Measurement of Z boson pair production and a search for the Higgs boson in e^+e^- collisions at LEP

Ivo Boely van Vulpen

Measurement of Z boson pair production and a search for the Higgs boson in e^+e^- collisions at LEP

ACADEMISCH PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR
AAN DE UNIVERSITEIT VAN AMSTERDAM
OP GEZAG VAN DE RECTOR MAGNIFICUS
PROF. MR. P.F. VAN DER HEIJDEN
TEN OVERSTAAN VAN EEN DOOR HET COLLEGE VOOR PROMOTIES
INGESTELDE COMMISSIE, IN HET OPENBAAR TE VERDEDIGEN
IN DE AULA DER UNIVERSITEIT
OP DONDERDAG 18 APRIL 2002, TE 12:00 UUR

door

Ivo Boely van Vulpen

geboren te Bilthoven

Promotor: **Prof. dr. J.J. Engelen**
Co-promotor: **Dr. J.J.M. Timmermans**

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

The work described in this thesis is part of the research programme of 'het Nationaal Instituut voor Kernfysica en Hoge-Energie Fysica' (NIKHEF) in Amsterdam, the Netherlands. The author was financially supported by the 'Stichting voor Fundamenteel Onderzoek der Materie' (FOM).

Contents

Chapter 1

Introduction

Throughout history it has been a great adventure for mankind to unravel the many mysteries of nature. By closely observing nature we try to find patterns or 'laws' that nature obeys. The models that have been constructed to represent nature often possess rather simple underlying principles or symmetries that can nevertheless describe its very complex behaviour. Nature is studied at many different scales and levels of complexity and in this thesis we focus on the part that studies the behaviour of the fundamental building blocks of matter and their interaction: particle physics.

In the twentieth century, physicists have gathered an enormous amount of knowledge about the world of elementary particles by performing a wide variety of experiments. At this moment all observed particles and their interactions can be accommodated in a single model called 'the Standard Model'. This model has been extremely successful in describing the results of all experiments (in some cases to below a per mill precision) over a wide range of energies. The success of this model reveals a beautiful symmetry and ordering at the most fundamental level of nature. However, in spite of its enormous success, there is a single essential piece of this model still missing: a particle called 'the Higgs boson'. This particle is a necessary ingredient of the theory for various reasons. The first reason is that it explains how particles acquire a mass. Most important in that respect is that it explains why the particles that mediate the weak force (one of the four fundamental forces in nature) are massive which in turn explains why this force only operates over small distances. In addition it also gives rise to the masses of the fundamental building blocks of matter, the quarks and leptons. The second reason that the Higgs boson is an essential part of the model is that without it the Standard Model is not renormalisable. This means that without this particle, the computations performed using this model yield meaningless results and the model has no predictive power. Including the Higgs boson in the theory solves these problems. Although there are many indications that this particle indeed exists, it has not been observed directly until this day, despite numerous searches and experiments. This thesis describes a search for the Higgs boson leading to one of the most competitive lower limits on its mass established to date.

We look for the Higgs boson in $\mathrm{e^+e^-}$ collisions where it is produced together with another fundamental particle: the Z boson. The high energies needed to create these two particles simultaneously can only be reached in large accelerators. In this thesis we study the collisions from electrons and positrons at the Large Electron Positron collider (LEP) of the European Laboratory of Nuclear Research (CERN) near Geneva.

The outline of this thesis is as follows. In Chapter 2 an introduction is given to the Standard Model of electroweak interactions and the characteristics of the Higgs boson are presented as a function of its unknown mass. Chapter 3 describes briefly both the LEP accelerator, where bunches of electrons and positrons are accelerated and collided and the DELPHI detector that records the particles that are produced in these collisions. We will focus on the class of events with a multi-jet (4-quark) final state. The selection of these events from all $\mathrm{e^+e^-}$ interactions together with the tools that are used to analyse these events are explained in chapter 4. In chapter 5 these tools are used to extract information about each of the observed events at different levels of complexity. This chapter also presents the method that is used to combine this experimental information with what is expected for various processes. Also the characteristics of (and the main differences between) the various event types that produce a 4-quark final state are discussed here. This information is then used to construct a compatibility measure for the observed event to originate from each of these different processes. In chapters 6, 7 and 8 this analysis method is then applied to three different measurements. Chapter 6 describes the measurement of Z boson pair production where both Z bosons decay into a quark anti-quark pair. This process is very similar to that of Higgs boson production and serves as a ’calibration’ of the method. In chapter 7 the data collected by the DELPHI detector at the highest centre-of-mass energies are used to look for the possible presence of a Higgs boson. Finally, in chapter 8, a more model independent analysis is presented by placing an upper limit on the cross section of a hadronically decaying scalar particle that is pair produced together with a Z boson.

Chapter 2

Theory

2.1 The Standard Model

The present understanding of the fundamental structure of nature is contained in the Standard Model of electroweak and strong interactions[1]. The Standard Model (SM) was developed by Glashow, Weinberg and Salam [1, 2, 3] and describes the interactions of the fundamental constituents of matter. In this section only a brief and general introduction of the SM is given, followed by a more detailed report on the subject of electroweak symmetry breaking as it is the focus of this thesis. For a more detailed introduction to the SM or gauge theories in general, the reader is referred to references like [4] or [5].

The fundamental particles of matter are the quarks and leptons. They are spin-1/2 particles (fermions) and are grouped into 'families' or 'generations'. The fundamental fermions in nature and their quantum numbers within the SM are shown in table 2.1. With the direct observation of tau neutrino interactions in the year 2000 [6], all matter particles have been discovered.

Fermions	Generations			I_3	Y	Q
Leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	+1/2 -1/2	-1 -1	0 -1
	e_R	μ_R	τ_R	0	-2	-1
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	+1/2 -1/2	+1/3 +1/3	+2/3 -1/3
	u_R	c_R	t_R	0	+4/3	+2/3
	d_R	s_R	b_R	0	-2/3	-1/3

Table 2.1: *The fermions of the Standard Model and some of their quantum numbers. In the SM the electric charge Q of a particle is related to the third component of the weak isospin (I_3) and the hypercharge Y by: $Q=I_3+\frac{1}{2}Y$.*

Interactions in the Standard Model are based upon the existence of local gauge symmetries in nature. Imposing these invariances gives rise to interactions between the fermions and predicts for

[1] Gravity, the fourth fundamental force, is not contained in the SM and extremely weak when compared to the other forces at our energy scales. It is therefore not discussed here.

each of these interactions a (set of) gauge boson(s) that mediates the corresponding force between the fermions. The underlying symmetry group of the SM is $SU(3)_{\mathrm{C}} \otimes SU(2)_{\mathrm{L}} \otimes U(1)_{\mathrm{Y}}$. The group $SU(3)_{\mathrm{C}}$ is associated to the strong force and Quantum Chromo Dynamics (QCD) describes the interaction between particles with colour charge (only quarks). The strong force is mediated by eight gluons. However, as we are primarily interested in the (breaking of the) electroweak symmetry in the SM, QCD is not discussed in any more detail here. The electroweak part of the Standard Model (a unified theory of the electromagnetic and the weak interaction) is described by the other symmetry group: $SU(2)_{\mathrm{L}} \otimes U(1)_{\mathrm{Y}}$ with the photon (γ), the Z and the W bosons as corresponding gauge bosons. In the next section this part of the Standard Model is discussed in a bit more detail where we focus on the question how it is possible that the photon is massless while the other three gauge bosons are massive.

In the SM quarks interact through the strong, weak and electromagnetic interactions and since leptons do not carry a colour charge, they interact only through the electromagnetic and weak force (for charged leptons) or the weak interaction (neutrinos). An overview of the interactions in the SM and their corresponding gauge bosons (spin-1 particles) are given in table 2.2.

Interaction	Bosons	Mass (GeV/c^2)
electromagnetic	γ	0
weak	W$^+$W$^-$	80.4
weak	Z	91.2
strong	gluon	0

Table 2.2: *The interactions and corresponding gauge bosons in the Standard Model.*

The down type quarks (d, s and b) given in table 2.1 are the eigenstates of the quarks under the weak interaction. They are not the same as the *mass* eigenstates of the quarks, that generate the hadrons that are observed in the detector. These two eigenstates are related ('quark mixing') via the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix as [7]:

$$\begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\mathrm{weak}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\mathrm{mass}} \sim \begin{pmatrix} 0.975 & 0.223 & 0.004 \\ 0.222 & 0.974 & 0.040 \\ 0.009 & 0.039 & 0.999 \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\mathrm{mass}} \tag{2.1}$$

Quantum field theories based on local gauge invariance are attractive because of their mathematical beauty, but also compelling because they are renormalisable theories. The fact that a non-Abelian gauge theory with local gauge invariance was renormalisable, was shown in the beginning of the 70's by 't Hooft and Veltman [8, 9]. One drawback however is that these theories allow only massless gauge bosons. As the hypothesised vector bosons connected to the weak force were experimentally shown to be massive, clearly the model had to be altered. One way for the model to remain renormalisable and in addition allow for massive gauge bosons is to introduce an additional scalar field that breaks the underlying electroweak symmetry. This is called spontaneous symmetry breaking. As will be shown in the next section, the additional scalar field will not only provide masses for the gauge bosons, it also predicts the presence of an additional scalar particle: the Higgs boson. This thesis is devoted to the search for this particle that is the only particle still missing in the Standard Model.

2.2 Electroweak symmetry breaking

The presence of the massive gauge bosons indicates that the $SU(2)_{\mathrm{L}} \otimes U(1)_{\mathrm{Y}}$ symmetry is broken[2]. However, since the photon is massless, the $U(1)_{\mathrm{Q}}$ symmetry should be respected, where Q is the electromagnetic charge.

Breaking the electroweak symmetry can be realised either by *explicit* or *spontaneous* symmetry breaking. Explicit symmetry breaking means adding mass terms in the Lagrangian by hand for the W and Z vector bosons. These mass terms are of the form $M_B^2 B_\mu B^\mu$ (where B is a field) and they therefore break the electroweak symmetry. However, they also make the model non-renormalisable. A more elegant mechanism that ensures renormalisability is so-called spontaneous symmetry breaking. The main idea is that an additional scalar field and potential is introduced in the theory in such a way that the Lagrangian still respects the $SU(2)_{\mathrm{L}} \otimes U(1)_{\mathrm{Y}}$ symmetry, but the electroweak vacuum does not. This mechanism [10, 11, 12], known as the Higgs mechanism, is discussed in the next section. Finally, in section 2.2.2, the consequences of this symmetry breaking are discussed: masses for the vector bosons and an additional scalar particle.

2.2.1 The Higgs mechanism

To break the electroweak symmetry of the vacuum, we first introduce a complex weak $SU(2)$ isospin scalar doublet Φ with Y_Φ=1 from four real scalar fields ϕ_i as:

$$\Phi \;=\; \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.2}$$

The gauge invariant Lagrangian corresponding to this field is constructed using the covariant derivative that is connected to the requirement of local $SU(2)_{\mathrm{L}} \otimes U(1)_{\mathrm{Y}}$ gauge symmetry as:

$$\mathcal{L}_{\mathrm{Higgs}} \;=\; (D^\mu\Phi)^\dagger(D_\mu\Phi) - V(\Phi), \tag{2.3}$$

with the covariant derivative D_μ explicitly given by:

$$D_\mu \;=\; \partial_\mu - \frac{1}{2}ig\vec{\tau}\cdot\vec{W}_\mu - \frac{1}{2}ig'YB_\mu, \tag{2.4}$$

In this expression g and g' are the couplings of the weak isospin and hypercharge group respectively and the τ_i's are the generators of the $SU(2)$ group ($\frac{1}{2}$ $\times$ the Pauli matrices). In the next section the first term in the Lagrangian of (2.3) is evaluated. Here we focus on the scalar potential $V(\Phi)$ that is given by:

$$V(\Phi) \;=\; \mu^2|\Phi|^2 + \lambda|\Phi|^4, \tag{2.5}$$

with $\lambda > 0$. There are no higher-order self-interactions of the field Φ (like for example $|\Phi|^6$), since that would make the theory non-renormalisable. There are now two possible shapes of the potential depending on the sign of μ^2. They are shown separately in figure 2.1 for $\mu^2 > 0$ (left plot) and $\mu^2 < 0$ (right plot).

[2] This means that the vacuum does not possess the same symmetry as the Lagrangian.

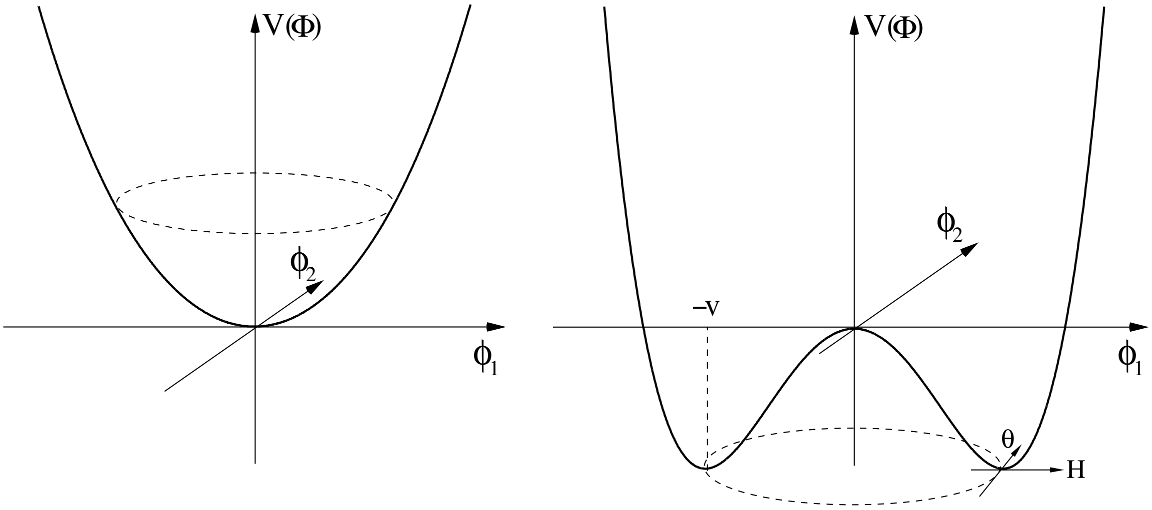


Figure 2.1: *The shape of the scalar potential V(Φ) as a function of ϕ_1 and ϕ_2 in case μ^2 is positive (left plot) and for μ^2 negative (right plot).*

If μ^2 is positive, the state of lowest energy in the theory is at Φ=0 and the Lagrangian from (2.3) describes a system of four scalar particles (the ϕ_i fields) each with a mass μ. As the theory remains invariant under $SU(2)_\mathrm{L} \otimes U(1)_\mathrm{Y}$ the gauge bosons remain massless.

The case $\mu^2 < 0$ is more interesting since at the minimum of the potential, the field acquires a non-zero expectation value. This minimum, that defines the vacuum, is given by the condition:

$$|\Phi_0|^2 = \Phi_0^\dagger \Phi_0 = -\frac{\mu^2}{2\lambda} \equiv \frac{1}{2}v^2 \tag{2.6}$$

where v is known as the vacuum expectation value. There are infinitely many solutions that satisfy (2.6). By choosing a specific solution to represent the electroweak vacuum, the $SU(2)_\mathrm{L} \otimes U(1)_\mathrm{Y}$ symmetry is broken. As remarked before, since the photon should remain massless, the vacuum is required to be still invariant under $U(1)_\mathrm{Q}$ symmetry. Therefore only ϕ^0 is given a vacuum expectation value and the electroweak vacuum is given by:

$$\phi_1 = \phi_2 = \phi_4 = 0 \quad \text{and} \quad \phi_3 = v \qquad \Rightarrow \qquad \Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.7}$$

This choice of Φ_0 has quantum numbers I_3=$-\frac{1}{2}$ and Y=1. It breaks the $SU(2)$ and $U(1)_\mathrm{Y}$ symmetry and electromagnetic gauge invariance is respected.

The resulting particle spectrum becomes apparent by expanding the scalar field around the minimum (see also figure 2.1) and substituting this expression back into the Lagrangian (2.3). The Φ_0 field near the minimum is given by:

$$\Phi(x) \; = \; \frac{1}{\sqrt{2}} e^{i\vec{\tau}\cdot\vec{\theta}(x)/v} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.8}$$

When substituting (2.8) back into the Lagrangian massless boson fields (the $\theta(x)$ fields) appear. These are the well-known Goldstone bosons that arise for each symmetry of the group that is spontaneously broken [13]. However, in the specific case of local gauge invariance the Goldstone bosons do not correspond to physical particles since via a specific $U(1)$ gauge transformation[3] these $\theta(x)$ fields can be 'absorbed' into longitudinal polarisation states of the Z and W bosons. In addition to the masses for these electroweak vector bosons (they will be computed in the next section) a massive neutral real scalar particle appears: the Higgs boson.

2.2.2 Massive electroweak gauge bosons

In the Standard Model, the first term in $\mathcal{L}_{\text{Higgs}}$ now contains the mass terms for the electroweak vector bosons as becomes clear when this term is written out explicitly using the covariant derivative as given in (2.4) and the electroweak vacuum given by Φ_0 from (2.7):

$$\begin{aligned}(D^\mu\Phi_0)^\dagger(D_\mu\Phi_0) &= \left|\left(-\frac{1}{2}ig\vec{\tau}\cdot\vec{W}_\mu - \frac{1}{2}ig'B_\mu\right)\Phi_0\right|^2 \\ &= \frac{1}{8}v^2g^2\left[(W^1_\mu)^2+(W^2_\mu)^2\right]+\frac{1}{8}v^2\left[gW^3_\mu - g'B^\mu\right]^2 \end{aligned} \tag{2.9}$$

Because the W bosons can be written as a function of the gauge fields W^1_μ, W^2_μ as

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \tag{2.10}$$

and the Z_μ and photon field as orthogonal combinations of the W^3_μ and B_μ field:

$$Z_\mu = \frac{gW^3_\mu - g'B_\mu}{\sqrt{g^2+g'^2}} \quad \text{and} \quad A_\mu = \frac{g'W^3_\mu + gB_\mu}{\sqrt{g^2+g'^2}}, \tag{2.11}$$

expression (2.9) can be written in terms of the electroweak vector bosons as:

$$(D^\mu\Phi_0)^\dagger(D_\mu\Phi_0) = \left(\frac{vg}{2}\right)^2 W^+_\mu W^{-\mu} + \frac{1}{2}\left(\frac{v}{2}\right)^2(g^2+g'^2)\,Z_\mu Z^\mu + \frac{1}{2}(0)^2\,A_\mu A^\mu \tag{2.12}$$

In general, the part of the Lagrangian corresponding to the masses of the gauge bosons is given by: $\mathcal{L}^{\text{gauge}}_{\text{mass}} = M^2_W W^2_\mu + \frac{1}{2}M^2_Z Z^2_\mu + \frac{1}{2}M^2_\gamma A^2_\mu$. Comparing this general form with (2.12) we see that the masses of the electroweak gauge bosons are given by:

$$M_{W^\pm} = \frac{1}{2}vg\,, \qquad M_Z = \frac{1}{2}v\sqrt{g^2+g'^2} \qquad \text{and} \qquad M_\gamma = 0. \tag{2.13}$$

The mixing between the B_μ and the W^3_μ field to obtain a massive Z boson and a massless photon is often expressed as a rotation using the Weinberg mixing angle θ_W. Using in addition the fact that the photon field couples to fermions and bosons with a strength corresponding to the electric charge the two free parameters g and g' (the gauge couplings) can be replaced by e and θ_W using the following relations:

$$e = g\sin\theta_W = g'\cos\theta_W \tag{2.14}$$

Therefore, at tree level, the relation between the masses of the massive vector bosons can also be expressed as $M_W = M_Z\cos\theta_W$.

[3] This specific gauge is known as the unitary gauge.

2.2.3 Massive fermions

The gauge bosons have acquired a mass through the spontaneous symmetry breaking via the Higgs mechanism. This mechanism has no influence on the mass of the fermions, but can in principle also be used to give fermions a mass. There are a priori two ways to give fermions a mass: either the masses are put in by hand or by imposing a Yukawa coupling between the scalar Higgs field and the fermions (with for each fermion a specific coupling constant g_f that is a free parameter). This effective interaction gives rise to a fermion mass of:

$$m_f = g_f \frac{v}{\sqrt{2}} \tag{2.15}$$

2.2.4 Mass of the Higgs boson

The mass of the Higgs boson is given by the Φ^2 term in V(Φ):

$$M_H = v\sqrt{2\lambda} \tag{2.16}$$

The value of v (246 GeV) is known through its relation to the Fermi coupling ($v = (\sqrt{2}G_F)^{-1/2}$) which is measured to high precision from muon decay measurements. The value of λ however is not fixed by low energy observables, leaving the Higgs mass as a free parameter in the Standard Model.

2.3 Bounds on the Higgs mass

Although the Higgs mass is not predicted within the minimal SM, there are theoretical upper and lower bounds on the mass of the Higgs boson if we assume there is no new physics between the electroweak scale and some higher scale called Λ. These limits are discussed here.

2.3.1 Unitarity

In the absence of a scalar field the amplitude for elastic scattering of longitudinally polarised massive gauge bosons (e.g. $\mathrm{W_L^+W_L^-} \rightarrow \mathrm{W_L^+W_L^-}$) diverges quadratically with the centre-of-mass energy when calculated in perturbation theory and at an energy of 1.2 TeV this process violates unitarity. In the Standard Model, the Higgs boson plays an important role in the cancellation of these high-energy divergences. Once diagrams involving Higgs bosons are introduced in the gauge boson scattering mentioned above, these divergences are no longer present and the theory remains unitary and renormalisable. This cancellation only works however if the Higgs boson is not too heavy. By requiring that perturbation theory remains valid an upper limit on the Higgs mass can be extracted. With the requirement of unitarity and using all (coupled) gauge boson scattering processes it can be shown that [14]:

$$M_H < \sqrt{\frac{4\pi\sqrt{2}}{3G_F}} \sim 700 \ \mathrm{GeV/c^2} \tag{2.17}$$

It is important to note that this does not mean that the Higgs boson can not be heavier than 700 GeV/c^2. It only means that for heavier Higgs masses, perturbation theory is not valid and the theory is not renormalisable.

2.3.2 Triviality and Vacuum stability

In this section, the running of the Higgs self-coupling λ with the renormalisation scale μ is used to put both a theoretical upper and a lower limit on the mass of the Higgs boson as a function of the energy scale Λ. Although these evolution functions (β functions) have been calculated for all SM couplings in the $\overline{\mathrm{MS}}$ scheme up to two loops [15], the arguments to obtain these mass limits are presented here by using only the one loop results. At one loop the quartic coupling runs with the renormalisation scale μ as [16]:

$$\mu\frac{d\lambda}{d\mu} \equiv \beta_\lambda = \frac{3}{4\pi^2}\left[\lambda^2 + \lambda g_t^2 - g_t^4 + \mathcal{B}(g,g')\right] \tag{2.18}$$

In this expression the dominant terms are the terms involving the Higgs self-coupling λ and the top quark Yukawa coupling g_t as given in (2.15). The contribution from the gauge bosons is small and explicitly given by $\mathcal{B}(g,g') = -\frac{1}{8}\lambda(3g^2+g'^2)+\frac{1}{64}(3g^4+2g^2g'^2+g'^4)$. This expression allows to evaluate the value of $\lambda(\mu)$ relative to the coupling at a reference scale which is taken to be $\lambda(v)$.

Triviality

For large values of λ and neglecting the effects from gauge interactions and the top quark the evolution of λ is given by the dominant term in equation (2.18) that can be easily solved for $\lambda(\mu)$:

$$\mu\frac{d\lambda}{d\mu} = \frac{3\lambda^2}{4\pi^2} \quad\Rightarrow\quad \lambda(\mu) = \frac{\lambda(v)}{1-\frac{3\lambda(v)}{4\pi^2}\ln\left(\frac{\mu^2}{v^2}\right)} \tag{2.19}$$

As μ increases, $\lambda(\mu)$ increases until at $\mu{=}v\exp(2\pi^2/3\lambda(v))$ there is a singularity, known as the Landau pole. At that point $\lambda(\mu)$ becomes infinite. If the SM is required to remain valid up to some cut-off scale Λ, i.e. if we require $\lambda(\mu) < \infty$ for all $\mu < \Lambda$ this puts a constraint (a maximum value) on the value of the Higgs self-coupling at the electroweak scale (v) and therefore on the Higgs mass. Taking $\lambda(\Lambda) = \infty$ and evolving the coupling downwards we find:

$$\lambda^{\max}(v) = \frac{4\pi^2}{3\ln\left(\frac{\Lambda^2}{v^2}\right)} \quad\Rightarrow\quad M_H^2 < \frac{8\pi^2 v^2}{3\ln\left(\frac{\Lambda^2}{v^2}\right)} \tag{2.20}$$

For Λ=10^{16} GeV the upper limit on the Higgs mass is 160 GeV/c^2. This limit gets less restrictive as Λ decreases. The upper limit on the Higgs mass as a function of Λ from a computation [17] that uses the two-loop β function and takes into account the contributions from top-quark and gauge couplings is shown in figure 2.2. In that analysis the mass of the top quark was taken to be 175 GeV/c^2.

Vacuum stability

For small λ a lower limit on the Higgs mass is found by the requirement that the minimum of the potential be lower than that of the unbroken theory and that the electroweak vacuum is stable. In equation (2.18) it is clear that for small λ the dominant contribution comes from the top quark through the Yukawa coupling ($-g_t^4$). Since this contribution is negative, there is a scale Λ for which $\lambda(\Lambda)$ becomes negative. If this happens, i.e. when $\lambda(\mu) < 0$ the potential is unbounded from below. As there is no minimum, no consistent theory can be constructed. The requirement that λ remains positive up to a scale Λ yields a lower limit on $\lambda(v)$ and therefore on the Higgs mass [18].

A related argument uses the fact that radiative corrections from the top quark change the effective scalar potential from (2.5) with terms of the form $-g_f|\Phi|^4 \ln|\Phi|^2$ [19]. For small λ the top quark corrections make that the Higgs vacuum expectation value is only a local minimum. The requirement that the Higgs vacuum is the global minimum below some cut-off scale Λ puts a lower limit on the Higgs self-coupling and therefore on the Higgs mass.

A detailed evaluation, taking into account these considerations has been performed [18, 19]. The region of excluded Higgs masses as a function of the scale Λ from this analysis (taking the mass of the top quark to be 175 GeV/c^2) is also shown in figure 2.2 by the lower excluded region.

Summary of the theoretical bounds on the Higgs mass

In figure 2.2 the theoretically allowed range of Higgs masses is shown as a function of Λ.

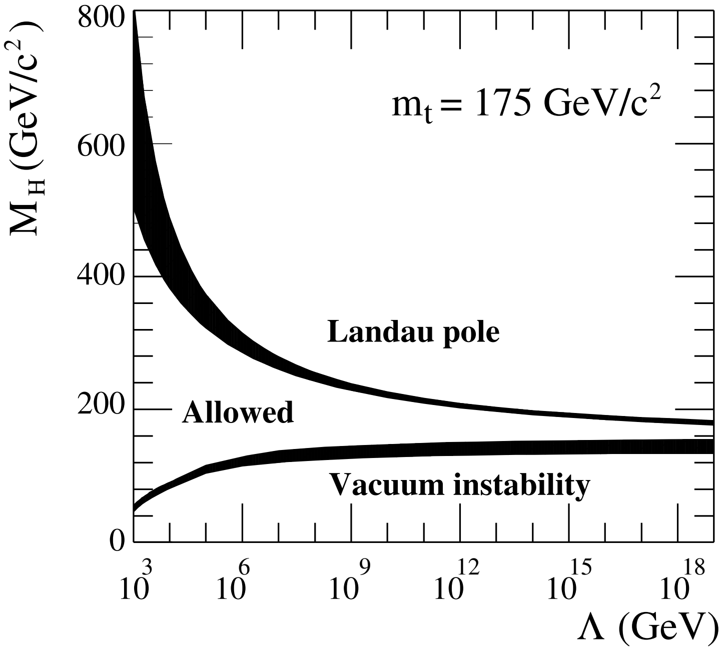


Figure 2.2: *Theoretical limits on the Higgs mass as a function the scale Λ. The region between the two bands is the allowed region. Higgs masses in the region labelled 'Landau pole' ('vacuum instability') are excluded because the Higgs quartic coupling at that energy scale and for that Higgs mass is infinite (negative).*

For a small window of Higgs masses around 160 GeV/c^2 the Standard Model is valid up to the Planck scale ($\sim 10^{19}$ GeV). For other values of the Higgs mass the Standard Model is only an effective theory at low energy and new physics has to set in at some scale Λ.

2.3.3 Indirect measurements

The electroweak gauge sector of the SM is described by only three independent parameters: g, g' and v. The predictions for electroweak observables, are often presented using three (related) variables that are known to high precision: G_F, M_Z and α_{QED}. To obtain predictions to a precision better than the experimental uncertainties (often at the per mill level) higher order loop corrections have to be computed. These higher order radiative corrections contain, among others, contributions from the mass of the top quark and the Higgs boson. Via the precision measurements one is sensitive to these small contributions and thereby to the masses of these particles.

An illustration of the possibility to estimate the mass of a heavy particle entering loop corrections is the very good agreement between the estimate of the top quark mass using only indirect measurements [20] and the direct observation [21]:

$$\text{Indirect estimate: } m_t = 181^{+11}_{-9}\ \text{GeV/c}^2 \qquad \text{Direct result: } m_t = 174.3 \pm 5.1\ \text{GeV/c}^2 \qquad (2.21)$$

The corrections connected to the top quark behave as m_t^2 whereas the sensitivity to the mass of the Higgs boson is unfortunately only logarithmic ($\sim \ln M_H$).

At present, one of the limitations in extracting an estimate of the Higgs mass is the accuracy to which the electromagnetic coupling is known at the Z mass scale. Since in the final fit to all electroweak measurements often two separate values for the corrections to $\alpha(\mathrm{M_Z})$ are shown, a bit more has to be said about these uncertainties. $\alpha(\mathrm{M_Z})$ is related to the value at zero-momentum, $\alpha(0)$, by vacuum polarisation loop corrections. The dominant uncertainty on $\alpha(s)$ comes from $\Delta\alpha^{(5)}_{\mathrm{had}}(s)$, which is the contribution of the (5) light quarks to the photon vacuum polarisation. To obtain a precise estimate of this value for the computations, experimental data from hadron production in $\mathrm{e^+e^-}$ annihilation at low energies is required. Using the new BES-II data [22] a value of 2761(36)$\cdot 10^{-4}$ is found [23]. A more theory driven extraction of $\Delta\alpha^{(5)}_{\mathrm{had}}(s)$ obtains a lower value with a much smaller error 2738(20)$\cdot 10^{-4}$ [24]. Since there is no clear favourite, the results from the two predictions are shown separately.

The results from such a global electroweak fit [20] with only the Higgs mass as a free parameter is shown in figure 2.3. The plot shows the $\Delta\chi^2$ distribution as a function of $\mathrm{M_H}$. The band indicates the remaining theoretical uncertainty in the fit.

The fit indicates a rather light mass for the Higgs boson and it can be summarised[4] by the central value with its one standard deviation and the one-sided (95% CL) upper limit:

$$\mathrm{M_H} = 88^{+53}_{-35}\ \text{GeV/c}^2 \quad \text{and} \quad \mathrm{M_H} < 196\ \text{GeV/c}^2 \quad \text{(at 95\% CL)}$$

In the next section it is explained how the Higgs boson is produced at LEP and how it decays since the real goal is to observe the Higgs boson directly.

[4] In general, as is done here, the results that summarise the combined electroweak fit correspond to the one where the value of $\Delta\alpha^{(5)}_{\mathrm{had}}(s)$ was extracted from the data.

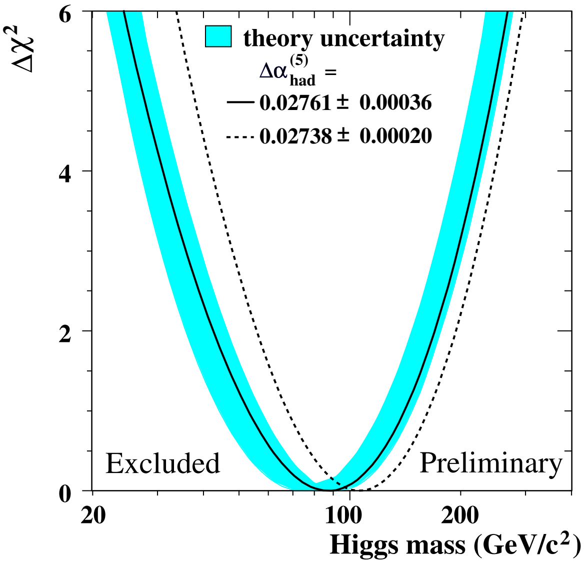


Figure 2.3: *The $\Delta\chi^2$ distribution as a function of M_H from a global fit (with only M_H as a free parameter) using the precision electroweak measurements from LEP, SLD, CDF, D0 and NuTeV.*

2.4 Higgs production and decay

In the scalar potential there are two free parameters: μ and λ. In the rest of the thesis the related variables v and $\mathrm{M_H}$ are used. Since v is known, the only free parameter left is the mass of the Higgs boson. The couplings of the Higgs boson to gauge bosons and fermions are completely determined by the Higgs mass in combination with the gauge couplings and fermion masses respectively. This allows, as is done in this section, to determine both the Higgs boson production cross section in $\mathrm{e^+e^-}$ collisions and its decay widths as a function of $\mathrm{M_H}$.

2.4.1 Higgs production at LEP

The coupling of electrons to the Higgs is very small, making the cross section for direct Higgs production through $\mathrm{e^+e^-}$ annihilation tiny. The dominant mechanism for the production of a SM Higgs boson at LEP is therefore through so-called Higgsstrahlung[5]. In this process a virtual Z is produced that subsequently radiates a Higgs boson:

$$\mathrm{e^+e^-} \to \mathrm{Z}^* \to \mathrm{ZH} \tag{2.22}$$

The corresponding Feynman diagram is shown in the left plot of figure 2.4 and at tree level the cross section is given by [25, 26]:

$$\sigma(\mathrm{e^+e^-} \to \mathrm{ZH}) = \frac{G_F^2 M_Z^4}{48\pi s}(1 - 4\sin^2\theta_W + 8\sin^4\theta_W)\,\sqrt{\lambda}\,\frac{\lambda + 12M_Z^2/s}{(1 - M_Z^2/s)^2} \tag{2.23}$$

In this expression $\lambda = (1 - (M_H^2 + M_Z^2)/s)^2 - (2M_H M_Z/s)^2$ is a two-particle phase space function. Although the electroweak corrections are small (only a few %), photon radiation [27]

[5] In addition there is also Higgs production through boson-boson fusion, but its contribution to the total cross section is an order of magnitude smaller than the Higgsstrahlung process for most Higgs masses.

significantly lowers the total cross section (see [28] for a review). In the right plot of figure 2.4 the cross section is shown as a function of the Higgs mass at three different centre-of-mass energies taking into account the various corrections. The ZH cross section is of the order of 0.5 pb, but drops rapidly as the sum of the two boson masses approaches the kinematic limit.

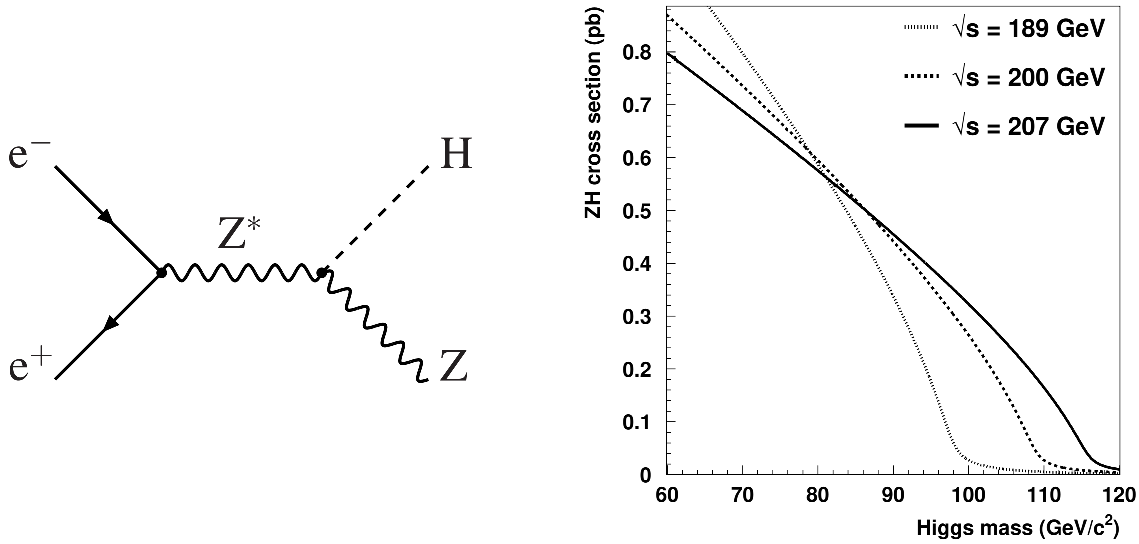


Figure 2.4: *The left plot shows the Higgsstrahlung Feynman diagram, the dominant Higgs production mechanism at LEP. The right plot shows the ZH cross section as a function of the mass of the Higgs boson at various centre-of-mass energies.*

2.4.2 Higgs decay

This section describes the decay of the Higgs boson into fermion pairs (leptons and quarks) and the decay into a pair of gauge bosons. The decays are shown in figure 2.5.

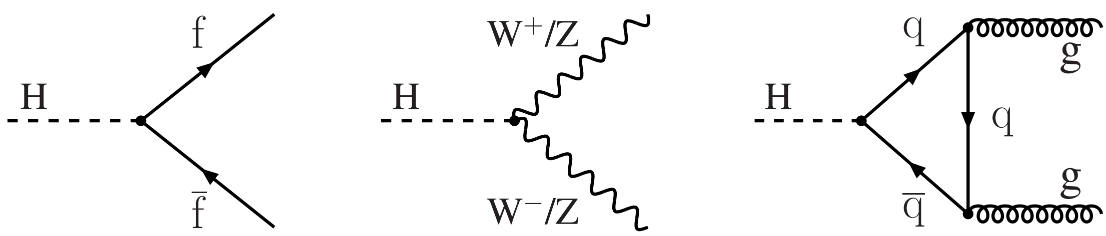


Figure 2.5: *This figure shows three decay modes of the Higgs. The left plot shows the decay into fermion pairs (leptons and quarks). The decay into a pair of electroweak gauge bosons is shown in the middle plot and the right plot shows the production of a gluon pair.*

Higgs decay into fermion pairs

In the Born approximation, the decay width into a pair of charged leptons is given by [29]:

$$\Gamma(H \to l^+l^-) = \frac{G_F m_l^2}{4\sqrt{2}\pi}\, M_H \beta_l^3 \tag{2.24}$$

In this expression m_l is the mass of the lepton and $\beta_l = \sqrt{1 - 4m_l^2/M_H^2}$, the velocity of the outgoing leptons. The Higgs boson decay into quarks is enhanced by a colour factor (N_c=3) and also receives large QCD corrections. Most of these corrections can be absorbed [30] by using the running quark mass instead of their pole mass in the expression for the width. In the $\overline{\mathrm{MS}}$ scheme for QCD with five flavours the expression for the width at NLO can be written as [31]:

$$\Gamma(H \to q\bar{q}) = \frac{3G_F m_q^2(M_H)}{4\sqrt{2}\pi}\, M_H\, \beta_q^3 \left[1 + \frac{17}{3}\left(\frac{\alpha_s(M_H)}{\pi}\right) + \mathcal{O}\left(\frac{\alpha_s(M_H)}{\pi}\right)^2\right] \tag{2.25}$$

In this expression, both the quark masses and the strong coupling constant should be evaluated at the scale given by M_H. The full NNLO expression can be found in [31]. These corrections are large as is illustrated by the fact that at full NNLO, the $\mathrm{b\bar{b}}$ decay width of a 100 GeV/c^2 Higgs boson is around half of what is expected using the quark pole-mass in equation (2.24). The branching fraction of the Higgs boson into various fermion pairs as a function of M_H is shown in the left plot of figure 2.6. For Higgs masses accessible at LEP, the dominant decay mode is clearly into a pair of b-quarks ($\sim$ 80-85%).

Higgs decay into boson pairs

The Higgs decay into gluons is mediated through quark loops as is shown in the right plot of figure 2.5. Due to its large mass the top quark provides the dominant contribution. At LEP, with Higgs masses well below the $\mathrm{t\bar{t}}$ threshold, the two-gluon width can be written as [32]:

$$\Gamma(\mathrm{H} \to \mathrm{gg(g)}, \mathrm{q\bar{q}(g)}) = \frac{G_F \alpha_s^2(M_H)}{36\sqrt{2}\pi^3}\, M_H^3 \left[1 + \frac{215}{12}\left(\frac{\alpha_s(M_H)}{\pi}\right)\right] \tag{2.26}$$

This expression takes both the virtual and real QCD corrections into account. Note that the QCD corrections to the gluonic width of the Higgs are large ($\mathrm{d}\Gamma/\Gamma \sim 60\%$) and positive. The branching fraction is however never more than 5.5% as is shown in the left plot of figure 2.6.

The Higgs decay into a pair of electroweak vector bosons, $\mathrm{H} \to \mathrm{VV}^*$, with V=W or Z can occur at masses well below $2M_V$ if one of the bosons is produced off-shell. The decay width into two W bosons as a function of the Higgs mass can be written as [33]:

$$\Gamma(\mathrm{H} \to \mathrm{WW}^*) = \frac{3G_F^2 M_W^4}{16\pi^3}\, M_H\, \mathcal{R}(x), \tag{2.27}$$

where $\mathcal{R}(x) = \frac{3(1-8x+20x^2)}{\sqrt{4x-1}} \arccos(\frac{3x-1}{2x^{3/2}}) - \frac{1-x}{2x}(2 - 13x + 47x^2) - \frac{3}{2}(1 - 6x + 4x^2)\ln(x)$ using $x = M_W^2/M_H^2$. The branching fractions of the Higgs into WW^* and ZZ^* as a function of M_H are shown in figure 2.6. Although this decay channel is dominant for Higgs masses above 135

GeV/c^2, for masses accessible at LEP the branching fraction is small[6]. Due to the large Z boson mass and the smaller neutral current couplings, the ZZ channel will always be smaller.

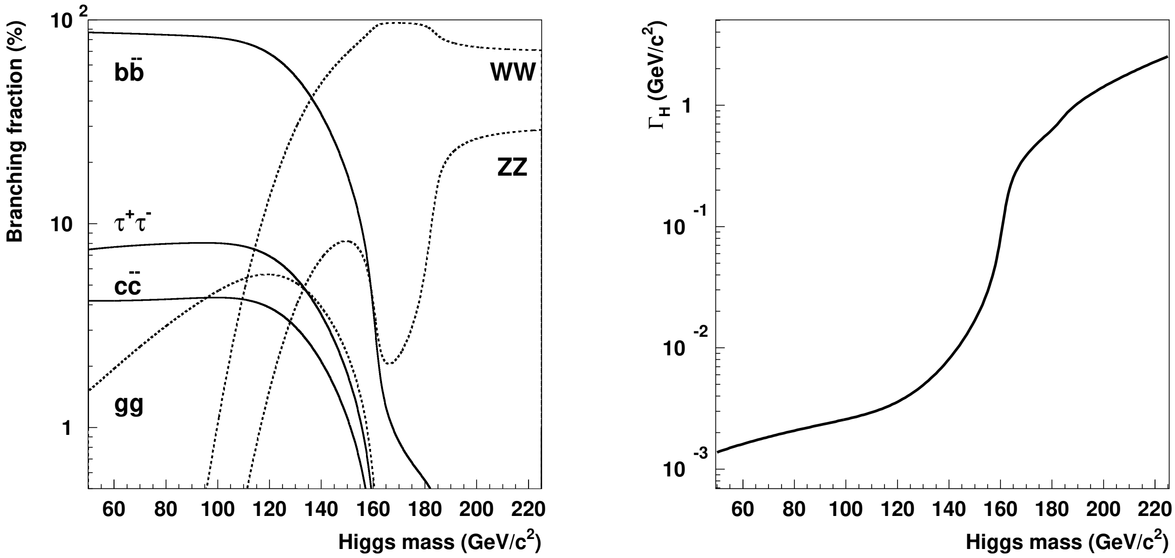


Figure 2.6: *The left plot shows the dominant branching fractions (in %) of the Higgs boson into fermion pairs (solid lines) and boson pairs (dashed lines) as a function of its mass. The right plot shows the total width of the Higgs boson as a function of its mass.*

Total width of the Higgs boson

The total width of the Higgs boson as a function of its mass is shown in the right plot of figure 2.6. For a mass between 70 and 115 GeV/c^2 the width is very small and ranges between roughly 2 and 3 MeV/c^2. This is much smaller than the typical experimental resolution for reconstructing the Higgs mass which is in the order of a few GeV/c^2. The large life-time of the Higgs boson compared to that of the Z boson (Γ_Z=2.49 GeV/c^2) also means that the Higgs and Z boson decay independently, i.e. the Higgs decays after the Z boson has finished its hadronisation.

2.4.3 ZH final states

The decay properties of the Z boson are predicted in the SM and have been measured to high precision at LEP1: the Z boson can decay either hadronically (BR(Z $\rightarrow$ q$\overline{\text{q}}$) = 69.9%), into a pair of neutrinos (BR(Z $\rightarrow \nu\overline{\nu}$) = 20.0%) or into a lepton pair (BR(Z $\rightarrow l^+l^-$) = 10.1%). As was shown in the previous section, the decay properties of the Higgs boson depend on the mass of the Higgs boson. To list the relative branching ratios into the various possible ZH final states we have taken as an approximation BR(H $\rightarrow$ q$\overline{\text{q}}$) = 92.5% and BR(H $\rightarrow \tau^+\tau^-$) = 7.5%. Combined with the numbers from the Z boson decay this results in the numbers shown in table 2.3.

[6] Although for a Higgs mass of 115.6 GeV/c^2 it is already the second largest decay channel.

Final state	$q\bar{q}q\bar{q}$	$q\bar{q}\nu\bar{\nu}$	$q\bar{q}\,l^+l^-$	$q\bar{q}\tau^+\tau^-$	other
Fraction (in %)	64.7	18.5	6.2	8.4	2.3

Table 2.3: *Relative branching fractions (in %) for final states present in ZH events. In the fourth column l stands for electron and muon only and the $q\bar{q}\ \tau^+\tau^-$ final state covers both the Z and Higgs decay into $\tau^+\tau^-$. The 'other' final states are $\tau^+\tau^-x\bar{x}$ final states ($x\bar{x}$ = all but $q\bar{q}$).*

Here we should also note that a Higgs decay into a gluon pair has an experimental signature that is very similar to that of a decay into a quark anti-quark pair. To obtain the numbers given in table 2.3, the small fraction of H→ gg events has been combined with the Higgs decays to $q\bar{q}$.

In this thesis the search for the Higgs boson is performed in the fully hadronic final state. It covers almost 2/3 of all ZH final states and its most important characteristic is that more than 90% of these events have at least two b-quarks in the final state. It is mainly this characteristic that generates sensitivity for the small ZH signal among the large SM 4-quark backgrounds that are described briefly in the next section.

2.5 Background processes

The background for the ZH search in the 4-quark final state can be divided into a 4-fermion and a 2-fermion background.

2.5.1 4-fermion background

The 4-fermion background corresponding to $e^+e^- \rightarrow\ 4f + n(\gamma)$ events is simulated using all 4-fermion diagrams (+interferences) using the `EXCALIBUR` [34] Monte Carlo programme. The total 4-fermion cross section is around 18 pb. The dominant diagrams for producing 4-quark final states are the CC03 and NC02 (sets of) diagrams that describe the resonant pair production of electroweak gauge bosons (W^+W^- and ZZ respectively). These two processes will be briefly discussed.

WW production

One of the consequences of the non-Abelian structure of the $SU(2)$ symmetry group is that there are interactions between the electroweak gauge bosons. When the term describing three boson vertices is written in terms of the mass eigenstates using expressions (2.10) and (2.11) a $W^+W^-\gamma$ and W^+W^-Z vertex can be identified. At LEP2, at centre-of-mass energies above the W^+W^- threshold, W bosons are therefore pair produced via the three diagrams (the CC03 diagrams) as shown in figure 2.7. As it involves two diagrams with three gauge boson couplings it has been an ideal process to firmly establish the non-Abelian structure of the SM. Combined with the precision measurement of the W mass, the physics involving W bosons is one of the main physics topics at LEP2.

The predicted cross section as a function of the centre-of-mass energy is shown in the left plot of figure 2.9 where it is compared to the LEP combined measurements. The band indicates the

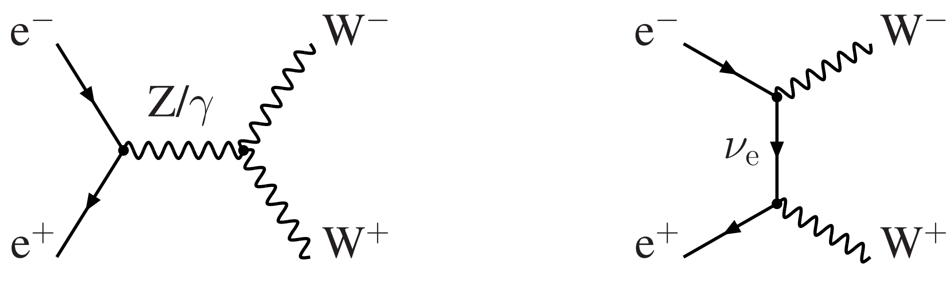


Figure 2.7: The CC03 diagrams describing W boson pair production.

theoretical uncertainty of 0.5% from the computations of `RACOONWW` [35] and `YFSWW` [36]. In the SM the coupling of the W boson to fermions is fixed and the W can either decay into a pair of quarks ($\sim$ 2/3) or into a lepton and a neutrino ($\sim$ 1/3). This results in three different possible final states in W^+W^- events as is shown in table 2.4.

Final state	$q\overline{q}q\overline{q}$	$q\overline{q}\, l\nu$	$l\nu\, l\nu$
Fraction (in %)	45.6	43.9	10.5

Table 2.4: *Relative branching fractions (in %) for W^+W^- final states.*

The W^+W^- background for the ZH search in the 4-quark final state ($W^+W^- \rightarrow q\overline{q}q\overline{q}$) is very large ($\sim$ 7.8 pb) when compared to the Higgs signal. One of the most important features of fully hadronic W^+W^- events, when considered as background for ZZ cross section measurements or ZH searches, is that there is hardly ever a b-quark present in the final state. The decay of the W into a quark anti-quark pair ($q_i\overline{q}_j$) can be written at Born level as: $\Gamma(W \rightarrow q_i\overline{q}_j) \propto |V_{ij}|^2$, where V_{ij} is the corresponding element of the CKM matrix. As there is not enough energy for the process $W \rightarrow b\overline{t}$ the main decay modes of the W boson are into $u\overline{d}$ or $c\overline{s}$ (the other diagonal elements of the CKM matrix). The rate of single b-quark production is proportional to the off-diagonal element V_{cb} ($\approx$ 0.04) squared, resulting in final states with b-quarks in only 0.3% of the fully hadronic W^+W^- events. As will be shown in the next chapters, demanding the presence of b-quarks in the final state provides a powerful tool to reduce the background from W^+W^- events.

ZZ production

Due to the structure constants of the $SU(2)$ group there is no ZZZ vertex at tree level in the SM. The pair production of resonant Z bosons is therefore described by the two diagrams shown in figure 2.8 (the NC02 diagrams).

In the SM, the couplings of the electroweak gauge bosons to fermions are completely fixed given $\sin^2\theta_W$. The expression for the total cross section for ZZ production in e^+e^- collisions in terms of these couplings and other properties of the Z boson was first derived in [37], followed by a more detailed evaluation (amplitudes for polarised beams and all combinations of polarisations of final state particles) in [38]. Using the precise measurements of the properties of the Z boson

Figure 2.8: The NC02 diagrams describing Z boson pair production.

at LEP1, the ZZ cross section can be accurately predicted. However, since the ZZ cross section is small, the LEP combined experimental uncertainty is at the level of 5%, the computation of the NC02 cross section has not received as much attention as has that of the CC03. A comparison [39] between various programmes and computations (for example `YFSZZ`[40] and `ZZTO`[39]) was done in the framework of the LEP2 Monte Carlo workshop where it was concluded that the current theoretical uncertainty is in the order of 2%, but that in principle a more detailed computation (similar to what was done for the double resonant $\mathrm{W^+W^-}$ (CC03) process) could be performed to bring the uncertainty down to around 0.5%. The cross section as a function of the centre-of-mass energy is shown in the right-hand plot of figure 2.9, where the band indicates the theoretical uncertainty.

Using the decay properties of the Z boson as given in the previous section, the relative branching fractions of the various final states possible in double resonant Z boson pair production can be easily computed. The results are given in table 2.5.

Final state	$\mathrm{q\bar{q}q\bar{q}}$	$\mathrm{q\bar{q}}\nu\bar{\nu}$	$\mathrm{q\bar{q}}\, l^+l^-$	$l^+l^-\nu\bar{\nu}$	$\nu\bar{\nu}\nu\bar{\nu}$	$l^+l^-\, l^+l^-$
Fraction (in %)	48.9	28.0	14.1	4.0	4.0	1.0

Table 2.5: *Relative branching fractions (in %) for final states present in ZZ events.*

The 4-quark final state represents about half of the ZZ final states and its cross section is therefore in the order of 0.5 pb, similar to that of the ZH cross section.

The measurement of the ZZ cross section is of particular interest for various reasons. The first reason is that in theories beyond the SM, the process $\mathrm{e^+e^-} \rightarrow \mathrm{Z^*} \rightarrow \mathrm{ZZ}$ is allowed when either the exchanged Z is different from the final state (SM) Z bosons or if there are anomalous neutral trilinear gauge boson couplings (NTGCs). In this thesis these aspects are not studied. The second reason is that it is a process that is very similar to that of the ZH process discussed in section 2.4. ZZ events are therefore often referred to as *irreducible* background. Showing that one can measure the ZZ cross section proves that one is sensitive to a possible ZH signal (and to any possible strong systematic effects related to this final state) as well. This aspect is studied in this thesis in chapter 6 where the measurement of the NC02(ZZ) cross section with 4 quarks in the final state is presented. Important to note is that in 38.7% of these ZZ final states there are at least 2 b-quarks, to be compared to (85-95)% in ZH events (see section 5.1.1).

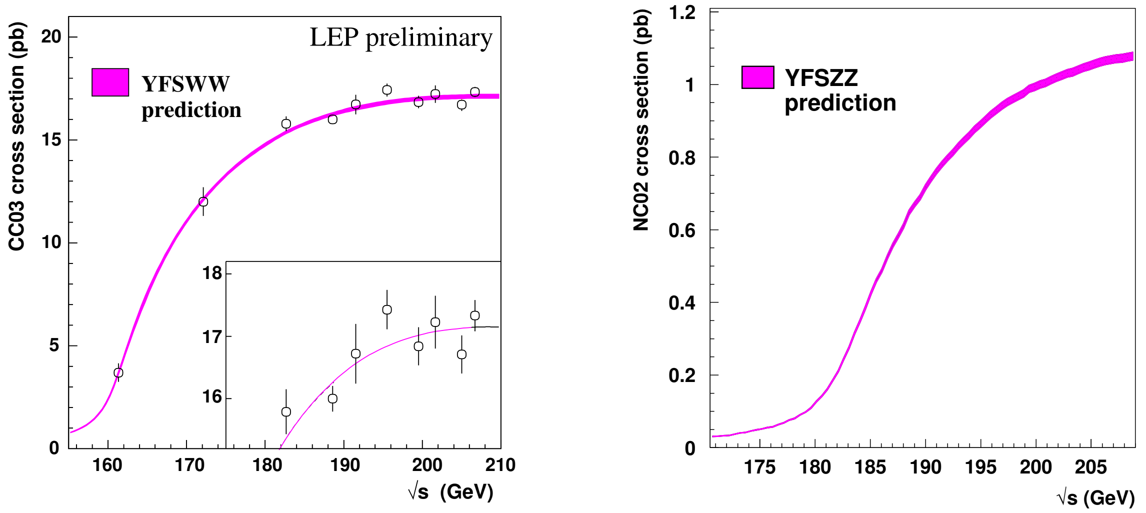


Figure 2.9: *The cross section for on-shell vector boson pair production as a function of the centre-of-mass energy for W^+W^- (left plot) and ZZ (right plot) production. For the W^+W^- cross section also the LEP combined measurements are shown.*

Other 4-fermion backgrounds

The only other remaining (small) background from 4-fermion processes are hadronic $Z\gamma^*$ and $\gamma^*\gamma^*$ events. These events are produced via diagrams as in figure 2.8, with one or two of the Z bosons replaced by virtual photons (γ^*), where the 'masses' of these photons can extend to above M_Z. Although there is a difference between the photon coupling and the Z coupling to quarks, for masses close to the Z mass the signature of these events is almost identical to that of ZZ events. For low γ^* masses (the bulk of these events), the characteristics of the final state are similar to those of the 2-fermion $q\overline{q}(\gamma)$ background that is described below.

2.5.2 2-fermion background

The process $e^+e^- \rightarrow Z^* \rightarrow q\overline{q}$ as shown in figure 2.10 has a large cross section (of the order of 100 pb at LEP2 centre-of-mass energies).

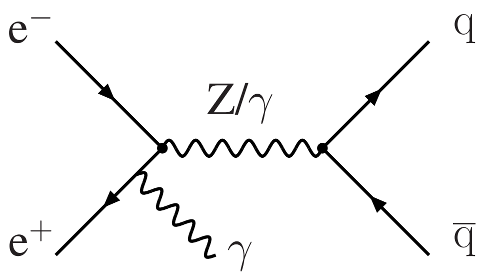


Figure 2.10: *Feynman diagram describing $q\overline{q}(\gamma)$ production in e^+e^- annihilations.*

The background for the fully hadronic ZZ and ZH analyses originates from the class of $q\overline{q}(\gamma)$ events that have no significant initial state photon radiation and in addition have an apparent 4-jet structure caused by radiation of hard gluons off the final state quarks and their subsequent fragmentation. Therefore only a small fraction ($\sim 2\%$) of the $q\overline{q}(\gamma)$ events remains finally as a real background. The aspects of quark fragmentation and initial state radiation are described in more detail in section 4.1 and 4.3.2 respectively. At the start of the analysis this background is smaller than that from W^+W^- events, but due to the presence of b-quarks in the final state ($\sim 16\%$, see section 5.1.1), the $q\overline{q}(\gamma)$ background is the dominant contribution to the total background at the final analysis level both in the measurement of the ZZ cross section and in the search for a high mass Higgs boson.

The computation of the cross section and the simulation of events is done using the PYTHIA [41] Monte Carlo programme. For centre-of-mass energies far above the Z pole this cross section behaves as $1/s$.

2.6 Summary

In this chapter the structure of the SM and the mechanism to obtain masses for the electroweak vector bosons has been discussed. This (Higgs) mechanism predicts in addition a massive scalar particle with unknown mass. The production cross section at LEP and decay of this last missing particle in the SM as a function of its mass were discussed in section 2.4 and in the last section of this chapter the SM background processes also resulting in a 4-quark final state were reviewed.

Chapter 3

Experimental set-up

The *Large Electron Positron* collider (LEP) [42] is the world's largest e^+e^- storage ring and part of the accelerator complex of CERN, the European Laboratory of Particle Physics near the city of Geneva. At four points along the ring the beams collide, where large experiments are set up to study the e^+e^- interactions: ALEPH [43], DELPHI [44], L3 [45] and OPAL [46]. In this chapter both the LEP accelerator and the DELPHI detector are described.

3.1 The LEP collider

LEP as accelerator and storage ring

With a circumference of 26.658 km and a maximum centre-of-mass energy of 209.2 GeV LEP is the largest and most powerful e^+e^- accelerator in the world. The ring has a quasi-circular shape consisting of 8 short (about 500 m long) straight sections connected by curved sections and is situated in a tunnel between 50 and 150 meters underground. In accelerating particles to the highest energies LEP is only the last step in the acceleration chain where the older accelerators of CERN function as an injector system.

Electrons are produced using an electron-gun and accelerated in the LEP Injector Linac (LIL-I) to 200 MeV and subsequently collided onto a Tungsten target. The positrons, produced by conversion in these collisions, are collected and, together with the electrons, further accelerated to 600 MeV (LIL-II) and stored (in bunches) in the Electron-Positron-Accumulator ring (EPA). Two further acceleration steps are done in the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) where the bunches are accelerated to energies of 3.5 and 22 GeV respectively and finally injected into LEP. The complete turnaround time (filling LEP and complete acceleration) takes typically around 45 minutes. An overview of the full accelerator chain is given in figure 3.1.

In LEP, electrostatic separators separate the e^- from the e^+ beam allowing them to be contained and accelerated in the same elliptically shaped vacuum beam-pipe. Guiding the beams through LEP requires a complex system of dipole (to bend) and quadrupole (to focus) magnets. Once LEP is filled, the electron and positron bunches are accelerated using (super-conducting) Radio-Frequency (RF) cavities located in the straight sections. Since charged particles undergoing centripetal acceleration (with radius r) radiate energy proportional to E^4/r (synchrotron radiation)

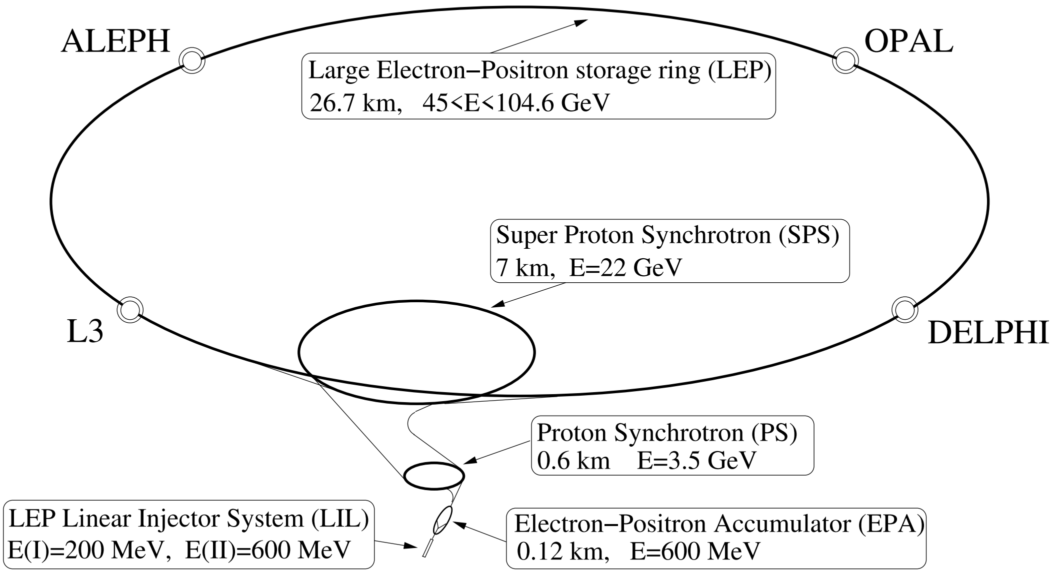


Figure 3.1: *An overview of the full CERN acceleration chain for the electrons and positrons.*

the cavities must, in addition to supplying power for the acceleration, compensate for the energy loss of the particles. For an electron in a circular accelerator the size of LEP the energy loss is:

$$\Delta E = 2.09 \times 10^{-8}\,\mathrm{E}^{\,4}_{\,\mathrm{beam}} \quad \text{GeV per turn} \tag{3.1}$$

In this expression the beam energy is given in GeV. Each electron (when LEP is operated at an energy of 100 GeV per beam) loses more than 2 GeV per turn. With $4 \cdot 10^{12}$ particles orbiting LEP at almost the speed of light the radiated power is around 20 MW.

After operating close to 91.2 GeV, the Z-pole, during the years 1989-1995 (LEP1), where data from millions of Z bosons were collected, the copper RF-cavities were replaced by super-conducting ones. In the years 1996-2000 (LEP2) the available RF-power increased each year. Starting at 130 GeV in 1995, LEP scanned a wide range of centre-of-mass energies reaching a maximum on the 10th of October 2000 with collisions at a centre-of-mass of 209.2 GeV [1].

Luminosity

Each beam (e^- and e^+) consists of 4 bunches each containing about $5 \cdot 10^{11}$ particles. At four points in the ring the beams are squeezed and collided head-on. In the collision, the interaction rate for a specific process depends on the cross section for that process (σ_{proc}) and the number of particles that traverse the interaction region per second per unit surface:

$$\frac{\mathrm{d}N_{\mathrm{proc}}}{\mathrm{d}t} = \mathrm{L}(t)\sigma_{\mathrm{proc}} = \frac{N_{\mathrm{e}^-} N_{\mathrm{e}^+} n_{\mathrm{bunch}} f}{4\pi\sigma_x\sigma_y}\,\sigma_{\mathrm{proc}} \tag{3.2}$$

[1] Only for 2 minutes and 43 seconds with an integrated luminosity of 0.0006 pb^{-1}, but nevertheless ... a record.

In this expression L(t) is called the *instantaneous luminosity*, N_{e^-} and N_{e^+} are the number of e^- and e^+ per bunch, n_{bunch} is the number of bunches and f is the bunch revolution frequency in LEP. The size of the beam at the interaction point in the plane perpendicular to the direction of motion is represented by σ_x and σ_y with typical values of 150 μm and 10 μm in x and y respectively. The instantaneous luminosity changes continuously according to the machine conditions (in turn this depends on parameters like the available RF-power, stability of electronics and the skill of the LEP operators) and at LEP2 it is typically a few times 10^{31} cm^{-2}s^{-1} and up to 10^{32} cm^{-2}s^{-1} at the beginning of the fill.

The total number of events produced by a specific process during a data-taking period can be determined by integrating the collected luminosity over this period:

$$N_{\mathrm{proc}} = \sigma_{\mathrm{proc}} \int \mathrm{L}(t)\mathrm{d}t = \sigma_{\mathrm{proc}}\,\mathcal{L} \tag{3.3}$$

The most precise estimate of the integrated luminosity $\mathcal{L}$ is not obtained by closely monitoring the machine parameters entering equation 3.2, but by using a reference process with a clear experimental signature and a large cross section that can be calculated to high precision. At LEP the process best suited is small angle $\mathrm{e}^+\mathrm{e}^-$ Bhabha scattering allowing the integrated luminosity to be computed by: $\mathcal{L} = N_{\mathrm{Bhabha}}/\sigma^{\mathrm{eff}}_{\mathrm{Bhabha}}$, where $\sigma^{\mathrm{eff}}_{\mathrm{Bhabha}}$ is the effective Bhabha cross section after corrections for acceptance and efficiency. The integrated luminosities collected by the DELPHI detector at the various centre-of-mass energies during the period 1996-1999 are given in table 3.1. $\mathcal{L}$s are given in units of pb^{-1} and have a typical uncertainty of 6 ‰.

year	1996		1997	1998	1999			
$\sqrt{s}$ (GeV)	161	172	183	189	192	196	200	202
$\mathcal{L}$ (pb^{-1})	10.0	10.0	54.7	158	25.9	76.9	84.3	41.1

Table 3.1: *Integrated luminosities collected by the DELPHI detector at various centre-of-mass energies during the period 1996-1999. The data collected at centre-of-mass energies below 180 GeV are not analysed in this thesis.*

Data taking in 2000

In the year 2000 no additional RF-power was available and the energy could not be increased significantly. In that year a special effort was made to optimise the sensitivity for a possible Higgs signal, sacrificing integrated luminosity for the highest centre-of-mass energy. A few tricks were used to increase the beam energy: enlarging slightly the bending radius of LEP, reinstalling old copper cavities etc., resulting in a 3.5 GeV increase. When operated at maximum energy (no spare RF-power) each trip of an RF-cavity is fatal and the beams are lost, resulting in an average beam life-time of only 14 minutes (compared to several hours during normal operations). Optimal sensitivity for the Higgs boson is however obtained by an operation scheme that combines a relatively long stable collision period (one or two spare klystrons) with a short period at the maximum energy available at that moment. This scheme is called the mini-ramp scheme [47]. The various centre-of-mass energies LEP operated at are shown in figure 3.2. For most analyses the 2000 data are divided in two regions: centre-of-mass energies below and above 205.5 GeV, identified as '205' and '207' respectively in the remainder of this thesis.

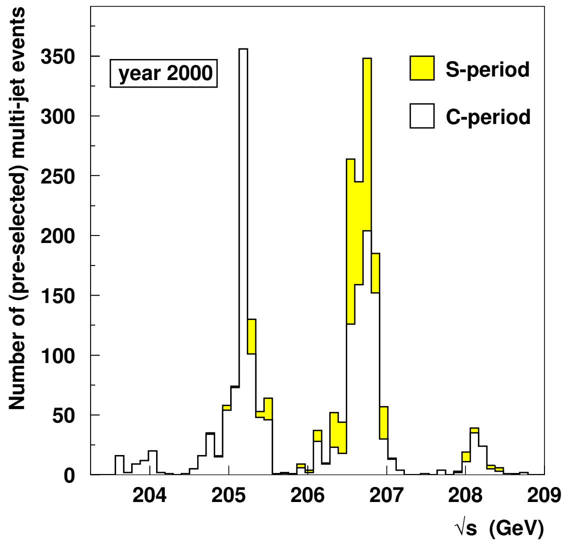


Figure 3.2: *Number of (pre-selected) multi-jet events collected at the various centre-of-mass energy ranges of LEP operation by the DELPHI detector in the year 2000.*

After the 1st of September DELPHI had to be operated without 1/12 of the central tracking detector. In the rest of the thesis the period with the full tracking system operational is called the C-period. The period after the 1st of September is labelled as the S-period. The integrated luminosity collected during the different periods is shown in table 3.2.

	$\mathcal{L}$ (pb^{-1})		
$\sqrt{s}$ (GeV)	<205.5 (205)	>205.5 (207)	all energies
C-period	75.6	87.8	163.4
S-period	6.3	54.3	60.7
year 2000	82.0	142.2	224.2

Table 3.2: *Collected integrated luminosity by the DELPHI detector in the year 2000. The integrated luminosity collected in the period before (C-period) and after (S-period) the 1st of September are given separately.*

Contrary to many other analyses, the impact of the distorted reconstructed energy flow in the event for multi-jet analyses is small. After a special effort of the DELPHI tracking group, the effect can be summarised as loosing on average a single charged track per event. The reconstructed energy remained nearly unaltered, due to a trade-off between the increased(decreased) reconstructed neutral(charged) energy. The effect on the final sensitivity of the analysis is found to be negligible.

3.2 The DELPHI detector

The DELPHI detector *(DEtector with Lepton, Photon and Hadron Identification)* is a detector surrounding one of the e^+e^- interaction regions. At LEP the laboratory system coincides with the e^+e^- centre-of-mass system resulting in detectors symmetric around the interaction point covering almost the full solid angle. The DELPHI detector consists of a cylindrical central part *(barrel)* covering polar angles between 40° and 140° and 2 circular 'plugs' *(end caps)* that allow particle detection down to polar angles of 2°. The DELPHI detector consists of various sub-detectors as can be seen in the schematic (cut-open) view of the detector in figure 3.3.

The coordinate system used in DELPHI has the z-axis pointing along the electron direction, the x-axis pointing towards the centre of LEP and the y-axis pointing upwards. The polar angle with respect to the z-axis is called θ, ϕ is the azimuthal angle and R=$\sqrt{x^2+y^2}$.

The sub-detectors form two sub-systems: the *tracking system* close to the interaction point to reconstruct the 3-dimensional trajectories, charge and momentum of charged particles and the *calorimeter system* to measure particle energies. In the next sections both systems and their characteristics will be briefly discussed. A more detailed description of the DELPHI detector and the performance of the sub-systems can be found elsewhere ([44] and [48]).

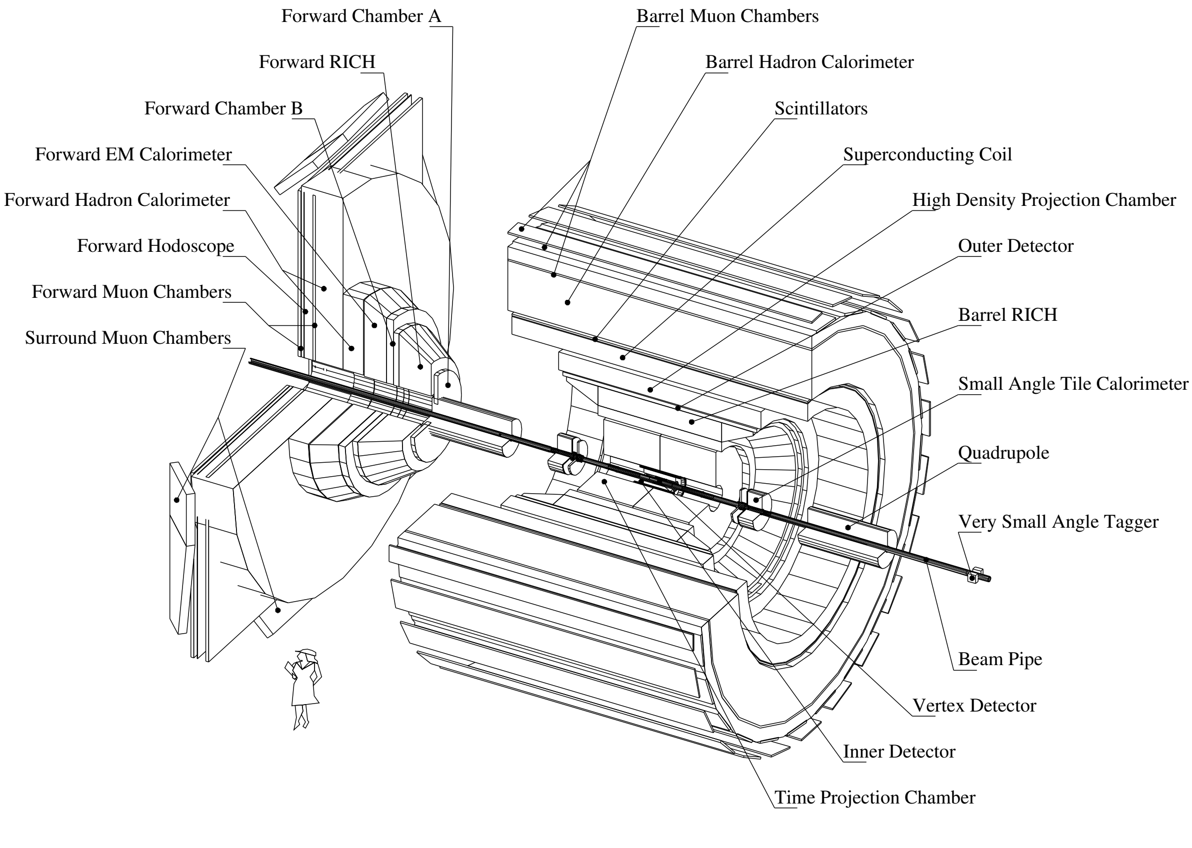


Figure 3.3: *Schematic view of the DELPHI detector. The barrel and a single end cap are shown.*

3.2.1 Tracking detectors

The tracking system reconstructs the 3-dimensional trajectory of charged particles. It consists of a set of detectors close to the interaction point inside a uniform magnetic field of 1.23 T parallel to the beam axis. This field is provided by a large super-conducting solenoid with a radius of 2.6 m. The trajectory of a charged particle can be reconstructed using its ionisation when traversing a medium and its reconstructed curvature allows determination of the charge and momentum of the particle. To minimise the influence on the direction of the particle tracking devices must be as transparent as possible. As a result the sensitive materials are gases or very thin layers of higher density materials. The various detectors that make up the tracking system in DELPHI are:

- Micro-vertex detector
  The detector closest to the interaction point is the micro-vertex detector [49] displayed in figure 3.4. It is composed of a Silicon Barrel (VD) and a Very Forward Tracker (VFT).

  The VD consists of 3 cylindrical layers of micro-strip silicon detectors located at 6.6, 9.2 and 10.6 cm from the beam axis, covering polar angles from 25° to 155°. The inner layer is composed of 2x20 modules and the outer layers are composed of 2x24 modules. All channels give either Rϕ measurements (in all three layers) or z (in two and a half layers). Within each plane of detectors and between different planes there is significant overlap between sensitive silicon regions, providing redundancy used in the (internal) alignment. The single hit precision for 1 layer in Rϕ is 8 μm. The single hit precision in z is a function of the incidence angle, reaching 9 μm for tracks incident at polar angles of 90°. The maximum number of hits on a track can be up to 6 in Rϕ (4 in z). The VD is an important detector when determining if particles originated from the primary interaction point, which is a key ingredient of analyses involving heavy quarks.

  The upgrade of the DELPHI vertex detector was completed in 1997 with the installation of the Very Forward Tracker (VFT) allowing track reconstruction at smaller polar angles. Each of the VFT end caps consists of two layers of silicon pixel detectors[2] and two layers of mini-strip detectors covering polar angles between 11° and 25°.

- Inner Detector (ID)
  The ID is composed of two parts: an inner high-resolution drift chamber (JET chamber, located at radii between 12 cm and 23 cm) surrounded by 5 concentric layers of straw-tubes (trigger layer), covering polar angles from 15° to 165°. The drift (wire) chamber is divided into 24 azimuthal sectors each providing up to 24 Rϕ measurements per track with a resolution of 85 μm each. The straw tubes provide additional (fast) Rϕ measurements used in the first level trigger. Due to the stereo angle in 2 layers also some z information can be extracted.

- Time Projection Chamber (TPC)
  The TPC is the main tracking device in DELPHI. It is located just outside the Inner Detector at radii between 35 cm and 111 cm and covers polar angles from 20° to 160°. The detector consists of two drift volumes separated by a high voltage plane at $z = 0$. Particles traversing

[2] DELPHI was the first collider experiment where pixels were used in the tracking.

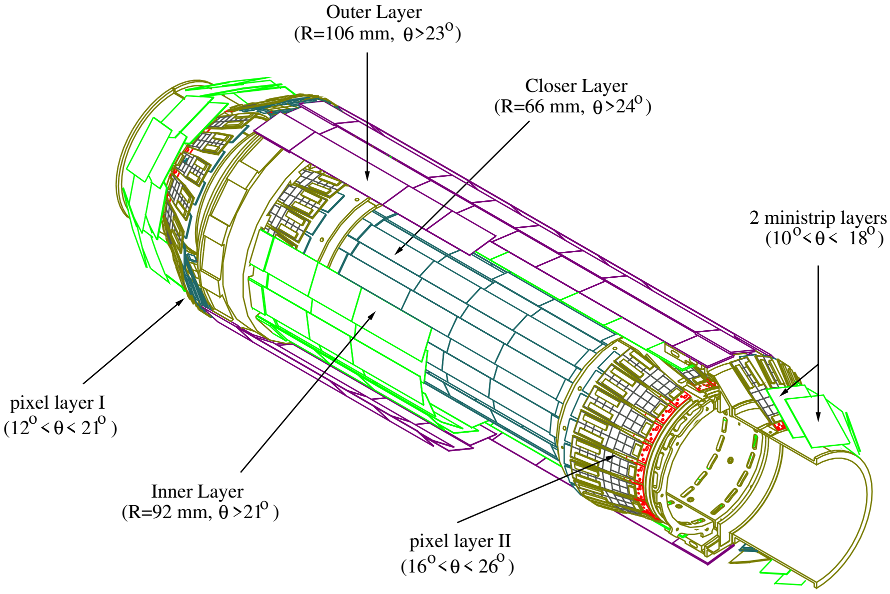


Figure 3.4: *The DELPHI vertex detector*

the chamber ionise the gas. The electrons knocked free in the gas drift towards the end-plates of the TPC due to the electric field parallel to the beam axis. Each end-plate consists of 6 multi-wire proportional chambers (MWPC's) that detect the drifting electrons. Each MWPC has 192 anode wires and 16 circular cathode pad rows which results in a maximum number of 16 3-dimensional position measurements with a resolution of 250 μm in Rϕ and 880 μm in z (using drift time information). In addition to position information the TPC performs a dE/dx measurement used in particle identification.

- Outer Detector (OD)
  The OD is located at about 2 meters away from the interaction point and consists of 5 layers of drift-tubes subdivided in 24 sectors, covering polar angles between 40° and 140°. The single point resolution is 110 μm in Rϕ and 3.5 cm in z. The OD is important in momentum estimation (see table 3.3) since it provides a large lever arm in the track reconstruction.

- Forward Chambers (FCA and FCB)
  In the forward region (down to 11°) 2 detectors provide tracking information. The FCA is located at $|z|$=160 cm and consists of 6 layers of drift chambers with combined track element precisions of $\sigma(x)$=290 μm and $\sigma(y)$=240 μm. The FCB is a drift chamber located at $|z|$=275 cm consisting of 3 groups of 4 wire planes with a track element resolution of $\sigma(x)$=$\sigma(y)$=150 μm.

Performance of the tracking system

When reconstructing the event the *track-elements* provided by each sub-detector are combined in a global track-fit, performed by the DELANA[50] programme. The combined tracking performance should not be shown *only* in the form of a simple list of the single hit precisions in each of the DELPHI tracking chambers. In this section the quality of the tracking system is demonstrated by two examples: first by the momentum resolution for isolated high momentum tracks and secondly by the achieved precision on the estimation of the primary vertex. The resolution on the distance between tracks and the primary vertex (impact parameter resolution) will be given in section 4.4 where it is discussed in relation to a quantity (impact parameter) used in the analysis of multi-jet final states.

1) single track momentum resolution:

The precision of the estimation of the momentum of the particle can be illustrated by the resolution on the momentum of muons produced by Z→ $\mu^+\mu^-$. For tracks containing information from all central tracking detectors (VD, ID, TPC and OD) the centre of the momentum resolution distribution can be described by a Gaussian of width:

$$\sigma(1/p) \ = \ 0.57 \ \times \ 10^{-3} \ (\mathrm{GeV/c})^{-1}$$

The tails of the distribution require a second Gaussian with a width of 1.04 $\times$ 10^{-3} (GeV/c)$^{-1}$. As can be seen in table 3.3 the resolution deteriorates when going down to smaller polar angles as could be expected by the increased influence of multiple scattering, the decreasing number of hits/detectors that provide information on the tracks and the fact that the momentum (curvature) determination is extracted via the component transverse to the magnetic field.

angular region	Detectors	$\sigma(1/p)$ (GeV/c)$^{-1}$
42°$< \theta <$ 90°	VD+ID+TPC+OD	0.6×10^{-3}
	ID+TPC+OD	1.1×10^{-3}
	VD+ID+TPC	1.7×10^{-3}
$\theta \leq$ 36°	VD+FCB included	1.3×10^{-3}
$\theta <$ 25°	FCB included	2.7×10^{-3}

Table 3.3: *Momentum resolution for 45.6 GeV/c muons in different detector configurations.*

2) Primary vertex reconstruction:

In each event the point where the positron and the electron collide is called *the primary vertex*. In DELPHI the primary vertex is reconstructed in every hadronic event using tracks from the event itself combined with the constraint that the vertex they form should be compatible with *the beam spot*, defined as the interaction region of the electron and positron beams. The beam spot itself is determined every 200 events using tracks from all events, it is small (typical size is 150 μm in x and 10 μm in y) and known to very good precision [51]. Using the beam spot as constraint the precision on the event-by-event primary vertex is:

$$\sigma(x, y, z) \sim (22{,}8{,}22)\ \mu\mathrm{m}$$

In both the reconstruction of the beam spot and primary vertex determination the influence of tracks from interactions with the detector material and of secondary vertices is minimised. Determining if a group of particles is compatible with the primary vertex is important when studying 4-quark final states. In section 4.4 the resolution on the (closest) distance between a track and the primary vertex will be investigated.

3.2.2 Calorimeter detectors

To measure the energy of a particle, calorimeters perform a destructive measurement by stopping particles in their volume. In a cascade of interactions (shower) the particle deposits its energy in the calorimeter, where the shower characteristics can be used to determine the energy of the incident particle. To stop particles calorimeters must consist of high density materials. Using high granularity detectors allows in addition to reconstruct the position and the direction of the particle. Different characteristics in showering of electrons and photons with respect to hadrons have resulted in a separate electromagnetic and hadronic calorimeter to optimally reconstruct their energy.

Electromagnetic Calorimeters

Electromagnetic calorimeters measure the energies of incident electrons and photons. For high energy (above 1 GeV) electrons and photons the dominant processes are bremsstrahlung and $\mathrm{e}^+\mathrm{e}^-$ pair production respectively. The secondary particles produced in these interactions are again mainly e^+, e^- and γs. This cascade develops through repeated similar interactions and can be characterised by the (material dependent) radiation length (X_0). The number of particles in the shower can be estimated and is a measure for the energy of the incident particle. In DELPHI a combination of three detectors ensures a good coverage over almost the full solid angle. After a short description of each, their main characteristics and resolutions are summarised in table 3.4. Hadrons traversing the electromagnetic calorimeters will deposit only a fraction of their energy.

- High density Projection Chamber (HPC)
  The HPC is a cylinder located in the barrel at radii between 208 and 260 cm (just before the DELPHI solenoid when seen from the interaction point), covering the angular regions[3] between 43° and 137°. It consists of 6 rings of 24 modules each containing layers of lead plates (18 X_0) inter-spaced with proportional gas chambers that detect the ionisation produced by electrons or positrons from the shower. The high granularity results in an angular precision for high energy photons of 1.0 mrad in θ and 1.7 mrad in ϕ.

- Forward ElectroMagnetic Calorimeter (FEMC)
  In each of the DELPHI end-caps the FEMC is the electromagnetic detector covering the angular region between 8° and 35° (172° and 155°). Each consists of 4532 lead glass blocks placed at $|z|$ = 284 cm with a depth of 40 cm (20 X_0). Charged particles inside the electromagnetic shower produce Cherenkov photons that are detected at the end of each block by a photo-multiplier (triode). For showers above 2 GeV the precision on the impact point in x and y is about 0.5 cm.

- Small angle TIle Calorimeter (STIC)
  At 2.2 m on either side of the DELPHI interaction region the STIC is a calorimeter covering the forward region between 29 mrad and 185 mrad. It consists of alternating layers of lead and scintillator. This detector is important in the determination of the integrated luminosity by identification of Bhabha events.

[3] Dead regions around θ=90° ($\Delta\theta$=2°) and along ϕ ($\Delta\phi$=0.2°) each 15°.

	Thickness	Energy resolution (E in GeV)
HPC	18 X_0	$\sigma(\mathrm{E})/\mathrm{E} = 0.043 \oplus 0.32/\sqrt{\mathrm{E}}$
FEMC	20 X_0	$\sigma(\mathrm{E})/\mathrm{E} = 0.03 \oplus 0.12/\sqrt{\mathrm{E}} \oplus 0.11/\mathrm{E}$
STIC	27 X_0	$\sigma(\mathrm{E})/\mathrm{E} = 0.0152 \oplus 0.135/\sqrt{\mathrm{E}}$

Table 3.4: *Characteristics (conversion depth and energy resolution) of the various sub-detectors forming the DELPHI electromagnetic calorimeter system.*

Hadronic Calorimeter

A hadronic shower process is more complex and is dominated by a succession of inelastic hadronic interactions, characterised by multi-particle production. The number of charged particles produced in the shower is a measure of the energy of the incident hadron. These charged particles are detected using for example scintillators inside the converting material which is required to be a very dense material like lead, iron or uranium.

- HAdron Calorimeter (HAC)
  The hadron calorimeter is installed in the return yoke of the DELPHI solenoid, covering a large part of the solid angle (11° to 169°). It consists of about 19,000 limited streamer tubes placed between the 5 cm thick iron plates. The tubes are connected to readout boards covering a fixed angular region of $\Delta\theta = 2.96^\circ$ and $\Delta\phi = 3.75^\circ$ consisting of up to 64 pads. In the barrel region the energy precision is found to be:

  $$\sigma(\mathrm{E})/\mathrm{E} \;=\; 0.21 \oplus 1.12/\sqrt{\mathrm{E}} \quad (\mathrm{E\ in\ GeV})$$

  The poor resolution is dominated by the large amount of material (magnet) the particles encounter before reaching the calorimeter.

Performance of the calorimeter system

The performance of the DELPHI calorimeter is investigated by studying the total reconstructed energy (and its resolution) in multi-jet events at LEP2. Figure 3.5 illustrates the fact that in a typical 4-quark event about 13% of the energy is not recovered. The energy resolution is also close to 12%. The tools that are used, on the level of reconstructed jets, to account for this bias and to improve the understanding of the energy-flow to beyond the detector resolution are discussed in section 4.3.

3.2.3 Remaining sub-detectors, trigger and event reconstruction

Remaining sub-detectors

The remaining sub-detectors that make up DELPHI are mostly used for particle identification. Although not used in this thesis they are described shortly:

- Muon chambers (MUB, MUS and MUF)
  The muon chambers (layers of wire chambers) are located outside the hadronic calorimeter. They are used for muon identification since the only charged particles that can reach them are muons (above 2 GeV).

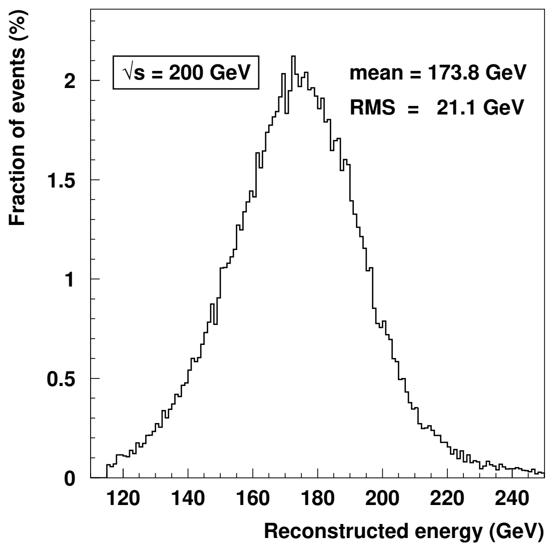


Figure 3.5: *Total reconstructed energy for simulated fully hadronic 4-fermion events at a centre-of-mass energy of 200 GeV.*

- Very small angle tagger (VSAT)
  The VSAT is a detector located far away from the interaction point very close to the beam-pipe to detect electrons and positrons coming from Bhabha scattering between 5 and 7 mrad.

- Ring Imaging Cherenkov detectors (RIB and RIF)
  DELPHI is the only LEP experiment that has large Cherenkov detectors (both in barrel and the forward region). They are used for particle identification.

- Hermeticity detectors and Taggers
  To achieve complete hermeticity for high energy photons scintillators are installed in order to cover 'blind' regions between the barrel and end-cap and between HPC modules.

- Time Of Flight detector (TOF) and Forward Hodoscope (HOF)
  Scintillator counters used to reject background events from cosmic radiation. They also provide a fast trigger for events from genuine $\mathrm{e^+e^-}$ interactions.

Trigger

In only a fraction of the beam crossings (beams cross every 22 μs) an $\mathrm{e^+e^-}$ interaction occurs. To store the information from all sub-detectors at each bunch crossing is impossible since reading out the information from all sub-detectors takes more time than the bunch crossing time. The goal of the DELPHI trigger system (described in detail in [52]) therefore is to select all events from $\mathrm{e^+e^-}$ interactions, while keeping the backgrounds from other processes at a low level. To cope with the high luminosity and large backgrounds the DELPHI trigger system is composed of four consecutive levels of increasing complexity.

The first two levels (T1 and T2) are **hardware** triggers synchronous with the beam cross overs (BCO) and are required to take decisions after 3.5 and 39 μs respectively. T3 and T4 are **software** filters where the full event information is used. Events that pass all triggers are written to storage tapes. At LEP2 the typical rate for this is a few Hz. For hadronic events the trigger efficiency is very close to 100%.

Event reconstruction

Using the raw information from each sub-detector the event is reconstructed using the DELANA [50] programme. DELANA performs pattern recognition and does a full track reconstruction linking information from different sub-detectors: track elements, pedestal and distortion corrections, calibrations etc. The reduced information on the reconstructed event is written in a Data Summary Tape (DST) format. The information on the DSTs is accessible to the users, for example within the SKELANA [53] framework.

Detector simulation

In order to compare the data with expectations the response of the DELPHI detector on simulated events needs to be modelled. The response is simulated using the DELSIM [54] (DELphi SIMulation) programme. It transforms the simulated particles into an event as it would be seen inside the DELPHI detector. The raw (reconstructed) information on the simulated event is then passed through the same reconstruction chain as the real data events and finally written to tape.

Chapter 4

Standard tools

In this section a few of the tools that are used when studying multi-jet final states are introduced and discussed shortly. The section ends with a description of the selection criteria used to select a pure sample of 4 quark final states from all e^+e^- interactions. This forms the starting point of both the fully hadronic ZZ cross section measurement and the search for the SM Higgs boson.

4.1 Production of multi-jet final states

The quarks that are produced in the e^+e^- interaction can not be observed freely in nature. Due to the colour confinement in QCD they will fragment and form jets of colour neutral particles, allowing quarks and gluons only to be observed through their remnants in the form of (collimated) jets of particles. The translation from jet-characteristics back to partons-characteristics is described in section 4.2. Inversely, the fragmentation from partons to jets, as shown schematically in figure 4.1, is shortly discussed below.

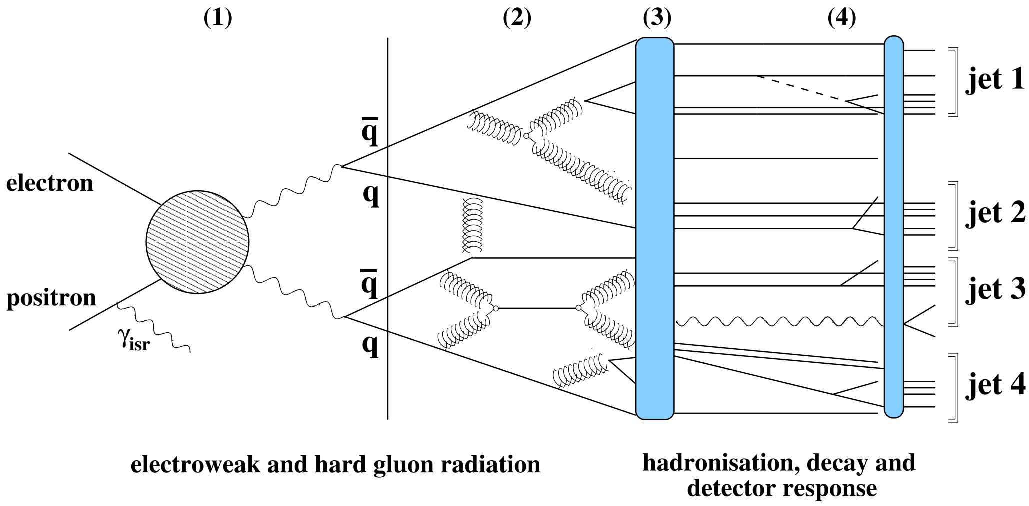


Figure 4.1: *Schematic representation of an* e^+e^- *interaction producing a multi-jet final state.*

In the process leading from the initial e^+e^- system to a set of (collimated) jets of particles, various physics processes and regimes are encountered. To obtain an optimal description of the full process, it is divided in parts that can either be computed accurately or have to be modelled. These different regimes are then matched together in a way similar to that shown in figure 4.1. Starting with the incident leptons (on the left) we can identify:

phase 1 & 2) producing a 4 quark final state and hard gluon radiation.

The first two steps in the process are well understood. The electroweak process describing the e^+e^- interaction producing a boson (pair) in the final state (either W^+W^-, ZZ or ZH) and their subsequent decay is well known and the matrix elements associated to specific final states can be computed to high precision. For boson pair production (4-fermion events) the `EXCALIBUR` [34] programme is used to generate events whereas for the annihilation processes ($e^+e^- \rightarrow Z \rightarrow q\overline{q}$), producing 2-fermion events, the `PYTHIA` [41] Monte Carlo programme is used.

The second step of the process describes subsequent hard gluon radiation of the initial partons. This region of large momentum transfers is the regime of perturbative QCD. The probabilistic evolution calculations, describing splitting of quarks and gluons, are performed inside the JETSET programme [41]. Hard gluon radiation is the reason that also 2-fermion events produce apparent 4-jet final states. The specific characteristics of the gluon radiation, infra-red and collinear divergences, allow to separate between 4-fermion (spherical) and 2-fermion (more cigar-like) events on a statistical basis. Such a separating variable is constructed in section 4.5.2 and is used both in the pre-selection of multi-jet final states and in the full analysis. This parton shower evolution is cut off when the virtuality of the partons is of the order of about 1 GeV. At that stage there are still only a limited number of partons in the final state.

phase 3 & 4) hadronisation, decay and detector response.

At scales of around 1 GeV QCD perturbation theory breaks down and the picture of QCD in terms of 'free' quarks and gluons no longer holds. This means that we have to use non-perturbative (phenomenological) models to make the transition from a few coloured partons to a jet consisting of many colourless hadrons. This is the step in the (multi-)jet production that can not be computed and there are different models that have been developed to describe the transition. Inside the JETSET programme for example the (Lund) string fragmentation model is used [41]. The hadronisation parameters in this model are tuned on the millions of Z events collected at LEP1 and the combination of parton showers and string fragmentation has been very successful in describing hadronic Z events.

Unstable hadrons formed during the hadronisation phase can decay before reaching the detector. Their decay properties are known from measurements and these decays are handled by the JETSET programme using tables of decays modes and branching fractions. Finally, the dynamics and behaviour of the stable particles inside the different detector components is simulated by the DELSIM [54] package which then provides the final topology as reconstructed in the DELPHI detector.

Experimentalists and theorists start from two different sides of figure 4.1 and, despite a not so well understood region called 'non-perturbative' QCD, the kinematics of the reconstructed jet is known to represent closely that of the underlying quarks or gluons. Only when testing very precise predictions on for example the exact particle content inside the jet or jet (sub-)structure, will the difficulties in the theoretical prediction manifest themselves. In this analysis only a 4-momentum estimate of the underlying fragmenting parton is required and since the overall energy flow of a high energy event is mainly determined by the perturbative process with only a minor additional smearing caused by the hadronisation step, this is expected to be well described by the Monte Carlo.

4.2 Jet clustering

Primary quarks and gluons produced in the interaction are identified as groups of particles in the detector as described in section 4.1. For an experimentalist the task is to traverse figure 4.1 in reverse order: reconstruct the dynamics of the underlying partonic structure given the detected event. Simplifying the event to a few partons allows to focus on a region where the theoretical predictions are well behaved (QED and perturbative QCD). To obtain the best estimate of the energy and momentum of each parton the goal therefore is to cluster all particles originating from the same parton fragmentation. At the end of the clustering the remaining groups of particles are called jets and their 4-momentum (the direct sum of all particles assigned to the jet) is used as an estimation of the parameters of the primary parton. In this section a general introduction on clustering algorithms is given followed by a short discussion on the DURHAM algorithm which is the algorithm used throughout the rest of this thesis.

Clustering algorithms in general

Most clustering procedures can be characterised by three definitions:

- **A distance (y_{ij}):** A measure defining 'how far' two particles i and j are apart. This measure is used to decide which two particles should be considered to be combined first.
- **A maximum distance (y_{cut}):** A resolution parameter defining the maximum distance for which two particles can be combined (and therefore the moment the clustering is ended).
- **A recombination scheme:** A prescription of how to construct the 4-momentum of the object that is formed when two particles are combined.

In order to cluster particles into jets most algorithms follow a similar procedure: iterative procedures starting for each event with all (n) measured energy flow objects (particles). For each pair of these final-state particles (i, j) their separation is calculated using the algorithm specific distance definition. In the next step of the procedure the two particles (i,j) with the smallest value of y_{ij} are combined and replaced by a 'pseudo-particle' with 4-momentum $\mathrm{p}^{\mu}_{(ij)}$ using a pre-defined recombination scheme. This clustering is performed only if y_{ij} is smaller than the maximum distance (or resolution) parameter y_{cut}. After clustering particles i and j the event now consists of n-1 clusters and again all distances between particles are calculated. This procedure is repeated until all pairs of particles have $y_{ij} > y_{\mathrm{cut}}$. Clusters of particles that remain at this stage are called jets. For a given value of y_{cut} the event is characterised by the number of reconstructed jets.

The original Jade clustering algorithm

One of the first clustering algorithms is the JADE [55] algorithm, where the distance between two particles is defined as:

$$y_{ij} = \frac{2\, E_i E_j (1 - \cos\theta_{ij})}{E_{\mathrm{vis}}^2} \tag{4.1}$$

Here θ_{ij} is the opening angle between the two particles, making the numerator identical to the invariant mass squared of the two particles in case both particles are massless. The distance y_{ij} is transformed into a dimensionless quantity by normalising to the total visible energy squared in the event. Although the recombination scheme originally suggested was adding of the 4-momenta a variety of schemes has been used in combination with the distance defined in equation 4.1:

E-scheme:	$E_{ij} = E_i + E_j$	and	$\vec{p}_{ij} = \vec{p}_i + \vec{p}_j$
E0-scheme:	$E_{ij} = E_i + E_j$	and	$\vec{p}_{ij} = (\vec{p}_i + \vec{p}_j)\, E_{ij}/\vert\vec{p}_i + \vec{p}_j\vert$
P-scheme:	$E_{ij} = \vert\vec{p}_{ij}\vert$	and	$\vec{p}_{ij} = \vec{p}_i + \vec{p}_j$

The reason why different recombination schemes were introduced were mostly given by theoretical considerations. The P- and E0-schemes for example violate energy and momentum conservation respectively, but have the advantage that the reconstructed jets are massless, which is similar to the approach in many theoretical calculations. An overview of the different considerations to decide on a different scheme, their performances and characteristics can be found in various review articles (like [56]).

The JADE clustering algorithm has some characteristics that cause problems in certain classes of events since two low-energetic clusters will be combined even when they have a large opening angle. In events containing many of these low-energetic clusters these can form a 'phantom' jet leading to an unnatural assignment of particles to jets. To overcome these problems other algorithms were developed that were less sensitive to soft gluon emission like the DURHAM clustering algorithm.

The DURHAM clustering algorithm

The DURHAM or $k_\perp$-algorithm [57] is the most widely used algorithm in multi-jet final states at LEP2. It is closely related to the JADE algorithm, but does not suffer from the problems discussed above by adopting a slightly different distance definition and the Lorentz-invariant E-scheme as recombination procedure:

$$y_{ij} = \frac{2\,\min(E_i^2, E_j^2)(1 - \cos\theta_{ij})}{E_{\mathrm{vis}}^2} \tag{4.2}$$

$$p_{ij}^{\mu} = p_i^{\mu} + p_j^{\mu} \qquad \text{(E-scheme)} \tag{4.3}$$

With θ_{ij} representing the opening angle between the two particles, the numerator represents the transverse momentum squared of the lowest energetic particle with respect to the most energetic particle. The distance is again normalised to the total visible energy in the event and has the advantage compared to the original JADE distance that it assigns low-energy particles to their closest high energy jet first and is therefore both experimentally and theoretically better behaved.

Another advantage of this clustering algorithm is the fact that the hadronisation of the parton final states can be shown to have, on average, little influence on the jet-rates. This can be evaluated on generator level by studying the differences between parton and hadron level. Differences vary significantly between clustering procedures and are shown to be smallest for the DURHAM clustering algorithm ([58, 59]). Although there is no single *best* clustering algorithm (every algorithm has specific qualities and problems), in multi-jet events at LEP2 the DURHAM clustering algorithm is generally accepted as one of the best in reconstructing the energy flow of the original partons and is therefore used throughout this thesis.

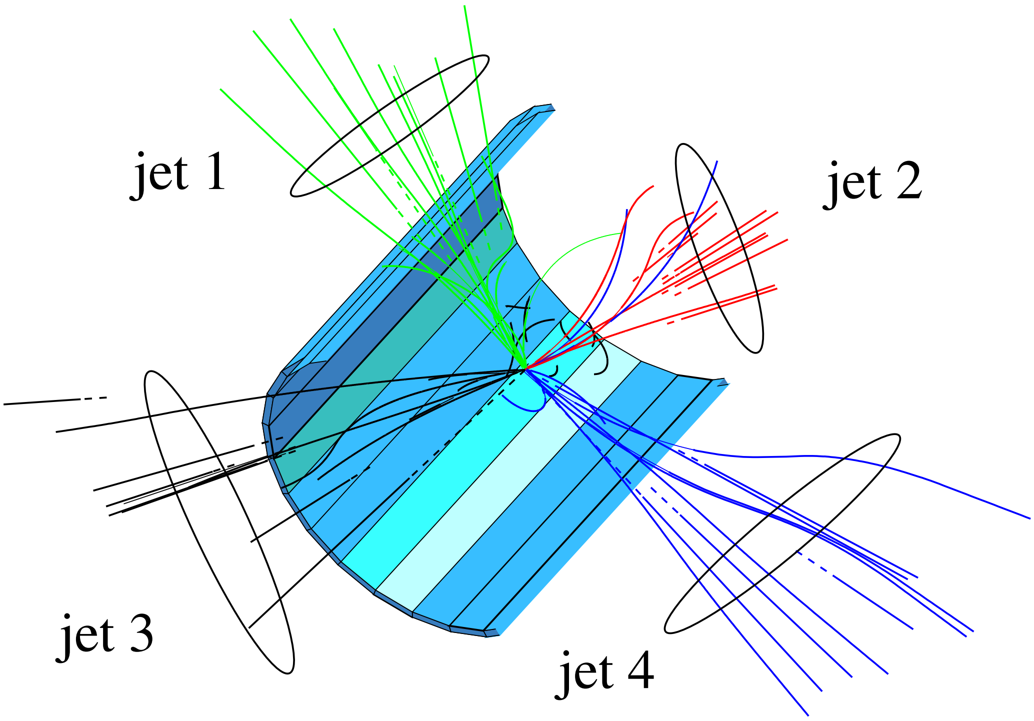


Figure 4.2: *Event display showing reconstructed charged tracks in the DELPHI detector. The DURHAM clustering algorithm has clustered the event into 4 jets.*

The DURHAM algorithm defines the procedure used in this thesis to cluster groups of detected particles into jets whose 4-momentum can be used as an estimate of the original parton 4-momentum. An event in which 4 jets have been reconstructed can be seen in figure 4.2.

4.3 Constrained fit

Estimating the 4-momenta of the partons using the reconstructed jets would be quite accurate (neglecting assignment of particles to the wrong jet) if all particles would be detected and their energy accurately measured. However, since neutrinos can be present in the final state and particles travelling through inefficient regions of the detector escape detection this is not always the case.

Comparing the reconstructed jet energy to that of the generated parton on Monte Carlo events it is known that, on average, about 15% of the jet energy is lost. The resolution on the jet energies is also about 15% due to both the intrinsic resolution of the calorimetry and the fact that some particles experience significant energy loss through interaction with the detector material. Both effects make accurate jet-energy measurements difficult. The estimation of the parton direction on the other hand is quite precise ($|\theta_{\rm jet} - \theta_{\rm parton}| \approx 3^\circ$) since this is a momentum weighted sum of all constituent particles and the resolution on angles and momentum of charged tracks is excellent.

The estimation of the dynamics of the partonic system can be improved by taking into account all known effects and biases. Combining this with the requirement that the reconstructed event should have specific properties (identical to the well-defined characteristics of the initial e^+e^- system) the jet properties can be improved beyond the detector resolution. This will significantly improve the resolution on di-jet invariant masses which is important in studying multi-jet final states. In section 4.3.1 the method (constrained fit) to implement these effects and simultaneously satisfy the constraints will be explained by describing the most general (4C) fit where energy and momentum conservation is demanded. In section 4.3.2 a different (3C) fit will be introduced that is used to provide information for the event selection.

4.3.1 Energy and momentum conservation (4C fit)

Unlike the final state, the initial state of the e^+e^- interaction is known to high precision: perfectly balanced in momentum and energy known to about 1‰ (assuming no initial state radiation (ISR)). Since these characteristics are preserved the reconstructed event has to satisfy 4 constraints (energy and momentum conservation):

$$\sum_{i=1}^{n_{\rm jets}} p_i^{({\rm fitted})\ \mu} = (0, 0, 0, \sqrt{s}) \tag{4.4}$$

In this fit there are 4 constraints to be satisfied and is therefore known as a 4C-fit. Several methods can be used to take into account the expected energy loss while simultaneously satisfying the different constraints. Here both a rescaling and a more complex constrained fit method are described.

Rescaling: One of the less frequent methods used is the rescaling method characterised by the fact that in the fit the jet directions are fixed and only their energies are allowed to vary. This uses the knowledge that the resolution on the jet-directions is much better than the resolution on the jet energies. This translates into one free parameter per jet and solving for the α_i's in $\sum_{i=1}^{n_{\rm jets}} \alpha_i\ p_i^{({\rm meas})\ \mu} = (0, 0, 0, \sqrt{s})$. The solutions of this procedure, found by simple matrix inversion, are however not very stable. Especially in planar events (all jets in the same plane) the problems are obvious: in such an event-topology one of the momentum constraints is automatically fulfilled and the solutions become degenerate. Using rescaling to estimate jet energies in events with high invariant di-jet masses, like ZZ or ZH events, introduces large biases making rescaling not the ideal tool to be used in multi-jet events at LEP2.

Constrained fit: To have a stable fit and to be able to take into account the errors on the jet directions the transverse components of the jet momentum must also be allowed to vary. Instead of one parameter this means three independent parameters are needed to parametrise the jet (the invariant mass of the jet is assumed to scale with the energy) in the following way:

$$\vec{\mathbf{p}}_i^{(\mathrm{fitted})} = e^{a_i} \vec{\mathbf{p}}_i^{(\mathrm{meas})} + b_i\ \hat{\mathbf{n}}_\mathbf{1} + c_i\ \hat{\mathbf{n}}_\mathbf{2} \tag{4.5}$$

In this expression $\vec{\mathbf{p}}^{(\mathrm{meas})}$ is the measured momentum and $\hat{\mathbf{n}}_\mathbf{1}$ and $\hat{\mathbf{n}}_\mathbf{2}$ are 2 orthogonal unit vectors in the plane perpendicular to the measured momentum direction $\vec{\mathbf{p}}^{(\mathrm{meas})}$. A schematic view of the different degrees of freedom can be seen in figure 4.3.

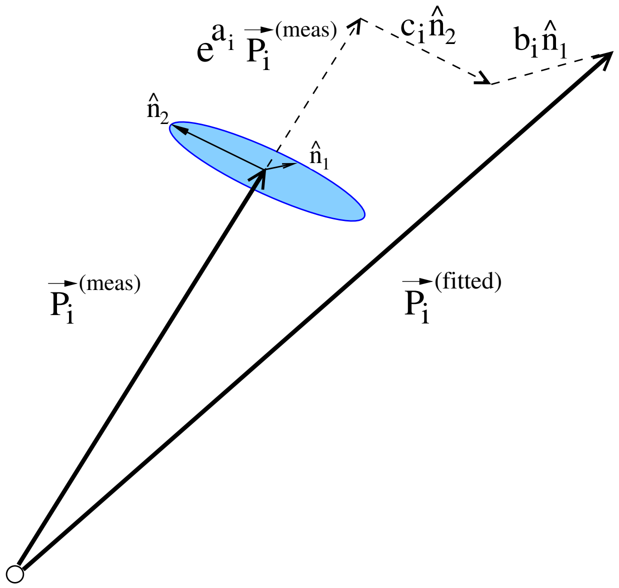


Figure 4.3: *Schematic view of the different components of the jet momentum (a_i, b_i and c_i) that are free to vary in the constrained fit. Both measured and fitted jets are shown.*

This specific parameterisation has been chosen instead of the maybe more obvious set (E,θ,ϕ) because the method used in the constrained fit requires a set of parameters whose errors are described by a Gaussian function. In the fit [60] a Lagrange multiplier technique is used to satisfy the constraints while minimising the following χ^2:

$$\chi^2_{4\mathrm{C}} = \sum_{i=1}^{n_{\mathrm{jets}}} \frac{(a_i - a_0)^2}{\sigma^2_{a_i}} + \frac{b_i^2}{\sigma^2_{b_i}} + \frac{c_i^2}{\sigma^2_{c_i}} \tag{4.6}$$

The parameters to be fitted are a_i, b_i and c_i. The energy loss parameter a_0 ($\sim 15\%$) and the errors on the other parameters (σ_{a_i}, σ_{b_i}, σ_{c_i}) are determined using Monte Carlo events where both the measured and true (parton) momenta are known.

When going to smaller polar angles the resolution on single tracks deteriorates and, due to the geometrical acceptance of the DELPHI detector, particles escape detection. For these reasons the resolutions on the jet-energy and the energy loss of the jet are also polar angle dependent. In addition, due to the undetected particles at small polar angles, there is a general bias to shift the measured jet directions away from the beam axis. The angular parametrisation of the errors has been tuned using Monte Carlo (2 jets at the Z-pole) events and has been found to be described by the following function and parameters [61]:

$$\begin{aligned} a_0 &= 0.15 + 0.40\cos^4(\theta_i) \\ \sigma_{a_i} &= 0.15 + 0.40\cos^4(\theta_i) \\ \sigma_{b_i} = \sigma_{c_i} &= 1.6 \;\; + 1.0 \;\; \cos^4(\theta_i) \end{aligned} \tag{4.7}$$

Using a constrained fit not only results in the best estimate of the parton 4-vectors, but also allows event-by-event errors on these parameters to be used in the analysis. The errors on the jet 4-momenta for example will be used in section 5.2.2 to test the compatibility of a di-jet invariant mass with a specific hypothesis.

4.3.2 Other constrained fits

In addition to the 4C-fit as described by equation (4.4) there are also 3C, 5C and 6C fits (corresponding to the number of constraints they put on the event) that are widely used. In this section the 3C-fit is described since it provides information used in the selection of high-energy multi-jet hadronic events (to reject a large fraction of the $\mathrm{q\overline{q}}(\gamma)$ events). The 5C and 6C-fits deal with the extraction of the di-jet mass information contained in the event and can only be discussed when related to a specific physics analysis or jet-pairing. The implementations and characteristics of these fits will be described and discussed in the relevant analysis sections.

3C-fit: Like a 4C-fit, but p_z is left free.
A fit assuming a single photon (γ_{isr}) was emitted along the z-axis. Its momentum is left free. The constraints used in the fit then take the following form:

$$\sum_{i=1}^{n_{\mathrm{jets}}} p_i^{(\mathrm{fitted})\ \mu} = (0, 0, p_z^{\gamma_{\mathrm{isr}}}, \sqrt{s} - |p_z^{\gamma_{\mathrm{isr}}}|) \tag{4.8}$$

Why a 3C fit: To estimate the effective-centre-of-mass energy ($\sqrt{s'}$).
The $\mathrm{e^+e^-}$ system can radiate a photon 'before' the interaction (initial state radiation or ISR) thereby reducing the effective centre-of-mass energy in the $\mathrm{e^+e^-}$ interaction. These photons are predominantly emitted at very small angles (well inside the beam pipe) and escape detection. Although the energy distribution of these photons is peaked at small energies the distribution has significant tails. In about 12% of fully hadronic 4-fermion events at a centre-of-mass energy of 206.7 GeV for example, the generated ISR energy is more than 10 GeV and for 5% of the events this is even more than 20 GeV.

Although the resolution on the reconstructed $\sqrt{s'}$ is not very good (a few GeV, see also [62]) the $\sqrt{s'}$-distribution can be used to reject events with significant photon radiation. This allows a large fraction of the $\mathrm{q\overline{q}}(\gamma)$ events at LEP2 (produced through $\mathrm{e^+e^-}$ annihilation) to be discarded already at pre-selection. Since the annihilation cross section is very large at the Z-peak (91.2 GeV/c^2), in about 50% of the interactions a large amount of energy is radiated such that the effective centre-of-mass energy of the $\mathrm{e^+e^-}$ system is brought down to the Z-peak. This photon energy emission pattern results in a characteristic double peak structure in the $\sqrt{s'}$ distribution. Since 4-fermion final states do not have this structure the evaluation of $\sqrt{s'}$ provides a powerful tool to eliminate a large fraction of the $\mathrm{q\overline{q}}(\gamma)$ background. The distributions of $\sqrt{s'}$ for different event types can be seen in the right plot of figure 4.10.

4.3.3 Performance of the 4C-fit

In this section the performance of the constrained fit is investigated by comparing, in Monte Carlo events (4-jet events at LEP2), the energy and momenta of the reconstructed jets to that of the generated initial partons. The resolution on the estimation of the angle of the partons is good as can be seen in the left plot of figure 4.4. For jets reconstructed in the central part of the DELPHI detector the performance on the parton energy estimation is shown in the right plot of the same figure. On average the jet-energy is overestimated. This can be explained by the fact that in the constrained fit the initial state radiation is neglected (although $<\mathrm{E_{isr}}>=3$ GeV for $\mathrm{W^+W^-}$ events at 200 GeV), hence overestimating the energy contained in the event (jets). This also explains the asymmetric distribution and enhanced (ISR) tail at higher (overestimated) jet energies.

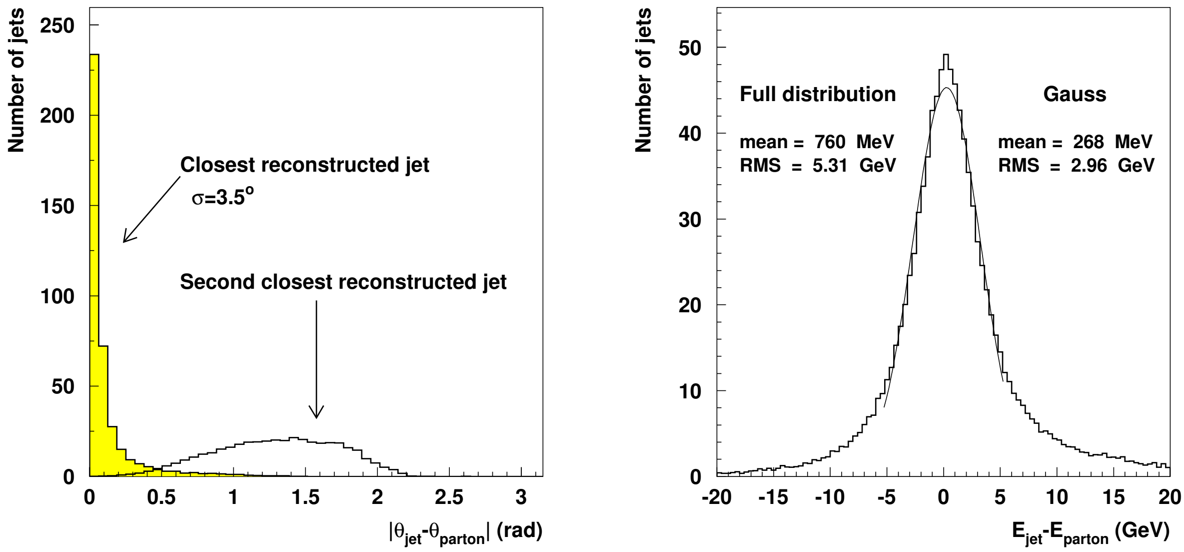


Figure 4.4: *The left plot shows the angle between the original parton and the closest and one but closest reconstructed jet. The right plot shows the energy difference between the reconstructed jet and the energy of the initial parton. These plots were constructed using 4-jet events.*

4.4 b-tagging

B-hadrons are formed in the hadronisation of b-quarks and have quite distinct properties allowing jets of particles to be identified ('tagged') as coming from the fragmentation of a b-quark. The b-tag configuration of different 4-quark final states at LEP2 can be quite specific for a given process. They can therefore be used to select or reject various physics processes (like $\mathrm{W^+W^-}$ events where there is hardly ever a b-quark present in the final state). Two complementary methods have been developed by DELPHI to identify b-quark jets. They have been combined to obtain optimal sensitivity. A detailed description of the method can be found in [63].

Since a b-quark produced in the interaction can not exist freely in nature it will fragment and in the hadronisation phase a B-hadron will be produced. These hadrons decay under the weak interaction via Cabibbo suppressed transitions (at tree level) of a b-quark to either a c- or u-quark under the emission of a W boson. The very small value of the CKM matrix element $|V_{cb}|$ ($\sim$0.04) and the even smaller value of $|V_{ub}|$ ($\sim$0.0035) lead to long life-times of B-hadrons of around 1.6 ps. In combination with the Lorentz boost they acquire at LEP the B-particles will travel several millimetres in the detector before they decay, giving them a characteristic displaced vertex topology as can be seen in figure 4.5. The most powerful technique to identify b-jets is to use the excellent resolution of the vertex detector to search for reconstructed tracks that do not originate from the primary interaction point. This is called the life-time or impact parameter method. However, since hadrons containing a c-quark can also have long life-times this method has limitations.

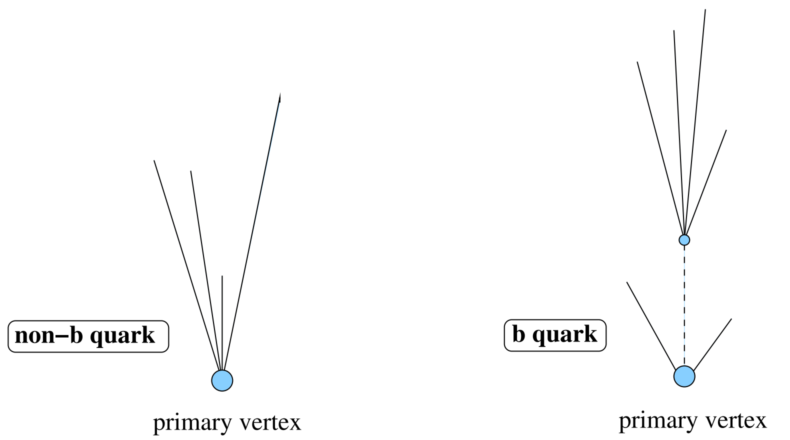


Figure 4.5: *Schematic view of quark fragmentation topologies for non-b (left) and b-quarks (right).*

To improve the b-identification performance beyond the life-time information alone, a second set of variables is introduced, using the specific characteristics of the B-hadron decay itself: the B-hadron is a heavy particle ($\approx$ 5 GeV/c^2), it has a large decay multiplicity and the fraction of the jet energy carried by the charged particles at the secondary vertex is larger for b-quarks than for light (u,d,s,c)-quarks.

Impact parameter or life-time method

Life-time based b-tagging requires testing if (a group of) particles originate from the primary vertex. In that procedure an accurate determination of the primary vertex is of course crucial (see section 3.2.1). Once the primary vertex is reconstructed an *impact parameter* can be defined for each charged particle as being the closest distance between the extrapolated track and the primary vertex itself. The sign of the impact parameter is defined with respect to the jet direction. It is positive if the extrapolated track crosses the axis of the jet to which it belongs before the primary vertex and negative otherwise.

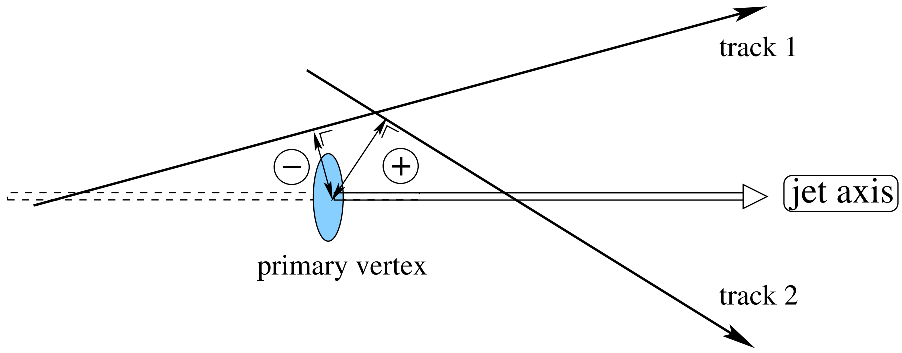


Figure 4.6: *The definition of the (sign of) impact parameter for tracks. Track 1(2) has a negative(positive) impact parameter.*

The resolution on the impact parameter is a measure of the quality of the tracking system. The impact parameter resolution involves a measurement error and a contribution from multiple-scattering. In DELPHI the impact parameter resolutions in the Rϕ-plane and Rz-plane are parametrised as:

$$\sigma^2_{IP\ R\phi} = \left(\frac{\alpha_{\rm ms}}{p\sin^{3/2}\theta}\right)^2 + (\sigma_{0,R\phi})^2 \ , \qquad \sigma^2_{IP\ Rz} = \left(\frac{\alpha'_{\rm ms}}{p\sin^{5/2}\theta}\right)^2 + (\sigma_{0,Rz})^2$$

In this expression p and θ are the momentum and polar angle of the track and the measurement error is given by: $\sigma_{0,R\phi}$ = 20 μm. The multiple scattering coefficients in the Rϕ-plane and Rz-plane are given by: $\alpha_{\rm ms}$ ($\alpha'_{\rm ms}$) = 65 (71) μm · GeV/c [48]. The signed distribution of the impact parameter for data and Monte Carlo on the Z-data can be seen in figure 4.7.

Ideally, for particles originating from the primary vertex, the impact parameter distribution is expected to be Gaussian with a width corresponding to the detector resolution. Unfortunately, due to mis-reconstruction and interaction of particles with the detector material, the distribution has large non-Gaussian tails. For particles originating from the primary vertex the distribution is expected to be symmetric. Particles originating from decay of long-lived particles will have (large) positive impact parameters of typically a (few) hundred μm. They will therefore contribute to the positive side of the impact parameter distribution. Before using the impact parameter distribution (as shown in figure 4.7) as a tool to identify particles from B-decay it must be verified that the full detector response is understood and correctly modelled. This is done by tuning the Monte Carlo to describe the negative side of the impact parameter distribution (containing only the combined detector and mis-reconstruction related effects). The tuning procedure [64] includes alignment of detector elements and removal of tracks in the Monte Carlo. It is performed each year using data taken during short dedicated Z runs where large statistics are collected to investigate detector effects and alignment problems that might differ from year to year.

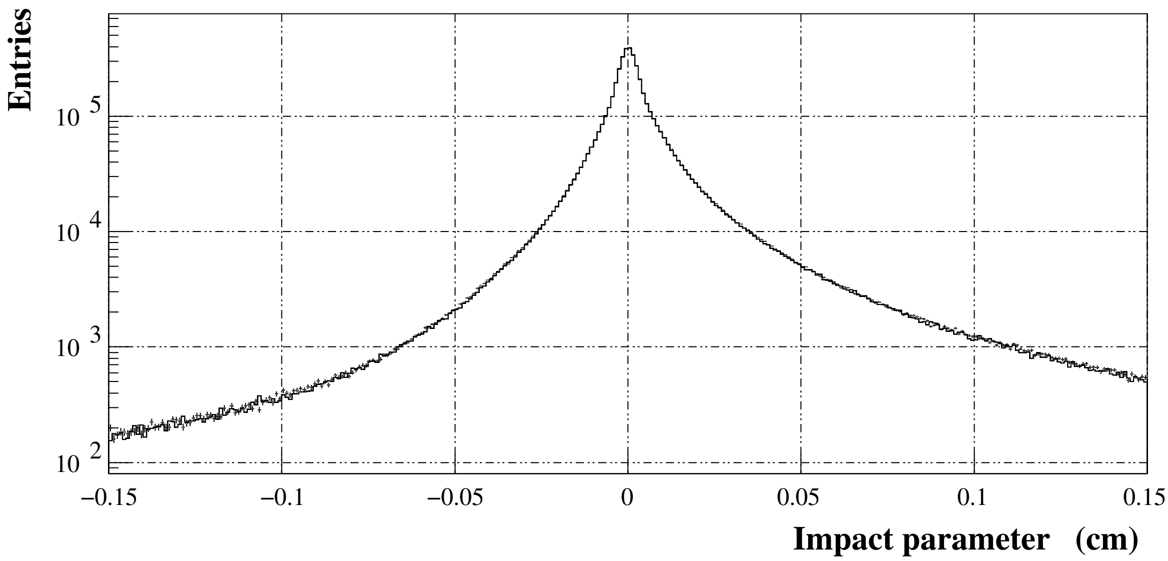


Figure 4.7: *The signed life-time impact parameter distribution after corrections. The points with error bars represent the data, the histogram is simulation.*

Since the precision on the track parameters can differ significantly from track to track it makes more sense to work with the *significance* of a track, defined as the ratio between the impact parameter and its error. Again, the (normalised) distribution of the significance from particles with negative impact parameters, $f(s)$, reflects all combined detector and mis-reconstruction effects and is used as a base to construct the *track probability function* $P(s_0)$, which is the probability for a track from the primary interaction to have a significance with absolute value larger or equal to s_0:

$$P(s_0) = \begin{cases} \int_{s\leq s_0} f(s)ds & \text{if } s_0 \leq 0. \\ P(-s_0) & \text{if } s_0 > 0. \end{cases} \tag{4.9}$$

For tracks originating from the primary vertex, the distribution of $P(s_0)$ is, per definition, a flat distribution between 0 and 1. Tracks from the decay of long-lived particles however have large positive values for s_0 and therefore a small value of $P(s_0)$. This reflects the small probability for tracks from the primary vertex to have such large positive impact parameters. Using the track probability function and the reconstructed significances for any group of N particles the probability can be calculated [65] that they are compatible with originating from the primary vertex:

$$P_N = \Pi \cdot \sum_{j=0}^{N-1} (-\ln \Pi)^j / j! \quad , \text{ where } \quad \Pi = \prod_{i=1}^{N} P(s_i) \tag{4.10}$$

For jets of particles coming from light quark fragmentation the distribution of P_N is expected to be flat, but for jets originating from B-hadron decay this N track probability calculated using formula (4.10) is usually very small due to the significant impact parameters of the decay products of the B-particle. Since a small value of the compatibility with the primary vertex indicates a long life-time this is often referred to as a life-time probability and is the most powerful tool to identify b-jets. The first global variable that gives information on the b-likeliness will therefore be a life-time one:

- **The jet life-time probability ($\mathbf{P_j^+}$)**
  For each jet in a hadronic event a probability is calculated (using formula (4.10)) that it is compatible with the primary vertex using only tracks in the jet that have positive impact parameters. The distribution of the life-time probability is shown in figure 4.8a for different quark flavours. Due to the non-zero life-time of D-mesons -$\log_{10} \mathrm{P_j^+}$ can also be large for c-jets limiting the performance of the life-time tag.

Secondary vertex or B-decay method

To improve the performance beyond the life-time information the characteristics of the B-hadron decay itself can be used. This requires removing the tracks in the jet that do not belong to the B-hadron (see figure 4.5), but are fragments of the b-quark fragmentation in the primary vertex. This is realised by performing a search for a secondary vertex [66] within each jet (the vertex is required to satisfy some quality cuts like the number (or type) of tracks that make up this vertex). In events where a secondary vertex is found additional variables can be defined using exclusively the tracks that make up the secondary vertex.

- **The invariant mass of the secondary vertex particles ($\mathrm{M_s}$)**
  The invariant mass of the particles from the secondary vertex for c-jets is limited by the mass of D-mesons. The probability for a secondary vertex to come from a c-jet decreases sharply above $\mathrm{M_s}$ = 1.8 GeV/c^2, while for b-jets the distribution extends up to 5 GeV/c^2 as can be seen in figure 4.8b.

- **The rapidity with respect to the jet direction of tracks from the secondary vertex ($\mathbf{R_s^{tr}}$)**
  Although a B-hadron in a b-jet has on average a higher energy than a D-meson from a c-jet the rapidities (defined as $1/2\ln[(E{+}p_z)/(E{-}p_z)]$) of particles from B-decay are on average less than those from D-decay. This can be explained by the larger mass of the B-hadron and the larger multiplicity of its decay [67]. The 'fake' secondary vertices in light quark jets are due to wrongly measured tracks caused by multiple scattering in the detector and interactions with detector material causing the tracks to be shifted to even smaller values. The distribution of this variable can be seen for the three quark classes in figure 4.8c.

- **The fraction of charged energy in the secondary vertex compared to the total jet energy ($\mathbf{X_s^{ch}}$)**
  The total energy of charged particles from the secondary vertex in light quark jets is much smaller than in b-quark jets. In case of b-jets the distribution of $\mathrm{X_s^{ch}}$ is determined by the fragmentation function $f(b \to B)$ whereas for a c-quark the distribution is determined by $f(c \to D)$ which is softer. The distribution of $\mathrm{X_s^{ch}}$ can be seen for the three quark types in figure 4.8d.

- **Transverse momentum with respect to the *b*-jet for identified leptons**
  In case the B decays semi-leptonically (Br(B$\to Xl^-\bar{\nu}_l) \approx 11\%$) the transverse momentum of the lepton with respect to the B hadron direction [68] can be exploited. Due to the large

mass of the b-quark this transverse momentum is expected to be larger for B-hadrons than for hadrons containing only light quarks. In absence of life-time information this is an excellent separating variable.

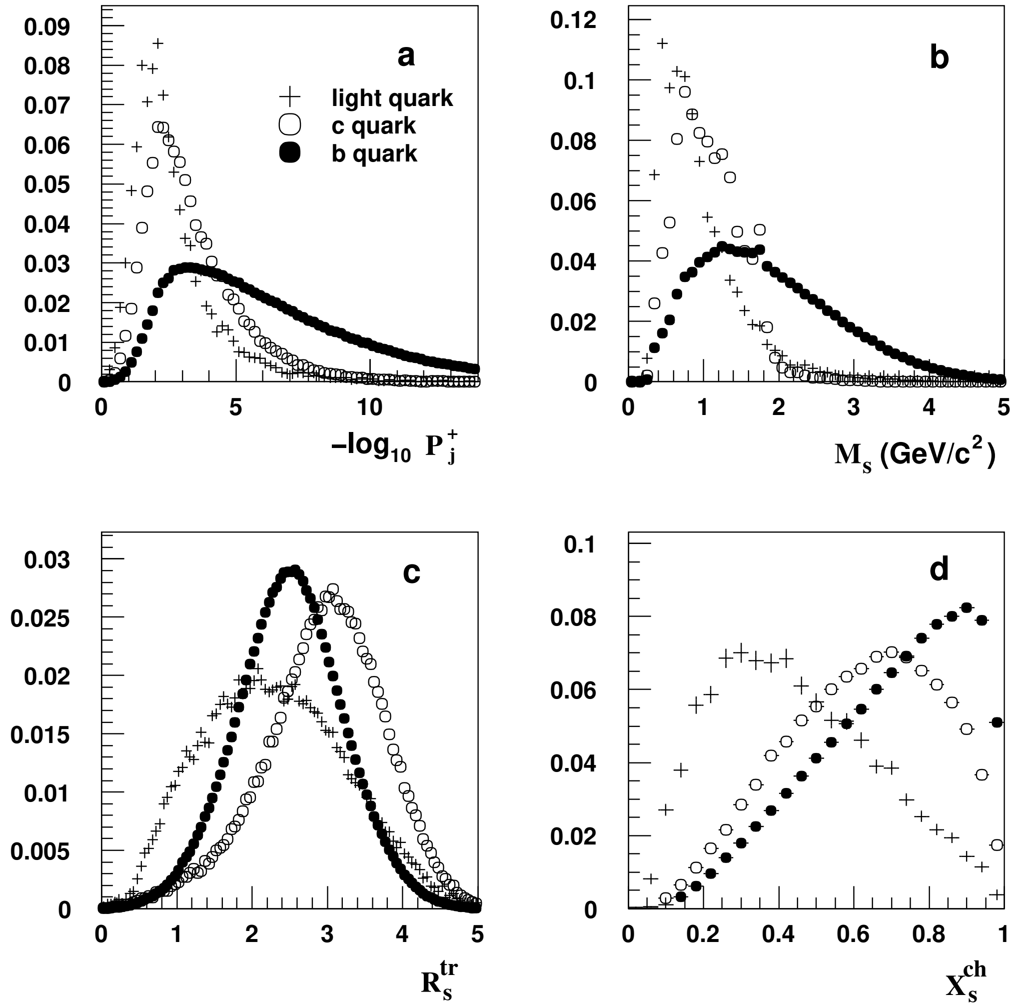


Figure 4.8: *Distribution of the discriminating variables used to construct the combined b-tagging variable. The figures show for different quark types the distributions of: a) the jet life-time probability; b) the invariant mass of particles included in the secondary vertex ; c) the rapidity of particles included in the secondary vertex and d) the fraction of the jet energy carried by the charged particles at the secondary vertex.*

Combining all information and performance

The b-tagging performance is optimised by combining all information on the jet. This is done by defining for each discriminating variable (i) a variable y_i as the ratio of the probability density functions for background and signal respectively: $y_i = f_i^B(x_i)/f_i^S(x_i)$. Since these tagging variables are only weakly correlated, a combined likelihood ratio is defined by simply taking their product: $y = \prod_{i=1}^{n} y_i$.

The backgrounds (corresponding to either c-quarks or u,d,s-quarks) have different behaviour in the variables entering the global b-tagging variable described above. The performance can be improved by taking this into account. This results in the following combined b-jet tagging variable:

$$\rho = n_c \prod \frac{f_i^c(x_i)}{f_i^b(x_i)} + n_q \prod \frac{f_i^q(x_i)}{f_i^b(x_i)} = n_c \prod_i y_i^c + n_q \prod y_i^q \tag{4.11}$$

In this expression n_c and n_q (q=u,d,s) are the normalised numbers of c- and u,d,s-jets with a reconstructed secondary vertex in $c\overline{c}$ and $q\overline{q}$ events respectively (n_c+n_q=1), and $f_i^q(x_i)$, $f_i^c(x_i)$ and $f_i^b(x_i)$ are the probability density functions of the variable x_i in u-,d-,s-,c- and b-quark jets. The jet is tagged as containing a b-quark if $\rho \geq \rho_0$, where the value ρ_0 can be varied to select the desired purity or efficiency of the tagging. The performance of the b-tagging in events with a secondary vertex can be seen in figure 4.9.

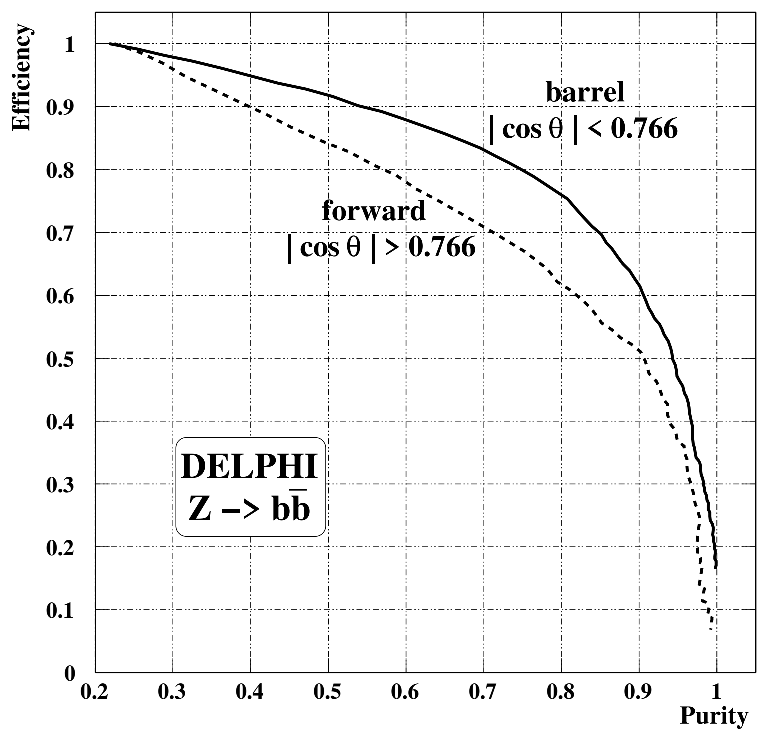


Figure 4.9: *B-tagging (hemisphere) efficiency versus purity.*

$b\overline{b}$ separation

Using more variables (like jet-charge, lepton sign etc.) it is possible to differentiate between jets coming from b- or $\overline{b}$-jets. This could be used to help resolve pairing ambiguities in case there are 2 or more reconstructed b-quarks in the event ($b\overline{b}$ allowed and bb forbidden). In the analysis presented here this possibility is not exploited, but, as will be shown in chapter 5, the structure of the analysis allows inclusion of any new or additional variables in a transparent way.

4.5 Hadronic event selection

In this thesis properties of 4-quark final states are studied. In order to remain as general and flexible as possible there is no optimisation for a specific physics hypothesis[1] during the event selection. Selecting a relatively pure sample of hadronic events requires only a loose pre-selection since its signature is so specific that with only a few cuts (retaining high efficiency for 4-quark final states) backgrounds from other processes are small.

After describing the track selection, the criteria used to select hadronic events are discussed by showing the distributions of variables used, the cuts applied and the data/Monte Carlo agreement before the cuts. In the last section the selection efficiencies for the different SM processes are presented.

4.5.1 Track selection

The observed event is reconstructed using the DELANA programme. A general track fit is performed using charged track segments supplied by all sub-detectors. In this analysis the tracks are required to satisfy the following conditions (default within the SKELANA analysis framework):

- Track momentum ($|p|$) $>$ 100 MeV/c
- Track momentum ($|p|$) $<$ 1.5 $\cdot \mathrm{E}_{\mathrm{beam}}$
- Relative momentum error (dp/p) $<$ 1.

The track should also be roughly compatible with the average primary vertex:

- Impact parameter in $R\phi$ $<$ 4 cm
- Impact parameter in $z \cdot \sin(\theta)$ $<$ 4 cm

There are various classes of tracks that are allowed or rejected. These definitions include the number of hits and the various sub-detectors supplying hits to the track.

A neutral track is defined as an energy deposition in the calorimeter above a specific energy threshold (typically around 350 MeV) that has no charged track pointing to it. To reject beam related background caused by off-momentum electrons there is an additional cut on the polar angle of neutral clusters at 3°.

4.5.2 Event selection

Selection criteria

As a first step the event is required to have a high charged track multiplicity, a large effective centre-of-mass energy ($\sqrt{s'}$: see section 4.3.2) and a large fraction of the available centre-of-mass energy observed in the detector:

[1] In ZH and ZZ analyses a large fraction of the $\mathrm{W}^+\mathrm{W}^-$ background is often rejected already at pre-selection level by applying an explicit cut on the b-tag variable.

- **cut 1** Number of charged tracks > 17
- **cut 2** $\sqrt{s'}$ $> 0.80\,\sqrt{s}$
- **cut 3** Total visible energy $> 0.69\,\sqrt{s}$

The distribution of the number of charged tracks and the effective centre-of-mass energy for various event types can be seen in figure 4.10. These cuts cost almost no efficiency (selection efficiency is 96.0% for fully hadronic 4-fermion events at 200 GeV), but reduce the background significantly (selection efficiency is 16.0 % for non-$\mathrm{q\bar{q}q\bar{q}}$ 4-fermion events and 20.2% for $\mathrm{q\bar{q}}(\gamma)$ (2-fermion) events at 200 GeV). After these cuts the remaining background is mostly due to $\mathrm{q\bar{q}}(\gamma)$ events with hard gluon radiation but also a few non-$\mathrm{q\bar{q}q\bar{q}}$ 4-fermion events remain.

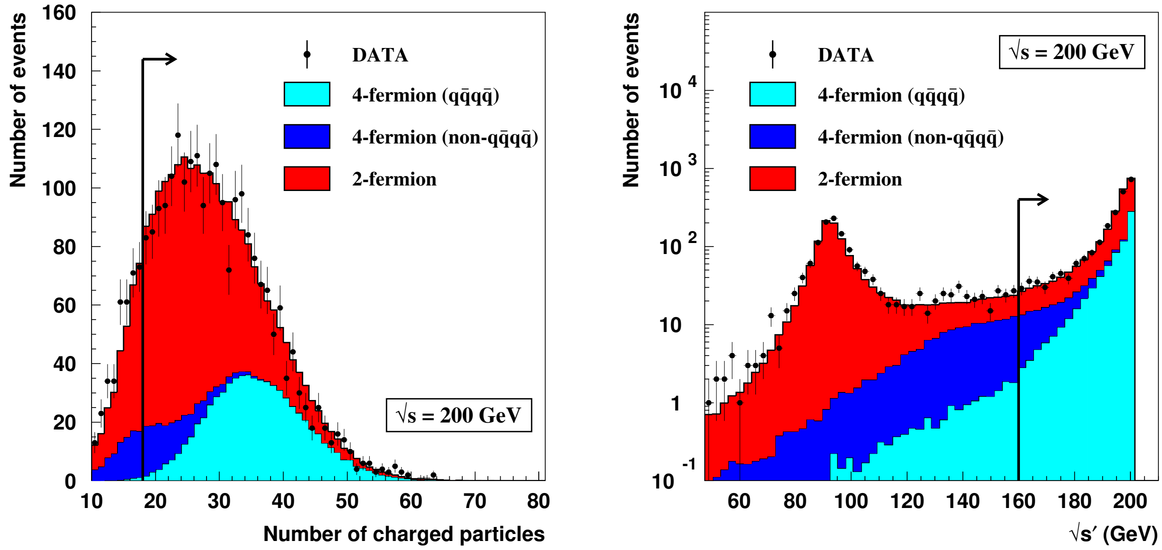


Figure 4.10: *Distributions of variables, at a centre-of-mass energy of 200 GeV, used in the preselection of fully-hadronic 4-fermion events. The plot shows the total charged multiplicity (left plot) and the effective centre-of-mass energy (right plot) when the other two criteria (of the first three) have been satisfied. The line and arrow indicate the region that is selected.*

The events are then clustered into jets using the DURHAM clustering algorithm (y_{cut}=0.001) and events are selected if there are 4 or more reconstructed jets[2]. Due to the kinematics of gluon radiation off quarks ($\sim 1/k_{\mathrm{T}}$) $\mathrm{q\bar{q}}(\gamma)$-events have only rarely 4 (or more) well separated reconstructed jets. In addition the knowledge that a quark manifests itself as a (large) number of particles in the detector can be used to reject jets composed of only one (or a few) particle(s) that are often isolated leptons or photons. To reduce the contamination from such events the following jet-quality cuts are introduced:

- **cut 4** Number of reconstructed jets ≥ 4
- **cut 5** Jet quality cuts:
  - Minimum jet multiplicity ≥ 4
  - Minimum jet mass > 1.0 GeV/c^2

[2] Jets with 6 or more reconstructed jets are forced to 5 jets in this analysis.

The distribution of the number of reconstructed jets and the minimum jet multiplicity for various event types can be seen in figure 4.11.

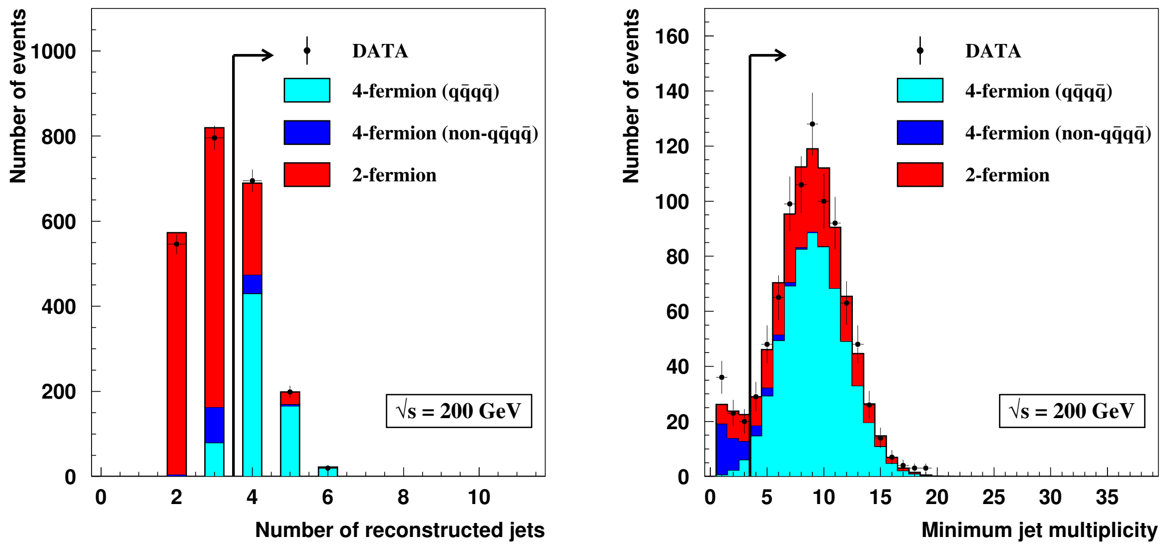


Figure 4.11: *Distributions of variables, at a centre-of-mass energy of 200 GeV, used in the pre-selection of fully hadronic events. The plot shows the number of reconstructed jets (left plot) and the minimum jet multiplicity (right plot) when all other criteria have been satisfied. The line and arrow indicate the region that is selected.*

The final pre-selection requirement uses the fact that the topological structure of the jets (separation between the jets in combination with their energies) in an event is different for 2-fermion ($q\overline{q}(\gamma)$) events with hard gluon radiation and genuine 4-fermion events as described already in section 4.1. The variable used to describe the topology of the event is defined as:

$$\mathrm{D_{pur}} = E_1\theta_1\sqrt{E_2\theta_2}/10 \tag{4.12}$$

In this expression E_1 (E_2) is the smallest (one but smallest) energy jet and θ_1 (θ_2) is the smallest (one but smallest) opening angle (in radians) between two jets. Events are required to satisfy the following condition:

- **cut 6** $\quad \mathrm{D_{pur}} \quad \geq \quad 7.5 \quad \mathrm{GeV}^{3/2}\ \mathrm{rad}^{3/2}$

In this section the topology of the event is used to reject some of the clear 2-fermion events, but in section 5.3 it is described in more detail how the full distribution of this variable is used inside the analysis to distinguish between the different event types. In that section also the advantages of using this specific topological variable will be discussed. The distribution of $\mathrm{D_{pur}}$ for the various event types at a centre-of-mass energy of 200 GeV can be seen in figure 4.12.

Pre-selection efficiencies

The selection efficiencies for 4-quark final states after all pre-selection cuts is 80% to 90% while most backgrounds are almost completely removed leaving only $q\overline{q}(\gamma)$ events ($\epsilon \approx 2.1\%$) with

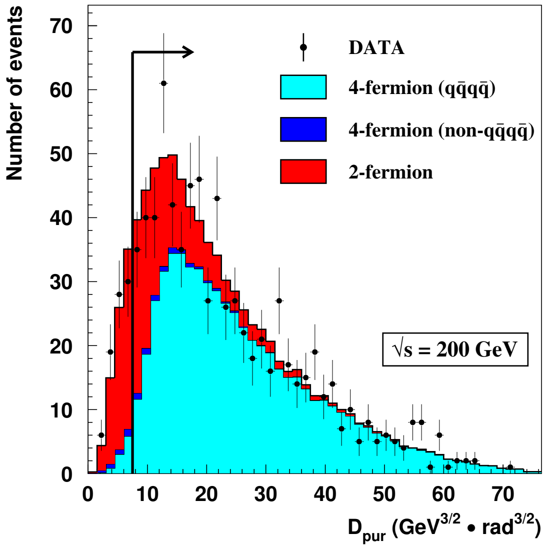


Figure 4.12: *Distributions of the topological (event) variable D_{pur} for the various event types at a centre-of-mass energy of 200 GeV. The line and arrow indicate the region that is selected.*

hard gluon radiation as significant remaining background. The pre-selection efficiencies for the various SM final states are shown in table 4.1.

$\sqrt{s}$	4-fermion $\mathrm{q\bar{q}q\bar{q}}$	4-fermion non-$\mathrm{q\bar{q}q\bar{q}}$	2-fermion $\mathrm{q\bar{q}}(\gamma)$	ZZ $\mathrm{q\bar{q}q\bar{q}}$
183 GeV	82.2	1.28	2.12	86.5
189 GeV	82.1	1.29	2.10	88.2
192 GeV	82.5	1.11	1.87	89.3
196 GeV	81.8	1.07	2.00	88.2
200 GeV	81.3	0.91	2.09	87.7
202 GeV	81.7	0.93	2.05	88.2
205 GeV	81.0	0.81	2.10	87.0
207 GeV	80.7	0.93	2.10	87.2

Table 4.1: *Pre-selection efficiencies (in %) for different event types at all LEP2 centre-of-mass energies.*

The pre-selection efficiencies for a two (heavy) boson final state decreases with increasing centre-of-mass energy. This is due to the increasing Lorentz boost the bosons acquire, resulting in an enhanced probability to lose particles in the beam pipe. This in turn results in a slightly enhanced probability for the reconstructed event to be (wrongly) reconstructed as 3 jets instead of 4 or more as can be seen in figure 4.13. This also implies that for a given centre-of-mass energy the pre-selection efficiency for a ZH event depends on the mass of the Higgs boson. For high Higgs masses efficiencies above 90% are obtained. The remaining background from $\mathrm{Z}\gamma^*$ events is small and in signature similar to $q\overline{q}(\gamma)$ events. They will therefore be included in the $q\overline{q}(\gamma)$ numbers in all future plots and tables.

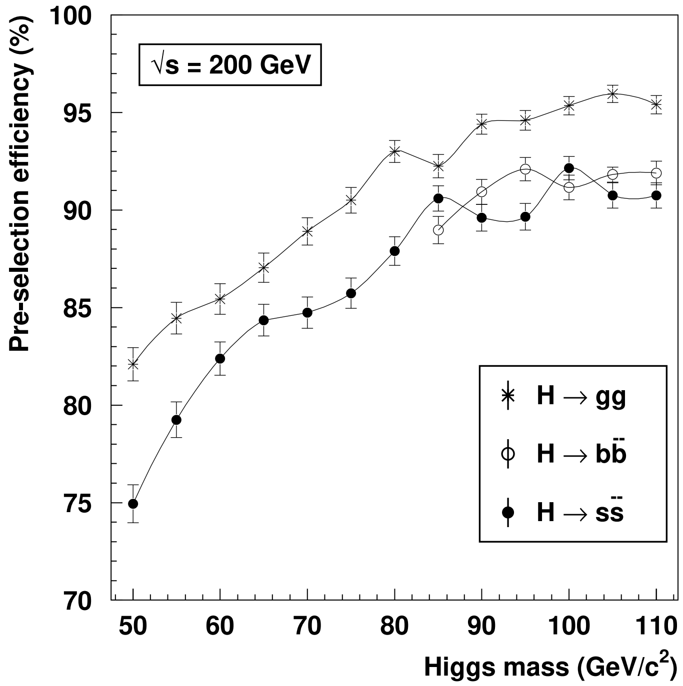


Figure 4.13: *Higgs pre-selection efficiencies as a function of the Higgs mass for different Higgs decay modes. As expected the selection efficiency drops as the Higgs mass decreases since the selection is optimised for heavy boson pair-production.*

Table 4.2 gives a breakdown of the number of expected events from the various physics processes after the pre-selection at the different energies. These numbers are compared with the number of observed events. There is overall good agreement.

$\sqrt{s}$	$\mathcal{L}$	ZZ	W^+W^-	$q\overline{q}(\gamma)$	expected	observed
183 GeV	54.7	6.0	330.5	122.2	458.7	473
189 GeV	158.0	43.6	1015.5	349.9	1409.0	1422
192 GeV	25.9	8.7	169.3	48.9	226.9	239
196 GeV	76.9	29.6	505.6	146.0	681.2	703
200 GeV	84.3	35.2	559.1	160.9	755.2	743
202 GeV	41.1	17.6	271.7	74.4	363.7	368
205 GeV	82.0	34.4	538.3	148.7	721.4	761
207 GeV	142.2	65.2	934.2	250.5	1249.9	1212
total	665.1	240.3	4324.2	1301.5	5866.0	5921

Table 4.2: *The number of expected events after the pre-selection for various physics processes at various centre-of-mass energies. Also the observed number of events at each energy is given.*

Chapter 5

Ingredients in the analysis

In this chapter a method is presented to extract information from events with a 4-quark final state. The idea is to compute the relative probabilities for a given observed event to originate from one of the processes leading to 4 or more reconstructed jets of particles: W^+W^-, ZZ, $q\overline{q}(\gamma)$ or a Higgs boson of any given mass produced together with a Z. The computation combines experimental information from the event with the theoretical expectations from the different hypotheses that will be described analytically where possible. In the first three sections the variables that provide information (in increasing levels of complexity) are introduced and discussed. Finally, section 5.4 describes the method to combine the information on the observed event. The method preserves all ambiguities and thereby uses the full information contained in the variables used.

5.1 Single-jet information

The b-tag configuration of the final states of various processes producing a 4-quark final state at LEP2 is quite different and can be quite distinct for a specific process. The final state produced by a pair of W bosons will for example hardly ever contain a b-quark whereas the most predominant decay channel of the Higgs boson is into a pair of b-quarks. In the first part of this section the predicted number of b-quarks present in final states (and their configurations) for the different physics processes is investigated. From an experimental point of view the goal is to test how well an observed jet is compatible with the hypothesis that it originated either from a b-, or from a non-b-quark fragmentation. In section 5.1.2 such a compatibility is constructed. This can then be converted into a quantitative statement on the compatibility of a jet-pairing with a given hypothesis as is described in section 5.1.3.

5.1.1 Theoretical information

b-quarks in 4-fermion events

In the framework of the electroweak theory the branching ratio of the heavy gauge bosons into quark anti-quark pairs can be computed allowing the relative occurrence of the different b-tag configurations to be predicted (see table 5.1). For a specific process the relative occurrence of these configurations can be quite distinct and is therefore an important ingredient in the analysis. It might also, once a measurement is performed, serve as a cross check to positively identify the underlying physics process and mechanism(s) producing the events. Although the branching ratio of the Higgs boson into a $\mathrm{b\overline{b}}$-pair is dependent on the mass of the Higgs itself, for a given Higgs mass this branching ratio is precisely predicted inside the SM. The relative frequency of various 4-quark (b-tag) configurations in heavy boson pair production and for three different Higgs boson masses are given in table 5.1.

	Br(4 quarks)	$(\mathrm{q\overline{q}})\,(\mathrm{q\overline{q}})$	$(\mathrm{q\overline{q}})\,(\mathrm{b\overline{b}})$	$(\mathrm{b\overline{b}})\,(\mathrm{q\overline{q}})$	$(\mathrm{b\overline{b}})\,(\mathrm{b\overline{b}})$	$(\mathrm{b\overline{q}})\,(\mathrm{q\overline{q}})$+$(\mathrm{q\overline{q}})\,(\mathrm{b\overline{q}})$
ZZ	48.9	61.3	17.0	17.0	4.7	0.0
$\mathrm{W^+W^-}$	45.6	99.7	0.0	0.0	0.0	0.3
ZH(80)	64.2	6.3	72.0	1.8	19.9	0.0
ZH(100)	64.1	8.5	69.8	2.4	19.3	0.0
ZH(115)	63.8	15.2	63.1	4.2	17.5	0.0

Table 5.1: *The fraction of fully hadronic (4-quarks) final states in % for the heavy boson pair-production processes at LEP2. In addition the relative occurrence of the b-tag configurations present in these fully hadronic final states is shown* ($\mathrm{q\overline{q}}$ *means non-*b *quark pair).*

The cross section for b-quark production in $\mathrm{W^+W^-}$ events is tiny since it involves the small CKM matrix elements $\mathrm{V_{cb}}$ and $\mathrm{V_{ub}}$ and is about 0.024 pb. During the generation of 4-fermion events with `EXCALIBUR` an implementation was used where the CKM matrix was assumed to be the unit matrix, thereby explicitly preventing b-quarks to be generated. Although small, the cross section is still about 60% of the cross section of ZH production at 206.7 GeV when the Higgs mass is 115 GeV. To investigate the effect of neglecting b quarks in generating $\mathrm{W^+W^-}$ events on the determination of the $\mathrm{W^+W^-}$ background in different phase space regions (especially where b-quarks are present) dedicated samples have been produced to simulate these events. Due to the use of the anti-$\mathrm{W^+W^-}$ mass information and the presence of long-lived c-quarks in the tails of the b-tag distributions possible effects (underestimation of the $\mathrm{W^+W^-}$ background in the pure ZH region) are diluted.

b-quarks in 2-fermion events

$\mathrm{e^+e^-}$ annihilation is described by both Z and γ exchange (and their interference). The fraction of primary b-quarks produced in $\mathrm{q\overline{q}}(\gamma)$ events at LEP2 is therefore different from that of a Z boson decaying into a $\mathrm{b\overline{b}}$-pair, which has been measured to high precision ($R_\mathrm{b}^0 = \Gamma(\mathrm{Z} \rightarrow \mathrm{b\overline{b}})/\Gamma(\mathrm{Z} \rightarrow \mathrm{had}) = 0.216$, see [66]). Although different from R_b^0 the relative production rate of $\mathrm{e^+e^-}$ into a pair of b-quarks compared to the total hadronic cross section ($R_\mathrm{b} = \sigma(\mathrm{e^+e^-} \rightarrow \mathrm{b\overline{b}})/\sigma(\mathrm{e^+e^-} \rightarrow \mathrm{had})$) is well predicted in the SM and is measured at each centre-of-mass energy by the LEP experiments. The expected value for R_b varies only weakly as a function of the centre-of-mass energies around 200 GeV and is 0.165 [69].

R_b reflects only the fraction of direct $\mathrm{b\overline{b}}$-production and this could be different for the specific set of events that have an apparent 4-jet structure and pass the event pre-selection presented in section 4.5. Corrections are required that reflect the differences between light and b-quarks, for example differences in the amount of gluon radiation of heavy quarks compared to light quarks and the difference in decay multiplicity. From pure kinematics the radiation of gluons from b-quarks is expected to be suppressed compared to radiation off light quarks due to the larger mass of the b-quark. These kind of asymmetries are known, understood and implemented in the generation of background events. A Monte Carlo study of $\mathrm{q\overline{q}}(\gamma)$ events at LEP2 shows that in the selected sample of 4-jet events the fraction of events with two primary b-quarks is 16.1 $\pm$ 0.2 %.

gluon splitting

There are additional issues in the description of $\mathrm{q\overline{q}}(\gamma)$ events at LEP2 that have been a source of many theoretical and experimental debates during the last years [70]. One of the important points relevant for the description of apparent 4-jet $\mathrm{q\overline{q}}(\gamma)$ events is the fraction of (secondary) quark anti-quark pairs produced through gluon splitting. It was found that the gluon splitting into a $\mathrm{b\overline{b}}$-pair was underestimated in PYTHIA [41]. In order to get a correct estimate of the $\mathrm{q\overline{q}}(\gamma)$ background level in regions where the presence of b-quarks is expected events containing these specific splittings were assigned an increased weight. The effect on the fraction of expected jets from $\mathrm{q\overline{q}}(\gamma)$ events that have a specific value of the combined b-tag probability per jet is given in table 5.2. The b-quarks from gluon splitting are softer than the primary b-quarks and therefore do not contribute to the extreme tail of the combined b-tag variable.

combined b-tag per jet	fraction of jets (no rescaling)	fraction of jets (rescaling)
$>$0.0	6.9	7.4
$>$1.0	3.2	3.5
$>$2.0	2.0	2.1

Table 5.2: *The fraction of jets (in %) in* $\mathrm{q\overline{q}}(\gamma)$ *events passing the pre-selection that have a specific value of the combined b-tag variable per jet before and after rescaling of the gluon splitting rate in PYTHIA.*

The effect of remaining uncertainties in background description and discrepancies between data and Monte Carlo for each analysis are incorporated in the systematic uncertainties.

5.1.2 Experimental information

To establish the compatibility between the observed event and a specific b-tag configuration requires a quantitative statement about the '*b-likeliness*' for each reconstructed jet. In this section such a variable will be defined using the combined b-tag variable [63] per jet as described in section 4.4. The distribution of this variable (for selected two jet events at a centre-of-mass energy close to the Z mass) is shown in figure 5.1 for all jets, where the contributions from b-quarks and non-b-quarks are plotted separately.

Once normalised, the distributions of the combined b-tag variable per jet for each jet-type f_i (b-quark or non-b-quark) can be interpreted as a probability density function:

$P_{f_i}(\rho_i)$ The probability density (function) for a quark of flavour f_i (b or non-b) to produce a jet with a value ρ_i for the combined b-tag variable.

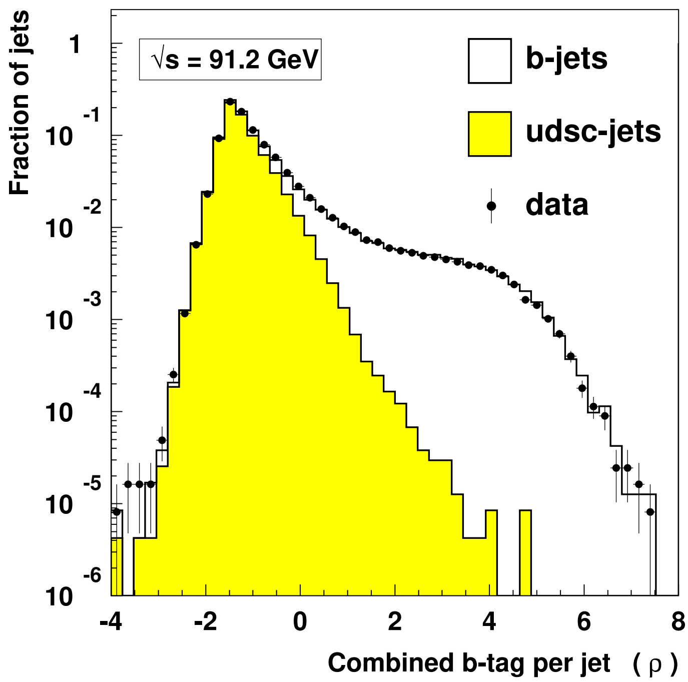


Figure 5.1: *Combined b-tag variable for jets in Z-decays. The data are represented by the dots and the two Monte Carlo histograms represent the different types of jets (b- and non-b jets).*

In addition, the b-tag likelihood ratio per jet for jet i is defined as:

$$\frac{\mathrm{P_b}(\rho_\mathrm{i})}{\mathrm{P_q}(\rho_\mathrm{i})}, \tag{5.1}$$

where q stands for any type of light quark (u,d,s,c). This ratio tells, for each value of the combined b-tag variable ρ_i, *how much more likely* it is that this reconstructed jet originated from a b-quark fragmentation relative to a non-b-quark fragmentation.

The likelihood ratio per jet is constructed using Z^0 Monte Carlo events, providing a clean environment (two jets back-to-back) in which the few million Z-decays at LEP1 have ensured that the Monte Carlo's have been tuned to describe the data to great precision. An experimental feature to take into account when constructing the likelihood ratio is that the discriminating power between b- and non-b-jets deteriorates when going to smaller reconstructed jet-angles. The combined b-tag variable uses track-by-track errors and since these are dependent on the polar angle of the track, the separation is less performant for jets in the forward region. To account for these differences in separation power the b- over non-b ratio is parametrised in three different angular regions: (0° to 30° & 150° to 180°), (30° to 40° & 140° to 150°) and (40° to 140°). The two probability density functions in each region, the corresponding likelihood ratio and its parametrisation (polynomials) are shown in figure 5.2.

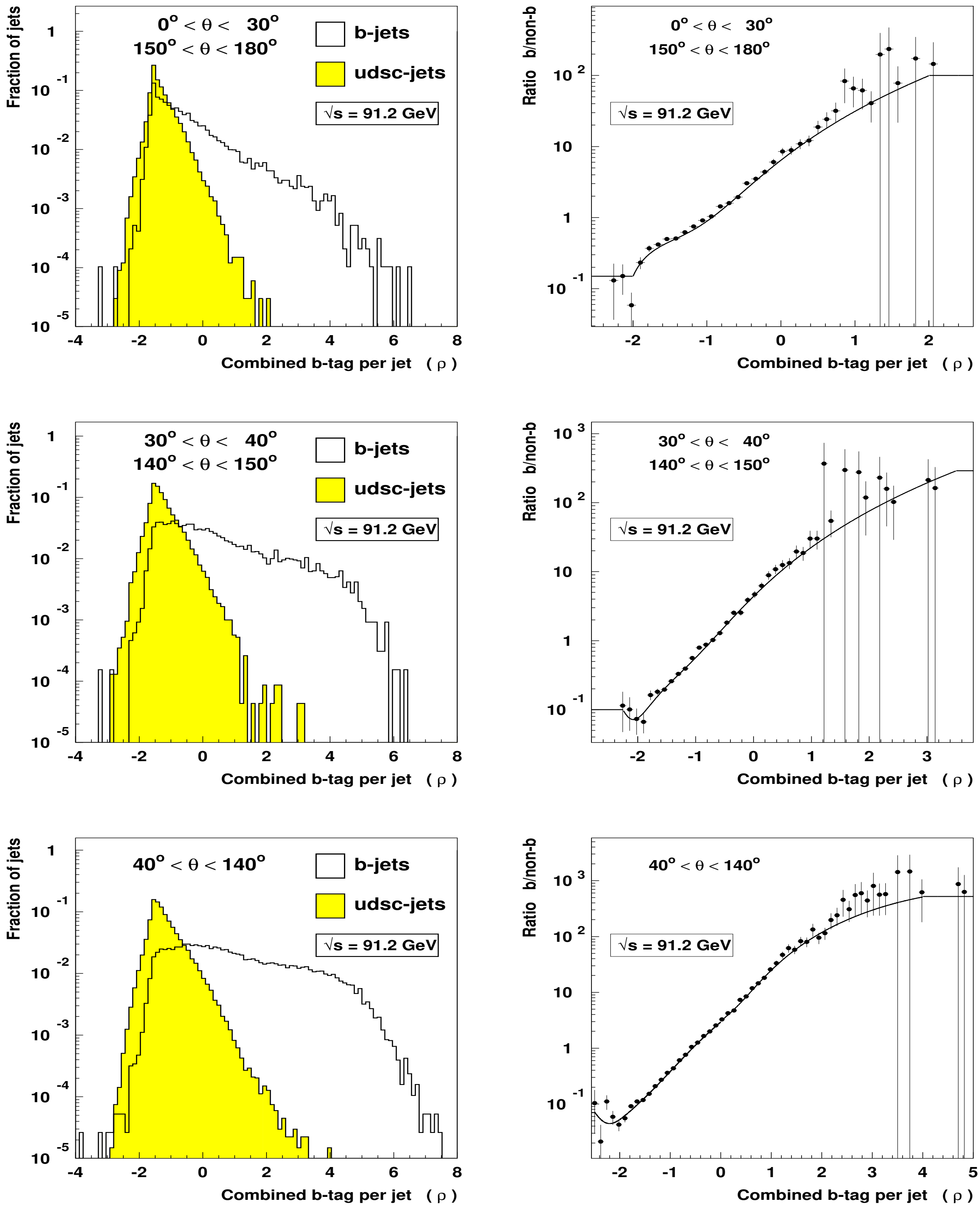


Figure 5.2: *The probability density functions for the combined b-tag variable for jets from Z-decays reconstructed in three angular regions for b- and non-b quarks separately (left plots) and the ratio of the two distributions (likelihood ratio) (right plots). The solid line in the right plots indicates the parametrisation (4th order polynomials) used in the analysis.*

5.1.3 Application in a multi-jet environment at LEP2

In a multi-jet event this allows (from an experimental point of view) to compute the relative probabilities of specific b-tag configurations relative to the hypothesis that they are all non-b quarks by multiplying the individual likelihood ratios per jet. An example: the probability that in an observed 4-jet event jet 1 and 2 are b's and jet 3 and 4 are non-b's relative to the probability that all 4 jets are non-b quarks can be written explicitly as:

$$\frac{P_{\mathrm{b}}(\rho_1)P_{\mathrm{b}}(\rho_2)P_{\mathrm{q}}(\rho_3)P_{\mathrm{q}}(\rho_4)}{P_{\mathrm{q}}(\rho_1)P_{\mathrm{q}}(\rho_2)P_{\mathrm{q}}(\rho_3)P_{\mathrm{q}}(\rho_4)} = \frac{P_{\mathrm{b}}(\rho_1)}{P_{\mathrm{q}}(\rho_1)} \cdot \frac{P_{\mathrm{b}}(\rho_2)}{P_{\mathrm{q}}(\rho_2)} \tag{5.2}$$

Multi-jet events at LEP2 have different characteristics than 2-jet events at the Z: higher track multiplicities, a large spread in jet-energies, overlapping jets. Using the constructed parametrisations at the Z in multi-jet events might cause systematic effects since distributions entering the combined b-tagging analysis might be strongly dependent on the energy of the jets and might involve non-trivial extrapolations to different energies. To first order no large differences are expected since, as an approximation, all variables in the combined b-tagging either scale with the jet-energy (like life-time information) or are (almost) energy independent (like jet mass, jet multiplicity and the energy of the B-hadron compared to that of the initial b-quark). Secondary effects like a higher mis-assignment rate of particles to jets caused by the higher multiplicity and the overlap of jets are (partly) corrected for by tuning the impact parameters (as described on page 43) on tracks from real 4-jet events.

To give an idea of the agreement between data and Monte Carlo at high energy the distribution of the combined b-tag variable per jet (ρ) is plotted for all reconstructed jets in events passing the pre-selection in figure 5.3. In this plot all centre-of-mass energies at LEP2 have been combined. The ratio of the efficiency for the data and Monte Carlo is shown in the right plot of that figure, where the efficiency for a given ρ_0 is defined as the fraction of jets with $\rho > \rho_0$. A good agreement between data and Monte Carlo is observed.

5.2 Di-jet information

Although a strong indication of a specific b-tag configuration can eliminate many physics hypotheses, additional information is needed to test the compatibility of the other event characteristics with that dictated by the hypothesis. In case there are 4 reconstructed b-quarks in the event for example angular distributions, cross sections and mass information decide on the relative probability of either a ZZ or ZH hypothesis.

This section describes the extraction of information on a different level of complexity in the event: invariant mass information from pairs of reconstructed jets. Reconstructed mass distributions can, just as specific b-tag configurations, be used as a fingerprint of a certain physics hypothesis. The expected mass distributions are known analytically (see section 5.2.1), but from an experimental point of view there are some difficulties. In a 4(5)-jet event there are 3(10) possible jet-jet pairings. In addition the quality of the energy flow reconstruction can vary significantly from

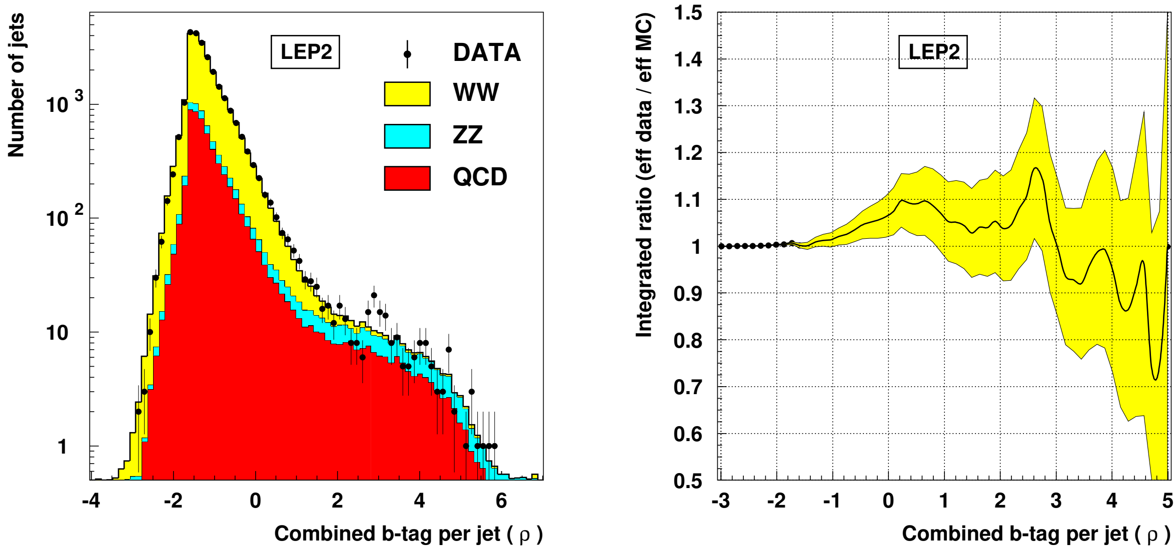


Figure 5.3: *The left plot shows the combined b-tag probability per jet for all reconstructed jets after pre-selection in the data set 183-207 GeV. The right plot shows the integrated ratio of the efficiency for data and the efficiency for Monte Carlo. The gray area in the right plot indicates the statistical uncertainty.*

event to event. The goal is to implement this knowledge in the interpretation of the reconstructed masses. In general these problems can be solved by introducing a tool to test the compatibility of the reconstructed di-jet masses with a specific set of mass hypotheses using the event-by-event errors. This is shown in section 5.2.3.

5.2.1 Expected mass distributions

The mass of a single heavy gauge boson is described by a Breit-Wigner shape [26] determined by the pole mass (m_0) and the intrinsic width (Γ_0) of the heavy boson:

$$BW(m_1 \mid m_0, \Gamma_0) \propto \frac{m_1^2}{(m_1^2 - m_0^2)^2 + (\frac{\Gamma_0}{m_0} m_1^2)^2} \tag{5.3}$$

In an event containing two heavy bosons the probability density function for the distribution of the two boson masses (m_1 and m_2) in a single event can be described by the (normalised) product of two Breit-Wigners times a phase space factor:

$$P^{\text{mass}}_{\text{theory}}(m_1, m_2 \mid \text{hyp}) \;\propto\; BW(m_1 \mid m_{01}, \Gamma_{01})\; BW(m_2 \mid m_{02}, \Gamma_{02})\; PS(m_1, m_2, s), \tag{5.4}$$

where the phase space factor $PS(m_1, m_2, s)$ equals $\sqrt{(1 - m_1^2/s - m_2^2/s)^2 - 4m_1^2 m_2^2/s^2}$. In this expression s is the available centre-of-mass energy in the collision and 'hyp' indicates the underlying physics hypothesis given by $(m_{01}, \Gamma_{01}, m_{02}, \Gamma_{02})$. To be interpreted as a probability density function equation (5.4) is normalised by requiring the integral over the allowed phase space to be equal to unity:

$$\int_{M_{\rm min}}^{\sqrt{s}} dm_1\; BW(m_1 \mid m_{01}, \Gamma_{01}) \int_{M_{\rm min}}^{\sqrt{s}-m_1} dm_2\; BW(m_2 \mid m_{02}, \Gamma_{02}) PS(m_1, m_2, s) = 1 \qquad (5.5)$$

A few additional remarks should be made:

- Picking a specific value for the jet resolution parameter ($y_{\rm cut}$) implies a minimum in the jet-jet masses that can be obtained. At a $y_{\rm cut}$ of 0.001 (Durham) for example this minimum mass ($M_{\rm min}$) is about 13 GeV/c^2.
- Expression (5.4) describes the probability density function for heavy boson masses and therefore the expected mass distribution in the correct pairing. The generated mass spectra (m_1, m_2) for the combinatorial background in heavy boson pair production, arising from the additional 2(9) 'wrong' pairings in the reconstructed 4(5) jets event, is parametrised as a flat distribution (multiplied by the additional phase space factor).
- For all pairings in $q\overline{q}(\gamma)$ events the (m_1, m_2) spectrum is parametrised as a flat distribution (multiplied by the additional phase space factor).

As an example, the mass distribution of a single W-boson and of the two masses in events where W-bosons are pair-produced are shown in figure 5.4. The analytic parametrisations (5.3) and (5.4) used in the analysis to represent these distributions are also shown.

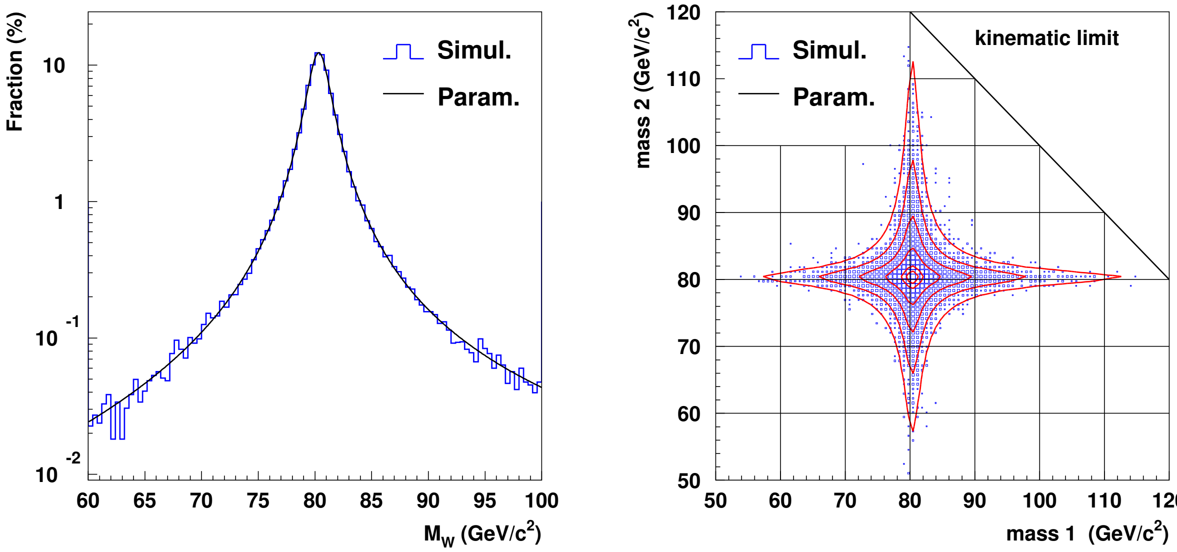


Figure 5.4: *The generated mass distributions for a single W-boson (left plot) and the 2-dimensional mass distribution in events where W-bosons are pair-produced at a centre-of-mass energy of 200 GeV. The solid lines show the analytic parametrisation used in the analysis to represent these distributions.*

5.2.2 Reconstructed mass distributions

For a given jet-pairing in a multi-jet event, the mass information is given to first order by only two numbers: the invariant masses of the two di-jet systems (m_1^{4C}, m_2^{4C}), computed using the estimates for the parton four-momenta from the 4C-fit as described in section 4.3. In this section an effort is made to try and quantify our experimental knowledge on the compatibility of these di-jet masses with masses that are different from the 4C masses, using the event-specific errors taken from the constrained fit. These errors can vary significantly from event to event reflecting the quality of the reconstruction of the energy-flow. Testing the compatibility for a specific pairing with a specific set of masses (m_1, m_2) contains two additional constraints apart from energy and momentum conservation and is therefore called a 6C-fit:

6C-fit: Like a 4C-fit and in addition for a given pairing **two** masses are known.
For a given pairing a compatibility (defined by the χ^2) of the reconstructed masses can be tested with a specific mass hypothesis where both masses are known (m_1, m_2):

$$\begin{aligned} \sum_{i=1}^{n_{\text{jets}}} p_i^{(\text{fitted})\ \mu} &= (0, 0, 0, \sqrt{s}) \\ \mathrm{M}_{\text{dijet 1}} &= m_1 \\ \mathrm{M}_{\text{dijet 2}} &= m_2 \end{aligned} \tag{5.6}$$

For each pairing the 6C-fit can be used to scan the full (m_1, m_2)-plane in order to test the compatibility with any specific mass hypothesis. The resulting $\Delta\chi^2$ distribution in the full (m_1, m_2)-plane can be interpreted as the probability that in this event and for this particular jet pairing the two (**observed**) heavy boson masses in the event were actually m_1 and m_2:

$$P_{\text{fit}}^{\text{mass}}(m_1, m_2) \propto e^{-\frac{1}{2}\Delta\chi^2(m_1, m_2)}, \tag{5.7}$$

where $\Delta\chi^2 = \chi^2 - \chi^2_{4C}$ and χ^2_{4C} is the χ^2 of the 4C-fit. In figure 5.5 a plot of the $\Delta\chi^2$ distribution for a single 4-jet event is shown. In this event one can clearly identify the three different jet pairings and the resolution on any di-jet mass in this event. In this figure also an important kinematic feature is evident: when the sum of the di-jet masses gets closer to the kinematic limit the error on the difference of the masses is much larger than the error on the sum of the masses. Figure 5.5 represents the full experimental mass information and allows to take into account all known kinematic correlations and di-jet mass resolutions on an event-by-event basis.

From an experimental point of view all pairings have equal weight and the pairing ambiguity intrinsic for the 4-quark final state can not be resolved. Only after a certain hypothesis is posed the different pairings can be assigned a certain weight in correspondence to the compatibility with the expectation. Such a compatibility is constructed in the next section.

5.2.3 Constructing a quantitative compatibility with a physics hypothesis

Once an experimental compatibility with any two masses is constructed for each pairing one can go one step further and assign a variable that tests the compatibility of the masses in this specific

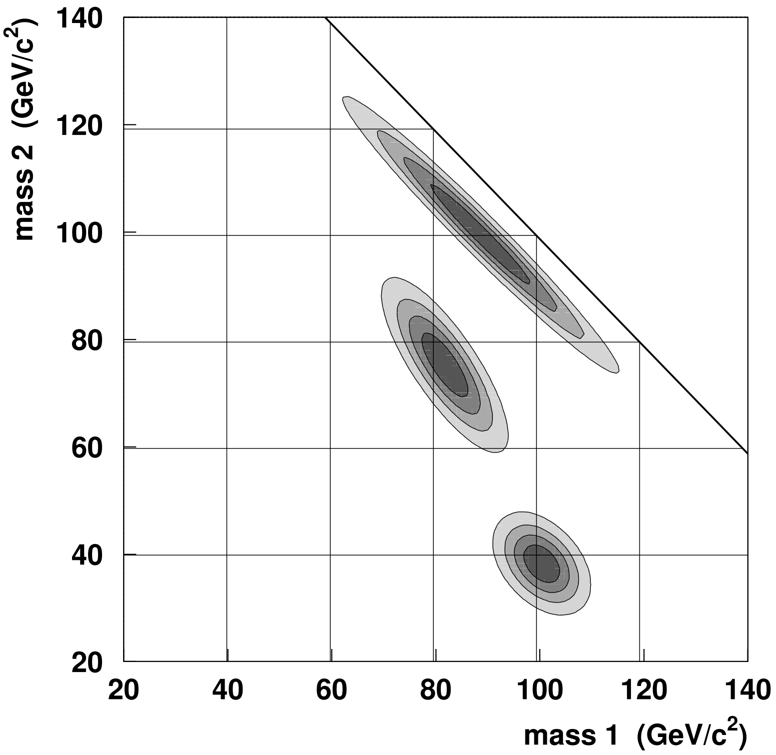


Figure 5.5: *2-dimensional $\Delta\chi^2$ distribution in a single 4-jet event. The three different pairings can be clearly identified. The different contours represent the 1, 2, 3 and 4 sigma contours.*

pairing with a predefined physics hypothesis.

When testing a physics hypothesis, in each point of the (m_1, m_2) plane there are now two distributions providing information on the probability of observing these two masses in an event. On the theoretical side one has the probability density function of the produced masses described by the product of two Breit-Wigners[1] multiplied by a phase space factor (equation (5.4)), whereas on the experimental side one has the $\Delta\chi^2$ distribution (equation (5.7)). Considering a single pairing at a time, the compatibility of the masses in this pairing with a posed hypothesis characterised by the two heavy boson masses and their widths $(m_{01}, m_{02}, \Gamma_{01}, \Gamma_{02})$, can be quantified by convoluting the two distributions:

$$P_{\text{hyp}}^{\text{mass (pair)}}(m_1^{4\text{C}}, m_2^{4\text{C}}, \Delta M) \; = \tag{5.8}$$

$$\int_{M_{\min}}^{\sqrt{s}} dm_1 \; BW(m_1 \mid \text{hyp}) \int_{M_{\min}}^{\sqrt{s}-m_1} dm_2 \; BW(m_2 \mid \text{hyp}) \cdot PS(m_1, m_2, s) \; e^{-\frac{1}{2}\Delta\chi^2(m_1,m_2)}$$

In this expression ΔM represents the experimental uncertainties on the reconstructed masses from the 4C fit. Using this expression, the compatibility for each pairing with specific hypotheses

[1] Or a flat distribution when testing the $\mathrm{q\overline{q}}(\gamma)$ hypothesis.

like W^+W^-, a $q\overline{q}(\gamma)$, a ZZ or a ZH hypothesis with known Higgs mass (m_H) can be computed. There are two possibilities to use these mass compatibilities per pairing in the analysis:

1 Use the information per pairing to choose the best or most likely pairing
Since maximally one jet-pairing represents the correct pairing and thereby the correct mass information on the two produced heavy bosons, many analyses try to find the most likely pairing among the 3(10) pairings in the reconstructed 4(5) jet event. Deciding on a single pairing automatically resolves the existing pairing ambiguity. 'Most likely' is analysis dependent as is illustrated in figure 5.5 where a W^+W^- cross section analysis and an analysis searching for a SM Higgs of 100 GeV/c^2 mass will choose a different pairing. By resolving the ambiguities both analyses discard relevant information (the masses in other pairings are compatible with a different hypothesis). Even though it is not optimal this is the most widely adapted approach.

2 Use all pairings, weighted by their compatibility with a certain mass hypothesis
By assigning a weight to each pairing, all pairings can be treated consistently and no ambiguities have to be resolved. This requires a special structure in the analysis as will be described in section 5.4. Extracting the mass information per pairing allows an optimal combination with other variables in a consistent way. This is the approach used in this thesis.

Example: Performance test of mass compatibilities for full events

The relative compatibility of events with the three SM hypotheses (ZZ, W^+W^- and $q\overline{q}(\gamma)$) can be studied using only mass information. This is done by computing for each jet-pairing the compatibility with the three different physics hypotheses using equation (5.9). The event probability is constructed by looping over all jet-pairings and adding the compatibilities. Finally, the mass-only event probabilities are normalised to the sum of all investigated hypotheses:

$$P_{\text{hyp}}^{\text{mass (event)}} = \frac{\sum_{\text{pairing}} P_{\text{hyp}}^{\text{mass (pair)}}(m_1^{4C}, m_2^{4C}, \Delta M)}{\sum_{\text{hyp=(ZZ,WW,QCD)}} \sum_{\text{pairing}} P_{\text{hyp}}^{\text{mass (pair)}}(m_1^{4C}, m_2^{4C}, \Delta M)} \tag{5.9}$$

The normalised distributions of $P_{\text{ZZ}}^{\text{event}}$, $P_{\text{WW}}^{\text{event}}$ and $P_{\text{QCD}}^{\text{event}}$ for Monte Carlo events with 4 reconstructed jets from three different physics processes (ZZ, W^+W^- and $q\overline{q}(\gamma)$ events) at a centre-of-mass energy of 200 GeV are shown in figure 5.6. For events with 4 jets a cut on the ZZ mass-only event compatibility ($P_{\text{ZZ}}^{\text{event}} > 0.80$) would result in a selection efficiency of 57.1% for ZZ, 6.1% for W^+W^- and 14.8% for $q\overline{q}(\gamma)$ 4-jet-events. These event probabilities using only mass information serve as inputs for the common DELPHI Higgs analysis in the 4-jet channel where the compatibilities are calculated in both the 4- and the 5-jet topology to reduce the W^+W^- and ZZ contamination [71].

From the available phase space it is expected that quite a large fraction of $q\overline{q}(\gamma)$ events has a di-jet mass combination near the W^+W^- or ZZ peak thereby mimicking a W^+W^- or ZZ event. Therefore, as can be seen in the right bottom plot of figure 5.7, less than half of the $q\overline{q}(\gamma)$ events can be clearly identified as a $q\overline{q}(\gamma)$ event when using only mass information.

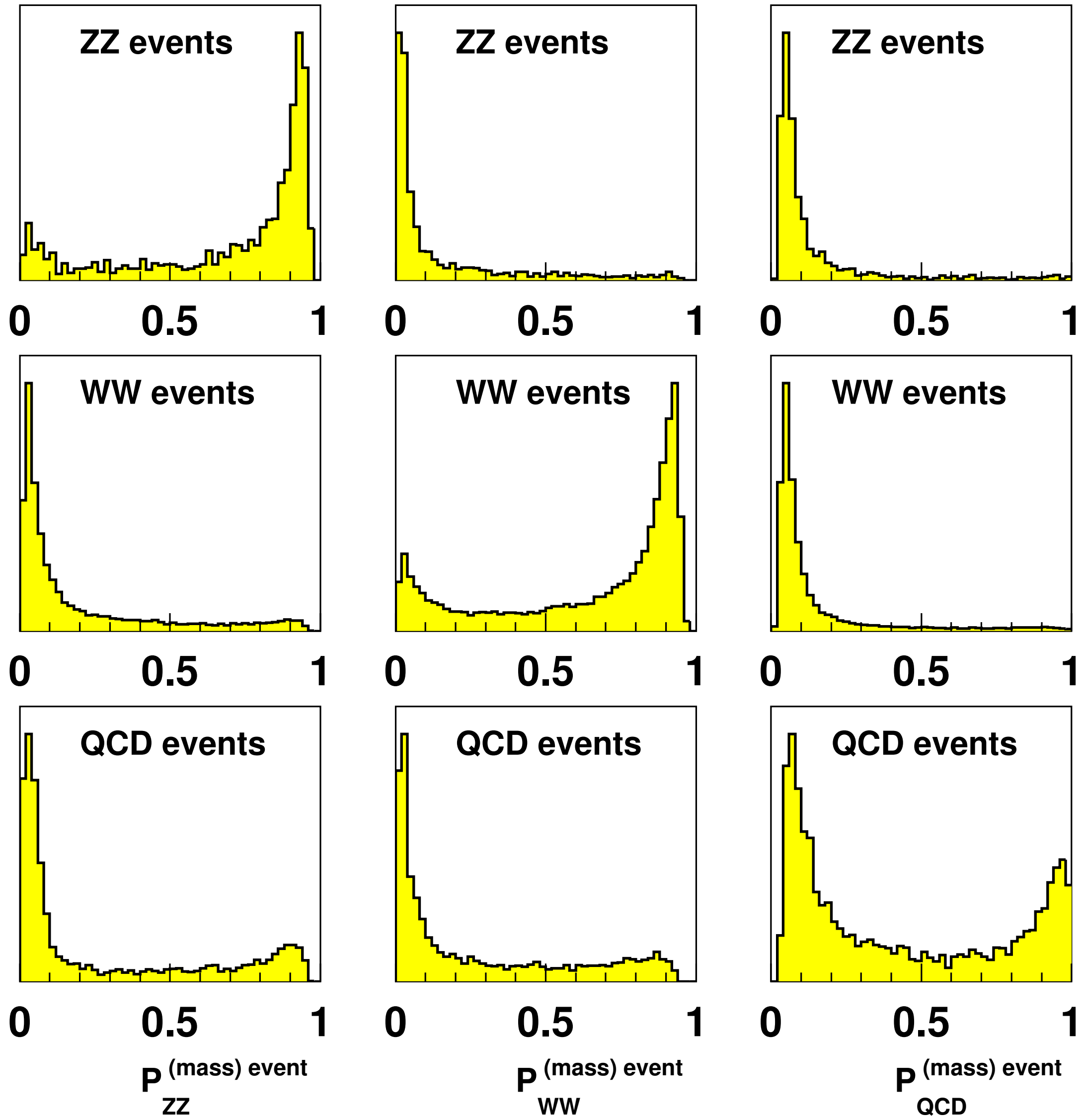


Figure 5.6: *Normalised probability density functions describing the compatibilities of the reconstructed masses in ZZ, W^+W^- and $q\overline{q}(\gamma)$ events (with 4 reconstructed jets) with three different hypotheses. The three top plots show the relative 'mass only' probabilities for ZZ events to be compatible with either a ZZ (left), a W^+W^-(middle) or a $q\overline{q}(\gamma)$(right) hypothesis. The middle three plots show the same distributions for W^+W^- events and the bottom plots for $q\overline{q}(\gamma)$ events. The Monte Carlo events for the different processes were generated at a centre-of-mass energy of 200 GeV.*

The comparison between data and Monte Carlo is shown in figure 5.7. The distributions of $\mathrm{P}_{\mathrm{ZZ}}^{\mathrm{mass\ (event)}}$ and $\mathrm{P}_{\mathrm{WW}}^{\mathrm{mass\ (event)}}$ are shown for all events in the data set 183-207 GeV that have 4 reconstructed jets and pass the event pre-selection. A good agreement between data and Monte Carlo is observed.

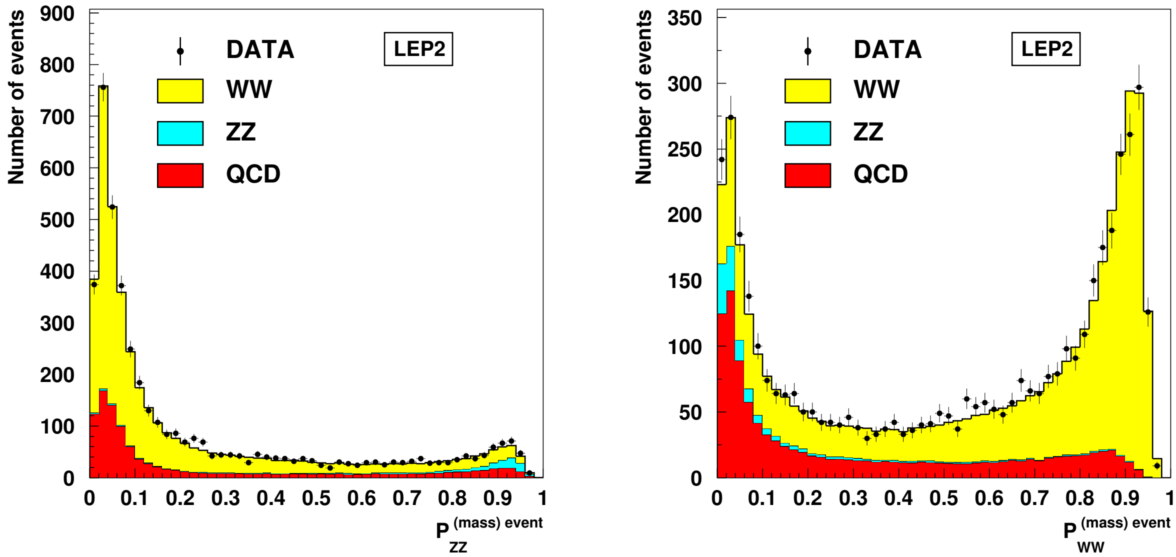


Figure 5.7: *The distribution of* $P_{\mathrm{ZZ}}^{\mathrm{mass\ (event)}}$ *(left plot) and* $P_{\mathrm{WW}}^{\mathrm{mass\ (event)}}$ *(right plot) for events that have 4 reconstructed jets after pre-selection in the data set 183-207 GeV.*

5.3 Event topological information

As described in a previous section, the topological structure of the jets in an event is different for 2-fermion ($\mathrm{q\overline{q}}(\gamma)$) events with hard gluon radiation and genuine 4-fermion events. Using the fact that the gluon emission has infrared (low energetic gluon jets) and collinear divergences (two jets close together) a topological variable can be constructed to separate spherical (4-fermion) from more cigar-like (2-fermion) events.

Optimal separation power can be achieved by using a combination of different event-shape variables. Many of these variables are strongly correlated with di-jet invariant masses in the event and using them to select events will restrict specific combinations of inter di-jet opening angles and jet-energies. To easily combine the topological information with the mass information later in the analysis a better approach is to sacrifice some separation power for simplicity as is done in this analysis by using a somewhat less performant single variable ($\mathrm{D_{pur}}$) as defined in equation (4.12). It combines the two most obvious differences in topology and has the advantage of being nearly independent of the value of the largest di-jet mass in the event.

To optimally discriminate between the different topologies the sample is further divided into events with 4 and 5 reconstructed jets: the relative number of $\mathrm{q\overline{q}}(\gamma)$ events with 5 reconstructed jets is smaller than that for genuine 4-fermion events.

The normalised distribution of $\mathrm{D_{pur}}$ is shown in figure 5.8 (for 4 and 5 jets events separately) for two different event types after pre-selection: 2-fermion events (generated using `PYTHIA`[41]) and 4-fermion events (generated using `EXCALIBUR` [34]) at a centre-of-mass energy of 200 GeV.

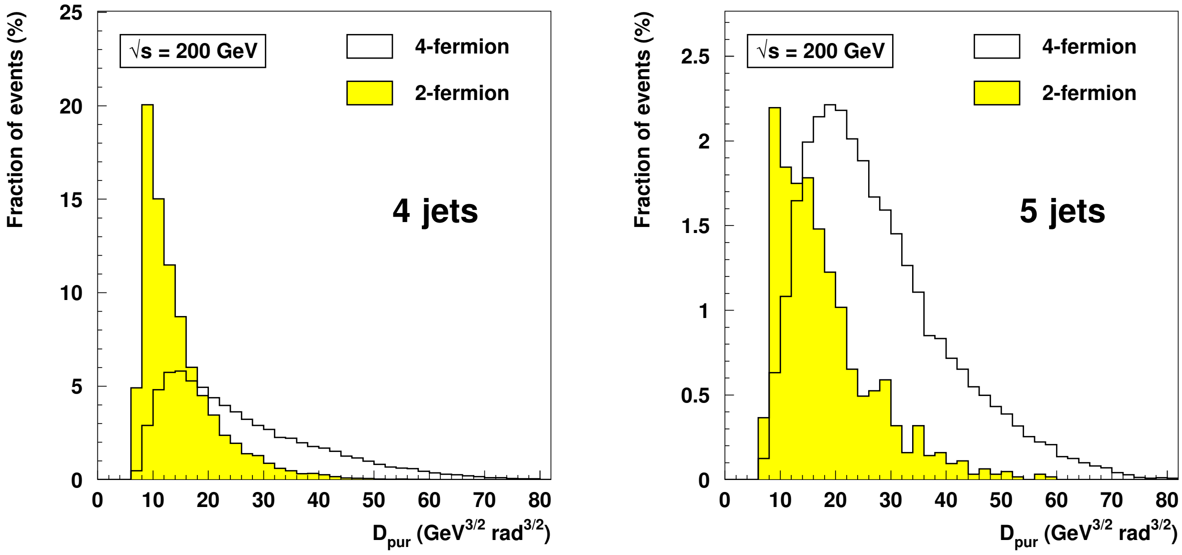


Figure 5.8: *Topological variable D_{pur} for reconstructed 4 (left plot) and 5-jet (right plot) events at a centre-of-mass energy of 200 GeV. The distribution for both 2-fermion and 4-fermion events is shown separately. For each event type the sum of the 4- and 5-jet distributions is normalised.*

These distributions can be interpreted as probability density functions.

$P^{\text{topo}}_{\text{hyp}}\,(\mathrm{D_{pur}})$ The probability for an event from the physics process hyp to have (4 or 5 reconstructed jets and) this value for $\mathrm{D_{pur}}$.

Parametrising the ratio of the two probability density functions allows computation of the relative probabilities that the observed event (given only topological information) originated from either a 2-fermion or a 4-fermion event:

$$\text{Topological likelihood ratio per event:} \quad \frac{P^{\text{topo}}_{\text{hyp}}\,(\mathrm{D_{pur}})}{P^{\text{topo}}_{\text{qcd}}\,(\mathrm{D_{pur}})} \tag{5.10}$$

The parametrisation (using polynomials) of the ratio is done at every centre-of-mass energy separately to account for small changes in kinematics. The description of the general topology of multi-jet events at LEP is well described by the Monte Carlo. The distribution of $\mathrm{D_{pur}}$ for all events after pre-selection in the data set 183-207 is shown in figure 5.9 for events with 4 and 5 reconstructed jets separately and shows very good agreement between data and Monte Carlo.

5.4 Computing event-by-event probabilities

For each process a probability can be computed to produce an event with characteristics identical to the ones in the observed event. In this section a general compatibility expression will be constructed for 4-quark final states by combining all sources of information. This allows a wide variety of applications as will be discussed in the following chapters.

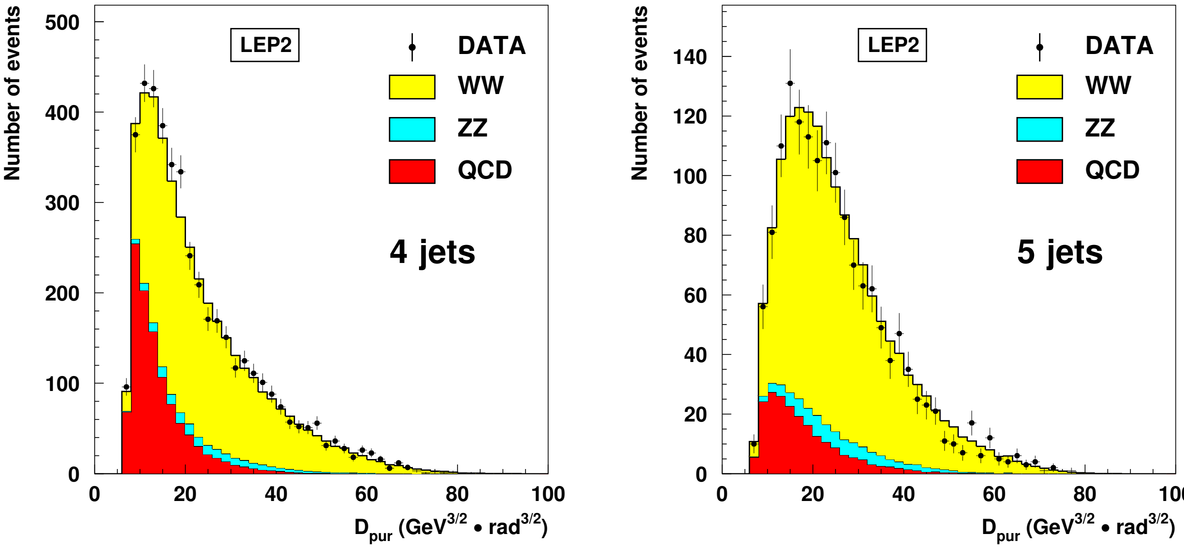


Figure 5.9: *Topological variable D_{pur} for events with 4 (left plot) and 5 (right plot) reconstructed jets after pre-selection in the data set 183-207 GeV.*

5.4.1 Initial probability expressions

In case there is no information contained in the event the probability[2] for this event to have been produced by a given process (p) can be expressed in terms of the cross section for that process, the branching ratios for that process into its possible final states (quarks, leptons, neutrinos) and the specific efficiency to pass the event pre-selection:

$$P_{\mathrm{process}} \sim \sigma_{\mathrm{p}} \sum_{\mathrm{fs}} \mathrm{Br(fs)}\, \epsilon_{\mathrm{p}}^{\mathrm{fs}} \tag{5.11}$$

Where the following symbols have been used:

p (process)	A certain physics process (for example ZZ or $\mathrm{W^+W^-}$)
σ_{p}	The cross section for this process
$\sum_{\mathrm{fs}}$	The sum over all final states
$\mathrm{Br(fs)}$	The branching ratio for this process p into this final state
$\epsilon_{\mathrm{p}}^{\mathrm{fs}}$	The probability for an event of this process and in this final state to pass the event pre-selection.

Since (almost) only events with 4 quarks in the final state are selected the probability expression can be simplified by eliminating the redundant terms concerning final states other than $\mathrm{q\bar{q}q\bar{q}}$. In doing this the following assumptions were made:

- The efficiencies for non-hadronic configurations to pass the event preselection are zero.
- The efficiencies for different hadronic (b-tag) configurations of a process to pass the event pre-selection are identical and equal to: $\epsilon_{\mathrm{p}}^{\mathrm{had}}$.

[2] All computations will be done in the framework of the SM.

This allows the probability expression to be written as a sum over the different hadronic (b-tag) configurations only:

$$P_{\text{process}} \; = \; \sigma_{\text{p}} \, \epsilon_{\text{p}}^{\text{had}} \sum_{\text{had.fs.}} \text{Br (had.fs.)} \tag{5.12}$$

Once an event is detected the measurement allows the characteristics of the event to be extracted in the form of observables. Each piece of information increases our knowledge on the underlying physics process producing this event: for each physics process the SM prediction for a given observable is known. Differences in these distributions for various physics processes allow to differentiate between them. In the preceding sections a few observables have been constructed providing information at different levels of complexity. These variables will be used in the expression of the probability for the observed event to originate from a particular physics hypothesis. All variables used in the computation of the probability are uncorrelated (at least to first order) and therefore enter linear in the expression for the probability.

5.4.2 Adding topological information

At the level of the full event, the event shape can be used to get information on the compatibility of the event to either originate from a 2-fermion or from a 4-fermion process using the probability density functions defined in equation 5.10. For each final state the probability to have a value of D_{pur} equal to the one in the observed event is known. Assuming that for a process p, the topological variable is identical for the different hadronic (b-tag) configurations and equal to $P_{\text{p}}^{\text{had}}(\text{D}_{\text{pur}})$ the probability expression can now be written as:

$$P_{\text{process}} \; = \; \sigma_{\text{p}} \, \epsilon_{\text{p}}^{\text{had}} \, P_{\text{p}}^{\text{had}} \, (\text{D}_{\text{pur}}) \sum_{\text{had.fs.}} \text{Br (had.fs.)} \tag{5.13}$$

5.4.3 Adding b-tag and mass information

In expression 5.13 the sum over the different hadronic final states is composed of the different b-tag configurations and jet-pairings and can be written more explicitly as:

$$\sum_{\text{had.fs.}} = \sum_{\text{btag}=1}^{16(31)} \sum_{\text{pair}=1}^{3(10)} \tag{5.14}$$

Writing it out in more detail allows to incorporate the b-tag and mass information as introduced in the previous sections:

- **b-tag configurations:**
  Each of the jets can be either a b- or a non-b-jet resulting (in a 4-jet event) in 16 different b-tag configurations. In a 5-jet event one of the jets is assumed to be a gluon jet (non-b), resulting in 31 different configurations.

  For each b-tag configuration the expected probability that this particular configuration occurs is known (see section 5.1.1) and this is written here as Br(had.fs.). On the experimental side we have the tool to quantify the compatibility of a given jet(-pair) with a given b-tag configuration (see section 5.1.2 and 5.1.3).

- **jet-pairings:**
  In case of 4(5) reconstructed jets there are 3(10) possible jet-pairings. In case of a ZH hypothesis one has for each jet-pairing the additional ambiguity for each pair to be either the Higgs or the Z, increasing the number of pairings by a factor 2.

  For each jet-pairing the compatibility of the reconstructed masses with the expectations from the given process is also known and is evaluated using expression (5.9).

$P_{\mathrm{p}}^{\mathrm{mass\ (pair)}}\,(m_1^{4C}, m_2^{4C}, \Delta M)$	The probability that the reconstructed masses in this event (and for this particular jet pairing) are compatible with the expected mass distribution for this process.

Incorporating all this information results in the following expression indicating the probability for a given process to produce an event with characteristics identical to that in the observed event:

$$\begin{aligned} P_{\mathrm{process}} &= \sigma_{\mathrm{p}}\, \epsilon_{\mathrm{p}}^{\mathrm{had}}\, P_{\mathrm{p}}^{\mathrm{had}}(\mathrm{D_{pur}}) \sum_{\mathrm{btag}=1}^{16(31)} \mathrm{Br\ (had.fs.)} \\ & \sum_{\mathrm{pair}=1}^{3(10)} P_{f_1}(\rho_1) P_{f_2}(\rho_2) P_{f_3}(\rho_3) P_{f_4}(\rho_4) \quad P_{\mathrm{p}}^{\mathrm{mass\ (pair)}}(m_1^{4C}, m_2^{4C}, \Delta M) \end{aligned} \tag{5.15}$$

5.4.4 Using relative probabilities

As was already remarked in the previous sections, the probability for a process p to have exactly the characteristics as in the observed event is zero, and only relative probabilities (likelihood ratios) are well defined quantities. To classify events only these relative probabilities are needed so the full expression in 5.15 can be multiplied by an arbitrary number which can conveniently be chosen to be:

$$\frac{1}{P_{\mathrm{qcd}}^{\mathrm{had}}(\mathrm{D_{pur}})} \cdot \frac{1}{P_q(\rho_1) P_q(\rho_2) P_q(\rho_3) P_q(\rho_4)} \tag{5.16}$$

This is the inverse of the probability that the event originated from a $\mathrm{q\overline{q}}(\gamma)$ event (looking only at topological information) and that all jets are light (non-b) quarks. This normalisation results in an expression where the likelihood ratios as constructed in the previous sections appear:

$$\begin{aligned} P_{\mathrm{process}} &= \sigma_{\mathrm{p}}\, \epsilon_{\mathrm{p}}^{\mathrm{had}}\, \frac{P_{\mathrm{p}}^{\mathrm{had}}(\mathrm{D_{pur}})}{P_{\mathrm{qcd}}^{\mathrm{had}}(\mathrm{D_{pur}})} \sum_{\mathrm{btag}=1}^{16(31)} \mathrm{Br\ (had.fs.)} \\ & \sum_{\mathrm{pair}=1}^{3(10)} \frac{P_{f_1}(\rho_1) P_{f_2}(\rho_2) P_{f_3}(\rho_3) P_{f_4}(\rho_4)}{P_q(\rho_1)\ P_q(\rho_2)\ P_q(\rho_3)\ P_q(\rho_4)} \quad P_{\mathrm{p}}^{\mathrm{mass\ (pair)}}(m_1^{4C}, m_2^{4C}, \Delta M) \end{aligned} \tag{5.17}$$

5.5 Conclusions

Using this simple algorithm to classify 4(5)-jet events, only composed of three variable types, we are able to compute absolute event-by-event probabilities for various physics hypotheses. The analysis does not need extensive tuning on Monte Carlo samples and can be used to study 4-quark final states at any centre-of-mass energy. All information contained in the event is extracted and treated in a consistent way, which has the advantage that the analysis is simultaneously optimised for selecting fully hadronic W^+W^-, ZZ and $q\overline{q}(\gamma)$ events. Using analytical expressions for the expected mass distributions and the use of the full mass information in the event ensures that the analysis is also optimised for ZH events of any given Higgs mass. The analysis can be improved by including additional information of which angular information on the boson production angles and the quark anti-quark separation for b-quarks are the most obvious. The final probability distributions provide optimal sensitivity to extract the corresponding cross sections by means of a maximum likelihood fit.

In this thesis a few applications will be discussed, starting with the measurement of the fully hadronic ZZ (NC02) cross section in chapter 6 followed by the search for the Standard Model Higgs boson in chapter 7. The method could be used to look for charged Higgses, pair production of techno-pions etc. As an example of these possibilities a search for a Higgs boson with properties different from that of the SM is described in chapter 8.

Chapter 6

NC02 (ZZ) cross section measurement

In 1997 LEP reached a centre-of-mass energy of 183 GeV, allowing for the first time on-shell Z bosons to be pair produced. In this chapter the measurement of the ZZ production cross section (through NC02 diagrams, see figure 2.8) is presented in case both Z bosons decay hadronically.

Measuring this small ($\sim$ 0.5 pb) and well predicted cross section at all centre-of-mass energies is interesting for various reasons. It allows to test the SM prediction, and, since both the cross section and the angular distribution of the produced Z bosons are sensitive to contributions from physics beyond the SM it also allows to set direct limits on parameters like for example anomalous neutral trilinear gauge couplings. Another important reason however to obtain a good estimation of the cross section is that ZZ events provide an environment that is both in terms of experimental signature[1] and in cross section similar to a possible Higgs signal. All the ingredients in the Higgs search (b-tagging, heavy boson mass reconstruction and topological variables) also play an important role in this analysis. By performing the measurement of the ZZ cross section the existing Higgs analyses can be tested for sensitivity to such a signature and important systematic effects involved in examining final states with these specific characteristics will show up in this channel as well.

In section 6.1 the ZZ signal is defined on the level of matrix elements and the characteristic (experimental) properties of its 4-quark final state are discussed in section 6.2. In section 6.3 all the information from the measured event is combined and the distribution of the ZZ-probability is shown. The performance and separation power together with the procedure used to extract the ZZ cross section from this distribution are described in section 6.4. An evaluation of the main systematic effects contributing to the uncertainty on the extracted ZZ cross section (using the full LEP2 data set) is described in section 6.4.5. Finally the measured ZZ cross sections are translated into pure NC02 cross sections and compared to the SM expectation in section 6.5.

A dedicated ZZ $\rightarrow \mathrm{b\overline{b}q\overline{q}}$ measurement is performed in section 6.6 to illustrate the flexibility of the event-by-event probability computations and because this particular final state is almost identical to a possible Higgs signal (a heavy object decaying into a $\mathrm{b\overline{b}}$ pair recoiling against a Z boson). The chapter ends with an indication of possible improvements in the analysis and a conclusion.

[1] At LEP the Higgs boson is almost always pair-produced together with a Z boson.

6.1 ZZ signal definition

In the SM the possible 4-quark final states originating from pure NC02 diagrams also receive contributions from other diagrams (non-NC02 diagrams and interferences). To measure the cross section for on-shell Z boson pair production it is necessary to define an unambiguous experimentally accessible NC02-like (ZZ) signal for generated 4-fermion final states (`EXCALIBUR`). Once this ZZ cross section is experimentally extracted it can be extrapolated to a pure NC02 cross section (see the right plot of figure 6.1) using the relationship between the predicted ZZ and NC02 (YFSZZ calculations [40]) cross sections. In the analysis ZZ (W^+W^-) events are defined using for each generated event both the full matrix element (all diagrams) and that calculated using only the NC02 (CC03) diagrams:

$$\mathcal{R}_{\mathrm{NC02}} = \frac{|\mathcal{M}_{\mathrm{NC02}}|^2}{|\mathcal{M}_{\mathrm{full}}|^2} \qquad \mathcal{R}_{\mathrm{CC03}} = \frac{|\mathcal{M}_{\mathrm{CC03}}|^2}{|\mathcal{M}_{\mathrm{full}}|^2} \tag{6.1}$$

$$\mathrm{ZZ} = \mathcal{R}_{\mathrm{NC02}} \geq 0.5 \qquad \mathrm{W^+W^-} = \mathcal{R}_{\mathrm{CC03}} \geq 0.5 \tag{6.2}$$

The distribution of $\mathcal{R}_{\mathrm{NC02}}$ for generated 4-fermion events with 4 quarks in the final state at a centre-of-mass energy of 200 GeV is shown in the left plot of figure 6.1.

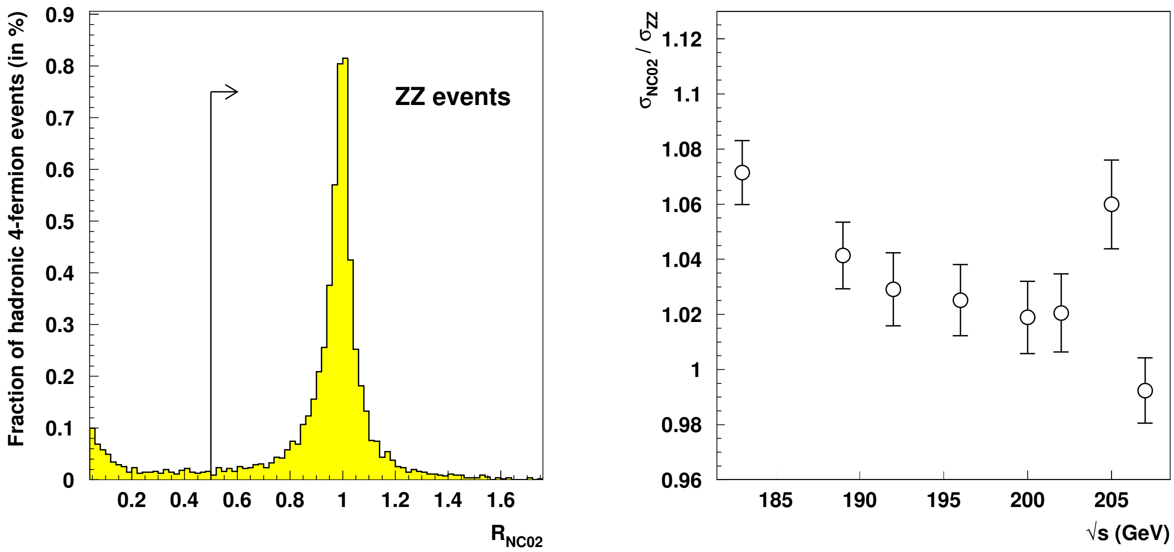


Figure 6.1: *The left plot shows the distribution of R_{NC02} for generated (*`EXCALIBUR`*) 4 quarks final states at a centre-of-mass energy of 200 GeV. The right plot shows the ratio of σ_{NC02} and σ_{ZZ}.*

Events are now classified as either ZZ, W^+W^- or, if none of the two conditions in equation 6.2 is satisfied, as Z/γ^* [2]. The pre-selection efficiencies for hadronic ZZ events are evaluated in section 4.5 for different centre-of-mass energies and are close to 90%.

[2] Due to the interference between (sets of) diagrams, a small fraction of the 4-fermion events ($\sim 5 \cdot 10^{-5}$) is simultaneously ZZ and W^+W^- according to the definition presented in equation 6.2. In this analysis these events are defined to be ZZ.

6.2 Characteristics of the ZZ final state

The ZZ fully hadronic final state has some characteristics that are quite distinct. To show their potential in separating the ZZ signal from the W^+W^- and $q\overline{q}(\gamma)$ background the two most important ones are discussed: a pair of b-quarks present in the final state and one of the jet-pairings has two reconstructed di-jet masses close to M_Z.

B-tag configurations

The decay of the Z boson has been studied extensively using the millions of Z bosons collected at LEP1. This resulted not only in a precise measurement of the hadronic decay ratio, but also of the relative branching fraction of the Z boson into quark-pairs of various flavours. In short: about half (48.9%) of the ZZ decays is predicted to result in 4 quarks in the final state and in 38.7% of those fully hadronic events there are at least 2 b-quarks in the final state (see table 5.1).

In section 5.1 it was shown that the b-tag information is quite powerful in identifying a jet from b-quark fragmentation. As an illustration of the separation power of the b-tag information the distribution of the combined b-tag *per event* (defined as the sum of the combined b-tag per jet of the two most b-tagged jets in the event) is shown in figure 6.2 for the two main backgrounds and two ZZ final states: with and without b-quarks. There is clear separation between events with and without a pair of b-quarks and in the distribution for $q\overline{q}(\gamma)$ events the contribution from b-quark pair production ($\sim$ 16%, see section 5.1.1) can be clearly identified. The agreement between the data and predictions in the full 183-207 GeV data set is shown in the right plot of figure 6.2.

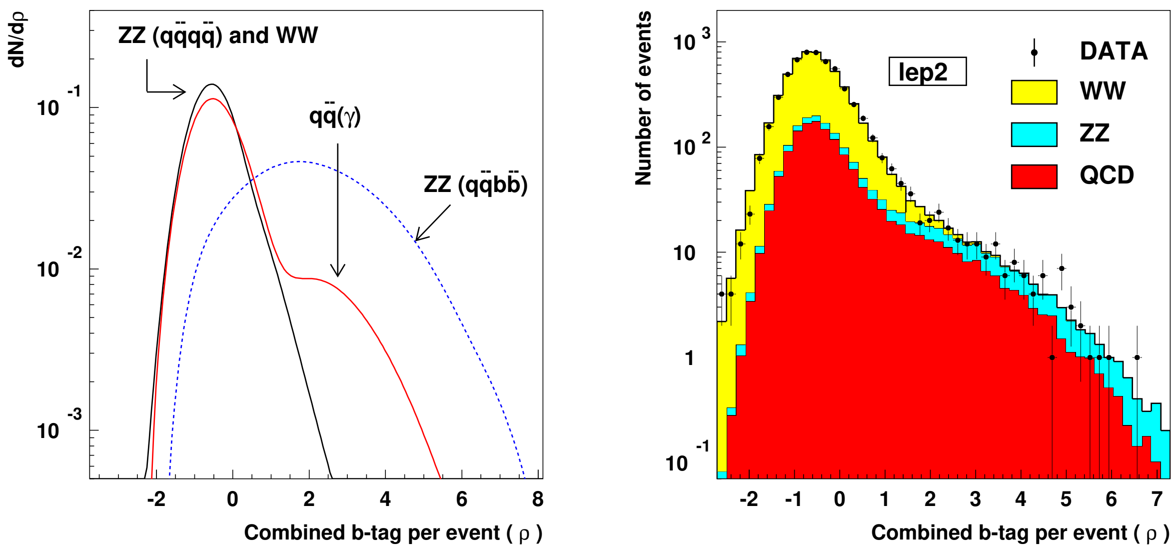


Figure 6.2: *The left plot shows the (normalised) distribution of the combined b-tag per event (sum of the combined b-tag per jet of the two most b-tagged jets in the event) for various event types. The right plot shows the comparison between data and Monte Carlo using the full 183-207 GeV data set.*

Mass distributions

The distribution of the masses of the two Z bosons in a ZZ event is given by equation (5.4), with M_Z=91.2 GeV/c^2 and Γ_Z=2.49 GeV/c^2. The method to assign a compatibility with a ZZ hypothesis for each jet-pairing has been described in section 5.2 and the separation power in identifying ZZ events using only the mass information is shown in figure 5.7. As another illustration of the separation power from the mass information alone, the procedure used by most analyses at LEP to extract mass information from processes where two heavy bosons are pair-produced is discussed: a 5C(a)-fit to define the so-called *equal mass* in the event.

5C(a)-fit: Like a 4C-fit and in addition for a given pairing both di-jet masses are equal.
In addition to requiring energy and momentum conservation (4C-fit) on the reconstructed event, an additional constraint is imposed that for a given pairing both di-jet masses are equal:

$$\sum_{i=1}^{n_{\text{jets}}} p_i^{(\text{fitted})\ \mu} = (0,0,0,\sqrt{s}) \quad \& \quad \text{for this pairing } \mathrm{M}_{ij} = \mathrm{M}_{kl} \tag{6.3}$$

Why a 5C(a)-fit: To find the equal mass (pairing) in the event
The pairing in the event with the best χ^2, once fitted with constraints given by equation (6.3) is defined as the equal mass pairing. Once this pairing is chosen, the *equal mass* is defined as the di-jet mass in that pairing after the fit. The equal mass distribution for ZZ events and the two physics background processes at a centre-of-mass energy of 200 GeV can be seen in figure 6.3. For ZZ events a Breit-Wigner fit indicates the peak position and width of the distribution.

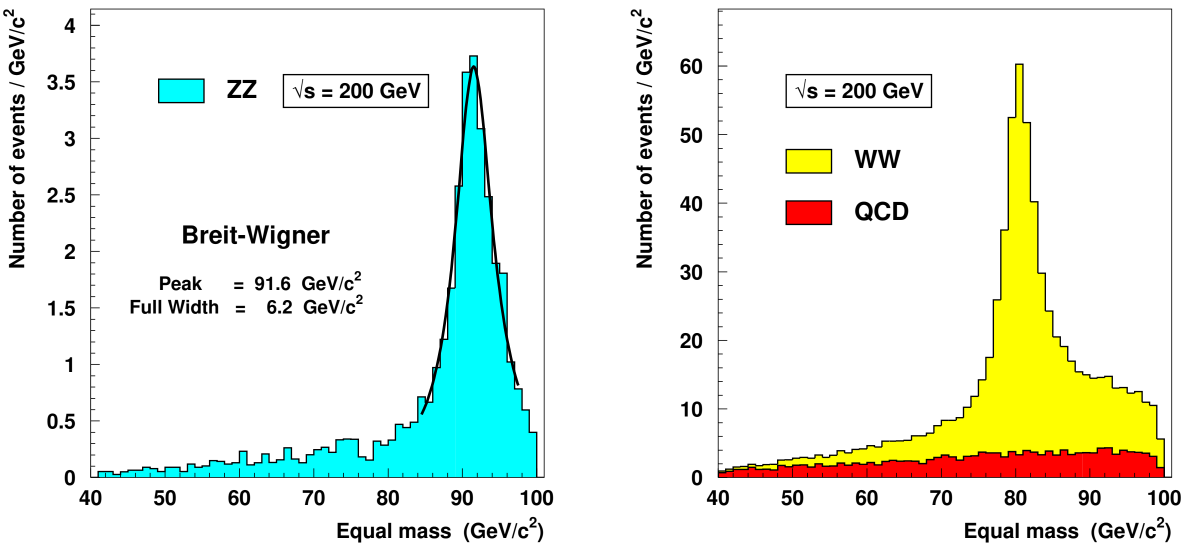


Figure 6.3: *The number of expected events for various event types as a function of the equal mass at a centre-of-mass energy of 200 GeV. The left plot shows the distribution for ZZ events and the right plot for W^+W^- and $q\overline{q}(\gamma)$ events.*

Figure 6.3 shows a few important features. The fact that the equal mass can not exceed $\frac{1}{2}\sqrt{s}$ is clear from energy conservation, but to understand the rising distribution of the $\mathrm{q\overline{q}}(\gamma)$ background requires re-investigating figure 5.5. The error contours on the reconstructed masses of the three pairings in that event illustrate the feature that when the sum of both di-jet masses in a given jet-pairing gets close to the kinematic limit the error on the sum of the masses gets much smaller than the error on the difference of the masses. This results in a bias towards pairings with large (m_1+m_2) when selecting the equal mass pairing. Combined with the available phase-space this translates into a rising distribution for the equal mass distribution from a 'flat' (m_1, m_2) spectrum.

The distribution for W^+W^- events has in addition a large tail towards higher masses[3]. This is due to the fact that the ISR is neglected in the constrained fit, resulting in an overestimation of the jet energies as explained in section 4.3.3. A fit to the central part of the distribution yields slightly higher values than the pole mass as is shown in the left plot of figure 6.3 for ZZ events (Z boson mass is 91.2 GeV/c^2). Even though the W and the Z mass are more than 10 GeV/c^2 apart there is a significant fraction of W^+W^- events with an equal mass close to M_Z.

6.3 ZZ-probability distributions

To obtain optimal sensitivity for the ZZ signal the information extracted from the measured event is compared with the SM expectation following the procedure described in section 5.4 and an event-by-event ZZ compatibility ($\mathcal{P}_{ZZ}$) is computed. Relative probabilities are obtained by normalising the compatibilities to the sum of the three SM hypotheses. In the remainder of this section the ZZ-probability is therefore defined as:

$$\mathrm{P}_{ZZ} = \frac{\mathcal{P}_{ZZ}}{\mathcal{P}_{ZZ} + \mathcal{P}_{WW} + \mathcal{P}_{QCD}} \tag{6.4}$$

The normalised distributions of $\ln(\mathrm{P}_{ZZ})$ for ZZ events and the combined background for the W^+W^- and $q\overline{q}(\gamma)$ processes at a centre-of-mass energy of 200 GeV are shown in figure 6.4.

The double peak structure in the ZZ-probability distribution for ZZ events reflects the two distinct final states (with and without a b-quarks pair). From the last two sections it is clear that the separation strength of various variables is different and this is reflected in the structure of the distribution. The real separation power is obtained when both the signal and backgrounds are scaled to equal luminosity (the expected number of background events is about 20 times larger than the number of ZZ events). In the left plot of figure 6.5, the number of expected events at a centre-of-mass energy of 200 GeV is shown as a function of P_{ZZ} for all SM processes separately.

To indicate the contribution of the ZZ events, the ZZ purity (per bin) is calculated as a function of P_{ZZ} and shown in the right plot of the figure. As can be seen, the calculated probability is almost a real purity, a property that is shared by all centre-of-mass energies. The advantage of this is clear when combining different centre-of-mass energies since distributions can be simply added without loosing sensitivity. It is this property that is used in section 6.4 to add the distributions from the 183-207 GeV data set to create a 'LEP2' ZZ-probability distribution. This allows to obtain good control of both the data/Monte Carlo agreement and a statistically more significant understanding and estimation of the systematic uncertainties.

[3] For ZZ events as well when they are produced well above threshold.

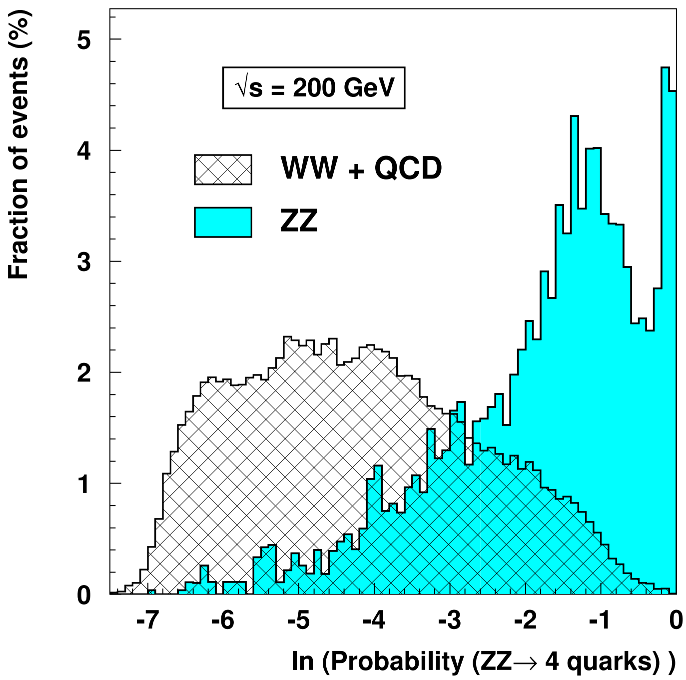


Figure 6.4: *Normalised distributions of the logarithm of the ZZ-probability for ZZ events and for the combined W^+W^- and $q\overline{q}(\gamma)$ background at a centre-of-mass energy of 200 GeV.*

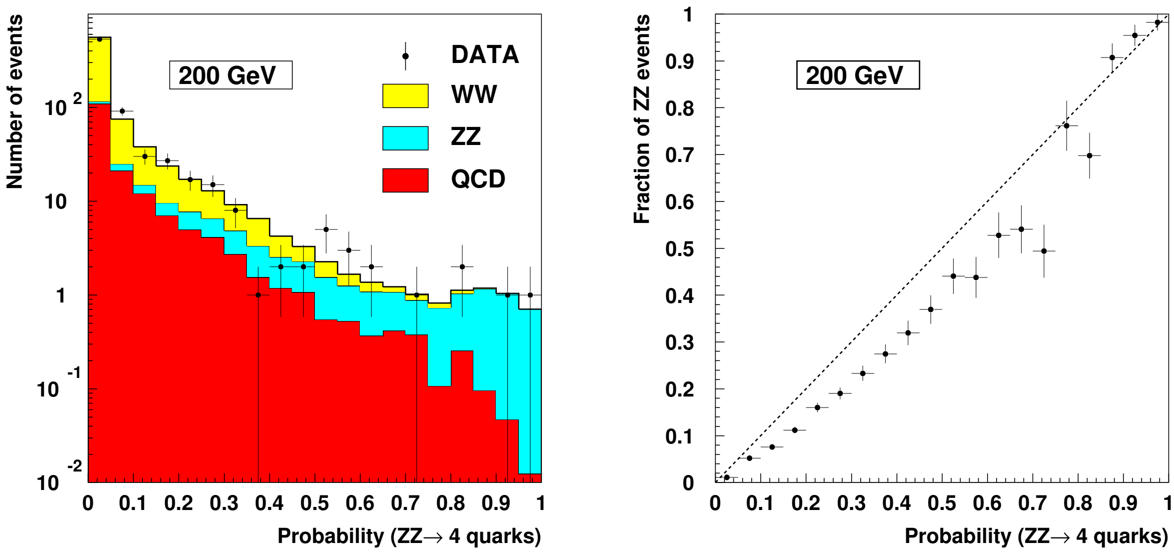


Figure 6.5: *The distribution of the ZZ-probability at a centre-of-mass energy of 200 GeV. The data are represented by the points and SM expectations by the solid histograms. In the right plot the ZZ purity (per bin) is calculated as a function of the computed ZZ-probability.*

The agreement between data and expectation over the full range of ZZ probabilities is indicated by plotting the number of events (data and that expected from the SM) at increasingly stringent levels of purity (for ZZ events with a 4-quark final state). This plot, as shown in figure 6.6, allows in addition to study the background composition at any given efficiency (purity). Note that neighbouring bins are fully correlated.

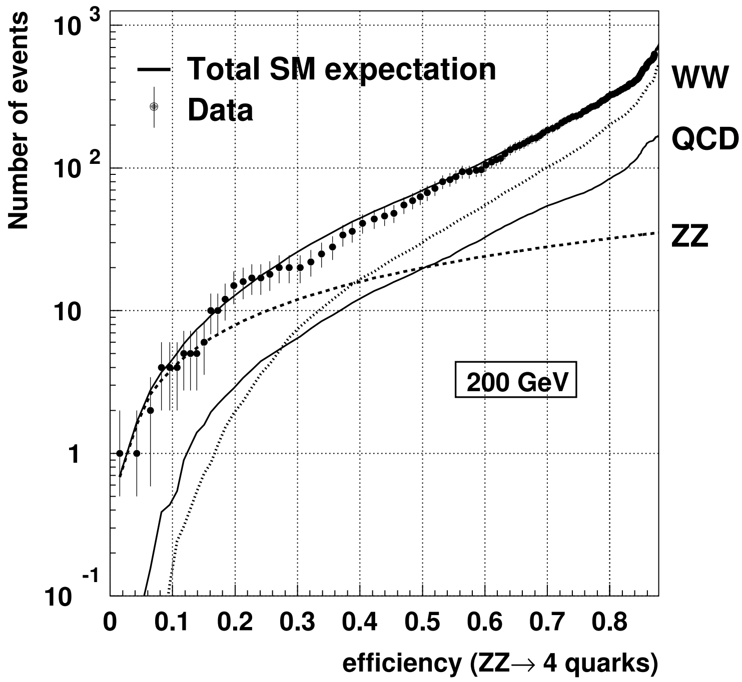


Figure 6.6: *Number of expected events left for various processes as a function of the efficiency for ZZ-events with a 4-quark final state at a centre-of-mass energy of 200 GeV. The number of events observed in the data is represented by the points.*

6.4 Extraction of the ZZ cross section

The ZZ cross section is extracted by means of a binned maximum likelihood fit to the ZZ-probability distribution, thereby exploiting the differences in the shape of the distribution for signal and background. In the fit only the contribution from ZZ events is varied.

First, the effect of the dominant systematic uncertainty on the extracted cross section (the uncertainty on the background $\mathrm{q\overline{q}}(\gamma)$ cross section) is evaluated to define the optimal region of the ZZ-probability distribution to be used in the fit. Then, the method used to extract the ZZ cross section is defined and the results are given for all centre-of-mass energies. After an evaluation of the dominant systematic errors the ZZ cross sections are finally (section 6.5) translated into a pure NC02 cross section and compared to the SM prediction.

6.4.1 Selecting the fit-region

The distribution of the purity (per bin) of ZZ events as a function of the computed ZZ-probability (figure 6.5) is identical for all centre-of-mass energies. This property allows to add the histograms of different energies (adding bins of equal purity) and construct a 'LEP2' ZZ-probability distribution as shown in figure 6.7. The increase in statistics will allow a more precise determination of the optimal region to be used for the ZZ cross section fit.

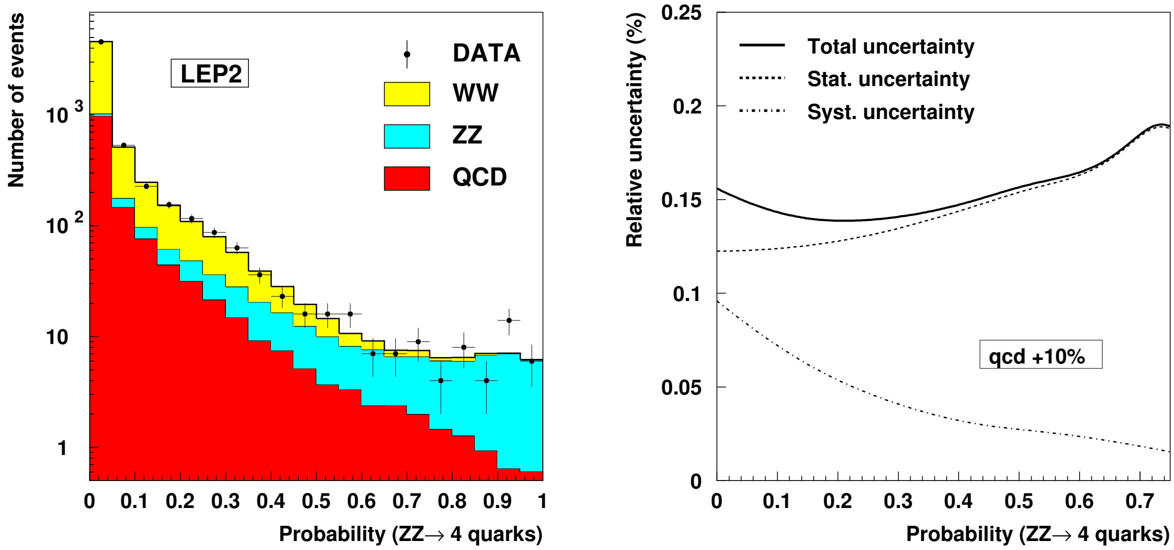


Figure 6.7: *The left plot shows the distribution of the combined ZZ-probability distribution for the full 183-207 GeV data set. The data are represented by the points and SM expectations by the solid histograms. The right plot shows the distribution of the two main relative uncertainties on the extracted ZZ cross section as a function of a cut on the ZZ-probability: the statistical uncertainty and the effect from the uncertainty on the background estimation.*

The statistical error is minimal when using the full distribution and increases when a more pure sample is used. This is expected and illustrated by the dashed line in the right plot of figure 6.7. The dominant systematic error on the extracted ZZ cross section is the uncertainty on the background estimation and this has a reverse behaviour: the effect on the extracted ZZ cross section is maximal when the full distribution is used (using also the large region where the background dominates) and decreases when going towards a more pure sample. The optimal region of the ZZ-probability distribution to use for the cross section fit is the one in which the quadratic sum of the statistical and systematic error is minimal. From figure 6.7 it is clear that the optimal region is the one above $P_{ZZ} = 0.25$ since at that point the value of the combined uncertainty is minimal.

6.4.2 Performance

The number of expected ZZ events after a cut on P_{ZZ} at 0.25 and the efficiency for hadronic ZZ events is shown in table 6.1. For most centre-of-mass energies this cut corresponds roughly to an efficiency just above 40% for the ZZ signal. As can be seen in figure 6.6 (for a centre-of-mass energy of 200 GeV), the contribution of the expected W^+W^- and $q\bar{q}(\gamma)$ events to the combined background is almost similar at that efficiency and are also comparable to the ZZ expectation.

$\sqrt{s}$	ϵ(ZZ) (4 quarks)	ZZ (4 quarks)	SM (all)	observed
183 GeV	13.7	0.95	2.14	2
189 GeV	28.3	13.98	35.51	36
192 GeV	36.3	3.53	9.98	14
196 GeV	41.1	13.78	38.66	55
200 GeV	41.7	16.73	48.53	43
202 GeV	45.2	9.02	25.76	28
205 GeV	43.2	17.07	50.30	53
207 GeV	44.7	33.40	95.65	85
total	40.1	108.46	306.53	316

Table 6.1: *The expected number of ZZ events and the full expectation from the SM above P_{ZZ} = 0.25 at various centre-of-mass energies. As an indication of the performance also the efficiency (in %) for fully hadronic ZZ events is given. The number of expected events is compared to the number of events observed in the data.*

As mentioned before, the double peak structure for ZZ events in figure 6.4 is caused by the two classes of ZZ fully hadronic final states: with a pair of b-quarks (the right most signal like peak, due to their more unique signature) and without a pair of b-quarks. Using a region above P_{ZZ} = 0.25, a large fraction of the ZZ events in the fit contain a pair of b-quarks. Their fraction is somewhere between 55 and 60% implying that there is also a non-negligible contribution from ZZ(light quarks) via the mass information.

As an illustration of the obtained sensitivity and the data/Monte Carlo agreement for a very pure sample of ZZ events, the distribution of the reconstructed equal mass is shown in figure 6.8 for events with $P_{ZZ} > 0.50$. At this level the purity of the signal is 64% to be compared with a purity of 35% for the sample selected with $P_{ZZ} > 0.25$.

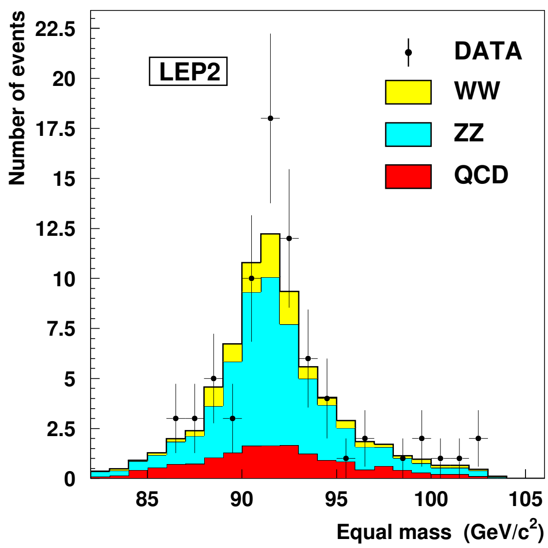


Figure 6.8: *Reconstructed equal mass distribution for ZZ-like events ($P_{ZZ} > 0.50$).*

6.4.3 Extraction of the ZZ cross section: the rules

In this section a quick overview is given of the procedure to obtain information on the ZZ cross section using the ZZ-probability distribution. As an illustration of the definitions, the results for the measurement at a centre-of-mass energy of 200 GeV are also given.

Constructing the likelihood function

When small numbers of events are expected (as is the case in the fit of the ZZ-probability distribution) the probability to observe n events when λ are expected is given by the Poisson distribution:

$$\mathrm{P(n|\lambda)} = \frac{\lambda^{\mathrm{n}}\, \mathrm{e}^{-\lambda}}{\mathrm{n!}} \tag{6.5}$$

In each bin of the distribution the probability that (n_i) events were observed while λ_i were expected can now be computed. Since the background cross sections from $\mathrm{W^+W^-}$ and $\mathrm{q\overline{q}}(\gamma)$ are fixed, the expected number of events is directly related to the ZZ cross section. In the rest of the section the number of ZZ events is varied and the cross section in the fit relative to the expected cross section from the SM is defined as ζ_{zz}:

$$\sigma_{\mathrm{zz}}^{\mathrm{in\ fit}} = \zeta_{\mathrm{zz}} \cdot \sigma_{\mathrm{zz}}^{\mathrm{SM}} \tag{6.6}$$

For the full measurement $\mathcal{M}$, given by the expected distribution of $\mathrm{P_{ZZ}}$ and the distribution of the observed events in the data, the total likelihood $\mathcal{L}$ is defined as the product of the individual likelihoods per bin:

$$\mathcal{L}_i(n_i|\zeta_{\mathrm{zz}}) = \frac{\lambda_i(\zeta_{\mathrm{zz}})^{n_i}\, e^{-\lambda_i(\zeta_{\mathrm{zz}})}}{n_i!} \qquad \text{and} \quad \mathcal{L}(\mathcal{M}|\zeta_{\mathrm{zz}}) = \Pi_i\, \mathcal{L}_i \tag{6.7}$$

In the expression for the full likelihood the product runs only over the bins of the distribution selected for the fit. By varying ζ_{zz}, a likelihood curve can be constructed that describes the relative likelihoods of this measurement for different values of ζ_{zz}. As an example, the likelihood curve (relative to the maximum likelihood: $-\Delta\ \ln\ \mathcal{L}(\mathcal{M}|\zeta_{\mathrm{zz}})$) corresponding to the measurement at a centre-of-mass energy of 200 GeV is given in the left plot of figure 6.9. The likelihood distributions for all centre-of mass energies are given in appendix A.1.

Extracting information on ζ_{zz}

The likelihood curve provides only information on the likelihood of this measurement given a certain value of ζ_{zz}. This is not what is needed, since it does not say anything about ζ_{zz} itself. To determine the ZZ cross section the exact opposite is needed: "Given the measurement $\mathcal{M}$, what inference can be made on the value of ζ_{zz} ?". This question can be answered by constructing the probability density function for ζ_{zz} using Bayes' Theorem [72]:

$$\mathrm{P}(\zeta_{\mathrm{zz}}|\mathcal{M}) = \alpha_0 \mathrm{P}(\mathcal{M}|\zeta_{\mathrm{zz}}) \cdot \mathrm{P}(\zeta_{\mathrm{zz}}) \tag{6.8}$$

In this expression $\mathrm{P}(\mathcal{M}|\zeta_{\mathrm{zz}})$ is the likelihood function from equation (6.7), α_0 is a normalisation constant with α_0^{-1} explicitly given by $\int_0^{+\infty} \mathrm{P}(\mathcal{M}|\zeta_{\mathrm{zz}})\mathrm{P}(\zeta_{\mathrm{zz}})\ \mathrm{d}\zeta_{\mathrm{zz}}$ and $\mathrm{P}(\zeta_{\mathrm{zz}})$ is the *a priori* distribution of the possible values of ζ_{zz} and is called the Bayesian prior function. Since the NC02 cross section is a quantity that is positive definite, the (normalised) prior function is chosen as:

$$\mathrm{P}(\zeta_{\mathrm{zz}}) = \begin{cases} 0. & \zeta_{\mathrm{zz}} < 0. \\ 1./\zeta_{\max} & 0. \leq \zeta_{\mathrm{zz}} \leq \zeta_{\max} \end{cases} \tag{6.9}$$

If $\zeta_{\max}$ is chosen to be large it will not introduce a bias in the determination of the extracted values for ζ_{zz}. In this definition the (normalised) likelihood function can be directly interpreted as a probability density function for ζ_{zz}. Since there is no unique procedure to summarise the measurement (given by the likelihood function) in a central value and a confidence interval the recommendations of the workshop on confidence levels [73] should be followed:

"An experiment should provide for each measurement the full likelihood function and a detailed description of the procedure used to extract the central value and the confidence intervals."

In this thesis the 'rules' are defined as follows:

- **Central value:** The median value of ζ_{zz} ($\zeta_{\mathrm{zz}}^{\mathrm{m}}$) defines the central value.

- **Upper and lower value for the CL interval:** The lower(upper) value of the CL interval is determined by the value of ζ_{zz} for which 15.87% of the integrated probability distribution is contained below(above) this value.

- **Upper limit:** If less than 31.74% of the probability density function is located below the value of ζ_{zz} for which the probability density function is maximal ($\zeta_{\mathrm{zz}}^{\max}$) an upper limit is given. The upper limit is defined as the value of ζ_{zz} for which 5% of the integrated probability distribution is contained above this value (95% CL).

Using these definitions the median value for ζ_{zz} extracted from the measurement is always above $\zeta_{\mathrm{zz}}^{\max}$. This can be understood from an interesting general feature of a Poisson distribution given a number of observed events n:

$$\begin{array}{lrcl} \text{Most likely value of } \lambda\text{:} & \lambda_{\max} & = & n \\ \text{Expectation value of } \lambda\text{:} & <\lambda> & = & n+1 \end{array}$$

In the next section the results from the cross section measurements at the different centre-of-mass energies are given.

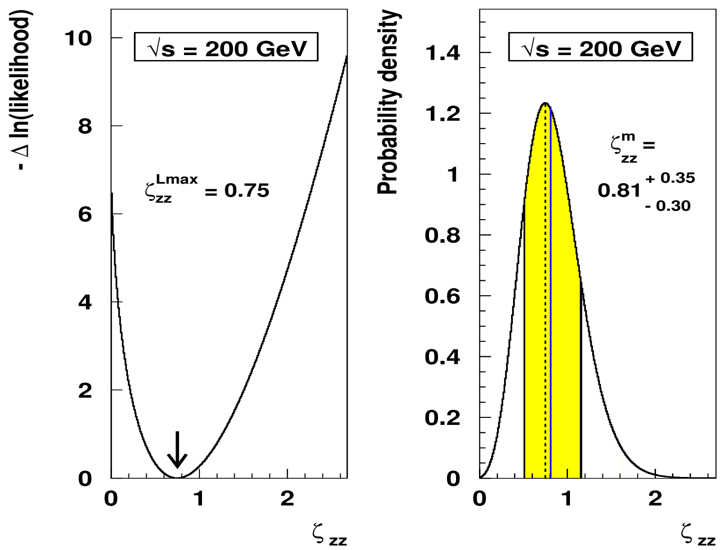


Figure 6.9: *The left plot shows the (delta)-likelihood distribution of the measurement as a function of the ZZ cross section in the fit at a centre-of-mass energy of 200 GeV. The right plot shows the corresponding probability density function of ζ_{zz}. The dotted line indicates the value of ζ_{zz} for which the likelihood was maximum and the solid line indicates the median value (ζ_{zz}^{m}). The two white areas under the distribution each represent 15.87% of the full integral.*

6.4.4 Results from the fit

The (delta) likelihood distributions as a function of ζ_{zz} are shown, together with the corresponding probability density functions for ζ_{zz} in appendix A.1 for each centre-of-mass energy. In table 6.2 the fitted values for ζ_{zz} obtained at all centre-of-mass energies in the full 183-207 GeV data set are presented. The measurement at 183 GeV allows only an upper limit to be given.

$\sqrt{s}$	ζ_{zz}^{m}	δ_-	δ_+
183 GeV	<5.29		
189 GeV	1.06	0.34	0.41
192 GeV	1.75	0.88	1.10
196 GeV	1.74	0.48	0.54
200 GeV	0.81	0.30	0.35
202 GeV	1.15	0.48	0.56
205 GeV	1.26	0.35	0.40
207 GeV	1.03	0.22	0.25

Table 6.2: *Results from the ZZ(4 quarks) cross section fit at all centre-of-mass energies in the 183-207 GeV data set. The confidence interval is given by:* $\zeta_{zz}^{m} - \delta_- \leq \zeta_{zz} \leq \zeta_{zz}^{m} + \delta_+$.

6.4.5 Systematic uncertainties

From table 6.2 it is clear that the statistical errors on the extracted ZZ cross sections at each centre-of-mass energy are large. The limited number of 4-jet events at LEP2 does not allow a precise determination of the systematic errors associated to this measurement. But since the systematic effects are not (strongly) energy dependent some of the systematic errors can be evaluated to a much better precision by using the combined statistics from the full 183-207 GeV data set.

- **background cross sections**
  The uncertainty on the $q\overline{q}(\gamma)$ 4-jet cross section is conservatively taken to be in the order of 10%. The effect of the uncertainty on the $q\overline{q}(\gamma)$ background cross section on the extracted value of the ZZ cross section has been evaluated in section 6.4.1. Varying the $q\overline{q}(\gamma)$ cross section in the fit with $\pm$ 10% results in a variation of the ZZ(NC02) cross section of $\mp$ 4.6% (as can be seen in the right plot of figure 6.7 at P_{ZZ}=0.25). A similar exercise varying the more precisely known W^+W^- cross section within $\pm$ 2% has an effect of $\mp$ 0.9%. The combined uncertainty on the extracted ZZ(NC02) cross section is therefore 4.7%

- **gluon splitting**
  The effect of the uncertainty on the gluon splitting rate into $b\overline{b}$, as explained in section 5.1.1, on the determination of the ZZ cross section is mainly caused by the increase in the $b\overline{b}$ component of the $q\overline{q}(\gamma)$ cross section. A 4.6% effect is observed when the ZZ cross section is extracted without the improved gluon splitting. Since the uncertainty on the correction to the splitting rate is estimated to be about half as large as the effect itself [70], the uncertainty on the extracted ZZ cross section is estimated to be 0.5·4.6%= 2.3%.

- **conversion factor**
  Due to the uncertainty on the theoretical prediction from YFSZZ ($\pm$ 1%) and the limited 4-fermion Monte Carlo statistics the conversion factor to translate the ZZ cross section into a pure NC02 cross section has an uncertainty in the order of 2% (depending on the centre-of-mass energy).

- **b-tag likelihood ratios**
  The parametrisation of the separation between b-jets and non-b jets as defined in section 5.1.2 is varied within the uncertainties associated to the b-tag per jet. The b/non-b likelihood ratio is modified such that for b-like jets (initial likelihood ratio $>$1) the likelihood ratio is lowered by 7.5%. The likelihood ratio for non-b like jets on the other hand is enhanced. The effect on the extracted ZZ-cross section is found to be small and in the order of 1-2%.

A conservative estimate of the total systematic uncertainty is given by the quadratic sum of the individual contributions and amounts to 6%.

6.5 NC02 cross section results

The measured ZZ cross sections can be translated into a pure NC02 cross section using the conversion factors as shown in the right plot of figure 6.1. The corrections are of the order of a few % (with a typical uncertainty of around 2%). The quoted uncertainties on the NC02 cross sections are first statistical and then systematic. In a graphical representation (figure 6.10) the measured fully hadronic NC02 cross sections are compared to the SM expectation.

$$
\begin{aligned}
\sigma_{\mathrm{NC02}}(183\ \mathrm{GeV}) &< 0.71\ \mathrm{pb} \quad (95\%\ \mathrm{CL}) \\
\sigma_{\mathrm{NC02}}(189\ \mathrm{GeV}) &= 0.35\ ^{+0.13}_{-0.11} \pm 0.02\ \mathrm{pb} \\
\sigma_{\mathrm{NC02}}(192\ \mathrm{GeV}) &= 0.69\ ^{+0.43}_{-0.34} \pm 0.04\ \mathrm{pb} \\
\sigma_{\mathrm{NC02}}(196\ \mathrm{GeV}) &= 0.78\ ^{+0.24}_{-0.22} \pm 0.05\ \mathrm{pb} \\
\sigma_{\mathrm{NC02}}(200\ \mathrm{GeV}) &= 0.39\ ^{+0.17}_{-0.15} \pm 0.02\ \mathrm{pb} \\
\sigma_{\mathrm{NC02}}(202\ \mathrm{GeV}) &= 0.57\ ^{+0.28}_{-0.24} \pm 0.03\ \mathrm{pb} \\
\sigma_{\mathrm{NC02}}(205\ \mathrm{GeV}) &= 0.62\ ^{+0.20}_{-0.17} \pm 0.04\ \mathrm{pb} \\
\sigma_{\mathrm{NC02}}(207\ \mathrm{GeV}) &= 0.54\ ^{+0.13}_{-0.12} \pm 0.03\ \mathrm{pb}
\end{aligned}
$$

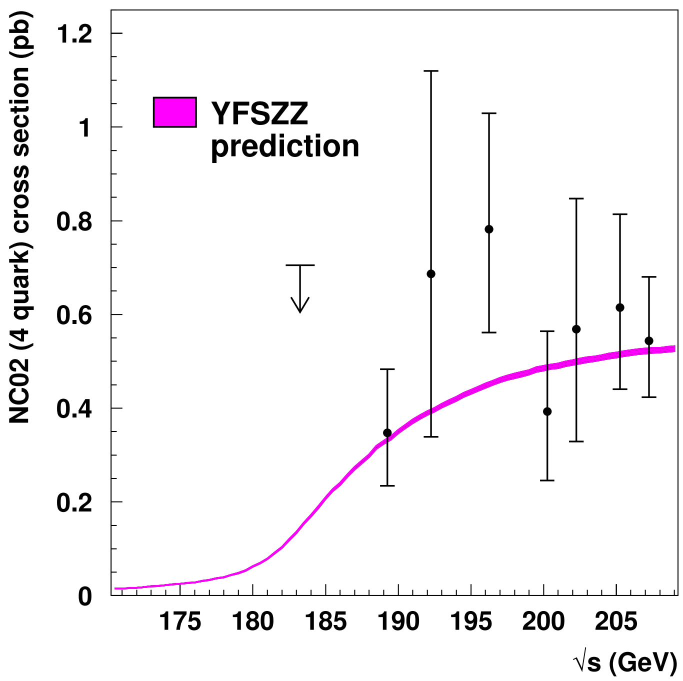


Figure 6.10: *Measured NC02 (4 quarks) cross section as a function of $\sqrt{s}$.*

6.6 ZZ($\mathrm{b\overline{b}q\overline{q}}$) cross section measurement

For events in which one of the Z bosons decays into a $\mathrm{b\overline{b}}$ quark-pair, the final state is almost identical to that of a Higgs signal: a heavy object decaying into a $\mathrm{b\overline{b}}$ pair recoiling against a Z boson. The probability computation for this process will be performed following the procedure defined in section 5.4. The event topology and mass information is identical to that of the full ZZ process; the only difference being the possible final state b-tag configurations and the cross section (38.7% of the fully hadronic ZZ cross section). Investigating this process requires more emphasis on the b-content of the event where the use of the b-tag likelihood ratios (section 5.1) incorporates the knowledge on the fraction of light-quark jets that are reconstructed with a high value of the combined b-tag probability per jet.

The double peak structure for ZZ events (with and without b-quarks in the final state), is also present in the ZZ($\mathrm{b\overline{b}q\overline{q}}$) probability distribution. In this case the peaks are further separated as can be seen in figure 6.11: the normalised distributions of ln($\mathrm{P_{ZZ}(b\overline{b}q\overline{q})}$) for ZZ($\mathrm{b\overline{b}q\overline{q}}$) events, for the background from ZZ(light quarks) and the combined background for the $\mathrm{W^+W^-}$ and $\mathrm{q\overline{q}}(\gamma)$ processes at a centre-of-mass energy of 200 GeV.

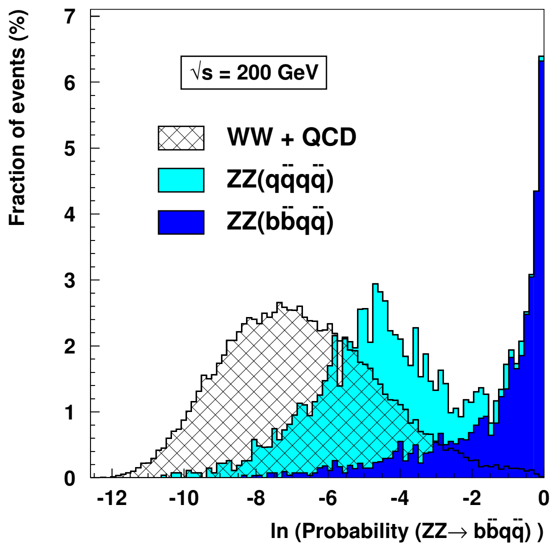


Figure 6.11: *Normalised distributions of the logarithm of the ZZ-probability for ZZ($b\overline{b}q\overline{q}$) events, ZZ(light quarks) and combined background separately at a centre-of-mass energy of 200 GeV.*

The distribution of the ZZ($\mathrm{b\overline{b}q\overline{q}}$)-probability for each of the centre-of-mass energies has the property that in each bin of $\mathrm{P_{ZZ}(b\overline{b}q\overline{q})}$, the probability is equal to the fraction of ZZ($\mathrm{b\overline{b}q\overline{q}}$) in that bin. The agreement between data and expectation is shown in the left plot of figure 6.12 where the probability distributions from all energies in the 183-207 GeV data set have been combined (as was done for the full ZZ cross section). A good agreement between the observed and expected number of events is observed as is the case in the cumulative distribution (right plot of figure 6.12). In that plot the numbers of events (data and expected from the SM) at decreasing efficiencies for ZZ($\mathrm{b\overline{b}q\overline{q}}$) are shown, allowing in addition the evalution of the background composition at any given efficiency (purity). In both plots the ZZ($\mathrm{b\overline{b}q\overline{q}}$) signal is not given separately, but only the full ZZ expectation is given. In the right plot the efficiencies are explicitly correct for the ZZ($\mathrm{b\overline{b}q\overline{q}}$) signal only. As expected, the most important background is $\mathrm{q\overline{q}}(\gamma)$ (with b-quarks) while the $\mathrm{W^+W^-}$ background is quickly eliminated.

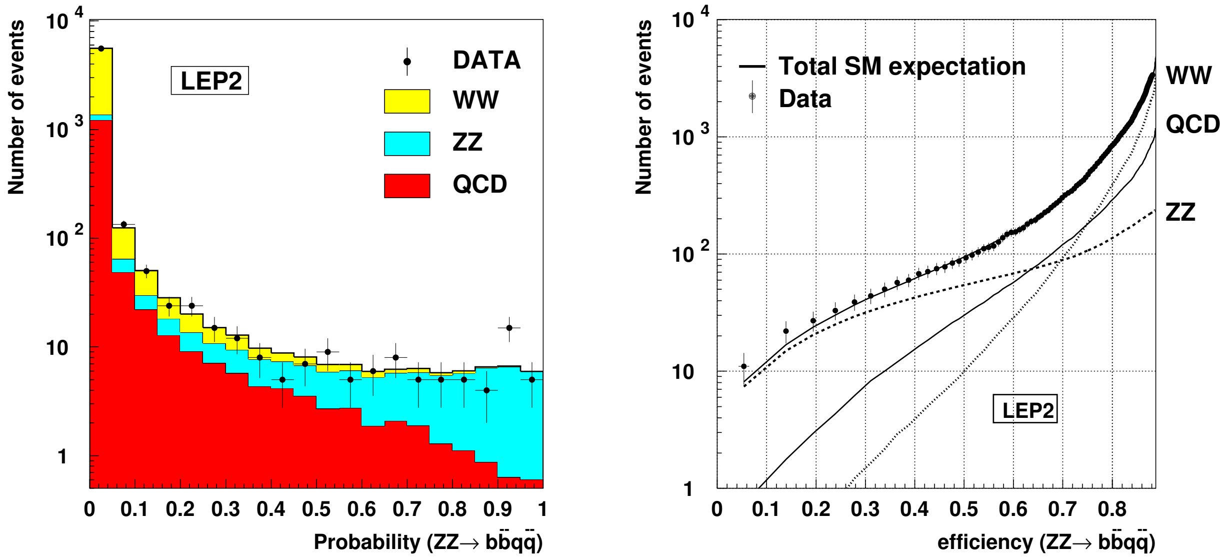


Figure 6.12: *The left plot shows the distribution of the ZZ($b\overline{b}q\overline{q}$) probability for the full 183-207 GeV data set. The data are represented by the points and SM expectations by the solid histograms. The right plot shows the number of expected events left (data and SM expectation) as a function of the efficiency for selecting ZZ($b\overline{b}q\overline{q}$) events.*

6.6.1 Performance and fit results

To extract the ZZ($\mathrm{b\overline{b}q\overline{q}}$) cross section, a procedure is used that is similar to the one used for the full ZZ cross section. In the fit the region above $\mathrm{P_{ZZ}}(\mathrm{b\overline{b}q\overline{q}})$ = 0.20 is used. The number of ZZ($\mathrm{b\overline{b}q\overline{q}}$) events and the total number of expected events from the SM at that point are given in table 6.3, where they are compared to the number of observed events. The region $\mathrm{P_{ZZ}}$ ($\mathrm{b\overline{b}q\overline{q}}$) $>$ 0.20 corresponds to an efficiency of almost 60% (for ZZ($\mathrm{b\overline{b}q\overline{q}}$)) and the fraction of ZZ($\mathrm{b\overline{b}q\overline{q}}$) compared to the total number of ZZ events is well above 90%. The (delta) likelihood distributions as a function of ζ_{zz} [4] are shown, together with the corresponding probability density functions for ζ_{zz} in appendix A.2 for each centre-of-mass energy. As a summary, the expectation values for ζ_{zz} obtained at all centre-of-mass energies in the full 183-207 GeV data set are presented in table 6.4.

These results are transformed into a pure NC02 cross section using for each centre-of-mass energy the conversion factors from figure 6.1. The extracted cross sections are compared to the SM expectation in figure 6.13, where a good agreement between the SM prediction and measurement is observed. Note that the measurements of the cross section for the full ZZ and ZZ($\mathrm{b\overline{b}q\overline{q}}$) are strongly correlated.

[4] Meaning: $\zeta_{\mathrm{b\overline{b}q\overline{q}}}$, since the contribution from ZZ($\mathrm{b\overline{b}q\overline{q}}$) events is varied.

$\sqrt{s}$	ϵ(ZZ) ($\mathrm{b\bar{b}q\bar{q}}$)	ZZ ($\mathrm{b\bar{b}q\bar{q}}$)	SM (all)	observed
183 GeV	37.0	0.96	2.55	2
189 GeV	51.3	9.81	22.24	24
192 GeV	59.4	2.28	5.43	6
196 GeV	56.7	7.40	16.95	18
200 GeV	58.0	9.20	20.80	26
202 GeV	59.4	4.58	10.65	8
205 GeV	59.0	8.66	20.46	17
207 GeV	60.0	17.50	38.89	37
total	57.7	60.39	137.98	138

Table 6.3: *The expected number of events above P_{ZZ} ($b\bar{b}q\bar{q}$) = 0.20 at all centre-of-mass energies. The total number of expected events is compared to the observed number of events in the data.*

Fit results

$\sqrt{s}$	ζ^{m}_{zz}	δ_-	δ_+
183 GeV	<4.83		
189 GeV	1.18	0.39	0.47
192 GeV	<3.12		
196 GeV	0.92	0.41	0.53
200 GeV	1.26	0.43	0.52
202 GeV	0.86	0.44	0.59
205 GeV	1.14	0.38	0.45
207 GeV	1.18	0.28	0.32

Table 6.4: *Results from the ZZ($b\bar{b}q\bar{q}$) cross section fit at all centre-of-mass energies in the 183-207 GeV data set*

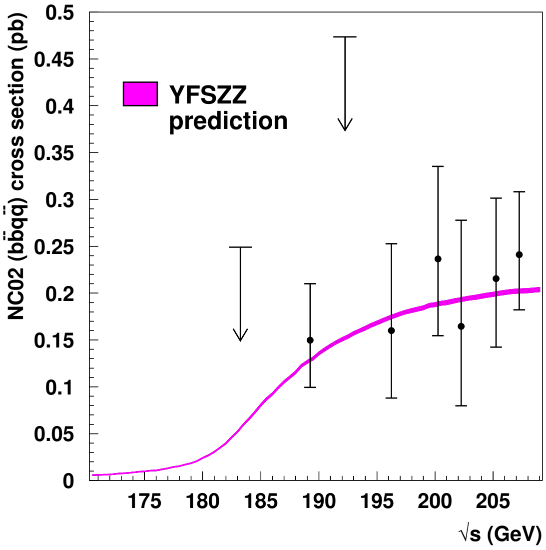


Figure 6.13: *Measured NC02($b\bar{b}q\bar{q}$) cross section as a function of $\sqrt{s}$.*

Looking at the errors associated with these measurements one can argue that with these large statistical errors a possible SM Higgs signal (which has a similar cross section and efficiency) can never be clearly identified. One should realise that in 'measuring' the Higgs cross section the full luminosity (all centre-of-mass energies) is used. As an example, one can interpret the ZZ($\mathrm{b\bar{b}q\bar{q}}$) measurements as the 'search' for on-shell Z boson pair production where one Z boson always decays into a $\mathrm{b\bar{b}}$-pair. Using the data in the full 183-207 GeV data set a 'discovery' can be claimed:

$$\sigma^{\mathrm{measured}}_{\mathrm{NC02}}\ (\mathrm{b\bar{b}q\bar{q}})/\sigma^{\mathrm{SM}}_{\mathrm{NC02}}\ (\mathrm{b\bar{b}q\bar{q}})\ =\ 1.05\ ^{+0.17}_{-0.15}(\mathrm{stat}) \tag{6.10}$$

6.7 Possible further improvements

The method of computing event-by-event probabilities allows additional information extracted from the event (in the form of likelihood ratios) to be easily included in the ZZ-probability computation. The main improvements that can be made to enhance the sensitivity of the analysis to the ZZ signal are the following:

- **Improved rejection of $\mathrm{q\overline{q}}(\gamma)$ events:**
  The pre-selection efficiency for $\mathrm{q\overline{q}}(\gamma)$ events is at the level of 2%, but due to its $\mathrm{b\overline{b}}$ component this is the dominant background for both the ZZ and ZZ($\mathrm{b\overline{b}q\overline{q}}$) cross section measurements. An improved separation between 4-fermion and 2-fermion events could be achieved by using a more sophisticated (topological) separating variable than $\mathrm{D_{pur}}$ (see equation (4.12)), the variable used in this analysis.

- **(Charged) angular distributions of the bosons:**
  Another improvement, at the level of jet-pairings, could be achieved by using the predicted angular distribution (the NC02 diagrams have only t-channel exchange of an electron) of the produced Z bosons in ZZ final states. The (differential) cross section results from this analysis have been used [74, 75] to put limits on neutral TGC's (through possible s-channel contributions), but could be used directly in the ZZ-probability distribution as well.

  An improvement connected to this is the use of jet-charge information (a momentum-weighted sum of the charge of all particles in a jet) to obtain an estimate of the charge of the heavy object defined by the jet-pairing. This would allow pairings where the bosons have both non-zero charge (like $\mathrm{W^+W^-}$ events in which the difference in charge is 2) to be disfavoured compared to a jet-pairing in which both are compatible with being neutral.

- **B-(flavour) tagging:**
  Using improved (flavour) b-tag information, jets originating from b-jet fragmentation can be separated from $\mathrm{\overline{b}}$-quark fragmentation jets. This information can help in solving ambiguities on the level of jet-pairings, especially in events with 4 b-quarks ($\sim$ 5% of the fully hadronic ZZ cross section), since $(\mathrm{b\overline{b}b\overline{b}})$ pairings can then be favoured over combinations that resemble situations not allowed in the SM: $(\mathrm{\overline{b}\overline{b}bb})$.

6.8 Combined results, summary and conclusion

The results from the DELPHI experiment on the extracted NC02 cross sections (in case both Z bosons decay hadronically) have been presented. The cross section results are in good agreement with the SM prediction as is shown in figure 6.10.

The DELPHI combined results on the NC02 cross section as a function of the centre-of-mass energy, using all ZZ decay modes can be found in the DELPHI reports on the NC02 cross section measurements [76, 77, 78]. In the combination, the weight of the three most sensitive channels to the final precision are the following: $\mathrm{q\overline{q}q\overline{q}}$($\sim$43%), $\mathrm{q\overline{q}}\nu\overline{\nu}$($\sim$35%) and $\mathrm{q\overline{q}}l^+l^-$($\sim$22%). The weight of the last two channels is enhanced (compared to their fraction of the full NC02 cross section) as expected, since $\mathrm{q\overline{q}}\nu\overline{\nu}$ is 28% of the cross section, but has no pairing ambiguity and

$q\overline{q}l^+l^-$ is only 9% of the full cross section (l= μ or e), but there is no pairing ambiguity and in addition there is no significant background.

Also the results from the 4 LEP experiments are combined [79]. The combined measured NC02 cross sections for all centre-of-mass energies at LEP2 are shown in figure 6.14. A good agreement with the SM prediction is observed.

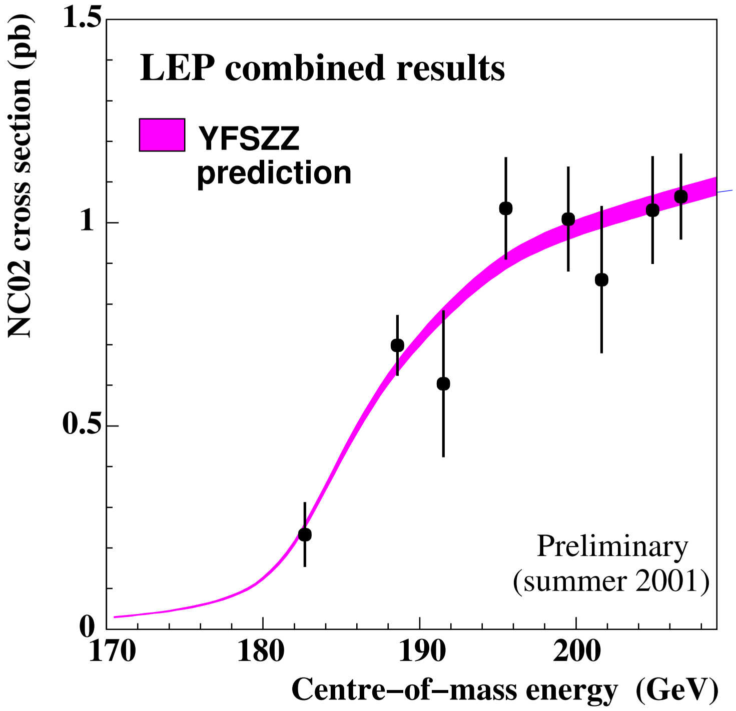


Figure 6.14: *Measured LEP combined NC02 cross section as a function of* $\sqrt{s}$*.*

Finally, the results on the combined $ZZ(b\overline{b}q\overline{q})$ cross section measurement performed in this thesis show that a Higgs boson, if the cross section is comparable to the $ZZ(b\overline{b}q\overline{q})$ cross section, will not escape detection.

Chapter 7

Higgs production (SM Higgs decay)

In this chapter a search for the SM Higgs boson is presented in case both the Higgs and the Z boson decay into a quark anti-quark pair. The analysis is an extended version of the event-by-event probability computation as introduced in section 5.4 and the chapter starts with a discussion of these various extensions and their implications on the probability computation. In section 7.2 two techniques used to extract mass information from the event are discussed in detail. This is done to illustrate the advantage of using the SM prediction for the Higgs boson's (changing) properties as a function of its mass.

The Higgs probability distributions and their characteristics are presented in section 7.3, while in section 7.4 the statistical procedure to define the sensitivity limit of the analysis is introduced in quite some detail. This procedure is then used in section 7.5 to summarise the measurement as a lower limit on the mass of the SM Higgs boson. In that section also an evaluation of the systematic uncertainties and possible improvements to the analysis are presented. The combined result from the LEP experiments is discussed and in the last two sections the search for the Higgs boson in the near future is quickly reviewed before ending with a summary and conclusion.

7.1 Extending the probability computation

Although the mass of the Higgs boson is not predicted within the SM, the SM does predict the properties of the Higgs boson as a function of its mass. As described in section 2.4 for a given Higgs mass not only the branching fraction into a pair of b-quarks is precisely predicted, but also the production cross section is known once a specific centre-of-mass energy is defined. Using the generated Monte Carlo samples, also the hadronic pre-selection efficiencies for the event selection described in section 4.5.2 are known at each centre-of-mass energy as a function of the Higgs boson mass. For Higgs masses close to the kinematic limit these are around 90% as can be seen in figure 4.13 for a centre-of-mass energy of 200 GeV. All properties of the Higgs signal needed as inputs in the expression to compute the compatibility with a Higgs signal, equation (5.15) are therefore known.

In this analysis these dependencies of final state characteristics on M_{H} are taken into account, analytically where possible. This results in an optimal sensitivity to a possible signal by incorporating in a consistent way the known (varying) differences between the ZH signal and SM background for a varying Higgs mass hypothesis.

An extension of the existing probability computation, compared to that used to measure the NC02 cross section (chapter 6), is required since in ZH production there are two bosons with different characteristics in the final state. The ambiguity that is present in a 4-jet ZZ or W^+W^- event, due to the possible jet-pairings, now doubles since in each pairing there is an additional ambiguity: each di-jet can be either the Higgs or the Z boson. Before discussing the performance of the analysis in separating the Higgs signal from the SM background in section 7.3, first the use of mass information is discussed.

7.2 Extracting mass information

The method to assign a compatibility for a pair of di-jets with a Higgs hypothesis using only mass information is described in section 7.2.2 and is a natural extension of the ideogram method described in section 5.2.3. To introduce some of the difficulties connected to mass reconstruction in these final states and to justify the choice for using the ideogram method, first the 'standard' method to extract the Higgs mass information per event (a single Higgs mass estimator) is presented in section 7.2.1.

7.2.1 Fixing the Z mass to 91.2 GeV/c^2

In most analyses the procedure to find *the* Higgs mass in the event is the following: first decide on the correct jet-pairing[1] (and assignment) representing the H and the Z (often the pairing with the two most b-quark like jets is defined as the Higgs) and then use the knowledge of the Z mass to improve the resolution on the reconstructed Higgs mass. In this section some of the properties and features of the last step in this procedure are discussed. To eliminate the effect of the first step in this procedure a Monte Carlo study (without ISR radiation) is performed in the (ideal) case in which both the correct pairing and assignment are known. The results on the reconstructed Higgs masses using this procedure are compared to those where the Z mass is left free.

The effect of the constrained fit (requiring the reconstructed event to satisfy energy and momentum conservation as described in section 4.3) on the estimation of the Z boson mass is shown in figure 7.1. The ZH events used to make this distribution were generated at a centre-of-mass energy of 200 GeV for a Higgs mass of 85 GeV/c^2. After this 4C-constrained fit, the starting point of each analysis, not only is the mean value of the reconstructed Z mass close to M_Z (91.2 GeV/c^2), but also the resolution is significantly improved compared to the one using directly the reconstructed jets. At this stage the resolution on the Z mass is however still much larger than the intrinsic width of the Z (Γ_Z=2.49 GeV/c^2).

The (standard) procedure to improve the precision on the estimation of the Higgs mass is to fix the Z boson mass to its pole mass of 91.2 GeV/c^2 and fit the event again. This type of fit is called a **5C(b)-fit**, where in addition to requiring energy and momentum conservation (4C-fit), there is

[1] Note that these problems are only present in the 4 quark and (part of) the $q\overline{q}\tau^+\tau^-$ final states.

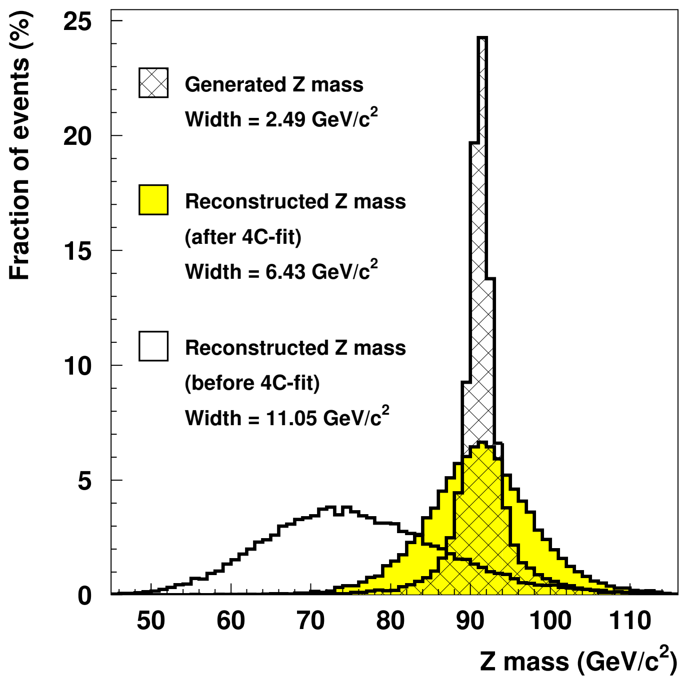


Figure 7.1: *The reconstructed Z boson mass (before and after the 4C-constrained fit) compared to the generated Z boson mass in ZH events in the ideal case that the correct jet-pairing is known. Events were generated at a centre-of-mass energy of 200 GeV and the Higgs mass was 85 GeV/c^2. The widths for the reconstructed Z mass distributions are defined as the σ's from a Gaussian fit.*

an additional constraint: for a given pairing one of the di-jet masses is known and equal to $\mathrm{M_{hyp}}$ (in this case 91.2 GeV/c^2).

$$\sum_{i=1}^{n_{\rm jets}} p_i^{(\rm fitted)\ \mu} = (0,0,0,\sqrt{s}) \quad \& \quad \text{for this pairing } \mathrm{M}_{ij} = \mathrm{M_Z} \tag{7.1}$$

The 5C(b)-fit is an ideal tool to estimate and improve the resolution on the unknown Higgs mass in case the Z boson mass in the event is described by a Breit-Wigner with a pole mass of 91.2 GeV/c^2 (hatched distribution in figure 7.1). For a wide range of (light) Higgs masses this is indeed the case, but when the sum of the Higgs mass and the Z mass gets close to the kinematic limit, the requirement of a Z mass distribution according to a Breit-Wigner is no longer fulfilled. The distribution of the mass of the Z boson in the event is truncated at $\mathrm{M_Z}$=$\sqrt{s}-\mathrm{M_H}$ as is illustrated in the left plot of figure 7.2. In that plot the (normalised) Z mass distribution for two different values of the Higgs mass is shown at a centre-of-mass energy of 200 GeV.

The effect of the varying shape of the Z mass distribution on the reconstructed Higgs mass using the 5C(b)-fit is shown in the right plot of figure 7.2. For Higgs masses (much) smaller than $\sqrt{s}$−91.2 GeV/c^2, fixing the Z mass to its pole mass results indeed in an increased precision in the estimation of the Higgs boson mass. For a Higgs mass close to (or above) $\sqrt{s}$−91.2 GeV/c^2 however, this procedure not only results in a bias on the reconstructed Higgs mass (the sum of the Z and the Higgs mass can not exceed $\sqrt{s}$), but also the reconstructed mass distribution is very asymmetric and its width can no longer be interpreted as the resolution on the true Higgs mass.

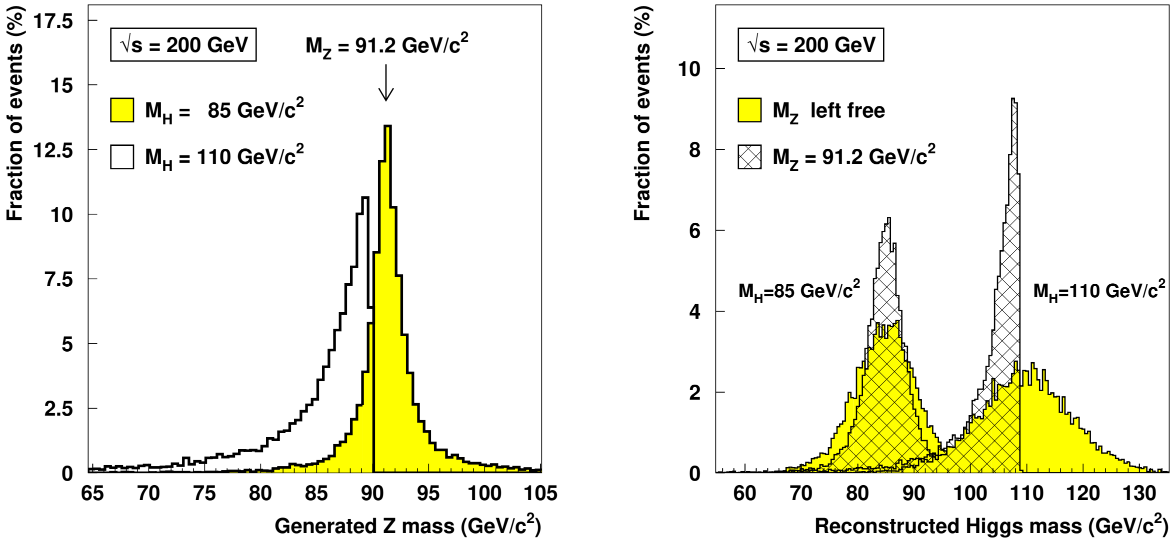


Figure 7.2: *The left plot shows the (normalised) mass distribution for the generated Z boson when pair-produced with a Higgs boson for two different Higgs masses. The right plot shows the reconstructed Higgs boson mass (for the same two Higgs masses) using two methods: the shaded histogram represents the method when the Z mass is not used at all and the hatched histogram the one where the Z mass is fixed to to 91.2 GeV/c^2.*

To follow the effect of this kinematic feature as a function of the true Higgs mass, the (mean) reconstructed Higgs mass is shown as a function of the true Higgs mass in the left plot of figure 7.3. When the Z mass is left free there is no bias on the reconstructed Higgs mass over the full range of true Higgs masses as is indicated by the shaded band. Using the fixed Z mass as constraint (represented by the hatched band), there is an increased bias in the mean of the reconstructed Higgs mass distribution as the true Higgs mass increases. For Higgs masses above $\sqrt{s}-91.2$ GeV/c^2 the bulk of the reconstructed masses remains very close to the kinematic limit (illustrated by the dashed line representing the most likely reconstructed Higgs mass), but the reconstructed mass distribution has a more prominent tail[2] resulting in an apparent 'reflection' off the kinematic limit for the mean reconstructed mass.

The resolution on a reconstructed Higgs mass versus the resolution on the true Higgs mass

The resolution on the reconstructed Higgs mass for a specific Higgs mass (distributions as in the right plot of figure 7.2) should not be confused with the resolution on the true Higgs mass given a specific reconstructed Higgs mass. These two numbers are identical only for light Higgs masses. When the sum of the Z mass and the Higgs mass is close to the kinematic limit, the uncertainty on the sum(difference) of the Z mass and Higgs mass becomes very small(large) as was shown already in section 5.2.2. Fixing the Z mass in these events will result in an artificially (too) small

[2] Reflecting the flattening (Breit-Wigner) tail of the Z boson mass distribution.

error estimate on the reconstructed Higgs mass. The uncertainty becomes (much) smaller than Γ_Z and can no longer be interpreted as the resolution on the true Higgs mass.

An illustration of this feature is shown in the right plot of figure 7.3, where the probability density function[3] for the true Higgs mass is shown for reconstructed Higgs masses between 103.0 and 104.0 GeV/c^2. Even for this intermediate Higgs mass the distribution already has a non-Gaussian shape. These effects will only increase for higher reconstructed masses as can also be understood from the left plot of figure 7.3, where the dashed line shows that for any Higgs mass above $\sqrt{s}$−91.2 GeV/c^2 the most likely reconstructed Higgs mass will always be close to the kinematic limit.

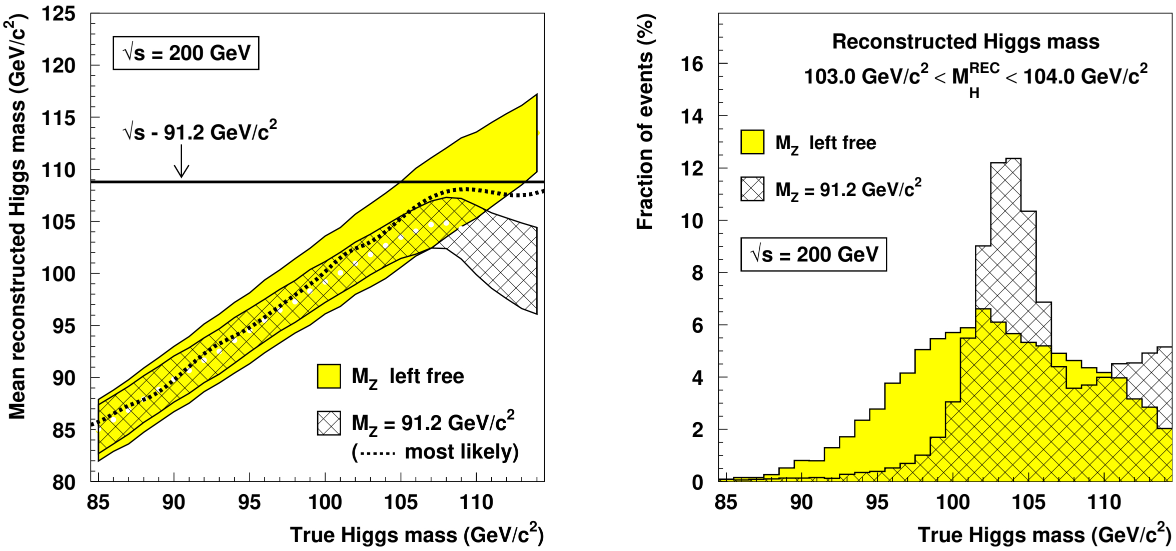


Figure 7.3: *The left plot shows the mean reconstructed Higgs mass as a function of the true Higgs mass using two different methods (the width of the bands corresponds to the RMS of the reconstructed mass distribution). The shaded band represents the method where the Z mass is left free and the hatched band represents the procedure where the Z mass is fixed to 91.2 GeV/c^2 (the most likely reconstructed Higgs mass in this last procedure is indicated by the dashed line. The right plot shows the probability density function for the true Higgs mass (assuming M_H <115 GeV/c^2) when the reconstructed Higgs mass is close to 103.5 GeV/c^2.*

When an excess in the reconstructed Higgs mass spectrum is observed close to the kinematic limit it is not the position of the peak, but the number of events in the peak that is most suitable to distinguish between various Higgs mass hypotheses since the relative ZH production cross sections for a Higgs boson with a mass of 110-112.5-115 GeV/c^2 (at a centre-of-mass energy of 206.7 GeV) are roughly 3:2:1.

7.2.2 Using Ideograms

To account for the varying expected mass distribution in this analysis the (ideogram) procedure defined in section 5.2.3 is used. This allows to compute per pairing (and assignment) a com-

[3] Assuming the Higgs mass is below 115 GeV/c^2.

patibility with a specific Higgs mass hypothesis since instead of fixing the Z mass to its pole the (analytically known) distribution of the Z boson mass in ZH events is used in combination with the event-by-event experimental errors on the reconstructed masses. These assignment-compatibilities are the ideal way to take into account the increased ambiguity connected to the mass information. Since also the other event variables are predicted given a specific pairing and assignment, the mass information can easily be incorporated in the probability computation allowing a consistent compatibility of the event with various hypotheses to be constructed.

In ZH events (for a Higgs mass $\mathrm{M_H}$) the probability density function of the masses of the two bosons can be described by the (normalised) product of a Breit-Wigner for the Z boson and a delta function for the Higgs boson[4]. Adding a phase space factor the distribution is as in equation (5.4):

$$P^{\mathrm{mass}}_{\mathrm{theory}}(m_1, m_2 \mid \mathrm{hyp}) \;\propto\; BW(m_1 \mid m_{01}, \Gamma_{01})\; \delta(m_2 - \mathrm{M_H})\; PS(m_1, m_2, s), \tag{7.2}$$

To illustrate the separation power between ZH events and events originating from SM background processes, the distributions of event probabilities using *only* mass information are shown for the different event types in figure 7.4 for a Higgs mass of 90 GeV/c^2 (left plot) and 110 GeV/c^2 (right plot). These event 'mass-only' probabilities $P^{\mathrm{mass\ (event)}}_{\mathrm{ZH(M_H)}}$ are defined as in equation 5.9 at the end of section 5.2.3, where similar definitions were used to study the separation power between ZZ, $\mathrm{q\bar{q}}(\gamma)$ and $\mathrm{W^+W^-}$ events (see figures 5.6 and 5.7). Since ZZ events have a similar mass distribution as ZH(90) events, the event probability can almost never exceed 0.5.

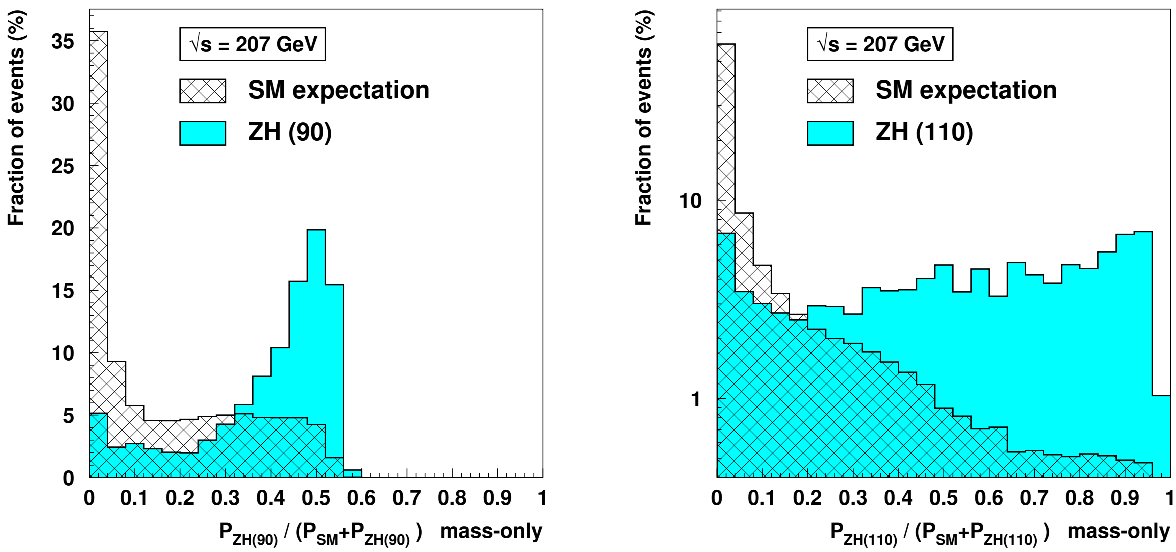


Figure 7.4: *The normalised distributions of the relative probability per event to be compatible with a ZH event when using only mass information for SM background processes (hatched histogram) and signal (solid histogram). The left(right) plot is for a 90 (110) GeV/c^2 Higgs mass hypothesis.*

[4] At masses below 120 GeV/c^2 the width of the Higgs boson is below 5 MeV/c^2 as can be seen in figure 2.6 and can be approximated by a delta function.

7.3 Probability distributions: characteristics & performance

Using the mass information as described in section 7.2.2 the event-by-event compatibility with a hypothesis of a Higgs boson with a specific mass can now be computed using the procedure described in section 5.4. Here, the characteristics and performance in the separation between signal and background are discussed.

In figure 7.5 the normalised distributions of the logarithm of the event-by-event probability are shown for a Higgs mass hypothesis of 90 GeV/c^2 (left plot) and 110 GeV/c^2 (right plot) for ZH events and for those originating from SM background processes. The additional distribution of the ZH(110)-probability for ZH(90) events (dashed line in the right plot) illustrates the fact that indeed the ZH(110)-probability computation is optimised for ZH(110) events, since the separation between the possible signal events and those originating from the SM background processes is clearly deteriorated.

As in the case of the ZZ analysis (see figure 6.6), for a given Higgs mass hypothesis, the background composition changes when going to lower efficiencies. In examining ZH events however, the dominant background (at a fixed efficiency) depends also strongly on the mass of the Higgs boson in the hypothesis as is illustrated in the two plots of figure 7.6. The agreement between observed and expected number of events is examined over a large range of ZH efficiencies by plotting the number of events (data and that expected from the SM with and without the hypothetical ZH signal) as a function of the corresponding ZH efficiency. These evolution plots are shown for a Higgs boson of 90 GeV/c^2 (left plot) and 110 GeV/c^2 (right plot). As expected, for a Higgs mass close to the Z mass the dominant background comes from ZZ events whereas for a Higgs mass of 110 GeV/c^2 the main background is from $\mathrm{q\overline{q}}(\gamma)$ events.

For both the 90 GeV/c^2 and 110 GeV/c^2 Higgs hypothesis there is a good agreement between the number of observed events and the number of predicted events when there is no additional Higgs signal assumed to be present. As an indication of the separation power at the various Higgs hypotheses and the agreement between the number of expected and observed events (in the year 2000 data set) a few characteristic numbers are given in table 7.1 for a range of Higgs masses between 90 and 115 GeV/c^2 at the point where the product of efficiency and purity is maximal.

From table 7.1 it is clear that the observed data in the 205-207 GeV data set are best explained by the SM predictions without an additional Higgs signal over the full range of Higgs boson masses, although for a Higgs signal at 115 GeV/c^2 there is not enough separation power to reject the hypothesis that there was actually a signal present. The next section is devoted to describing the method used by all the LEP experiments to determine which hypothetical signal can be excluded at a predefined confidence level and thereby allowing, within the framework of the SM prediction, the measurement to be summarised by a lower limit on the mass of the Higgs boson.

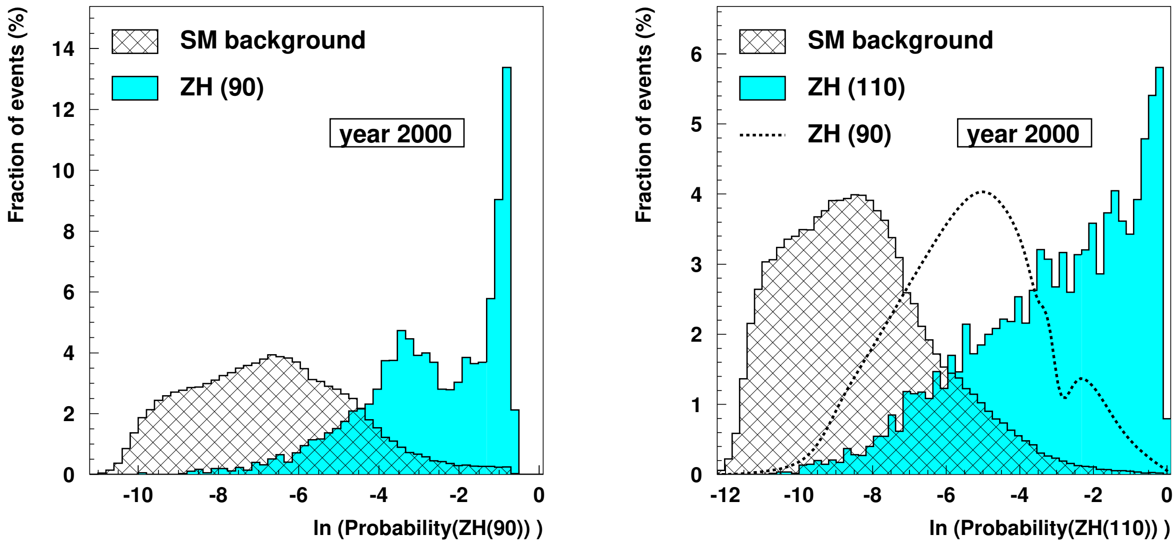


Figure 7.5: *Expected normalised distributions of the ZH-probability at a centre-of-mass energy of 206.7 GeV for ZH events (solid histograms) and the SM background processes (hatched histograms). The left plot shows the separation between ZH signal and SM background for a Higgs of 90 GeV/c^2 whereas the right plot shows the distributions for a 110 GeV/c^2 Higgs. The dashed line in the right plot shows the distribution for a ZH(90) signal.*

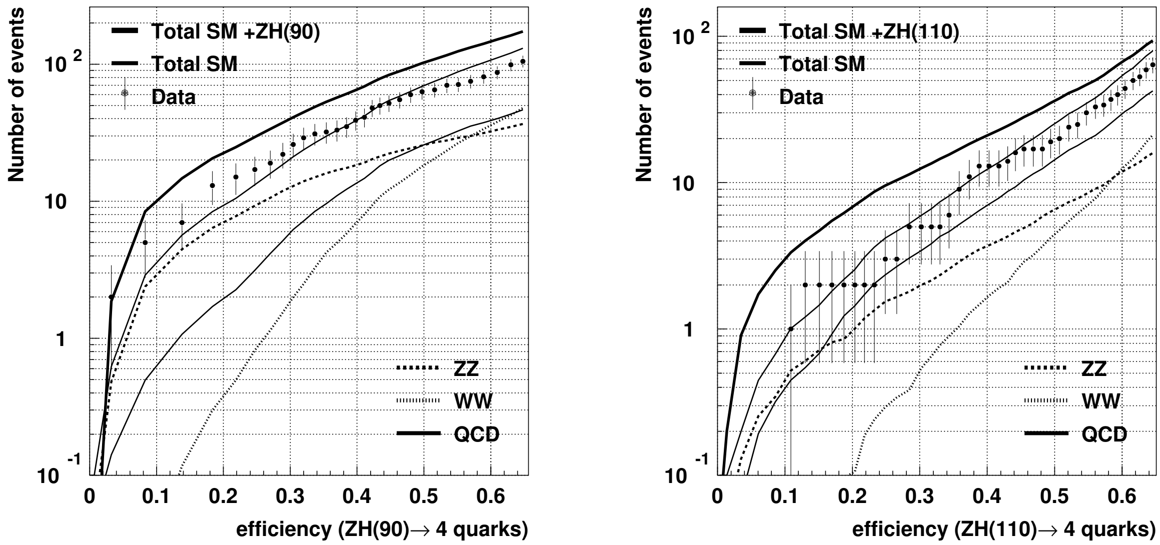


Figure 7.6: *Number of expected events left for various processes as a function of the efficiency for ZH events with a 4-quark final state in the combined year 2000 data set. The number of events observed in the data is represented by the points and compared to the expected number of events with and without an additional Higgs signal. The left plot is for ZH(90) events whereas the right plot shows the distributions for a 110 GeV/c^2 Higgs.*

M_H (GeV/c^2)	ln(P_{ZH})	ϵ(ZH)	ZH	SM(no ZH)	observed
90	-2.22	41.3	27.02	41.90	41
95	-3.13	57.2	31.74	80.57	69
100	-2.22	46.4	21.34	31.21	28
105	-2.12	43.7	15.22	19.68	18
110	-2.43	41.8	9.09	13.78	13
115	-3.03	28.4	1.64	4.82	4

Table 7.1: *The number of observed events with a ZH-probability above the value for which the product of efficiency and purity is maximal, is compared to the SM prediction (efficiencies are given in %) . Note that the probabilities are the ones optimised for the corresponding Higgs mass hypothesis. The number of predicted ZH events is also given. A good agreement between data and the SM prediction (without an additional ZH contribution) is observed for all mass hypotheses.*

7.4 Probing the sensitivity limits

Probing the sensitivity limits and interpreting the (negative) result from a search as a limit on a hypothetical physics parameter is common practise in search analyses performed in (particle) physics experiments. In the case of the search for the SM Higgs boson for example the result of the measurement, as an answer to the question “Is there a Higgs boson ?”, can not be summarised with a simple ’yes’ or ’no’, since (as can be seen from table 7.1) the experiment has no sensitivity to positively exclude nor to confirm the presence of a signal of a heavy Higgs boson.

In this section the procedure to determine the sensitivity limit of the analysis is introduced and applied to the analysis of the search for the SM Higgs boson. The result from a single measurement (for a specific Higgs mass) will be analysed in the form of hypothesis testing, where the null hypothesis is the one in which the signal is absent (background-only) and the alternate hypothesis is the one where the Higgs signal is present (signal+background). The method used to interpret the measurements is then the CL_s (or *modified frequentist*) procedure [80] that is used by the LEP collaborations to analyse[5] and summarise their combined results.

In section 7.4.1 a general introduction of the method is given and the *test statistic*, used to rank experiments from most background-like to most signal+background-like, is defined. Then the method to quantify the degree to which various hypotheses are favoured or excluded, given a specific measurement, is formally introduced. Within this framework a definition of the sensitivity reach of the measurement is given. In section 7.4.2 the method is applied to summarise the results from the search for the SM Higgs boson in the fully hadronic final state. The obtained sensitivity limit is finally interpreted (within the SM) as a lower limit on the mass of the Higgs boson in section 7.5.1.

7.4.1 General definition of the test statistic and confidence levels

The first step in assessing the separation power of the analysis in distinguishing between the two hypotheses is to construct a single variable that can be used to rank experiments from most

[5] The ALEPH collaboration uses a slightly different method [81] to present its individual result.

background-like to most signal+background-like. This variable is called the *test statistic*. A sensible choice is to define it as (a function of) the likelihood ratio between the two hypotheses since such a definition possesses the property that it allows the measurement to confirm as strongly as possible a signal when it exists or to exclude it as strongly as possible in its absence. The *test statistic* (X) is here defined as:

$$X = -2\ln(Q) \qquad \text{, with } Q = \frac{\mathcal{L}_{s+b}}{\mathcal{L}_b}, \tag{7.3}$$

where $\mathcal{L}_{s+b}$ is the likelihood for the signal+background hypothesis and $\mathcal{L}_b$ is the likelihood for the background-only hypothesis. In this way, X is constructed to increase monotonically for increasingly background-like measurements: low-X is signal+background-like and high-X is background-like. The probability density function for X for 'gedanken' background-only experiments is defined as $\mathrm{P_b}(X)$ and for signal+background experiments the probability density function is given by $\mathrm{P_{s+b}}(X)$.

To visualise the various definitions derived from the distributions of $\mathrm{P_b}(X)$ and $\mathrm{P_{s+b}}(X)$ the two probability density functions are shown in figure 7.7. The sensitivity of a measurement is given by the effective separation of the two distributions which is given by a combination of the intrinsic separation power of the analysis and the amount of collected statistics. Note that the experimental data produce only **a single** value for -2ln(Q) denoted as X_{obs}. For this value one can define compatibilities with either of the two hypotheses.

Compatibility with the background-only hypothesis

For a given experimental observation (X_{obs}) the *confidence level in the background hypothesis* ($\mathrm{CL_b}$) is given by the probability that a background-only experiment has a value for X that is larger or equal to X_{obs}. Therefore the probability that a background-only experiment is more background-like than the observed measurement can be formally written as:

$$\mathrm{CL_b} = \mathrm{P_b}(X \geq X_{\mathrm{obs}}) = \int_{X_{\mathrm{obs}}}^{\infty} \mathrm{P_b(X)}\,\mathrm{d}X \tag{7.4}$$

From this definition (see also figure 7.7) it is clear that 1-$\mathrm{CL_b}$ defines the probability that a background-only experiment looks more signal+background-like than the one observed in this measurement (the probability to observe an even larger fluctuation towards the signal region in case there is no signal present). For background-only experiments, by definition, 1-$\mathrm{CL_b}$ is on average $\frac{1}{2}$ whereas for the average signal+background experiment 1-$\mathrm{CL_b}$ is (very) small (in case the two distributions are well separated).

Compatibility with the signal+background hypothesis

Similar, for a given experiment the *confidence level in the signal+background hypothesis* is given by the probability that the test-statistic X from a signal+background experiment is larger than X_{obs}. Again, this can be interpreted as the probability that an experiment, in case there is indeed a signal present, looks more background-like than the observed measurement. Analogous to equation (7.4) it can be written as:

$$\mathrm{CL_{s+b}} = \mathrm{P_{s+b}}(X \geq X_{\mathrm{obs}}) = \int_{X_{\mathrm{obs}}}^{\infty} \mathrm{P_{s+b}(X)}\,\mathrm{d}X \tag{7.5}$$

For the average signal+background experiment, by construction, $\mathrm{CL_{s+b}}$ is $\frac{1}{2}$ whereas for the average background-only experiments this value will be very small (again under the assumption that the two distributions are well separated) as can be understood from figure 7.7.

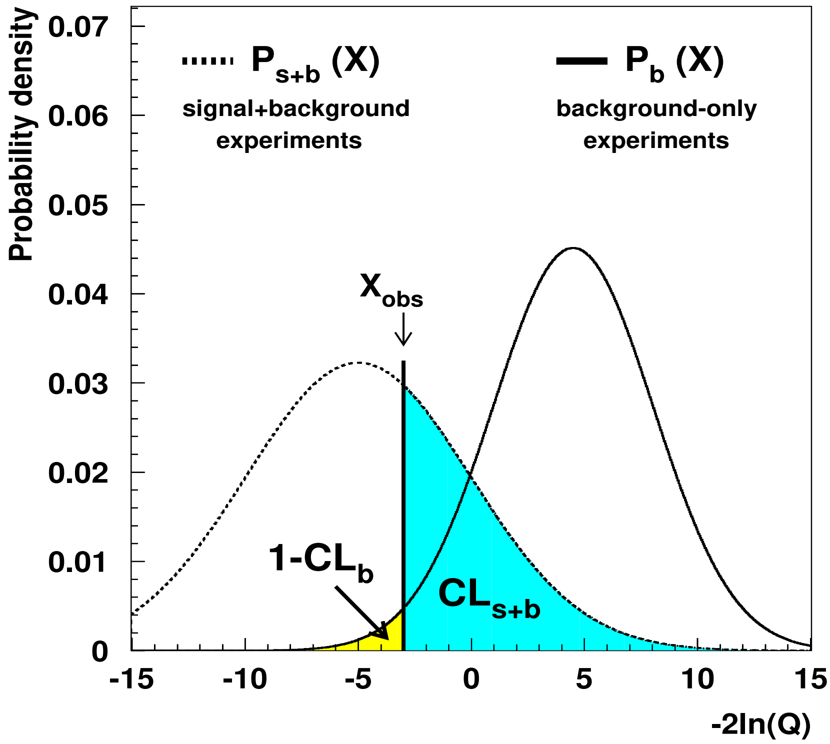


Figure 7.7: *Definition of the various confidence levels used in the analysis of testing the compatibilities with both a background-only and a signal+background hypothesis.*

Compatibility with the signal hypothesis

Interpretation of the difference between the number of expected and observed events as an estimation of the signal rate can lead to controversial results in measurements where there is no clear separation between the background and signal+background hypotheses (e.g. in case of a very low expected signal rate). When for example less events are observed than expected from background this can be interpreted as a negative signal cross section even to the point where a zero cross section is excluded at the 95% CL when the 'standard' methods of inference are applied. To deal with these difficult regions (the region where the limit is to be set is by definition the region where the experiment loses sensitivity) the confidence level in the signal+background hypothesis ($\mathrm{CL_{s+b}}$) is modified by normalising it to the confidence level for the background-only hypothesis ($\mathrm{CL_b}$). The (modified frequentist) *confidence level in the signal* is therefore defined as:

$$\mathrm{CL_s} = \mathrm{CL_{s+b}} \;/\; \mathrm{CL_b} \tag{7.6}$$

Note that $\mathrm{CL_s}$ is not a real confidence, but rather a ratio of confidences. Using this modified definition of the signal confidence allows to obtain sensible exclusion limits even when the observed rate is so low that the background hypothesis is called into question. The exclusion limits obtained using this modified definition are therefore always conservative. This method of assessing the sensitivity limit is known as the *Modified Frequentist* method.

Rules for discovery and exclusion

For a discovery the null hypothesis has to be rejected. To do that the 'incompatibility with the background-hypothesis' (1-$\mathrm{CL_b}$) is used. It is the probability that a background-only experiment fluctuates towards the signal region as much as in the observed measurement or more. The values for 1-$\mathrm{CL_b}$ can be transformed into standard deviations using a (one-sided) Gaussian approximation (see table 7.2). A discovery can be claimed when 1-$\mathrm{CL_b}$ is smaller than 5.7×10^{-7}, a so-called 5σ-effect.

1-$\mathrm{CL_b}$	size of effect
2.7×10^{-3}	3σ (evidence)
5.7×10^{-7}	5σ (discovery)

Table 7.2: *Conversions of values for 1-$\mathrm{CL_b}$ into one-sided Gaussian fluctuations.*

A signal hypothesis will be considered excluded at a given CL when:

$$1 - \mathrm{CL_s} < \mathrm{CL} \tag{7.7}$$

The lower limit on the Higgs boson mass obtained using the $\mathrm{CL_s}$ method should be interpreted as the sensitivity boundary of the experiment. This means, the boundary of a region where one would not have expected the signal (in case the signal exists) to fluctuate downward as much as in the actual measurement. In case of the search for the Higgs boson the lower limit on the Higgs mass is defined as the highest mass whose corresponding ZH signal can still be excluded at the 95% CL. This can be interpreted as the Higgs mass for which the probability that the signal fluctuated downward to the measured value (or more) is still smaller than 5% in case the signal exists.

Once the definitions of exclusion and discovery are given the method can be applied on the analysis of the search for the Standard Model Higgs boson as presented in this thesis.

7.4.2 Application of the $\mathrm{CL_s}$ method to a single channel Higgs search

Applying the $\mathrm{CL_s}$ method on the results from the analysis presented in section 7.3 requires the construction of a test statistic as in equation (7.3). For a given Higgs mass the best separation between the background-only and background+signal hypothesis is obtained by using the probability distributions as a separating variable (where signal and backgrounds have been normalised to equal luminosity).

One of the possibilities to construct the likelihood ratio of the two hypotheses is to perform a simple (Poisson) counting experiment for events above the value of the ZH-probability that provides optimal statistical sensitivity (see table 7.1). An improved sensitivity is obtained however by using the shape of the probability distribution for the two hypotheses. Using this approach the likelihood ratio ($Q = \mathcal{L}_{s+b}/\mathcal{L}_b$) between the signal+background and the background-only hypothesis can be written as:

$$Q = \frac{e^{-(s+b)}(s+b)^{n_{\text{obs}}} \; / \; n_{\text{obs}}!}{e^{-b} \; b^{n_{\text{obs}}} \; / \; n_{\text{obs}}!} \cdot \frac{\Pi_{i=1}^{n_{\text{obs}}} \frac{s\; S(p_i)+b\; B(p_i)}{s+b}}{\Pi_{i=1}^{n_{\text{obs}}} B(p_i)} \tag{7.8}$$

$S(p)$ and $B(p)$ are the probability density functions for the ZH-probability for signal and background events respectively. The most background-like part of the ZH-probability distribution is not used when computing Q and in expression (7.8) therefore n_{obs} is the number of observed events and s and b are the number of expected signal and background events with ln(p_i)>-3.5. Restricting to a reasonably pure part of the probability distribution does not reduce the sensitivity to the signal and is done to avoid potential large systematic effects from background processes (as was shown in section 6.4.1 for the ZZ cross section measurement). Expression (7.8) can be simplified to:

$$Q = e^{-s} \; \Pi_{i=1}^{n_{\text{obs}}} \left(1 + \frac{s\; S(p_i)}{b\; B(p_i)}\right) \tag{7.9}$$

and the test statistic -2ln(Q) (as defined in (7.3)) can be expressed as a sum of weighted events:

$$-2\ln(Q) = 2s \; - 2 \sum_{i=1}^{n_{\text{obs}}} \; \ln\left(1 + \frac{s(p_i)}{b(p_i)}\right) \tag{7.10}$$

In this expression $1 + \frac{s(p_i)}{b(p_i)}$ is the relative probability of an event with ZH-probability p_i to have originated from an experiment where there was signal+background to a background-only experiment: the likelihood ratio.

Computing the distributions for -2ln(Q) for 'gedanken' background-only and signal+background experiments and obtaining the value of -2ln(Q) in the data (X_{obs}) requires the correct weight for each event to be determined. The weight for each event is parametrised as a function of the ZH-probability (p_i) using the ZH-probability distributions from a large number of Monte Carlo background and signal events. As an example this is shown in the right plot of figure 7.8 for a Higgs boson of 110 GeV/c^2. In the left plot the original ZH-probability is shown for the two hypotheses in the pure ZH region (solid histogram is background-only and the open histogram represents the expectation in case there is also a ZH signal present).

To obtain the probability density functions for the signal+background and the background-only 'gedanken' hypotheses, a large number of 'gedanken' experiments have been performed by drawing a (Poisson) number of events from the distribution around the expectation corresponding to the luminosity as collected in the data.

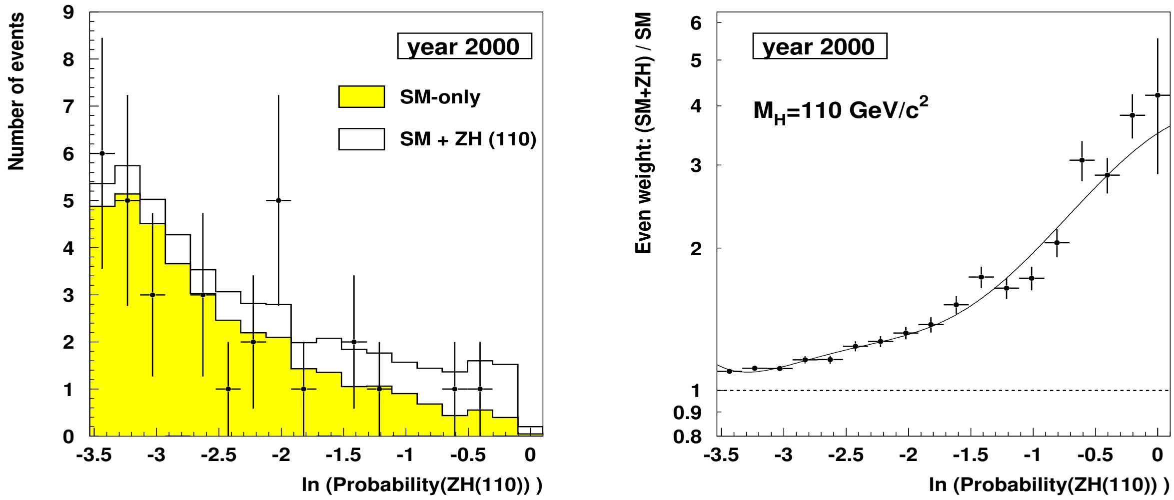


Figure 7.8: *The left plot shows the distribution of the logarithm of the ZH(110)-probability above -3.5 for SM background (solid histogram) and the expectation with the additional ZH(110) contribution (open histogram). The right plot shows the ratio of the signal+background and the background-only distribution (event weight) and its parametrisation (solid line).*

7.4.3 Example: results for a Higgs boson with a mass of 110 GeV/c^2

For a Higgs mass of 110 GeV/c^2 the probability density functions for the test-statistic for background-only and signal+background experiments, obtained from a large number of 'gedanken' experiments, are given in figure 7.9. For the background-only experiments the one and two standard deviations are given and also the observed value of -2ln(Q) is shown. The observation is clearly background like. The characteristics of these two distributions are given in the first two rows of table 7.3 as expected results for the average background-only and signal+background experiment.

The expected 1-CL$_\mathrm{b}$ for the average signal+background experiment is 0.003, far away from 5.7×10^{-7} indicating that if the signal is there it is not expected to be discovered (a 5σ effect as defined in table 7.2). On the other hand, since the CL$_\mathrm{s}$ for the average background-only experiment is smaller than 0.05 (0.016, see first row of table 7.3) the signal corresponding to this Higgs boson mass is expected to be excluded. With an observed CL$_\mathrm{s}$ of 0.008 it is also excluded (at 95% CL).

	1-CL$_\mathrm{b}$	CL$_\mathrm{b}$	CL$_\mathrm{s+b}$	CL$_\mathrm{s}$
average background-only experiment	0.500	0.500	0.008	0.016
average signal+background experiment	0.003	0.997	0.500	0.502
observed in year 2000 data	0.698	0.302	0.002	0.008

Table 7.3: *A summary of the characteristics for the analysis searching for a Higgs boson with a mass of 110 GeV/c^2. Apart from the expected numbers for the average background-only and signal+background 'gedanken' experiment also the observed values in the year 2000 data set are given.*

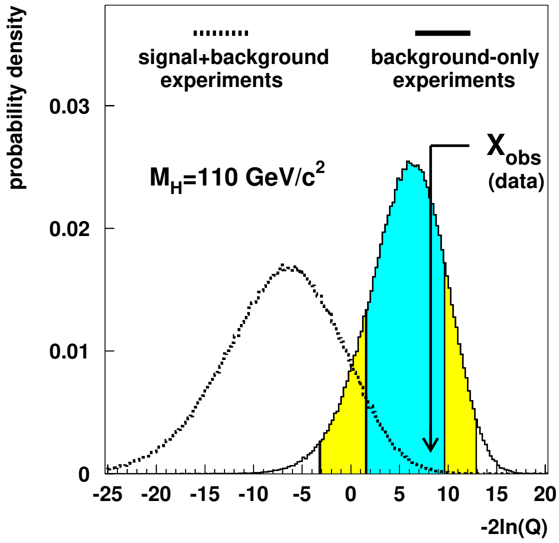


Figure 7.9: *The probability distribution for the test-statistic for background-only and background+signal (ZH(110)) experiments. The shaded areas indicate the one and two sigma regions for background-only experiments. The observed value for -2ln(Q) in the year 2000 data set is also indicated.*

7.5 Results

In this section the results from the search for the SM Higgs boson are presented. The result from the 4-quark final state is given in section 7.5.1. Since the main focus is on the maximum Higgs mass that can be excluded, only the year 2000 data set is used. The sensitivity of the data sets at centre-of-mass energies below 205 GeV is marginal for large Higgs masses, therefore the Higgs compatibilities and corresponding results are evaluated up to Higgs masses of 115 GeV/c^2. The (combined) results from the four LEP experiments are presented in section 7.5.3. These results use the full LEP2 data set and hence provide the most stringent (expected) limit on the mass of the Higgs boson.

7.5.1 Results for the 4-quark final state

The distribution of the probability density functions of the test-statistic for the two hypotheses at Higgs masses of 100, 105, 110 and 115 GeV/c^2 are constructed using a procedure similar to the one defined in the previous section. These distributions are shown in figure 7.10 where also the observed values for -2ln(Q) in the data are indicated. In order for the background-only hypothesis to be compared to the signal+background hypothesis for an arbitrary Higgs mass (M_H) a procedure has to be defined to construct the distributions for the test-statistic at that specific mass. In this section both this method and the one used to determine the evolution of the observed values in the data are discussed. At the end of this section the analysis is summarised by interpreting the results as the (expected and observed) lower limit on the mass of the SM Higgs boson.

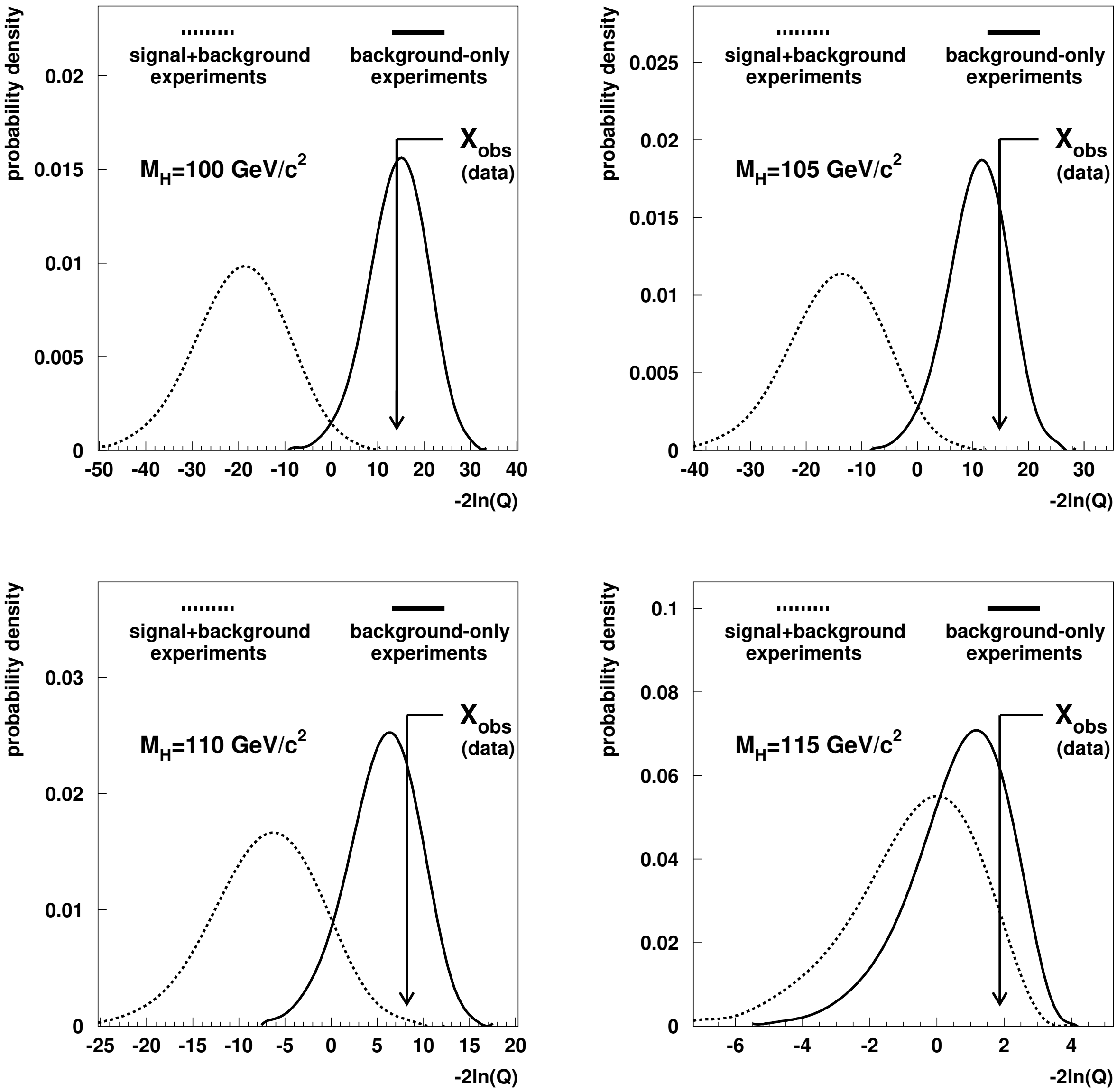


Figure 7.10: *The probability density functions for the test-statistic for background-only and signal+background experiments for hypotheses where the Higgs mass is 100 GeV/c^2 (top left plot), 105 GeV/c^2 (top right plot), 110 GeV/c^2 (bottom left plot), 115 GeV/c^2 (bottom right plot). The observed values at each of these Higgs mass hypotheses, using the year 2000 data set, are indicated by* X_{obs}.

Expected -2ln(Q) distribution as a function of $\mathrm{M_H}$

The expected -2ln(Q) distributions for a given Higgs mass ($\mathrm{M_H}$) are interpolated from those corresponding to the neighbouring two mass hypotheses that are known from Monte Carlo studies (figure 7.10). This is done for background-only and signal+background experiments separately. The shape of the test-statistic for an intermediate Higgs mass is constructed as a weighted average of the shapes of the two neighbouring distributions, with weights inversely proportional to the distances (in mass) between the intermediate Higgs mass and that of the neighbouring distributions.

Since the interpolation is only over a few GeV/c^2, the mean position of the distribution behaves to first order similar to the change in effective ZH cross section in the selected pure-ZH region. Using the known ZH cross section and the assumption that the selection efficiency changes smoothly over this mass range the mean value is found. The resulting expected -2ln(Q)(M_H) distributions are shown in the left plot of figure 7.11. In that plot the central position of the background-only (signal+background) distribution is indicated by the solid(dashed) line. For the background-only hypothesis the width of the distribution is indicated by the shaded areas (1 and 2 sigma).

Observed -2ln(Q) distribution as a function of M_H

The observed values for -2ln(Q) for an arbitrary Higgs mass hypothesis can not be obtained from the observed values at the neighbouring mass points for which the full ZH-probability distributions have been computed. In order to set a 'confident' confidence limit, a precise determination of the observed values in the (mass) region where the limit is expected is needed. These values can be obtained without having to repeat the full analysis at many different Higgs masses as is described below. For both expectation and observation the -2ln(Q)(M_H) distributions are translated into 1-CL_b(M_H) values. The results are shown in the right plot of figure 7.11.

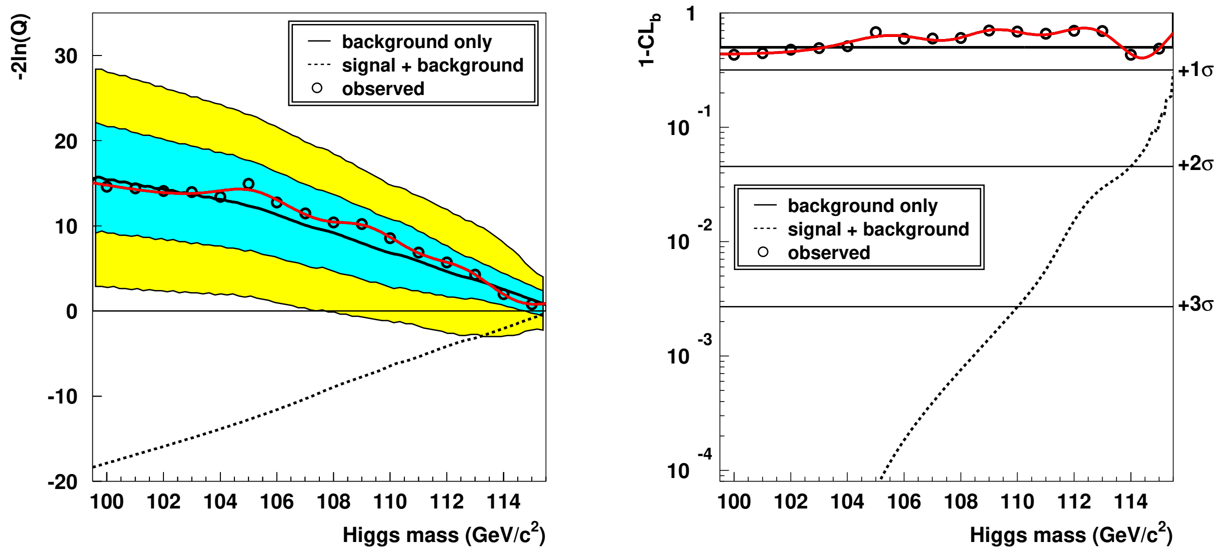


Figure 7.11: *Results from the search for the SM Higgs boson in the fully hadronic final state using the year 2000 data set. The left plot shows the evolution of the -2ln(Q) distribution for the background-only (solid line) and the signal+background-hypothesis (dashed line). The observed values in the data are indicated by the connected dots. The right plot shows the corresponding distribution of 1-CL_b for both hypotheses and the data.*

For Higgs masses between 100 and 115 GeV/c^2 extra ideogram compatibilities (mass-only information compatibilities per-pairing and assignment) have been computed at 1 GeV/c^2 intervals. Using these mass compatibilities the ZH probabilities can be computed for the corresponding Higgs masses since the other components entering the expression for the ZH-probability (as in expression 5.17) are either known as a function of the Higgs mass or are not (or only slightly) dependent on the Higgs mass. Each of these ZH-probability distributions for the data can be converted into a single observed value for -2ln(Q) without the use of dedicated Monte Carlo samples to study the background and signal distributions in the ZH-probability distribution. This conversion is possible since the probabilities are defined in such a way that the parametrisation used to compute the event weights ($1+\frac{s(p_i)}{b(p_i)}$ as is shown in figure 7.8 for a Higgs mass of 110 GeV/c^2) is almost identical for all Higgs masses. This characteristic allows the event weight to be computed for each data event. The sum of the weights for the events in the pure-ZH region can be converted into a value for -2ln(Q) using formula 7.10. The observed values of -2ln(Q)($\mathrm{M_H}$) in the data are shown in figure 7.11 were they are indicated (as in future plots) by connected points.

$\mathrm{CL_s}$ as a function of $\mathrm{M_H}$ and extracting a lower limit on the Higgs mass

The distribution of the signal confidence as a function of the Higgs mass, $\mathrm{CL_s}(\mathrm{M_H})$, is shown in figure 7.12 for the background-expectation (solid line) and the signal+background (dashed line) hypothesis. The 1σ uncertainty for the background-only experiments is indicated by the shaded area. As defined at the end of section 7.4.1 a Higgs mass hypotheses for which $\mathrm{CL_s} < 0.05$ can be excluded at 95%CL.

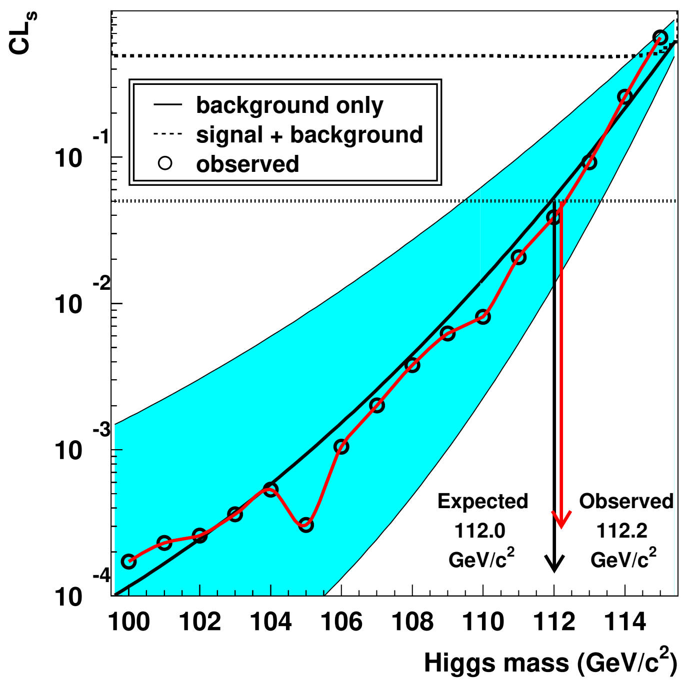


Figure 7.12: *Distribution of the signal confidence CL_s for the background-only and the signal+background hypotheses. The shaded area indicates the 1σ uncertainty for the background-only experiments. The observed values in the year 2000 data set are also shown.*

Using the intersection point at 0.05 the full measurement can be summarised as a 95% CL lower limit on the mass of the SM Higgs boson:

$$\text{Expected:} \quad \mathrm{M_H} > 112.0 \;\; \mathrm{GeV/c^2} \quad \text{(at 95\% CL)}$$

$$\text{Observed:} \quad \mathrm{M_H} > 112.2 \;\; \mathrm{GeV/c^2} \quad \text{(at 95\% CL)}$$

7.5.2 Systematic uncertainties and possible further improvements

In this section the uncertainty on the predicted background level is estimated and its effect on the expected and observed lower bound of the mass of the SM Higgs boson is evaluated. Finally some suggestions are given to improve the sensitivity of the analysis to a possible Higgs signal.

Systematic uncertainties

The search for the Higgs boson is a cross section measurement of a (very) small signal. In the region where the experiment loses sensitivity effectively only a handful of signal and background events remain. The statistical fluctuations are large and are absorbed in the definition of the lower limit on the mass of the Higgs boson. The systematic uncertainties are often small compared to the statistical uncertainties and are not expected to strongly affect the results. For counting experiments an analytic treatment to incorporate systematic uncertainties in an upper/lower limit is discussed in [82], but for more complex analyses the effect is evaluated using a more rigorous Monte Carlo approach.

As was shown in section 7.3, the background decomposition at high signal purity changes as a function of the Higgs mass in the hypothesis. For a 90 GeV/c^2 Higgs, the main background originates from Z boson pair production, but above 100 GeV/c^2 the dominant background is from $q\overline{q}(\gamma)$ events. This is illustrated in figure 7.13 where the relative background decomposition is shown as a function of the Higgs mass hypothesis at the point of optimal statistical sensitivity.

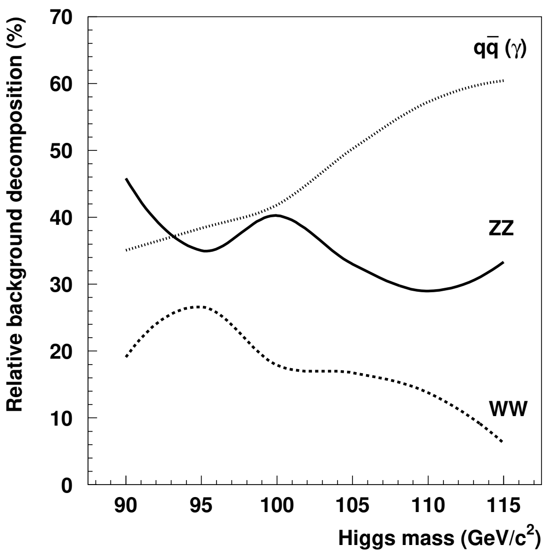


Figure 7.13: *This plot shows the (relative) background decomposition as a function of the Higgs mass in the hypothesis at the point where the product of efficiency and purity is maximal.*

The uncertainty on the predicted SM background level is therefore mainly coming from the uncertainty in the $q\overline{q}(\gamma)$ cross section (7%). The effect from the uncertainty on the gluon splitting rate into a $b\overline{b}$-pair on the background cross section is 3% and the uncertainty in the parametrisation of the b-tag likelihood ratios per jet (using a similar procedure as in 6.4.5) induces a maximum change in the background level of 6%. The combined uncertainty on the background level is therefore conservatively estimated to be around 10%.

The procedure defined in section 7.5.1 to find the maximum Higgs mass that can be excluded at 95% CL is repeated with a background cross section that has been increased by10%. This corresponds to a reduced sensitivity of 350(100) MeV/c^2 for the expected(observed) lower limit on the mass of the SM Higgs boson.

Possible further improvements

The analysis presented in this thesis is simultaneously optimised to identify ZH, ZZ, W^+W^- and $q\overline{q}(\gamma)$ events. It is therefore not only good in identifying signal-like events, but also in rejecting those events that are also compatible with a background hypothesis. Possible improvements to increase the sensitivity for a Higgs signal are presented below. They are similar to those suggested to improve the ZZ cross section measurement (see section 6.7), although with different emphasis:

- **Improved rejection of $q\overline{q}(\gamma)$ events**:
  An improved rejection of $q\overline{q}(\gamma)$ events (the dominant background) could be achieved by using a more complex topological variable than the D_{pur} variable. Implementation of a (more) complex topological variable induces correlations with the extracted mass information, but these effects are expected to be small compared to the gain in sensitivity due to the elimination of a large fraction of the $q\overline{q}(\gamma)$ events.

- **Quark-flavour tagging**:
  An improvement in solving pairing ambiguities can be achieved by using a more complex b-tagging algorithm that allows to (partially) separate b-quarks from anti-b-quarks. Using this information ($b\overline{b}$,$b\overline{b}$)-pairings can be favoured over (bb,$\overline{bb}$) ones. Resolving these ambiguities in a 4 b-quark event is more important for ZH events than for ZZ events, since in almost 20% of ZH events there are 4 b-quarks present compared to only 5% for ZZ.

- **Using boson angular distributions**
  Since the Higgs is a scalar particle, the angular distribution of the two bosons in ZH events is different from that produced in ZZ or W^+W^- events. This is shown in figure 7.14 where the absolute value of the cosine of the polar angle of both bosons is shown for ZZ, W^+W^- and ZH events at a centre-of-mass energy of 207 GeV. At the level of the pairing the corresponding angles of the constructed bosons can be used to disfavour an angular region that is more compatible with either a W^+W^- or ZZ region. Note that in the case of W^+W^- events also the charge of the bosons can be used to reduce their contribution to the total background.

7.5.3 Combined LEP results

Using the CL_s method, the results from the fully hadronic final state can easily be combined with those from other channels, since additional search channels can be incorporated in the expression for the likelihood ratio. The general expression for Q can be written (see equation 7.9 for a single channel) in the most general way as:

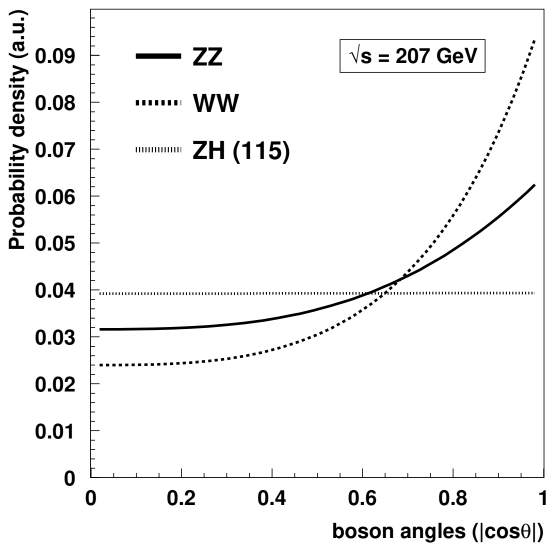


Figure 7.14: *The distribution of the absolute value of the cosine of the polar angle for bosons in ZZ, W^+W^- and ZH(115) events at a centre-of-mass energy of 207 GeV.*

$$Q = e^{-s}\, \Pi_{j=1}^{n_{\rm chan}} \Pi_{i=1}^{n_{\rm obs}} \left(1 + \frac{s_j\ S(p_{ij})}{b_j\ B(p_{ij})}\right) \tag{7.11}$$

In this expression j defines the search channel[6]. The test-statistic -2ln(Q) can therefore again be written as a sum of weighted events as in equation (7.10), where each channel contributes to the final result with the combined weight of its selected events. The results from the four LEP experiments can be combined in a similar way as they can simply be seen as separate search channels. A detailed analysis of the results obtained using the full LEP2 data set can be found in various publications of the LEP Higgs Working Group (for example [83] for the results up to the summer of 2001). A short summary of the main results[7] is given below.

The combination of the results from all LEP experiments, using the full LEP2 data set and all search channels, results in a distribution of -2ln(Q) as a function of the Higgs mass as shown in the left plot of figure 7.15, displaying the LEP combined sensitivity to a Higgs signal. Contrary to the results from the analysis presented in this thesis, the LEP combined results show an excess in number of observed events compared to the background expectation. This excess is compatible with a ZH signal where the mass of the Higgs boson is equal to 115.6 GeV/c^2. The two -2ln(Q) distributions corresponding to this specific Higgs mass (a slice in the left plot at this specific mass) are shown in the right plot of figure 7.15. Given these distributions, the probability that a background-only experiment results in a similar value of -2ln(Q) or even more signal-like is 3.4%: 1-CL$_{\rm b}$ = 0.034.

[6] The search channels for ZH events are all decay modes where the Higgs boson decays hadronically and the Z boson either decays into a quark anti-quark pair ($\sim$ 70%), a pair of neutrinos ($\sim$ 20%) or a lepton pair ($\sim$ 10%). Also the channel in which the Higgs decays into a pair of tau leptons and the Z boson decays hadronically is used.

[7] All results quoted here are preliminary and represent the status at the time of the 2001 summer conferences.

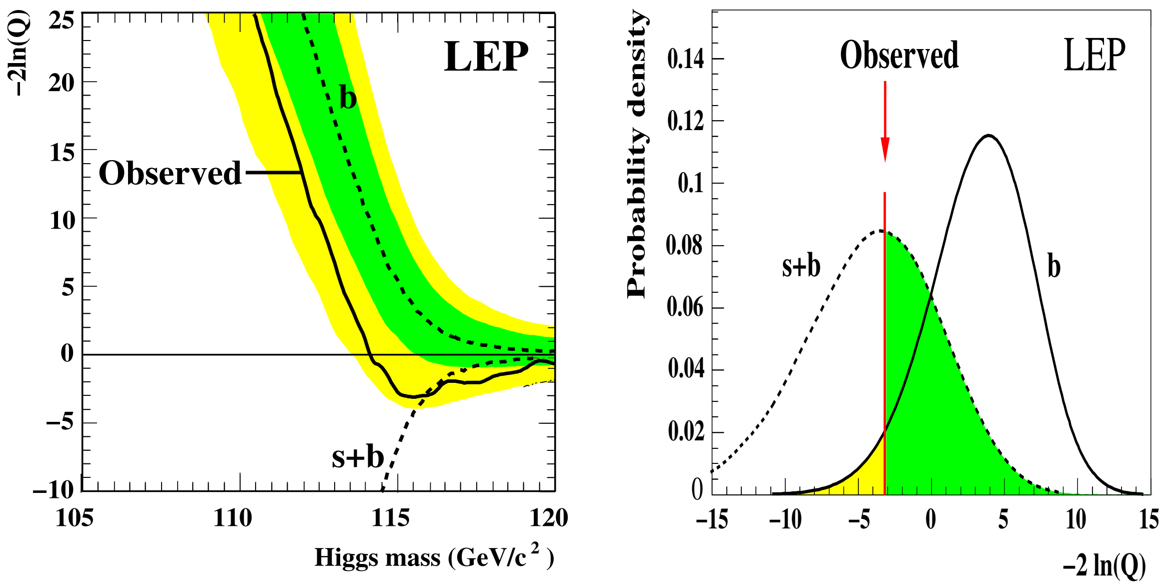


Figure 7.15: *LEP combined results from the search for the SM Higgs boson. The left plot shows the evolution of the -2ln(Q) distribution for the background-only (dashed line with 1 and 2 sigma bands) and the signal+background-hypothesis (lower dashed line). The observed values are indicated by the thick solid line. The right plot shows the -2ln(Q) distributions for the two hypotheses for a Higgs mass of 115.6 GeV/c^2.*

With a 1-CL$_\mathrm{b}$ of 0.034 (a 2.1σ effect), the combined LEP measurement indicates a hint of a possible Higgs signal being present in the data. However, this signal compatibility is not the combined effect of four experiments each supporting a signal+background hypothesis. From the two most sensitive experiments, ALEPH and DELPHI, the results from DELPHI clearly favour a background-only explanation of the observed data whereas the ALEPH data are very signal+background-like. As an indication of the sensitivity of the individual LEP experiments in combination with their actual measurements, the expected limits and observed 1-CL$_\mathrm{b}$'s are shown in table 7.4 for various combinations of channels and experiments.

From these numbers it is clear that the signal-like interpretation of the data is most clearly favoured by the results from the (fully hadronic final state in the) ALEPH experiment. It should be noted that this channel is one of the most sensitive channels in LEP and that, if the measurement is interpreted as a Higgs signal, the obtained result in the ALEPH experiment is compatible with almost a 2σ upward fluctuation of a ZH(115.6) signal (as is indicated by a CL$_\mathrm{s+b}$ of 0.94). The DELPHI experiment on the other hand observed less events than expected from background alone. This downward fluctuation is quite strong and the probability that a Higgs signal, if it exists, would result in a measurement that was even more background-like is only 2%. When all LEP experiments are combined, still a preference for a signal+background interpretation remains. A clear, but to many people unsatisfactory result: when DELPHI is left out from the combination there is a strong preference for a signal+background interpretation of the data (1-CL$_\mathrm{b}$=3.7$\times$10^{-3} which is almost a 3σ effect, see table 7.2). On the other hand, when ALEPH is left out of the

	expected sensitivity lower limit	observed lower limit	Observations at M_H=115.6 GeV/c^2 1-CL_b	CL_{s+b}
ALEPH	113.8	111.5	2.0×10^{-3}	0.94
DELPHI	113.5	114.3	0.87	0.02
L3	112.7	112.2	0.24	0.47
OPAL	112.6	109.4	0.22	0.47
DLO	114.9	114.8	0.49	0.07
ALO	–	–	3.7×10^{-3}	0.83
Four-jet	–	–	0.016	0.74
Missing energy	–	–	0.40	0.26
All but four-jet	–	–	0.34	0.19
LEP	115.4	114.1	0.034	0.44

Table 7.4: *The first two columns show the expected and observed sensitivity in the search for the SM Higgs boson in various combinations of experiments and search channels. The last two columns present the observations for a Higgs mass hypothesis of 115.6 GeV/c^2. DLO(ALO) represents the combined LEP results when ALEPH (DELPHI) is left out from the combination. The results from the ALEPH, DELPHI and OPAL experiment (and therefore those of the LEP combination) are preliminary.*

combination, the results are in perfect agreement with what is expected from a background-only hypothesis (1-CL_b=0.49).

Interpreting the combined measurement as a lower limit on the mass of the SM Higgs boson will therefore result in a lower limit that is smaller than the one that was expected. The combined LEP results can be summarised as follows:

$$\text{Expected:}\quad M_H > 115.4 \ \text{GeV/c}^2 \quad \text{(at 95\% CL)}$$

$$\text{Observed:}\quad M_H > 114.1 \ \text{GeV/c}^2 \quad \text{(at 95\% CL)}$$

At this point an important remark should be made on how (not) to interpret a lower bound on the Higgs mass. A lower bound on the Higgs mass says nothing on the probability of the Higgs mass to be higher or lower than some value. To be allowed to make such a statement the direct results must first be folded with a prior probability density distribution for the Higgs mass itself. In the next section the results from such an exercise are described.

7.5.4 Combining the indirect with the direct measurements

In this section the indirect measurement of the Higgs mass (see section 2.3.3) is combined with the direct exclusion limits from LEP to obtain the probability density function for the Higgs mass[8].

[8] Of course the Higgs mass (if the Higgs exists) has a single unknown value, but in the Bayesian framework the mass is treated as a stochastic variable.

Given the exclusion limit from the direct Higgs searches at LEP2 at 114.1 GeV/c^2, the additional information contained in the likelihood curve (figure 2.3) for the Higgs mass from the electroweak precision measurements below around M_H=110 GeV/c^2 is extremely small. Above 116 GeV/c^2 however, the direct searches have no sensitivity for a Higgs signal and all sensitivity to the Higgs mass comes from the indirect measurements. While the 'statistical' properties of a particle with unknown mass are ill defined, the likelihood ratios associated to these measurements can be combined in a rigorous way. The main results from such an exercise [84] and p.d.f of the Higgs mass are summarised here. In general, given a certain data set, the probability density function of the Higgs mass can be constructed using Bayes' formula:

$$\mathrm{p(M_H|world\ data) \propto p(world\ data|M_H)\ p(M_H)} \tag{7.12}$$

In this expression $\mathrm{p(world\ data|M_H)}$ describes the combined likelihood of the electroweak measurements within the SM as a function of the Higgs mass and $\mathrm{p(M_H)}$ is the *a priori* probability density function for values of the Higgs mass. They are combined to give the probability density function for the Higgs mass: $\mathrm{P(M_H|world\ data)}$.

$\mathrm{p(world\ data|M_H)}$ is a combination of the likelihood curve for the Higgs mass obtained from electroweak precision measurements and the results from the direct searches at LEP2. The exercise of [84] is based on the combined electroweak results in the summer of 2000 [85] and the direct Higgs search results at the time of the LEP fest (October 2000). To obtain the probability density function for the Higgs mass also a prior distribution for the Higgs mass, $\mathrm{p(M_H)}$, has to be defined. A commonly accepted choice, as is the case in the exercise whose results are discussed here, is to define $\mathrm{p(M_H)}$ to be flat in ln($\mathrm{M_H}$). The information on the Higgs mass from the indirect measurements ensures a properly normalised posterior probability density function as is shown in figure 7.16.

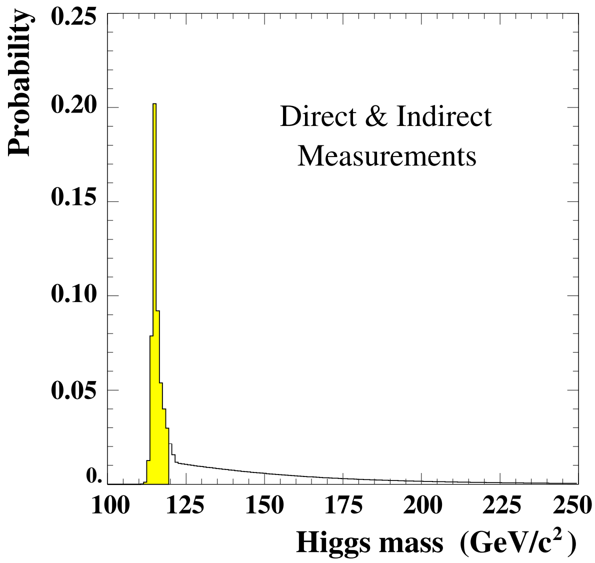


Figure 7.16: *This plot shows the obtained probability density function for the mass of the Higgs boson when the results from the direct and indirect measurements are combined. The two areas both contain 50% of the full integral.*

The distribution has a sharp peak around 116 GeV/c^2 reflecting the excess observed in the direct searches at LEP. Some numbers characterising this distribution are:

$$\mathrm{M_H = 119\ \ GeV/c^2\ \ (median\ mass)\quad and\quad M_H < 205\ \ GeV/c^2\quad at\ 95\%\ CL}$$

7.6 Higgs physics in the near future

The result of the direct search for the SM Higgs boson at LEP is a hint for a signal that is however not significant enough to claim a discovery. In the next years, the discovery of the Higgs boson and the study of its properties therefore remains one of the main topics in high-energy physics.

If the Higgs boson exists and has a mass in the range as indicated by the electroweak precision measurements (roughly from 100 to 250 GeV/c^2) it will be discovered in the coming 5-8 years at one of the two main hadron colliders: the Tevatron at Fermilab or the Large Hadron Collider at CERN. The sensitivity of the multi-purpose experiments at the Tevatron (D0 and CDF) and LHC (ATLAS and CMS) have been investigated in terms of their discovery potential for a Higgs boson. A quick summary of the main results is presented here.

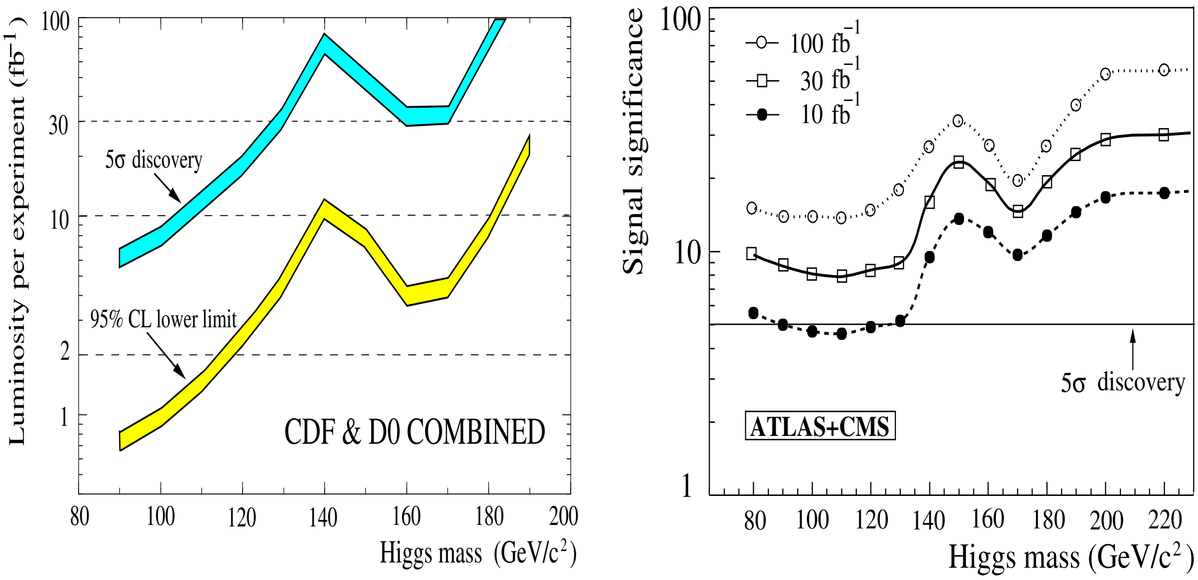


Figure 7.17: *The left plot shows the discovery and exclusion potential for the CDF and D0 combined SM Higgs search as a function of the Higgs mass and the collected luminosity per experiment. The right plot shows the ATLAS and CMS combined signal significance as a function of the Higgs mass for three different integrated luminosities. The 5σ discovery mark is also indicated.*

The RUN2 of the Tevatron (a proton anti-proton collider with a centre-of-mass energy of 2 TeV) has already started. The main results from the Higgs sensitivity studies [86] are summarised in the left plot of figure 7.17 where the luminosity needed (per experiment) for the Tevatron to either exclude the signal at 95% CL or to claim a discovery is given as a function of the Higgs mass. With around 10 fb^{-1} of luminosity per detector (this luminosity is expected to be reached at the end of 2006 [87]), the Higgs mass can be excluded up to a mass of 180 GeV/c^2, but for a large range of masses many more years of data-taking are needed to claim a discovery. The Tevatron predictions were made using a simplified detector simulation and the impact of systematic uncertainties can only be estimated once real data is collected and analysed.

The big adversary of the Tevatron is the LHC collider at CERN (a proton-proton collider with a centre-of-mass energy of 14 TeV) that is scheduled to start in 2006. It will deliver about 10 fb^{-1}

per year (at low luminosity running) to the individual LHC detectors. In the right plot of figure 7.17 the ATLAS[88] and CMS[89] combined signal significance as a function of the Higgs mass is shown at three different values of the integrated luminosity. For masses between 200 and 500 GeV/c^2 the LHC sensitivity is about equal for all masses, and above that value the significance drops only very slowly. If the Higgs boson is above 140 GeV/c^2 it will be discovered within one year. For a light Higgs boson in the region below 125 GeV/c^2 it will require slightly more data. As is the case for the Tevatron predictions, the numbers assume complete and well understood detectors in addition to well understood physics backgrounds, so the actual discovery claim might take a bit longer.

From these numbers and dates it is clear that there will be an exciting race for the discovery of the Higgs boson between the various experiments and colliders in the coming decade. In the more distant future (2010+) a high luminosity linear e^+e^- collider (TESLA) will allow a very precise determination of the properties of the Higgs boson like its mass, width, spin and the coupling to various quark types and gauge bosons to test the (non-)SM predictions. Although TESLA will not be constructed to discover the Higgs boson it is interesting to note that a single day of data taking would be more than enough to claim a discovery.

7.7 Summary and conclusions

In this chapter the analysis method as defined in chapter 5 has been applied to the search for the SM Higgs boson in case both the Z boson and the Higgs decay into a quark anti-quark pair. For each Higgs mass hypothesis the analysis is optimised by using (analytically where possible) the expected properties of both the ZH signal and the SM background processes. A detailed description of the method to define the sensitivity region of the analysis is given and applied to the results in the fully hadronic channel. In the year 2000 DELPHI data set no hint for an additional Higgs signal was observed and the observed lower limit on the Higgs mass was close to the one expected. The analysis can be summarised as follows:

$$\text{Expected lower limit:} \quad M_H > 112.0 \ \text{GeV/c}^2 \quad \text{(at 95\% CL)}$$

$$\text{Observed lower limit:} \quad M_H > 112.2 \ \text{GeV/c}^2 \quad \text{(at 95\% CL)}$$

The combined measurement from the four LEP experiments (using all search channels and the full LEP2 data set) resulted in a hint for a signal at a (most likely) Higgs mass of 115.6 GeV/c^2. The significance of the excess (2.1σ) was however not enough to claim a discovery.

After the shutdown of LEP, the search for the Higgs boson and the determination of its properties will remain one of the main topics in high-energy physics in the coming years. In particular the next 5-8 years will be very exciting at the two main hadron colliders since an important discovery is within reach. Despite the more complex detectors and analyses compared to LEP, in a sense these experiments have it more easy. With their enormous discovery (or exclusion) potential they are in a win-win situation, since an absence of the Higgs is probably even more interesting than discovering it.

Chapter 8

Higgs production (Hadronic decay)

In the SM the properties of the Higgs boson are well defined as is described in section 2.4. For Higgs boson masses accessible in direct searches at LEP2 (below 115 GeV/c^2) the Higgs boson decays predominantly into either a pair of b-quarks ($\sim 82\%$) or into $\tau^+\tau^-$ ($\sim 8\%$). Since also the cross section is predicted as a function of the mass the result from the direct searches (a cross section measurement) is presented as a lower limit on the mass of the Higgs boson. Such a procedure is model dependent and does not maximally exploit the information contained in the data.

There are many models in which the Higgs coupling to down-type fermions (like b-quarks) is suppressed. This can for example be realised in a two Higgs doublet model, in super-symmetric models or in more exotic hypotheses like composite models [90]. In such models the Higgs will decay predominantly into a $\mathrm{c\bar{c}}$ pair or, through top-quark loops, into gluon pairs. The existing Higgs analyses, like the one presented in the previous chapter, have a significantly reduced sensitivity to these signals as they rely heavily on the identification of b-quarks in the final state. In this chapter an analysis is presented to derive a general upper limit on the cross section of a hadronically decaying scalar particle that is pair produced with a Z boson. Such a *flavour independent* cross section limit allows parameter space of specific models to be excluded. A search for the Higgs boson without using b-tag information to identify the Higgs would in addition allow an independent cross check of the background behaviour in terms of mass and topological information in the search for the SM Higgs boson.

In section 8.1 an overview is given of the Monte Carlo signal samples and data set. In section 8.2 the characteristics of the probability distributions are discussed and an overview of the performance of the analysis in separating the signal from the background at various Higgs mass hypotheses is given. Then, in section 8.4, the method used to set a cross section limit is defined and applied to the results from the measurements on this specific data set. The chapter then continues with an overview of the combined LEP results and ends with a summary and conclusion.

8.1 Signal generation and strategy

The cross section exclusion measurement is performed using the data collected in the years 1999 and 2000, with centre-of-mass energies ranging from 192 to 209 GeV. The corresponding integrated collected luminosities by DELPHI in this period are given in table 3.1 and 3.2. They amount to a total integrated luminosity of 452.4 pb^{-1}. The HZHA [91] programme was used to generate signal Monte Carlo events at each centre-of-mass energy for Higgs masses from 50 to 110 GeV/c^2 at 5 GeV/c^2 intervals. For each of these hypotheses the Higgs was made to decay explicitly into a quark-pair ($s\overline{s}$, $c\overline{c}$ or $b\overline{b}$) or into a gluon pair. For each of these Higgs decays the exclusion potential is determined and an upper limit on the cross section is computed.

Differences in sensitivity between the various quark-pair final states are expected to be small, except for gluon decays, where the jet structure is significantly different from that of a quark-jet [92]. The main differences originate from the enhanced (charged) particle multiplicity for gluon jets. The impact on the total charged multiplicity in a ZH event is indicated in the left plot of figure 8.1. As a result, the pre-selection efficiency for ZH events in which the Higgs decays into a gluon pair is about 5% higher than for an $s\overline{s}$ decay for all Higgs masses as is shown in figure 4.13. This increase is not caused by the explicit cut on the charged multiplicity ($>$17 as described in section 4.5.2), but is the result of the enhanced average reconstructed jet-multiplicity. The increase in particle multiplicity and jet-broadness (a gluon jet is also broader than a quark jet) allows a jet to be reconstructed even if the original parton direction was directed towards a less efficient region of the detector and/or a few particles remain undetected. The fraction of $q\overline{q}gg$ events that is reconstructed as a 3-jet event is therefore less than that for $q\overline{q}q\overline{q}$ events, resulting in the observed enhanced pre-selection efficiency since an event is required to have at least 4 reconstructed jets.

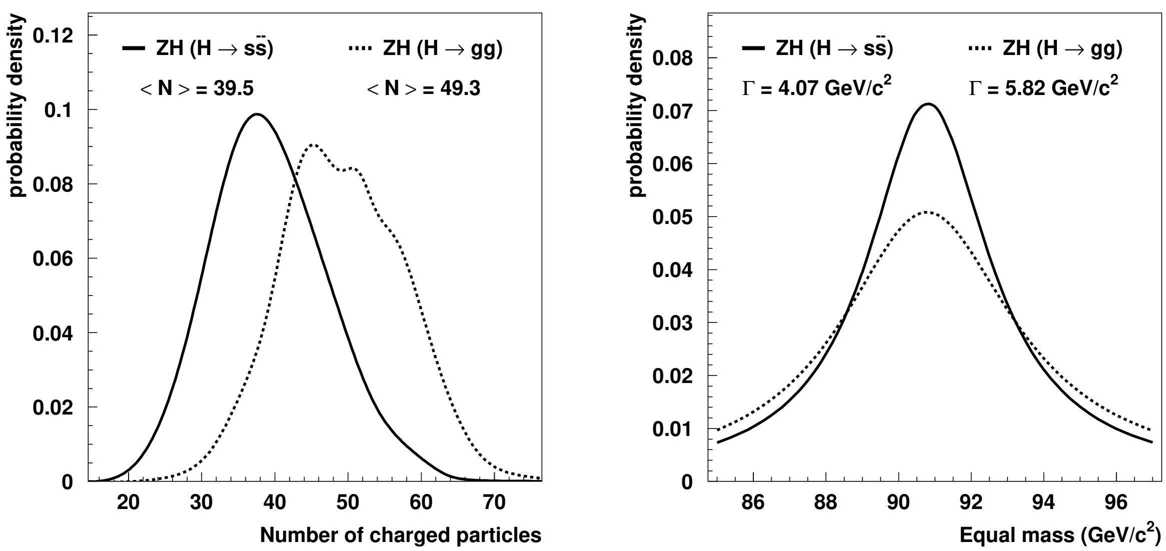


Figure 8.1: *Differences in the ZH(90) 4-quark final state when the Higgs decays into either $s\overline{s}$(solid lines) or a gluon-pair (dashed lines). The left plot shows the charged particle multiplicity and the right plot shows the fit to the reconstructed equal-mass distribution.*

Another effect of the increased jet multiplicity and broadening is the fact that the uncertainties in the estimation of the four-momentum of the initial parton are larger in case the jet originated from a gluon relative to a quark fragmentation. A consequence of this is that the di-jet mass resolution is deteriorated. This is shown in the right plot of figure 8.1, where a fit to the reconstructed *equal mass* distribution is shown for ZH(90) events.

The impact of these effects on the separation between signal and background varies as a function of the Higgs mass in the hypothesis as is shown in the next section. The upper limit on the hadronic Higgs cross section is therefore conservatively defined for each mass hypothesis as the limit obtained from the Higgs boson decay for which the analysis was least sensitive.

8.2 Probability distributions and performance

In this section the performance of the analysis is investigated as a function of the Higgs mass in the hypothesis. The Higgs mass hypotheses that are considered run from 50 to 110 GeV/c^2 and for each mass hypothesis the performance for two decay modes is evaluated: Higgs decays to either $s\bar{s}$ or a gluon pair. First it is shown that for a given Higgs mass hypothesis (M_H), the analysis is indeed automatically optimised for a ZH(M_H) signal. This is followed by the evaluation of the performance at each mass point for both decay modes. This is done by comparing the number of predicted events from the SM and SM+ZH(M_H) to the number of observed events after a cut in the ZH-probability distribution at the point where the value of the product of efficiency and purity of the corresponding optimised analysis is maximal. When a counting experiment would be performed a cut imposed at this value of the ZH-probability distribution results in the smallest statistical uncertainty of the cross section measurement.

Optimised analyses. An example: a 95 GeV/c^2 Higgs analysis

Since many characteristics of the Higgs signal are used analytically in the expression for the event-by-event probability, the analysis is automatically optimised once a specific hypothesis is chosen, as was the case for the SM Higgs. A few characteristics of this optimisation are illustrated using the analysis for a 95 GeV/c^2 Higgs boson. The corresponding (normalised) distribution of the logarithm of the event-by-event probability is shown in figure 8.2 for both the SM background expectation (hatched histogram) and for three different Higgs signals with masses of 80 GeV/c^2 (dashed line), 95 GeV/c^2 (solid histogram) and 110 GeV/c^2(dotted line). Maximal intrinsic separation between Higgs signal and background is indeed achieved when the Higgs is 95 GeV/c^2.

The degradation in this intrinsic separation power is indicated in the left plot of figure 8.3 where the mean of the ZH(95)-probability distribution is shown for signals with a varying Higgs mass between 50 and 110 GeV/c^2. In the right plot of that figure, the effective separation power is shown by plotting the maximum value of the product of efficiency and purity that can be obtained by making the cut on the probability distribution after scaling the signal and backgrounds to equal luminosity. For light Higgs masses, although the intrinsic separation power is small, the signal cross section is much larger which results in an almost similar effective sensitivity to that of the signal for which the analysis is optimised. It can also be noted that for the optimised signal, a 95 GeV/c^2 Higgs, the analysis is more sensitive to a Higgs decaying into $s\bar{s}$ than it is for a Higgs decaying into a gluon pair.

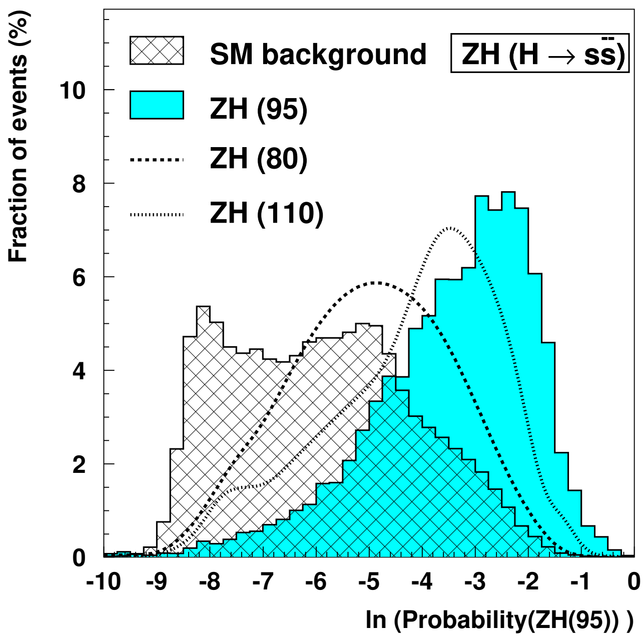


Figure 8.2: *This plot shows the logarithm of the ZH(95)-probability density for the combined 1999-2000 data set for the SM expectation (hatched histogram) and ZH(95) signal events (solid histogram). The distributions for Higgs signals where the Higgs boson is 80 GeV/c^2 (dashed line) and 110 GeV/c^2 (dotted line) are also shown.*

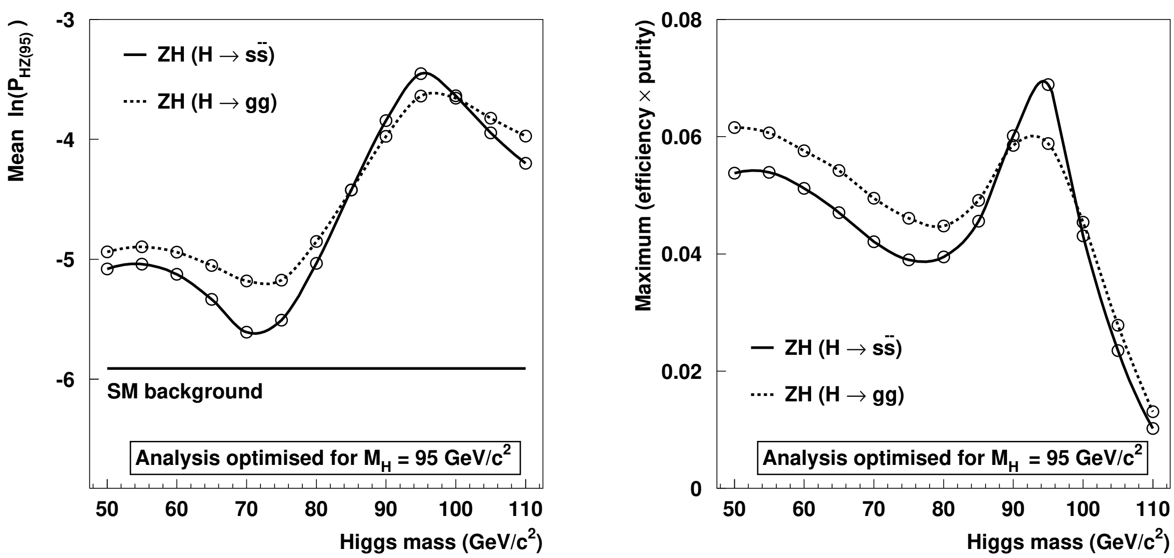


Figure 8.3: *The left plot shows the* <u>*intrinsic*</u> *separation power of the ZH(95) analysis. The mean of the logarithm of the ZH(95)-probability distribution is shown for the SM background and Higgs signals of varying mass. The solid(dotted) line indicates a Higgs decaying into $s\bar{s}(gg)$. The right plot shows the* <u>*effective*</u> *separation power (the backgrounds and signals are scaled to equal luminosity) by the maximum of the product of efficiency and purity.*

Performance as a function of M_H

As is the case for the SM Higgs boson, the observations at different centre-of-mass energies are combined. The performance of the optimised analysis for a given Higgs mass (M_H) and decay is shown in figure 8.4. In that plot the maximum product of efficiency and purity is given as a function of M_H, where for each mass hypothesis an optimised analysis was used. For Higgs masses below 85 GeV/c^2 the analysis is most sensitive to a Higgs boson decaying into a gluon pair. For heavier Higgs bosons, the analysis is more sensitive to $s\overline{s}$ Higgs final states.

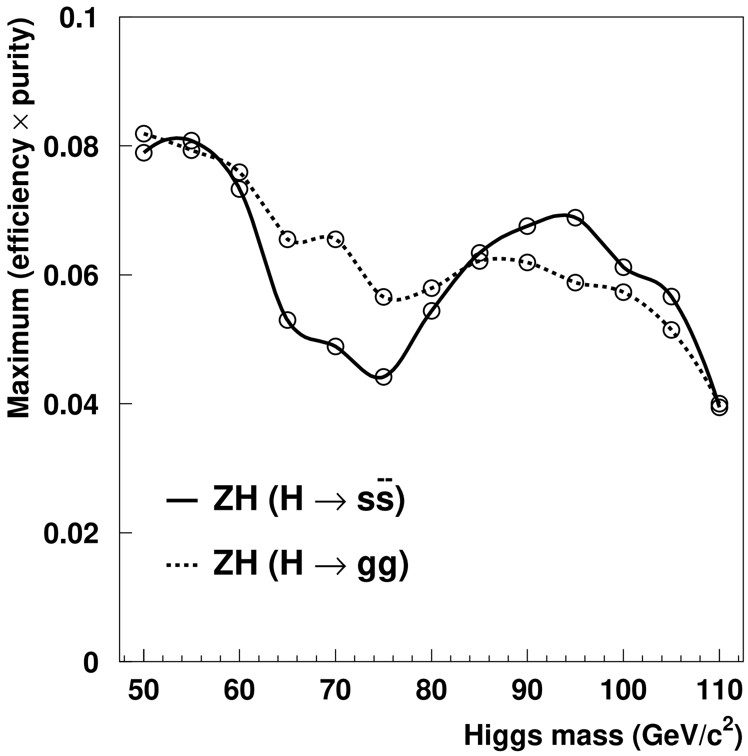


Figure 8.4: *This plot shows the maximum of the product of efficiency and purity for optimised analyses for Higgs masses ranging from 50 to 110 GeV/c^2. The solid(dashed) line represents the performance for a Higgs decaying into $s\overline{s}$ (a gluon pair).*

At this optimal point (the cut on the ZH-probability where a counting experiment would yield the statistically most precise measurement), the number of events predicted from SM background processes and for the Higgs signal (both decay modes) are shown in table 8.1 for a range of Higgs masses. In addition the corresponding value of the ZH-probability at the optimal cut and the efficiencies for the Higgs signal are given. The uncertainty on the efficiencies are around 1%. For a large range of (light) Higgs mass hypotheses an excess of events is observed compared to the SM expectation. It is therefore likely to be due to a general underestimation of the $q\overline{q}(\gamma)$ background. An evaluation of the systematic uncertainties involved in the background estimation in that particular kinematic region is made in section 8.3.2.

I: Higgs decay into $s\bar{s}$

M_H (GeV/c^2)	ln(P_{ZH})	ϵ(ZH)	ZH	SM(no ZH)	observed
50	-3.23	41.3	118.26	500.03	534
55	-3.03	40.4	109.44	438.21	470
60	-3.85	55.2	140.62	919.09	982
65	-4.68	64.8	154.89	1739.17	1853
70	-4.30	62.0	137.63	1609.33	1714
75	-4.33	64.1	129.20	1745.87	1860
80	-3.88	53.9	97.75	870.68	921
85	-3.63	53.1	85.31	628.94	661
90	-3.15	45.4	62.69	358.16	380
95	-2.98	40.9	46.01	227.00	235
100	-3.23	41.7	34.95	203.44	216
105	-3.55	44.3	23.80	162.46	155
110	-4.20	44.2	12.01	122.38	105

II: Higgs decay into a gluon pair

M_H (GeV/c^2)	ln(P_{ZH})	ϵ(ZH)	ZH	SM(no ZH)	observed
50	-3.58	48.0	137.59	668.78	713
55	-3.90	56.0	151.47	917.25	990
60	-3.98	59.2	150.67	1023.63	1088
65	-4.25	62.6	149.72	1281.04	1372
70	-4.15	65.3	144.91	1299.96	1393
75	-4.20	67.3	135.59	1477.27	1576
80	-4.20	66.4	120.37	1259.71	1345
85	-4.03	62.9	101.05	920.90	941
90	-3.98	62.0	85.63	771.05	800
95	-4.15	63.3	71.22	694.93	730
100	-4.23	61.2	51.27	496.00	500
105	-4.28	55.5	29.81	291.92	269
110	-4.35	47.3	12.85	138.83	119

Table 8.1: *The number of observed events at the value of the cut on ln(P_{ZH}) for which the product of efficiency and purity is maximal (efficiencies are given in %), is compared to the prediction of the SM. Note that the analyses used are the ones optimised for the corresponding Higgs mass hypothesis. The number of predicted events from a ZH signal is also given. Table I(II) shows the results for a Higgs boson decaying into $s\bar{s}$ (a gluon pair).*

8.3 Excluding the hadronic ZH cross section

In this section the sensitivity of the DELPHI experiment is presented by excluding a range of hadronic ZH cross sections using the analysis described in this thesis. The excluded cross sections as a function of the Higgs mass are presented as exclusions relative to the (hadronic) SM ZH cross section using the notation $\zeta_{\mathrm{ZH}}(\mathrm{M_H})$, where

$$\zeta_{\mathrm{ZH}} = \sigma_{\mathrm{ZH}}/\sigma^{\mathrm{SM}}_{\mathrm{ZH(hadronic)}} \tag{8.1}$$

The method used to exclude a cross section for a particular Higgs mass hypothesis is the $\mathrm{CL_s}$ method that has been discussed in detail in section 7.4. As an example, first an upper limit on the cross section for a 80 GeV/c^2 Higgs boson decaying into $\mathrm{s\bar{s}}$ is determined. This is followed by the full hadronic exclusion limit for Higgs masses from 50 to 110 GeV/c^2 in section 8.3.2.

8.3.1 Example: Excluding $\sigma_{\mathrm{ZH(80)}}$ for a Higgs decaying into $\mathrm{s\bar{s}}$

The ZH(80) event-by-event probability distribution for the SM background and the ZH(80) signal is used to construct the weights of individual events w_i=($1 + s_i/b_i$) that are subsequently used to construct the test-statistic for a large number of 'gedanken' background-only and signal+background experiments using expression (7.10). These distributions are shown in the left plot of figure 8.5 for background-only and signal+background experiments for three different values of $\zeta_{\mathrm{ZH(80)}}$: 0.5, 1.0 and 1.5. The observed value in the data (X_{obs}) is also shown. As only the intrinsic separation between the two distributions is important here, and this section serves an example, the mean position of the distribution of the test-statistic for background-only experiments and X_{obs} are given an arbitrary offset and fixed. The distributions for the three different signal+background experiments (with a varying cross section) are then shown relative to this fixed distribution as an indication of the changing separation power.

For a median background experiment, the confidence in the signal $\mathrm{CL_s}$ is computed while varying the signal cross section (0< ζ_{ZH} <3). The distribution of this running $\mathrm{CL_s}$ for a median background-only experiment and the observation is shown in the right plot of figure 8.5. Cross sections for which $\mathrm{CL_s}$ is smaller than 0.05 are excluded at 95% CL and therefore the ZH(80) cross section measurement can be summarised as:

Expected: $\zeta_{\mathrm{ZH(80)}} < 0.49$ (at 95% CL)

Observed: $\zeta_{\mathrm{ZH(80)}} < 0.76$ (at 95% CL)

As there is an excess of events in the data compared to the SM expectation the constraint on the hadronic ZH(80) cross section is weaker than was expected. In addition, since the performance of the analysis for $\mathrm{s\bar{s}}$ and gluon pair decay is similar as can be seen in figure 8.4, the numbers quoted above also apply for the gluon pair final state and can therefore be interpreted as a general upper limit on a ZH(80) cross section where the Higgs decays hadronically.

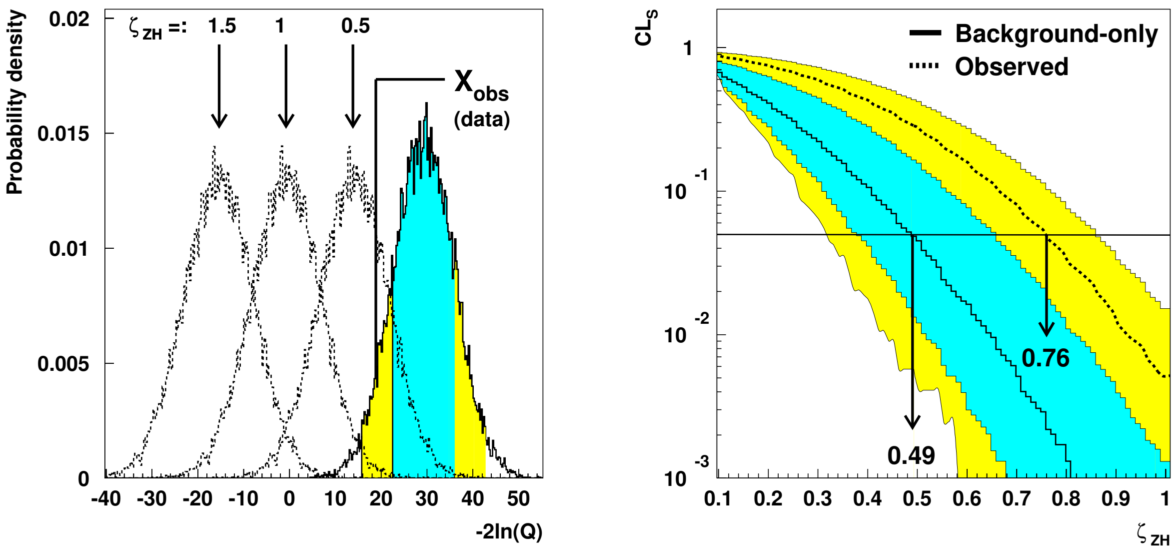


Figure 8.5: *The left plot shows the probability density function for the test-statistic for background-only (fixed) and background+signal (ZH(80)) experiments for three different signal cross sections. The shaded areas (in both plots) indicate the 1 and 2 sigma bands on the background-only. The (fixed) observed value for -2ln(Q) in the year 2000 data set is also indicated. The right plot shows the running confidence in the signal hypothesis CL_s as a function of the signal cross section. The expected and observed excluded (at 95%CL) ζ_{ZH} values are also shown.*

8.3.2 Hadronic cross section limit for Higgs masses from 50 to 110 GeV/c^2

For each Higgs mass hypothesis from 50 to 110 GeV/c^2 and for each decay mode, the expected and observed excluded hadronic ZH cross sections are determined as in the previous section. To set a conservative cross section limit for each Higgs mass, the worst limit obtained from the two decay modes is taken: below(above) 80 GeV/c^2 these are the exclusions from the $s\bar{s}$ (gluon pair) decay. This 'worst-case' distribution is shown in figure 8.6, summarising the full measurement by means of an excluded hadronic cross section region.

An additional interesting exercise is to compare the sensitivity relative to that of the SM Higgs boson. This is done by excluding a Higgs mass range when assuming a SM cross section:

If $\sigma_{ZH} = \sigma_{ZH}^{SM}$ then: Expected: $M_H > 103.9$ GeV/c^2 (at 95% CL)

Observed: $M_H > 106.9$ GeV/c^2 (at 95% CL)

Due to a deficit at Higgs mass hypotheses exceeding 100 GeV/c^2 the observed limit for a SM cross section is higher than expected. As mentionned before, for a range of light Higgs masses there is an excess observed in the data. At a small region of Higgs mass hypotheses around 70 GeV/c^2, the excess is such that not even a SM cross section could be excluded at 95% CL although the measurement is sensitive to cross sections half that size.

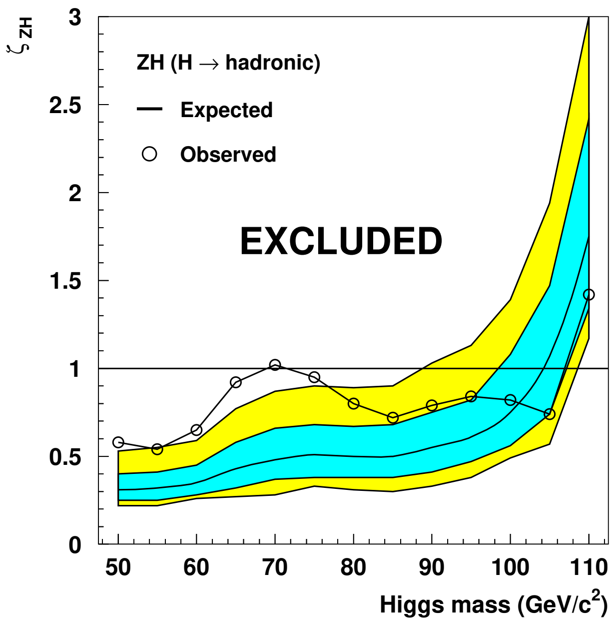


Figure 8.6: *This plot shows the most conservative expected and observed excluded cross section as a function of the Higgs mass for the fully hadronic final state in DELPHI.*

Systematic uncertainties for light Higgses

When interpreted as a ZH signal, the excess of events for Higgs mass hypotheses below 80 GeV/c^2 is most compatible with a ZH(70) signal whose cross section is almost half of what is expected for a SM Higgs boson. One should realise that for Higgs masses between 65 and 75 GeV/c^2 there is a huge almost irreducible background from W^+W^- events, causing the signal over background ratio ($\sim$ 0.1) to be smaller than at other masses. The effects from the uncertainties in the background processes on the excluded signal cross section are therefore also expected to be significantly larger.

When the $q\overline{q}(\gamma)$ background cross section is increased by 5(10)% the median expected excluded ζ_{ZH} increases only slightly from 0.48 to 0.49(0.50). The observed excluded scaling however goes down from 1.02 to 0.90(0.82). A similar effect is observed when the region of the ZH-probability distributions that is used in the evaluation of the exclusion limit is varied. When using the region ln(p_i)>-3.25 instead of ln(p_i)>-4.25 in the evaluation, the small loss in sensitivity causes the median expected excluded ζ_{ZH} to increase to 0.52, but again the observed exclusion goes down (0.86). In addition, the excess does not seem consistent between different centre-of-mass energies, although the statistics is too small to make a firm statement.

Concluding we can say that, although a large excess is observed in the data and studies are still ongoing, the effects discussed above, unfortunately hint more to a problem in the modelling of the $q\overline{q}(\gamma)$ background than to the discovery of a Higgs signal.

8.3.3 Combined DELPHI (and LEP) cross section limits

In addition to the hadronic decay of the Z boson also the other decay modes ($\nu\overline{\nu}$ and l^+l^-) are used in the DELPHI experiment to obtain maximum sensitivity in the search for a hadronically decaying Higgs boson. The different channels have a varying expected sensitivity as can be seen from their expected limits on the Higgs mass assuming a SM cross section: 81.4 GeV/c^2 for $\mathrm{q\overline{q}}l^+l^-$and 77.7 GeV/c^2 for $\mathrm{q\overline{q}}\nu\overline{\nu}$. The combined expected(observed) DELPHI Higgs mass exclusion potential[1] is shown in table 8.2 and is 108.8(109.6) GeV/c^2.

In table 8.2 also the Higgs mass exclusion potentials for the three other LEP experiments are given. The results from the four LEP experiments are also combined [93] and in figure 8.7 the preliminary combined result on excluded ZH(hadronic Higgs decay) cross sections is shown. The LEP experiments had the sensitivity to discover a Higgs signal for masses below 107 GeV/c^2 (mass for which the expected median 1-CL$_\mathrm{b}$ <5.7 10^{-7}), but no deviation from the SM expectation is observed. The observed lower limit on the Higgs mass of 112.9 GeV/c^2 is therefore very close to the expected one of 113.0 GeV/c^2.

Assuming $\sigma_{\mathrm{ZH}} = \sigma_{\mathrm{ZH}}^{\mathrm{SM}}$

	expected lower limit	observed lower limit
ALEPH	108.4	109.3
DELPHI	108.8	109.6
L3	109.3	111.6
OPAL	108.5	109.4
LEP	113.0	112.9

Table 8.2: *The preliminary expected and observed sensitivity (in GeV/c^2) for the mass of a hadronically decaying Higgs boson assuming its cross section is identical to the SM one.*

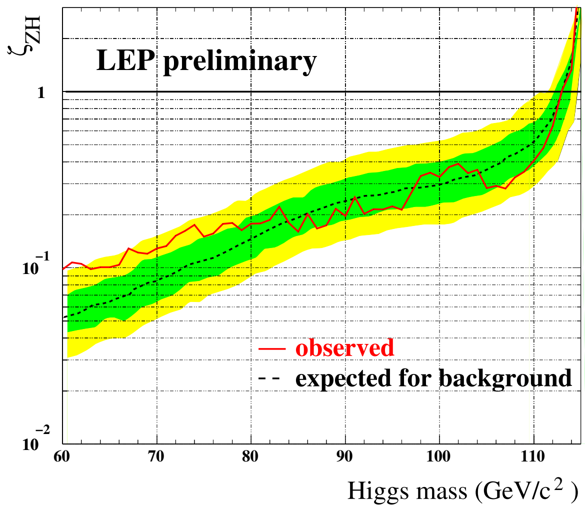


Figure 8.7: *This plot shows the LEP combined expected and observed excluded hadronic ZH cross section as a function of the Higgs mass.*

8.4 Summary and conclusions

A search for a hadronically decaying Higgs boson pair produced with a Z boson has been presented in the channel where the Z boson also decays hadronically. The results from the search have been given as a cross section limit for Higgs masses between 50 and 110 GeV/c^2. In the DELPHI experiment, in the fully hadronic final state, an excess of events compared to SM expec-

[1] Assuming the Higgs decays hadronically and the ZH cross section is similar to that of the SM Higgs.

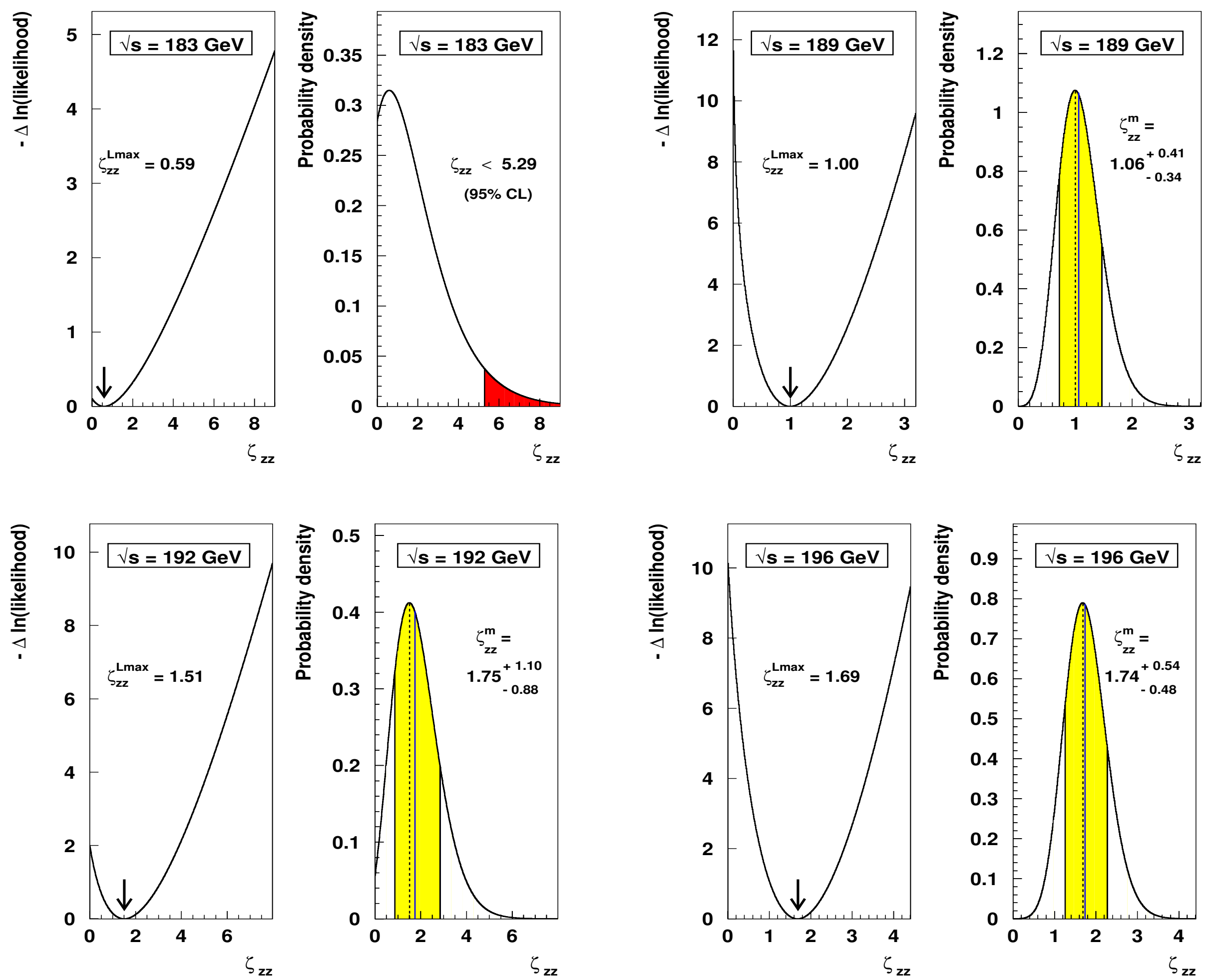
tations has been observed for all Higgs mass hypotheses below 80 GeV/c^2. The more sensitive LEP combination, using various final states, shows no significant deviations from the SM expectation and the expected and observed lower limit on the Higgs mass are very close: 113.0 and 112.9 GeV/c^2 respectively. These limits are only slightly ($\sim$ 2 GeV/c^2) lower than the ones obtained for the SM Higgs assuming the predominant decay into $\mathrm{b}\overline{\mathrm{b}}$. It is important to note that with these measurements more information is extracted from the LEP2 data set by placing (stringent) model independent limits on Higgs cross sections.

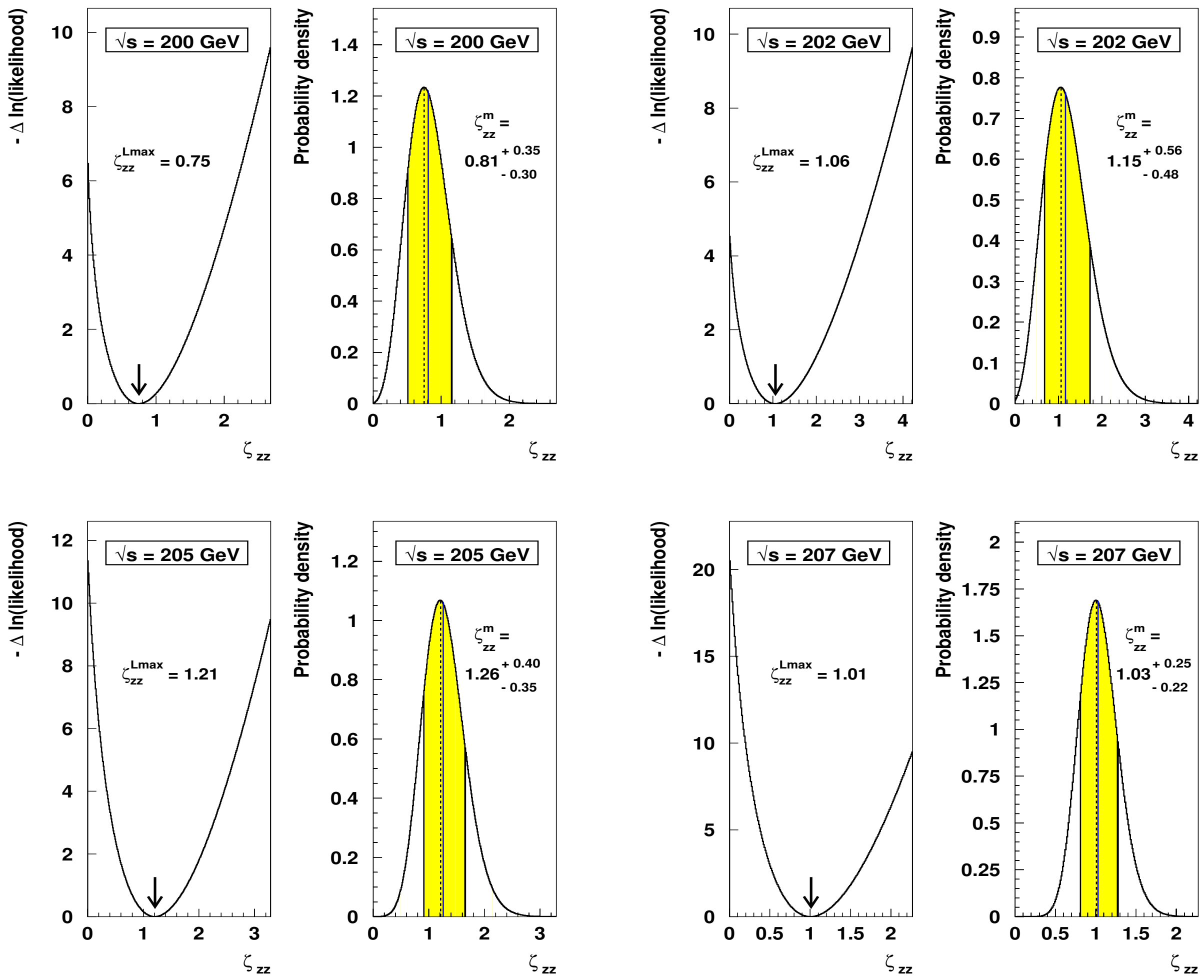
Appendix A

Cross section likelihood curves

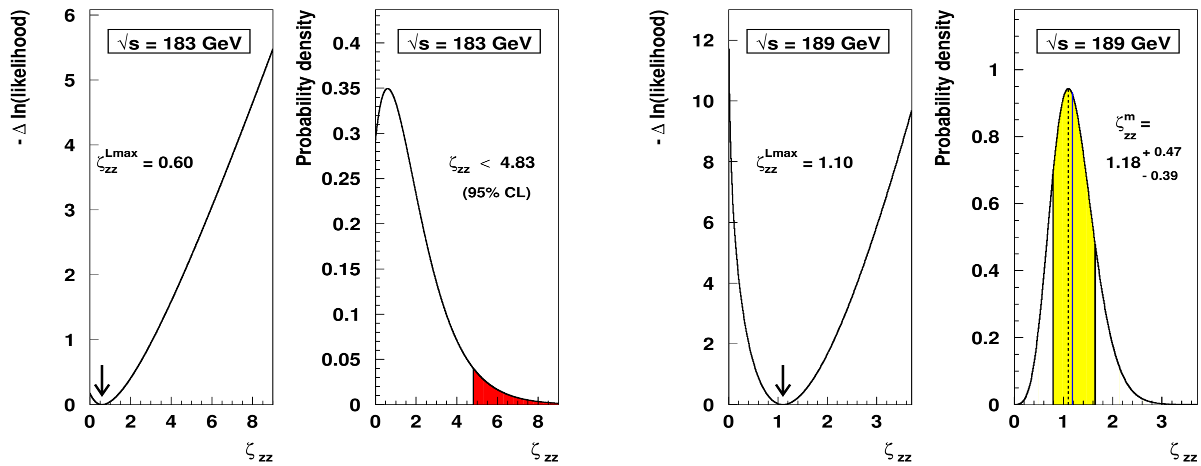
The full likelihood curves and the corresponding probability density functions for ζ_{zz} from the ZZ (4 quarks) and ZZ($b\bar{b}q\bar{q}$) cross section fits at all centre-of-mass energies.

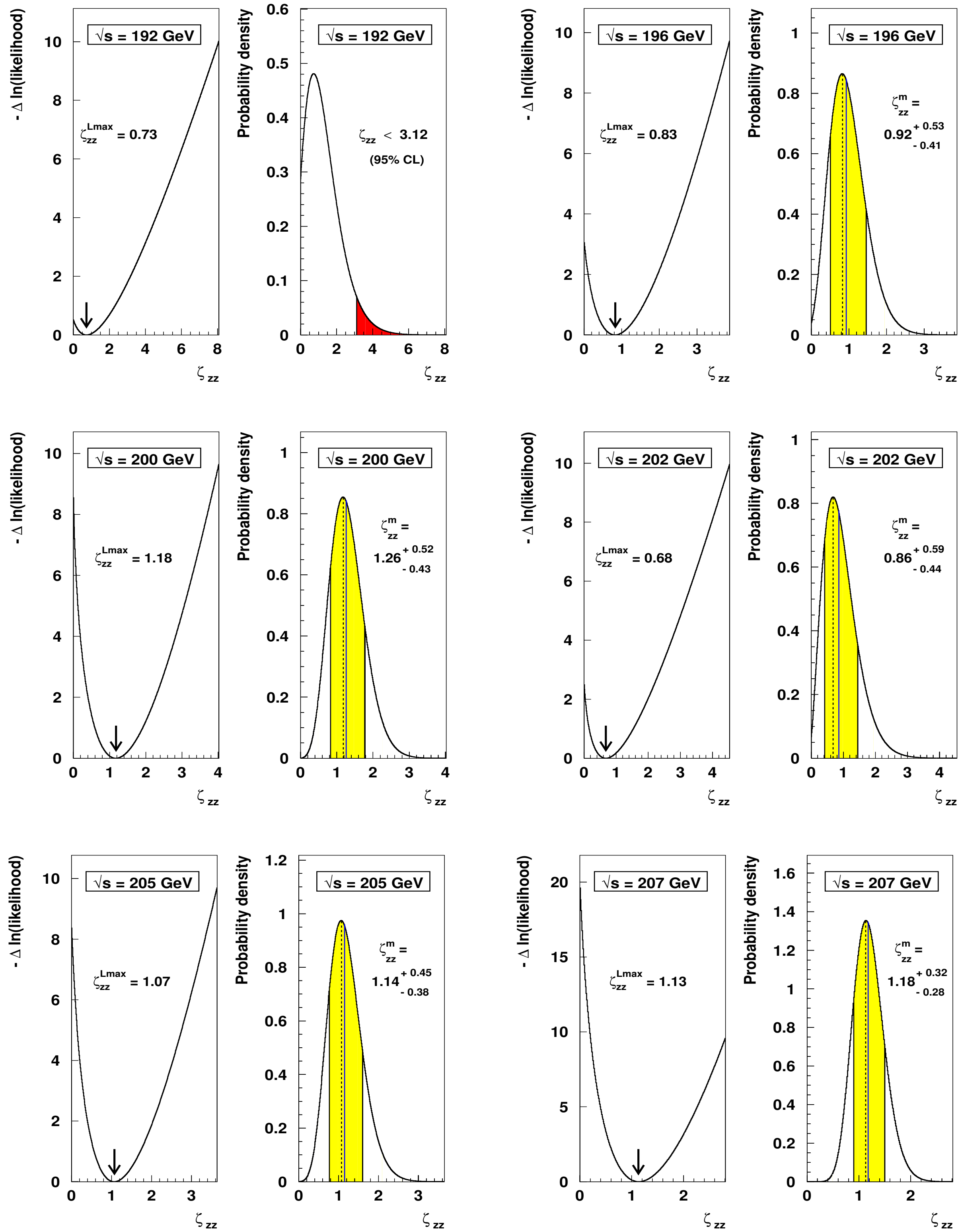
A.1 Details from ZZ (4 quarks) cross section measurements





A.2 Details from ZZ ($b\bar{b}q\bar{q}$) cross section measurements




- Δ ln(likelihood)
Probability density
ζ zz
√s = 192 GeV
ζ_zz^Lmax = 0.73
ζ_zz < 3.12
(95% CL)
√s = 196 GeV
ζ_zz^Lmax = 0.83
ζ_zz^m = 0.92 + 0.53 - 0.41
√s = 200 GeV
ζ_zz^Lmax = 1.18
ζ_zz^m = 1.26 + 0.52 - 0.43
√s = 202 GeV
ζ_zz^Lmax = 0.68
ζ_zz^m = 0.86 + 0.59 - 0.44
√s = 205 GeV
ζ_zz^Lmax = 1.07
ζ_zz^m = 1.14 + 0.45 - 0.38
√s = 207 GeV
ζ_zz^Lmax = 1.13
ζ_zz^m = 1.18 + 0.32 - 0.28

References

[1] S. Glashow, *Partial symmetries of weak interactions*. Nucl. Phys. **22**, 579 (1961).

[2] S. Weinberg, *A model of leptons*. Phys. Rev. Lett. **19**, 1264 (1967).

[3] A. Salam, *Weak and electromagnetic interactions*. Originally printed in Svartholm: Elementary Particle Theory, proceedings of the Nobel symposium held in 1968 at Lerum, Sweden, Stockholm 1968, 367-377.

[4] F. Halzen and A. Martin, *Quarks & Leptons*. John Wiley & Sons, Inc., 1984.

[5] C. Quigg, *Gauge Theories of the Strong, Weak and Electromagnetic Interactions*. The Benjamin/Cummings Publishing Company, Inc., 1983.

[6] K. Kodama et al., *Observation of tau-neutrino interactions*. Phys. Lett. **B504**, 218 (2001), hep-ex/0012035.

[7] D. Groom et al., *Review of particle physics*. Eur. Phys. J. **C15**, 1 (2000).

[8] G. 't Hooft, *Renormalizable lagrangians for massive Yang-Mills fields*. Nucl. Phys. **B35**, 167 (1971).

[9] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*. Nucl. Phys. **B44**, 189 (1972).

[10] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*. Phys. Rev. Lett. **13**, 321 (1964).

[11] P. Higgs, *Broken symmetries, massless particles and gauge fields*. Phys. Lett. **12**, 132 (1964).

[12] G. Guralnik, C. Hagen and T. Kibble, *Global conservation laws and massless particles*. Phys. Rev. Lett. **13**, 585 (1964).

[13] J. Goldstone, *Field theories with 'superconductor' solutions*. Nuovo Cim. **19**, 154 (1961).

[14] J. Gunion, H. Haber, G. Kane and S. Dawson, *The Higgs Hunter's Guide*. Addison-Wesley Publishing Company, 1990.

[15] C. Ford, D. Jones, P. Stephenson and M. Einhorn, *The effective potential and the renormalization group*. Nucl. Phys. **B395**, 17 (1993), hep-lat/9210033.

[16] S. Dawson, *Introduction to electroweak symmetry breaking*, 1998. hep-ph/9901280.

[17] T. Hambye and K. Riesselmann, *Matching conditions and Higgs mass upper bounds revisited*. Phys. Rev. **D55** (1997), hep-ph/9610272.

[18] G. Altarelli and G. Isidori, *Lower limit on the Higgs mass in the Standard Model: An update*. Phys. Lett. **B337** (1994).

[19] J. Casas, J. Espinosa and M. Quiros, *Standard Model stability bounds for new physics within LHC reach*. Phys. Lett. **B382**, 374 (1996), hep-ph/9603227.

[20] The LEP Collaborations, the LEP Electroweak Working Group, the SLD Heavy Flavour, Electroweak Working Group, *A Combination of Preliminary Electroweak Measurements and Constraints on the Standard Model*, 2001. LEPEWWG/2001-02,hep-ex/0112021.

[21] L. Demortier et al., *Combining the top quark mass results for run 1 from CDF and D0*, 1999. FERMILAB-TM-2084.

[22] J. Bai et al., *Measurements of the cross section for* $\mathrm{e^+e^-} \rightarrow$ *hadrons at center-of-mass energies from 2-GeV to 5-GeV*, 2001. hep-ex/0102003.

[23] H. Burkhardt and B. Pietrzyk, *Update of the hadronic contribution to the QED vacuum polarization*. Phys. Lett. **B513**, 46 (2001).

[24] A. Martin, J. Outhwaite and M. Ryskin, *A new determination of the QED coupling* $\alpha(\mathrm{m_z^2})$ *lets the Higgs off the hook*. Phys. Lett. **B492**, 69 (2000), hep-ph/0008078.

[25] B. Lee, C. Quigg and H. Thacker, *Weak interactions at very high-energies: the role of the Higgs boson mass*. Phys. Rev. **D16** (1977).

[26] G. Altarelli et al., *Physics at LEP2*, 1995. CERN Yellow Report 96-01, Volume 1.

[27] F. Berends, W. van Neerven and G. Burgers, *Higher order radiative corrections at LEP energies*. Nucl. Phys. **B297**, 429 (1988).

[28] E. Gross, G. Wolf and B. Kniehl, *Production and decay of the Standard Model Higgs boson at LEP200*. Z. Phys. **C63**, 417 (1994), hep-ph/9404220. Erratum-ibid.C66, 321-322 (1995).

[29] L. Resnick, M. Sundaresan and P. Watson, *Is there a light scalar boson ?* Phys. Rev. **D8**, 172 (1973).

[30] E. Braaten and J. Leveille, *Higgs boson decay and the running mass*. Phys. Rev. **D22**, 715 (1980).

[31] A. Kataev and V. Kim, *The effects of the QCD corrections to* $\Gamma(H0 \rightarrow \mathrm{b\bar{b}})$. Mod. Phys. Lett. **A9**, 1309 (1994).

[32] A. Djouadi, M. Spira and P. Zerwas, *QCD corrections to hadronic Higgs decays*. Z. Phys. **C70**, 427 (1996), hep-ph/9511344.

[33] W. Keung and W. Marciano, *Higgs scalar decays: $H \to W^{\pm}X$.* Phys. Rev. **D30**, 248 (1984).

[34] F. A. Berends and R. Pittau and R. Kleiss, *Excalibur: A Monte Carlo program to evaluate all four fermion processes at LEP-200 and beyond.* Comput. Phys. Commun. **85**, 437 (1995), hep-ph/9409326.

[35] A. Denner et al., *Electroweak radiative corrections to* $\mathrm{e}^+\mathrm{e}^- \to W^+W^- \to$ *4fermions in double-pole approximation: The RACOONWW approach.* Nucl. Phys. **B587**, 67 (2000), hep-ph/0006307.

[36] S. Jadach et al., *The Monte Carlo event generator YFSWW3 version 1.16 for W pair production and decay at LEP2/LC energies.* Comput. Phys. Commun. **140**, 432 (2001), hep-ph/0103163.

[37] R. Brown and K. Mikaelian, *W^+W^- and Z^0Z^0 pair production in* $\mathrm{e}^+\mathrm{e}^-$ *, p p, p anti-p colliding beams.* Phys. Rev. **D19**, 922 (1979).

[38] K. Gaemers and G. Gounaris, *Polarization amplitudes for* $\mathrm{e}^+\mathrm{e}^- \to W^+W^-$ *and* $\mathrm{e}^+\mathrm{e}^- \to$ *ZZ.* Zeit. Phys. **C1**, 259 (1979).

[39] M.W. Grunewald, et al., *Four-fermion production in electron positron collisions*, hep-ph/0005309,2000. Contribution to the LEP2 Monte Carlo Workshop: *Report of the Working Groups on Precision Calculations for LEP2 Physics*, 2000, CERN Yellow Report, CERN-2000-009.

[40] S. Jadach, W. Placzek and B. F. L. Ward, *Gauge invariant YFS exponentiation of (un)stable Z pair production at and beyond LEP2 energies.* Phys. Rev. **D56**, 6939 (1997), hep-ph/9705430.

[41] T. Sjöstrand, *High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4.* Comput. Phys. Commun. **82**, 74 (1994).

[42] S. Myers, E. Picasso, *The design, construction and commissioning of the CERN large electron positron collider.* Contemp. Phys. **31**, 387 (1990).

[43] ALEPH Collaboration (D. Decamp et al.), *ALEPH: a detector for electron - positron annihilations at LEP.* Nucl. Instrum. Meth. **A294**, 121 (1990).

[44] DELPHI Collaboration (P. Aarnio et al.), *The DELPHI detector at LEP.* Nucl. Instrum. Meth. **A303**, 233 (1991).

[45] L3 Collaboration, *The construction of the L3 experiment.* Nucl. Instrum. Meth. **A289**, 35 (1990).

[46] OPAL Collaboration (K. Ahmet et al.), *The OPAL detector at LEP.* Nucl. Instrum. Meth. **A305**, 275 (1991).

[47] P. Janot, *Priorities for LEP in 2000* . Writeup of talk at the Chamonix-X Workshop, January 2001, http://home.cern.ch/janot/.

[48] DELPHI Collaboration (P. Abreu et al.), *Performance of the DELPHI detector*. Nucl. Instrum. Meth. **A378**, 57 (1996).

[49] P. Chochula et al. (DELPHI Silicon Tracker Group), *The DELPHI silicon tracker at LEP2*. Nucl. Instrum. Meth. **A412**, 304 (1998).

[50] DELPHI Collaboration, *DELPHI Data Analysis Program (DELANA) User's Guide* . DELPHI-89-44, internal DELPHI-note, 1989.

[51] W.J. Murray, *Measurement of the beam position in DELPHI* . DELPHI-96-6, internal DELPHI-note, 1996.

[52] V. Bocci et al. (DELPHI Trigger Group), *Architecture and performance of the DELPHI trigger system*. Nucl. Instrum. Meth. **A362**, 361 (1995).

[53] F. Cossutti et al., *Improvements to SKELANA for Version 2.0* . DELPHI-99-175, internal DELPHI-note, 1999.

[54] DELPHI Collaboration, *DELSIM, DELPHI event generation and detector simulation-Reference Manual* . DELPHI-89-68, internal DELPHI-note, 1989.

[55] S. Bethke et al., *Experimental investigation of the energy dependence of the strong coupling strength*. Phys. Lett. **B213**, 235 (1988).

[56] B. R. Webber, *Jets in perturbation theory* . Talk given at Workshop on QCD: 20 Years Later, Aachen, Germany, June 1992.

[57] S. Catani and Yu. L. Dokshitzer and M. Olsson and G. Turnock and B. R. Webber, *New clustering algorithm for multi - jet cross-sections in* $\mathrm{e}^+\mathrm{e}^-$*annihilation*. Phys. Lett. **B269**, 432 (1991).

[58] S. Moretti, L. Lonnblad and T. Sjöstrand, *New and old jet clustering algorithms for electron positron events*. JHEP **08**, 001 (1998), hep-ph/9804296.

[59] S. Bethke, *Jets in* Z^0 *decays* . Talk given at Workshop on QCD: 20 Years Later, Aachen, Germany, June 1992.

[60] N. Kjær and R. Møller, *Reconstruction of Invariant Masses in Multi-Jet Events* . DELPHI-91-17, internal DELPHI-note, 1991.

[61] M. Mulders, *Direct measurement of the W boson mass in* $\mathrm{e}^+\mathrm{e}^-$*collisions at LEP*. Ph.D. Thesis, University of Amsterdam, 2001.

[62] P. Abreu et al., *The estimation of the effective centre of mass energy in* $q\overline{q}(\gamma)$ *events from DELPHI*. Nucl. Instrum. Meth. **A427**, 487 (1999), hep-ex/9809008.

[63] G. Borisov, *Combined b-tagging*. Nucl. Instrum. Meth. **A417**, 384 (1998).

[64] G. Borisov and C. Mariotti, *Fine tuning of track impact parameter resolution of the DELPHI detector*. Nucl. Instrum. Meth. **A372**, 181 (1996).

[65] ALEPH Collaboration (D. Buskulic et al.), *A precise measurement of* $\Gamma(Z \to b\bar{b})/\Gamma(Z \to$ *hadrons)*. Phys. Lett. **B313**, 535 (1993).

[66] P. Abreu and others, *A precise measurement of the partial decay width ratio* $R_b^0 = \Gamma_{b\bar{b}}/\Gamma_{had}$. Eur. Phys. J. **C10**, 415 (1999).

[67] DELPHI Collaboration (P. Abreu et al.), *Hadronization properties of b quarks compared to light quarks in* $\mathrm{e}^+\mathrm{e}^- \to q\overline{q}$ *from 183-GeV to 200-GeV*. Phys. Lett. **B479**, 118 (2000), hep-ex/0103022.

[68] M. Boonekamp, *b-tagging with high* $\mathrm{p_T}$ *leptons* . DELPHI-98-54, internal DELPHI-note, 1998.

[69] E. Brodet, P. Bruckman, P. Collins, M. Elsing, T. Lesiak, W. Liebig, G. Wilkinson, *A measurement of the Cross Section Ratio* R_b *and the Forward-Backward Asymmetry* $\mathrm{A_{FB}^b}$ *for* $\mathrm{b\overline{b}}$ *Events with the DELPHI Detector at LEP2* . DELPHI-2001-016, internal DELPHI-note, 2001.

[70] A. Ballestrero et al., *Report of the QCD Working Group*, hep-ph/0006259,2000. Contribution to the LEP2 Monte Carlo Workshop: *Report of the Working Groups on Precision Calculations for LEP2 Physics*, 2000, CERN Yellow Report, CERN-2000-009.

[71] DELPHI Collaboration (P. Abreu et al.), *Search for the Standard Model Higgs boson at LEP in the year 2000*. Phys. Lett. **B499**, 23 (2001), hep-ex/0102036.

[72] T. Bayes, *An Essay Toward Solving a Problem in the Doctrine of Chances*, 1763. Philos. Trans. R. Soc. London 53, page 370-418.

[73] F. James, Y. Perrin and L. Lyons (editors.), *Confidence limits. Proceedings, 1st Workshop, Geneva, Switzerland, January 17-18, 2000*. CERN-2000-005.

[74] P. Bambade et al., *Study of Trilinear Neutral Gauge Boson Couplings*. Abstract 346, International Europhysics Conference on High Energy Physics July 2001, Budapest, Hungary. DELPHI-note 2001-097 CONF 525.

[75] V. Verzi, *Study of the anomalous couplings ZZZ, ZZ*γ *and Z*$\gamma\gamma$ *at LEP with the DELPHI detector*. Ph.D. Thesis, University of Milano, 2000.

[76] P. Abreu, et al., *Measurement of the ZZ cross-section in* $\mathrm{e}^+\mathrm{e}^-$*interactions at 183-GeV - 189-GeV*. Phys. Lett. **B497**, 199 (2001).

[77] G. Borisov et al., *Update of ZZ Production Measurements in* $\mathrm{e}^+\mathrm{e}^-$*interactions using data at 192-202 GeV*. Abstract 659, paper 2, XXXth International Conference on High Energy Physics, July 2000, Osaka, Japan. DELPHI note 2000-145 OSAKA CONF 444.

[78] G. Borisov et al., *Update of the ZZ cross-section measurement in* $\mathrm{e}^+\mathrm{e}^-$*interactions using data collected in 2000*. Abstract 354, International Europhysics Conference on High Energy Physics, July 2001, Budapest, Hungary. DELPHI note 2001-105 CONF 533.

[79] Aleph, Delphi, L3 and Opal collaborations, *LEP W-pair, Z-pair and Single W Cross Section Results for the Summer 2001 Conferences*, 2001. LEPEWWG/XSEC note 2001-03.

[80] A.L. Read, *Modified frequentist analysis of search results (The CL_s method)*, 2000. CERN-OPEN-2000-205.

[81] S. Jin, *The signal estimator limit setting method.* Nucl. Phys. **A675**, 88c (2000).

[82] R. D. Cousins and V.L. Highland, *Incorporating systematic uncertainties into an upper limit.* Nucl. Instrum. Meth. **A320**, 331 (1992).

[83] LEP Higgs Working Group for Higgs boson searches, *Search for the Standard Model Higgs boson at LEP*, 2001, hep-ex/0107029.

[84] J. Erler, *The probability density of the Higgs boson mass.* Phys. Rev. **D63**, 071301 (2001), hep-ph/0010153.

[85] LEP Electroweak working group and the SLD Heavy Flavour and Electroweak Working Group, *A Combination of Preliminary Electroweak Measurements and Constraints on the Standard Model*, 2001, CERN-EP-2001-021, hep-ex/0103048.

[86] M. Carena et al., *Report of the Tevatron Higgs working group*, 2000. hep-ph/0010338.

[87] S. Holmes, *talk given at the October 1999 meeting of the International Committee for Future Accelerators (ICFA).* See also http://conferences.fnal.gov/icfa99/icfasem_agenda_3oct.htm.

[88] ATLAS Collaboration, *ATLAS detector and physics performance, Technical Design Report Vol. 2.* CERN-LHCC-99-15.

[89] CMS Collaboration, *Technical proposal.* CERN-LHCC-94-38.

[90] X. Calmet and H. Fritzsch, *The Higgs boson might not couple to b quarks.* Phys. Lett. **B496**, 190 (2000), hep-ph/0008252.

[91] P. Janot, *The HZHA generator in: "Physics at LEP2"*, 1996. CERN 96-01, Volume 2, page 309.

[92] P. Abreu et al., *Energy dependence of the differences between the quark and gluon jet fragmentation.* Z. Phys. **C70** (1996).

[93] Aleph, Delphi, L3 and Opal collaborations, *Flavour Independent Search for Hadronically Decaying Neutral Higgs Bosons at LEP*, 2001. LHWG note 2001-07,hep-ex/0107034.

Summary

In this thesis an analysis is presented that studies high energy e^+e^- collisions in which four (or more) jets of particles are created. These collisions are produced by the Large Electron Positron collider (LEP) at CERN with centre-of-mass energies ranging from 183 to 209 GeV and correspond to a total integrated luminosity of 665 pb^{-1}. The particles that are produced in these collisions are recorded by the DELPHI detector.

The main goal of this thesis is to search for the possible presence of interactions (events) that can be attributed to the last missing particle of the Standard Model, the Higgs boson. This particle is a crucial ingredient in the model that describes the world of elementary particles as it not only explains how particles acquire mass, but in addition ensures that the model can provide precise predictions for their interactions. Although the mass of the Higgs boson itself is not predicted, all its further properties are. At LEP, if kinematically allowed, the Higgs boson is mainly produced together with a Z boson and in this thesis we are interested in the events in which both the Z boson and the Higgs boson decay into a quark anti-quark pair, which results in four quarks in the final state: $e^+e^- \rightarrow ZH \rightarrow q\bar{q}q\bar{q}$. These fully hadronic Higgs events ($\sim$ 2/3 of all possible final states) have very distinct properties. Unfortunately, even if the Higgs is produced, only a small number of ZH events are expected, with the additional complication that the heavier the Higgs boson is, the less events are produced. The main challenge is therefore to isolate possible candidates from the large sample of 'background' multi-jet events.

An analysis is presented that is sensitive to both Higgs events and other Standard Model processes that produce four quarks in the final state. This is achieved by comparing, on an event-by-event basis, the experimental information extracted from the event with what is expected from each of these different event types. In this comparison the predicted properties are used analytically where possible which allows to optimally take into account the characteristics of the Higgs boson as a function of its (unknown) mass. For each event a compatibility measure with all hypotheses is computed. The analysis presented in this thesis differs from other Higgs analyses mainly in the way the mass information is treated and in the method that is used to combine event information from various levels of complexity. The main advantage of this method is that the inherent ambiguity in multi-jet events connected to the pairing of jets does not need to be resolved. This allows the construction of a more complete and consistent compatibility measure with Higgs events (and other hypotheses).

The process of Z boson pair production where both Z bosons decay into a quark anti-quark pair ($e^+e^- \rightarrow ZZ \rightarrow q\bar{q}q\bar{q}$) is very similar both in experimental signature and cross section to that of Higgs events. Although an interesting process in itself, it is a background in the search for the

Higgs boson and therefore an ideal process to 'calibrate' the analysis method. The cross section for this process has been measured at all LEP2 centre-of-mass energies and is found to be in good agreement with the Standard Model expectation. The related measurement of the (small) fully hadronic ZZ cross section in which there are at least two b-quarks present ($\sim$ 0.2 pb) is summarised by the good agreement between measurement and expectation when the combined LEP2 data set is used:

$$\sigma^{\mathrm{measured}}_{\mathrm{NC02}}\ (\mathrm{b\bar{b}q\bar{q}})/\sigma^{\mathrm{SM}}_{\mathrm{NC02}}\ (\mathrm{b\bar{b}q\bar{q}})\ =\ 1.05\ ^{+0.17}_{-0.15}(\mathrm{stat})$$

This result shows that the analysis is indeed sensitive to a signal that is very similar to ZH events. However, in the data set that was analysed, no evidence for the presence of the Standard Model Higgs boson was found. Since the cross section of ZH production gets smaller as the Higgs mass increases, at very large Higgs masses the analysis is no longer sensitive to a possible Higgs signal and the analysis can no longer exclude the presence of such a (small) signal. The search for the Higgs boson is summarised by the largest mass of the Higgs boson that can still be excluded at 95% confidence level. The expected and observed lower limit on the mass of the Standard Model Higgs boson in the 4-quark channel are:

$$\text{Expected:}\quad \mathrm{M_H}\ > 112.0\ \ \mathrm{GeV/c^2}\quad \text{(at 95\% CL)}$$

$$\text{Observed:}\quad \mathrm{M_H}\ > 112.2\ \ \mathrm{GeV/c^2}\quad \text{(at 95\% CL)}$$

The clearest signature of a ZH event is the presence of a pair of b-quarks in the final state as this is the dominant decay mode of a Standard Model Higgs boson and jets of particles originating from the fragmentation of a b-quark can be clearly identified experimentally. There are, however, also models in which the Higgs boson coupling to b-quarks is significantly lower than in the Standard Model and the Higgs might decay exclusively to a pair of light quarks or gluons. To be able to confront specific models with the data from $\mathrm{e^+e^-}$ collisions, an upper limit on the cross section is derived for a scalar object that decays hadronically and that is produced together with a Z boson. These cross section limits are presented as a function of the mass of this scalar particle and present a more complete and model independent analysis of the collected multi-jet event sample. If the cross section is assumed to be identical to that in the Standard Model, the measurement can also be translated into a lower limit on the mass of such a scalar particle. Using the analysis presented in this thesis an expected(observed) lower limit of 103.9(106.9) GeV/c^2 is obtained at 95% CL.

As the LEP experiments can not claim to have observed the Higgs boson, the search for this elusive particle continues. With the new proton (anti-)proton colliders at Fermilab and CERN it is very likely that, if the Higgs boson exists, it will be discovered in the coming five to eight years.

Samenvatting

In dit proefschrift wordt een analyse beschreven van elektron-positron botsingen bij hoge energie, waarbij vier (of meer) sproeiers van deeltjes ('jets') gecreëerd worden. Deze botsingen, met een zwaartepuntsenergie tussen de 183 en 209 GeV, worden bewerkstelligd door de grote Europese deeltjesversneller LEP bij CERN in Genève en de deeltjes die in deze botsingen worden geproduceerd worden gedetecteerd door de DELPHI detector. De geïntegreerde luminositeit is 665 pb^{-1}.

Het doel van dit proefschrift is om botsingen (gebeurtenissen) te ontdekken die toegeschreven kunnen worden aan het enige nog ontbrekende deeltje in het Standaard Model, het Higgs boson. Dit deeltje speelt een cruciale rol in dit model, dat de wereld van de elementaire deeltjes beschrijft. Het Higgs deeltje verklaart niet alleen hoe deeltjes hun massa verkrijgen, maar bovendien zorgt het ervoor dat met dit model nauwkeurige voorspellingen kunnen worden gedaan over hun gedrag. Hoewel de massa van het Higgs boson niet voorspeld wordt in het model, worden verdere eigenschappen ervan dat wel. Als het kinematisch mogelijk is, wordt bij LEP het Higgs boson voornamelijk samen met een Z boson geproduceerd. In dit proefschrift zijn we vooral geïnteresseerd in de gebeurtenissen waarbij zowel het Z boson als het Higgs boson in een quark anti-quark paar uiteenvalt en waarbij er dus vier quarks in de eindtoestand aanwezig zijn: $\mathrm{e^+e^-} \rightarrow \mathrm{ZH} \rightarrow \mathrm{q\overline{q}q\overline{q}}$. Deze volledig hadronische Higgs gebeurtenissen ($\sim$ 2/3 van alle mogelijke eindtoestanden) hebben zeer specifieke eigenschappen. Helaas, zelfs als de Higgs geproduceerd wordt, verwachten we slechts een klein aantal van deze gebeurtenissen. Daarbij komt nog de extra moeilijkheid dat hoe zwaarder het Higgs boson is, hoe minder van deze speciale botsingen er plaatsvinden. De voornaamste uitdaging is dan ook om kandidaatgebeurtenissen te selecteren uit een groot aantal 'normale' multi-jet botsingen.

Een analyse wordt gepresenteerd die tegelijkertijd gevoelig is voor Higgs gebeurtenissen en voor andere Standaard Model processen die vier quarks produceren. Dit wordt bereikt door, voor iedere gebeurtenis apart, de eigenschappen die experimenteel zijn gemeten te vergelijken met wat verwacht wordt voor de verschillende processen. In deze vergelijking worden de verwachte eigenschappen waar mogelijk analytisch beschreven, hetgeen het mogelijk maakt om optimaal de veranderende eigenschappen van het Higgs boson als functie van de (onbekende) massa mee te nemen in de analyse. Voor elke gebeurtenis wordt een compatibiliteitsmaat met elk van de mogelijke hypotheses vastgesteld. De beschreven analyse-methode verschilt van andere Higgs analyses voornamelijk in de manier waarop de massa-informatie gebruikt wordt en de manier waarop experimentele informatie met een zeer uiteenlopende mate van complexiteit gecombineerd wordt. Het voornaamste voordeel van deze methode is dat de inherente ambiguïteit die in multi-jet botsingen aanwezig is bij het paarsgewijs samenvoegen van de jets niet opgelost hoeft te worden. Dit maakt het mogelijk een vollediger en consistenter beeld van een botsing te verkrijgen en daarmee gevoeliger te zijn voor Higgs (en andere) gebeurtenissen.

Z boson paarproductie, waarbij beide Z bosonen in een quark anti-quark paar uiteenvallen ($e^+e^- \rightarrow ZZ \rightarrow q\overline{q}q\overline{q}$) is zowel in experimenteel opzicht als in werkzame doorsnede vergelijkbaar met Higgs gebeurtenissen. Dit proces is daardoor uitermate geschikt om de analyse-methode te 'calibreren', hoewel het voor de zoektocht naar het Higgs boson als achtergrond wordt gezien. De werkzame doorsnede voor dit proces is voor elke LEP2 zwaartepuntsenergie apart bepaald en er is vastgesteld dat deze zeer goed in overeenstemming is met de voorspelling vanuit het Standaard Model. De (gerelateerde) meting van de volledig hadronische werkzame doorsnede van ZZ productie, waarbij er minstens twee b-quarks in de eindtoestand aanwezig zijn (~ 0.2 pb), toont een goede overeenstemming tussen de voorspelling en de meting wanneer de volledige LEP2 data-set gebruikt wordt:

$$\sigma^{\mathrm{gemeten}}_{\mathrm{NC02}}\ (\mathrm{b\overline{b}q\overline{q}})/\sigma^{\mathrm{SM}}_{\mathrm{NC02}}\ (\mathrm{b\overline{b}q\overline{q}})\ =\ 1.05\ ^{+0.17}_{-0.15}(\mathrm{stat})$$

Dit resultaat laat zien dat de analyse inderdaad gevoelig is voor gebeurtenissen die sterk op ZH botsingen lijken. In de data die geanalyseerd zijn is echter geen aanwijzing gevonden voor de aanwezigheid van een extra signaal afkomstig van een Standaard Model Higgs boson. Omdat de werkzame doorsnede voor ZH productie kleiner is voor een Higgs boson dat zwaarder is, is de analyse niet langer gevoelig voor een Higgs signaal als de Higgs massa erg groot is. Er kan dan niet meer met zekerheid gezegd worden dat zo'n (klein) signaal niet aanwezig is. De zoektocht naar het Higgs boson wordt daarom samengevat door de grootste Higgs massa te geven waarvan het signaal nog uitgesloten kan worden met 95% betrouwbaarheid. De verwachte en geobserveerde ondergrens op de massa van het Standaard Model Higgs boson in het 4-quark kanaal is:

Verwacht:	$\mathrm{M_H} > 112.0$ GeV/c^2	(met 95% betrouwbaarheid)
Geobserveerd:	$\mathrm{M_H} > 112.2$ GeV/c^2	(met 95% betrouwbaarheid)

Het duidelijkste kenmerk van een Higgs gebeurtenis is de aanwezigheid van twee b-quarks, omdat dit het dominante vervalskanaal van het Higgs boson is en omdat jets die afkomstig zijn van de fragmentatie van een b-quark experimenteel goed te identificeren zijn. Er bestaan echter ook modellen waarin de koppeling van het Higgs boson aan b-quarks aanzienlijk zwakker is dan in het Standaard Model en waarin het Higgs boson mogelijk uitsluitend vervalt in twee lichte quarks of gluonen. Om specifieke modellen met de resultaten uit e^+e^- botsingen te kunnen confronteren is een bovengrens bepaald op de werkzame doorsnede voor de productie van een scalair deeltje dat hadronisch vervalt en dat samen met een Z boson geproduceerd wordt. Deze limieten voor de werkzame doorsnede worden gepresenteerd als functie van de massa van dit deeltje. Hiermee wordt een vollediger en meer model-onafhankelijke samenvatting gegeven van de geobserveerde data. Als de werkzame doorsnede gelijk wordt verondersteld aan die van het Standaard Model Higgs boson kan de meting vertaald worden naar een ondergrens op de massa van zo'n deeltje. Een verwachte(geobserveerde) ondergrens van 103.9(106.9) GeV/c^2 is verkregen met 95% betrouwbaarheid.

De LEP-experimenten kunnen niet claimen dat ze een onomstotelijk bewijs voor het Higgs boson in de data hebben waargenomen en de zoektocht naar dit laatste ontbrekende deeltje gaat verder. Met de nieuwe proton (anti-)protonversnellers bij Fermilab en CERN is het zeer waarschijnlijk dat, wanneer het Higgs boson bestaat, het in de komende vijf tot acht jaar ontdekt zal worden.

Acknowledgements / Dankwoord

De afgelopen jaren zijn behoorlijk enerverend geweest en ik kijk er nu met veel plezier op terug. In deze periode heb ik veel mensen leren kennen die ieder op hun eigen manier hebben bijgedragen aan het tot stand komen van dit proefschrift. Het is onmogelijk iedereen te bedanken en dat ga ik dan ook niet doen. Een aantal mensen die (soms ongewild) dicht bij het hele proces zijn betrokken wil ik toch even persoonlijk bedanken.

Als eerste wil ik mijn promotor Jos Engelen bedanken. Niet alleen voor de zetjes in de rug tijdens het schrijven van dit proefschrift, maar vooral ook voor het vertrouwen dat hij me heeft gegeven tijdens mijn hele promotie. Daarnaast wil ik mijn co-promotor Jan Timmermans bedanken voor de mogelijkheid om mijn promotieonderzoek bij DELPHI te verrichten en voor zijn nauwgezet commentaar bij het doorlezen van eerdere versies van dit proefschrift.

The English part of the acknowledgement I start with a special thanks to Niels Kjær who introduced me to the many subtleties of a multi-jet analyses and who got me 'on track' with the likelihood ratio analysis. Although we have quite different personalities we worked (I think) very well together. I learned a lot and I hope you are also satisfied with the result. In this context I would also like to thank Martijn Mulders who has been a great help and 'vraagbaak' for the tools that have been developed within the W mass team and that have been used extensively in this thesis. The enthusiasm and energy of Philip Bambade have also influenced me. He has put in a large effort to understand the analysis and always suggested improvements during the discussions we had. It was also he who initialised the non-SM Higgs analysis group in DELPHI.

As a hardware task I worked (only for a small fraction of my time I must confess) in the RICH group of DELPHI. A group with a very friendly atmosphere and without the patience of Olav Ullaland, Erich Albrecht and Valerio Gracco I would certainly have 'snapped' when faced with one of the problems that for some reason almost always occur in the middle of the night.

Twee Nederlanders in DELPHI die ik wil bedanken zijn Jan van Eldik en André Augustinus. Als kamergenoot van Jan heb ik niet alleen mijn sporttrivia behoorlijk kunnen uitbreiden, maar zag ik ook alle problemen in de computing binnen DELPHI voorbijkomen en kon ik ze met zijn hulp mooi omzeilen. Als 'guinnea pig' kon ik daarnaast profiteren van de nieuwste aanwinsten in de grote trukendoos. André kan, als een van de weinige personen in de DELPHI collaboratie, in panieksituaties onwaarschijnlijk rustig blijven en heeft daarmee niet alleen mij persoonlijk, maar ook DELPHI meermaals gered.

Tijdens de reis over de woelige baren van de promotie deelde ik het schuitje met de heren van de PVDD (Maarten Bruinsma, Ernst-Jan Buis, Rutger van der Eijk, Wouter Hulsbergen en Niels Tuning) en we hebben in die tijd behoorlijk wat zin en onzin gedeeld. Delen van dit manuscript zijn ook gelezen door Aart Heijboer en Stefan Groot Nibbelink. Stefan heeft me een paar keer mooi door de mand laten vallen met de theorie, maar me daarna gelukkig ook weer opgevangen en door de discussies met Aart is het statistiek verhaal (mij) een stuk duidelijker geworden. Martin Blom heeft als mijn laatste kamergenoot het afgelopen jaar de voltooiing van dit proefschrift van dichtbij meegemaakt en hij heeft naast het voorkomen van een mentale instorting ook als een esthetische eenmanscommissie gefungeerd.

Zonder het dagelijkse E-mail contact met Jaap Voets was ik het contact met de echte wereld al lang verloren en ik heb genoeg voorbij zien komen om zeker te weten dat ik in ieder geval niet de raarste persoon op aarde ben. Eelco Schillings heeft veel van de stappen in mijn natuurkundetijd meegemaakt. Allebei ideaal om af en toe heerlijk tegenaan te zeuren en/of (bij Eelco meestal en) met een paar biertjes even de natuurkunde te vergeten. Heren, ik vind het een fijn idee dat jullie mijn paranimfen zijn.

Tenslotte wil ik mijn ouders bedanken die me altijd gestimuleerd en gesteund hebben in mijn leven en natuurlijk ook mijn kleine broertje Rody. Sjuul, jij hebt als enige de hele rit met alle pieken en dalen van dichtbij meegemaakt (en je houdt niet eens van achtbanen). Zonder jouw steun was het allemaal nooit gelukt.