceled.*

larger than the one of the single run. A similar stability was observed, also on longer time intervals, at the 2002 test beam [70] and it is well within the CMS requirements.

4.5.3 The determination of the α parameter

As it was discussed in section 3.3.2, the relation between the response to the electromagnetic showers $S = S(t)$ and the response to the laser light $R = R(t)$ can be written as

$$\left(\frac{S}{S_0}\right) = \left(\frac{R}{R_0}\right)^{\alpha} \tag{4.10}$$

where S_0 and R_0 are the responses to the beam and to the laser before the irradiation. The α coefficient includes the contribution of the absorption cross section and of the path followed by the light. Since the scintillation light and the laser light follow different paths within the crystal, some logarithmic terms should be added as already discussed. Equation 4.10 is anyway effective if the relative loss

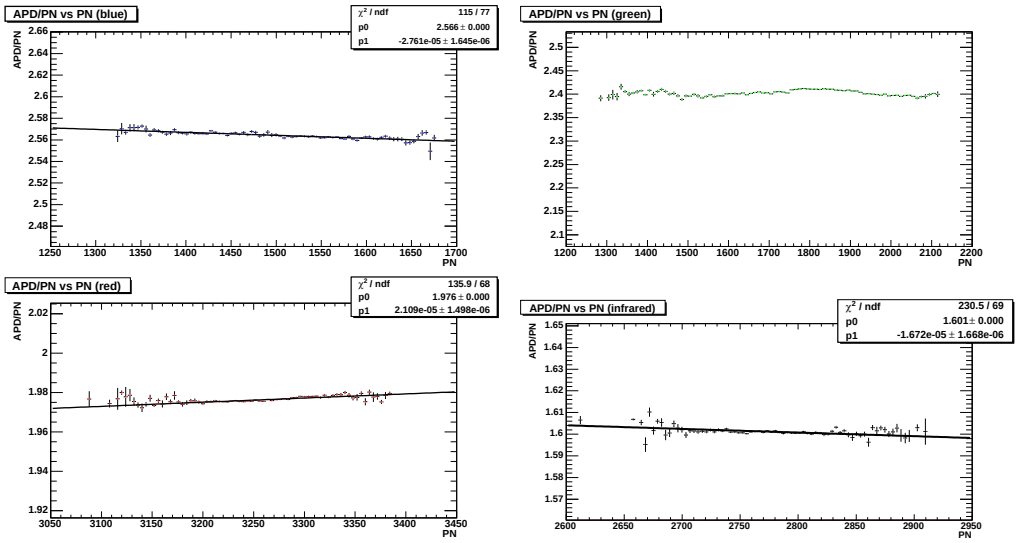

Figure 4.23: *APD/PN versus PN distributions. The statistics corresponds to many laser runs, for the four available wavelengths. Superimposed is the result of the linear fit. The data refer to the crystal number 1097 of SM0 before the irradiation.*

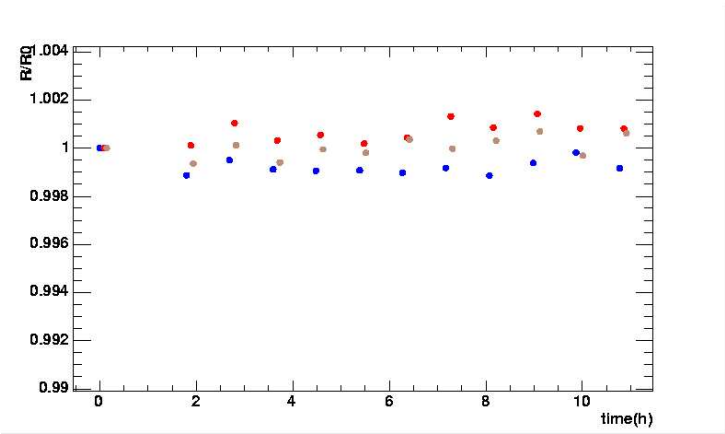

Figure 4.24: *APD/PN ratio versus time after the correction for the non linearity of equation 4.9; the time interval is similar to that of an irradiation study. The brown spots refer to the infrared laser.*

$\Delta S = S - S_0$ is kept at the level of some percents.
In order to determine the α coefficient, equation 4.10 can be written as

$$lnS = \alpha lnR + K \tag{4.11}$$

with

$$K = lnS_0 - \alpha lnR_0 \tag{4.12}$$

In this way, the initial values lnS_0 and lnR_0 do not affect the determination of α and they modify only the constant K. For each irradiated crystal the parameter α can be determined with a linear fit of lnS versus lnR.
Since the laser and the electron beam runs could not be taken at the same time, the response to the laser R at the same dose as the electron data point S has been determined interpolating the response measured in the two laser runs immediately before and after the electron beam run. A linear interpolation based on the absorbed dose has been used. This is an approximation since the change in the response versus the dose is not linear; anyway, such approximation is accurate enough when the time interval between the two laser runs is short.
The uncertainties on the crystal response have been evaluated *a posteriori*. All the laser points have been considered without error and the same experimental uncertainty σ_S was assumed for the response of each crystal to the beam, so being σ_{lnS} = σ_S/S. First, an arbitrary value $\tilde{\sigma_S}$ was given to σ_S. The error was then scaled until the sum $\tilde{\chi^2}$ of the χ^2 of the whole data sample was equal to the total number of degrees of freedom, using the relations

$$\beta = \sqrt{\frac{\tilde{\chi^2}}{ndf}} \tag{4.13}$$

$$\sigma_S = \beta \cdot \tilde{\sigma_S} \tag{4.14}$$

The rescaling was done separately for each wavelength. With this procedure an estimate uncertainty on each data point of about 0.12% has been found for the study with the blue laser.

Irradiation with 120 GeV electrons

Figure 4.25 shows the evolution of the response to the electron beam and to the laser light as a function of the mean dose absorbed by the crystals during the irradiation with electrons. The losses for the different crystals, here normalized to the response before the irradiation, are summarized in table 4.2. For the monitoring with an electron beam, the losses range from about 1.5% to 5%. For the laser, the losses are lower and they range from about 1% to 3% for the blue light and from 0.2% to 0.5% for the infrared laser, being the crystal transparency almost inaffected by the color center formation process in the infrared region.
The results of the analysis, after the correction for the non linearity between the APD and the PN, are reported in table 4.3. Since the cross sections for the absorption by the color centers change from one wavelength to another, the α parameters for the four lasers are different. An example of the linear correlation

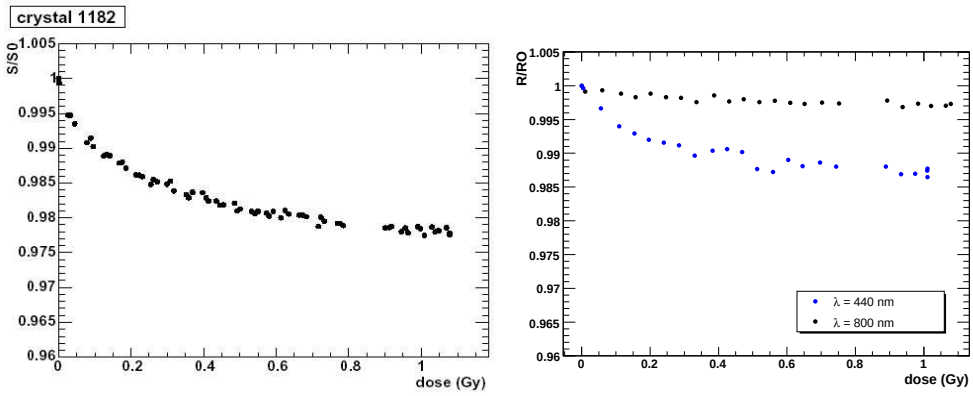

Figure 4.25: *Response to electrons (left) and to the laser (right) versus the received dose, both normalized to the first point. The responses to the blue and the infrared laser are shown. The data refer to the crystal number 1182 of SM0.*

Irradiation with electrons			
Crystal	Δ_{beam}	Δ_{laser} (λ=440 nm)	Δ_{laser} (λ=800 nm)
1522	1.9%	1.2%	0.17%
1097	4.8%	3%	0.5%
1607	1.7%	1.2%	0.17%
1602	3.3%	2.6%	0.37%
1182	2.2%	1.3%	0.27%

Table 4.2: *Relative losses in the response to the electrons and to the laser lights after the irradiation with 120 GeV electrons.*

α coeffi cients, irradiation with electrons				
Crystal	α (λ=440 nm)	α (λ=495 nm)	α (λ=709 nm)	α (λ=800 nm)
1522	1.435±0.047	1.266±0.09	8.547±0.55	13.64±0.82
1097	1.666±0.012	1.591±0.04	6.721±0.13	10.51±0.18
1607	1.352±0.050	-	-	7.01±0.48
1602	1.434±0.024	-	-	8.24±0.23
1182	1.609±0.042	-	-	7.55±0.34
average	1.50±0.06	1.43±0.16	7.63±0.91	9.39±1.22

Table 4.3: α *coefficients computed for the irradiations with electrons.*

between the relative response loss to the electrons $\Delta lnS = \Delta S/S_0$ and to the laser $\Delta lnR = \Delta R/R_0$ is shown in figure 4.26 for the four wavelengths; the linear fit is superimposed.

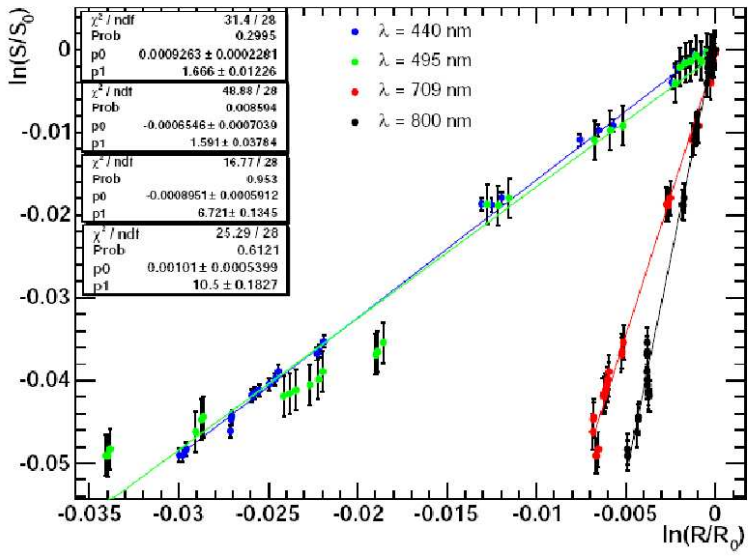

Figure 4.26: *Response to the electrons (S) versus response to the laser (R) for the four wavelengths. The logarithms are plotted. Each point in the figure corresponds to a beam run (or subrun) and the laser response has been interpolated. The linear fit is superimposed. The data refer to crystal 1097*

The loss in the response during the irradiation can differ from one crystal to the other due to the different radiation hardness. The α parameter is essentially a geometrical factor which depends on the crystal geometry, the optical emittance and the position of the monitoring fiber. Since all these factors are well reproducible or do not vary much, α is expected to be universal but small variations from crystal to crystal can however arise. The S versus R relation has to be well known to apply a correction procedure. The universality of the relation was tested during the 2002 test beam [70] on 20 crystals all belonging to the same production batch, that were found to have an intrinsic dispersion of the α parameters around 5% (only the monitoring with blue light was done). The data reported in table 4.3, which refer to crystals from different production batches, show a dispersion around 9% in the case of the blue laser, which is lowered to about 8% when the errors on the α parameters are deconvoluted. The mean value determined from this study, α = 1.50 $\pm$ 0.06, is compatible with the value α = 1.53$\pm$0.09 found

during the 2002 test beam. The results obtained at the test beam therefore showed no experimental evidence of problems in assuming an universal α value. In any case, first studies [71] have shown that α can be fitted with data recorded *in-situ* even at low luminosity and therefore the universality of α is not strictly necessary.

Irradiation with 120 GeV pions

At LHC all the crystals will undergo both electromagnetic and hadron irradiation and it is therefore necessary to check the effectiveness of the monitoring system for the different radiation profiles. The same procedure described for the irradiation with electrons has been applied to study the effect of the irradiation with pions. Figure 4.27 shows the relative loss in the response monitored with an electron beam and with the laser versus the absorbed dose for a reference crystal. The losses for the whole sample of irradiated crystals are given in table 4.4. They range from 2% to 7% and they are comparable with those observed during the irradiation with electrons for comparable doses.
The results of the analysis are reported in table 4.5 and they are shown in figure 4.28 for a reference crystal.

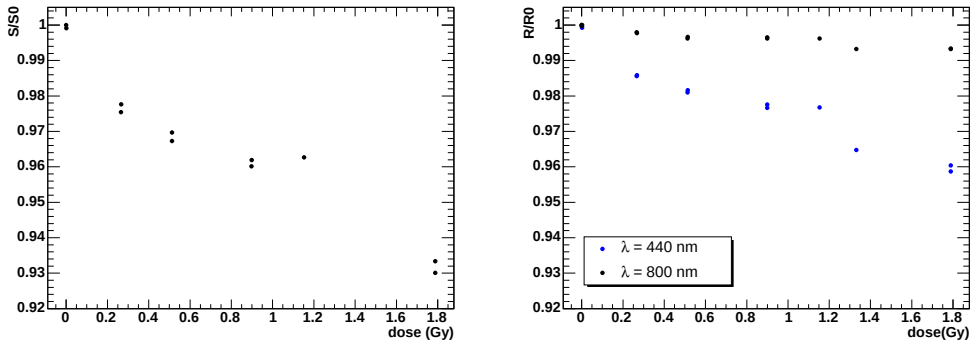

Figure 4.27: *Response to electrons (left) and to the laser (right) versus the received dose from pions, both normalized to the first point. The two spots at each dose correspond to the two laser or beam runs done between the two irradiations with pions. The responses to the blue and the infrared laser are shown. The data refer to the crystal number 1097 of SM0.*

A good agreement has been found among the values of the α coefficients determined under pion and electron irradiation with the blue laser source, as it is shown in figure 4.29. The radiation profiles of electrons and pions are really different, since an electron beam at 120 GeV has its maximum at about $8X_0$ while a pion beam deposits its energy mainly towards the end of the crystal. The consistency of

irradiation with pions			
Crystal	Δ_{beam}	Δ_{laser} (λ=440 nm)	Δ_{laser} (λ=800 nm)
1097	6.7%	3.8%	0.7%
1607	1.8%	1.2%	0.1%
1353	2.2%	1.4%	0.4%
1182	3.2%	1.9%	0.3%

Table 4.4: *Relative losses in the response to the electrons and to the lasers after the irradiation with 120 GeV pions.*

α *coeffi cients, irradiation with pions*		
Crystal	α (λ=440 nm)	α (λ=800 nm)
1097	1.772±0.024	10.47±0.57
1607	1.299±0.059	13.87±2.90
1353	1.585±0.051	5.355±0.79
1182	1.637±0.033	10.96±0.95
average	1.57±0.10	10.2±1.77

Table 4.5: α *coefficients computed for the irradiations with pions.*

the α parameters therefore suggests that the response loss is almost independent of the profile of the received dose and it is rather related to the integrated effect of the number of color centers which are created. The statistics of the present analysis is too low to allow definitive conclusions and the results need to be confirmed with more data. At this stage, anyway, the results obtained with the blue laser on electrons and pions show an agreement within few percents.

4.5.4 The monitoring with the infrared laser

Some more considerations are necessary about the analysis performed with the infrared laser, which turned out to be much less accurate notwithstanding the very stable functioning of the infrared laser during all the test period.
According the model introduced in chapter 3, the crystal response to the light decreases exponentially and the time constant depends on the absorption coefficient. The α parameter for a given wavelength can be therefore compared with the absorption coefficient μ at the same wavelength. It can be easily demonstrated that the relation among them is

$$\frac{\alpha(\lambda_1)}{\alpha(\lambda_2)} = \frac{\mu(\lambda_2)}{\mu(\lambda_1)}. \tag{4.15}$$

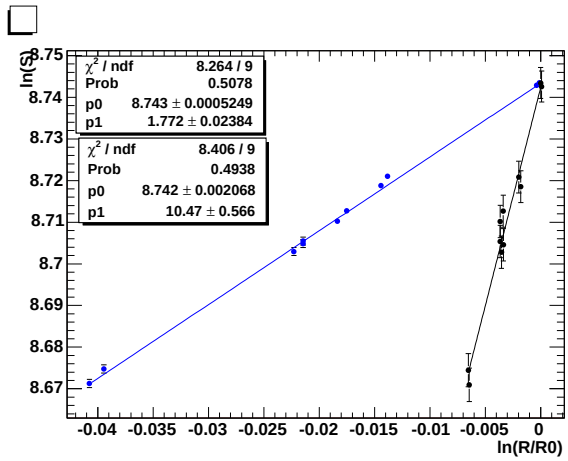

Figure 4.28: *Response to the electrons (S) versus response to the laser (R) for the two available wavelengths during the irradiation with pions. The logarithms are plotted. Each point in the figure corresponds to a beam run and the laser response has been interpolated. The linear fit is superimposed. The data refer to crystal 1097 of SM0.*

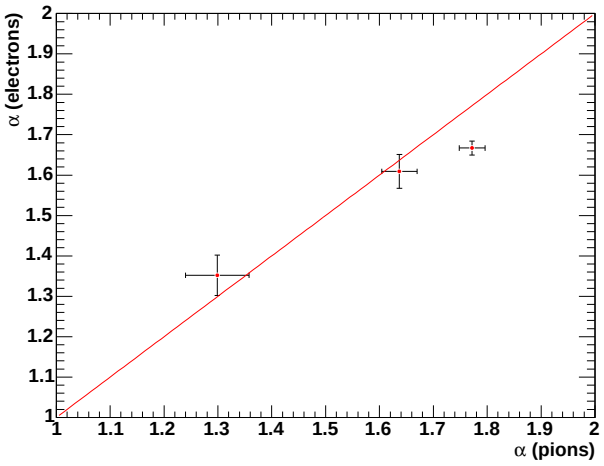

Figure 4.29: *Correlation between the α coefficients obtained with data refering to the irradiation with electrons and with pions with the blue laser.*

To check the validity of this relation, the α values computed in the analysis presented here have been used to derive the corresponding absorption coefficients, which are compared with the ones deduced from laboratory measurements [72] in

figure 4.30. Since the light path is not well known, neither for the electron data nor for the laser, the absorption coefficients can not be estimated from the α parameters. Equation 4.15 has been therefore used here only as a relative prediction and the points in figure 4.30 have been computed as

$$\mu(\lambda) = \mu(\lambda_{ref}) \times \frac{\alpha(\lambda_{ref})}{\alpha(\lambda)} \tag{4.16}$$

being λ_{ref} = 440 nm and $\mu(\lambda_{ref})$ the absorption coefficient at 440 nm quoted in [72]. The agreement is good, even if some points which are not understood are present.
According to relation 4.15 and due to the small value of $\mu_{infrared}$ (which is about a factor 7 lower than μ_{blue}), the infrared laser is almost insensitive to the radiation damage. The small changes in the crystal response monitored with the infrared laser are comparable to the fluctuations which are present also if the crystal is not irradiated and the uncertainties are therefore large. In this sense it is interesting to underline that the crystals with the less precise estimation are those with low relative loss.
Another source of uncertainty for the analysis with the infrared laser is the fact that the APD response to the infrared laser is not well defined. Since the attenuation length for light with λ = 800 nm in the silicon is about 30 μm, the infrared light is not completely absorbed by the 5 μm thick p^+ window at the APD entrance and it can partially enter the moltiplication region. As a conseguence, the gain is lower and not stable. The effect of possible fluctuations which can affect the APD gain in the readout chain or in the external conditions (ie. the temperature) is therefore different for the infrared laser and for the blue one [73].
On the basis of all these considerations, the blue light is the best mean to monitor the effect of the radiation damage. A comparison between the response to the blue laser and to the infrared one could in principle be used to get informations about possible sources of instabilities not related with the radiation damage.

4.5.5 The correction procedure

The informations collected from the monitoring system can be used to correct the crystal response for the loss due to the radiation damage.
From equation 4.10 it follows that the corrected crystal response is

$$S_{corr} = S(t)_{raw} \left(\frac{R_0}{R(t)} \right)^{\alpha} \tag{4.17}$$

being $S(t)_{raw}$ the direct measure at time t. To test the precision that can be achieved with this procedure, all the data collected with the blue light in the 2003

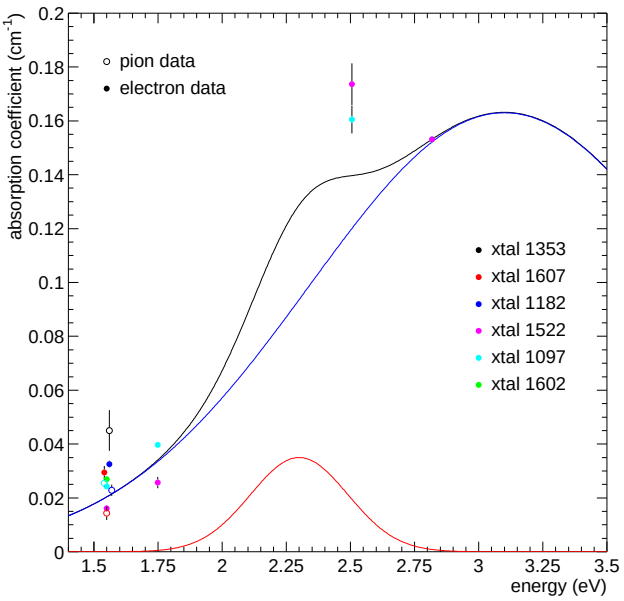

Figure 4.30: *Absorption coefficients versus the photon energy for dose rates $\leq$ 1 Gy/h. (Y/Nb and Y doped crystals). The blue and red lines represent the color centers. The black line represents the absorption coefficient μ.*

year have been corrected with the mean value computed for α during the 2002 test beam (α = 1.53). It must be noticed that the value of the first point now affects the correction procedure, differently with respect to the determination of the α coefficient; to determine it correctly, the mean of the crystal responses in many runs before the irradiation has been considered.
Figure 4.31 reports the corrected and the raw response for a sample crystal during the irradiation with electrons and it shows that the accuracy of the correction does not depend on time and on dose. The same figure also presents the distribution of the residuals $(S_{corr} - S_0)/S_0$ for all the crystals that underwent the irradiation with electrons; the accuracy after the correction is at the 0.15% level. It is worth mentioning here that the measurement of the α parameters *in-situ* will improve the correction precision, so the results discussed here are a superior limit to the ultimate precision which can be reached.
Similar results have been obtained on the crystals irradiated with pions; the accuracy of the correction is at the 0.18% level for the average of all the irradiated crystals and again it is time and dose independent.
The effectiveness of the correction procedure during the whole irradiation period

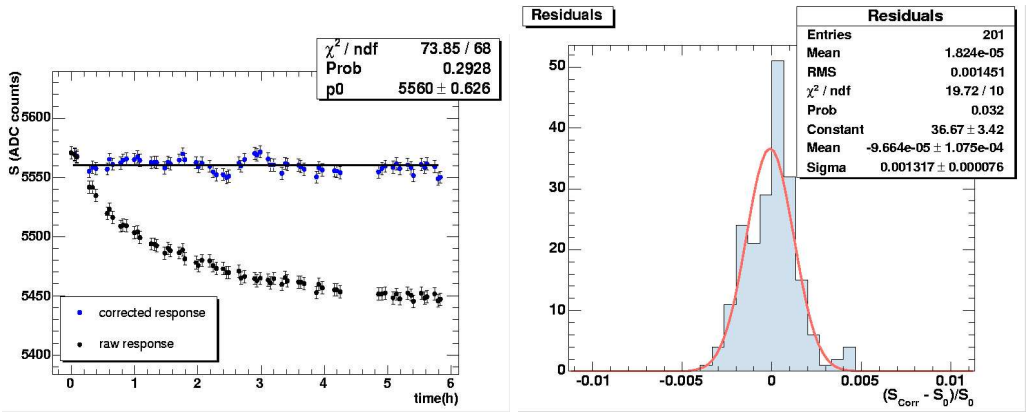

Figure 4.31: *The effectiveness of the correction procedure relative to the irradiation with electrons. Left: dependence of the corrected response on the time for a reference crystal. Right: distribution of the residuals* $(S_{corr} - S_0)/S_0$ *for all the irradiated crystals.*

with electrons and pions was finally tested. Figure 4.32 on the left shows the corrected response as a function of the time for a crystal taken as example. On the

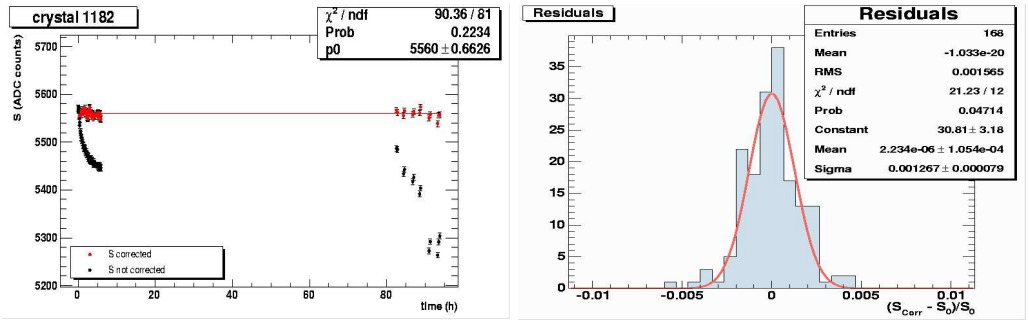

Figure 4.32: *The effectiveness of the correction procedure relative to the whole irradiation period both with electrons and pions. Left: dependence of the corrected response on the time for a reference crystal. Right: distribution of the residuals* $(S_{corr} - S_0)/S_0$ *for the three crystals that underwent both the irradiation with pions and with electrons.*

right the average behaviour of the crystals which underwent both the irradiations is reported. The stability of the response is again at the 0.16% level, in agreement with the results for both the irradiations. Since the R_0 value in equation 4.17 is determined before the first irradiation, the agreement means that the laser effectively tracks the recovery of the crystals response between the two irradiation.

4.6 Conclusions

In this chapter, the procedure for the intercalibration of the ECAL crystals at the test beam has been presented. The robustness of the method was checked in great detail and the intercalibration precision has been shown to be affected by the many possible systematics only at the 0.3% level. The study of the statistical precision of the intercalibration has also shown that only about 1000 events per crystal are enough to keep the statistical accuracy better than 0.1%. The proposed strategy can be therefore usefully exploited for an extensive intercalibration. The results discussed here have been confirmed with a much larger statistics during the 2004 test beam, when a full supermodule was tested.

Since the response of the crystals is affected by the irradiation, the evolution of the crystals response when irradiated with electrons and pions has been studied, also to check its effect on the calorimeter calibration. The reliability of the laser monitoring system has been demonstrated. The proposed correction procedure allows to stabilize the response of each single crystal with an accuracy better than 0.2%.

The results discussed in this chapter confirm that the constant term of the energy resolution can be kept under the 0.5% value which is the ECAL target.

Chapter 5

The electron reconstruction in CMS

The electron reconstruction is a task of primary importance at CMS and the comprehension of the electron behaviour within the detector is crucial to get the best possible information from it for many physics channels. To achieve this goal, the informations coming both from the electromagnetic calorimeter and from the tracker have to be exploited in the best way.
In this chapter the electron reconstruction in CMS is discussed. First, after a quick overview of the CMS software, the algorithms which are currently used are presented. One of the main problems to deal with when considering electrons is the large bremsstrahlung emission due to the tracker material. Here its effects on the reconstruction are discussed and a tracking algorithm to take care of the bremsstrahlung is introduced. Finally some criteria to distinguish between real and fake electrons are presented.
The study is based on the full simulation of the CMS detector. Both the proposed tracking algorithm and the electron identification strategy will be applied in the next chapter to the study of the Higgs decay in the channel $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$. The kinematical characteristics of this decay channel (in particular the presence of 'low' p_T and almost collinear electrons) make indeed a detailed study of the electron reconstruction necessary.

5.1 The CMS simulation and reconstruction software

The analysis discussed both in this chapter and in the next one are based on the full simulation of the CMS detector.
The CMS reconstruction software ORCA (**O**bject-oriented **R**econstruction for **C**MS **A**nalysis) [74] is implemented in C++ and it is based on the Objected-Oriented technology. It consists of the general framework COBRA [75] and of a set of packages which are used for the reconstruction of the different subdetectors.

COBRA provides both the basic computing services and the tools and the utilities which are common to the many subdetectors. In the ORCA packages, the tools for the various subdetectors analysis tasks are developed.
The ORCA software is part of the CMS chain for the simulation and the digitisation of the events, together with some other specific programs.
The event generation uses CMKIN [76], a FORTRAN interface for different event generators which is usually interfaced to the PYTHIA [77] generator. The generated data are then simulated, i.e. propagated through the different subdetectors materials. The simulation is done with OSCAR [78], a C++ program based on GEANT4 [79] which describes both the active areas and the dead zones of each subdetector. OSCAR has recently replaced the old simulation program CMSIM [80], written in FORTRAN and based on GEANT3 [81]. Both signal and pileup events are generated and separately simulated. They are then merged in the digitisation phase, when the simulation of the noise and of the electronics is added together with the selective readout algorithms. The pileup events are selected in a random way from the pileup database according to the wanted luminosity. The ORCA software is then used to digitize the events.

Different samples have been employed for the electron reconstruction analysis presented here. The tracker and calorimetric performance study is based on samples of back-to-back electrons, with flat distribution in ϕ and η ($\mid \eta \mid < 2.5$) and with fixed transverse momentum (p_T = 5, 10, 30 GeV/c). CMSIM 133 has been used for the simulation and ORCA_7_6_1 for the digitisation, without the addition of pile-up events. For the electron identification analysis a sample of electrons coming from the $H \rightarrow WW^{(*)}$ decay has been used. CMSIM 133 and ORCA_7_6_1 have been used respectively for the simulation and the digitisation, with the add of pile-up events corresponding to the low luminosity phase. The sample of fake electrons was taken from QCD di-jets events simulated with OSCAR_2_4_5 and digitised with ORCA_7_6_1 (low luminosity pile-up). Three different samples divided according the transverse momentum exchanged in the parton interaction $\widehat{p}_T$ have been used, covering the region $\widehat{p}_T > 25$ GeV/c (25 GeV/c< $\widehat{p}_T$ <50 GeV/c, 50 GeV/c< $\widehat{p}_T$ <170 GeV/c and $\widehat{p}_T$ >170 GeV/c). The jets samples have been strongly biased at the generation since only a fraction of the original QCD jets has been retained to enhance the electromagnetic component of the samples.

5.2 The standard electron reconstruction in ORCA

The default electron reconstruction in ORCA relies on the combination of both tracker and calorimetric informations. Three main steps can be identified in the

electron reconstruction procedure: the energy clustering in ECAL, the matching between the pixel hits and the ECAL clusters to provide the track seeds and finally the inward-outward track reconstruction. The algorithms for these three steps are implemented in the ORCA *Electron Photon* package and they are currently employed both for the HLT and for the offline analysis. Some details are given in the following.

5.2.1 The calorimetric energy reconstruction

The clustering of the energy deposits in the electromagnetic calorimeter is the first step of the electron reconstruction procedure.
A single electron generates a shower which develops in more than one crystal and the bremsstrahlung emission is responsible for a further spread of the energy in the ϕ direction. It is therefore necessary to collect the energy of all the crystals which are involved.
The starting point is the search for crystals with energy above a certain threshold, the so called *seeds*. The seeds are ordered with decreasing energy and only the most energetic seed is kept among the adjacent ones. Starting from each seed, the energy deposits are then collected using two different reconstruction algorithms [82]: the *Hybrid* algorithm, used by default in the barrel, and the *Island* algorithm, employed in the endcaps.
The **Island algorithm** moves in both directions in ϕ starting from the seed position and it collects all the crystals until it finds an energy rise or a hole. Then it moves one step in η and makes another ϕ search. The same conditions used in the ϕ search also stop the search in η; the algorithm then comes back to the seed and it moves in the opposite η direction. A bremsstrahlung recovery procedure is then applied to create the so called superclusters, groups of clusters which collect the energy released both by the electron and by the emitted photons. The Island algorithm includes each crystal in one cluster only, to avoid double countings of energy. In this way, the energy deposits in the crystals which are below the threshold may remain unclustered, with a significative energy loss. On the other hand, small deposits of energy due to the noise or to pileup events are not clusterized either.
The **Hybrid algorithm** exploits the knowledge of the lateral shower shape in η and searches for bremsstrahlung energy in ϕ. First rows of 3 or 5 crystals in η (the so called 'dominoes') are created and collected, then the dominoes are clustered in ϕ. At each stage, the clusters are required to have energy above a certain threshold. The Hybrid algorithm, a sort of super-clustering algorithm with different clustering steps, is designed to reconstruct relatively high energy electrons in the barrel; small deposits of energy are better described by the island clusters.

After the clustering, the electron energy is computed as the sum of the deposits in the crystals belonging to a supercluster. A correction factor depending on the number of crystals in the supercluster is applied, to minimize the residual dependence of the energy scale both on the energy and on the pseudorapidity.

5.2.2 The pixel matching

The superclusters which are reconstructed in the ECAL are used as starting points for the electron track reconstruction [83]. First the supercluster position and energy are used to compute the electron transverse momentum. The expected hit position in the innermost pixel layer is estimated propagating the electron candidate from the supercluster position to the nominal vertex, taking into account its transverse momentum and the magnetic field. A search area is constructed around the expected point using the error in the ϕ measurement of the cluster and the spread of the vertex z. If a compatible hit is found in such region, it is used together with the supercluster to estimate the true vertex z position. The track is then propagated from this point to the second layer and if a second compatible hit is found the supercluster is kept as an electron. This kind of search is performed for both charge hypothesis. The pixel detector consists of three rings and two disks and two out of the three possible hits are requested to start the electron reconstruction. The output of the search are the seeds which are used to start the track reconstruction.

5.2.3 The track reconstruction

Once the track seed has been identified, the full track reconstruction is performed starting from there and going towards ECAL. The Kalman Filter is used as default in this procedure. In order to guarantee a good momentum resolution and high efficiency, together with the speed which is required for HLT purposes, some parameters involved in the reconstruction of electronic tracks are modified with respect to the default tracking. The minimum number of hits required to keep the track is relaxed from 5 to 3 and the χ^2 cut to check the compatibility of the hit which is found at each layer with the one which is expected is set to 5 (30, in the tracking for muons).
A detailed description of the track reconstruction in ORCA will be given in the following, when a different tracking procedure will be discussed.

5.2.4 The electron reconstruction efficiency

The electron reconstruction efficiency is shown in figure 5.1 as a function of the pseudorapidity for electrons with different transverse momenta.

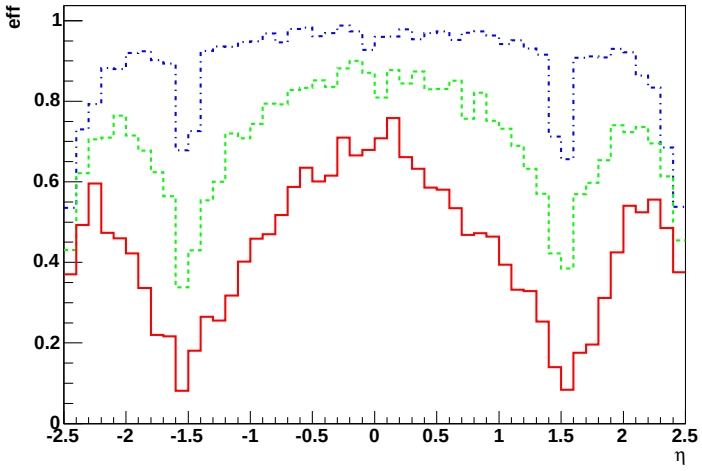

Figure 5.1: *Electron reconstruction efficiency as a function of the pseudorapidity. Single electrons at p_T = 5 (red, solid line), 10 (green, dashed line) and 30 (blue, dash-dotted line) GeV/c have been used. The E_T threshold for the superclusters has been lowered to 1 GeV.*

Here the efficiency is defined as the number of reconstructed electrons over the number of generated ones and it is taken equal to one if more than one electron is reconstructed. No explicit matching with the MonteCarlo truth is done, but this was found not to influence the result in the simple case of single electrons. To deal with low p_T electrons, the E_T threshold to build a supercluster has been set to 1 GeV, while the default in ORCA is 4 GeV.
The electron reconstruction efficiency as a function of the pseudorapidity presents a drop at $\mid \eta \mid = 1.5$ and another one at large pseudorapidity values ($\mid \eta \mid > 2.4$). The first fall corresponds to the transition between the ECAL barrel and the endcaps, while the second one is due to the lack of coverage by the tracker endcap disks. More important, a strong decrease for low values of the electron transverse momentum is evident.
The average electron finding efficiencies are around 45% for 5, 70% for 10 and 90% for 30 GeV/c p_T electrons. When lowering the E_T threshold to build a supercluster, the supercluster reconstruction efficiency is almost 100% and the electron reconstruction inefficiency mostly comes from the tracking. Several sources can be at the origin of such inefficiency. A first source of inefficiency is the lack of full coverage of the detector itself. A second problem is related to the inability of the tracking algorithm to follow the changes of curvature related to the bremsstrahlung emission, as will be discussed in the following. The main source

of inefficiency at low energy is anyway the current track seeding, which requires a very precise matching between the innermost pixel layers and the ECAL. The efficiency decreases when decreasing the transverse momentum of the track since the track curvature, and hence the probability to miss the energy cluster, is higher. It is worth to mention that the current electron reconstruction has been optimized at high energy (p_T = 35 GeV/c) for HLT purposes. Furthermore, the pixel matching selection cuts have been chosen to reject background to triggering electrons at the HLT Level-2.5. The algorithms are therefore not optimized for the low p_T region; some changes in the parameters and in the seeding window could easily help in recovery efficiency in the low energy range. A revisiting of the algorithm to better deal with off-line electrons is on-going.

5.3 The bremsstrahlung problem

The main problem related to the electron reconstruction is the bremsstrahlung radiation in the tracker, which strongly affects both the momentum and the energy measurements. The material budget before the electromagnetic calorimeter, which is shown in fig 5.2 in term of radiation lengths, varies as a function of the pseudorapidity and it has its maximum of about 1.4 X_0 around $|\eta|$ = 1.5. Such large

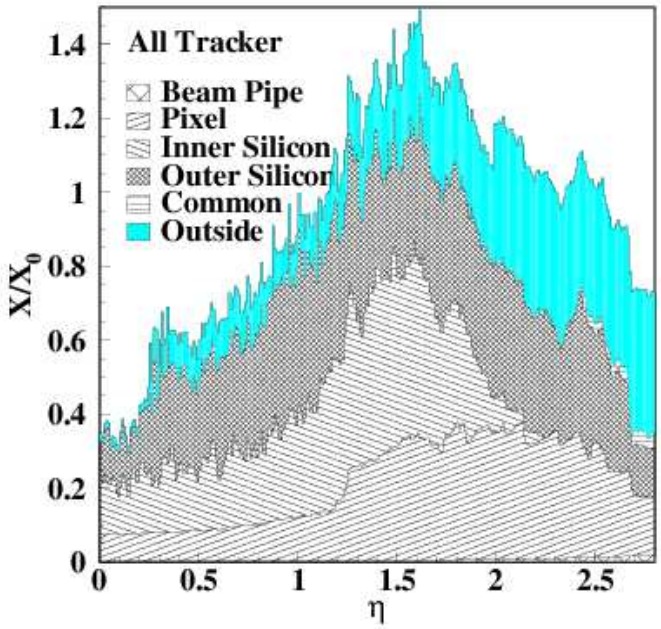

Figure 5.2: *Tracker material budget.*

amount of material, combined with the 4T magnetic field, reflects in a significative bremsstrahlung emission. The importance of the effect is quantified in figure 5.3, which shows the fraction of energy emitted by bremsstrahlung in the tracker

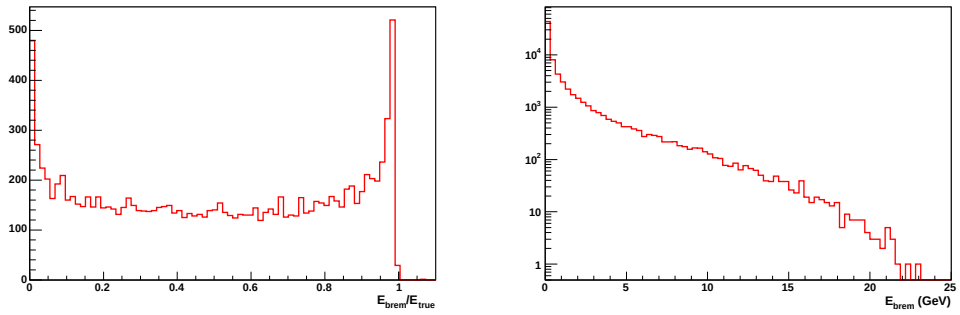

Figure 5.3: *Bremsstrahlung emission for 10 GeV/c p_T electrons in the ECAL barrel. Left: fraction of the electron energy emitted by bremsstrahlung. Right: secondary particles energy. Only photons with energy higher than 10 MeV have been generated.*

for 10 GeV/c p_T electrons in the ECAL barrel. About 50% of the electrons lose more than 50% of their energy and about 9% lose more than 95%. Figure 5.3 also shows the energy of the secondary particles, which is often quite high. The same plots for 30 GeV/c p_T electrons are given in fig 5.4. As can be expected, such

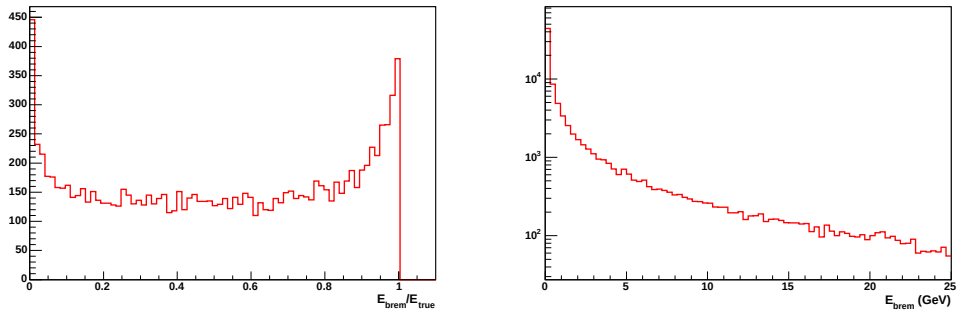

Figure 5.4: *Bremsstrahlung emission for 30 GeV/c p_T electrons in the ECAL barrel. Left: fraction of the electron energy emitted by bremsstrahlung. Right: secondary particles energy. Only photons with energy higher than 10 MeV have been generated.*

large bremsstrahlung emission is directly related to the tracker material. In fig

5.5 the number of photons which are emitted is plotted as a function of the electron pseudorapidity for 30 GeV/c p_T electrons. The pattern clearly reproduces the tracker material budget.

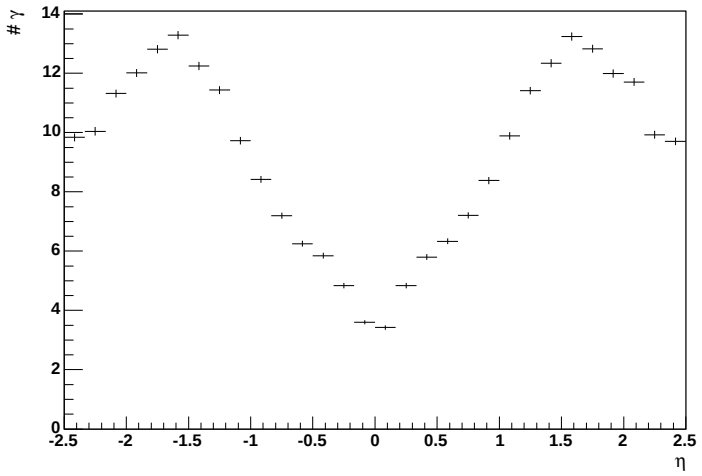

Figure 5.5: *Number of bremsstrahlung photons emitted as a function of the electron pseudorapidity for 30 GeV/c p_T electrons. Only photons with energy higher than 10 MeV have been generated.*

The bremsstrahlung radiation affects both tracker and calorimetric measurements. Due to the kinematics of the bremsstrahlung emission, the photons move along the tangent to the electron trajectory. Their trajectory is a straight line while the electrons curve in the magnetic field and this results in a spread in the ϕ direction of the energy reaching ECAL. The impact points of the photon and of the electron on the ECAL surface are more distant if the photon is emitted at the beginning of the electron trajectory with respect to the case of late emission. The distance between the two impact points also decreases for high transverse momenta.
From the ECAL point of view this affects the electron energy reconstruction. Figure 5.6 on the left shows the reconstructed over the true energy distribution for electrons with fixed p_T in the ECAL barrel. The long tail at low energy is the result of the bremsstrahlung emission, while the Gaussian part of the distribution corresponds to the energy that would be reconstructed in absence of bremsstrahlung. The effect of the bremsstrahlung emission is more evident at low energy due to the largest distance between the electron and the photon impact points on ECAL. In figure 5.7 the dependence of the ratio E/E_{true} on the first 'hard' ($E \geq 100$ MeV)

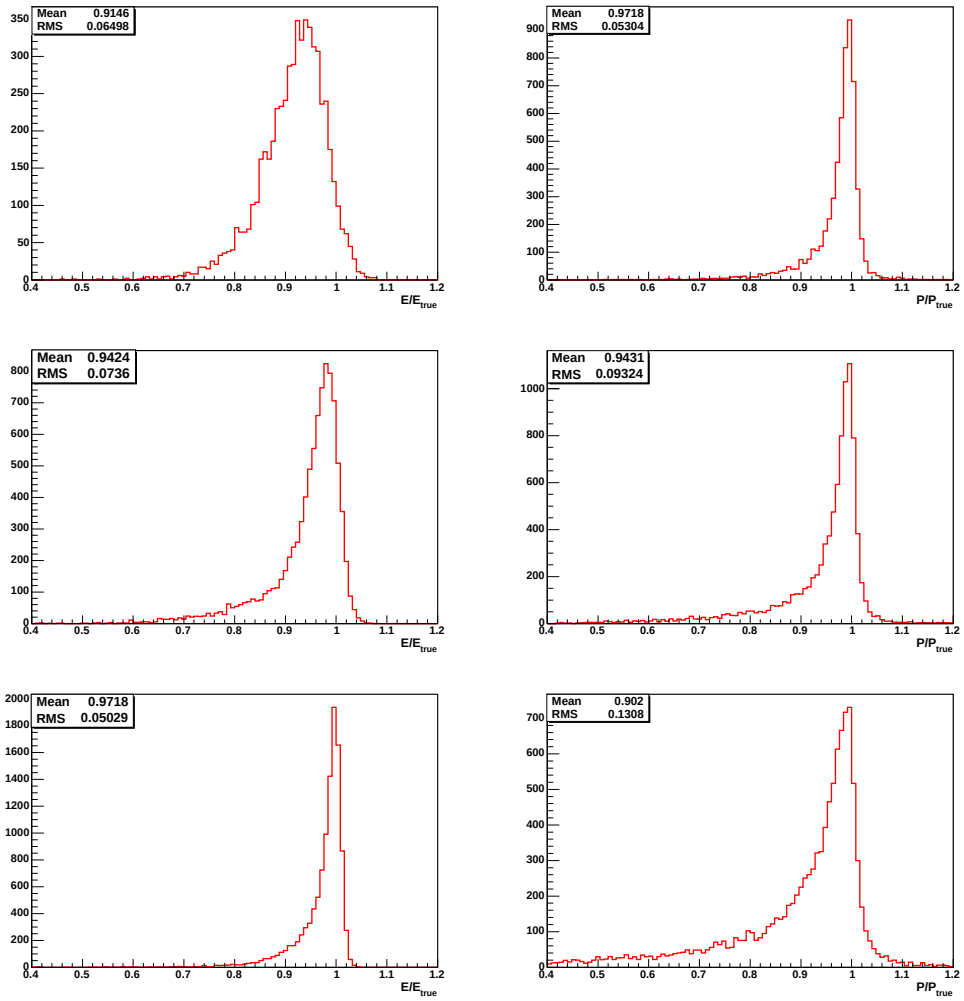

Figure 5.6: *Left: reconstructed over generated energy. Right: reconstructed over generated momentum. Samples of p_T = 5 GeV/c (on the top), 10 GeV/c (in the middle) and 30 GeV/c (on the bottom) electrons in the ECAL barrel region have been used. The supercluster threshold has been lowered to E_T = 1 GeV.*

photon emission point is shown. In case of late radiation the photon cluster partially merges with the electron one, so the loss in the energy measured with ECAL is lower than in case of early radiation.
For what concerns the tracker, when a photon is emitted the track gets more curved than predicted from the most probable value, hence biasing the estimation towards lower p_T values. This creates the tails in the distributions in figure 5.6 on the right, showing the ratio of the reconstructed to the generated momentum. The effect de-

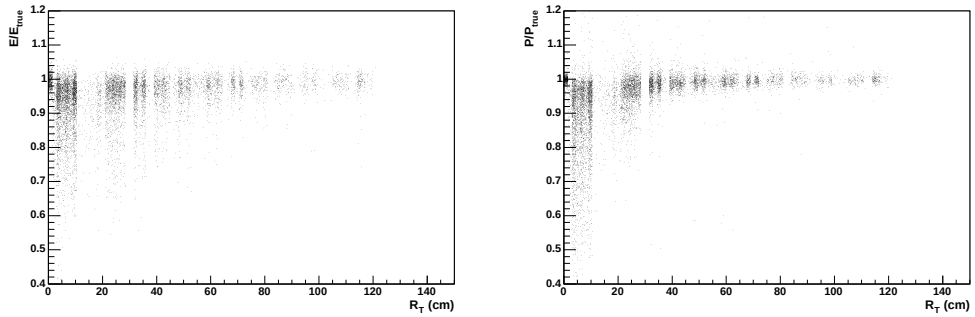

Figure 5.7: *Reconstructed over generated energy (left) and momentum (right) as a function of the transverse radius of the first hard (E≥100 MeV) bremsstrahlung photon. A sample of p_T = 10 GeV/c electrons in the ECAL barrel region has been used. The supercluster threshold has been lowered to E_T = 1 GeV.*

pends on the hardness of the photon which is radiated, so the tail is more evident for higher p_T tracks. Again, a late radiation has only a small effect on the reconstructed track parameters, as it is shown in figure 5.7 on the right. Another important consequence of the bremsstrahlung emission on tracker measurements is the low number of hits in the tracks, which are often stopped by the hard χ^2 cut in the track updating when hard photons are emitted. This is clear looking at figure 5.8, which shows the number of reconstructed hits as a function of the electron pseudorapidity. The average number of hits depends of course on the number of possible layers. In the region where the overall tracker material is larger, anyway, the probability to have bremsstrahlung emissions increases. Such emissions stop the trajectory, which has therefore a smaller number of hits. The pattern in figure 5.8 reproduces the tracker material distribution.
To quantify the effect of the tails in the distributions of figure 5.6, a comparison between the mean value of the histograms and the position of the peak is given in tables 5.1 and 5.2 for the $\mathrm{E}/\mathrm{E}_{true}$ and the $\mathrm{P}/\mathrm{P}_{true}$ distributions respectively. It is worth noticing that the origin of the tails is mainly the sequence of soft bremsstrahlung emissions and not the emission of single hard photons. From the ECAL point of view, for instance, energetic clusters due to very hard photons are easy to be recovered by the superclustering algorithms, therefore a good energy estimation is possible despite such emissions.

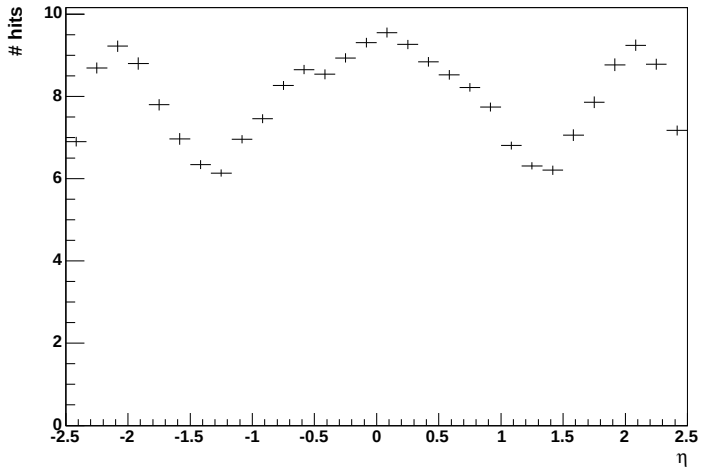

Figure 5.8: *Number of reconstructed track hits as a function of the electron pseudorapidity for p_T = 30 GeV/c electrons.*

E/E_{true}			
	p_T = 5 GeV/c	p_T = 10 GeV/c	p_T = 30 GeV/c
mean value	0.915	0.942	0.972
peak position	0.940	0.980	0.996

Table 5.1: *Comparison between the peak position and the mean value of the E/E_{true} distributions.*

P/P_{true}			
	p_T = 5 GeV/c	p_T = 10 GeV/c	p_T = 30 GeV/c
mean value	0.972	0.943	0.902
peak position	0.996	0.996	0.996

Table 5.2: *Comparison between the peak position and the mean value of the P/P_{true} distributions.*

5.4 An algorithm for the electron track reconstruction

It follows from what was discussed in the previous sections that the bremsstrahlung emission has to be carefully handled to have good performances in the electron re-

construction both from the tracker and the calorimeter point of view. Bremsstrahlung recovery has to be efficiently done when reconstructing the energy in the ECAL; an important role is played also by the tracking algorithm which is used to reconstruct the tracks.
In this thesis we mainly deal with this second aspect. The choice was done in view of the analysis which will be discussed in the next chapter, on which a precise momentum measurement has a major impact that the calorimetric aspect. In this section a method to reconstruct electron tracks based on the Gaussian Sum Filter is proposed and the results obtained on single electrons samples are presented.

5.4.1 Track fitting methods

The Kalman Filter

The default algorithm for the track reconstruction in CMS in the Kalman Filter (KF) [84]. Throughout the filter, the tracks are described by a five dimensional state vector **x** containing the information about the momentum, the direction and the position at each point of the trajectory.
The state vector can be written as a function of a coordinate $\mathbf{x} = \mathbf{x}(z)$ and its evolution as a function of z can be described by a set of differential equations. It is anyway sufficient to consider the state vector in a discrete set of points, like the intersections of the track with the detector. In this way the problem reduces to a discrete system of equations

$$\mathbf{x}(z_k) \equiv \mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \tag{5.1}$$

where $\mathbf{f}_{k-1}$ is the track propagator from the detector $k-1$ to the detector k and the random variable $\mathbf{w}_{k-1}$ is the process noise which incorporates a random disturbance of the track between z_{k-1} and z_k.
In general the state vector is not directly observed and the quantities $\mathbf{m}_k$ measured by the detector are functions of the state vector with a distortion due to the measurement noise ϵ_k

$$\mathbf{m}_k = \mathbf{h}_k(\mathbf{x}_k) + \epsilon_k \tag{5.2}$$

The track fitting with the Kalman Filter proceeds throughout three kinds of operations:

- filtering : the estimate of the state vector with a local measurement
- prediction: the estimate of the state vector in the future
- smoothing : the estimate of the state vector in the past, using all the measurements collected up to the present time

The track vector is extrapolated from the detector $k-1$ to detector k by means of the track model, then the extrapolated state vector is updated with the measurements on the detector k. The covariance matrix of the extrapolated state vector is computed by error propagation and the covariance matrix of the process noise between the detector $k-1$ and k is added to the propagated one [85]. The smoothing of the estimate state vector can be done running two filters both inward-outward and outward-inward and then combining both the predictions with the measured value.
The Kalman Filter is a least square estimator. It is the optimal filter when the system is linear and both $\mathbf{w}_k$ and ϵ_k are Gaussian random variables, otherwise non linear filters can do a better job.

The Gaussian Sum Filter

The process noise **w** accounts both for the multiple scattering and the energy loss. In case of high energy electrons, the ionization loss can be neglected with respect to the energy loss due to bremsstrahlung. The latter can be described using the Bethe-Heitler [86] model, in which the probability density function $f(z)$ of the electron energy loss is

$$f(z) = \frac{[-lnz]^{c-1}}{\Gamma(c)} \tag{5.3}$$

Here $c = t/ln2$, being t the thickness in units of radiation length of the material that the electron crosses, and z is the fraction of energy remaining after the material layer is traversed. The Bethe-Heitler pdf, which does not depend on the particle energy, is shown in figure 5.9 [87].
While the multiple scattering can be well described with a single Gaussian, such approximation doesn't hold for the Bethe-Heitler distribution. The Gaussian Sum Filter [88] [89] is a non linear generalization of the Kalman Filter which approximates the distribution of the process noise by a mixture of Gaussians. The resulting filter is a weighted sum of Kalman Filters running in parallel, with weights which depend on the observations. Each KF corresponds to one of the components of the mixture.

The Gaussian Sum Filter (GSF) approach has already been applied to the track reconstruction in ORCA [87]. Since it requires the approximation of non Gaussian distributions by Gaussian mixtures, the performances of the filter strongly rely on the quality of such approximation. For each component of the mixture the weights, the mean values and the variances are determined by minimizing the

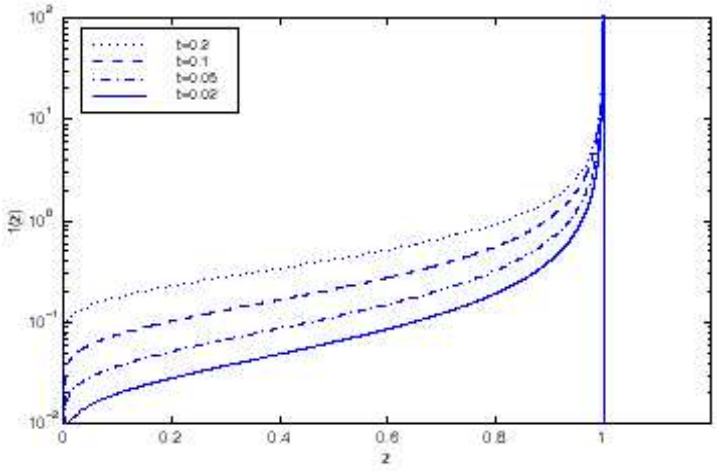

Figure 5.9: *Probability density function f(z) for different values of the path length; z is the fraction of energy remaining after the material layer is traversed [87].*

distances

$$D_{cdf} = \int_{-\infty}^{+\infty} \mid F(z) - G(z) \mid dz$$
$$D_{kl} = \int_{-\infty}^{+\infty} ln\left(\frac{f(z)}{g(z)}\right) f(z)dz \tag{5.4}$$

where $f(z)$ and $F(z)$ are respectively the pdf (Probability Density Function) and the cdf (Cumulative Density Function) of the Bethe-Heitler distribution and $g(z)$ and $G(z)$ are the pdf and the cdf of the Gaussian mixture. The computation being quite fast, the calculation of the mixture is done in CMS on the fly during the reconstruction. The effective thickness of a detector layer, from the knowledge of the incident angle, enters the minimization of the distances in equation 5.4.

5.4.2 Electron reconstruction using the Gaussian Sum Filter

The Gaussian Sum Filter algorithm described in the previous section has been employed to develop a new method for the reconstruction of electron tracks in CMS. It combines the electron track seed finding in the pixel detector, the building and the fitting of the charged tracks using the GSF and the geometrical matching with energy deposits clustered in the electromagnetic calorimeter. Particular attention is devoted to recover efficiency in the low p_T region and to follow the tracks up to the end even if high energy photons are emitted.

Electron track reconstruction

In the ORCA *TrackerReco* package, the track reconstruction is decomposed into four modular components. First, the track seeds are looked for with the *Seed Generator*. Then the *Trajectory Builder* constructs all the possible trajectories for a given seed. With the *Trajectory Cleaner* the ambiguities among the possible trajectories are solved and a maximum number of track candidates is kept. Finally, the final fit of the track is performed with the *Trajectory Smoother*, which uses all the compatible hits to estimate the track parameters at each layer through a backward fit.
In the following our implementation of the first two steps is described. The default cleaning implemented in ORCA is used. Finally, for the smoothing step, the GSF is used instead of the KF. To have the complete information both at the vertex and at the outermost state, the GSF is used in both the inward-outward and in the outward-inward fit.

Seed generation
In order to build tracks starting from inward and going outward, a seed is created when two hits compatible with a given beam spot are found in the pixel detector. To avoid having too many combinatorics, the search for seeds is restricted to a region compatible with a supercluster in the ECAL detector, but the supercluster is then not directly used to determine the seed parameter. The method described in section 5.2.1 is used to build superclusters, with the supercluster E_T threshold lowered to 1 GeV to improve efficiency at low p_T. For the present analysis, the vertex is assumed within a cylinder of radius equal to 300 μm and length of 15 cm around the nominal vertex; the hits are required to match a supercluster within a loose cone with $\Delta\phi = 0.25$ and $\Delta\eta = 0.15$. All the seeds for the different superclusters are collected together and then cleaned to avoid redundancy. Finally a minimum transverse momentum of 3 GeV/c is required.
In table 5.3 the seeding efficiencies integrated over the whole pseudorapidity range for different values of the electron transverse momentum are reported. Thanks to the relaxed parameters, high efficiency is obtained also for the low p_T momentum range.

p_T = 5 GeV/c	p_T = 10 GeV/c	p_T = 30 GeV/c
0.86	0.95	0.96

Table 5.3: *Seeding efficiency on single electrons with the GSF tracks.*

Trajectory building

Starting from the seed, a trajectory is created. At each step, compatible hits are looked for on the next compatible layers, then the track parameters and the error matrix are propagated using the Bethe-Heitler modeling of the energy losses. All the material effects are neglected between the track layers and concentrated there. This procedure is iterated until the end of the tracker layers or when no hit is found on a compatible layer, with a tolerance of one layer without hits. The trajectory state at each layer is computed using both the predicted state and the measured hit. The compatibility is defined using the χ^2 of the residuals of the extrapolated track parameters on the new surface and the reconstructed hit on the same surface. If the χ^2 value exceeds a certain limit the hit is not taken. Not to lose efficiency at this stage, it was chosen to loose the tracker quality reconstruction requirements and to set the χ^2 cut to a very large value (100000). If many hits are found on a compatible layer, many candidate trajectories are grown in parallel. At the end only a maximum number of candidates per layer is kept (2 in this study). The GSF is used also in this forward fit, to have all the necessary information both at the vertex and at the last hit. A minimum of 5 hits is finally required to create a track.

Figure 5.10 shows the global track reconstruction efficiency as a function of the pseudorapidity, again taken equal to one when more than one track is recon-

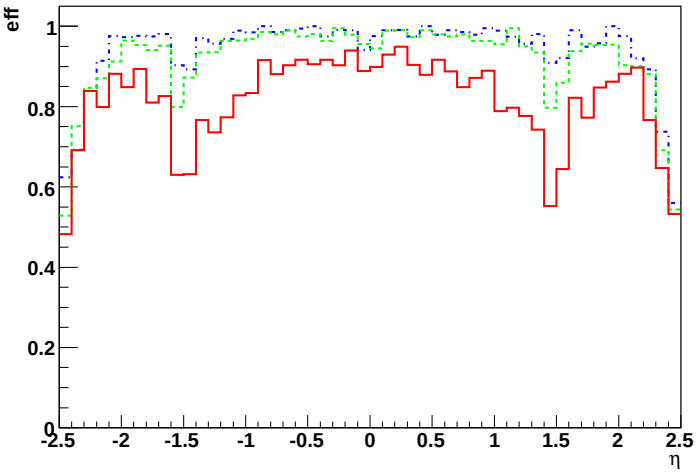

Figure 5.10: *Electron track reconstruction efficiency as a function of the track pseudorapidity for fixed p_T = 5 (red, solid line), 10 (green, dashed line), 30 (blue, dash-dotted line) GeV/c electrons. The results are obtained using GSF tracks. The E_T threshold for the superclusters has been lowered to 1 GeV.*

structed. The algorithm is quite efficient in the full pseudorapidity range, with the

two drops at $|\eta| = 1.5$ and at $|\eta| = 2.4$ already discussed in par 5.2.4.
The efficiencies integrated over the whole range in η for different values of the electron transverse momentum are given in table 5.4 and shown in figure 5.11; at low p_T, a very good efficiency is obtained for the proposed algorithm. Also shown for a comparison in figure 5.11 is the efficiency obtained with the method discussed in 5.2 and currently in use in ORCA (these tracks will be referred to as *Egamma* tracks in the following). It is worth anyway remembering that the method has been mainly optimized for triggering electrons, as already discussed.

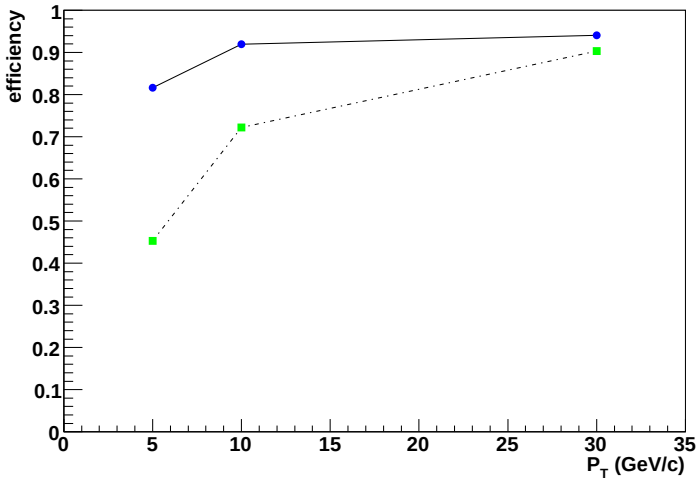

Figure 5.11: *Electron track reconstruction efficiency as a function of the the electron transverse momentum. The results are given for the GSF tracks (blue, solid line) and for the Egamma tracks (green, dash-dotted line). In both cases, the E_T threshold for the superclusters has been lowered to 1 GeV.*

p_T = 5 GeV/c	p_T = 10 GeV/c	p_T = 30 GeV/c
0.82	0.92	0.94

Table 5.4: *Seeding plus tracking efficiency on single electrons with the GSF tracks.*

Table 5.5 reports the fraction of events in which N tracks have been reconstructed. The fraction of events with more than 2 reconstructed tracks gives the order of magnitude for the fake rate in the GSF tracks reconstruction on the study sample with single electrons. To have a better feeling of the possible fake rate, some checks were also done on real events from Higgs samples (with pile-up added).

Provided the use of an electron identification criterion, the fake rate was found to be below the percent level, so well under control.

p_T (GeV/c)	N = 0	N = 1	N = 2	N >= 3
5	0.03	0.30	0.66	0.004
10	0.01	0.14	0.83	0.02
30	0.004	0.10	0.85	0.04

Table 5.5: *Fraction of events with N reconstructed tracks when 2 electrons are generated.*

The bremsstrahlung losses are better taken into account with the Gaussian Sum Filter than with the Kalman Filter, so the emission of hard photons does not stop the track. This allows to follow the track up to the end in most of cases, as shown in figure 5.12 for 10 GeV/c p_T electrons. The peak of the distribution is at the expected value, confirming that the gain in efficiency is not related to fake tracks. For a comparison, the number of reconstructed hits is shown also for the *Egamma* tracks and for the tracks built using the default *Combinatorial Track Finder* of the ORCA *TrackerReco* package which exploits the Kalman Filter (which will be referred to as *KF* tracks in the following).

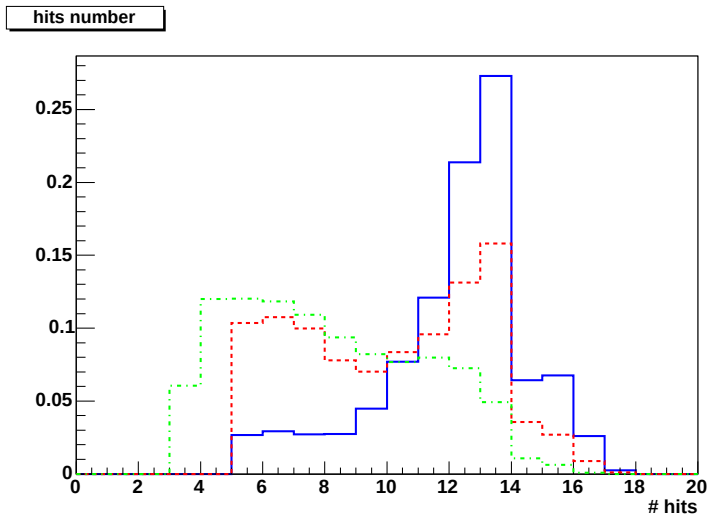

Figure 5.12: *Number of reconstructed track hits for p_T = 10 GeV/c electrons. The results are given for the GSF tracks (blue, solid histogram), the KF tracks (red, dashed histogram) and the Egamma tracks (green, dash-dotted histogram).*

5.4.3 Results on single electron tracks in the CMS detector

When using the Gaussian Sum Filter to fit the track, the parameters of all the gaussians which enter the mixture are available at each hits. A method to compute the track parameters is, given the track state on each layer, to compute the weighted mean of all the components. An alternative way is to consider only the most probable value of the pdf, so giving more importance to the highest weight component. Quite different results are obtained in the two cases, both bringing informations. The choice between the two possible strategies depends on whether one is interested on optimizing the average behaviour or rather on having the best possible parameter estimation for the tracks with small amount of emitted bremsstrahlung For the analysis described in the next chapter the second possibility has been chosen.
The distributions for the reconstructed transverse momentum at the vertex are shown in fig 5.13 for the p_T = 10 and 30 GeV/c case. When taking the mean (left) of the components, a gaussian distribution with a tail which is more pronounced on the right is obtained. When taking the mode (right), the distribution is instead well peaked at the correct value with the long left tail typical of the bremsstrahlung losses. This behaviour is quite similar to the one obtained by tightening the track to follow a non radiating expectation, as done in the *Egamma* tracks procedure.
The track direction estimation gives similar results with both methods.
The possibility to choose between the two estimates is one of the main advantages of the proposed strategy. Once the trajectory building is decorrelated from the fit (which can not be done ie with the *Egamma* reconstruction procedure, since high p_T resolution can only be obtained by truncating the trajectory), any fitting procedure could be in principle used to optimize the momentum estimate. The scheme proposed here, using the mode of the distribution only, is a consistent possibility.

Track parameters at the vertex

In this section some comparisons between the reconstructed track parameters at the vertex and the MonteCarlo truth are presented. Figure 5.14 shows the difference between the electron pseudorapidity extracted from the reconstructed momentum at the vertex and the generated one for 30 GeV/c p_T electrons. Figure 5.15 gives the comparison for the ϕ coordinate.
Again, the results are reported for the three track reconstruction algorithms and for the GSF tracks both the results obtained using the mode and the mean value of the pdf are shown. The results obtained with the different algorithms are well compatible both for the η and the ϕ coordinate.
A Gaussian fit in the central part for the ϕ coordinate shows small differences, as expected from the different bremsstrahlung treatments, with a slightly better reso-

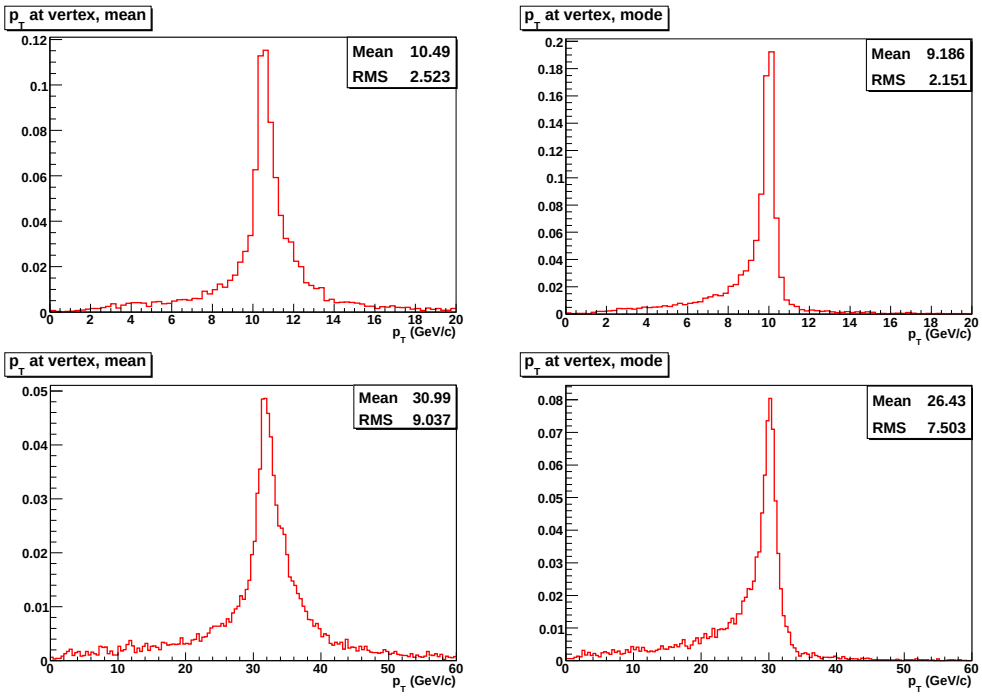

Figure 5.13: *Reconstructed p_T distribution for fixed p_T = 10 GeV/c (on the top) and p_T = 30 GeV/c (on the bottom) electrons. Left: all the components of the gaussian mixture are used to evaluate the track momentum. Right: the mode of the pdf is used.*

lution for the GSF case with respect to the other ones. The ϕ coordinate is indeed sensitive to the bremsstrahlung emission and a gain can come from having more hits providing informations. The effect is anyway very small, probably due to the fact that the momentum resolution mostly comes from the pixel hits, the other hits having a smaller weight in the ϕ estimation. Figure 5.16 shows the comparison between the *Egamma* and the GSF results at p_T = 10 and 30 GeV/c. The effect is higher at high transverse momentum, where from the pure tracker point of view, the resolution is worse.

In figure 5.17 a comparison between the reconstructed electron momentum at the vertex and the generated one for the three tracking algorithms is given. The *Egamma* tracks have the best peak resolution but with a sizable bias, plus a long tail on the left hand side. Due to the combination of the gaussians of the mixture, part of the tail moves to the right in the GSF tracks also when considering the mode of the mixture. In this case the distribution has a bulk of almost the same width with respect to the *Egamma* tracks and a larger spread. This is partially due

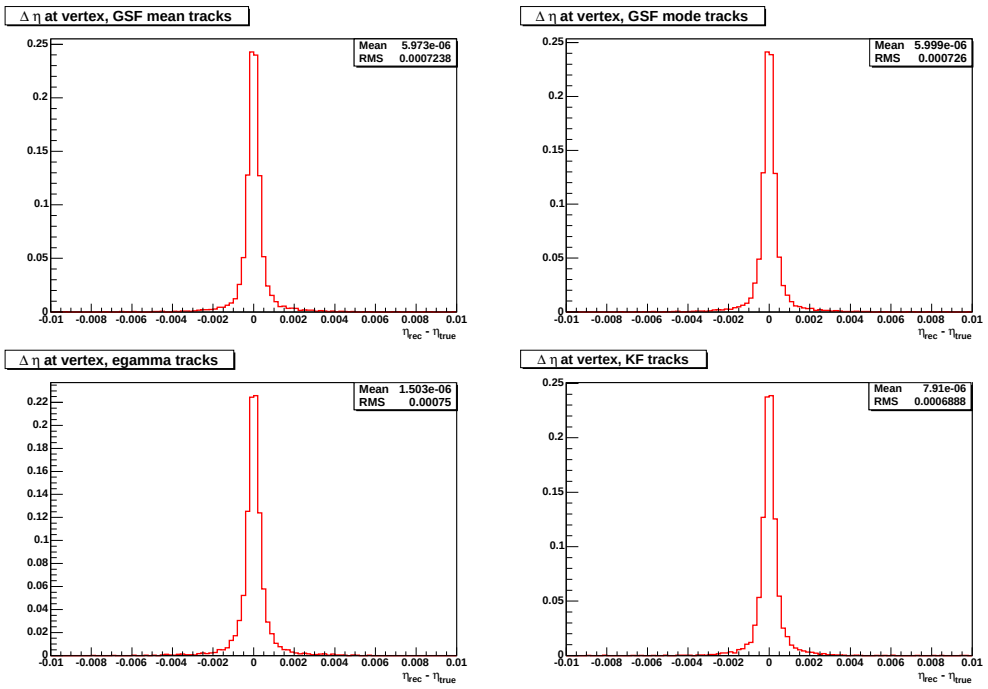

Figure 5.14: *Difference between the pseudorapidity of the reconstructed track at the vertex and the generated one, for p_T = 30 GeV/c electrons. On the top: GSF tracks. The momentum is computed using the mean value (left) and the mode (right) of the pdf. On the bottom: Egamma tracks (left) and KF tracks (right).*

to the events which are not reconstructed with the *Egamma* algorithm and which are here recovered, which are often the most difficult cases. Finally, for what concerns the KF tracks, it is worth mentioning that the algorithm was tuned to reconstruct muons and not electrons. A simple test done using the Bethe-Heitler energy loss parameterization both for the building and the fitting has shown results similar to those obtained with the GSF using the mean of the pdf. Again, the advantage of the GSF is in the possibility to choose the better way of reconstructing the momentum.

Track parameters at the outermost state

Since the tracks arrive almost in front of the calorimeter, a good estimation of the track parameters at the ECAL entrance is possible. This gives both the possibility of improving the matching between the tracker and the calorimeter and of estimating the bremsstrahlung fraction using the tracker alone.

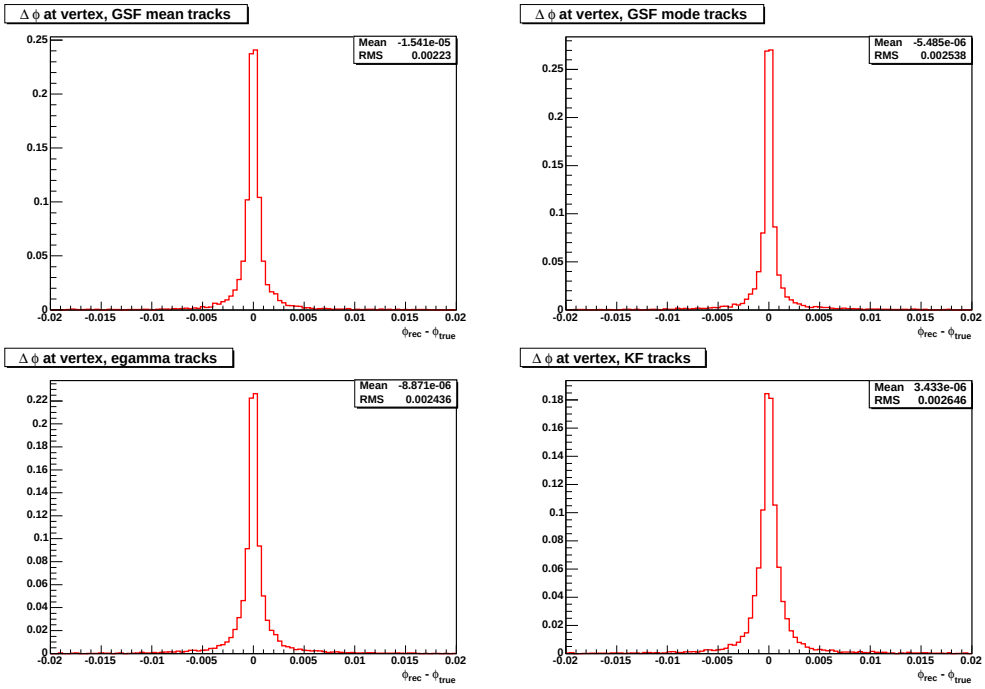

Figure 5.15: *Difference between the ϕ coordinate of the reconstructed track at the vertex and the generated one, for p_T = 30 GeV/c electrons. On the top: GSF tracks. The momentum is computed using the mean value (left) and the mode (right) of the pdf. On the bottom: Egamma tracks (left) and KF tracks (right).*

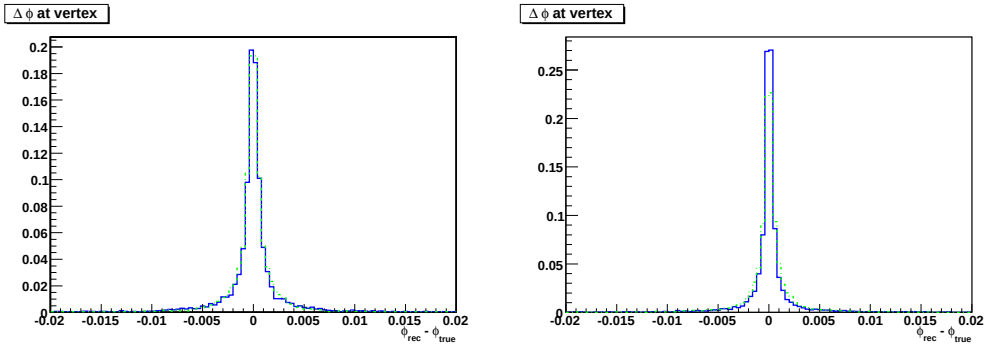

Figure 5.16: *Difference between the ϕ coordinate of the reconstructed track at the vertex and the generated one, for p_T = 10 GeV/c (left) and p_T = 30 GeV/c electrons (right). The results obtained with GSF tracks (blue, solid line) are compared with those obtained with Egamma tracks (green, dash-dotted line).*

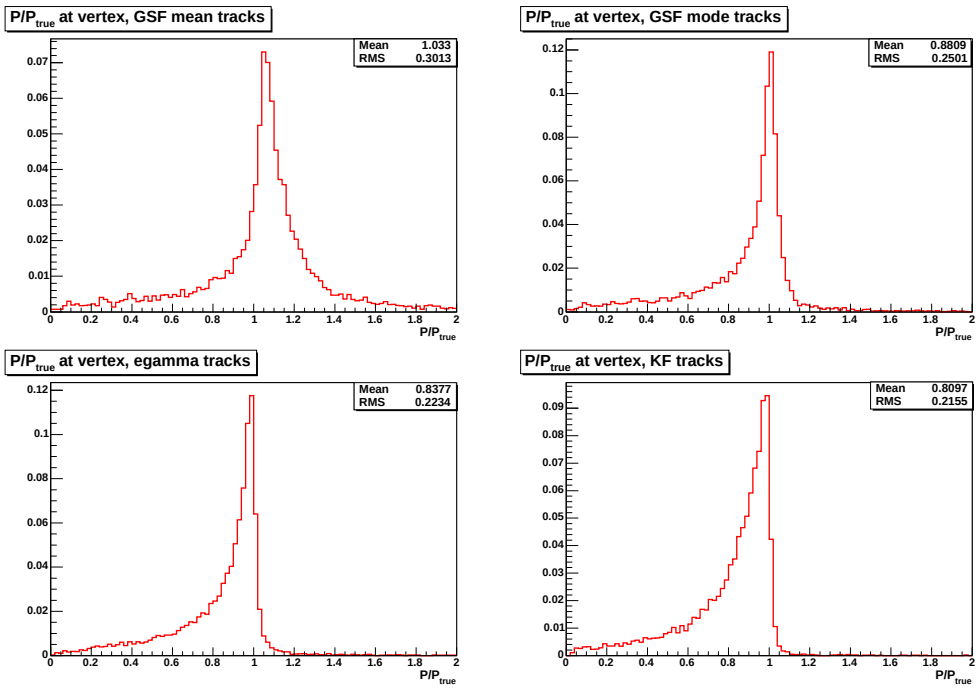

Figure 5.17: *Comparison between the reconstructed momentum at the vertex and the generated one, for p_T = 30 GeV/c electrons. On the top: GSF tracks. The momentum is computed using the mean value (left) and the mode (right) of the pdf. On the bottom: Egamma tracks (left) and KF tracks (right).*

Matching between tracks and ECAL

Due to the bremsstrahlung emission, the matching between the tracker and the ECAL is often done using the track parameters at the vertex, which are known with better precision thanks to the backward fit. This is possible due to the simple kinematics of the bremsstrahlung emission, where the average calorimeter position is the position the track would have reached without bremsstrahlung. Of course this doesn't hold in case of more complicated topologies, as for instance when a bremsstrahlung photon converts. The track parameters at the outermost state are known with larger uncertainty with respect to the ones at the vertex, but they can still be used for the matching.

Figure 5.18 shows the ratio E/P computed both using the mode and the mean value of the pdf. On the left, the momentum evaluated at the vertex is used to compare with the supercluster energy; on the right, the momentum is taken at the outermost state and the energy comes from the seed of the supercluster. In both cases a tail is present. When using the track parameters at the vertex, a tail on the left can be seen and this can be due both to the events in which the momentum

is overestimated and to a reconstructed energy which underestimates the real one. When using the track parameters at the outermost state, instead, the tail is on the right. This is mainly due to events in which part of the bremsstrahlung photons contributes to the seed energy together with the electron but they are not taken into account when estimating the momentum at the end. An extreme case of that is when the seed cluster corresponds to the photon rather to the electron shower, which appears to be the case in the current superclustering algorithm when the emitted brem takes more than half of the electron energy. Uncertainties in the momentum estimation are also responsible for this.
The results obtained using the momentum estimated using the mode and the mean of the pdf are qualitatively similar.

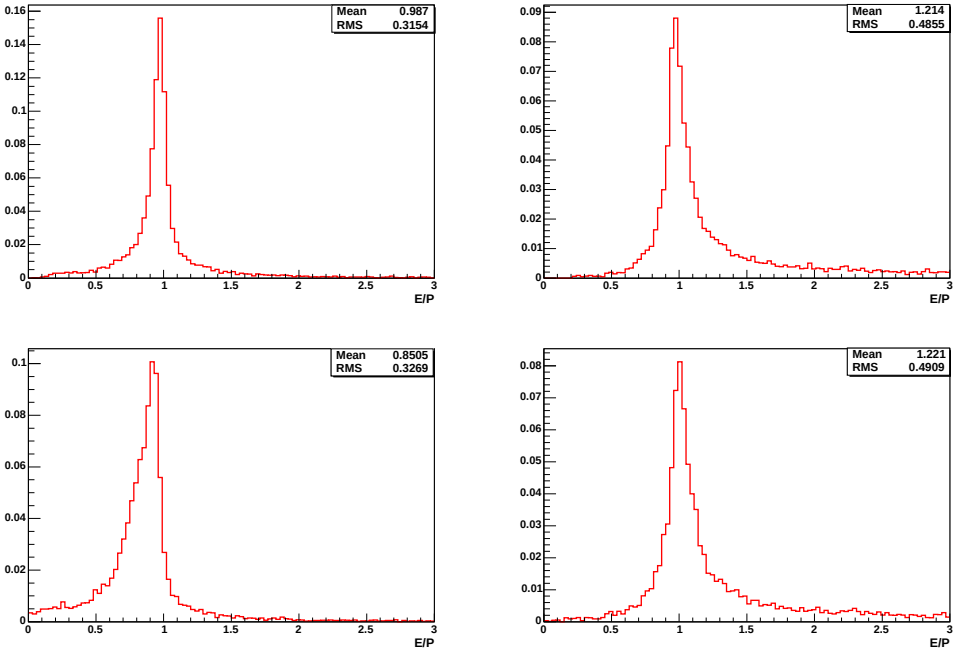

Figure 5.18: *Ratio E/P for p_T = 10 GeV/c electrons. On the left: track momentum at the vertex over the supercluster energy. On the right: track momentum at the last hit over the supercluster seed energy. The plots on the top are obtained using the mode of the pdf to compute the momentum and the plots on the bottom using the mean value of all the components of the gaussian mixture.*

The bremsstrahlung fraction estimation

The knowledge of the track momentum at the outermost state gives the possibility of estimating the fraction of energy which is lost by bremsstrahlung emission.

With the standard tracking, as soon as an hard photon is emitted the track is stopped, so the magnitudes of the momentum at the vertex and at the outermost state are comparable. With the proposed method they are much more different, since the tracks now include also hits after significant bremsstrahlung losses. A comparison between the momentum at the vertex and at the outermost state is given in figure 5.19 for GSF tracks and for *Egamma* tracks.

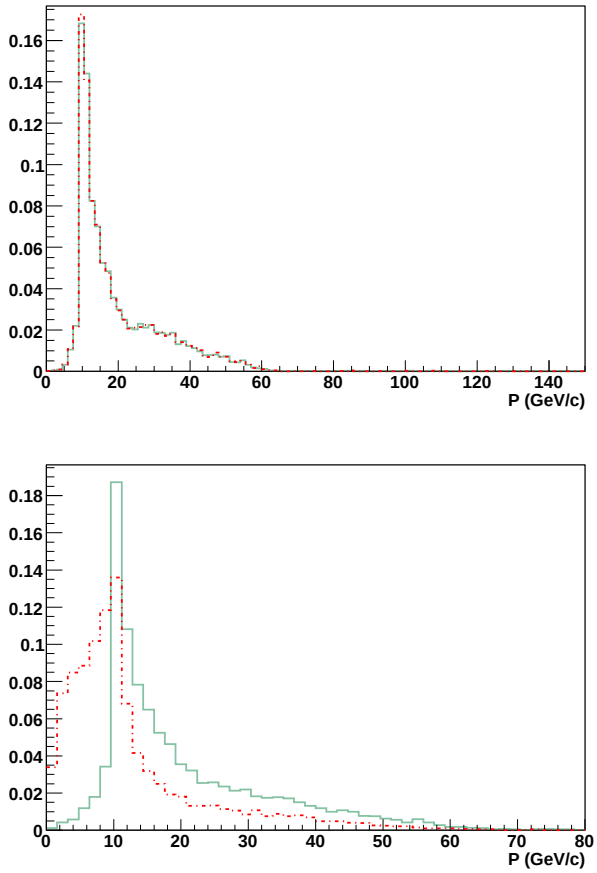

Figure 5.19: *Comparison between the track momentum at the vertex (green, solid histogram) and at the outermost hit (red, dashed-dotted) for 10 GeV/c p_T electrons. Top) Egamma tracks: the two distributions are very similar. Bottom) GSF tracks: a large difference is visible.*

The difference among these two quantities is a measure of the bremsstrahlung fraction. Figure 5.20 on the left shows the difference between the momentum magnitude at the vertex and at the last hit versus the energy of the generated photons which are emitted by bremsstrahlung. A correlation can be observed, even if

many fluctuations are present. Part of the fluctuations are canceled out when the calorimetric information is combined with the tracker one. This is shown in figure 5.20 on the right, in which only the electrons for which at least one bremsstrahlung cluster separated from the seed is reconstructed in the calorimeter are considered. The correlation is more evident [90].
The difference between the momentum at the vertex and at the last hit can be therefore used to select electrons which have emitted only a small fraction of their initial energy. This can be usefully exploited especially for calibration purposes.

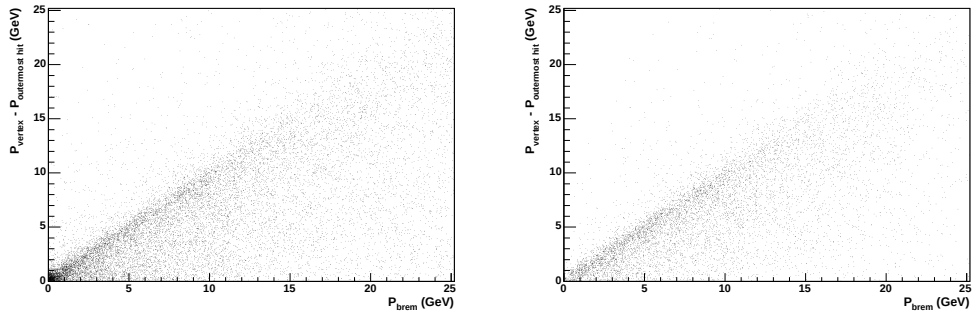

Figure 5.20: *Difference between the track momentum at the vertex and at the outermost state versus the generated energy emitted by bremsstrahlung photons. Left: all the events are considered. Right: only electrons with at least one bremsstrahlung cluster in the ECAL are considered.*

5.5 Electron identification

To build electrons, the electron tracks described in the previous sections are geometrically matched with superclusters in the electromagnetic calorimeter, then some more criteria have to be applied to define a robust electron identification strategy. At hadron colliders a huge QCD jets background is expected. Besides, simple electrons could be faked by processes such as $\pi^0\pi^+$ overlap and π^0 conversions or multiple electron candidates corresponding to a single supercluster could result when tracks from underlying events lie close to the electron track. Here a method to identify electrons is discussed.
An electron is characterized by a narrow cluster in ECAL and a track pointing towards it. Different variables could be used to separate real electrons from fakes. Two main different approaches are currently in use in CMS to identify electrons. The first one selects a certain number of discriminating variables and then applies

some cuts on them [91]; the second approach exploits a likelihood method to combine the variables [92]. For this study the latter was preferred since it should in principle give better results by combining all the available informations at the same time. It is worth mentioning anyway that such approach is more sensitive to the accuracy in the MonteCarlo simulation.
The likelihood for an electron candidate to be consistent with the hypothesis Ψ is built using a vector $\vec{x} = (x_1, ...x_n)$ of discriminating variables as

$$L(\vec{x}, \Psi) \; = \; P(\vec{x}, \Psi) \; = \; \prod_{i=1}^{n} P_i(x_i, \Psi) \tag{5.5}$$

$P_i(x_i, \Psi)$ gives the probability for the variable i to assume value x_i under the hypothesis Ψ. The latter equation is valid under the assumption of uncorrelated variables.
The likelihood ratio to discriminate between real and fake electrons is given by

$$eleID \; = \; \frac{L(\vec{x}, ele)}{L(\vec{x}, ele) + L(\vec{x}, jet)} \tag{5.6}$$

The PDFs are built in the form of binned histograms.
The samples used in this study are those discussed in section 5.1. The reference histograms for real electrons have been built using the electron candidates matching the true ones coming from the two Ws in the Higgs samples. For the backgrounds, the jet samples have been used; since the jet backgrounds cover a large $\widehat{p}_T$ spectrum, each sample was given an opportune weight.

The different variables which could be used to discriminate between real and fake electrons mostly rely on the differences between the shower shapes in the calorimeter and on the combination of tracker and ECAL informations. The following variables have been chosen to build the likelihood for the analysis:

- H/E, the ratio between the energy which is deposited in the electromagnetic and the hadron calorimeter. Electrons tend to fully deposit their energy in the electromagnetic calorimeter, while hadrons leave a large fraction of their energy in the HCAL. The ratio is therefore close to zero for electrons and it is higher for hadrons.

- E9/E25 and $\sigma_{\eta\eta}$, which exploit the fact that electromagnetic showers are narrower than hadronic ones. Since the shower from a true electron is almost completely contained in a 3×3 array of crystals, the ratio E9/E25 between the energy of the 3×3 and the 5×5 crystals arrays around the seed has a more pronounced peak around one in the electron case. The variable $\sigma_{\eta\eta}$

$$\sigma^2_{\eta\eta} \; = \; \frac{\sum_i (\eta_i - \eta_{seed})^2 E_i}{\sum_i E_i},$$

with i running on the 5×5 matrix around the seed, is instead a measure of the shower spread in η. Since the latter is not affected by bremsstrahlung, this variable allows a better discrimination between electrons and fakes with respect to the analogous for ϕ and it is almost independent on the energy.

- $\mid \eta_{track} - \eta_{sc} \mid$ and $\mid \phi_{track} - \phi_{sc} \mid$. Electron tracks directly point on the associated supercluster. These two variables compare the track direction with the supercluster position which is measured with an energy weighted mean of the crystals positions. In the case of $\Delta\phi$, the track is extrapolated from the vertex to the cluster position with the inclusion of the magnetic field only. Like the ratio E9/E25, $\Delta\phi$ is affected by bremsstrahlung emission and it is therefore energy dependent.

- $E_{seed}/P_{outermost}$, the ratio between the seed energy and the track momentum at the outermost state. As discussed in precedence, this variable is expected to be close to one for real electrons.

Notwithstanding the dependence on the energy of some of the chosen variables, it was decided not to divide the electron candidates in different p_T regions, since this was found *a posteriori* not to give a significatively better identification power. Different distributions have been instead done for electrons in the barrel region and in the endcaps. The hypothesis of uncorrelation between the variables has been checked. In most of cases, no important correlation has been found.
The distribution of the eleID variable is shown in figure 5.21 for real electrons coming from the Higgs tree and for fake electrons reconstructed in jets. The efficiency on signal electrons and on electrons coming from jets is given in figure 5.22 as a function of the eleID cut. It is defined as the number of electrons passing the electron identification cut with respect to the total reconstructed candidates.
The selection efficiency for the signal and the three background samples is given in table 5.6 for different values of the eleID cut. The latter has to be chosen

jet samples	eleID> 0.15	eleID> 0.50	eleID> 0.90	eleID> 0.98
25 GeV/c $< \widehat{p}_T <$ 50 GeV/c	8%	4%	1.7%	0.8%
50 GeV/c $< \widehat{p}_T <$ 170 GeV/c	7%	3%	1.2%	0.5%
$\widehat{p}_T >$ 170 GeV/c	5%	3%	1.0%	0.4%
signal	97%	95%	91%	85%

Table 5.6: *Electron identification efficiency for signal electrons and fake electrons reconstructed in the three jets sample for different values of the electron ID cut.*

according the specific analysis. If high efficiency is required, a loose cut can be

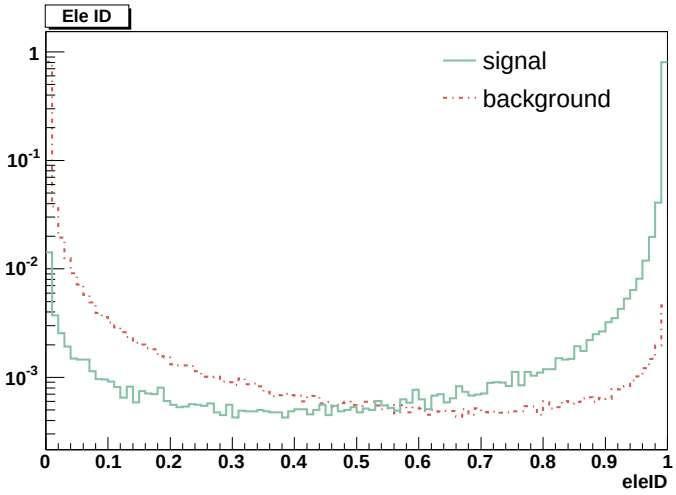

Figure 5.21: *EleID distribution for electron candidates which are reconstructed in the signal and in the background samples.*

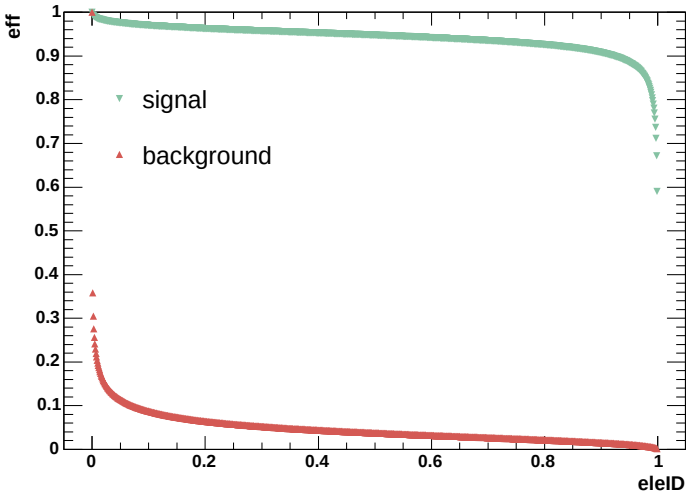

Figure 5.22: *Selection efficiency on the signal and the background as a function of the eleID cut.*

applied. If the electrons purity is more important than the selection efficiency instead a tight cut is needed. This is the case of the $H \rightarrow WW^{(*)}$ analysis which will be presented in the next chapter.

5.6 Conclusions

In this chapter the problem of the electron reconstruction and identification at CMS has been addressed.
The bremsstrahlung emission due to the tracker material has been shown to affect both the tracker and the calorimeter performances, in term of reconstruction efficiency and precision. An overall tracking strategy for the seed finding, the trajectory building and the fitting of the track parameters with the use of the Gaussian Sum Filter has been introduced. It has been shown that the proposed algorithm allows to build tracks with high efficiency and that the track hits can be in this way collected up to the outermost tracker layers without introducing a noticeable fake rate. All hits being collected, it is then possible to optimize the track parameter estimation. The use of the mode of the gaussian mixture rather than the weighted mean has been shown to give a better precision for the electron tracks emitting only a small momentum fraction through bremsstrahlung. The momentum resolution which is obtained is comparable to the one obtained using the default CMS procedure. The proper treatment of the non gaussian radiation losses was also shown to provide a meaningful estimate of the momentum at the outermost hit, giving in turn both an estimate of the bremsstrahlung fraction using the tracker only and the possibility to improve the matching between the tracker and the calorimeter. Finally, an electron identification strategy based on a likelihood method has been shown to give both good efficiency on real electrons and high rejection power on fake ones.

Chapter 6

Study of the decay channel $\mathbf{H \rightarrow WW^{(*)} \rightarrow e^+e^-\nu_e\overline{\nu_e}}$

It was discussed in the first chapter how for Higgs boson masses close to 170 GeV/c^2 the significance for the channel $H \rightarrow ZZ^{(*)} \rightarrow 4l$ is reduced due to the suppression of the ZZ^* branching ratio as the WW decay mode opens up. In this mass region, the $H \rightarrow WW$ decay channel is the dominant one.
In this chapter, the discovery potential of the CMS detector in this channel is investigated, using a method based on that suggested in [93]. The presence of the two neutrinos in the final state makes the reconstruction of a mass peak not possible. An excess of events may be anyway observed and then used to identify the presence of a Higgs boson signal and to extract information on its mass.
The study presented here is based on the full simulation of the CMS detector. The two electrons in the final state make this analysis a good test for the algorithms which have been introduced in the previous chapter.
After an introduction on the signal and the main backgrounds, the event reconstruction and selection are discussed. The significance which can be obtained for different Higgs masses is presented. Finally, some hints to the possibility to extract information on the Higgs mass are given.

6.1 The signal and the main backgrounds

In this section an overview of the signal and of the principal backgrounds is given. The cross sections and the main characteristics of the processes considered in the analysis are discussed.

6.1.1 The signal

As subject of this thesis, the channel $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$ has been chosen. The diagram for the Higgs boson production via gluon fusion and the following decay into two W bosons is shown in figure 6.1. The final state of this process is characterized by the presence of two isolated electrons, which bring a clear and

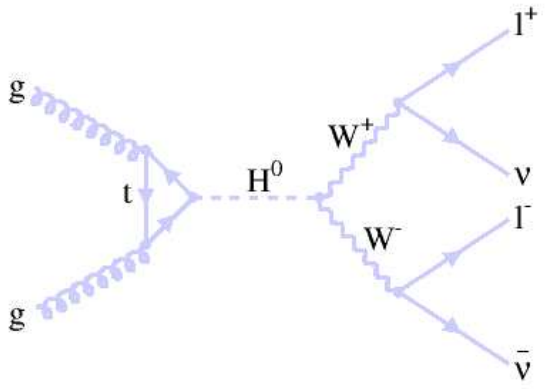

Figure 6.1: *Diagram for the signal process. The Higgs boson is produced via gluon fusion and it decays into two W bosons, which then originate 2 electrons and 2 neutrinos.*

detectable signal in the detector, and two neutrinos, which are responsible for a large missing energy. No hadron activity is present.
The Higgs decay into two W bosons has been computed at leading order in ref [94], then the result at one loop was given in ref [95]. The NLO cross section values for the main Higgs production mechanisms are given for different Higgs masses at the center of mass energy $\sqrt{s}$ = 14 TeV in table 6.1, together with the branching ratios for the H$\rightarrow$WW$^{(*)}$ decay. The cross sections have been computed with M. Spira's program [96] using the CTEQ6M structure functions and assuming m_{top} = 175 GeV/c^2; the branching ratios have been obtained with HDECAY [97]. Only the Higgs production through gluon-gluon fusion and vector boson fusion are implemented in PYTHIA and have been considered in the analysis described here.
The two last columns of table 6.1 show the total cross sections times the branching ratios for the W decay into electrons and the number of events which are expected at LHC with an integrated luminosity of 10 fb^{-1}. The latter varies between $\sim$700 and $\sim$2800 in the considered range of masses without any acceptance cut. The numbers in table 6.1 clearly show the importance of this decay channel in the region of masses between $2m_W$ and $2m_Z$, where both the W bosons can be created

Mass (GeV/c^2)	$\sigma(gg \to H)$ (pb)	$\sigma(VV \to H)$ (pb)	$\sigma(q\overline{q} \to VH)$ (pb)	$\sigma(q\overline{q}, gg \to Hq\overline{q})$ (pb)
m = 120	36.5	4.47	2.66	0.67
m = 130	31.7	4.14	2.07	0.53
m = 140	27.8	3.83	1.63	0.43
m = 150	24.6	3.56	1.30	0.35
m = 160	21.9	3.32	1.05	0.29
m = 170	19.7	3.09	0.85	0.24

	BR ($H \to WW^{(*)}$)	σ_{tot} (pb) × BR ($H \to 2e2\nu$)	Events (for $\int L$ = 10 fb^{-1})	
	0.1331	0.068	680	
	0.2888	0.127	1270	
	0.4854	0.188	1880	
	0.6832	0.234	2340	
	0.9016	0.275	2750	
	0.9654	0.265	2650	

Table 6.1: *Higgs production NLO cross sections and $H \to WW^{(*)}$ branching ratios for the Higgs masses used in this analysis. The number of events which are expected for an integrated luminosity of 10 fb^{-1} without any preselection is indicated in the last column.*

on-shell but the two Z bosons can not. Despite the lower cross sections, the low mass region could be usefully exploited. In this mass region, indeed, the Higgs boson is expected to be discovered mainly via its ZZ^* or $\gamma\gamma$ decay modes, but the WW^* decay mode remains also interesting for its complementarity, as the Higgs couplings are accessible through the $H \to ZZ^*$ and $H \to WW^*$ cross sections ratio, in which common systematics cancel out. In this study, we have concentrated on the mass region between 120 GeV/c^2 and 170 GeV/c^2.

6.1.2 The backgrounds

The branching ratio for the H→WW$^{(*)}$ decay is nearly unity for Higgs boson masses above the real W pair production threshold and below the real Z pair one, then it quickly decreases. In this discovery region, therefore, the main difficulty with detecting the Higgs boson in this channel is not the production rate, but the large backgrounds. Since the signal is purely leptonic, the irreducible backgrounds are of electroweak strength and they are therefore less severe that the

QCD ones. However the information which is carried away from the two neutrinos makes the reconstruction of the Higgs mass from the invariant mass of the decay products impossible. This is therefore a counting experiment and a good background control together with a high signal to background ratio is needed.
All the processes with two electrons plus missing energy in the final state have to be considered as possible sources of background. The most important ones are the continuum WW production and two backgrounds involving the top quark, the $t\bar{t}$ pair production and the single top production Wtb. They are described in detail in the following.
Some hints to the other processes which can mimic the signal final state, but which have not been considered here, are also given. Among them, the W+jets background is analysed as an example of process in which only one electron is present in the final state and the other one comes from a misidentified hadron.

WW production

The WW production represents the dominant irreducible background for the channel under study. Both the final and intermediate states are very similar to those of the signal, so care has to be taken to the search for the proper kinematical selections to reduce it without too largely affecting the signal.
At hadron colliders the $WW^{(*)}$ system can be produced through the processes $q\overline{q} \rightarrow$WW$^{(*)}$ and $gg \rightarrow WW^{(*)}$, from quarks or gluons in the initial state. The Feynman diagrams at leading order are shown in figure 6.2. Only the $q\overline{q} \rightarrow$WW$^{(*)}$

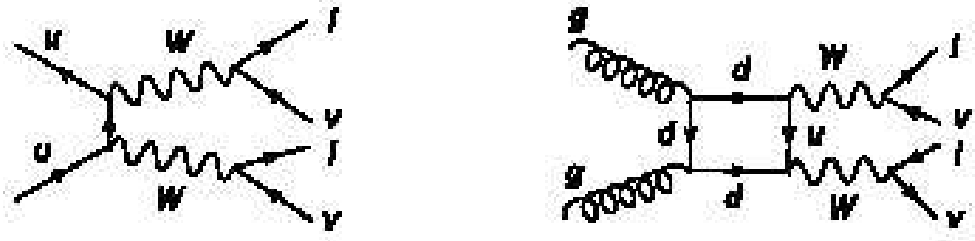

Figure 6.2: *Tree level diagrams for the $q\overline{q} \rightarrow WW^{(*)}$ and the $gg \rightarrow WW^{(*)}$ processes with two charged leptons and two neutrinos in the final state.*

process is implemented in PYTHIA and has been generated. The NLO calculations [98][99] give a K-factor around 1.6 at LHC with a larger effect of higher order corrections if the Ws have high transverse momentum. In this analysis the value σ^{NLO} = 114.3 pb has been used; it was computed with MCFM [100] using

the CTEQ6M probability functions.
The process $gg \rightarrow$WW$^{(*)}$ has not been included since it is not yet implemented in PYTHIA. The LO computations for such process have become available only recently [101][102] and higher order corrections are presently unknown. The process contributes at $O(\alpha_S^2)$ relative to $q\overline{q}$ annihilation, but its importance is enhanced by the large gluon flux at LHC. The contribution provides a 5% correction to the inclusive W-pair production cross section at LHC. After the introduction of some of the cuts which enter this analysis, anyway, the gluon fusion process becomes significative and it increases the theoretical WW background by a factor around 30%. Its introduction is therefore necessary for the further development of this study.

$t\bar{t}$ production

The production of $t\bar{t}$ pairs is one of the dominant reducible backgrounds. It is separable from the signal due to the presence of additional particles in the final state, however the very large cross section can make it difficult to suppress it efficiently. $t\bar{t}$ pairs are produced at hadron colliders via the gluon fusion process $gg \rightarrow t\bar{t}$ or via the quark annihilation $q\overline{q} \rightarrow t\bar{t}$. The Feynman diagrams at LO for the two processes are shown in figure 6.3. The NLO cross section has been com-

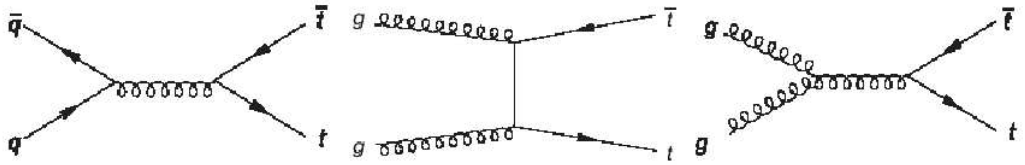

Figure 6.3: *Tree level diagrams for the $t\bar{t}$ production.*

puted [103] and the results for LHC are given in [104]. For this analysis the value $\sigma^{NLO} = 840$ pb has been used, which was computed with the MRST PDF set [105]. The scale uncertainty is $\pm$5% and the PDF uncertainty is $\pm$3%. No up to date number for the cross section at NLO+NNL is currently available with the most recent PDFs.
The top quark decays almost exclusively ($\sim$ 98%) into a W boson and a b quark. To speed up the MonteCarlo generation, the two Ws coming from the tops have been forced to decay into electrons. In most of cases therefore the two reconstructed electrons come from the W and the contribution of the electrons coming from the semileptonic decays of the B hadrons (which have a softer p_T spectrum and which are less isolated) is secondary.

Wtb production

At LHC top quarks are mainly produced in pairs, but there is also a significative number of top quarks which are produced singly via weak interactions. The process of interest for this analysis is $bg \rightarrow tW$, which is shown in fig 6.4, where the top quark is produced in association with a real W boson and one of the initial partons is a b quark in the proton sea.

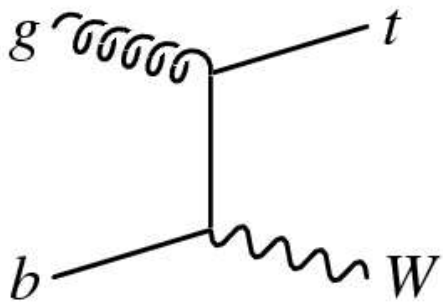

Figure 6.4: *Tree level diagram for the $bg \rightarrow tW$ process.*

The total cross section for the process has been computed at LO [107] and only very recently [108] the full NLO computation has been made available, with the inclusion of the full spin correlations. An important feature of the single top production in association with a W boson is the fact that one of the diagrams which contribute to the initial gluon correction ($gg \rightarrow Wt\overline{b}$) corresponds to the $t\overline{t}$ pair production at leading-order ($gg \rightarrow t\overline{t}$) followed by the decay $\overline{t} \rightarrow W\overline{b}$. To disentangle these two processes it is necessary to subtract the contributions in which the $\overline{t}$ is on-shell. Different procedures are available to do that; in [108] a method which exploits a b-jet veto is introduced.
The NLO cross section, computed with MCFM, is $\sigma^{NLO} = 33.1$ pb. The main theoretical uncertainties are related to the choice of the PDF set ($\sim$ 5%) and to the renormalization and factorization scales ($\sim$ 10%). The analysis of the channel $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$ performed in [108] shows how the effect of NLO corrections is strongly reduced when a jet veto is applied as in this analysis.

Table 6.2 summarizes the total cross sections and the branching ratios for the three considered backgrounds. The expected number of events for an integrated luminosity of 10 fb^{-1} is also shown. From a comparison with the numbers in table 6.1, it is clear that a strong reduction factor has to be achieved with properly tuned cuts. This will be discussed in the following.

Process	σ_{tot} (pb)	$\sigma_{tot} \times BR\,(W \to e\nu)$ (pb)	Events (for $\int L$ = 10 fb^{-1})
WW	114.3	1.3	13135
$t\bar{t}$	840	9.7	96531
Wtb	33.1	0.4	3804

Table 6.2: *Main backgrounds production NLO cross sections. The number of events which are expected for an integrated luminosity of 10 fb^{-1} without any preselection requirement is indicated in the last column.*

Other possible background sources

Beyond the main backgrounds which have been discussed in the previous paragraphs, there are other possible background sources which have not been taken into account in detail yet, mainly for pratical reasons. In most of cases, some cuts which are expected to reduce them have been introduced in the analysis, despite the absence of a detailed simulation. In some cases such backgrounds have been checked to be of minor importance with a fast simulation of the detector or with preliminary studies at the generator level.
Vector bosons productions like ZZ or WZ are potential background sources which are expected not to give a significative contribution to the final result. The background from Z boson pair production for instance is expected to be strongly reduced by rejecting events in which the invariant mass of the electron pair is equal to m_Z. In a small fraction of events the Z boson which decays into 2 electrons is off-shell and it is therefore not rejected by the above mentioned cut, however these events are largely suppressed by the Z boson propagator.
The ordinary Drell-Yan production of lepton pairs is not accompanied by missing energy, so it should be possible to reduce it by imposing a cut on the minimum missing transverse momentum.
$Wb\bar{b}$ and direct $b\bar{b}$ production containing one or two electrons from semileptonic b-decays are important sources of reducible background. Due to the large cross sections, a huge statistics would be necessary to deal with such backgrounds, which have not been included in the analysis yet. They have been shown anyway to be strongly reduced by lepton identification and isolation requirements in the study of this same channel done for the $\mu\mu$ final state within the CMS collaboration [109].
Finally processes with only one electron could also contribute in presence of an additional electron coming from other sources. An example is given by the W+jets background [110], where one jet is misidentified as an electron. Again, due to the high cross section, a very large statistics would be necessary to perform a detailed analysis on such background. This was not possible due to data-

accessing problems, anyway some hints will be given in the following.

6.2 Event generation

In this study, both the signal and the backgrounds have been fully simulated using the production chain presented in the previous chapter.
The datasets from the official CMS production have been used. The event generation was done with PYTHIA for the signal and the WW production. TopReX [106], which takes into account the spin correlations in the top quark decay, has been used to generate the $t\bar{t}$ and the Wtb samples, then PYTHIA was used for the hadronization. The programs used for the generation and the number of generated events for each sample is shown in table 6.3. In all cases, the W boson was forced

Channel	Dataset	Event generation	Generated events
signal (different masses)	hg03b_hww_2l_120	PYTHIA	2003
	hg03b_hww_2l_130		2110
	hg03b_hww_2l_140		2116
	hg03b_hww_2l_150		1845
	hg03b_hww_2l_160		5264
	hg03b_hww_2l_170		5024
$qq \rightarrow WW$	hg03_ww_2l	PYTHIA	20158
$t\bar{t}$	eg03_tt_2l_topr	TOPREX	21575
Wtb	eg03_wt_2l_toprex	TOPREX	19617

Table 6.3: *Datasets used for the analysis and number of used fully simulated events.*

to decay into electrons only. No preselection has been applied at the generator level.
All the data have been simulated with CMSIM133. The low luminosity phase only has been considered, therefore all the samples used for the analysis have been digitised superimposing the pile-up which is expected at low luminosity, which is 3.5 events per bunch crossing on average.

6.2.1 The reweighting technique

Higher order corrections can not be neglected at LHC, since they have a large influence on both the cross sections and on the event kinematics.
The data used in this analysis have been generated with PYTHIA, which computes the cross sections at leading order and accounts for higher order contributions only

through initial and final state QED and QCD showering processes. The scaling of the LO cross sections to the NLO ones with the use of the theoretically computed K-factors introduced in chapter one is justified only if the corresponding kinematical variables spectra agree. In the study of the $H \to WW^{(*)}$ channel, some cuts have been developed in order to improve the signal over background ratio, and it will be discussed in the following that these cuts are particularly sensitive when the Higgs boson is produced with small transverse momentum. In particular, a jet veto is applied to reduce the size of the $t\bar{t}$ background. Since the effect of higher order QCD corrections is reduced if a jet veto is applied [112], a simple K-factor can not be used to rescale the cross sections.
For this analysis, the reweighting procedure introduced in [113] has been applied. An effective experimental K-factor can be computed by comparing the distributions obtained with a LO and a NLO MonteCarlo for a kinematical variable **X**. With an integration over the whole possible range of **X**, the number of events for a given integrated luminosity L at NLO is given by

$$\frac{N}{L} = \int \frac{d\sigma^{NLO}}{d\mathbf{X}} d\mathbf{X} = \int K(\mathbf{X}) \frac{d\sigma^{LO}}{d\mathbf{X}} d\mathbf{X} \tag{6.1}$$

and the **X**-dependent K-factor is defined as

$$K(\mathbf{X}) = \left(\frac{d\sigma^{NLO}(\mathbf{X})}{d\mathbf{X}} \right) / \left(\frac{d\sigma^{LO}(\mathbf{X})}{d\mathbf{X}} \right) \tag{6.2}$$

Such K-factor can be used to correct the spectrum which is generated with a LO MonteCarlo, by reweighting each event in such a way that the new spectrum for the variable **X** matches the spectrum obtained with the correct QCD calculations. The K-factors used in this analysis have been computed using PYTHIA as LO MonteCarlo and MC@NLO [114] for the NLO computations [1]. The latter includes NLO corrections and spin correlations and it performs the showering with HERWIG [115].
The reweighting technique has been applied to the signal and the WW background. In the signal case, the procedure has been applied only to the events in which the Higgs boson was produced via gluon fusion and the Higgs p_T has been used as **X** variable. The comparison between the Higgs p_T spectrum before and after the reweighting is given in figure 6.5 on the top (m_H = 160 GeV/c^2), which shows how PYTHIA provides a softer Higgs spectrum with respect to the NLO computations. For what concerns the WW background, instead, the WW system p_T has been used as variable for the reweighting procedure. The comparison between the LO and NLO shape is given in figure 6.5 (bottom). Again, the LO distribution is softer and differences arise also for high values of $p_T(WW)$.

[1] A special thank to G.Davatz and V.Drollinger who provided the MC@NLO distributions

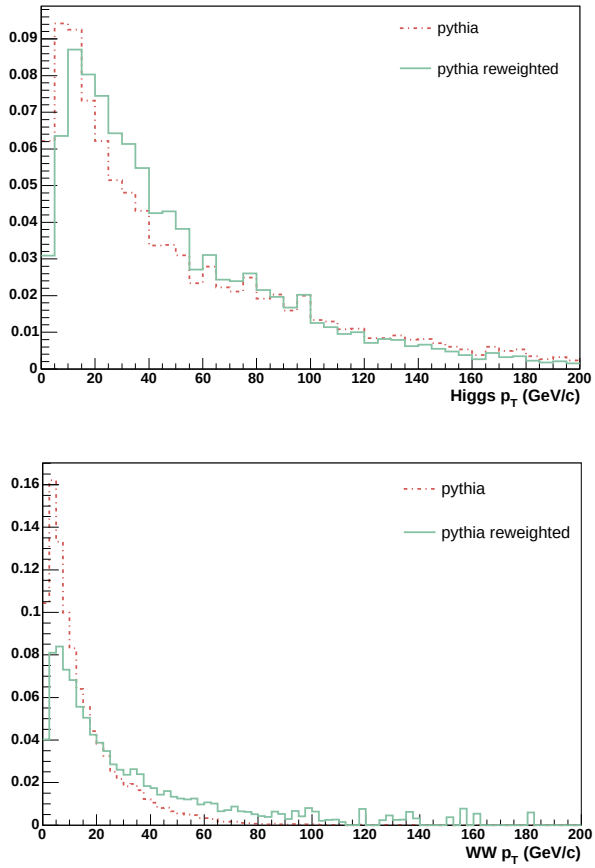

Figure 6.5: *Comparison between the PYTHIA distributions before (red, dash-dotted histogram) and after (green, solid histogram) the reweighting. On the top: Higgs p_T spectrum (m_H = 160 GeV/c^2). On the bottom: WW p_T spectrum.*

For the signal produced through vector boson fusion, the $t\bar{t}$ and the Wtb background, the PYTHIA spectra have been used without any reweighting and the corresponding cross sections have been normalized using constant K-factors.

6.3 Event reconstruction and preselections

Chapter five was devoted to the electron reconstruction, which plays a central role in the analysis which is presented here. Other important issues for the channel under study are the reconstruction of jets and missing energy, which are described here together with the results of the electron reconstruction when applied to this

channel. A set of preselections to select events with two isolated electrons within the detector acceptance is also discussed.

6.3.1 Electron reconstruction

The electron reconstruction is done in two steps, online and offline. Since the final state of the $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$ channel is characterized by the presence of two isolated electrons, the trigger tables which have been used here to select the events are the single electron trigger and the double electron trigger. The trigger represents therefore the first step in the electron reconstruction procedure. The electrons in the events which pass the Level1 Trigger and the High Level Trigger have been then reconstructed as it was discussed in the previous chapter.

Online electron reconstruction

The CMS HLT electron reconstruction follows the steps which are described in paragraph 5.2. At the first step, 'Level-2.0', the energy in the calorimeter is clusterized. 'Level-2.5' corresponds to the matching between the supercluster and the track seed, then at 'Level-3.0' the full track reconstruction is performed. To be accepted by the single electron trigger, the event has to contain one isolated electron with p_T >26 GeV/c, while the threshold for the double electron trigger is p_T = 14.5 GeV/c for both the two isolated electrons which are required in the final state.
In this analysis, the HLT response has been defined as the logical OR of the two mentioned trigger responses. Figure 6.6 gives the HLT efficiencies for the signal at the different masses and for the main backgrounds; the separate contributions of the single and double trigger are also shown.

For what concerns the signal, a drop in the efficiency is evident for low Higgs masses, which correspond to lower p_T electrons. It is worth noticing that most of the events pass the single electron trigger and that the contribution of events which pass only the double electron trigger is marginal. This means that the possible introduction of asymmetric p_T thresholds for the two electrons could not help in increasing the trigger efficiency, notwithstanding the favourable kinematics of the signal events, which often present a big imbalance between the p_T of the two electrons in the final state.

Offline electron reconstruction and preselections

In the offline electron reconstruction procedure, the tools which have been introduced in the previous chapter have been employed. Electrons have been built asking for a supercluster and a track matching within $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.15$,

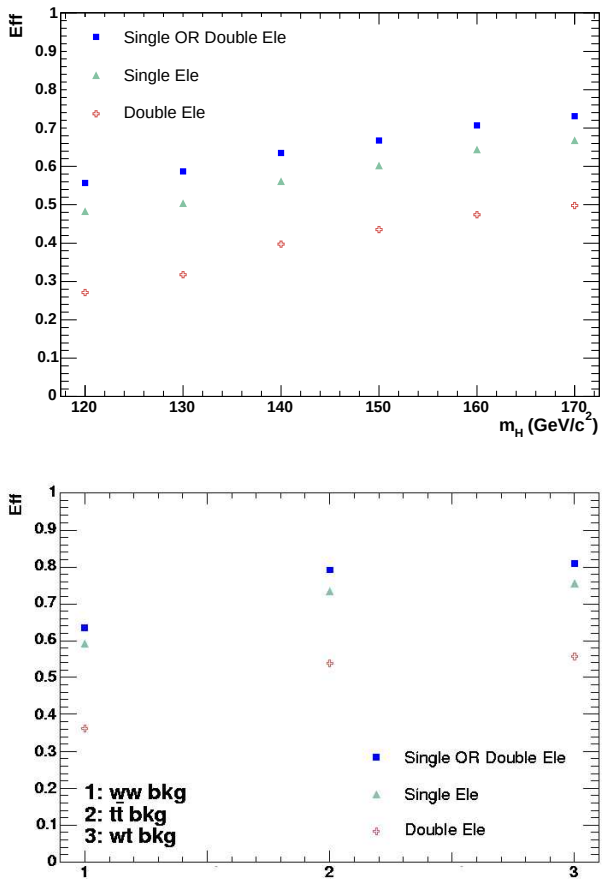

Figure 6.6: *Electron HLT efficiency for the signal (on the top) and the backgrounds (on the bottom). Together with the global response, the separate contributions of the single electron trigger and of the double electron trigger are shown.*

with the eleID variable higher than a certain threshold. If more than two electrons were reconstructed, the 2 ones with the highest transverse momentum have been kept for the analysis.
Figure 6.7 shows the reconstruction efficiency for at least one e^+e^- pair in the final state for the signal and the main backgrounds for different hardnesses of the cut on the electron identification variable. The efficiency is defined with respect to those events which pass the HLT and no detector acceptance restriction is imposed.
In the following, the cut on eleId is put to 0.98. Such high value guarantees the high purity which is necessary to keep the backgrounds coming from jet misiden-

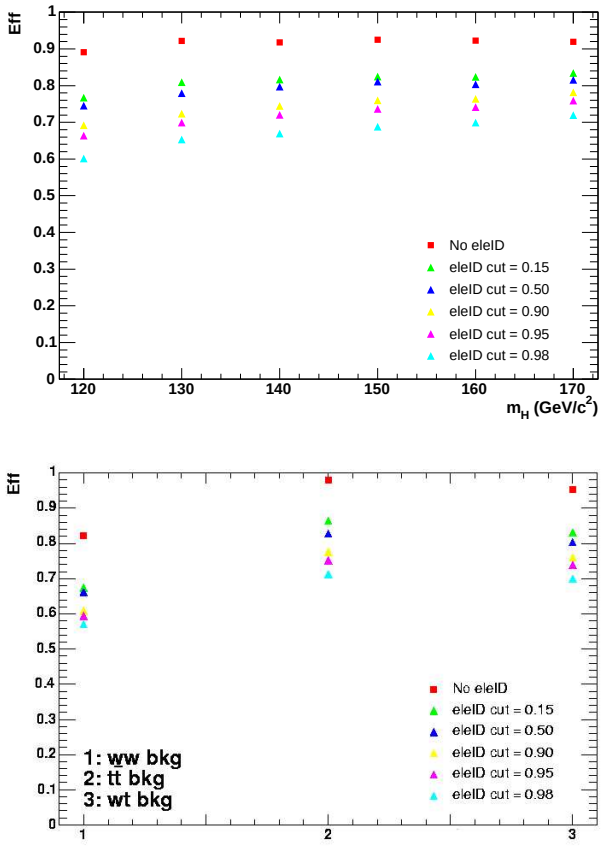

Figure 6.7: *Efficiency of finding at least one e^+e^- pair for the signal (on the top) and the backgrounds (on the bottom) with respect to the events passing the HLT. The efficiency is shown for different cuts on the electron identification variable.*

tification under control. In this way, the efficiency on the signal is lower but the signal over background S/B ratio is not reduced. With this cut, the fraction of events in which the two electrons are not correctly identified (ie the reconstructed electrons do not match those from the Higgs boson decay) is about 2%. Among them, about one half comes from events in which more than 1 e^+e^- pair has been reconstructed and the highest p_T criterion has been applied to select the two leptons for the analysis.

The final state of channel $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$ is characterized by two isolated electrons with high transverse momentum. The starting point for the analysis are the events passing the HLT in which at least two electrons are reconstructed.

Since relatively high energetic electrons are involved, a p_T threshold at 10 GeV/c has been set. Besides, the electrons have been requested to fall within the pseudorapidity region covered by both ECAL and the tracker and therefore a cut at $|\eta|$ = 2.5 has been imposed. With these preselections the number of events to be analyzed is reduced without loosing efficiency on the significative events.
In figure 6.8 the comparison between the reconstructed and the generated track parameters at vertex are shown for the electrons which are reconstructed in the events passing these criteria. A good momentum estimation is crucial for this analysis, since mismeasured tails in the distribution could badly affect the S/B ratio. Both the pseudorapidity (left) and ϕ (right) are well reconstructed and the resolution is comparable to that obtained on the single electrons study sample.

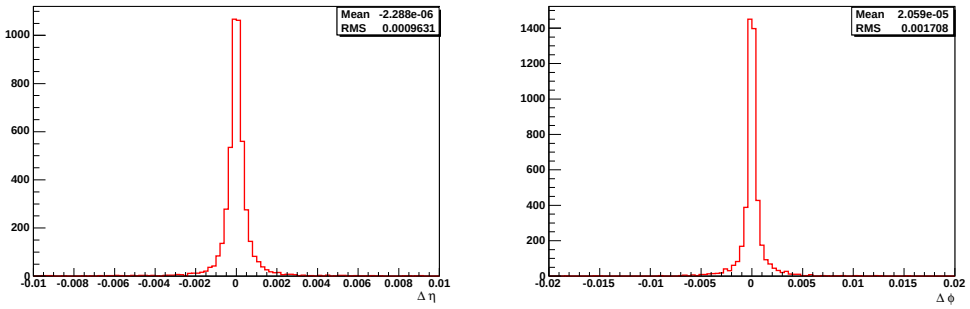

Figure 6.8: *Comparison between the reconstructed and the generated electrons in the signal sample with m_H = 160 GeV/c^2. On the left: $\Delta\eta$ at the vertex. On the right: $\Delta\phi$ at the vertex.*

Finally, an isolation criterion has been defined, since the presence of two isolated electrons is one of the main characteristics of the signal. Together with the use of the electron identification variable, the electron isolation is useful to reject fake electrons which are part of jets. Besides, it plays an important role against the backgrounds from the top quark, which have many particles in the final state.
The tracker information has been used to define the electron isolation. The sum of the transverse momenta of the tracks falling within a cone of fixed ΔR= $\sqrt{\Delta\phi^2 + \Delta\eta^2}$ around the electron has been considered. To avoid the inclusion of the track associated with the electron itself in the sum, a veto cone with ΔR = 0.015 around the electron has been defined. As it will be discussed in the next paragraphs, in a large fraction of the signal events the two signal electrons tend to be collinear, therefore the two cones built around them can partially overlap. To avoid double countings of tracks and the inclusion of one electron in the cone built

around the other one, the approach proposed in [92] has been followed. An isolation region has been defined as the sum of the cones around the two electrons and the sum of the transverse momenta of the tracks falling in such region has been considered. Only the tracks coming from the signal vertex have been taken into account with the request $\mid z_{track} - z_{vertex} \mid < 0.4$ and $\mid d_{track} - d_{vertex} \mid < 0.1$, being d the transverse impact parameter.
In figure 6.9 the efficiency on the signal of the isolation cut is shown for different values of the cone radius as a function of the threshold, which is expressed using

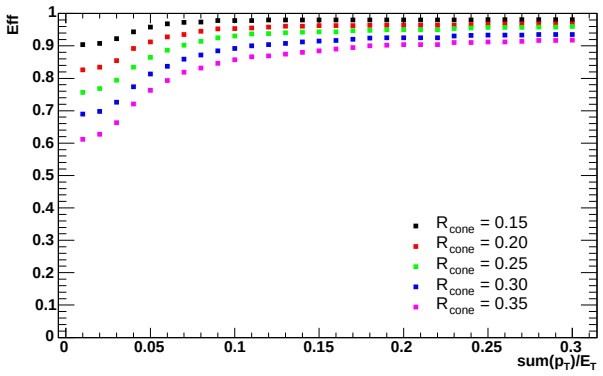

Figure 6.9: *Efficiency of the isolation cut on the signal ($m_H = 120$ GeV/c^2) for different cone sizes as a function of the threshold. The efficiency is defined with respect to the events with two reconstructed electrons passing the HLT.*

the ratio between the sum of the tracks p_T and the mean value of the two electrons transverse energy, $\sum p_T^{tr} / E_T^{mean}$. The efficiency is defined with respect to the events which pass the HLT and with two reconstructed electrons.
Unfortunately the available background samples do not allow to study in detail the problem since the W was forced to decay leptonically and the two electrons are therefore in most of cases isolated. In the case of $t\bar{t}$ production, the isolation cut efficiency is only a few percents lower with respect to the signal.
For what concerns the cone size, small radii have to be preferred. ΔR = 0.15 and ΔR = 0.20 give similar results and the latter was at the end chosen. The threshold has been set looking at the signal over background ratio in the region of high efficiency on the signal. The value $\sum p_T^{tr} / E_T^{mean} = 0.1$ has been used for the analysis.

A summary of these preselection efficiencies with respect to the triggered events is given for the signal and the main backgrounds in table 6.4.

	m_H = 120 GeV/c^2	m_H = 130 GeV/c^2	m_H = 140 GeV/c^2
at least 1e^+e^-pair	60.2	65.3	67.0
isolation	95.2	94.7	95.1
p_T >10 GeV/c, $\mid \eta \mid$<2.5	90.6	93.4	97.0
total efficiency (%)	51.94±1.14	57.76±1.08	61.82±1.00
	m_H = 150 GeV/c^2	m_H = 160 GeV/c^2	m_H = 170 GeV/c^2
at least 1e^+e^- pair	68.8	70.0	72.0
isolation	95.9	94.9	94.3
p_T >10 GeV/c, $\mid \eta \mid$<2.5	97.1	97.9	98.2
total efficiency (%)	64.07±1.04	64.96±0.59	66.66±0.59
	WW	$t\overline{t}$	Wtb
at least 1e^+e^- pair	57.1	71.1	70.5
isolation	98.5	93.1	96.0
p_T >10 GeV/c, $\mid \eta \mid$<2.5	97.8	97.6	97.8
total efficiency (%)	55.02±0.11	64.59±0.07	66.21±0.17

Table 6.4: *Efficiencies of the preselections. Each number refer to the events passing the previous cut. The total efficiency is given with respect to the triggered events. The quoted errors include the statistical uncertainty only.*

6.3.2 Jet reconstruction

Different algorithms are available in ORCA to reconstruct jets and different choices can be done for the jet constituents. For this analysis, the Iterative Cone Algorithm [116] has been chosen and the energy of the calorimetric towers has been used as input to create the jets.
The Iterative Cone Algorithm searches for the maximum transverse energy 'object' (a tower in ECAL or HCAL with energy above a certain threshold, in this case) and throws a $\eta - \phi$ cone around its direction. Any tower within the cone is merged to form a proto-jet. The direction of the proto-jet is computed from the energy weighted directions of the constituents, then a new cone is thrown around the new direction. The procedure is iterated until both the jet energy and direction remain stable between two following steps, then the jet is created and the constituents are removed from the 'objects' list. This is done until all the towers with energy above the threshold are used.
For this analysis, the cone size has been set to 0.5. The calorimeter towers with $E_T > 0.5$ GeV and $E > 0.8$ GeV have been used as jets constituents and reconstructed jets have been required to have $E_T > 10$ GeV. Since the jets enter the analysis only through a jet veto to reject the top backgrounds, a calibration is not strictly needed, provided the veto is tuned on the raw energy. Due to some non-

linearities between the raw and the calibrated jet energy for E_T^{jet} <25 GeV, it was decided to use not calibrated jets.
Figure 6.10 (signal, m_H = 160 GeV/c^2) shows the η distribution of the reconstructed jets and the fraction of jets which match a generated one as a function of the jet transverse energy. Generated jets are built at the generator level using the same algorithm as in the reconstructed case, but without the detector simulation. A large number of fake jets is reconstructed both at large pseudorapidity and at low transverse energy [117]. To get rid of the large number of fake jets seeded

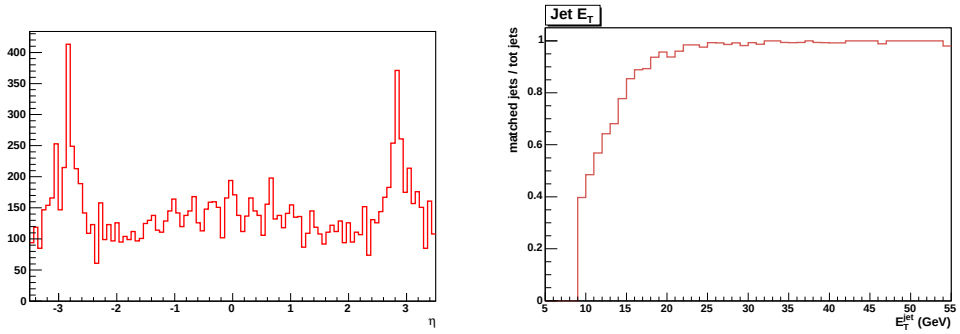

Figure 6.10: *Jet reconstruction in signal (m_H = 160 GeV/c^2) events. Left) η distribution. Right) Fraction of reconstructed jets matching a generated one as a function of the jet transverse energy.*

by noise at high pseudorapidity, only jets with $\mid \eta \mid < 2.5$ have been considered. The information from the tracker was instead used to deal with the low transverse energy jets, in most of cases coming from minimum bias events. To determine the jets from the signal vertex, the sum of the p_T of the charged tracks within ΔR = 0.5 around the jet axis and having the same vertex as the signal electrons was computed for each jet and the variable $\alpha = \frac{\sum p_T^{tracks}}{E_T^{jet}}$ was defined [118]. For a perfect detector, α should be around 0.66, since about 2/3 of a jet is made by charged particles. In a fake jet, underlying events contain many low p_T particles which are not seen by the tracker, giving to α a value around zero. The parameter can therefore be used to discriminate between real and fake jets.

6.3.3 Missing energy reconstruction

As for the jet reconstruction, many different algorithms are available in ORCA to reconstruct the missing transverse energy (MET). In this analysis, the MET is reconstructed by adding vectorially the transverse energies in each calorimetric cell and then correcting with the information from the muon chambers.

Figure 6.11 gives a comparison between the reconstructed MET and the transverse energy of the two neutrinos coming from the Ws (signal, m_H = 160 GeV/c^2). The

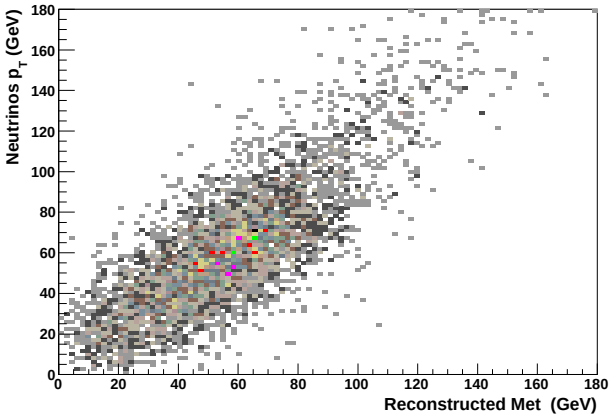

Figure 6.11: *Comparison between the reconstructed transverse missing energy and the transverse energy of the two neutrinos coming from the Ws (signal, m_H = 160 GeV/c^2).*

difficulty in the MET measure, especially at low energy, is one of the main experimental problems of the channel, leading to a significative associated uncertainty.

6.4 Event selection

The event topology of the decay channel $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$ is characterized by the presence of two isolated electrons and a significative missing energy due to the two undetected neutrinos. To discriminate between the signal and the several processes with the same final state, some selection cuts can be applied.
In the development of this analysis many variables have been considered. In the following a set of cuts is proposed to obtain both a good signal efficiency and a sufficient background suppression. Particular care was devoted to this second aspect, since the theoretical uncertainties on the backgrounds are quite large and therefore a good rejection is necessary to keep the systematics under control.
A systematic variation of cuts through a minimization procedure to further increase the signal significance was also tried out, but the limited available statistics made the procedure too strongly dependent on fluctuations. To get a good set of cuts, anyway, each cut was tried several times and the order of the cuts was also changed, to detect for possible correlations among the selection variables. The final set of selections was chosen looking at the signal over background value after

each cut.
Since some of the cuts are correlated among them, the distributions of the kinematical variables are often distorted by the cuts coming before the one under analysis. Despite this problem, each distribution is shown here after the preselections, since the application of all the cuts would reduce too strongly the available statistics, making the variables behaviour not clear.

6.4.1 Central jet activity

The $H \rightarrow WW^{(*)}$ signal is characterized by the absence of central jet activity. Even if this characteristics can be partially spoiled by the initial and final state radiations which can produce jets, a jet veto can be usefully exploited to reduce the backgrounds coming from the top quarks without affecting too largely the signal. Events are therefore selected only if no jet in the pseudorapidity region $\mid \eta \mid < 2.5$ has a transverse energy higher than a certain threshold.
Figure 6.12 shows the transverse momentum of the highest p_T jet for the signal

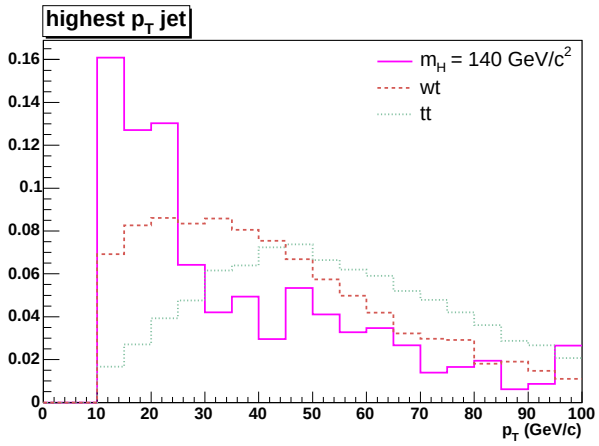

Figure 6.12: *Transverse momentum distribution for the highest p_T jet in the signal (m_H = 140 GeV/c^2, solid histogram), the $t\bar{t}$ (dashed hostogram) and the Wtb (dash-dotted histogram) background. The distribution is shown for events passing the preselections on electron reconstruction and isolation. For a better visibility the plots are normalized to the same area.*

in the case of m_H =140 GeV/c^2 and for the $t\bar{t}$ and Wtb productions. The signal behaviour depends only slightly on the Higgs mass and it is similar to that of the WW irreducible background. Top backgrounds are instead characterized by the presence of high p_T jets, coming from the b quark, which in average tend to be

more central than the jets from the signal. A central jet veto can be therefore usefully employed to reduce them.
The effect of the jet veto is presented in figure 6.13, where the efficiency on the

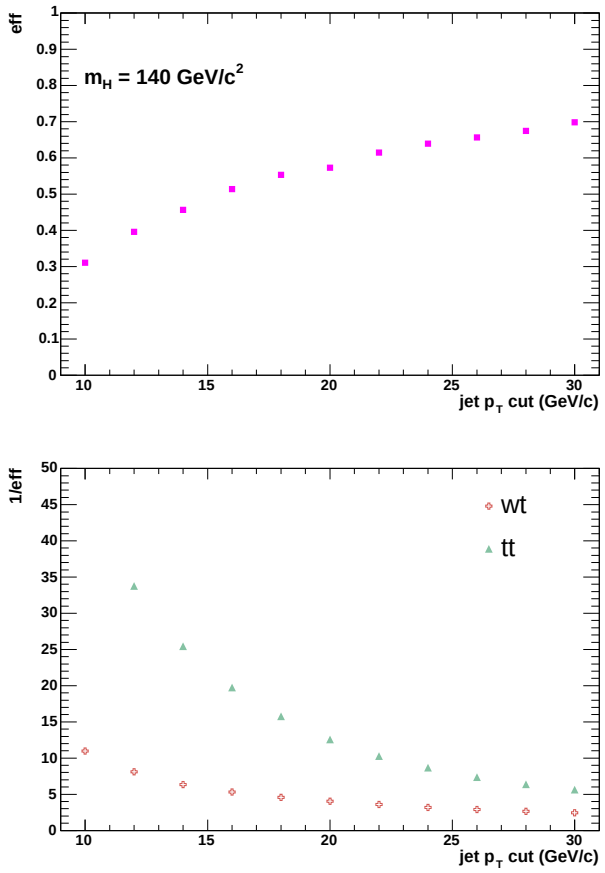

Figure 6.13: *Efficiency on the signal (top) and rejection factor on the backgrounds (bottom) as a function of p_T threshold for the jet veto. No α cut is imposed. The results are shown for events passing the preselections on electron reconstruction and isolation.*

signal and the rejection factors on the two top backgrounds are shown. The efficiency is defined with respect to those events which pass the preselections on electron reconstruction and isolation.
As a compromise between the necessity of selecting a large fraction of the signal and a having a good rejection power, a p_T cut at 20 GeV/c has been chosen. To further increase the efficacy of the jet veto and at the same time not to undergo an efficiency loss due to the fakes from minimum bias events, also the events with at

least one jet having p_T between 15 GeV/c and 20 GeV/c and $\alpha_{jet} > 0.2$ have been rejected. As already discussed, only jets within $\mid \eta \mid < 2.5$ have been considered. The jet veto acceptance, normalized to the preselection acceptance, for the signal samples and the three main backgrounds is given in tables 6.7 - 6.12. The power in the rejection of the $t\bar{t}$ background is clearly visible.

6.4.2 Angular separation between the two leptons

In the Standard Model the Higgs boson has spin 0 and the W boson has spin 1, therefore to conserve the angular momentum the spins of the W bosons produced in $H \rightarrow WW^{(*)}$ decay have to be anticorrelated.
We call [119] z the decay axis of the WW system in the Higgs rest frame and we consider the longitudinal (L) and transverse (T) polarisations with respect to such axis. In the Higgs rest frame, the conservation laws make only the decays $H \rightarrow W_T^+W_T^-$ and $H \rightarrow W_L^+W_L^-$ possible and they forbid the decay $H \rightarrow W_T^{\pm}W_L^{\mp}$. The decay rate of $W_T^+ \rightarrow e^+\nu$ is proportional to $(1+cos\theta)^2$ where θ is the angle between the positron direction and the W_T^+ spin, therefore the right handed positron is emitted in most of cases in the same direction as the W_T^+ spin. The opposite happens for the left handed electron, which is emitted in the opposite direction with respect to the W_T^- spin since its decay follows a $(1-cos\theta)^2$ distribution. Being the two Ws anti-correlated, the electrons are mostly emitted in the same direction. Similar considerations also apply to the case of longitudinally polarised Ws and in this case also the two electrons are mainly emitted in the same direction. A schematic picture is given in fig.6.14.
In the case of the WW production, the initial state is unpolarized, therefore the combinations $q\overline{q} \rightarrow W_T^+W_T^-$, $q\overline{q} \rightarrow W_L^+W_L^-$ and $q\overline{q} \rightarrow W_T^{\pm}W_L^{\mp}$ are all allowed. At $m(WW)$ = 165 GeV/c^2, the last case almost corresponds to half of the production rate. Since in this third case the two Ws spins are not anti-correlated, the directions of the two electrons are also not correlated.
A small opening angle between the two electrons is therefore a good discriminating variable to separate the signal from the background. It is mainly important to reduce the WW production, which is the irreducible background to deal with. The opening angle between the two electrons in the plane transverse to the beam is shown in figure 6.15 for the signal (m_H = 160 GeV/c^2) and the backgrounds for events passing the preselection cuts. The two electrons tend to be less collinear in case of low masses, since one of the two Ws is virtual and the quantum numbers are therefore not well defined.
The rejection factors on the backgrounds as a function of the opening angle cut and compared with the efficiency on the signal are shown in figure 6.16. The request $\Delta\phi < 100^\circ$ significatively suppresses the backgrounds and at the same keeps the efficiency on the signal at the 80% level.

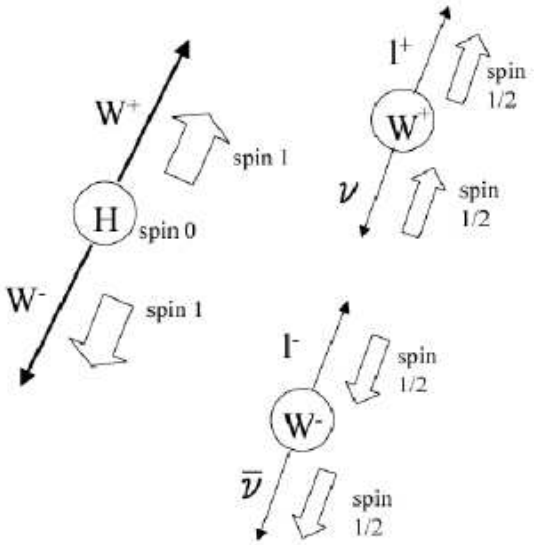

Figure 6.14: *Schematic picture of the spin correlations which characterize the $H \rightarrow WW^{(*)} \rightarrow 2l2\nu$ decay. The two leptons are mainly emitted in the same direction.*

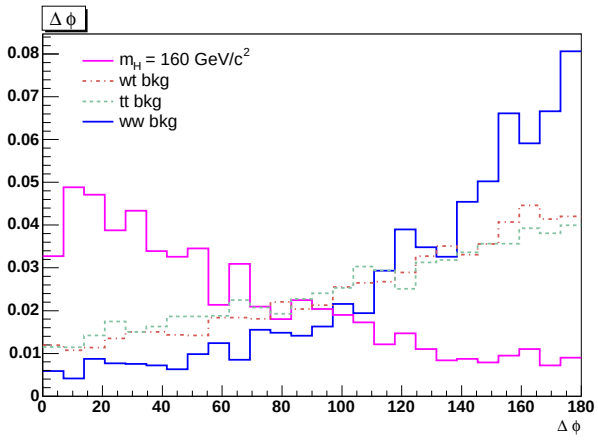

Figure 6.15: *Distribution of the azimuthal opening angle $\Delta\phi$ between the two electrons for the different channels. The distribution is shown for events passing the preselections on electron reconstruction and isolation. For a better visibility the plots are normalized to the same area.*

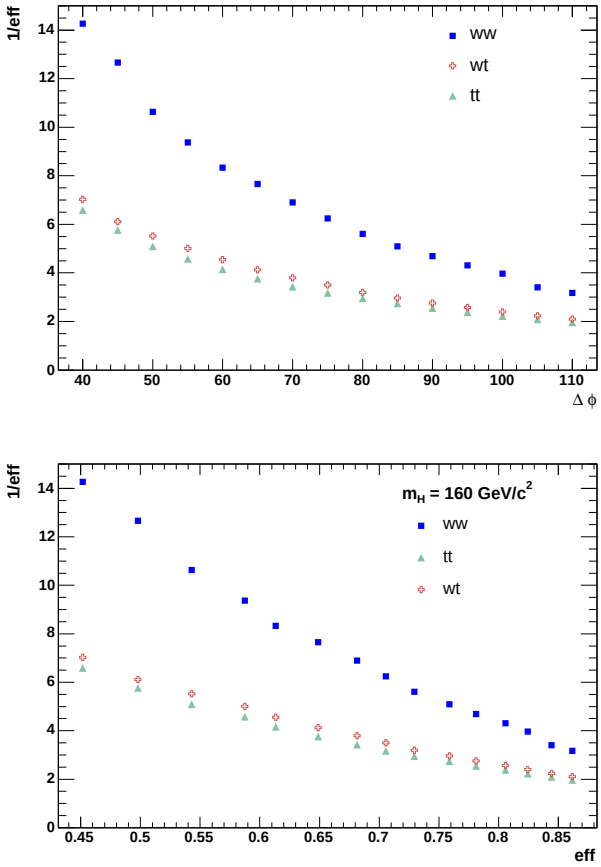

Figure 6.16: *Top) Rejection factor on the backgrounds as a function of the cut on the azimuthal opening angle between the two electrons. Bottom) Rejection factor on the backgrounds versus the efficiency on the signal (m_H = 160 GeV/c^2). The results are shown for events passing the preselections on electron reconstruction and isolation.*

6.4.3 Other selections

Electrons transverse momenta

The electrons coming from the Higgs boson have a quite different p_T spectrum with respect to those coming from the backgrounds.
Figure 6.17 shows the p_T distributions for the two signal electrons sorted in p_T decreasing order after the preselection. The distribution is only slightly dependent on the Higgs mass for the highest p_T electron, but a strong dependence can be observed for the second electron, especially for low masses. This is mainly due

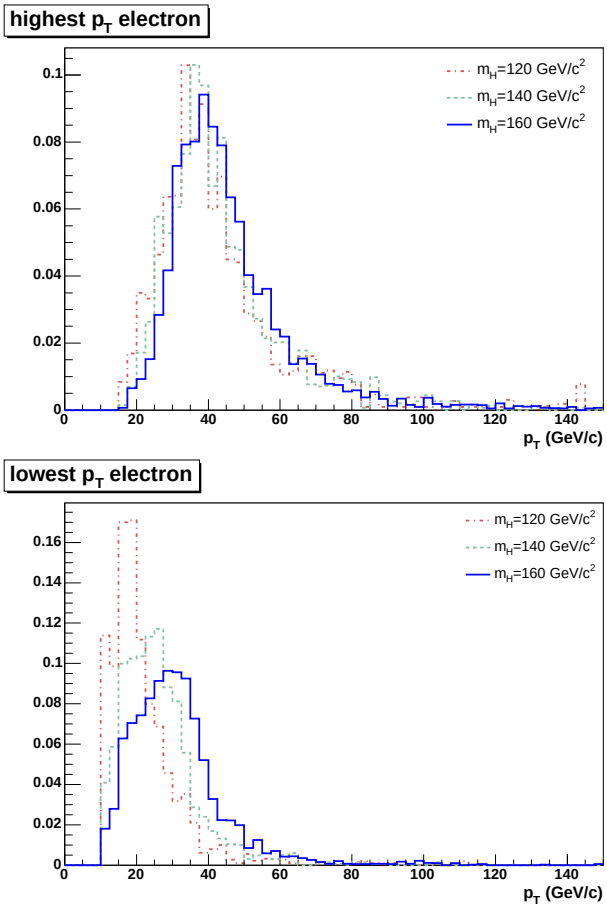

Figure 6.17: *Spectrum of the two electrons transverse momenta for the signal at different values of the Higgs masses. The distribution is shown for events passing the preselections on electron reconstruction and isolation. For a better visibility the plots are normalized to the same area.*

to the fact that for high Higgs masses both the 2 Ws tend to be produced on shell and the second electron can have a large transverse momentum as well.
For a comparison, figure 6.18 shows the same distributions for the three backgrounds. For what concerns the highest p_T electron, there is no particular difference between the low p_T part of the spectrum in the signal and in the background case. The background distributions anyway present a longer queue which can be usefully exploited to separate the signal from the backgrounds through a cut on the upper value. A cut on low p_T values for the second electron can be also exploited; it is particularly effective for high values of the Higgs mass, since for low values high rejection power can be obtained only with a large reduction of the

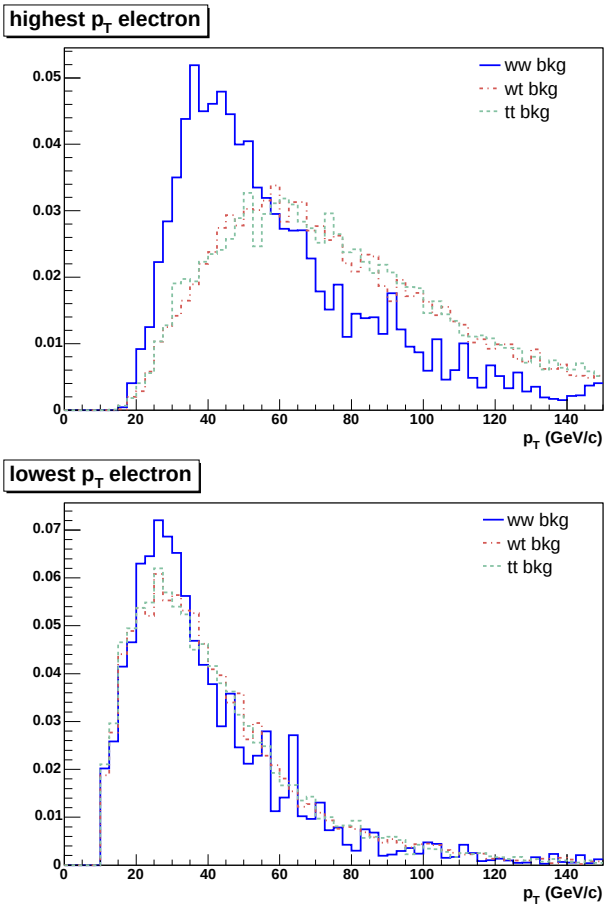

Figure 6.18: *Spectrum of the two electrons transverse momenta for the three backgrounds. The distribution is shown for events passing the preselections on electron reconstruction and isolation. For a better visibility, the plots are normalized to the same area.*

signal statistics.
The proposed selection cuts are summarized in table 6.5

Two electrons invariant mass

A source of background which has not been considered in this analysis is the ZZ production. To reduce it a cut on the two electrons invariant mass can be applied, by requiring $m(ee) < m_Z$.
The two electrons invariant mass distribution is shown in figure 6.19 for the signal and the backgrounds under study here. In the case of low Higgs mass the two Ws

m_H (GeV/c^2)	inf 1st (GeV/c)	sup 1st (GeV/c)	inf 2nd (GeV/c)
120	25	50	15
130	25	50	15
140	25	50	15
150	25	55	15
160	30	55	25
170	30	55	25

Table 6.5: *Values for the cuts on the two electrons transverse momentum.*

of the signal are produced close to the threshold and they are almost at rest in the Higgs frame. As a result, the neutrino and the electron coming from a single W are almost back-to-back in this frame and they have equal energies. This implies that the invariant masses of the two charged leptons and of the two neutrinos are approximately equal and they can not exceed half of the Higgs boson mass. Besides, the region of small dilepton invariant mass is also favoured by the small angle between the two charged electrons.
Due to all these reasons, the upper cut on the dielectron system invariant mass can be usefully lowered in order to improve the signal over background ratio. Even if a certain dependence in the signal efficiency on the Higgs mass can be seen, a common cut at $m(ee) = 40$ GeV/c^2 was found to give the best result in term of signal over background efficiency and it was therefore chosen.
A lower cut on the invariant mass is also necessary to suppress the events with two electrons coming from Υ and other hadronic resonances decays. Such cut has been set to 12 GeV/c^2. For low Higgs masses, it is responsible for signal efficiency loss, but it can not be further lowered to take the hadronic resonances background under control.

Missing transverse energy

As already mentioned the presence of large missing energy due to the two neutrinos in the final state is one of the main characteristics of the $H \rightarrow WW^{(*)}$ decay. The missing transverse energy distribution is shown in figure 6.20 for the signal and the backgrounds. A cut on the low value has been imposed at 35 GeV and it is expected to be efficient against the Drell Yan production. An upper cut can be also exploited to improve the significance, due to the different spectrum of the signal with respect to the backgrounds, in the region of low masses; such cut is not particularly relevant for the intermediate mass region.
The proposed selection cuts on the missing transverse energy are summarized in

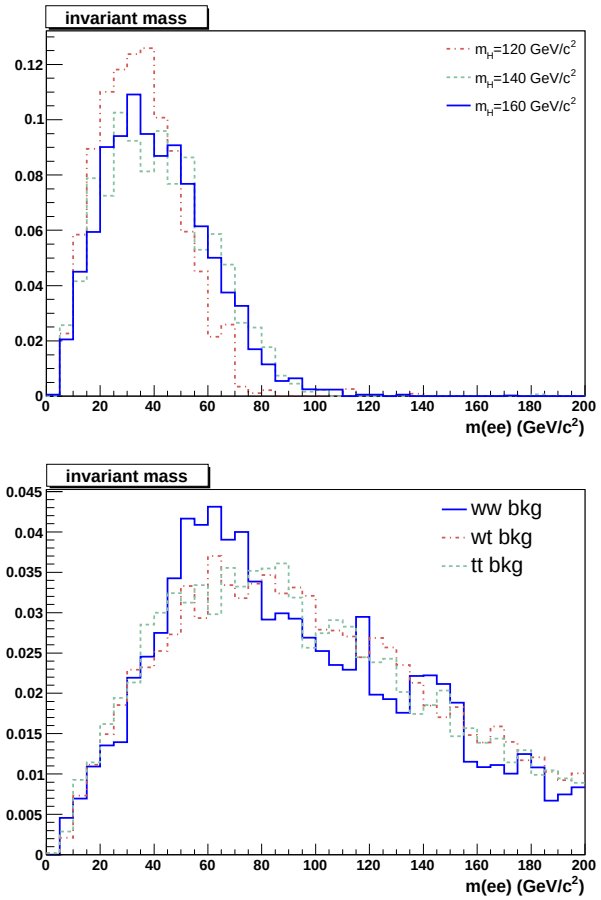

Figure 6.19: *Di-electron system invariant mass distribution for (top) the signal at different values of the Higgs mass and (bottom) the backgrounds. The distributions refer to events passing the preselections on electron reconstruction and isolation. For a better visibility the plots are normalized to the same area.*

table 6.6.

Transverse mass

Because of the two undetected neutrinos in the decay of the two Ws, the invariant mass of the Higgs decay products can not be reconstructed. It is anyway possible to reconstruct a transverse mass using the leptons and the missing energy as

$$m_T = \sqrt{2p_T^{ll} E_T^{miss} (1 - \cos\Delta\theta)} \tag{6.3}$$

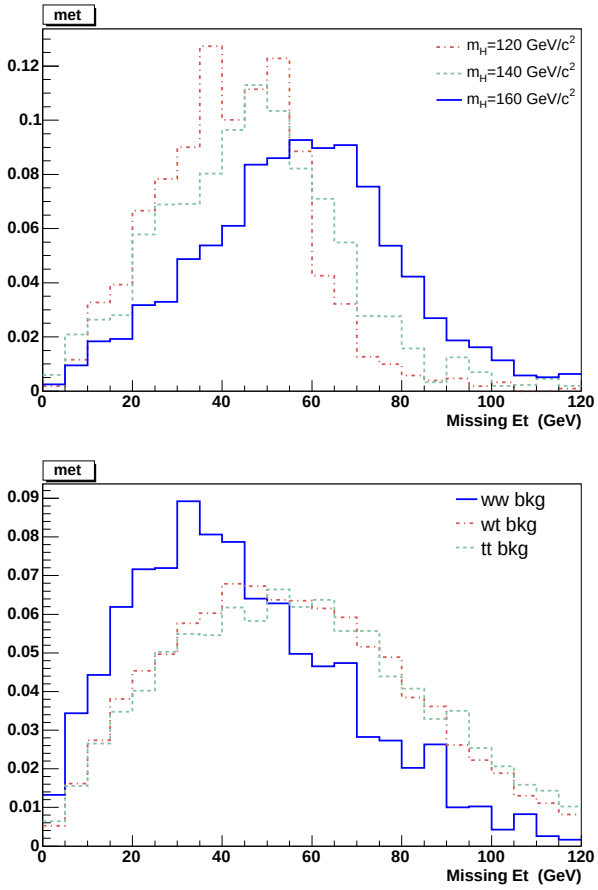

Figure 6.20: *Missing transverse energy distribution for (top) the signal at different values of the Higgs masses and (bottom) the backgrounds. The distributions refer to events passing the preselections on electron reconstruction and isolation. For a better visibility the plots are normalized to the same area.*

m_H (GeV/c^2)	inf (GeV)	sup (GeV)
120	35	70
130	35	75
140	35	80
$>$ 150	35	100

Table 6.6: *Values for the cuts on the transverse missing energy.*

The transverse mass distribution is shown in figure 6.21 for the events passing the preselections. To further suppress the background, the condition $m_H/2 < m_T <$

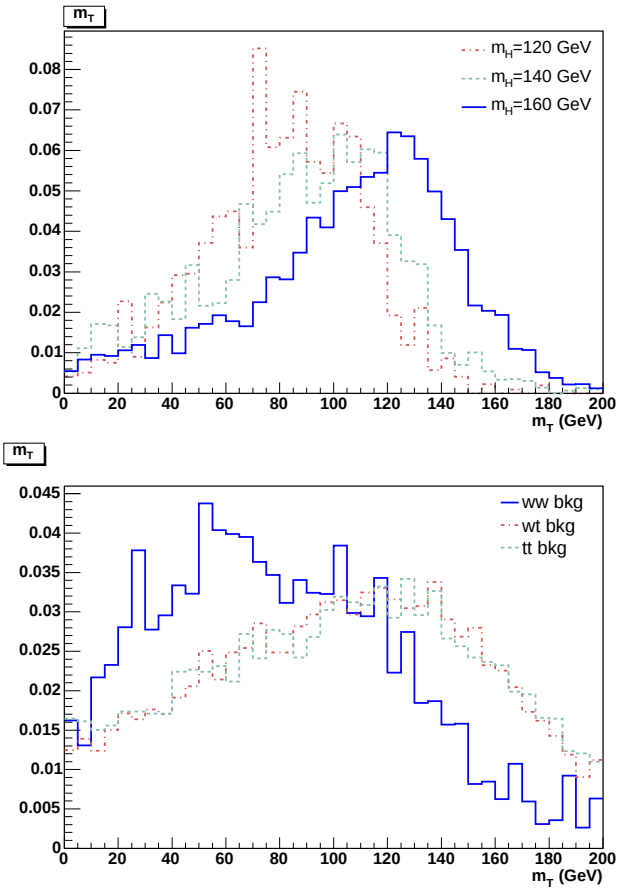

Figure 6.21: *Transverse mass distribution (eq.6.3) for (top) the signal at different values of the Higgs masses and (bottom) the backgrounds. The distributions refer to events passing the preselections on electron reconstruction and isolation. For a better visibility the plots are normalized to the same area.*

m_H has been imposed.

6.5 Selection efficiency and signal visibility

The number of selected events after the cuts which have been discussed in the previous section are reported in tables 6.7 to 6.12 for the different Higgs mass selections. The single cut efficiency is defined with respect to the events passing

the previous cut. The global efficiencies and the signal over background ratio are also reported.
Over the considered mass range, the efficiency on the signal varies between ∼4% and ∼6%. Since the cuts depend on the mass hypothesis which is done, the backgrounds efficiencies are also mass-dependent. Thanks to the hard jet veto, the $t\overline{t}$ background is almost completely reduced despite the large cross section and it is at a level between 20% and 50% of the WW background over the considered range of masses. The latter remains as dominating background contribution.
Figure 6.22 gives the signal over background ratio as a function of the Higgs mass. For Higgs masses in the discovery region the ratio is well above one, in the low mass region the ratio is below 0.5. In the low mass region, the significance could be enhanced by exploiting the clear signature of the vector boson fusion process, by requiring two forward jets with a large pseudorapidity difference [120].

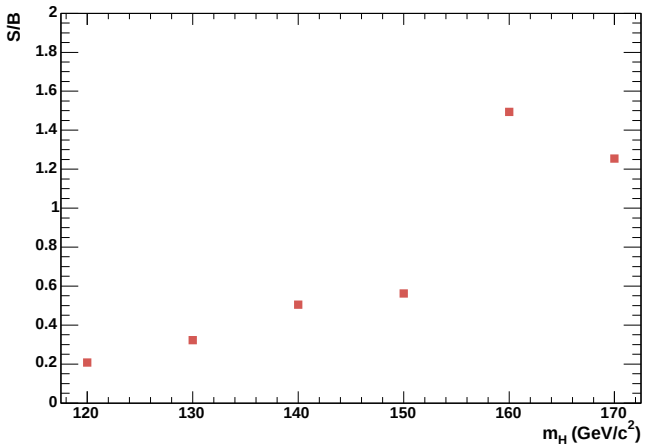

Figure 6.22: *Signal over background ratio as a function of the Higgs mass.*

It is worth noticing here how the applied cuts are particularly sensitive to the low Higgs p_T region as it was mentioned at the beginning of the chapter. This is quite evident looking at figure 6.23, which shows the Higgs p_T distribution before and after the cuts. This is the proof of the necessity of reweighting the events.

To quantify the probability to observe a signal in a future experiment different significance estimators can be used. They are usually classified as event counting and likelihood methods. For a given luminosity L the expected number of events is $n_S = \sigma_S L$ for the signal and $n_B = \sigma_B L$ for the background. Therefore $< n > = n_B + n_S$ events are expected in total on average in presence of signal.
Event counting methods compare n_S and n_B to define the significance in term

	m_H = 120 GeV/c^2	WW	$t\bar{t}$	Wt
$\sigma^{LO}_{Pythia} \times BR$	0.038 pb	0.83 pb	5.9 pb	0.62 pb
$\sigma^{NLO} \times BR$	0.063 pb	1.31 pb	9.7 pb	0.38 pb
HLT 1e/2e	35 (56%)	834 (64%)	7696 (80%)	308 (81%)
at least 2 leptons	21 (60%)	476 (57%)	5473 (71%)	217 (70%)
isolation	20 (95%)	469 (98%)	5093 (93%)	208 (96%)
$E_T^{sc} > 10$ GeV,$\mid \eta \mid < 2.5$	18 (91%)	459 (98%)	4971 (98%)	204 (98%)
$\Rightarrow$ Preselection eff (%)	51.94$\pm$ 1.14	55.02$\pm$ 0.11	64.59$\pm$ 0.07	66.21$\pm$ 0.17
CJV + α-cut	10 (55%)	301 (66%)	237 (5%)	36 (18%)
25$< p_T^{max,\,l} <$ 50 GeV/c	7.7 (78%)	142 (47%)	52 (22%)	8.3 (23%)
$p_T^{min,\,l} > 15$ GeV/c	6.1 (78%)	132 (93%)	46 (87%)	7.7 (93%)
35 $<$ MET $<$ 70 GeV	4.1 (67%)	59 (45%)	17 (38%)	3.9 (51%)
$\Delta\phi_{ll} < 100^\circ$	3.2 (80%)	36 (62%)	8.5 (49%)	2.2 (57%)
$12 < m_{ll}^{inv} <$ 40 GeV/c^2	2.7 (84%)	13 (35%)	2.2 (26%)	0.7 (32%)
$m_H/2 < m_T^{WW} < m_H$	2.5 (93%)	10 (79%)	1.3 (60%)	0.5 (68%)
$\Rightarrow$ Selection eff (%)	14.00$\pm$ 1.10	2.22$\pm$ 0.11	0.03$\pm$ 0.06	0.24$\pm$ 0.16
$\Rightarrow$ Total eff (%)	4.05 $\pm$ 0.33	0.78 $\pm$ 0.05	0.01 $\pm$ 0.01	0.13 $\pm$ 0.03
$\Rightarrow S/B$	0.211 $\pm$ 0.046			

Table 6.7: *Mass hypothesis: 120 GeV/c^2. Number of expected events for an integrated luminosity of 1 fb^{-1} after each selection. The relative efficiencies with respect to the previous cut are given in percent within the brackets. The preselection efficiencies are computed with respect to the events passing the HLT, the selection efficiencies with respect to the events passing the preselections. The quoted errors include the statistical uncertainty only.*

	m_H = 130 GeV/c^2	WW	$t\bar{t}$	Wt
$\sigma^{LO}_{Pythia} \times BR$	0.073 pb	0.83 pb	5.9 pb	0.62 pb
$\sigma^{NLO} \times BR$	0.119 pb	1.31 pb	9.7 pb	0.38 pb
HLT 1e/2e	70 (59%)	834 (64%)	7696 (80%)	308 (81%)
at least 2 leptons	46 (65%)	476 (57%)	5473 (71%)	217 (70%)
isolation	43 (95%)	469 (98%)	5093 (93%)	208 (96%)
$E_T^{sc} > 10$ GeV,$\mid \eta \mid < 2.5$	40 (93%)	459 (98%)	4971 (98%)	204 (98%)
$\Rightarrow$ Preselection eff (%)	57.76$\pm$ 1.08	55.02$\pm$ 0.12	64.59$\pm$ 0.07	66.21$\pm$ 0.18
CJV + α-cut	22 (54%)	301 (66%)	237 (5%)	36 (18%)
25$< p_T^{max,\ l} <$ 50 GeV/c	17 (79%)	142 (47%)	52 (22%)	8.3 (23%)
$p_T^{min,\ l} >$ 15 GeV/c	13 (78%)	132 (93%)	46 (87%)	7.7 (93%)
35 $<$ MET $<$ 75 GeV	9.4 (70%)	62 (47%)	19 (42%)	4.2 (54%)
$\Delta\phi_{ll} < 100^\circ$	7.6 (80%)	39 (63%)	9.8 (51%)	2.4 (59%)
12 $< m_{ll}^{inv} <$ 40 GeV/c^2	5.3 (70%)	14 (35%)	3.1 (32%)	0.8 (31%)
$m_H/2 < m_T^{WW} < m_H$	5.0 (96%)	12 (89%)	2.7 (86%)	0.6 (85%)
$\Rightarrow$ Selection eff (%)	12.50$\pm$ 0.95	2.65$\pm$ 0.11	0.05$\pm$ 0.07	0.31$\pm$ 0.17
$\Rightarrow$ Total eff (%)	4.24 $\pm$ 0.34	0.93 $\pm$ 0.05	0.03 $\pm$ 0.01	0.17 $\pm$ 0.03
$\Rightarrow S/B$	0.325 $\pm$ 0.053			

Table 6.8: *Mass hypothesis: 130 GeV/c^2. Number of expected events for an integrated luminosity of 1 fb^{-1} after each selection. The relative efficiencies with respect to the previous cut are given in percent within the brackets. The preselection efficiencies are computed with respect to the events passing the HLT, the selection efficiencies with respect to the events passing the preselections. The quoted errors include the statistical uncertainty only.*

	m_H = 140 GeV/c^2	WW	$t\bar{t}$	Wt
$\sigma^{LO}_{Pythia} \times BR$	0.105 pb	0.83 pb	5.9 pb	0.62 pb
$\sigma^{NLO} \times BR$	0.176 pb	1.31 pb	9.7 pb	0.38 pb
HLT 1e/2e	112 (64%)	834 (64%)	7696 (80%)	308 (81%)
at least 2 leptons	75 (67%)	476 (57%)	5473 (71%)	217 (70%)
isolation	71 (95%)	469 (98%)	5093 (93%)	208 (96%)
$E_T^{sc} > 10$ GeV,$\mid \eta \mid < 2.5$	69 (97%)	459 (98%)	4971 (98%)	204 (98%)
$\Rightarrow$ Preselection eff (%)	61.82$\pm$ 1.00	55.02$\pm$ 0.13	64.59$\pm$ 0.08	66.21$\pm$ 0.20
CJV + α-cut	35.3 (51%)	301 (66%)	237 (5%)	36 (18%)
25$< p_T^{max,\,l} <$ 50 GeV/c	29 (83%)	142 (47%)	52 (22%)	8.3 (23%)
$p_T^{min,\,l} >$ 15 GeV/c	26 (88%)	132 (93%)	46 (87%)	7.7 (93%)
35 $<$ MET $<$ 80 GeV	18 (69%)	64 (49%)	22 (49%)	4.5 (58%)
$\Delta\phi_{ll} < 100^\circ$	15 (84%)	41 (64%)	12 (54%)	2.7 (61%)
$12 < m_{ll}^{inv} <$ 40 GeV/c^2	9.5 (64%)	14 (35%)	3.6 (30%)	0.9 (33%)
$m_H/2 < m_T^{WW} < m_H$	9.2 (97%)	14 (95%)	3.6 (100%)	0.9 (94%)
$\Rightarrow$ Selection eff (%)	13.29$\pm$ 0.89	2.99$\pm$ 0.13	0.07$\pm$ 0.08	0.42$\pm$ 0.19
$\Rightarrow$ Total eff (%)	5.22 $\pm$ 0.37	1.04 $\pm$ 0.06	0.04 $\pm$ 0.01	0.22 $\pm$ 0.04
$\Rightarrow S/B$	0.508 $\pm$ 0.065			

Table 6.9: *Mass hypothesis: 140 GeV/c^2. Number of expected events for an integrated luminosity of 1 fb^{-1} after each selection. The relative efficiencies with respect to the previous cut are given in percent within the brackets. The preselection efficiencies are computed with respect to the events passing the HLT, the selection efficiencies with respect to the events passing the preselections. The quoted errors include the statistical uncertainty only.*

	m_H = 150 GeV/c^2	WW	$t\bar{t}$	Wt
$\sigma^{LO}_{Pythia} \times BR$	0.129 pb	0.83 pb	5.9 pb	0.62 pb
$\sigma^{NLO} \times BR$	0.221 pb	1.31 pb	9.7 pb	0.38 pb
HLT 1e/2e	148 (67%)	834 (64%)	7696 (80%)	308 (81%)
at least 2 leptons	102 (69%)	476 (57%)	5473 (71%)	217 (70%)
isolation	97 (96%)	469 (98%)	5093 (93%)	208 (96%)
$E_T^{sc} > 10$ GeV,$\mid \eta \mid < 2.5$	95 (97%)	459 (98%)	4971 (98%)	204 (98%)
$\Rightarrow$ Preselection eff (%)	64.07$\pm$ 1.04	55.02$\pm$ 0.13	64.59$\pm$ 0.08	66.21$\pm$ 0.19
CJV + α-cut	49 (52%)	301 (66%)	237 (5%)	36 (18%)
25$< p_T^{max,\ l} <$ 55 GeV/c	42 (85%)	166 (55%)	67 (28%)	10 (29%)
$p_T^{min,\ l} >$ 15 GeV/c	39 (93%)	156 (94%)	59 (88%)	9.7 (93%)
35 $<$ MET $<$ 100 GeV	30 (77%)	79 (51%)	40 (67%)	6.6 (68%)
$\Delta\phi_{ll} < 100^\circ$	23 (78%)	50 (62%)	22 (55%)	4.1 (62%)
$12 < m_{ll}^{inv} <$ 40 GeV/c^2	13 (58%)	17 (33%)	8.5 (39%)	1.3 (31%)
$m_H/2 < m_T^{WW} < m_H$	13 (96%)	15 (93%)	6.7 (79%)	1.1 (86%)
$\Rightarrow$ Selection eff (%)	13.70$\pm$ 0.93	3.34$\pm$ 0.13	0.13$\pm$ 0.08	0.54$\pm$ 0.19
$\Rightarrow$ Total eff (%)	5.86 $\pm$ 0.41	1.17 $\pm$ 0.06	0.07 $\pm$ 0.01	0.29 $\pm$ 0.05
$\Rightarrow S/B$	0.560 $\pm$ 0.061			

Table 6.10: *Mass hypothesis: 150 GeV/c^2. Number of expected events for an integrated luminosity of 1 fb^{-1} after each selection. The relative efficiencies with respect to the previous cut are given in percent within the brackets. The preselection efficiencies are computed with respect to the events passing the HLT, the selection efficiencies with respect to the events passing the preselections. The quoted errors include the statistical uncertainty only.*

	m_H = 160 GeV/c^2	WW	$t\bar{t}$	Wt
$\sigma^{LO}_{Pythia} \times BR$	0.146 pb	0.83 pb	5.9 pb	0.62 pb
$\sigma^{NLO} \times BR$	0.261 pb	1.31 pb	9.7 pb	0.38 pb
HLT 1e/2e	185 (71%)	834 (64%)	7696 (80%)	308 (81%)
at least 2 leptons	129 (70%)	476 (57%)	5473 (71%)	217 (70%)
isolation	123 (95%)	469 (98%)	5093 (93%)	208 (96%)
$E_T^{sc} > 10$ GeV,$\mid \eta \mid < 2.5$	120 (98%)	459 (98%)	4971 (98%)	204 (98%)
$\Rightarrow$ Preselection eff (%)	64.96$\pm$ 0.59	55.02$\pm$ 0.23	64.59$\pm$ 0.14	66.21$\pm$ 0.34
CJV + α-cut	59 (49%)	301 (66%)	237 (5%)	36 (18%)
30$< p_T^{max,\,l} <$ 55 GeV/c	47 (80%)	148 (49%)	60 (25%)	9.8 (27%)
$p_T^{min,\,l} >$ 25 GeV/c	32 (68%)	102 (69%)	35 (58%)	6.1 (62%)
35 $<$ MET $<$ 100 GeV	28 (88%)	50 (49%)	25 (72%)	4.1 (68%)
$\Delta\phi_{ll} < 100^\circ$	25 (89%)	28 (57%)	12 (48%)	2.5 (61%)
$12 < m_{ll}^{inv} <$ 40 GeV/c^2	15 (58%)	6.4 (22%)	3.1 (26%)	0.6 (24%)
$m_H/2 < m_T^{WW} < m_H$	14 (94%)	6.0 (95%)	2.7 (86%)	0.6 (97%)
$\Rightarrow$ Selection eff (%)	11.56$\pm$ 0.49	1.32$\pm$ 0.21	0.05$\pm$ 0.12	0.29$\pm$ 0.30
$\Rightarrow$ Total eff (%)	5.31 $\pm$ 0.23	0.46 $\pm$ 0.04	0.03 $\pm$ 0.01	0.15 $\pm$ 0.03
$\Rightarrow S/B$	1.491 $\pm$ 0.200			

Table 6.11: *Mass hypothesis: 160 GeV/c^2. Number of expected events for an integrated luminosity of 1 fb^{-1} after each selection. The relative efficiencies with respect to the previous cut are given in percent within the brackets. The preselection efficiencies are computed with respect to the events passing the HLT, the selection efficiencies with respect to the events passing the preselections. The quoted errors include the statistical uncertainty only.*

	m_H = 170 GeV/c^2	WW	$t\bar{t}$	Wt
$\sigma^{LO}_{Pythia} \times BR$	0.141 pb	0.83 pb	5.9 pb	0.62 pb
$\sigma^{NLO} \times BR$	0.253 pb	1.31 pb	9.7 pb	0.38 pb
HLT 1e/2e	185 (73%)	834 (64%)	7696 (80%)	308 (81%)
at least 2 leptons	133 (72%)	476 (57%)	5473 (71%)	217 (70%)
isolation	125 (94%)	469 (98%)	5093 (93%)	208 (96%)
$E_T^{sc} > 10$ GeV,$\mid \eta \mid < 2.5$	123 (98%)	459 (98%)	4971 (98%)	204 (98%)
$\Rightarrow$ Preselection eff (%)	66.66$\pm$ 0.59	55.02$\pm$ 0.23	64.59$\pm$ 0.14	66.21$\pm$ 0.34
CJV + α-cut	61 (50%)	301 (66%)	237 (5%)	36 (18%)
30$< p_T^{max,\ l} <$ 55 GeV/c	47 (78%)	148 (49%)	60 (25%)	9.8 (27%)
$p_T^{min,\ l} >$ 25 GeV/c	33 (70%)	102 (69%)	35 (58%)	6.1 (62%)
35 $<$ MET $<$ 100 GeV	29 (88%)	50 (49%)	25 (72%)	4.1 (68%)
$\Delta\phi_{ll} < 100^\circ$	26 (91%)	28 (57%)	12 (48%)	2.5 (61%)
12 $< m_{ll}^{inv} <$ 40 GeV/c^2	13 (49%)	6.4 (22%)	3.1 (26%)	0.6 (24%)
$m_H/2 < m_T^{WW} < m_H$	12 (97%)	6.1 (96%)	3.1 (100%)	0.6 (97%)
$\Rightarrow$ Selection eff (%)	10.02$\pm$ 0.46	1.33$\pm$ 0.19	0.06$\pm$ 0.11	0.29$\pm$ 0.28
$\Rightarrow$ Total eff (%)	4.88 $\pm$ 0.23	0.47 $\pm$ 0.04	0.03 $\pm$ 0.01	0.15 $\pm$ 0.03
$\Rightarrow S/B$	1.254 $\pm$ 0.170			

Table 6.12: *Mass hypothesis: 170 GeV/c^2. Number of expected events for an integrated luminosity of 1 fb^{-1} after each selection. The relative efficiencies with respect to the previous cut are given in percent within the brackets. The preselection efficiencies are computed with respect to the events passing the HLT, the selection efficiencies with respect to the events passing the preselections. The quoted errors include the statistical uncertainty only.*

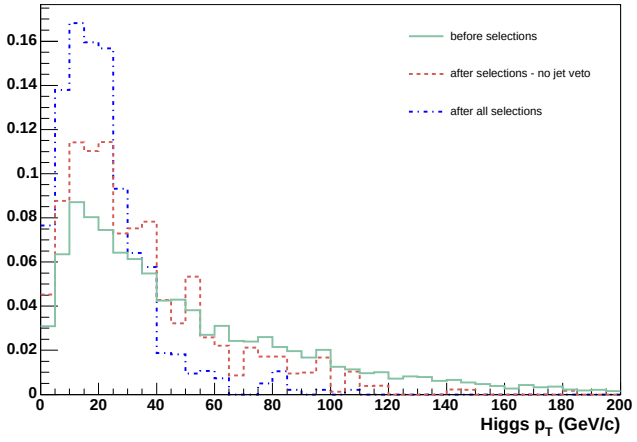

Figure 6.23: *Higgs transverse momentum distribution at the generator level and after the application of the jet veto and of all the cuts (m_H = 160 GeV/c^2). The cuts affect the spectrum in a Higgs p_T dependent way. For a better visibility, the plots are normalized to the same area.*

of the number of standard deviations an observed signal is above the expected background fluctuations. The most commonly used estimators are

$$S_1 = n_S/\sqrt{n_B}$$

$$S_{12} = 2(\sqrt{n_S + n_B} - \sqrt{n_B})$$

$$S_2 = n_B/\sqrt{n_S + n_B}$$

Likelihood methods consider the ratio of the null hypothesis H_0 (background only, $< n >$ = n_B) to an alternative hypothesis H_1 (the signal is present, $< n >$ = $n_B + n_S$). The significance is defined as

$$S_L = \sqrt{2lnQ} = \sqrt{2(lnL_1 - lnL_0)}$$

where L_0 and L_1 are the likelihood values in the null hypothesis and in the alternative one. Assuming Poisson statistics and considering

$$P_0(n; n_B) = \frac{1}{n!} n_B^n e^{-n_B}$$

$$P_1(n; n_B + n_S) = \frac{1}{n!}(n_B + n_S)^n e^{-(n_B + n_S)}$$

the previous expression can be expressed as

$$S_{cL} = \sqrt{2lnQ} \quad with \quad Q = \left(1 + \frac{n_S}{n_B}\right)^{n_B+n_S} e^{-n_S}$$

S_{cL} is the significance estimator which has been chosen by the CMS collaboration. Among its advantages there is the fact it is a likelihood estimator based on event counting and it is the simplest method to be used in presence of few background events. The signal significance as a function of the mass is given in figure 6.24 for an integrated luminosity of 10 fb^{-1}. For a comparison, the significance is also expressed with S_1. The agreement between the two estimators is in general

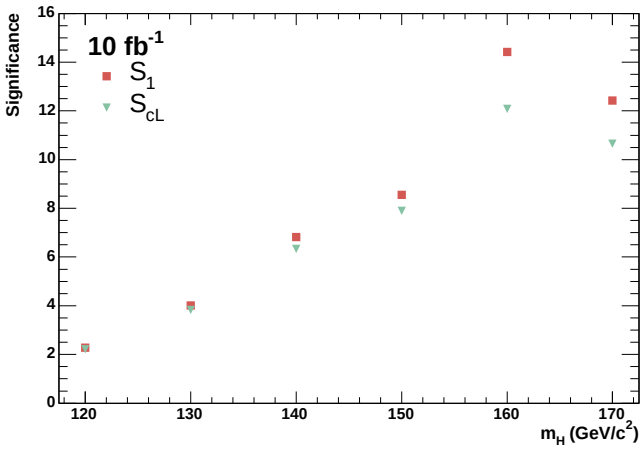

Figure 6.24: *Signal significance for an integrated luminosity of 10 fb^{-1} as a function of the Higgs mass.*

good, differences arise for high masses where the background over the signal is smaller and S_1 is not the correct estimator to use.

In this analysis only statistical fluctuations have been considered and a detailed study of systematic uncertainties has not yet been included.
Possible sources of theoretical uncertainties are related to the scale dependency and to the PDF choice. Experimental factors involved in the event selection and reconstruction have also to be included. The background normalization is a particularly serious topic, since the theoretical uncertainties are in many cases too large and the MonteCarlo programs background estimates can not be trusted before a tuning with the real data. A possible way to overcome the problem is the use of normalization regions obtained with slightly modified cuts where the background

is enhanced. From the number of measured events in the normalization region, the expected number of events in the signal region can be estimated as

$$N_{signal}^{estimated} = \frac{N_{signal}^{MC}}{N_{normalization}^{MC}} N_{normalization}^{measured} \quad (6.4)$$

Systematic errors cancel in the ratio if similar cuts are used to define the signal and the normalization region.
Following the proposition in [121], the normalization region for the $t\bar{t}$ background could be defined by removing the jet veto and replacing it with the request of two b-jets in the final state. For the WW background, the normalization region could be obtained imposing a looser $\Delta\phi$ cut.
This study is left for the future. On the basis of the results which have been obtained within the collaboration in the μe channel, systematic errors around 20% are expected on the background renormalization.

6.5.1 Higgs mass estimate

The presence of two neutrinos in the final state makes the reconstruction of a mass peak not possible. It is anyway important to note that some constraints on the Higgs boson mass can be extracted in this channel also.
The most sensitive variable to the Higgs mass is the transverse mass defined in equation 6.3, whose distribution for the sum of the backgrounds only and the signal plus background is shown in figure 6.25. The cuts introduced in the previous sections has been applied in a relaxed way, to reduce their dependence on the Higgs mass. The distributions, corresponding to 300 fb^{-1}, have been obtained with the fast simulation of the CMS detector FAMOS [122], after the agreement with the results obtained with the full simulation chain was checked. The pile-up contribution was not taken into account.
The transverse mass distribution, and its upper edge in particular, shows a clear dependence on the Higgs mass. Such dependence is anyway spoiled by the experimental resolution, which is dominated by the hadronic part. Besides, the uncertainties related to the background normalization and shape can be important.
The possibility to extract informations on the Higgs mass using a likelihood minimization method has been investigated. Reference distributions have been compared with 'experiments' at different masses and the compatibility was checked using the likelihood method. Since many approximations have been introduced in the analysis, only some hints can be gotten on the available resolution. The results obtained suggest the possibility to distinguish among two different mass hypothesis with a resolution order of few GeV, provided high statistics is collected (300 fb^{-1}).

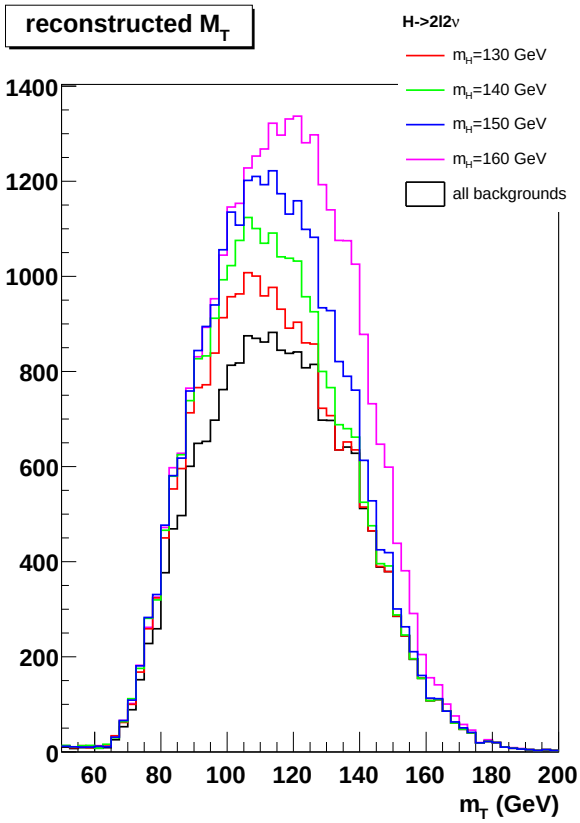

Figure 6.25: *Transverse mass distribution for the backgrounds and the signal plus backgrounds. The plots refer to an integrated luminosity L = 300 fb^{-1}.*

This could be usefully exploited to confirm the possible discovery of the Higgs boson at a given mass in another channel, ie.$ZZ^{(*)}$.

6.6 Further backgrounds: the W+jet production

As it was mentioned at the beginning of the chapter, the W+jet production is an example of process with the same final state as the signal in case one jet is misidentified as an electron. It has not been included into the analysis since the limited available statistics does not allow at this stage to take final conclusions on it. Some considerations have to be done anyway.
First, some comments are necessary on the samples generation. As it is discussed in [110], the inclusive W production can be generated using PYTHIA process number 2, based on 2$\rightarrow$1 matrix elements. The initial state radiation produces

jets which give the W a non zero transverse momentum. To get a correct description of the hard part of the p_T(W) spectrum, matrix element corrections are implemented in PYTHIA. However, since the cross section for the process quickly decreases with p_T(W), the hard part of the spectrum can not be efficiently generated in this way. W+jets are therefore generated using the 2→2 matrix element process W+1jet with the application of a lower cut on p_T(W). Since the cut is applied before the parton shower, the initial state radiation tends to shift the value of p_T(W). The samples produced in this way can be only used in the p_T(W) regions which match the inclusive distribution. Following the results obtained in [110] and shown in figure 6.26, the three p_T(W) regions in table 6.13 have been

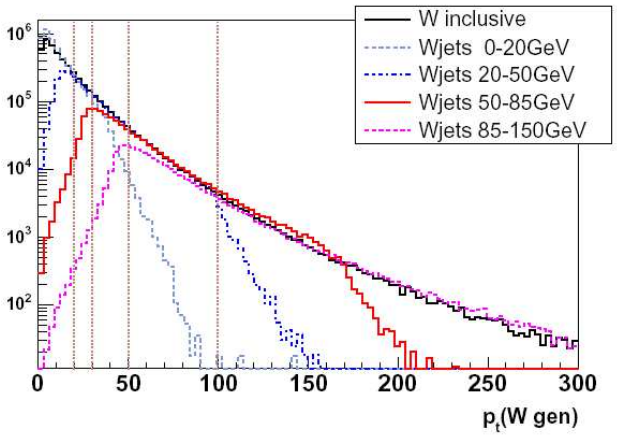

Figure 6.26: *The p_T(W) distribution for the different background samples together with the inclusive PYTHIA W production. The vertical lines show the separation between the different p_T(W) intervals studied [110].*

considered. The total inclusive W production cross section at NNLO times the

p_T(W) interval	$\sigma^{NNLO} \times BR\,(W \to l\nu)$	samples
20 GeV/c - 30 GeV/c	7.6 nb	Wjet0_20 + Wjet20_50
30 GeV/c - 50 GeV/c	6.3 nb	Wjet20_50
50 GeV/c - 100 GeV/c	3.1 nb	Wjet20_50 + Wjet50_85

Table 6.13: *The different W+jets NNLO cross sections times the leptonic branching ratio for the three p_T(W) intervals considered, together with the used samples.*

leptonic branching ratio is 60.9±3.0 nb for the MRST PDF set [111]. The cross sections for each region have been obtained rescaling it with the fraction of events

falling in each region.
W+jet is a background for the $H \rightarrow WW$ channel only if two isolated electrons with opposite charge are reconstructed in the detector which pass the kinematical selections. The fraction of W+jets events passing the preselections introduced in the previous paragraphs are given in table 6.14 for the three p_T(W) regions. These numbers represent the fraction of events which pass the trigger in which at least two isolated electrons with p_T >15 GeV/c and $\mid \eta \mid < 2.5$ are reconstructed.

p_T(W) interval	$\sigma^{NNLO} \times BR\,(W \rightarrow l\nu)$	eff (2 electrons)	σ_{2ele}
20 GeV/c - 30 GeV/c	7.6 nb	4.e-05	339 fb
30 GeV/c - 50 GeV/c	6.3 nb	8.e-05	511 fb
50 GeV/c - 100 GeV/c	3.1 nb	2.e-04	678 fb

Table 6.14: *Preselection efficiency and selected cross section times the branching ratio after having required two opposite charge isolated electrons in the final state.*

Even if the background is strongly reduced by the preselections applied here, the signal over background ratio is largely below one. The kinematical cuts should also be applied to get a final result but the limited available statistics makes such application not possible, since only few events survive for each p_T(W) interval.
The problem is therefore left open. The W+jets production seems at this stage to be a quite serious background for the channel under analysis.
A first study has shown that the imposition of hard cuts on the two electrons p_T makes it completely negligible, confirming the results given in [110]. Figure 6.27 shows the ratio signal over background and the signal significance as a function

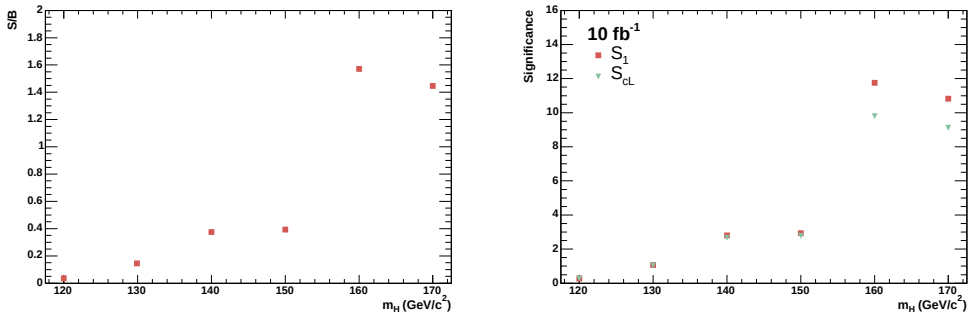

Figure 6.27: *Signal over background ratio (left) and signal significance for an integrated luminosity of 10 fb^{-1} (right) as a function of the Higgs mass. The kinematical cuts proposed in [110] have been applied.*

of the Higgs mass after the cuts proposed in [110] have been applied (the W+jets background can be considered negligible after such cuts, therefore it was not introduced in the analysis). Since such cuts have been tuned on the high mass region and they would spoil the signal significance at low Higgs masses, it was preferred not to apply them until it is not proved - with larger statistics - that no other solution can be found. Alternative selections are currently under study, together with a different isolation definition tuned on this kind of background and a better electron identification strategy. Preliminary results have shown that in this way the efficiencies in table 6.14 can be reduced by a factor larger than 10, while leaving the signal almost inaffected. Combined with the kinematical cuts, this makes it possible to keep the W+jets background well under control - without a strong reduction of the signal over background ratio. The study of this background source is anyway still on-going.

6.7 Conclusions

In this chapter, the study of the decay channel $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$ has been performed. A set of selections has been developed in order to obtain a significative background reduction. It has been shown that a good signal over background ratio can be obtained in the whole considered range of masses and in the region between $2m_W$ and $2m_Z$ in particular. A 5σ significance can be obtained with only 10 fb^{-1} in large part of the considered region. Finally, the possibility to extract informations on the Higgs mass despite the absence of a narrow mass peak has been investigated. Preliminary results, which do not take into account the possible systematics, suggest the possibility to get it with a resolution of few GeV.

Conclusions

This thesis summarizes the results of the work I did within the CMS collaboration in the last years.

Two main streams can be identified. The first part of the work shows the contribution I gave to the development of the electromagnetic calorimeter and in particular to some aspects related to its calibration. The second part describes the analysis work that was done to develop electron reconstruction algorithms. The study of the decay $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$, with two electrons in the final state, is also discussed as an example of decay channel in which a good electron reconstruction strategy is important.

A central role in the electron reconstruction is played by the electromagnetic calorimeter, which is required to have a very good energy resolution for the physics CMS plans to do. A direct contribution to the constant term of the energy resolution is given by the intercalibration of the calorimeter channels. Several aspects related to the intercalibration have been discussed using the 2003 test beam data. The precision which can be obtained on the intercalibration at test beam, including both statistical and systematic uncertainties, has been analysed. The contribution of several systematics which could affect the intercalibration was found to be lower than 0.3% and the statistical precision better than 0.1% with about 1000 events per crystal. The study of the effects of the irradiation with electrons and pions on the crystals was then performed. Since the effect changes from crystal to crystal, it could spoil the initial intercalibration. The irradiation affects the light transmission of the crystals only and not the scintillation mechanism, therefore it is possible to monitor the crystals response evolution with a laser system. The reliability of the monitoring system has been demonstrated. A procedure to correct for the radiation effects has been discussed and it has been found to stabilize the response of the crystals with an accuracy better than 0.2%.

The electron reconstruction currently in use in CMS has been discussed. The algorithms and their characteristics have been analysed and the main problems

have been identified. The large bremsstrahlung emission due to the tracker material has been shown to affect both the tracker and the calorimeter performances, in term of reconstruction efficiency and precision. A tracking strategy to better deal with it has been proposed. The method, which exploits the Gaussian Sum Filter for the track fitting, has been shown to allow for high reconstruction efficiency even at low p_T and for the collection of the track hits up to the outermost tracker layers without the introduction of a noticeable fake rate. The proper treatment of the non gaussian radiation losses was also shown to provide a meaningful value of the momentum at the outermost hit, giving both an estimate of the bremsstrahlung fraction and the possibility to improve the matching between the tracker and the calorimeter. An electron identification procedure based on a likelihood method has been also discussed.

The proposed electron reconstruction algorithms have been employed for the analysis of the decay channel $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$. The study of the signal significance in the mass range 120 GeV/c^2 $< m_H <$170 GeV/c^2 has been performed. A set of selections has been developed in order to obtain a significative background reduction. With the proposed cuts the signal over background ratio has been found to be higher than one in the mass region between $2m_W$ and $2m_Z$. A 5σ statistical significance have been obtained with only a 10 fb^{-1} integrated luminosity in most of the considered region. Finally, the possibility to extract informations on the Higgs mass despite the absence of a narrow mass peak has been investigated. The analysis discussed here takes into account the statistical fluctuations due to the finite MonteCarlo samples only. The study of the systematic uncertainties has still to be done and only some hints to a possible procedure to normalize the backgrounds have been given.

Riassunto

Attualmente uno dei problemi principali della fisica delle alte energie è l'origine della massa di fermioni e bosoni. Il meccanismo di rottura spontanea della simmetria elettrodebole viene comunemente invocato per spiegare le masse delle particelle nell'ambito del Modello Standard. Esso tuttavia prevede l'esistenza di almeno una particella scalare, il bosone di Higgs, non ancora rivelata sperimentalmente.
La ricerca del bosone di Higgs è uno dei principali obiettivi del *Large Hadron Collider* (LHC), un nuovo acceleratore attualmente in costruzione al CERN di Ginevra che comincerà la presa dati nel 2007. LHC è un collisionatore protone-protone con energia nel centro di massa di 14 TeV e luminosità di 10^{34} cm^{-2} s^{-1}. Quattro esperimenti raccoglieranno dati a LHC: CMS ed ATLAS, a carattere generale, LHCb, per lo studio della fisica del quark b ed ALICE, per lo studio della fisica degli ioni pesanti.
Il rivelatore *Compact Muon Solenoid* (CMS) si propone un vasto programma di fisica, attraverso la ricerca del bosone di Higgs e di nuova fisica oltre il Modello Standard e attraverso misure di precisione di quest'ultimo. La principale caratteristica di CMS è l'elevato campo magnetico solenoidale di 4T; i sottorivelatori che costituiscono il sistema sono, a cominciare dalla linea del fascio, il tracciatore, i due calorimetri (elettromagnetico ed adronico) e le camere a muoni.

L'argomento di questa tesi è lo studio del potenziale di scoperta del rivelatore CMS per il canale di decadimento $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$. Il calorimetro elettromagnetico è essenziale per la rivelazione degli elettroni, che caratterizzano lo stato finale di tale canale. Un contributo alla sua calibrazione e lo sviluppo di algoritmi per la ricostruzione di elettroni sono altresì trattati in questa tesi.

Il calorimetro electromagnetico di CMS ECAL è un calorimetro omogeneo a cristalli scintillanti di tungstato di piombo $PbWO_4$ a struttura modulare. Il disegno di ECAL è stato guidato dalle richieste dettate dal canale di decadimento $H \rightarrow \gamma\gamma$, il principale canale per la ricerca di un bosone di Higgs 'leggero' ($m_H \lesssim 130$ GeV/c^2) per il quale ottima risoluzione angolare ed energetica sono

necessarie. La risoluzione energetica di un calorimetro elettromagnetico è comunemente parametrizzata come $\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$, dove a è un termine stocastico e b un termine di rumore. Al termine costante c, che domina ad alte energie, contribuiscono diversi fattori tra i quali la stabilità delle condizioni di lavoro (voltaggio dei fotorivelatori, temperatura...), l'uniformità della risposta dei cristalli e la precisione della calibrazione.
La procedura di calibrazione dei cristalli di ECAL è organizzata in tre tappe e consiste di una precalibrazione prima dell'inizio della presa dati, una calibrazione *in-situ* con eventi fisici e un monitoraggio della stabilità con un sistema laser. La precalibrazione, che fornisce i valori di partenza per la calibrazione *in-situ*, è attualmente in corso e sfrutta sia misure fatte in laboratorio che test con raggi cosmici e fasci di elettroni. L'ambiente controllato del test su fascio consente di ottenere la massima precisione. Questa tesi mostra come il metodo qui discusso sia stabile e robusto e come le diverse fonti di incertezza sistematica abbiano un contributo massimo di pochi permille. La precisione statistica è migliore dello 0.1% con 1000 eventi circa per cristallo. L'elevato tasso di radiazioni previsto a LHC potrebbe degradare la precisione dell'intercalibrazione, influenzando la trasparenza dei cristalli attraverso la formazione di centri di colore in modo diverso da cristallo a cristallo. Un sistema laser è previsto in CMS per monitorare la perdita di trasparenza dei cristalli. I risultati qui discussi mostrano come sia possibile relazionare la perdita di trasparenza dei cristalli misurata con il laser a quella misurata con un fascio di elettroni e di conseguenza correggere la risposta dei cristalli con una precisione migliore del 2 permille. Si è anche dimostrato, pur se con poca statistica, che la relazione tra laser e fascio è universale e non dipende dal particolare cristallo. Questi risultati soddisfano le richieste di ECAL, per il quale l'obiettivo è mantenere il termine costante della risoluzione energetica inferiore a 0.5%.

L'utilizzo di algoritmi di ricostruzione di elettroni appositamente sviluppati è importante per poter sfruttare al meglio le eccellenti proprietà del calorimetro elettromagnetico. La ricostruzione di elettroni è stata studiata in questa tesi con una simulazione dettagliata del rivelatore. In CMS la ricostruzione di elettroni utilizza lo stesso algoritmo a livello di trigger e per l'analisi off-line. Un elettrone è identificato da una traccia puntante verso un deposito di energia nel calorimetro. Il principale problema per la ricostruzione di elettroni viene dall'emissione di fotoni per radiazione di bremsstrahlung dovuta al materiale del tracciatore, che influenza sia l'efficienza che la precisione di ricostruzione. In questa tesi è discusso un algoritmo per la ricostruzione di tracce elettroniche basato sul *Gaussian Sum Filter*, in cui la perdita di energia per bremsstrahlung è approssimata con una serie di gaussiane invece che con una sola. L'algoritmo è stato testato su elettroni simulati a momento trasverso fissato. Il metodo proposto per la ricostruzione del

seme della traiettoria consente di migliorare l'efficienza di ricostruzione rispetto a quella dell'algoritmo usato per la ricostruzione on-line di elettroni; per elettroni di $p_T = 10$ GeV/c il guadagno è circa del 25%. La migliore parametrizzazione della perdita di energia consente la ricostruzione di un numero superiore di hit nel tracciatore e conseguentemente una stima realistica del momento della traccia all'ingresso del calorimetro. Tale stima può essere utilizzata per determinare la frazione di energia persa dall'elettrone nel propagare attraverso il tracciatore, un'informazione questa utile ad esempio per le procedure di calibrazione. L'informazione sui parametri di traccia all'ingresso del calorimetro consente anche una migliore efficienza di identificazione degli elettroni. Il metodo di massima verosimiglianza qui discusso permette di avere un'efficienza di ricostruzione di elettroni a livello del 97% per una misidentificazione di jet elettromagnetici dell'ordine di 85%; con un'efficienza di identificazione di elettroni di circa 85%, la misidentificazione di jet è circa 0.5%

I risultati ottenuti sulla ricostruzione di elettroni sono stati utilizzati per lo studio del canale di decadimento $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$. La massa del bosone di Higgs è un parametro libero del Modello Standard, ma argomenti teorici e dati sperimentali suggeriscono un bosone di massa intermedia. Nella regione di masse inferiori a $2m_Z$ uno dei principali canali di ricerca del bosone di Higgs a LHC è $H \rightarrow WW^{(*)}$, con il successivo decadimento dei due W in leptoni, elettroni o muoni. Tale canale è il principale canale di scoperta per un bosone di Higgs di massa compresa tra $2m_W$ e $2m_Z$, poichè in tale regione il branching ratio dell'Higgs in $2W$ è quasi unitario. Nella regione di masse compresa tra 120 e 155 GeV/c^2 circa, il decadimento $H \rightarrow WW^*$ può essere sfruttato per lo studio degli accoppiamenti del bosone di Higgs attraverso rapporti di sezioni d'urto. Come argomento di questa tesi si è scelto lo stato finale con due elettroni e due neutrini. La presenza dei due neutrini rende la ricostruzione del picco di massa dell'Higgs a partire dai prodotti di decadimento non possibile, pertanto lo studio si risolve in un esperimento di conteggio in cui è ricercato un eccesso di eventi rispetto a quanto atteso in presenza di solo fondo. In questa analisi i principali fondi riducibili $t\bar{t}$ e Wtb e il fondo irriducibile WW sono stati studiati con una simulazione dettagliata del rivelatore ed è stata discussa una serie di tagli per migliorare il rapporto segnale su fondo in funzione della massa dell'Higgs. Con il metodo proposto, nella regione di masse studiate il rapporto segnale su fondo varia tra 0.2 e 1.5 e la significanza statistica del segnale tra 2 e 14 con una luminosità integrata di 10 fb^{-1}.

Résumé

Actuellement, une des questions principales dans le domaine de la physique des hautes énergies est l'origine de la masse des particules. Le mécanisme de brisure de symétrie électrofaible est communément utilisé pour expliquer les masses des fermions et des bosons vecteurs. Il prédit l'existence d'au moins une particule scalaire, le boson de Higgs, cependant les recherches directes effectuées auprès de l'accélérateur LEP n'ont pas permis sa découverte. La recherche du boson de Higgs est une des motivations pour le projet d'un nouvel accélérateur, le *Large Hadron Collider* (LHC), en construction au Laboratoire Européen pour la Physique des Particules (CERN) et qui commencera à prendre des données en 2007.
Le LHC est un collisionneur proton-proton avec une énergie dans le centre de masse égale à 14 TeV et une luminosité de $10^{34}\text{cm}^{-2}\text{s}^{-1}$. Quatre détecteurs sont prévus auprès du LHC: ATLAS et CMS (deux détecteurs à vocation généraliste), LHCb (pour l'étude de la physique des mésons B) et ALICE (pour l'étude de la physique des ions lourds).
Le détecteur *Compact Muon Solenoid* (CMS) sera utilisé pour explorer la physique accessible au LHC: recherche du boson de Higgs, recherche de nouvelle physique au-delà du Modèle Standard et mesures de précision du Modèle Standard. La caractéristique principale du détecteur CMS est le très fort champ magnétique solenoïdal de 4T, genéré par un aimant supraconducteur. Les sous-systèmes de détection sont le trajectomètre, les deux calorimètres (électromagnétique et hadronique) et le système de détection des muons.

Le sujet de cette thèse est l'étude des possibilités de détection du boson de Higgs dans CMS pour le canal $H \to WW^{(*)} \to 2e2\nu$. Le calorimètre électromagnétique est essentiel pour la détection des électrons. Nos contributions à sa calibration ainsi qu'au développement des algorithmes de reconstruction des électrons sont également presentées dans cette thèse.

Le calorimètre électromagnétique de CMS (ECAL) est un calorimètre homogène composé de cristaux scintillants de $PbWO_4$. La définition de ses caractéristiques a été guidée par les demandes du canal de désintégration du boson

de Higgs en deux photons $H \rightarrow \gamma\gamma$. C'est le canal principal pour la recherche d'un boson de Higgs léger ($m_H \lesssim 130$ GeV/c^2) et il nécessite de très bonnes résolutions en énergie et position. La résolution en énergie d'un calorimètre électromagnétique peut être exprimée par $\frac{\sigma E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$, où a est le terme stochastique et b est le terme de bruit. Nombreux facteurs contribuent au terme constant c, qui est prédominant à haute énergie; parmi eux, l'un des plus importants est la précision de la calibration relative. La procédure de calibration des cristaux de l'ECAL est organisée en trois étapes. Une précalibration, exploitant les mesures optiques en laboratoire, les mesures de test en faisceau et les mesures de rayons cosmiques, est en cours pour obtenir les premiéres valeurs. En début de prise de données, une calibration relative et une calibration absolue seront réalisées avec des événements de physique ($Z \rightarrow e^+e^-$, $W \rightarrow e\nu$ ...). En plus, un système laser sera utilisé pour le contrôle de la stabilité de la calibration. Nous avons participé à la précalibration des cristaux pendant le test en faisceau. La procédure consiste à envoyer des électrons avec un énergie connue sur le crystal considéré et à corriger la réponse du cristal en fonction du point d'impact de l'électron; ensuite les réponses des différents cristaux au même signal sont normalisées. Nous avons démontré que la procédure est très robuste et que les nombreuses sources d'incertitudes systématiques ont une contribution maxime de l'ordre du millième. La précision statistique est meilleure de 0.1% avec 1000 événements. La précision de la calibration relative des cristaux peut être dégradée par l'effet des radiations. En effet, celles-ci sont très importantes au LHC et influencent la transparence des cristaux par la création de centres de couleur dans le cristal. Étant différent d'un cristal à l'autre, cet effet peut influencer la calibration relative. Un système laser a été prévu pour contrôler la perte de transparence. En effet, la perte de transparence mesurée avec le laser est corrélée avec celle mesurée avec un faisceau des particules. Nous avons démontré que la relation entre les deux est universelle et ne dépend pas du cristal, et qu'il est possible de corriger la réponse des cristaux pour la perte de transparence avec une précision meilleure que 2 millièms. Les résultats satisfont les requêtes de CMS, demandant pour la résolution en énergie un terme constant inférieur à 0.5%.

L'utilisation des algorithmes spécialement développés pour la reconstruction des électrons est très importante pour exploiter au maximum les propriétes du calorimètre électromagnétique. La reconstruction des électrons a été étudiée grâce à une simulation détaillée du détecteur CMS. Un électron est défini par l'association d'une trace avec un agrégat dans le calorimètre électromagnétique. Dans CMS, la reconstruction des électrons au niveau du déclerchement est effectuée en utilisant le même algorithme que pour les électrons off-line au niveau de l'analyse. Cet algorithme est ajusté sur les électrons d'énergie assez élevée ($p_T \sim 30$ GeV/c). Le plus gros problème dans la reconstruction des électrons provient du rayonnement

de bremsstrahlung dans la matière du trajectomètre. Dans cette thèse nous avons étudié un algorithme pour la reconstruction des traces électromagnétiques. Celui-ci cherche le germe de la trace dans une région très large pour ne pas perdre en efficacité et reconstruit ensuite la trace en utilisant l'algorithme *Gaussian Sum Filter* (GSF), qui effectue une approximation de la perte d'énergie de l'électron avec une série de fonctions gaussiennes au lieu d'une seule gaussienne, afin d'obtenir une meilleure paramétrisation. L'algorithme a été testé avec des événements de simulation MonteCarlo pour la reconstruction des traces d'électrons avec impulsion transverse fixe. Pour des électrons de 10 GeV/c de p_T, la méthode de recherche du germe donne un efficacité de reconstruction supérieure de $\sim$25% par rapport à celle de la méthode utilisée pour la reconstruction des électrons dans le trigger. Les résolutions en énergie et en direction sont équivalentes entre les deux algorithmes. La meilleure paramétrisation de la perte d'énergie de l'électron permet de collecter un grand nombre de hits dans le trajectomètre. Ainsi, une meilleure estimation des paramètres de la trace au niveau du calorimètre est possible. Ceci peut être utilisé pour un estimation de la perte d'énergie dans le trajectomètre, qui est très utile pour la calibration du calorimètre. L'exploitation de l'information sur la trace au niveau de calorimètre est également utilisée pour améliorer l'efficacité de l'identification des électrons. La méthode de maximisation de vraisemblance qui est proposée pour l'identification permet d'avoir une efficacité d'identification au niveau de 97% pour une mauvaise identification des jets électromagnétiques d'environ 7%; avec une efficacité d'identification de 85%, la mauvaise identification est au niveau de 0.5%.

Les résultats de la reconstruction des électrons ont été exploités pour l'étude du canal $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$. Bien que la masse du Higgs soit un paramètre libre du Modèle Standard, des arguments théoriques et les données expérimentales de mesure électrofaible privilégient un boson de Higgs de masse intermédiaire. Dans le domaine de masses qui s'étende d'environ 120 GeV/c^2 jusqu'à $2m_Z$, un des principaux canaux de recherche du Higgs est $H \rightarrow WW^{(*)} \rightarrow 2l2\nu$, où les leptons sont des électrons ou des muons. Ce canal est le canal principal pour la découverte d'un boson de Higgs de masse entre $2m_W$ et $2m_Z$, car dans cette région le rapport de branchement du Higgs en WW est presque unitaire; dans la région de masse entre 120 GeV/c^2 et 155 GeV/c^2, la désintégration $H \rightarrow WW^*$ peut être exploitée pour l'étude des couplages du boson de Higgs. Dans cette thèse, nous avons étudié le canal avec deux électrons dans l'état final. En raison de la présence des deux neutrinos, la reconstruction du pic de masse du Higgs à partir des produits de désintégration n'est pas possible. L'étude de ce canal est donc une expérience de comptage et la connaissance des bruits de fond est très importante. Les bruits de fonds principaux (WW, $t\bar{t}$ et Wtb) ont été étudiés en détail ainsi que le signal. Les sections efficaces et les rapports de branchement ont

été calculés grâce à des programmes appliquant les calculs théoriques les plus récents, incluant les corrections QCD et QED aux ordres supérieurs à l'ordre dominant. Les particules de l'état final ont été produites par le programme de simulation MonteCarlo PYTHIA, avec des sections efficaces calculées à l'ordre dominant. Les événements ont été pondérés pour obtenir la concordance des variables cinématiques avec des modèles théoriques plus rigoureux dans le cas du signal produit par fusion de gluons et du bruit de fond WW; pour les deux bruits de fond $t\bar{t}$ et Wtb les sections efficaces ont été normalisées sans la pondération des événements. Nous avons étudié l'optimisation des coupures de l'analyse en fonction de la masse de Higgs. Une coupure d'isolation et la requête de n'avoir aucun jet dans la région centrale suppriment efficacement les bruits de fond réductibles $t\bar{t}$ et Wtb. Pour supprimer le bruit de fond irréductible WW, la corrélation de spin entre les deux W a été exploitée. Des coupures sur le p_T des électrons , leur masse invariante et sur la masse transverse des WW ont également été utilisées pour améliorer le rapport signal sur bruit. Avec la méthode proposée, le rapport signal sur bruit est entre 0.2 et 1.5 et la significance du signal est entre 2 et 14 avec un luminosité intégrée de 10 fb^{-1}, dans la région de masse entre 120 et 170 GeV/c^2.

Chapitre 1.

Nous donnons un bref résumé de la partie du Modèle Standard relative aux interations électrofaibles et au secteur de Higgs. Les limites théoriques et expérimentales pour la masse du boson de Higgs sont discutées, ainsi que les caractéristiques principales du secteur de Higgs dans les théories au-delà du Modèle Standard. Un bref survol des principaux mécanismes de production du Higgs au LHC est donné et les canaux de recherche sont discutés.

Chapitre 2.

Nous donnons une brève description du *Large Hadron Collider* et du détecteur *Compact Muon Solenoid* avec ses sous-systèmes de détection. Une attention particulière est donnée au calorimètre électromagnétique.

Chapitre 3.

La procédure du calibration du calorimètre électromagnétique est décrite. Des détails sont donnés sur la précalibration en utilisant les mesures optiques en laboratoire et les mesures avec les rayons cosmiques, ainsi que sur la calibration initiale qui sera effectuée au début de la pris de données avec des événements de physique. Le problème de l'endommagement des cristaux par radiations est présenté dans la deuxième partie du chapitre, de même que le fonctionnement du système de contrôle avec un laser.

Chapitre 4.
Les résultats de la précalibration des cristaux pendant le test en faisceau sont donnés, de même que l'étude du contrôle de la perte de transparence avec un laser. Nous avons démontré la stabilité de la procédure pour la précalibration et nous avons étudié les sources d'incertitudes systématiques et la précision statistique. La mesure de la perte de transparence avec le laser a été correlée avec celle mesurée avec un faisceau des particules. L' étude de la relation entre les deux et de la possibilité de corriger la réponse des cristaux pour la perte de transparence est donnée.

Chapitre 5.
Nous donnons un description des algorithmes utilisés pour la reconstruction des électrons dans CMS et des problèmes qui affectent la reconstruction, comme le rayonnement de bremsstrahlung dans la matière du trajectomètre. Un algorithme pour la reconstruction des traces électromagnétiques en utilisant l'algorithme *Gaussian Sum Filter* est développé et l'étude de son efficacité de reconstruction et de sa résolution est décrite. L'information sur la trace qui vient de cette methode est aussi utilisée pour améliorer l'efficacité de l'identification des électrons.

Chapitre 6.
L'étude du canal $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$ est présentée dans le domaine de masses s'étendant de 120 GeV/c^2 jusqu'à 170 GeV/c^2. Les bruits de fonds principaux (WW, $t\bar{t}$ et Wtb) et le signal ont été étudiés en détail ainsi que l'optimisation des coupures de l'analyse en fonction de la masse de Higgs. Le rapport signal sur bruit et la significance du signal sont présentés.

Bibliography

[1] F.J.Yndurain 'The Theory of Quarks and Gluon Interactions',
Springer-Verlag, Berlin 1993

[2] S.Weinberg, Phys.Rev.Lett. 19 (1967) 1264;
S.Glashow, Nucl.Phys.22 (1961) 579;
S.Weimberg, Phys.Rev.Lett. 19 (1967) 1264;
A.Salam, 'Elementary Particle Theory: Relativistic groups and Analiticity' (Nobel Symposium N.8), edited by Svartholm (Almqvist and Wiksell, Stokholm, 1969), pag 367

[3] P.W.Higgs, Phys.Rev.Lett. 12 (1964) 132; Phys.Rev.Lett. 13 (1964) 508;
Phys.Rev. 145 (1966) 1156;
F.Euglert et al., Phys.Rev.Lett.13 (1964) 321;
G.S.Guralnik et al., Phys.Rev.Lett.13 (1964) 585

[4] 'Precision Tests of the Electroweak Interactions at the Z Pole',
M.Martinez et al., CERN-EP/98-27

[5] S.P.Martin, arXiv:hep-ph/9709356

[6] H.P.Nilles, Phys.Rep. 110 (1984) 1

[7] G.Altarelli et G.Isidori, Phys. Lett. B337 (1994),
J.A.Casas et.al., Phys. Lett. B342 (1995) 171, B382 (1996) 374,
T.Hambye et K.Riesselmann, Phys. Rev. D55 (1997) 7255

[8] G.Ridolfi, 'An Introduction to the Standard Model of Electroweak Interactions', available at http://www.ge.infn.it/ ridolfi/

[9] The LEP Electroweak Working Page,
http://lepewwg.web.cern.ch/LEPEWWG/, updated with summer 2005 results

[10] 'Search for the Standard Model Higgs Boson at LEP', ALEPH, DELPHI, L3 and OPAL Collaborations, The LEP Working Group for Higgs Boson Searches, hep-ex/0306033, Phys.Lett.B565 (2003) 61

[11] M.Spira, hep-ph/9705337

[12] H.Georgy et al., Phys. Rev. Lett. 40 (1978) 692

[13] M.Spira et al, Nucl. Phys B453 (1995) 17,
D.Graudenz et al., Phys. Rev. Lett. 70 (1993) 1372

[14] S.Catani, D. de Florian et M.Grazzini, JHEP 0105 (2001) 025,
R.V.Harlander et W.B.Kilgore, Phys.Rev.D 64 (2001) 013015;
C.Anastasiou et K.Melnikov, Nucl. Phys. B646 (2002) 220,
V.Ravindran, J.Smith and W.L.van Neerven, Nucl.Phys. B665 (2003) 325

[15] R.N.Cahn and S.Dawson, Phys.Lett.B136 (1984) 196,
G.Altarelli, B.Mele and F.Pitolli, Nucl. Phys. B287 (1987) 205,
T.Han, G.Valencia and S.Willenbrock, Phys. Rev. Lett. 69 (1992) 3274

[16] S.L.Glashow, D.V.Nanopoulos and A.Yildiz, Phys. Rev. D18 (1978) 1724,
Z.Kunzst, Z.Trocsanyi and W.J.Striling, Phys. Lett. B271 (1991) 247

[17] T.Han and S.Willenbrock, Phys. Lett. 71 (1993) 2699

[18] Z.Kunszt, Nucl. Phys. B247 (1984) 339;
J.F.Gunion, Phys. Lett. B253 (1991) 269;
W.J.Marciano and F.E.Paige, Phys. Rev. Lett. 66 (1991) 2433;
Z.Kunszt, S.Moretti and W.J.Stirling, Z.Phys C74 (1997) 479

[19] R.N.Cahn, Rep. Prog. Phys. 52 (1989) 389

[20] LHC Design Report Volume I.

[21] 'Collider Physics: LHC', F.Gianotti, CERN-open-2000-293

[22] 'ATLAS Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN', CERN/LHCC 94-43, LHCC/P2, Dec 15, 1994

[23] 'The Compact Muon Solenoid', CERN/LHCC 94-38, LHCC/P1, Dec 15, 1994

[24] 'LHCb Technical Proposal', CERN/LHCC 98-004, LHCC/P1, Feb 20, 1998

[25] 'ALICE Technical Proposal' TDR, CERN/LHCC 2001-021, Oct 3, 2001

[26] 'The Tracker System Project Technical Design Report', CERN/LHCC 98-06, CMS TDR 5, Apr 15, 1998
Addendum to the CMS Tracker TDR, CERN/LHCC 2000-016, CMS TDR 5, Addendum 1, Feb 21, 2000

[27] R.Wunstorf, T.Rohe, A.Rolf, Nucl.Inst.& Methods A 388 (1997) 308-313.

[28] B.C.MacEvoy, G.Hall, K.Gill, Nucl.Inst.& Methods A 374 (1996) 12-26.

[29] 'The performance of the CMS tracking', N. De Filippis, CMS CR 2004/024.

[30] 'The Electromagnetic Calorimeter Project Technical Design Report', CERN/LHCC 97-33, CMS TDR 4, Dec 15, 1997

[31] 'Development of a Uniformisation Procedure for the PbWO4 Crystals of the CMS Electromagnetic Calorimeter', E.Auffray et al., CMS NOTE 2001/004.

[32] 'Uniformisation Techniques for Crystals in the ECAL Endcaps', D.Britton et al., CMS NOTE 2004/006.

[33] 'Progress on Avalanche Photodiodes as photon detectors for $PbWO_4$ crystals in the CMS experiment', F.Cavallari, CMS CR 1997/10

[34] 'Low Voltage Supply System for the Very Front End Readout Electronics of the CMS Electromagnetic Calorimeter', W.Lustermann, Proceedings of the 9th Workshop on Electronics for the LHC Experiments, Amsterdam Sep.2003

[35] 'The MGPA Electromagnetic Calorimeter Readout Chip for CMS', M.Raymond, Proceedings of the 9th Workshop on Electronics for the LHC Experiments, Amsterdam Sep.2003

[36] 'The CMS ECAL Very Front End Electronics: production and tests', C.Biino et al., Proceedings of the 10th Workshop on Electronics for LHC and future Experiments, Boston

[37] 'The Hadron Calorimeter Project Technical Design Report', CERN/LHCC 97-31, CMS TDR 2, Jun 20, 1997

[38] 'The Magnet Project Technical Design Report', CERN/LHCC 97-10, CMS TDR 1, May 2, 1997

[39] 'CMS Muon Technical Design Report', CERN/LHCC 97-32, CMS TDR 3, Dec 15, 1997

[40] 'The Trigger and Data Acquisition Project, Volume I: The Level-1 Trigger Technical Design Report', CERN/LHCC 2000-38, CMS TDR 6.1, Dec 15, 2000

[41] 'The Trigger and Data Acquisition Project, Volume II: Data Acquisition and High Level Trigger Technical Design Report', CERN/LHCC 02-26, CMS TDR 6.2, Dec 15, 2002

[42] E.Auffray et al., Nucl.Inst.& Methods A 456 (2001) 325

[43] S.Baccaro et al., Nucl.Inst.& Methods A 459 (2001) 278

[44] S.Baccaro et al., Nucl.Inst.& Methods A 385 (1997) 69

[45] 'Correlation between Light Yield and Longitudinal Transmission in $PbWO_4$ crystals and impact on the precision of the crystal intercalibration', L.Barone et al., CMS RN 2004/005

[46] 'The Calibration Strategy of CMS Electromagnetic Calorimeter', P.Meridiani, CMS CR 2004/045.

[47] 'Test of the Feasibility of Pre-intercalibration of ECAL Supermodules with Cosmic Rays', K. Deiters et al., CMS IN-2004/023

[48] ' Feasibility of Intercalibration of CMS ECAL Supermodules with Cosmic Rays', W. Bertl et al., CMS NOTE-2004/036

[49] 'Pre-calibration of CMS Electromagnetic Calorimeter with cosmic rays', M.Bonesini et al., CMS NOTE in preparation.

[50] 'Intercalibration of the CMS Electromagnetic Calorimeter Using Jet Trigger Events', D.Futyan, CMS NOTE 2004/007.

[51] 'In-Situ Calibration of the CMS Electromagnetic Calorimeter', D.I.Futyan, CMS CR 2003/005.

[52] 'Calibration of the L3 electromagnetic calorimeter in electron beam', L3BGO Collaboration, L3 Note 1712, 1995

[53] 'Use of $Z \to ee$ events for ECAL calibration', P.Meridiani et al, CMS AN 2005/054.

[54] 'A glossary for Numeric analysis', S.A.Householder, Oak Ridge National Laboratory, Oak Rigde, Tenn., 1958

[55] 'ECAL Barrel Inter-Calibration with Isolated Electrons', G.Daskalakis et al., CMS AN 2005/009.

[56] 'ECAL Endcaps Inter-Calibration with Isolated Electrons', P.Govoni et al., CMS AN 2005/055; P.Govoni, private comunication

[57] A.Fyodorov et al., Radiation Measurements, Vol 26, No1, 1996, 107

[58] P.Lecoq et al, Nucl.Inst.& Methods A 365 (1995) 291

[59] Xiangdong Qu et al, Nucl.Inst.& Methods A 480 (2002) 468

[60] 'Improvement of Several Properties of Lead Tungstate Crystals with Different Doping Ions', E.Auffray et al., CMS NOTE 1997/054

[61] 'The ECAL calibration: Use of the light monitoring system, Version 2.0', P.Bonamy et al., CMS NOTE 1998/013

[62] 'A beam hodoscope based on square scintillating fibers', L.Dobrzynsky et al., CMS NOTE in preparation

[63] 'Design and Performance of the Cooling System for the Electromagnetic Calorimeter of CMS', P.Baillon et al., CMS CR 2004/030.

[64] 'Pulse amplitude reconstruction in the CMS ECAL using the weights method', P.Paganini et al., CMS NOTE 2004/025.

[65] 'Test-beam results on the performance of two matrices of $PbWO_4$ crystals for the CMS ECAL and comparison with laboratory measurements', E.Auffray et al., CMS IN 2001/033.

[66] 'Study of coherent noise', M.Dejardin, presentation given at the ECAL M0' Testbeam Meeting, Sept 25 2002. Available at: http://agenda.cern.ch/fullAgenda.php?ida=a021427

[67] 'Energy Resolution of the CMS ECAL Barrel Super-Module Using MGPA Electronics', G.Dewhirst et al., CMS RN 2004/004.

[68] 'Performance of the electromagnetic calorimeter of CMS and search of extra dimensions in the ADD model', A.Ghezzi, PhD Thesis, Milano 2004

[69] 'Test beam irradiation catalogue for 1997 and 1998', C.Seez, CMS IN 1998/026

[70] ‘Performance and measurements of the light monitoring system for CMS-ECAL from 2002 test beam data’, A.Van Lysebetten et al., CMS RN 2004/001.

[71] ‘On the ratio between electron signal and monitoring losses due to irradiation’, P.Bloch, CERN, September 16, 2004

[72] ‘Radiation induced color centers and light monitoring for lead tungstate crystals’, X.Qu et al., CMS NOTE 1999/069

[73] ‘The electromagnetic calorimeter of CMS and its sensitivity to non-pointing photons’, G.Franzoni, PhD Thesis, Milano 2004

[74] http://cmdoc.cern.ch/orca

[75] ‘CMS Software Architecture Software Framework, Services and Persistency in High Level Trigger, Reconstruction and Analysis’, V.Innocente et al., CMS NOTE 2000/047

[76] http://cmdoc.cern.ch/cms00/projects/CMKIN/index.html

[77] T.Sjostrand et al. Comp.Phys.Comm., 135 (2001) 238

[78] http://cmdoc.cern.ch/OSCAR

[79] http://cern.ch/geant4

[80] http://cmdoc.cern.ch/cmsim/cmsim.html

[81] GEANT - Detector description and simulation tool, Version 3.21, CERN, 1993

[82] ‘Electron Reconstruction in the CMS Electromagnetic Calorimeter’, E.Meschi et al., CMS NOTE 2001/034

[83] ‘Jet rejection using the pixel matching for the low and the high luminosity’, G.Daskalakis et al., CMS NOTE 2002/039

[84] R.Fruhwirth, Nucl.Inst.& Methods A 262, 1987, 444

[85] D.Stampfer et al., Comp.Phys.Comm. 79 (1994) 157

[86] H.Bethe and W.Heitler, Proc.R.Soc.London A 146 (1934) 83

[87] ‘Reconstruction of electrons with the Gaussian Sum Filter in the CMS tracker at LHC’, W.Adam et al., CMS NOTE 2005/001

[88] R.Fruhwirth, Comp.Phys.Comm.100 (1997) 1

[89] R.Fruhwirth et al., Comp.Phys.Comm.110 (1998) 80

[90] F.Ferri, talk given at the CMS Ecal E/γ Meeting, March 15 2005.
Available at: http://agenda.cern.ch/fullAgenda.php?ida=a051548#s13

[91] 'Identifying central electrons in the CMS detector', M.Dittmar et al.., CMS IN 2004/036

[92] 'Optimisation of the discovery potential of the Higgs Boson in the decay channel $H \to ZZ^{(*)} \to 2e^+2e^-$ with the CMS detector', P.Meridiani, PhD Thesis, Roma 2004

[93] M.Dittmar et al., Phys.Rev. D55, 167 (1997)

[94] T.Rizzo, Phys.Rev.D 22, 722 (1980);
W.Y.Keung and W.J.Marciano, *ibid.* 30, 248 (1984)

[95] J.Fleischer and F.Jegerlehner, Phys.Rev.D 23, 2001 (1981);
B.A.Kniehl, Nucl.Phys.B357, 439 (1991)

[96] M.Spira, http://people.web.psi.ch/spira/proglist.html

[97] A.Djouadi et al., Comp.Phys.Commun, 108 (1998) 56

[98] S.Frixione, Nucl.Phys.B 410 (1993) 280

[99] J.Ohnemus, Phys.Rev.D 44 (1991) 1403

[100] http://mcfm.fnal.gov/

[101] T.Binoth et al., hep-ph/0503094

[102] M.Duhrssen at al., hep-ph/0504006

[103] P.Nason et al., Nucl.Phys.B 327 (1988) 49;
R.Bonciani et al., Nucl.Phys.B 529 (1998) 424

[104] M.Beneke et al., hep-ph/0003033

[105] F.Maltoni, talk given at 'Hadron Collider Physics Simposium 2005', Les Diablerets, 4-9 july 2005.
Available at $http://cmsdoc.cern.ch/\ anikiten/cms-higgs/sm_cross-sections.txt$;

[106] $http://sirius.ihep.su/\ spitsky/toprex/toprex.html$

[107] A.S.Belyaev et al., Phys.Rev.D 59, 075001 (1999), hep-ph/9806332;
T.M.P.Tait, Phys.Rev.D 61, 034001 (2000), hep-ph/9909352;
A.S.Belyaev et al., Phys.Rev.D 63, 034012 (2001) hep-ph/0003260

[108] J.Campbell et al., hep-ph/0506289

[109] M.Zanetti, talk given at the CMS Higgs Meeting, 30th September 2004.
Available at http://agenda.cern.ch/fullAgenda.php?ida=a044165,

[110] 'Jets faking leptons', A.S.Giolo-Nicollerat, CMS IN 2005/028

[111] A.D.Martin et al., Eur.Phys.J.C 14 (2000) 133, hep-ph/9907231

[112] S.Catani et al., JHEP 0201 (2002) 015

[113] G.Davatz et al., hep-ph/0402218

[114] S.Frixione et al., JHEP 0308 (2003) 007, hep-ph/0305252

[115] G.Corcella et al., JHEP 0101 (2001) 010, hep-ph/0210213

[116] http://cmsdoc.cern.ch/orca/Manuals/ORCA_8_7_3/doc/UserGuide/html/UserGuide.html

[117] G.Davatz et al, talk given at the CMS Higgs Meeting, 10th May 2005.
Available at http://agenda.cern.ch/fullAgenda.php?ida=a052713#s0,

[118] 'Association of jets with signal vertex', N.Ilina et al.., CMS AN 2005/006

[119] M.Dittmar et al, hep-ph/9703401

[120] N.Akchurin et al, CMS NOTE 2002/010

[121] N.Kauer, hep-ph/0404045

[122] http://cmsdoc.cern.ch/famos/

Acknowledgments

During the last three years I had the opportunity to meet many people who helped me, contributing to my personal and professional growth. I'd like now to thank them all.

Vorrei ringraziare l'intero gruppo di ricerca di CMS Milano, per avermi dato la possibilità di partecipare all'esperimento CMS lavorando in un ambiente stimolante e sereno. Un ringraziamento particolare lo devo a Stefano, il mio tutore, e a Tommaso e Marco, per le numerosi discussioni, i consigli e il supporto dimostratomi. Grazie a Tonino, sempre un'importante figura di riferimento all'interno del gruppo. Grazie a Pietro, per i numerosi consigli e la saggezza dispensata. Un sentito ringraziamento va ai giovani del gruppo con i quali ho condiviso l'esperienza di questi anni. Grazie a tutti, per essere stati sempre amici prima che compagni di lavoro. Un ringraziamento particolare a Martina, Pietro e Roberto, con i quali ho avuto il piacere di lavorare direttamente. Grazie a Gemma e Sara, per la loro vicinanza. Un grazie davvero speciale lo devo ad Alessio e Giovanni, per la preziosa collaborazione, l'infinita pazienza e i numerosi consigli, lavorativi e non.

I had the opportunity to spend many months working with the CMS group at the Ecole Polytechnique. This was for me an important and pleasant experience. I'd like to thank for this Philippe, who accepted to be my tutor and always supported me, and Ludwik, who welcomed me in his group. A special thanks to Claude, for his infinite patience and his continuous help, suggestions and encouragements, and to Yves, for the useful discussions which really made me learn a lot. I'd also like to thank Gerard, for his efforts in teaching me French.... Finally, a very special thanks to Evelyne, with whom I had the opportunity and the pleasure to collaborate, and to all the young people in the group, who very nicely welcomed me.

I'd like to thank the many people I had the opportunity to work with within the CMS collaboration. Many thanks to Wolfgang Adam, his continuous help and his

experience were very precious during the work on electron reconstruction. Thanks to Paolo Meridiani, for his great help on software and reconstruction issues, for the discussions about physics and the nice time at LLR. Thanks to the group of young people working on the $H \rightarrow WW$ channel, for the many nice and really useful discussions. And thanks to all the many people I met and I worked with, I can not mention them all but their collaboration was really precious for me and for this work.

Thanks to Philippe Bloch and Francesco Navarria, for having accepted to read this thesis and for the many interesting suggestions and remarks.

Da ultimo, siccome questi anni sono stati anche e soprattutto un'occasione di crescita personale, vorrei ringraziare tutte le persone che mi hanno sostenuto e supportato. Il ringraziamento più grande è per i miei genitori. Grazie a tutti gli amici e alle persone che mi sono state vicine.