



Université catholique de Louvain  
Faculté des Sciences  
Département de Physique

---

**Study of  $WW$  decay  
of a Higgs boson with  
the ALEPH and CMS detectors**

---

Christophe Delaere  
Promoteur: Professeur V. Lemaitre

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

---

July 2005



## Quelques remerciements

Mettre la touche finale à une thèse, c'est aussi jeter un bref regard en arrière, et réaliser tout ce que l'on doit à d'autres. Je voudrais donc exprimer ici toute ma gratitude envers ceux qui ont contribué à l'élaboration de ce travail.

Je remercie avant tout mon promoteur, le Professeur Vincent Lemaître, pour m'avoir soutenu, guidé et conseillé durant ces quatre années de recherche. Je remercie, en outre, les Professeurs Catherine Vander Velde, Denis Favart, Jean-Marc Gérard, Giacomo Bruno, Jean-Pierre Antoine et Roberto Tenchini pour avoir accepté d'être membres de mon jury et pour leurs commentaires judicieux et constructifs.

Je tiens tout particulièrement à remercier Olivier van der Aa. Il est vite devenu naturel de travailler ensemble, car cela rimait avec complémentarité, efficacité et souvent complicité, au grand bénéfice des divers projets techniques entrepris.

Lorsque commençait cette thèse, c'est au sein d'un petit groupe que je m'insérais... un petit groupe qui était appelé à grandir. Ce sont les post-doctorants et post-doctorantes qui font avancer bien des choses. Merci donc à Guillaume, pour son œil critique, à Muriel, entre autres pour avoir relu et commenté ce texte, et à Vincent qui m'a beaucoup aidé ces derniers mois pour la validation du HLT. Merci aussi à Alain, Fabian et Thomas, notre "staff informatique". Ce que j'ai appris ces dernières années en informatique, c'est à leur contact, et parce qu'ils ont su me laisser chipoter (non sans risque) avec l'environnement informatique professionnel de FYNU. Merci aussi à tous ceux en FYNU et en FYMA sans qui cette entreprise n'aurait pas été la même, qu'ils aient contribué activement, ou simplement par leur présence amicale, à ce que cette thèse prenne forme.

Au-delà de ces années passées "au cyclotron", cela fait enfin plus de vingt-six ans déjà que mes parents me conseillent et me soutiennent. Je leur en suis très reconnaissant.



# CONTENTS

<b>I</b>	<b>Theoretical background</b>	<b>5</b>
<b>1</b>	<b>The Standard Model</b>	<b>7</b>
1.1	Gauge Symmetries . . . . .	8
1.2	Particle content . . . . .	9
1.3	The Standard-Model Lagrangian . . . . .	10
1.4	Giving mass to particles . . . . .	12
<b>2</b>	<b>The Higgs particle</b>	<b>15</b>
2.1	Higgs boson properties . . . . .	16
2.2	Theoretical limits . . . . .	17
2.3	Experimental limits . . . . .	21
2.4	Indirect measurements . . . . .	27
<b>II</b>	<b>ALEPH - ZH associated production</b>	<b>31</b>
<b>3</b>	<b>An Apparatus for LEP Physics</b>	<b>33</b>
3.1	The LEP accelerator . . . . .	33
3.2	The ALEPH detector . . . . .	35
3.3	Software aspects . . . . .	42
<b>4</b>	<b>Search for ZH associated production</b>	<b>47</b>
4.1	Introduction . . . . .	47
4.2	Signal and background . . . . .	48
4.3	Event classification . . . . .	54
4.4	Confidence levels . . . . .	55
4.5	Study of events with a single hard or soft lepton . . . . .	58
4.6	Study of events with no lepton, or more than one hard lepton . . . . .	70
4.7	Results . . . . .	73

---

<b>III CMS - WH associated production</b>	<b>87</b>
<b>5 The Compact Muon Solenoid</b>	<b>89</b>
5.1 The Large Hadron Collider . . . . .	89
5.2 The CMS detector . . . . .	94
<b>6 The High-Level Trigger</b>	<b>107</b>
6.1 Level-1 streams . . . . .	109
6.2 Event reconstruction at HLT . . . . .	111
6.3 HLT streams . . . . .	117
6.4 The HLT steering code . . . . .	119
6.5 HLT rate measurements . . . . .	125
6.6 Trigger evaluation sequence and timing issues . . . . .	129
<b>7 Search for WH associated production</b>	<b>133</b>
7.1 Introduction . . . . .	133
7.2 Signal . . . . .	136
7.3 Background . . . . .	137
7.4 Trigger issues . . . . .	143
7.5 Signal discrimination . . . . .	149
7.6 Systematic uncertainties . . . . .	163
7.7 Results . . . . .	165
7.8 Fermiophobic Higgs boson . . . . .	172
7.9 Concluding remarks . . . . .	174
<b>IV Conclusion and Appendices</b>	<b>177</b>
<b>A High-Level Trigger tree representation</b>	<b>183</b>
<b>B Computing Farm</b>	<b>189</b>

## Introduction

You only need to look at the number of Higgs-related preprints published on the scientific web servers to sweep out any doubt about the interest that the Higgs particle arises within the physics community. All this excitement is explained by the dominance of one fundamental theory and the often unsuccessful search for disagreements with the data. This *Standard Model* is one of the big achievements of the last century. All known particles are perfectly described as leptons, quarks or bosons. All forces (but gravity) are included as gauge bosons exchanges.

More than an index of the particles and their interactions, the Standard Model is a mathematical description of the very nature of these entities, now seen as relativistic quantum fields. It is based on the concept of symmetries and has an important predictive power. Until recently, all the predictions were satisfied, and precision measurements were in agreement with expectations. The evidence of neutrino oscillations has only altered the Standard Model in a minor conceptual way. The main issue is therefore the non-observed Higgs boson. This predicted particle is a key feature of the theory, as it is a direct consequence of the mechanism giving their mass to weak gauge bosons. If the Higgs boson was not there, or if it had unexpected properties, the Standard Model would have to be replaced by a maybe not yet imagined alternative.

The elaboration, confirmation and test of the Standard Model took place all over the world in particle physics laboratories, where many new particles were discovered and studied in the second part of the 20th century. Since its creation in 1954, the European Organization for Nuclear Research (CERN) has been playing a leading role in the search for new physics. After the discovery of neutral currents (1973) and the discovery of the W and Z bosons (1983), it turned to the search for the Higgs boson. In 1989, the Large Electron Positron Collider (LEP) was brought into operation. During 11 years, scientists have looked at the Standard Model, validating its predictions to a large extent, but being unable to discover the desired particle.

The ALEPH detector was built to measure the events created by electron-positron collisions in LEP, in the energy range around 91 GeV (to study the

Z particle) and above the threshold of W pair production ( up to 200 GeV). It was one of four detectors built for that purpose, together with DELPHI, L3 and OPAL. The event rate at the peak of the Z is below 1 Hz and at least a factor hundred smaller at the highest energies. The ALEPH detector was therefore designed to accumulate as much information as seemed practical, keeping each event on tapes. In the last years of running, attention was focused on Higgs boson searches, in the Standard Model and beyond. Even if no evidence has been found, all those studies contributed to put stringent limits on alternative models and theories. One such study, developed during this thesis and presented as part of the present dissertation, is the search for a Higgs boson decaying into a pair or W bosons. In the mass range accessible at LEP, this channel is suppressed in the Standard Model but can contribute significantly if a fermiophobic Higgs boson is present in Nature. This study instigated software developments both of the ALEPH analysis environment (ALPHA++) and of new tools (likelihood-ratio method, neural networks, ...) for the analysis package used in High-Energy physics (ROOT).

CERN is now involved in the construction of the next generation of accelerators and detectors. The Large Hadron Collider (LHC) will, in conjunction with the general purpose ATLAS and CMS detectors, scrutinize the still allowed part of the Higgs boson mass range. The CMS detector is optimized for the identification and reconstruction of muons, electrons and photons in a large energy range and at high luminosity. The main objectives of the collaboration are to study physics within the Standard Model and beyond, like top physics, Higgs boson, supersymmetry or large-extra-dimensions models. The challenging LHC beam energy (7 TeV) and luminosity ( $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  in the first phase) should allow the measurement of a Higgs boson up to  $1000 \text{ GeV}/c^2$ . The observation of a light Higgs boson should nevertheless be challenging as the background will be higher, and difficult to control. In such a context, it is very important to study a wide variety of scenarios and decay channels for a light (from  $110 \text{ GeV}/c^2$  to  $220 \text{ GeV}/c^2$ ) Higgs boson, as this will allow to confirm any possible discovery. The detailed study of branching ratios will indeed be the only way to disentangle the proposed (exotic) Higgs sectors of the theory.

If the Higgs boson mass is lower than twice the W boson mass, the dominating decay mode is the production of a  $b\bar{b}$  pair, while the main production channel is the fusion of gluons from the protons. Nevertheless, direct production via gluon fusion followed by a decay to  $b\bar{b}$  cannot be observed as the QCD background is overwhelming. For this reason, the favorable channels do involve leptons (in associated production) or photons (for  $H \rightarrow \gamma\gamma$ ).

Beyond background reduction, interesting events - whatever the definition of an interesting event - must be accepted by the CMS trigger system, whose aim is to retain one event per million. The situation is here completely different to that at LEP, where the role of the trigger system of an experiment was "only" to record all interactions in the detector. The large cross-section at a hadron machine obliges us to reduce the rate to a level compatible with com-

puting resources available to store and analyze them. This reduction will be achieved in two stages. First, the hardware “Level-1” trigger will use rough information quickly collected by some dedicated electronics on the detector to reduce the event rate by a factor 1000. Then, the “High-Level” software-based trigger will use the full information to retain those valuable events and send them to mass storage. A lot of work has therefore to be done at the trigger level to ensure a good selectivity for events containing a Higgs boson, in a large variety of channels.

When coming into contact with the CMS software at the beginning of this thesis, it quickly appeared that the existing code was not adapted to the large-scale deployment of trigger strategies elaborated until then. There was a lack of steering code to handle various reconstruction and selection strategies for electrons, muons, taus, jets or missing transverse momentum. This missing part of the online and offline software has been developed in the context of this thesis. The HLT steering software is now massively used in preparation of the CMS Physics Technical Design Report (Physics-TDR).

Using this new piece of code, a search for Higgs boson decay into a pair of W bosons, in the associated production channel WH, was performed. This channel contributes at relatively low masses, and is of particular importance to measure branching ratios of the Higgs boson to W bosons and top quarks (by comparing this channel with the equivalent direct production mechanism). Technical developments were necessary to put in place the required computing infrastructure. This includes both hardware (a 72 CPU computing cluster was put together) and software (the CMS software was installed, and used for production of Monte-Carlo events and for analysis) activities.

This thesis is divided into three parts. First, the theoretical framework will be presented. After a brief description of the Standard Model (Chapter 1), the Higgs mechanism will be sketched. Emphasis will be put on existing limits on the Higgs boson mass, whether theoretical or experimental (Chapter 2). As explained before, the not-yet-observed Higgs boson is the subject of lots of attention. It is in this exciting context that a search for a (fermiophobic) Higgs boson was carried out at LEP, using data from the ALEPH detector. This work will be presented in Section Two (Chapter 4), after a brief description of the ALEPH detector (Chapter 3). Finally, a physics-readiness study of the CMS detector and software will be developed in the third Section, together with the developments within the context of the HLT steering software. The detector (Chapter 5), the High-Level Trigger (Chapter 6) and a search for the Higgs boson (Chapter 7) will be discussed. Some technical developments that are direct consequences of my thesis work are finally presented in appendices.



## **Part I**

# **Theoretical background**



## The Standard Model

In this chapter, the basics of the Standard Model will be introduced. Most of the work presented in this thesis is motivated by the success (and shortcomings) of this theory. The objective is to underline the key role of the electroweak symmetry breaking mechanism in the present theoretical construction. What is often called the *Higgs particle* will then be discussed in Chapter 2. This is one of the major predictions of the Standard Model.

According to the Standard Model, the ordinary matter is made up of elementary spin-1/2 particles (*fermions*). One knows the electron, observed in cathodic rays and as a cloud around the nucleus of atoms. The muon and the tau have similar properties, except for their mass which is each time larger. A massless neutrino is associated to each of these three leptons; this makes three *families*. Then there are the six quarks, again distributed in three families: (u,d), (s,c) and (b,t). The observation of the  $\Omega^-$  baryon, understood as a sss bound state, led to the introduction of color charge for quarks, which is now the basis of our understanding of mesons and baryons phenomenology.

The fermions interact via the electromagnetic, weak, strong and gravitational fields<sup>1</sup>. The quantified description of the fundamental interactions is made in terms of exchanges of *gauge bosons*. For instance, charged particles exchange photons, the muon decays emitting a W boson, or the quarks are bound by gluons to form a proton. The Standard Model accounts for key experimental observations made in the last century. These are for example the violation of parity observed in beta decays, related to the “V-A” structure of electroweak currents, the observation of neutral currents, first in bubble chambers, or the vast baryon spectrum. W and Z bosons are also observed to be massive particles, which led to the introduction of the Brout-Englert-Higgs mechanism. The properties of interactions led to the introduction of

---

<sup>1</sup>Gravity is not part of the theory, and is described by General Relativity. Integrating coherently gravity into the Standard Model is one of the long-term goals of physicists. This will not be discussed here.

gauge bosons, whose properties are related to the mathematical structure of the gauge theory. For this reason, the mathematical description is central in our understanding of Nature.

## 1.1 Gauge Symmetries

The Standard Model is a relativistic quantum field theory based on the groups of symmetries  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ . These three local symmetry groups correspond to the three interactions just described. The  $SU(3)_C$  group is related to the strong interaction between quarks, governed by QCD (Quantum Chromodynamics), and implies the existence of eight colored gluons. The  $SU(2)_L \otimes U(1)_Y$  groups are associated to the electroweak interaction. After dynamic symmetry breaking, this allows to describe the quantum electrodynamics ( $U(1)_{e.m.}$ ) and the weak interaction responsible for the beta decay.

The description of interactions relies on the work of Yang and Mills, who generalized the gauge structure of electromagnetism [1]. Making a symmetry  $G$  local (one says gauging the symmetry), one has to introduce a covariant derivative  $D_\mu$  that generalizes the usual derivative  $\partial_\mu$ , for the Lagrangian to remain invariant under the gauged symmetry. This can be done by introducing a set of real vector fields  $A_\mu^a$  so that  $D_\mu = \partial_\mu + igA_\mu^a T^a$ , where  $g$  is the gauge coupling, and  $T^a$  are the generators of group  $G$ . In order to give a dynamic to these gauge fields, a gauge tensor  $F_{\mu\nu}$  has to be introduced. It must be antisymmetric in its two spatial indices, so it is natural to define  $igF_{\mu\nu} = [D_\mu, D_\nu]$ . The Lagrangian of the Yang-Mills theory is then:

$$\mathcal{L}_{YMT} = i\bar{\eta}\gamma_\mu D^\mu \eta - \frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu}, \quad (1.1)$$

where  $\eta$  denotes here a system of scalar and spinorial fields.

The eight gauge bosons of the  $SU(3)_C$  group of QCD are denoted  $G_\mu^a$ , with the index  $a = 1, \dots, 8$ . The associated invariant gauge tensor is defined as

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_S f^{abc} G_\mu^b G_\nu^c, \quad (1.2)$$

where  $g_S$  is the coupling constant of strong interactions and  $f_{abc}$  are the structure constants of  $SU(3)_C$ . In the same spirit, the three gauge bosons of the  $SU(2)_L$  group are denoted  $W_\mu^i$  with  $i = 1, 2, 3$ . The associated gauge tensor is defined as

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g \epsilon^{ijk} W_\mu^j W_\nu^k, \quad (1.3)$$

where  $g$  is the coupling constant, and  $\epsilon^{ijk}$  the corresponding structure constants. The same development can be done for  $U(1)_Y$ . The coupling constant of this gauge group is  $g'$ .  $B_\mu$  is the gauge boson, and the gauge tensor is defined as

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \quad (1.4)$$

Table 1.1: The bosons of the Standard Model. To each symmetry correspond a boson and a tensor. The fourth column shows to which well-known physical state it corresponds to.

Group	Boson	Tensor	Physical state
$SU(3)_C$	$G_\mu^i$	$G_{\mu\nu}^i$	gluons
$SU(2)_L$	$W_\mu^i$	$W_{\mu\nu}^i$	$W^+$ , $W^-$ , $Z$ , photon
$U(1)_Y$	$B_\mu$	$B_{\mu\nu}$	photon, $Z$

To close this discussion, the charged weak bosons are directly defined from the  $SU(2)_L \otimes U(1)_Y$  gauge bosons as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \quad (1.5)$$

in order to have physical states of determined electric charge. At the same time, the neutral weak boson ( $Z_\mu$ ) and the photon ( $A_\mu$ ) are defined from  $W_\mu^3$  and  $B_\mu$  as

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3, \quad (1.6)$$

$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3, \quad (1.7)$$

with  $\cos\theta_W = g/\sqrt{g^2 + g'^2}$  and  $\sin\theta_W = g'/\sqrt{g^2 + g'^2}$ . This definition is motivated by the form of the  $SU(2)_L \otimes U(1)_Y$  covariant derivative, and results from the diagonalization of the gauge mass matrix obtained after symmetry breaking. All this is summarized in Table 1.1.

## 1.2 Particle content

Fermions are described by solutions of the Dirac equation. The equations of motion are derived from the Dirac Lagrangian

$$\mathcal{L} = \bar{\psi} i \gamma_\mu \partial^\mu \psi - m \bar{\psi} \psi \quad (1.8)$$

where  $\psi$  is a 4-component complex field (Dirac spinor). It represents a free fermion field of mass  $m$ . Using the Pauli representation of Dirac matrices, spinors can be expressed as a pair of right-handed and left-handed two-component complex fields

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \quad (1.9)$$

Fields of distinct chirality can be handled separately as long as the fermion is massless. Any mass term will nevertheless mix the two components, since

the mass term obtained introducing (1.9) into (1.8) is proportional to  $(\overline{\Psi}_L \Psi_R + \overline{\Psi}_R \Psi_L)$ . The  $SU(2)_L$  interaction only acts on left-handed doublets, while the  $U(1)$  interaction acts on both right-handed singlets and left-handed doublets. This means that fields are invariant under the following transformation:

$$\Psi_L \rightarrow e^{i\alpha^j(x)\tau_j + i\beta(x)Y} \Psi_L \quad (1.10)$$

$$\Psi_R \rightarrow e^{i\beta(x)Y} \Psi_R \quad (1.11)$$

where  $\tau_j$  are the three generators of  $SU(2)_L$ , the weak isospin, and  $Y$  is the generator of  $U(1)_Y$ , the weak hypercharge.  $SU(2)_L$  and  $U(1)_Y$  cannot be considered separately, since the two components of the doublets have distinct electric charges, which also means that  $U(1)_Y$  is distinct from the electromagnetic interaction  $U(1)_{e.m.}$ . This translates into the Gell-Mann-Nishijima relation:

$$Q = T_3 + \frac{Y}{2}, \quad (1.12)$$

where  $T_3$  is the third component of the weak isospin and  $Y$  is the weak hypercharge.

Spinors describing matter are of two types. *Leptons* are distinct from *quarks* as they are not sensitive to the strong interaction. The left-handed particles are arranged in  $SU(2)$  doublets while right-handed particles are  $SU(2)$  singlets. Each fermion has right-handed and left-handed representations, except for neutrinos, for which only the left-handed component is present in the theory. A right-handed neutrino cannot couple to gauge and matter fields, and cannot be observed as such. Furthermore, the neutrino is massless by construction since any mass term couples right-handed and left-handed components<sup>2</sup>. We are therefore left with the representations of Table 1.2. Each representation will be written with a prime to stress that they are interaction eigenstates, possibly distinct from mass eigenstates.

### 1.3 The Standard-Model Lagrangian

The Standard-Model Lagrangian is built so as to respect the aforementioned symmetries<sup>3</sup>. The possible terms are further restricted by imposing the renormalizability of the Lagrangian. It can be divided into four terms:

$$\mathcal{L} = \mathcal{L}_{gauge} + \mathcal{L}_{Higgs} + \mathcal{L}_{fermions} + \mathcal{L}_{Yukawa}. \quad (1.13)$$

In this expression, the gauge term is constituted by the previously introduced gauge tensors. It gives their dynamic to gauge bosons and induces quartic

<sup>2</sup>The neutrino is now known to be massive, and neutrino oscillations have been observed. This can be considered by introducing right-handed neutrinos, which extends the Standard Model. This subject will not be discussed further in the present dissertation as it has no impact on the phenomenology of the Higgs sector.

<sup>3</sup>We will follow the approach adopted for example in reference [2].

Table 1.2: The fermions of the Standard Model. Fermions are grouped in generations (or families) to form singlets or doublets of  $SU(2)_c$ . The first column shows the tensor notation conventionally used. The next columns show the third component of weak isospin, the weak hypercharge, the charge and the particle content.

Notation	$T_3$	$Y$	$Q$	Content
$L'_L$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	-1	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$\begin{pmatrix} \nu'_{eL} \\ e'_L \end{pmatrix}$ $\begin{pmatrix} \nu'_{\mu L} \\ \mu'_L \end{pmatrix}$ $\begin{pmatrix} \nu'_{\tau L} \\ \tau'_L \end{pmatrix}$
$E'_R$	0	-2	-1	$e'_R$ $\mu'_R$ $\tau'_R$
$Q'_L$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\frac{1}{3}$	$\begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} u'_L \\ d'_L \end{pmatrix}$ $\begin{pmatrix} c'_L \\ s'_L \end{pmatrix}$ $\begin{pmatrix} t'_L \\ b'_L \end{pmatrix}$
$U'_R$	0	$\frac{4}{3}$	$\frac{2}{3}$	$u'_R$ $c'_R$ $t'_R$
$D'_R$	0	$-\frac{2}{3}$	$-\frac{1}{3}$	$d'_R$ $s'_R$ $b'_R$

and trilinear couplings that are subject to experimental tests. It is given by

$$\mathcal{L}_{gauge} = -\frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} - \frac{1}{4}W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu}. \quad (1.14)$$

On the other hand, the minimal Higgs sector, discussed more extensively in Section 1.4, introduces a new complex scalar (Higgs) doublet  $\phi$ . It has no color charge and will therefore not affect the  $SU(3)_C$  sector:

$$\phi = \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix}. \quad (1.15)$$

The lower component is neutral by convention. This scalar doublet enters the Lagrangian as:

$$\mathcal{L}_{Higgs} = |(\partial_\mu + igW_\mu^i \frac{\tau^i}{2} + ig'B_\mu \frac{1}{2})\phi|^2 - \mu^2 |\phi|^2 - \lambda(|\phi|^2)^2. \quad (1.16)$$

This term, also sometimes called the symmetry breaking term ( $\mathcal{L}_{SB}$ ), is responsible for the dynamical symmetry breaking of  $SU(2)_L \otimes U(1)_Y$ . It describes the evolution of the  $\phi$  field in a quadratic potential. This opens the possibility for a non-zero vacuum expectation value for  $\phi$ , if the potential has a global minimum for the right value of the field, which in turn will give their mass to gauge bosons via the covariant derivative.

Fermions are introduced in a series of terms that describe their interaction with gauge fields. These terms are built from covariant derivatives obtained using the gauge tensors introduced in Table 1.1. In this construction, leptons are set to be sensitive to the electroweak interaction only, by terms involving  $W_\mu$  and  $B_\mu$ , while quarks are also subject to the strong interaction, since gluon

fields  $G_\mu$  enter the corresponding covariant derivative.

$$\begin{aligned}
\mathcal{L}_{fermions} = & i\overline{L}'_L \gamma^\mu (\partial_\mu + igW_\mu^i \frac{\tau^i}{2} + ig'B_\mu \frac{-1}{2}) L'_L \\
& + i\overline{E}'_R \gamma^\mu (\partial_\mu + ig'B_\mu \frac{-2}{2}) E'_R \\
& + i\overline{Q}'_L \gamma^\mu (\partial_\mu + ig_s G_\mu^a \frac{\lambda^a}{2} + igW_\mu^i \frac{\tau^i}{2} + ig'B_\mu \frac{1}{6}) Q'_L \\
& + i\overline{U}'_R \gamma^\mu (\partial_\mu + ig_s G_\mu^a \frac{\lambda^a}{2} + ig'B_\mu \frac{2}{3}) U'_R \\
& + i\overline{D}'_R \gamma^\mu (\partial_\mu + ig_s G_\mu^a \frac{\lambda^a}{2} + ig'B_\mu \frac{-1}{3}) D'_R.
\end{aligned} \tag{1.17}$$

At this level, right-handed and left-handed fields are introduced separately. It can be seen from Equation (1.8) that any mass term mixes right- and left-handed components, so that fermions are still massless. This is not the case anymore when introducing the Yukawa term,

$$\begin{aligned}
\mathcal{L}_{Yukawa} = & -\overline{L}'_l \lambda_l \phi E'_R - \overline{E}'_R \lambda_l^\dagger \phi^\dagger L'_L \\
& -\overline{Q}'_d \lambda_d \phi D'_R - \overline{D}'_R \lambda_d^\dagger \phi^\dagger Q'_L \\
& -\overline{Q}'_u \lambda_u (i\tau_2 \phi^*) U'_R - \overline{U}'_R \lambda_u^\dagger (i\tau_2 \phi^*)^\dagger Q'_L.
\end{aligned} \tag{1.18}$$

All of these Yukawa terms involve the Higgs field, or its complex conjugate. It is a particularity of the dimension 2 representation of  $SU(2)_L$ , which is pseudo-real. It means that the complex conjugate transforms also as a doublet, if one applies a unitary transformation, which is here  $i\tau_2$ . This is a key feature of the theory, since this term will give, after spontaneous symmetry breaking, their mass to leptons and quarks. The same doublet can therefore give their mass to up-type and down-type fermions.

## 1.4 Giving mass to particles

Since right-handed fermions are  $SU(2)$  singlets while left-handed fermions are grouped in  $SU(2)$  doublets, Standard-Model symmetries impose zero fermion masses, as they impose zero boson masses. In order to allow massive particles, one has to dynamically break the  $SU(2)_L \otimes U(1)_Y$  symmetry in such a way that the photon remains massless. The Higgs mechanism[3], introduced simultaneously by Brout, Englert and Higgs, is the simplest and most natural way to restore the experimentally observed masses. This justifies the scalar doublet introduced in the theory in Equation (1.16). The Higgs mechanism will generate a non-invariant vacuum and consequently give mass to W and Z bosons.

For  $\mu^2 < 0$  and  $\lambda > 0$ , the potential introduced in Equation (1.16) has the “bottle-end” form depicted on Figure 1.1. It has a circle of non-zero degener-

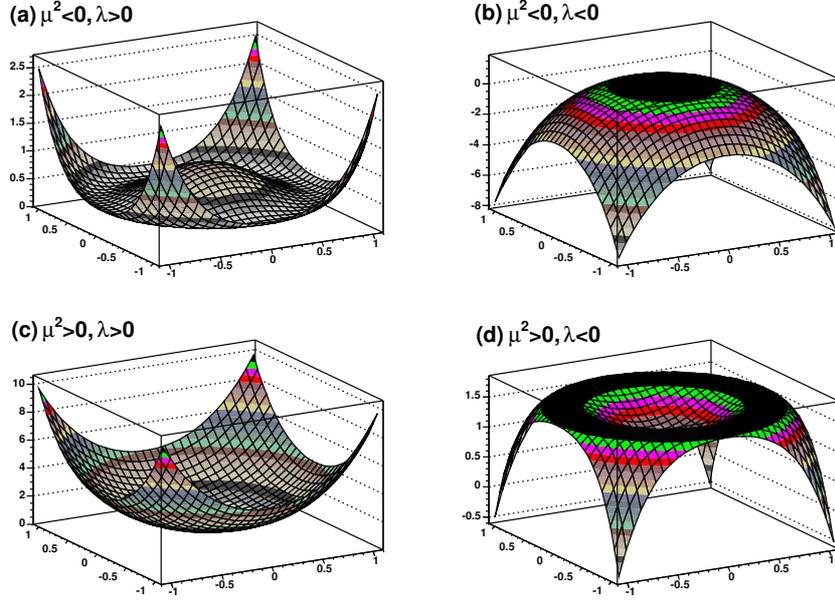


Figure 1.1: Scalar Higgs potential from Equation (1.16) as a function of the components of a complex scalar field. The four sign combinations are shown. (a) corresponds to the Standard-Model choice. There is a circle of minima at  $|\phi| = v/\sqrt{2}$ . Plot (b) shows a vacuum stability problem, while (c) does not break the symmetry as the minimum is at the origin. Case (d) combines the drawbacks of (b) and (c).

ated minima at

$$|\phi| = \sqrt{\frac{-\mu^2}{2\lambda}} \equiv \frac{1}{\sqrt{2}}v, \quad (1.19)$$

where  $v$  governs the vacuum value of the scalar field. The ground state of the Higgs field can be chosen as

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (1.20)$$

so that only the neutral component of the Higgs field features a non-zero mean value. This leaves the  $U(1)_{e.m.}$  symmetry unbroken.

The boson mass spectrum is extracted from the derivative part of the Higgs Term (1.16) of the Lagrangian:

$$D_\mu \langle \phi \rangle = \frac{i}{2\sqrt{2}} \begin{pmatrix} \sqrt{2}gvW_\mu^+ \\ -(gW_\mu^3 - g'B_\mu)v \end{pmatrix}. \quad (1.21)$$

Putting in the bosonic physical states definitions of Equations (1.5) to (1.7),

one reads the boson masses

$$m_W = \frac{1}{2}gv, \quad m_Z = \frac{1}{2}\sqrt{g^2 + g'^2}v, \quad m_\gamma = 0. \quad (1.22)$$

Once introduced, this field can enter the Standard-Model Lagrangian in terms of Form (1.18) and give mass to fermions. Substituting the mean field  $\langle\phi\rangle$  from Equation (1.20) into Term (1.18) of the Lagrangian, fermion masses are found to be proportional to Yukawa couplings:

$$m_f = \frac{1}{\sqrt{2}}v\lambda_f. \quad (1.23)$$

Yukawa couplings  $\lambda_f$  result from the diagonalization of the  $\lambda_{(u,d,l)}$  complex 3x3 matrices in Expression (1.18), that can be performed via a biunitar transformation. For leptons, transformation matrices can be chosen to cancel each other, since all three neutrinos are massless. For quarks, this will generate the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which relates interaction to mass eigenstates. The phase of this complex matrix will also induce a CP violation. A consequence of Expression (1.23) is that the mass hierarchy observed is explained by a corresponding hierarchy of the Yukawa couplings. The hierarchy of couplings remains unexplained, which can be seen as an unnatural and unpleasant feature of the theory. In addition, one *a-priori* undetermined parameter,  $\lambda_i$ , is introduced for each fermion mass.

## The Higgs particle

Another consequence of the Higgs mechanism that has just been introduced is the appearance of a new neutral spin-0 particle, the Higgs boson. Infinitesimal variations around the ground state of the Higgs field  $\phi(x)$  can be parametrized as

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_a(x) + i\phi_b(x) \\ v + H(x) + i\phi_0(x) \end{pmatrix}. \quad (2.1)$$

Out of the four newly introduced degrees of freedom, three are absorbed to give mass to gauge bosons. Only  $H(x)$  remains as a massive Higgs mode. Introducing (2.1) into the Standard-Model Lagrangian Higgs Term (1.16) yields the tree-level Higgs boson mass:

$$m_H^2 = -2\mu^2 = 2\lambda v^2. \quad (2.2)$$

While the  $v$  parameter is measured via the gauge boson masses (e.g. at LEP via the  $W$  mass measurement), and is found to be  $v = 246$  GeV [4], the  $\lambda$  parameter remains undetermined. The Higgs boson mass is consequently a free parameter of the theory.

Other properties of the Higgs boson will be quickly reviewed in Section 2.1. This includes the Higgs boson (partial) width(s) and couplings. Three types of constraints will then be discussed. On the one hand, some theoretical limits based on unitarity, triviality and vacuum stability concepts can be put. They will be addressed in Section 2.2. On the other hand, the negative outcome of direct searches based on collider data allows us to put a lower limit on the Higgs boson masses. Limits obtained at LEP will be discussed in Section 2.3. Finally, precision measurements during the last years, combined with (next-to-)next-to-leading-order ((N)NLO) calculations, permit us to put stringent limits on Higgs boson mass within the Standard Model. This electroweak fit will be briefly covered in Section 2.4.

## 2.1 Higgs boson properties

### Couplings to fermions and bosons

The coupling of the Higgs boson to fermions is obtained by introducing Higgs Field (2.1) into the Yukawa part (Equation (1.18)) of the Lagrangian. One then reads

$$g_{Hff} = \frac{m_f}{v}. \quad (2.3)$$

The coupling of the Higgs boson to gauge bosons arises from the derivative part of (1.16). Substituting the Higgs field by means of (2.1) and concentrating on terms linear in  $H(x)$ , it is obtained that:

$$g_{HVV} = \delta_V \frac{m_V^2}{v}, \quad (2.4)$$

where  $V$  stands for  $W$  or  $Z$ ,  $\delta_W = 2$  and  $\delta_Z = 1$ .

It follows from the above expressions that fermions' decay modes will contribute in the low-mass region ( $m_H < 150 \text{ GeV}/c^2$ ) where the Higgs boson will mainly decay into  $b$  or  $c$  quarks and taus. Once the decay into pairs of bosons is allowed, it quickly dominates, as the coupling to vector bosons goes like the mass squared. This behavior is shown in Figure 2.1, where NLO QCD and QED corrections are taken into account.

The Higgs boson does not couple to photons or gluons at tree level. Such a decay can arise via fermion or  $W$  loops and contributes in the low mass region.

### Branching ratios

Since the couplings of the Higgs boson to fermions are proportional to the fermions' mass, the decay branching ratio to any fermion  $f$  is proportional to  $m_f^2$ . At tree level, the partial width of fermionic channels is therefore

$$\Gamma(H \rightarrow f\bar{f}) = \frac{N_c g^2 m_f^2}{32\pi m_W^2} \beta_f^3 m_H, \quad (2.5)$$

where  $N_c$  is a color factor that is 1 for leptons and 3 for quarks, and  $\beta_f^2 = 1 - 4m_f^2/m_H^2$  is the phase-space factor.

The partial width of bosonic channels is

$$\Gamma(H \rightarrow VV) = \frac{g^2}{128\pi} \frac{m_H^3}{m_W^2} \delta_V (1 - \tau_V + \frac{3}{4}\tau_V^2) \beta_V, \quad (2.6)$$

where  $\tau_V = 4m_V^2/m_H^2$ .

In the high mass regime, the total decay width of  $H$  will be dominated by  $\Gamma(H \rightarrow WW)$  and  $\Gamma(H \rightarrow ZZ)$ . Also,  $\tau_V \rightarrow 0$  and  $\beta_V \rightarrow 1$ , so that it can be approximated as

$$\Gamma_{tot} \simeq \Gamma(H \rightarrow WW) + \Gamma(H \rightarrow ZZ) \simeq \frac{3m_H^3}{32\pi v^2}. \quad (2.7)$$

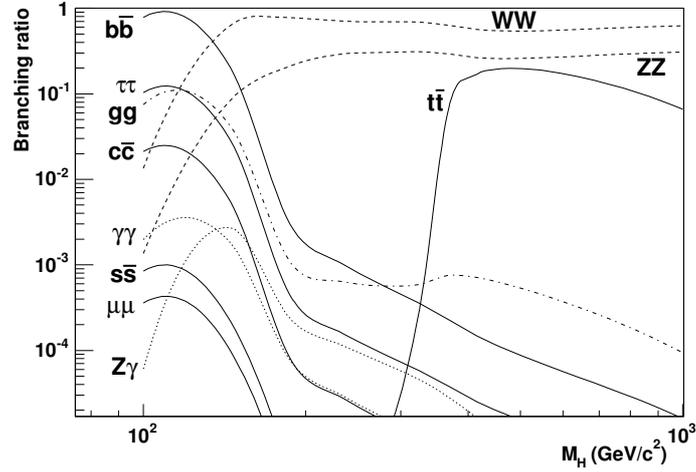


Figure 2.1: Branching ratios of the Standard Model Higgs boson, as a function of the Higgs boson mass. Curves include NLO QCD and QED corrections. Obtained with hdecay[5].

This gives a total decay width of  $\Gamma_{tot} \simeq m_H$  when  $m_H \simeq 1.4 \text{ TeV}$ . It then becomes experimentally very difficult to separate the Higgs resonance from the  $VV$  continuum, and difficult to interpret the Higgs field excitations as particles. It will be shown in the next section that such a Higgs boson is also less relevant. The behavior of  $\Gamma_{tot}$  is depicted in Figure 2.2.

## 2.2 Theoretical limits

A first limit on the Higgs boson mass can be made by looking at the scattering of weak gauge bosons in the high energy limit. Using the current value for the vacuum value  $v$  of the Higgs sector, this calculation shows that without the Higgs boson, this amplitude would violate unitarity.

Following the idea sketched by J. Bagger et al. [6], this can be easily understood. The complete set of Feynman diagrams for  $WW$  scattering are presented Figure 2.3. In the high energy limit, when both the Higgs boson mass and the center-of-mass energy are larger than the  $W$  and  $Z$  boson masses, diagrams involving the Higgs boson will dominate. The scattering matrix element can therefore be written in terms of the Mandelstam variables,  $s$ ,  $t$  and  $u$ , as

$$\mathcal{M}(W_L^+ W_L^- \rightarrow W_L^+ W_L^-) = A(s) + A(t), \quad (2.8)$$

$$\mathcal{M}(W_L^+ W_L^- \rightarrow W_L^- W_L^+) = A(s) + A(u), \quad (2.9)$$

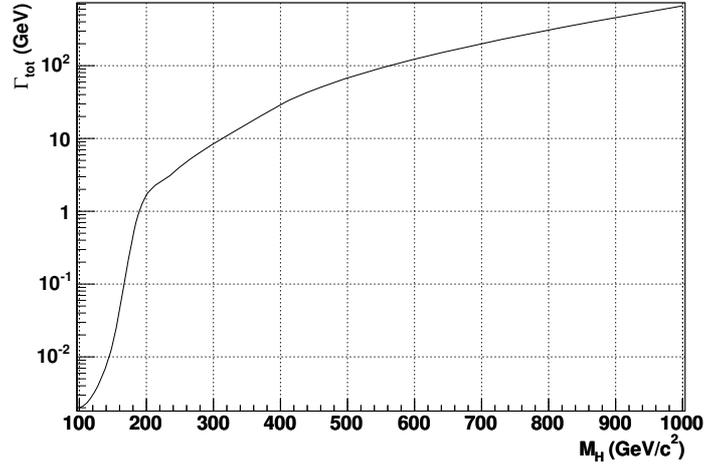


Figure 2.2: Higgs boson decay width as a function of the Higgs boson mass. In the high mass regime, the curve shows the typical  $m_H^3$  behavior presented in the text. Obtained with hdecay[5].

$$\mathcal{M}(Z_L Z_L \rightarrow Z_L Z_L) = A(s) + A(t) + A(u), \quad (2.10)$$

$$\mathcal{M}(W_L^+ W_L^- \rightarrow Z_L Z_L) = A(s), \quad (2.11)$$

with the scattering amplitude given by

$$A(s) = -\frac{m_H^2}{v^2} \left( 1 + \frac{m_H^2}{s - m_H^2 + i \frac{3m_H^4}{32\pi v^2} \Theta(s)} \right). \quad (2.12)$$

In this expression proportional to the Higgs boson coupling to  $WW$  squared, one sees the Higgs boson propagator. The complex part of the denominator is the product of the width (from Equation (2.7)) and the mass. The  $\Theta(s)$  function is equal to 1 for  $s > 0$  and equal to 0 otherwise. It is there to remove the width effect in the  $t$  (and  $u$ ) channel. This can be re-expressed in the weak isospin basis using Clebsch-Gordan coefficients,

$$T(0) = 3A(s) + A(t) + A(u), \quad (2.13)$$

$$T(1) = A(t) - A(u), \quad (2.14)$$

$$T(2) = A(t) + A(u). \quad (2.15)$$

In order not to violate unitarity, it is required that the scattering amplitude remains below one for all weak isospin states and all partial waves. The simple form of the lowest order Legendre polynomial makes the computation of

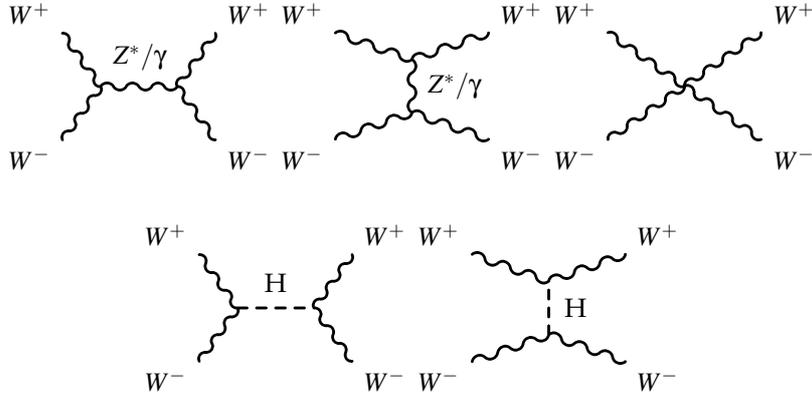


Figure 2.3: Complete set of Feynman diagrams for WW scattering.

order-0 partial waves easy. Defining

$$a_l^J = \frac{1}{64\pi} \int_{-1}^1 d(\cos\theta) P_l(\cos\theta) T(J), \quad (2.16)$$

the first partial wave is found to be

$$a_0^0 = -\frac{3m_H^2}{32\pi v^2} \left( 1 + \left( \frac{s}{m_H^2} - 1 + i\frac{3m_H^2}{32\pi v^2} \right)^{-1} \right). \quad (2.17)$$

Requiring the partial wave unitarity<sup>1</sup>  $|a_0^0| < 1/2$  for all values of  $s$  sets the limit

$$m_H < \sqrt{\frac{16\pi}{3}} v \sim 1.0 \text{ TeV}/c^2 \quad (2.18)$$

Nevertheless, this limit cannot be taken too strictly as a perturbative partial-waves development has been driven into a non-perturbative regime where  $|a_0^0| = \mathcal{O}(1)$ .

More stringent limits on the Higgs boson mass can be applied by studying the running of the self-interaction  $\lambda$ . Using the renormalization group equations, the running of the coupling  $\lambda$  is found to be<sup>2</sup>

$$\frac{d\lambda}{dt} = \frac{1}{16\pi^2} (12\lambda^2 + 6\lambda g_t^2 - 3g_t^4 - \frac{3}{2}(3g^2 + g'^2) + \frac{3}{16}(2g^4 + (g^2 + g'^2)^2)) + \dots, \quad (2.19)$$

<sup>1</sup>Unitarity of the two-body elastic scattering matrix translates into  $|a_l^J - i/2| = 1/2$ . See for example [6].

<sup>2</sup>This can be found for example in [7], where it is presented together with coherent conventions for the vacuum expectation value (Equation (1.20)), which is not the case for all publications in the field.

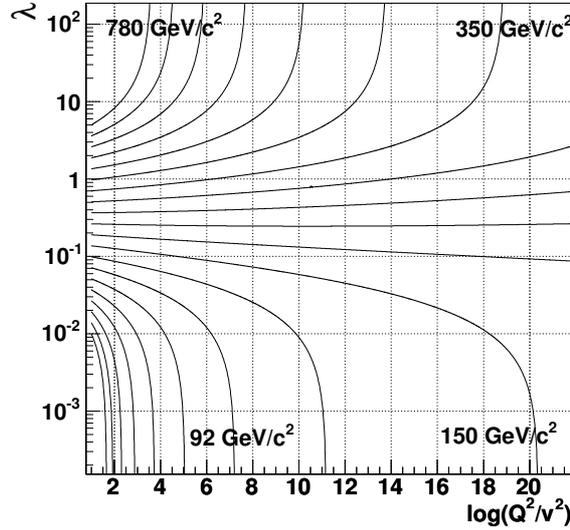


Figure 2.4: Running of  $\lambda$  from Equation (2.19) for various initial values. One can see how low initial values of  $\lambda$  which correspond to the low Higgs boson masses have a zero for large energy scales and how higher Higgs boson masses lead to a Landau pole. The mass quoted at the end of some curves corresponds to the corresponding initial Higgs boson mass.

where  $t$  is the normalized energy scale ( $t = \log(Q^2/v^2)$ ) and  $g_t$  is the top coupling constant ( $g_t = m_t/v$ ). Dots stand for higher-order contributions. According to this equation, the behavior of  $\lambda$  at high energy will depend of its initial value  $\lambda(t=0)$ .

If this initial value is too high, a Landau pole will appear at some energy scale  $Q$ . There must be some new physics at a lower scale so that the evolution equation gets modified, and the self-interaction remains finite. This is what is called the triviality limit. On the other hand, if the initial  $\lambda$  value is too low, the derivative becomes negative, and  $\lambda$  itself becomes negative at some scale  $Q_0$ . As it has been emphasized before, a negative value of  $\lambda$  does not allow for vacuum stability, as the electroweak symmetry breaking potential has no minimum anymore. This behavior is depicted Figure 2.4.

If the Standard Model is supposed to remain valid up to an energy scale  $\Lambda$ , there must be no Landau pole below that scale, and  $\lambda$  must remain positive. This translates respectively into an upper and lower bound on  $\lambda(t=0)$  and therefore on the Higgs boson mass. Depending on the energy scale  $\Lambda$  at which new physics enters, a different domain is allowed for the Higgs boson mass. If this new-physics scale is set at 1 TeV, the Higgs boson mass

must be in the range  $[55 \text{ GeV}/c^2, 700 \text{ GeV}/c^2]$ . If, on the contrary, the Standard Model is said to be valid up to the Planck scale, it is even more constrained on  $[130 \text{ GeV}/c^2, 190 \text{ GeV}/c^2]$ , as can be seen in Figure 2.5.

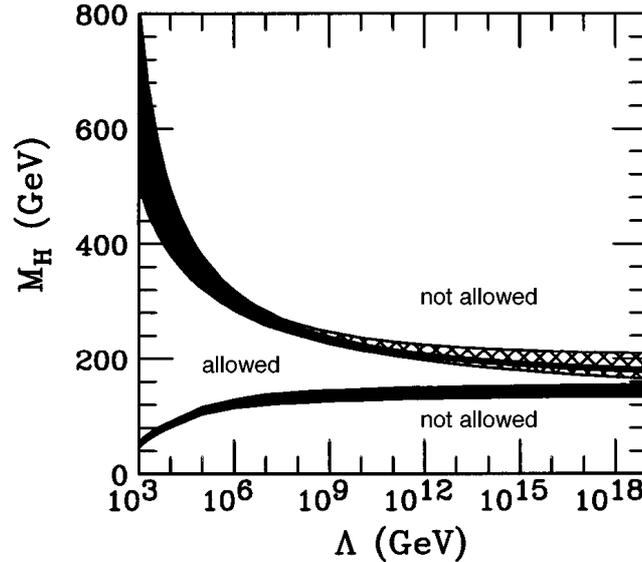


Figure 2.5: Allowed Higgs boson mass as a function of the energy at which the Standard Model breaks down (from [8]). The solid area indicates the sum of theoretical uncertainties in the  $m_H$  upper bound when keeping  $m_t = 175 \text{ GeV}/c^2$  fixed. The cross-hatched area shows the additional uncertainty from  $m_t$ .

## 2.3 Experimental limits

### 2.3.1 Standard Model

There is also much activity on the experimental front. Only the direct observation of the Higgs boson will constitute a proof of its existence. The LEP Higgs Working Group combines results of the direct searches for the Higgs boson at LEP. These direct searches lead to the conclusion that the Higgs boson, if it presents the properties predicted by the Standard Model, must be heavier than  $114.4 \text{ GeV}/c^2$  at the 95% confidence level [9]. It is worth looking more closely into the details of this limit.

Data used to obtain this result span the range of center-of-mass energies from  $189 \text{ GeV}$  to  $209 \text{ GeV}$ . The integrated luminosity of the data samples is given in Table 2.1. The main production channel at LEP, in the energy range corresponding to the last years of running, is the Higgsstrahlung mechanism  $e^+e^- \rightarrow HZ$ . In this process, depicted Figure 2.6(a), a virtual Z boson emits

Table 2.1: Integrated luminosities of the data samples of the four experiments and their sum (LEP). The subsets taken at energies exceeding 206 GeV and 208 GeV are listed separately.

Integrated luminosities ( $pb^{-1}$ )					
	ALEPH	DELPHI	L3	OPAL	LEP
$\sqrt{s} \geq 189$ GeV	629	608	627	596	2461
$\sqrt{s} \geq 206$ GeV	130	138	139	129	536
$\sqrt{s} \geq 208$ GeV	7.5	8.8	8.3	7.9	32.5

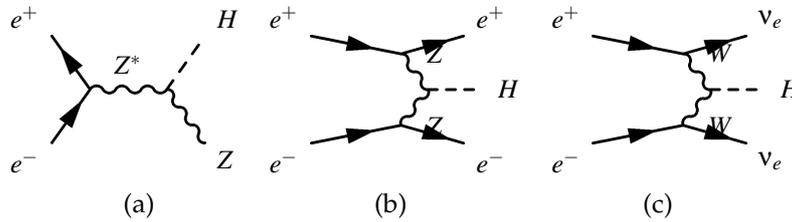


Figure 2.6: Feynman diagrams entering Higgs production at LEP. (a) shows the dominant Higgsstrahlung; (b) and (c) show boson-fusion processes.

a Higgs boson. The final state therefore contains decay products of both a Z boson and a Higgs boson. Contributions from boson-fusion processes (represented in Figure 2.6(b,c)) remain small. For Higgs boson masses kinematically available at LEP, the Higgs boson is expected to decay predominantly into a  $b\bar{b}$  pair, so that the final state will contain two b-jets and either two leptons, two (b)jets or missing momentum. Decays of the Higgs boson into  $\tau^+\tau^-$  pairs were also considered.

Various analyses from the four LEP collaborations (ALEPH, DELPHI, L3, OPAL) were combined using a semi-Bayesian likelihood ratio method [9]. Each data configuration is tested in both the background-only and the signal-and-background hypothesis. From these two hypotheses, a test statistic  $Q$  is constructed. This method has been used in this thesis and will be described in more detail in Section 4.4.

Figure 2.7 shows the behavior of the test statistic  $-2\ln Q$  as a function of the Higgs boson mass, for combined LEP data. This curve is compared to the expected behavior in both the background-only and the signal-and-background hypothesis. These two expected curves, and the associated uncertainty bands, are obtained by replacing data by Monte-Carlo generated events in a large number of toy experiments. It is interesting to note that a general upward variation of the background is observed, with a broad minimum of  $-2\ln Q$  around  $115 \text{ GeV}/c^2$ . Back in 2000, this excess of events has been at the origin of many discussions, when it was attributed to a Higgs boson signal. It was

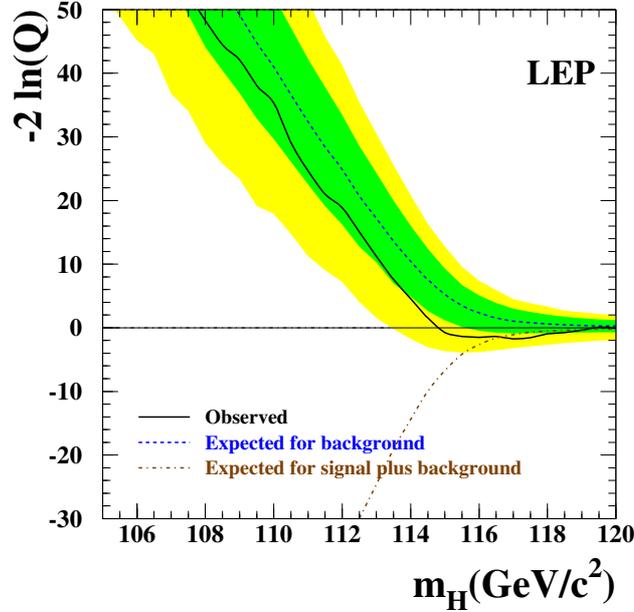


Figure 2.7: Observed and expected behavior of the test statistic  $-2\ln Q$  as a function of the test mass  $m_H$ , obtained by combining data of the four LEP experiments. The full curve represents the observation; the dashed curve shows the median background expectation, and is surrounded by the 68% and 95% probability bands. The dash-dotted curve indicates the expectation for  $-2\ln Q$ , given the signal plus background hypothesis. From [9].

estimated to be over three standard deviations, but has been reestimated to less than two since then.

The expected and observed lower bounds on the Standard Model Higgs boson mass obtained at the 95% confidence level (C.L.) by combining the four LEP experiments are respectively  $115.3 \text{ GeV}/c^2$  and  $114.4 \text{ GeV}/c^2$ .

It is important to notice that this limit is extracted in the context of the Standard Model. There are extensions, or alternative models, where the observability of a Higgs boson is modified, either because of the production cross-section, either because of different couplings. One example is found in so-called two-Higgs-doublets models.

### 2.3.2 Two-Higgs-doublets models

The simplest extension of the Higgs sector of the Standard Model is obtained by adding a second Higgs doublet, with the same quantum numbers as the first one (2HDM). This is the most theoretically favorable non-minimal scalar

sector, as this naturally keeps the ratio  $\rho_0 \equiv m_W^2 / (m_Z^2 \hat{c}_Z^2 \hat{\rho}) \approx 1$ , where  $\hat{c}_Z^2$  is the cosine of  $\theta_W$  obtained in the  $\overline{\text{MS}}$  scheme, and  $\hat{\rho} = 1 + \rho_t$  contains the quadratic  $m_t$  dependence [10]. Indeed, the general formula at tree level is

$$\rho = \frac{\sum_{T,Y} [4T(T+1) - Y^2] |v_{T,Y}|^2 c_{T,Y}}{\sum_{T,Y} 2Y^2 |v_{T,Y}|^2}, \quad (2.20)$$

where  $v_{T,Y}$  is the vacuum expectation value of the field with hypercharge  $Y$  and isospin  $T$ , and  $c_{T,Y}$  is a multiplicative factor equal to 1 for complex representations and 1/2 for the real ones. From Equation (2.20),  $\rho$  is automatically equal to 1 at tree level, for any number of Higgs doublets (and singlets).

A first motivation for introducing a second doublet in the model comes from the hierarchy of couplings observed in Equation (1.18). In the third generation of leptons, there is a large ratio between the top mass and the bottom mass. As both masses must come from the same Higgs doublet, the hierarchy translates into a hierarchy of Yukawa couplings. If bottom and top quarks acquire their masses from different doublets, the hierarchy of their Yukawa couplings could be more natural. Another motivation comes from the possibility to introduce CP violation in the Higgs sector, either spontaneously or explicitly. However, we shall restrict our discussion to a CP-conserving framework. A last motivation comes from the natural controlled appearance of flavor changing neutral currents (FCNC). The addition of a second doublet induces tree-level FCNC which are forbidden within the Standard Model. This makes such a 2HDM peculiarly suited to the study of such rare processes. On the other hand, FCNC can be avoided in the general framework of multi-doublet models by coupling all fermions of a given electric charge to one and only one Higgs doublet [11].

As explained above, a second doublet with the same quantum numbers is introduced. It is then convenient to generalize (2.1) to

$$\begin{aligned} \phi_1(x) &= \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_{1a}(x) + i\phi_{1b}(x) \\ v_1 + h_1(x) + i\phi_{10}(x) \end{pmatrix}; \\ \phi_2(x) &= \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_{2a}(x) + i\phi_{2b}(x) \\ v_2 e^{i\xi} + h_2(x) + i\phi_{20}(x) \end{pmatrix}. \end{aligned} \quad (2.21)$$

In that expressions,  $\xi$  is a phase that may be a source of CP-violation in some circumstances. It will be assumed to be zero as we will concentrate on the CP-conserving case<sup>3</sup>.

Higgs mass eigenstates and Higgs boson self-interactions are determined by the very structure of the Higgs potential, which is not unique, even if it is heavily constrained by renormalizability and CP invariance. In all cases, a Higgs model incorporating two doublets of complex scalar fields generates

<sup>3</sup>A comprehensive overview of the conditions for the 2HDM to be CP-invariant can be found in [12].

five scalar Higgs bosons, three of which are neutral ( $A^0$ ,  $H^0$  and  $h^0$ ).  $H^0$  and  $h^0$  are CP-even states very similar to the Standard-Model Higgs boson, while  $A^0$  is a CP-odd states that originates from the complex part of the neutral component of each doublet. These mass eigenstates are obtained from the gauge eigenstates defined in (2.21) by a linear transformation

$$\begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}}(\phi_{1a}(x) + i\phi_{1b}(x)) \\ \frac{1}{\sqrt{2}}(\phi_{2a}(x) + i\phi_{2b}(x)) \end{pmatrix} = \begin{pmatrix} G^+ \\ H^+ \end{pmatrix} \quad (2.22)$$

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} H^0 \\ h^0 \end{pmatrix} \quad (2.23)$$

$$\begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \phi_{10} \\ \phi_{20} \end{pmatrix} = \begin{pmatrix} G^0 \\ A^0 \end{pmatrix} \quad (2.24)$$

where  $\tan\beta = v_2/v_1$ , and  $\alpha$  is a mixing angle for CP-even Higgs bosons that depends on the Higgs potential and can be expressed in function of its parameters. The kinetic sector is the natural extension of the Standard-Model:

$$\mathcal{L}_{Higgs}(2HDM) = \sum_{f=1,2} |(\partial_\mu + igW_\mu^i \frac{\tau^i}{2} + ig'B_\mu \frac{1}{2})\phi_f|^2 - V(\phi_1, \phi_2). \quad (2.25)$$

Gauge bosons masses and couplings with Higgs bosons are obtained by substituting Higgs fields mass eigenstates from Equations (2.21) to (2.24) into Equation (2.25). If  $v_1^2 + v_2^2 = v^2$ , the expressions for the gauge boson masses coincide with the Standard-Model ones. As  $v^2$  is a known parameter, a tree-level constraint is obtained on  $v_1$  and  $v_2$ . Developing the derivative term in (2.25), the couplings between the CP-even Higgs bosons and massive gauge bosons are found to be given by

$$g_{h^0VV} = \sin(\beta - \alpha)\delta_V \frac{m_V^2}{v}; \quad g_{H^0VV} = \cos(\beta - \alpha)\delta_V \frac{m_V^2}{v}. \quad (2.26)$$

That has to be compared to expression (2.4). An important consequence of this is that these couplings are suppressed with respect to the Standard-Model ones. If  $\cos(\beta - \alpha) = 0$ ,  $h^0$  becomes a Standard-Model-like boson, while  $H^0$  decouples from the bosons. The phenomenology of the 2HDM strongly depends of the Yukawa sector. We will restrict ourself to models avoiding FCNC, as it is experimentally strongly constrained. Four variants of two-Higgs-doublets models (2HDM) are then present. Table 2.2 shows the way Higgs doublets couple to fermions for each of those four categories. The numbers (1 or 2) show which doublet couples to which fermion type.

In type-I models, only one Higgs doublet ( $\phi_2$ ) couples to fermions, so the Yukawa term of the Lagrangian (Equation (1.18)) becomes

$$\begin{aligned} \mathcal{L}_{Yukawa}(\text{type I}) = & -\overline{Q}'_L \lambda_d \phi_2 D'_R - \overline{D}'_R \lambda_d^\dagger \phi_2^\dagger Q'_L \\ & -\overline{Q}'_L \lambda_u (i\tau_2 \phi_2^*) U'_R - \overline{U}'_R \lambda_u^\dagger (i\tau_2 \phi_2^*)^\dagger Q'_L \\ & + \text{leptonic sector}. \end{aligned} \quad (2.27)$$

Table 2.2: Higgs doublets coupling to fermions for each of the four considered types of 2HDM. For each type, the numbers in the first column (1 or 2) show which doublet couples to which fermion type. The second column shows the fermion couplings of the lightest CP-even Higgs boson, relative to those in the Standard Model.

	Model I		Model I'		Model II		Model II'	
	$\phi$	coupl.	$\phi$	coupl.	$\phi$	coupl.	$\phi$	coupl.
u-type quarks	2	$\frac{\cos\alpha}{\sin\beta}$	2	$\frac{\cos\alpha}{\sin\beta}$	2	$\frac{\cos\alpha}{\sin\beta}$	2	$\frac{\cos\alpha}{\sin\beta}$
d-type quarks	2	$\frac{\cos\alpha}{\sin\beta}$	2	$\frac{\cos\alpha}{\sin\beta}$	1	$-\frac{\sin\alpha}{\cos\beta}$	1	$-\frac{\sin\alpha}{\cos\beta}$
charged leptons	2	$\frac{\cos\alpha}{\sin\beta}$	1	$-\frac{\sin\alpha}{\cos\beta}$	1	$-\frac{\sin\alpha}{\cos\beta}$	2	$\frac{\cos\alpha}{\sin\beta}$

The couplings of Higgs bosons ( $h^0, H^0$ ) to fermions are obtained by expanding this in terms of mass eigenstates. The result for  $h^0$  is given in Table 2.2 relatively to the Standard-Model couplings. If the CP mixing angle  $\alpha$  tends to  $\pi/2$ , the heaviest CP-even neutral scalar provides mass only to the fermions and the other couples exclusively to the bosons, i.e. is a *fermiophobic* Higgs boson. In addition, if we further assume that  $\tan(\beta) \simeq 0$ ,  $H^0$  becomes also bosophobic, while  $h^0$  acquires Standard-Model-like couplings to vector bosons. One interesting scenario is obtained noticing that in the limit where  $\tan(\beta) \rightarrow 0$ ,  $m_W = gv_1/2$  while  $m_e = \lambda_e v_2/\sqrt{2}$ . It may be natural to postulate  $\lambda_e = g$ , so that  $\tan\beta \simeq \sin\beta = \sqrt{2}m_e/m_W = 8.6 \times 10^{-6}$ .

In models of type II, for some choices of parameters, one doublet couples exclusively to up-type quarks while the other couples only to down-type quarks and charged leptons:

$$\begin{aligned}
\mathcal{L}_{Yukawa}(\text{type II}) = & -\overline{Q}'_L \lambda_d \phi_1 D'_R - \overline{D}'_R \lambda'_d \phi_1^\dagger Q'_L \\
& -\overline{Q}'_L \lambda_u (i\tau_2 \phi_2^*) U'_R - \overline{U}'_R \lambda'_u (i\tau_2 \phi_2^*)^\dagger Q'_L \\
& + \text{leptonic sector.}
\end{aligned} \tag{2.28}$$

This Lagrangian can be reexpressed in a form invariant under the *supersymmetry*, which is not the case for the type I Lagrangian. A nice consequence of this model is that the hierarchy of couplings becomes more natural if  $\tan\beta \sim 35$  and  $\tan\alpha \sim 1$ . Amongst the other consequences of two-doublets models (both type I and type II), the coupling to fermions is enhanced, due to the smaller vacuum values of the Higgs fields.

Higgs boson mass limits strongly depend on the model hypothesis. The previously mentioned limit is only valid within the Standard Model; limits in 2HDM are generally weaker. Limits in the minimal supersymmetric model (MSSM) are very sensitive to the scenario. In the so-called *no mixing* scenario, the LEP limit is  $93.3 \text{ GeV}/c^2$  (strictly speaking for  $\tan\beta \geq 0.7$ , Figure 2.8) [13]. In this particular scenario, the Higgs boson masses are naturally low, so that limits are stringent. On the other hand, the limit on a fermiophobic Higgs boson

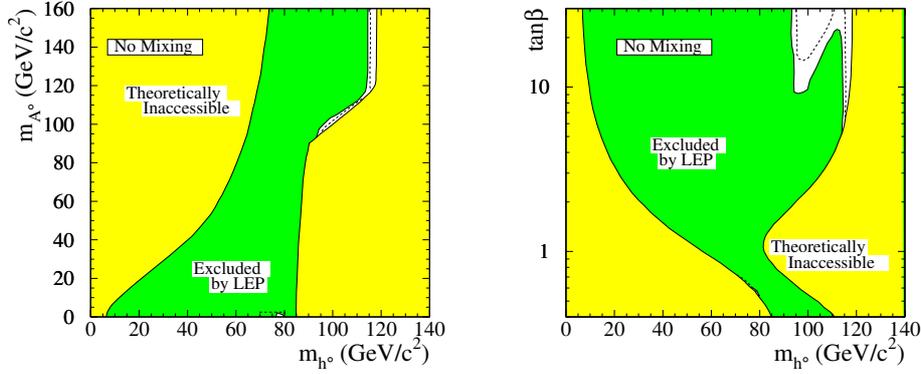


Figure 2.8: MSSM parameter exclusions, at 95% C.L., for the no-mixing benchmark scenario, with  $m_t = 179.3 \text{ GeV}/c^2$ . The figure shows the excluded and theoretically inaccessible regions as a function of the MSSM parameters in two projections:  $(m_h, m_A)$  and  $(m_h, \tan\beta)$ . The dashed lines indicate the boundaries of the regions expected to be excluded on the basis of Monte-Carlo simulations with no signal. From [13].

has been derived from decays of Higgs boson into two photons. The combined preliminary LEP limit obtained in reference [14] is  $109.7 \text{ GeV}/c^2$  (Figure 2.9). It is obtained considering a benchmark model where the Higgs boson couples only to bosons, with couplings equal to the Standard Model ones (meaning  $\alpha = \pi/2$ ,  $\tan\beta \simeq 0$  for a type I 2HDM). This limit could be improved by incorporating decays of the Higgs boson into a  $WW$  pair. This is the subject of Chapter 4.

## 2.4 Indirect measurements

The combinations from the LEP electroweak working group are used to perform stringent tests on the Standard Model, in particular on the Higgs mechanism, by comparing precise measurements with theory predictions. Data used include mainly LEP measurements but also other well established results, such as the top mass from Tevatron or SLD results at the Z peak.

It turns out that the Standard Model describes experiments pretty well. Furthermore, the masses of heavy fundamental particles such as the top quark and the W boson can be indirectly extracted from other quantities via the calculation of radiative corrections. Compared with direct measurements, this checks the validity of the theory.

The constraint on the mass of the Higgs boson is shown in Figure 2.10. It shows, as a function of the Higgs boson mass, the increase of  $\chi^2$  with respect to the best fit value, for a fit of precision electroweak measurements. Performed in the context of the Standard Model, this fit suggests a Higgs boson mass of  $117 \text{ GeV}/c^2$ , just above the limit from direct measurements. This

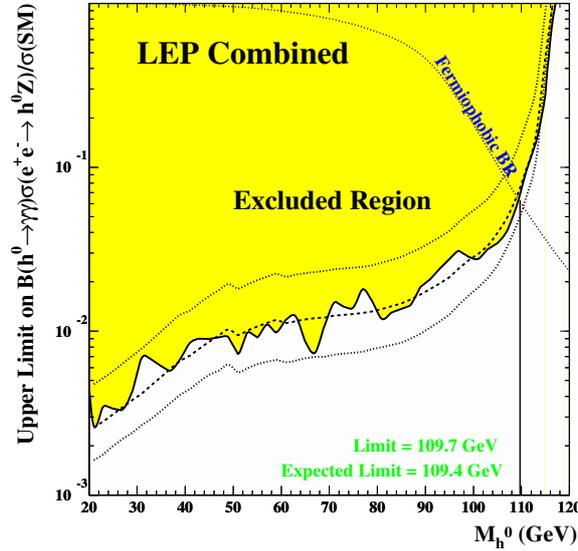


Figure 2.9: Combined LEP experimental limits for Higgs bosons decaying into diphotons. The 95% C.L. upper limits on  $B(h^0 \rightarrow \gamma\gamma) \times \sigma(e^+e^- \rightarrow h^0 Z^0)/\sigma(SM)$  is shown as a function of Higgs boson mass. Also shown (dotted line) is the branching fraction obtained for the benchmark fermiophobic model. The median expected limits and the  $\pm 2\sigma$  confidence level region are denoted by the dashed curves. The combined limit is indicated by the vertical line. From [14].

also gives an upper limit at 95% C.L. of  $260 \text{ GeV}/c^2$ . Theoretical uncertainties originating from uncalculated higher-order corrections are materialized by the shaded band. Also shown is the result (dashed curve) obtained when using an alternative recent estimation of the contribution of light quarks to the photon vacuum polarization ( $\Delta\alpha_{had}^{(5)}(m_Z^2)$ ). Among the experimental inputs which are used in the global fit, the  $W$  mass ( $m_W$ ) and the leptonic effective electroweak mixing angle ( $\sin^2 \theta_{eff}^{lept}$ ) have the largest impact on the extracted value of  $m_H$ . The dependence of the Higgs mass on both quantities can be observed in Figure 2.11. Other relevant inputs are the  $Z$  width and various asymmetries. The combined value for  $\sin^2 \theta_{eff}^{lept}$  includes hadronic forward-backward asymmetries in addition to the leptonic asymmetries from LEP and SLD. This is justified within the Standard Model as the hadronic asymmetries have a reduced sensitivity to the corrections specific to the quark vertex. Hadronic- and leptonic-based predictions differ by 2.9 standard deviations, which translates into a bad  $\chi^2$  (the corresponding probability is 8.4%).

The top quark mass ( $178.0 \pm 4.3 \text{ GeV}/c^2$ ) enters the global fit as a parameter; its value and its error have a strong impact on the extracted  $m_H$  value and presently induce the dominant uncertainty on the Higgs boson mass. The  $D0$  collaboration recently updated their top quark mass estimate from RUN I,

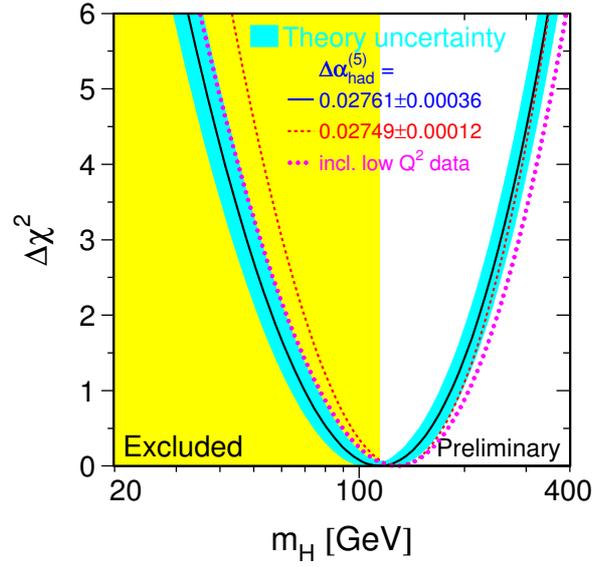


Figure 2.10:  $\Delta\chi^2 = \chi^2 - \chi_{min}^2$  vs.  $m_H$  curve. The line is the result of the fit using all precision electroweak data available in summer 04. The band represents an estimate of the theoretical error due to missing higher-order corrections. The vertical band shows the 95% C.L. exclusion limit on  $m_H$  from the direct search. The dashed curve is the result obtained using the evaluation of  $\Delta\alpha_{had}^{(5)}(m_Z^2)$ . From [15].

inducing a shift in the world average and therefore in the Higgs boson best-fitted mass [16]. This 21 GeV/ $c^2$  shift in the best fit Higgs boson mass is significant enough to make the best fit compatible with direct measurements, while previous data favored an excluded Higgs boson mass. This is included in Figure 2.10.

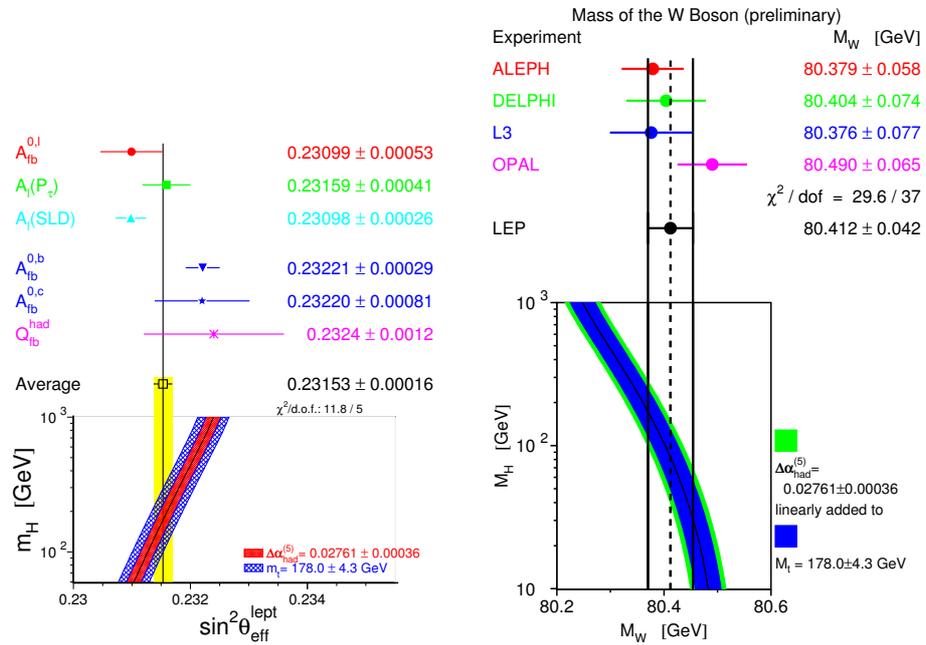


Figure 2.11: (a) Comparison of  $\sin^2 \theta_{\text{eff}}^{\text{lept}}$  measured from asymmetries with the Standard Model prediction as a function of  $m_H$ . (b) Comparison of  $m_W$  measured at LEP-2 and  $p\bar{p}$  colliders with the Standard Model prediction as a function of  $m_H$ . The measurement with its error is shown as the vertical band. The width of the Standard Model band is due to the uncertainties in  $\Delta\alpha_{\text{had}}^{(5)}(m_Z^2)$ ,  $\alpha_S(m_Z^2)$ ,  $m_Z$  and  $m_t$ . The total width of the band is the linear sum of these effects. From [15].

**Part II**

**ALEPH - ZH associated  
production**



## An Apparatus for LEP Physics

### 3.1 The LEP accelerator

LEP stands for Large Electron Positron Collider. With a 26.66 km circumference, this was the world's largest particle accelerator. It was located about 100 meters under the France-Switzerland border, near Geneva (Figure 3.1).

In order to bring electrons and positrons to their nominal energy, LEP operators benefit from CERN's previous accelerator chain. Electrons were first brought to 200 MeV by the *Linear Accelerator of LEP* (LIL). Some of them were shot on a tungsten target to produce positrons. Both electrons and positrons were then accelerated to 600 MeV by the *Electron-Positron Accumulator* (EPA). The *Proton Synchrotron* (PS) and the *Super Proton Synchrotron* (SPS) were then successively used to produce 23 GeV particles. After their injection, electrons and positrons were circulating in opposite directions. Bunches were crossing each other in four sites on the ring. Detectors were placed at the collision points to record all desired information.

The magnet configuration of an  $e^+e^-$  circular accelerator like LEP consists of dipoles and quadrupoles. Dipoles bend the beam, while quadrupoles (re)focus the bunches. Additional sextupoles and octupoles then contribute to beam stability, allowing for small corrections that keep the magnets' configuration away from resonant instability regimes. Acceleration is provided by a set of cavities distributed around the ring. This is necessary not only to bring electrons and positrons from the injection energy to the nominal energy, but also to compensate for the significant loss by synchrotron radiation that continuously damps the electrons' momentum on their path.

The number and technology of elements in the ring evolved as the goals were reevaluated in terms of beam energy. In the initial configuration, referred to as LEP1, beam energy was of the order of 46 GeV. This allowed scanning of the Z resonance peak. The LEP2 upgrade extended the center-of-mass energy, first to 161 GeV (this is the W era, dedicated to the study of the W boson pro-

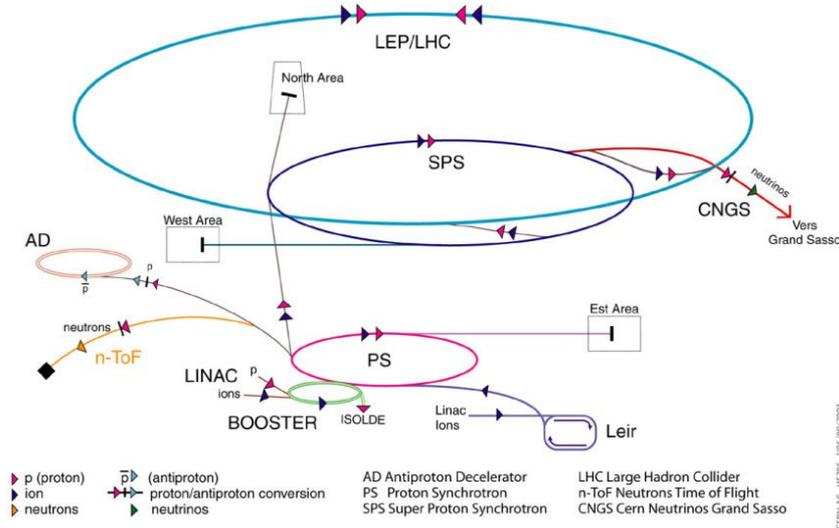


Figure 3.1: The CERN accelerators chain. The LEP is the main light-blue loop on top, in which electrons and positrons used to circulate. It is now being replaced by the LHC, discussed in Chapter 5.

duced in pairs), then up to 209 GeV in order to study new physics (in addition to WW physics).

As can be seen in Table 3.1, these upgrades are characterized by an increasing number of radio-frequency cavities, also shifting from a copper technology to a superconducting NbTi technology. This is the direct consequence of the increasing synchrotron radiation, which limits the accelerator performances.

The total radiation rate is obtained by the relativistic extension of the Larmor formula[17], which is for a circular accelerator:

$$P = \frac{dE}{dt} = -\frac{1}{4\pi\epsilon_0} \frac{2}{3} \frac{e^2}{m^2 c^3} \left( \frac{dP_\mu}{d\tau} \cdot \frac{dP^\mu}{d\tau} \right) \quad (3.1)$$

$$= \frac{1}{6\pi\epsilon_0} \frac{e^2 c}{\rho^2} \beta^4 \gamma^4, \quad (3.2)$$

where  $d\tau = dt/\gamma$  is the proper time element,  $P^\mu$  is the electron Lorentz-vector, and  $\rho$  is the radius of curvature. The energy loss per turn is therefore

$$\delta E = \frac{2\pi\rho}{c\beta} P = \frac{1}{3\epsilon_0} \frac{e^2}{\rho} \beta^3 \gamma^4 \quad (3.3)$$

$$\delta E(\text{MeV}) = 8.85 \cdot 10^{-2} \frac{(E(\text{GeV}))^4}{\rho(\text{m})}, \quad (3.4)$$

where the last expression holds for high-energy electrons, where  $\beta \rightarrow 1$ . Hence, the energy loss per turn evolves quickly with the energy of electrons. At LEPI,

Table 3.1: Main LEP elements. Dipoles' value is approximate as it does not include those in injection regions, or the small number of particular elements.

Phase	Year	Beam Energy	Dipoles	RF cavities		
				Cu	NbCu	Nb
I	1989	45.6 GeV	~520	128	0	0
II	1996	86 GeV		120	160	32
III	1997	92 GeV		56	192	32
IIIb	1998	94.5 GeV		52	208	32
IV	1999	96 GeV		52	240	32

for 46 GeV electrons, 93.6 MeV were lost in a turn. This went up to 2.4 GeV at the end of the LEPII phase in 2000, and would have reached 5 GeV for a center-of-mass energy of 250 GeV. Pushing forward the beam energy was not possible at a reasonable cost. This is one of the main motivations for dismantling LEP in 2000, and installing the LHC, described in Chapter 5.

With the energy, the (integrated) luminosity was the second key characteristic of the machine. Luminosity depends on the bunch frequency, density and transverse size at the crossing point, as

$$\mathcal{L} = \frac{N^- N^+ f}{4\pi\sigma_x\sigma_y}, \quad (3.5)$$

with  $N^-$  ( $N^+$ ) being the number of electrons (positrons) in each bunch,  $f$  the bunch frequency, and  $\sigma_x, \sigma_y$  the transverse dimensions of the beam. With about 4 mA of current per beam, and initially four bunches in each beam (which, given the LEP's circumference, implies a frequency of 11.246 kHz), LEP bunches contained some  $54 \cdot 10^{10}$  particles. Since the transverse beam spot was  $145\mu\text{m} \times 5\mu\text{m}$ , LEP luminosity was approximately  $2 \cdot 10^{31} \text{ cm}^{-2}\text{s}^{-1}$ .

In 2000, the machine was pushed to its highest frequency at the price of beam life-time. In order to recover part of the penalty, beams were first accelerated to 102 GeV, and the RF was then shifted during so-called "mini-ramps" to gain a few GeV. This resulted in many different beam energies in 2000, as well as periods of uncertain energy, corresponding to the mini-ramps. Beam Energy distribution for 2000 is shown in Figure 3.2.

## 3.2 The ALEPH detector

As it is the case for most of the High-Energy physics detectors designed nowadays, ALEPH is a multi-purposes detector built to record all the events with a  $4\pi$  angular acceptance, with an optimum trade-off between performances and feasibility.

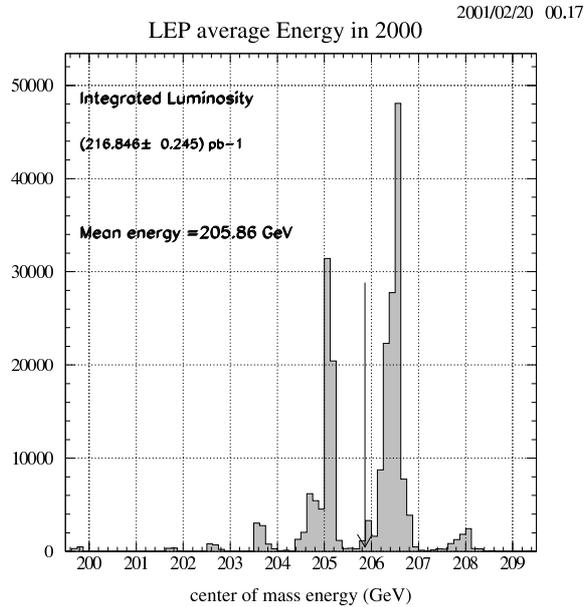


Figure 3.2: Distribution in energy of the luminosity delivered by the LEP accelerator during experiments carried out in 2000. The arrow marks the mean energy (205.86 GeV).

ALEPH stands for “Apparatus for LEP Physics”. It is a cylindrical device made of a main central section (the “barrel”) closed by two “endcaps” (Figure 3.3). It is approximately 11 meters long and 11 meters in diameter. Following a typical onion-like layered structure, several subdetectors surround the beam pipe. The ALEPH detector is described in detail in Ref. [18]. Only a brief description of some of the key features will be given here. A more exhaustive account of its performance and a description of the standard analysis algorithms can be found in Ref. [19].

Subdetectors will be categorized into two types of devices. The tracking devices dedicated to the detection of charged particles are placed in the central region (with the major exception of the muon chambers, which are in the very peripheral regions). In association with the solenoidal magnetic field, this allows a momentum determination. The second type includes the calorimeters. They will contribute to particle identification and resolve the energy.

Nearest to the beam pipe is a silicon strip vertex detector (VDET) installed in 1990 and upgraded during 1996 for the LEP2 phase. It plays a key role in heavy flavor physics as it measures the primary vertex with an accuracy of  $15\mu\text{m}$ . Surrounding this is a multilayer axial-wire drift chamber called the inner tracking chamber (ITC). It gives  $(r, \phi)$  measurements used for tracking and enters the Level-1 trigger system, as the detector response can be given

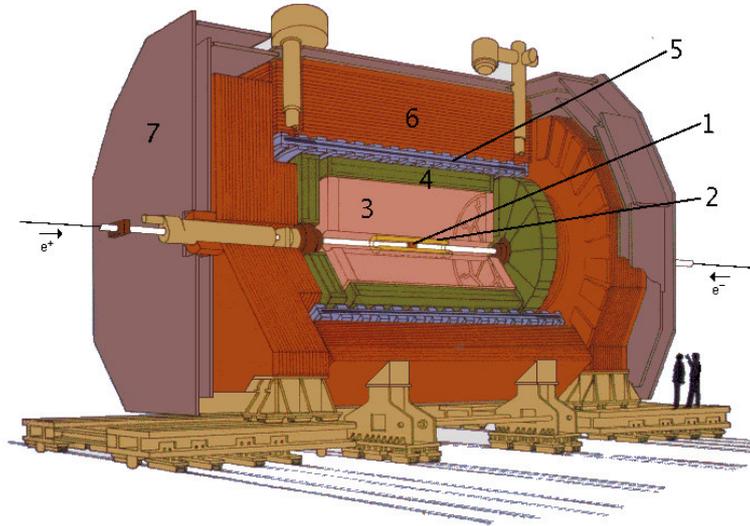


Figure 3.3: The ALEPH detector with all its subdetectors. 1. VDET; 2. ITC; 3. TPC; 4. ECAL; 5. Magnet; 6. HCAL; 7. Muon system. See the text for more details and definitions of the acronyms.

in less than  $2 \mu\text{s}$ . Next is the primary ALEPH tracking device, a large time projection chamber (TPC). The TPC is also used to identify particles by  $dE/dx$  measurement. This is surrounded by a highly-segmented 22-radiation-length-thick electromagnetic calorimeter (ECAL). It measures the energy of photons and charged particles and gives a significant contribution to particle identification by determination of the (transverse and longitudinal) shower shape. The superconducting magnet provides a uniform 1.5 T magnetic field in the axial direction, which is returned by a large iron structure instrumented to form the hadronic calorimeter (HCAL). Finally, most of the solid angle is covered by muon chambers, which measure three-dimensional coordinates for penetrating charged particles [20].

Small-angle luminosity monitors complement the device by improving the angular coverage in the very-forward and -backward regions down to 34 mrad from the beam axis. These will not be covered here.

### 3.2.1 Vertex detector

The vertex detector was installed in 1990, and upgraded for the LEP2 phase in 1996. It consists of two layers of double-sided silicon micro-strips sensors, with average radii of 63 mm and 110 mm from the beam-pipe axis [21]. The inner layer presents nine faces while the outer layer contains 15 faces. Each face is 400 mm long, 50 mm wide and  $300 \mu\text{m}$  thick. A technical drawing of the

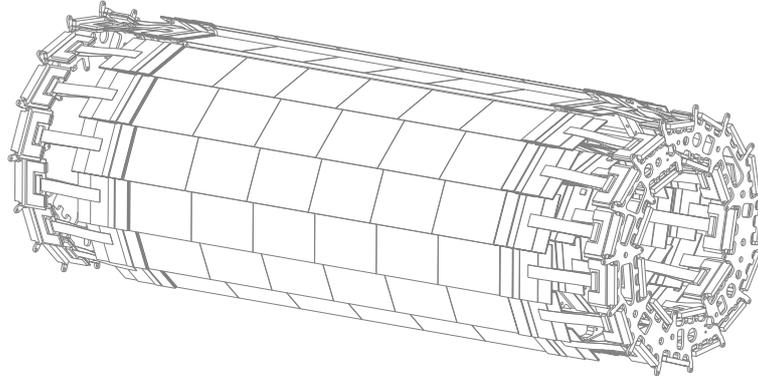


Figure 3.4: The Vertex detector, showing faces mounted on support structures. From [21]

VDET is shown in Figure 3.4.

The VDET is used to accurately identify and position primary or secondary vertices. The resolution is as good as  $10\mu\text{m}$  in the  $\phi$  direction, and varies from  $15\mu\text{m}$  to  $50\mu\text{m}$  in the  $z$  direction. This is very important to tag decays of hadrons containing  $b$  or  $c$  quarks. One of the motivations of this upgrade is the search for a Higgs boson, whose main decay consists in a pair of  $b$  jets, in the mass range kinematically accessible at LEP2.

### 3.2.2 Inner tracking chamber

The inner tracking chamber is a conventional two-meters long multi-wire drift chamber, with 960 sense wires parallel to the beam pipe. There are eight layers of sense wires, organized in 96 cells for the inner four layers and 144 cells for the outer four layers. The chamber extends to a radius of 28.8 cm, and is filled with a (80% Ar, 20%  $\text{CO}_2$ ) gas mixture.

Three-dimensional coordinates are extracted.  $(r, \phi)$  coordinate pairs are obtained from the drift time, while the  $z$  coordinate comes from the difference in arrival time between both ends of the wires. Resolutions are roughly  $150\mu\text{m}$  and 7 cm respectively for these two measurements.

The ITC plays an important role in the ALEPH trigger, as it provides the only tracking information at the first trigger level. There are  $(r, \phi)$  (decision in 500 ns) and  $(r, \phi, z)$  (decision in  $2\mu\text{s}$ ) fast trigger processors with special bits to indicate the number of tracks, hits per layer and back-to-back tracks.

### 3.2.3 Time projection chamber

The time projection chamber is the main tracking device of the ALEPH detector. It is a 4.6 meter-long gas (91% Ar, 9%  $\text{CH}_4$ ) volume, exposed to a lon-

gitudinal electric field of 11 kV/m. At each end of the cylinder, proportional chambers collect and measure charges deposited into the volume.

Charged particles crossing the chamber induce electron-ion pairs into the gas. Electrons from this ionization process drift along the electric field to the end-plates where their charge is recorded. Each end-plate has 18 sectors giving three-dimensional coordinate measurements ( $r\phi$  from pad position,  $z$  from drift time) as well as a  $dE/dx$  determination.

The TPC outputs up to 21 position measurements of the helix corresponding to the trajectory of charged particles in the strong magnetic field. The  $dE/dx$  information is extracted from the pulse height on the sense wire, sampled 340 times on a track. This measurement is important for particle identification of low-energy electrons. Using the dependence of the mean  $dE/dx$  on velocity, and combining this information with the momentum obtained from the track fit, the particle mass can be derived.

### 3.2.4 Electromagnetic calorimeter

The aim of the electromagnetic calorimeter (ECAL) is to determine the energy of charged particles and photons. The ALEPH ECAL is a lead/wire-chamber sampling device whose thickness is at least 22 radiation lengths. There are 45 layers, each composed of a wire chamber and of a lead plate. The wire chamber is Al extruded, and covered by graphite-coated mylar (Figure 3.5(a)). The gas mixture is made of 80% Xe and 20% CO<sub>2</sub>. The total deposited energy is measured using small (approximately  $3 \times 3 \text{ cm}^2$ ) cathode pads as well as anode wires, internally connected to form towers and storeys. Data reading is such that the transverse and longitudinal shapes of showers are recorded and can be used for particle identification.

The calorimeter is arranged in a barrel and two endcaps. Similar detector design and electronics are used for both of these elements. Barrel and endcaps are divided into 12 modules (Figure 3.5(b)), each representing 30° in azimuthal angle  $\phi$ . Endcaps are rotated by 15° with respect to the barrel.

The energy resolution is

$$\frac{\Delta E}{E} = 0.009 + \frac{0.18}{\sqrt{E(\text{GeV})}}, \quad (3.6)$$

for an electron detected at 90° in the barrel. It is slightly degraded in the overlap region and in the forward edge of the calorimeter. There are also 6% cracks in the barrel and 2% in the endcaps. The angular resolution is

$$\Delta\phi = \frac{\Delta\theta}{\sin\theta} = 0.32 + \frac{2.7}{\sqrt{E(\text{GeV})}} \text{mrad}. \quad (3.7)$$

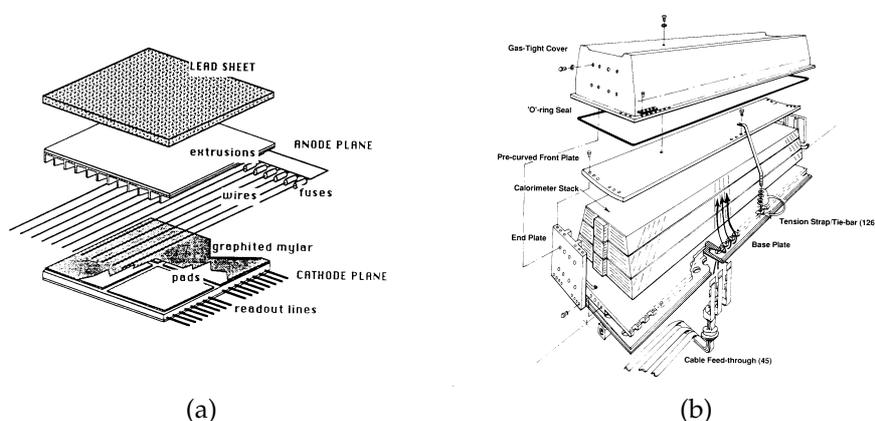


Figure 3.5: Details of the electromagnetic calorimeter; (a) typical ECAL stack layer; (b) exploded view of a barrel module. The calorimeter stack is made of 45 layers. From [20].

### 3.2.5 Magnet system

The magnet system consists of a superconducting solenoid coil, producing a uniform 1.5 Tesla magnetic field parallel to the beam pipe, and of a fully calorimetrized iron yoke.

The yoke is a very massive object made of iron (2580 tons in total for the barrel and the endcaps) which has to support all the other subdetectors. In addition, this is the converter for the hadronic calorimeter, detailed hereafter.

The superconducting coil is a helium-cooled aluminum conductor with a NbTi/Cu core. The field is produced by 1712 loops carrying 5000 A.

The superconducting coil was designed in order to minimize the material budget. This led to massive use of aluminum (or aluminum alloy), and the implementation of innovative indirect cooling techniques. The coil represents 1.6 radiation lengths.

### 3.2.6 Hadronic calorimeter and muon chambers

The return yoke just described is instrumented with streamer tubes in order to become the hadronic calorimeter (Figure 3.6).

As for ECAL, the HCAL is made of 12 modules. The total thickness of the calorimeter is 120 cm of iron, corresponding to 7.16 interaction lengths at 90 degrees, and increases up to 199 cm at the end of the barrel. The streamer tubes are PVC structures containing eight graphite-coated cells. Each cell contains a 100  $\mu\text{m}$  wire set at high voltage. Signal is read from copper pads on one side

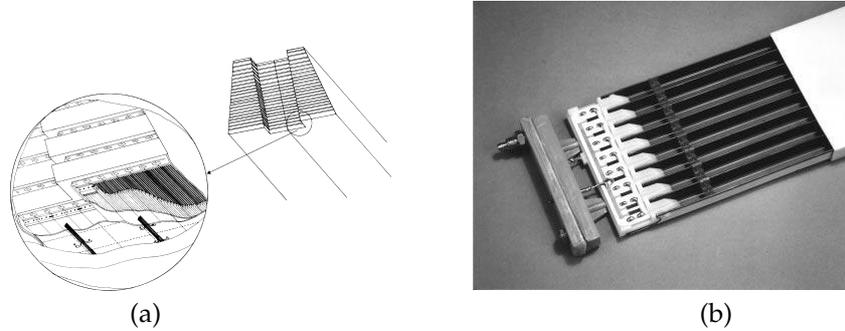


Figure 3.6: The hadronic calorimeter. (a) schematic view showing the positioning of the streamer tubes between the iron plates of a barrel module. (b) photograph of one end of a barrel tube. From [20].

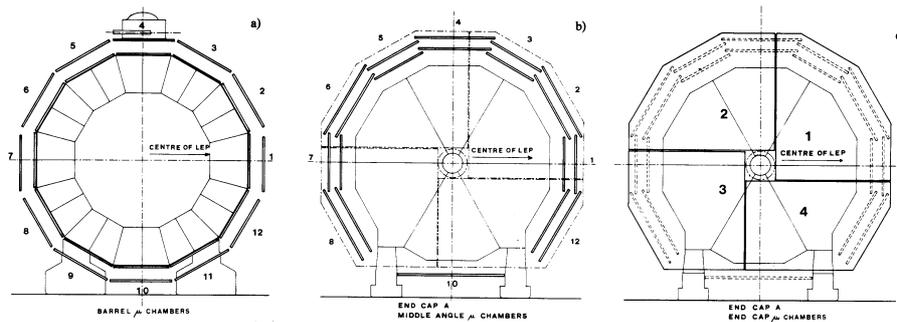


Figure 3.7: Cross-sectional view of the muon chambers. (a) barrel, (b) middle-angle, (c) endcap. From [20].

of the tube. The resolution of HCAL for pions is

$$\frac{\Delta E}{E} = \frac{0.85}{\sqrt{E(\text{GeV})}}. \quad (3.8)$$

On the outer side of the iron, two layers of streamer tubes are provided in order to detect muons, the only known charged particles that can go through the whole hadronic calorimeter (Figure 3.7).

Information from these two last layers complements the tracking in the hadronic calorimeter to identify certainly muons in complex events. The 50 cm separating the two layers allow angle and position measurement (with a resolution of 3.5 mm) for the crossing muon.

### 3.2.7 Trigger and event reconstruction

The purpose of the trigger system is to collect all interesting physical events, while keeping the “fake” rate due to noise or gas interactions “low”.

Such a background rate reduction is important for three main technical reasons. First, the open-time for the TPC has to be reduced; then the dead-time due to read-out of the electronics has to be lowered to an acceptable value; finally the disk space available is neither infinite nor free, and each rejected event could be expressed as saved money.

The trigger has been designed in three levels. The first level provides a decision in  $5 \mu\text{s}$ . This decision is based on information from the ITC and from both calorimeters. After a positive Level-1 decision, the second level verifies the decision, incorporating the TPC into the process. If the event is still accepted, the whole detector is read out and all the collected information is used in a Level-3 software-based trigger. This will reduce the rate to 1-2 Hz.

Events selected by the trigger system are fully reconstructed. Main aspects of ALEPH reconstruction are tracking in the central detectors, clustering in the calorimeter and association of these to identify charged and neutral particles.

Track reconstruction combines points recorded in VDET, ITC and TPC. Measurements compatible with a helix built from outside the TPC and propagated inwards are grouped. If at least four points are found, scattering is taken into account, and a compatible hit is looked for in the ITC, and finally in the VDET. Measured using  $e^+e^- \rightarrow Z^0 \rightarrow \mu^+\mu^-$  events, the resolution obtained this way is:

$$\frac{\Delta P_t}{P_t^2} = 0.610^{-3} (\text{GeV}/c)^{-1}, \quad (3.9)$$

which is half the resolution from the TPC alone.

## 3.3 Software aspects

### 3.3.1 Reconstructed physics objects

The role of the ALEPH reconstruction algorithms is to deliver a list of particles seen by the detector, while optimizing the energy and position resolutions. This is achieved by combining information from each subdetector. Electrons are identified using the shape of the energy deposit in ECAL and information from the TPC, combining  $dE/dx$  information,  $E/p$  measurement and the mean longitudinal position, while muon detection is based on signals in muon chambers, as well as on muon tracking in HCAL. A muon is identified as a track associated to some typical ionization pattern in HCAL, and to hits in muon chambers.

Apart from electrons and muons, advanced event reconstruction uses an energy flow algorithm. Information from calorimeters is first considered. After a cleaning procedure used to reduce the impact of noise from the front-

end electronic, from the acquisition system or from the reconstruction chain, calorimeter clusters are associated to tracks from charged particles to form a “calorimeter object”. Calorimeter objects associated to a valid track are counted as charged energy, assuming they are pions. Electrons and muons, reconstructed as previously explained, are removed from the collection by removing the track and the corresponding calorimetric deposit. For muons, a maximum of 1 GeV is removed from ECAL, as well as a maximum of 400 MeV per HCAL plane. Photons or neutral pions reconstructed by a dedicated algorithm are then removed from the calorimeter. Once all of the charged particles have been removed, all of the neutral hadron remain. All clusters exceeding 500 MeV and not associated to a track are assumed to be neutral hadrons. This hypothesis is used to reweight the ECAL and HCAL deposits so as to improve the overall resolution. The procedure just sketched results in a set of “particles”: electrons, muons, photons, charged or neutral hadrons. Global event quantities (such as total energy and missing momentum) are determined from these objects.

This method yields a resolution of

$$\sigma(E) = (0.59 \pm 0.03) \sqrt{E/\text{GeV}} + (0.6 \pm 0.3) \text{ GeV} \quad (3.10)$$

for hadronic decays of the  $Z$ , while the same quantity from raw calorimetric measurements is found to be  $\sigma(E) = 1.2 \sqrt{E/\text{GeV}}$ . Another more qualitative demonstration of ALEPH performances is given in Figure 3.8. It shows the total invariant mass versus the number of reconstructed particles for events measured at the  $Z$  peak, and shows how hadronic, leptonic and  $\gamma\gamma$  events are clearly identified.

### 3.3.2 ALEPH software suite

The full reconstruction of ALEPH events is performed by one single program called JULIA (Job to Understand Lep Interactions in Aleph). It is performed just after a run has been written to tapes, and can be divided into a set of distinct steps: track reconstruction, primary vertex determination, calorimeter clustering, energy flow analysis and particle identification. The resulting information is stored on tapes, together with the raw data, and delivered to physicists for analysis.

The ALEPH Physics Analysis package (ALPHA) is intended to simplify Fortran programs for physics analysis. ALPHA provides easy access to physical variables (momentum, energy, ...), to allow to write physics analysis programs without detailed knowledge of the ALEPH data structure. An extensive set of utility routines (e.g., kinematics, event shape, secondary vertex finding, b-tagging, etc.) is available as part of the ALPHA package.

In order to match the new trends in computing, the ALPHA++ project was launched. The ALEPH object oriented analysis package (ALPHA++) is intended to simplify programs for physics analysis by the use of C++. The main objectives are presented in the ALPHA++ manual [22]:

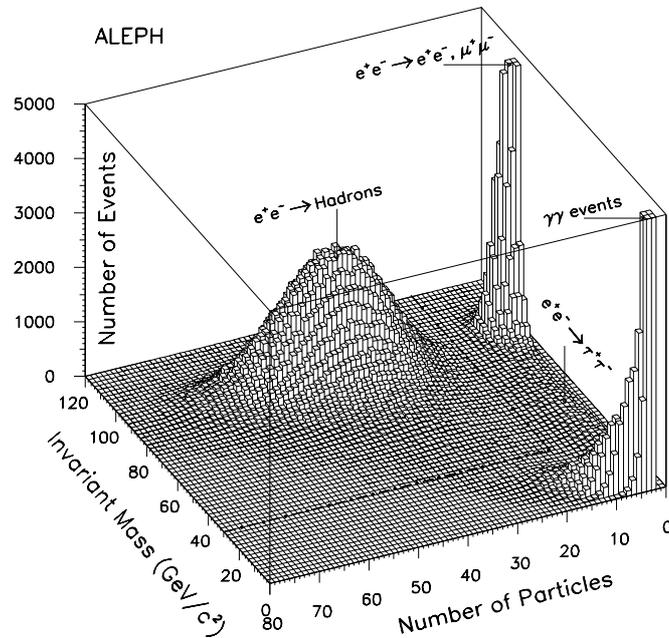


Figure 3.8: Invariant mass as a function of the number of reconstructed particles (energy flows) for all events with at least two charged particle tracks. The two low-multiplicity peaks have been truncated. From [19].

- Rewrite a mini version of the ALEPH analysis package ALPHA in an object-oriented computing language.
- Test the software engineered by the RD45 and ANAPHE projects, which evaluate data storage and analysis options for the LHC experiments.
- Compare performance with regard to efficient access to the data.
- Convert the ALEPH data (preferably LEP2 data) from BOS bank style into persistent objects and write them to an object oriented database.
- Provide some input and experience for the LEP archive project.
- Give an opportunity to learn object oriented analysis, programming and design.

During the last years of ALEPH analysis, we were responsible of the ALPHA++ development. To be able to perform every type of analysis, several improvements and additions where needed. Among other developments, the

---

offline lepton selection has been ported from ALPHA to ALPHA++, an interface has been provided to give access to vertices information, a wrapped FORTRAN b-tag routine has been included, etc. Other modifications to the code involve a general cleaning, the addition of comments, or modification of the driver part to enhance usability.

In order to guarantee availability of ALEPH data for the future, the ROOT package [23] has been interfaced with ALPHA++ at the driver level. That new system allowed us to convert the ALEPH database into a new set of ROOT files. The new database contains additional metadata, so that the new database can in principle be read without specialized software. We performed several cross-checks with ALPHA for previous analyses[24] so that ALPHA++ is now fully validated.



## Search for ZH associated production

### 4.1 Introduction

As mentioned in Section 2.3.2, an electroweak symmetry-breaking sector incorporating two doublets of complex scalar fields generates five scalar Higgs bosons, three of which are neutral. For certain choices of the model's parameters, one of these neutral scalars provides mass only to the fermions and the other couples exclusively to the bosons, i.e. is a "fermiophobic" Higgs boson.

The search for a fermiophobic Higgs boson was primarily carried out by the four LEP experiments in the  $H \rightarrow \gamma\gamma$  channel, in which the Higgs boson couples to photons via a W loop [25, 26, 27, 28]. A benchmark fermiophobic Higgs boson is defined by considering Standard-Model couplings to bosons, and null couplings to fermions. Such a fermiophobic Higgs boson is excluded up to  $109.7 \text{ GeV}/c^2$  by combining the four LEP experiments. For fermiophobic Higgs bosons heavier than  $90 \text{ GeV}/c^2$ , the predicted  $H \rightarrow \gamma\gamma$  branching ratio becomes small relative to the predicted  $H \rightarrow WW$  branching ratio (Figure 4.1), motivating a search in this new channel. Such an analysis was already conducted by the L3 collaboration [29] and is performed here with data collected by the ALEPH detector.

The main production processes at  $e^+e^-$  colliders for a fermiophobic Higgs boson are  $e^+e^- \rightarrow Z^* \rightarrow HZ$  (Higgsstrahlung), and WW or ZZ fusion. The cross-section for the boson fusion production processes are notably smaller than for the Higgsstrahlung process for LEP center-of-mass energies, as shown in Figure 4.2. In the Higgs boson mass range accessible at LEP (typically for  $M_H < 115 \text{ GeV}/c^2$ ), we are far below the W-pair production threshold. Hence, it is important to distinguish one W from the other, since one will be produced quasi on-shell, and the other will be off-shell. Decay products from the off-shell W boson will be significantly softer than those from the on-shell W boson.  $ZH \rightarrow ZW^+W^-$  decays will produce six-fermion final states, containing a total of 96 different channels depending on the decays of the Ws and

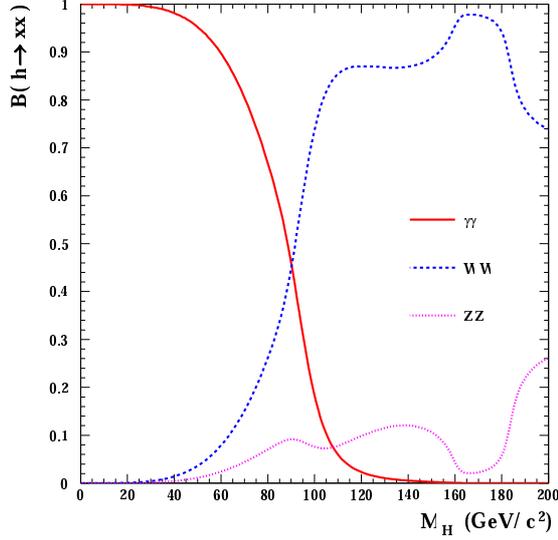


Figure 4.1: Branching fraction of benchmark fermiophobic Higgs boson into boson pairs as calculated by a modified version of HDECAY2 [30].

the Z. The analysis is performed on data taken in the years 1999 and 2000 at center-of-mass energies ranging from 191 to 209 GeV. The corresponding luminosities and center-of-mass energies are shown in Table 4.1.

## 4.2 Signal and background

Samples of signal events were generated using HZHA [31] with a Higgs boson mass between 90 and 117 GeV/c<sup>2</sup> and for seven different center-of-mass energies: 192, 196, 200, 202, 205, 207 and 208 GeV.

In the HZHA code, there is no spin correlation between the W bosons coming from the Higgs boson decay. The signal events are therefore re-weighted to take into account this spin correlation in the final state. To do so, the differential width has to be calculated for the full process  $H \rightarrow (WW) \rightarrow 4f$  in one go. The differential width for this process can be written as

$$d\Gamma_{H \rightarrow (WW) \rightarrow 4f} = \frac{1}{2M_H} |\mathcal{M}_{H \rightarrow (WW) \rightarrow 4f}|^2 d\text{Lips}_4(H, f1, f2, f3, f4), \quad (4.1)$$

where the four-particle phase-space element  $d\text{Lips}_4(H, f1, f2, f3, f4)$  contains 8 degrees of freedom and can be symbolically expressed as a function of Jacobians of transition from momenta to uniformly distributed scalars:

$$d\text{Lips}_4(H, f1, f2, f3, f4) = \frac{1}{4\pi^2} \mathcal{J}_{HWW} \mathcal{J}_{W^+} \mathcal{J}_{W^-} \mathcal{J}_{W^+2f} \mathcal{J}_{W^-2f} dr_1 \dots dr_8, \quad (4.2)$$

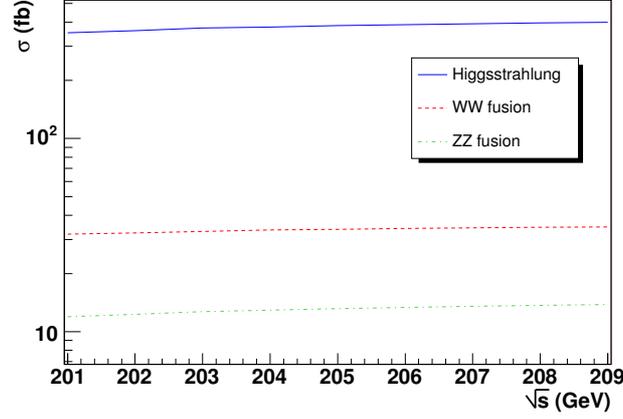


Figure 4.2: Cross-section for the Higgs boson production at LEP, obtained with HZHA. Boson fusion curves include interference terms. Therefore, Higgsstrahlung does not include  $H\epsilon\epsilon$  and  $H\nu\nu$ .

where  $\mathcal{J}_n$  are the Jacobians and  $r_1\dots r_8$  are uniformly distributed scalars. The  $\mathcal{J}_{HWW}$  Jacobian corresponds to the two decay angles of W bosons in the Higgs boson rest frame,  $\mathcal{J}_{W^+}$  and  $\mathcal{J}_{W^-}$  correspond to the  $W^\pm$  masses, and  $\mathcal{J}_{W^\pm 2f}$  to the decay angles of fermions in the  $W^\pm$  rest frames.

This has to be compared to the corresponding expression for the uncorrelated process computed in two steps, factorizing the Higgs boson decay from the decay of W bosons. It is possible to write down the partial widths taking

Table 4.1: Integrated luminosities, center-of-mass energy ranges and mean center-of-mass energy values for data collected by the ALEPH detector for years 1999 and 2000.

Year	Luminosity ( $\text{pb}^{-1}$ )	Energy range (GeV)	$\langle\sqrt{s}\rangle$ (GeV)
2000	7.3	207 – 209	208.0
	125.9	206 – 207	206.6
	79.7	204 – 206	205.2
1999	44.4	201 – 203	201.7
	87.2	199 – 201	199.5
	79.9	195 – 197	195.5
	28.9	191 – 193	191.6

into account the fact that both W bosons can be produced off-shell:

$$d\Gamma_{H \rightarrow WW} = \frac{1}{2M_H} |\mathcal{M}_{H \rightarrow WW}|^2 d\text{Lips}_2(H, W, W), \quad (4.3)$$

$$= \frac{1}{2M_H} |\mathcal{M}_{H \rightarrow WW}|^2 \mathcal{J}_{HWW} \frac{\mathcal{J}_{W^+}}{D(W^+)} \frac{\mathcal{J}_{W^-}}{D(W^-)} dr_1 \dots dr_4, \quad (4.4)$$

$$d\Gamma_{W^\pm \rightarrow 2f} = \frac{1}{2\sqrt{W^\pm 2}} |\mathcal{M}_{W^\pm \rightarrow 2f}|^2 d\text{Lips}_2(W, f_{1/3}, f_{2/4}), \quad (4.5)$$

$$= \frac{1}{2\sqrt{W^\pm 2}} |\mathcal{M}_{W^\pm \rightarrow 2f}|^2 \mathcal{J}_{W^\pm 2f} dr_{5,7} dr_{6,8}. \quad (4.6)$$

In the expressions above,  $W^\pm$  stands for the Lorentz momentum of corresponding weak bosons, and  $D(X)$  is defined from the Breit-Wigner propagators:

$$D(X)^{-1} = \frac{1}{(X^2 - M_X^2)^2 + \Gamma_X^2 M_X^2}. \quad (4.7)$$

It can be seen from Equations (4.1) to (4.6) that the correction weight defined as the ratio between correlated and uncorrelated differential widths is [32]

$$w = \frac{1}{\pi^2} \sqrt{W^{+2}} \sqrt{W^{-2}} \frac{|\mathcal{M}_{H \rightarrow (WW) \rightarrow 4f}|^2 D(W^+) D(W^-)}{|\mathcal{M}_{H \rightarrow WW}|^2 |\mathcal{M}_{W^+ \rightarrow 2f}|^2 |\mathcal{M}_{W^- \rightarrow 2f}|^2}. \quad (4.8)$$

It is interesting to note that the Breit-Wigner factors cancel similar factors in the full (correlated) matrix element. With such a normalization, the weight has a unit mean value (Figure 4.3) so that it can be directly used in the analysis to re-weight events.

The main effect of the re-weighting is expected to be seen in the distribution of the angle between positively charged fermions from the  $W^+$  and negatively charged fermions from the  $W^-$ , ( $\theta_{+-}$ ). This quantity is presented in Figure 4.4. It is observed that the angle ( $\theta_{+-}$ ) is only slightly modified by the re-weighting procedure.

Event samples of all Standard-Model background processes relevant for the Higgs boson search were also generated with the appropriate Monte-Carlo programs. At the center-of-mass energy relevant at LEP2, four-fermion processes become relevant as the kinematic limit is reached, as seen in Figure 4.5. The production of  $e^+e^-$  pairs (also known as the Bhabha process) was simulated with `BHWIDE` [33]. Muon and tau pair production was handled with the `KORALZ` [34] generator, the  $\gamma\gamma$  processes, where each  $e^\pm$  from the beam emits a photon that interacts later on, with `PHOT02` [35], and WW pair-production with `KORALW` [36]. The  $q\bar{q}$  events were simulated with `KK2F` [37], and the remaining four-fermion processes with `PYTHIA` [38]. This gives an accurate description of all events observed by the detector prior to the signal selection. Only very soft  $\gamma\gamma$  processes were ignored, as the processes are easy to distinguish from the signal but difficult to simulate. Details about the Monte-Carlo samples are presented in Table 4.2.

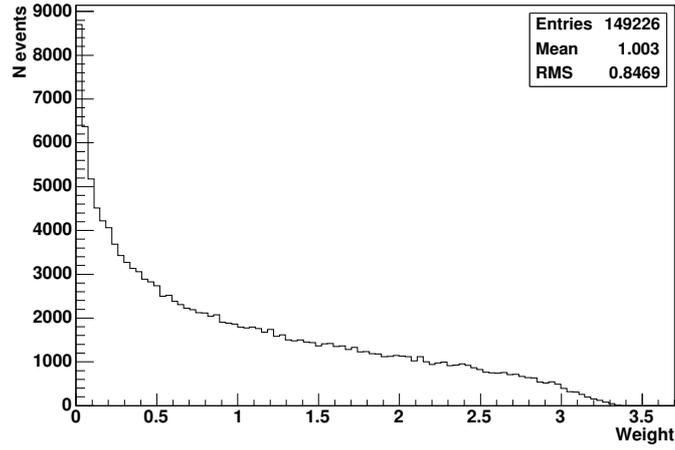


Figure 4.3: Distribution of weights to apply to the events generated by the HZHA code to take into account the spin correlation of the two  $W$ s coming from the Higgs boson decay.

Table 4.2: Monte-Carlo samples used for this analysis. Cross-sections are given at 207 GeV.

Process	Generator	Cross-section (pb)	Number of events
$q\bar{q}$	KK2F	81.007	4000000
$\gamma\gamma$ processes	PHOTO2	10711	30075000
$e^+ e^-$	BHWIDE	806.827	12600000
$\mu^+ \mu^-$	KORALZ	7.561	1525000
$\tau^+ \tau^-$	KORALZ	6.742	660000
$W e \nu$	PYTHIA	0.885	1150000
$WW$	KORALW	17.501	2625000
$Z e e$	PYTHIA	8.470	1300000
$Z \nu \nu$	ZNUNU	0.017	320000
$ZZ$	PYTHIA	2.34	1250000
$HZ, H \rightarrow WW$	HZHA	—	$7 \times 300000$

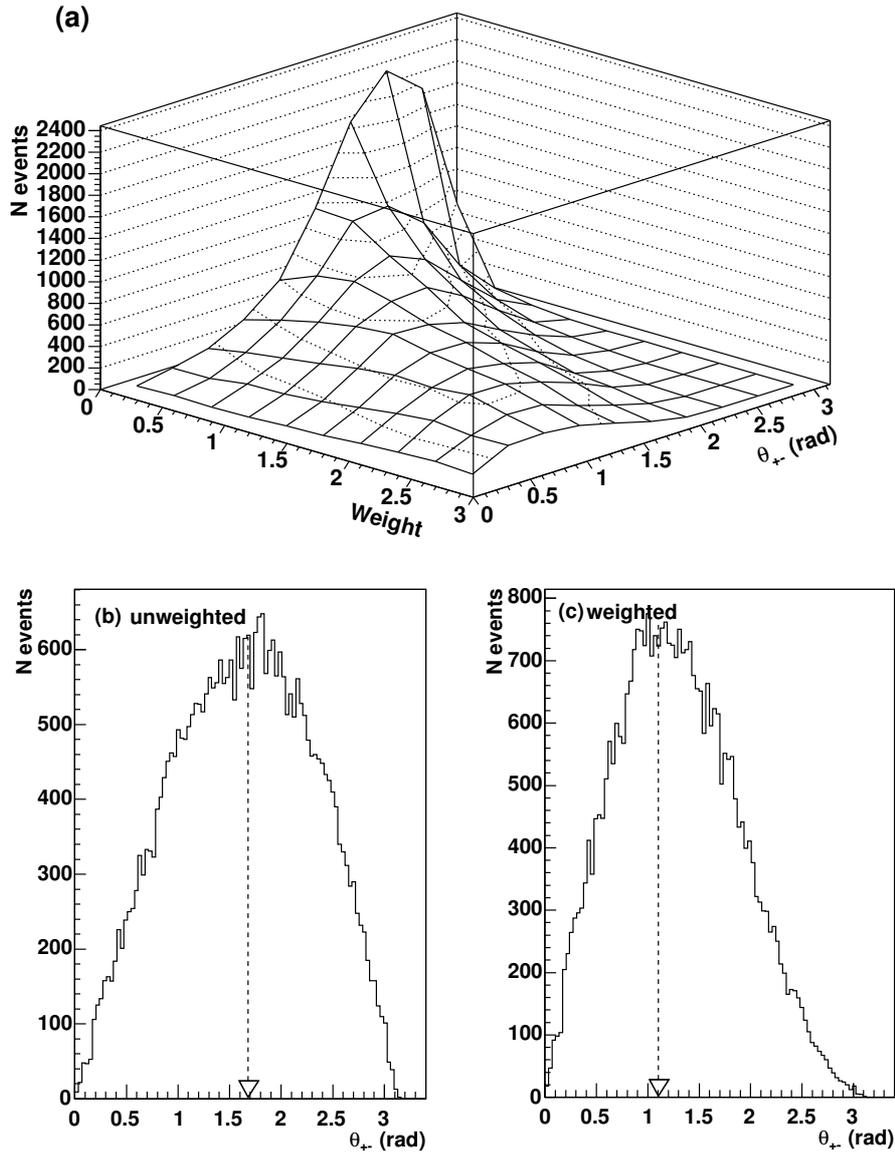


Figure 4.4: Effect of spin correlation correction on the angle between positively charged fermions from the  $W^+$  and negatively charged fermions from the  $W^-$  at generation level ( $\theta_{+-}$ ): (a) number of events as a function of the weight and as a function of  $\theta_{+-}$ ; the  $\theta_{+-}$  distribution before (b) and after (c) the spin correlation correction.

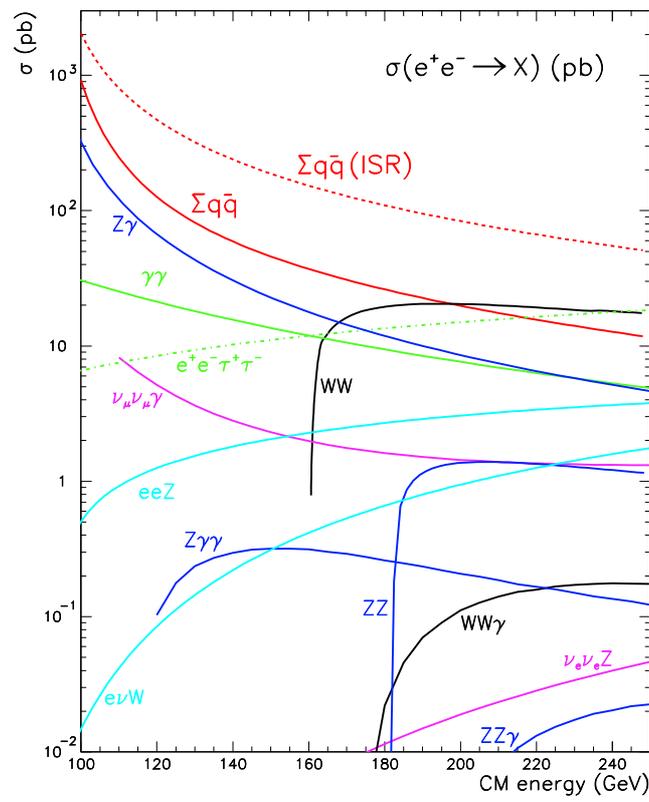


Figure 4.5: Cross-section for the main Standard-Model backgrounds at LEP, as a function of the center-of-mass energy.

A detailed GEANT[39] simulation of the detector response was applied to both background and signal events in order to reproduce the response of each subdetector. Particles were reconstructed from this information and the analysis was performed with ALPHA++[22].

### 4.3 Event classification

Four exclusive classes are defined according to the leptons in the final state, in such a way as to minimize overlap between the classes. The first class contains events with no lepton. Signatures with more than one hard lepton, including all leptonic Z decays, are purposes in the second class. The third class contains events with a single hard lepton, while the fourth contains those with only one soft lepton. For simplicity and conciseness, the term “lepton” is used here to denote the very final state and does not include taus, which decay into electrons, muons or hadrons prior to their detection. Taus are only specifically studied for a subsample of class two, where hadronic taus can be efficiently separated from the hadronic W boson decays. The topological classification is given in Table 4.3. The notation convention is as follows: the decay of the Z is given first, followed by the decays of the quasi-on-shell and the off-shell Ws, in that order. The table also gives the ZWW decay branching ratio corresponding to each class, with and without channels where at least one W decays into a tau and a neutrino.

To separate the classes, the energies of the two most energetic leptons in the final state are used. The first lepton is labeled as “hard” if its energy is higher than 25 GeV. Events with a hard lepton are set in class 2 if the second most energetic lepton has more than 20 GeV, and in class 3 otherwise. To separate events between classes 1 and 4, a linear discriminant ( $D_{14}$ ) is built with the energy of the most energetic lepton ( $E_{L1}$ ), the total missing three-momentum ( $\vec{P}$ ), and the isolation of the most energetic lepton ( $I_{L1}$ ):

$$D_{14} = 2.3E_{L1} + \vec{P} + 4.8\ln(|I_{L1}|), \quad (4.9)$$

where the lepton isolation is defined from the angle with the closest charged particle track ( $\theta_{L1T}$ ):

$$I_{L1} = 1 - \cos(\theta_{L1T}). \quad (4.10)$$

This discriminant will take a higher value for events in class 4, characterized by a soft isolated lepton accompanied by a neutrino, which generate missing transverse momentum. Events are classified as purely hadronic (class 1) if the discriminant is lower than 13, and are set in the fourth class otherwise. The corresponding fractions of events are presented in Table 4.4.

In Section 4.5, the analysis of events containing a single hard lepton or a single soft lepton will be discussed, and then combined to other classes in Section 4.7. Other parts of the analysis were performed by our collaborators and are presented briefly in Section 4.6. More details can be found in reference [40].

## 4.4 Confidence levels

In the search for new physics, two distinct hypotheses are considered. The background-only hypothesis, where there is nothing new to be observed, is opposed to the signal-plus-background hypothesis. The aim is to decide if there is an excess of events with respect to the background-only hypothesis, and if not, to determine whether or not the signal-plus-background hypothesis can be excluded.

The approach used throughout this work is the modified frequentist confidence level method presented in [41]. This method is suited to low-statistics backgrounds, as it relies on the Poisson distribution. A binned discriminant variable is used as input, which gives more sensitivity with respect to simple event-counting methods. Finally, it is easy to incorporate systematic effects in the evaluation of limits.

Let's consider  $n$  independent channels. For each channel  $i$ , one defines the number of signal events ( $s_i$ ), the number of background events ( $b_i$ ) and the number of events observed in data ( $d_i$ ). The expected signal and background levels are obtained from Monte-Carlo simulations. From these quantities,

Table 4.3: Final state topologies and branching ratios corresponding to the four classes. The decay of the  $Z$  is given first, followed by the decays of the quasi-on-shell and the off-shell  $W$ s, in that order. Branching ratios are given with and without considering the taus.  $l$  stands for electron or muon.

Class	Channels	BR without (with) taus
1: Fully hadronic	$q\bar{q}q\bar{q}q\bar{q}$ (1a)	0.328
	$\nu\bar{\nu}q\bar{q}q\bar{q}$ (1b)	0.094
		0.422 (0.422)
2: >1 Hard lepton	$l^+l^-q\bar{q}q\bar{q}$ (2a)	0.032
	$l^+l^-q\bar{q}lv$ (2b)	0.010
	$l^+l^-lvq\bar{q}, l^+l^-\tau\nu q\bar{q}$ (2c,T)	0.010
	$l^+l^-lvlv$ (2d)	0.003
		0.054 (0.100)
3: 1 Hard lepton	$q\bar{q}lvq\bar{q}$ (3a)	0.101
	$q\bar{q}lvlv$ (3b)	0.031
	$\nu\bar{\nu}lvlv$ (3c)	0.029
	$\nu\bar{\nu}lvq\bar{q}$ (3d)	0.008
		0.171 (0.284)
4: 1 Soft lepton	$q\bar{q}q\bar{q}lv$ (4a)	0.101
	$\nu\bar{\nu}q\bar{q}lv$ (4b)	0.029
		0.130 (0.195)
	Total:	0.777 (1.000)

Table 4.4: Distribution of candidates in each class for each signal topology (“signals” correspond to the ideal separation based on the true leptons at the generator level). Numbers are computed for a Higgs mass of 110 GeV/c<sup>2</sup>.

	Class 1	Class 2	Class 3	Class 4
Signal 1	<b>37.4%</b>	0%	0.3%	3.4%
Signal 2	0.9%	<b>4.7%</b>	1.5%	1.2%
Signal 3	5.3%	1.2%	<b>17.5%</b>	5.9%
Signal 4	7.1%	0.3%	0.6%	<b>12.2%</b>

the likelihood ratio between the signal-plus-background and the background-only hypothesis can be defined as:

$$Q = \prod_{i=1}^n Q_i, \quad (4.11)$$

where the likelihood ratio for each channel is obtained from the Poisson statistic:

$$Q_i = \frac{e^{-(s_i+b_i)}(s_i+b_i)^{d_i}}{d_i!} \frac{d_i!}{e^{-b_i}b_i^{d_i}} = e^{-s_i} \left(1 + \frac{s_i}{b_i}\right)^{d_i}. \quad (4.12)$$

It is then straightforward to obtain a statistic that can be used to distinguish the two hypotheses:

$$-2 \ln Q = 2s - 2 \sum_{i=1}^n d_i \ln \left(1 + \frac{s_i}{b_i}\right), \quad (4.13)$$

where  $s$  is the total signal ( $s = \sum_{i=1}^n s_i$ ). This is nothing else than a weighted sum of the events observed in the data. This quantity has a distribution that can be interpreted as a probability density, if  $d_i$  are random variables following Poisson distributions. There is one probability density function (PDF) in the background-only hypothesis (where the most probable value of  $d_i$  is  $b_i$ ), and one other PDF in the signal-plus-background hypothesis (where the most probable value of  $d_i$  is  $s_i + b_i$ ), as shown in Figure 4.6. The confidence level to exclude the signal-plus-background hypothesis is defined from the probability to observe a given value of the likelihood ratio in the presence of the signal:

$$\text{CL}_{s+b} = P_{s+b}(Q < Q_{obs}), \quad (4.14)$$

where  $Q_{obs}$  is obtained directly from Equation (4.12) using the actual values for  $d_i$ . In the same way, the confidence level for the background-only hypothesis is computed as:

$$\text{CL}_b = P_b(Q < Q_{obs}). \quad (4.15)$$

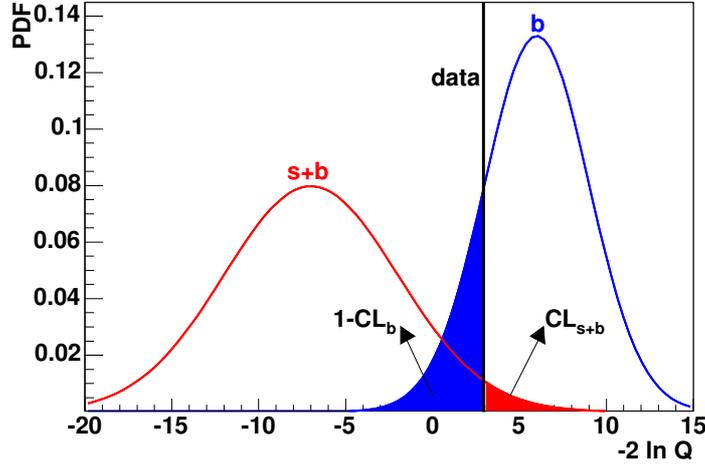


Figure 4.6: Probability density functions for the likelihood ratio in background-only and signal-plus-background hypothesis, and related quantities.

$CL_b$  is used to quantify a possible discovery, as it measures the probability for the background processes to give more candidates than observed. If observations match perfectly the predictions ( $d_i = b_i$ ),  $CL_b$  is equal to 0.5, and rises if there is an excess of events. Deviations from 0.5 are usually assumed to be Gaussian and expressed in terms of standard deviations.

$CL_{s+b}$  could be used to test the signal-plus-background hypothesis, but is very sensitive to downward fluctuations of the background and could give too optimistic limits. It is therefore normalized to  $CL_b$ :

$$CL_s = CL_{s+b}/CL_b. \quad (4.16)$$

One will say that the 95% confidence level (C.L.) is reached if  $1 - CL_s > 0.95$ . Expected values  $\langle CL_s \rangle$  and  $\langle CL_b \rangle$  are directly obtained from the PDF under the hypothesis that  $d_i = b_i$ .  $\langle CL_b \rangle$  is by definition equal to 0.5, except for very low numbers of background events ( $\sum_{i=1}^n b_i < 1$ ).

Practically, the expected confidence level  $\langle CL_s \rangle$ , the observed confidence level  $CL_s$  and the background confidence level  $CL_b$  are computed using a toy Monte-Carlo method with the algorithm implemented by us within the ROOT package [23], using as inputs the distribution of a variable chosen for its discriminant power between signal and background. In this procedure, pseudo-data distributions are generated in either the background-only or the signal-plus-background hypothesis, and distributions of  $Q$  are built (as in Figure 4.6). Each bin of the discriminant variable distribution is considered as an independent channel in the sense of Equation (4.11).

One advantage of this method is that systematic uncertainties are naturally included. Systematic uncertainties are taken into account by varying  $s_i$  and  $b_i$

around their expected values for each Monte-Carlo experiment, which affects the width of the PDFs. Uncertainties from the Monte-Carlo statistics are considered by varying independently each bin, while other systematic sources are handled as fully correlated variations of the histogram content.

After a first signal preselection step, designed to reduce the background that can be easily suppressed without tuning, all the selection cuts are optimized in order to minimize the expected confidence level  $\langle \text{CL}_s \rangle$ . For the signal, a Higgs boson mass hypothesis of  $110 \text{ GeV}/c^2$  is chosen for the optimization process.

Only simulated events accounting for the year 2000 luminosity are used during the optimization process. Data taken in 1999 and corresponding simulated events are therefore not included in tables and plots presented in Sections 4.5 and 4.6. They are however included for the final results of Section 4.7.

## 4.5 Study of events with a single hard or soft lepton

We will here consider final state topologies characterized by the presence of a single hard or soft lepton. Topologies with only one hard lepton in the final state are studied in class 3. This can only come from the leptonic decay of the on-shell W boson. Consequently, a class 3 event is characterized by missing energy due to the presence of at least one neutrino in the final state. The Z boson can decay either into jets or neutrinos and the off-shell W into jets or into a lepton and a neutrino. Events in class 3 are therefore distributed in four subclasses, as already presented in Table 4.3. The separation between the hadronic and invisible decays of the Z boson is easily achieved by cutting on the total invariant mass of the event. A cut on the energy of the second most energetic lepton is then used to distinguish between hadronic and leptonic off-shell W decay.

Final state topologies with only one soft lepton are studied in class 4. Class 4 is significantly different from class 3 since all the jets will be relatively hard, while two of them (coming from the off-shell W boson) are rather soft in class 3. It may be noticed that analysis of these events relies on the correct reconstruction, identification and measurement of soft non-isolated leptons. It therefore depends on the high-quality ALEPH energy flow algorithm presented in Section 3.3.1. The high granularity of the calorimeters, both transversally and longitudinally, is also crucial in this respect. The Z boson can decay either into jets or neutrinos, which justifies subdivision of class 4 into 4a (for  $ZH \rightarrow q\bar{q}q\bar{q}lv$ ) and 4b (for  $ZH \rightarrow \nu\bar{\nu}q\bar{q}lv$ ).

Due to the presence of one or several neutrinos in the final state, the exact reconstruction of the Higgs boson mass  $M_H$  is not possible. However, for the  $HZ \rightarrow q\bar{q}lvlv$  and  $HZ \rightarrow \nu\bar{\nu}lvlv$  channels, a good estimate of the Higgs boson mass is obtained with twice the sum of the lepton energies:

$$M_H = 2(E_{L1} + E_{L2}). \quad (4.17)$$

This results from the fact that all three ( $W$  or  $Z$ ) bosons are more or less decaying at rest. For  $HZ \rightarrow \nu\bar{\nu}l\nu q\bar{q}$ , the Higgs boson mass can be approximated as

$$M_H = \sqrt{(E_{tot} + E_{L1})^2 - (\vec{P}_{tot} - \vec{P}_{L1})^2}, \quad (4.18)$$

with  $E_{tot}$  ( $E_{L1}$ ) and  $\vec{P}_{tot}$  ( $\vec{P}_{L1}$ ) being respectively the total (leading lepton) energy and momentum. For the  $HZ \rightarrow q\bar{q}l\nu q\bar{q}$  channel, the mass of the  $Z$  boson has to be subtracted. The Higgs boson mass reconstructed as explained above is used as a discriminant variable for all subclasses, except for  $HZ \rightarrow q\bar{q}l\nu q\bar{q}$ , where the visible mass obtained ignoring the lepton ( $M_{tot} (hadr)$ ) is more powerful. For class 4, the discriminating variable used for the estimation of the confidence level is the off-shell  $W$  mass, constructed from the missing momentum, the missing energy, the lepton momentum and the lepton energy as follows:

$$M_W = \sqrt{(|\vec{P}| + E_{L1})^2 - (\vec{P} + \vec{P}_{L1})^2}, \quad (4.19)$$

where  $E_{L1}$  and  $\vec{P}_{L1}$  are the energy and momentum of the lepton, and  $\vec{P}$  is the missing momentum.

Preselection cuts are applied in order to strongly reduce the  $\gamma\gamma$ ,  $q\bar{q}$  and  $l^{+1-}$  backgrounds.  $\gamma\gamma$  and  $l^{+1-}$  backgrounds are characterized by a low multiplicity, so that they are very well rejected imposing  $N_{ch} > 3$ . This cut is used for all classes, except for  $HZ \rightarrow \nu\bar{\nu}l\nu l\nu$  (class 3c) where only two tracks are expected. All three backgrounds are well balanced, so that acollinearity<sup>1</sup> is close to  $2\pi$ . For the same reason,  $P_t/\sqrt{s}$  is expected to be small. Hence, one asks  $P_t/\sqrt{s} > 0.05$  in all classes. In addition, a cut on the acollinearity ( $\theta_{aco} < 170^\circ$ ) is used for signal preselection in class 3c, to compensate the absence of cut on  $N_{ch}$ . The energy fraction within twelve degrees of the beam line ( $E_{12}$ ) must be lower than 40% of the center-of-mass energy to reject most of the radiative  $\gamma\gamma$  events, as well as  $e^+e^-$  events at low angle. This is complemented by a cut on the azimuthal angle of the missing momentum ( $\cos(\theta_{miss})$ ) that removes also events with unmeasured particles lost in the beam pipe. One asks  $|\cos(\theta_{miss})| < 0.95$ . The event total invariant mass ( $M_{tot}$ ) and total transverse momentum ( $P_t$ ) must satisfy  $M_{tot}/\sqrt{s} > 0.2 - 6P_t/\sqrt{s}$ , to protect against radiative  $Z$  returns. For classes 3b, 3c and 3d, there is an additional cut on the energy of the most energetic photon ( $E_{\gamma 1}$ ) at the preselection level in order to reject radiative events.

In the following results and figures, the number of signal events is obtained assuming a Standard-Model production cross-section and a unit branching ratio for the Higgs boson decay into a  $W$  pair.

<sup>1</sup>Acollinearity ( $\theta_{aco}$ ) is defined as the angle between the two jets when two jets are forced in the event using the Durham jet algorithm [42]. Similarly, the acoplanarity ( $\Phi_{aco}$ ) is defined as the angle between these two jets, this time in the transverse plane.

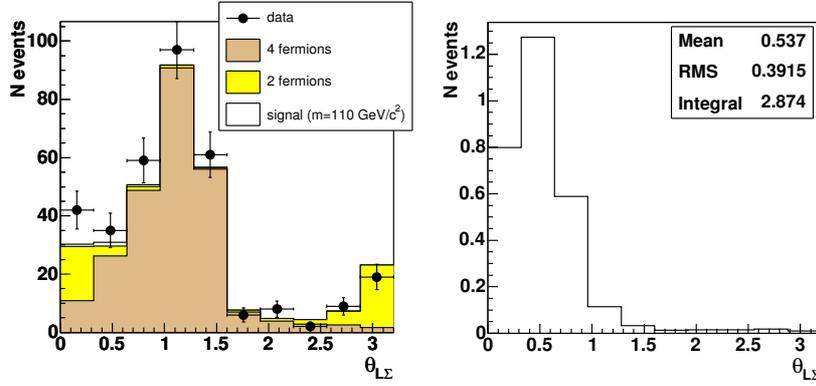


Figure 4.7: Distribution of the angle between the hard lepton and the total momentum for class 3a events, after preselection cuts and the cut on  $N_{ch}$ . The distributions is obtained from 2000 data.

#### 4.5.1 Class 3a - hadronic Z and hadronic off-shell W

The channel  $HZ \rightarrow q\bar{q}l\nu q\bar{q}$  is separated from the other events in class 3 by requiring a total invariant mass higher than 40% of the center-of-mass energy, a missing mass ( $M_{miss}$ ) less than  $95 \text{ GeV}/c^2$  and no second lepton with an energy exceeding  $10 \text{ GeV}$ .

A first set of cuts is used to eliminate the  $\gamma\gamma$ ,  $l^+l^-$  and  $q\bar{q}$  backgrounds. To reflect the presence of hadrons from the Z boson decay, there must be more than 20 charged particle tracks. The angle between the hard lepton and the total momentum ( $\theta_{L\Sigma}$ ) is then considered. This angle approximates the supplement of the angle between the lepton and the neutrino. Assuming that the on-shell W boson is produced at rest, it is expected to be small, as seen in Figure 4.7. The optimization procedure suggests that  $\theta_{L\Sigma}$  must be less than  $41^\circ$ . The mass of the hard W boson, that can be identified to the invariant mass built from the hard lepton momentum and the missing momentum ( $M_{L\cancel{p}}$ ), has then to be larger than  $55 \text{ GeV}/c^2$ . For events with no W boson decaying into a lepton-neutrino pair, this quantity is indeed expected to be small, as shown in Figure 4.8.

The following cuts are applied to suppress the remaining backgrounds, mainly WW events. The hardest lepton must have an energy  $E_{L1}$  less than  $55 \text{ GeV}$ . The hadronic acollinearity must be larger than  $137^\circ$ . Both cuts account for the absence of other neutrino that the one compensated by the lepton, and for the fact that W and Z bosons are produced more or less at rest.

One quantity that is particularly useful when analyzing hadronic events is the thrust [43]. Given the momentum of all particles, here identified to energy flows, one puts a straight line through their common origin, projects all the

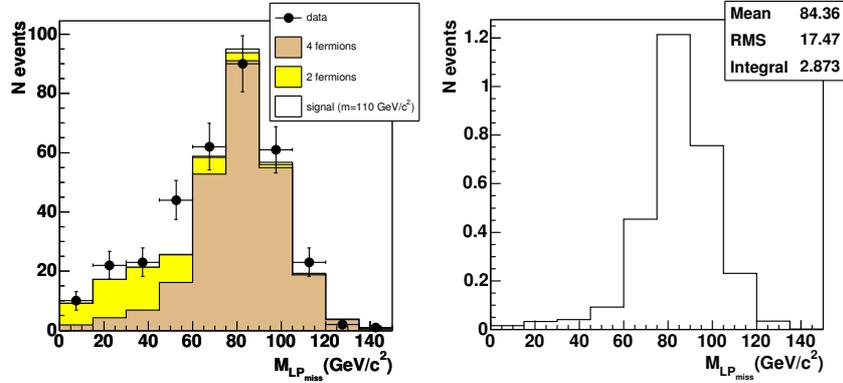


Figure 4.8: Distribution of the invariant mass of the hard lepton and the missing momentum for class 3a events, after preselection cuts and the cut on  $N_{ch}$ . The distributions is obtained from 2000 data.

momenta onto this line, and sums up the absolute values of these projections. The line which maximizes this sum is called thrust axis, and the projection, normalized to the scalar sum of momenta, is called thrust. This quantity gives a measure of the “two-jet-likeness” of an event. In this analysis, the thrust, when computed ignoring the hard lepton ( $T_{no\ lept}$ ), must be lower than 0.93, since it is expected to be reduced by the soft hadronic activity from the off-shell W.

The thrust is complemented by three other global topological quantities computed from the event shape [44]. For that purpose, the momentum tensor is defined as

$$T^{\alpha\beta} = \sum_i (\delta^{\alpha\beta} \vec{P}_i^2 - P_i^\alpha P_i^\beta), \quad (4.20)$$

where the sum is performed on all particles, and  $\alpha$  and  $\beta$  take values from 1 to 3. From the eigenvalues ( $\lambda_i$ ) ordered in decreasing order, the sphericity ( $S$ ), the aplanarity ( $A_{pl}$ ) and the planarity ( $P_{lan}$ ) are defined as:

$$S = \frac{3}{2}(1 - \lambda_1), \quad (4.21)$$

$$A_{pl} = \frac{3}{2}\lambda_3, \quad (4.22)$$

$$P_{lan} = \frac{\lambda_3}{\lambda_2}. \quad (4.23)$$

By construction, these quantities are defined in the range  $[0, 1]$ . The sphericity approaches 1 for events with large multiplicity and isotropic phase-space particle distributions. The planarity approaches 1 if all particles are bounded in a single plane. For this subclass, all these three quantities were considered, but

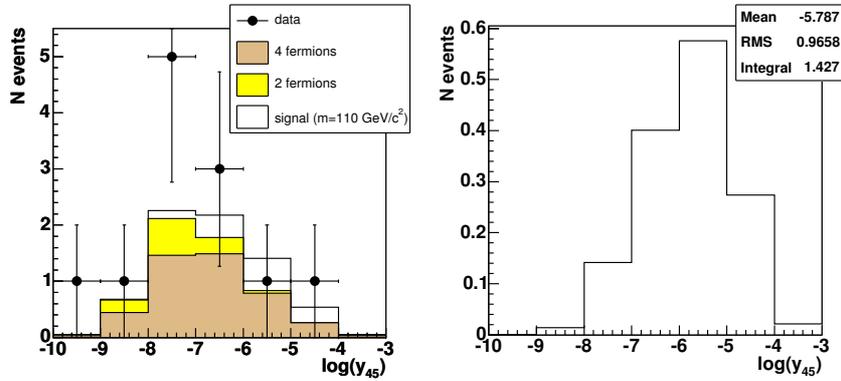


Figure 4.9: Distribution of the natural logarithm of the jet transition value  $y_{45}$  for events in class 3a, after all cuts but the cuts on this quantity and on the total transverse momentum. The distributions is obtained from 2000 data.

were not found to allow a good background rejection. Sphericity and aplanarity are used in other subclasses.

Jets resulting from the hadronization of final-state quarks and gluons are reconstructed using the Durham algorithm [42]. The Durham algorithm relies on the definition of a jet metric  $y$ , inspired from the invariant mass between two objects (we will use energy flows):

$$y = \frac{2(\min(E_1, E_2))^2 \left(1 - \frac{\vec{P}_1 \cdot \vec{P}_2}{|\vec{P}_1| |\vec{P}_2|}\right)}{E_{tot}^2}, \quad (4.24)$$

with  $E_{tot}$  being the total visible energy. The algorithm proceeds by merging the pair of objects for which this distance is minimum. The new object is obtained by summing the Lorentz momentums of the two merged objects. The procedure is repeated until the minimal value of  $y$  is larger than a given value, or until a given number of objects (then called jets) is obtained. The  $y$  value at which the topology changes from  $j$  to  $i$  jets is defined to be  $y_{ij}$ . In this analysis, the jet transition value  $y_{45}$  (Figure 4.9) must be larger than 0.0014 (i.e.  $\ln(y_{45}) > -7.2$ ), since five-jet topologies are favored. Cutting on the total transverse momentum finally reduces the WW background.

Figures 4.10a and 4.10b give the distributions of the discriminant variable ( $M_{tot}(hadr)$ ) after the preselection and after all cuts, respectively. The selection criteria together with the numbers of signal, background and data events for the  $HZ \rightarrow q\bar{q}l\nu q\bar{q}$  channel are summarized in Table 4.5.

After applying these cuts, 6.6 background events remain, for 1.27 signal events (for years 1999 and 2000 together). The expected value of the confidence level,  $\langle CL_s \rangle$ , is 0.47. Thirteen candidates are observed, which corresponds to a  $2\sigma$  excess ( $CL_b = 0.96$ ), and  $CL_s = 0.84$ . Note that the difference in

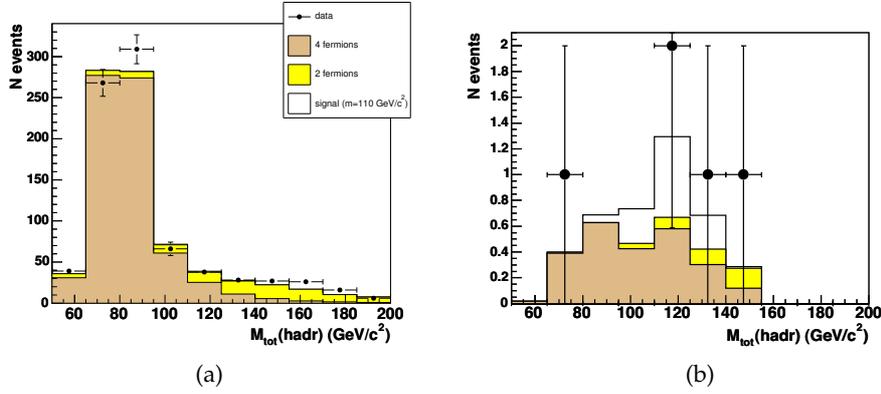


Figure 4.10: Distribution of the discriminant variable  $M_{tot}(hadr)$  for the decay  $HZ \rightarrow q\bar{q}'lvq\bar{q}'$  plotted for the class 3a events after the preselection (a) and after the final selection cuts (b). Both distributions are obtained from 2000 data.

number of events with respect to those in Figure 4.10 comes from the inclusion of data from 1999.

#### 4.5.2 Class 3b - hadronic Z and leptonic off-shell W

The  $HZ \rightarrow q\bar{q}'lv$  events are separated from the other events in class 3 requiring a total invariant mass higher than 40% of the center-of-mass energy, a missing mass less than  $95 \text{ GeV}/c^2$  as well as a second lepton with more than 10 GeV. The selection procedure for  $HZ \rightarrow q\bar{q}'lv$  is very similar to that performed for  $HZ \rightarrow q\bar{q}'lvq\bar{q}'$ , and is shown together with the number of signal, background and data events after each selection step in Table 4.5.

The  $\gamma\gamma$ ,  $l^+l^-$  and  $q\bar{q}$  background events are rejected by strengthening the cuts on the number of charged particle tracks and on the transverse momentum, and requesting a low acollinearity. To further suppress  $q\bar{q}$  events, cuts on the angle between the lepton and the total momentum as well as on the event sphericity are applied. Finally, most of the WW background events are removed by four last cuts. The lepton energy is smaller for signal events than for WW events, so that an upper energy cut is applied; the hadronic acollinearity ( $\theta_{aco}(no\ lept)$ ), the jet metric value computed without the lepton ( $y_{12}(no\ lept)$ ), and  $y_{34}$  are the last sensitive quantities used. Remaining background events are mainly semi-leptonic WW decays with a soft lepton produced in a jet. The distribution of the discriminant variable  $M_H$  (Equation (4.17)) after the preselection and after all cuts is given in Figures 4.11a and b respectively. The selection criteria together with the numbers of signal, background and data events for the  $HZ \rightarrow q\bar{q}'lv$  channel are summarized in Table 4.5.

After all cuts, there remains two background events for 1.4 signal events

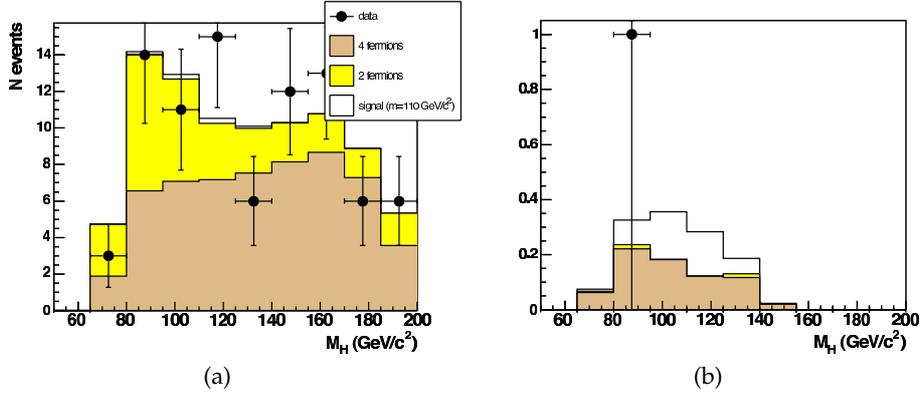


Figure 4.11: Distribution of the discriminant variable  $M_H$  (Equation (4.17)) for the decay  $HZ \rightarrow q\bar{q}'lvlv$  plotted for the class 3b events after the preselection (a) and after the final selection cuts (b). Both distributions are obtained from 2000 data.

(for years 1999 and 2000 together). Three events are observed in data. While the expected confidence level  $\langle CL_s \rangle$  is 0.62, the observed  $CL_s$  value is 0.66.  $CL_b$  is equal to 0.65, which is in accord with expectations, within  $1\sigma$ .

### 4.5.3 Class 3c - invisible Z and leptonic off-shell W

$HZ \rightarrow \nu\bar{\nu}lvlv$  is a simple signature since there are only two acoplanar leptons in the final state, but it has a very small branching ratio and it suffers from the very large fully leptonic  $WW$  background.

Only events with exactly two charged particle tracks are kept for this analysis. Signal events are selected by requiring a total invariant mass lower than 40% of the center-of-mass energy as well as a second lepton with an energy higher than 10 GeV. The signal is characterized by a significant acollinearity between the two leptons.  $\gamma\gamma$  events are therefore rejected by a cut on the acollinearity and on the transverse momentum. Then, the total mass and the acoplanarity are used to reject  $l^+l^-$  events as well as part of the  $WW$  background. A final cut on the total missing mass and on  $\theta_{L\Sigma}$  removes most of the remaining  $WW$  and  $ZZ$  background. For the signal, this angle is indeed expected to be small, because of the spin correlation between the leptons. The distribution of the discriminant variable  $M_H$  (Equation (4.17)) after the preselection and after all cuts is given in Figure 4.12. The selection criteria together with the numbers of signal, background and data events for the  $HZ \rightarrow \nu\bar{\nu}lvlv$  channel are summarized in Table 4.5.

The remaining background is still 30 times higher than the reference Higgs signal (i.e.  $M_H = 110 \text{ GeV}/c^2$  and  $BR=1$ ). After applying all cuts, 0.15 signal events are expected for 6.6 background events, leading to  $\langle CL_s \rangle = 0.95$  (for

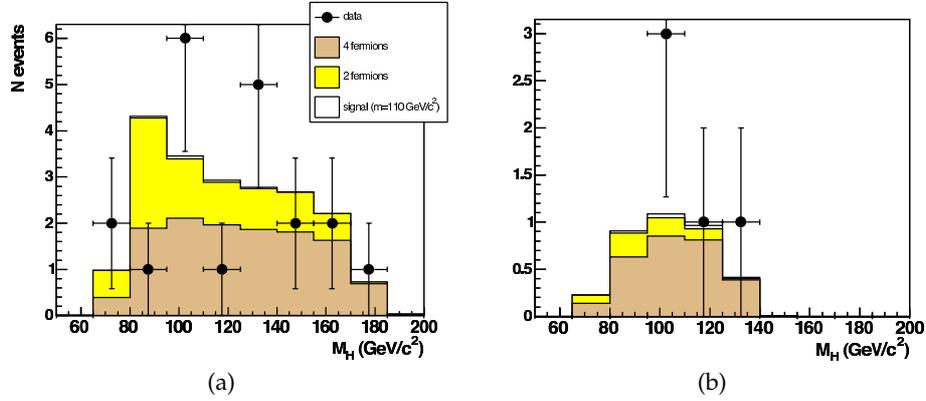


Figure 4.12: Distributions of the discriminant variable  $M_H$  (Equation (4.17)) for the decay  $HZ \rightarrow \bar{\nu}l\nu l\nu$  plotted for the class 3c events after the preselection (a) and after the final selection cuts (b). Both distributions are obtained from 2000 data.

years 1999 and 2000 together). Fourteen candidates are observed in the data, which corresponds to a  $2\sigma$  excess over the expected background ( $CL_b=0.97$ ). This leads to  $CL_s = 0.99$ .

#### 4.5.4 Class 3d - invisible Z and hadronic off-shell W

The last sub-channel,  $HZ \rightarrow \bar{\nu}l\nu q\bar{q}$ , being the complement of the three others is selected by requiring a total invariant mass lower than 40% of the center-of-mass energy as well as no second lepton with more than 10 GeV. Events are characterized by a single hard lepton (around 30 GeV) with some soft hadronic activity.

Both  $\gamma\gamma$  and  $l^+l^-$  events are rejected by requiring a significant aplanarity ( $A_{pl} > 0.00034$ ). The dominant WW background is then reduced in two stages. First, the angle between the hard lepton and the total momentum and the energy of both the first and second most energetic leptons are used. A final rejection is then achieved by a cut on the sphericity, on the hadronic invariant mass  $M_{tot}(hadr)$  (expected to be high for the WW background) and on  $y_{12}(no\ lept)$ , the jet metric value corresponding to the 1- to 2-jet transition computed without the lepton. Distribution of the discriminant variable  $M_H$  (Equation (4.18)) after the preselection and after all cuts is given in Figures 4.13 a and b respectively. The selection criteria together with the number of signal, background and data events for the  $HZ \rightarrow \bar{\nu}l\nu q\bar{q}$  channel are summarized in Table 4.5.

After the final selection, 0.99 background events and 0.89 signal events are expected (for years 1999 and 2000 together). The corresponding confidence level  $\langle CL_s \rangle$  is 0.44. No candidate is observed, and  $CL_s$  is observed to be equal to 0.43.  $CL_b$  is then 0.49, which is in perfect agreement with expectations.

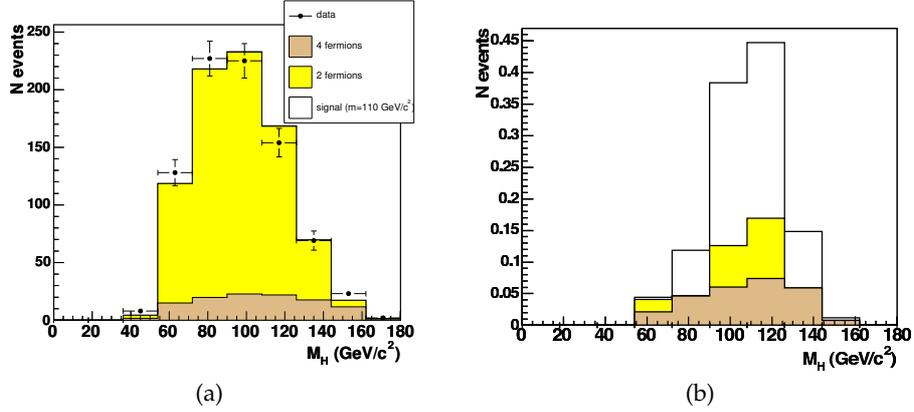


Figure 4.13: Distribution of the discriminant variable  $M_H$  (Equation (4.18)) for the decay  $HZ \rightarrow \nu \bar{\nu} l \nu q \bar{q}$  plotted for the class 3d events after the preselection (a) and after the final selection cuts (b). Both distributions are obtained from 2000 data.

Table 4.5: Selection criteria for each channel of class 3. The number of signal ( $N_s$ ), background ( $N_b$ ) and data ( $N_d$ ) events is given in the table for the year 2000 used in the optimization phase. Energies, momenta and masses are expressed in GeV, GeV/c and GeV/c<sup>2</sup>, respectively. Angles are in degrees.

Selection	Cuts	$N_s$	$N_b$	$N_d$
Class 3	$E_{L1} > 25$ $E_{L2} < 20$			
Class 3a $HZ \rightarrow q\bar{q}l\nu q\bar{q}$	$M_{miss} < 95$ $M_{tot}/\sqrt{s} > 0.4$ $E_{L2} < 10$			
Preselections	$E_{12}/\sqrt{s} < 0.005$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_1/\sqrt{s})$ $N_{ch} > 3$ $P_1/\sqrt{s} > 0.05$ $ \cos(\theta_{miss})  < 0.9$	2.7	793.4	823
anti-qq, ll, $\gamma\gamma$	$N_{ch} > 20$ $M_{L1} > 55$ $\theta_{L2} < 41$	1.57	37	43
anti-WW	$E_{L1} < 55$ $T_{no lept} < 0.93$ $\theta_{acc}^{(no lept)} > 137$	1.42	5.7	12
anti-WW	$\ln(y_{45}) > -7.2$ $P_1/\sqrt{s} < 0.25$	1.24	2.9	5
Selection	Cuts	$N_s$	$N_b$	$N_d$
Class 3	$E_{L1} > 25$ $E_{L2} < 20$			
Class 3b $HZ \rightarrow q\bar{q}l\nu l\nu$	$M_{miss} < 95$ $M_{tot}/\sqrt{s} > 0.4$ $E_{L2} > 10$			
Preselections	$E_{12}/\sqrt{s} < 0.4$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_1/\sqrt{s})$ $N_{ch} > 3$ $P_1/\sqrt{s} > 0.05$ $E_{\gamma 1} < 40$ $ \cos(\theta_{miss})  < 0.9$	0.85	86.96	86
anti-qq, ll, $\gamma\gamma$	$N_{ch} > 5$ $P_1/\sqrt{s} > 0.08$ $\theta_{acc} < 178$	0.79	55.40	52
anti-qq, WW	$\theta_{L2} < 41$ $\ln(S) > -2.2$	0.59	3.63	5
anti-WW	$y_{12}^{(no lept)} > 0.23$ $E_{L1} < 55$ $\ln(y_{34}) > -6$ $\theta_{acc}^{(no lept)} > 135$	0.49	0.76	1
Selection	Cuts	$N_s$	$N_b$	$N_d$
Class 3	$E_{L1} > 25$ $E_{L2} < 20$			
Class 3c $HZ \rightarrow \nu\bar{\nu}l\nu l\nu$	$M_{miss} > 95$ $M_{tot}/\sqrt{s} < 0.4$ $E_{L2} > 10$			
Preselections	$E_{12}/\sqrt{s} < 0.005$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_1/\sqrt{s})$ $\theta_{acc} < 170$ $P_1/\sqrt{s} > 0.05$ $ \cos(\theta_{miss})  < 0.9$ $E_{\gamma 1} < 15$	0.17	19.9	20
anti- $\gamma\gamma$	$N_{ch} = 2$ $\theta_{acc} < 138$ $P_1/\sqrt{s} > 0.09$	0.15	11.16	14
anti-ll, WW	$M_{tot}/\sqrt{s} < 0.26$ $\Phi_{acc} > 160$	0.14	7.9	9
anti-WW, ZZ	$M_{miss} > 127$ $\theta_{L2} < 32$	0.12	3.5	5
Selection	Cuts	$N_s$	$N_b$	$N_d$
Class 3	$E_{L1} > 25$ $E_{L2} < 20$			
Class 3d $HZ \rightarrow \nu\bar{\nu}l\nu q\bar{q}$	$M_{miss} > 95$ $M_{tot}/\sqrt{s} < 0.4$ $E_{L2} < 10$			
Preselections	$E_{12}/\sqrt{s} < 0.15$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_1/\sqrt{s})$ $N_{ch} > 3$ $E_{\gamma 1} < 15$ $P_1/\sqrt{s} > 0.05$ $ \cos(\theta_{miss})  < 0.9$	1.22	829.8	836
anti- $\gamma\gamma$ , ll	$\ln(A_{\mu}) > -8$ $\theta_{L2} < 57$	0.94	4.92	3
anti-WW	$E_{L1} < 60$ $E_{L2} < 6$	0.90	3.39	1
anti-WW	$M_{tot}^{(hadr)}/\sqrt{s} < 0.17$ $-5 < \ln(S) < -2$ $\ln(y_{12}^{(no lept)}) > -4$	0.79	0.65	0

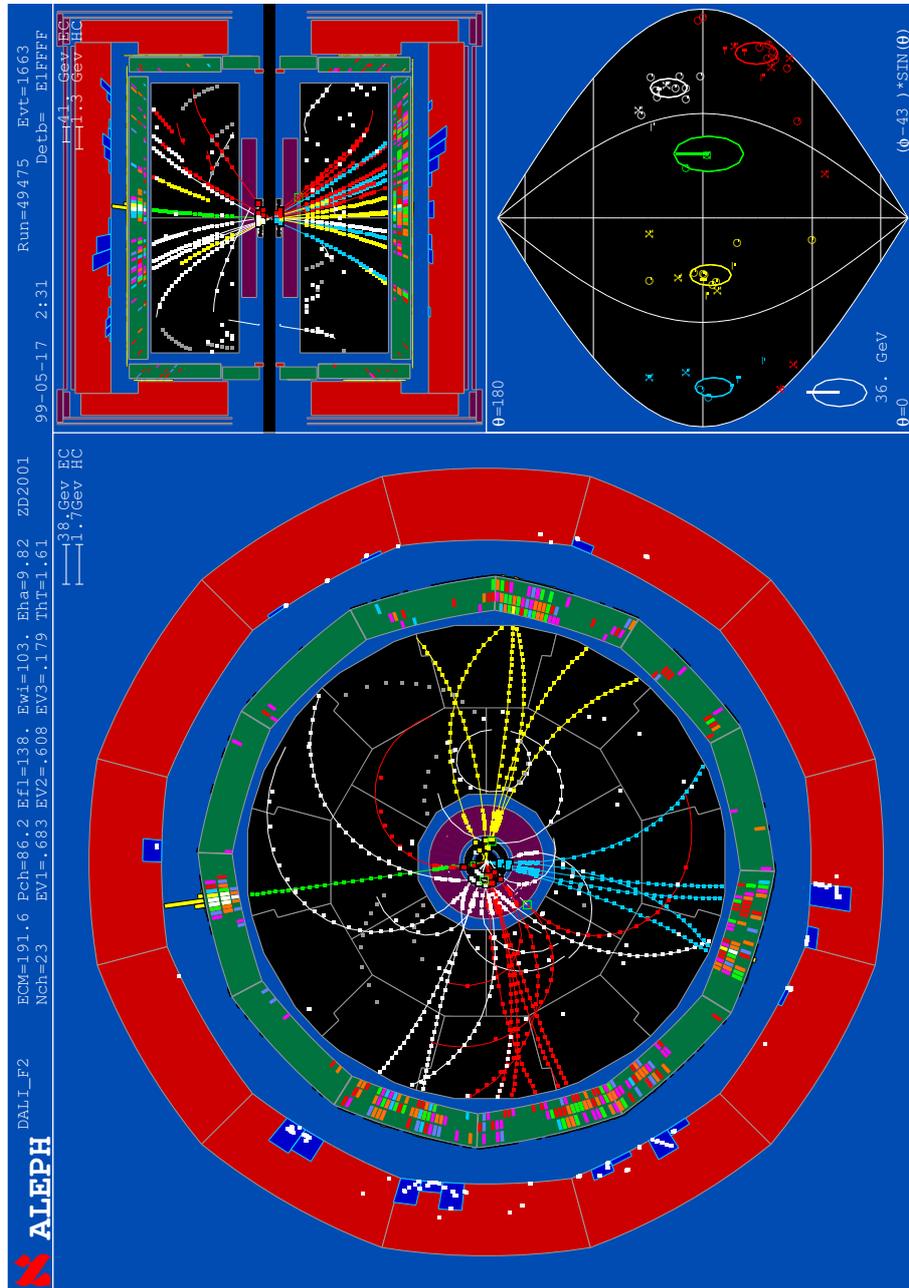


Figure 4.14: An outstanding event from class 3a selection. One sees four jets and one lepton.

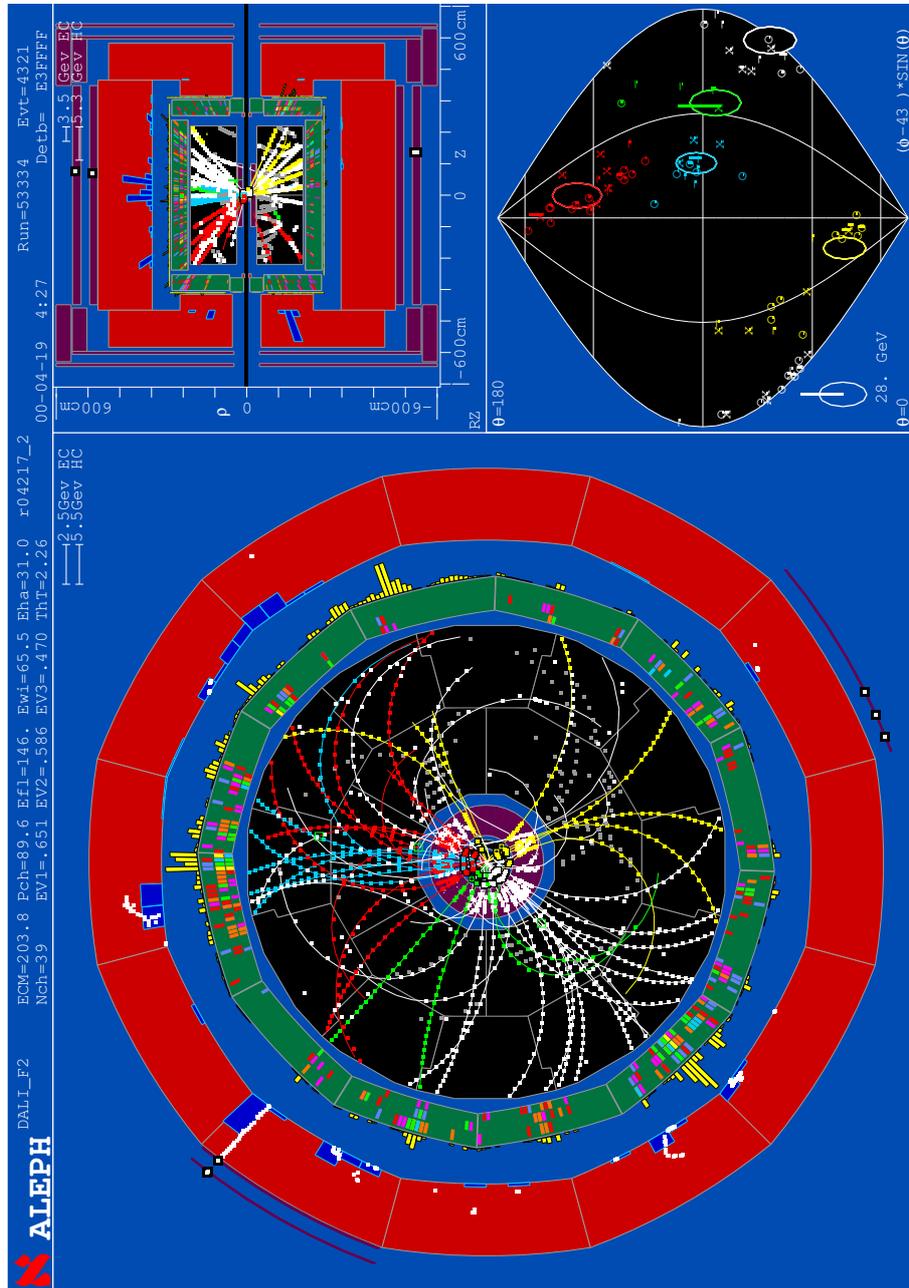


Figure 4.15: An outstanding event from class 3a selection. One sees four jets and one lepton.

### 4.5.5 Class 4a - hadronic Z boson

Class 4a corresponds to the decay channel  $ZH \rightarrow q\bar{q}q\bar{q}l\nu$  and considers topologies with four jets plus one soft lepton in the final state. The event selection is made by the use of  $y_{45}, y_{34}$  (*no lept*) and the isolation of the most anti-parallel lepton with respect to the missing momentum. Though signal efficiency is reduced by 42%, the background is reduced by 84%, which makes this variable very powerful.

A  $\chi^2$  is then built which takes into account the W and Z boson masses, reconstructed from the four jets in the event. For this, jets are paired and the mass of each pair is compared either to the W boson mass or to the Z boson mass. The jet pairing that minimizes the  $\chi^2$  is retained:

$$\chi^2 = \frac{(M_{12} - M_Z)^2}{\sigma^2} + \frac{(M_{34} - M_W)^2}{\sigma^2}, \quad (4.25)$$

where  $M_{12}$  and  $M_{34}$  are the masses of each jet pair and  $\sigma^2$  is the measured resolution on the reconstructed W and Z masses, obtained from signal events. The background is finally reduced by constraining the value of this quantity, the total hadronic mass and the lepton energy. All details of the event selection are given in Table 4.6. The distribution of the discriminant variable, after the preselection and after the final selection is presented in Figures 4.16a and b respectively.

After the final selection, one expects nine background events and 1.41 signal events (for years 1999 and 2000 together). The expected confidence level  $\langle CL_s \rangle$  is 0.55. Eight candidates are actually observed, giving a confidence level of  $CL_s = 0.40$ . The observed deficit of events gives  $CL_b = 0.17$ , corresponding to a  $1.3\sigma$  deviation.

### 4.5.6 Class 4b - invisible Z boson

Class 4b corresponds to the decay channel  $ZH \rightarrow \nu\bar{\nu}q\bar{q}l\nu$ . Topologies with two jets, a soft lepton and missing momentum in the final state are therefore studied. The two jets and the single soft lepton are selected by cutting on  $y_{12}$  and  $y_{34}$ , on the number of tracks from charged particles and on the lepton energy. The remaining background, still 75 times larger than the signal, is reduced by additional cuts on the total hadronic mass, the lepton isolation and on the hadronic acollinearity. Since at least three neutrinos are expected in this channel, a cut on the missing mass ( $M_{miss}$ ) is used in the final selection. The dominating background after the selection is WW events.

The distributions of the discriminant variable, the off-shell W mass (Equation 4.19), after the preselection and after the final selection are presented in Figures 4.17a and b respectively. The expected number of events after the signal selection is shown in Table 4.6. One expects 18.3 background events and 0.85 signal events while 19 events are observed in the data (for years 1999 and 2000 together). This leads to a confidence level  $CL_s$  of 0.67, while 0.77 was

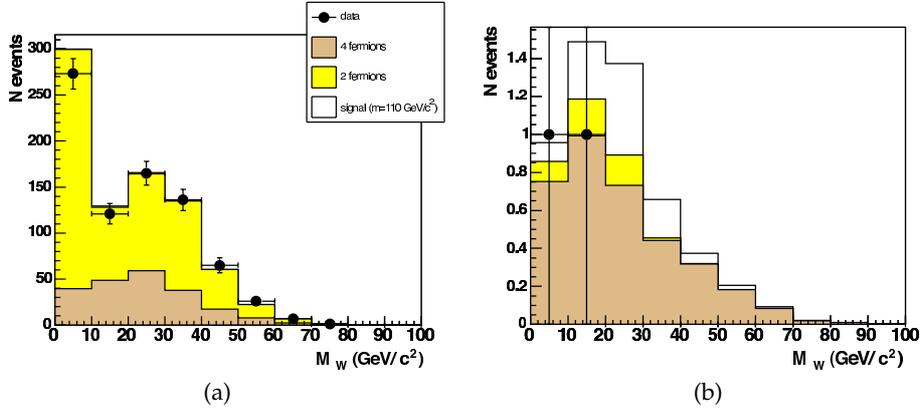


Figure 4.16: Distribution of the discriminant variable for the decay  $HZ \rightarrow q\bar{q}q\bar{q}l\nu$  plotted for the class 4a events after the preselection (a) and after the final selection cuts (b). Both distributions are obtained from 2000 data.

expected. The observed deficit of events gives  $CL_b = 0.22$ , corresponding to a  $1.2\sigma$  deviation.

## 4.6 Study of events with no lepton, or more than one hard lepton

As mentioned previously, the analysis of events with no leptons or more than one hard lepton (classes 1 and 2) was performed by our collaborators, and only a short review of the procedure and results is given. These results will be combined afterward to those of classes 3 and 4, in Section 4.7.

### 4.6.1 Class 1 - Fully hadronic channels

The final state topology where the Higgs boson decays hadronically into four jets via the 2  $W$ s is studied in this class. Depending on the hadronic or invisible decay of the  $Z$  boson, two channels are analyzed:  $ZH \rightarrow q\bar{q}q\bar{q}q\bar{q}$  (class 1a) and  $ZH \rightarrow \nu\bar{\nu}q\bar{q}q\bar{q}$  (class 1b). The signal final state of class 1a is characterized by six jets and is balanced in energy and momenta. For class 1b, the final state is composed of four jets and two neutrinos, leading to an important energy and mass imbalance. These two channels are easily separated by a selection on the final state missing mass. In both channels, the background comes from  $WW$ ,  $q\bar{q}$  and  $ZZ$  production (with a small contribution from the  $W\text{e}\nu$  process in the missing mass channel).

Jet counting, track multiplicity and other topological variables are used to isolate the  $ZH \rightarrow q\bar{q}q\bar{q}q\bar{q}$  events. The selection cuts are detailed in Table 4.7.

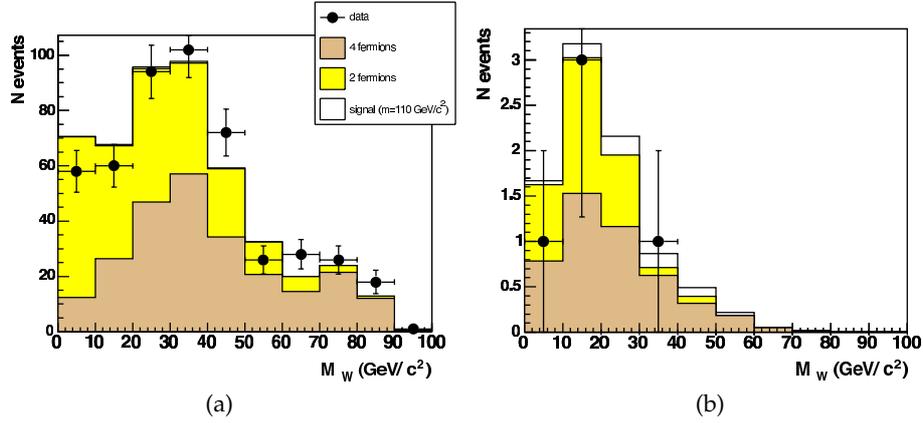


Figure 4.17: Distribution of the discriminant variable for the decay  $HZ \rightarrow \nu\bar{\nu}q\bar{q}l\nu$  plotted for the class 4b events after the preselection (a) and after the final selection cuts (b). Both distributions are obtained from 2000 data.

Table 4.6: Selection criteria for each channel of class 4. The number of signal ( $N_s$ ), background ( $N_b$ ) and data ( $N_d$ ) events is given in the table for the year 2000, used in the optimization phase. Energies, momenta and masses are expressed in GeV, GeV/c and GeV/c<sup>2</sup>, respectively. Angles are in degrees.

Selection	Cuts	$N_s$	$N_b$	$N_d$
Class 4	$D_{l4} > 13$ $E_{l1} < 25$			
Class 4a ZH $\rightarrow q\bar{q}q\bar{q}l\nu$	$M_{tot}/\sqrt{s} > 0.6$			
Preselections	$E_{l2}/\sqrt{s} < 0.1$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_l/\sqrt{s})$ $N_{th} > 10$ $ \cos(\theta_{miss})  < 0.95$	4.6	818	794
Group 1	$\ln(y_{l4}) > -6$ $\ln(y_{l4}^{(no\ lept)}) > -4.6$ $\ln(I_{L4}) > -3.5$	1.4	7.4	8
Group 2	$\chi^2 < 0.02$ $M_{tot}^{(hadr)}/\sqrt{s} < 0.95$ $E_{L4} > 7$	1.17	4	2
Class 4b ZH $\rightarrow \nu\bar{\nu}q\bar{q}l\nu$	$M_{tot}/\sqrt{s} < 0.6$			
Preselections	$E_{l2}/\sqrt{s} < 0.4$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_l/\sqrt{s})$ $N_{th} > 3$ $P_l/\sqrt{s} > 0.05$ $ \cos(\theta_{miss})  < 0.95$	2.4	478	485
Group 1	$\ln(y_{l2}^{(no\ lept)}) > -7$ $\ln(y_{l23}^{(no\ lept)}) > -2.6$ $N_{th} > 9$ $E_{L4} > 7$	1.35	99.7	82
Group 2	$M_{tot}^{(hadr)} \in [50, 103]$ $\ln(I_{L4}) > -4.5$ $\theta_{acc}^{(no\ lept)} > 140$ $M_{miss} > 65$	0.69	8.0	5

Table 4.7: Selection criteria for each channel of class 1. The number of signal ( $N_s$ ), background ( $N_b$ ) and data ( $N_d$ ) events are given in the table for the year 2000. Energies, momenta and masses are expressed in GeV, GeV/c and GeV/c<sup>2</sup>, respectively. Angles are in degrees.

Selection	Cuts	$N_s$	$N_b$	$N_d$
Class 1	$D_{14} < 13$ $E_{L1} < 25$			
Class 1a	$M_{miss} < 60$			
ZH $\rightarrow$ qq $\bar{q}\bar{q}$ qq				
Preselection				
	$N_{ch} > 25$ $M_{tot}/\sqrt{s} > 0.6$ $\ln(y_{34}) > -5.3$ $0.1 < S < 0.85$ $P_T/\sqrt{s} < 0.15$	10.9	1452.0	1401
Topological	$S > 0.13$ $N_{ch} > 32$ $\ln(y_{34}) > -7.2$ $\ln(y_{12} + y_{34} + y_{56}) > -0.83$ $N_{ch,FB}^{min} > 0$ $(\theta_{FB}^{min} + \theta_{FB}^{max}) > 190$	10.7 9.17 8.18 8.07 6.28 6.21	1374.1 866.9 494.5 476.4 231.5 212.3	1325 845 481 456 247 229
Masses	$M_1 < 117$ $M_3 > 13$	6.10 5.74	186.7 163.7	200 168

Selection	Cuts	$N_s$	$N_b$	$N_d$
Class 1	$D_{14} < 13$ $E_{L1} < 25$			
Class 1b	$M_{miss} > 60$			
ZH $\rightarrow$ $\nu\bar{\nu}$ qq $\bar{q}\bar{q}$				
Preselection	$\cos(\theta_{miss}) < 0.9$ $E_{12}/\sqrt{s} < 0.05$			
	$N_{ch} > 12$ $M_{tot}/\sqrt{s} > 0.35$ $\ln(y_{34}) > -7.6$ $0.03 < S < 0.8$ $P_T/\sqrt{s} < 0.2$	3.60	190.4	211
Topological	$N_{ch} > 24$ $\ln(y_{34}) > -6.6$	1.77 1.73	31.2 28.2	46 43
Anti-qq	$\Phi_{acc} < 179$ $P_T/\sqrt{s} > 0.035$	1.69 1.63	26.1 22.0	38 28
Anti-WW	$T > 0.74$ $P_T/\sqrt{s} < 0.12$ $N_{ch,FB}^{min} > 0$ $\theta_{FB}^{min} > 23$ $(\theta_{FB}^{min} + \theta_{FB}^{max}) > 170$	1.33 1.25 1.23 1.16 1.11	16.3 12.1 11.3 8.61 8.18	23 19 18 15 13

The channel  $ZH \rightarrow \nu\bar{\nu}q\bar{q}q\bar{q}$  is characterized by a final state with a large missing mass from the undetected neutrinos coming from the decay of the Z. This can be exploited to tag those events. It is used together with  $P_T/\sqrt{s}$  and the thrust to tag signal events. The selection cuts are detailed in Table 4.7.

#### 4.6.2 Class 2 - Channels with more than one hard lepton

Class 2 is characterized by final states with at least two “hard” leptons from the leptonic Z decay. Requiring the third most energetic electron or muon to have an energy lower than 8 GeV separates the two-lepton channel from the topologies involving more than two leptons (soft hadronic taus are not considered in the study). The  $ZH \rightarrow l^{+1-}q\bar{q}lv$  channel is identified by requiring that the total hadronic energy  $E_{had}$  is greater than 60 GeV.  $ZH \rightarrow l^{+1-}lvlv$  and  $ZH \rightarrow l^{+1-}lvq\bar{q}$  both require  $E_{had} < 60$  GeV and are separated from each other by the requirement that the former has exactly four tracks from charged particles.  $ZH \rightarrow l^{+1-}lvlv$  (l being an electron or a muon) is then uniquely distinguished from  $ZH \rightarrow l^{+1-}lvq\bar{q}$  since it contains strictly four charged particle tracks in the final state. Within the two-lepton channel, one finds events with taus decaying hadronically. The separation of these events from those where all Ws decay hadronically is achieved with a cut on the total hadronic energy. Again, the selection procedure will only be sketched. Details can be found in Table 4.8 and in reference [40].

To select the  $ZH \rightarrow l^{+1-}q\bar{q}q\bar{q}$  channel, one requires that the energy of the third lepton is less than 8 GeV and that the total hadronic energy ( $E_{had}$ ) is more than 60 GeV. The dominant backgrounds after preselection are  $q\bar{q}$ , semi-

leptonic WW as well as ZZ events where one of the gauge bosons decays into hadrons and the other into charged leptons. The rejection of  $q\bar{q}$  and WW events is achieved by applying a cut on the variable  $y_{45}$  and requiring the reconstructed mass of the two leptons to be in a 14 GeV window around the Z mass. The remaining ZZ background is reduced by requiring the two leptons to be back to back and the sum of their energy to be less than 95 GeV.

The  $ZH \rightarrow l^+l^- q\bar{q}lv$  channel is required to have a third lepton with more than 8 GeV and a total hadronic energy of more than 60 GeV. The selection proceeds in a similar way as in the  $ZH \rightarrow l^+l^- q\bar{q}q\bar{q}$  channel by reconstructing the Z peak. The logarithm of the isolation for the lepton that is the most anti-parallel to the missing momentum is used in the selection. This variable,  $\ln(I_{L_A})$ , is sensitive to the isolation of the soft lepton emitted by the off-shell W. The lepton tends to be more isolated for the signal than for the background, where the third lepton is often produced in the hadronization of heavy quarks.

The  $ZH \rightarrow l^+l^- lvq\bar{q}$  channel is required to have a third lepton with more than 8 GeV and a total hadronic energy of less than 60 GeV. Compared to the channel  $ZH \rightarrow l^+l^- q\bar{q}lv$ , the missing transverse momentum is higher. The preselection rejects most of the background by requiring a total transverse momentum greater than 1% of the total center-of-mass energy. The cuts that typically reduce the semi-leptonic WW background do not improve the background rejection much. The mass window for the reconstructed Z mass is also broader than in the  $ZH \rightarrow l^+l^- q\bar{q}lv$  case since the third lepton is sometimes misidentified as one of the leptons coming from the Z. Finally a cut on the  $P_t$  of the Z boson is applied to further reduce the ZZ background.

The  $ZH \rightarrow l^+l^- lvlv$  channel is characterized by a small branching fraction (5% of the events in class 2) but with a clear topology: four leptons in the final state, one of them being soft. This channel is then selected by cutting on the energy of the third most energetic lepton ( $E_{L3} > 8$  GeV) and on the total hadronic energy. One also requires strictly four charged-particle tracks. To further reduce the  $l^+l^-$  background, cuts on the missing transverse momentum, the acoplanarity and the thrust are applied. Finally the remaining WW events are rejected by requiring that the most anti-parallel lepton with respect to the missing momentum be well isolated.

## 4.7 Results

### 4.7.1 Systematic uncertainties

Systematic errors stem from uncertainties and inaccuracies in the Monte-Carlo simulation. The main contribution comes from the Monte-Carlo statistics. The uncertainties from the production cross-sections, from the simulation of the calorimeter energy scale and from hadronization processes are also taken into account. These are handled as explained in Section 4.4. Systematics were applied on the background only. This is a realistic and conservative approach

Table 4.8: Selection criteria for each channel of class 2. The number of signal ( $N_s$ ), background ( $N_b$ ) and data ( $N_d$ ) events is given in the table for the year 2000. Energies, momenta and masses are expressed in GeV, GeV/c and GeV/c<sup>2</sup>, respectively. Angles are in degrees.

Cuts	$N_s$	$N_b$	$N_d$
Class 2	$E_{L1} > 25$		
	$E_{L2} > 20$		
Class 2a	$E_{L3} < 8$		
ZH $\rightarrow l^+ l^- q\bar{q} q\bar{q}$	$E_{had} > 60$		
Preselections	$E_{12}/\sqrt{s} < 0.4$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_1/\sqrt{s})$ $N_{ch} > 8$	1.03	48.8 39
(anti WW,qq)	$ M_{Z_{tot}} - M_{Z_0}  < 14$ $\ln(\gamma_{45}) \geq -7$	0.62	6.29 2
(anti ZZ)	$E_{L1} + E_{L2} < 95$ $\theta_{L1L2} > 135$	0.54	0.67 2

Cuts	$N_s$	$N_b$	$N_d$
Class 2	$E_{L1} > 25$ $E_{L2} > 20$		
Class 2T	$E_{L3} > 8$		
ZH $\rightarrow l^+ l^- \tau\nu q\bar{q}$	$E_{had} < 60$		
Preselection	$E_{12}/\sqrt{s} < 0.4$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_1/\sqrt{s})$ $P_1/\sqrt{s} > 0.002$	0.08	2.59 2
Selection	$P_{T2} < 60$ $ M_{Z_{tot}} - M_{Z_0}  < 23$	0.06	0.46 1

Cuts	$N_s$	$N_b$	$N_d$
Class 2	$E_{L1} > 25$ $E_{L2} > 20$		
Class 2b	$E_{L3} > 8$		
ZH $\rightarrow l^+ l^- q\bar{q} l\nu$	$E_{had} > 60$		
Preselections	$E_{12}/\sqrt{s} < 0.4$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_1/\sqrt{s})$ $N_{ch} > 7$	0.29	10.7 12
(anti q $\bar{q}$ , WW)	$ M_{Z_{tot}} - M_{Z_0}  < 20$ $\ln(\gamma_{45}) > -8$ $\ln(L_{41}) > -7$	0.19	1.08 0
(anti ZZ)	$E_{L1} + E_{L2} < 98$ $\theta_{L1L2} > 142$	0.18	0.16 0

Cuts	$N_s$	$N_b$	$N_d$
Class 2	$E_{L1} > 25$ $E_{L2} > 20$		
Class 2c	$E_{L3} > 8$		
ZH $\rightarrow l^+ l^- l\nu q\bar{q}$	$E_{had} < 60$		
Preselection	$E_{12}/\sqrt{s} < 0.4$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_1/\sqrt{s})$ $P_1/\sqrt{s} > 0.01$ $N_{ch} \geq 4$	0.23	1.73 2
(anti q $\bar{q}$ , WW)	$ M_{Z_{tot}} - M_{Z_0}  < 23$ $\ln(\gamma_{41}) > -11$ $\ln(\gamma_{45}) > -11$	0.20	0.18 0
(anti ZZ)	$P_{T2} < 60$	0.20	0.17 0

Cuts	$N_s$	$N_b$	$N_d$
Class 2	$E_{L1} > 25$ $E_{L2} > 20$		
Class 2d	$E_{L3} > 8$		
ZH $\rightarrow l^+ l^- l\nu l\nu$	$E_{had} < 60$ $N_{ch} = 4$		
Preselection	$E_{12}/\sqrt{s} < 0.4$ $M_{tot}/\sqrt{s} > (0.2 - 6 * P_1/\sqrt{s})$ $P_1/\sqrt{s} > 0.11$	0.10	3.13 3
(anti ZZ)	$T < 0.98$ $\Phi_{230} < 176$	0.09	0.93 1
(anti WW)	$\ln(L_{41}) > -9$	0.09	0.58 0

as long as no signal is observed, since only the background can move upwards and make the observation of the signal more difficult.

For the four different classes, after applying the full selection, the main remaining background comes from the  $W$  pair production with NLO QCD effects producing a six-fermion-like final state. The cross-section uncertainty on this process is taken to be 2%, which corresponds to the uncertainty on  $\alpha_s$ , from [4]. The systematic errors related to the simulation of the calorimeter are evaluated by modifying the calorimeter calibration within the uncertainties, and lead to a mean uncertainty of 3% on the remaining background level. The uncertainties on the hadronization process are evaluated by comparing  $WW$  events produced with the JETSET[45] Monte Carlo and  $WW$  events from the ARIADNE[46] Monte Carlo. The associated mean systematic error is 6%. Both the hadronization and calorimetric uncertainties are evaluated separately for each sub-channel and are presented in Table 4.9. When extracting these results, the effect of statistical uncertainties (due to the fact that different Monte-Carlo samples were used) has been evaluated and subtracted. The effect on the presented uncertainties is small in most cases.

Table 4.9: Levels of systematic uncertainties considered for each sub-channel.

Channels	Calorimetry	Hadronization
1a	0%	2%
1b	6%	8%
2	0%	0%
3a	2%	4%
3b	0%	14%
3c	2%	8%
3d	12%	8%
4a	2%	0%
4b	0%	9%

### 4.7.2 Combined limits

The number of signal, background and data events for the years 1999 and 2000, for each channel, is given in Table 4.10. The Higgs boson mass hypothesis of  $110 \text{ GeV}/c^2$  chosen for the optimization process is used for the signal. The combined expected values of the signal and background confidence levels are 0.08 and 0.50 respectively. Such a  $110 \text{ GeV}/c^2$  fermiophobic Higgs boson is hence expected to be excluded at 92% C.L. . The observed values for the signal and background confidence levels are 0.26 and 0.87, respectively. This shows that there is an excess of events with respect to the predicted background.

Until now, we have been working in a naive scenario, where the Higgs boson decays exclusively into  $WW$  pairs. This is not very realistic, since the

Table 4.10: Number of signal ( $N_s$ ), background ( $N_b$ ) and data ( $N_d$ ) events, as well as the value of the expected and observed confidence levels for each selection. Data from 1999 and 2000 are considered together.

Class	$N_s$	$N_b$	$N_d$	$\langle \text{CL}_s \rangle$	$\text{CL}_s$
1a : $\text{HZ} \rightarrow q\bar{q}q\bar{q}q\bar{q}$	$6.80 \pm 0.06$	$372.9 \pm 1.7$	360	0.599	0.503
1b : $\text{HZ} \rightarrow \nu\bar{\nu}q\bar{q}q\bar{q}$	$1.35 \pm 0.03$	$18.0 \pm 0.4$	20	0.577	0.790
Class 1 combined	$8.15 \pm 0.07$	$390.9 \pm 1.7$	380	0.44	0.53
2a : $\text{HZ} \rightarrow l^+l^-q\bar{q}q\bar{q}$	$0.64 \pm 0.02$	$2.41 \pm 0.07$	5	0.57	0.91
2T : $\text{HZ} \rightarrow l^+l^- \tau\nu q\bar{q}$	$0.070 \pm 0.006$	$1.21 \pm 0.08$	1	0.94	0.96
2b : $\text{HZ} \rightarrow l^+l^-q\bar{q}l\nu$	$0.21 \pm 0.01$	$0.56 \pm 0.04$	2	0.81	0.83
2c : $\text{HZ} \rightarrow l^+l^-l\nu q\bar{q}$	$0.24 \pm 0.01$	$0.31 \pm 0.05$	0	0.79	0.79
2d : $\text{HZ} \rightarrow l^+l^-l\nu l\nu$	$0.113 \pm 0.008$	$1.33 \pm 0.12$	1	0.91	0.90
Class 2 combined	$1.24 \pm 0.03$	$5.78 \pm 0.17$	9	0.43	0.73
3a : $\text{HZ} \rightarrow q\bar{q}l\nu q\bar{q}$	$1.27 \pm 0.03$	$6.6 \pm 0.2$	13	0.47	0.84
3b : $\text{HZ} \rightarrow q\bar{q}l\nu l\nu$	$1.37 \pm 0.02$	$2.01 \pm 0.12$	3	0.62	0.66
3c : $\text{HZ} \rightarrow \nu\bar{\nu}l\nu l\nu$	$0.150 \pm 0.009$	$6.6 \pm 0.2$	14	0.95	0.99
3d : $\text{HZ} \rightarrow \nu\bar{\nu}l\nu q\bar{q}$	$0.89 \pm 0.02$	$0.99 \pm 0.15$	1	0.44	0.43
Class 3 combined	$3.08 \pm 0.04$	$17.0 \pm 0.4$	32	0.23	0.47
4a : $\text{HZ} \rightarrow q\bar{q}q\bar{q}l\nu$	$1.41 \pm 0.03$	$9.0 \pm 0.2$	8	0.55	0.40
4b : $\text{HZ} \rightarrow \nu\bar{\nu}q\bar{q}l\nu$	$0.85 \pm 0.02$	$18.3 \pm 0.4$	19	0.78	0.68
Class 4 combined	$2.23 \pm 0.03$	$29.3 \pm 0.5$	27	0.50	0.30
All combined	$14.7 \pm 0.09$	$442.8 \pm 1.9$	448	0.08	0.26

Higgs boson should at least couple to Z bosons in order to be produced by the Higgsstrahlung process. It is nevertheless useful as a reference. As mentioned in the introduction, a benchmark model is defined by considering Standard-Model couplings to bosons, and null couplings to fermions. A modified version of HDECAY (Figure 4.1) is used for that purpose. Motivations for this scenario come from the 2HDM type-I models discussed in Section 2.3.2, where we are then concerned with the lightest CP-even Higgs boson (h). Since results are presented in a model-independent way, we will nevertheless stick to the Standard-Model notations, and present our results as a function of  $m_H$ .

The compatibility between data and background with and without the signal is evaluated with the frequentist log-likelihood ratio method presented in Section 4.4. Figure 4.18 shows the evolution of  $-2\ln Q$  as a function of the Higgs boson mass for each of the four classes. Class 3 has clearly the highest discriminative power, as the signal hypothesis is up to three standard deviations away from the background-only hypothesis. Results from the four classes are then combined, and the final log-likelihood ratio, depicted in Figure 4.19, is computed as a function of the Higgs boson mass. After combination, a  $1.5\sigma$  excess of events with respect to the expected background level

is observed. This comes mainly from channel 3, where a  $2\sigma$  excess of events is observed in 3a and 3c. Among other kinematical variables, the final  $P_t/\sqrt{s}$  (Figure 4.20) distribution is not compatible with the signal hypothesis. Since this excess does not depend on the hypothetical mass, and since no other indication of a signal is found in the data, this excess is attributed to a statistical fluctuation, and an upper limit on the signal cross-section is set as a function of the Higgs boson mass.

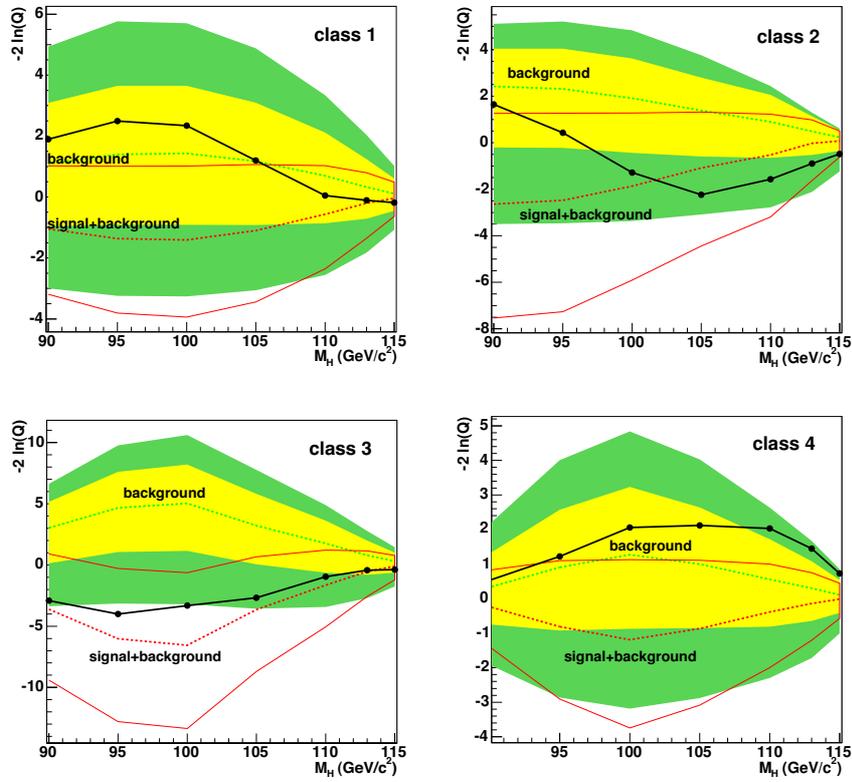


Figure 4.18: Log-likelihood ratio,  $-2\ln Q$ , as a function of the test mass  $M_H$  with all data taken from 192 GeV to 209 GeV for each of the 4 classes. The solid line is the result obtained from the data. The expected background-only and signal-plus-background likelihoods are indicated by the dashed lines; the light and dark shaded bands around the background expectation contain 68% and 95% of the simulated background-only experiments, respectively. The one-sigma bands for the signal hypothesis are also shown.

For each class in the analysis, an upper limit on the cross-section production at a given Higgs boson mass is deduced. Figure 4.21 presents the resulting

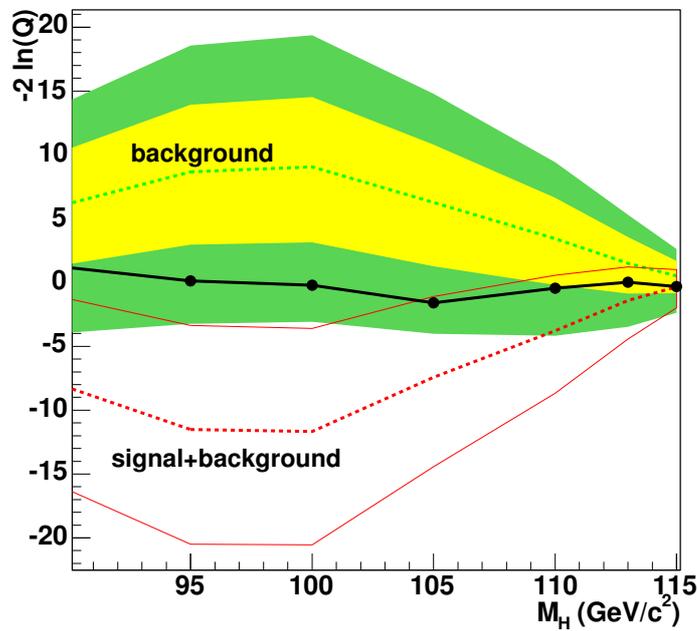


Figure 4.19: Log-likelihood ratio,  $-2\ln Q$ , as a function of the test mass  $M_H$  with all data taken from 192 GeV to 209 GeV. The solid line is the result obtained from the data. The expected background-only and signal-plus-background likelihoods are indicated by the dashed lines; the light and dark shaded bands around the background expectation contain 68% and 95% of the simulated background-only experiments, respectively. The one-sigma band for the signal hypothesis is also shown.

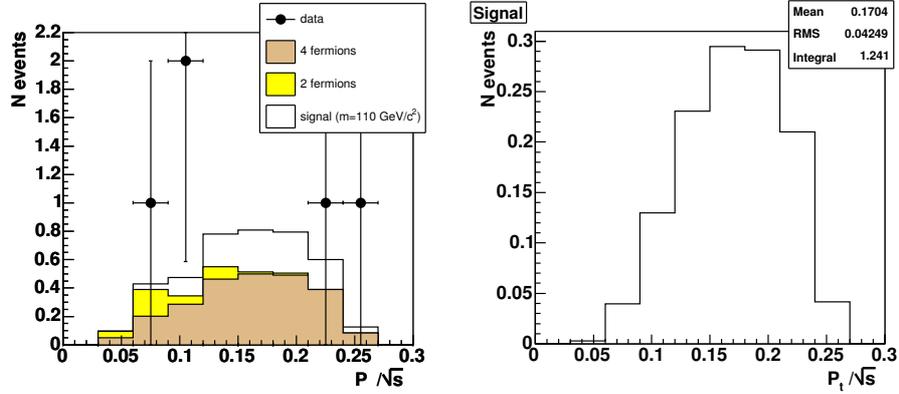


Figure 4.20: Distribution of the total transverse momentum for events of class 3a, after all selection, obtained for the year 2000.

95% C.L. exclusion region in the plane defined by  $\xi^2 = B(H \rightarrow WW)\sigma(e^+e^- \rightarrow Hff)/\sigma^{SM}(e^+e^- \rightarrow Hff)$  and the Higgs boson mass. The third class has the highest reach, as it provides the best compromise between a clean signature and signal sensitivity. The other three channels are nevertheless as important in order to get a competitive limit. Combining all four analyses, the 95% C.L. upper limit on  $\xi^2$  as a function of the Higgs boson mass is given in Figure 4.22. The value of  $\xi^2$  for the benchmark fermiophobic Higgs model is drawn as a function of the Higgs boson mass hypothesis as the full line in the figure. The aforementioned excess of events in channel 3 gives a final observed limit lower than the one expected.

In the context of a given model, mass exclusion limits can be extracted from upper limits on the cross-section. In the naive  $B(H \rightarrow WW) = 1$  case, the expected limit is  $107.5 \text{ GeV}/c^2$ . In the benchmark fermiophobic Higgs boson scenario, one expects to put a limit on a Higgs boson with a mass between  $97.5 \text{ GeV}/c^2$  and  $104 \text{ GeV}/c^2$ . Due to the excess of events observed, no limit at 95% C.L. on the presence of a Higgs boson can be obtained.

### 4.7.3 Combined results with previous ALEPH studies

In addition, a model-independent limit can be derived by scanning the  $H \rightarrow \gamma\gamma$  and  $H \rightarrow WW$  branching fractions. This can be conveniently parametrized as

$$\text{BR}_{\text{bosons}} = \text{BR}_{H \rightarrow \gamma\gamma} + \text{BR}_{H \rightarrow WW} + \text{BR}_{H \rightarrow ZZ}, \quad (4.26)$$

$$R_{\gamma\gamma} = \text{BR}_{H \rightarrow \gamma\gamma} / \text{BR}_{\text{bosons}}, \quad (4.27)$$

where  $R_{\gamma\gamma}$  represents the fraction of bosonic decays into photons pairs, and ranges from zero to one.  $\text{BR}_{\text{bosons}}$  represents the total bosonic branching fraction. The best limit is obtained by combining the present results with the one

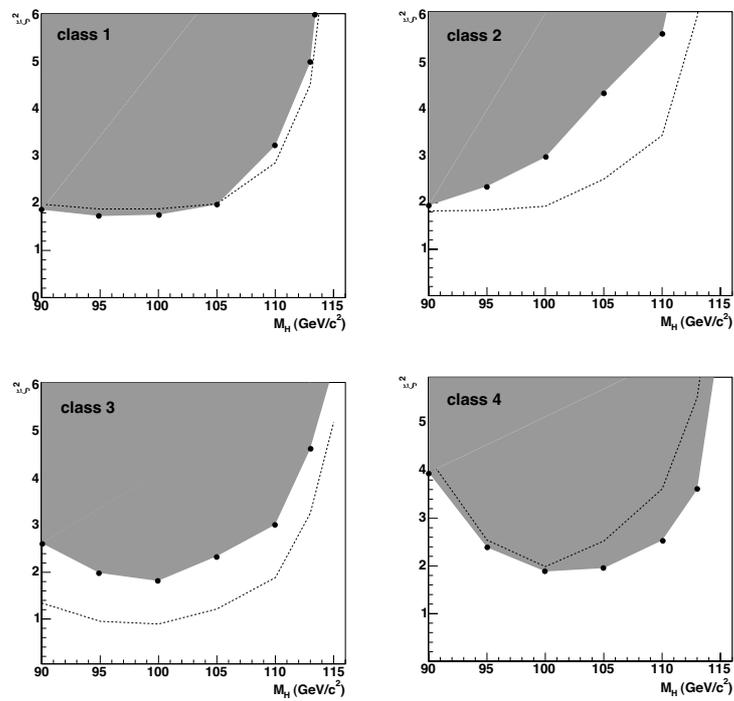


Figure 4.21: Limit on  $\xi^2$  (defined in the text) as a function of the Higgs boson mass, for the four different classes. The dotted line corresponds to the expected limit while the 95% C.L. excluded region is shown by the gray area.

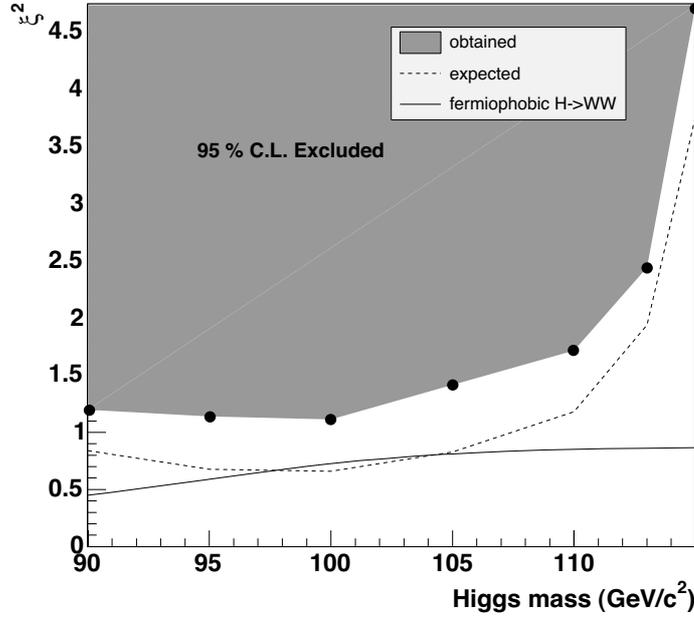


Figure 4.22: Limit on  $\xi^2$  as a function of the Higgs boson mass. The dotted line corresponds to the expected limit while the 95% C.L. excluded region is shown by the gray area. The  $\xi^2$  curve corresponding to the benchmark model described in the text is also presented.

previously published by ALEPH exploiting  $H \rightarrow \gamma\gamma$  (Figure 4.23) [27]. The 95% C.L. limit on  $\text{BR}_{\text{bosons}}$  is determined at each point of the  $m_H$  versus  $R_{\gamma\gamma}$  plane, resulting in the exclusion curves of Figure 4.24. In the benchmark fermiophobic scenario, the fermiophobic Higgs boson is excluded up to 105.8  $\text{GeV}/c^2$ . It represents a significant improvement with respect to the limit obtained considering  $H \rightarrow \gamma\gamma$  alone. This limit is found to be 104.5  $\text{GeV}/c^2$  using the same likelihood ratio method [47], even if ALEPH published a limit of 105.4  $\text{GeV}/c^2$  using another statistical method. It must be noticed that the limit would have been stronger without the excess of events observed in the  $WW$  channel. The expected combined limit is found to be 111.4  $\text{GeV}/c^2$ . The combined limit is also less sensitive to systematics. Not taking systematic uncertainties into account only improves the limit by 200  $\text{MeV}/c^2$ .

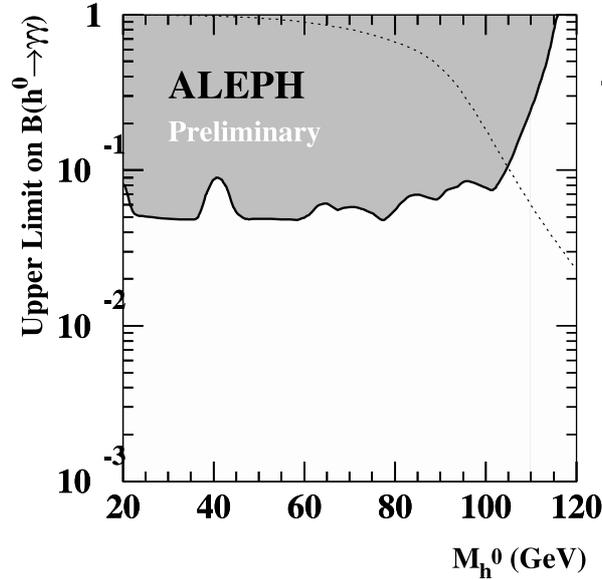


Figure 4.23: ALEPH experimental limit on fermiophobic Higgs bosons for the photonic decay mode obtained using the  $CL_s$  method. The 95% confidence level upper limit on the di-photon branching fraction is shown as a function of Higgs mass; the shaded regions are excluded. Also shown (dotted line) is the branching fraction obtained for the benchmark fermiophobic model. From [47].

#### 4.7.4 Comparison with other studies

The search for a fermiophobic Higgs boson in the WW channel has also been considered by the L3 collaboration [29]. In this paper, the L3 collaboration considers data collected at center-of-mass energies from 200 GeV to 209 GeV. It concentrates on the six dominant final states. The analyses adopt a common preselection procedure, followed by a selection using a neural network approach. The L3 collaboration also considered ZZ decays of the Higgs boson in some of the analyses, which enhances the signal expected in the benchmark scenario. This introduces new theoretical assumptions (namely on the  $BR(WW)/BR(ZZ)$  ratio), but enhances the expected signal, and hence the final observability.

We also considered the use of neural networks to isolate signal over the background. Neural networks are more and more used in various fields for data analysis and classification, both for research and commercial applications. More than 50% of neural networks are multilayer perceptrons, simple feed-forward networks made of neurons, characterized by a local bias and weighted links in between. A layer of neurons receives, normalizes and forwards inputs to a first hidden layer. Each neuron in any subsequent layer first

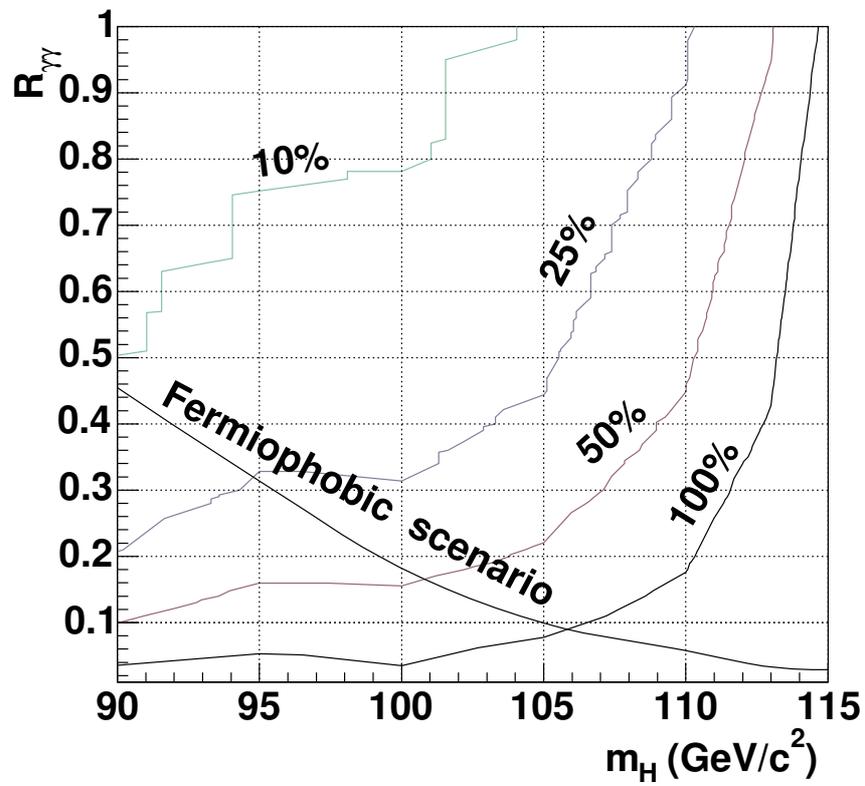


Figure 4.24: The 95% C.L. limit for  $\text{BR}_{\text{bosons}}$  as a function of  $m_H$  and  $R_{\gamma\gamma}$ . The solid lines indicate the upper limit of exclusion regions. The crossing point between the “ $\text{BR}_{\text{bosons}} = 100\%$ ” line and the “Fermiophobic scenario” line provides the lower limit on the Higgs boson mass in the benchmark scenario:  $m_H > 105.8 \text{ GeV}/c^2$ .

computes a linear combination of the outputs of the previous layer. The output of the neuron is then a function  $f$  of that combination, with  $f$  being linear for output neurons or a sigmoid for hidden layers. Such a network must be trained, so as to optimize all the internal parameters to best fit the desired output. This computation is called “back-propagation of the errors”. In the trend to provide tools to perform complete analysis with C++ tools, neural network algorithms were ported from PAW (from the MLPFIT[48] package) to ROOT. Compared to the simpler cut-based selection procedure, no significant advantage has been found in using neural networks, while new problems were arising. Indeed, neural networks require far more statistics for the training and validation of the test function. They also introduce difficulties in controlling systematic effects, as strong non-linear dependences on some variables may be hidden deeply. Finally, a neural-network based analysis will often be very specific to a given signal at a given mass, as it exploits many more correlations between observables. It implies that the analysis has to be retuned for the different mass points, which we did not want.

L3 results are presented in Figure 4.25. The expected limit is comparable to our result in the mass range we considered (for masses of  $90 \text{ GeV}/c^2$  and higher), but no excess of events is observed so that the obtained limit is close to the one expected. In the benchmark scenario, the fermiophobic Higgs is limited at 95%C.L. to have  $m_H > 104.6 \text{ GeV}/c^2$ , compared to an expected limit of  $m_H > 108 \text{ GeV}/c^2$ . The analysis is then combined to the L3 fermiophobic Higgs boson search in the  $\gamma\gamma$  channel (Figure 4.26). In the benchmark scenario, the fermiophobic Higgs is limited at 95%C.L. to have a mass  $m_H$  above  $108.3 \text{ GeV}/c^2$ , compared to an expected limit of  $110.7 \text{ GeV}/c^2$ , when combined to the search in  $\gamma\gamma$  channel. The limit expected by the L3 collaboration is thus slightly lower than the expected ALEPH limit, but the final result is better, due to the excess of events in the ALEPH  $H \rightarrow WW$  analysis. Looking more closely at Figure 4.26, it is interesting to compare it to our combination (Figure 4.24). The two plots are similar at first glance. A significant difference is observed in the behavior of the “100%” curve, which crosses the  $R_{\gamma\gamma} = 0$  line for L3 but not for ALEPH. Therefore, L3 is able to put an absolute limit independent of  $R_{\gamma\gamma}$ , that ALEPH cannot. This reflects the fact that L3 obtained an exclusion region with the WW search alone, that we were not able to obtain due to the small excess of events in the data.

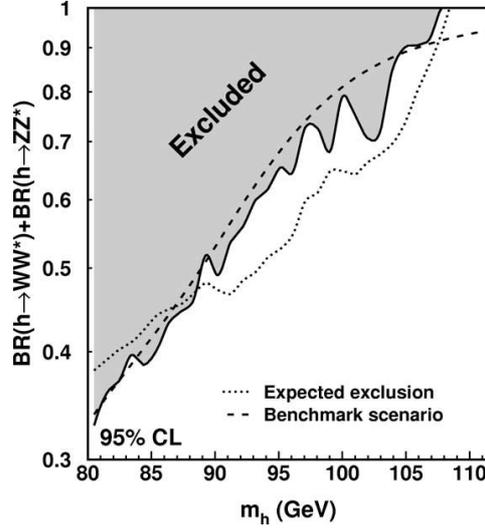


Figure 4.25: L3 excluded values at 95% C.L. of  $\text{BR}(H \rightarrow WW) + \text{BR}(H \rightarrow ZZ)$  as a function of the Higgs boson mass (solid line) in assumption of the Standard-Model production cross-section. The expected limit (dashed line) and the fermiophobic benchmark model prediction (dotted line) are also presented. From [29].

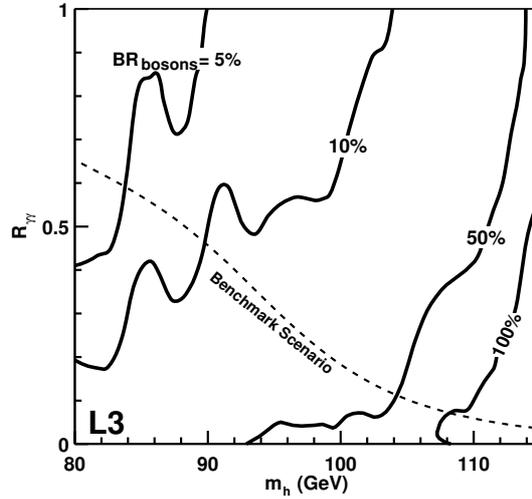


Figure 4.26: The 95% CL limit for  $\text{BR}_{\text{bosons}}$  obtained by the L3 collaboration as a function of  $m_H$  and  $R_{\gamma\gamma}$ . The solid lines indicate the borders between regions of exclusion. The crossing point between the  $\text{BR}_{\text{bosons}} = 100\%$  line and the dashed line provides the lower limit on the Higgs boson mass in the benchmark scenario:  $m_H > 108.3 \text{ GeV}/c^2$ . The minimum value of  $m_H$  on the  $\text{BR}_{\text{bosons}} = 100\%$  contour gives a lower mass limit for any model where the Higgs decays exclusively into electroweak boson pairs:  $m_H > 107 \text{ GeV}/c^2$ . From [29].



## **Part III**

# **CMS - WH associated production**



## The Compact Muon Solenoid

### 5.1 The Large Hadron Collider

Electron-positron colliders are intrinsically limited in energy by synchrotron radiation losses. As it has been emphasized in Section 3.1, it was not realistic to push the existing  $e^+e^-$  accelerator far beyond 100 GeV per beam, as the number of accelerating cavities was already large. In order to go higher in energy, and unveil new physics, another technology had to be considered.

A natural extension of  $e^+e^-$  colliders is muons colliders. It is however unrealistic at present to produce, accelerate and collide muons with a decent luminosity, as they are difficult to produce and decay shortly afterward. Developments are ongoing (see for example [49]), but will not lead to actual machines before decades.

Another way to avoid synchrotron radiation is to consider linear accelerators. Since particles only travel once inside a linear accelerator, it is not easy to accelerate electrons to high enough energies over realistic distances. Recently, the Tesla technology [50] demonstrated good performances, and one expects to have in the future a linear collider using this technology. This will nevertheless not occur before several years.

Another approach that has been considered is to use hadrons colliders. Protons (and anti-protons) are heavy charged particles, well suited to accelerators and colliders. History shows how successful a  $p\bar{p}$  collider can be. But producing anti-protons is not an easy task and limits luminosity. As rare processes are targeted by physicists, luminosity is a key feature, and the European physicists' community decided to build a proton-proton collider to be placed at CERN. The use protons and not anti-protons has also a more limited impact at high energy because gluons start to be more relevant, and because anti-quarks can be picked up from the sea of  $q\bar{q}$  pairs inside the proton. One drawback of a hadron machine is that it is far less cleaner than an  $e^+e^-$  collider, because of the hadronization of partons from the proton that do not

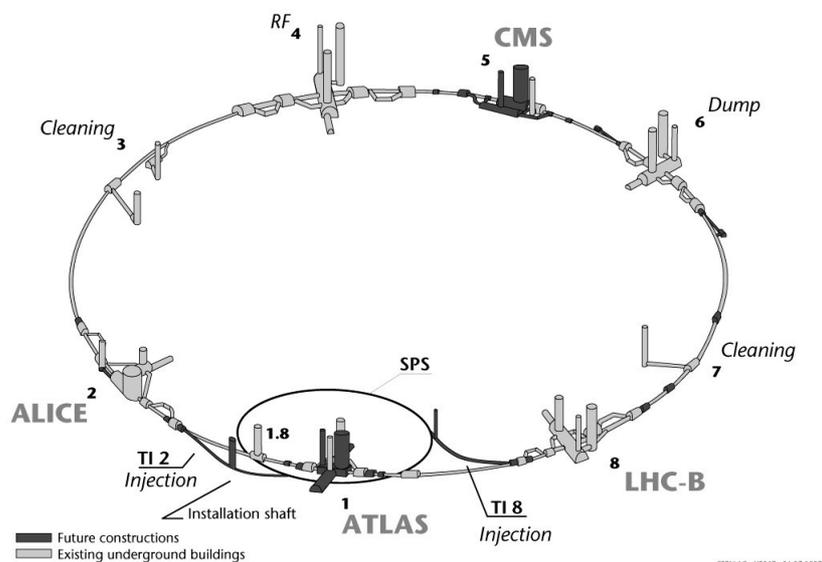


Figure 5.1: The LHC ring at CERN. LHC is being built in the LEP tunnel, which has to be adapted. New caverns and tunnels (red) are being built around existing infrastructures (gray).

participate to the main process.

The Large Hadron Collider (LHC) project was approved by the CERN Council in December 1994. It is now in advanced construction phase, and will be a proton-proton collider operating in 2007 at a nominal energy of 14 TeV in the center-of-mass<sup>1</sup>. The LHC is being built in the “old” LEP tunnel, and will therefore have a similar circumference (26659 m).

Five experiments will be located around the LHC beam pipe (Figure 5.1). Atlas[51] and CMS[52] are general multi-purposes detectors. Aims are to study/look for the Higgs sector, supersymmetry, Standard Model top sector, and more. LHCb[53] is dedicated to CP-violation studies, while ALICE[54] will cover quark-gluon-plasma in heavy-ion collisions. Finally, the smaller TOTEM[55] experiment will probably be installed in the CMS region (Point 5) and will then study small angle scattering, which could also allow for more precise luminosity measurements.

The advantage of a pp collider is evident in the light of the energy loss per turn (Equation (3.2)). In the LHC, each proton will lose 6.71 keV per turn, which has to be compared to the 2.4 GeV lost per turn by 103 GeV electrons at LEP2. This small quantity allows to reduce drastically the number of accelerating cavities with respect to what it was at LEP, down to four Cu/Nb cavities.

<sup>1</sup>For physics, the center-of-mass energy of partons is more important. It depends on the process considered and is of the order of 1 TeV.

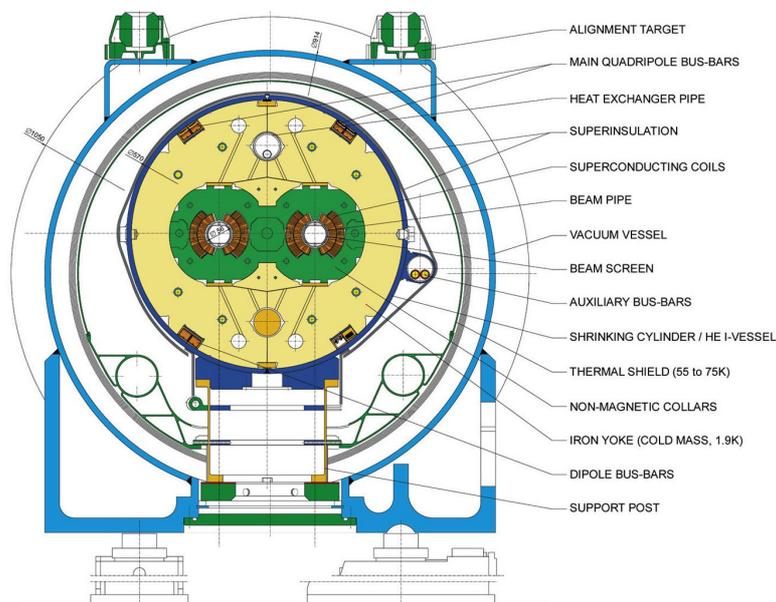


Figure 5.2: Cross-section of cryodipole. The two beam lines are visible in the center part, surrounded by the austenitic steel collars and by an iron yoke. Cryogenic and electric services are also represented. From [56].

Most of the LHC effort goes into bending dipole magnets. There will be 1232 14.3 meters-long dipoles in the LHC ring, each operating at a field strength of 8.33 T.

Another difficulty comes from the acceleration of same-sign particles in opposite directions. It is no longer possible to convey particles in a unique vacuum tube. The LHC is therefore a two-ring superconducting accelerator and collider. Only an approximately 130 m long common beam pipe is shared by the two beams near the interaction regions. There was not enough space in the LEP tunnel to install two separate rings of magnets. Therefore the LHC uses twin bore magnets, which consist of two sets of coils and beam channels sharing the same mechanical structure.

This unique design of magnets required an extensive R&D program, summarized with other LHC developments in the LHC design report[56]. The position and nature of conductors and bus lines, the materials used and cooling technology for superconductors are some of the items studied and optimized under constraints both from physics, in terms of beam quality, and from engineering, in terms of costs and existing infrastructures (mainly the LEP tunnel).

The cryodipole is depicted in Figure 5.2. In order to obtain large magnetic fields in the magnetic elements, superfluid helium is used. By lowering the temperature down to 1.9 K, an extra 1.5 T can be gained over standard super-

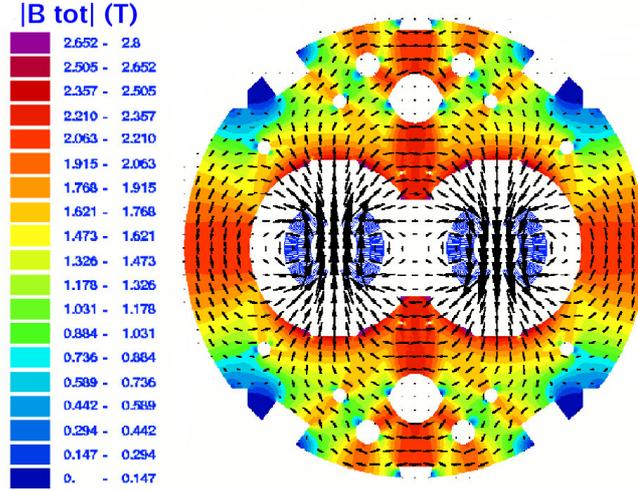


Figure 5.3: Dipole magnetic field plot. From [56].

conducting elements. This represents a 20% gain in beam energy. An 8.33 T magnetic field induces huge mechanical constraints on the surrounding material. In order to counter these forces, non-magnetic austenitic steel collars are used to maintain the conductors in place.

Electromagnetic design is another key feature of dipoles and quadrupoles. The dipole field is generated by a set of 40 cables distributed in six blocks around the beam screen. Cables in the inner layer have 28 strands, while cables in the outer layer have 36 strands. This design guarantees an homogeneous dipolar field and minimizes higher order multipolar contributions naturally arising in the two-in-one collar design. The magnetic-flux map obtained for the dipoles is presented in Figure 5.3.

Luminosity is the main challenge of the LHC. The LHC will reveal new physics if the designed luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$  can be sustained so as to accumulate  $100 \text{ fb}^{-1}$  per year. Such challenging running conditions will of course not be obtained from day 1, and a “low luminosity” regime, corresponding to a  $2 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  luminosity, is planned for the first few years.

The LHC luminosity is not constant over a physics run, but decays exponentially due to the degradation, mainly in intensity, of the circulating beams. The collisions themselves significantly limit the beam life-time, as the total cross-section is large. Indeed, the decay time of the bunch intensity due to this effect is

$$\tau_{\text{col}} = \frac{N_{\text{tot},0}}{\mathcal{L} \sigma_{\text{tot}} k}, \quad (5.1)$$

where  $N_{\text{tot},0}$  is the initial beam population,  $\mathcal{L}$  the luminosity,  $\sigma_{\text{tot}}$  the total cross section ( $\sigma_{\text{tot}} = 10^{-25} \text{ cm}^2$  at 14 TeV), and  $k$  the number of interaction points. For

nominal LHC conditions, this gives a decay time of 44.85 h. Time required to reach a  $1/e$  luminosity is thus given by  $t_{1/e} = (\sqrt{e} - 1)\tau$ , yielding  $\tau_{\text{col},1/e} = 29\text{h}$ .

Other sources of luminosity loss are intra-beam scattering (IBS) and beam-gas interactions. Since all these effects induce an exponential luminosity drop, the net luminosity lifetime can be approximated as

$$\frac{1}{\tau_{\mathcal{L}}} = \frac{1}{\tau_{\text{IBS}}} + \frac{2}{\tau_{\text{gas}}} + \frac{1}{\tau_{\text{col},1/e}}, \quad (5.2)$$

which gives  $\tau_{\mathcal{L}} = 14.9\text{ h}$ .

Estimating the integrated luminosity also requires hypotheses about the number of luminosity fills per year. From the PS and SPS cycling time, the number of SPS cycles needed to fill the LHC, and the time needed to ramp magnets up and down, experts predict a theoretical LHC turnaround time (i.e. the net machine cycle time, excluding the physical run) of 70 minutes. Experience from the HERA proton ring tells us that even with a good knowledge of the machine, the effective turnaround time could be six times higher. Two scenarios are hence considered. A minimal theoretical turnaround time of 1.2 hours (case 1) and a 6 times longer one of 7 hours (case 2) are therefore considered.

Integrating the luminosity over one run yields for 200 days of running per year,

$$L_{\text{tot}} = \frac{200 \times 24}{T_{\text{run}}[\text{h}] + T_{\text{turnaround}}[\text{h}]} \mathcal{L}_0 \tau_{\mathcal{L}} (1 - e^{-T_{\text{run}}/\tau_{\mathcal{L}}}), \quad (5.3)$$

where  $T_{\text{run}}$  is the total length of the luminosity run, period that can be used for physics, and  $\mathcal{L}_0$  is the initial luminosity. Equation (5.3) can be used to optimize the total length of the luminosity run, to 5.5 h and 12 h for cases 1 and 2 respectively. This gives a total luminosity per year between  $80\text{ fb}^{-1}$  and  $120\text{ fb}^{-1}$ , which meets the original objectives of the machine. It can also be noticed that the optimum  $T_{\text{run}}$  does not depend on the initial luminosity, so that the total luminosity per year at low luminosity will be five times lower. In practical terms, more problems are expected in the early phase of running, so that  $10\text{ fb}^{-1}$  is generally assumed for one year of running at low luminosity.

The high LHC luminosity also has important consequences on experimental conditions. Inelastic cross-section at LHC is  $0.6 \times 10^{-25}\text{ cm}^2$ , which implies an event rate of 600 MHz at nominal luminosity. With 2808 bunches in the ring, the mean bunch frequency is 31.6 MHz (the difference with the 40 MHz LHC clock arises from aborted bunches). This means there are 19.02 inelastic events per bunch crossing (3.8 at low luminosity). This superposition of events in the detector is commonly called “pile-up”, and has to be considered in each analysis of LHC data.

To this pile-up, one must add effects from machine-induced backgrounds. Each beam in a hadron collider is accompanied by stray particles outside the beam pipe. These particles, when entering experimental caverns, will cause additional background in the detectors. Conversely, particles scattered from

the interaction point to the LHC tunnel could deposit some 200W into the cold mass of the accelerator and induce unwanted quenches in the first dipole. To reduce both effects, absorbers will be installed, combining copper, iron and boron-loaded concrete.

## 5.2 The CMS detector

As mentioned in the previous section, the Compact Muon Solenoid (CMS) is one of the two general purpose experiments at LHC. It has a conventional cylindrical onion-like structure. The detector (Figure 5.4) has an overall length of 24 m and a diameter of 14.6 m. The total weight is 14500 tons. It is made of a barrel and two endcaps. Typically, the barrel covers a pseudorapidity<sup>2</sup> range  $|\eta| < 1.5$  while endcaps cover the range  $1.1 < |\eta| < 3.0$ . Precise limits depends on the subdetector considered. Like the ALEPH detector, CMS has a solenoid that generates in this case a four Teslas magnetic field.

The CMS design has been optimized for the LHC environment and physics goals. In the hostile environment that will be generated by the LHC, accurate reconstruction of leptons (mainly muons and electrons) is very important. Muons are particles with great penetrating power, which make them particularly suitable for observation (single muons were already signatures for the W discovery by the UA1 experiment [57], together with single electrons). Good performances in muon reconstruction are a challenging task that has driven the CMS design.

Since particles have a higher energy at LHC than at LEP, the CMS detector is globally larger than the LEP detectors, and presents a stronger magnetic field. This is also justified by the precise particle-momentum measurement aimed by the collaboration. It can be shown that the relative error on the particle  $P_t$  measurement by a tracker goes like

$$\frac{\Delta P_t}{P_t} \sim \frac{1}{BR^2}, \quad (5.4)$$

where B is the magnetic field in a direction orthogonal to the particle momentum, and R is the tracker radius. Resolution (especially for muons) can therefore be improved either by building a larger detector (this is the ATLAS choice), or by opting for a high magnetic field. This latter choice has been made by CMS. Tracking performances also relies on precise position measurements, mainly at each end of the tracks [58].

Near the beam pipe, a silicon pixel vertex detector allows precise impact parameter measurement. This is important as the identification of B mesons is critical, notably for top and Higgs physics. Next are the silicon tracker, that measures tracks from charged particles, and the electromagnetic and hadronic

<sup>2</sup>The pseudorapidity  $\eta$  is defined as  $\eta = -\ln(\tan(\theta/2))$ . It is an approximation of the rapidity, defined as  $y = \frac{1}{2} \ln\left(\frac{E+P_z}{E-P_z}\right)$ .

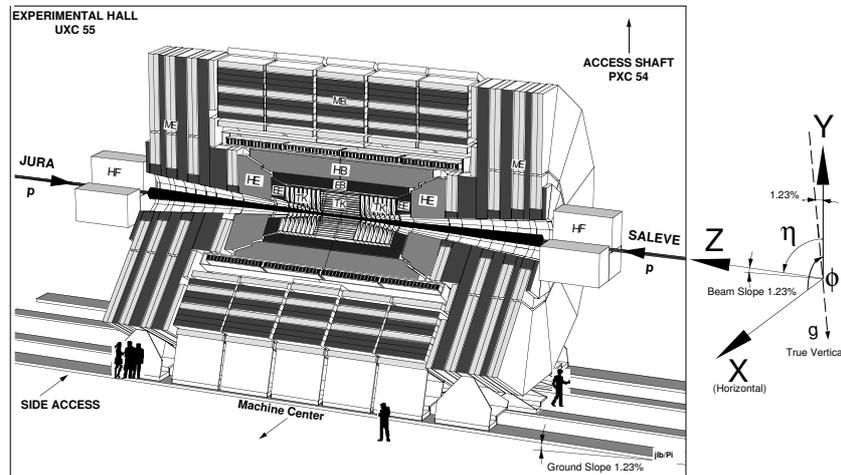


Figure 5.4: The CMS detector. The CMS reference system is shown on the right. From [59].

calorimeters. The muon system is the only set of detectors sitting outside the solenoid. Placed in the return yoke, these are various tracking devices dedicated to the measurement of muons. Because of this configuration, muon tracks have a typical bi-curved shape.

Very forward calorimeters extend the pseudorapidity coverage down to about five. Other calorimeters (e.g. CASTOR) extend down to  $\eta = 8.5$ , while the Roman pots from the TOTEM collaboration will cover part of the pseudorapidity range from 10 to 13. Very forward calorimeters will not be considered here as they do not contribute explicitly to the analysis presented in Chapter 7.

Details about the CMS detector can be found in [52]. In the following paragraphs, only a brief description of each subdetector will be given, according to the needs of the description of the reconstruction algorithms used in Chapters 6 and 7.

### 5.2.1 The tracking system

The CMS tracking system is an important piece of hardware that is concerned with the reconstruction in several ways. The tracker precisely measures coordinates of charged particles along their path in the central part of the detector. Hence, information from the tracker can be used to determine momentum of all charged particles. For electrons and muons, combining this information with measurements obtained either in the electromagnetic calorimeter or in the muon system provides a unique identification.

The tracker plays a key role in two other applications. Secondary vertices

can be identified by measuring the impact parameter of tracks. This information can be used to tag decays of B mesons produced in the hadronization of bottom quarks. High multiplicity events can also be studied using information from the tracker. Isolation and multiplicity estimators are very important in data analyses.

The CMS tracker is made of two subsystems [60]. At the smallest radii, the interaction region is surrounded by two or three layers of *silicon pixel detectors*. The *cell* size is  $100\ \mu\text{m}$  by  $150\ \mu\text{m}$ , which gives a resolution of  $15\ \mu\text{m}$  using analog readout. Moreover, the efficiency to find three pixel hits on a track (in the three-layer geometry) is above 90% for  $|\eta| < 2.2$ . At higher radii, ten layers of *silicon strip sensors* are installed. The outer radius of the CMS tracker extends up to 110 cm. The total length is approximately 540 cm. The tracker is further organized into a central barrel and two endcaps. In the barrel, sensors form *rods*, while endcaps are made of *petals* (Figure 5.5).

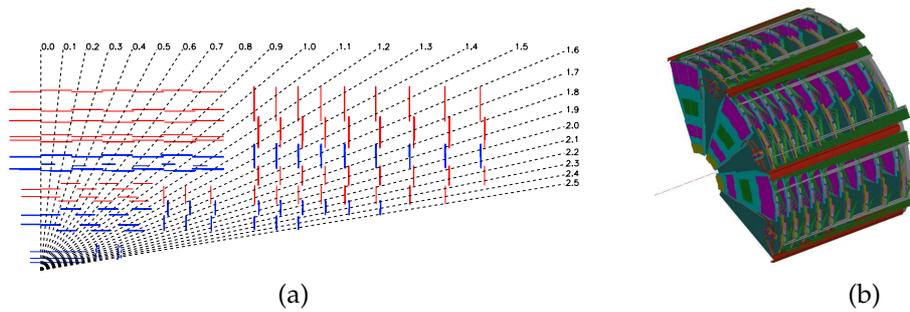


Figure 5.5: CMS silicon strip tracker layout. (a) Schematic view of quadrant of a tracker. The inner barrel (TIB) has four layers, the outer barrel (TOB) has six layers. The inner (TID) end cap is made of three small disks on each side of the inner barrel. The outer end-cap (TEC) is made of nine big disks on both sides of the tracker. The pixel detectors are also depicted. (b) Part of an endcap assembly, with several petals. From [61].

The silicon strip detector is made of *sensors* built as  $p^+$ -type strips implanted in an  $n$ -type bulk. Strips are capacitively coupled to analog readout chips. The charge collected from the strips is sampled at 40 MHz and stored in analog pipelines. When needed, pulse height data are transferred to the counting room via optical fibers, where digitization takes place. Analog readout improves the spatial resolution as information from several strips can be combined.

The pitch was determined from the targeted position resolution and varies from  $180\ \mu\text{m}$  in the inner part to  $45\ \mu\text{m}$  in the outer part. Radiation hardness requirements put constraints on the sensors' thickness, which was fixed at  $320\ \mu\text{m}$  in the tracker inner barrel (TIB) and at  $500\ \mu\text{m}$  in the tracker outer barrel (TOB). In the tracker endcaps (TEC), sensors are of both types, depending of

the radius. The expected resolution using simulated isolated muons is:

$$\frac{\Delta P_t}{P_t^2} = 0.2510^{-3} (\text{GeV}/c)^{-1}, \quad (5.5)$$

for 30 GeV/c muons, which is a factor two better than the ALEPH performances (Equation (3.9)). The CMS detector is furthermore able to reconstruct particles with a similar precision for particles with much higher energy. Studies have shown that performances even reach  $\Delta P_t/P_t^2 = 10^{-4} (\text{GeV}/c)^{-1}$  for 1 TeV/c muons.

To protect the silicon detectors from aging due to the high radiation flux, which causes an increase in leakage current, the full silicon tracker will be operated at  $-10^\circ\text{C}$ . A thermal shield placed outside of the tracker volume provides insulation while a cooling system extracts 60 kW of heat dissipated by the the front-end electronics.

### 5.2.2 The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) will play a significant role in exploiting the physics potential of hadron collisions. Excellent resolution on photon energy is one of the design goals of CMS, as decays of Higgs bosons may generate very clean signatures involving photons or electrons.

If the Higgs boson is light ( $115 \text{ GeV}/c^2 < m_H < 140 \text{ GeV}/c^2$ ), the golden discovery channel consists of the decay into a photon pair. There is nevertheless a huge background, and the search for a mass peak above that background will be needed. An excellent resolution is therefore required so as not to smear such a peak. If the Higgs boson was slightly higher in mass, decays into four leptons (via a Z pair) would become significant. Energy resolution will drive the discovery potential in this channel too.

A crystal-based scintillating calorimeter (as opposed to sampling calorimeters) offers the best performances for energy from electrons or photons, since most of the energy is deposited inside the active volume. The crystal has to have a high density and a small Molière radius. The CMS collaboration has chosen lead tungstate ( $PbWO_4$ ) crystals. Lead tungstate crystals have a Molière radius of 2.19 cm, and a density of  $8.28 \text{ g}/\text{cm}^3$ , which meets the constraints [59]. The radiation length (0.89 cm) is therefore significantly lower than for previously built crystal calorimeters<sup>3</sup>. Moreover, it is a fast scintillator compatible with the high event rate at LHC, it is very transparent and can be produced industrially.

The electromagnetic calorimeter (Figure 5.6) is built from 82728 individual crystals, mechanically organized into modules and supermodules. Each crystal has a transverse size compatible with the Molière radius, so as to have an optimal spatial resolution. Deposited energy is proportional to the light collected at the end of the crystal. Optical properties of crystals are therefore

<sup>3</sup>For comparison, L3 CsI crystals' radiation length was 1.85 cm.

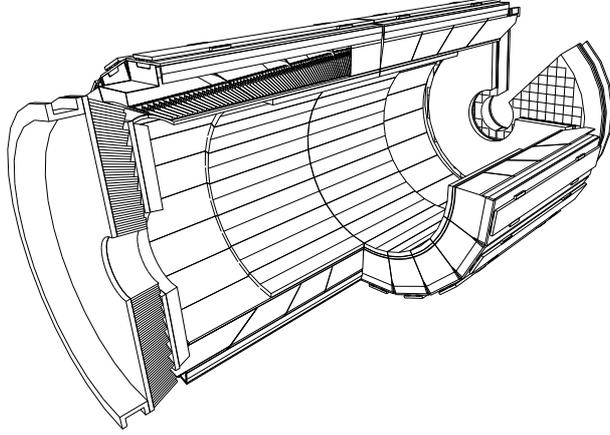


Figure 5.6: Schematic view of the electromagnetic calorimeter (ECAL). Crystals are partly visible, as well as their organization into modules and supermodules. Preshower modules in front of each endcaps are also shown. From [59].

important parameters to control. A laser system will be installed to monitor transparency and correct measurements accordingly. It has been shown that irradiation does not affect the scintillation mechanism, so that energy resolution will not decrease with time.

In front of most of the endcaps ( $|\eta| > 1.653$ ) there is a three-radiation-length preshower, consisting of a lead radiator and two parallel planes of silicon strip detectors. Strips from one plane are orthogonal to those of the second plane, which gives a two-dimensional position measurement with a precision of  $300\mu\text{m}$ . The main function of the endcap preshower is to make the distinction between photons and neutral pions, since pions will decay into two closely-spaced photons that cannot be resolved by the ECAL in that rapidity region.

In the range from 25 GeV to 500 GeV, the energy resolution can be expressed as

$$\left(\frac{\Delta E}{E}\right)^2 = (a/\sqrt{E})^2 + (\sigma_n/E)^2 + c^2, \quad (5.6)$$

where the energy is expressed in GeV,  $(a/\sqrt{E})^2$  is the stochastic term,  $(\sigma_n/E)^2$  is the noise, and  $c^2$  is a constant term. The stochastic term includes contributions from photostatistics and fluctuations in the shower containment. The noise term incorporates contributions from the electronics readout and pile-up, and therefore depends on the luminosity. Test-beam studies have shown that  $a = 2.7\%$  ( $a = 5.7\%$ ) in the barrel (endcap). The constant term accounts for effects like the non-uniformity in the light collection or intercalibration errors, and is found to be  $c = 0.55\%$ . Finally, the noise  $\sigma_n$  factor is approximately

equal to 200 MeV. Performances are four times better than the performances of the ALEPH electromagnetic calorimeter for a 50 GeV energy deposit.

The mass resolution for a 100 GeV/ $c^2$  Higgs boson decaying into two photons is found to be 650 MeV at low luminosity and 690 MeV at high luminosity, which is still high with respect to the width of a Higgs boson of this mass (1 MeV, as seen in Figure 2.2). This resolution combines effects from the energy and position resolution in ECAL, and uncertainties in the determination of the interaction vertex, using tracks from the proton remnants.

### 5.2.3 The hadronic calorimeter

The hadronic calorimeter (HCAL) is designed to measure the direction and energy of hadrons produced either as remnants of the proton, during the hadronization of quarks and gluons from the hard process, or from tau decays [62]. By combining all calorimetric activity, a missing transverse energy flow ( $\cancel{E}_T$ ) is also determined<sup>4</sup>. This is one of the only variables sensitive to neutrinos and other weakly interacting particles that could be encountered in the search for new physics. Missing transverse energy measurements are one of the motivations for the large pseudorapidity coverage ( $|\eta| < 5.3$ ).

HCAL is built as a barrel and two endcaps. The barrel is 9 m long and covers a pseudorapidity region  $|\eta| < 1.48$ . Endcaps are 1.8 m thick, and cover the pseudorapidity range  $1.48 < |\eta| < 3$ . The thickness corresponds to at least ten nuclear interaction lengths, which is enough to contain hadronic showers. A set of (very) forward calorimeters completes the acceptance of the detector.

The hadron calorimeter barrel is a sampling calorimeter with copper absorber plates interlaced with plastic scintillators. Copper was chosen for its relatively low  $Z$ , which minimizes multiple scattering for muons. It is also a non-magnetic metal, which is important since the calorimeter is placed within the 4 T solenoid. Light from scintillating tiles is collected by wavelength-shifter optical fibers. Test beam results yield a resolution in the range from 30 GeV to 1 TeV given by

$$\frac{\Delta E}{E} = \frac{1.22}{\sqrt{E(\text{GeV})}} + 0.05. \quad (5.7)$$

One may notice that this is slightly worse than the ALEPH HCAL performances. It is nevertheless compensated by the higher mean energy considered.

### 5.2.4 The magnet system

Equation (5.4) shows that a strong magnetic field allows a compact design compatible with the CMS goals. The 4 T magnetic field is generated by a

<sup>4</sup>Since the total longitudinal momentum is not constrained in hadron collisions, only transverse quantities (e.g. transverse momentum or transverse mass) are generally considered.

superconducting 13-meter-long solenoid[63]. Tracking and calorimetric sub-systems are completely contained in the solenoid.

The coil must also support tracker and calorimeters. Care has been taken to ensure very good mechanical properties and to resist to stress induced by the magnetic field. The same is true for the return yoke, which directly supports muon chambers, and indirectly all subdetectors. This massive steel structure closes the magnetic loop. From the outer radius of the solenoid to the outer radius of CMS, a 2 T magnetic field therefore bends muons trajectories in the reverse direction.

At the nominal current of 20 kA, the coil will store 2.7 GJ in the magnetic field. Ramping to this current will take five hours, while discharge will last up to 18 hours. In case of emergency, this energy has to be evacuated quickly or both the coil and the detectors could be damaged. A fast discharge in 50 m $\Omega$  resistors can be used; the current decay time is then 280 s.

### 5.2.5 The muon system

As discussed earlier, the muon system shaped the whole detector [64]. The return yoke represents 16 interaction lengths before the last muon station. This is enough to clearly identify muons, but the muon system also has to measure precisely the position of pass-through muons (Figure 5.7) and must deliver a quick response that can be used by the trigger, already in the first decision level. Detectors must sustain the 2 T magnetic field in the return yoke. These constraints led to a redundant design that combines three technologies (Figure 5.8).

In the barrel, *drift tubes* constitute the main muon system device. Drift tubes are suited to the barrel as the magnetic field is mainly contained by the return yoke. The relatively low particle rate in the central region also makes this technology very interesting. Drift tubes are aluminum tubes with a stainless steel wire in the middle. Drift tubes are assembled in *drift chambers* containing 12 layers of tubes. There are four layers of chambers, interleaved in the iron yoke plates. The twelve planes of drift tubes present in every chamber are organized in three independent subunits made up of four planes with parallel wires. Two subunits measure the coordinate in the bending plane, the third measures the track coordinate along the beam. Measurements in a chamber are combined to form an oriented *segment* used later on for track reconstruction.

In both endcaps, drift tubes are replaced by *cathode strip chambers*. Cathode strip chambers are multiwire proportional chambers where the cathode is subdivided into strips perpendicular to the anode wires. The forward environment is indeed very different from the central one, since the high particle flux requires a finer granularity and a faster response. Moreover, cathode strip chambers can sustain the highly varying magnetic field present in this transition region between the solenoid and the return yoke.

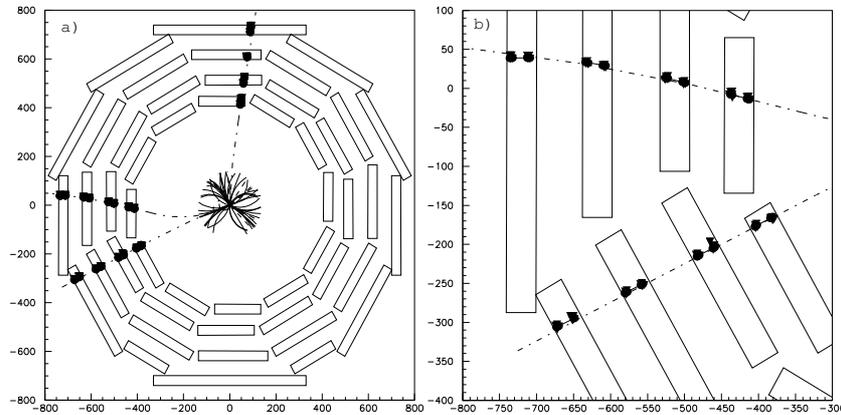


Figure 5.7: a. Side view of a  $t\bar{t}$  event generated by Pythia. Dashed-dotted lines show extrapolated tracks from the muon system to the inner tracker. b. Enlarged view of the same event; hits in the chamber super-layers (dots) are associated to form a track candidate, materialized by dotted lines. Full line segments show track elements in the barrel muon chambers. From [64].

Additional *resistive plate chambers* (RPC) are used in both the barrel and the endcaps. They are slightly faster, more robust, but have a less good spatial resolution. A RPC consists of two parallel resin plates, with a high bulk resistivity, separated by a gas-filled gap of a few millimeters. On the outside, electrodes are made of conductive graphite paint. Avalanches in the gas induce a fast charge on the cathodes, that can be exploited without costly electronics. RPCs complement drift tubes and CSCs, acting as additional sensitive planes in higher trigger levels and offline reconstruction (the CMS trigger system will be described in Chapter 6).

The efficiency for the reconstruction of muons tracks is above 90% for 100 GeV muons, which ensures a very good reconstruction of some important physical signals featuring muons. The best  $p_t$  resolution is obtained by combining information from the tracker to information from the various muon detectors. For 10 GeV muons, resolution ranges from 1% to 1.5%, depending on the pseudorapidity. For 1 TeV muons, where curvature is far more difficult to measure, it ranges from 6% to 17%. It is noticeable that such a precise measurement is already obtained at Level-3 trigger level.

### 5.2.6 The Data Acquisition system

As discussed in Section 5.1, each beam crossing at LHC will result in about 20 inelastic pp events. This corresponds to about 1 MB of zero-suppressed data every 25 ns, far too much to be handled and stored with present-day technol-



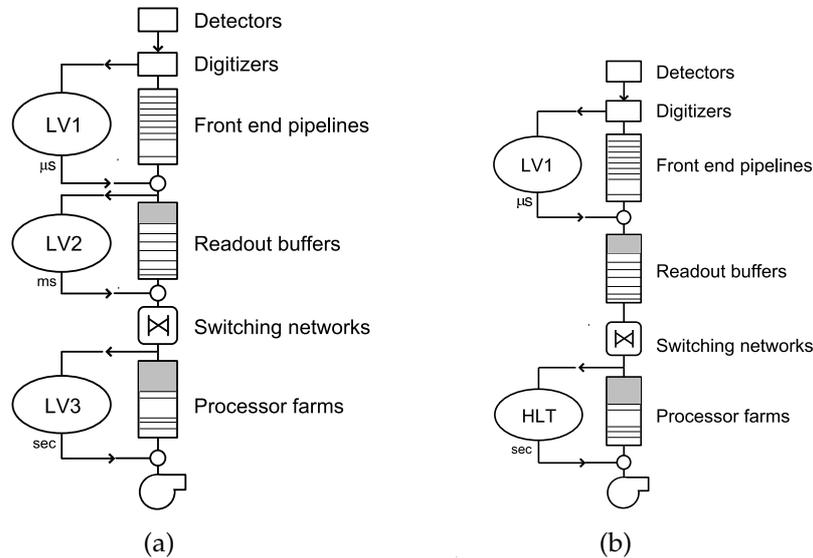


Figure 5.9: Data flow in a trigger and data acquisition system; (a) “historical” choice of many past and present experiments; (b) the CMS choice: a unique High-Level trigger. From [61].

on both the builder network, that carries data from the detector front-ends to the filter farm for the HLT processing, and the filter farm itself. The bet is that network speed and CPU power will scale with time as witnessed over the last twenty years, since the present technology does not allow to match the needs within the financial budget.

Another important aspect is modularity. The innovative CMS TriDAS design makes it possible to benefit from new technologies developing with time and from evolving hardware standards, which improves adaptability of the system to changes in accelerator and detector parameters or goals. One also wants to be able to make the system evolve. It is expected that initial luminosity will be lower, so that the full bandwidth and computing power will not be needed. At low luminosity, it is expected that only half of the full bandwidth will be available. Delaying the purchase of part of the system will result in substantial savings. All these considerations led to the three-dimensional design presented Figure 5.10. The daq system is organized in eight slices of 12.5 kHz. Each slice is totally independent from the others, can be installed independently, and can even cope with different technologies.

Also shown in Figure 5.10 are the various DAQ components. At the top, are the detector front-ends, where data is stored in pipelines scattered on the detector or close by. Front-end drivers are about 700 modules that read and forward data from the detectors upon the reception of a Level-1 accept. Part of them are specific to subdetectors, while a common interface to readout sys-

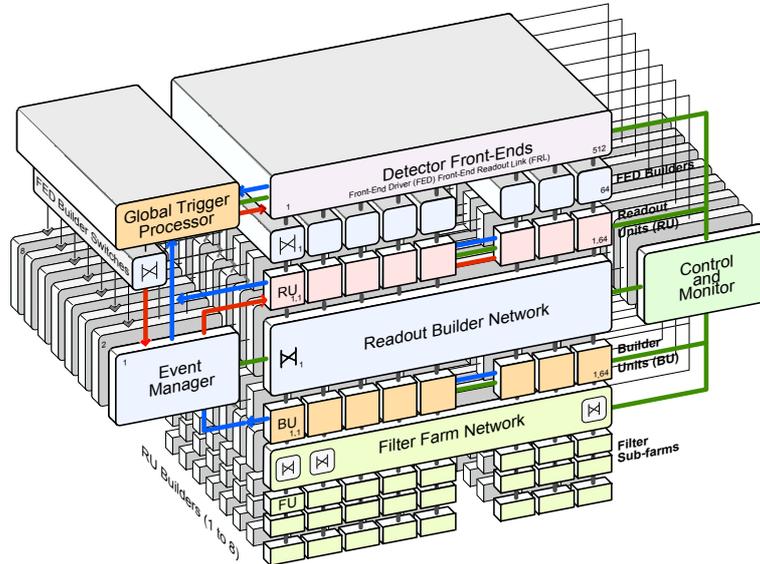


Figure 5.10: Three-dimensional view of the TriDAS system. Data from the detector front-end (top) are sent to readout-units (RU) independently for each event on receipt of an L1 accept. The readout Builder Network ships event fragments to the Builder Units (BU) where events are assembled. Events are then sent to the Filter Units (FU) to be processed by the HLT. The design is made of eight independent slices. From [61].

tems is imposed. There are then 512 readout units (RU), that are responsible for buffering event data after a Level-1 accept. Each readout unit collects information from several front-end drivers related to a certain type of subdetector, in a given portion of the detector. At this point, data from one event is distributed in 64 RU, within a single DAQ slice. This step decouples global event building, performed in “pull mode”, from detector readout, performed in “push mode”. Event fragments are read from the readout units and transferred through the builder network to builder units (BU), where the final event building is performed. The readout builder network is a large  $512 \times 512$  switching fabric that can sustain 100 GB/s. This nominal bandwidth limits the Level-1 output rate to 100 kHz. Events are then pulled to the filter farm, where the High-Level trigger selection is performed. Details about the High-Level trigger are discussed in Chapter 6.

Even if it was designed as a nearly no-dead-time system, this assembly is subject to small dead-time sources. The main loss of efficiency comes from technical constraints at the read-out level, which induces a  $\mathcal{O}(1\%)$  dead-time. For example, two subsequent L1 accept cannot be associated to subsequent bunch-crossings. There are second-order dead-time sources in the event building and HLT processing that are expected to be small.

## The High-Level Trigger

As explained at the end of chapter 5, the CMS data acquisition system has to reduce the event rate from about 30 MHz down to 100 Hz. This is achieved in two stages, a hardware Level-1 Trigger using rough information from the detector and a software High-Level Trigger, which uses the full granularity from all subdetectors. Time available to make a decision is driven by the buffer size on various front-ends electronics (at Level-1) and by the number of CPUs available in the filter farm (for the HLT). It was decided to limit the Level-1 decision time to  $\mathcal{O}(1 \mu\text{s})$ . This is why the Level-1 is designed on Field Programmable Gate Arrays (FPGA) and uses part of the information from the detector. Time allowed to select or reject one event at High-Level Trigger is virtually distributed amongst  $\mathcal{O}(1000)$  computing nodes. Since the L1 output rate is 100 kHz, each CPU has to handle one event in  $\mathcal{O}(10 \text{ ms})$ .

The pp cross-section is dominated by inelastic scattering. This process has a cross-section of 60 mb and generates rather soft jets. On the contrary, interesting processes involving gauge bosons or new physics are characterized by a somewhat low cross-section but more energetic, transversally produced particles. Interesting examples are the  $100 \text{ GeV}/c^2$  Standard-Model Higgs boson production (about 10 pb), the single  $W$  production (about 20 nb in the Standard Model), or the  $1 \text{ TeV}/c^2$   $Z'$  production (10 pb in some models). The cross-sections for these processes are shown in Figure 6.1 as a function of the center-of-mass energy. In the same figure, one sees that supersymmetric processes do have a typical cross-section of the order of 1 pb.

It is helpful to see with a couple of examples the usefulness of the HLT in selecting key signals. The most important signal involving photons is the decay of a light ( $< 125 \text{ GeV}/c^2$ ) Higgs boson into a pair of photons. The benchmark signal, corresponding to a  $115 \text{ GeV}/c^2$  Higgs boson decaying into photons is selected with an efficiency of 77% ( $L1 \wedge HLT$ ) if first and second photons are respectively above 40 GeV and 25 GeV. Using these cuts at HLT will induce a total rate from other processes of 5 Hz. Another striking example

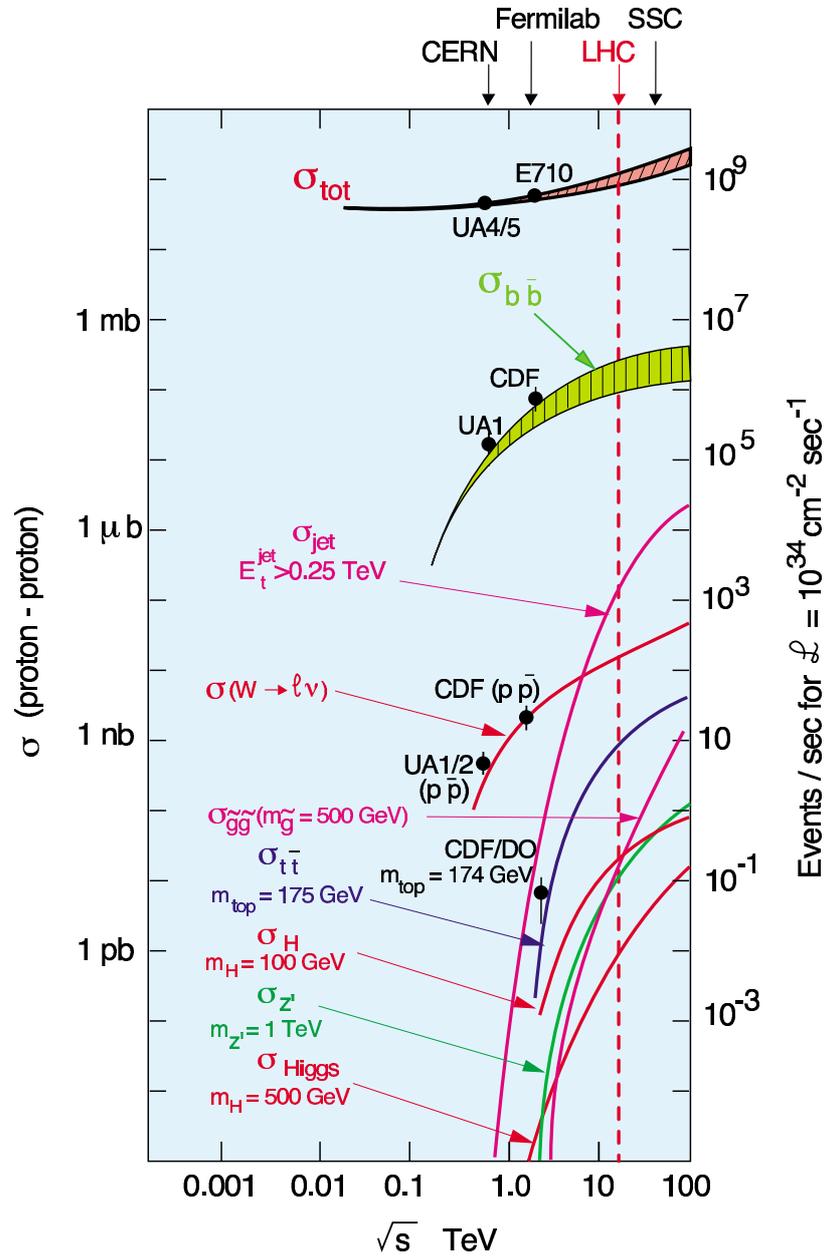


Figure 6.1: Cross-section of some known or predicted processes at LHC as a function of the center-of-mass energy. The corresponding event rate for a luminosity of  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  is also given.

is the selection of supersymmetric signatures with large missing momentum and energetic jets arising in cascade decays of charginos and neutralinos. Performances depend on the choice of parameters for the supersymmetric model, but a typical efficiency of 30% is obtained by requesting a jet above 180 GeV as well as at least 123 GeV of missing transverse energy. Using these cuts at HLT will induce a total rate from QCD processes of 5 Hz [61].

The common strategy regarding all triggers will therefore consist in a cut on transverse momentum or transverse energy of particles reconstructed in the time allowed to the trigger. Other more specific selections can be based on the topological properties of a specific signal. Such exclusive triggers are not yet implemented but are being considered, for example for the selection of  $J/\psi$  events. Each trigger strategy will be exploited in parallel when processing events, and will determine the subsequent handling of events. The term *stream* is therefore often used for a subtrigger, or a set of related subtriggers.

## 6.1 Level-1 streams

The idea of regional event reconstruction is one of the core principles driving the HLT design. In many cases, the decision whether or not an event should be kept offline is taken on basis of an object identified at Level-1 and validated in High-Level processing from the information collected around that candidate. In that process, the Level-1 candidate defines a *region* whose extent depends on the type of candidate.

This approach drastically reduces the CPU time needed to make a decision, but an object that has been “missed” (a lepton, jet or missing transverse energy) at Level-1 cannot be recovered later on. Hence, Level-1 is a decisive prerequisite to the HLT. The 50 kHz (100 kHz) L1 bandwidth available at low (high) luminosity has been optimized in order to ensure the widest possible physics reach by dispatching the total bandwidth into four categories of objects : electromagnetic deposits, muons, tau-jets and global objects based on calorimeters’ information (jets, missing transverse energy ( $E_T$ ), total energy ( $E_{tot}$ ), ...). It must be noted that there is no way of differentiating electrons from photons without tracker information, which is absent at this level. Thresholds on  $P_t$  for single, double or multi-object triggers were determined so as to match the allocated bandwidth. The resulting Level-1 trigger menu at low luminosity is shown in Table 6.1. The total event rate corresponds to the 50 kHz available at low luminosity, reduced by a factor of three that accounts for uncertainties in the simulation of physics processes and detectors. Level-1 trigger thresholds are expressed in Table 6.1 as the  $P_t$  or  $E_t$  value for which 95% of the asymptotic trigger efficiency is reached. This value is slightly higher than the actual cut applied, due to resolution effects.

The output of the L1 trigger is a list of trigger bits determining whether or not the corresponding subtrigger selected the event of interest. This is used by the HLT to select algorithms that have to be executed to perform the final

Table 6.1: Level-1 trigger table at low luminosity ( $2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$ ). Thresholds correspond to values with 95% of the asymptotic trigger efficiency. The upper part corresponds to what was used in the DAQ Technical Design Report (TDR) [61], while the second part consists in additional triggers used for the Physics TDR, for which no recent rate measurement is available. Grayed lines correspond to L1 triggers that are not effectively used due to a lack of subsequent HLT algorithm.

Trigger	Threshold (GeV or GeV/c <sup>2</sup> )	Rate (kHz)	Cumulative Rate (kHz)
Single electron/photon	29	3.3	3.3
Di-electrons/Di-photons	17	1.3	4.3
Single muon	14	2.7	7.0
Di-muons	3	0.9	7.9
Single tau-jet	86	2.2	10.1
Di-tau-jets	59	1.0	10.9
1,2, 3 or 4 jet(s)	177,130,86,70	3.0	12.5
Jet $\wedge$ $\cancel{E}_T$	88 * 46	2.3	14.3
Electron $\wedge$ Tau-Jet	14 * 52	0.8	15.1
Tri-tau-jets	40	-	-
Quadri-tau-jets	59	-	-
Electron $\wedge$ Jet	21 * 45	-	-
Muon $\wedge$ Electron	5 * 15	-	-
Muon $\wedge$ Tau-jet	5 * 25	-	-
Muon $\wedge$ jet	5 * 30	-	-
Muon $\wedge$ $\cancel{E}_T$	5 * 45	-	-
Electron $\wedge$ $\cancel{E}_T$	21 * 75	-	-
Tau-Jet $\wedge$ $\cancel{E}_T$	21 * 75	-	-
$E_T^{\text{sum}}$	600	-	-
$\cancel{E}_T$	140	-	-

selection (in the following, this will be denoted vetoing mechanism). The L1 trigger also delivers four candidates for each type of object. These candidates can be used in the HLT as seeds for advanced reconstruction.

## 6.2 Event reconstruction at HLT

### 6.2.1 Electrons and photons

Reconstruction of electrons and photons proceeds in three steps, respectively called level 2, level 2.5 and level 3. This internal division of the HLT processing is aimed at saving computing time, as the streams can be aborted as soon as one single step fails.

#### L2 clustering

At level 2, candidates are built from the calorimeter information using a clustering algorithm. Electron reconstruction is complicated by the presence of tracker material in front of the calorimeter. Electrons will radiate a significant fraction of their energy as bremsstrahlung as they cross the tracker layers. For 35 GeV/c electrons, this represents on average 43.6% of the initial energy. Due to the magnetic field, this energy is deposited in a  $\phi$  slice and has to be recovered by the clustering algorithm. Different algorithms are used in the barrel and endcap regions[59], characterized by a very different geometry.

In the barrel, the *hybrid* algorithm starts from the most energetic crystal above 1 GeV in the regions identified by the L1 electron trigger. Alignments of crystals (called dominoes) are aggregated in the  $\phi$  direction, as shown in Figure 6.2. These domino's are made of three or five crystals along  $\eta$ , depending on the energy deposited in the central crystal. When 10 alignments are obtained on both sides (in  $\phi$ ) of the seed, those with less than 0.1 GeV are discarded. The set of subclusters that is then obtained is filtered, requiring a central crystal with more than 0.35 GeV. Surviving subclusters are combined to form the electron candidate.

In the endcap, the *island* algorithm is used to form clusters, as it has been shown to give better performances in that region. The algorithm starts from the most energetic crystal in the L1 region and aggregates adjacent crystals as long as a crystal with no energy (after zero-suppression) or an energy higher than the one in the previous crystal is found. This creates subclusters, as depicted in Figure 6.3. The island algorithm takes the preshower into account by constructing subclusters in each plane of silicon, starting from the center of each ECAL subcluster. All subclusters are combined in a narrow  $\eta$  window to form the electron candidate.

The reconstruction performances in terms of resolution in the barrel and in the endcap is shown in Figure 6.4. This was determined using simulated single-electron events.

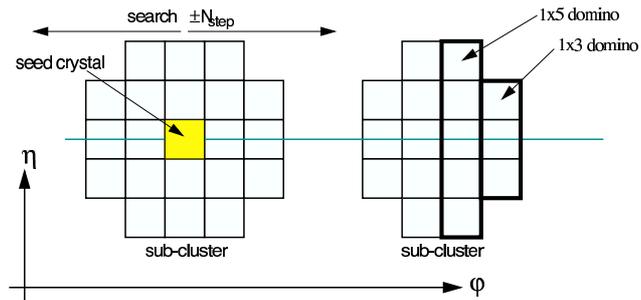


Figure 6.2: Schematic view of the Hybrid clustering algorithm. From [65].

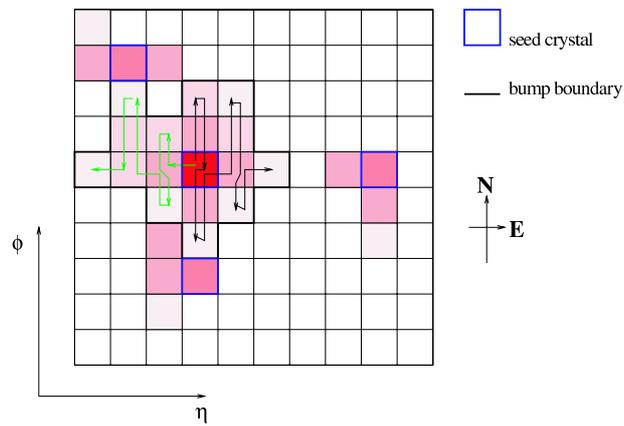


Figure 6.3: Schematic view of the Island clustering algorithm. From [65].

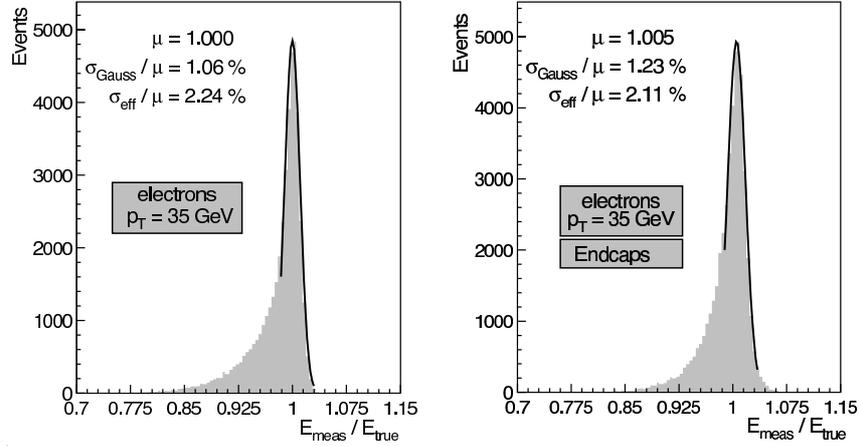


Figure 6.4: Distribution of the ratio of the reconstructed and simulated energy for 35 GeV electrons in single-electron events [61], for electrons in the endcap (right) and for electrons in the barrel (left).  $\sigma_{eff}$  corresponds to the RMS of the distribution, while  $\sigma_{Gauss}$  is the width of the fitted Gaussian.

### L2.5 matching with hits in the pixel detector

After this clustering in the calorimeter, electrons and photons are still not differentiated. Since a very small amount of material lies between the interaction point and the pixel detector (there is only a small contribution from the beam pipe), it is unlikely that photons will convert before reaching this detector. The presence of hits in the pixel detector can therefore be used to differentiate electrons and photons.

From the cluster position, regions in the pixel detector are obtained by back-propagating the would-be electron or positron, assuming the vertex at  $(0,0,0)$ , once for each charge hypothesis. Hits are looked at in the first pixel layer in a  $\eta - \phi$  region defined by  $(\Delta\phi = 40 \text{ mrad}, \Delta z = \infty)$ . If a hit is found, it is included in the electron fit, and a compatible hit is searched for in the second and third pixel layers. If no hit is found in the first layer, the same procedure is performed starting with the second layer, and searching for a compatible hit in the third pixel layer only. The L2 candidate will be marked as an electron if at least two compatible hits are found.

In the search for the first hit, electron and positron regions can overlap, but when a compatible hit is propagated from one layer to the other, the search areas only overlap for 1 TeV particles, and beyond.

### L3 photon selection and electron track reconstruction

The selection of photons at level 3 is straightforward. A stronger cut on the transverse momentum of the candidate is applied to reduce the rate, and isolation is imposed using segments reconstructed in the pixel detector.

The selection of electrons at level 3 is based on full reconstruction using information from the tracker detector. The electron track is reconstructed inside-out from the seed obtained at level 2.5, and extrapolated from the silicon tracker to the calorimeter. Information from the calorimeter and from the tracker are then compared, and a cut on the  $E/P$  ratio between the energy measured in the ECAL and the momentum from the tracker is applied ( $E/P < 1.5$ ). The distance between the track extrapolated at the surface of the calorimeter and the cluster position also has to be small (typically less than  $0.0025$  in  $\eta$ ).

In the endcaps, a cut on the energy found behind the cluster, in the HCAL, expressed as a fraction of the cluster energy ( $H/E < 0.028$ ), is found to give useful additional rejection.

## 6.2.2 Muons

### L2 muon reconstruction

Muons are selected in two steps. In a first step, at level 2, all the information from drift tubes, RPCs and CSCs is used. Tracks are built from segments and hits obtained in the muon detectors, using regions where muons were found at level 1. The trajectory is built inside-out, using first the momentum estimate from L1, and updating this estimate at each step. Each time, the predicted state (position and direction) at the next measurement surface is compared with actual measurements, and updated accordingly using a Kalman filter technique. This propagation of muons within the iron, where magnetic field effects and multiple scattering must be taken into account, is very time-consuming. Once the outermost layer is reached, the track fitting is performed from the outermost muon station inwards, including the nominal interaction region  $(0,0,0)$  with the uncertainty ( $\sigma_r = 15\mu\text{m}$  and  $\sigma_z = 5.3\text{cm}$ ). The resulting track parameters (mainly the track  $P_t$ ) are used to accept or reject the muon candidate for further L3 processing. During both steps, segments are rejected if their effect on the track  $\chi^2$  is larger than 25. If there are less than three compatible hits or segments, the muon candidate is rejected.

A calorimeter isolation cut based on the energy found in a  $0.2$  ( $\eta, \phi$ )-cone around the muon candidate is applied. This significantly reduces soft ( $< 30$  GeV) muons originating predominantly from the decay of b-mesons, c-mesons, K, and  $\pi$  decays. These muons are generally accompanied by other nearby particles and are thus suppressed by isolation cuts. This cut is important as these muons dominate the total muon rate.

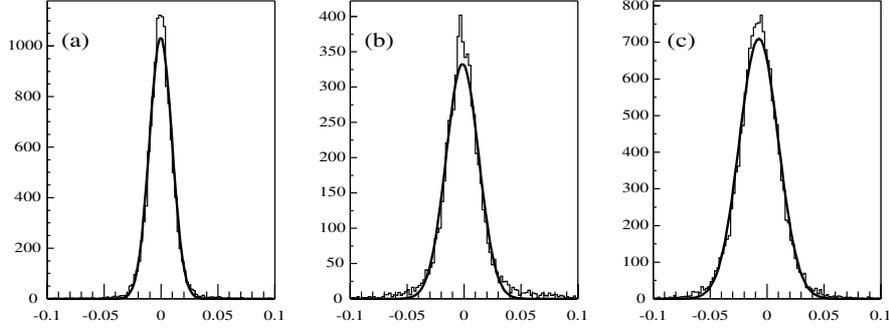


Figure 6.5: Distribution of  $(1/PT_{rec} - 1/PT_{gen})/(1/PT_{gen})$  where  $PT_{gen}$  and  $PT_{rec}$  are the generated and Level-3 reconstructed transverse momenta, respectively, shown in three pseudorapidity intervals: (a)  $|\eta| < 0.8$ , (b)  $0.8 < |\eta| < 1.2$ , and (c)  $1.2 < |\eta| < 2.1$ . Resolution is 0.010, 0.014 and 0.017, respectively in these three regions. From [61].

### L3 muon reconstruction

At level three, tracks are built in the tracker, in the region-of-interest defined by the L2 candidate. Tracker hits are combined to hits in the muon system to improve resolution (Figure 6.5). In the region of interest, initial candidates for the muon trajectory are built from pairs of reconstructed hits. The two hits forming a seed must come from two different tracker layers, and all combinations of compatible pixel and double-sided silicon strip layers are used in order to achieve high efficiency. A track reconstruction algorithm based on the Kalman filter technique then proceeds in three steps. Tracks are first built from the seeds (trajectory building). A cleaning procedure subsequently removes multiple tracks from a single seed, using the number of hits and the track  $\chi^2$  (trajectory cleaning). In the last step (trajectory smoothing), selected tracks are reconstructed once again, including hits in the muon system. The final track is selected on basis of the  $\chi^2$  of the corresponding fit. The muon is then selected by cuts on its transverse momentum, imposing at least five hits in the tracker, and applying a cut on the transverse position at the vertex ( $\Delta r < 0.02cm$ ).

A tracker-based isolation cut is applied that uses the total  $P_t$  of tracks in a  $(\eta, \phi)$ -cone around the muon. Since only tracks originating from the same vertex are used, this technique is less sensitive to the pile-up than the calorimeter isolation used at L2. Threshold on the summed  $P_t$  varies from 2 GeV/c to 3 GeV/c for typical cone size of 0.2.

### 6.2.3 Taus

Identification of isolated tau leptons is expected to play a key role in signatures of supersymmetry at large  $\tan\beta$ , as couplings to charged leptons are enhanced.

As the tau decays in 68% of the cases via 1-prong vertices (excluding muon decays) and in 15% of the cases via 3-prong vertices, the final-state signature is a narrow jet-like cluster associated with one or three tracks.

### L2 tau-jet reconstruction

The tau jet reconstruction starts by applying an iterative cone jet algorithm [66] in the regions of the four tau-jet candidates identified by the Level 1 trigger. For each Level-1 candidate, the iterative cone algorithm proceeds by summing all energy deposits around a seed, obtained from the most energetic calorimetric tower(s) in the region of interest. The weighted mean position resulting from this sum defines a new direction (jet axis), and the procedure is repeated until the jet axis stabilizes. For the double tau trigger, the first central jet region identified at Level 1 is also considered if no second tau has been found at Level 1.

Energy deposited in both the ECAL and the HCAL is taken into account. Once the jet is identified, energy deposited in crystals of a ring around the jet is summed and used as an isolation variable. Inner and outer radii of this ring are respectively 0.13 and 0.4 in the  $\eta - \phi$  plane, and the collected energy has to be below 5.6 GeV for the tau candidate to be accepted.

#### L2.5 tau isolation

At level 2.5, tau isolation is performed with information from the tracking devices. This can be either the pixel detectors, or the full silicon-strip tracker. Only tracks of more than 3 GeV/c and pixel lines from particles of more than 1 GeV/c are considered.

A *matching cone* is built around the direction of the L2 tau-jet. The most energetic track in the matching cone is used as leading track, and a *signal cone* is built around it. Either one or three tracks, including the leading track, must be found in the signal cone. An *isolation cone* is also built around the direction of the tau candidate. No other track must be found in that cone. This isolation scheme and all three cones are depicted in Figure 6.6.

## 6.2.4 Jets and missing transverse momentum

Jet reconstruction is a quick task that is achieved at level 2. It does not exploit regional reconstruction, but searches for jets in the whole calorimeter using a simple iterative cone algorithm using all calorimeter towers.

Due to the high jet rates at LHC, the thresholds for the event selection based only on jets are very high, but can be reduced to an acceptable level with additional trigger conditions, such as a lepton or missing transverse momentum ( $\cancel{p}_T$ ). Missing transverse momentum is the only way to identify neutrinos and other undetectable particles.  $\cancel{p}_T$  is computed as the vector sum of tower energies above 500 MeV in the whole calorimeter.

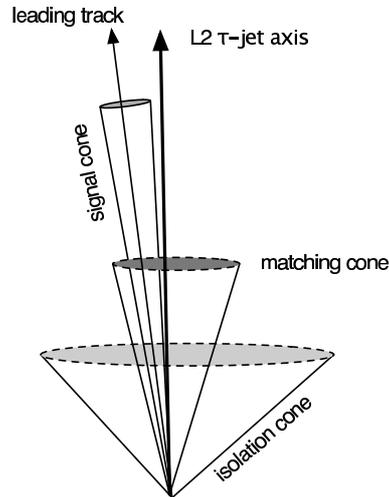


Figure 6.6: Schematic view of the cones used in the isolation of a tau-jet candidate. The cone apertures are:  $R_{match} = 0.1$ ,  $R_{signal} = 0.07$ ,  $R_{isol} = 0.35$ , respectively for the matching cone, the signal cone and the isolation cone. From [61].

Another way to reduce the jet rate is to concentrate on jets from B mesons. This can be done by exploiting the large impact parameter of tracks with respect to the production vertex in such a jet. If more than two tracks have an impact parameter above three standard deviations of the vertex resolution (computed from tracks from the primary vertex), the jet is tagged as a b-jet. Preliminary studies show that this reduces the jet rate by a factor 30 [67], and the jet threshold can be adjusted accordingly. More advanced algorithms have been developed to tag b-jets, but are too slow to be used at HLT.

### 6.3 HLT streams

The strategies considered for the High-Level Trigger (Table 6.2) reflect strategies at Level-1. Most of the bandwidth is allocated to single particle streams: inclusive electron and inclusive muon. These are the golden trigger strategies, followed by di-electron and di-muon triggers, photon streams or (multi-)jet strategies. The trigger menu also includes an exclusive trigger designed for the study of the  $t\bar{t}H$  channel. As it is now, this triggers is a L3 muon trigger, with tuned  $P_t$  and isolation cuts. Due to the vetoing mechanism that explicitly links the Level-1 trigger and the High-Level trigger, L1 trigger streams that are not complemented at HLT are not useful at all. This way,  $(\mu \wedge e)$ ,  $(\mu \wedge \text{Jet})$ ,  $(\mu \wedge \cancel{E}_T)$  and  $(e \wedge \text{Jet})$  L1 triggers streams are virtually disabled, as no HLT algorithm will be attempted if only these contributed to the L1 decision.

Table 6.2: High-Level trigger table at low luminosity ( $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ ). Muon thresholds correspond to values with 90% efficiency for muons, while other values stand for the actual cut applied. The single-electron stream includes 10 Hz of single-W events in addition to 23 Hz of QCD events. The second part of the table shows triggers that are new since the DAQ TDR, and for which no rate estimates are available.

Trigger	Threshold ( GeV or GeV/c)	Rate (Hz)	Cumulative Rate (Hz)
Single electron	26	33	33
Di-electrons	14.5		
Relaxed di-electrons	21.8	1	34
Single photon	80	4	38
Di-photons	20,35		
Relaxed di-photons	21.8,35	5	43
Single muon	19	25	68
Di-muons	7	4	72
Di-tau-jets	0,50	1	73
Jet $\wedge$ $\cancel{E}_T$	180 * 123	5	78
1, 3 or 4 jet(s)	572,195,80	9	86
Electron $\wedge$ Tau-Jet	16 * 52	2	88
1, 3 or 4 jet(s) $\wedge$ 1 b-tag	200,150,55	5	93
Muon $\wedge$ Tau-Jet	15 * 40	—	—
$\cancel{E}_T$ $\wedge$ Tau-Jet	67 * 20	—	—
Inclusive $t\bar{t}H$	15 (muon)	—	—

Ongoing developments will introduce new multi-object triggers and advanced topological strategies dedicated to specific channels. There are also many other uncertainties like the effect of new jet calibrations. Bandwidth distribution may also vary. It is probable that the bandwidth allocated to jet triggers will be enlarged, at least in the beginning, so as to allow the study of QCD multi-jet processes. The HLT steering code must therefore be flexible to accommodate changes in the trigger scheme.

## 6.4 The HLT steering code

Algorithms sketched in Section 6.2 were first presented in the DAQ Technical Design Report (TDR) [61]. Trigger strategies based on electrons, photons, muons, taus, jets and missing transverse momentum were implemented and discussed there. These algorithms must now be used by a large community of physicists in order to prepare the Physics TDR, which will demonstrate what physics we can do and how we will do it. There is a need for a simple steering code common to all algorithms involved in the event selection, allowing the use of High-Level Trigger selection as a preliminary step in the forthcoming activities. It is also necessary to provide a simple interface where the implementation of existing and future selection algorithms can be put. Integration of new triggers and study of their complementarity will only be possible if such an interface is in place.

### 6.4.1 Principle

A compilation of requirements for the HLT steering code can be found in [68]. We collected these requirements between January and June 2003, with input from people involved in software development, data acquisition and physics analysis. In addition to already mentioned prerequisites, it is worth noticing that the HLT definition must be dynamic. Trigger modifications must be allowed between runs as luminosity decreases, without the need to recompile the code or restart the program. To summarize, the HLT steering code should:

- provide a single placeholder for the implementation of HLT conditions (single electron, single muon, ...),
- allow a dynamic HLT definition, which can be redefined without the need to even restart the program,
- allow optimization of the HLT evaluation sequence,
- produce a global boolean HLT decision,
- deliver a full HLT response as a bit pattern (as it is the case for L1),
- provide off-line checking facilities, such as candidates at several stages of the selection procedure, intermediate decision status, etc.

- be based on L1 output and existing reconstruction code,
- minimize the computing overhead.

Based on these requirements, we have designed and implemented the CMS HLT steering code. It is common practice to consider the trigger as a logical “OR” of several trigger conditions. More generally, a trigger is a logical equation, and can therefore be represented as a binary tree. This allows for more flexibility and features commonly found in triggers of past experiments, like automatic prescale or combined subtriggers.

The design of a binary tree is such that navigation between trigger conditions becomes natural and effective, making it possible to use recursivity techniques for the evaluation and configuration of the HLT. The binary tree corresponding to the default trigger for low luminosity is shown in Appendix A.

## 6.4.2 Implementation

### Binary trigger tree

The *UML* diagram shown in Figure 6.7 shows how the various classes we defined in the context of this thesis are related, and how they integrate into ORCA, the CMS reconstruction software suite [69]. Here are some more explanations.

The HLT selection is constructed from elementary building blocks called *HLT Elements*. Each element is intended to be either a physical subtrigger (single electron, single muon, ...) or a logical node (AND, OR, ...). The selection code is implemented in the element by using the reconstruction code. The end-product of the element is a boolean value and a bit vector. The boolean value gives the outcome of the HLT element and the bit vector specifies with more details the decision of each trigger stream. A global trigger response bit pattern results from the sequential evaluation of a tree of such elements.

In order to form a binary tree<sup>1</sup>, an HLT element has two pointers to *parents nodes*. Since the information is pulled from the extremities of the tree, corresponding to the earliest step in the HLT evaluation, to the root, corresponding to the final decision step, there is no need for a reverse pointer. The High-Level Trigger elements are described in a set of classes deriving from `HighLevelTriggerElement`, each implementing a pure virtual method of the base class (in the Design-Patterns terminology [70], this are strategies of `HighLevelTriggerElement`). The `HighLevelTriggerElement` derives from both `RecAlgorithm` (the class from the CMS reconstruction infrastructure that represents a generic algorithm) and `LazyObserver` (the class from the CMS reconstruction infrastructure that represents any part of algorithm that has to be run on demand).

---

<sup>1</sup>A binary tree, is a structure composed of nodes, where each node is linked to exactly two other nodes.

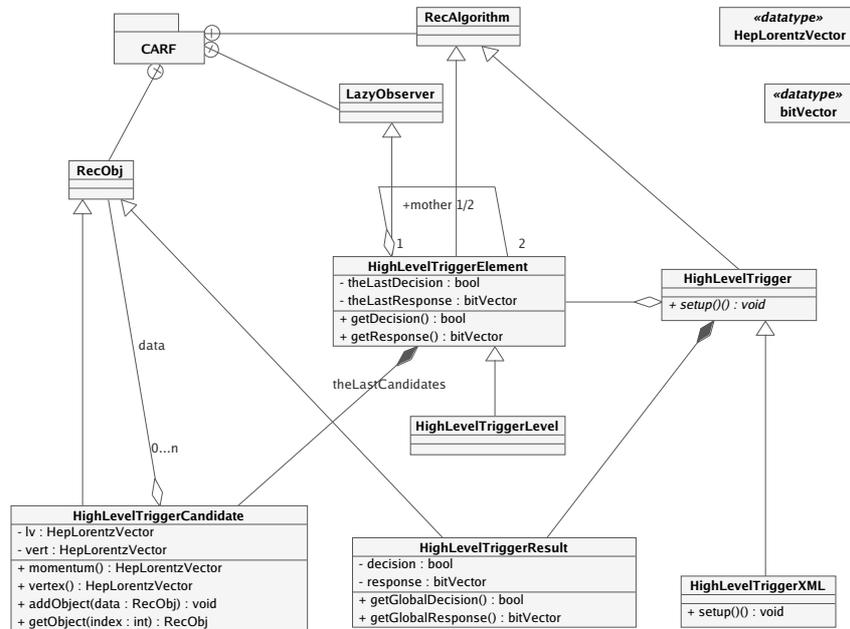


Figure 6.7: UML diagram of the HLT steering prototype. Only the main methods and attributes are shown. There are 23 classes deriving from `HighLevelTriggerElement` and `HighLevelTrigger` (The detailed list of available element is shown in Table A.1). The link to the CMS reconstruction framework is done via three classes from the CARF subsystem: `RecObj`, `LazyObserver` and `RecAlgorithm`.

`HighLevelTrigger` is the base class used for the logical tree construction. It is a `RecAlgorithm`, so that it integrates in the CMS reconstruction infrastructure (CARF). It has a handle on the logical tree root and acts as a facade in front of the HLT mechanics. The `HighLevelTrigger::setup()` method is abstract and has to be implemented to define the logical tree. It is in this method that each trigger element is instantiated and connected to form the tree which represents a specific HLT selection scheme. A concrete implementation that reads the tree definition from an XML file is discussed later on.

The method `HighLevelTrigger::getResult()` returns an instance of the `HighLevelTriggerResult` class, which is a `RecObj`, a reconstructed object that can be handled by CARF and made persistent. Persistency of the global trigger response is important as evaluating the HLT response is a time consuming task that we want to avoid repeating each time. This simple object contains not only the boolean decision that is the “green light” for an event to

be forwarded to off-line facilities but also the full trigger response as a vector of bits, one or more bit per trigger element. The bit-pattern finds its use in at least two applications. Firstly, it is a key feature for the debugging of the High-Level Trigger, and is already used in this sense in preliminary studies and analyses. Secondly, as recommended in the CMS Computing Model[71], the online (HLT) system will classify events in  $\mathcal{O}(10)$  online streams of similar size (within 10%). These streams should be completely determined by the HLT trigger bits. After full reconstruction in the offline infrastructure, events will be further classified in  $\mathcal{O}(50)$  offline streams. The discussion on whether or not these offline streams should also be based on HLT bits, or could use information from more advanced reconstruction, is ongoing.

Finally, instances of `HighLevelTriggerCandidate` are produced by each `HighLevelTriggerElement` during the evaluation of the local response. This class represents a generic reconstructed object, characterized by a Lorentz momentum, a production vertex, and a collection of objects from which it has been built. This gives summary information that can be used to study the HLT response regardless of the type of trigger involved.

### Dynamical trigger definition

The design of the HLT prototype enables to dynamically build the tree of elements at runtime. This feature is used in the `HighLevelTriggerXML` class.

`HighLevelTriggerXML` derives from `HighLevelTrigger` and implements the method `HighLevelTriggerXML::setup()` in a way that allows to build the HLT from the description found in an XML file<sup>2</sup>. When the HLT is built (at the beginning of the first event evaluation), the XML file is read and translated into a binary tree. Changing the file and executing the same program is thus possible. This allows for a fast analysis cycle since no compilation is needed (e.g. to replace the pixel tau validation with the tracker tau validation only requires to change the XML file).

There are lots of advantages using XML. The DTD can be used to restrict the allowed connections between trigger elements, and there are simple tools available to check the validity of XML files. In the online farm, it is expected that a remote process control mechanism will couple the definition of the trigger tree to a running conditions database. Trigger menus will be driven by both instantaneous luminosity and physics goals.

---

<sup>2</sup> The Extensible Markup Language (XML) is a standard for creating special-purpose markup languages. Its primary aim is to facilitate the sharing of data across different systems. Languages based on XML are themselves described in a formal way via the Document Type Definition (DTD), allowing applications to modify and validate XML documents without prior knowledge of their form. See for example [72].

### 6.4.3 Validation

The various steps of HLT reconstruction and selection were first validated for the DAQ TDR[61]. The various codes involved have however been modified, improved and adapted to the HLT steering code since then. In preparation for the Physics TDR, it was necessary to validate the current version of the code (ORCA 8.7.2). In this prospect, efficiencies, rates and timing are three aspects to be considered.

It was not conceivable to reproduce all results from the DAQ TDR in this context, the main difficulties being often related to missing Monte-Carlo samples. It was decided to concentrate our efforts on some chosen quantities:

1. resolution in energy and momentum of electron and muon candidates,
2. efficiencies for some channels,
3. partial rate for each trigger stream,
4. total rate.

Part of this work was carried out with an older version of ORCA (ORCA 8.3.0), and it was then checked that no modification of the related code had been done between then and ORCA 8.7.2. Rate measurements were all carried out directly with the final L1/HLT trigger tables of ORCA 8.7.2. This information is gathered in an internal CMS note[73].

#### Electron and photon streams

Electron reconstruction was tested in the DAQ TDR on a sample of back-to-back electrons and positrons with  $P_t = 35$  GeV/c, without pile-up. The distribution of  $E_{meas}/E_{true}$  is presented in Figure 6.8. This result has been obtained with ORCA 8.3.0. This has to be compared with the DAQ TDR result shown previously, obtained with an earlier ORCA version (Figure 6.4). The resolution is found to be similar in the barrel, and slightly worse in the endcaps (new measurements indicate a resolution of 1.8% in the endcaps, while 1.23% was obtained for the DAQ TDR). In the same way, angular resolutions were shown to be compatible with previous results. As seen in Table 6.3, the angular resolutions at L3 are better by a factor three with respect to L2, due to the use of full track reconstruction.

The  $W \rightarrow e\nu$  channel was used as a benchmark for single electron efficiency. This efficiency is determined in the fiducial region defined as  $|\eta| < 2.5$  with the region  $1.4442 < |\eta| < 1.566$  excluded, for electrons above 29 GeV. The region where tracker cables screen the calorimeter is thus avoided in the calculation of efficiency. Efficiency for the double electron stream was checked using a  $Z \rightarrow ee$  sample, while  $H \rightarrow \gamma\gamma$  was used to validate the double photon stream. Results are presented in Table 6.4. Differences observed for the  $H \rightarrow \gamma\gamma$  channel with respect to the DAQ TDR results are understood as the effect of extensive modifications and improvements in the reconstruction code.

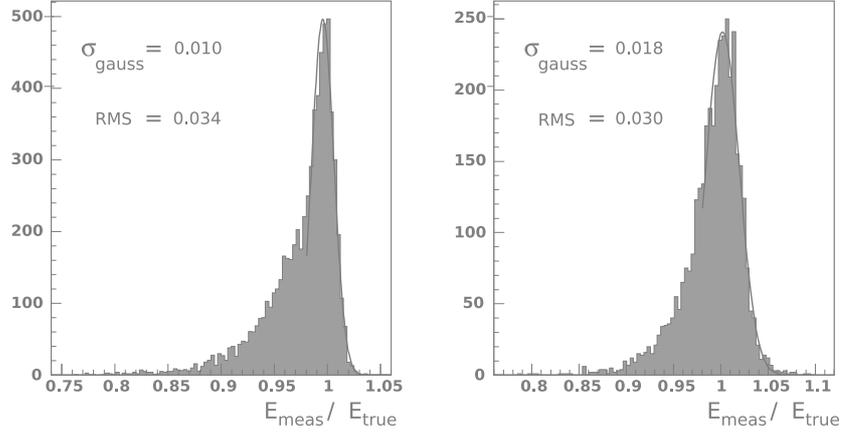


Figure 6.8: Distribution of  $E_{meas}/E_{true}$  for 35 GeV/c Level-2 electrons, (a) in the barrel, (b) in the endcap.

Table 6.3: Energy and angular resolution for the barrel ECAL using 35 GeV/c electrons simulated without pile-up. Values from the DAQ TDR [61] are shown in brackets.

	Energy resolution		Angular resolution		
	$\sigma_{Gauss}(E)/\mu$	$\sigma_{eff}(E)/\mu$	$\sigma_{Gauss}(\eta)$	$\sigma_{Gauss}(\phi)$	$\sigma_{eff}(\phi)$
L2	1.03%	2.6%	$1.2 \cdot 10^{-3}$	1.6 mrad	1.7 mrad
	(1.1%)	(2.2%)	( $1.1 \cdot 10^{-3}$ )	(1.7 mrad)	(2.5 mrad)
L3	idem	idem	$0.4 \cdot 10^{-3}$	0.4 mrad	0.4 mrad

Table 6.4: Efficiency for the benchmark channels used to validate electron and photon HLT streams. Values in brackets are from the DAQ TDR [61]. Efficiencies are given relative to the previous trigger level, and the cumulative efficiency is given at the end.

stream	$W \rightarrow e\nu$		$Z \rightarrow ee$		$H \rightarrow \gamma\gamma$	
	single e		double e		double $\gamma$	
L1	0.874	(0.872)	0.751	(0.753)	0.869	(0.908)
L2	0.997	(0.994)	0.979	(0.977)	0.995	(0.987)
L2.5	0.946	(0.936)	0.868	(0.855)	0.948	(0.996)
L3	0.825	(0.820)	0.99	(—)	0.996	(0.92)
TOTAL	0.68	(0.67)	0.63	(0.63)	0.816	(0.77)

### Muon streams

The resolution for muons was measured using a  $W \rightarrow \mu\nu$  sample at low luminosity. Figure 6.9 shows the resolution of the inverse transverse momentum determined by the level-2 and the level-3 calculations. On the horizontal axis, the quantity  $(1/P_T^{rec} - 1/P_T^{gen})/(1/P_T^{gen})$  is reported, where  $P_T^{gen}$  and  $P_T^{rec}$  are the generated and the reconstructed transverse momenta respectively. The distributions are divided into three pseudorapidity intervals: barrel ( $|\eta| < 0.8$ ), overlap ( $0.8 < |\eta| < 1.2$ ) and endcap ( $1.2 < |\eta| < 2.1$ ).

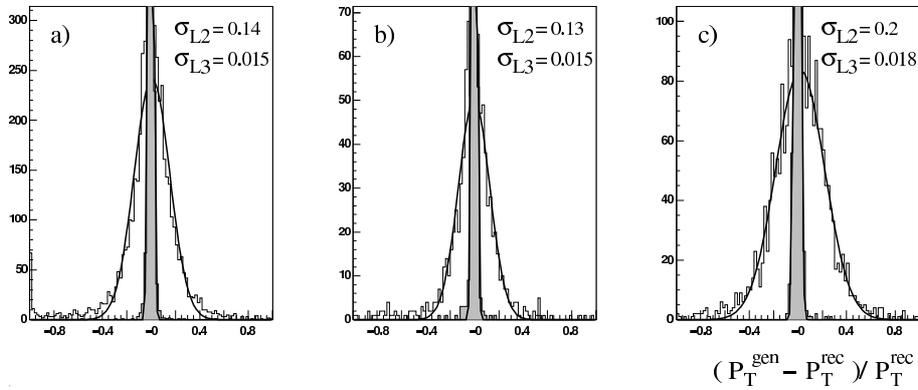


Figure 6.9: Distribution of  $(1/P_T^{rec} - 1/P_T^{gen})/(1/P_T^{gen})$  for the muons from the  $W$  decays, where  $P_T^{gen}$  and  $P_T^{rec}$  are the generated and the reconstructed transverse momenta respectively. This distribution is shown for level 2 and level 3 in three pseudorapidity intervals: (a)  $|\eta| < 0.8$ , (b)  $0.8 < |\eta| < 1.2$  and (c)  $1.2 < |\eta| < 2.1$ .

Efficiencies were measured on single  $W$  events (for single muons) and in  $t\bar{t}$  events (where both single and double trigger streams contribute). In both cases, the obtained efficiency is greater by about 10% to what was obtained in the DAQ TDR, and is respectively 79% and 82% for  $W$  and  $t\bar{t}$  events. This difference is not understood, but must be related to changes in the details of muon reconstruction.

## 6.5 HLT rate measurements

Rates already presented in Table 6.2 were obtained in preparation of the DAQ TDR 3 years ago. Changes in resolution and/or tuning of the thresholds may have affected these results, which motivated a validation in the context of the steering software development. This validation goes beyond the validation of the steering code and also concerns the specific trigger table that has been implemented so far. This validation work was pursued with ORCA 8.7.2 and

default trigger tables at low luminosity, which corresponds to the following entries in the configuration file:

- L1GlobalTrigger:TriggerMenu = lumi2x1033
- HighLevelTriggerXML:XMLfile = 2x1033HLT.xml

Unfiltered QCD samples generated in several  $P_T$  bins from 20 GeV/c to 1400 GeV/c were used for a total cross-section close to 1 mb (Table 6.5). Even if some missing contributions to the total rate are not negligible (like  $W \rightarrow e\nu$  events), this represents the main fraction of the total rate. Results of the rate measurement are summarized in Figure 6.10.

Table 6.5: Distribution of QCD events in Pt bins, together with the corresponding cross-section, number of simulated events, and number of events passing the HLT.

Bin Pt	Cross-section (mb)	Simulated	Triggered	Total Rate (Hz)
20-30	$7.819 \times 10^{-01}$	23400	14	936
30-50	$1.849 \times 10^{-01}$	19800	26	486
50-80	$2.433 \times 10^{-02}$	20000	89	217
80-120	$3.359 \times 10^{-03}$	18200	338	125
120-170	$5.654 \times 10^{-04}$	20000	1762	100
170-230	$1.163 \times 10^{-04}$	7200	2495	81
230-300	$2.812 \times 10^{-05}$	10000	6916	39
300-380	$7.848 \times 10^{-06}$	7000	6279	14
380-470	$2.396 \times 10^{-06}$	5000	4868	4.7
470-600	$9.249 \times 10^{-07}$	5000	2415	1.8
600-800	$2.903 \times 10^{-07}$	5000	2932	0.6
800-1000	$5.541 \times 10^{-08}$	2900	2900	0.1
1000-1400	$1.83 \times 10^{-08}$	4900	4900	0.04
Total	0.995	148400	40549	2003

Compared to the DAQ TDR results, it should be noted that:

- The “ $t\bar{t}H$ ” trigger, which was not there at the time of the TDR, contributes far too much to the trigger rate (754 Hz). Alone, it represents already more than seven times the allowed bandwidth.
- The rate is then dominated by the electron and photon streams. Rates are typically a factor 10 higher than expected. This large difference has to be understood.
- The relaxed electron and photon streams contribute largely to the rate and should maybe be reconsidered.

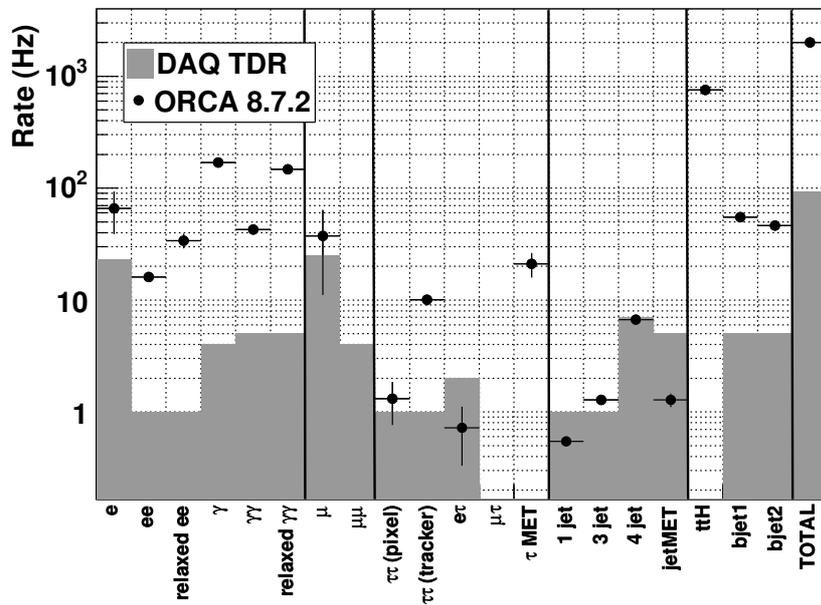


Figure 6.10: Rate observed for each subtrigger, as well as the total (inclusive) HLT rate for the default trigger table implemented in ORCA 8.7.2 . Large differences between new and old results are obtained.

- The rate for the muon triggers is consistent with the DAQ TDR. More statistic is required to draw final conclusions.
- Pixel-based double tau and electron-tau triggers are behaving as expected. The rate for the tracker-based double tau trigger is too high by a factor 10.
- The rate for the  $\tau \wedge$  MET trigger is 20 Hz. There was no such trigger in the DAQ TDR.
- The rate for the 1-jet, 3-jets and 4-jets is compatible with the values presented in the DAQ TDR (9 Hz).
- The rate for the MET  $\wedge$  jet is 1.3 Hz. This is lower than the value presented in the DAQ TDR (5 Hz).
- The rate for the b-jet triggers is much higher than the DAQ TDR (it reaches 55Hz). This should be resolved by putting stronger Pt cuts on the jets at L2.
- The total rate is 1 kHz, which is a factor 10 larger than the DAQ TDR specifications. Ignoring the  $t\bar{t}H$  contribution, one obtains 454Hz, which is still fairly high, mainly due to the excessive rate of electron and photon triggers.

The main outcome of this validation is thus the excessive rate of the HLT, as implemented in ORCA 8.7.2. The precise understanding of these differences will certainly need more work in collaboration with the experts of each trigger stream, as the precise understanding of rates and thresholds is needed for most of the analyses handling low-Pt objects.

Electron and photon subtriggers have a rate surprisingly high, which has to be understood. Moreover, the difference between standard and relaxed streams was not expected by the consulted experts. We shall check that the sample does not introduce any bias to the rate measurement, but isolation and Pt cuts will probably have to be retuned.

The rate measured for the muon streams is compatible with expectations but suffers from very low statistics. A dedicated prefiltered sample is certainly needed to achieve better results for these trigger streams.

The rate for the tracker-based tau triggers is observed to be 10 times too high. This is true for the double tau trigger, and for the  $\tau \wedge$  MET trigger, which also uses tracker-based validation. On the contrary, tau streams based on pixel isolation (double tau and  $e \wedge$  tau triggers) present correct rates. The isolation criteria in the L2.5 tracker tau validation has certainly to be controlled.

Moreover, the exclusive  $t\bar{t}H$  trigger is not acceptable. It is a relaxed muon trigger, with isolation and Pt cuts tuned to improve the trigger efficiency for the  $t\bar{t}H$  channel. Its rate is high, and is computed for all events, disregarding the L1 trigger bits.

## 6.6 Trigger evaluation sequence and timing issues

### Time optimization

As already mentioned, the HLT response is recursively evaluated, and the tree leaves correspond to the first step in the HLT evaluation. More precisely, the evaluation request is propagated recursively from the root element up to all leaves. The actual evaluation then starts from the leaves down to the root, going through all the HLT steps and stopping as soon as one element in the chain fails. There are three ways to optimize this sequence[74].

1. When two elements are combined with a logical “OR”, there is no need to evaluate the outcome of both elements if the first is true.
2. When two elements are combined with a logical “AND”, there is no need to evaluate the outcome of both elements if the first is false.
3. It is preferable to first evaluate the element that gives the quickest response, or that is often decisive.

These mechanisms are implemented in the HLT prototype and activated as an option of the algorithm. We introduced all three methods in the HLT steering code in the context of this thesis. Points 1 and 2 are straightforward. Point 3 proceeds as follows: Let  $p_i$  be the probability for the  $i^{\text{th}}$  ( $i = 1, 2$ ) subtrigger to give a positive outcome.  $ta_i$ , ( $tr_i$ ) is the corresponding accept (reject) time. For an “OR”, the mean time to accept an event depends on the evaluation sequence, and is given by

$$\langle Ta \rangle_{ij} = p_i ta_i + (1 - p_i) p_j (tr_i + ta_j), \quad (6.1)$$

where trigger  $i$  is evaluated first, then trigger  $j$ . The ordering that gives the smallest mean evaluation time has to be retained. In the case of an “AND” element, the mean reject time can be optimized in a similar way. It is given by:

$$\langle Tr \rangle_{ij} = (1 - p_i) tr_i + p_i (1 - p_j) (ta_i + tr_j). \quad (6.2)$$

An extension to the case of an “OR” with more than two subtriggers has been proposed in [75].

The collaboration does not plan, at least in a first phase, to optimize the trigger sequence. This is justified since rejection cannot be optimized. Time can be gained only for accepted events, which represents only 1/1000 of the input. Moreover, optimizing the trigger sequence could induce difficulties to define streams.

An optimized trigger sequence may nevertheless find its justification in the early phase of data-taking at very low luminosity, when the trigger will not just keep all events but start to reject some events. With a rejection factor of 50%, the effect of time optimization may be significant. One possibility would be to use the optimized trigger in the online farm, and to reevaluate the

HLT without optimization in the CERN offline computing facility to build the offline streams. The drawback of such a scheme is that it complicates operations, notably as the trigger menus would have to be synchronized between online and offline facilities. It would also have a cost in CPU, as some trigger elements would have to be evaluated twice.

### HLT timing

The actual time needed to evaluate the default HLT menu was measured on unbiased QCD events using ORCA 8.7.2. The total computing time distribution is shown Figure 6.11. The mean evaluation time lasts one second, which is six times the targeted evaluation time (result obtained on a 2.4 GHz Intel Xeon CPU, rated at 825 SpecInts). It must be noted that this result has been obtained as a byproduct of rate measurements, on the uncontrolled environment of a computing cluster at Fermilab, and disregarding of eventual other jobs running on the same machine. Timing is nevertheless observed to be stable in time. A careful study of the timing for each subtrigger showed that, as expected, the main contribution comes from the  $t\bar{t}H$  trigger, which is the only one not vetoed by L1 bits. In its present form,  $t\bar{t}H$  is a tuned L3 muon trigger, which obviously cannot be run for each event, but has to be vetoed by L1 and preceded by a corresponding L2 trigger element.

The time distribution is also shown to be compatible with a log-normal, which is required for correct operation of the CMS filter farm. More tails would require a special treatment so as to avoid “locking” of computing resources, which is a source of dead-time. The confirmation of this hypothesis is important as it was only checked per subtrigger in the DAQ TDR<sup>3</sup>.

Estimating the time consumed by the HLT without the  $t\bar{t}H$  trigger is difficult, since there are cross-talks at the reconstruction level with other subtriggers. It has not been possible to rerun all jobs without  $t\bar{t}H$ , since many QCD events have to be processed to have enough statistics. Estimates show that the mean time excluding this trigger is of the order of 350 ms. To this reconstruction and selection time, one must add the time needed to convert data from the raw readout format to the format suited for reconstruction. From these estimates, it follows that a single 825-specInts CPU can handle 2.85 Hz from the L1 trigger output. Considering that the L1 output will be 50 kHz during the first year, 18000 such CPUs are found to be needed in the filter farm. Since the Moore law[76] predicts a factor 2 improvement in CPU performances every 18 months, one obtains a typical number of 4500 CPUs in 2007. The present estimation has to be compared to the results presented in the DAQ TDR. In December 2002, the mean HLT computing time has been evaluated to be 300 ms on a 1GHz Pentium III. This is twice as good as the result presented in this work. If our result can be trusted, and if no solution

<sup>3</sup>It is nevertheless expected that this approach is well justified, as cross-talks between the considered subtriggers are small, and would anyway reduce the total time, so that it is conservative.

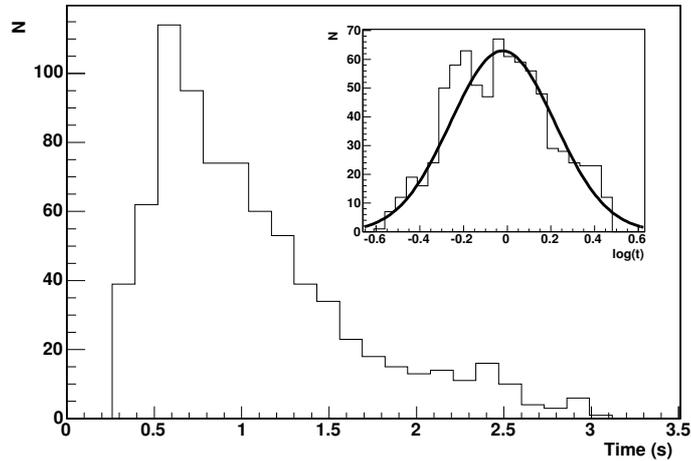


Figure 6.11: Total HLT computing time for QCD events passing the L1 trigger. The distribution is well described by a log-normal, as seen in the inset.

is found to improve the evaluation time, this will have a strong impact on the HLT performances, since it is not expected to have more than about 1500 dual-CPU computers in the filter farm. It is therefore important to cross-check our result on a more controlled environment, were no other resource-consuming task influences the performances. Finally, the current effort to improve software performances must certainly be continued.



## Search for WH associated production

### 7.1 Introduction

While the latest analyses of LEP data are being performed either to refine Standard-Model observables or to search for new physics in the uncovered channels, analyses are being developed for the long awaited LHC experiments. The philosophy is very different. While LEP studies are the result of several years of thought analysis and benefit from a good understanding of Standard-Model processes, the work at LHC is still in an early stage. Having handled ALEPH data in the first part of this dissertation (Chapter 4), we wanted to turn to CMS and contribute to the preparation of data taking and analysis.

The Standard-Model cross-section for each Higgs boson production mechanism is shown Figure 7.1, as a function of the Higgs boson mass. The dominant contribution to Higgs boson production at LHC comes from gluon-fusion processes over the mass range from  $100 \text{ GeV}/c^2$  up to  $1 \text{ TeV}/c^2$ . The boson-fusion mechanism is less important but still contributes significantly, especially for Higgs boson masses of the order of  $1 \text{ TeV}/c^2$ . The Higgsstrahlung processes, where the Higgs boson is produced in association with a W or Z boson, is far less important, as an anti-quark from the proton sea must enter the diagram. Cross-sections are a factor 20 to 1000 lower, depending on the Higgs boson mass. The  $t\bar{t}H$  associated production has a similar cross-section. The leading diagrams corresponding to these processes are presented in Figure 7.2.

Many dominant channels have been covered by the CMS collaboration [78]. They often focus on gluon-fusion processes since the cross-section is higher. In this chapter, we present the analysis we performed, considering the WH associated production, with a subsequent decay of the Higgs boson into a W pair. The cross-section for this process exhibits a maximum near the WW resonance, due to the combined behavior of the production cross-section

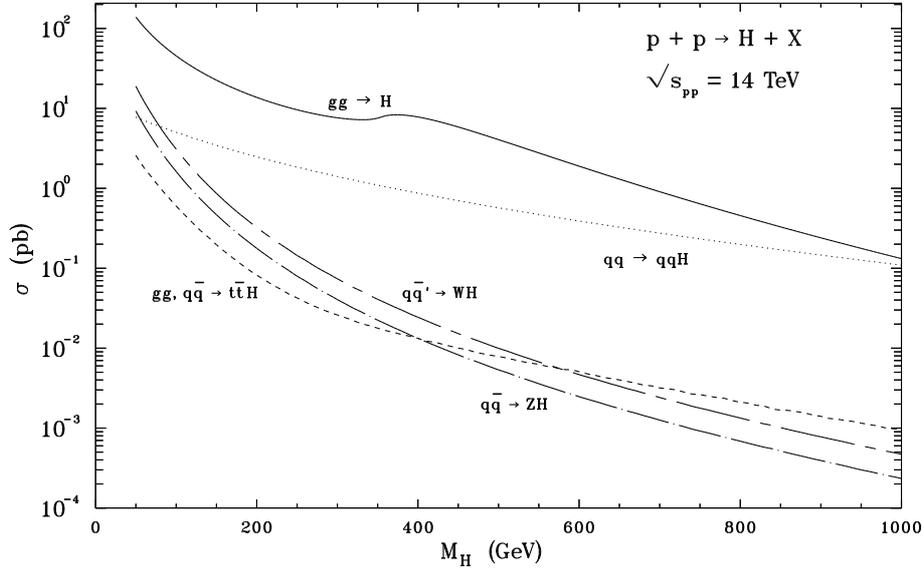


Figure 7.1: Next-to-leading order cross-section calculations for the Standard-Model Higgs boson at LHC [77].

(Figure 7.1) and the Higgs boson branching ratio (Figure 2.1). As it can already be seen in Figure 7.3, we will concentrate on the intermediate mass region between  $115 \text{ GeV}/c^2$  and  $190 \text{ GeV}/c^2$ .

Motivations for studying this channel are twofold. First, the corresponding diagram contains the  $g_{HWW}$  coupling constant twice. It could therefore be precisely measured. Moreover, it is one of the few processes that are still allowed for a fermiophobic Higgs boson, along with some boson-fusion processes. Our contribution to the HLT software also finds a natural continuation in this analysis. As mentioned previously, CMS analyses will be based on trigger streams, and it is important to adequately choose the streams for each analysis. The contribution of each stream has therefore been controlled before and after the signal selection. This is presented in Section 7.5.7.

This analysis is not intended to be a LEP-like analysis, but a proof-of-principle of what can be done with the present software and expertise. It is therefore a software readiness demonstration and is performed in the spirit of the *CMS Physics TDR* that will be published by the collaboration in April 2006. The present analysis was performed with ORCA 8.7.2. Default reconstruction methods used for the official Physics-TDR DSTs were used when not stated otherwise.

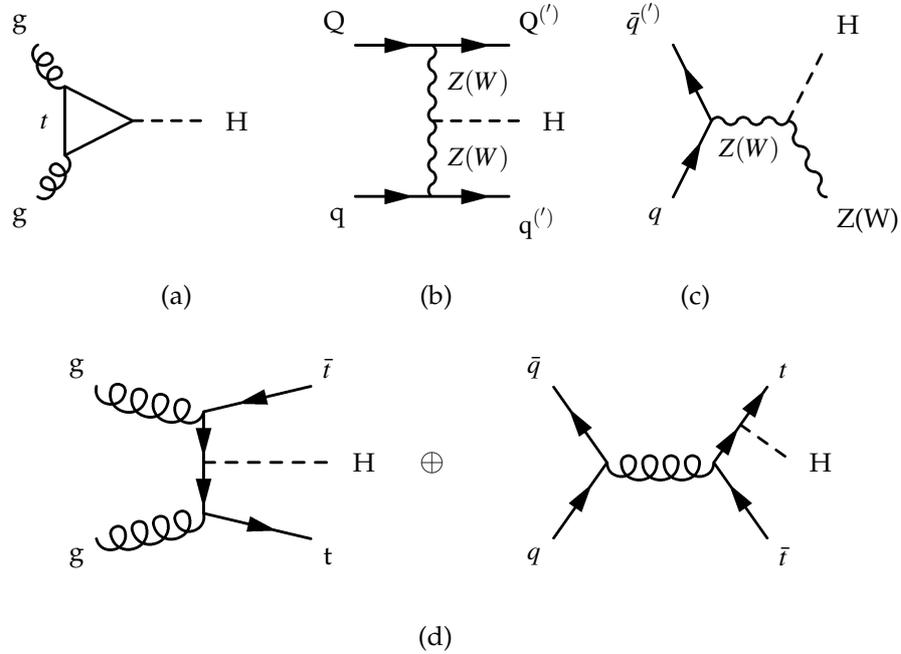


Figure 7.2: Representative Feynman diagrams describing the main Higgs boson production mechanisms at LHC. (a) gluon fusion, (b) boson fusion, (c) Higgsstrahlung, (d) associated production with a  $t\bar{t}$  pair.

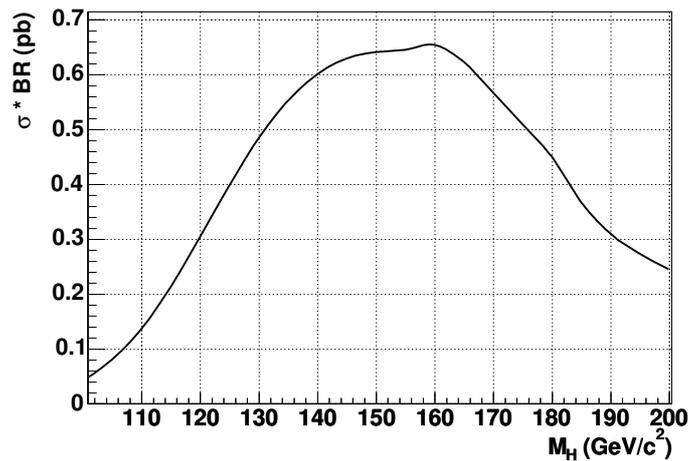


Figure 7.3: Production cross-section time branching ratio for WH associated production, with a subsequent decay of the Higgs boson into a W pair. Obtained with HDECAY[5] and V2HV [79].

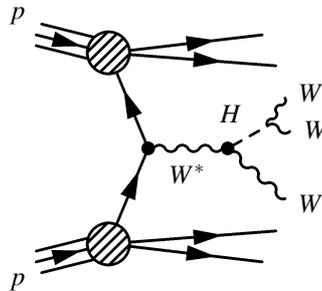


Figure 7.4: Feynman diagram for the WH associated production, with a subsequent decay of the Higgs boson into a W pair. This process is the subject of the present analysis.

## 7.2 Signal

As we have just announced, we are interested in the WH associated production, with a subsequent decay of the Higgs boson into a W pair. Since three W bosons are produced in this process (Figure 7.4), the final state is characterized by six fermions in addition to soft remnants from the protons. Topologies therefore range from six jets to three charged leptons. Amongst these topologies, some are very clear while others are very difficult to disentangle. This is generally determined by the number of leptons. Two clear signatures are the same-sign lepton-pair channel and the three-lepton channel. Previous fast-simulation ATLAS studies [80, 81, 82] show that these two event classes can give a significant contribution to the discovery potential. In this thesis, we decided to investigate events containing three leptons in the final state, concentrating on electrons and muons (including leptonic tau decays).

The next-to-leading order (NLO) QCD corrections to WH production are known to increase the cross-section by about 30%. NNLO QCD corrections are expected to be small. Accuracy is estimated to reach 15%, a level at which electroweak corrections become significant. These electroweak NLO corrections were recently calculated and led to a  $\mathcal{O}(5)\%$  decrease in cross-section [83]. In the following sections, the cross-section including only QCD NLO effects will be used to keep coherency with other CMS studies.

Events are generated using PYTHIA, which includes spin-correlation effects. At this level, W bosons are forced to decay into electrons, muons or taus. The resulting files are processed through a detailed GEANT simulation of the CMS detector. Events are *digitized*<sup>1</sup> at low-luminosity ( $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ ), which means that the electronic response is simulated taking into account pile-up ef-

<sup>1</sup>More details about the CMS simulation chain and the hardware used for simulation (and analysis) are given in Appendix B.

fects due to 3.5 simultaneous events in a bunch crossing. The effect of five bunch-crossings before and three bunch-crossings after the nominal bunch crossing is also considered as it affects the energy determination in calorimeters as well as the muon system. A small number of events was generated with CompHep[84] to control effects of spin-correlation.

Starting at  $115 \text{ GeV}/c^2$ , we also considered Higgs bosons at 125, 130, 140, 150, 160, 170, 180 and  $190 \text{ GeV}/c^2$ . This covers the complete mass range accessible in this channel. Ten thousand events were generated for each of the nine simulated Higgs boson masses. Simulated events are shown in Figures 7.5 and 7.6.

Figure 7.7 shows the  $P_T$  spectrum of leptons in signal events. These leptons are relatively soft and the spectrum has a significant contribution at very low transverse momenta. This has an impact on reconstruction and selection efficiency. In 60.5% of the events, all three leptons are produced inside the fiducial region defined as  $|\eta| < 2.4$ . Three neutrinos are also produced and generate missing transverse momentum ( $\cancel{p}_T$ ). This is nevertheless significantly reduced by mutual cancellation effects between the three neutrinos. The missing transverse momentum distribution is presented in Figure 7.8.

Spin correlation between the Higgs boson decay products is an important effect which significantly modifies the event's topology. The process is identical to what was presented in Chapter 4 for the analysis of LEP events, and will not be repeated here. As before, it will tend to reduce the angle between leptons originating from the Higgs boson decay chain. Even so, the angle between leptons is influenced by the Higgs boson boost along the z axis ( $\langle E_H \rangle = 576 \text{ GeV}$  for a Higgs boson mass of  $140 \text{ GeV}/c^2$ ), which hides part of spin correlation effects. The influence on the acoplanarity between leptons is on the contrary still clearly visible. In order to check the magnitude of this effect, the topology of events generated with PYTHIA and with CompHep was compared. The result of this comparison is shown in Figure 7.9, which presents the angle between leptons originating from the Higgs boson for events simulated with PYTHIA and CompHep. Spin correlation clearly favors smaller  $|\Delta\phi|$ .

There is no strong dependence of the quantities shown in this section on the Higgs boson mass. This justifies the development of the event selection independently from the mass hypothesis. In the following discussion, a mass of  $140 \text{ GeV}/c^2$  has been used.

## 7.3 Background

All Standard-Model processes likely to produce three leptons must be considered as background for this analysis. This includes events where three leptons are actually produced (e.g. WWW or WZ) but also events with two or four leptons, i.e. with a fake lepton or a missed lepton. One particular case is the production of leptons in the semi-leptonic decay of a B meson. In the present

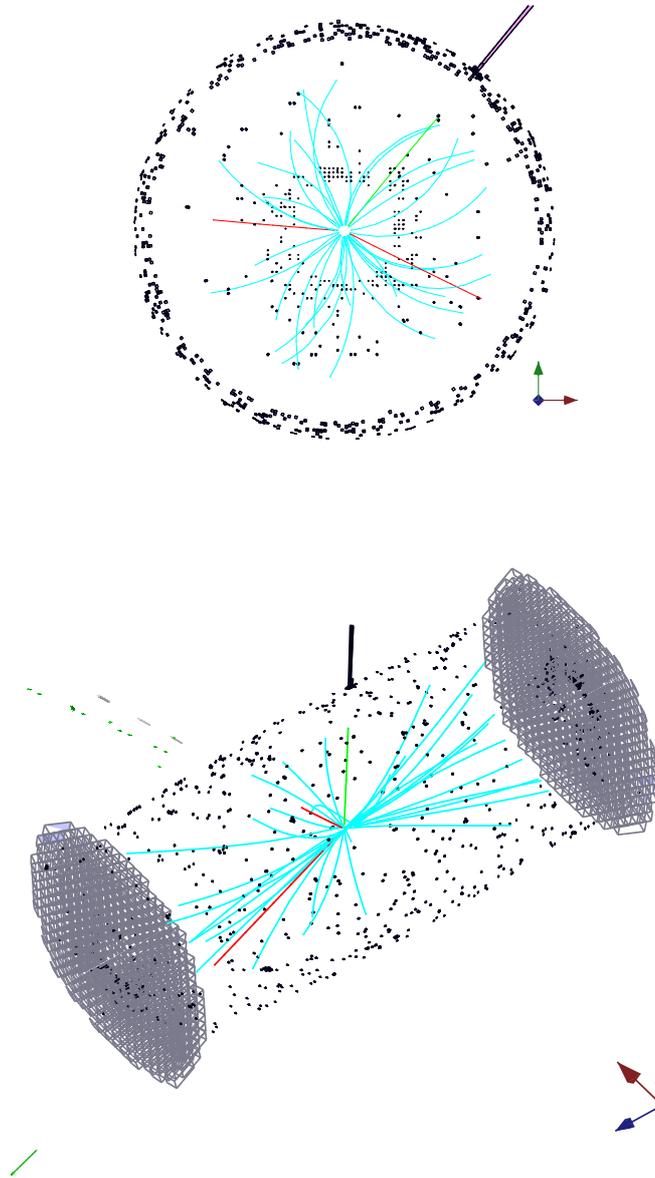


Figure 7.5: Two views of the same event where there are two muons (red) and one electron (green), for a Higgs boson mass of  $115 \text{ GeV}/c^2$ . The top part shows a transverse view of the event. The electron deposit in ECAL is seen at 45 degrees from the vertical. The bottom part shows the same event together with ECAL endcaps. One sees the electron deposit in the central part, and segments in the muon chambers outside the calorimeter's volume.

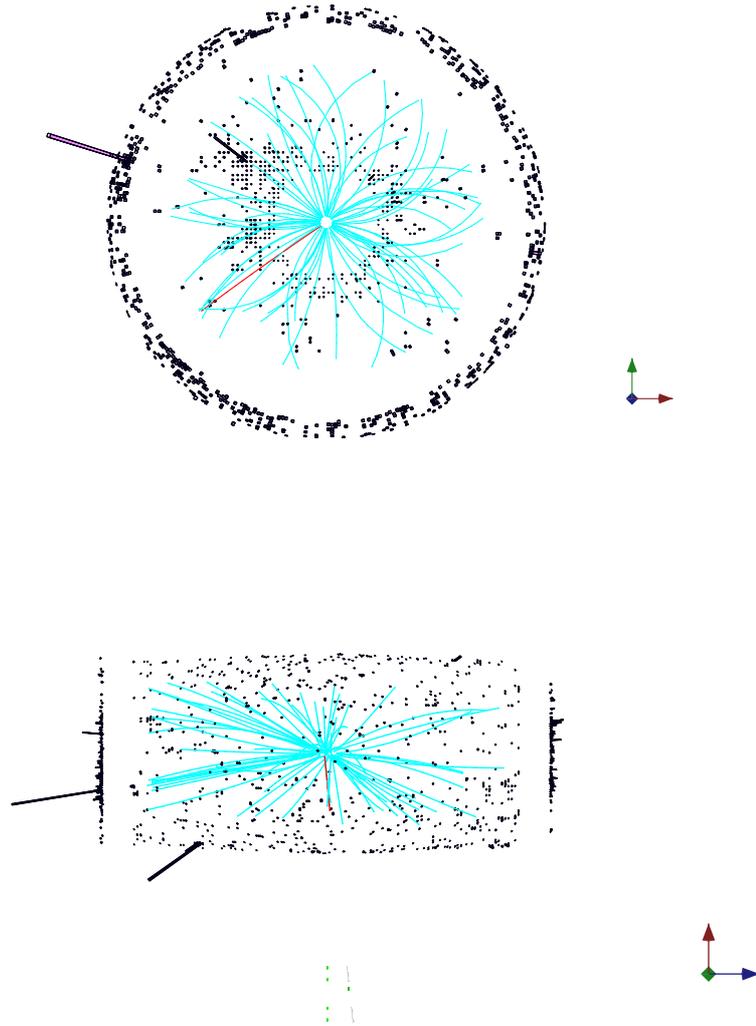


Figure 7.6: Two views of the same event where there are two electrons (green) and one muon (red), for a Higgs boson mass of  $115 \text{ GeV}/c^2$ . Electrons are more easily identified by the energy deposited in the calorimeter. In the top diagram, one electron is seen, while the second hits the endcaps. Both electrons are seen in the bottom view.

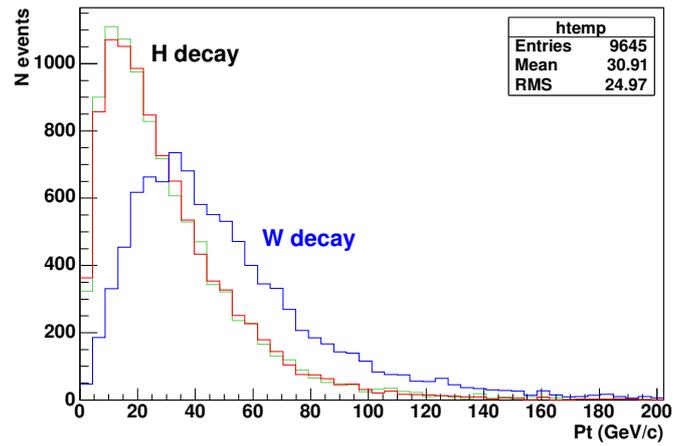


Figure 7.7:  $P_t$  distribution for leptons in signal events ( $m_H = 140 \text{ GeV}/c^2$ ). The two narrower distributions correspond to leptons originating from the Higgs boson decay, while the broader distribution corresponds to leptons coming from the spectator W boson.

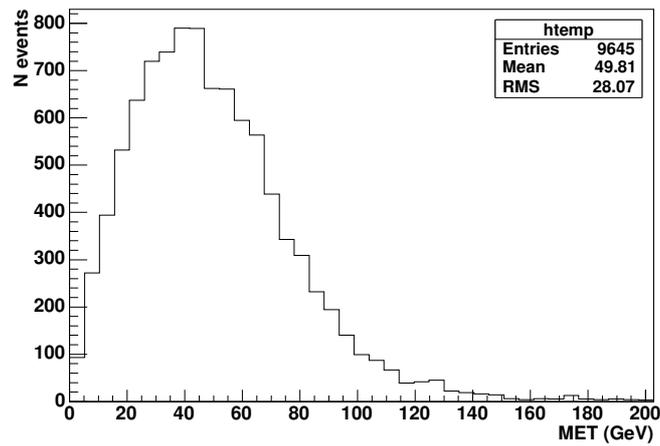


Figure 7.8: Missing transverse energy ( $\cancel{E}_T$ ) distribution for signal events for a Higgs boson mass of  $140 \text{ GeV}/c^2$ .

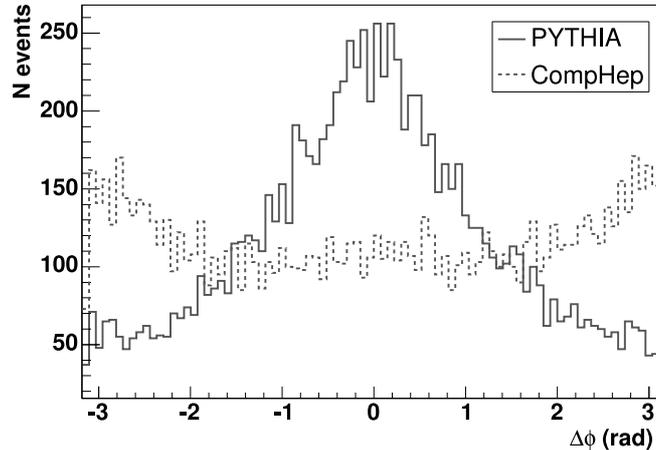


Figure 7.9: Acoplanarity ( $\Delta\phi$ ) between leptons originating from a  $140 \text{ GeV}/c^2$  Higgs boson for events simulated with PYTHIA (continuous) and CompHep (dashed). One sees the effect of spin correlation, only taken into account in PYTHIA. The normalization is arbitrary.

analysis, we considered the production of  $WWW$ ,  $WZ$ ,  $ZZ$ ,  $t\bar{t}$ , and  $Wt$ . Most of the processes used are simulated with PYTHIA, except for  $WWW$ , which is generated with CompHep, and  $Wt$  generated with TopRex [85]. In all cases, PYTHIA is used for the hadronization step. The dedicated generator AlpGen [86] is used to compute the cross-section of  $Zb\bar{b}$ ,  $Wb\bar{b}$  and  $Z$ +jets. Details are given in Table 7.1. It gives the cross-section (including branching ratios of  $W,Z$  into leptons), the generator used and the number of events produced. One sees a large difference between irreducible backgrounds with three leptons, for which the cross-section is similar to the signal's, and reducible backgrounds that have a larger cross-section.

Some background sources were ignored for practical reasons.  $Wb\bar{b}$ ,  $Zb\bar{b}$  or single top are potential background sources and should ideally be integrated in the study but are not expected to contribute much to the final number of events. We will show how the implemented selection procedure efficiently suppresses  $t\bar{t}$ , and  $Wt$  backgrounds, which are larger and potentially more dangerous.  $b\bar{b}$  events could contribute to the background if one of the  $B$  mesons decays in chain. Due to the large cross-section, handling this background would require a huge statistic, and has not been included explicitly in the study. Previous ATLAS studies show it is indeed negligible (after requiring three isolated leptons). The  $Z$ +jets events could further enlarge the background if a particle from the jet is misidentified as an electron. Such an effect is expected to be low in offline reconstruction, and greatly reduced by isolation cuts. Here also, a serious consideration of the background would require

Table 7.1: Background processes considered into the present analysis. The cross-section includes the decay of W and Z bosons into leptons. The generator and the number of events processed are also shown together with the corresponding weight for a luminosity of  $1fb^{-1}$ . For comparison, the signal varies, with the same hypotheses, from 5 fb to 20 fb in the considered mass range.

Background	Cross-section	Generator	MC statistic	weight ( $1fb^{-1}$ )
WWW( $3l^{\pm}$ )	4.95 fb	CompHep	10000	$5.19 \times 10^{-4}$
WZ( $3l^{\pm}$ )	1.71 pb	Pythia	50000	$3.46 \times 10^{-2}$
ZZ( $4l^{\pm}$ )	0.17 pb	Pythia	50000	$3.67 \times 10^{-3}$
$t\bar{t}(l^+l^-bb)$	90.9 pb	Pythia	100000	0.93
Wt( $l^+l^-b$ )	5.25 pb	TopRex	50000	0.11
Zbb( $l^+l^-bb$ )	6.52 pb	AlpGen	-	-
Wbb( $l^{\pm}bb$ )	9.3 pb	AlpGen	-	-
b $\bar{b}$	0.33 mb	Pythia	-	-
Z+jets	786 pb	AlpGen	-	-

a very large statistics, which would be incompatible with present computing resources.

The WWW background cannot be generated with PYTHIA as it corresponds to a six-fermion final state. It was therefore generated with CompHep, hadronized by PYTHIA and then followed the usual CMS simulation chain. It corresponds to the production of three W bosons via quartic W couplings, via a  $Z^*/\gamma^*$ , or directly in the t channel. Diagrams contributing to the WWW production are presented in Figure 7.10. It must be noted that interferences are important when computing the cross-section. This background is *a priori* difficult to distinguish from the signal, as the final state differs only by the effect of spin correlation, and cannot be separated event by event. The cross-section ( $4 \times 10^{-2}$  pb) is nevertheless lower than what is predicted for the signal.

Another interesting background is the (associated) single top production Wt. It enters as a background when, in addition to leptons from W decays, a lepton is seen in the jet from the B meson, be it either a fake electron or a real lepton from a semi-leptonic B decay. Two diagrams contribute to this process, respectively for the s and t channel, as shown in Figure 7.11. The cross-section (51 pb) is large compared to that of the signal, but is reduced to 659 fb when one has three electrons or muons in the final state, taking into account tau and B decays. It can already be noted in these events that at least one lepton is expected to be less isolated, and that the hadronic activity is significantly higher than for the signal. As shown later, in Section 7.5, it is used to suppress this background (and the  $t\bar{t}$  background) with respect to the signal. Compared to Wt, the  $t\bar{t}$  pair production is more easily suppressed since there is a second b-jet in the event, but it has a larger cross-section. The weight associated to

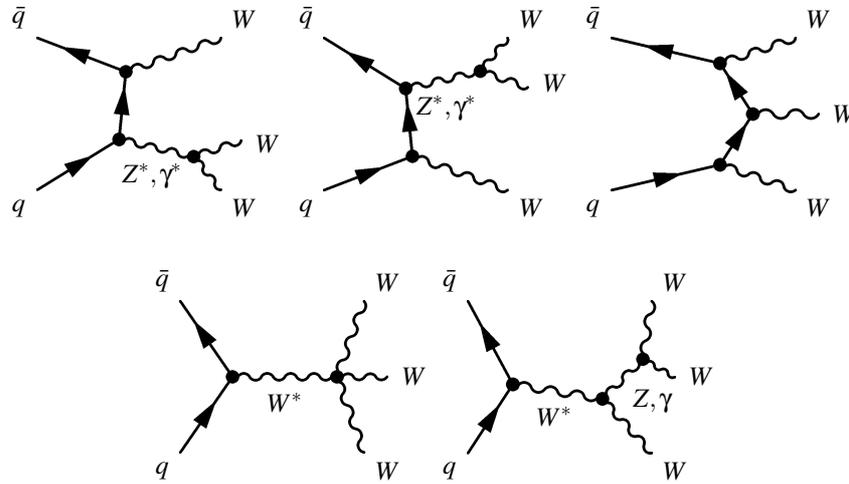


Figure 7.10: Leading order diagrams for the production of three W bosons at LEP.

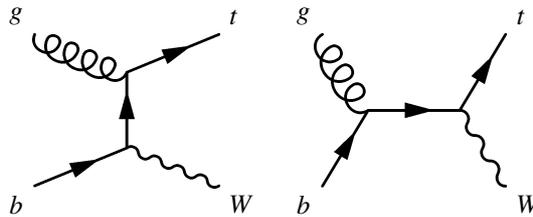


Figure 7.11: Leading order diagrams for the production of  $Wt$  at LHC.

this background is large, as it would require more statistics in the simulated sample. Computing resources did not allowed to generate and analyze more events.

## 7.4 Trigger issues

Trigger is an important aspect of the analysis in several ways. Analyzed events must first pass the trigger chain, from L1 to HLT. This step is designed to be as efficient as possible for events containing at least one energetic object, but may induce a significant loss for soft processes. Then, at analysis time, one will probably have to choose analysis streams to search for new signal. Streams will mainly depend (if not exclusively, as mentioned in Chapter 6) on trigger patterns. In order to make this choice in a reasonable way, trigger efficiencies are studied both before and after event selection.

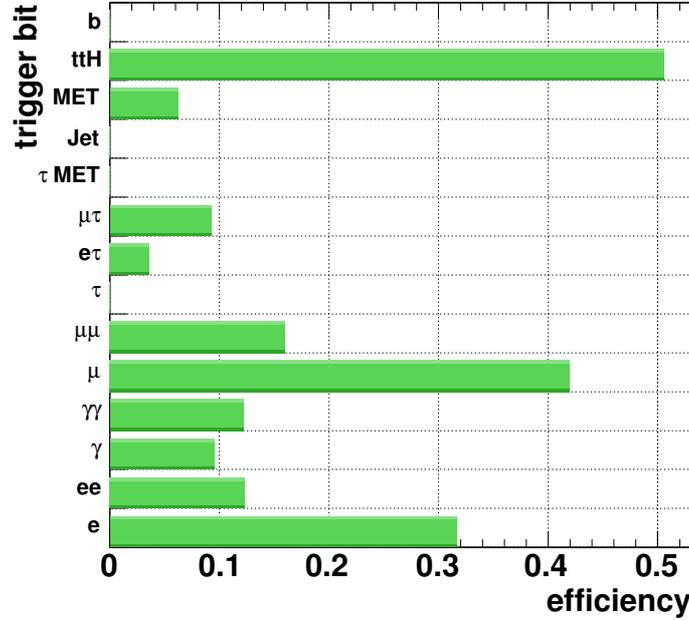


Figure 7.12: Distribution of the trigger decision in subtriggers for signal events containing a  $140 \text{ GeV}/c^2$  Higgs boson. The subtrigger is shown on the vertical axis; the horizontal axis shows the efficiency calculated as  $N_x/N_{HLT}$ , where  $N_x$  is the number of events triggered by the algorithm  $x$ , and  $N_{HLT}$  is the number of events triggered by the High-Level Trigger.

#### 7.4.1 Triggering on signal events

Trigger tables at low luminosity were discussed in Chapter 6. The global (cumulative) trigger efficiency after L1 and HLT is found to reach 72% for a  $140 \text{ GeV}/c^2$  Higgs boson using the full trigger table. Individual contributions to this efficiency are shown in Figure 7.12. The main contribution comes from the dedicated  $t\bar{t}H$  trigger. In the present form, this is a relaxed muon trigger that has a suspicious rate. It will be ignored in the following. Main contributions then come from single and double leptonic ( $e$  and  $\mu$ ) triggers. There is a small contribution from the missing transverse energy trigger ( $\cancel{E}_T$ ) and from combined ( $e \wedge \tau$ ) and ( $\mu \wedge \tau$ ) triggers. Finally, it is interesting to note that (b)-jet triggers are completely inefficient in this channel. This is a direct consequence of the small overall hadronic activity. The correlation between these subtriggers is shown in Figure 7.13. One observes that the  $t\bar{t}H$  trigger is strongly correlated to the standard single muon trigger. Photons and combined tau triggers are also shown to be correlated with other triggers.



For this analysis, we restricted ourselves to affordable triggers known to have the highest impact on the total efficiency: single- and double-electron and muon triggers. This choice is made at HLT, but implicitly limits the L1 to the corresponding trigger bits due to the vetoing mechanism. The correlation matrix between these triggers is:

$$\rho(i, j) = \left( \begin{array}{c|cccc} \text{bit} & e & ee & \mu & \mu\mu \\ \hline e & 1 & 0.378 & -0.427 & -0.299 \\ ee & 0.378 & 1 & -0.322 & -0.244 \\ \mu & -0.427 & -0.322 & 1 & 0.381 \\ \mu\mu & -0.299 & -0.244 & 0.381 & 1 \end{array} \right). \quad (7.1)$$

The presence of a single-electron trigger favors the double-electron trigger, as expected, and disfavors muon-based triggers. Conversely, muon-based triggers disfavor electron-based triggers. The correlation factors for  $e$  vs.  $\mu\mu$  and  $ee$  vs.  $\mu$  only differ because of efficiency and threshold effects. The correlation coefficient for  $e$  vs.  $ee$  is also very close to the symmetric  $\mu$  vs.  $\mu\mu$  coefficient. These are consequences of the equal distribution of electron and muon topologies for this analysis.

When restricting ourselves to key triggers, the global (cumulative) trigger efficiency after L1 and HLT is found to be 65% for a 140 GeV/ $c^2$  Higgs boson. There are 7% “lost”, which were coming from  $\cancel{E}_T$  and  $t\bar{t}H$  triggers. Figure 7.14 shows the efficiency for each (exclusive) trigger pattern, given the above choice of interesting bits. The loss of efficiency with respect to the full trigger is also shown as a separate entry. As will be demonstrated later on, the impact of ignored trigger bits is further reduced by the event selection, which favors multi-leptonic patterns. Trigger patterns after the analysis are discussed in Section 7.5.7.

Trigger efficiency depends on the final state lepton types. Details about the efficiency for each type of event (defined from the number of muons, electrons and taus in the event) are given in Figure 7.15. Events containing one or more muons are more easily retained (efficiency reaches 85% for events with three muons) while tau events are only marginally selected (efficiency: 12%). To conclude this discussion of trigger performances for the signal, we will consider the mass dependence of the trigger efficiency (Figure 7.16). Efficiency rises slightly with the Higgs boson mass, from 58% at 115 GeV/ $c^2$  to 74% at 190 GeV/ $c^2$ .

## 7.4.2 Triggering on background events

Table 7.2 shows the trigger efficiency for each source of background. Efficiency, when restricting ourselves to single- and double-electron and muon triggers, varies from 64% to 73%, which is the same magnitude as the trigger efficiency for signal events. This restricted trigger is 15% (for  $t\bar{t}$ ) to 5% (for ZZ) less efficient than the inclusive High-Level trigger. Trigger patterns are domi-

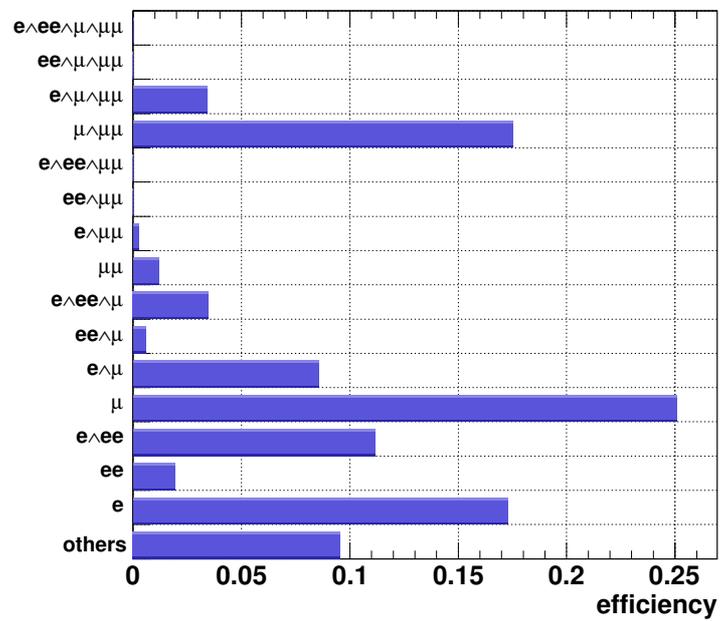


Figure 7.14: Trigger efficiency by trigger pattern, for the signal. Efficiency is calculated as  $N_x/N_{HLT}$ , where  $x$  is one of the  $2^4$  exclusive trigger classes. “Others” stands for unconsidered trigger patterns.

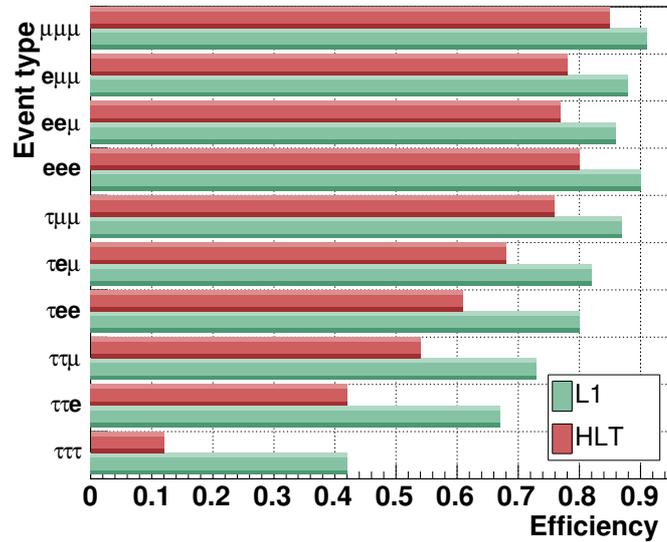


Figure 7.15: Trigger efficiency for each class of Monte-Carlo events. Results are given after L1 and after HLT. Efficiency is computed as the ratio between the number of triggered events and the total number of generated events.

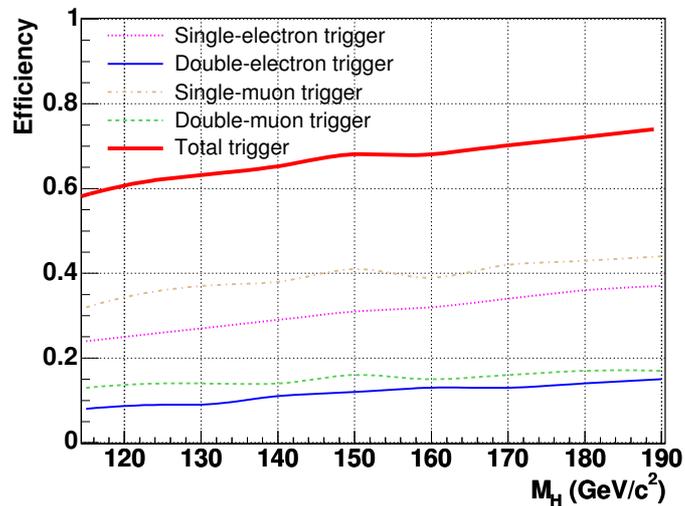


Figure 7.16: Total and partial trigger efficiency ( $L1 \wedge HLT$ ) as a function of the hypothetical Higgs boson mass. Efficiency is given separately for each of the single- and double-electron or muon triggers, as well as for the total trigger combining each of these.

Table 7.2: Trigger efficiency for each source of background. Efficiency at HLT with the restricted trigger set used in the present analysis is also presented.

Background	L1 efficiency	HLT efficiency	$e, ee, \mu, \mu\mu$ HLT efficiency
WW( $3l^\pm$ )	0.87	0.79	0.73
Wt( $1^{+1^-}b$ )	0.88	0.78	0.67
WZ( $3l^\pm$ )	0.8	0.72	0.65
ZZ( $4l^\pm$ )	0.78	0.69	0.64
$t\bar{t}(1^{+1^-}b\bar{b})$	0.91	0.79	0.65

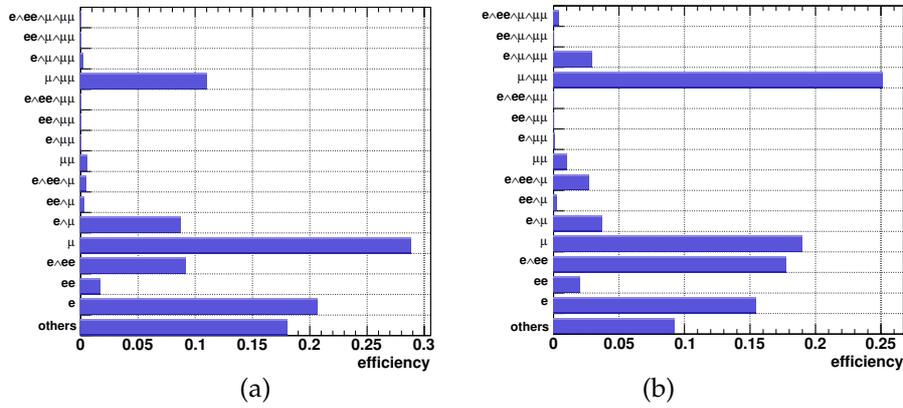


Figure 7.17: Trigger efficiency by trigger pattern for  $t\bar{t}$  and  $Wt$  events (a) and for other backgrounds (b). Efficiency is calculated as  $N_x/N_{HLT}$ , where  $x$  is one of the  $2^4$  exclusive trigger classes. “Others” stands for unconsidered trigger patterns.

nated by single-electron and muon triggers. The detailed contribution of each bit is shown in Figure 7.17.

## 7.5 Signal discrimination

Events that have passed the HLT can be handled with offline-quality algorithms. Electrons, muons, taus, (b)jets and  $\cancel{E}_T$  are reconstructed with dedicated algorithms without the time constraints that have driven the HLT design. Signal events are selected using that information.

### 7.5.1 Preselection and topology reconstruction

A first set of selection criteria is applied to select signal-like topologies.

- Three and only three leptons are required in the event.
- The total charge of these three leptons is required to be either +1 or -1.

This contributes to reject badly reconstructed events. Among events with a total of three electrons and muons in the detector ( $|\eta| < 2.4$ ), 77% are reconstructed with three leptons. This corresponds to a global efficiency of 92% per lepton in the considered region and is compatible with expected CMS performances. Efficiency rises with lepton energy.

It is important to assign reconstructed leptons either to the Higgs boson decay or to the decay of the W boson that does not come from the Higgs boson. This is achieved by choosing the two closest opposite-sign leptons. The third lepton must then come from the associated W boson.

Preselection cuts we have just discussed naturally arise from the topology, and were not optimized. The following selection cuts were optimized to maximize  $\sigma = s/\sqrt{b}$ , where  $s$  and  $b$  are the total number of signal and background events, respectively. That optimization is performed on an independent sample at  $m_H = 140 \text{ GeV}/c^2$ . Each possible cut is considered several times in an iterative procedure. When it is necessary (e.g. for cuts affecting mainly low cross-section backgrounds) only a subset of background has been used. The order in which each cut is applied has also been varied during the optimization process to better take into account correlations among them. All following plots are obtained with trigger and preselection cuts applied.

### 7.5.2 Lepton selection

Electrons and muons are reconstructed using default offline-quality reconstruction methods. These are `OfflineElectronReco::defaultQuery()` and `RecQuery("IsolatedGlobalMuonReconstructor")`, used respectively for electrons (`ElectronCandidates`) and muons (`RecMuons`). To avoid some fake leptons originating from reconstruction errors, the minimal angle between pairs of leptons is required to be higher than 0.1 rad. For electrons, additional quality cuts are applied: the energy measured by ECAL and the momentum obtained by the tracker must agree within 50%, and the ratio of energy measured by HCAL and ECAL must be lower than 0.15.

As all leptons come from a W boson decay, they are expected to be more energetic than “fake” leptons or leptons produced in a jet. Only leptons above 14 GeV/c are thus retained.

The angle between leptons attributed to the Higgs boson can be used to distinguish signal and background. Both the acollinearity ( $\theta_{aco}$ ) and the the acoplanarity ( $\Delta\phi$ ) between the leptons are used, as they provide complementary information. Typically, the acollinearity must be lower than 0.75 rad, and

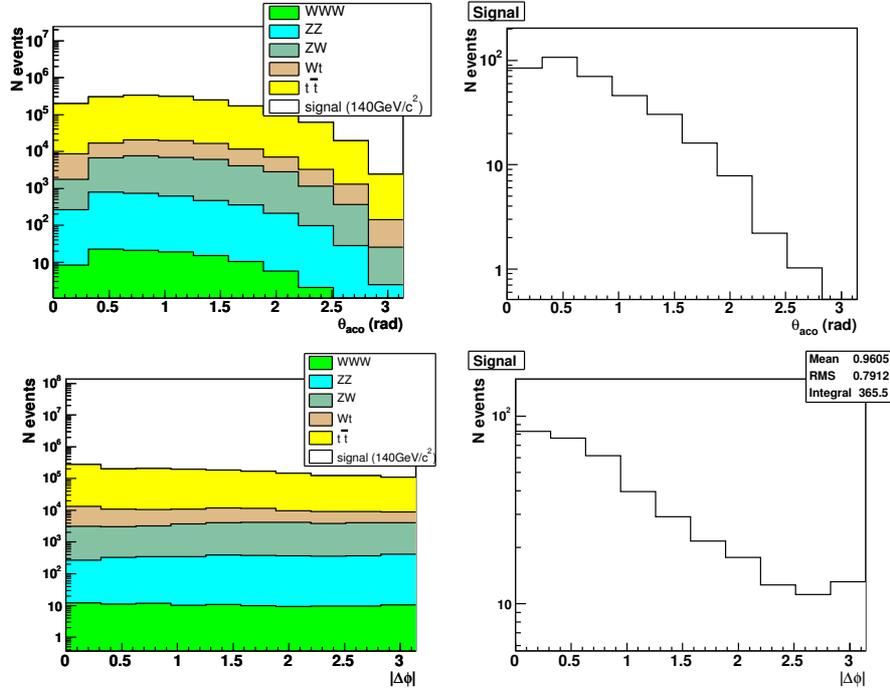


Figure 7.18: Angles between the two leptons assigned to the Higgs boson decay, for signal and background events. The upper plot shows the acollinearity, while the lower plot focuses on the acoplanarity. The normalization corresponds to  $100fb^{-1}$ .

the acoplanarity lower than 1.75 rad. These two cuts reduce the background by a factor 3.6, while the efficiency for the signal is 55%. One of the major effects of these cuts is to reduce the WWW background, which is otherwise similar to the signal. These two quantities are shown in Figure 7.18.

Most of the background contains a “fake” lepton produced in a jet, often from a semi-leptonic B meson decay. In such events, one of the leptons is less isolated. Different isolation methods have been considered. Isolation can rely either on tracker or on calorimeter information.

- The sum of the transverse momentum of charged particles associated to tracks reconstructed in the silicon tracker, in a cone around the lepton’s track, is computed. If the total  $P_T$  exceeds a given value, the lepton is said to be non-isolated.
- Another technique consists in the measurement of the angle between the lepton track and the closest track above a certain  $P_T$ .
- Using the calorimeter information, an isolation measure can be built by collecting all the energy in a cone around the lepton’s direction. If this sum exceeds a given threshold, the lepton is said to be non-isolated.

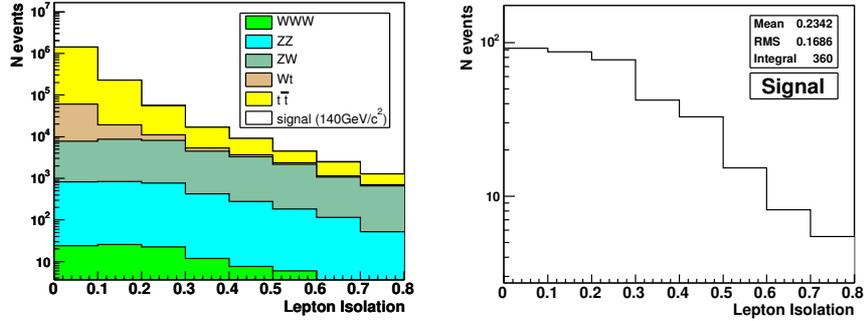


Figure 7.19: Lepton isolation by method (b) (as defined in the text). The distance (in  $\eta, \phi$ ) to the closest track above 3 GeV is shown for the less isolated lepton in any event. The normalization corresponds to  $100fb^{-1}$ .

In the present study, isolation method (b) has been chosen ( $I_L^b$ ). This isolation method allows to reach a large background rejection factor, for a reasonable signal efficiency (Figure 7.19). A typical cut corresponding to a cone radius of 0.2 gives a rejection factor of 95% for a signal efficiency of 49%. The other two methods have been shown to be interesting when the signal efficiency is more crucial than the background rejection, since the (signal) isolated leptons form a narrow distribution, but with an irreducible background contribution in the same region.

The isolation cut is complemented by a vertex-based selection. A cut on the z-spread of the reconstructed vertex is applied to ensure that all of the three leptons are coming from the same interaction. This quantity is defined as the largest difference between the z position of the impact point for each of the reconstructed leptons. Such a cut eliminates fake coincidences of leptons produced in pile-up interactions.

### 7.5.3 Jet veto

Signal events are ideally free of any hadronic activity, so that the presence of jets could be used to distinguish these leptonic events from the large  $t\bar{t}$  and  $Wt$  backgrounds with lots of hadronic activity. It is nevertheless seen in Figures 7.5 and 7.6 that remnants and pile-up induce significant activity in the calorimeters.

Central jets are first reconstructed using the iterative cone [61] or inclusive  $K_{\perp}$  [88] algorithms. The inclusive  $K_{\perp}$  algorithm relies on the definition of a resolution variable  $d_{ij}$  between two objects  $i$  and  $j$ . For the purpose of this work, this distance is defined as

$$d_{ij} = \min(P_{t_i}^2, P_{t_j}^2)((y_i - y_j)^2 + (\phi_i - \phi_j)^2). \quad (7.2)$$

Calorimeter tower combining ECAL and HCAL information are first put in a collection of “objects”. For each pair of objects, the resolution variable is computed, and closest objects are merged. If, for one object, the distance to the beam-pipe (defined as  $P_t^2$ ) is smaller than the distance between any pair of objects in the collection, the object is defined to be a jet, and removed from the collection of objects to be merged. This procedure is repeated until no object remains in the collection. The recombination scheme defines how two objects  $i$  and  $j$  are merged into a single object. In the  $E$  recombination scheme, a simple Lorentz-vector addition is performed:

$$P_{ij}^\mu = P_i^\mu + P_j^\mu, \quad (7.3)$$

which results in massive final-state jets. In the  $E_t$  recombination scheme, the transverse energy is summed:

$$\begin{aligned} E_{t(kl)} &= E_{t(k)} + E_{t(l)}, \\ \eta_{kl} &= \frac{\eta_k E_{t(k)} + \eta_l E_{t(l)}}{E_{t(kl)}}, \\ \phi_{kl} &= \frac{\phi_k E_{t(k)} + \phi_l E_{t(l)}}{E_{t(kl)}}, \end{aligned} \quad (7.4)$$

which results in massless final-state jets.

Four jet reconstruction methods have been tried:

1. Iterative cone with  $E_t$  recombination scheme, cone size 0.5
2. Iterative cone with  $E_t$  recombination scheme, cone size 0.7
3. Inclusive  $K_\perp$  jets with  $E_t$  recombination scheme
4. Inclusive  $K_\perp$  jets with  $E$  recombination scheme

Only jets above 10 GeV/c in the central region ( $|\eta| < 2.1$ ) are kept. Built from ECAL plus HCAL towers containing all the energy deposited in either of the two calorimeters, jets contain the energy deposited by any electron already identified. Ideally, the corresponding energy should be subtracted from the calorimeter prior to jet reconstruction, but such a procedure is not possible within the CMS software (yet). Hence, electrons are subsequently subtracted from reconstructed jets. Figure 7.20 shows the  $P_t$  spectrum of the leading central jet, before and after subtraction of electrons. This procedure clearly lowers the average hadronic activity reconstructed in the event.

Backgrounds like  $Wt$  and  $t\bar{t}$  are characterized by at least one or two additional jets from the leading-order interaction. We therefore reject a significant part of these events by cutting on the transverse momentum of the leading jet. The iterative cone with a cone size of 0.7 is the best algorithm to discriminate signal over background, as the difference between the energy of fake jets in signal events (from pile-up and calorimeter threshold effects) and jets in  $t\bar{t}$  or  $Wt$  backgrounds is more significant (Figure 7.21). Considering this cut alone, the best signal significance ( $s/\sqrt{b}$ ) is obtained for a cut at 25 GeV/c.

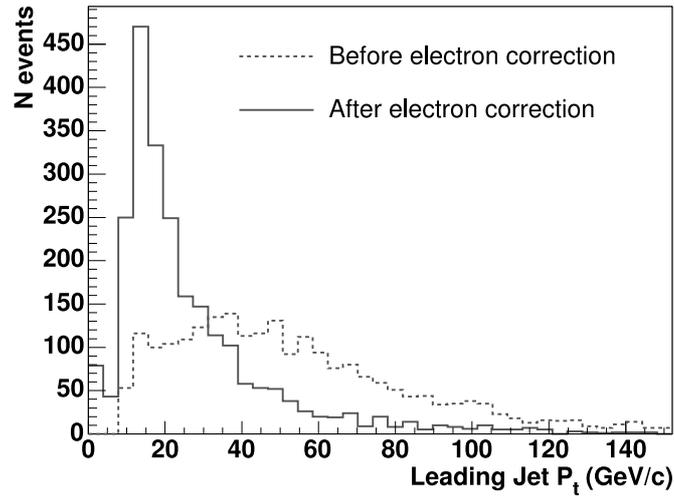


Figure 7.20:  $P_t$  spectrum of the leading central jet, before (dashed) and after (plain) subtraction of electrons for signal events with  $M_H = 140 \text{ GeV}/c^2$ . Jets are reconstructed using the iterative cone algorithm with the  $E_t$  recombination scheme and a cone size of 0.5. The normalization is arbitrary.

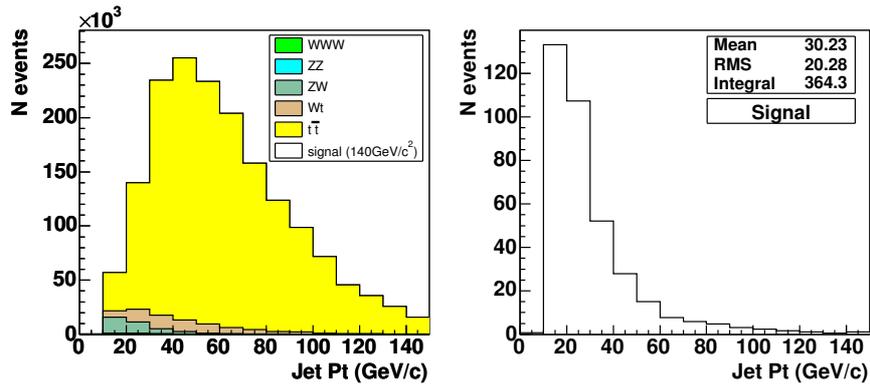


Figure 7.21: Leading jet  $P_t$  distribution, obtained using the iterative cone algorithm with a cone size of 0.7. The normalization corresponds to  $100fb^{-1}$ .

### 7.5.4 Identification of B mesons

In addition to the lepton isolation cut, one can try to directly identify jets from B decays, and to reject any tagged event. This procedure would ideally benefit from soft lepton B-tag algorithms not yet available. In this study, the *combined B-tag* algorithm has been used [87]. This algorithm combines impact parameter and secondary vertices measurements to build a probability for jets (built using the algorithm “1” just presented) to result from the decay of a B meson. Default parameters for this algorithm are tuned by the collaboration for the best efficiency, while keeping the fake identification rate low. Typical values are an efficiency of 60% for a fake identification rate of 2%.

We also tried the *track counting* algorithm with different sets of parameters. This was motivated by the following observation: for this study, where B-tag is used to reject background, the fake rate can be set to be higher and parameters re-tuned to privilege efficiency. It has been shown however that the retuned track counting algorithm was not better than the default combined B-tag algorithm. It could be interesting to perform a similar retuning for the combined B-tag algorithm.

Requiring no jet tagged as originating from a B meson by the combined b-tag algorithm, 95% of the signal is kept, while 77% of the  $t\bar{t}$  background is rejected.

### 7.5.5 Z veto

In ZZ and WZ backgrounds, there is a lepton pair that finds its origin in the Z boson decay. A cut on the invariant mass of any pair of leptons compatible with this hypothesis (via charge and flavor constraints) is used to reject these events. Figure 7.22 shows this invariant mass for ZZ, WZ, and signal events. Rejecting the masses between 65 and 115 GeV/c<sup>2</sup> clearly favors signal over background. Considering only WZ and ZZ backgrounds, these numbers are the optimal values for this cut alone. The fraction of background events kept is then 0.11 and 0.17, for WZ and ZZ backgrounds respectively, while signal efficiency amounts to 0.80.

### 7.5.6 Other topological cuts

After applying these cuts, the remaining background contains mainly WZ and ZZ events, where all three leptons are well reconstructed. Only more elaborate kinematical cuts can distinguish them from the signal.

One natural quantity that has not been considered so far is the missing transverse momentum ( $\cancel{E}_T$ ). Figure 7.8 shows that the missing transverse momentum is not very large. Its shape after all mentioned cuts (Figure 7.23) is nevertheless modified, and it can be used to remove part of the ZZ background, where no missing transverse momentum is expected. A typical cut

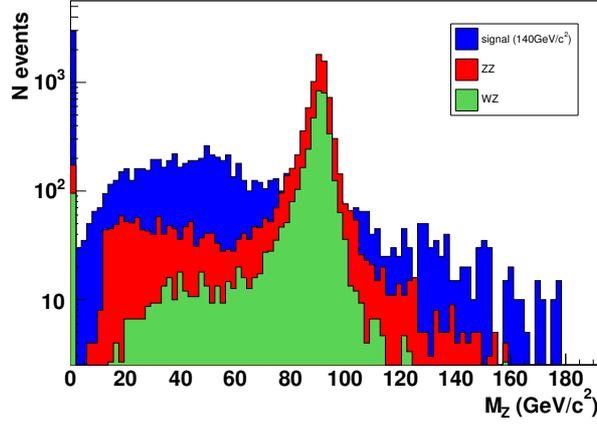


Figure 7.22: Invariant mass of any pair of leptons compatible with the Z hypothesis, for ZZ, WZ, and signal events. When more than one pair is possible, the pair with an invariant mass closest to  $M_Z$  is used. Events in the first bin are those for which no pair is compatible with the Z hypothesis. The normalization is arbitrary.

at 50 GeV/c is applied for that purpose. Applied after all already mentioned cuts, it removes 76% of the ZZ background while keeping 73% of the signal.

The  $\cancel{E}_T$  can then be combined with the transverse momentum of leptons to build the transverse mass of the would-be W bosons, defined as

$$M_T(W_i) = \sqrt{2 * P_T^i \cancel{E}_T (1 - \cos \Delta\phi_{l\cancel{E}_T})}, \quad (7.5)$$

where  $P_T^i$  is the transverse momentum of one of the three leptons,  $\cancel{E}_T$  is the missing transverse momentum, and  $\Delta\phi_{l\cancel{E}_T}$  the polar angle between the lepton and the missing transverse momentum. By definition, this transverse mass is lower than the W or Z boson mass, respectively when  $\cancel{E}_T$  results from a single neutrino (WZ events) or a missed lepton (ZZ events). For signal events, a significant fraction of energy is carried by the three neutrinos. The reconstructed transverse mass is therefore expected to be larger. The reconstructed transverse mass from Equation (7.5) is shown Figure 7.24.

Similarly, the Higgs boson transverse mass is computed from the two chosen leptons and from the missing transverse momentum:

$$M_T(H) = \sqrt{M_T^{ll^2} + 2E_T^{ll} \cancel{E}_T - 2P_T^{ll} \cancel{E}_T \cos \Delta\phi_{ll\cancel{E}_T}}, \quad (7.6)$$

where  $M_T^{ll}$  and  $E_T^{ll}$  are here respectively the transverse mass and the transverse momentum of the lepton pair, and  $\Delta\phi_{ll\cancel{E}_T}$  is the polar angle between the missing transverse momentum and the summed momentum of the lepton pair. This is only an approximation as the third neutrino, originating from the associated W boson, cannot be subtracted from the missing momentum. It is

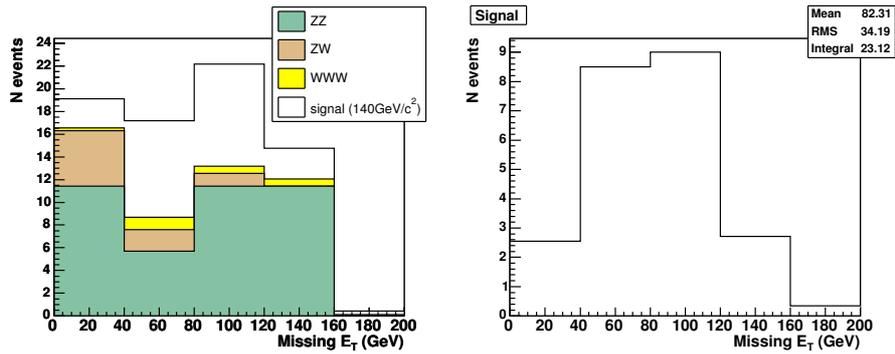


Figure 7.23: Missing transverse energy for signal and background events, after all cuts but the topological cuts (see text). The normalization corresponds to  $100fb^{-1}$ .

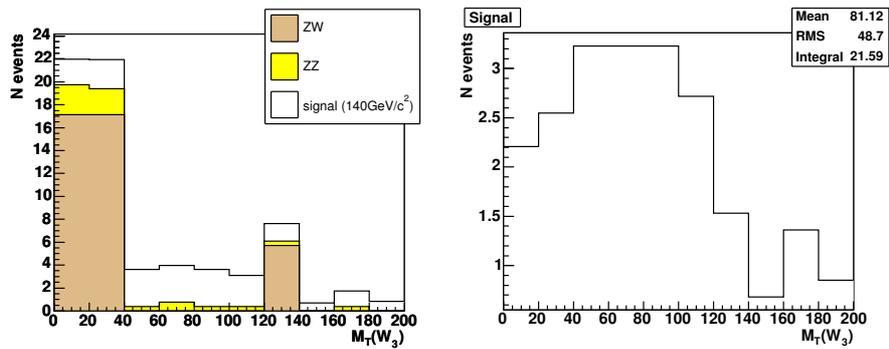


Figure 7.24: Reconstructed  $W$  mass  $M_T(W_3)$  (from Equation (7.5)) for signal and background events, after all cuts but the topological cuts (see text).  $ZW$  events are clearly characterized by a low transverse mass. The normalization corresponds to  $100fb^{-1}$ .

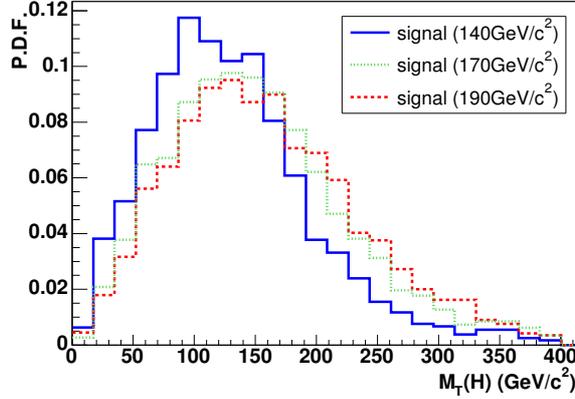


Figure 7.25: Reconstructed Higgs boson transverse mass from Equation (7.6). The normalization corresponds to  $100fb^{-1}$ .

nonetheless correlated to the Higgs boson mass and it will be used as a discriminant variable in the evaluation of likelihood ratios. This evaluated Higgs boson transverse mass is shown Figure 7.25.

Quantities like aplanarity, sphericity or thrust, constructed either from the calorimeter information or from tracks of charged particles, were also considered. No discriminating power has been observed. In order to further protect against the effect of Z+jets background that is not considered, the ATLAS cut on the transverse momentum of the summed Lorentz vector of the three leptons ( $P_t(\sum P_i)$ ) is used. A cut at  $40 \text{ GeV}/c$  on this quantity is also found to be sensible to reduce the ZZ background, as it can be seen in Figure 7.26.

### 7.5.7 Summary of optimized selection cuts

Optimized cuts are summarized in Table 7.3. Evolution of efficiency with the selection for the signal and each background source is represented in Figure 7.27. The starting point corresponding to 100% efficiency is chosen to be all events passing the HLT. Figure 7.28 shows the cumulated efficiency (including trigger and event selection) as a function of the Higgs boson mass. The curve has a maximum at the WW resonance. Beyond the WW production threshold, efficiency drops since W bosons start to be boosted in the Higgs boson frame, which influences the angular distribution of leptons. Efficiency in that region could certainly be improved by optimizing the analysis for a Higgs boson mass of  $190 \text{ GeV}/c^2$ . The mass-dependence of the selection efficiency varies from 0.5% to 1.3%.

We can now have an insight into the trigger patterns contributing with highest efficiency after offline selection. Before selection, the dominating triggers were the single electron and the single muon, with a smaller contribution

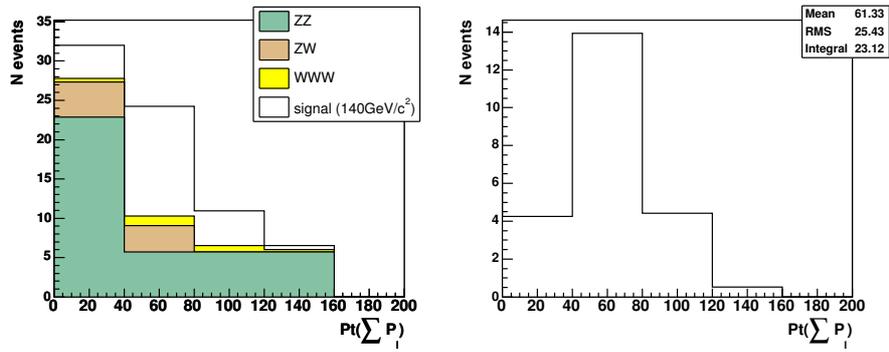


Figure 7.26: Transverse momentum of the summed Lorentz vector of the three leptons for signal and background events, after all cuts but the topological cuts (see text). The normalization corresponds to  $100fb^{-1}$ .

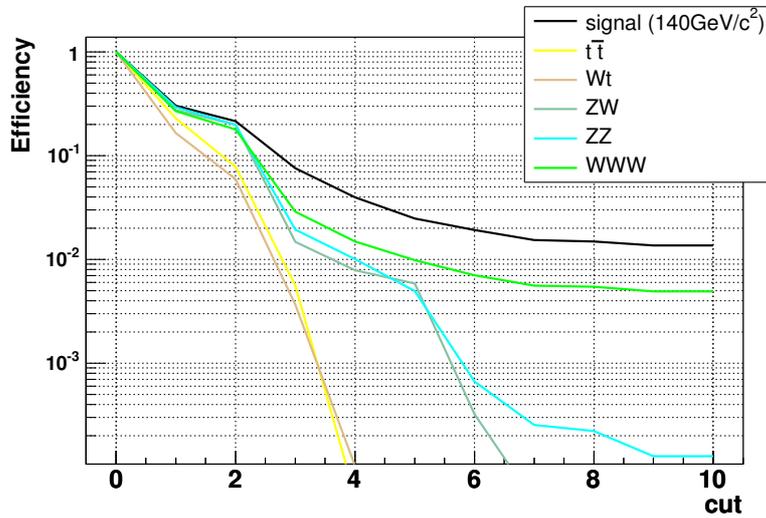


Figure 7.27: Evolution of efficiency with the selection for the signal and each background source. The starting point corresponding to 100% efficiency is chosen to be all events passing the trigger. Cut numbering used is the same as in Table 7.3.

Table 7.3: Summary of the optimized selection cuts. The cross-section for the signal and backgrounds, for each step in the selection, is given in fb. An upper limit for the  $Wt$  and  $t\bar{t}$  cross-sections is given when no simulated event remains.

Id.	Cut Type	Selection	Signal (fb) 140 GeV/c <sup>2</sup>	Background (fb)				
				$t\bar{t}$	$Wt$	ZW	ZZ	WWW
0	Trigger	L1 $\wedge$ HLT	12.24	72067	4115.8	1238.4	118.438	3.91
1	Reconstruction	$N_{lept} = 3$ $ \sum charge  = 1$	3.81	16432.7	680.0	339.4	34.65	1.05
2	Lepton cuts for electrons:	$\Delta_Z < 0.15cm$ $\Delta_R < 0.015cm$ $Pt(l_i) > 16 GeV/c$ $0.9 < E/P < 1.75$ $H/E < 1.5$	2.67	5629.1	245.3	245.9	23.53	0.70
3	Angular cuts	$ \Delta\phi  < 1.75rad$ $0.1 <  \Delta\Omega (rad) < 0.75$	0.87	400.6	15.0	18.3	2.29	0.11
4	B veto	$N_{b-jet} = 0$ $I_b^* > 0.17$	0.43	3.85	0.42	9.77	1.19	0.06
5	Jet veto	$Pt_{jet1} < 25 GeV/c$	0.27	< 1.93	0.31	7.26	0.58	0.04
6	Z veto	$!(65 GeV/c^2 < M_Z < 115 GeV/c^2)$	0.21	< 1.93	0.21	0.40	0.08	0.03
7-9	Topological	$M_T(W_3) > 40 GeV/c^2$ $Pt(\sum P_i) > 40 GeV/c$ $\cancel{E}_T > 30 GeV$	0.13	< 1.93	< 0.11	0.06	0.01	0.02
10	Subtrigger	$e, ee, \mu, \mu\mu$	0.13	< 1.93	< 0.11	0.06	0.01	0.02

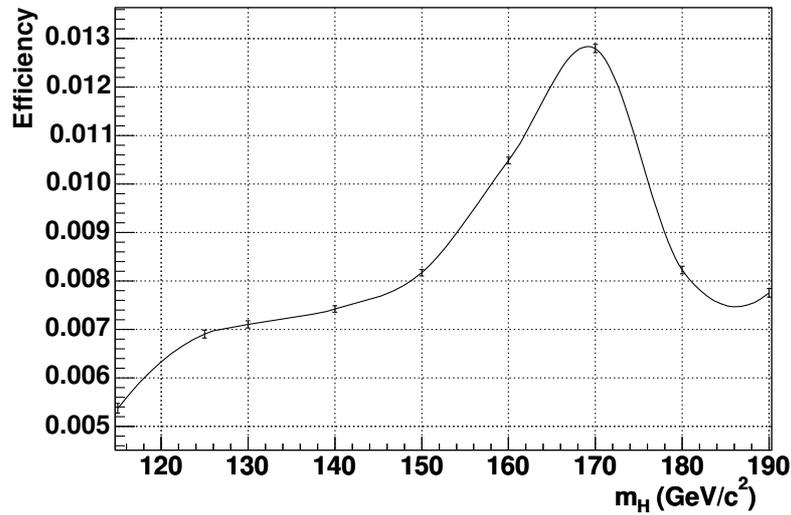


Figure 7.28: Cumulated efficiency (including trigger and event selection) as a function of the Higgs boson mass.

from double triggers. After selection, the situation is more or less identical to what was presented in Section 7.4. Even if the impact of multi-leptonic patterns is enhanced, as seen in Figure 7.29, double triggers are less predominant, as leptons are required to be above the single-lepton threshold during the selection procedure. The double-muon trigger is not needed anymore, while the double-electron trigger alone only represents 1% of the selected events. The choice of subtriggers to be considered for this analysis therefore includes single-electron and single-muon triggers. Both double-electron and double-muon triggers should also be kept, since their importance could be highly dependent on the selection cuts on the leptons' transverse momentum.

It must also be noted that the trigger process and the subsequent reconstruction cuts have privileged muonic events, as muons are easier to detect and isolate (Figure 7.30). When a ratio of 1.1 is originally expected after HLT between fully muonic and electronic classes, this ratio is 1.6 after selection. It could be used to check the coherence of any signal observed. Tau channels are suppressed, as only leptonic tau decays are considered in the analysis. Leptonic tau decays also present a slightly softer electron and muon spectrum due to the additional neutrinos.



## 7.6 Systematic uncertainties

Systematic sources considered in this study are related to the normalization of backgrounds, to the reconstruction, the event selection, the luminosity and the structure functions of protons.

One way to roughly evaluate *a priori* the uncertainty on the background normalization is to compare leading-order and next-to-leading-order calculations. For events of interest for this study, these differences are large: 45% for  $t\bar{t}$ , 40% for WZ and 25% for ZZ events. This is a very pessimistic way to estimate the uncertainties, but the scale-dependence of the NLO cross-sections, usually used for this purpose, is not known. Such large uncertainties are certainly not affordable and must be overcome. Before tuning them with real data, and thus especially at low integrated luminosity, it will be impossible to trust background estimates from Monte-Carlo programs. Hence, background will be normalized to signal-free regions of the phase-space, or to measurements performed independently by the CMS collaboration. By slightly modifying the selection cuts, it is possible to enrich some backgrounds, while staying in a phase-space close to the one considered previously. This is particularly true for WZ and  $t\bar{t}$ . The WZ background can be measured in an almost signal-free region by using cuts 0, 1, 2, 4, 5 and 10, and by requiring a pair of leptons compatible with a Z boson hypothesis, and having an invariant mass between  $65 \text{ GeV}/c^2$  and  $115 \text{ GeV}/c^2$ . For  $100 \text{ fb}^{-1}$ , 500 WZ events are expected with a very good purity (Figure 7.31(a)), so that an accuracy of 5% can be obtained on this background level. There is no hope to measure accurately the ZZ background while keeping only three (reconstructed) lepton final states, as it will always remain one order of magnitude lower than WZ. Imposing a small missing transverse momentum reduces the WZ contribution, but not in a sufficiently significant way. We will have to rely on measurements from events with four reconstructed leptons, which is expected to be very accurately done in the context of the  $H \rightarrow ZZ$  analysis. The  $t\bar{t}$  background is easily measured by using cuts 0, 1, 2, 6, 7, 8, 9 and 10. For  $100 \text{ fb}^{-1}$ , 19485 events are obtained, with a small contamination from  $Wt$  events (Figure 7.31(b)). This contamination can be reduced by requiring two b-tagged jets. For  $100 \text{ fb}^{-1}$ , 4643 events are then obtained, with 99.5% of  $t\bar{t}$  and only 0.5% of  $Wt$  events. It allows in principle to measure the  $t\bar{t}$  background with an accuracy of 2%. None of the two remaining backgrounds (WWW and  $Wt$ ) can be measured accurately. The WWW background always remains one order of magnitude smaller than the signal. As demonstrated, it remains as the only dominant background after the analysis, which induces some concerns. The signal is nevertheless higher than the theoretical uncertainties. Finally, a detailed study of the  $Wt$  background is certainly needed to control the effect of this process. Such a study is being considered within CMS by the Standard-Model group in the one- or two-lepton channels. In the following, an uncertainty of 5% on the background level will be considered.

Reconstruction and selection uncertainties mainly arise from the jet veto,

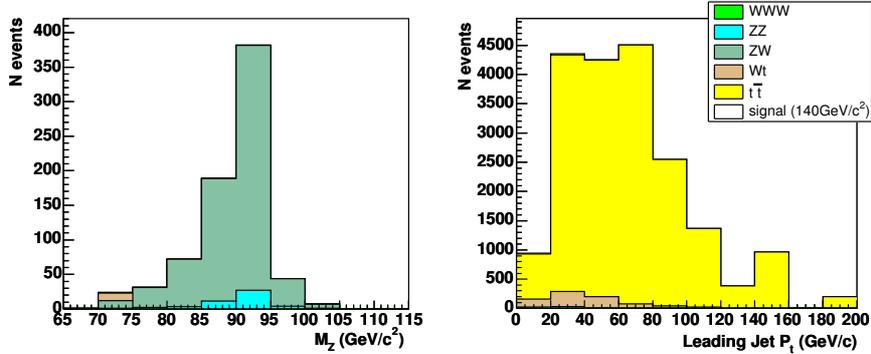


Figure 7.31: (left) Invariant mass of the pair of leptons compatible with the Z hypothesis for a set of cuts tuned to enhance the WZ background;(right) distribution of the leading jet  $P_t$  for a set of cuts tuned to enhance the  $t\bar{t}$  background.

the b veto and lepton reconstruction. Experience from Tevatron tells us that a typical 2% uncertainty on lepton reconstruction efficiency has to be considered, while 3% uncertainty comes from lepton isolation. Since three leptons are present in our analysis, this gives 8% uncertainty from lepton reconstruction and selection. The additional uncertainties from the jet veto and the b veto will be assumed to be 5% each, which also seems reasonable in the light of Tevatron data.

The last considered uncertainty comes from the product of the luminosity and the proton structure functions, known as the parton luminosity. Considering these two quantities separately, 5% uncertainty on the luminosity is generally assumed, while the uncertainty from the proton PDF is taken to be 4% [89]. This latter uncertainty is reduced by the process considered, for which the mid-x region (where uncertainties are small) dominates. This gives an uncertainty of 6.4% on the parton luminosity. Since this parton luminosity is the actual quantity involved in the interaction physics, an alternative approach consists in measuring it directly [90]. The electron and muon pseudorapidity distributions, originating from the decay of weak bosons, contains information about both the PDF and the proton luminosity. It provides a key to measure the parton luminosity with an accuracy of about 1%. It nevertheless requires to extrapolate probability density functions to lower x and higher  $Q^2$ , which may induce additional uncertainties. Other emerging ideas cope with the (proton) luminosity issue by measuring new well predicted processes. Using elastic scattering in the Coulomb region (which requires instrumentation of the beam with Roman pots at 240m)[91] or photoproduction of muon pairs[92], an uncertainty of 1-2% is predicted. In the present work, a conservative uncertainty of 6.4% on the parton luminosity is considered, keeping in mind that this can probably be improved. Considered uncertainties are summarized in Table 7.4.

Table 7.4: Levels of systematic uncertainties considered for this study.

Source	Level
Parton Luminosity	6.4%
Background normalization	5%
Lepton reconstruction	8%
Signal selection	2%
Jet veto	5%
B tag	5%
Total	14%

## 7.7 Results

### 7.7.1 Observability

Following the recommendations of the CMS collaboration, the signal significance after selection is estimated using the  $S_{L2}$  estimator from [93]:

$$S_{L2} = \sqrt{2 \ln Q}, \text{ with } Q = \left(1 + \frac{N_s}{N_b}\right)^{N_{obs}} e^{-N_s}. \quad (7.7)$$

Setting the expectation value of  $N_{obs}$  to  $N_s + N_b$ , a likelihood estimator based on event counting is obtained. This is the simplest method that can be used to evaluate signal significance in the presence of less than about 50 background events, when the central limit theorem does not hold. The resulting significance curve is presented in Figures 7.32 and 7.33. For an integrated luminosity of  $100 fb^{-1}$ , a the signal can be observed with a  $5\sigma$  significance for masses ranging from  $154 \text{ GeV}/c^2$  to  $176 \text{ GeV}/c^2$ . For an asymptotic integrated luminosity of  $300 fb^{-1}$ , the whole range from  $126 \text{ GeV}/c^2$  to  $188 \text{ GeV}/c^2$  is covered by the analysis.

This estimator does not take into account the effect of systematic uncertainties and does not exploit the discriminative power from the transverse mass distribution (Figure 7.34). In order to integrate both aspects, the log-likelihood method introduced in Section 4.4 is used. Figure 7.35(a) shows the luminosity needed to obtain a  $5\sigma$  significance using this method, with systematics only, with Monte-Carlo statistical uncertainties, or with both effects considered. The main effect comes from the limited Monte-Carlo statistics. Compared to Figure 7.33, the needed luminosity is found to be similar in the most favorable mass region, while it increases faster at low and high masses, where the number of expected signal events decreases. This behavior is expected, as small-statistic signals are more easily mixed up with background fluctuations related to the systematic uncertainties. Figure 7.35(b) shows the luminosity needed to exclude a Higgs boson at 95% C.L. if no excess is observed, using the same method. Less than  $50 fb^{-1}$  are required in most of the

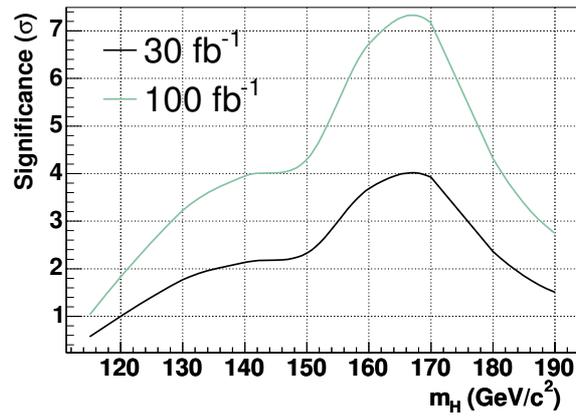


Figure 7.32: Significance of the signal over background, obtained using formula (7.7) as a function of the Higgs boson mass. Significance is shown for an integrated luminosity of  $30 \text{ fb}^{-1}$  and  $100 \text{ fb}^{-1}$ .

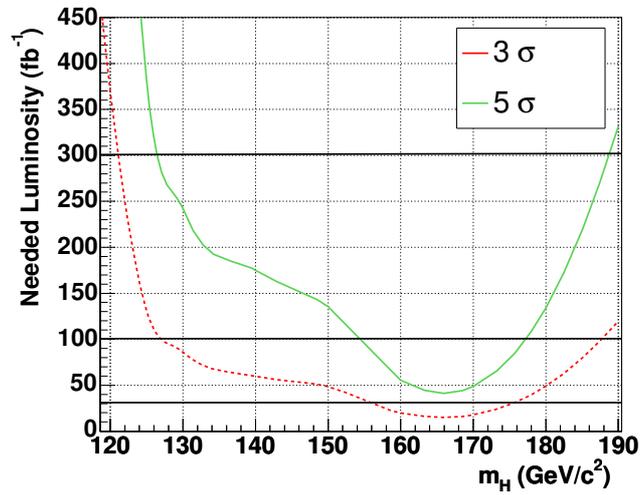


Figure 7.33: Luminosity necessary to obtain a  $3\sigma$  or  $5\sigma$  excess as a function of the Higgs boson mass.

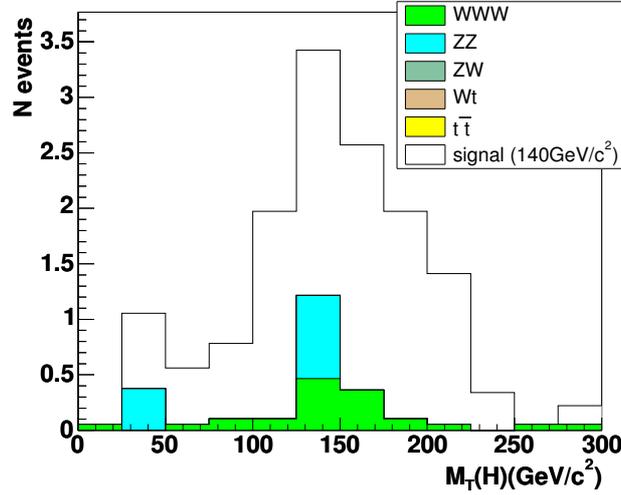


Figure 7.34: Reconstructed transverse mass from Equation (7.6) for a  $140 \text{ GeV}/c^2$  Higgs boson.

mass range, while only  $20 \text{ fb}^{-1}$  are needed at  $170 \text{ GeV}/c^2$ .

### 7.7.2 Cross-section and branching ratio

If a new physical state is observed by a Higgs boson search, it will be crucial to study the properties of this signal to further check whether this is a Higgs boson or not, and to which symmetry-breaking sector it corresponds. One of the best ways to perform such a test is to measure and compare the branching ratios of the particle. These branching ratios can also allow to distinguish Higgs bosons from different models. Non-minimal Higgs models often predict different coupling constants, depending on the models' parameters.

After the measurement of the rate in one channel, a precise theoretical prediction is needed to determine the branching ratio. Such a precise prediction is often unavailable because of uncertainties related to the proton PDF or to NLO effects. Two types of approaches have been considered in the literature. A model-independent analysis where ratios of couplings are extracted has been performed by the ATLAS collaboration [94]. In this context, the channel studied here plays a key role in the determination of the ratio  $\Gamma_W/\Gamma_t$ , which gives an indirect insight into the top coupling in a Higgs boson mass window where it does not dominate:

$$\frac{\sigma \times \text{BR}(\text{WH} \rightarrow \text{WWW})}{\sigma \times \text{BR}(\text{H} \rightarrow \text{WW})} = \frac{\Gamma_W \Gamma_W}{\Gamma_g \Gamma_W} * C_{QCD} \propto \frac{\Gamma_W}{\Gamma_t} * C_{QCD}, \quad (7.8)$$

where  $C_{QCD}$  is the ratio of the K-factors between associated and direct pro-

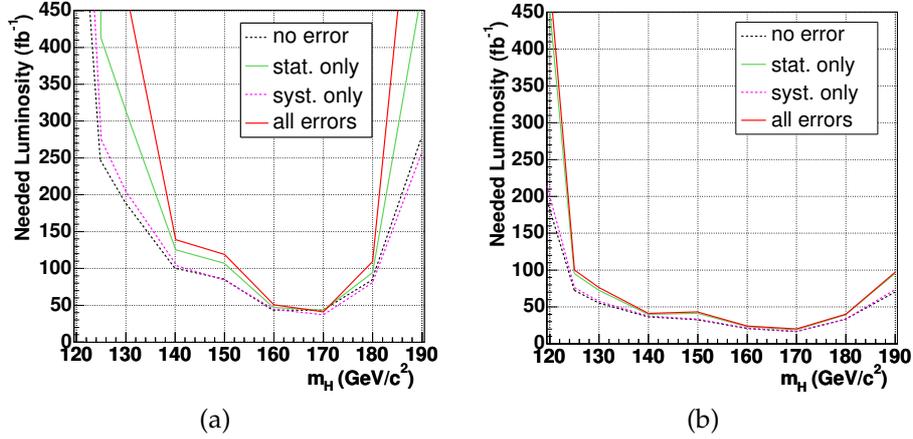


Figure 7.35: (a) Luminosity needed to obtain a  $5\sigma$  significance using the method from Section 4.4, with systematics only, Monte-Carlo statistical uncertainties only, or with both effects considered; (b) luminosity needed to exclude a Higgs boson at 95% C.L. if no excess is observed, using the same method.

duction. Another approach consists in imposing theoretically motivated constraints on the  $g_{HWW}$  and  $g_{HZZ}$  couplings to extract directly all Higgs boson couplings as well as the total width. This approach that involves a combined study of all LHC analyses has been considered in reference [95]. In both cases, the uncertainty on the total width after  $100fb^{-1}$  is found to be of 20%.

In the present study, the partial decay width of the Higgs boson into W bosons,  $\Gamma_W$ , is determined with a good accuracy. This is possible using the narrow width approximation and benefiting from the fact that only the  $g_{HWW}$  coupling is involved in the Feynman diagram for the signal. This shows the potential and the importance of this channel in the global approaches considered in the two mentioned papers.

From the number of observed events, and assuming a good description of the background by the Monte-Carlo simulations, the product of the signal cross-section and the branching ratio of the Higgs boson into WW can be computed as:

$$\sigma \times BR = \frac{N_s}{\varepsilon \mathcal{L}} = \frac{N - N_b}{\varepsilon \mathcal{L}}, \quad (7.9)$$

with  $\varepsilon$  representing the combined trigger and selection efficiency. The uncertainty on this measurement is then found to be

$$\left( \frac{\Delta(\sigma \times BR)}{\sigma \times BR} \right)^2 = \frac{(\Delta N)^2}{N_s^2} + \frac{(\Delta N_b)^2}{N_s^2} + \frac{(\Delta \mathcal{L})^2}{\mathcal{L}^2} + \frac{(\Delta \varepsilon)^2}{\varepsilon^2}. \quad (7.10)$$

In this expression,  $\Delta N$  corresponds to the statistical uncertainty on  $N$ , and will

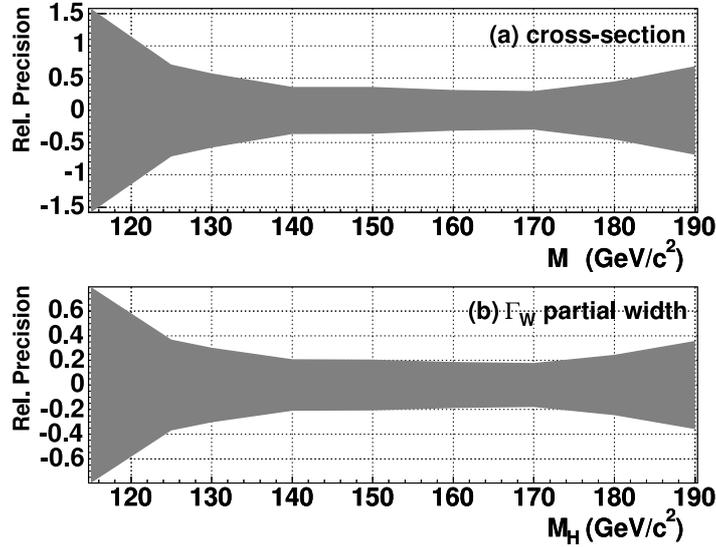


Figure 7.36: (a) Resolution on the product of cross-section and branching ratio into  $WW$  as a function of the Higgs boson mass, for an integrated luminosity of  $100fb^{-1}$ . (b) Resolution on the  $\Gamma_W$  partial width as a function of the Higgs boson mass, for the same integrated luminosity.

be assumed to be  $\sqrt{N}$ . The other three  $\Delta N_b$ ,  $\Delta \mathcal{L}$  and  $\Delta \epsilon$  quantities are dominated by systematic error sources described in Section 7.6.

In the narrow width approximation, the measured product of signal cross-section and branching ratio of the Higgs boson into  $WW$  can also be expressed from partial and total widths of the Higgs boson:

$$\sigma \times BR = \frac{\sigma^{SM} \Gamma_W}{\Gamma_W^{SM}} \times \frac{\Gamma_W}{\Gamma_{tot}} \equiv T \Gamma_W^2. \quad (7.11)$$

This allows us to write the uncertainty on the Higgs boson width into  $W$  bosons as

$$\left( \frac{\Delta \Gamma_W}{\Gamma_W} \right)^2 = \frac{1}{4} \left( \frac{(\Delta \sigma \times BR)^2}{(\sigma \times BR)^2} + \frac{(\Delta T)^2}{T^2} \right). \quad (7.12)$$

In this expression, the uncertainty on the cross-section is the quantity obtained in Equation (7.10), and the uncertainty on the  $T$  factor comes from theoretical uncertainties and expected experimental uncertainties on  $\Gamma_{tot}$ . Assuming  $100fb^{-1}$  luminosity and the aforementioned systematic sources, the uncertainty on the cross-section is found to be 36% for a  $150 \text{ GeV}/c^2$  Higgs boson. The corresponding uncertainty on the  $\Gamma_W$  partial width is 21%. These results depend on the Higgs boson mass, and are presented Figure 7.36.

The uncertainty on the cross-section can be compared with the uncertainty predicted by the ATLAS collaboration in [94]. This paper predicts an uncertainty on the cross-section of 33%, again for a Higgs boson of  $150 \text{ GeV}/c^2$  after  $100fb^{-1}$ . This is apparently better than the result that has been obtained here, but in fact only reflects a less conservative estimation of systematics. The quoted number of signal and background events (15.1 and 9.9 respectively for a  $150 \text{ GeV}/c^2$  Higgs boson after  $100fb^{-1}$ ) is indeed very similar to our results (14.5 and 9.0 respectively in the same conditions). The uncertainty obtained on the  $\Gamma_W$  partial width can be compared to the uncertainty on the same partial width obtained in the compilation of ATLAS studies [95]. This paper predicts an uncertainty on the partial width of 20%, for a Higgs boson of  $150 \text{ GeV}/c^2$  after  $100fb^{-1}$ , which is in perfect agreement with our results.

### 7.7.3 Mass resolution

Since there are three neutrinos in the event, no mass peak can be extracted over the background. The most sensible variable to determine the mass is the Higgs boson transverse mass computed in Equation (7.6). Unfortunately, the mass dependence of this quantity is not marked. In order to extract the mass and the resolution, a likelihood fit must be performed. This method is inspired by the work presented in reference [94], but differs as binned probability density functions are used here.

For a given mass, a pseudo-data distribution is generated from the binned Monte-Carlo probability density function after analysis. Inspired by Poisson statistics, the value of the likelihood for each mass hypothesis is then computed as

$$-2\ln(\mathcal{L}) = 2 \sum_{i=1}^N n_i - d_i + d_i \ln\left(\frac{d_i}{n_i}\right), \quad (7.13)$$

where the sum is performed over the bins of the transverse mass distribution,  $d_i$  are the pseudo-data, and  $n_i$  the predicted number of events. The obtained log-likelihood distribution is fitted by a second order polynomial around the minimum, and the minimum of the parabola is taken as mass estimate. This procedure is repeated several times to produce a mass distribution, the width of which gives the mass resolution (Figure 7.37).

Due to the lack of statistics, the mass cannot be determined from a low integrated luminosity. Using  $100fb^{-1}$ , the reconstructed Higgs boson mass distribution is not Gaussian and presents a large tail. While the fitted width is  $2.5 \text{ GeV}/c^2$ , the RMS is  $5.4 \text{ GeV}/c^2$ . Using  $300fb^{-1}$ , a resolution ranging from  $2 \text{ GeV}/c^2$  to  $3 \text{ GeV}/c^2$  can be obtained. This procedure cannot be extended towards the edge of the mass region of interest, since samples with a simulated mass on each side of the studied mass are needed. It would require unavailable samples with a mass between  $100 \text{ GeV}/c^2$  and  $115 \text{ GeV}/c^2$ , as well as between  $190 \text{ GeV}/c^2$  and  $210 \text{ GeV}/c^2$ . The resolution obtained as a function of the Higgs boson mass is shown in Figure 7.38.

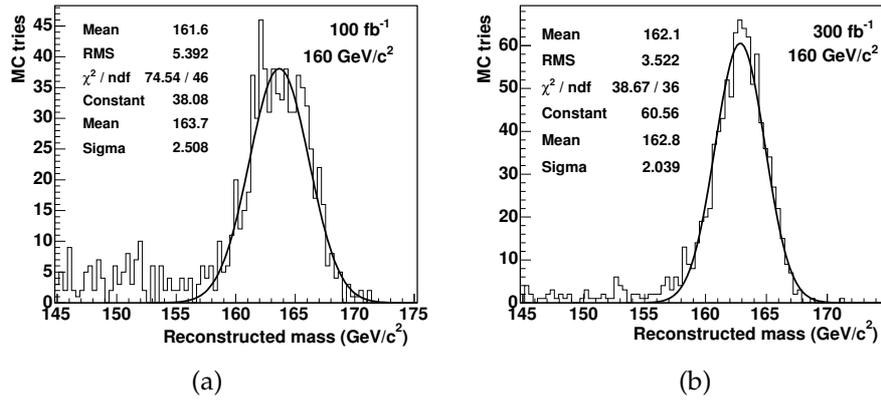


Figure 7.37: Mass distribution obtained by minimizing the Higgs boson (simulated at  $160 \text{ GeV}/c^2$ ) transverse mass likelihood function for several pseudo-experiments, for  $100 \text{ fb}^{-1}$  and  $300 \text{ fb}^{-1}$ .

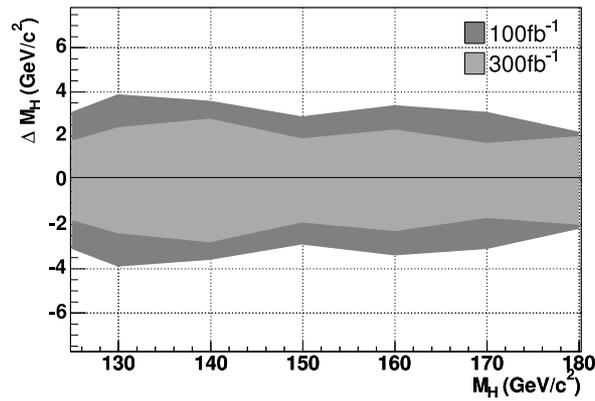


Figure 7.38: Resolution obtained on the Higgs boson mass in the range covered by the  $WH, W \rightarrow WW$  channel, after  $100 \text{ fb}^{-1}$  and  $300 \text{ fb}^{-1}$ . After  $100 \text{ fb}^{-1}$ , the uncertainty has to be considered with caution, as important tails are observed.

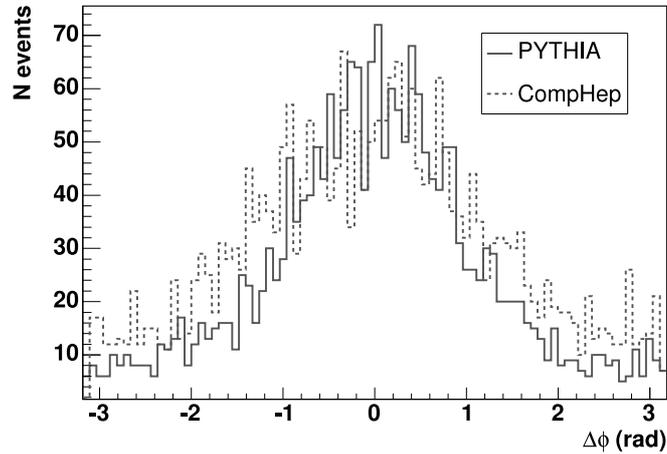


Figure 7.39: Reconstructed polar  $\phi$  angle between leptons originating from the Higgs boson for events simulated with PYTHIA (continuous) and CompHep (dashed). The normalization is arbitrary.

#### 7.7.4 Higgs boson spin

As already mentioned, reconstructed leptons are assigned either to the Higgs boson decay or to the decay of the  $W$  boson that does not come from the Higgs boson by choosing the two closest opposite-sign leptons. The third lepton must then come from the associated  $W$  boson. One drawback of this way of assigning leptons to the Higgs boson decay is that it biases the reconstructed angle between leptons. Figure 7.39 shows the reconstructed acoplanarity between leptons originating from the Higgs boson for events simulated with PYTHIA and CompHep. Distributions are very similar, which is rather different from what was obtained at generator level in Figure 7.9. Several alternative methods to select leptons from the Higgs boson decay have been tried but have not been more successful. These are the  $P_i$  ordering, the angular position with respect to the missing transverse momentum or the  $\eta$  ordering. Other methods always introduce additional smearing that washes the spin-correlation effect out. Hence, the acoplanarity between leptons can probably not be used to measure the Higgs boson spin, as it is the case for the inclusive  $H \rightarrow WW$  (via gluon fusion) channel.

### 7.8 Fermiophobic Higgs boson

One important motivation for studying this channel is also that it is one of the only allowed signatures for a fermiophobic Higgs boson model, as we already

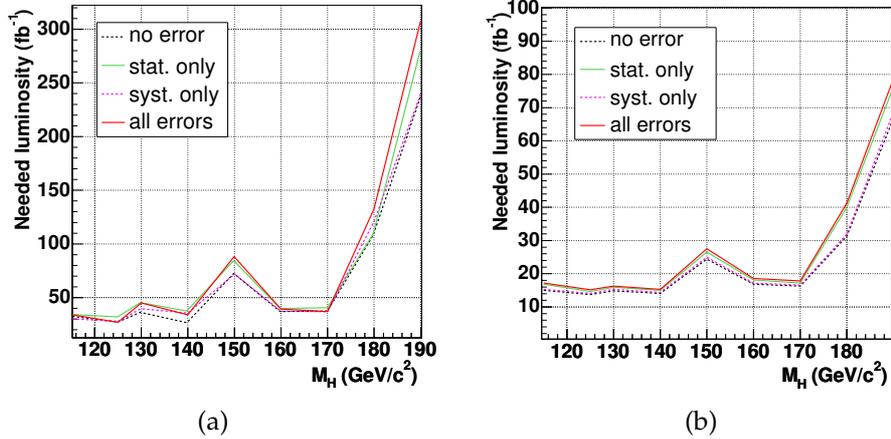


Figure 7.40: Results obtained using the benchmark fermiophobic model; (a) Luminosity needed to obtain a  $5\sigma$  significance using the method from Section 4.4, with systematics only, Monte-Carlo statistical uncertainties only, or with both effects considered; (b) luminosity needed to exclude a Higgs boson at 95% C.L. if no excess is observed, using the same method.

considered in Chapter 4. If the Higgs boson does not couple to fermions, the usual gluon-fusion diagrams are indeed forbidden, as well as  $b\bar{b}$  decays. A fermiophobic Higgs boson will present a large cross-section at low mass, as the branching ratio does not drop down as in the Standard Model.

Figure 7.40(a) shows the luminosity needed to obtain a  $5\sigma$  significance for a fermiophobic Higgs boson. Compared to Figure 7.35, the needed luminosity is found to be similar in the most favorable mass region for the Standard Model (around 170  $\text{GeV}/c^2$ ) and above, but far better results are obtained in the low mass region. After  $100\text{fb}^{-1}$ , all masses between the LEP limit and 175  $\text{GeV}/c^2$  will be covered by this analysis alone. Figure 7.40(b) shows the luminosity needed to exclude a fermiophobic Higgs boson at 95% C.L. if no excess is observed. Less than  $30\text{fb}^{-1}$  are required to reject any fermiophobic Higgs boson up to 175  $\text{GeV}/c^2$ . It must be noted that the presence of a Standard-Model Higgs boson would compromise this exclusion limit.

Since the number of expected events is larger at low mass, the expected precision on the partial branching ratio into W bosons is better. Assuming a  $100\text{fb}^{-1}$  luminosity and the aforementioned systematic sources, the uncertainty on the cross-section is found to be 30% for a 150  $\text{GeV}/c^2$  Higgs boson. The corresponding uncertainty on the  $\Gamma_W$  partial width is 20% (Figure 7.41). At this level, the uncertainty is dominated by the uncertainty on the total Higgs boson width.

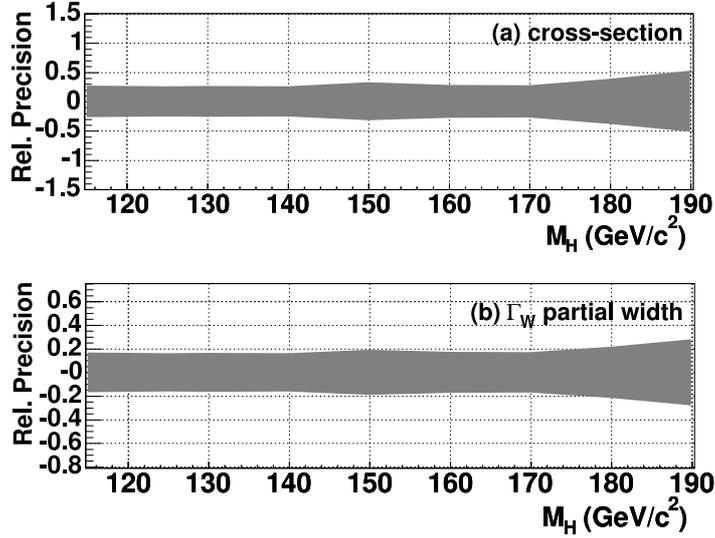


Figure 7.41: Results obtained using the benchmark fermiophobic model; (a) Resolution on the product of cross-section and branching ratio into  $WW$  as a function of the Higgs boson mass, for an integrated luminosity of  $100fb^{-1}$ . (b) Resolution on the  $\Gamma_W$  partial width as a function of the Higgs boson mass, for the same integrated luminosity.

## 7.9 Concluding remarks

The associated production channel  $WH$ , with  $W \rightarrow l\nu$  and  $H \rightarrow WW \rightarrow l\nu l\nu$ , provides an interesting possibility to observe a Standard-Model Higgs boson in the intermediate mass range. For an integrated luminosity of  $100fb^{-1}$  a significance of more than  $5\sigma$  can be observed in the mass interval between 155 and  $178 \text{ GeV}/c^2$ . This channel therefore complements a Higgs boson discovery in other channels. In that mass range, these are ( $qqH$ , with subsequent  $H \rightarrow \tau\tau, \gamma\gamma$ ), ( $H \rightarrow \gamma\gamma$ ), ( $H \rightarrow WW$ ) and ( $H \rightarrow ZZ$ ) [78]. While it is generally assumed to be a channel used after a discovery to determine Higgs boson properties, the present analysis shows that it can also be considered as a discovery channel in the center of the considered mass range. Performances compare with those in  $qqH$  and  $H \rightarrow \gamma\gamma$  channels. Compared to the  $WW$  channel, the associated production channel studied here has the advantage that a better signal to background ratio can be obtained due to the third lepton. It is nevertheless disadvantaged by the low cross-section. In the context of a fermiophobic Higgs boson model, where gluon-fusion diagrams are not allowed, this channel also provides one of the only chances of discovery in that mass range.

The transverse mass spectrum can be used to provide some (limited) information on the Higgs boson mass. From the shape of the measured transverse mass distribution, the Higgs boson mass can be determined with an uncer-

tainty of  $3 \text{ GeV}/c^2$  for an integrated luminosity of  $300 \text{ fb}^{-1}$ . One drawback of this channel is that it does not provide as clear a signature for the Higgs boson spin as expected, due to the ambiguity from the third lepton. More work is needed to determine if this can be overcome by more advanced methods.

The present analysis suffers from a low statistic for some backgrounds. It would be useful to process more events for the high cross-section backgrounds which are heavily suppressed, like  $t\bar{t}$  and  $Wt$ . Second-order backgrounds, like  $b\bar{b}$ ,  $W(Z)b\bar{b}$  single top and  $Z + jets$  should also be included in the analysis. It would certainly require a strong generator-level preselection and a large statistic. A reliable fast-simulation framework would certainly allow a more complete study of these problematic background sources, but it has to reproduce adequately difficult areas of trigger and reconstruction, such as b-tagging and vertexing.



## **Part IV**

# **Conclusion and Appendices**



## Conclusions

Leaving the LEP era, the particle physics community working at CERN is in a transition period, since physicists are concluding the LEP studies and preparing the LHC analyses. The Brout-Englert-Higgs mechanism is a key feature of the Standard Model. The Higgs boson has not yet been observed, and Tevatron will probably have difficulties to pin it down within a few years. This is therefore one of the open subjects that we hope to revolutionize with the new tools that will be offered to us. In this dissertation, we concentrated on the study of  $WW$  decays of a Higgs boson produced in associated production either at LEP (with the ALEPH detector) or at LHC (with the CMS detector).

At LEP, the Higgs boson could have been produced via the Higgsstrahlung mechanism, and be accompanied by a  $Z$  boson. If the Higgs boson subsequently decays into a pair of  $W$  bosons, it generates a six-fermion final state. All potential decay channels have been considered in a set of 13 independent analyses. In the context of a given model, mass exclusion limits have been extracted from upper limits on the cross-section. If the Higgs boson is supposed to decay exclusively in  $WW$ , the expected limit is  $107.5 \text{ GeV}/c^2$ . In a benchmark fermiophobic Higgs boson scenario, where Higgs boson couplings to  $W$  and  $Z$  bosons remain unchanged with respect to the Standard Model, one expects to put a limit on a Higgs boson with a mass between  $97.5 \text{ GeV}/c^2$  and  $104 \text{ GeV}/c^2$ . Due to a small excess of events, no limit at 95% C.L. on the presence of a Higgs boson can be obtained, but interesting constraints on such models are set. This is evident when combining this analysis with the ALEPH search for a fermiophobic Higgs boson in the  $\gamma\gamma$  channel. A lower limit on the Higgs boson mass in the benchmark fermiophobic Higgs boson scenario is set at  $105.8 \text{ GeV}/c^2$ , which is significantly better than the limit obtained in the  $\gamma\gamma$  channel alone ( $104.5 \text{ GeV}/c^2$ ).

The full set of selections that we have performed for that purpose also scrutinizes the Standard Model in parts of the phase space that are not often considered. It exploits the detailed knowledge of many processes and confirms the absence of unknown signal in many topologies. If the LEP studies were not able to unveil new physics, they participated to the deep understand-

ing of the Standard Model via many precision measurements. The search for a Higgs boson that we performed in this thesis certainly benefits from this precise understanding, and brings its humble contribution to that beautiful construction.

This study instigated software developments both of the ALEPH analysis environment (ALPHA++) and of new analysis tools for the analysis package used in High-Energy physics. The LEP likelihood-ratio method has been implemented within the ROOT package, and is now part of the official distribution. It is presently used by several running experiments. In the same spirit, neural networks facilities were ported to the ROOT framework and are now officially distributed.

In 2007, the Large Hadron Collider (LHC) will start its operation at CERN, with an wide-ranging experimental program that encompasses many searches for the Higgs boson. In preparation of the LHC operation, we considered the WH associated production of a Higgs boson decaying into a W pair. Three-lepton final states have been considered, and have been shown to provide interesting possibilities in the intermediate mass range. For an integrated luminosity of  $100\text{fb}^{-1}$ , a significance of more than  $5\sigma$  can be obtained in the mass interval between 155 and 178  $\text{GeV}/c^2$ . In the context of a fermiophobic Higgs boson model, where gluon-fusion diagrams are not allowed, this channel also provides one of the only chances of discovery in that mass range. We showed how some key properties of the Higgs boson can be measured in this channel. The partial decay width into W boson pairs can be extracted with an accuracy of 20% with an integrated luminosity of  $100\text{fb}^{-1}$ . The Higgs boson mass can be measured with a 3  $\text{GeV}/c^2$  accuracy for an integrated luminosity of  $300\text{fb}^{-1}$ . Measurement of Higgs boson properties is indeed a challenging task that will follow any possible discovery. Analyses like the one developed in the present work will hopefully be at the center of the LHC physics program.

We have shown the importance of the High-Level Trigger for physics with the CMS detector. The CMS High-Level Trigger is an advanced piece of software dedicated to the selection of sensible events. Present performances are already full of promises for 2007. In addition, physicists developing analyses in the forthcoming years will have to consider the trigger as an integral part of the selection process, as it will probably be the starting point for the definition of offline physics streams. This is particularly true for channels coping with relatively soft objects, and complex topologies combining several types of leptons and/or jets.

These studies have led to many original technical developments concerning not only the computing infrastructure but also the software tools. A computing cluster has been installed and set up. In parallel, we have designed a High-Level Trigger steering software that allows to easily build a trigger menu from simple trigger primitives. It provides the flexibility needed for both on-line operations and detailed offline studies, like the one presented here. It also offers a natural framework for designing alternative trigger strategies based

on combined objects or topological properties of specific signatures.

The aim of the CMS physics technical design report is to demonstrate the software and people's readiness for the "day 1" of data taking. The present study shows that we are indeed pretty close to that objective. It is clear that the CMS software is not yet in a shape that allows all the refinements of LEP studies. Important parts for analysis that are still lacking are a complete and reliable fast simulation framework, jet and missing transverse momentum calibrations and an energy-flow algorithm optimally combining information from various subdetectors. Bringing CMS into operation will anyway require more work from the time of the first beam, as the detector will first have to be understood before the first real search for new physics. Existing software nevertheless already allows the selection of even low-statistic signatures, such as the one considered here, which makes me very optimistic for the future.



## High-Level Trigger tree representation

It is common practice to consider the trigger as a logical “OR” of several trigger elements. More generally, a trigger (e.g. the HLT) is a logical equation, and can therefore be represented as a binary tree. This allows for more flexibility, and features like automatic prescale or combined subtriggers, commonly found in trigger systems of past experiments. The design of a binary tree is such that navigation between elements becomes natural and effective, making it possible to use recursivity techniques for the evaluation and configuration of the HLT.

According to boolean algebra, there are various equivalent descriptions of the same expression. Exploiting associativity, distributivity and other common properties, it is possible to arrange elements to improve other properties, like the mean evaluation time. The present (as in ORCA 8.7.3) official trigger table emphasizes the readability of the tree, as it has to be understood, discussed and improved by as many collaborators as possible.

The list of available trigger elements evolves continuously, as new strategies are designed, but tends to stabilize. It consists in a set of basic elements, concentrating on single triggers based on elementary particles. More elaborated trigger strategies are then built from there using logical operators. The list of trigger elements implemented in ORCA 8.7.3 is presented in table [A.1](#).

The trigger menu used in this work at low luminosity ( $2 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ ) is presented in Figures [A.1](#), [A.2](#), [A.3](#) and [A.4](#).

L2EleTrigger	Level 2 single electron trigger
L25EleTrigger	Level 2.5 single electron trigger
L25PhotonTrigger	Level 2.5 single photon trigger
L3EleTrigger	Level 3 single electron trigger
L3PhotonTrigger	Level 3 single photon trigger
L2MuTrigger	Level 2 single muon trigger
L3MuTrigger	Level 3 single muon trigger
L2JetTrigger	1,2,3,4 Jet trigger (calorimetric)
L25BJetTrigger	Level 2.5 b tagger (based on pixel lines)
L3BJetTrigger	Level 3 b tagger (based on full tracks)
L2MetTrigger	Missing Et trigger
FamosL2JetTrigger	Jet trigger for fast reconstruction
FamosL2MetTrigger	Missing Et trigger for fast reconstruction
L2TauTrigger	Level 2 single tau trigger
L25PixelTauTrigger	Level 2.5 tau validation with pixel
L25TrackerTauTrigger	Level 2.5 tau validation with tracker
L1vetoForHLT	Level 1 condition (based on a bit subset)
dummyHLTElement	Termination condition in the tree recursion
CombinatorialAndTrigger	Logical and between 2 triggers
ANDTrigger	Logical and between 2 triggers
NOTTrigger	Logical not
ORTrigger	Logical or between 2 triggers
RandomTrigger	Random trigger (for tests or pre-scale)

Table A.1: Elements implemented in the HLT steering code. CombinatorialAndTrigger is a logical “AND” with the additional restriction that the candidates are not in the same region. More generally, the CombinatorialAndTrigger can be used to build combined triggers of non exclusive objects (like an electron and a tau).

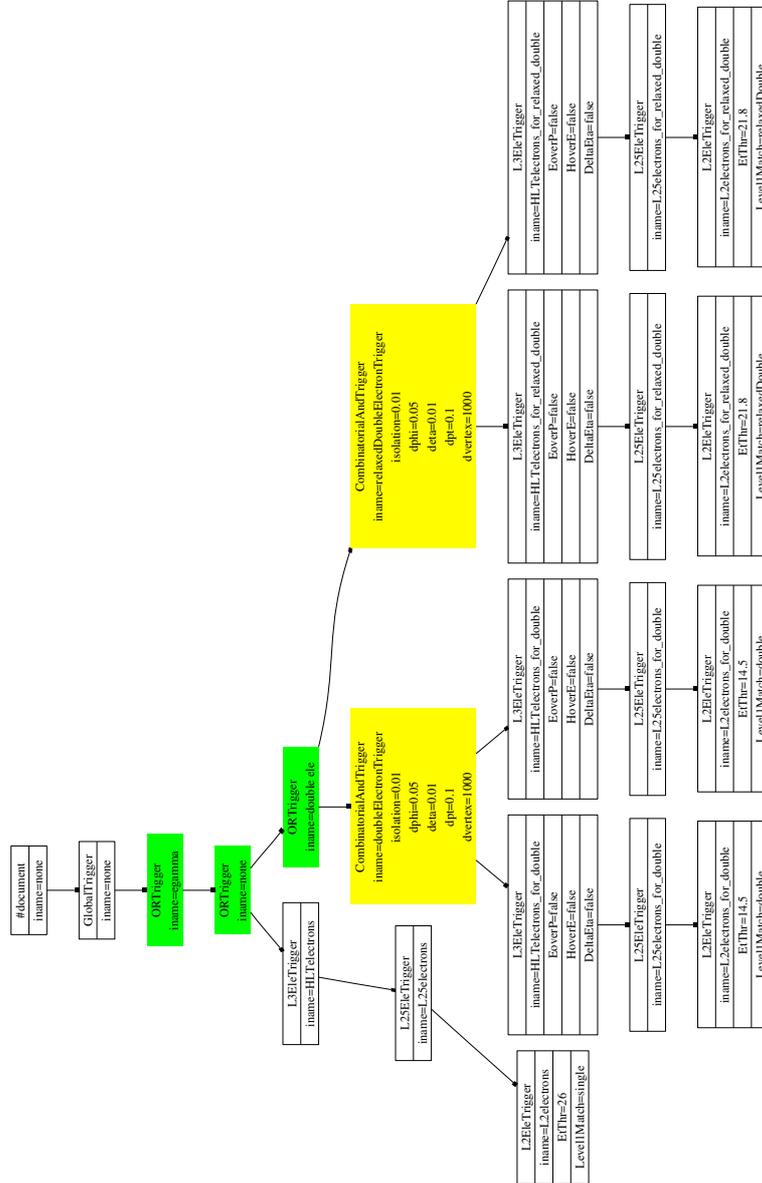


Figure A.1: Part 1/4 of the trigger tree at low luminosity ( $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ ). Here are the electron triggers (single, double, and double relaxed). Only parameters different from the default values are specified in the XML file from which this tree is produced. One can see how the L3 selection of electrons in the (relaxed) double trigger is simple, as E/P, H/E and  $\Delta\eta$  cuts are disabled. One sees also the stronger Pt cuts for the relaxed stream, based on non-isolated L1 candidates.





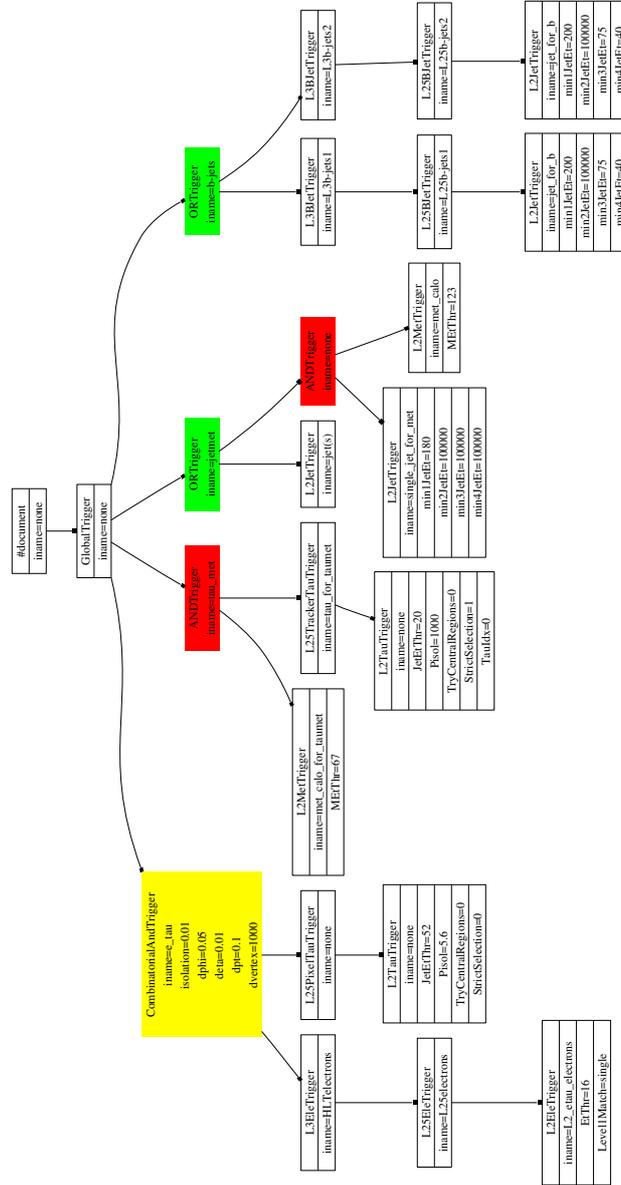


Figure A.4: Part 4/4 of the trigger tree at low luminosity ( $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ ). Combined electron\* Tau trigger, jetMET triggers and B triggers are presented. Subtriggers involved in this part demonstrates well the gain in factorizing the trigger, as many of the elements have already been used in previous parts. This avoids most of the potential code duplication.

## Computing Farm

Because university computing resources are limited, particle physics research is concentrated at several institutions worldwide, which share the data as well as the work of preparing, constructing, and running experiments. Analysis of ALEPH data presented in Chapter 4 was performed in such a scheme, and jobs were run at CERN on public batch nodes. Such an approach does not scale with the new “gigantic” collaborations like CMS, which has to produce and analyze huge amounts of events. The analysis presented in Chapter 7 required both advanced hardware and software tools, and we decided to install a local low-cost computing cluster at UCL, where more than 20 scientists are already contributing to the CMS experiment. The scientific motivation of such an investment is twofold:

- to test how efficient a proposed analysis may be once the detector is finished,
- to check the most crucial part of the entire software chain: the automatic discarding of well-understood background events by the trigger (L1 and HLT).

The priority needs are above all in a relatively large CPU power, and then significant storage space. Typically, a simulated event represents 1 MB on disk, and takes  $\mathcal{O}(2)$  kSi.min to be generated. Since a standard analysis requires  $\mathcal{O}(10^4)$  signal events per mass bin and  $\mathcal{O}(10^5)$  events per background source (as larger background samples common to several analyses are centrally produced, in a “tier 1” center or at CERN),  $\mathcal{O}(500)$ GB are needed as well as  $\mathcal{O}(20000)$ kSi.h. To fulfill the needs of several analyses performed in parallel, the resources needed are then  $\mathcal{O}(1)$ TB of storage and  $\mathcal{O}(40)$ kSi of CPU power.

In the context of this thesis, we have chosen to develop a cluster of commercial computers in four steps. The first two steps were achieved in 2002 with the installation of a RAID (Redundant Array of Independent Disks) stor-

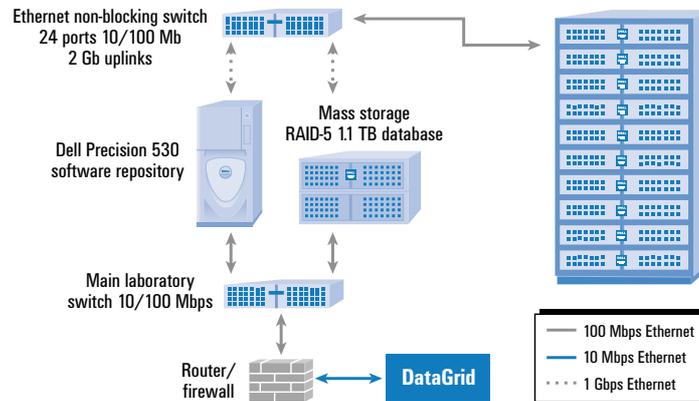


Figure B.1: Overview of the cluster architecture, with the 24 machines cluster rack shown on the right and the Ethernet switch with its two uplinks to the workstation and the mass storage on the top. From [96].

age device as well as a manager and software repository and with the proof-of-concept of the cluster using height existing machines at the institute. In 2003, the third step was to acquire, set up and bring into operation the first 12 servers, while the fourth step was achieved in 2004 with the installation of 12 new computing nodes. The present configuration is as follows (Figure B.1):

- 24 Computing nodes: 12 dual Intel PIII-S 1.4GHz, 512 KB of L2 cache, 512MB RAM; 12 dual Xeon 2.4GHz, 512 KB of L2 cache, 1GB ECC RAM.
- One Ethernet Switch: 24 ports 10/100Base-TX and 2 copper Gigabit uplinks. Switch fabric speed of 9.6 Gb.
- 1 Workstation for software repository and cluster management: Dual Intel Xeon 2GHz, 512 KB L2 cache, 512MB RAM.
- 1 Mass Storage: 2 RAID10 arrays of 630GB each and 2 RAID1 arrays of 157 GB each.

The total computing power available is estimated at 45 kSi, even if no official estimate for our machines is available. In addition, a 10TB RAID5 disk array was being configured at the time of writing. It will significantly improve productivity, as the 1.5TB available up to now are already completely filled with simulated events ( $\mathcal{O}(10)$  different analyses are being pursued in parallel).

The computing nodes are connected to the mass storage and the software repository using 100Mb Fast-Ethernet links. The use of Gigabit-Ethernet links from these two sensitive devices to the switch allows concurrent access. This



Figure B.2: Photograph of the new computer center refurbished in the context of this project. The mass storage and the software repository are placed on the table on the right, while computing nodes sits in the (opened) white rack. The second rack regroups servers of the institute, the commodity network and the link to the Internet.

is motivated by the data throughput, which exceeds 100 Mb/s during Monte-Carlo generation and digitization with pile-up, when all the 24 computing nodes are contributing to the simulation. The workstation is used as IP router and secondary DHCP server to guarantee network connectivity between the cluster nodes and the client machines submitting production jobs. A photograph of the cluster can be seen in Figure B.2.

We imagined and carried out a stress procedure in order to test the mass storage arrays from the cluster nodes. We transferred simultaneously 1GB files from the computing nodes to the RAID arrays. A rate up to 75MB/s was reached and maintained during the whole transfer when using a combination of RCP for data transfer and RAID0 for storage. It should be noted that switching either to NFS or RAID5 led to problems respectively with the NFS daemon cache and the RAID firmware. NFS is nevertheless successfully used in production as it eases management. All necessary softwares for the simulation of CMS are exported this way, so that a single directory has to be mounted to be able to run the full software chain.

The Table B.1 presents a detailed list of these softwares, together with the latest version installed. It differentiates packages developed by the CMS collaboration from other HEP-specific softwares. The Table B.2 presents an exhaustive list of “external” softwares installed to meet dependencies. The complete chain of CMS softwares spans processes starting with event generation,

followed by the simulation of detector effects, simulation of the electronic response, reconstruction of physical quantities, and finally, analysis. One can divide this chain into three independent projects: CMKIN, OSCAR and ORCA.

CMKIN is an interface to the Monte-Carlo tools available on the market. Since these tools are often written in FORTRAN, CMKIN remains written in this language, while all other tools are written in C++. Available back-ends are Pythia, Herwig, CompHep and TopRex. All generators compatible with Pythia, or providing a Pythia interface, are also supported.

The generated events are then fed to OSCAR. This Geant4 application simulates the energy deposited in the matter by the particles produced in the final state. For each sensitive volume, the energy deposited is computed and stored as "hits". The last simulation step is performed by ORCA and consists in the simulation of the whole electronic read-out chain. It produces an electronic footprint (called "digis") of what would have been recorded in real conditions.

Event reconstruction, which stands for the interpretation of digis in terms of particles and physical quantities, is then performed later on by another ORCA job. ORCA is also used by all physicists for event analysis.

Administrative tasks are performed with the help of the Cluster Command and Control (C3)[97] tool suite. Among other services, it allows the simultaneous execution of a command on all cluster nodes, e.g. to start, reconfigure or shut down a service.

The monitoring of the farm is achieved by SNMP, and the gathered data can be visualized on the web with the CACTI[98] interface. In this way, temperature, network load and CPU usage are logged, which eases the tracking of problems.

The CONDOR[99] batch system is used as the scheduler for the execution of CMS simulation and reconstruction jobs. Both machines in the cluster and desktop machines of the institute are included in a virtual distributed batch cluster so that analysis jobs can be run indifferently on the cluster or on desktop machines, exploiting CPU cycles that are unused at the time. Desktop machines represent 36 kMips that have to be shared with interactive use of the resources. Simulation jobs are run exclusively on the cluster to benefit from the private network we have just presented.

More than  $285 \times 10^3$  CPU hours were accumulated by the cluster at the time of writing (4/5/2005); in addition to CMS applications, it was used for computations of vortex dynamics in supra-conductors, two-dimensional conformal field theory studies and numerical integration of one-loop Feynman diagrams (see for example reference [100]).

Software	Version	Description
cmkin	1.3.3	Pythia application for Monte-Carlo event generation
cmsim	1.3.3	GEANT 3 Simulation for Cms Analysis and Reconstruction
geometry	1.14.0	CMS detector geometry
cobra	8.1.0	Coherent Object-oriented Base for Reconstruction, Analysis and simulation.
ignominy	2.3.1	Tools for package dependencies management
orca	8.7.1	Object Oriented Reconstruction for CMS Analysis
oscar	3.6.5	Object oriented Simulation for Cms Analysis and Reconstruction
iguana	5.3.4	Modular C++ toolkit for interactive visualization
iguanacms	1.12.0	CMS specific visualization routines
pax	1.02.02	C++ toolkit for advanced physics analyses
pi	1.2.5	Physicist Interface to analysis
pool	1.8.1	Common persistency framework for the LCG
seal	1.4.3	Common Software infrastructure for the LHC experiments
dcap	1.2.26	POSIX-like commands for dCache storage
aida	3.2.1	Abstract interfaces for common physics analysis objects
geant4	6.2	Toolkit for the simulation of the passage of particles through matter
root	3.10.02	Object-Oriented Data Analysis Framework
clhep	1.9.1.2	HEP-specific foundation and utility classes
pythia	2003	Program for the generation of high-energy physics events
geant	2003	Toolkit for the simulation of the passage of particles through matter
edg-rls-client	2.3.3	Protocol to access the data in the LRC through a set of secure SOAP-RPC calls

Table B.1: CMS (part 1) and other HEP-specific (part 2) softwares installed for the simulation and analysis of CMS events.

Software	Version	Description
Linux	3.0.3	Scientific Linux CERN, Operating system.
Condor	6.6.1	Job Submission system
SNMP	4.2.5	Simple Network Management Protocol
CVS	1.11.1	Concurrent Version System
ViewCVS	0.9.2	Web viewing tool of the CVS repository.
Cacti	0.8.2a	Web monitoring tool
otl	4.0.67	Oracle, Odbc and DB2-CLI Template Library
unixodbc	2.2.6	Open specification to access Data Sources
myodbc	3.51.06	MySQL Connector/ODBC
wxpython	2.4.2.4	GUI toolkit for Python
rfio	2003	Hierarchical storage management (HSM) system (CASTOR)
xerces-c	2.3.0	Validating XML parser
condldb	0.2.0	Tools to manipulate condition data
mysqlpp	1.7.9	C++ wrapper for MySQL's C API
mysql	4.0.18	Open source database
sqlite	3.0.8	Self-contained SQL database engine
gsl	1.4	GNU Scientific Library
qt	3.3.2	C++ application development framework
pcre	4.4	Perl Compatible Regular Expressions
qutxmllrpc	0.1	XML-RPC library for Qt
soqt	1.2.0	Coin interface for Qt
coin3d	2.3.0	Libraries for creating 3D graphics applications
openssl	0.9.7d	Secure Sockets Layer protocol
uuid	1.32	Universal Unique Identifier
boost	1.31.0	Portable C++ source libraries
gccxml	0.6.0p1	XML output extension to GCC
python	2.3.4	Interpreted object-oriented programming language
zlib	1.1.4	Compression Library
bz2lib	1.0.2	High-quality data compressor
graphviz	1.9	Graph drawing tools
gcc	3.2.3	GNU Compiler Collection

Table B.2: Service softwares installed to meet dependencies for the simulation and analysis of CMS events. The first part of the table presents tools and services installed to ease administration. The second part gives an exhaustive list of packages installed to meet dependencies.

## BIBLIOGRAPHY

- [1] C.N. Yang and R.L. Mills, *Phys. Rev.* **96**,96 (1954).
- [2] Jan Govaerts, UCL-IPN-95-P02 (1995), Proceedings of the Ecole Internationale Joliot-Curie de Physique Nucleaire, 84p.
- [3] P.W. Higgs, *Phys. Lett.* **12** (1964) 132; F. Englert and R. Brout, *Phys. Rev. Lett.* **13** (1964); G. Guralnik and C.K. Hagen, *Phys. Rev. Lett.* **13** (1964) 585.
- [4] Particle Data Group, Review of Particle Physics, *Phys. Lett.* **B592** (2004)
- [5] A. Djouadi, J. Kalinowski, M. Spira, HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension, hep-ph/970448 (1997).
- [6] J.Bagger et al., *Phys. Rev.* **D49** (1994) 1246-1264.
- [7] Gordon Kane, John F. Gunion, Sally Dawson, Howard E. Haber, Higgs Hunter's Guide, Lightning Source Inc (2000).
- [8] T. Hambye and K. Riesselmann, *Phys. Rev.* **D55** (1997) 7255.
- [9] ALEPH, DELPHI, L3 and OPAL Collaborations and The LEP Working Group for Higgs Boson Searches, *Phys. Lett.* **B565** (2003) 61-75
- [10] D.A. Ross and M. Veltmann, *Nucl. Phys.* **B95** (1975) 135.
- [11] S.L. Glashow and S. Weinberg, *Phys. Rev.* **D15** (1977) 1958.
- [12] J.F. Gunion and H.E. Haber, hep-ph/0207010; H.E. Haber, hep-pj/9501320.
- [13] ALEPH, DELPHI, L3 and OPAL Collaborations and The LEP Working Group for Higgs Boson Searches, LHWG-Note 2004-01 (2004), paper contributed to ICHEP'04, Beijing (China), 25p.
- [14] LEP Higgs working group, LHWG Note 2002-02 (2002), paper contributed to ICHEP'02, Amsterdam (Netherlands), 8p.

- [15] The LEP collaborations, the LEP electroweak working group and the SLD Electroweak and Heavy Flavour Groups, hep-ex/0412015 (2004), 223p, Updated for 2005 winter conferences, April 2005, <http://www.cern.ch/LEPEWWG>.
- [16] D0 Collaboration, hep-ex/0407005 (2004), 6p.
- [17] John D. Jackson, Classical Electrodynamics, John Wiley & Sons, Inc., 3-rd edition, 1999.
- [18] ALEPH Coll., ALEPH: A Detector for Electron-Positron Annihilations at LEP, Nucl. Instrum. and Methods **A294** (1990) 121.
- [19] ALEPH Coll., Performance of the ALEPH Detector at LEP, Nucl. Instrum. and Methods **A360** (1995) 481.
- [20] C. Bowdery et al. (The ALEPH Collaboration), The ALEPH Handbook, Vol 1 and 2, CERN, Geneva, Switzerland, (1995 and 1997)
- [21] D. Creanza et al., The New ALEPH Silicon Vertex Detector, Nucl. Instrum. and Methods **A409** (1998) 157-160.
- [22] R. Cavanaugh, C. Delaere, G. Dissertori, K. Huettmann, V. Lemaître, O. van der Aa, ALPHA++ User's Manual. ALEPH OO Physics Analysis Package version 3.4, CERN-ALEPH-2002-033; CERN-ALEPH-SOFTWARE-2002-002. (2002) 49p.
- [23] Rene Brun and Fons Rademakers, ROOT - An Object Oriented Data Analysis Framework, Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Instrum. and Methods **A389** (1997) 81-86. See also <http://root.cern.ch/>.
- [24] C. Delaere and V. Lemaître, An Alternative Invisible Higgs Search Performed with ALPHA++, CERN-ALEPH-2002-035; CERN-ALEPH-SOFTWARE-2002-003. (2002) 32p. and V. Lemaître, C. Delaere, Comparison Between an Event Selection Program Performed with ALPHA and ALPHA++, CERN-ALEPH-2000-093; CERN-ALEPH-SOFTWARE-2000-034. (2000) 49p.
- [25] OPAL Coll., G. Abbiendi et al., Search for Higgs Boson and Other Massive States Decaying into Two Photons in  $e^+e^-$  Collisions at 189 GeV, Phys. Lett. **B464** (1999) 311-322
- [26] L3 Coll., P. Achard et al., Search for a Higgs Boson Decaying into Two Photons at LEP, Phys. Lett. **B534** (2002) 28-38
- [27] ALEPH Coll., Search for  $\gamma\gamma$  Decays of a Higgs Boson in  $e^+e^-$  Collisions at  $\sqrt{s}$  up to 209 GeV, Phys. Lett. **B544** (2002) 16-24.

- [28] DELPHI Coll., Search for Fermiophobic Higgs Bosons in Final States with Photons at LEP 2, *Eur. Phys. J.* **C35** (2004) 313-324
- [29] L3 Coll., P. Achard et al., Search for a Higgs Boson Decaying to Weak Boson Pairs at LEP, *Phys. Lett.* **B568** (2003) 191-204
- [30] A. Djouadi, J. Kalinowski and M. Spira, *Comput. Phys. Commun.* **108** (1998) 56 and The LEP Higgs working group, LHWG note 2001-08.
- [31] G. Ganis and P. Janot, "The HZHA Generator in "Physics at LEP2", Eds. G. Altarelli, T. Sjöstrand and F. Zwirner, CERN 96-01 (1996), Vol 2, 309.
- [32] Tomasz Pierzchala, private communication. November 2003.
- [33] S. Jadach, W. Placzek and B.F.L. Ward, *Phys. Lett.* **B390** (1997) 298.
- [34] S. Jadach and Z. Wąs, *Comput. Phys. Commun.* **36** (1985) 191.
- [35] J.A.M. Vermaseren, in Proceedings of the IVth International Workshop on Gamma Gamma Interactions, Eds. G. Cochar and P. Kessler, Springer Verlag, 1980.
- [36] M. Skrzypek, S. Jadach, W. Placzek and Z. Wąs, *Comput. Phys. Commun.* **94** (1996) 216.
- [37] S. Jadach, B.F.L. Ward, and Z. Wąs, *Comput. Phys. Commun.* **130** (2000) 260.
- [38] T. Sjöstrand, *Comput. Phys. Commun.* **82** (1994) 74.
- [39] GEANT Detector Description and Simulation Tool, CERN Program Library, CERN-W5013 (1993).
- [40] B. Clerbaux, C. Delaere, V. Lemaitre, O. van der AA, Search for WW decay of a Higgs boson produced in association with a fermion pair in e+e- collisions at LEP. Aleph 2003-018. (2003) Geneva; and Aleph publication in preparation.
- [41] T. Junk, *Nucl. Instrum. and Methods* **A434** (1999) 435.
- [42] S. Catani, Yu.L. Dokshitzer, M. Olsson, G. Turnock and B.R. Webber, *Phys. Lett.* **B269** (1991) 432.
- [43] S. Brandt et al. *Phys. Lett.* **12** (1964) 57; E. Farhi, *Phys. Rev. Lett.* **39** (1977) 1587.
- [44] J.D. Bjorken and S.J. Brodsky, *Phys. Rev.* **D1** (1970) 1416, Hanson G et al, *Phys. Rev. Lett.* **35** (1975) 1609, C. Berger et al. (PLUTO collaboration), *Phys. Lett.* **B82** (1979) 449.
- [45] T. Sjöstrand, *Comput. Phys. Commun.* **67** (1994) 74.

- [46] T. Sjöstrand, *Comput. Phys. Commun.* **71** (1992) 15.
- [47] The ALEPH, DELPHI, L3 and OPAL Collaborations, The LEP Higgs Working Group, Searches for Higgs Bosons Decaying into Photons: Preliminary Combined Results Using LEP Data Collected at Energies up to 209 GeV, LHWG Note 2001-08 (2001).
- [48] Jerome Schwindling, the MLPFIT package, <http://schwind.home.cern.ch/schwind/MLPfit.html>
- [49] Proposal to the Rutherford Appleton Laboratory, An International Muon Ionization Cooling Experiment, MICE-Note 21, <http://www.mice.iit.edu/mnp/MICE0021.pdf>
- [50] TESLA Technical Design Report, DESY 2001-011; ECFA 2001-209; March 2001.
- [51] ATLAS Technical Proposal, CERN/LHCC/94-43; LHCC/P2; December 1994.
- [52] CMS Technical Proposal, CERN/LHCC/94-38; LHCC/P1; December 1994.
- [53] LHCb Technical Proposal, CERN/LHCC/98-4; February 1998.
- [54] ALICE Technical Proposal, CERN/LHCC/95-71; December 1995.
- [55] TOTEM Technical Proposal, Total Cross Section, Elastic Scattering and Diffractive Dissociation at the LHC, CERN/LHC 99-7, March 1999.
- [56] O. Bruning et al., LHC design report vol.1 : the LHC Main Ring, CERN-2004-003; 2003.
- [57] UA1 Collaboration, Experimental Observation of Isolated Large Transvers Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540$  GeV, *Phys. Lett.* **B122** (1983) 103-116; UA1 Collaboration, Observation of the Muonic Decay of the Charged Intermediate Vector Boson, *Phys. Lett.* **B134** (1984) 469-476.
- [58] V. Karimaki, Explicit Covariance Matrix for Particle Measurement Precision, *Nucl. Instrum. and Methods* **A410** (1998) 284-292.
- [59] CMS collaboration, The Electromagnetic Calorimeter Technical Design Report, CERN/LHCC 97-33; December 1997.
- [60] CMS collaboration, Tracker Technical Design Report, CERN/LHCC 98-6; April 1998; and CMS collaboration, Addendum to the CMS Tracker TDR, CERN/LHCC 2000-016, February 2000.

- [61] CMS collaboration, The TriDAS Project Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger, CERN/LHCC 02-26; December 2002.
- [62] CMS collaboration, The HCAL Technical Design Report, CERN/LHCC 97-31; June 1997.
- [63] CMS collaboration, The Magnet Project Technical Design Report, CERN/LHCC 97-10; May 1997.
- [64] CMS collaboration, Muon Technical Design Report, CERN/LHCC 97-32; December 1997.
- [65] E. Meschi, T. Monteiro, C. Seez, P. Vikas, Electron Reconstruction in the CMS Electromagnetic Calorimeter, CMS Note **2001/034** (2001).
- [66] J.E. Huth et al., "Toward a Standardization of Jet Definition", Proceedings of Research Directions For The Decade, Snowmass 1990, Snowmass, July 1990, ed. E.L. Berger (World Scientific, Singapore, 1992), p. 134.
- [67] Marcel Vos, private communication. (2004)
- [68] C. Delaere et al., HLT Steering Code Prototype, CMS IN 2003/052 (2003).
- [69] see <http://cmsdoc.cern.ch/orca>
- [70] Erich Gamma, Richard Helm, Ralph Johnson, John Vlissides, Design Patterns: Elements of Reusable Object-Oriented Software, Copyright 1998 by Addison Wesley Longman, inc.
- [71] C. Grandi, D. Stickland, L. Taylor, CMS Computing Model, CMS-NOTE/2004-031
- [72] Elliotte Rusty Harold, W. Scott Means, XML in a Nutshell, 2nd Edition. A Desktop Quick Reference
- [73] CMS internal note in preparation.
- [74] O. van der Aa, C. Delaere, The High Level Trigger Software for the CMS experiment, proceedings of CHEP2004 Vol 1, 242-245 (2005).
- [75] O. van der Aa, CMS trigger strategies for the selection of MSSM Higgs bosons using electron tau-jet decay modes, FYNU-THESIS-2005-015 (2005).  
<http://cds.fynu.ucl.ac.be/CDS/search.py?recid=118>
- [76] G.E. Moore, Cramming more components onto integrated circuits, Electronics, Volume 38, Number 8, April 19, 1965.
- [77] Z. Kunszt, S. Moretti and W. J. Stirling; Z. Phys. **C74** (479) 1997 and hep-ph/9611397.

- [78] S. Abdullin et al., Summary of the CMS Potential for the Higgs Boson Discovery, CMS NOTE 2003/033.
- [79] see <http://people.web.psi.ch/spira/v2hv>. This code uses the formula presented in T. Han and S. Willenbrock, Phys. Lett. **B273** (1991) 167.
- [80] V. Cavasinni, D. Costanzo, Search for  $WH \rightarrow WWW \rightarrow l^\pm \nu l^\pm \nu$  jet using like-sign leptons, ATLAS-PHYS-2000-013.
- [81] W. Bonivento, First look at a SM Higgs boson search with a tri-lepton signal, ATLAS-PHYS-2000-011.
- [82] K. Jakobs, A study of the associated production  $WH, H \rightarrow WW \rightarrow l\nu l\nu$ , ATLAS-Phys-2000-008.
- [83] M.L. Ciccolini, S. Dittmaier and M. Kramer, Electroweak Radiative Corrections to Associated WH and ZH Production at Hadron Colliders, hep-ph/0306234 (2003).
- [84] E.E. Boos et al., CompHEP: Specialized Package for Automatic Calculations of Elementary Particle Decays and Collisions, SNUTP report 94-116, Seoul, 1994 (hep-ph/9503280).
- [85] S. Slabospitsky and L. Sonnenschein, Comput. Phys. Commun. **148** (2002) 87, hep-ph/0201292.
- [86] M.L. Mangano et al., JHEP 0307:001, 2003, hep-ph/0206293.
- [87] G. Segneri, F. Palla, Lifetime based b-tagging with CMS, CMS NOTE 2002/046 (2002).
- [88] S. Catani, et al., Nucl. Phys. **B406** (1993) 187.
- [89] A. Djouadi and S. Ferrag, PDF Uncertainties In Higgs Production At Hadron Colliders, hep-ph/0310209.
- [90] M. Dittmar, F. Pauss, D. Zurcher, Towards a Precise Parton Luminosity Determination at the CERN LHC, Phys. Rev. **D56** (1997) 7284-7290.
- [91] ATLAS Collaboration, CERN/LHCC/2004-010
- [92] A.G. Shamov, V.I. Telnov, Precision Luminosity Measurement at LHC Using Two-photon Production of  $\mu^+\mu^-$  pairs. hep-ex/0207095, and D. Bocian, K. Piotrkowski, Very Forward Two-Photon  $e^+e^-$  Production and Luminosity Measurement for Ion Collisions at the LHC, Acta Phys. Polon. **B35** 2417-2424 (2004).
- [93] V. Bartsch, G. Quast, Expected Signal Observability at Future Experiments, CMS IN 2003/039 (2003)

- 
- [94] M. Hohlfeld, On The Determination Of Higgs Parameters In The ATLAS Experiment At The LHC, ATL-PHYS-2001-004.
- [95] M. Duhrssen et al., Extracting Higgs Boson Couplings From LHC Data, hep-ph/0406323.
- [96] C. Delaere, T. Keutgen, G. Leibenguth, V. Lemaitre, A. Ninane, O. van der Aa, Solving High-Energy Physics Using a Dell HPC Cluster at the Université catholique de Louvain, Dell Power Solutions, (August 2003)
- [97] Cluster Command & Control (C3) Tool Suite, 3rd Austrian-Hungarian Workshop on Distributed and Parallel Systems (DAPSYS 2000) in conjunction with EuroPVM/MPI 2000, September 10-13, 2000, Balatonfured, Lake Balaton, Hungary. and <http://www.csm.ornl.gov/torc/C3/>
- [98] see <http://www.cacti.net>
- [99] Jim Basney, Miron Livny, and Todd Tannenbaum, *High Throughput Computing with Condor*, HPCU news, Volume 1(2), June 1997
- [100] G. Stenuit et al, Temperature Dependence of Penetration and Coherence Lengths in Lead Nanowires, *Supercond. Sci. Technol.* 18 (2005) 174-182.