# Observation of an excess in the search for the Standard Model Higgs boson using cuts based analyses at ALEPH

John Alan Kennedy

UNIVERSITY
*of*
GLASGOW

November 2001

A thesis submitted for the degree of Doctor of Philosophy of the University of Glasgow

# Abstract

In 2000, the final year of LEP running, the LEP accelerator produced its highest ever collision energies with centre of mass energies reaching 210 GeV. In total 217 $pb^{-1}$ of integrated luminosity were collected by the ALEPH detector. The collected data have been analysed using cuts based analyses for evidence of Standard Model Higgs boson production. An excess of high mass candidate events is observed, corresponding to a significance of $3.1\sigma$ above the standard model background expectation, with much of this excess originating in the four jets final state. The production of a Higgs boson with a mass of $\approx 115$ GeV$/c^2$ is favoured if the observed excess is attributed to the production of a Standard Model Higgs boson.

# Acknowledgements

Acknowledgements often seem to be the thing that you do last but they are quite probably the most important part of any thesis. I think I am who I am because of all the people who give me the strength to be me.

At Glasgow I'd like to thank the whole of the PPE group, especially those who managed to make the Friday 5 o'clock meetings in Tennents. I'd like to thank Jim Lynch and Peter Negus for supervising me and David Saxon for taking me as a student, I must also say thank you to the other staff for providing such a good group of people to work with. All the usual suspects from the Detector Development group and boys (even girls this year) from the lab deserve thanks for their friendship and constant bickering ab out football. Thanks, especially of late, to Keith Mathieson who has conned many a beer from this betting fools pocket and Gavin McPants who gave his own job so that I could be employed. Big Stan deserves much praise for still speaking to me and even mana ging to buy me a Budweiser when his many attempts to convert me to real ale failed.

I must thank Steven Devine for his constant (mental) friendship and also simply because he went through the same stress, pain and shite that I did, it's nice to know you're not alone when things get a little too much. Thanks also to the other guys in my y ear Will Bell and Steven Passmore for allowing me to win prizes.

Gordon Moorhouse deserves much thanks for his many attempts to teach a poor experimentalist about the theoretical (dark) side of particle physics.

Now to CERN. I met many people and have many friends (honest), most of who I'll prob forget to thank here. Thanks if you were at cern and put up with my company/shared a beer with me/once saw me in the corridor, it was good while it lasted. Thanks especia lly to Luke, Jerremy, B.Camanzi, Tim, Valeria and all others who once drank in an Irish bar in Geneva (believe me you need real (rich) friends to buy a round in Geneva).

The people I worked with on ALEPH and especially in the Higgs TaskForce made things a little easier and life a little better (generally at the BBQ's). Special thanks go to David Smith and Pedro Teixeira-Dias for their continued friendship and guidance thr ough my years as their understudy, attempting to keep your pace has improved me immensely. I also owe much to the Wisconsin group simply for being utter physics ninjas, helping us of lesser manpower and generally being a good bunch of blokes (blokets) to work with.

Two people who made a huge impact on my life and made me see I really had a life are Barbara and Tatjana. I can't put into words what you gave me, I only know that you both made me a much stronger, happier person and I will owe you for this always. Barbar a managed to teach me that we are supposed to stay out of school for those two strange days at the end of the week and Tatjana has been such a close, and loved, friend and grown to know me as well an anyone could (or would want to anyway).

I'd also like to thank Monique for being such a good friend for so many years and all the people in Hull and Manchester who made life worthwhile.

Finally I'd like to thank my family, my mother and father who have stood by me and often held me up as I was about to fall. My asymmetric sister and her ever increasing number of children (child 1,2 and 3) deserve thanks for making Christmas fun and gener ally turning home into a zoo.

If I missed you out then I'm sorry, I promise it isn't personal.

johnk - Nov 2001 - Glasgow

# Declaration

Except where explicit reference is made to the work of others, this dissertation is the result of my own work. None of this material has been submitted for any other degree at the University of Glasgow or any other institution.

John Alan Kennedy

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

The nature of the universe has intrigued mankind for many centuries. In ancient Greece natural philosophers such as Democritus laid the foundations of modern particle physics with their postulations that the universe is simply a void filled with an almost infinite number of invisible particles [1]. These particles were called "atoms", originating from the Greek word meaning indivisible. Thus began our search for the fundamental building blocks of the universe. Today we find ourselves much progressed with the observation of atomic phenomena and the discovery of quantum mechanics leading to new and unexpected heights. However many questions remained to be answered and our present models of the universe are far from completion. This thesis describes the search for one of the fundamental particles which we currently believe to make up our universe, the Higgs boson. The observation of the Higgs boson would further bind our observations to the theoretical model currently held to describe our universe and represent a major triumph for modern particle physics.

This thesis presents an overview of modern particle physics and the Higgs boson in chapter 2 while a description of the apparatus used to search for its existence is given in chapter 3. Chapter 4 describes the four jets cuts based analysis, produced prior to the authors work, and chapter 5 describes the application of this analysis by the author to the 2000 data set complete with a treatment of the statistical uncertainties associated with the analysis. Chapter 6 provides a summary of the observed results from the described analysis and the combined

Higgs boson searches results at ALEPH when analyses other than the authors are combined.

Chapter 7 Presents a detailed description of the author's work on the development of a cleaning algorithm to reduce beam induced noise.

Finally a summary with outlook for future Higgs boson searches is presented in chapter 8.

There also exist an extensive set of appendices to provide further information on specific aspects of the presented analysis and other relevant information which is not extensively covered in the main body of the thesis. In particular the author's work on jet pairing studies is presented in appendix A.

# Chapter 2

# Theoretical overview

...you can hardly ever simplify and unify something just because somebody wants you to.
- The Catcher in the Rye, J.D Salinger

## 2.1 Introduction

This chapter presents a theoretical overview of the Standard Model (SM), the Higgs mechanism therein and a phenomenological discussion of the Higgs boson and backgrounds relevant to searches at LEP. More detailed information about the SM and underlying Quantum Field Theory may be found in references [2–4] while the physics of the Higgs boson is discussed in depth in [5].

## 2.2 The Standard Model

The behaviour of the fundamental particles and their interactions is described in terms of Quantum Field Theory (QFT). The SM, developed by Glashow [6], Weinberg [7] and Salam [8] is based on QFT and describes the interaction of fermionic fields representing quarks and leptons via the exchange of gauge bosons. Quarks and leptons are thought to be the fundamental constituents of matter and are grouped together into three generations or families. The properties of the known quarks and leptons are summarised in table 2.1.

The interactions of the fermions is governed by three fundamental forces known as the strong, weak and electromagnetic interactions[1]. Quarks are observed to interact through all three forces while the leptons interact through the electromagnetic and weak force for the charged leptons and the weak force only for the neutrinos.

| Leptons | | | Quarks | | |
|---|---|---|---|---|---|
| Flavour | $Q$ ($q_e$) | $m_l$ ( MeV/$c^2$) | Flavour | $Q$ ($q_e$) | $m_q$ (bare) ( GeV/$c^2$) |
| e | -1 | 0.510999 | down | -1/3 | $\sim$0.006 |
| $\nu_e$ | 0 | <0.000003 | up | 2/3 | $\sim$0.003 |
| $\mu$ | -1 | 105.65836 | strange | -1/3 | $\sim$0.13 |
| $\nu_\mu$ | 0 | <0.19 | charm | 2/3 | $\sim$1.3 |
| $\tau$ | -1 | 1777 | bottom | -1/3 | 4 |
| $\nu_\tau$ | 0 | <18.2 | top | 2/3 | 174 |

**Table 2.1:** The Standard Model fermions [9].

Each interaction has an associated set of bosonic fields which represent the fundamental force-carrying particles. The properties of these force-carrying bosons are summarised in table 2.2. Original successes in QFT came through the development of Quantum Electro Dynamics (QED). QED describes the interactions of charged particles via the exchange of gauge bosons, in this case the photon, and has been successfully tested within experiments to very high precision.

The development of the weak force as a Quantum Field Theory and its unification with the electromagnetic force to form the electroweak force proved to be a great step forward in the construction of the SM.

Finally the introduction of the colour force mediated by gluons in the form of Quantum Chromo Dynamics (QCD) completes the SM as it is implemented today.

### 2.2.1 Fundamental particles and gauge symmetries

The identification of conserved quantities in the interactions of fundamental particles and the association of these quantities with space-time or internal symme-

[1]The fourth force, gravity, has yet to be satisfactorily described at the quantum level and is omitted from the discussion presented here.

| Bosons | | | |
|---|---|---|---|
| | Interaction | Q ($q_e$) | M ( GeV/$c^2$) |
| $\gamma$ | Electro(weak) | 0 | 0 |
| Z | (Electro)weak | 0 | 91.19 |
| $W^{\pm}$ | (Electro)weak | $\pm 1$ | 80.4 |
| g | Strong | 0 | 0 |

**Table 2.2:** The Standard Model Interactions [9].

tries forms the basis of the SM.

When considering a space-time independent transformation of a fermion field

$$\psi \rightarrow e^{i\alpha}\psi \tag{2.1}$$

where $\alpha$ is a real constant, it is found that the Lagrangian for the free propagation of a field with mass $m$,

$$\mathcal{L} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi, \tag{2.2}$$

is invariant under such a transform by considering that

$$\partial_{\mu}\psi \rightarrow e^{i\alpha}\partial_{\mu}\psi, \tag{2.3}$$

and

$$\bar{\psi} \rightarrow e^{-i\alpha}\bar{\psi}. \tag{2.4}$$

Such an invariance is known as a "global gauge" invariance. Noether's theorem [10] implies the existence of a conserved current for every continuous symmetry of a Lagrangian. Following from this the fundamental particles are said to possess "conserved gauge charges". By introducing locality and allowing $\alpha$ to vary from space-time point to space-time point, that is $\alpha \rightarrow \alpha(x)$, we find that

$$\psi(x) \rightarrow e^{i\alpha(x)}\psi(x), \tag{2.5}$$

where $\alpha(x)$ is now dependent on space-time in a completely arbitrary manner. This type of transformation is known as a "local gauge" transform. The Lagrangian

$$\mathcal{L} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi, \tag{2.6}$$

is however not invariant under such a transform. Considering $\bar{\psi} \to e^{-i\alpha(x)}\bar{\psi}$ the last term of the Lagrangian is indeed invariant yet the invariance of the derivative term does not follow since

$$\partial_\mu \psi \to e^{i\alpha(x)} \partial_\mu \psi + i e^{i\alpha(x)} \psi \partial \alpha, \tag{2.7}$$

with the second term breaking the invariance of the Lagrangian.

The gauge invariance of the Lagrangian may be recovered by replacing the derivative with the covariant derivative where

$$D_\mu \psi \to e^{i\alpha(x)} D_\mu \psi, \tag{2.8}$$

To form a covariant derivative such that the unwanted terms from the local gauge transform are cancelled, and the gauge symmetry revived, we must introduce a vector field $A_\mu$ with the appropriate transformation properties. The covariant derivative $D_\mu$ is thus constructed as

$$D_\mu = \partial_\mu - ieA_\mu \tag{2.9}$$

where the vector field $A_\mu$ transforms as

$$A_\mu \to A_\mu + \frac{1}{e}\partial_\mu \alpha. \tag{2.10}$$

The invariance of the Lagrangian is now satisfied since replacing $\partial_\mu$ with $D_\mu$ gives

$$\begin{aligned} \mathcal{L} &= i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi \\ &= \bar{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi + e\bar{\psi}\gamma^\mu \psi A_\mu \end{aligned} \tag{2.11}$$

Thus to maintain the gauge invariance of the system we have been forced to introduce a vector field $A_\mu$, called a gauge field, which couples to the fermion fields $\psi$.

The interpretation of this new field as a physical particle requires the addition of the corresponding kinetic energy term to the Lagrangian. To remain gauge invariant this kinematic term must be formed from the gauge invariant field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.12}$$

This leads to the Lagrangian

$$\mathcal{L} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi + e\bar{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{2.13}$$

This Lagrangian is that of Quantum Electro Dynamics (QED) in which the electrically charged fermions interact via the exchange of photons.

In summary: By imposing the requirement of local gauge invariance on the free fermion field Lagrangian we are led to the introduction of a vector field and finally to the theory of QED.

Note that the introduction of a mass term for the gauge boson ($\frac{1}{2}mA_{\mu}A^{\mu}$) violates the gauge invariance which we have sought to achieve. Thus the gauge particle, in this case the photon, must remain massless.

### 2.2.2 The Higgs Mechanism

Quantum Field Theories based upon local gauge invariance are desirable not only due to their aesthetic beauty but also since they present us with renormalisable theories which avoid infinite divergences and are thus predictive. The flip side of the coin however, is their lack of accommodation for massive gauge bosons. The explicit introduction of mass terms for gauge boson fields not only destroys the gauge invariance of the Lagrangian, as noted above, but also leads to the loss of renormalisability.

The problem we now face is how to reconcile the requirement of gauge invariance, which requires massless gauge bosons, with the observations of massive weak gauge bosons, i.e. the $W^{\pm}$ and $Z^0$.

The weak gauge bosons must be massive since they are observed to operate on extremely short distance scales and their masses have indeed been measured to be of the order 80-90 GeV$/c^2$ [9].

A solution to this apparent dichotomy is found in the Higgs Mechanism [11–14]. The Higgs mechanism provides a way to generate masses for the gauge bosons while retaining the local gauge invariance which is central to the theory.

The Higgs mechanism generates mass for the gauge bosons by means of spontaneous symmetry breaking. A given symmetry is said to be spontaneously broken if the vacuum does not possess the same symmetry as the Lagrangian. To

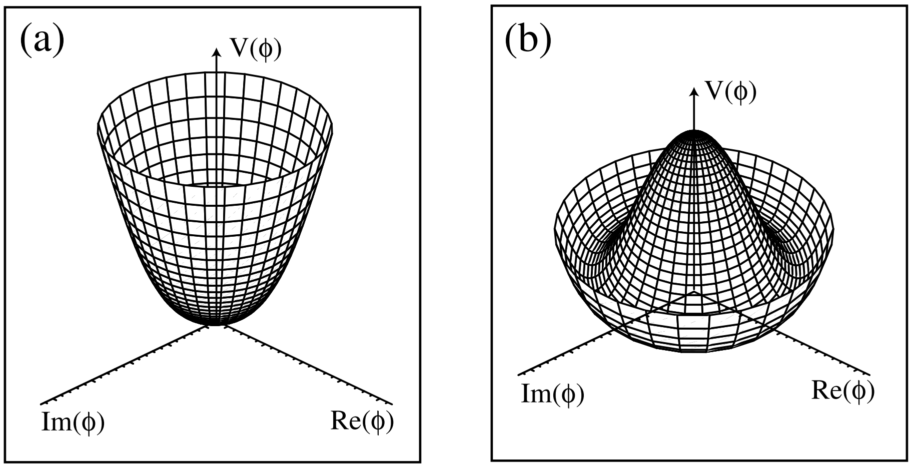


**Figure 2.1:** The potential $V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$ for (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$.

illustrate this we consider the case of a U(1) locally gauge invariant Lagrangian describing the interaction of a scalar field $\phi(x)$ with a gauge field $A_\mu(x)$

$$\mathcal{L} = (D_\mu\phi)^\dagger(D_\mu\phi) - V(\phi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.14}$$

where $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2), D^\mu = \partial^\mu + igA^\mu$ and the scalar potential $V(\phi)$ is described by $V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$ with $\lambda > 0$.

Two possible solutions exist to the form of the scalar potential $V(\phi)$. Choosing $\mu^2 > 0$ gives a minimum at $\phi_1 = \phi_2 = 0$, illustrated in figure 2.1a, and returns the original QED Lagrangian plus a $\phi^4$ self-interaction term.

The choice of $\mu^2 < 0$, see figure 2.1b, however defines a ring of minima situated at

$$|\phi|^2 = \frac{1}{2}\phi_1^2 + \phi_2^2 = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2} \tag{2.15}$$

The symmetry is broken by a specific choice of minimum (eg, $\phi_1 = \frac{v}{\sqrt{2}}, \phi_2 = 0$). Thus the vacuum does not exhibit the gauge invariance which the Lagrangian possesses.

By translating the field $\phi$ to a true ground state and expanding about the vacuum in terms of the fields $\eta, \xi$ according to

$$\phi(x) = \sqrt{\frac{1}{2}}[v + \eta(x) + i\xi(x)] \tag{2.16}$$

the Lagrangian becomes

$$\begin{aligned}\mathcal{L}' = &\frac{1}{2}(\partial_\mu \xi)^2 + \frac{1}{2}(\partial_\mu \eta)^2 - v^2\lambda\eta^2 + \frac{1}{2}e^2v^2A_\mu A^\mu \\ &- evA_\mu\partial^\mu\xi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \text{Interaction terms...}\end{aligned} \tag{2.17}$$

The particle spectrum belonging to the Lagrangian $\mathcal{L}'$ appears to be composed of a massless Goldstone boson[2] ($\xi$), a massive scalar ($\eta$) and the long sought-after massive vector boson $A_\mu$, with the following mass spectrum

$$\begin{aligned} m_\xi &= 0, \\ m_\eta &= \sqrt{2\lambda v^2}, \\ m_A &= ev. \end{aligned} \tag{2.18}$$

However, although we appear to have generated a mass for the gauge boson, we now encounter the problem of the occurrence of a massless Goldstone boson. The problem lies in the interpretation of the Lagrangian $\mathcal{L}'$. By producing a massive vector boson, $A_\mu$, we have raised the number of degrees of freedom by one. Now consider the alternate substitution of a set of real fields h, $\theta$ and $A_\mu$ where

$$\phi \rightarrow \sqrt{\frac{1}{2}}(v + h(x))e^{-i\theta(x)/v} \tag{2.19}$$

and

$$A_\mu \rightarrow A_\mu + \frac{1}{ev}\partial_\mu\theta \tag{2.20}$$

With $\phi$ chosen such that h is real we find

$$\begin{aligned}\mathcal{L}'' = &\frac{1}{2}(\partial_\mu h)^2 - \lambda v^2 h^2 + \frac{1}{2}e^2v^2A_\mu^2 - \lambda\mu h^3 - \frac{1}{4}\lambda h^4 \\ &+ \frac{1}{2}e^2A_\mu^2h^2 + ve^2A_\mu^2h - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\end{aligned} \tag{2.21}$$

In this interpretation the Goldstone boson does not appear. The Lagrangian describes two interacting massive particles, a vector boson ($A_\mu$) and a scalar h, which is known as the Higgs particle.

[2]Goldstone's theorem [15,16] states that the spontaneous breaking of a continuous global symmetry is always accompanied by one or more massless scalar particles which are thus referred to as "Goldstone bosons".

### 2.2.3 The electroweak interaction

The electroweak model of Glashow, Salam and Weinberg [6–8]is based on the symmetry group $SU(2)_L \otimes U(1)_Y$. Following from the case of electromagnetism in which

$$-iej^{em}_{\mu}A^{\mu} = -ie(\bar{\psi}\gamma_{\mu}Q\psi)A^{\mu} \tag{2.22}$$

where Q is the electric charge operator and $j^{em}_{\mu}$ represents the conserved electromagnetic current, we find that the electroweak process requires two basic interactions, firstly an iso-triplet of weak currents $J_{\mu}$ coupled to three vector bosons $W^{\mu}$

$$-gJ_{\mu}W^{\mu} = -ig\bar{\chi_L}\gamma_{\mu}TW^{\mu}\chi_L \tag{2.23}$$

and secondly a weak hypercharge current coupled to a fourth vector boson $B^{\mu}$

$$\frac{-ig'}{2}J^{Y}_{\mu}B^{\mu} = -g'\bar{\psi}\gamma_{\mu}\frac{Y}{2}\psi B^{\mu} \tag{2.24}$$

The operators T and Y are the generators of the SU(2)$_L$ and U(1)$_Y$ groups of gauge transforms respectively, while $g$ and $g'$ represent the two coupling constants.

Taken together the transformations of the left handed and right handed components of $\psi$ are

$$\chi_L \rightarrow \chi'_L = e^{i\alpha(x)T+i\beta(x)Y}\chi_L \tag{2.25}$$

and

$$\psi_R \rightarrow \psi'_R = e^{i\beta(x)Y}\psi_R \tag{2.26}$$

where left handed fermions form isospin doublets $\chi_L$ and right handed fermions form iso-singlets $\psi_R$.

For example, in the first generation we have

$$\chi_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \quad \text{with} \quad T = \frac{1}{2}, Y = -1 \tag{2.27}$$

$$\psi_R = e_R^- \quad \text{with} \quad T = 0, Y = -2 \tag{2.28}$$

$$\chi_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \qquad \psi_R = u_R \text{ } or \text{ } d_R \tag{2.29}$$

Imposing $SU(2)_L \otimes U(1)_Y$ invariance results in a Lagrangian of the form

$$\begin{aligned} \mathcal{L}_1 = & \bar{\chi_L}\gamma^\mu[i\partial_\mu - g\frac{1}{2}\tau W_\mu - g'(\frac{-1}{2})B_\mu]\chi_L \\ & + \bar{e_R}\gamma^\mu[i\partial_\mu - g'(-1)B_\mu]e_R \\ & - \frac{1}{4}W_{\mu\nu}W^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \end{aligned} \tag{2.30}$$

where $\tau$ are the Pauli spin matrices and we have inserted the hypercharge values $Y_L$=-1, $Y_R$=-2. $\mathcal{L}_1$ embodies both the weak isospin and hypercharge interactions. The final two terms represent the kinematic energy and self coupling of the $W_\mu$ fields and the kinetic energy of the $B_\mu$ field where

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.31}$$

and

$$W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - gW_\mu \text{x} W_{nmu} \tag{2.32}$$

with the last term in 2.32 arrises due to the non-abelian nature of the W field.

The gauge symmetry of $\mathcal{L}_1$ is broken if we attempt to directly introduce a mass term for the boson or the fermion fields. To generate mass within the electroweak sector we again turn to the Higgs mechanism. The Higgs mechanism is here formulated such that the $W^\pm$ and $Z^0$ bosons become massive while the photon remains massless. To achieve this, four real scalar fields $\phi_i$ are introduced by adding the gauge invariant Lagrangian for these fields

$$\mathcal{L}_2 = |(i\partial_\mu - gTW_\mu - g'\frac{Y}{2}B_\mu)\phi|^2 - V(\phi) \tag{2.33}$$

to the electroweak Lagrangian $\mathcal{L}_1$, where the fields $\phi$ belong to a $SU(2)_L \otimes U(1)_Y$ multiplet and $|x|^2 = x^\dagger x$.

The minimal choice is to arrange the four fields in an isospin doublet with weak hypercharge $Y$=+1:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.34}$$

The Higgs potential is chosen such that

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \tag{2.35}$$

where, in the case of $\mu^2 < 0$ and $\lambda > 0$, this has a minimum at a finite value of $|\phi|$ defined by

$$|\phi|^2 = \frac{-\mu^2}{2\lambda} \neq 0 \tag{2.36}$$

The ground state we have chosen is degenerate and has no preferred direction in weak isospin space as a consequence of the SU(2)$_L$ symmetry. We are thus free to chose the value of the phase $\phi$.

By choosing

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.37}$$

we leave the U(1)$_{em}$ symmetry unbroken. That is $Q\phi_0 = 0$ such that

$$\phi_0 \rightarrow \phi_0' = e^{i\alpha(x)Q} \phi_0 = \phi_0 \tag{2.38}$$

for any value of $\alpha(x)$.

The vacuum remains invariant under a U(1)$_{em}$ transform and the gauge boson associated with the electromagnetic force, the photon, remains massless. Now by expanding $\phi(x)$ about the chosen vacuum we find that $\phi$ may be expressed in terms of the one remaining scalar field, which is here denoted as the Higgs field h:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \mathrm{h}(x) \end{pmatrix} \tag{2.39}$$

The massive gauge bosons are identified by substituting the vacuum expectation value $\phi_0$ for $\phi(x)$ into the Lagrangian $\mathcal{L}_2$ giving:

$$(\frac{1}{2} v g)^2 W_\mu^+ W_\mu^- + \frac{1}{8} v^2 (W_\mu^3, B_\mu) \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^{\mu 3} \\ B^\mu \end{pmatrix} \tag{2.40}$$

where $W^{\pm} = \frac{1}{\sqrt{2}}(W^1 \mp W^2)$. The comparison of the first term with the expected form of a mass term for a charged boson, $m_W^2 W^+ W^-$, leads to the identification of the W boson mass as

$$m_W = \frac{1}{2} v g \tag{2.41}$$

The second term is off diagonal in the $(W_\mu^3, B_\mu)$ basis. Diagonalisation of the 2x2 mass matrix yields

$$0 \left( g' W_\mu^3 + g B_\mu \right)^2 + \frac{1}{8} v^2 \left( g W_\mu^3 - g' B_\mu \right)^2 \tag{2.42}$$

These two orthogonal terms represent the physical fields $A_\mu$ and $Z_\mu$ with their respective masses $m_A$ and $m_Z$ given by comparison of the expected mass terms for neutral vector bosons $\frac{1}{2} m_A^2 A^2$ and $\frac{1}{2} m_Z^2 Z^2$

$$A_\mu = \frac{g' W_\mu^3 + g B_\mu}{\sqrt{g^2 + g'^2}} \qquad m_\mathrm{A} = 0, \tag{2.43}$$

$$Z_\mu = \frac{g W_\mu^3 - g' B_\mu}{\sqrt{g^2 + g'^2}} \qquad m_Z = \frac{v}{2} \sqrt{g^2 + g'^2}. \tag{2.44}$$

This result may be recast in terms of $\theta_W$ by considering

$$\tan \theta_W = \frac{g'}{g}. \tag{2.45}$$

Thus we find

$$A_\mu = cos\theta_W B_\mu + sin\theta_W W_\mu^3 \tag{2.46}$$

and

$$Z_\mu = -sin\theta_W B_\mu + cos\theta_W W_\mu^3 \tag{2.47}$$

and we have

$$\frac{m_W}{m_Z} = cos\theta_W \tag{2.48}$$

The inequality $m_Z \neq m_W$ originates from the mixing between the $W_\mu^3$ and $B_\mu$ fields. The application of the Higgs mechanism to the electroweak sector of the Standard Model has led to the generation of a massless photon, $A_\mu$ and massive $Z_\mu$ and $W^{\pm}$ fields with $m_Z > m_W$. The relationship between $m_Z$ and $m_W$ is a prediction of the Standard Model and the Higgs sector as we have described it. Tests of this relationship form an indirect probe into the exact nature of the Higgs mechanism which is at work in the electroweak sector.

### 2.2.4 The Generation of Fermion Masses

It was previously noted that the direct introduction of a fermion mass term, $-m\bar{\psi}\psi$, into the electroweak Lagrangian (2.30) leads to the loss of gauge invariance. One of the most attractive features of the Higgs mechanism, as we have applied it, is that the same Higgs doublet which generates mass for the W and Z bosons is also sufficient to give mass to the leptons and quarks.

As an example we will consider the case of the electron. To generate the electron mass the following $SU(2)_L \otimes U(1)_Y$ gauge invariant term is added to the Lagrangian,

$$\mathcal{L}_3 = -G_e \left[ (\bar{\nu_e}\bar{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e_R}(\phi^-, \bar{\phi^0}) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right], \tag{2.49}$$

here the Higgs doublet is found to have the exact $SU(2)_L \otimes U(1)_Y$ quantum numbers to couple to $\bar{e_L}e_R$.

Once spontaneous symmetry breaking takes place we find

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \mathrm{h}(x) \end{pmatrix} \tag{2.50}$$

and, on substituting for $\phi$ in equation (2.49), we obtain

$$\mathcal{L}_3 = -\frac{G_e}{\sqrt{2}} v(\bar{e_L}e_R + \bar{e_R}e_L) - \frac{G_e}{\sqrt{2}} v(\bar{e_L}e_R + \bar{e_R}e_L)\mathrm{h} \tag{2.51}$$

now choosing $m_e$ such that

$$m_e = \frac{G_e v}{\sqrt{2}} \tag{2.52}$$

the Lagrangian may be recast in the form

$$\mathcal{L}_3 = -m_e\bar{e}e - \frac{m_e}{v}\bar{e}e\mathrm{h}. \tag{2.53}$$

Here $G_e$ is arbitrary and so the mass of the electron is not predicted in the model we present. In addition to the desired mass term the Lagrangian includes an interaction term which couples the Higgs to the electron. The small size of this coupling term, $m_e/v = m_e/246$ GeV, however means that any effects produced by this term are very small and as yet no detectable effect has been observed in electroweak interactions.

We now turn our attention to the generation of mass for the quarks. Although the quark masses are generated in the same manner as the lepton masses the generation of mass for the upper member of a quark doublet requires the construction of a new Higgs doublet from $\phi$.

The new Higgs doublet is

$$\phi_c = i\tau\phi^* = \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} \xrightarrow{\text{broken}} \frac{1}{\sqrt{2}} \begin{pmatrix} v + \mathrm{h}(x) \\ 0 \end{pmatrix} \tag{2.54}$$

where $\phi_c$ transforms identically to $\phi$ but has the opposite weak hypercharge (ie, $\phi \Rightarrow$ Y=1, $\phi_c \Rightarrow$ Y=-1).

This new Higgs doublet may then be used to construct a further gauge invariant term which is added to the Lagrangian,

$$\begin{aligned} \mathcal{L}_4 = & -G_d(\bar{u}, \bar{d})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_R - G_u(\bar{u}, \bar{d})_L \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} u_R \\ & + \text{hermitian conjugate...} \end{aligned} \tag{2.55}$$

$$\mathcal{L}_4 = -m_d\bar{d}d - m_u\bar{u}u - \frac{m_d}{v}\bar{d}d\mathrm{h} - \frac{m_u}{v}\bar{u}u\mathrm{h} \tag{2.56}$$

The above Lagrangian is formulated in terms of the $(u, d)_L$ quark doublet. The weak interaction however operates on $(u, d')_L$, $(c, s')_L$ and $(t, b')_L$ doublets where the primed states are linear combinations of the flavour eigenstates. Once this is taken into account the Lagrangian becomes

$$\begin{aligned} \mathcal{L}_4 = & -G_d^{ij}(\bar{u}_i, \bar{d}'_j)_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_{jR} - G_u^{ij}(\bar{u}_i, \bar{d}'_j)_L \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} u_{iR} \\ & + \text{hermitian conjugate...} \end{aligned} \tag{2.57}$$

where $i, j$ = 1,...N and N is the number of quark doublets. The Lagrangian may be written in the diagonal form,

$$\mathcal{L}_4 = -m_d^i\bar{d}_i d_i \left(1 + \frac{\mathrm{h}}{v}\right) - m_u^i\bar{u}_i u_i \left(1 + \frac{\mathrm{h}}{v}\right), \tag{2.58}$$

again we find that the mass terms depend upon the arbitrary parameter $G_{u,d}$ and as such the quark masses remain free parameters in the SM. One desirable outcome of the Lagrangian in this form is that the Higgs coupling is flavour conserving.

The minimal choice of a single Higgs doublet, as applied above, is sufficient to generate the masses of both the gauge bosons and the fermions. However the fermion masses are not predicted by the theory and remain as free parameters which must be input into the SM.

The Higgs coupling to fermions is proportional to their mass and provides a testable prediction of the Higgs mechanism as applied here. The mass of the Higgs boson itself is also a free parameter which may be searched for experimentally.

From the effective potential

$$V(\phi) = \mu^2 \phi \dagger \phi + \lambda(\phi \dagger \phi)^2 \tag{2.59}$$

it is found that

$$m_h^2 = 2v^2\lambda \tag{2.60}$$

and since $v$ is fixed, larger values of $m_h$ will lead to correspondingly large values of $\lambda$, thus increasing the self coupling term and eventually leading to the failure of perturbative QFT.

## 2.3 The MSSM and associated Higgs sector

In addition to the SM interpretation of the Higgs sector, as discussed above, the Higgs may also be defined in the context of the Minimal Supersymmetric Model (MSSM) [17]. Supersymmetry is an attractive extension to the SM which has been extensively studied theoretically although no experimental evidence has yet been forthcoming. In supersymmetric models each matter fermion has a scalar "susy" partner (e.g. quark ⇔ squark) and each gauge boson has a fermionic "susy" partner (e.g. photon ⇔ photino). Supersymmetry cannot be an exact symmetry since it predicts that the supersymmetric particles have the same mass as their SM partners. The lack of observation of any such supersymmetric particles means that if supersymmetry exists it must be a broken symmetry. The MSSM interpretation of supersymmetry presents the minimal form of the supersymmetric extension to the SM. In the MSSM the Higgs sector requires two Higgs doublets to give mass to the up-type and down-type quarks separately [5]. In this model the Higgs sector thus consists of five physical states with three being neutral, two CP-even (h and H) and one CP-odd (A), and a pair of charged bosons ($H^{\pm}$).

The form of the Higgs potential is

$$\begin{aligned} V(\Phi_1, \Phi_2) =& \lambda_1(\Phi_1^\dagger\Phi_1 - v_1^2)^2 + \lambda_2(\Phi_2^\dagger\Phi_2 - v_2^2)^2 \\ &+\lambda_3[((\Phi_1^\dagger\Phi_1 - v_1^2) + (\Phi_2^\dagger\Phi_2 - v_2^2)]^2 \\ &+\lambda_4[(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) - (\Phi_1^\dagger\Phi_2)(\Phi_2^\dagger\Phi_1)] \\ &+\lambda_5[Re(\Phi_1^\dagger\Phi_2) - v_1 v_2]^2 \\ &+\lambda_6[Im(\Phi_1^\dagger\Phi_2)]^2. \end{aligned} \tag{2.61}$$

where $\lambda_i$ are real parameters and $\Phi_1$ and $\Phi_2$ denote two complex fields

$$\begin{aligned} <\Phi_1>=& \begin{pmatrix} \phi_1^{0*} \\ -\phi_1^- \end{pmatrix}, \quad Y = -1 \\ <\Phi_2>=& \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix}, \quad Y = 1, \end{aligned} \tag{2.62}$$

that each form a doublet under $SU(2)_L$ with opposite hypercharge.

If the parameters $\lambda_i$ are not negative then the following minima for the potentials are found

$$<\Phi_1>= \begin{pmatrix} 0 \\ v_1 \end{pmatrix}, <\Phi_2>= \begin{pmatrix} 0 \\ v_2 \end{pmatrix}, \tag{2.63}$$

where $v_1$ and $v_2$ are the vacuum expectation values[3] of the Higgs fields.

Once the Goldstone bosons have been gauged away the five physical states arise. The MSSM Higgs sector is described by six independent parameters: The four Higgs boson masses, the ratio of vacuum expectation values $v_2/v_1 = \tan\beta$ and $\alpha$, the mixing angle in the CP-even sector, described below.

The physical Higgs states in the charged sector are

$$H^{\pm} = -\Phi_1^{\pm}\sin\beta + \Phi_2^{\pm}\cos\beta \tag{2.64}$$

with a mass $m^2_{H^\pm} = \lambda_4(v_1^2 + v_2^2)$.

In the CP-odd sector we have one boson A,

$$A = \sqrt{2}(-Im(\Phi_1^0)\sin\beta + Im(\Phi_2^0)\cos\beta) \tag{2.65}$$

[3]The vacuum expectation values $v_1$ and $v_2$ are related to the W mass via $m_W^2 = g^2(v_1^2 + v_2^2)/2$.

with a mass $m_A^2 = \lambda_6(v_1^2 + v_2^2)$
In the CP-even sector two physical Higgs scalars mix through the following mass-squared matrix

$$\mathcal{M} = \begin{pmatrix} 4v_1^2(\lambda_1+\lambda_3)+v_2^2\lambda_5 & (4\lambda_3+\lambda_5)v_1v_2 \\ (4\lambda_3+\lambda_5)v_1v_2 & 4v_2^2(\lambda_2+\lambda_3)+v_1^2\lambda_5 \end{pmatrix} \quad , \tag{2.66}$$

with the physical mass eigenstates

$$\begin{aligned} \mathrm{H}^0 =& \sqrt{2}[(Re(\Phi_1^0)-v_1)\cos\alpha + (Re(\Phi_2^0)-v_2)\sin\alpha], \\ \mathrm{h}^0 =& \sqrt{2}[-(Re(\Phi_1^0)-v_1)\sin\alpha + (Re(\Phi_2^0)-v_2)\cos\alpha] \end{aligned} \tag{2.67}$$

and the corresponding masses

$$m^2_{\mathrm{H}^0,\mathrm{h}^0} = \frac{1}{2}[\mathcal{M}_{11} + \mathcal{M}_{22} \pm \sqrt{(\mathcal{M}_{11}-\mathcal{M}_{22})^2 + 4\mathcal{M}_{12}^2}]. \tag{2.68}$$

The sixth free parameter, the mixing angle of the CP-even fields $\alpha$ is obtained from

$$\begin{aligned} \sin(2\alpha) =& \frac{2\mathcal{M}_{12}}{\sqrt{(\mathcal{M}_{11}-\mathcal{M}_{22})^2 + 4\mathcal{M}_{12}^2}}, \\ \cos(2\alpha) =& \frac{\mathcal{M}_{11}-\mathcal{M}_{22}}{\sqrt{(\mathcal{M}_{11}-\mathcal{M}_{22})^2 + 4\mathcal{M}_{12}^2}} \end{aligned} \tag{2.69}$$

The couplings of the Higgs bosons to gauge bosons and fermions determine production cross sections as well as decay rates.

The production cross sections for the processes $\mathrm{e^+e^-} \to \mathrm{Zh}$ and $\mathrm{e^+e^-} \to \mathrm{hA}$ are

$$\begin{aligned} \sigma(\mathrm{e^+e^-} \to \mathrm{hZ}) =& \sin^2(\beta-\alpha)\sigma_{SM} \\ \sigma(\mathrm{e^+e^-} \to \mathrm{hA}) =& \cos^2(\beta-\alpha)\bar{\lambda}\sigma_{SM}. \end{aligned} \tag{2.70}$$

The factor $\bar{\lambda}$ is defined as $\lambda_{\mathrm{Ah}}^{3/2}/[\lambda_{\mathrm{Zh}}^{1/2}(12m_Z^2/s + \lambda_{\mathrm{Zh}})]$ with the 2-particle phase space factor $\lambda_{ij} = (1-(m_i+m_j)^2/s)(1-(m_i-m_j)^2/s)$. $\sigma_{SM}$ is the cross section for the hZ process within the minimal Standard Model.

## 2.4 Higgs boson phenomenology

This section presents the phenomenological issues relevant to the Higgs boson searches at LEP2. The interaction of the Higgs boson with the other particles of the Standard Model may be described fully as a function of the Higgs boson mass. Therefore, for a given Higgs boson mass, all relevant processes(e.g. , production cross sections and decay branching ratios) of the Higgs boson may be determined. The results of these calculations may be used to provide a framework for developing direct Higgs boson searches and interpreting the results of such searches.

### 2.4.1 Constraints on the Higgs boson mass

The Higgs boson mass, although a free parameter in the SM, can be constrained by both theoretical and experimental methods. Theoretical arguments based on the self consistency of the Standard Model may be used to derive upper bounds on the mass of the Higgs boson while experimental results from electroweak data and direct searches may be used to set indirect and direct mass bounds respectively.

Theoretical arguments based on unitarity may be used to place upper bounds on the mass of the Higgs boson. The absence of a fundamental scalar field, the Higgs field, causes the amplitude for longitudinally polarised WW scattering to diverge quadratically in energy when calculated perturbatively [18]. This ultimately leads to the violation of unitarity. The inclusion of the Higgs boson in the theory suppresses this behaviour and unitarity remains unviolated provided that the Higgs boson has a mass less than $\approx 1$ TeV/$c^2$.

The Higgs boson mass may also be constrained by precision measurements of electroweak data [19]. Electroweak processes are sensitive to the mass of the Higgs boson due to its contribution to radiative corrections through loop diagrams. The mass of the Higgs boson may be constrained by comparing the experimental measurements from electroweak data to the predictions given for various Higgs boson masses. Similar methods were successfully used to constrain the mass of the top quark before its discovery in 1995. Precision electroweak measurements have been combined by the LEP Electroweak Working Group with a global fit to these data being performed as a function of the Higgs boson mass.

The results to the fit, expressed in the form $\Delta\chi^2 = \chi^2 - \chi^2_{min}$, are shown in figure 2.2 [19, 20]. The resulting fit determines the mass of the Higgs boson to be in the range [19]

$$98^{+58}_{-38}\ \mathrm{GeV}/c^2 \tag{2.71}$$

while masses above 212 GeV/$c^2$ are excluded at the 95% confidence level.

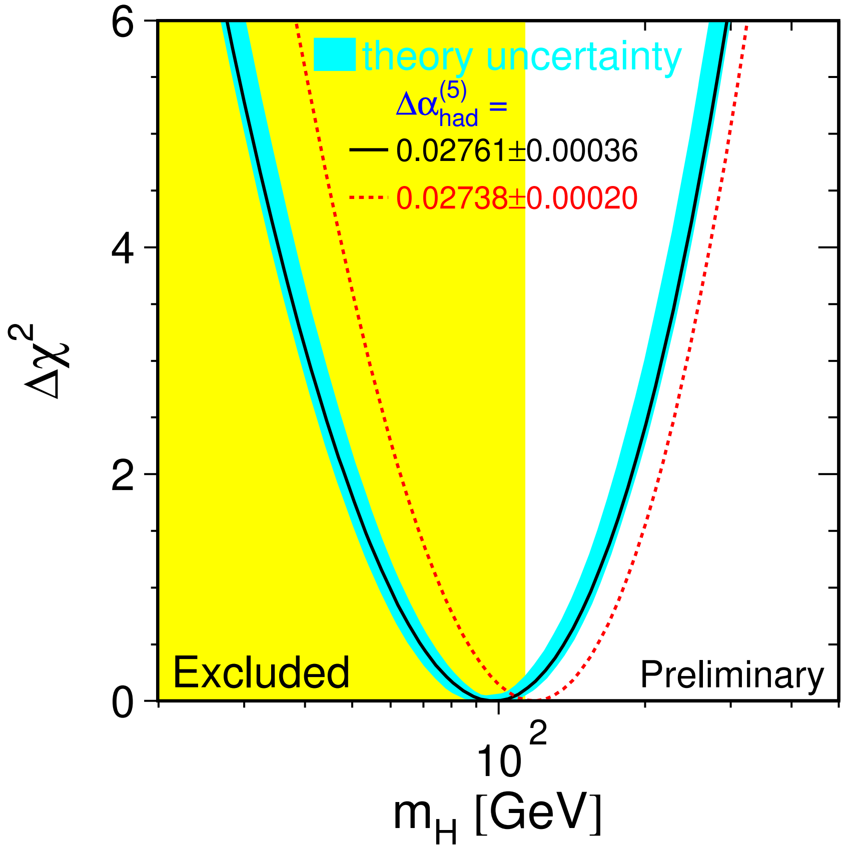


**Figure 2.2:** The distribution of $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ for a global fit to electroweak data as a function of $m_h$. The shaded bands around the curve indicates the theoretical uncertainty while the solid yellow region represents the masses excluded at 95% confidence level by direct searches.

Direct searches for the Higgs boson by the four experiments at the LEP collider provide stringent lower bounds on the mass of the Higgs boson. The combination of the data from all four LEP experiments with centre of mass energies up to

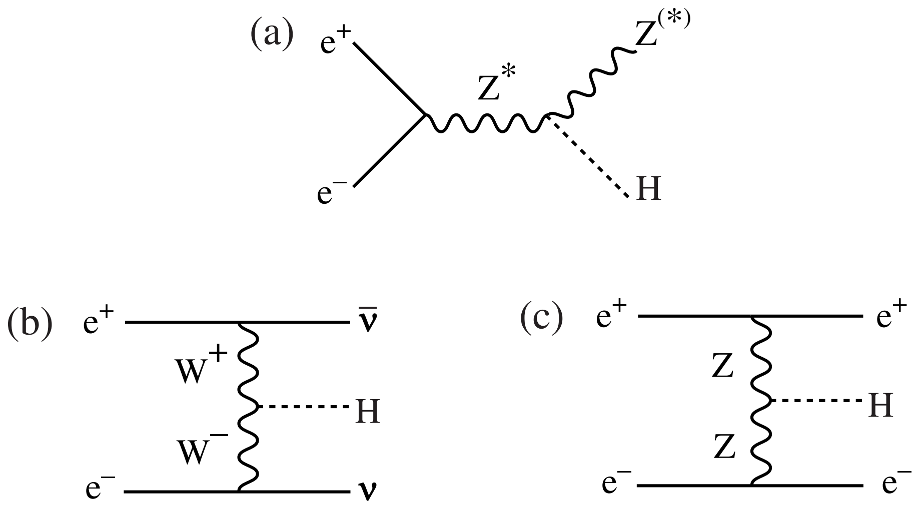


**Figure 2.3:** Higgs boson production mechanisms at LEP 2.(a) Higgsstrahlung, (b) $W^+W^-$ fusion and (c) ZZ fusion.

202 GeV by the LEP Higgs Working Group leads to an exclusion of a Higgs boson with mass below 107.9 GeV/$c^2$ at the 95% confidence level [21].

### 2.4.2 Higgs boson production at LEP

The dominant production mechanisms for the Higgs boson at the LEP accelerator are the so called Higgsstrahlung and boson-fusion processes, figure 2.3. Direct Higgs production via $e^+e^- \rightarrow h$ is suppressed due to the very small electron mass and thus small eeh coupling term. The Higgsstrahlung process in which the Higgs boson is radiated from a virtual Z with the Z returning to a value close to its mass shell provides the majority of the Higgs production cross section. However the cross section for the Higgsstrahlung process falls sharply as the hypothetical Higgs mass reaches and exceeds threshold, $m_\mathrm{h} > \sqrt{s} - m_Z$, in which case the final state Z boson is required to be off shell. In the region of kinematic threshold the WW and ZZ fusion processes, figure 2.3, contribute a much more significant fraction of the total Higgs boson production cross section [22]. The Higgs production cross section for $\sqrt{s}$=206 GeV is shown in figure 2.4. A centre of mass energy of 206 GeV represents the luminosity-weighted LEP average for the data taken in 2000 and corresponds to a kinematic threshold for the Higgs mass of 114.8 GeV/$c^2$.

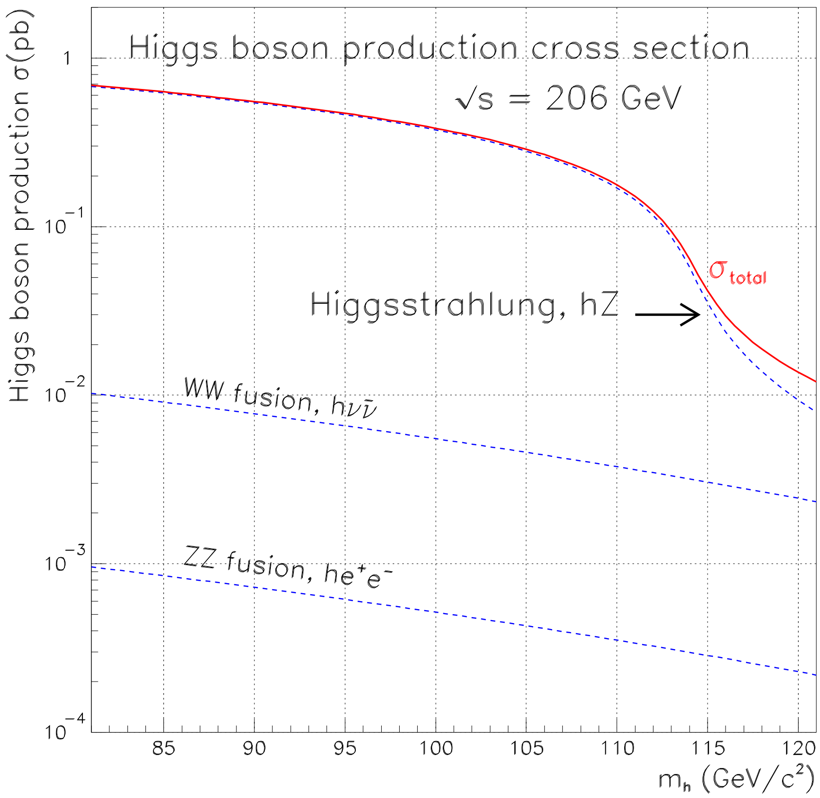


**Figure 2.4:** The Higgs production cross section for the three possible production mechanisms shown in figure 2.3 both individually and as a combined total.The cross sections are calculated at $\sqrt{s} = 206$ GeV using the HZHA03 [23] [24] generator.

### 2.4.3 Higgs boson decay and detection

The detection of the Higgs boson at LEP is directly related to the decay of the Higgs system and signatures from this decay.

The Higgs decay width is predicted to be very narrow, ≈3 MeV, for a Higgs boson of mass ≈ 114 GeV/$c^2$ [5]. The width of the Higgs boson is, as such, too small to be resolved experimentally. The main decay modes of the Higgs in the mass ranges relevant to LEP2 are shown in table 2.3.

| Decay Mode Mode | Branching Fraction (%) |
|---|---|
| $b\bar{b}$ | 74.78 |
| $\tau^+\tau^-$ | 7.33 |
| WW | 6.90 |

**Table 2.3:** The main decay modes of the Higgs boson of mass 114 GeV/$c^2$.

The partial decay width of the Higgs to fermions is given by [5]

$$\Gamma(\mathrm{h} \to \mathrm{f\bar{f}}) = \frac{N_c g^2 m_f^2}{32\pi m_W^2}(\beta^3 m_\mathrm{h}) \tag{2.72}$$

where $N_c = 1$ for leptons and 3 for quarks and $\beta^2 = 1 - 4m_f^2/m_\mathrm{h}^2$. Higher order QCD corrections produce the branching ratios shown in table 2.3. The partial width is proportional to the square of the fermion mass when $m_\mathrm{h} >> m_f$ and so in the LEP 2 region we expect the Higgs boson to decay most strongly to the heaviest kinematically available fermion, in this case the b quark.

The partial width of the decay to gauge bosons is suppressed for Higgs boson masses in the range $m_\mathrm{h} < 2m_W$ and as such the fermionic decay $\mathrm{h} \to \mathrm{b\bar{b}}$ dominates.

The Higgs branching ratios for each final state as a function of Higgs mass are shown in figure 2.5.

The $b\bar{b}$ decay channel is clearly dominant while $\tau^+\tau^-$ and WW final states contribute roughly equal amounts for a Higgs boson of mass ≈ 114 GeV/$c^2$.

The Higgs boson search strategy at LEP is to focus on distinctive final state topologies which may be defined by considering the decay mechanisms of the

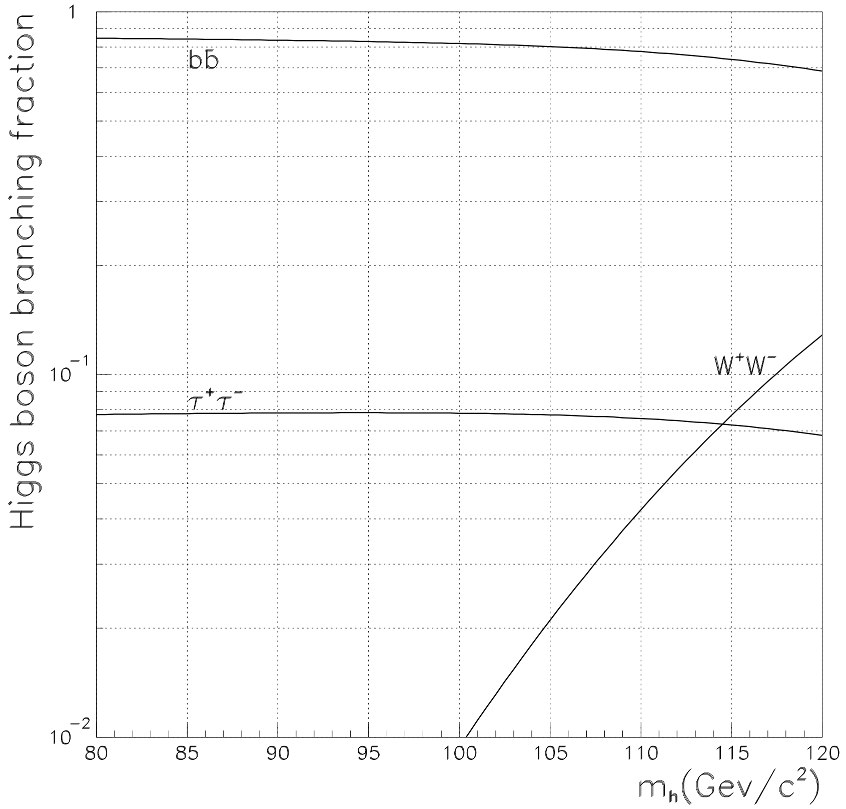


**Figure 2.5:** The branching fractions of the Higgs boson to the decay modes searched for at LEP 2, expressed as a function of the Higgs boson mass. The fractions are calculated using the HZHA03 [23] [24] generator.

Higgs in the relevant mass range. The channels used, in decreasing order of power, are $\mathrm{hq\bar{q}}$, $\mathrm{h}\nu\overline{\nu}$, $\mathrm{h}\ell^{+}\ell^{-}$ (where $\ell = e, \mu$) and $\mathrm{q\bar{q}}\tau^{+}\tau^{-}$.

These channels are described in more detail in appendix C.

## 2.5 Background processes

There exist several processes within the Standard Model which may mimic a Higgs boson event and, as such, these processes present a background for any search for a possible Higgs boson signal. It is imperative that the backgrounds are correctly understood and minimised within each individual search channel.

The backgrounds originate from numerous processes each of which mimics a true Higgs decay in a given channel to a greater or lesser extent depending on the similarities between the signal and background topologies. The SM background processes can be broken down into two distinct groups, the two fermion (section 2.5.1) and four fermion processes (section 2.5.2- 2.5.5).

The decay channels treated at LEP, with the exception of $\tau^{+}\tau^{-}\ell^{+}\ell^{-}$, are all multi-hadronic in nature and as such only multi-hadronic background processes are relevant.

There follows a brief description of the background processes which are relevant to the Higgs searches performed at LEP2.

### 2.5.1 Two fermion processes

The relevant background from the two fermion process may be split into two components.

Firstly, the major contribution is from the production of a $\mathrm{q\bar{q}}$ pair from a Z or virtual photon with an associated Initial State Radiation (ISR) photon, $\mathrm{e^{+}e^{-}} \rightarrow \mathrm{Z}^{*}/\gamma^{*} \rightarrow \mathrm{q\overline{q}}(\gamma)$, figure 2.6b. The ISR photon escapes detection down the beam pipe leading to a visible energy below $\sqrt{s}$. It is possible for one of the final state quarks to radiate a hard gluon thus forming an event with more than two jets (eg, $\mathrm{q\bar{q}}g$, $\mathrm{q\bar{q}}gg$, $\mathrm{q\bar{q}q\bar{q}}$). This allows the two fermion process to mimic a four jet final state, especially in cases where two of the final jets are b quarks.

A smaller contribution to the two fermion final state arises from hard $e^+e^-$ annihilation to $q\bar{q}$. In this process no ISR photon is radiated and the visible mass is closer to the centre of mass energy.

### 2.5.2 Two photon process

The two photon or $\gamma\gamma$ process refers to events in which virtual photons from the initial state $e^+e^-$ form a fermion/antifermion pair. This process, shown in figure 2.6a, represents a four fermion final state as the original $e^+e^-$ are still present. The majority of these events are however classed as "untagged" where the final state $e^+e^-$ escape detection by continuing on their path down the beam pipe. This process is relevant in the case where the fermion pair produced is $q\bar{q}$ or $\tau^+\tau^-$. The high production cross section for this process is countered by the low visible mass and particle multiplicity which make it easy to distinguish as background and remove in the early stages of analysis. This process does not contribute significantly to the backgrounds for the Higgs search analyses but is considered for the $h\nu\nu$ analysis.

### 2.5.3 The $W^+W^-$ process

The pair production of $W^{\pm}$ bosons can occur via one of three possible processes, depicted in figures 2.6c. As the W can decay to $q\bar{q}'$ or $\ell\nu$ these processes can produce final states with leptons and jets and is a source of c quark jets which may be misidentified as b jets in Higgs searches. The inability of these processes to produce real b jets does however present a method of suppressing the majority of this background and thus reduces its impact on the Higgs bosons searches.

### 2.5.4 The $ZZ$ process

The ZZ process actually refers to numerous processes ( $e^+e^- \rightarrow \gamma^*\gamma^*$, $Z^{(*)}\gamma^*$, $Z^{(*)}Z^{(*)}$) and produces four fermion final states. The possible decay modes of the Z boson $q\bar{q}$(70%), $\ell^+\ell^-$(10%), $\nu\bar{\nu}$(20%) allow this decay process to form multiple final states which overlap with different Higgs boson signals. The ZZ background process, depicted in figure 2.6d, presents a major challenge for the Higgs boson searches since it may produce final states with $b\bar{b}$ and $\tau^+\tau^-$ pairs which are effectively indistinguishable from true Higgsstrahlung final states. As a result of this, the ZZ process is often referred to as an irreducible source of background.

### 2.5.5 The single W and Z processes

The production of a single $W^\pm$ via $e^+e^- \rightarrow W^\pm e\nu_e$ (denoted We$\nu$) and single Z bosons via $e^+e^- \rightarrow Ze^+e^-$ (denoted as Zee) contribute to the four fermion final state. Examples of their production processes are shown in figures 2.6e and 2.6f respectively. In reality the Z in the Zee process represents either $Z^{(*)}$ or $\gamma^*$ as is the case in ZZ processes discussed earlier.

A characteristic which is common to both the Zee and We$\nu$ processes is the typical escape of one of the initial $e^+e^-$ pair down the beam pipe. The events therefore have a large component of missing longitudinal momentum which may be used to identify them as background events. The events are also characterised by possessing a visible mass which is near that of the produced W or Z boson rather than at the centre of mass energy as would be expected for the production of a Higgs boson in most channels.

The production of single Z bosons is also possible via the $e^+e^- \rightarrow Z\nu\bar{\nu}$ process which follows exactly that of Higgs production via WW fusion but with a Z boson produced in the final state rather than a Higgs boson. This process however has a very low cross section and so does not contribute significantly[4]. The cross sections of the associated background processes at $\sqrt{s} = 206$ GeV are summarised in table 2.4

Each background process will contribute with a different degree of significance to each of the search analyses. The significance of the background is determined by the production cross section and the overlap of the event characteristics between the given background and the Higgs signal hypothesis for that specific channel.

At LEP each search channel first applies a loose pre-selection to the data to eliminate unmodelled backgrounds and the majority of the most distinguishable

[4]Single Z production is however considered as a source of background for the h$\nu\nu$ search channel, as described in appendix C.

| Background Process | Cross-Section( $pb^{-1}$) |
|---|---|
| ZZ | 2.81 |
| WW | 17.54 |
| $q\bar{q}$ | 80.61 |
| Zee | 7.18 |
| $We\nu$ | 0.884 |
| $Z\nu\bar{\nu}$ | 0.0183 |

**Table 2.4:** The production cross-sections of various background processes.

background sources. The production of pre-selection cuts also allows a comparison of data and the Monte Carlo simulation to determine the accuracy of the simulation. There then follows a tighter selection (eg, optimised cuts, Neural Network, Likelihood) in which the power of the analysis is optimised to increase the possibility of observing a Higgs boson signal. The selection analysis is designed to reduce the more difficult background sources and lead to an optimal analysis.

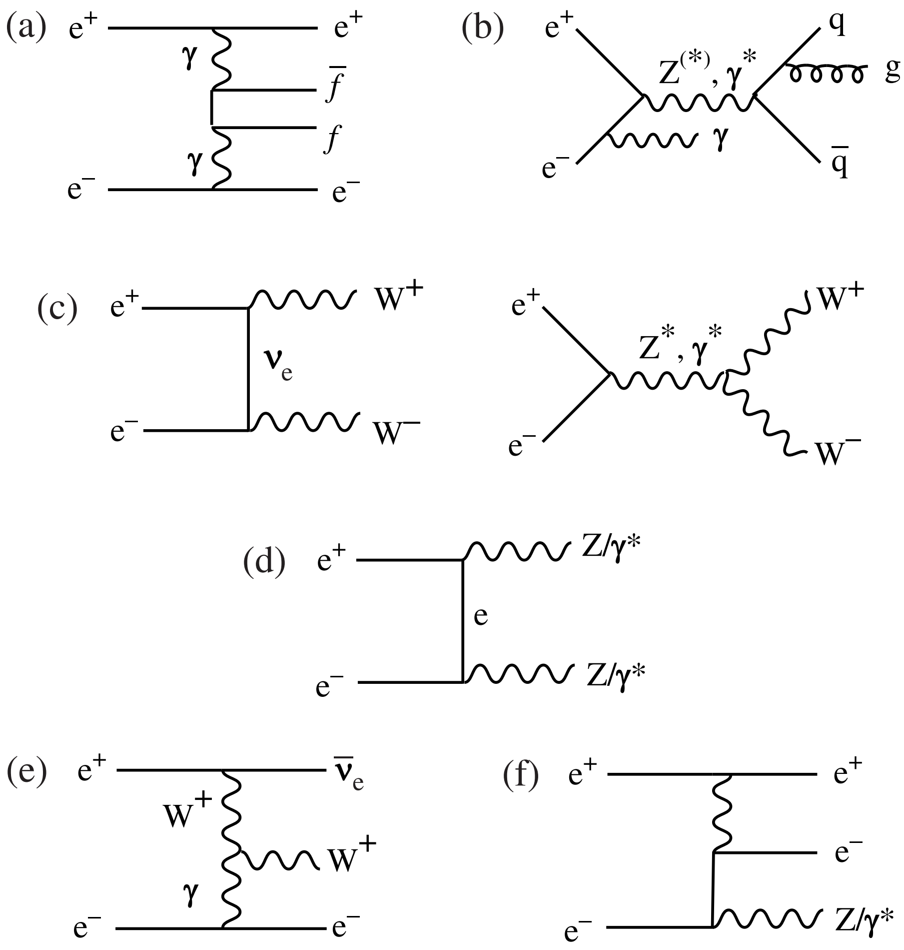


**Figure 2.6:** Standard Model background processes relevant to the Higgs searches at LEP. (a) two photon($\gamma\gamma$), (b)$q\bar{q}(\gamma)$ ,(c) W pair production (WW), (d) $Z^{(*)}/\gamma^{*}$ pair production (ZZ), (e) example of single W production (We$\nu$), (f) example of single Z production (Zee).

# Chapter 3

# LEP and The ALEPH detector

## 3.1 Introduction

The Large Electron Positron (LEP) [25] collider at CERN, the European Laboratory for Particle Physics, is the end product of an ambitious project to extend our knowledge of the fundamental particles which make up our universe. The LEP collider is situated underground at an average depth of $\approx$100 m and is formed from eight straight and eight curved sections which together form a ring measuring 26.7 km in circumference. The ALEPH detector [26–29], described below, is one of four multi-purpose particle detectors which are located at the four interaction points on the ring, the other detectors being DELPHI [30], L3 [31] and OPAL [32]. The purpose of the LEP collider and the detectors situated there, is to explore the physics of the SM and to search for any signs of physics beyond this model. This exploration involves both the study of known particles such as the massive Z and W bosons of the electroweak sector and the search for any signs of the existence of yet unseen particles such as the Higgs boson. The initial part of this chapter describes the LEP accelerator and its performance. The latter part of the chapter describes the ALEPH detector which was used to collect the data for the analysis described within this thesis.

## 3.2 The LEP collider

The LEP collider shown in figure 3.1 collects and accelerates counter-rotating high energy bunches of electrons and positrons, finally colliding them at four equidis-

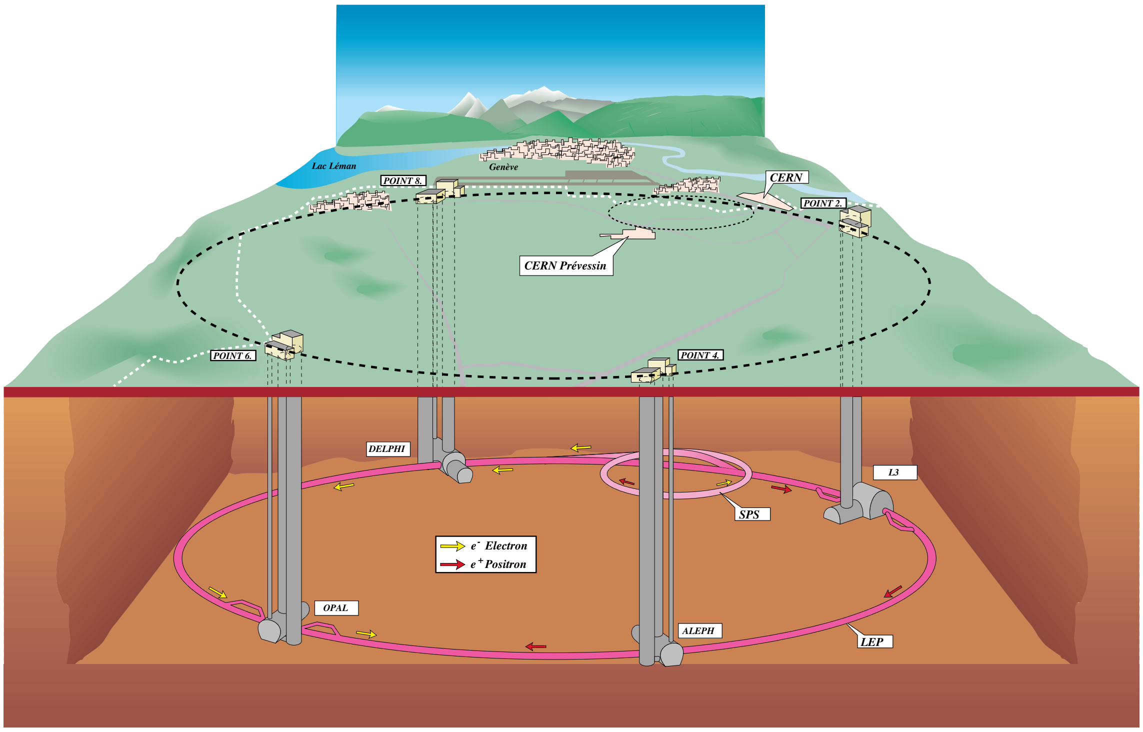


**Figure 3.1:** A schematic view of the LEP accelerator showing the four experiments placed at the interaction points.

tant interaction points (IP's) located on the straight sections of the storage ring.

The electrons and positrons are accelerated by radio frequency cavities installed on the straight sections and are guided around the arcs by dipole bending magnets. The beams are contained in an evacuated pipe with a vacuum maintained at the order of $10^{-9}$ torr with the integrity of the beam being maintained by a complex system of quadruples, sextupoles and correcting magnets. The diameter of the LEP collider is large to reduce the effect of energy loss due to synchrotron radiation. Any charged particle moving along a curved path will radiate photons and thus lose energy, this is known as synchrotron radiation. The energy loss for each revolution of the storage ring is proportional to $\frac{E^4}{m^4\rho}$, where $E$ and $m$ are the energy and mass of the particle and $\rho$ is the bending radius of the ring. For beam energies of 100 GeV, the theoretical maximum for which LEP was designed, the energy loss for each electron or positron is approximately 3 GeV per rotation. This lost energy must be replenished by the accelerating RF cavities to maintain the beam energy.

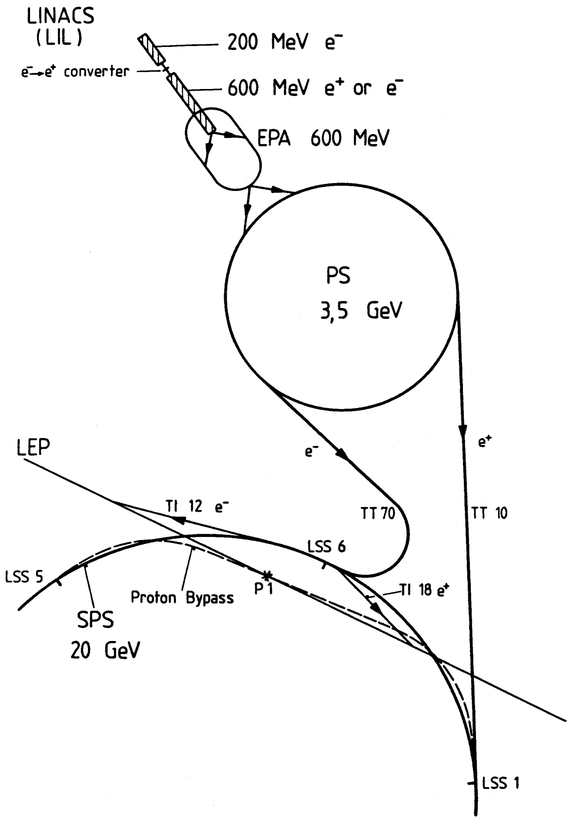


**Figure 3.2:** Schematic overview of the machines used in the acceleration process. As the beams accelerate they are focused and injected into increasingly larger accelerators until they reach the energy needed to enter the LEP ring.

A series of smaller accelerators and storage rings are used to accelerate and finally inject the electrons and positrons into the LEP ring, see figure 3.2. Initially electrons are produced by a pulsed electron gun and then are accelerated to 200 MeV by a linear accelerator (LINAC). A fraction of these accelerated electrons are then collided with a fixed tungsten target, known as the $e^- \to e^+$ converter, to produce positrons via the pair production process. A second linac is then used to accelerate the electrons and positrons to 600 MeV. Both the electrons and positrons are then fed into the Electron Positron Accumulator (EPA). The EPA is a 0.12 km storage ring which separates the particles into bunches and stores them until an intensity of approximately $10^{10}$ particles is achieved. The EPA acts as a buffer between the dedicated LEP pre-injectors and the multi-purpose CERN Proton-Synchrotron (PS). Once a sufficient intensity of electrons and positrons has been accumulated in the EPA the particles are injected into the PS where they are accelerated to 3.5 GeV. The bunches are then fed into the Super Proton Synchrotron(SPS) in which they are accelerated to their final pre-LEP injection energy of 22 GeV. Once injected into LEP the particle bunches are then accelerated to the LEP collision energy through a series of steps. The beams of electrons and positrons in LEP circulate at a frequency of approximately 11,000 Hz which corresponds to a time of around 90$\mu$s per revolution. The particles are typically collected into four bunches of each type with approximately 22$\mu$s between the circulating bunches. Each bunch is approximately 2.0 cm long, 0.1 cm wide in the vertical direction and 0.4 cm wide in the horizontal direction when not in collision and is focused down to 190$\mu m$ wide and 4$\mu m$ tall when the beams are brought into collision. Typical beam currents are of the order of 2.5 mA which corresponds to approximately 1x$10^{12}$ particles [33]. The instantaneous event rate for the colliding beams at a given IP is

$$\frac{dN}{dt} = \sigma \mathcal{L} \tag{3.1}$$

here $\sigma$ refers to the cross-section of the process of interest and $\mathcal{L}$ is the instantaneous luminosity given by:

$$\mathcal{L} = \frac{N_{e^+} N_{e^-} N_{bunch} \mathcal{F}}{4\pi \sigma_x \sigma_y} \tag{3.2}$$

where $N_{e^\pm}$ is the number of $e^\pm$ per bunch, $N_{bunch}$ is the number of bunches , $\mathcal{F}$ is the revolution frequency and $\sigma_x$ and $\sigma_y$ are the RMS beam sizes in the $x$ and $y$ directions at the interaction point. Typical instantaneous luminosities in 2000 were of the order of $10^{31}$ cm$^{-2}$s$^{-1}$.

The LEP project was split into two data taking stages 1989-1995 (LEP 1) and 1995-2000 (LEP 2). These two stages were separated by a shutdown period in which the LEP accelerator was upgraded to provide higher energies and increased luminosities with some components of the detectors also upgraded. During the LEP 1 stage, LEP accelerated $e^+$ and $e^-$ beams to energies of approximately 45 GeV thus allowing collisions at centre of mass energies at or near to the Z peak.

In this first phase an integrated luminosity of approximately 200 $\text{pb}^{-1}$ was delivered to each of the four experiments. The first phase of LEP allowed many detailed studies to be performed including precision measurements of electroweak parameters and investigations into aspects of QCD including the production and decay of hadrons containing heavy b and c quarks.

In the second phase, LEP 2, increasing centre of mass energies allowed further tests of the Standard Model to be performed and the search for heavier undiscovered particles to be undertaken. In 1997 the achievement of 161 GeV centre of mass energy allowed the production of $W^+W^-$ pairs and the study of the triple gauge coupling mechanism via $Z \rightarrow W^+W^-$. Increases in the centre of mass energy above 161 GeV allowed increasing reach in the search for new particles, including Supersymmetric particles and the Higgs boson.

The nature of the data set recorded in 2000 is quite different from those recorded in previous years. In previous years the LEP 2 accelerator was operated at one set centre of mass energy or in the case of 1999 at four specific centre of mass energies. The change in the operating procedure during 1999 was an attempt to gain the highest possible centre of mass energy thus aiding the search for new particles. This goal was continued in 2000 with a slight difference in the operation of the LEP accelerator in that no specific preferred centre of mass energies were chosen. The goal during 2000 was to gain a large amount of integrated luminosity at the highest possible centre of mass energy [34]. During 2000 the run procedure was to initiate physics runs at the highest possible energy with 2 RF cavities in reserve to ensure that the circulating beams were not lost in the case of an RF trip. After a stable period of running, (e.g. $\approx$ 1hour), the energy of the accelerated beams was raised to its maximum with no RF cavities left in reserve, thus the beams were lost at the first RF trip after this so called "mini-ramp".

The resulting performance of the LEP accelerator in 2000 may be shown by considering the distribution of the collected luminosity versus the centre of mass energy in figure 5.1. This clearly shows that, unlike the previous years, the data obtained in 2000 is recorded over a continuous range of centre of mass energies.

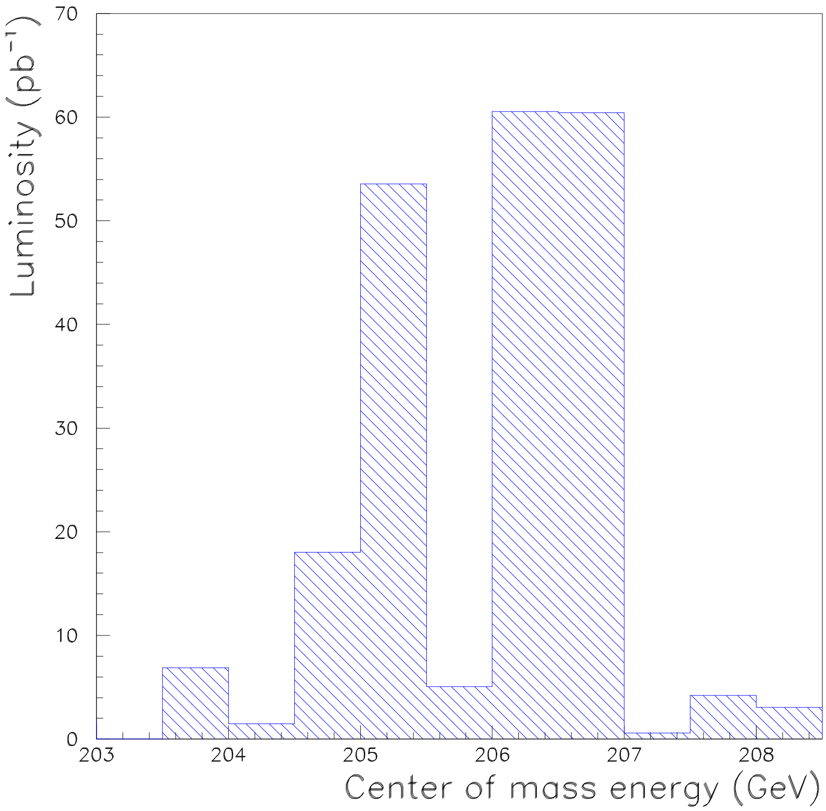


**Figure 3.3:** The distribution of luminosity collected in 2000 as a function of centre of mass energy.

## 3.3 The ALEPH detector

The ALEPH (Apparatus for LEp PHysics) detector was designed to study all types of SM processes accessible at LEP and to search for any new phenomena which may arise. To meet these goals ALEPH was designed to gather as much information as possible from each $e^+e^-$ interaction which it observed. ALEPH was thus designed to cover as much solid angle as possible with a high degree of hermeticity and granularity. A near full coverage of $\approx 3.9\ \pi$Sr was finally achieved. The ALEPH detector consists of several sub-detectors which are shown in the cut away view in figure 3.4.

The major sub-detectors are arranged in six cylindrical layers situated around the IP. Radially outwards from the IP the first three detectors are charged particle

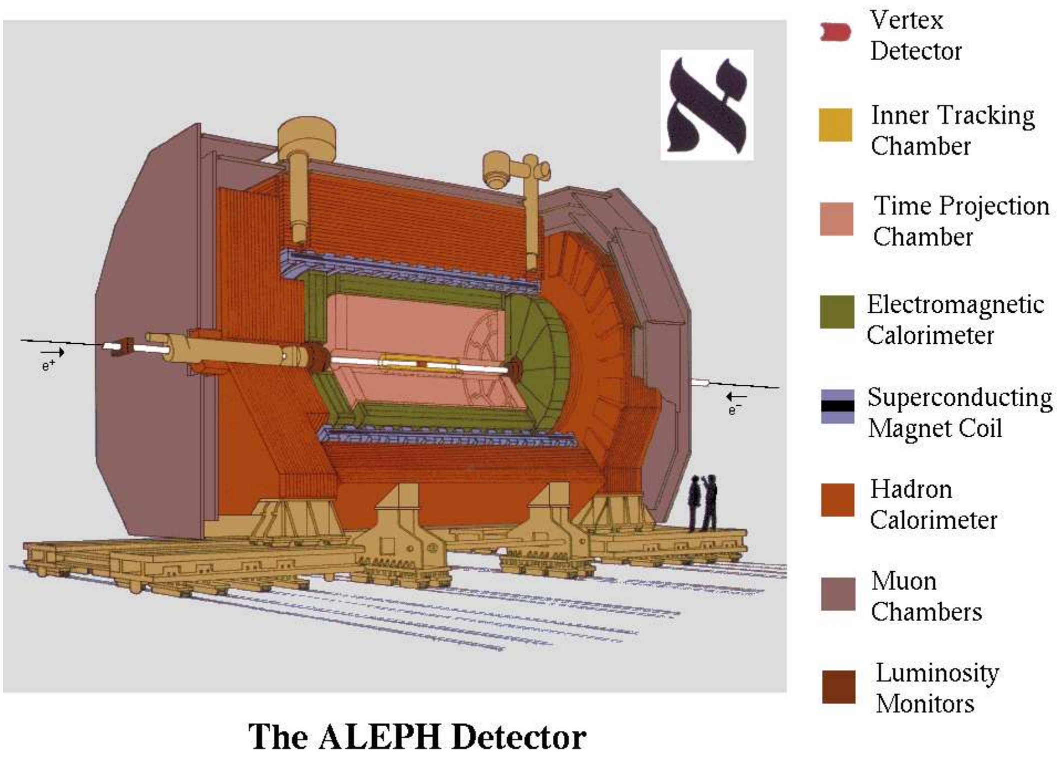


**Figure 3.4:** A cut away view of the ALEPH detector with the different sub-detectors highlighted and people added for scale.

tracking detectors and the fourth and fifth form calorimeter detectors to measure the particle energy whilst the last layer aids with muon particle identification($PID$).

Radially outwards from the IP the sub-detectors are

- VDET, a vertex detector designed to provide high precision spatial co-ordinates for charged particles very close to the interaction point;
- ITC, Inner Tracking Chamber formed from a multi-wire drift chamber designed for fast tracking and trigger information;
- TPC, Time Projection Chamber, a large drift chamber to provide accurate three dimensional tracking information for charged particles and ionisation information for $PID$;
- ECAL, the electromagnetic calorimeter to identify and measure the energy of electrons, positrons and photons;

- HCAL, the hadronic calorimeter to determine the energies of hadrons and aid with the identification of muons;
- Muon chambers, to detect and provide positional information for muons

The ALEPH co-ordinate system is shown in figure 3.5. The co-ordinate system has its origin at the theoretical beam crossing point which is defined by the midpoint of the straight section between the two nearest LEP quadrupoles. The positive $z$-axis is along the $e^-$ beam direction and makes an angle of +3.5875 mrad (upwards) with the local horizontal. The positive $x$-axis points towards a vertical line through the LEP centre and is horizontal by definition. The $y$-direction is orthogonal to $x$ and $z$ and points upwards. Spherical co-ordinates are defined as

$$x = rsin\theta cos\phi,$$
$$y = rsin\theta sin\phi,$$
$$z = rcos\theta.$$

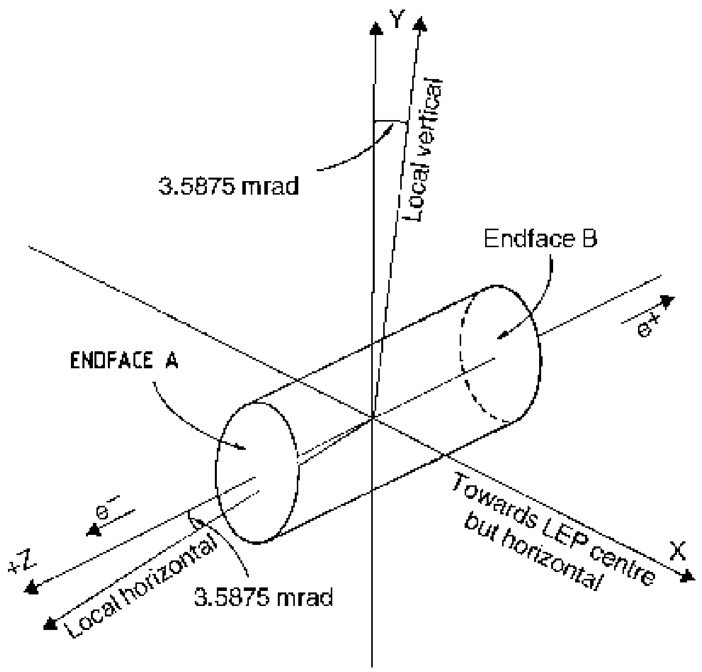


**Figure 3.5:** A schematic display of the reference system of ALEPH co-ordiantes.

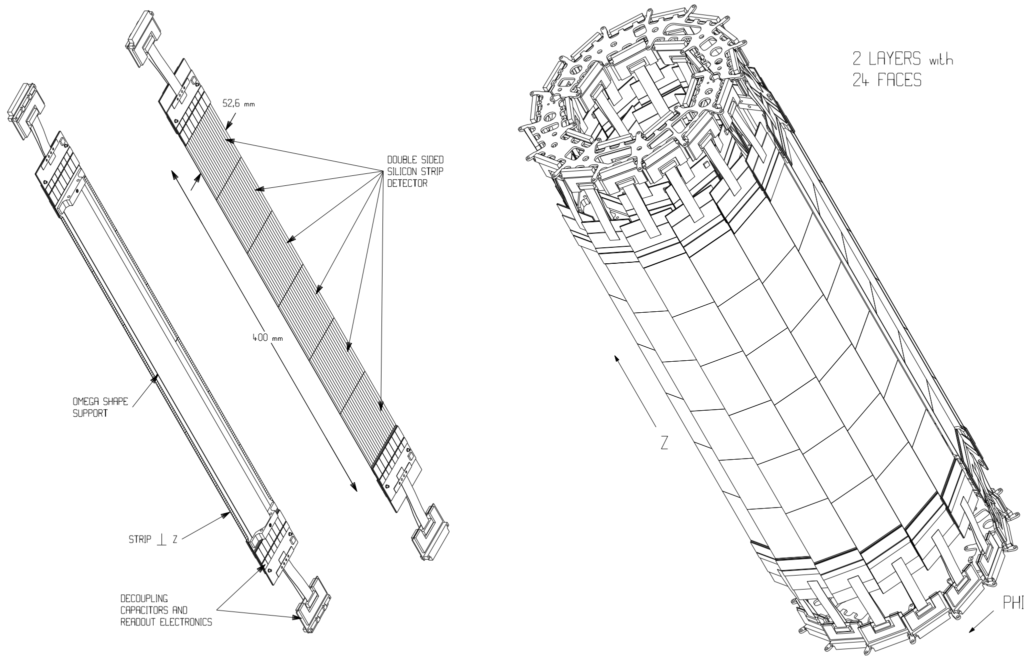


**Figure 3.6:** The VDET showing the mounting of the faces with their overlapping structure and an individual face in detail.

### 3.3.1 The vertex detector

The Vertex DETector (VDET) [35], shown in figure 3.6, is the closest of the main sub-detectors to the interaction point and is used to provide high precision measurements of charged tracks within an event. These high precision measurements are needed to enable the reconstruction and identification of short lived hadrons containing heavy c and b quarks. The upgraded VDET installed in 1995 provides extended angular coverage with less passive material and improved radiation hardness. The VDET consists of two concentric layers, at radii of 6.3 and 11.0 cm, of silicon wafers with double sided readout. Adjacent faces are arranged to overlap 0.2 cm in the $\phi$ direction to avoid any loss of coverage due to possible transition of particle through the wafer edges. The inner and outer layers contain 9 and 15 faces respectively, each of which lies parallel to the beam axis and consists of six 6.5 cm long 300 $\mu m$ thick silicon wafers. An angular acceptance of $|cos\theta| \leq 0.95$ is achieved for tracks which have at least one hit in the VDET. The wafers are divided into strips of 25 $\mu m$ in pitch with each strip acting as a reverse biased p-n diode for detecting the passage of charged particles. The motion of the electron hole pairs produced by the passage of charged particles induces signals over several strips due to capacitive coupling. Therefore only every second strip is equipped with readout electronics and the coordinates for the charged particle are obtained by interpolation with negligible loss in precision. The p junction sides provide $r\phi$ coordinates whilst the $n^+$ side provides the $z$ coordinate. Every second strip is read out thus a readout pitch of 50 $\mu m$ in $r\phi$ and 100 $\mu m$ in $z$ is achieved. The use of interpolation techniques gives a final $r\phi$ and $z$ resolution of 10 $\mu m$ and 15 $\mu m$ respectively for tracks perpendicular to the Si wafers. The $r\phi$ and $z$ measurements are independent and are paired together to form a 3D point during the offline event reconstruction. Figure 3.7 shows the resolution as a function of $\theta$.

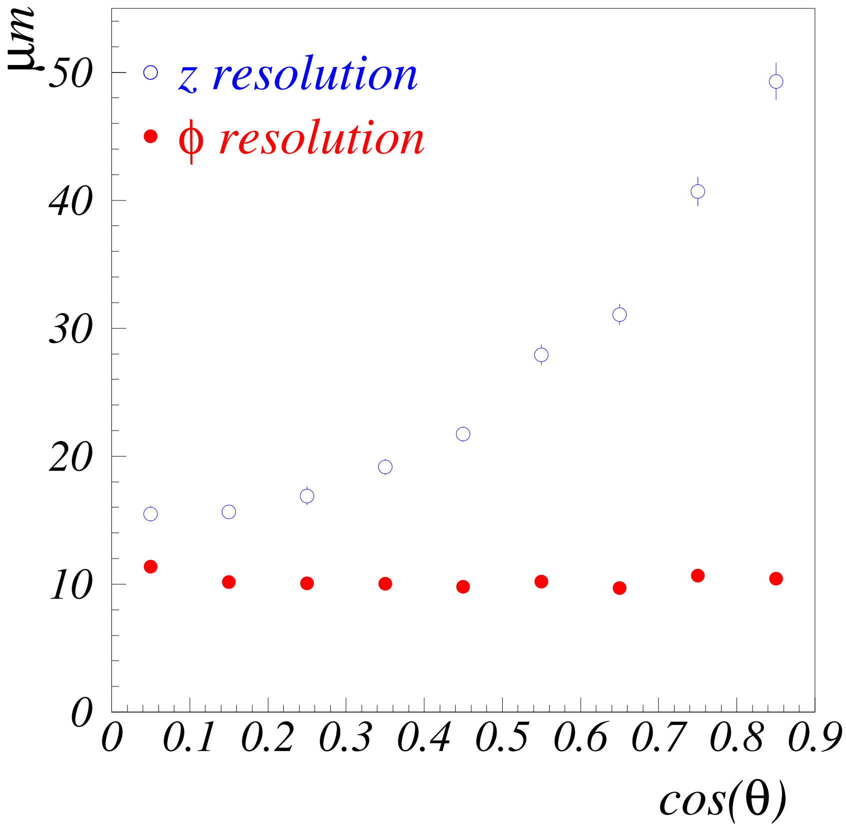


**Figure 3.7:** The single-hit resolution of the VDET as a function of $\theta$.

### 3.3.2 The inner tracking chamber

The Inner Tracking Chamber (ITC) [36] forms the second layer of the ALEPH tracking system. The ITC is a cylindrical multi-wire drift chamber with inner and outer radii of 12.8 cm and 28.8 cm and an active length of 2 m. The chamber is filled with a gas mixture of 80% argon and 20% $CO_2$ at atmospheric pressure. A small amount of alcohol is also present since it has been found to slow the aging process of the chamber.

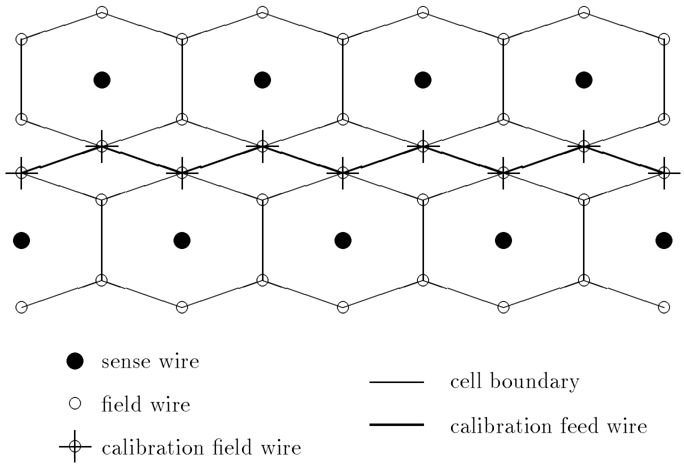


**Figure 3.8:** A schematic diagram of the ITC cell structure.

The ITC contains 960 anode sense wires running parallel to the beam axis. Each sense wire is held at a potential of approximately 1.8 kV and is surrounded by six earthed wires to form a hexagonal drift cell, see figure 3.8. The cells are arranged in eight concentric layers with 96 wires in each of the four inner layers and 144 wires in each of the four outer layers. Guard cages are formed between every two layers of drift cells by using copper/beryllium guard wires and aluminium hoops to catch any broken wires. An end on view of the wire configuration of the ITC can be seen in figure 3.9. The $r\phi$ resolution of the ITC is determined from the properties of the gas, e.g. drift velocity, with an average resolution of approximately 150 $\mu m$ being achieved. The $z$ coordinate of the track is determined from time of arrival of the signal at each end of the sense wire. The $z$ coordinate is thus determined with an average resolution of approximately 5 cm. A maximum of eight coordinates per track may be provided giving three dimensional information about the track trajectory. Tracks with $|cos\theta| < 0.97$ will traverse all eight drift cell layers and provide the maximum number of track coordinates.

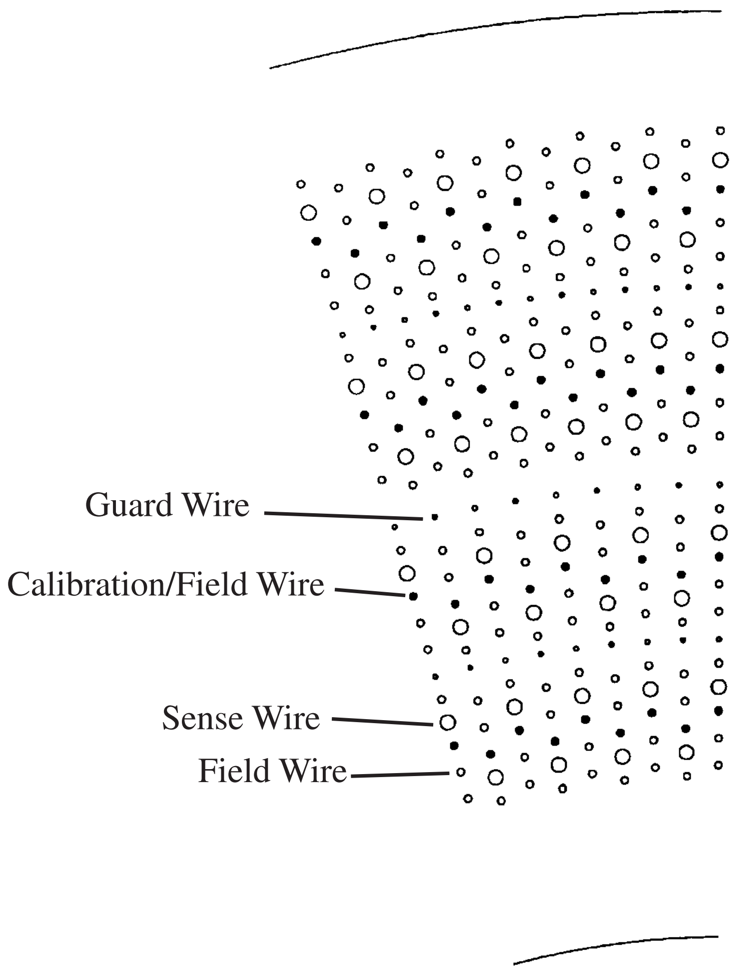


**Figure 3.9:** A schematic diagram of the ITC wire structure.

The small size of the drift cells, with a maximum drift distance of 6.5 mm, allows the tracking information provided by the ITC to be available extremely quickly. The ITC thus provides the only tracking information for the first level trigger, see section 3.4.1. Two dimensional tracking information in the $r\phi$ projection is available within 1 $\mu$s of the beam crossing while full 3D information is available within 2-3 $\mu$s.

### 3.3.3 The time projection chamber

The Time Projection Chamber (TPC), figure 3.10, forms the third layer of the ALEPH tracking system and is the main tracking detector in ALEPH. The TPC is a long cylindrical drift chamber of 4.7 m in length and covers a radial region between 31 cm and 1.8 m. The large dimensions of the TPC allow it to have excellent momentum resolution for particles including those of the highest energy. The TPC chamber is filled with a mixture of 91% argon, which provides desirable ionisation properties, and 9% methane to prevent run away avalanches.

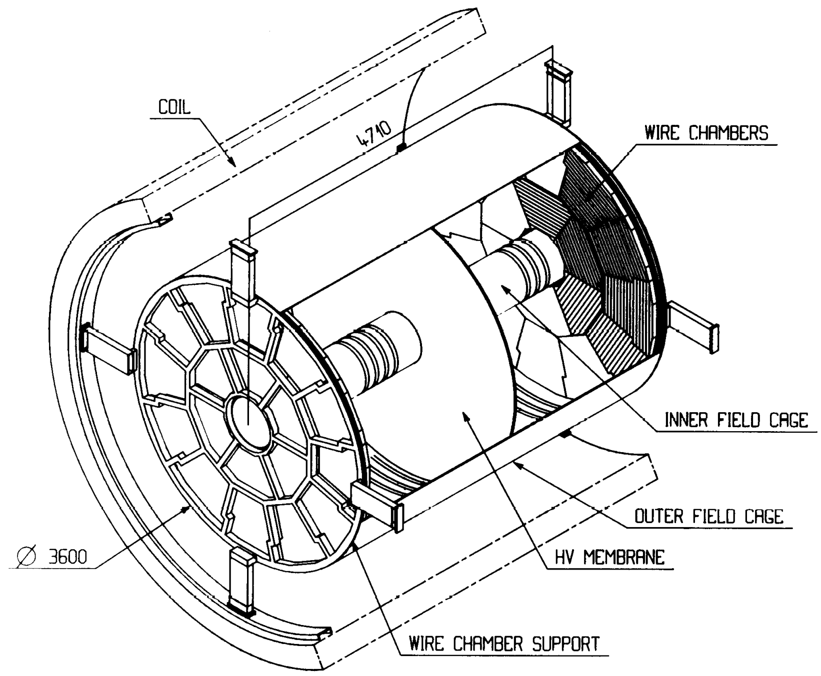


**Figure 3.10:** A cut away schematic of the Time Projection Chamber with the wire chambers, field cages and central membrane all indicated. The attachment of the TPC to the Solenoid is also shown.

The chamber is held at a slight overpressure between 8 and 12 mb above atmospheric pressure to prevent contamination from the atmosphere. A central membrane constructed from a 25 $\mu m$ thick mylar sheet, with a conductive graphite paint coating, is placed at the centre of the TPC. This central membrane is maintained at a potential of -27 kV whilst the ends of the TPC are held at a potential near ground. The surface of the inner and outer cylinders which form the TPC are referred to as field cages. These field cages contain ring shaped copper electrodes which are connected in series by resistor chains to the central membrane and are used to shape the electric drift field within the TPC. This arrangement forms a highly uniform electric field of magnitude 11 kV/m between the central membrane and the ends of the TPC with the electric field arranged so that it is parallel to the magnetic field from the superconducting solenoid (see section3.3.5).

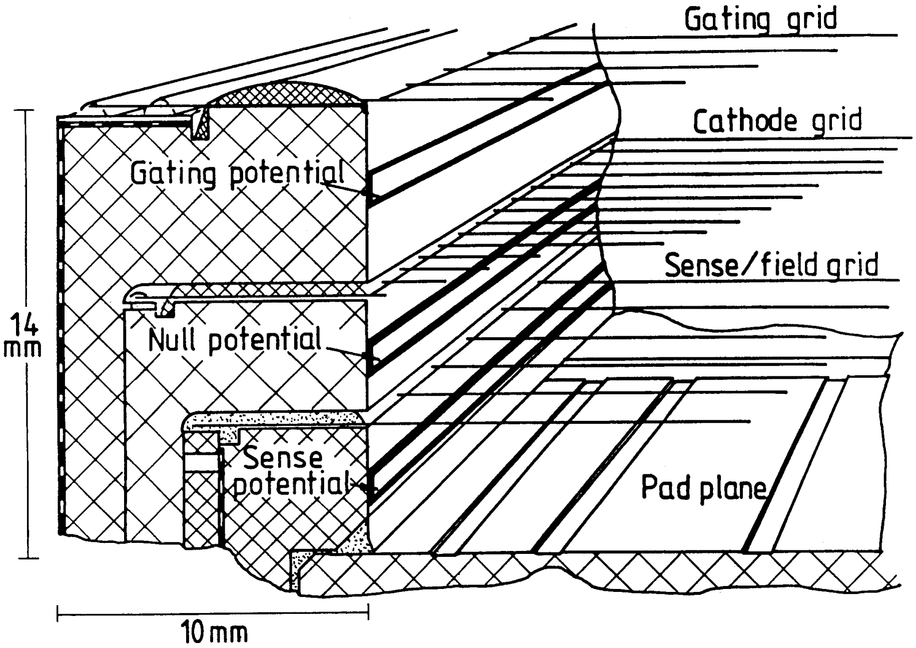


**Figure 3.11:** Schematic of a TPC sector edge showing the wire planes and the pad plane

Charged particles which traverse the TPC will leave ionisation trails in the gas. The ionisation electrons produced by the traversal of charged particles will drift towards the end plates of the TPC with a uniform velocity of 5.2 $\mu m.s^{-1}$. Each end of the TPC contains a system of proportional wire chambers arranged into 18 sectors to detect the signal produced by a charged track. The recorded signal is read out by a system of cathode pads arranged into 21 circular layers. In each sector these planes of wire chambers are arranged over the cathode pad plane as in figure 3.11. A cathode wire plane held at 0 V potential forms the boundary of the drift volume. Anode sense wires at +1250 V collect electrons from avalanches initiated by the ionisation electrons. The cathode pads underneath the sense wires collect an induced signal which is used to provide an accurate $r\phi$ track coordinate by the use of signal interpolation between different pads. The $z$ coordinate is deduced from the arrival time of the signal pulse and the known electron drift velocity in the TPC. The 21 concentric rings of pads allow up to 21 three dimensional coordinates to be provided for each charged particle which fully traverses the TPC. The layer of wires above the cathode wire plane is referred to as the gating grid. The purpose of the gating grid is to prevent positive ions, produced in the avalanche close to the sense wires, from entering the main chamber and distorting the electric field. In the "open" state the gating wire is held at a negative potential of -67 V making it transparent to the passage of charged particles. In the "closed" state alternate wires are set to potentials of $-67 \pm 100$ V so that the resulting dipole field renders the gate opaque to the passage of charged particles. The gate is opened 2 $\mu$s before each beam crossing and if the first level of the event trigger system (see section 3.4.1) returns a positive signal, after 5 $\mu$s, the gate is held open for 45 $\mu$s. The open time for the gate corresponds to the maximum drift time for the ionisation electrons in the TPC. If a negative signal is returned for the first level trigger the gate is closed.

The $r\phi$ coordinate resolution depends on the drift length and the angle of the track segment with respect to both the wires and the pads with an achieved resolution of 173 $\mu m$ for leptonic $Z^0$ decays. The resolution of the $z$ coordinate depends on the polar angle and an achieved resolution of 740 $\mu m$ is found for tracks with polar angles greater than $80^\circ$ to the beam axis.

The TPC is also equipped with a Nd-YAG laser calibration system. The laser provides thirty straight ionisation tracks which may be used to correct inhomogeneities in the electric and magnetic fields and to monitor drift velocities. The laser system is fired between beam crossings in order to monitor time dependent effects present during data taking, see figure 3.12.

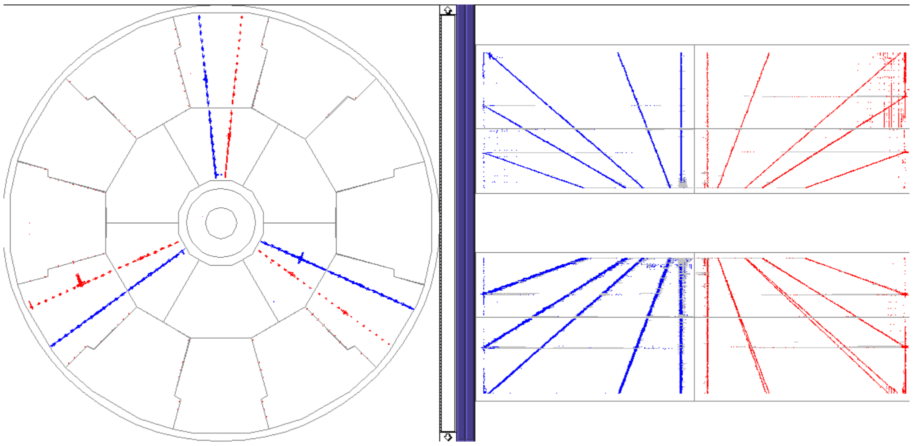

**Figure 3.12:** An example of a TPC laser event [37].

### 3.3.4 The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is a sampling calorimeter used primarily to detect the energy deposited by electrons, positrons and photons with some sensitivity to energy deposited by hadrons. The ECAL, shown in figure 3.13, consists of a 4.8 m long barrel placed around the TPC with an end-cap to cover each end of the detector and provide near full solid angle coverage of $\approx 3.9\ \pi$Sr. The barrel and end-caps consist of 12 modules, each of which covers 30° in $\phi$. Each of these modules consists of 45 interleaved layers of lead sheets and wire chambers with a total of approximately 40 cm of lead corresponding to 22 radiation lengths, $X_0$.

The angular coverage of the ECAL is slightly degraded due to the presence of so called "cracks" which exist between the ECAL modules. These ECAL cracks account for 2% of the inner surface of the barrel and 6% of the end-cap surface. To minimize the effect of these cracks the end-cap modules are off-set by 15° in $\phi$ with respect to the barrel and all modules are rotated by 1.9° in $\phi$ with respect to the hadronic calorimeter.

The structure of the ECAL layers is shown in figure 3.14. The proportional wire chamber is filled with a gas mixture of 80% xenon and 20% $CO_2$. The passage of charged particles through the ECAL causes electromagnetic showers to develop in the lead sheets which then causes ionisation avalanches to form near to the anode wires. These avalanches induce signals on small cathode pads which

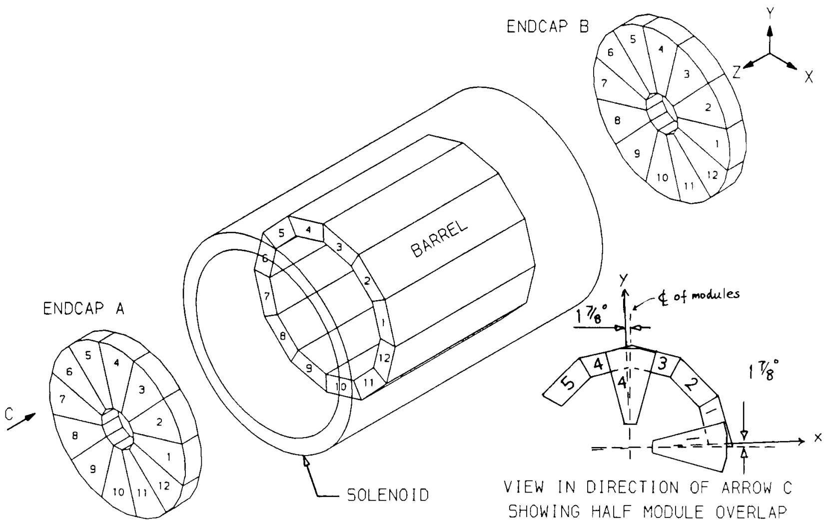


**Figure 3.13:** A schematic view of the ECAL barrel and end-caps with the surrounding solenoid. The rotation of the end-caps is also depicted to show the half module overlap.

are connected to form towers which point towards the interaction point and each has an angular coverage of 0.9° x 0.9°. The towers are divided into three sections along their length, known as storeys, corresponding to $4X_0$, $9X_0$ and $9X_0$ radially outwards respectively. These storeys allow the shower profile within the ECAL to be studied and thus further aid with particle identification.

There are a total of 49512 towers in the barrel and 24576 towers in both end-caps. The high number of towers provides a high granularity and thus aids the spatial resolution of the electromagnetic showers and also particle identification.

The energy resolution of the ECAL has been studied using Bhabha scattered electrons by comparing their measured ECAL energy with the track momentum and(or) beam energy. The resolution has been parameterised as

$$\frac{\sigma(E)}{E} = \frac{0.18}{\sqrt{E(GeV)}} + 0.009 \tag{3.3}$$

The angular resolution of the ECAL which is important for the identification of electrons, photons and neutral pions has been found to be

$$\sigma_{\phi,\theta} = 0.25 + \frac{2.5}{\sqrt{E(GeV)}} \text{ mrad.} \tag{3.4}$$

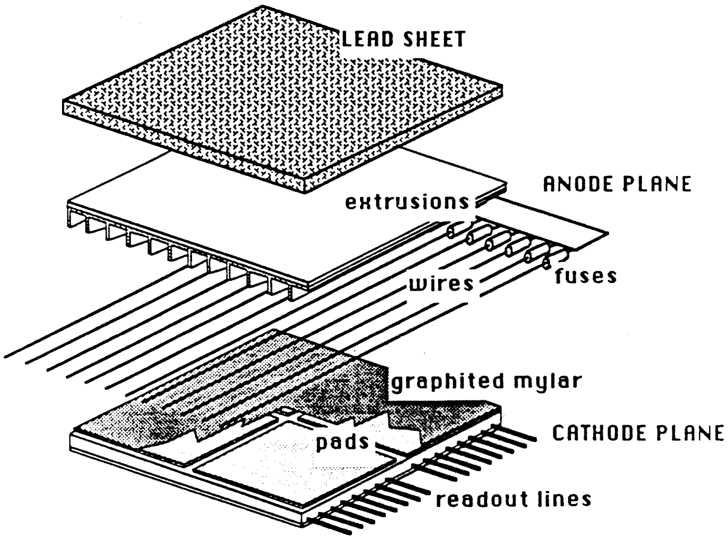


**Figure 3.14:** An expanded view of the composition of a lead/wire chamber layer of the ECAL. The lead sheet, aluminium extrusion, anode wires and cathode pad plane with readout are shown.

### 3.3.5 The superconducting solenoid

The magnetic field for the ALEPH detector is provided by a solenoidal superconducting coil positioned between the ECAL and HCAL, see figure 3.15. The superconducting coil is formed from a continuous winding of a NbTi/Cu conductor along a length of 6.35 m and is cooled to a temperature near 4.4 Kelvin using liquid helium. Compensating coils are located near to both ends of the main coil. The main coil has a full nominal current of 5000 A and produces a 1.5 T uniform axial field parallel to the beam axis. When operating with the compensating coils the field within the tracking volume has a homogeneity of $\frac{\Delta B_Z}{B_z} <$ 0.2% and radial and azimuthal field components of $\frac{B_r}{B_z} <$ 0.4% and $\frac{B_\phi}{B_z} <$ 0.04% respectively.

### 3.3.6 The hadronic calorimeter

The hadronic calorimeter (HCAL) is a sampling calorimeter that is 7.2 interaction lengths, $\lambda_{int}$, thick and comprises of 23 layers of iron and gas-filled streamer tubes. The HCAL is used to detect the energy deposited by hadrons, aid in the identification of muons and also serves as a return yoke for the magnetic field produced by the superconducting solenoid.

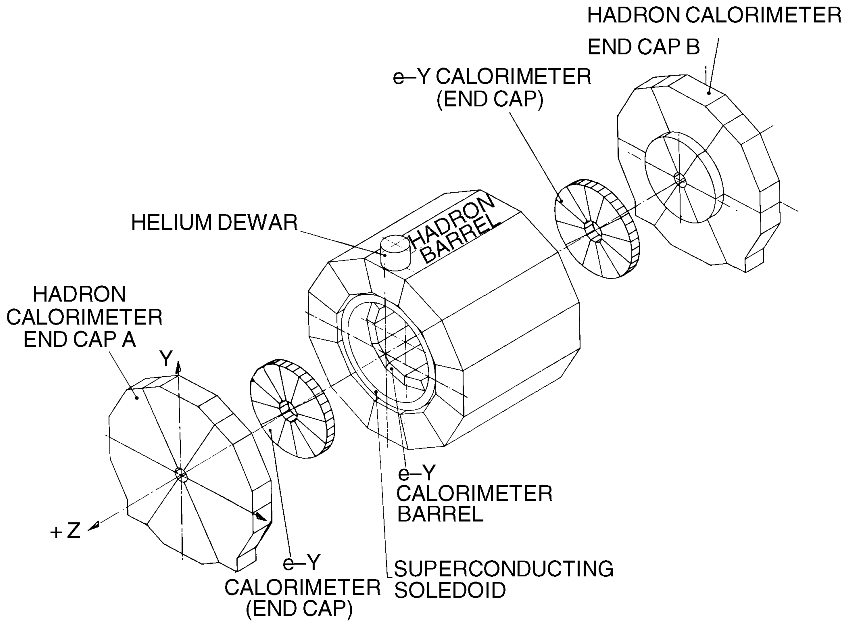


**Figure 3.15:** A view of the HCAL in place around the solenoid and ECAL with the end-caps of both calorimeters displayed.

The HCAL, as shown in figure 3.15, is similar in structure to the ECAL being formed from a central barrel with two end-caps to give a full coverage of approximately 99% of 4 $\pi$Sr. The barrel has inner and outer radii of 3.0 m and 4.7 m respectively and is 7.34 m in length. The barrel itself is segmented into twelve modules whilst the end-caps have only six.

The structure of the HCAL layers is shown in figure 3.16. The streamer tubes contain a gas mixture of 13% argon and 57% $CO_2$ and 30% butane. The streamer tubes are similar to proportional wire chambers but operate at higher voltages, thus the size of the signal is independent of the particle energy. The total charge collected is therefore proportional to the number of particles within the hadronic shower. Each streamer tube consists of eight cells with a square 9 mm x 9 mm cross-section separated by 1 mm thick walls. An anode wire held at a potential of 4250 V runs along the middle of each cell. As with ECAL, cathode pads are connected into projective towers that point towards the interaction point. The pads have an angular coverage of $\Delta\phi$ x $\Delta\theta$ = 3.7° x 3.0° in the barrel and 7.5° x 2.7° or 15° x 2.5° in the end-caps.

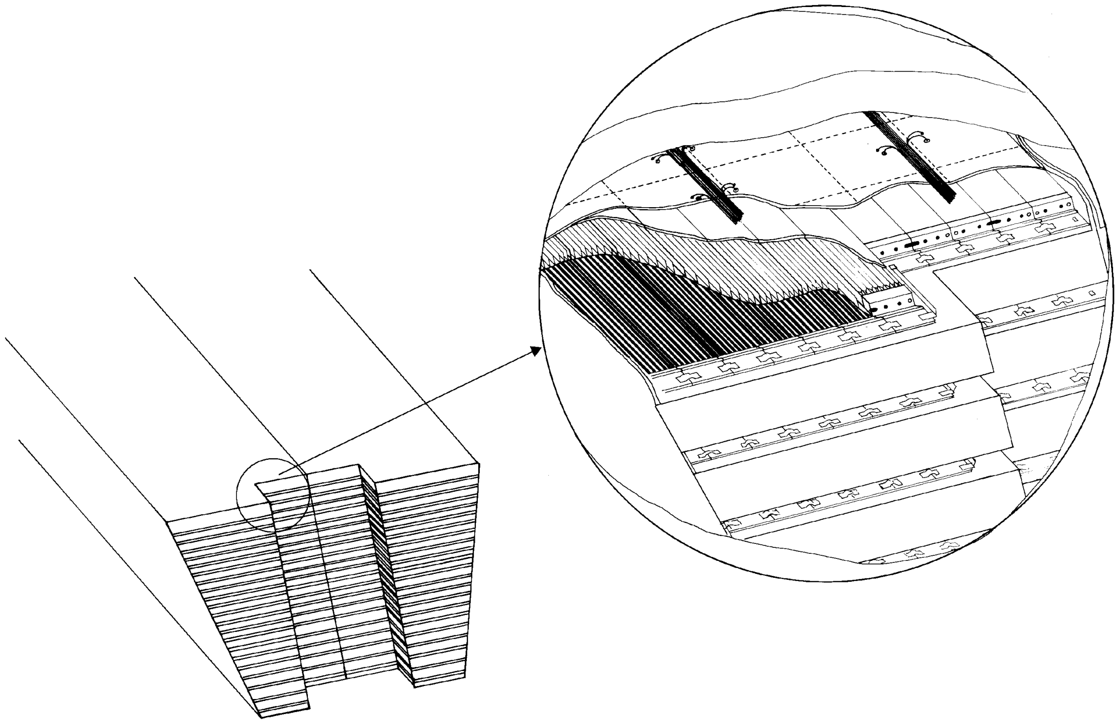

**Figure 3.16:** A view of the HCAL layers showing the position of the streamer tubes between the iron plates.

The energy resolution for pions at normal incidence is

$$\frac{\sigma(E)}{E} = \frac{0.85}{\sqrt{E(GeV)}} \tag{3.5}$$

On the cathode plane opposite to the pads a 4 mm wide strip runs parallel to each wire and provides a digital signal if a streamer tube is fired. This provides a projection of the development of a shower in the $r\phi$ plane and is useful in the identification of muons which pass through the HCAL leaving only minimum ionisation trails.

### 3.3.7 The muon chambers

To further aid the identification of muons two double layers of streamer tubes, referred to as muon chambers, are added outside the HCAL. The two double layers are separated by a distance between 40 and 50 cm. The placement of the muon chambers follows from the layout of the HCAL barrel and end-cap sections but in addition there are so called middle angle chambers placed over the outer edges of the petals thereby covering the region of overlap between the barrel and end-caps.

The muon chambers do not provide energy measurement but rather act as tracking detectors and give $x$ and $y$ coordinates for the tracks by means of strip detectors. The configuration of one of the muon chambers is shown in figure 3.17. Particles which traverse the muon chambers leave up to four hits, from which the track directions can be measured with an accuracy between 10 and 15 mrad.

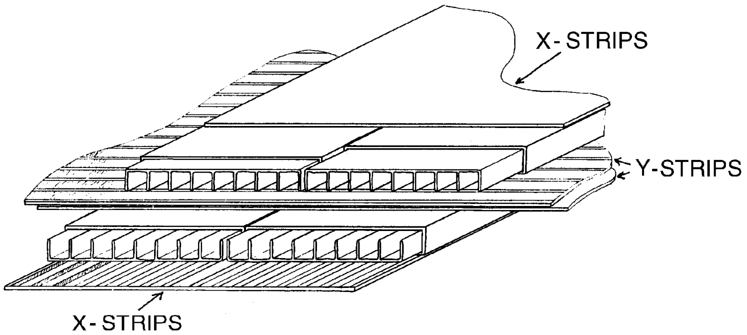


**Figure 3.17:** A view of one of the Muon Chamber double layers showing the orthogonal positioning of the readout strips over the streamer tubes.

### 3.3.8 The luminosity monitors

An accurate determination of the integrated luminosity delivered by LEP to ALEPH is essential for physics analyses. The luminosity is determined using elastic $e^+e^-$ (Bhabha) scattering since the cross-section is known theoretically to a high precision.

The rate of Bhabha scattering is measured by the simultaneous detection of the $e^+$ and $e^-$ on opposite sides of the detector. The luminosity is given by the ratio of the measured rate to the theoretical cross section.

The cross section for Bhabha scattering is strongly peaked at polar angles close to the beam axis. Following from this, three specialised calorimeters are installed close to the ALEPH beam pipe to provide the luminosity measurement.

The primary luminosity measurement for ALEPH in the LEP 2 era is provided by the luminosity calorimeter (LCAL). LCAL consists of two 38 layer lead/wire chamber sampling calorimeters and is similar in design to the ECAL. LCAL is situated at $\pm$2.62 m from the interaction point (on either side) and covers a region of radii between 10 cm and 52 cm. Constructed as such, LCAL can detect Bhabha electrons and positrons with scattering angles between 45 mrad and 160 mrad and has an energy resolution of $\frac{\sigma_E}{E} = \frac{15\%}{\sqrt{E}}$ and a positional resolution in $x$ and $y$ of 0.43 mm + $\frac{6.5mm}{\sqrt{E}}$. LCAL luminosity measurements have a typical uncertainty of 0.5%.

The Silicon Calorimeter (SICAL) is the most accurate of the luminosity detectors and provided the primary luminosity measurement during the LEP 1 era. SICAL consists of two detectors placed at 2.5 m on either side of the IP. Each detector is a sampling electromagnetic calorimeter consisting of 12 layers of tungsten alternating with layers of silicon pad detectors. SICAL has inner and outer radii of 6 and 15 cm respectively and a corresponding angular coverage between 25 and 58 mrad. SICAL achieves an energy resolution of $\frac{\sigma_E}{E} = \frac{34\%}{\sqrt{E}}$ with a radial position resolution of 150 $\mu m$ and an azimuthal angular resolution of 3.8 mrad. Luminosity measurements with SICAL have a typical uncertainty of 0.09%. The introduction of tungsten masks inside the beam pipe to shield the inner ALEPH detectors from the increased synchrotron radiation present at LEP 2 and a change in the bunch running mode for LEP 2 resulted in the reduction of the SICAL performance as a luminosity monitor. LCAL is thus used as the primary luminosity monitor in LEP 2.

The third luminosity monitor, the Bhabha Calorimeter (BCAL), has two modules situated at 7.7 m on either side of the IP. Each module is formed from 12 tungsten/scintillator layers. The rate of Bhabha events observed by BCAL is much higher that those of LCAL and SICAL (order 20 times) due to the low angle positioning of BCAL and the forward peak in the Bhabha cross-section. However the position of BCAL is close to a LEP focusing quadrupole. This leads to some uncertainty of the polar angle of the electrons and positrons since the field of the quadrupole changes during LEP operation and also to an increased rate of beam related noise. Therefore BCAL is calibrated relative to LCAL and is used to provide instantaneous on line luminosity measurements and also to monitor the beam background conditions.

The luminosity monitors are also used to extend the angular coverage of the ALEPH calorimeters down to 40 mrad.

## 3.4 Data collection and reconstruction

### 3.4.1 The ALEPH trigger system

The bunch crossing rate of electrons and positrons in LEP is of the order of 45 kHz. At this rate it is impractical, both in terms of physical data storage and detector dead times, to read all the information from the sub-detectors at each beam crossing. A trigger system is therefore used to filter out events resulting from genuine $e^+e^-$ interactions from background events caused by cosmic rays or the interaction of the LEP beams with residual gas within the beam-pipe. The ALEPH trigger system is formed from three levels each of which uses more information than the last to make a logical decision to keep the data from a particular event or not.

The three stages of the ALEPH trigger, as implemented in 2000 [33] (with the original LEP 1 figures also noted [26]), are briefly summarised in table 3.1.

| Stage | Decision Time ($\mu$s) | Rate (Hz) | Information Used |
|---|---|---|---|
| Level 1 | 5 $\mu$s | 10-13 (Few 100) | Hit Patterns in ITC<br>Pad/wire readout<br>from ECAL+HCAL+LCAL |
| Level 2 | 50 $\mu$s | 2.5-4.5 (10) | + TPC tracking |
| Level 3 | 62 $\mu$s | 2-3 (1-3) | All sub-detectors |

**Table 3.1:** Summary of the ALEPH trigger system showing the decision time and sub detector information used within the trigger decision. The rates in brackets refer to the original LEP 1 trigger rates.

The first two levels of the ALEPH trigger system are hardware based while the third, and final level, is software based.

The first level trigger uses information from the ITC,ECAL, HCAL and LCAL detectors to produce a logical yes/no decision within 5$\mu$s of a bunch crossing. A standard bunch crossing time of 22$\mu$s means that the level 1 trigger introduces no detector deadtime. A positive trigger at level 1 maintains the TPC gate in an open position for 45$\mu$s to allow all the TPC information from the triggered event to be collected.

The second level trigger uses the same information as level 1 except that the ITC information is replaced with the full TPC track information. The second level trigger analysis is performed during the 45$\mu$s drift time for the TPC ionisation electrons and allows a logical decision to be made within 50$\mu$s of a beam crossing. If an event is rejected the detector is reset and is ready to accept data two bunch crossings after the event which initiated the triggering of the detector. If an event is accepted at level 2 a full data acquisition readout is initiated.

The final software based level 3 trigger uses raw digitised information from the whole event and takes advantage of correlations in the sub-detectors. If a positive decision is made at level 3 the whole event is written to tape.

The triggering efficiency for multi-hadronic events is essentially 100%.

### 3.4.2 Track reconstruction

A typical hadronic event recorded by ALEPH may have approximately 20 charged tracks which, in total, may leave several hundred coordinate hits within the ALEPH tracking system (VDET, ITC, TPC). The combination of these three subdetectors allows the trajectory of each charged track to be traced with a string of up to 31 3-dimensional points. Pattern recognition algorithms attempt to group the coordinates together to form track candidates. Points in the TPC are connected by requiring that they are consistent with a helix hypothesis, which is then extended to the ITC and VDET and fitted to points measured there. The results of the track finding procedures are passed on to Kalman filtering algorithms [38] which perform the final global track fitting, see figure 3.18.

The presence of a 1.5 T uniform magnetic field within the central tracking system forces the charged tracks to follow helical trajectories. The radius of the trajectory is inversely proportional to the transverse momentum of the charged particle where the transverse momentum is the particles momentum component perpendicular to the field.

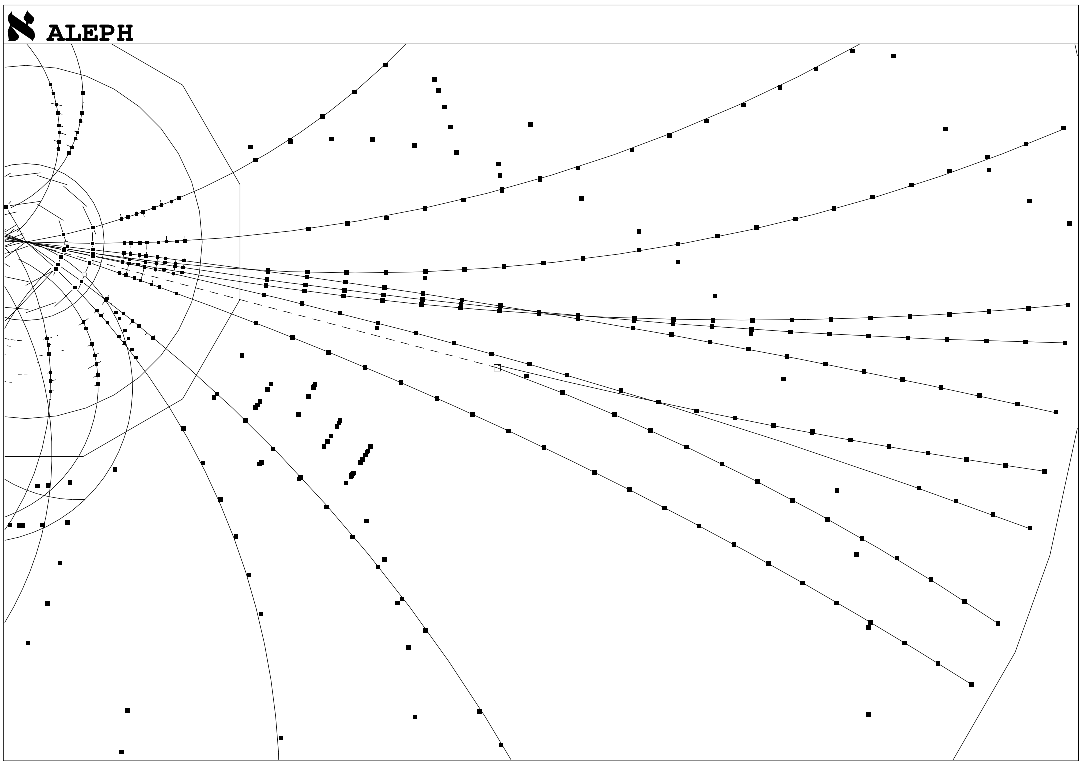


**Figure 3.18:** Track fits in ALEPH. The dots show the hits in the VDET,ITC and TPC while the solid lines represent the final result of the helix fit.

The following five parameters, illustrated in figure 3.19, are used to describe helical tracks within ALEPH:

- $\rho$ is the radius of curvature of the particle trajectory in the $xy$ plane. It is positive if the track bends counter-clockwise i.e. turning around the $z$-axis in a left handed sense.

- $\tan(\lambda)$ is the tangent of the angle of the particle trajectory with respect to the $xy$ plane ; $\lambda = 90^\circ - \theta$.

- $\theta_0$ is the angle in the $xy$ plane that the track makes with respect to the $x$ axis at the point of closest approach.

- $d_0$ is the distance of closest approach between the track and the nominal ALEPH origin in the $xy$ plane. It is signed according to the particle's angular momentum about the $z$ axis and that point.

- $z_0$ is the $z$ coordinate of the point of closest approach.

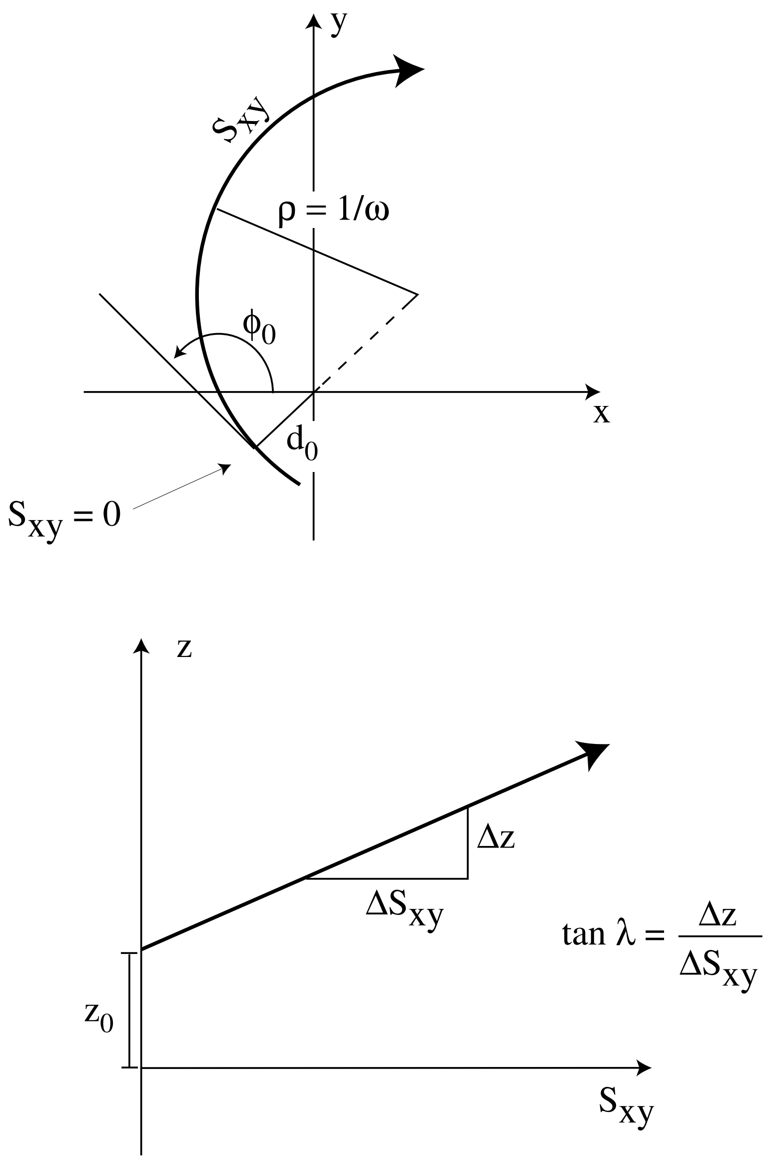


**Figure 3.19:** Definition of the helical track parameters used to describe charged particle tracks within ALEPH.

To evaluate the performance of the track reconstruction a sample of $Z \rightarrow \mu^+\mu^-$ events were used. Tracks are used if at least 19 TPC, 6 ITC and 1 VDET coordinates are found. The transverse momentum resolution from this sample is found to be [35]:

$$\sigma(p_t)/p_t = 0.6\text{x}10^{-3} \cdot p_t/\ \text{GeV/c} \tag{3.6}$$

The momentum resolution for the three tracking detectors is summarised in table 3.2. At low momentum a constant term of 0.005 is added to the resolution

| Detector | $\sigma(p_t)/p_t^2$ (% per GeV/c) |
|---|---|
| TPC only | 0.12 |
| TPC + ITC | 0.08 |
| TPC + ITC + VDET | 0.06 |

**Table 3.2:** Resolution of the track momentum as defined in equation 3.6. To illustrate, a track with $P_t$=45 GeV/c track will have a relative uncertainty in its momentum of 5.4% if only TPC information were used for the reconstruction [27].

due to the effect of multiple scattering. Since the error on the measurement of the polar angle is small, the relative error on the momentum coincides with the relative error on the transverse momentum. Most ALEPH analyses rely upon well measured tracks which satisfy the following criteria

- $|cos\theta| < 0.95$
- $\mathrm{N_{TPChits}} > 4$
- $|d_0| < 2.0$ cm
- $|z_0| < 10.0$ cm

### 3.4.3 $dE/dx$ measurements

The specific energy loss due to ionisation ($dE/dx$) for charged particles may be determined from TPC wire measurements. The wire signals may be associated with the reconstructed tracks. Tracks must be separated by at least 3 cm in the $z$ plane for a $dE/dx$ measurement to be resolved. Signals for which there is more than one compatible track are discarded. A $dE/dx$ resolution of 5.5% is achieved for minimally ionising pions with more than 150 wire measurements. Figure 3.20 shows the $dE/dx$ measurement for a given track sample and the average separation of particle types of a given momenta.

### 3.4.4 The energy flow algorithm

The energy flow algorithm [27] combines information about track momenta, calorimeter energy deposits and particle identification to provide improved energy resolution within an event. Energy flow utilises track momenta, TPC $dE/dx$, the shape

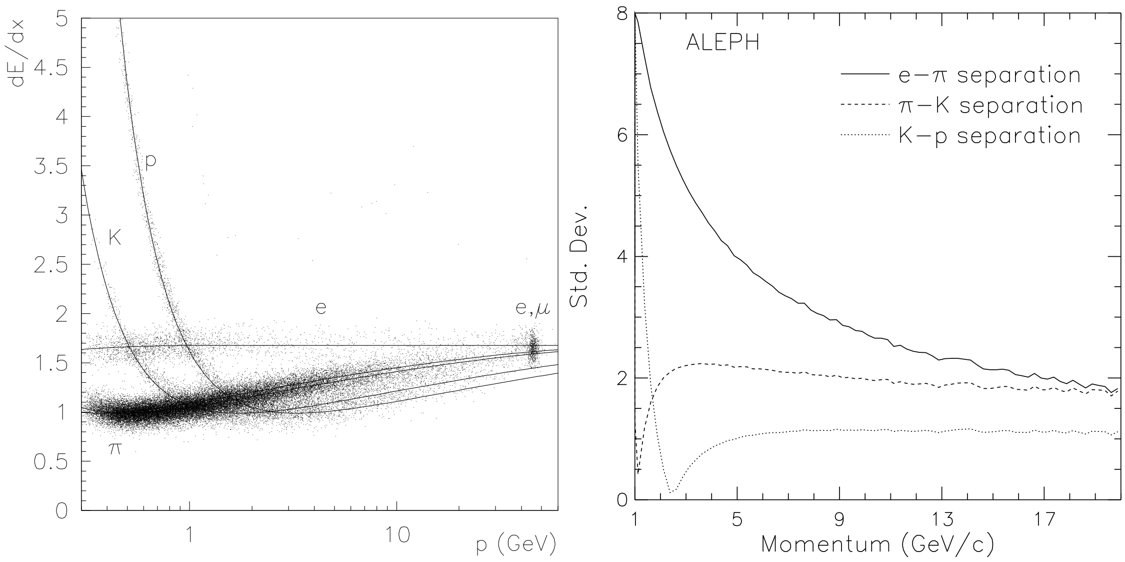


**Figure 3.20:** The left figure shows the measured $dE/dx$ versus particle momentum for a sample of about 40,000 tracks. Each track was required to have at least 150 $dE/dx$ measurements. The fitted parametrisation is shown for electrons, muons, pions, kaons and protons. The right figure shows the average $dE/dx$ separation in standard deviations between different particle types as a function of momentum. This has been computed using all tracks in hadronic $Z^0$ decays having at least 50 $dE/dx$ measurements. Taken from Reference [27].

of the energy showers deposited in the calorimeters, muon chamber information, and also energy deposited in the luminosity sub-detectors which enables the angular coverage to extend down to 40 mrad.

The algorithm uses this information to build a list of 'energy flow objects' (electrons, muons, photons, charged and neutral hadrons) which may be used for subsequent physics analysis.

The energy flow algorithm applies an initial event cleaning procedure to charged tracks and calorimeter clusters. The event cleaning is designed to reject badly reconstructed tracks and calorimeter noise before extrapolating good quality tracks to the calorimeters where they may be grouped with calorimeter energy deposits to form so called 'calorimeter objects'.

The charged tracks used within the energy flow algorithm must satisfy one of the following criteria:

1. The track must have at least four hits in the TPC and originate from a cylinder of length 20 cm and radius 2 cm coaxial with the beam and centred at

the nominal interaction point. Tracks with momentum exceeding 15 GeV require at least eight hits in the TPC and one hit in the ITC.

2. Charged particle tracks rejected by the above criteria may be recovered if they originate form a reconstructed $V^0$ [(1)] which is compatible with originating from the nominal IP within a cylinder of length 30 cm and radius 5 cm coaxial with the beam.

Tracks which fail the above criteria are subsequently ignored by the energy flow algorithm.

To clean the noisy channels of the electromagnetic and hadronic calorimeters noisy channels which appear systematically in many consecutive events are located prior to the event reconstruction and are not used in the cluster finding. The fake energy deposits due to occasional noise in the towers of the calorimeters are detected and removed when the corresponding signal is not compatible with the signal measured independently on the ECAL wire planes or on the HCAL streamer tubes.

After cleaning, the charged particle tracks are extrapolated to the calorimeters, and groups of topologically connected tracks and energy clusters are combined to form calorimeter objects.

Each of the calorimeter objects are processed according to the following steps:

- Charged particle tracks originating from the IP or belonging to a reconstructed $V^0$ are counted as charged energy assuming they are pions.
- Charged particle tracks identified as electrons are removed from the calorimeter objects list , together with the energy contained in the associated calorimeter towers. If the difference between the calorimeter energy and the track momentum is larger than three times the expected resolution the difference is assumed to be due to a bremsstrahlung photon and is counted as neutral electromagnetic energy.

[1]Here $V^0$ refers to the decay of short lived neutral particles (e.g. $K^0_s$, $\Lambda$ or photon conversions) within the detector volume.

- Charged particle tracks identified as muons are removed from the calorimeter object list, together with a maximum of 1 GeV from the closest associated electromagnetic energy cluster, if any exist, and a maximum of 400 MeV per plane fired around the extrapolation of the muon track from the corresponding hadronic calorimeter cluster.

- Identified photons and $\pi^0$'s are counted as neutral electromagnetic energy and are removed from the calorimeter object list.

- At this stage the only particles remaining in the calorimeter object list should be charged and neutral hadrons. The charged hadrons are easily identified by matching to the remaining charged tracks, where the pion mass is assumed. Neutral hadrons are identified as a significant excess of calorimeter energy. The sum of the energy remaining in the calorimeters, after scaling the ECAL contribution by the ratio of the response to electrons and pions, is compared to the energy of the remaining charged particle tracks. An excess, that is both larger than the expected calorimeter resolution and 0.5 GeV, is then counted as neutral hadronic energy.

## 3.5 Monte Carlo simulation

The development of the Higgs boson search analyses and interpretation of results from these searches relies heavily on the generation of large samples of simulated events referred to throughout this thesis as Monte Carlo (MC). These MC simulations are used in the training and optimisation of the search analyses, to parameterise the distributions of the discriminating variables and determining the event selection efficiencies which are used in the confidence level calculations.

The production of these simulated events is implemented in three stages. Firstly electron-positron interactions are simulated using one of several programs, known as generators, according to the particular physical process required. The results of this simulation is the kinematics of all the particles in the final state of the simulated event. The final state information is then passed to the second stage of the process in which the response of the ALEPH detector to these final state particles is simulated. This is achieved using the **GALEPH** program [39], which is based on the **GEANT** [40] package. **GALEPH** simulates the interaction between the final state particles and the sub-detectors and the corresponding response of these detectors. The output from the **GALEPH** program has the same form as a real data event recorded using the ALEPH detector allowing the simulated event to be reconstructed in exactly the same way as data events. The third and final part of the simulation, the reconstruction, uses the **JULIA** [41] package in both the case of data and simulation. The only difference between the MC simulations and real data is the presence of information about the true process, (e.g. the actual particle type,momentum,mass...), which is contained within the MC simulation. This information may be used to aid studies and optimisations of the analyses which are performed using the simulated MC events.

The **JETSET 7.4** package [42] was used to model the hadronization of the simulated partons. In the **JETSET** model a colour string is assumed to join two coloured partons and once this string is stretched past a particular energy threshold it breaks forming a new $q\bar{q}$ pair.

Numerous generator programs exist, those used in the simulation of the relevant backgrounds for the four jets cuts based search as presented in this thesis are summarised here. The $q\bar{q}$ background was generated using the **KORALZ 4.02** generator [43]. The $W^+W^-$ background[2] was generated using the **KORALW 1.21** generator [44]. The ZZ background was produced using the **PYTHIA 5.7** generator [42].

The signal simulation was performed using the **HZHA03** generator [24] which includes all possible final states arising from the decays of the Higgs and Z bosons. The interference between the Higgsstrahlung process and weak boson fusion process is also considered.

The simulated sample sizes typically correspond to many times the actual collected integrated luminosity, with the simulated samples generated for 2000 corresponding to at least 50 times the actual data set. The generated samples for 2000 are summarised in table 3.3. The background samples are typically generated in

[2]The generation of the $W^+W^-$ background was restricted to the CC03 diagrams

1 GeV steps in $\sqrt{s}$ with special high statistics samples generated at $\sqrt{s}$=206.7 GeV to coincide with the majority of the collected luminosity. The signal samples are generated in 2 GeV steps in $\sqrt{s}$ with an additional set at $\sqrt{s}$=206.7 GeV. The signal samples are generated at various values of $m_{\mathrm{h}}$ over a full range of Higgs boson masses as described in table 3.3, with extra samples generated for masses exactly at the kinematic threshold $m_{\mathrm{h}} = \sqrt{s} - m_{\mathrm{Z}}$.

| $\sqrt{s}$ | ZZ | $\mathrm{q\bar{q}}$ | WW | hZ SIGNAL (7.5k) |
|---|---|---|---|---|
| 204 | 50k | 100k | 250k | $m_{\mathrm{h}}$ =75-120 GeV/$c^2$ |
| 205 | 50k | - | 250k | - |
| 206 | 50k | 100k | 250k 250k | $m_{\mathrm{h}}$ =75-120 GeV/$c^2$ |
| 206.7 | 250k | 2M $\mathrm{b\bar{b}}$ , 500k $\mathrm{c\bar{c}}$ | - | $m_{\mathrm{h}}$ =80-130 GeV/$c^2$ |
| 207 | 50k | 400k | 250k, 1M, 250k | - |
| 208 | 50k | 100k | 250k | $m_{\mathrm{h}}$ =75-120 GeV/$c^2$ |
| 209 | 50k | - | 250k | - |
| 210 | 50k | 100k | 250k | $m_{\mathrm{h}}$ =75-120 GeV/$c^2$ |

**Table 3.3:** The MC simulation samples generated for use within the 2000 analysis.

## 3.6 b-quark jet tagging

The high branching ratio of the Higgs boson to b quarks (≈75% for a Higgs boson with $m_{\mathrm{h}} = 114$ GeV/$c^2$) means that any search for the Higgs boson requires a method of identifying ("tagging") b-quark jets. To reduce the impact of possible backgrounds any algorithm which is developed to tag these b-quarks must also have a small "mis-tag" rate, i.e. the rate to mis-identify a light quark (u,d,s,c) as a b-quark.

Within the following analyses an artificial Neural Network (NN) [45] is used to identify b-quark jets [46]. The inputs to the NN are chosen to reflect the nature of b-quark jets which arises through the high mass and long lifetime of the produced b hadrons. The first two inputs are lifetime based, the third is based on the transverse momentum of identified leptons and the last three are based on the jet shape properties.

1. $\mathcal{P}_{jet}$ (QIPBTAG): probability of the jet being a light quark (uds) jet based upon the impact parameters of the tracks in the jet [47].

2. $\Delta\chi^2$ (QVSRCH): the $\chi^2$ difference between fitting tracks in the jet assuming the existence of a secondary vertex and assuming only the primary vertex is present [48].

3. $p_t$: the transverse momentum of identified leptons with respect to the jet axis.

4. Boosted Sphericity: the sphericity of energy flow objects within the rest frame of the jet.

5. $\sum p_t^2$: the sum of the transverse momenta squared of each of the energy flow objects with respect to the jet axis.

6. Jet multiplicity / $\ln E_{jet}$: the energy flow multiplicity of the jet divided by the logarithm of the jet energy. Normalising by $\ln E_{jet}$ removes the expected energy dependence of the multiplicity.

The NN is a fully-connected multi-layer feed-forward network with four layers and is based upon the JETNET 3.4 [49]. The NN is trained on a sample of radiative return $Z \rightarrow q\overline{q}(\gamma)$ events which provide a large sample of events independent to the training of the event selection analyses and thus independent to the selected data set. The output of the NN for jets in Z peak data recorded compared to MC prediction is shown in figure 3.21 while the efficiency for tagging b jets vs the rejection factor for light quark jets is shown in figure 3.22.

The training and performance of the NN is discussed in more detail in [50] [51].

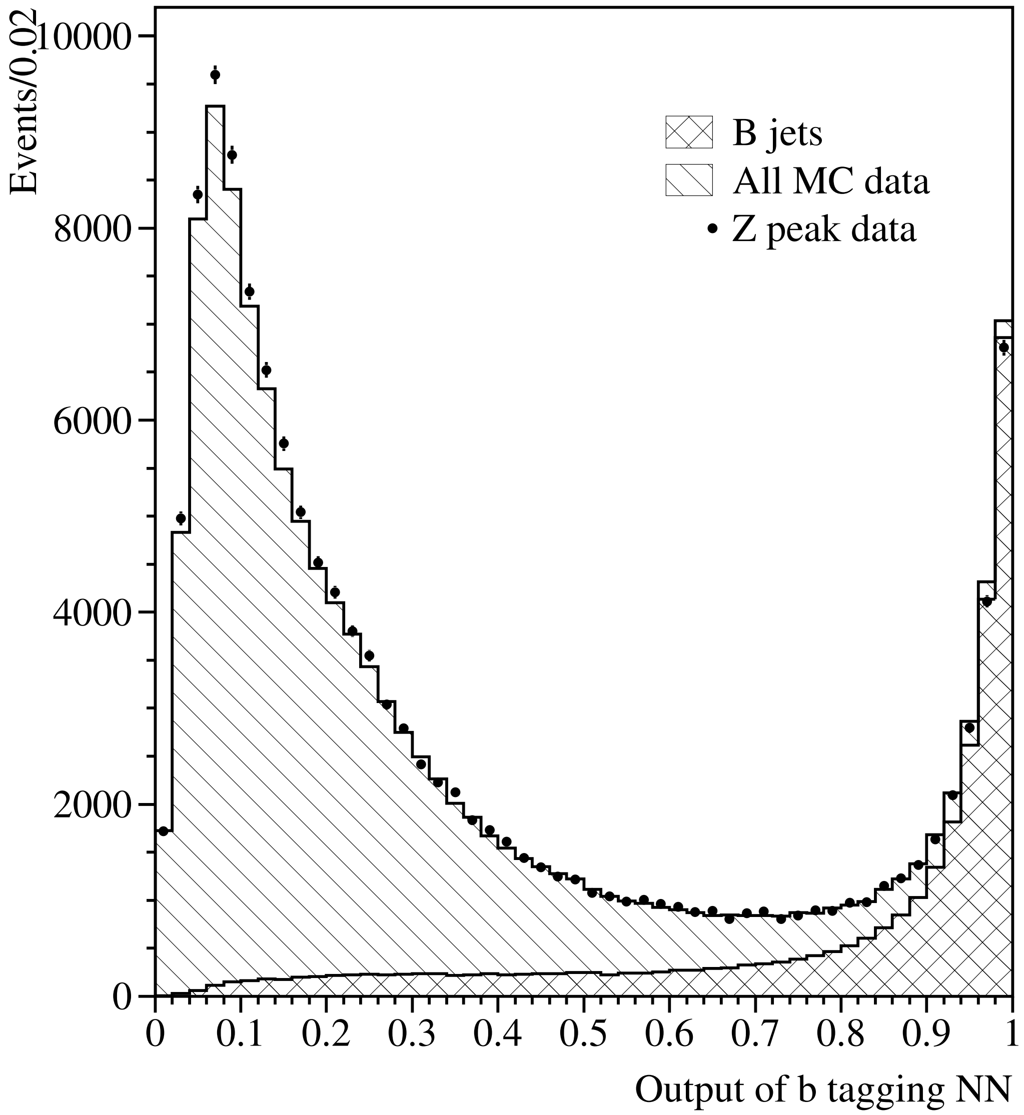


**Figure 3.21:** The b-tagging NN output for jets from MC simulation and data at the Z peak. The b quark content is clearly shown to peak towards 1.

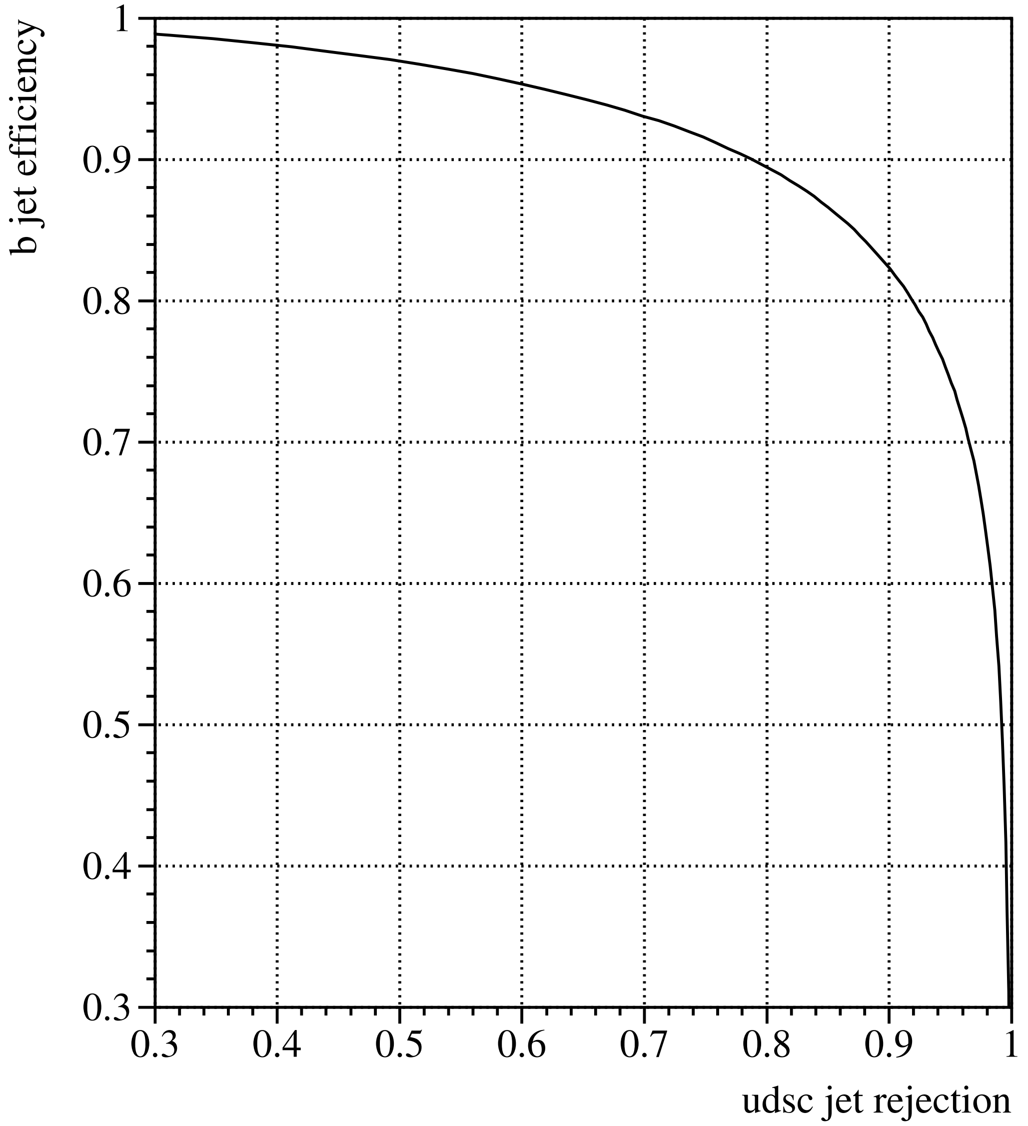


**Figure 3.22:** The b-tagging efficiency vs light quark rejection rate for jets from MC simulation at the Z peak.

# Chapter 4

# The hZ four jets event selection

## 4.1 Introduction

The aim of the hZ four jets analysis is to maximise the selection efficiency for signal events within this channel whilst minimising the selection of background processes. The analysis applies a set of optimised cuts and focuses on the rejection of the three main background processes to the fully hadronic hZ decay. These backgrounds are:

- $e^+e^- \rightarrow ZZ \rightarrow b\bar{b}b\bar{b}, b\bar{b}q\bar{q}, q\bar{q}q\bar{q}$
- $e^+e^- \rightarrow W^+W^- \rightarrow q\bar{q}'q\bar{q}'$
- $e^+e^- \rightarrow q\bar{q}g(g), q\bar{q}q\bar{q}$

The hadronic decay mode of the ZZ background process has been shown in its three possible components to highlight its overlap with a possible signal. Indeed the final states of both the hZ and ZZ hadronic decay modes may be identical. This fact leads the ZZ background to be the dominant background within this search channel and its possible identical nature with the hZ signal lead to it being referred to as an irreducible background. The event selection proceeds in two stages, see figure 4.1. Firstly, through a loose pre-selection which is designed to find four jet multi-hadronic events and secondly through a set of optimised cuts which form the main event selection. For each selected event the value of the reconstructed Higgs mass is then calculated. The reconstructed Higgs mass, denoted as $m_{\text{Reco}}$, is used as the single discriminating variable for the four jets cuts-based analysis. The value of the discriminating variable is used to provide a measure of the signal likeness, in effect an event weight, when interpreting the experimental data in the form of confidence levels (see appendix B).

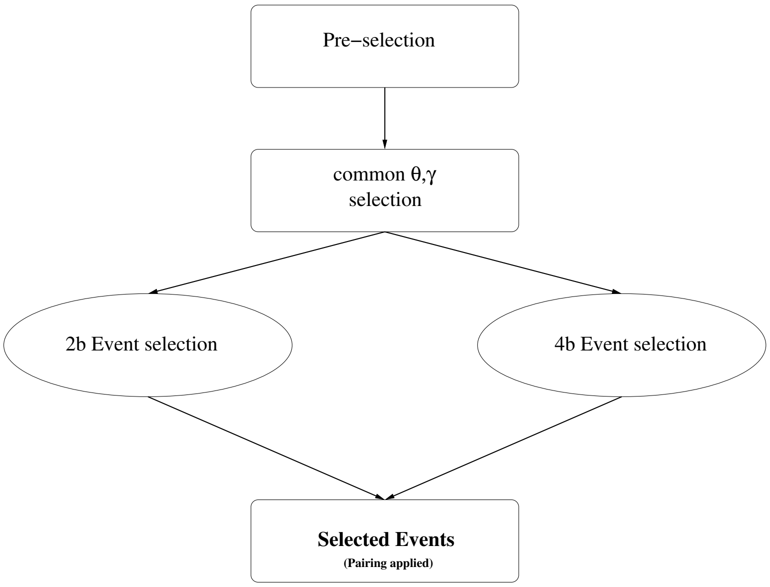


**Figure 4.1:** Shematic depiction of the four jets event selection

## 4.2 Pre-selection

The main purpose of the event pre-selection is to select hadronic events. The pre-selection is designed to search for criteria which are common to good multi-hadronic final states at LEP 2. Initially the events are selected by placing a requirement on the number of good charged tracks and the amount of charged energy within the event. A good charged track is defined as one which meets the following criteria: the track is required to originate from the interaction point (IP). This is ensured by requiring that the radial and longitudinal displacement from the IP lie within the region $|d_0| \leq 2cm$, $|z_0| \leq 10cm$. Furthermore the track is required to be reconstructed from at least 4 TPC coordinates and must fulfil the angular requirement of $|cos\theta| < 0.95$.

Events which contain eight or more such good tracks with the sum of their energy equal to or greater than 0.1 $\sqrt{s}$ are selected. This hadronic requirement is designed to cut the number of two photon and leptonic four fermion processes from the event stream.

Selected events are then clustered into four jets using a jet clustering algorithm. Jet clustering algorithms cluster the charged and neutral final state objects in an event together in an attempt to form the individual underlying jets. These clustered jets are then taken to approximate the direction and energy of the initial partons of the final state.

The jet clustering algorithm used within the four jets analysis is a variant of a standard jet clustering algorithms known as the DURHAM E scheme [52]. The DURHAM scheme has several advantages, one of which is a reduced sensitivity to soft gluons.

Initially all energy flow objects within the event are considered and each possible pairing of these objects is found. For each possible pairing the quantity $y_{ij}$ is calculated where, in the DURHAM scheme,

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

with $i$ and $j$ referring to the two energy flow objects and $E_{vis}$ refers to the total visible energy within the event.

The two energy flow objects which give the lowest value of $y_{ij}$ are then combined and form a new “pseudo-particle” which replaces the two original objects. In the DURHAM E scheme the two energy flow objects are combined by simply summing the four momenta of the objects, ie, $E_{pseudo} = E_i + E_j$ and $\vec{P}_{pseudo} = \vec{P}_i + \vec{P}_j$ where E and $\vec{P}$ refer to the energy and 3-momentum respectively.

This process is continued iteratively using the energy flow objects and pseudo-particles until no more pseudo-particles with a value of $y_{ij}$ below a given cut can be found. Alternatively the iterative clustering process is continued until the desired number of pseudo-particles is left. The pseudo-particles are then referred to as jets. At this point the associated value of the $y_{ij}$, in this case $y_{34}$, may be taken as a measure of how well the event fits the the four jet topology. Events in which there are two or three true jets, e.g. those originating from $q\overline{q}(\gamma)$ and $q\overline{q}(g)$, may be clustered to form four jets when using the above method. These events however will have small $y_{34}$ values in contrast to those from true four jet events which are readily clustered to form four jets. Following this all events with a $y_{34}$ value which is less than 0.004 are rejected. That is,

$$y_{34} > 0.004 \tag{4.2}$$

Figure 4.2 shows the distribution of $y_{34}$ in simulated signal and real data. Further to the above described cuts each jet is required to contain at least one good charged track.

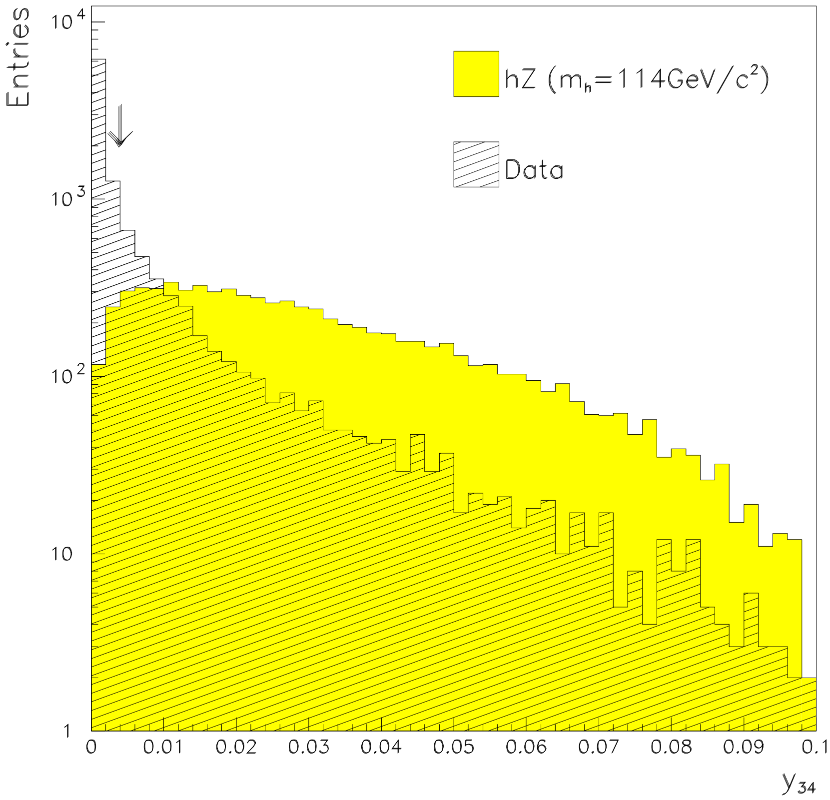


**Figure 4.2:** The distribution of the $y_{34}$ variable in data and hZ signal with $m_{\mathrm{h}} = 114\,\mathrm{GeV}/c^2$. The signal distribution is shown (with arbitrary normalisation) for comparison purposes. The applied cut is indicated by the vertical arrow.

Although events which pass the multi-hadronic criteria are consistent with true four jet events there still exists a sizeable quantity of radiative return events. Radiative return events refer to events following the reaction $e^+e^- \rightarrow Z^0(\gamma) \rightarrow q\overline{q}(\gamma)$. In this process one of the initial $e^+e^-$ pair emits an initial state radiation (ISR) photon, typically with an energy of approximately 80 GeV for $\sqrt{s} \approx$ 200 GeV. The emission of the ISR photon leaves the initial $e^+e^-$ system with a centre of mass energy $\approx m_Z$. The annihilation of this $e^+e^-$ state to a $Z^0$ and the subsequent decay of this to $q\bar{q}$ provides a final state which is hadronic in nature. The emitted ISR photon may be observed within the detector or it may escape undetected down the beam-pipe. In either case the detection or lack of detection of this ISR photon provides an event characteristic which may be used to distinguish such ISR events from true four jet events. Radiative return events in which the ISR photon escapes undetected down the beam-pipe have a large amount of missing momentum along the direction of the beam, $P_z$, and a visible mass, $m_{vis}$, near to that of the Z boson, $m_Z$. The visible mass, denoted $m_{vis}$, is defined as

$$m_{vis} = \sqrt{{E_{vis}}^2 - |\overrightarrow{P_{vis}}|^2} \tag{4.3}$$

where $E_{vis}$ and $\overrightarrow{P_{vis}}$ and the visible energy and momentum within the event. To remove events in which an ISR photon is emitted down the path of the beam pipe a cut is placed to ensure that $m_{vis}$ is larger than 90 GeV and also require that $P_z$ is small. This requirement is

$$|P_z| < 1.5(m_{vis} - 90) \tag{4.4}$$

Figure 4.3 shows the $|Pz|$ versus $m_{vis}$ variables which form the anti radiative return cut. The figure compares recorded data and a hypothetical Higgs boson signature with $m_h = 114$ GeV/$c^2$. The cut is indicated by the dividing line with the selected and rejected regions marked. The impact of the applied anti radiative return cut may be seen when considering the distributions of the visible mass both before and after the application of the cut. Figures 4.4 show exactly this for (a) the data and (b) the hypothetical Higgs boson signal. The cut has a visible effect on the data sample by removing the radiative return events which have $m_{vis} \approx 90$ GeV/$c^2$ whilst there is little effect on the signal case.

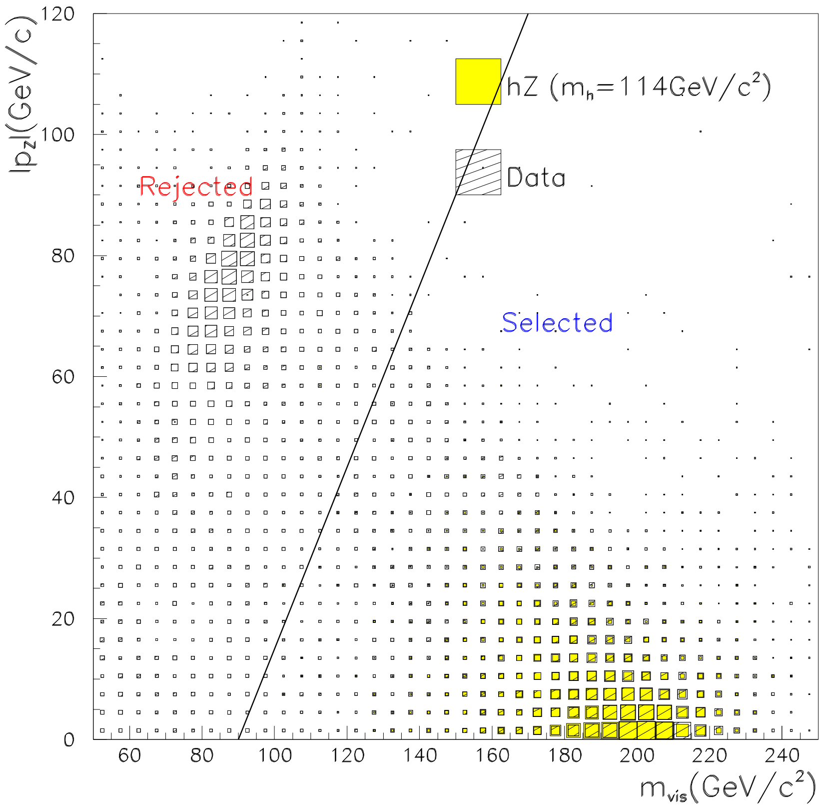


**Figure 4.3:** The distribution of the $|Pz|$ versus $m_{vis}$ variables which form the anti radiative return cut. Data are compared to a hypothetical Higgs signal with the cut highlighted by the dividing line. The distributions are arbitrarily normalised for comparison purposes.

Events in which an ISR photon is observed within the detector are rejected when the fraction of electromagnetic energy in a $1^\circ$ cone around any one track in an event is found to be too high. The ISR photon may convert to an $e^+e^-$ pair via pair production in matter. To allow for this possibility a search is undertaken for all electromagnetic energy-flow objects within the jet. The energy flow objects are flagged as electromagnetic if they are identified as electrons, photons or photon conversions.

Further to this, energy deposits in ECAL below the TPC acceptance (i.e. $|cos\theta|$ $> 0.95$) and LCAL, SiCal or HCAL deposits behind ECAL module cracks are also flagged as electromagnetic. The largest fraction of electromagnetic energy within a $1^\circ$ cone around an electromagnetic energy-flow object within any jet $(X_\gamma)_{max}$

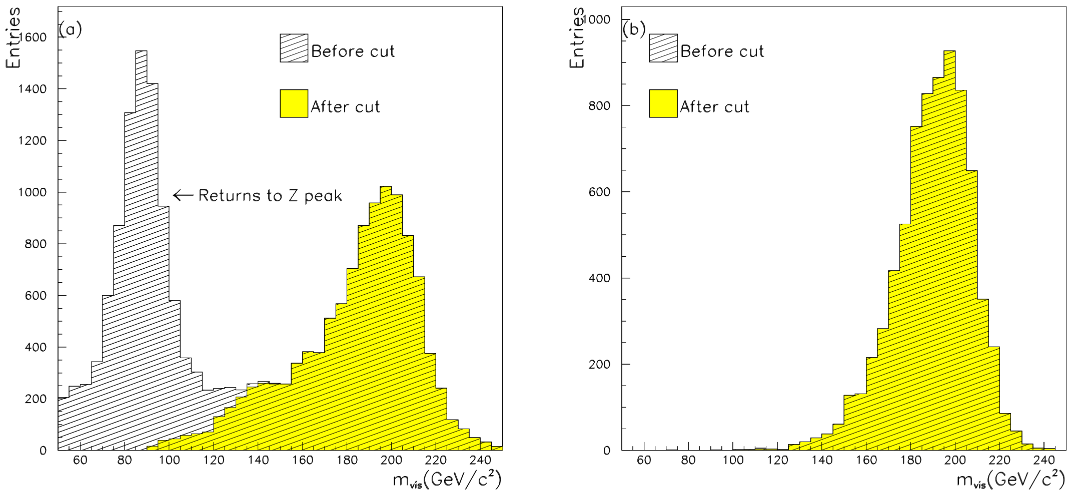


**Figure 4.4:** The $m_{vis}$ distributions for (a) data and (b) hypothesised Higgs boson signal with $m_h = 114$ GeV$/c^2$ both before and after the application of the radiative return cut, equation 4.4.

must be below 0.80. The $X_\gamma$ distribution is shown in figure 4.5.

A four-constraint (4C) fit [53] is applied to the events to ensure the conservation of energy and momentum. The 4C fit is applied in place of the previous four constraint energy rescaling technique [54] [55] used up to 1999. The 4C fit allows the jet angles to be fitted as well as the jet energies, thus improving the di-jet mass resolution, particularly in the region of the kinematic threshold. The kinematic fit was adopted for the final analysis of the 1999 data set [56] with no observable biases or degradation of the analysis performance.

## 4.3 Final selection

The composition of events passing to the final selection is predominantly true four jet events with a small fraction of $q\overline{q}(g)$ events.

The following cuts selection [50] is designed to provide the greatest possible rejection of the $q\overline{q}(g)$ and four fermion (four jet) backgrounds.

The $q\overline{q}(g)$ final state in which a gluon is radiated from one of the final state quarks, typically has a kinematic topology of one hard jet recoiling against three softer jets. A sample event which depicts this distinct topology is shown in figure

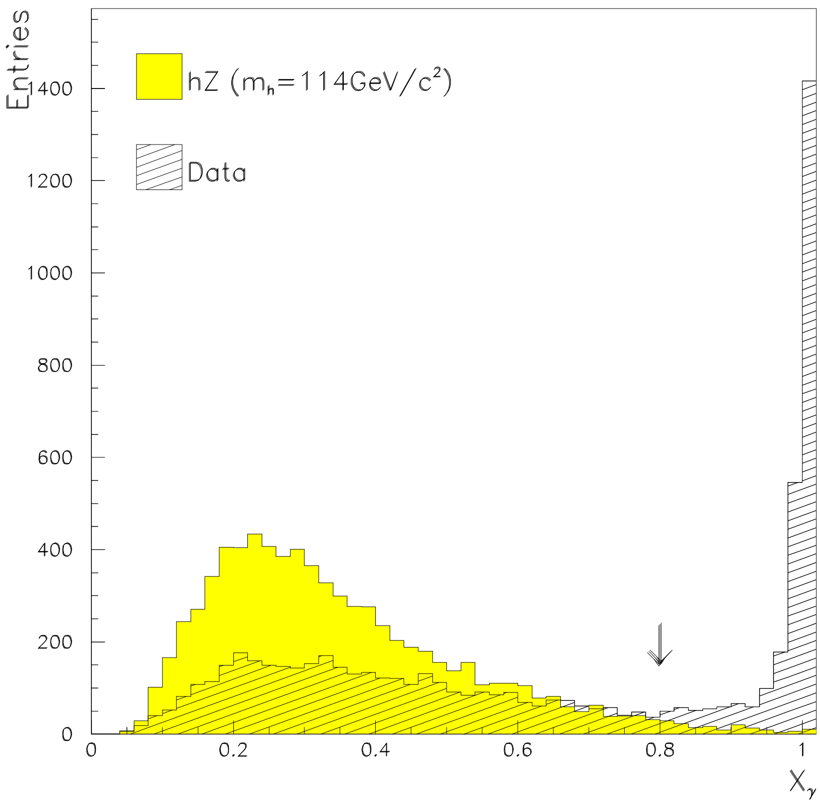


**Figure 4.5:** The distribution of the $X_\gamma$ variable in data and hZ signal with $m_{\rm h} = 114\,\text{GeV}/c^2$. The signal distribution is shown (with arbitrary normalisation) for comparison purposes.

4.6. To utilise this topology in the rejection of $q\overline{q}(g)$ events the quantity $\Theta$ is defined. Here $\Theta$ refers to the sum of the four smallest inter-jet angles within an event. The $\Theta$ distribution for background and signal MC as well as a comparison to data is given in figure 4.7. Events which pass the selection are required to have a $\Theta$ value of $350^\circ$ or greater.

The production of a Higgs boson close to the kinematic threshold ($\sqrt{s} - m_Z$) is of particular importance within the search as it defines the ultimate limit of sensitivity. The nature of the Higgs boson searches at LEP 2, with increasing centre of mass energies each year, means that a Higgs boson below the kinematic threshold, by some 10 GeV or more, would have been observed in previous years. The Higgs boson production cross section with respect to a fixed centre of mass

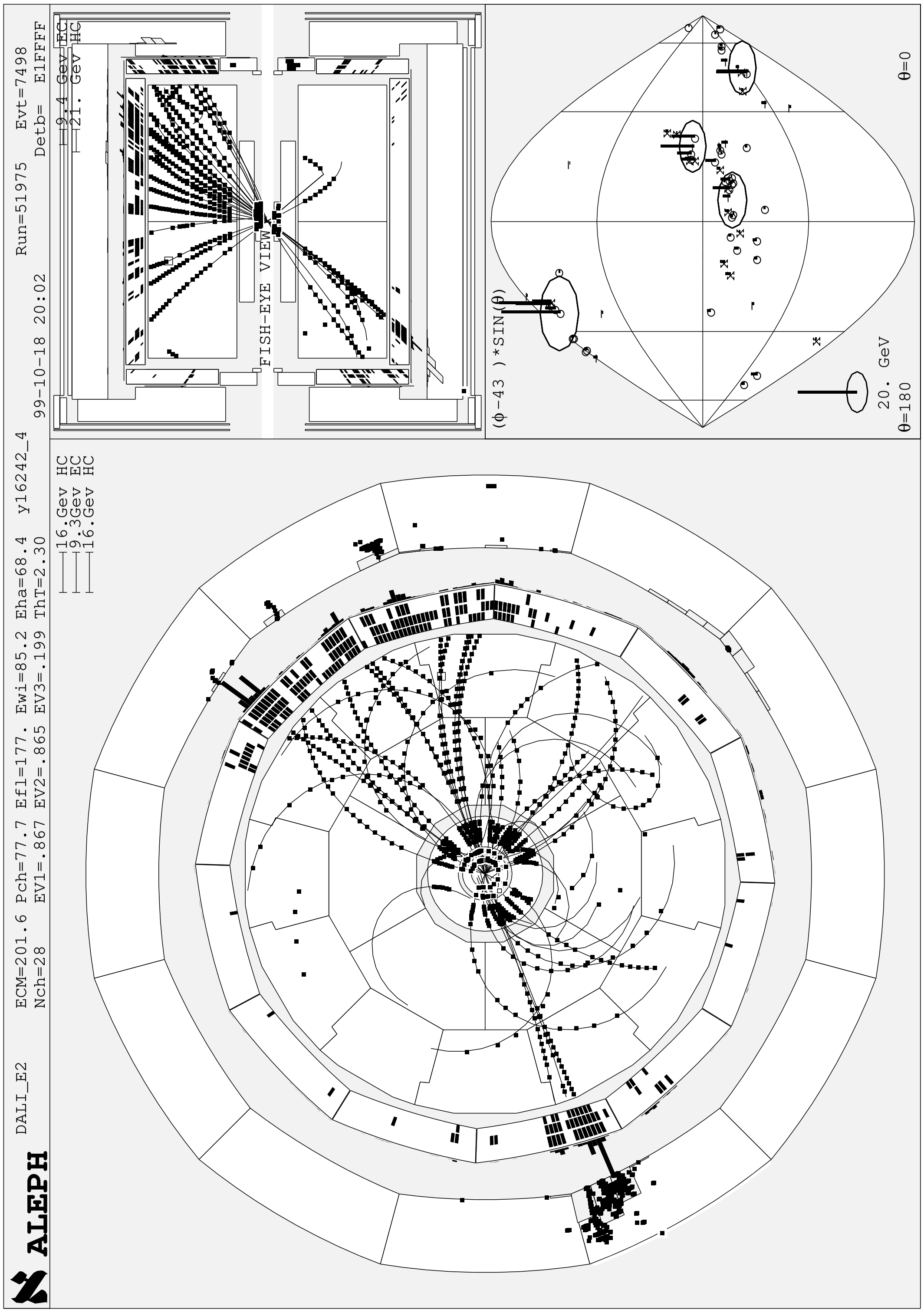


**Figure 4.6:** An example of a $q\overline{q}(g)$ event which typifies the $\Theta$ variable characteristic used to reject such events [50] [57].

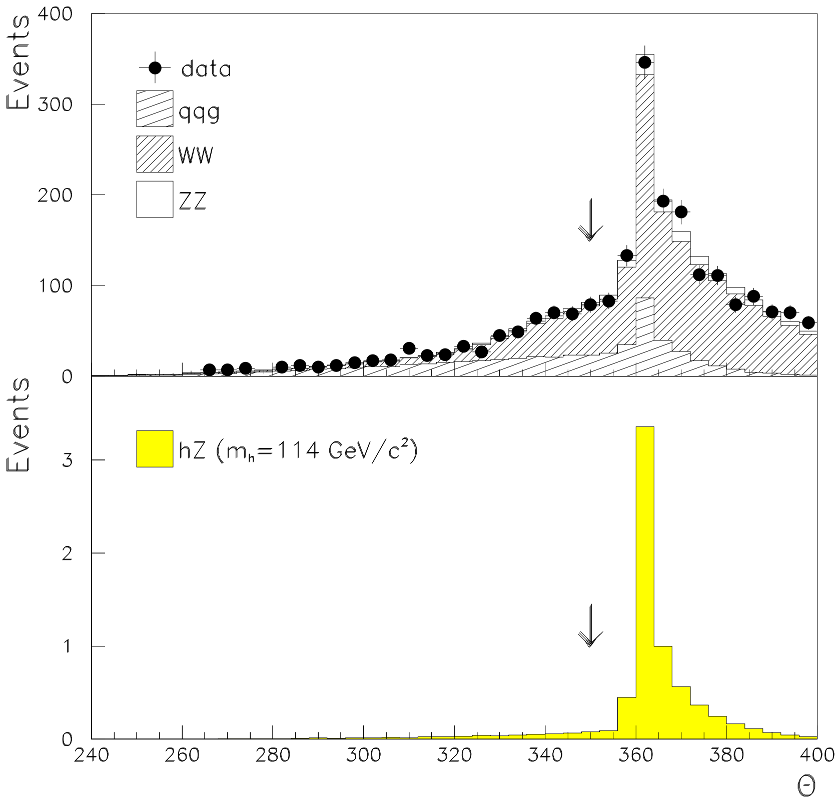


**Figure 4.7:** The distribution of the Θ variable for Monte Carlo background, data and a hypothetical signal with $m_{\mathrm{h}} = 114\,\mathrm{GeV}/c^2$. The applied cut is indicated by the vertical arrow.

energy rises considerably for Higgs bosons with a mass several GeV below the kinematic threshold. The production of a Higgs boson close to the kinematic threshold yields the characteristic topology of two pairs of back to back jets. A variable which is very descriptive of this topology is the sum of the cosines of the di-jet angles. In an event close to threshold the quantity

$$\gamma = \min(cos\theta_{ij} + cos\theta_{kl}), \tag{4.5}$$

where all possible permutations of $i, j, k, l$ are considered, tends towards -2. Figure 4.8 shows the distribution of $\gamma$ for data in comparison to background and signal MC. An example of a four jet event which typifies this back to back topology is shown in figure 4.9. The correlation between $\gamma$ and the reconstructed Higgs boson mass means that a tight cut on the value of $\gamma$ would also impose a tight cut on the value of the reconstructed mass associated with the event. Such a tight cut on the reconstructed mass is undesirable as this would hinder searches in the alternate scenarios in which lower Higgs boson masses have a greater importance. Following this consideration it is required that

$$\gamma < -1.3 \tag{4.6}$$

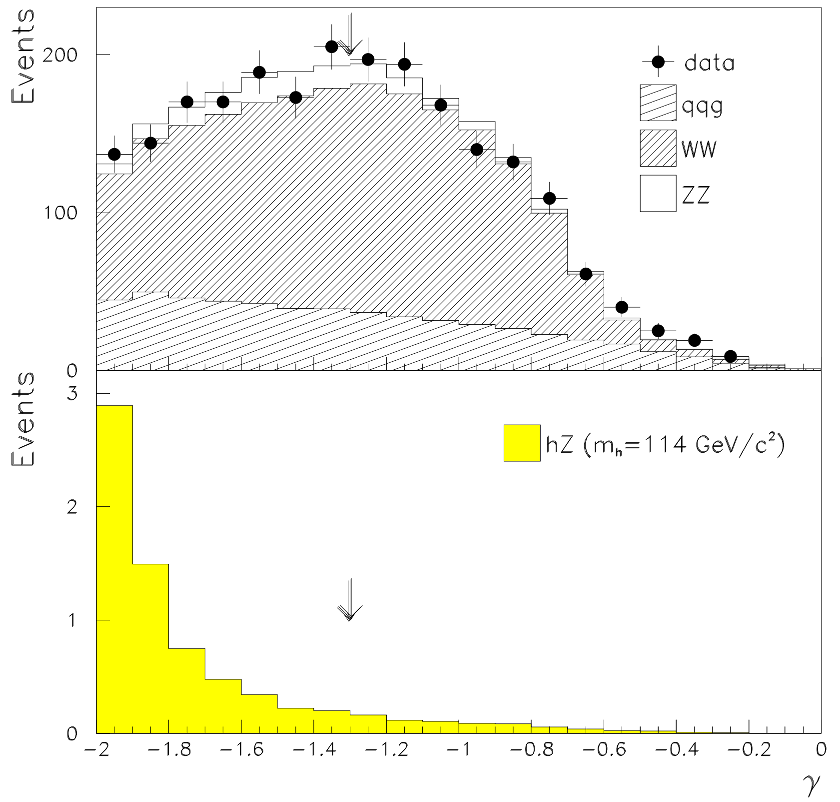


**Figure 4.8:** The distribution of the $\gamma$ variable for Monte Carlo background, data and a hypothetical signal with $m_{\mathrm{h}} = 114\,\mathrm{GeV}/c^2$. The applied cut is indicated by the vertical arrow.

### 4.3.1 The 2b and 4b event selections

From this point onwards the hZ analysis is split into two parallel streams of cuts. The two streams, denoted as the "2b" and "4b" branches, are optimised to search

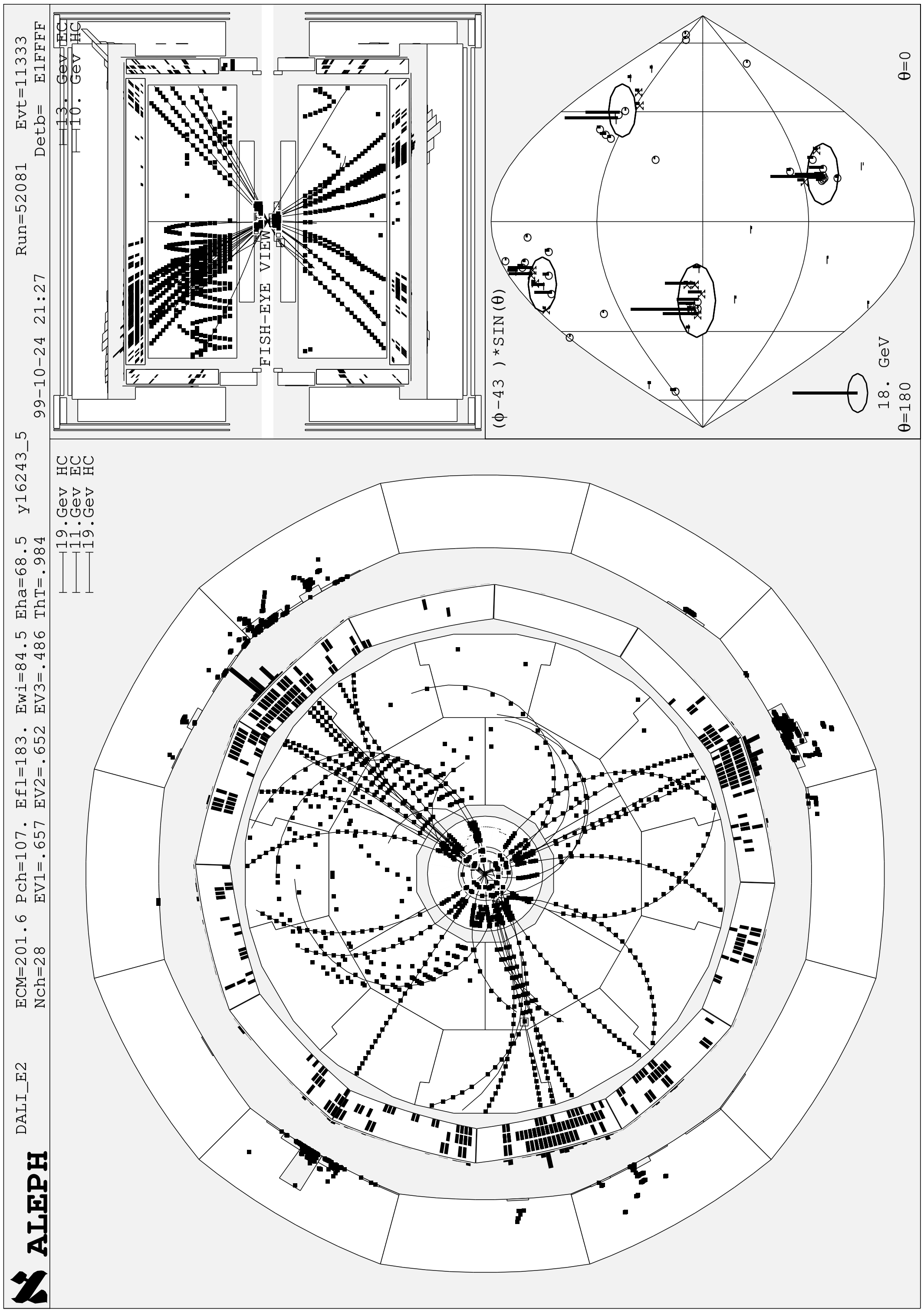


**Figure 4.9:** An example of a four jet event which displays the back to back topology which the $\gamma$ variable is designed to select [50].

for hZ $\rightarrow$ $b\bar{b}q\bar{q}$ and hZ $\rightarrow$ $b\bar{b}b\bar{b}$ respectively. All events which pass one or both of the branches are selected by the analysis as hZ$\rightarrow$four jets candidates. An optimisation procedure is used to determine the value of the cuts used in the two branches is discussed in section 4.5.

### 4.3.2 The 2b event selection

The 2b event selection is formed from five cuts which place constraints on the event topology, di-jet (boson candidate) masses and the b-tagging of the jets. Here b-tagging refers to a measure of how "b-like" a candidate jet appears. The first constraint to be applied is the requirement of a greater cut on the $y_{34}$ value than was originally chosen for the pre-selection purposes. The tightened requirement is

$$y_{34} > 0.008 \tag{4.7}$$

This increased cut on $y_{34}$ ensures that the four jets within the event are well isolated and increases the probability that the selected event is indeed of a true four jet nature. Figure 4.10 shows the distribution of the $\log(y_{34})$ variable for the three main backgrounds versus the observed data and a hypothetical Higgs boson signal with $m_{\mathrm{h}} = 114$ GeV$/c^2$.

The remaining cuts which are applied within this branch are dependent on the di-jet pairing within the event. This allows cuts to be applied to the individual h and Z boson candidates, based on their specific expected characteristics (e.g. $m_Z$), and not simply on the overall event characteristics (e.g. $y_{34}$). Within each four jet event the actual four jets themselves must be paired into di-jet (boson) candidates. Once two of the jets are paired to form the h or Z boson candidate the other two jets are automatically paired to form the opposing boson candidate. Each four jet event has six such possible pairing configurations, also called combinations. This means that each candidate has six possible interpretations each of which must be considered when the following cuts are applied. Indeed each of the following pairing dependent variables may take a different value for each of the six different interpretations of the event. In the following the convention is to label the pair

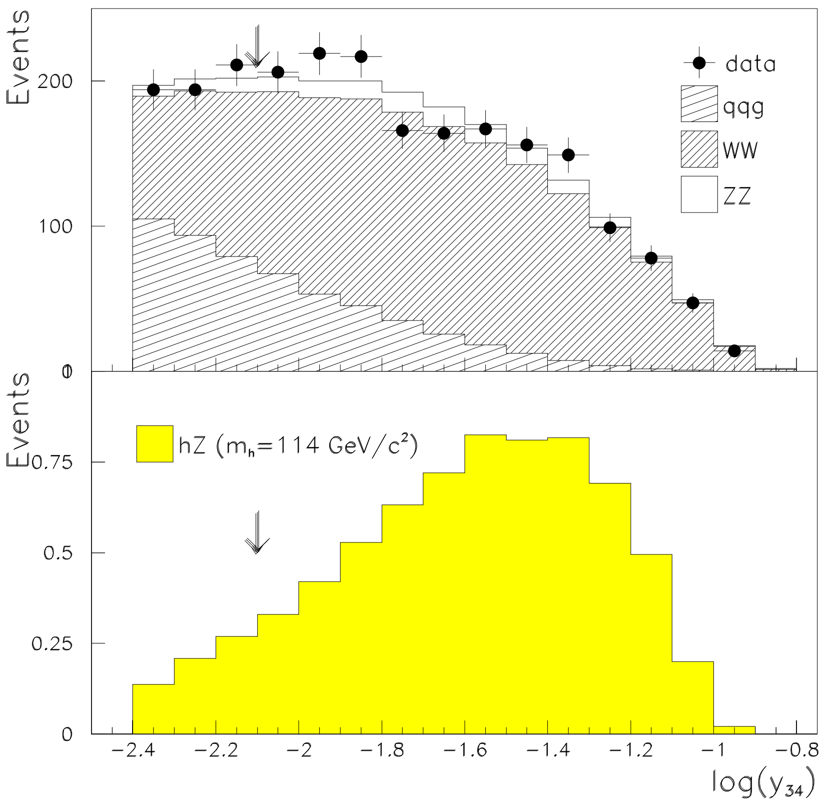


**Figure 4.10:** The distribution of the $\log(y_{34})$ variable for simulated background, data and a hypothetical signal with $m_h = 114\ \mathrm{GeV}/c^2$. The applied cut is indicated by the vertical arrow.

of jets associated with the Z boson as jets 1 and 2 and those associated with the Higgs boson as jets 3 and 4.

Firstly cuts are applied to the di-jet masses of the h and Z boson candidates. The boson mass is defined as the invariant mass of the di-jet system which is associated with it. The mass cuts are applied to all six possible interpretations and only those configurations in which both cuts are satisfied are selected.

The mass cuts

$$m_{12} > 77\ \mathrm{GeV}/c^2 \tag{4.8}$$

$$m_{34} > 55\ \mathrm{GeV}/c^2 \tag{4.9}$$

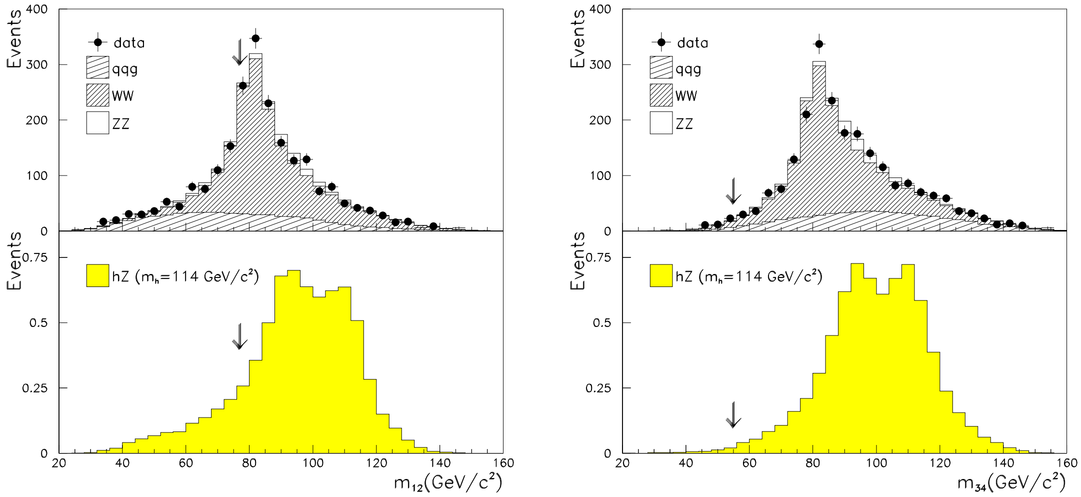


**Figure 4.11:** The distributions of the $m_{12}$ and $m_{34}$ variables for Monte Carlo background data and a hypothetical signal with $m_\mathrm{h} = 114\,\mathrm{GeV}/c^2$. The applied cut are indicated by the vertical arrows.

are imposed upon each possible interpretation of the hZ system. These mass cuts require the Z boson candidate to be in a mass range which is compatible with the measured value of $m_\mathrm{Z}$ ($\approx$90 GeV/$c^2$) whilst allowing more freedom in the case of the h boson candidate mass so that the sensitivity to lower h boson masses is not reduced. Figure 4.11 shows the distributions of $m_{12}$ and $m_{34}$.

The final two cuts which are applied in the 2b branch are placed on the b-tagging content of the two jets which form the Higgs boson candidate (i.e. jets 3 and 4). For each jet a measure of the b-likeness, b-content, of the jet is calculated using a six variable artificial neural network (6VNN), see section 3.6. The six variables which are used to form the 6VNN are summarised in table 4.1.

The Neural Network b-tag outputs for the two jets associated with the h boson candidate are denoted $\eta_3$ and $\eta_4$ corresponding to jets 3 and 4 respectively. Two cuts are placed on the b-tagging content of the Higgs boson candidate jets. Firstly a cut is placed on the minimum b-tag of either of the two jets. This ensures that both jets are at least modestly b-tagged. The second, and more stringent, cut ensures that the combined di-jet system is itself well b-tagged. These two cuts are

$$\min(\eta_3, \eta_4) > 0.35 \tag{4.10}$$

| | Input Variable |
|---|---|
| 1 | $\mathcal{P}_{jet}$ (QIPBTAG)<br>Generic ALEPH b-tagging variable using impact parameters. |
| 2 | $\Delta\chi^2$ (QVSRCH)<br>Measure of secondary vertex quality. |
| 3 | Largest $p_{\mathrm{T}}$ of identified leptons |
| 4 | Boosted sphericity of jet |
| 5 | Sum of $p_{\mathrm{T}}^2$ of all particles in the jet |
| 6 | Jet multiplicity / ln $E_{jet}$ |

**Table 4.1:** The six variables used within the construction of the 6V b-tagging Neural Network.

$$(1-\eta_3)(1-\eta_4) < 4.8\text{x}10^{-3} \tag{4.11}$$

where $\eta_i$ is the NN output for the $i^{th}$ jet.

The distributions of these two variables are seen in figure 4.12. From these two figures the discriminating power available in the b-tag information is clearly visible.

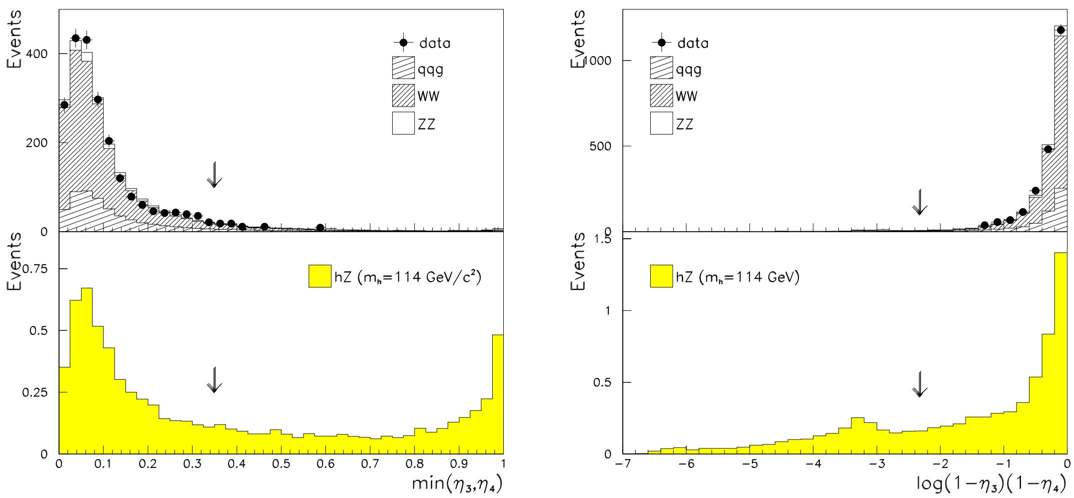


**Figure 4.12:** The distributions of the two b-tagging variables for Monte Carlo background, data and a hypothetical signal with $m_{\mathrm{h}} = 114\ \mathrm{GeV}/c^2$. The applied cuts are indicated by the vertical arrows.

### 4.3.3 The 4b event selection

The 4b final state, arising from hZ $\rightarrow b\bar{b}b\bar{b}$, accounts for $\approx$ 20% of the fully hadronic decay of the hZ system. Although it accounts for a smaller fraction of the decay than the 2b state the presence of four b quarks means that it possesses much less background. The 4b final state of the hZ decay can be selected by requiring the presence of four well isolated jets with a high b content. To this end a linear discriminant [58] is formed using the sum of the four neural network b-tags and the value of $y_{34}$. The cut placed on the linear discriminant is

$$9.5 \cdot y_{34} + \sum_{i=1}^{4} \eta_i > 3.2 \tag{4.12}$$

This optimised linear discriminant ensures that an improved rejection of the background is obtained compared to cutting on the values of the two variables independently. This cut forms the only cut within the 4b event selection. The distribution of the linear discriminant variable can be seen in figure 4.13.

Events which are selected by either branch pass the selection and are treated as possible hZ candidates. Events which pass both 2b and 4b branches are automatically interpreted as 4b candidates due to the high amount of b-tagging within the event and thus the ambiguity in the selection (di-jet) pairing. Table 4.2 shows the fractional distribution of candidate events between the 2b and 4b selections.

| | 2b branch (%) | 4b branch (%) | 2b + 4b branch (%) |
|---|---|---|---|
| hZ | 67.53 | 14.83 | 17.63 |
| WW | 91.77 | 7.45 | 0.77 |
| ZZ | 78.48 | 10.47 | 11.03 |
| $q\bar{q}$ | 89.41 | 6.56 | 4.01 |
| Tot bgd | 83.58 | 8.86 | 7.51 |
| Data | 90.90(30) | 3.03(1) | 6.06(2) |

**Table 4.2:** The fraction of selected events, in %, entering each branch of the analysis. The number of recorded data events is shown in brackets.

### 4.3.4 Jet pairing

In each event that passes the selection there may be more than one possible jet pairing combination which satisfies the above cuts. The combination indepen-

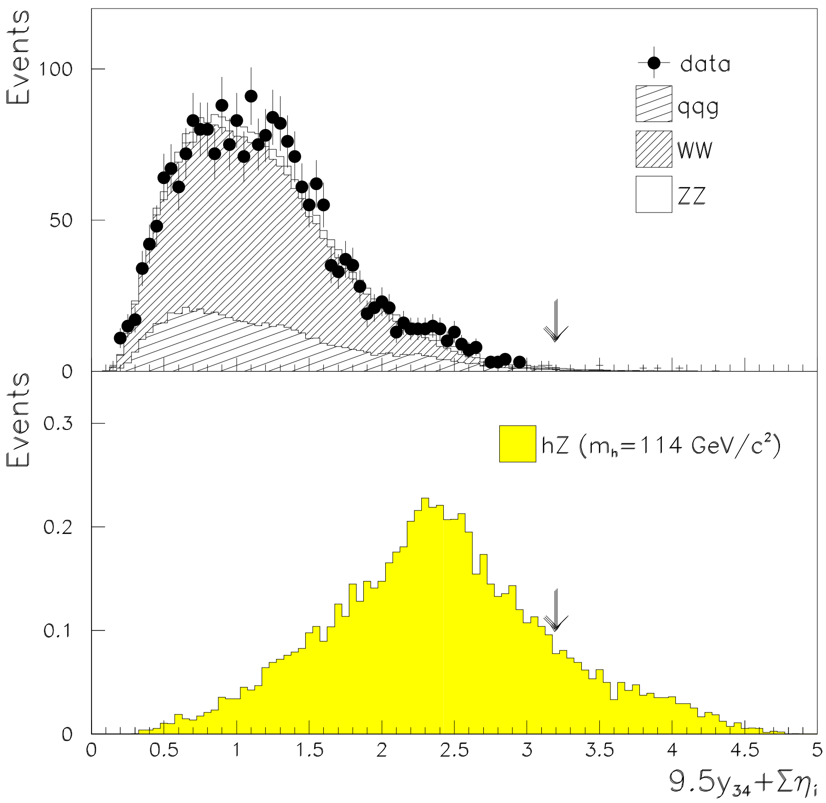


**Figure 4.13:** The distribution of the linear discriminant variable for Monte Carlo background, data and a hypothetical signal with $m_{\mathrm{h}} = 114\,\mathrm{GeV}/c^2$. The applied cut is indicated by the vertical arrows

dent cuts in the four b branch ensure that all 6 possible combination are selected as possible interpretations. The combination dependent cuts of the 2b branch, on the other hand, mean it is possible for less than six combinations to be selected by this branch. Table 4.3 shows the fraction of events passing through the analysis which need pairing while table 4.4 shows the fraction of events passing through the 2b branch which have a given number of selected combinations.

For each selected event the value of the discriminating variable, in this case the reconstructed Higgs mass, is calculated. The reconstructed Higgs mass, defined as

$$m_{\mathrm{Reco}} = \mathrm{m}_{12} + \mathrm{m}_{34} - 91.2\ \mathrm{GeV}/c^2 \tag{4.13}$$

is clearly a combination dependent variable. Following from this it is required that only one of the selected pairings be finally chosen[1]. To determine which combination is chosen the decay angles for the h and Z boson candidates are calculated for each of the selected pairings. The decay angle, $\theta_d$, is the angle between the di-jet axis from the h or Z boson and the direction of motion of the boson in the lab frame, as measured in the rest frame of the boson. The decay angle is depicted diagrammatically in figure 4.14.

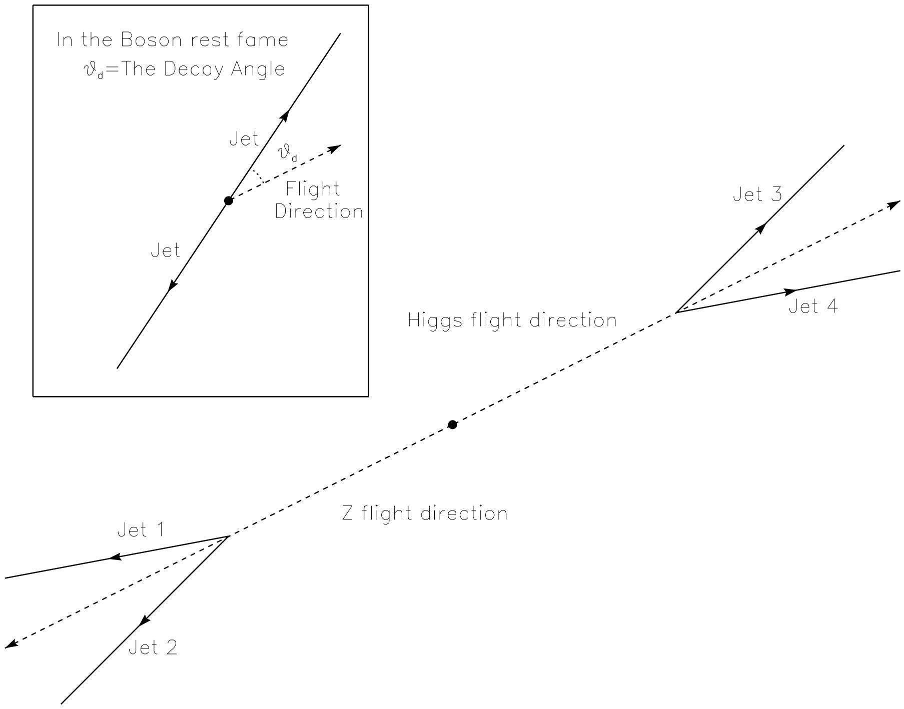


**Figure 4.14:** A pictorial depiction of the decay angles variable. The dashed lines represent the flight direction of the bosons whilst the solid lines represent the flight direction of the jets produced in the boson decay. The diagram is set in the rest frame of the $e^+e^-$ interaction while the inset is set in the rest frame of the decaying boson.

The decay angle($\theta_d$) from MC simulations was used to produce two likelihoods, one each for the h and Z boson candidate within the event. The likelihoods were formed using the following equation.

[1]A summary of jet pairing studies is provided in appendixA

$$\mathcal{L}_{h/Z} = \frac{pdf_{(correct,signal)}}{pdf_{(wrong,signal)} + pdf_{(allbackgrounds)}} \tag{4.14}$$

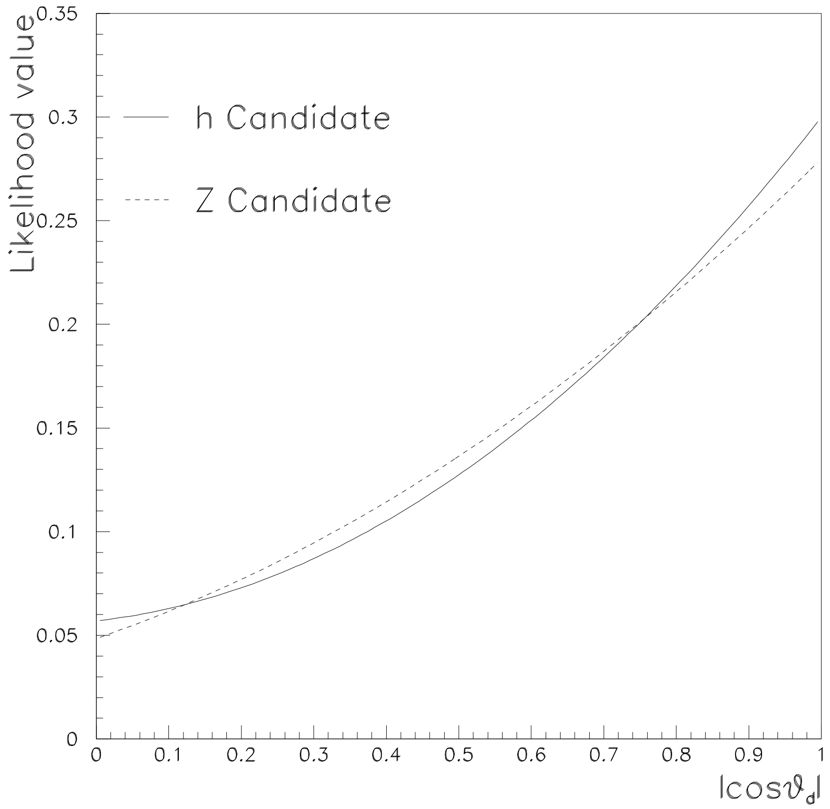


**Figure 4.15:** The distribution of the decay angles likelihoods for the h and Z boson candidates.

where the pdf for the wrong signal pairings takes into account all the possible incorrect signal pairings and the pdf for the backgrounds takes into account all possible pairings for all possible backgrounds. The form of the two likelihoods is displayed in figure 4.15 and the actual decay angles are shown in figure 4.16 .For all selected combinations the value of the likelihood for both the h and Z boson is calculated. The chosen pairing is the one in which the highest likelihood output for either the h or Z boson is found (i.e. max($\mathcal{L}_h$, $\mathcal{L}_Z$) ). The likelihood distributions are peaked towards values of 1 due to the nature of the background decay angles which is peak at 0. For Higgs masses close to threshold the efficiency of the decay angles pairing is of the order of 90%. The intrinsic pairing efficiency of the 2b analysis, i.e. events passing through the 2b branch with only one possible pairing, is also of the order of 90%.

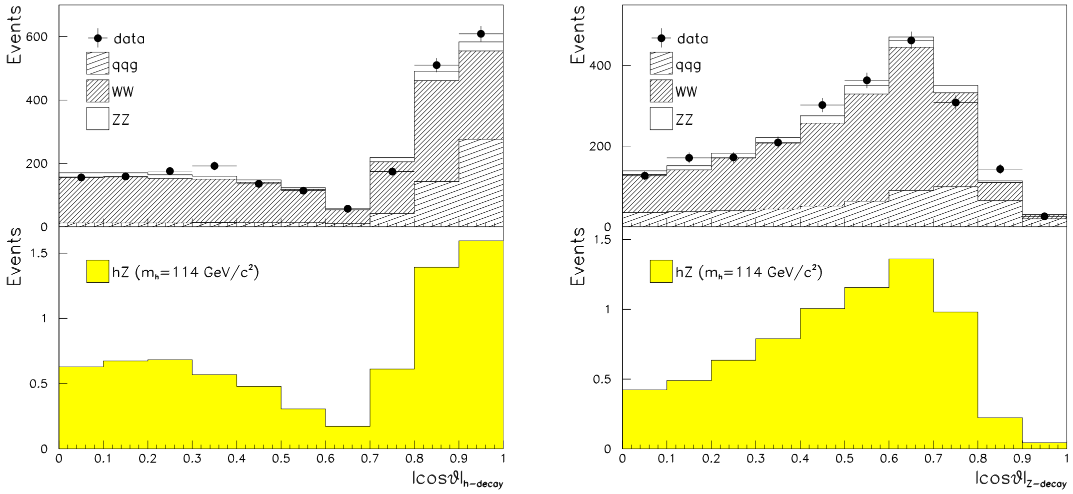


**Figure 4.16:** The distribution of the decay angles for the h and Z bosons for Monte Carlo background, data and a hypothetical signal with $m_{\mathrm{h}} = 114\,\mathrm{GeV}/c^2$.

| Event type | Fraction of events requiring pairing (%) |
|---|---|
| hZ | 42.42 |
| WW | 13.97 |
| ZZ | 30.44 |
| $\mathrm{q\bar{q}}$ | 25.00 |
| Tot bgd | 22.32 |
| Data | 21.21(7) |

**Table 4.3:** The fraction of selected events, in %, in each MC sample which require pairing. The number of recorded data events is shown in brackets.

The final analysis level is reached once all events are paired. At this point the value of the discriminating variable is calculated to provide information about each selected event.

The distribution of the discriminating variable, $m_{\mathrm{Reco}}$, for events passing through pre-selection with jet pairing performed is given in figure 4.17.

| | 1 comb | 2 combs | 3 combs | 4 combs | 5 combs | 6 combs |
|---|---|---|---|---|---|---|
| hZ | 85.26 | 14.38 | 0.29 | 0.06 | 0.0 | 0.0 |
| WW | 93.73 | 6.26 | 0.0 | 0.0 | 0.0 | 0.0 |
| ZZ | 88.61 | 11.16 | 0.21 | 0.0 | 0.0 | 0.0 |
| $q\bar{q}$ | 83.87 | 15.30 | 0.81 | 0.0 | 0.0 | 0.0 |
| Tot bgd | 87.81 | 11.78 | 0.37 | 0.0 | 0.0 | 0.0 |
| Data | 86.66(26) | 13.33(4) | 0.0 | 0.0 | 0.0 | 0.0 |

**Table 4.4:** The fraction of selected events, in %, passing through the 2b branch only which have "n" combinations. The number of recorded data events is shown in brackets.

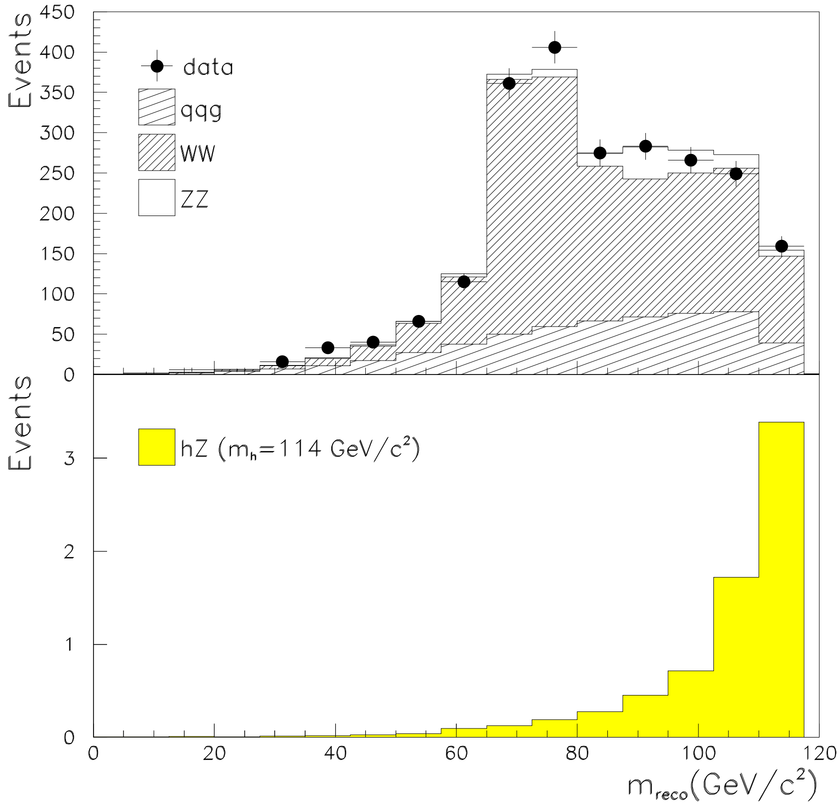


**Figure 4.17:** The distribution of the $m_{\mathrm{Reco}}$ for Monte Carlo background, data and a hypothetical signal with $m_{\mathrm{h}} = 114\ \mathrm{GeV}/c^2$.

## 4.4 Treatment of analysis overlaps

The four jets hZ cuts based analysis described here is one of four independent analyses which are used to search for each of the different final states of the hZ decay system [2]. The different analyses are designed to search for four different decay topologies of the hZ system and are known as decay channels. Although each decay channel is distinct from the others, it is still possible for one event to be selected by more than one of the four analyses. The treatment of the overlaps between analyses, when considering their application to MC, proceeds in two ways. In both cases the overlap between the analysis channels is expected to be very small and as such the problem of overlaps is treated in a conservative manner rather than attempting a full and rigorous treatment. Firstly, in the case of the application of the analysis to hZ signal MC, only simulated events corresponding to a defined particular state are used to determine the efficiency of that final state. This provides a conservative estimate (i.e. an underestimate) of the signal selection efficiency within that final state as no possible overlap from other final state MC is allowed to increase the expected number of final events. Secondly, to ensure a conservative estimate of the background efficiency within each search channel, all overlaps are ignored. In this case double counting of the backgrounds is allowed and this will lead to the subsequent over-estimation of the backgrounds in each channel. This double counting thus leads to a conservatively higher estimate of the expected background number compared to that which would be obtained by treating the background rigorously.

The overlap between the four jets and the $\mathrm{h}\ell^+\ell^-$ analysis is an exception to this rule and is correctly treated [59] [60]. The $\mathrm{h}\ell^+\ell^-$ analysis is designed to select final states in which the Z boson decays to an electron or muon pair. To eliminate the overlap between the two analyses, any candidate which is to be selected by the four jets analysis must have been rejected by the $\mathrm{h}\ell^+\ell^-$ analysis.

The method for treating the overlap within the data recorded is to apply the analyses sequentially in the order $\mathrm{h}\ell^+\ell^-$, $\mathrm{h}\nu\overline{\nu}$, $\tau^+\tau^-\mathrm{Z}$, $\mathrm{hq\bar{q}}$ . Any candidate which is selected by one of the given analysis is then removed from the data events stream which is passed to the subsequent analysis. This treatment ensures that no data candidate is selected by more than one analysis. When applied to the 2000 data set this treatment removes one data candidate from the four jets selection which is also selected by the $\tau^+\tau^-\mathrm{q}\overline{\mathrm{q}}$ analysis.

[2]Other channels are described in appendix C

## 4.5 Cuts optimisation and working point determination

The four jets event selection is optimised for a signal of chosen $m_{\mathrm{h}}$ and $\sqrt{s}$. The background selection efficiency is optimised as a function of the signal selection efficiency for this chosen Higgs mass. The Higgs boson mass is chosen to be just below the kinematic threshold of the $\sqrt{s}$ at which the optimisation is performed. This mass is chosen to ensure that the analysis is most performant in the region where an exclusion limit is expected to be set in the absence of a signal. This choice of mass also enhances the chance of finding any Higgs boson if it exists close to the limit of the experimental sensitivity.

The optimisation is achieved by performing a scan over a number of the cuts variables to find a minimum background selection efficiency as a function of the signal selection efficiency. A scan over the complete set of selection cuts within the four jets analysis would be impractical and would also result in the production of a set of cuts which are strongly dependent on the chosen optimisation mass. Although it is inevitable that performing an optimisation for one specific Higgs mass will indeed lead to some mass bias, a strong mass bias can be avoided, allowing sensitivity to a larger range of possible Higgs boson masses. Three of the four optimisation variables are from the 2b event selection whilst the fourth is the 4b cut variable itself. The choice of variables from both of the two selection ensures that the optimisation is performed for each selection and no loss in sensitivity is incurred. The four chosen variables are $\mathrm{m}_{12}$, $min(\eta_3, \eta_4)$ and $(1-\eta_3)(1-\eta_4)$ from the two b branch of the analysis and $9.5\mathrm{y}_{34} + \sum_{i=1}^{4} \eta_i$ from the four b branch. An extensive scan is performed over the four variables producing a set of optimal cut values for each possible selection efficiency.

Although these cuts define the optimal values for each selection efficiency the question of which selection efficiency is to be used for the analysis is still unanswered. The selection efficiency at which the analysis is to be performed is known as the working point. To determine the working point a full calculation of the confidence level for the exclusion of a signal is required (see Appendix B).

The selection efficiency which provides the greatest sensitivity for excluding a signal, in its absence, is chosen as the working point.

The optimisation was performed using $m_{\mathrm{h}}$=106.5 GeV/$c^2$ and $\sqrt{s}$=199.5 GeV for application to the 1999 data set [50]. To avoid possible bias the optimisation was frozen before the data taking period in 2000.

# Chapter 5

# Analysis application in 2000

## 5.1 Introduction

The operational mode of LEP 2 during the 2000 data taking period, as described earlier in Chapter 3, resulted in the collection of a continuum of data collected across a range of centre of mass energies, see figure 5.1.

This spread of data presents a minor problem when predicting the performance of the applied analysis both in the number of expected candidates and the distributions of the discriminating variables. To overcome this problem MC simulation samples were generated at numerous specific centre of mass energies. Interpolation methods were then used to allow the results gained from these centre of mass energy samples to be transformed into general results for any centre of mass energy. Table 5.1 describes the MC samples generated for the determination of the analysis performance and interpretation of the results during 2000.

This chapter describes the results of the application of the four jets cuts based analysis to the generated MC samples and thus the prediction of the analysis performance in 2000. The systematic uncertainties associated with the analysis performance are also detailed.

## 5.2 Analysis branches

The analysis of the four jets cuts channel is split in three statistically independent branches to allow the treatment of the overlap between the SM hZ→ four jets

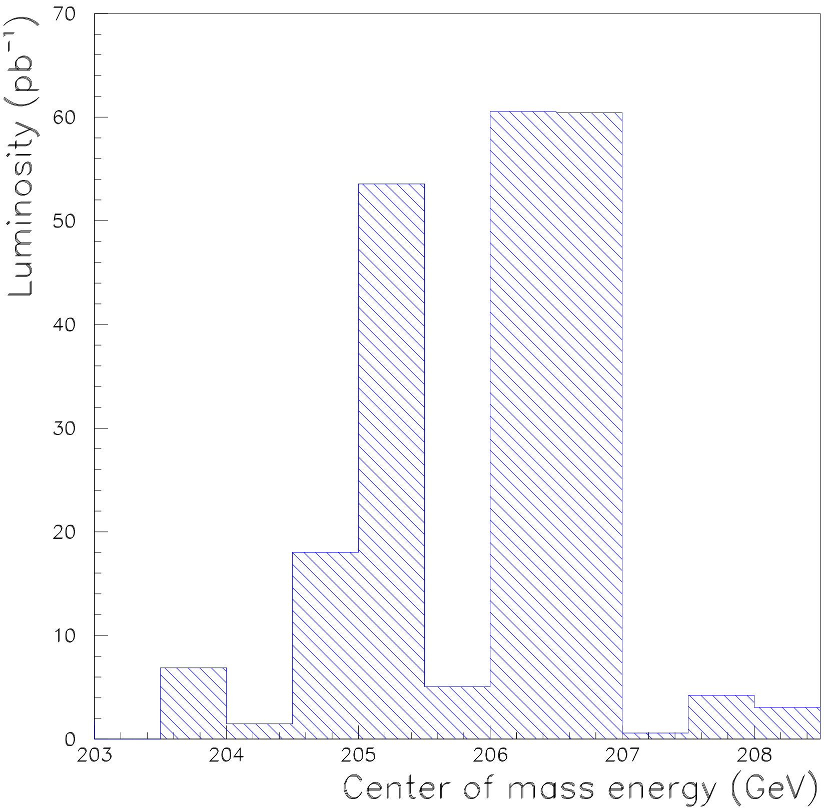


**Figure 5.1:** The distribution of luminosity collected in 2000 as a function of centre of mass energy.

and the MSSM hA$\rightarrow$ $b\bar{b}b\bar{b}$ analyses [60] which is detailed in appendix C.5. The three branches are referred to as hZ exclusive, hA exclusive and hZ/hA overlap since they contain events which are selected by only the hZ analysis, only the hA analysis and both analyses respectively. Figure 5.2 displays the construction of the branches. The interpretation of the analysis in terms of the hZ or hA searches is accomplished by considering the relevant exclusive branch and the overlap together. For example the hZ search considers the hZ exclusive + hZ/hA overlap branches and is referred to as the hZ stand-alone branch.

| $\sqrt{s}$ GeV | ZZ | $q\bar{q}$ | WW | hZ SIGNAL (7.5k) |
|---|---|---|---|---|
| 204 | 50k | 100k | 250k | 75-120 GeV/$c^2$ |
| 205 | 50k | - | 250k | - |
| 206 | 50k | 100k | 250k 250k$^\dagger$ | 75-120 GeV/$c^2$ |
| 206.7 | 250k | 2M $b\bar{b}$ , 500k $c\bar{c}$ | - | 80-130 GeV/$c^2$ |
| 207 | 50k | 400k | 250k, 1M, 250k$^\dagger$ | - |
| 208 | 50k | 100k | 250k | 75-120 GeV/$c^2$ |
| 209 | 50k | - | 250k | - |
| 210 | 50k | 100k | 250k | 75-120 GeV/$c^2$ |

**Table 5.1:** The MC simulation samples generated for use within the 2000 analysis. All of the detailed samples were used to determine the expected selection efficiencies while those marked with $^\dagger$ were omitted from the sample used to generate the probability density functions (pdf's),see section5.4.

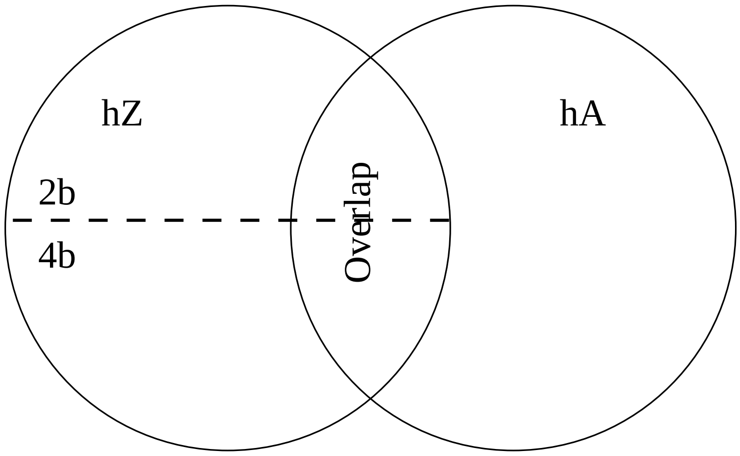


**Figure 5.2:** Depiction of the the three independent analysis branches.

## 5.3 Selection efficiencies

To determine the number of expected background and signal events requires three specific quantities at each centre of mass energy:

$$n = \sigma \cdot \mathcal{L} \cdot \epsilon \tag{5.1}$$

where $\sigma$ is the cross-section, $\mathcal{L}$ the collected luminosity and $\epsilon$ is the selection efficiency. The first two quantities are independent of the analysis, with the cross section being predicted theoretically and the collected luminosity being an observed quantity. The third value is dependent on the performance of the analysis itself. To determine the selection efficiency the analysis is applied to the MC samples for the background and signals. These MC samples are generated to cover the whole range of relevant backgrounds and possible signals that may be encountered within the search. The generation of the MC samples is dependent on various parameters. The background sources are simply dependent on the centre of mass energy at which they are generated. The signal MC on the other hand, in the cases of both hZ and hA, have extra dependencies. The search for a possible Higgs signal in both the hZ and hA cases requires a mass dependent production of MC samples, as a scan over a mass range must be performed and the selection efficiency over this range may not be constant. Thus both hZ and hA MC samples are generated across a range of $m_{\mathrm{h}}$ values for each centre of mass energy. The hA signal is also dependent on the value of $\tan\beta$ and, as such, MC samples are generated at two values of $\tan\beta$ these being 1 and 10 for $m_{\mathrm{h}} = m_{\mathrm{A}}$.

The selection efficiencies for the three main background sources as a function of centre of mass energy for the hZ exclusive analysis branch are shown in figure 5.4, while Table 5.2 shows the background selection efficiencies at $\sqrt{s} = 206$ GeV for the three analysis branches.

| Sample | hZ excl | hZ/hA Overlap | hA excl |
|---|---|---|---|
| ZZ | 1.526 | 0.614 | 0.456 |
| WW | 0.073 | 0.0076 | 0.0232 |
| $q\bar{q}$ | 0.029 | 0.0116 | 0.0246 |

**Table 5.2:** Background selection efficiencies(%) per analysis branch for the different background sources.

The interpolation for each background type in each analysis branch is performed by using a linear interpolation between the two nearest MC points, see figure 5.3. For example an efficiency for a centre of mass energy $\sqrt{S'}$ which lies somewhere between $\sqrt{S_1}$ and $\sqrt{S_2}$, with $\sqrt{S_2} > \sqrt{S_1}$, is found from

$$\sqrt{S'} = ((\sqrt{S_1} \cdot \Delta_2) + (\sqrt{S_2} \cdot \Delta_1))/(\Delta_1 + \Delta_2) \tag{5.2}$$

where $\Delta_1$ and $\Delta_2$ are the distances from $\sqrt{S'}$ to $\sqrt{S_1}$ and $\sqrt{S'}$ to $\sqrt{S_2}$ respectively. This method ensures that the resulting efficiency is a weighted average

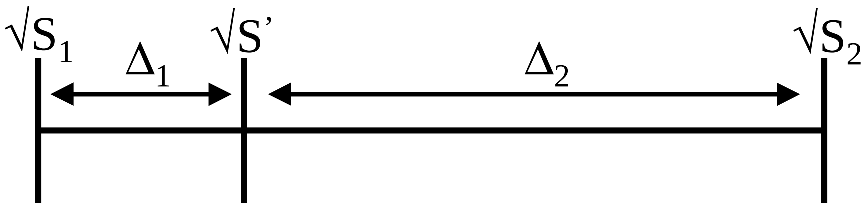


**Figure 5.3:** Diagrammatic depiction of the interpolation method for background samples.

of the two nearest MC samples and if the required point coincides with a MC sample then the result for that sample is returned.

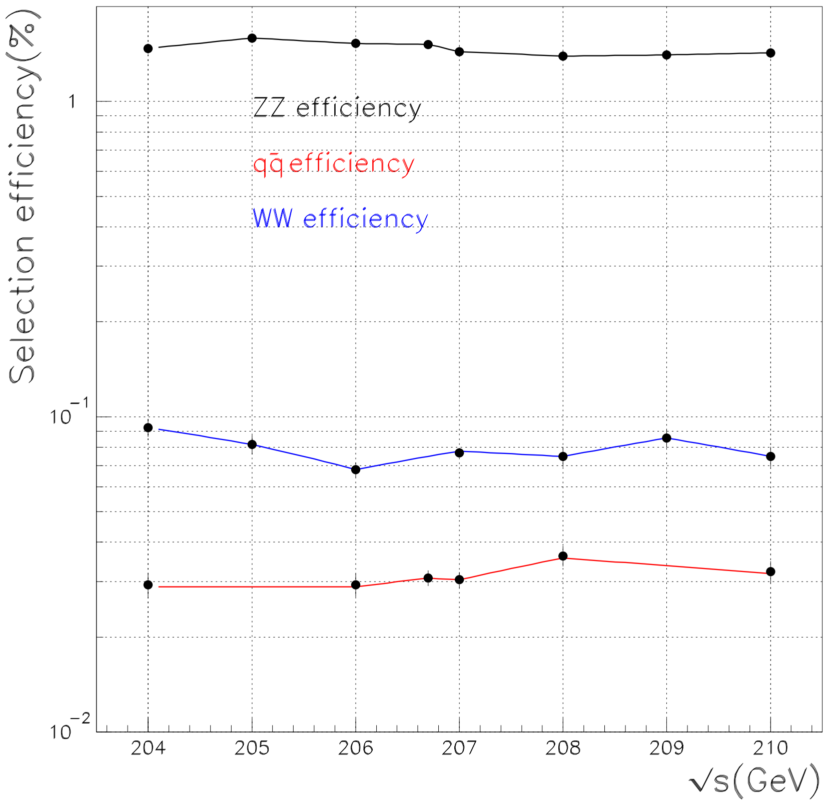


**Figure 5.4:** Background selection efficiencies in the hZ exclusive branch as a function of centre of mass energy. Solid lines show the interpolation while the points with errors show the MC samples.

The selection efficiencies for the hZ and hA signals are somewhat more complicated due to their dependence on the centre of mass energy and the mass of the Higgs boson in question. This leads to a selection efficiency which varies with $m_{\mathrm{h}}$ for each $\sqrt{s}$. This situation is greatly simplified by using the distance from threshold approximation [50] [61]. The distance from threshold approximation holds that any given Higgs boson of mass $m_{\mathrm{h}}$ which is a set distance from the kinematic threshold of a given centre of mass energy will have the same properties (e.g. selection efficiency) as a Higgs boson which is at that same distance from the threshold of its respective centre of mass energy. The kinematic threshold for a given centre of mass energy is defined as $kt = \sqrt{s} - m_{\mathrm{Z}}$.

Thus for example

$$\text{Efficiency}_{(m_{\mathrm{h}}=104,\sqrt{s}=204)} \equiv \text{Efficiency}_{(m_{\mathrm{h}}=108,\sqrt{s}=208)} \tag{5.3}$$

Following the distance from threshold approximation the values of $\sqrt{s}$ and $m_{\mathrm{h}}$ for each of the generated signal sample may be transformed into one variable, $kt$. The selection efficiencies may then be calculated as a function of this variable. Another advantage of this method is that the MC samples used at each $\sqrt{s}$, once translated into distance from threshold, may be added together linearly and averaged. This produces a much better approximation of the selection efficiency because of the increased amount of MC statistics available. All of the available signal MC may be combined to produce a general efficiency distribution as a function of $kt$. Figure 5.5 shows the selection efficiency as a function of Higgs mass for the hZ signal in the hZ exclusive branch of the four jets analysis at a centre of mass energy of 206.7 GeV. The figure shows a comparison between the actual efficiencies obtained from the MC generated at $\sqrt{s} = 206.7$ GeV and two interpolation methods which use the kinematic threshold approximation for MC samples at other $\sqrt{s}$'s to predict the nature at $\sqrt{s} = 206.7$ GeV. The two considered methods are firstly a polynomial fit to the selection efficiency at each $\sqrt{s}$ and secondly a linear interpolation. It is found that the linear interpolation describes the nature of the selection efficiency with greater accuracy than the polynomial fit method, particularly in the Higgs mass region close to threshold, and so the linear interpolation method is adopted.

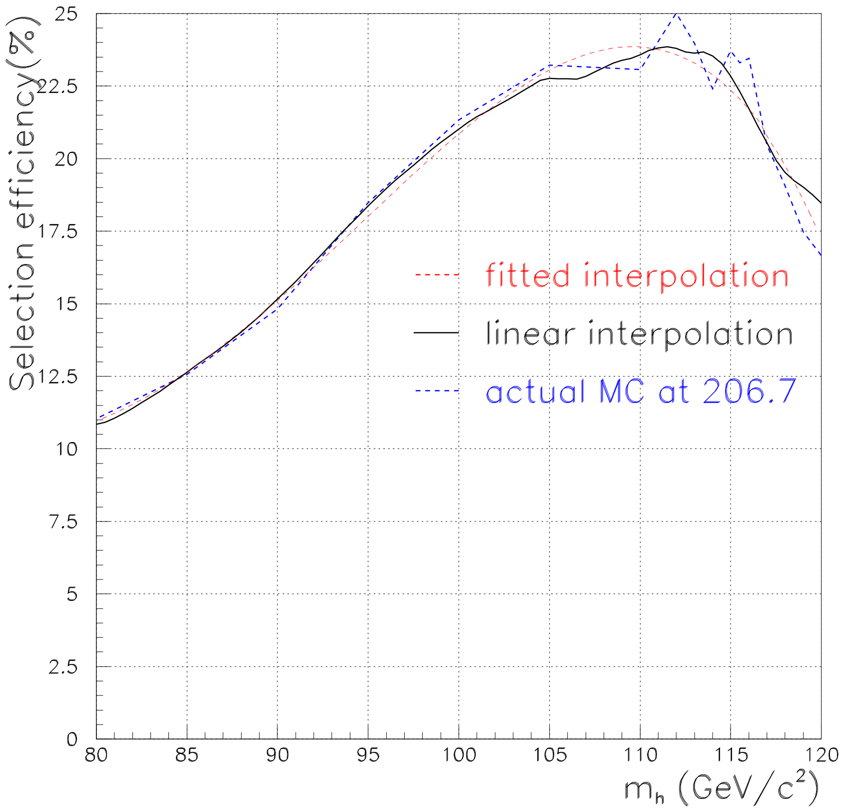


**Figure 5.5:** Signal selection efficiency as a function of $m_h$ for $\sqrt{s} = 206.7$ GeV in the hZ exclusive branch of the analysis. The linear and polynomial fit interpolations from MC samples excluding $\sqrt{s} = 206.7$ GeV are compared to the actual selection efficiencies at $\sqrt{s} = 206.7$ GeV.

To find the exact number of expected events the luminosity distribution for the recorded data is split into 110 bins of 100 MeV ranging from 200 GeV to 210 GeV. The integrated luminosity collected within each bin is then associated with the centre of mass energy in the centre of the bin. The relevant cross sections and selection efficiencies for each background and signal source are then used to calculate the expected number of events associated with the bin. Once all the bins have been accounted for, they are combined to give the total number of expected events.

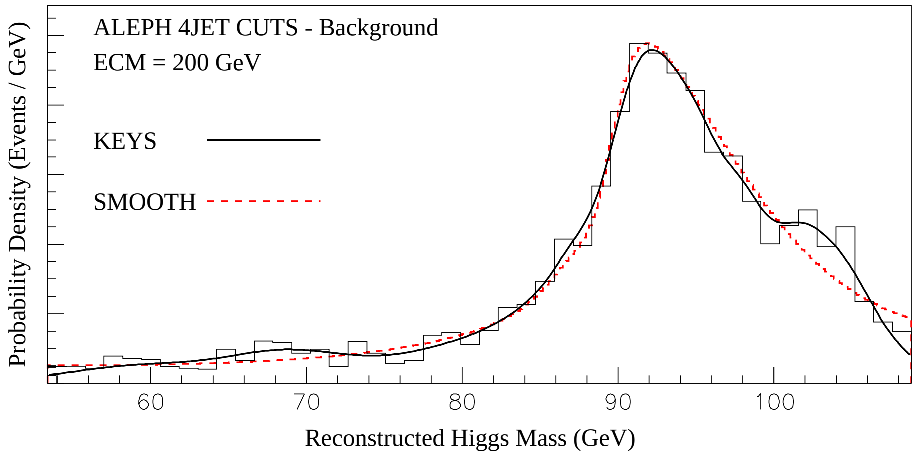


**Figure 5.6:** The pdf's obtained by KEYS(solid) and SMOOTH(dashed) for the background distribution of 4 jets cuts reconstructed Higgs mass at $\sqrt{s}$ = 200 GeV

## 5.4 The pdf's for event weighting

Each candidate event which is selected has associated with it a value of a discriminating variable, in this case the reconstructed Higgs mass ($m_{\text{Reco}}$). When the results of the search are finally analysed this variable is used to increase the ability to distinguish between signal and background. The probability that a given value of the discriminating variable originated from signal or background is used to provide a type of event weighting for the candidate (see appendix B). To correctly define these weights the expected population densities for the discriminating variable in both background and signal are required. These probability density functions (pdf's) are obtained by using the simulated MC event samples. Each MC sample is passed through the analysis and the resulting distribution for the discriminating variable is then formed from the selected events. The pdf's based on these selected events are formed by using an adaptive kernel based fit "KEYS" [62]. The pdf for the discriminating variable is required for each signal and background MC sample in each of the analysis branches, as is the case for the selection efficiencies. The use of the KEYS fitting method ensures that an unbiased and unbinned fit is formed from the sample of events selected by the analysis.

Figure 5.6 shows the resulting pdf's from two fitting methods when applied to a data set, in this case the complete four jets background in the hZ standalone branch for $\sqrt{s} = 200$ GeV. The compared methods are the KEYS and PAW SMOOTH [63] method which was used prior to the development of KEYS. The KEYS method is seen to provide a better fit to the data histogram at high reconstructed Higgs mass.

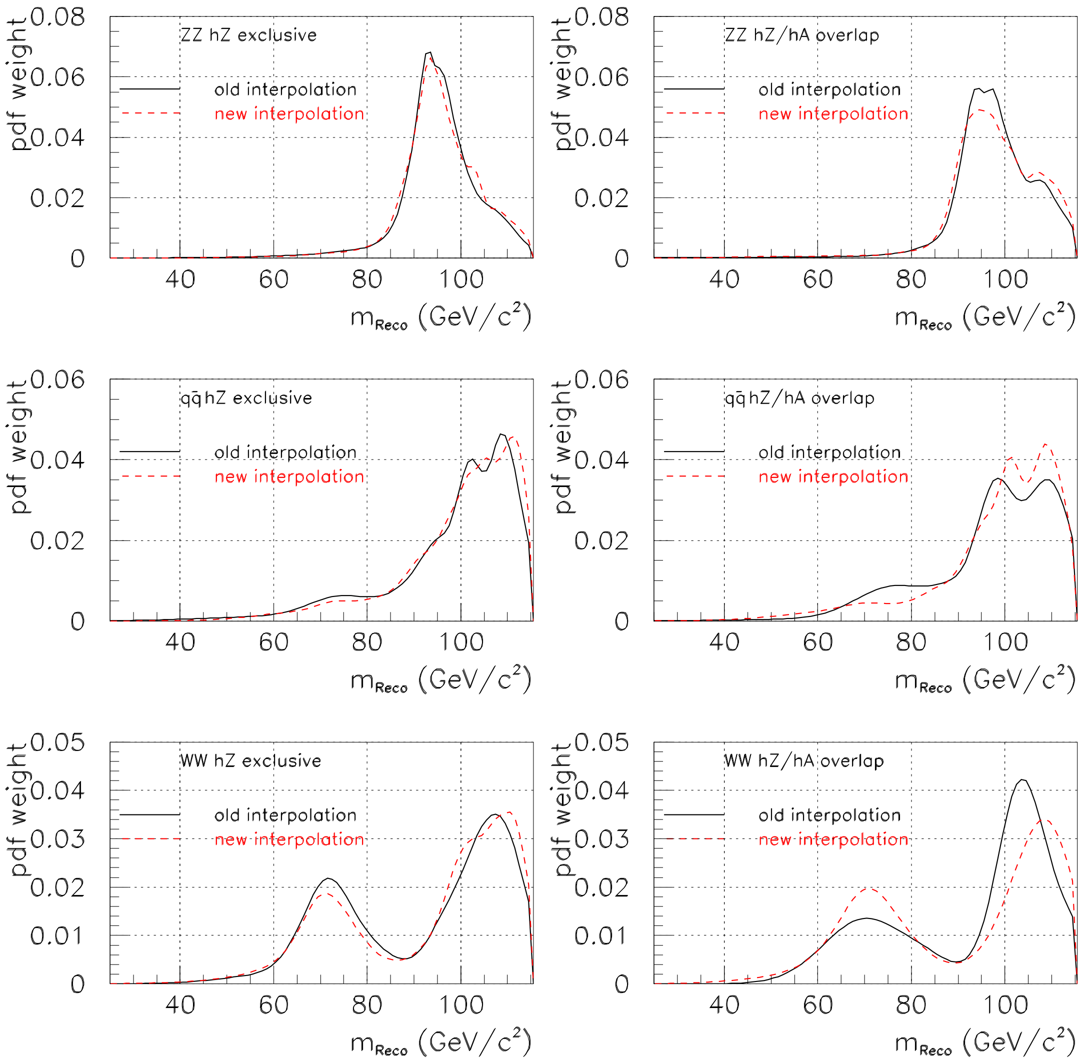


**Figure 5.7:** A comparison of interpolated pdf event shapes for the various backgrounds at $\sqrt{s} = 206.7$ GeV in the hZ exclusive and hZ/hA overlap branches.

Once all the pdf's have been fitted they are interpolated to provide a general result for the event weighting. The methods of interpolation for the background and signal pdf's are the same as those of the selection efficiencies. The background interpolation is again a simple linear interpolation across $\sqrt{s}$ and the signal interpolation is again achieved using the distance to threshold method [50] [61].

Figure 5.7 shows the results of background interpolations without (old) and with (new) several high statistics samples at and around $\sqrt{s} = 206.7$ GeV. The interpolations each predict the background distribution of the discriminating variable, $m_{\mathrm{Reco}}$, at $\sqrt{s} = 206.7$ GeV. The difference in the two interpolations can be seen as an indication of the accuracy of the interpolation method.

Figure 5.8 shows a comparison of interpolated event shapes versus actual event shapes from MC simulation for various Higgs signal masses at $\sqrt{s} = 206.7$ GeV. Good agreement is observed between the interpolation method and the true event shapes [64].

## 5.5 Systematic uncertainties

The selection efficiencies determined from the application of the analysis to MC sample are subject to possible systematic uncertainties.

Uncertainties arise due to the finite size of the MC samples used in the derivation of the selection efficiencies as well as inaccuracies in the modelling of the underlying physical events and the detector simulation. The size of the corresponding uncertainty is calculated and incorporated into the interpretation of the final results from the analysis.

The uncertainties arising from the finite statistics available for the calculation of the selection efficiencies are determined following the binomial approximation,

$$\sigma_{stat} = \sqrt{Np(1-p)}, \tag{5.4}$$

where N is the MC sample size, p = n/N and n is the selected number of events. The statistical components of the uncertainty are determined using equal sized MC samples at $\sqrt{s} = 206$ GeV and $\sqrt{s} = 208$ GeV with the final uncertainty

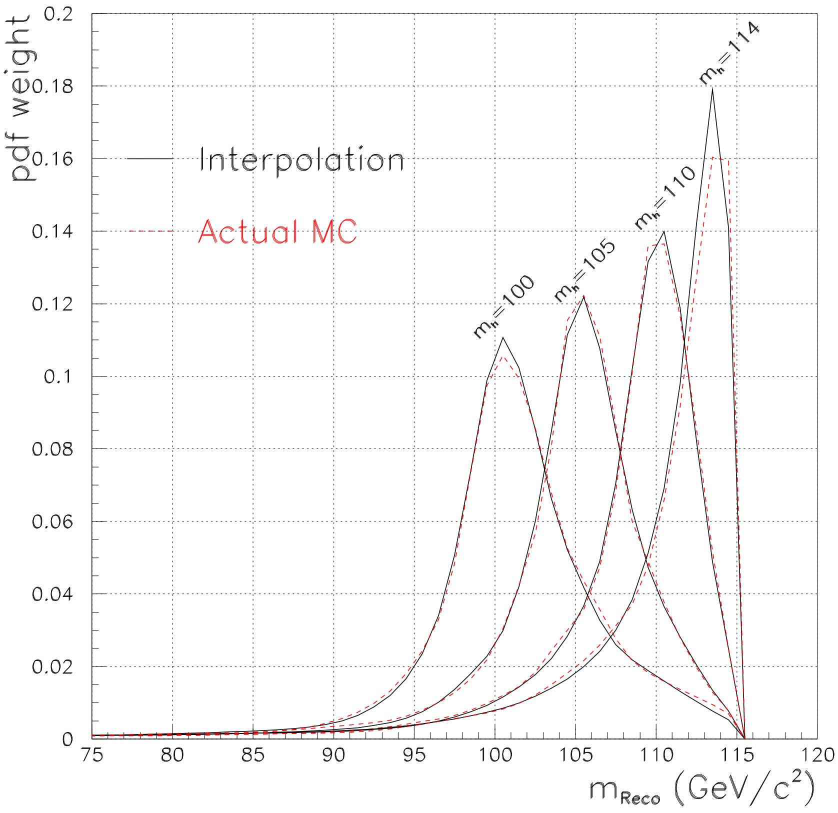


**Figure 5.8:** Comparison of interpolated and true MC pdf event shapes for various mass Higgs boson signals, in the hZ exclusive branch, as a function of $m_{\text{Reco}}$ at $\sqrt{s} = 206.7$ GeV.

| Analysis branch | hZ | $q\bar{q}$ | WW | ZZ | hA |
|---|---|---|---|---|---|
| hZ exclusive | ±2.10 | ±7.93 | ±6.27 | ±3.71 | ±5.13 |
| hA exclusive | ±6.58 | ±9.11 | ±10.75 | ±6.11 | ±1.54 |
| hZ/hA Overlap | ±2.68 | ±13.49 | ±21.47 | ±5.63 | ±1.63 |

**Table 5.3:** The uncertainty, $\sigma_{stat}$, expressed in relative %, associated with the selection efficiencies for signal and background, for each branch of the analysis.

then taken as the average of these two results. The uncertainties for various samples can be seen in table 5.3.

Other systematic errors can occur due to numerous sources. The largest expected sources of systematic error have been analysed using various methods. The systematic sources, once identified, are varied within regions which are representative of their uncertainties.

The resulting systematic uncertainties for the MC samples in each of the three analysis branches are shown in tables 5.4, 5.5 and 5.6 for the hZ exclusive, ha exclusive and hZ/hA overlap branches of the analysis respectively.

| Systematic Description | hZ | $q\bar{q}$ | WW | ZZ | hA |
|---|---|---|---|---|---|
| gluon splitting | 0.0 | $\pm$4.02 | 0.0 | 0.0 | 0.0 |
| $\alpha_s$ | 0.0 | $\pm$5.00 | 0.0 | 0.0 | 0.0 |
| B Smearing | 0.16 | 0.56 | -2.23 | 0.31 | 1.92 |
| Jet Smearing | -0.42 | 0.12 | -0.93 | -0.65 | -1.04 |
| B lifetime +1$\sigma$ | 0.53 | 0.44 | 0.17 | 0.83 | -0.63 |
| B lifetime -1$\sigma$ | -0.56 | -0.5 | -0.17 | -0.88 | 0.56 |
| B Multi +1$\sigma$ | 2.79 | 1.37 | -0.26 | 2.95 | -2.72 |
| B Multi -1$\sigma$ | -2.96 | -1.96 | 0.07 | -3.03 | 2.18 |
| B frag ($\varepsilon_b = 0.0025$) | -0.06 | -1.06 | 0.0 | -0.55 | 1.04 |
| B frag ($\varepsilon_b = 0.0037$) | 0.06 | 0.61 | 0.0 | 0.40 | -0.99 |
| C frag ($\varepsilon_c = 0.0302$) | 0.05 | -0.36 | 0.0 | 0.19 | -0.03 |
| C frag ($\varepsilon_c = 0.0376$) | -0.11 | 0.34 | 0.0 | -0.16 | -1.85 |
| $y_{34}^{\dagger}$ | -0.14 | -0.1 | -0.62 | -0.12 | 0.22 |
| $\gamma^{\dagger}$ | 1.54 | 0.27 | -0.28 | -0.77 | 1.29 |
| $\Theta^{\dagger}$ | -1.68 | -2.57 | -2.70 | -2.78 | -1.85 |
| $m_{12}{}^{\dagger}$ | -0.81 | -1.24 | -1.16 | -1.16 | -0.51 |
| $m_{34}{}^{\dagger}$ | -0.70 | 0.17 | -0.26 | -0.61 | -0.67 |
| Total | $\pm$3.89 | $\pm$7.82 | $\pm$3.88 | $\pm$4.53 | $\pm$4.34 |

**Table 5.4:** The systematic uncertainties on the event selection efficiencies, expressed in relative %, for background and signal in the hZ-exclusive branch of the analysis.

- **Gluon splitting:** The systematic error arising from the correction to the gluon splitting ratio to cc and bb quark pairs and also the correction to the strong coupling constant, $\alpha_s$, apply to the $q\bar{q}$ simulated events only. The incorrect weighting of gluon splitting events to $c\bar{c}$ and $b\bar{b}$ quark pairs in the generated MC is treated by re-weighting the MC to the values obtained from data [65]. The re-weighting factors are $c\bar{c}$=1.6 and $b\bar{b}$=1.9. The systematic error associated with this re-weighting is taken to be $\pm$50% of the re-weighting. The original MC sample and final selection number is calculated according to this re-weighting and the difference is assigned as a symmetrical systematic error. The gluon splitting systematic is calculated using the 206.7 GeV and 207 GeV $q\bar{q}$ samples due to their

| Systematic Description | hZ | $q\bar{q}$ | WW | ZZ | hA |
|---|---|---|---|---|---|
| gluon splitting | 0.0 | $\pm$9.89 | 0.0 | 0.0 | 0.0 |
| $\alpha_s$ | 0.0 | $\pm$5.00 | 0.0 | 0.0 | 0.0 |
| B Smearing | -2.36 | -7.78 | -10.22 | -2.82 | -0.46 |
| Jet Smearing | 1.23 | 0.78 | 2.75 | 2.36 | 0.34 |
| B lifetime +1$\sigma$ | -0.2 | 0.64 | 0.2 | 0.37 | 0.13 |
| B lifetime -1$\sigma$ | 0.14 | -0.72 | -0.21 | -0.42 | -0.37 |
| B Multi +1$\sigma$ | 1.63 | 1.88 | 0.62 | 3.84 | 0.44 |
| B Multi -1$\sigma$ | -1.30 | -3.06 | -0.46 | -3.10 | -2.97 |
| B frag ($\varepsilon_b = 0.0025$) | 0.32 | -1.8 | 0.0 | -0.51 | 0.17 |
| B frag ($\varepsilon_b = 0.0037$) | -0.05 | 1.31 | 0.0 | 0.45 | -0.06 |
| C frag ($\varepsilon_c = 0.0302$) | 0.20 | -0.16 | 0.0 | -0.45 | 0.02 |
| C frag ($\varepsilon_c = 0.0376$) | -0.17 | 0.13 | 0.0 | 0.41 | -2.08 |
| $y_{34}^{\dagger}$ | -0.84 | -1.58 | -0.85 | -0.76 | -0.29 |
| $\gamma^{\dagger}$ | 0.42 | 0.36 | -0.12 | 0.52 | 0.29 |
| $\Theta^{\dagger}$ | -1.9 | -2.46 | -1.76 | -0.76 | -2.05 |
| $m_{12}{}^{\dagger}$ | 2.53 | 2.17 | -0.24 | -0.54 | 1.75 |
| $m_{34}{}^{\dagger}$ | 6.96 | 5.62 | 0.79 | 1.55 | 4.39 |
| Total | $\pm$8.29 | $\pm$15.42 | $\pm$10.81 | $\pm$5.50 | $\pm$5.58 |

**Table 5.5:** The systematic uncertainties on the events selection efficiencies, expressed in relative %, in the hA-exclusive branch of the analysis.

large size and the final value of the systematic error is taken to be the average of these two values.

• $\alpha_s$**:** The correction to $\alpha_s$ is taken to be directly related to the $q\bar{q}$ cross section and a value of $\pm$5% is assigned to this systematic for the $q\bar{q}$ samples.

The following systematic errors are evaluated using the $\sqrt{s}$=206 GeV and $\sqrt{s}$= 208 GeV MC samples with the final systematic being the average of the two values obtained in each case.

Numerous systematic sources will affect the performance of the b-tagging within the Higgs search. The importance of b-tagging to the Higgs search requires that these systematics be treated rigorously. Uncertainties in b-lifetimes, jet multiplicities and heavy quark fragmentation are all expected to contribute to the overall systematic uncertainty associated with b-tagging and are treated as follows.

•**b-lifetime:** The lifetime of the b hadron produced within an event is directly related to the displacement of any secondary vertex observed within that event.

| Systematic Description | hZ | $q\bar{q}$ | WW | ZZ | hA |
|---|---|---|---|---|---|
| gluon splitting | 0.0 | $\pm$8.96 | 0.0 | 0.0 | 0.0 |
| $\alpha_s$ | 0.0 | $\pm$5.00 | 0.0 | 0.0 | 0.0 |
| B Smearing | -1.84 | -8.75 | -16.34 | -3.09 | -0.73 |
| Jet Smearing | 0.88 | 1.25 | -3.57 | 1.75 | 0.12 |
| B lifetime +1$\sigma$ | 0.64 | 0.7 | 0.02 | 0.97 | 0.66 |
| B lifetime -1$\sigma$ | -0.7 | -0.72 | -0.03 | -1.04 | -0.80 |
| B Multi +1$\sigma$ | 2.80 | -1.38 | -0.1 | 2.94 | 2.41 |
| B Multi -1$\sigma$ | -3.09 | 1.0 | -0.38 | -3.84 | -2.62 |
| B frag ($\varepsilon_b = 0.0025$) | -0.26 | -1.57 | 0.0 | -0.67 | 0.31 |
| B frag ($\varepsilon_b = 0.0037$) | 0.12 | 1.13 | 0.0 | 0.22 | -0.23 |
| C frag ($\varepsilon_c = 0.0302$) | 0.09 | 0.73 | 0.0 | 0.03 | 0.01 |
| C frag ($\varepsilon_c = 0.0376$) | -0.09 | -0.61 | 0.0 | -0.03 | -1.88 |
| $y_{34}^{\dagger}$ | 0.08 | -0.08 | 2.71 | -0.64 | -0.08 |
| $\gamma^{\dagger}$ | 1.35 | 0.36 | 0.0 | -0.65 | 0.99 |
| $\Theta^{\dagger}$ | -1.65 | -1.79 | -3.20 | -2.73 | -1.99 |
| $m_{12}{}^{\dagger}$ | -0.72 | -0.03 | -0.99 | -0.95 | -0.66 |
| $m_{34}{}^{\dagger}$ | -0.51 | 0.1 | 0.0 | -0.47 | -0.66 |
| Total | $\pm$4.32 | $\pm$13.82 | $\pm$17.27 | $\pm$5.89 | $\pm$3.77 |

**Table 5.6:** The systematic uncertainties on the event selection efficiencies, expressed in relative %, in the hZ/hA overlap branch of the analysis.

The variables $\mathcal{P}_{jet}$ and $\Delta\chi^2$ which are used within the b-tagging 6VNN, as described in section 3.6 , are directly dependent on the b-lifetime. To determine the systematic uncertainty associated with the lifetime of the b hadrons the values are varied within $\pm 1\sigma$ of the experimental uncertainty.

•**Jet multiplicity:** The typical high multiplicity of b decays is also used to discriminate between b-quark jets and light quark jets and is incorporated into the b-tagging NN as described in section 3.6. The systematic uncertainty associated with the multiplicity of b decays is estimated by varying the multiplicity values by $\pm 1\sigma$ of the experimental uncertainty.

• **Fragmentation:** The fragmentation of the heavy b and c quarks is parameterised by the Peterson function [66]

$$D_Q^H(z) \propto \frac{1}{z}\left(1 - \frac{1}{z} - \frac{\epsilon_Q}{1-z}\right)^{-2} \tag{5.5}$$

where $z$ is defined as

$$z = \frac{(E + p_{\parallel})_H}{(E + p)_Q} \tag{5.6}$$

with $p_{\parallel}$ representing the component of momentum along the direction of the primary quark and $\epsilon_Q$ is a parameter which is dependent on the heavy quark type and defines a probability density function which describes the nature of the fragmentation.

The fragmentation of the quarks is an important issue since the b-tagging performance is dependent on the momenta of the decay products of the heavy quark hadrons and also since increased momentum is expected to lead to an increase in the displacement of any secondary vertices associated with the decay of the b hadron.

The systematic uncertainty associated with the fragmentation is determined by calculating a weighting function from the ratio of the fragmentation pdf when using the value of $\epsilon_Q$ obtained from experiment and the value used in the generation of the simulated samples. The simulated events were generated with $\varepsilon_b$=0.0035 and $\varepsilon_c$=0.040 while the experimental values are found to be $\varepsilon_b = 0.0031 \pm 0.0006$ [67] and $\varepsilon_c = 0.0339 \pm 0.0037$ [68].

Figure 5.9 shows the distribution of the Peterson functions for the b quark flavour. The original value used to generate the MC sample is shown as well as the values corresponding to the observed result, within errors, and the weighting functions are also shown for the positive and negative errors on the function.

Smearing is applied to improve the agreement between the simulated samples and calibration runs performed at the Z peak. These smearing routines are treated as sources of systematic errors as follows.

• **b smearing:** The b smearing is performed on the track impact parameters ($d_0$,$z_0$) in simulated events to obtain improved agreement with distributions observed from calibration runs performed at the Z-peak. The associated systematic is taken to be half of the effect of the correction due to the application of the smearing.

• **Jet smearing:** The jet smearing improves the agreement between the mean jet energy and the jet energy resolution. The jet smearing is evaluated from a

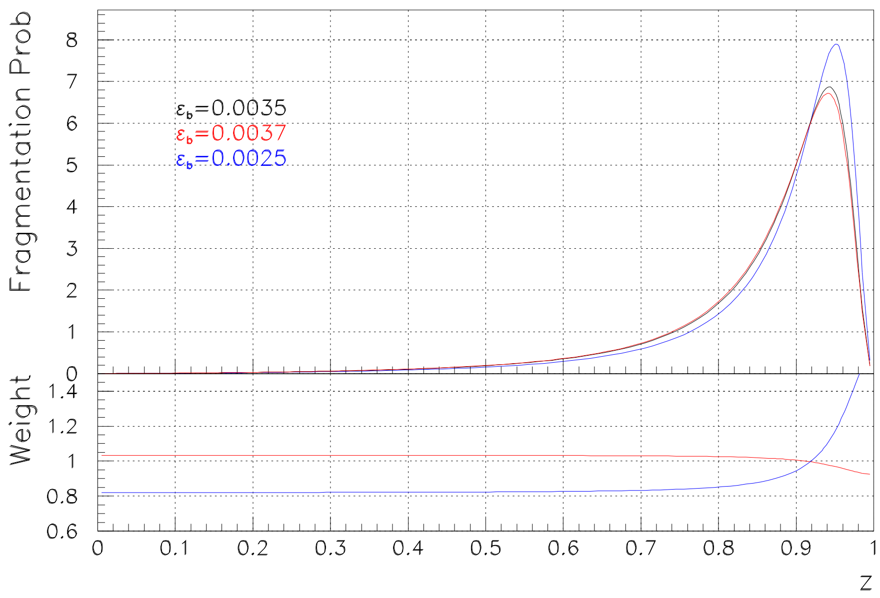


**Figure 5.9:** The Peterson fragmentation function for the b quark flavour for generated MC(black) and the experimental observed result within errors(red,blue). The weighting function is illustrated also.

sample of hadronic events collected during calibration runs at the Z peak. The associated systematic is taken to be half of the effect of the correction due to the application of the smearing.

• **Event selection variables ($y_{34}, \gamma, \Theta, m_{12}, m_{34}$):** The systematic uncertainties marked with a $^\dagger$ arise from variables which are used within the event selection. To account for any possible inaccuracies in the modelling of these variables a systematic error is determined by considering the agreement of the given variable between MC simulation and data. To determine the size of any associated systematic, a comparison is performed between the MC simulation and recorded data at the pre-selection level of the analysis, thus adequately large statistical samples are present. A binned histogram for each variable is considered for the comparison and a weight is assigned to each bin such that the weight represents the ratio of data to MC in that bin. Figure 5.10 shows a binned histogram for the $\gamma$ variable with data, MC and the corresponding weighting function. The comparison is performed at the pre-selection level to utilise the larger statistics available at this level and thus reduce any possible statistical effects. The presence of any signal at the pre-selection level is negligible and, as such, the comparison is performed between background and data. To further reduce the impact of statistical fluctuations the binned data and MC histograms are both normalised to the same unit area before the bin ratios are determined. This ensures that the comparison is based purely on the distribution of the event variable and is unaffected by any statistical fluctuations. Once the comparison is performed and the bin re-weighting factors are known the selection efficiencies for each MC sample are re-calculated with the relevant bin weighting applied to any candidate event which falls within that bin. Half of the observed shift in the selection efficiency is then taken to be a symmetric systematic associated with the selection variable.

Treating each source of systematic error independently the total systematic error associated with each MC sample in each analysis branch is calculated by combining all considered sources in quadrature. The total systematic error associated with each sample is combined in quadrature with the corresponding statistical error, $\sigma_{stat}$, to produce the total error associated with a given MC sample in a given analysis branch. The total combined errors for each event type in each of the three analysis branches is displayed in table 5.7.

| Analysis branch | hZ | $q\bar{q}$ | WW | ZZ | hA |
|---|---|---|---|---|---|
| hZ exclusive | 4.42 | 11.14 | 7.37 | 5.86 | 6.72 |
| hZ/hA overlap | 5.08 | 19.31 | 27.55 | 8.15 | 4.10 |
| hA exclusive | 10.58 | 17.91 | 15.25 | 8.22 | 5.79 |

**Table 5.7:** The total uncertainty, in relative %, for the different MC samples in the three different analysis branches.

## 5.6 Systematic uncertainty associated with pdfs

The pdf's generated for the event weighting and the interpolation method which is used to provide the general pdf from the finite samples are subject to systematic uncertainties due to the limited amount of MC statistics available for the generation of these pdf's. The systematic uncertainty associated with the pdf is determined by comparing the differences in pdf generated with a specific MC

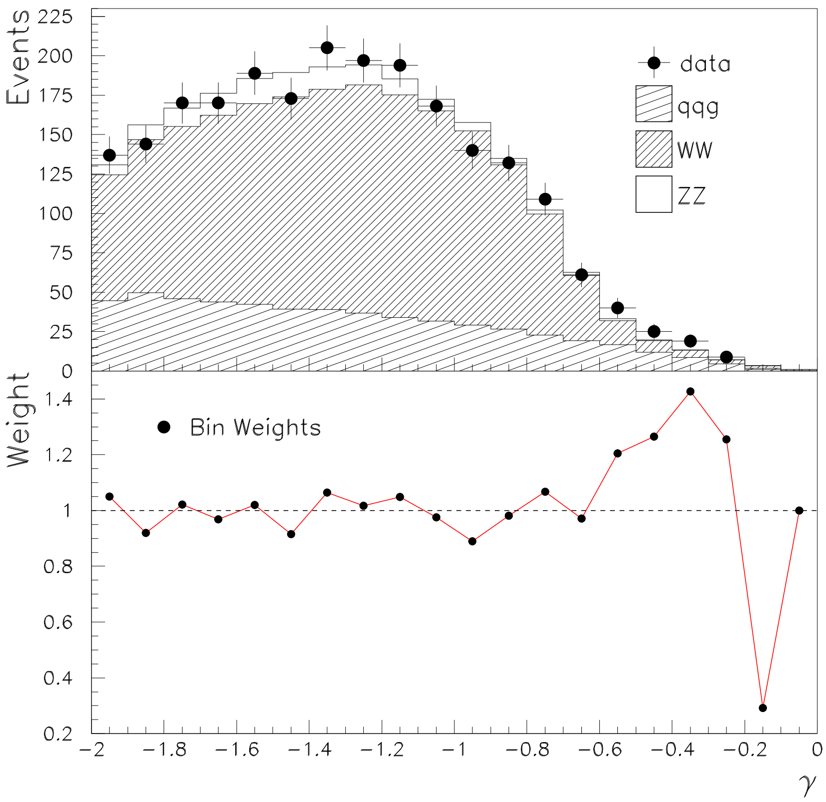


**Figure 5.10:** The data/MC comparison for the $\gamma$ variable with the individual bin weights illustrated.

sample and pdf's generated using the interpolation method. Figures 5.11 and 5.12 show the ratio of the actual pdf and interpolation for background and signal respectively [64]. The figures correspond to figures 5.7 and 5.8 which display the comparison of the pdf's. The systematic uncertainty is estimated to be half of the observed difference between the two pdf's in the 3 GeV/$c^2$ region leading up to the kinematic threshold. This is of the order of $\pm 10 - 20\%$ for the background samples and 1.5% for the signal samples.

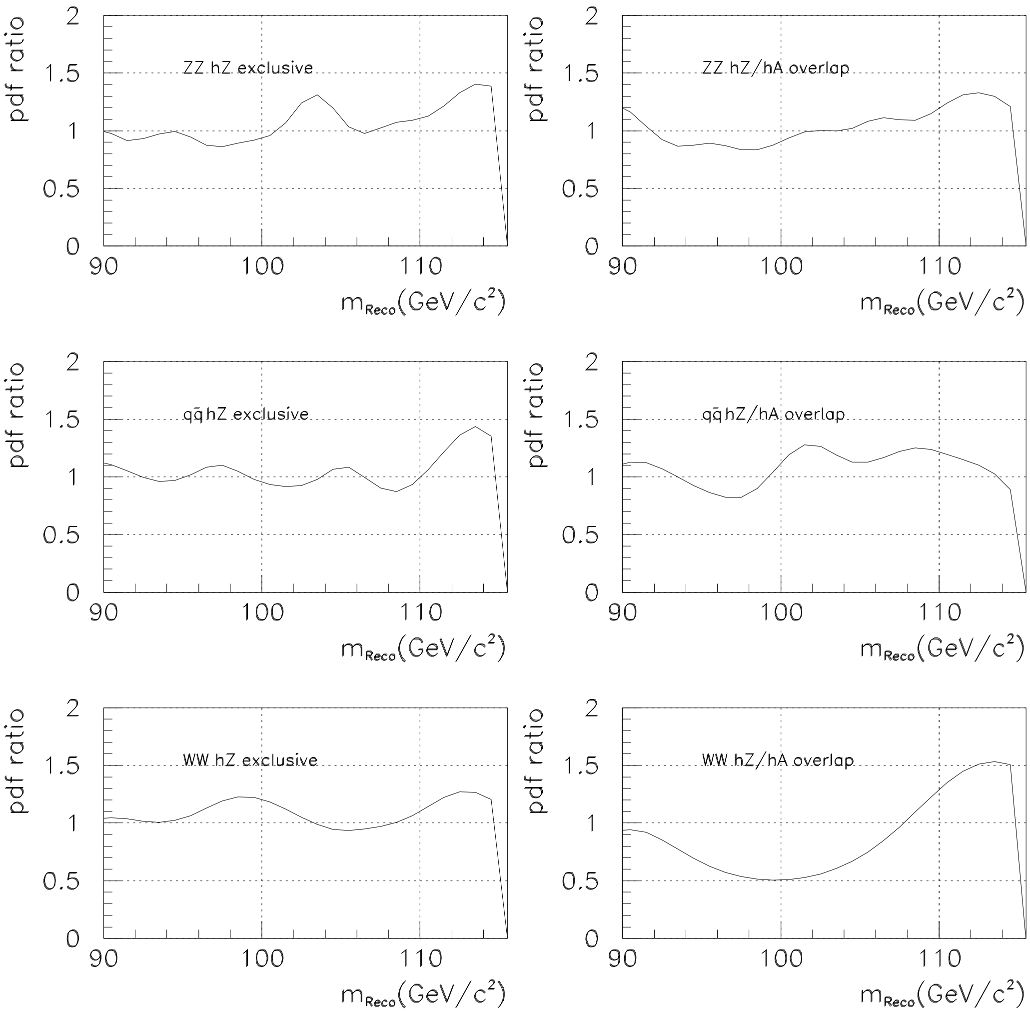


**Figure 5.11:** The ratio of the background pdfs from the MC simulation and interpolation to determine the systematic uncertainty. The pdfs for the hz exclusive and hZ/hA overlap branches are shown.

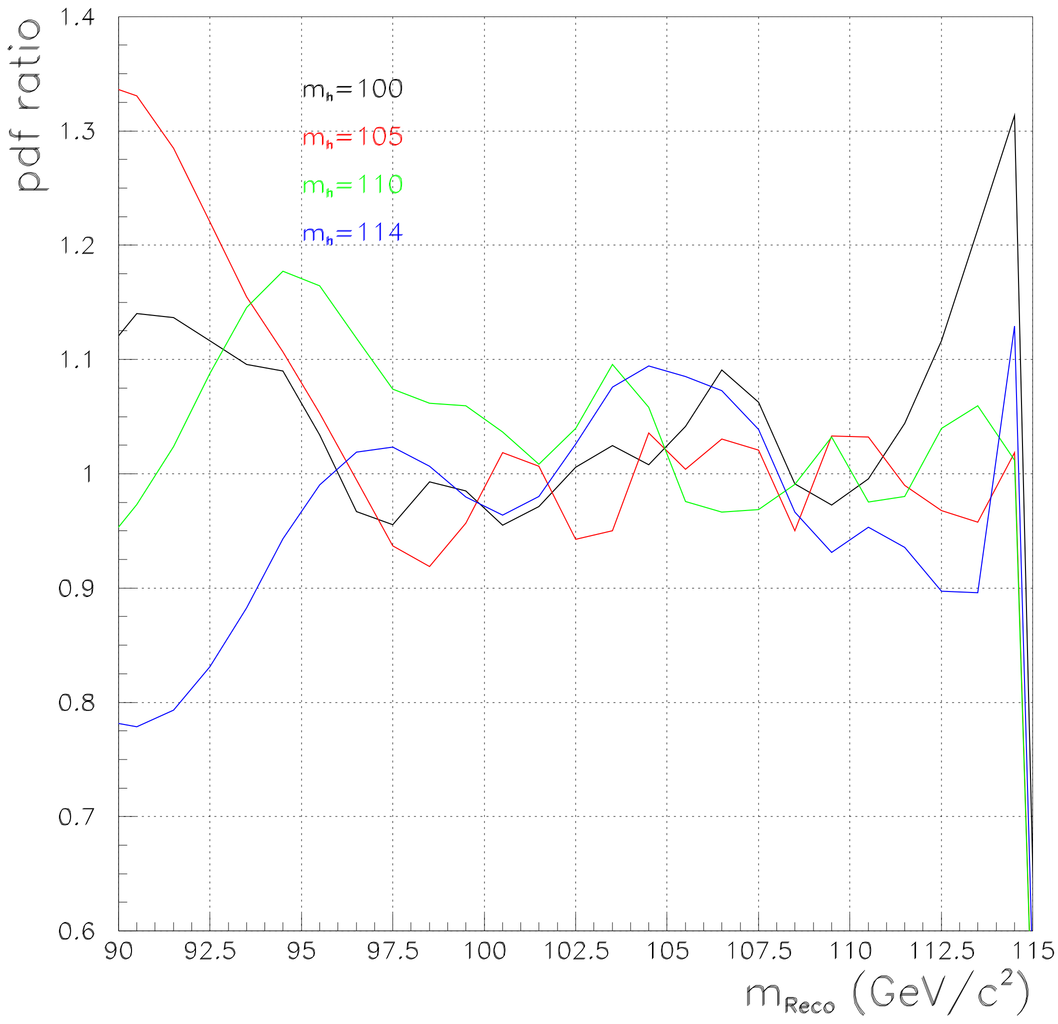


**Figure 5.12:** The ratio of the signal pdfs from the MC simulation and interpolation to determine the systematic uncertainty. The pdf are taken from the hZ exclusive branch.

# Chapter 6

# Experimental results

This chapter presents a review of the experimental results obtained from the four jets analysis when it is applied to the 2000 data set. Firstly the latest four jets results are presented and then the results form the latest combination of the ALEPH Higgs search channels [69] is covered. It is important to note that the inputs to the combination are from before data reprocessing and before the addition of the four jets systematic uncertainties determined for 2000. The combination does however use systematic uncertainties as determined in 1999 [50]. For this reason there exists a slight difference between the number of selected and expected events as summarised in the sections 6.1, which presents the latest numbers, and 6.2, which presents the data used during the latest full combination[1]. The chapter concludes with a more detailed description of some of the higher mass Higgs candidates and some discussion of the origin and credibility of the observed results.

## 6.1 Four jets results

The four jets cuts based analysis was applied to 217.18 $\text{pb}^{-1}$ of integrated luminosity recorded by the ALEPH detector in 2000. The data was recorded at centre of mass energies ranging from 200-209 GeV.

[1] The latest results are to be combined in the ALEPH Higgs search paper which is currently in preparation.

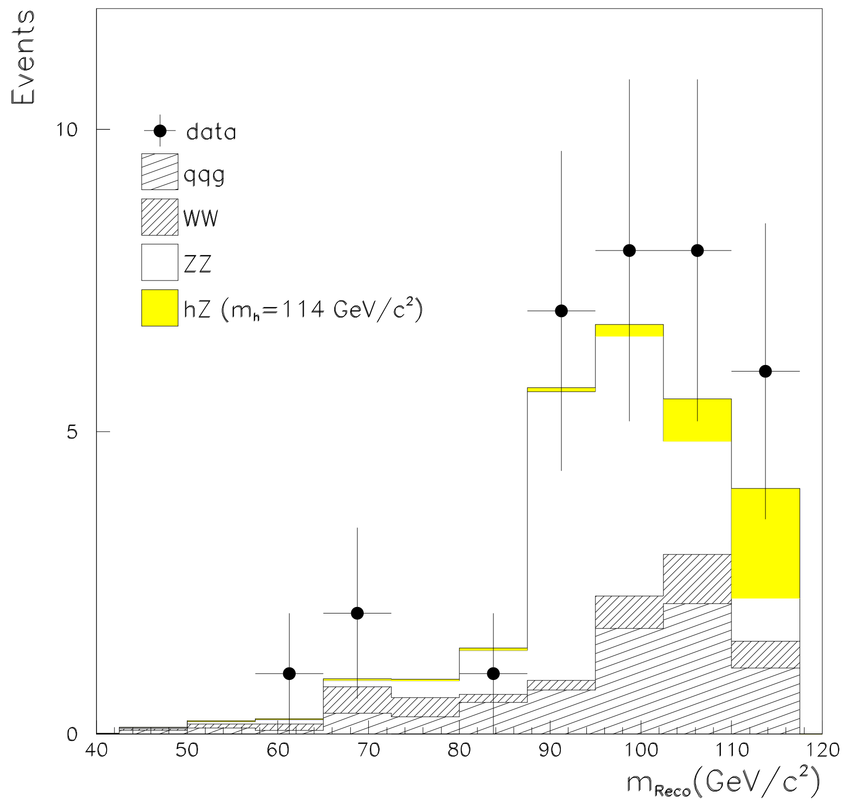


**Figure 6.1:** The distribution of the reconstructed Higgs boson mass for the expected background and signal and the observed data candidates for the hZ stand-alone branch of the four jets analysis.

| | hZ exclusive | hZ/hA overlap | hA exclusive |
|---|---|---|---|
| hZ | 1.725±0.076 | 1.182±0.060 | 0.237±0.025 |
| hA | 0.377±0.025 | 2.688±0.110 | 2.740±0.159 |
| $q\bar{q}$ | 5.148±0.573 | 1.952±0.377 | 3.985±0.714 |
| WW | 2.834±0.209 | 0.203±0.056 | 0.825±0.126 |
| ZZ | 9.033±0.529 | 3.899±0.318 | 2.897±0.238 |
| Tot bgd | 17.015±1.311 | 6.054±0.751 | 7.707±1.078 |
| Observed | 25 | 8 | 11 |

**Table 6.1:** The expected number of candidates, with associated systematic errors, and observed number of candidates events for the three analysis branches.

Table 6.1 shows the comparison of the expected and observed number of candidates in each of the three analysis branches. There is a visible excess in the number of observed candidates. Assuming the hypothesis of a SM Higgs boson and interpreting the search in terms of hZ, the hZ exclusive and overlap branches are combined to form the hZ stand-alone analysis branch. In this interpretation the combined results give a background expectation of 23.069±2.062 events compared to an observed 33 events. The distribution of the reconstructed Higgs mass for the hZ stand-alone interpretation of the 2000 data is shown in figure 6.1. Good agreement is seen at lower Higgs boson mass whilst an excess is observed in the high mass region of $m_{\mathrm{Reco}} > 110\ \mathrm{GeV}/c^2$.

The accuracy of the MC simulations, which are used to provide the expected number of candidates and discriminating variable ($m_{\mathrm{Reco}}$) distributions, is tested extensively in appendix D with both the number of expected events and the distributions of numerous variables being considered at several event selection levels.

Full details of all the selected data candidates are also presented in appendix E with the candidates broken down into the three selection branches and additional information presented about the alternate pairings for the high mass candidates.

## 6.2 Combined results

To optimise the sensitivity of the Higgs search within the ALEPH collaboration, independent analyses are performed on all possible decay topologies of the hZ (Higgsstrahlung) system. The decay channels are separated into the four jets ($\mathrm{hq\bar{q}}$), leptonic ($\mathrm{h}\ell^+\ell^-$), missing energy ($\mathrm{h}\nu\overline{\nu}$) and Tau ($\tau^+\tau^-\mathrm{Z}$) final sates with an independent analysis designed to search for each final state. The analysis of the four jets channel is described in chapter 4 and the remaining analyses are described in appendix C. The combination of the four final state analyses provides increased statistical power for the Higgs search. Table 6.2 details the expected performance of the different analyses in terms of both expected numbers of signal and background events and the expected significance in the presence of a $114\ \mathrm{GeV}/c^2$ Higgs boson signal. The number of data candidates observed for each analysis channel in the data collected in 2000 is also given. Figure 6.2 shows the distribution of the reconstructed Higgs mass for the 2000 data for both the expected events and observed data candidates. The figure shows good agreement for lower mass Higgs and a clear excess of candidates at high mass in the region around 114 GeV/$c^2$. The comparison of candidate number and the reconstructed mass distributions although providing an intuitive insight into the nature of the collected data by no means presents an optimal interpretation of the data. This comparison treats all candidates equally with no attention to the channel topology and $\sqrt{s}$ from which the candidate originated or the attributes of the candidates themselves. The interpretation of the individual candidates according to their analysis channel and their associated value of the relevant discriminating variable provide more statistical power for the final analysis.

| Search channel | Expected background | Expected Signal | Events Observed | Expected Significance($\sigma$) |
|---|---|---|---|---|
| 4-jet | 23.7 | 2.9 | 31 | 1.3 |
| h$\nu\nu$ | 19.7 | 1.3 | 20 | 0.7 |
| h$\ell\ell$ | 30.6 | 0.7 | 29 | 0.8 |
| $\tau\tau q\bar{q}$ | 13.6 | 0.4 | 15 | 0.4 |
| Total | 87.6 | 5.3 | 95 | 1.8 |

**Table 6.2:** The number of expected signal and background events for each analysis channel with the expected significance and number of observed candidates. The expected significane defines the significance in $\sigma$ which is expected in the presence of a 114 GeV/$c^2$ Higgs boson signal when the results are interpreted according to the statistical interpretation described in appendix B.

The likelihood ratio, equation 6.1, used in the analysis of the final data set incorporates the channel information and the discriminating variable value in addition to the expected number of events (Appendix B).

$$Q(m_{\mathrm{h}}) = \frac{\mathcal{L}_{\mathrm{s+b}}(m_{\mathrm{h}})}{\mathcal{L}_{\mathrm{b}}(m_{\mathrm{h}})} \tag{6.1}$$

To perform the complete statistical interpretation of the observed results data collected in 1998 [70] and 1999 [56] are combined with the 2000 data set.

The expected and observed distributions of -2 ln Q are shown in figure 6.3. The value of the likelihood ratio is usually expressed in this log likelihood form due to the relationship between the likelihood ratio and the chi-squared distribution. The logarithm of the likelihood ratio also has the property that individual candidate events will contribute as a sum of event weights, ln(1+s/b) where s/b is the signal to background ratio, which may then be treated individually. (see appendix B for a detailed discussion of the statistical interpretation of the results.)

In the presence of a deviation from the expected SM background the search results are interpreted in terms of the confidence in the background only hypothesis. Figure 6.4 shows the expected and observed distributions of $1 - \mathrm{CL_b}$ as a function of the hypothesised Higgs boson mass. A large deviation from the "median expected value of 0.5" is observed with a maximum deviation at a Higgs boson mass of $\approx$ 116 GeV/$c^2$. This deviation is consistent with an excess of events over the background hypothesis with the probability of observing such an excess being 1.1x$10^{-3}$. The significance of this excess when translated into Gaussian $\sigma$ is 3.1 $\sigma$ relative to the expected SM background.

The interpretation of the search results in the form of a lower limit on the Standard Model Higgs boson mass (see appendix B) leads to a 95% confidence level at Higgs mass of 110.6 GeV/$c^2$ with an expected limit of 113.8 GeV/$c^2$[2].

## 6.3 Individual candidate events

| | Run | Evt | $m_{\mathrm{Reco}}$ | $\mathrm{m}_{12}$ | $\mathrm{m}_{34}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ |
|---|---|---|---|---|---|---|---|---|---|
| a | 56698 | 7455 | 109.93 | 95.76 | 105.37 | 0.999 | 0.831 | 0.999 | 0.197 |
| b | 56065 | 3253 | 114.43 | 110.61 | 95.02 | 1.00 | 0.996 | 0.996 | 0.663 |
| c | 54698 | 4881 | 114.15 | 101.06 | 104.29 | 0.124 | 0.012 | 0.998 | 0.999 |
| d | 56366 | 955 | 114.38 | 78.55 | 127.03 | 0.201 | 0.051 | 0.998 | 0.956 |
| e | 55982 | 6125 | 114.42 | 80.00 | 125.62 | 0.071 | 0.306 | 0.499 | 0.998 |
| f | 58201 | 6835 | 111.75 | 80.66 | 122.30 | 0.096 | 0.277 | 0.965 | 0.870 |

**Table 6.3:** Details of the high impact four jet candidates which contribute a large event weight for high hypothetical Higgs masses. The di-jet masses are quoted as well as the reconstructed Higgs mass, b-tag for each jet and NN output from the alternate NN four jets analysis. Here jets 3 and 4 are associated with the Higgs boson. An RMS resolution in $m_{\mathrm{Reco}}$ of $\approx$ 9 GeV/$c^2$ is achieved for a MC hZ signal near to kinamatic threshold.

[2]The alternative Neural Network combination produces similar results for both the significance of the observed deviation from the SM background expectation and the observed lower limit [69].

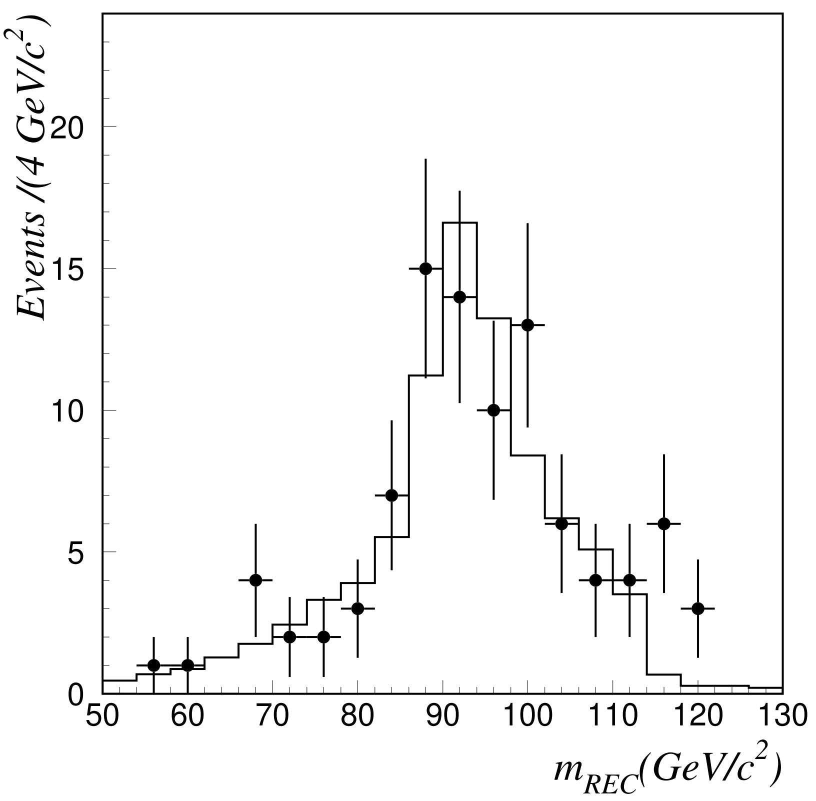


**Figure 6.2:** The distribution of the reconstructed Higgs boson mass for the data collected in 2000 (points with error bars) and the expected background (histogram) for all search channels [69].

The "Quality" of an individual candidate event may be assessed by two, quite different, methods.

Firstly, the impact of the candidate on the analysis can be determined in terms of its contribution to the test statistic, in this case the likelihood ratio $Q$, which is used to determine how signal-like or background-like the search results are. This contribution may be referred to as the weight of the candidate event. The event weights for several of the four jet candidates are displayed in figure 6.5[3] as a function of the hypothesised Higgs boson mass. Details of these candidates are given in table 6.3.

[3]The figure corresponds to the unreprocessed data. The reconstructed masses vary little between the reprocessed and unreprocessed events, generally of the order of $\approx$200 MeV$/c^2$, and so the figure is still valid.

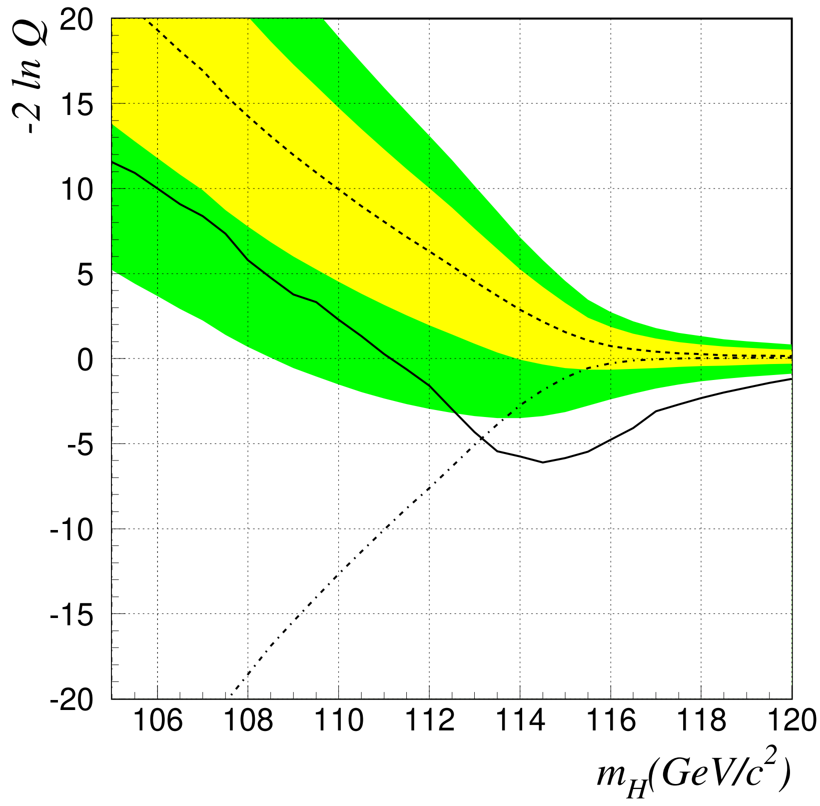


**Figure 6.3:** The distribution of the log-likelihood estimator -2ln$Q$ as a function of the hypothesised Higgs boson mass for the observed (solid) background only (dashed) and signal+background (dot-dash). The light and dark regions around the background only expectation represent the one and two sigma bands respectively [69].

The weight of the candidate peaks at a value of hypothesised Higgs mass close to the reconstructed Higgs mass associated with the candidate. In general candidates with a high reconstructed Higgs boson mass will have a larger weight than those with a smaller reconstructed mass due to the increased separation of signal and background sources at high mass. It should be noted that candidates selected in the hZ/hA overlap branch will in general have a greater event weight than a candidate from the hZ exclusive channel with the same reconstructed mass. This is due to the higher signal to background ratio associated with the hZ/hA overlap branch.

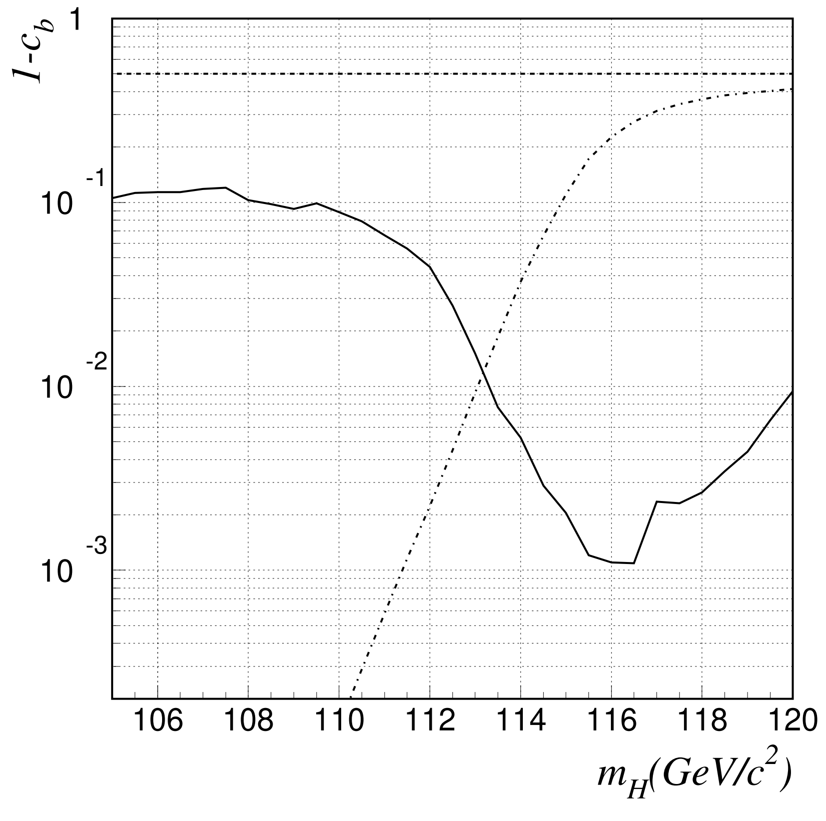


**Figure 6.4:** The observed (solid) and expected background only (dashed) CL curves for the background hypothesis as a function of the hypothesised Higgs boson mass. The dot-dash curve indicates the expected median CL for a signal+background hypothesis as a function of the Higgs boson mass [69].

The quality of any given candidate may also be assessed independently of its event weight. To determine the quality of a candidate in this manner the concept of "purity" is introduced. Purity is defined as the ratio of signal ($m_\mathrm{h}$ = 114 GeV/$c^2$) to background (s/b) for events with a reconstructed Higgs mass greater than 109 GeV/$c^2$, from here on denoted $(s/b)_{109}$. The purity is calculated using MC simulation and cuts selections with increasing purity achieved by tightening the cuts of the event selection analysis. Although the optimum working point for the analysis has been previously determined, see chapter 4.5, and no increase in sensitivity is expected to be gained by operation with different cut values, the tightening of the cuts and increasing purity can be used to perform checks on the analysis and also to aid the analysis of the individual candidates.

The determination of the working point involved the production of an optimal set of cut values for each possible signal selection efficiency. In this procedure the cuts were optimised to provide the lowest possible background for each signal selection efficiency.

Several of these selection efficiency points were chosen to provide a range of different purities for the analysis. Figure 6.6 shows the distribution of the reconstructed mass for expected background, expected signal and observed data for cuts of increasing purity as described in table 6.4. The cuts start at the selected working point with a purity of 0.68 and increase in steps through to higher purity selections. Table 6.4 gives the purity selection numbers for both signal, background and data and also the signal selection efficiency and the applied cuts.

| Purity Level | A | B | C | D |
|---|---|---|---|---|
| Exp Sig | 2.907 | 1.851 | 1.107 | 0.732 |
| Exp Bgd | 23.068 | 11.172 | 5.208 | 2.778 |
| Observed Data | 33 | 15 | 6 | 3 |
| Signal Eff(%) | 40 | 25 | 15 | 10 |
| $(s/b)_{109}$ | 0.68 | 1.26 | 1.47 | 1.64 |

**Table 6.4:** The number of expected background, signal and observed data events for the varying purity levels for the hz stand-alone branch. The $(s/b)_{109}$ value is quoted as well as the signal selection efficiency for the given purity level.

As the purity of the selection is increased the b-tag cut becomes very strict. This tightening of b-tagging is consistent with a search for the Higgs boson in which the major decay mode is to $b\bar{b}$. Increasing the purity of the selection leads to smaller number of expected signal and background events and also to a loss of data candidates. The presence of data candidates at high levels of purity is indeed consistent with a signal like hypothesis and the excess of high mass candidates is observed throughout the increasing purity cuts thus showing that the observed result is stable. Three data candidates are selected with high purity cuts. Two of the candidates which were highlighted due to their large impact on the analysis are found to pass the tightened selection cuts.

Although exercises in purity and event weights show which candidates have the largest impact on the analysis and also which candidates survive the tightest

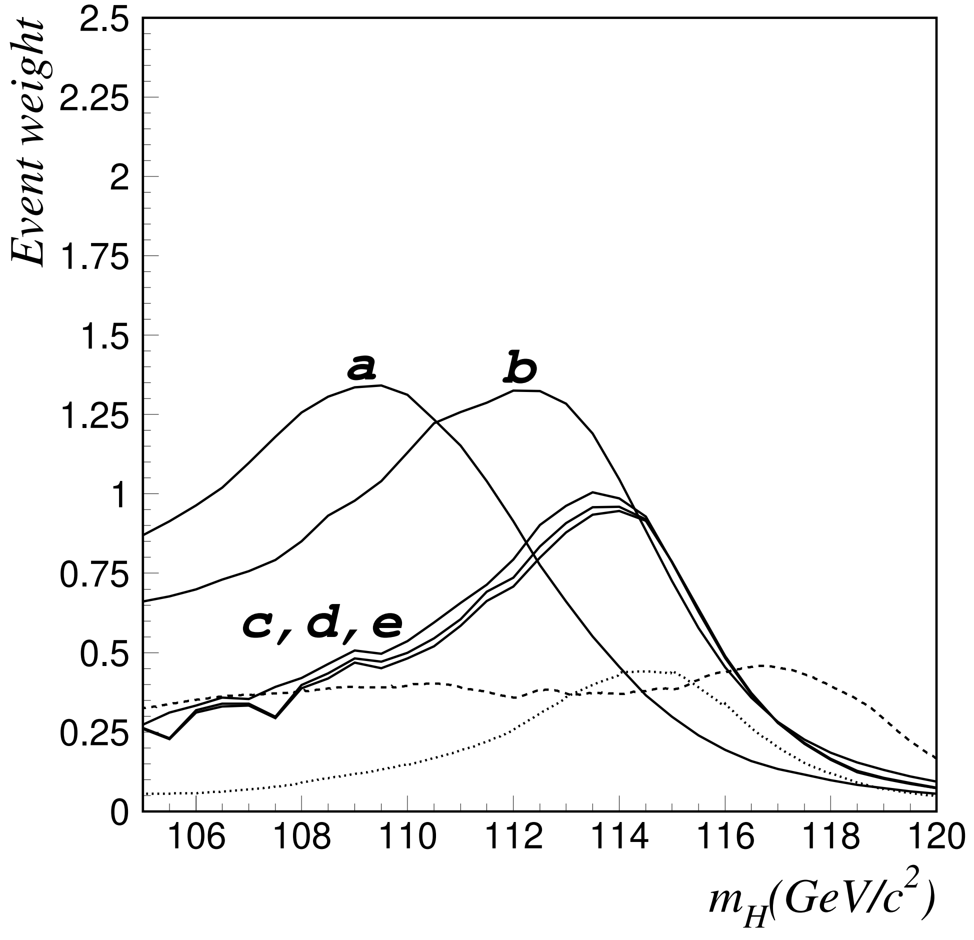


**Figure 6.5:** The event weight evolution as a function of the hypothesised Higgs boson mass for the high impact candidates which have a large event weight at high Higgs mass. The four jet candidates(solid) contribute the greatest with the highest contributing lepton(dashed) and tau(dotted) candidates highlighted for comparison [69].

cuts, they are only guides. The interpretation of the whole of the data set at the optimised working point still provides the final result for the analysis and the use of all the candidates at this working point is proven to be optimal.

Three of the "high impact" candidates highlighted in table 6.3 are found to possess a high NN output from the 19 variable Neural Network [71] [69] which is used to form an alternate four jets analysis to the cuts based analysis presented in this thesis. The NN four jets analysis uses the reconstructed Higgs boson mass and the NN output as discriminating variables and, as such, these three common candidates will have a high impact in both streams. These three candidates are thus described in more detail here. Two of the candidates are selected in the hZ/hA overlap branch and have a 4b nature with the final candidate selected in the hZ exclusive branch and appears to be 2b in origin.

The first candidate (run 54698, event 4881) occurs in the hZ exclusive branch of the analysis. The event, displayed in figure 6.7, was recorded at $\sqrt{s} = 206.6$ GeV. The candidate has four well separated jets two of which are very well b-tagged (0.998, 0.999) and form the $b\bar{b}$ pair associated with the Higgs boson decay. The secondary vertices associated with these b-tagged jets are clearly seen in the zoomed region of the interaction point in the bottom right of figure 6.7. One of the b-jet candidates also has an identified muon which is consistent with the hypothesis of a semi-leptonic b decay. The candidate has a reconstructed Higgs boson mass of 114.15 GeV/$c^2$ and a NN output of 0.997.

The second candidate (run 56698, event 7455) occurs in the hZ/hA overlap branch and is displayed in figure 6.8. The event has three well b-tagged jets with values of 0.999, 0.831 and 0.999, but the fourth jet is poorly b-tagged and has a value of only 0.179. From MC simulation the probability of a hZ $\rightarrow b\bar{b}b\bar{b}$ event possessing such a low b-tag within one of its jets is found to be 19% [69]. The event was recorded at $\sqrt{s} = 206.7$ GeV and has a reconstructed Higgs boson mass of 109.9 GeV/$c^2$ and NN output of 0.999. The interpretation of the candidate as a ZZ event, using an appropriate kinematic fit, leads to very large boson masses of 98.9 and 101.6 GeV/$c^2$.

The third candidate (run 56065, event 3253), displayed in figure 6.9, was collected at $\sqrt{s} = 206.7$ GeV/$c^2$. The event is selected in the hZ/hA overlap branch

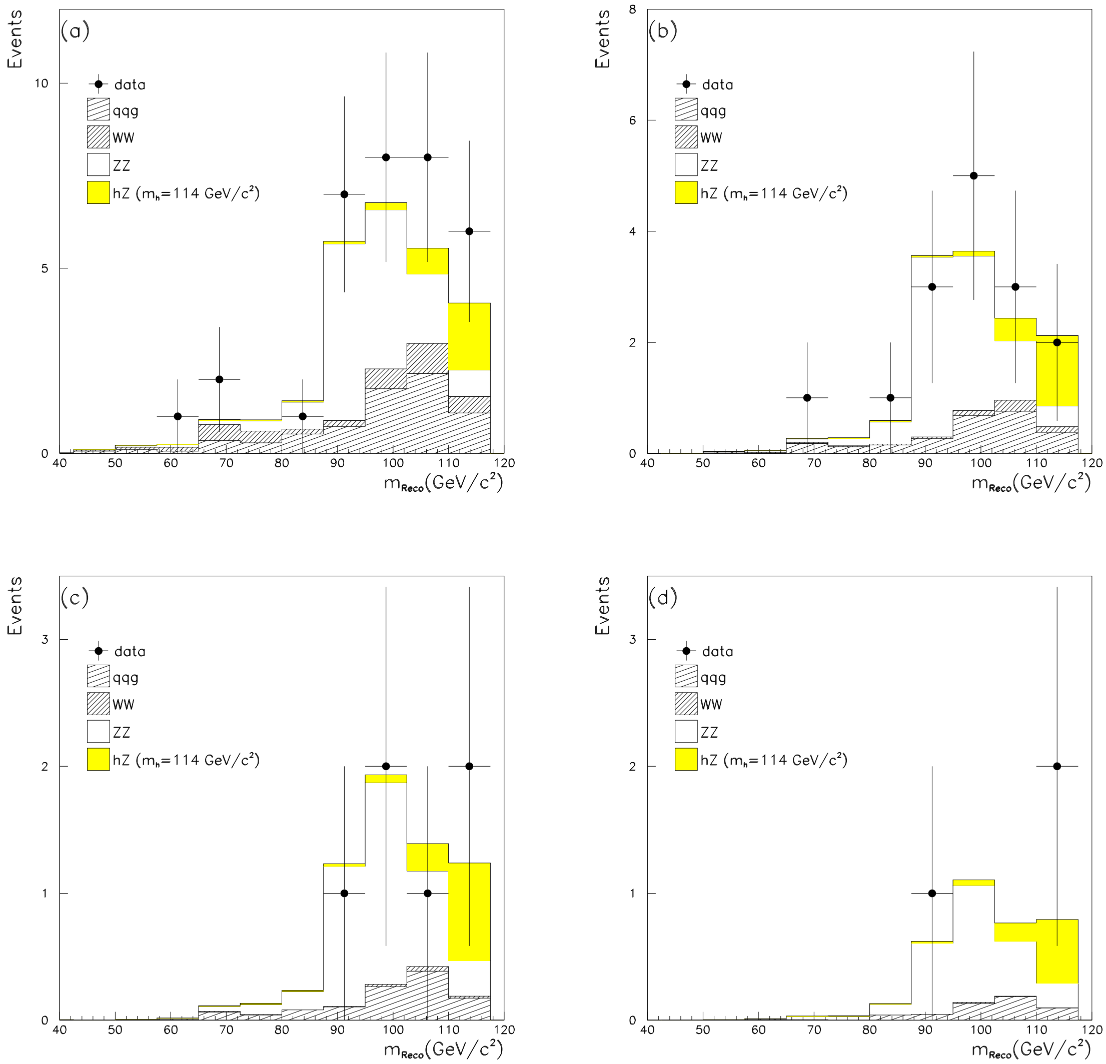


**Figure 6.6:** The distribution of the reconstructed Higgs boson mass for varying $(s/b)_{109}$ purity levels(a=0.68,b=1.26,c=1.47,d=1.64) with data (points with errors), background and signal histograms.

and has a high NN b-tag output for all four jets (0.996, 0.663, 0.992, 0.999) which indicates the event is unlikely to have originated from $W^+W^-$ background. The event has a very large visible energy of 252 GeV. This is partially attributed to the presence of a high energy deposit, some 22 GeV, which is observed at a very low $\theta$ angle in SICAL. A possible hypothesis for the observation of such an object is the presence of beam related noise. An extensive study of beam related noise was performed with the production of an event cleaning algorithm which is specifically designed to remove such background from data events. The study of the beam induced noise and development of the cleaning algorithm is described in detail in chapter 7. The application of the cleaning algorithm to the data candidate described above leads to a change in the reconstructed Higgs boson mass (NN output) from 112.8 GeV/$c^2$ (0.996) to 114.43 GeV/$c^2$(0.964). The small shift in the reconstructed Higgs boson mass reflects the dependence of the fitted masses on the jet direction rather than the jet energies.

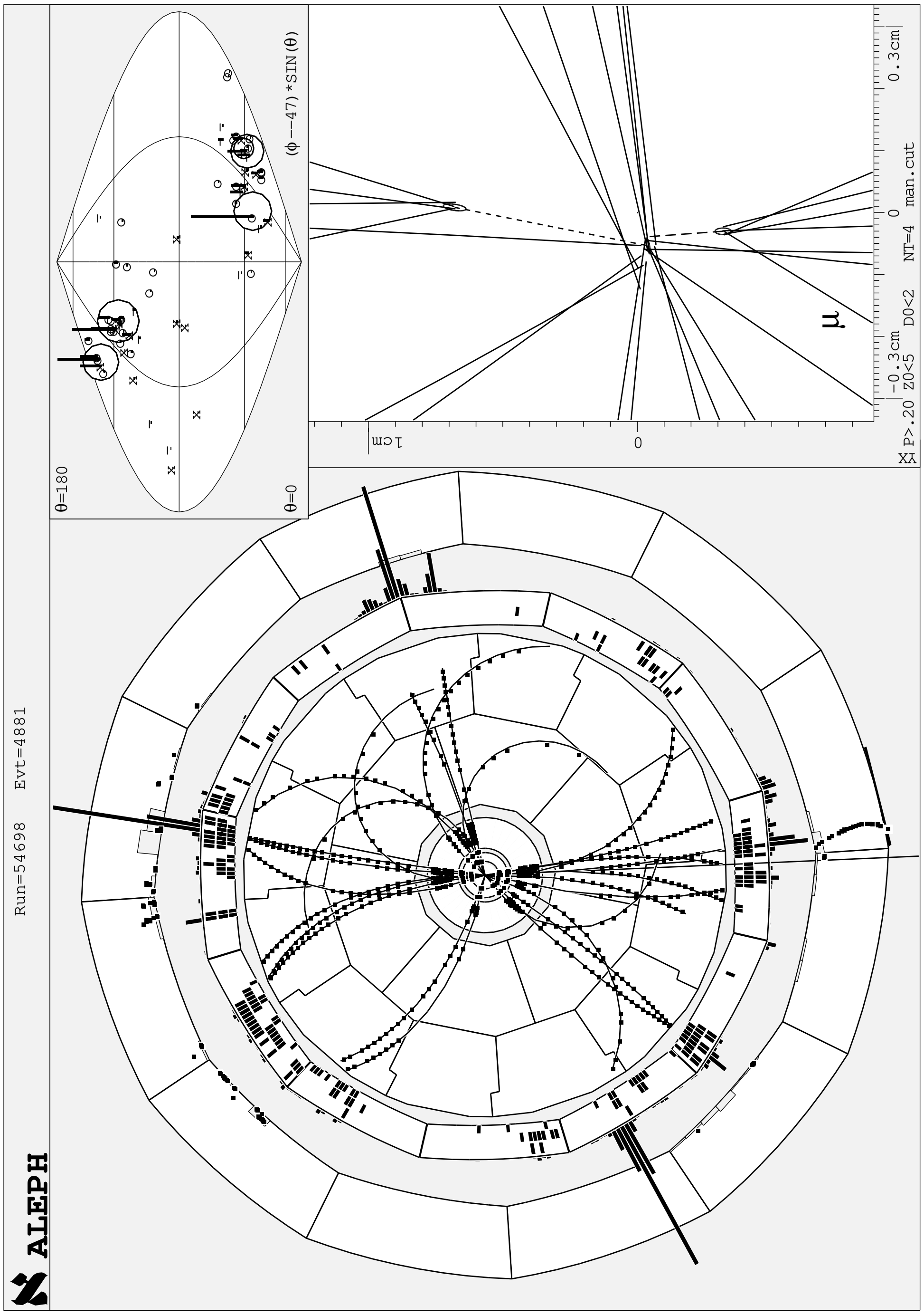


**Figure 6.7:** Event Display of high impact candidate Run=54698, Event=4881

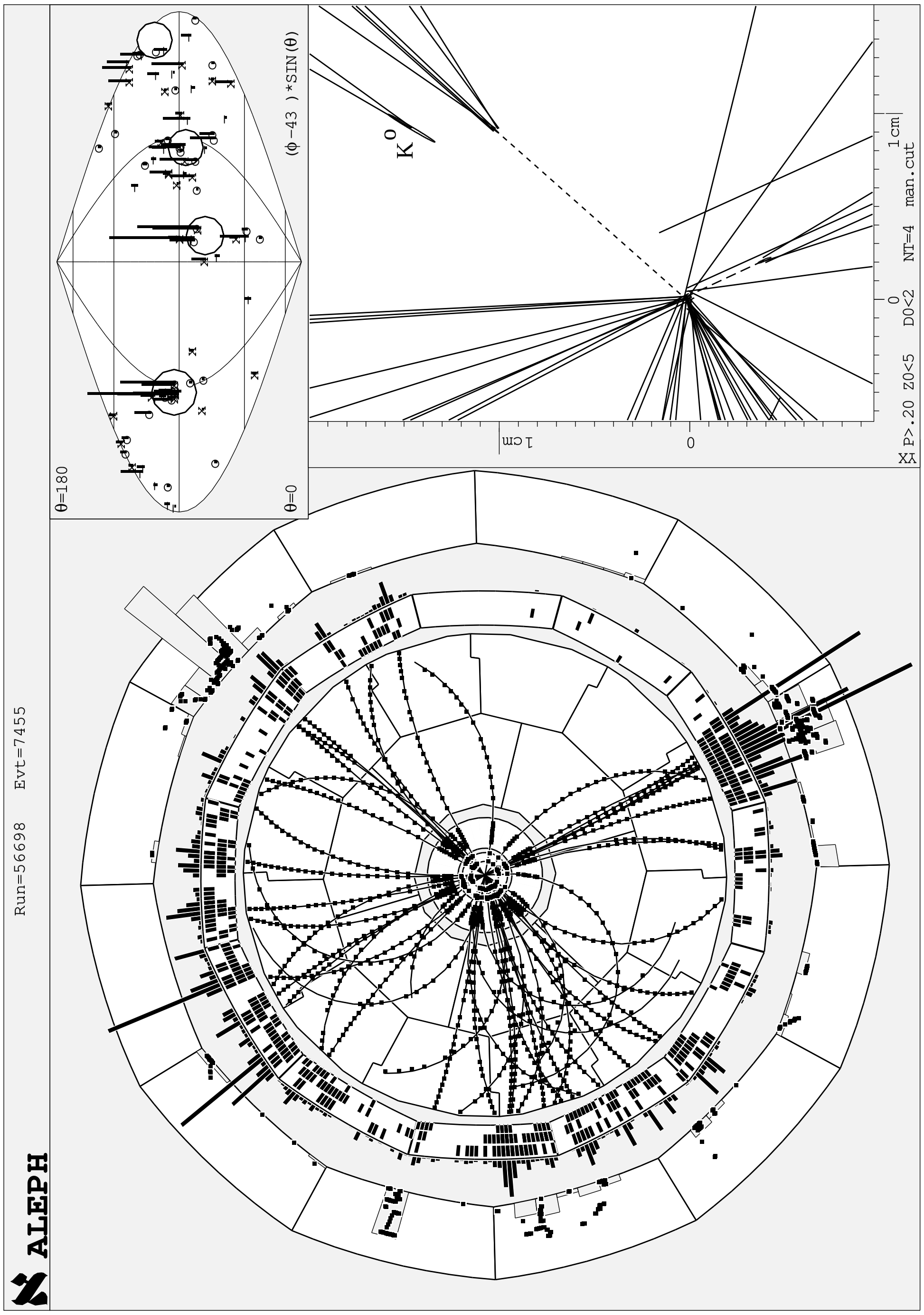


**Figure 6.8:** Event Display of high impact candidate Run=56698, Event=7455

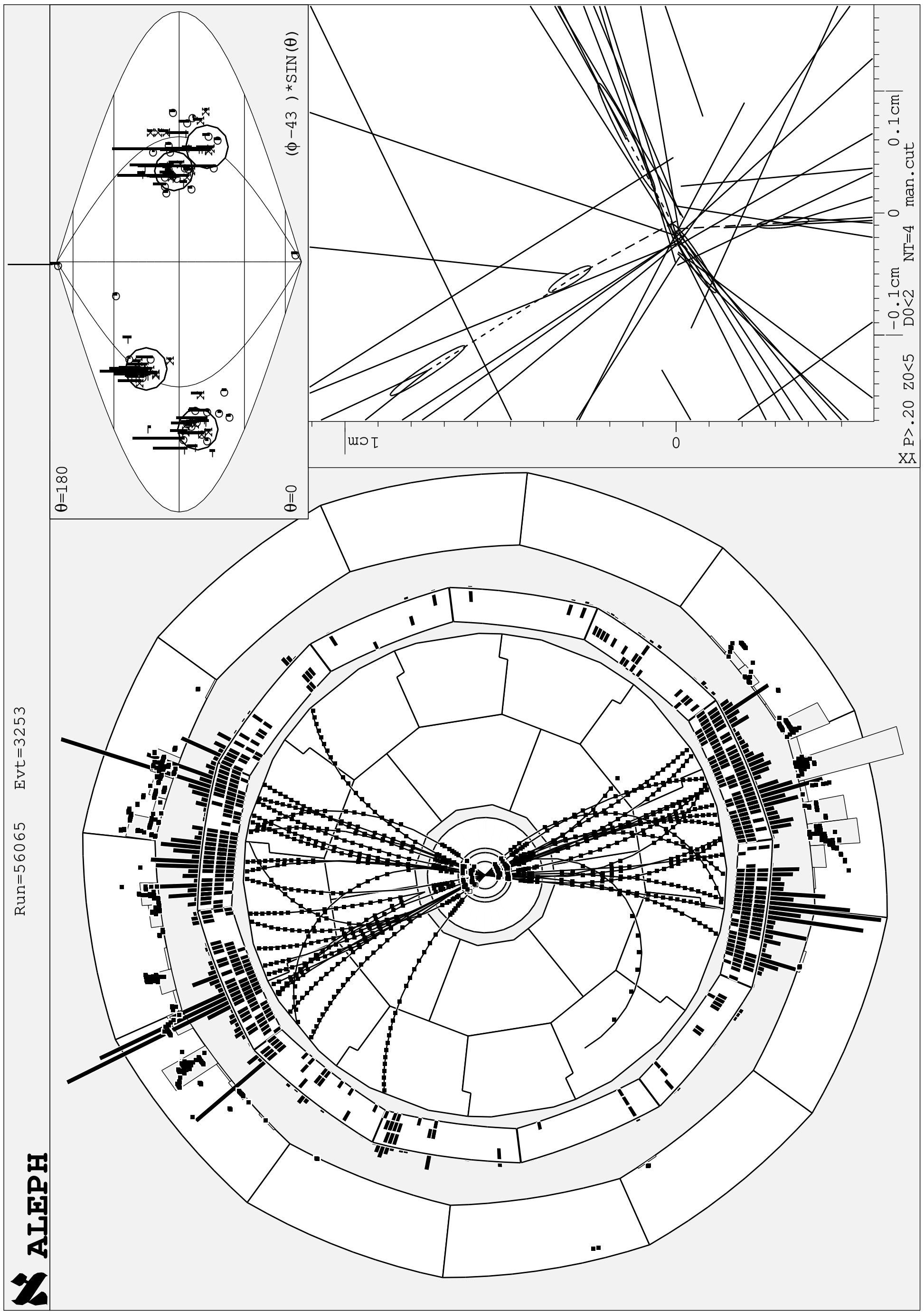


**Figure 6.9:** Event Display of high impact candidate Run=56065, Event=3253

# Chapter 7

# Beam related background

The nature of a collision at the LEP collider is such that an electron and positron from the crossing beams collide and annihilate or scatter by means of the exchange of a gauge boson.

In this process all the observed final state particles are assumed to originate from the electron positron interaction. However, the final state as recorded by the ALEPH detector is subject to physical backgrounds from extraneous sources such as cosmic rays and beam-induced particles.

Neither of these background sources are present in the MC simulation. This chapter investigates the effect of the beam induced backgrounds on the event selection efficiencies and event weighting pdf's of the 4 jets cuts-based analysis. A possible method for treating the beam induced background is also explored.

## 7.1 Random Trigger study

The beam induced background in LEP 2 is composed of synchrotron radiation photons and off-momentum beam particles. Synchrotron radiation photons can enter the detector either directly or after scattering from vacuum equipment on either side of the IP (Interaction Point) and have energies which range from several keV to many hundreds of keV. The off-momentum beam particles, often created by beam gas bremsstrahlung, reach the detector after being over focused in the strong low-beta quadrupoles. The off-momentum beam particles have a broad energy spectrum which is centred in the region of half of the beam energy [28].

To study the beam related backgrounds a sample of events in which no physical collision or scattering has occurred is needed. Such a data set exists in the form of random triggers (RT). Random trigger events are recorded when the ALEPH trigger is activated on a random bunch crossing even if no actual physics event has occurred. The low probability of a physics event occurring within any given beam crossing means that this sample accurately reflects the nature of the beam induced backgrounds, or any detector-related noise.

A sample of ≈217 $pb^{-1}$ of ALEPH random trigger data recorded during 2000 was analysed. This data set corresponds to 435272 random trigger events within a centre of mass energy range of 200 - 209 GeV. From this sample it is possible to establish the physical characteristics of the beam related backgrounds. The sample may then be used to determine the size of the effect of this unmodelled background on the 4 jets analysis and also to devise a method of treating events taking this background into account.

Figure 7.1 shows the energy distribution $E_{12}$, the energy within 12° of the beam axis, and the largest energy flow object within 12° of the beam axis. Both the $E_{12}$ and the largest energy-flow object distributions are strongly peaked towards zero. This shows that the outstanding majority of events have only a negligible amount of beam related noise present. Table 7.1 shows the fraction of random trigger events which contain an energy-flow object within 12° of the beam axis which is equal to or exceeds a given cut [64].

| Energy | Number of Events | Fraction (%) |
|---|---|---|
| All | 435272 | 100.00 |
| >3 GeV | 3855 | 0.886 |
| >5 GeV | 2685 | 0.617 |
| >10 GeV | 2086 | 0.479 |
| >20 GeV | 1557 | 0.358 |
| >40 GeV | 1024 | 0.235 |

**Table 7.1:** The number, and fraction, of RT events possessing an object which is above a specific energy cut.

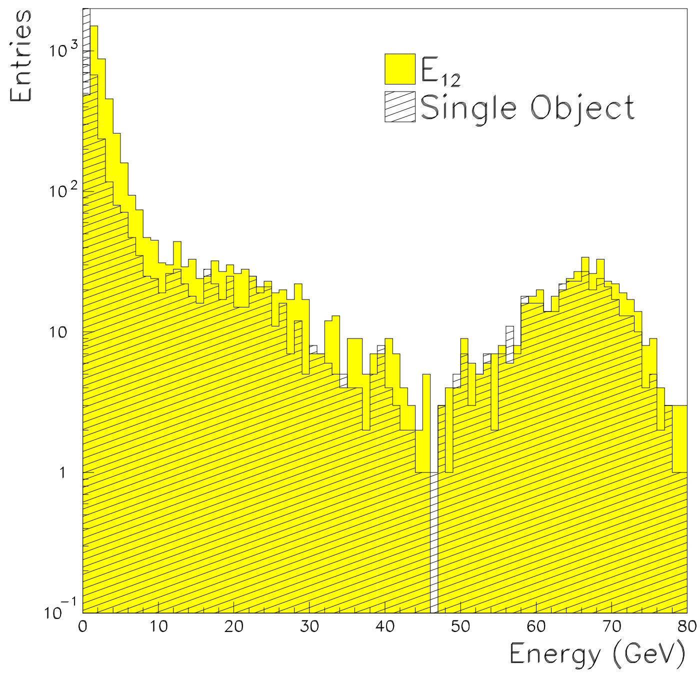


**Figure 7.1:** The energy distribution of $E_{12}$ and of the largest energy object within $E_{12}$ for the complete RT sample.

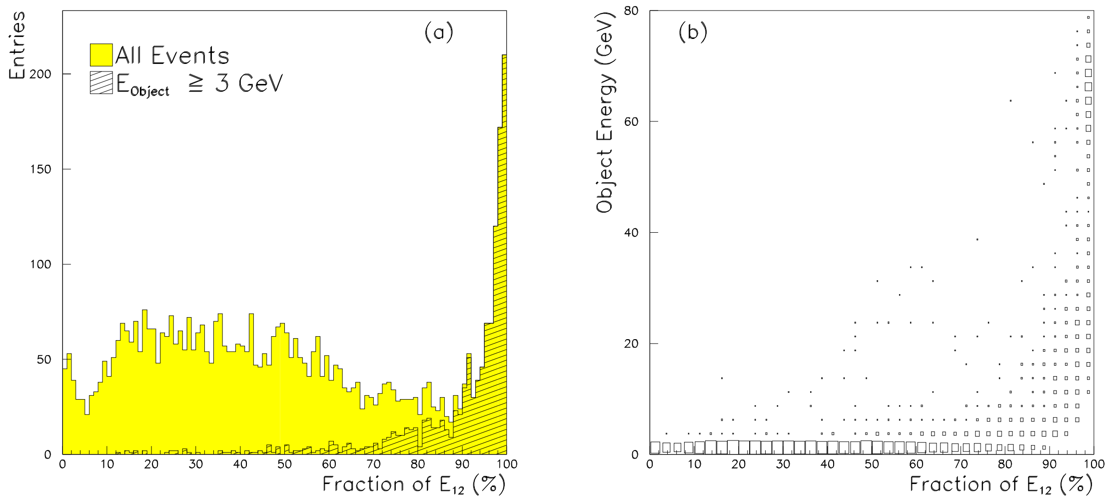


**Figure 7.2:** (a) The fraction of $E_{12}$ contained within the most energetic object for all events and for events with an object greater than or equal to 3 GeV. (b) The fraction of $E_{12}$ contained within the most energetic object versus the object energy.

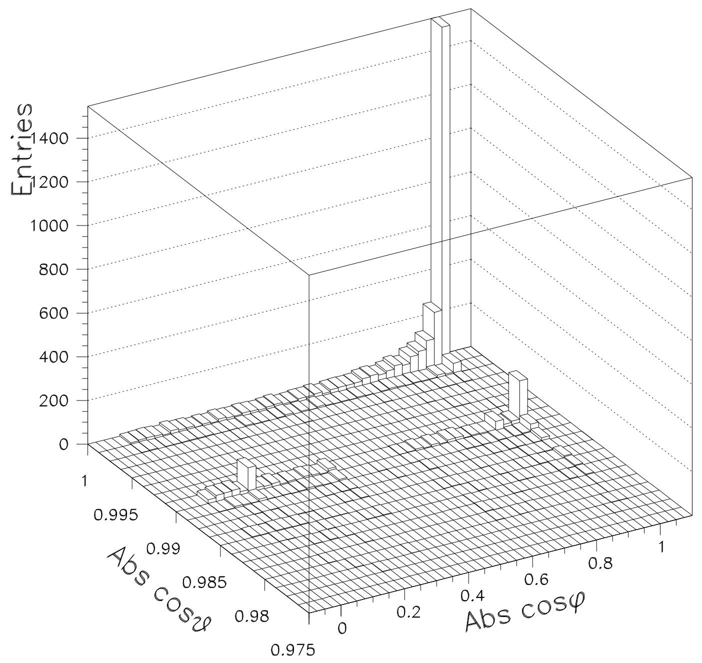


**Figure 7.3:** The Distribution of $\theta$ and $\phi$ for the largest energy flow object within $E_{12}$

## 7.2 Random Trigger overlay

Although the fraction of events containing a significant amount of beam induced noise is low, the true impact of this background on the 4 jets analysis may only be determined by correctly adding this background to the simulated physics events within the Monte Carlo. To simulate the effects of the beam related background a sub-sample of the random trigger events was superimposed onto the standard MC events. It is intuitive that low energy random trigger events are unlikely to have a significant impact on the analysis of the underlying MC events. Following from this only random trigger events with an amount of energy that exceed 3 GeV are used in the overlaying process.

A further approximation which aids the addition of the beam related backgrounds to the MC is the use of only the largest energy object within 12° of the beam axis. Figure 7.2a shows the fraction of $E_{12}$ contained within the most energetic track. The figure shows the fraction for all events and for a sub sample of events in which the object energy is greater than or equal to 3 GeV. It is clear that in events which have an object of 3 GeV or above, the majority of $E_{12}$ is contained within this object. This is further illustrated in figure 7.2b, which shows the fraction of $E_{12}$ contained within the highest energy object versus the energy of that object. The higher energy objects contain a large fraction of $E_{12}$. Following from the above arguments and considering that only high energy objects are likely to significantly alter the underlying physics events it was decided that a good approximation of the beam related background could be achieved by considering only the single objects with energy equal to or exceeding 3 GeV. The angular distribution of these object is shown in figure 7.3. The distribution is clearly forward peaked and favours the emission of a particle close to the beam pipe and in the horizontal plane.

A sample of 3855 objects containing information about the object 4-vectors and the object energy-flow type was created. These objects were cyclically added to the energy flow banks of the standard MC events with one object added per event.

The addition of these beam related objects to each of the MC events provides a good approximation of the actual beam induced backgrounds which are most likely to alter the underlying physics event. The addition of the object in this manner ensures that the energy and angular distribution characteristic of the background are retained. It is important to note that the objects in the overlaying sample represent only a small fraction of the actual beam background. The overwhelming majority of the beam induced background is of low energy and has been omitted from this sample. Therefore the effect from the addition of this sample will be an over exaggeration compared to that of real data given that it contains only the highest energy subset of the background. The effect of the addition of this sample must be re-weighted by considering the probability for the occurrence of one of these large energy object in order to provide an accurate simulation of real data.

Throughout this chapter the following MC samples will be used:

- Clean MC - The original MC sample, with no beam induced noise.

- Dirty MC - The MC sample with beam induced noise overlayed onto each event, thus making this sample a pure beam related noise sample
- Corrected MC - The correctly weighted addition of the Clean and Dirty MC to produce the conditions observed in data.

Although the overlaying sample size is quite small in comparison to the typical size of the MC onto which it is to be added, potential biases are reduced due to the following two factors. Firstly, as stated previously, the cyclic use of this sample ensures that the energy and angular distributions of the sample will be retained. Secondly, although the tracks will be used numerous times it is anticipated that the changing nature of the underlying physics events will reduce the effect of any potential bias. Although the object sample is finite in size it is reasonable to assume that the overlapping method provides an unbiased and practical method of adding the beam related background to the standard physics event simulation.

The random trigger objects were overlayed onto MC event samples which had centre of mass energies in the region of 206.7 GeV. The details of these MC samples are found in Table 7.2.

| MC Type | $\sqrt{s}$( GeV/$c^2$) | Sample size(k) |
|---|---|---|
| hZ($M_h$=114 GeV/$c^2$) | 206 | 7.5 |
| ZZ | 207 | 50 |
| WW | 207 | 250 |
| $q\bar{q}$ | 206 | 500 |

**Table 7.2:** The MC samples used in the RT overlay study.

The MC energies were chosen to provide a spread of samples around the centre of mass energy at which the majority of the 2000 data was collected.

Figures 7.4a-d show the distributions for various variables, a)$E_{12}$, b)$E_{vis}$, c)$C_{ang}$ and d)cos$\theta_{Iso}$ for the Clean versus Dirty MC samples[1]. The $E_{12}$ and $E_{vis}$ distributions clearly show that dirtying the MC with the RT events has lead to an increase

[1]The Clean and Dirty compared distributions are normalised to the same area.

in the event energy which is observed globally and in the region close to the beam pipe.

The $C_{ang}$ variable is the cosine of the angle between the beam background object and the total momentum of the event. The addition of the RT objects to the MC causes a large shift in the value of $C_{ang}$ towards the value of 1 which shows how the event momentum is following the added RT object.

The $\cos\theta_{Iso}$ variable defines the isolation of the largest energy object in $E_{12}$ from all other energy-flow deposits within the event. The distribution is clearly shifted when the RT objects are added thus highlighting their isolation from other particles from within the actual event.

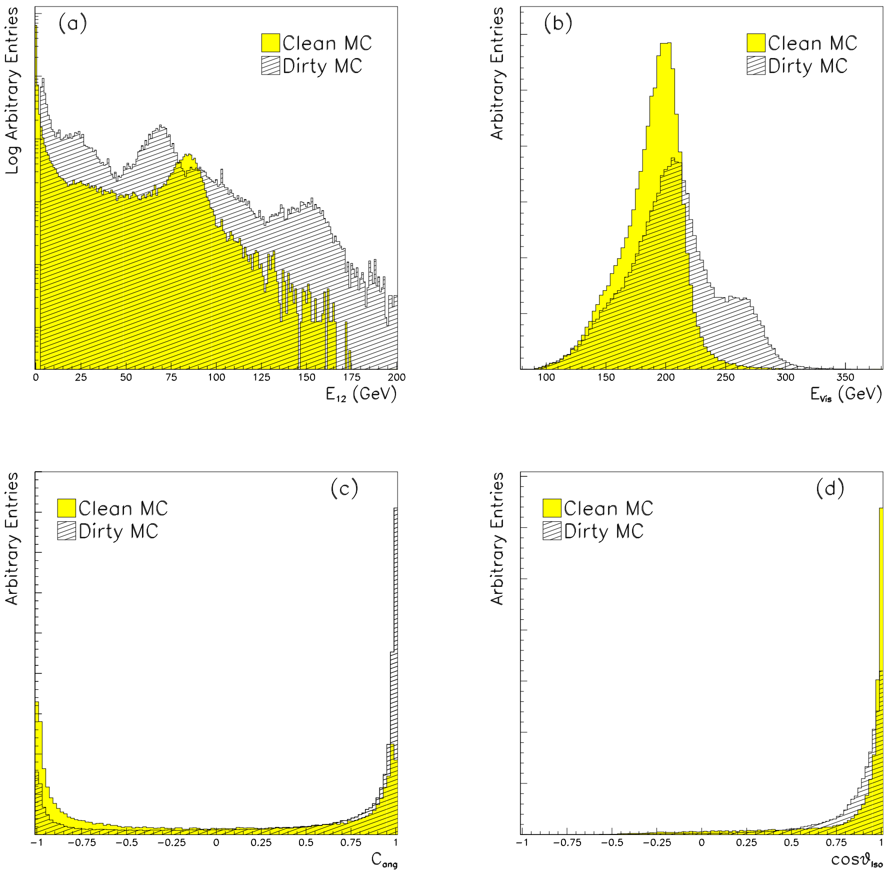


**Figure 7.4:** Distributions of various variables for Clean and Dirty MC.

## 7.3 Results of Random Trigger study

The addition of the beam induced background is expected to effect the four jets cuts analysis in two ways. Firstly the event selection efficiency may be altered. The standard Clean MC is used to determine the selection efficiencies for all the major background processes and from a possible Higgs boson signal. The expected number of events follows directly from this efficiency. The Dirty MC sample allows the efficiency to be calculated with this additional background included. Table 7.3 summarises the differences in both the signal and background efficiencies when the four jets analysis is applied to the corrected MC.

| MC Type | Pre-selection level (%) | Full selection hZ stand-alone (%) |
|---|---|---|
| hZ($m_{\mathrm{h}} = 114$ GeV$/c^2$) | -0.139 | -0.101 |
| ZZ | -0.150 | -0.161 |
| WW | -0.175 | +2.736 |
| $\mathrm{q\bar{q}}$ | -0.168 | -0.064 |
| Total background | -0.172 | +0.269 |

**Table 7.3:** The relative percentage change in the selection efficiency for the Corrected MC when compared to the Clean MC, at the pre-selection and full selection levels.

From table 7.3 it can be seen that the addition of the beam related noise produces a drop in the pre-selection efficiency which follows through to the selection level for all the samples except that of the WW background which has a marked increase in selection efficiency at the full selection level. To determine the origin of this increase the variables used within the 4 jets analysis were compared in the Dirty and Clean MC. The results of the comparison show that although the kinematic variables such as $\gamma$ and $\Theta$, $\mathrm{m}_{12}$, $\mathrm{y}_{34}$ show no significant change, (see figure 7.5) there is indeed a large shift in the b-tagging variables. Figure 7.6 shows the distributions of the b-tagging variables for the WW MC. The most stringent b-tagging variable used within the 2b branch, $(1-\eta_3)(1-\eta_4)$, is shown in figure 7.6. The maximum b-tag and minimum b-tag are also plotted. The $(1-\eta_3)(1-\eta_4)$ variable shows a significant increase in the signal like region for the Dirty MC. The $\sum \eta_i$ variable,used within the 4b branch of the analysis, however shows only a marginal rise in the signal like region. These increases in the b-tagging are consistent with one of the jets within the event showing an increased b-tag rather than a number of the jets.

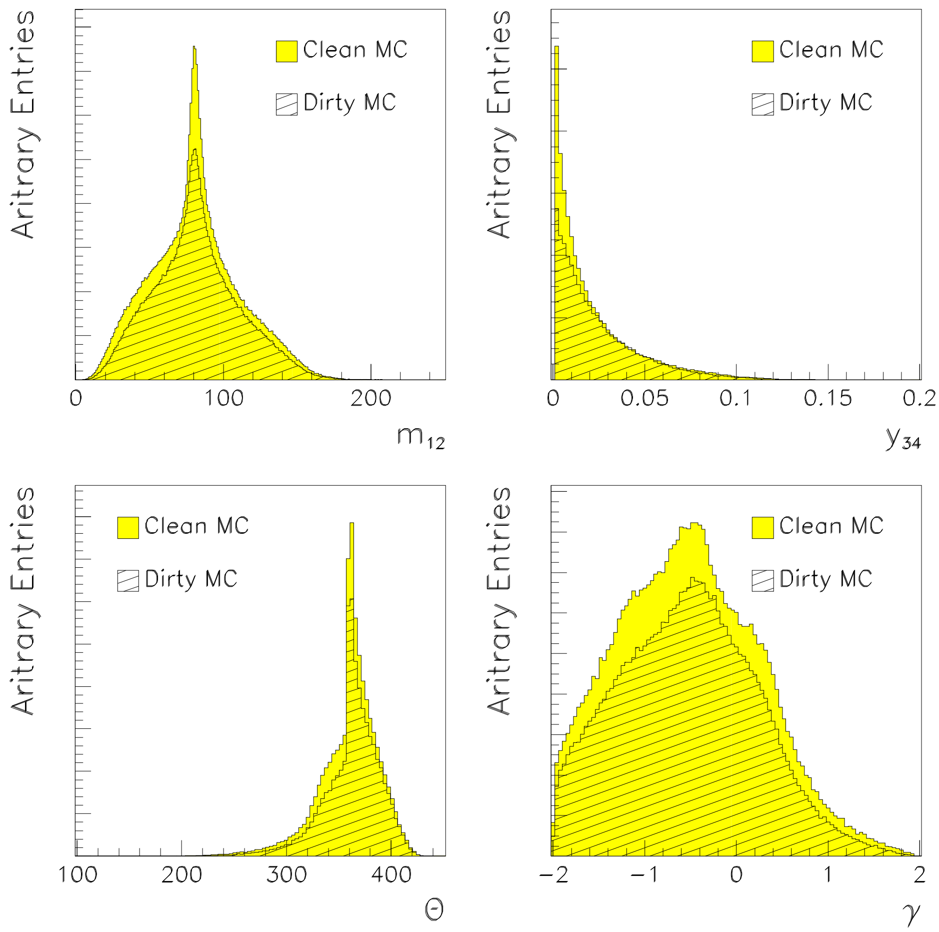


**Figure 7.5:** The distributions of $m_{12}$, $y_{34}$, $\Theta$ and $\gamma$, for the Clean and Dirty WW MC at the pre-selection level of the analysis

To determine the cause of the increase in the b-tag the six input variables to the NN b-tagger were compared both for Clean and Dirty MC. The distributions for the 6 NN input variables are shown in Figure 7.7. The figures show that the major impact of the RT overlay onto the WW MC is to produce a larger number of high $p_t$ leptons. This is evident in the lepton $p_t$ plot and also the $\sum p_t^2$ plot. This impact is consistent with the addition of a single new track which may have

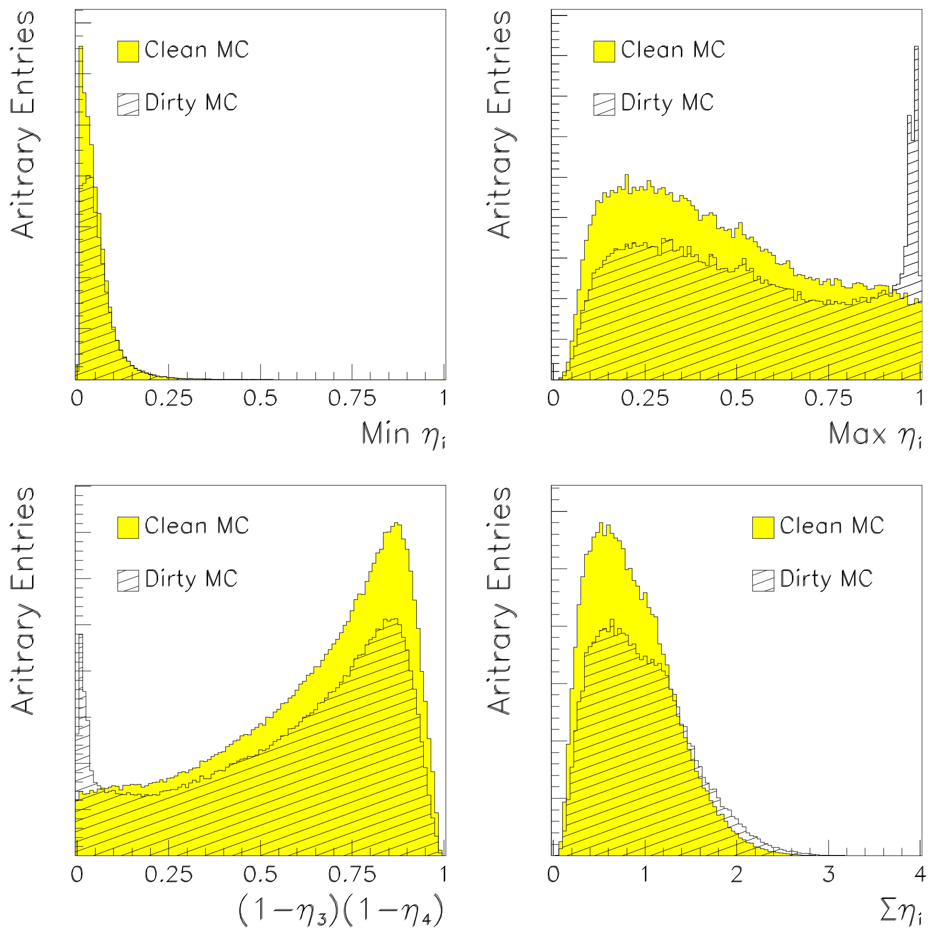


**Figure 7.6:** The distributions of various b-tagging variables for the Clean and Dirty WW MC at the pre-selection level of the analysis

high momentum. The other b-tagging variables are affected to a lesser extent. The increase in $p_t$ and $\sum p_t^2$ is consistent with an increased b-tagging output from the NN. The addition of a lepton candidate to the WW events may lead to the decaying c-quark from $W \rightarrow$ cs being interpreted as a decaying b-quark from b $\rightarrow W$c with $W \rightarrow \ell\nu$ thus leading to the same observed final state c$\ell\nu$.

Although a large increase in the selection numbers for the WW Monte Carlo is observed, the overall effect is somewhat reduced since at the final selection level this background is the least prominent. Indeed, when all backgrounds are considered with their correct weightings, it is found that the actual change in

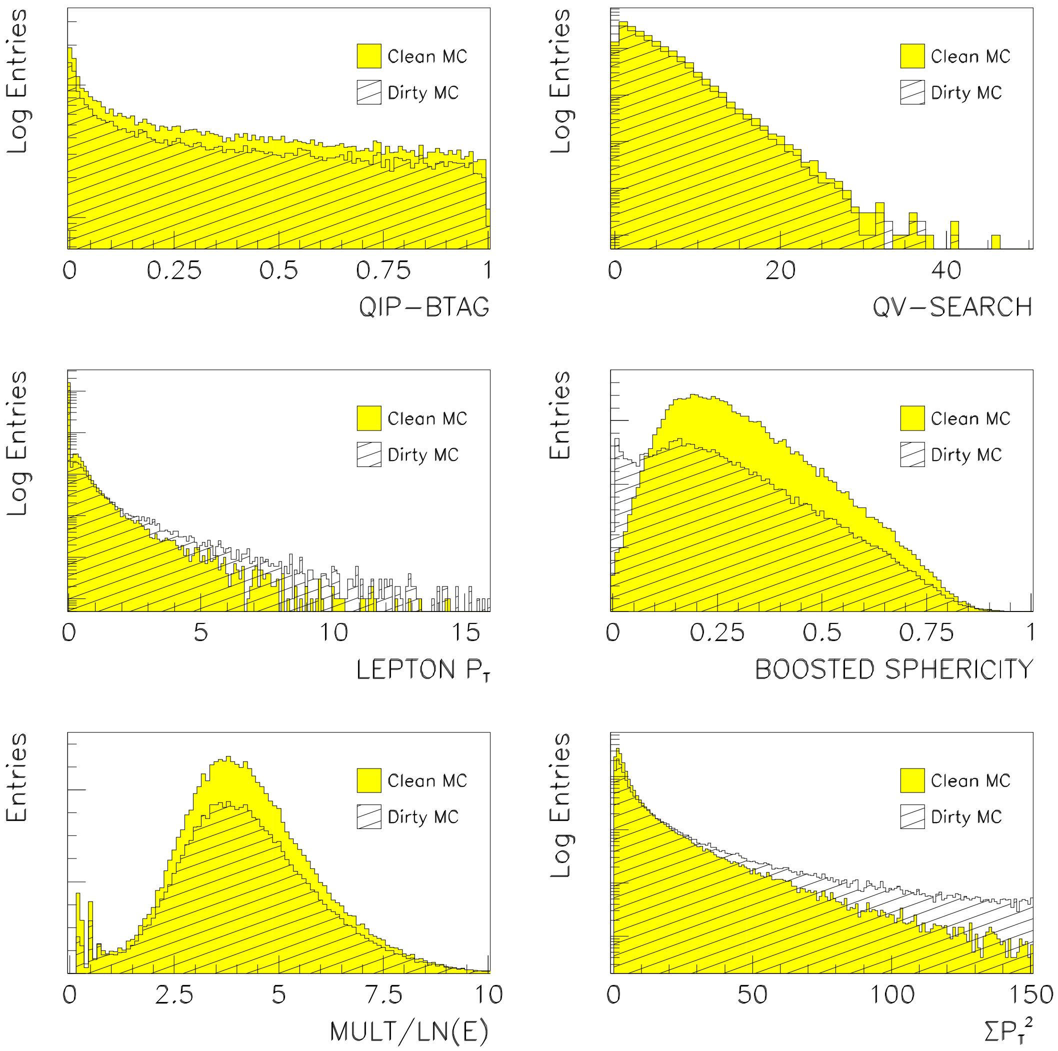


**Figure 7.7:** The distributions of the six variables used to for the 6VNN for both Clean and Dirty WW MC at the pre-selection level

selection efficiency for the combined background is +0.269% for the hZ stand-alone analysis as reported in table 7.3.

The second way in which the addition of the beam background may affect the cuts based analysis is by changing the distributions of the discriminating variable ($m_{\text{Reco}}$). The discriminating variable is used to assign a weight to the events when computing the likelihood ratio. The addition of beam related noise and as such extra energy, is expected to cause a shift in the reconstructed Higgs mass with a corresponding shift in the event weights. The pdf's from which the event weights are defined were constructed using the Clean MC. This means that the data events in which beam related background is present would be given an incorrect weight from these pdf's. The extent to which any bias in the discriminating variable is introduced when beam related noise is present in a given event must be determined. Figure 7.8 shows the comparison of Clean and Corrected MC shapes for the discriminating variable. The corrected distributions show imperceptible changes from the standard distributions as defined using the Clean MC.

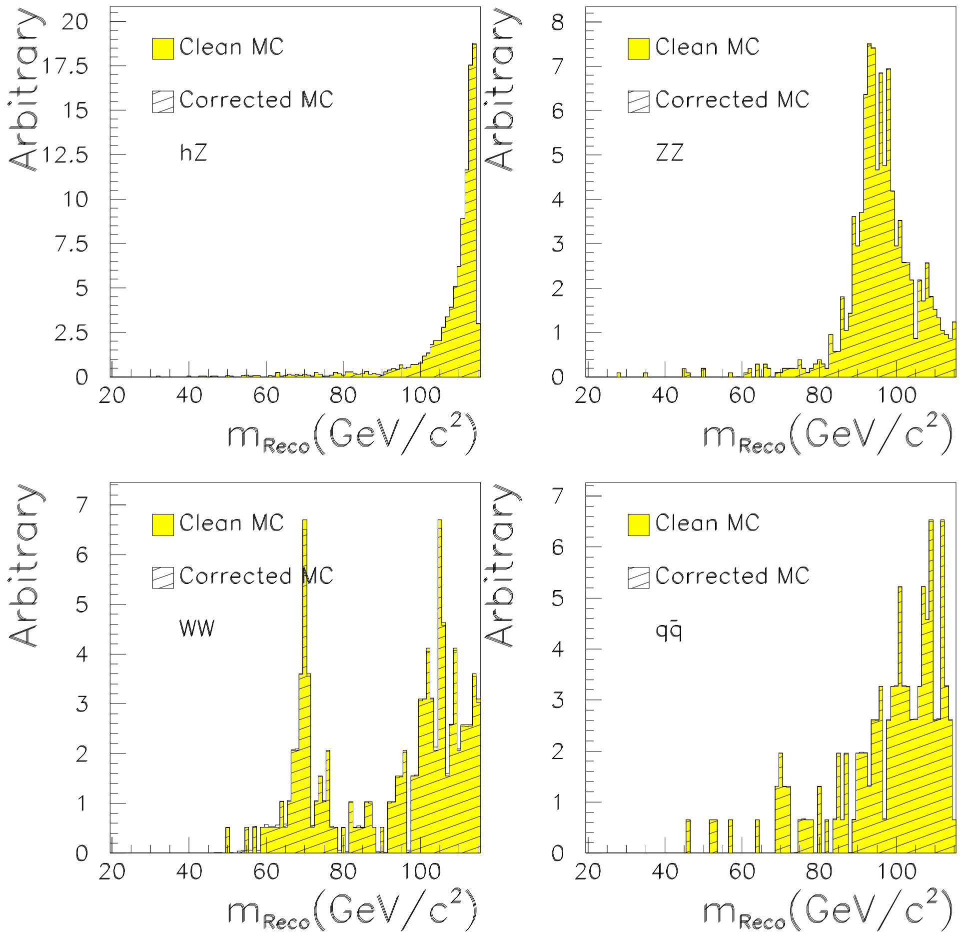


**Figure 7.8:** Comparisons of the reconstructed mass, $m_{\mathrm{Reco}}$, distributions in signal and background, for Clean MC and Corrected MC

One possible way to incorporate both the selection efficiency and changes in the discriminating variable into one quantity is to consider the affected number of events. The definition of an affected event is one which has either been lost/gained by the analysis or has undergone a change in $m_{\mathrm{Reco}}$ of 1 GeV/$c^2$ or greater. This definition allows the overall impact of the introduction of the beam background to be quantified. The number of affected events for the different backgrounds and the signal is summarised in table 7.4.

| MC Type | Original Selected number | Affected events Dirty MC | Fraction affected Corrected MC (%) |
|---|---|---|---|
| hZ($m_{\mathrm{h}} = 114$ GeV/$c^2$) | 2839 | 1800 | 0.56 |
| ZZ | 1015 | 923 | 0.81 |
| WW | 194 | 800 | 3.67 |
| q$\bar{\mathrm{q}}$ | 153 | 153 | 0.89 |
| Total background | - | - | 1.21 |

**Table 7.4:** The number of affected hZ stand-alone events when the Dirty MC is compared to the Clean MC. The correctly weighted fraction of events is also given.

Table 7.5 shows the number of selected events in common to both the Clean and Dirty MC final selections as a function of the mass shift. It is clear that the majority of the events have mass shifts of less than 2 GeV/$c^2$ but a non negligible number of events have larger shifts up to 10 and 20 GeV/$c^2$.

| Mass Shift GeV/$c^2$ | hZ events | ZZ events | WW events | q$\bar{\mathrm{q}}$ events |
|---|---|---|---|---|
| $> 1$ | 817 | 386 | 67 | 38 |
| $> 2$ | 355 | 138 | 28 | 22 |
| $> 5$ | 203 | 77 | 16 | 12 |
| $> 10$ | 117 | 35 | 10 | 6 |
| $> 15$ | 99 | 24 | 9 | 5 |
| $> 20$ | 87 | 16 | 9 | 3 |

**Table 7.5:** The number of hZ stand-alone candidate events common to both the Clean and Dirty MC final selections that have undergone a shift in the value of the reconstructed Higgs mass equal to or greater than the value indicated.

## 7.4 Event cleaning

In the previous section it was shown that the effects of the beam related background appear very small if not completely negligible. It must be stressed however that the effects detailed above are arrived at by considering the correctly weighted beam background fractions. The presence of a beam related object of 3 GeV+ within an event can cause an observable shift in the value of the discriminating variable for that event. Figure 7.9 shows the impact of the RT overlay on the 3 main backgrounds and a possible signal. In this figure the correct weighting of the beam background, due to its low frequency, has not been accounted for. The figure thus highlights the possible effect of the beam related background on a single event, such as an individual data candidate. The small number of expected data events, $\approx$ 20-30, means that the probability of any single event possessing a large beam related object is very low. This low probability is reflected in the small shifts of the discriminating variable and selection efficiency when the Corrected MC is considered. This small shift, although arguably correct for defining the shapes, does not reflect the large shifts that may occur in an individual event. Although the probability is low the possibility that a data candidate could contain a large amount of beam related noise and so undergo a large shift in $m_{\mathrm{Reco}}$ cannot be discounted. Such a large shift could cause the candidate weight to be considerably under/over estimated even if the corrected pdf shapes are used.

The possibility that a data candidate could be wrongly treated shows that it is not an adequate solution to simply re-calculated the pdf shapes and efficiencies using Corrected MC. Following this it is obvious that a method of treating data candidates, in which this background arises, must be devised.

Methods which reject events based on $E_{12}$ cuts or that remove all objects within $12^{\circ}$ of the beam axis are rejected due to the possible effect on the event selection efficiency and degradation of the mass resolution.

An ideal solution would be to retain the current selection efficiency whilst also ensuring that the beam background can be treated without any degradation of the mass resolution. The proposed method is to produce an event cleaning algorithm. As shown previously in figure 7.1 the majority of the beam related background

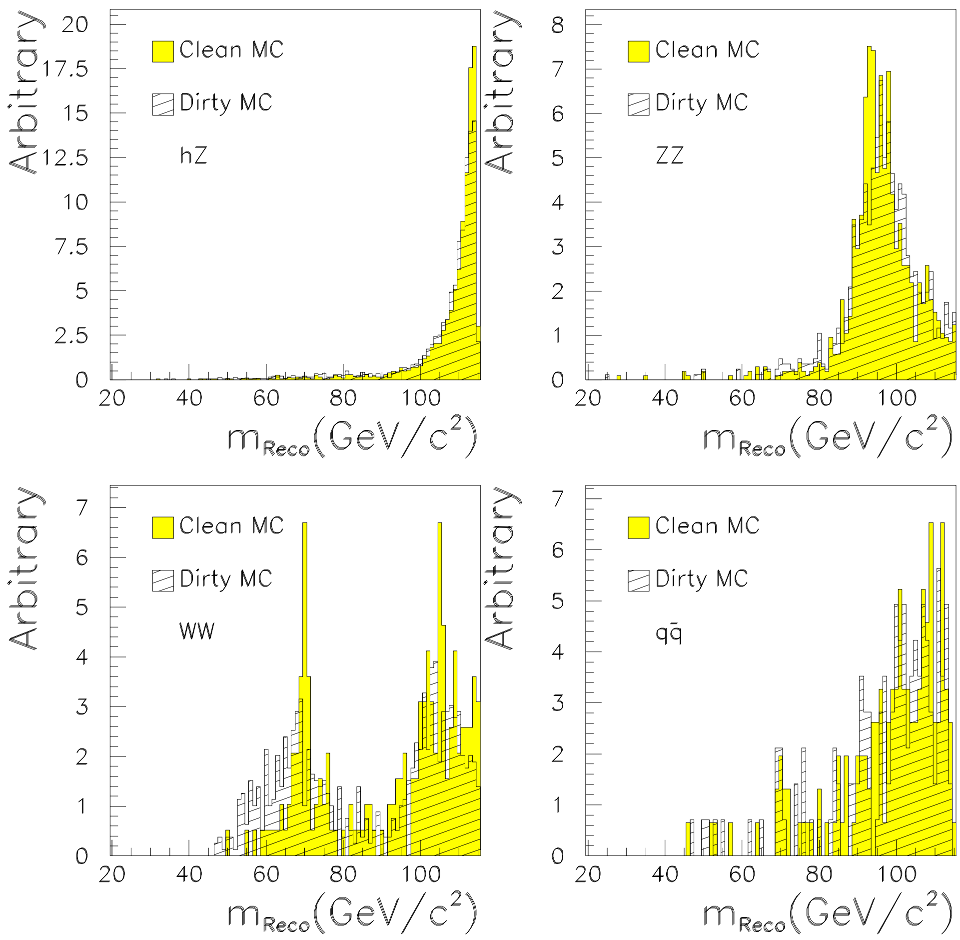


**Figure 7.9:** Comparisons of the pdf shapes for the discriminating variable, $m_{\text{Reco}}$, in signal and background cases for Clean MC and Dirty MC

associated with an event is contained within one energetic object at very low angle to the beam axis. The method of event cleaning is intended to select this beam related object and remove it from the event. The cleaned event may then be treated as a normal event. This method has the advantage that events which are correctly treated will be returned to a standard physics event as modelled in the MC. Such events may then be passed through the analysis as normal. The method also ensures that the individual data candidates are correctly treated.

## 7.5 Event cleaning algorithm

The cleaning algorithm has been developed to work in two stages. The first stage is to find possible candidate energy deposits which are consistent with the beam induced background. Information from the study of the random trigger events and from their overlay onto the physics MC is used to define the characteristics of the beam related background.

A modified version of an algorithm used in the ALEPH hadronic WW analysis [72] to find Initial State Radiation (ISR) photons is used to find these candidate objects. The algorithm identifies possible beam background candidate deposits which satisfy the following cuts

- $|\cos\theta| \geq 0.998$
- $E > 3$ GeV
- $\cos\theta_{Iso} > 0.99$

where $\theta_{Iso}$ is the isolation of the object from its nearest neighbour.

The highest energy object within 12° of the beam axis which satisfies the above criteria is taken as the beam related object.

Once an object has been selected it may be tested for its compatibility with the beam background hypothesis. To test if the hypothesised object is indeed beam related, and as such no part of the underlying event, a method which utilises the 4C kinematic fit [73] [74] has been derived.

Once a possible beam related object has been identified the event is twice clustered into 4 jets. Firstly with all tracks and then a second time with the hypothesised beam background candidate object omitted from the clustering. The two sets of fitted jets and the hypothesised object may then be used to form three possible event hypothesis to which the 4C kinematic fit is applied. The three hypotheses are:

1. The candidate object is part of the underlying event and should be included in that event. For this the original fitted four jets are used in which the beam related object is included.

2. The hypothesised object corresponds to ISR from one of the incoming $e^+/e^-$ pair. In this case the 4C fit is modified to provide for the ISR hypothesis by using the second clustering of jets and the object as the ISR photon.

3. The candidate object is beam related and is not part of the underlying event. In this case the second set of clustered jets is used for the 4C fit and the object is omitted from the fit entirely.

When each of the 3 event hypothesis are fitted a $\chi^2$ value is returned which quantifies the goodness of the fit. The values of these $\chi^2$'s may then be compared to see which hypothesis is most compatible with the observed event. One advantage of using this fit based algorithm rather than defining cuts on the kinematic variables within the event is that the kinematic fit should fold all of the information within the event into one variable.

A quantity call the $\chi^2_{Ratio}$ is defined to determine if the candidate beam background object is truly beam related or simply part of the actual physics event. The $\chi^2_{Ratio}$ is defined as:

$$\chi^2_{Ratio} = \frac{Min(\chi^2_{Norm}, \chi^2_{ISR})}{\chi^2_{BB}} \tag{7.1}$$

where $\chi^2_{Norm}$, $\chi^2_{ISR}$ and$\chi^2_{BB}$ are the 4C kinematic fit $\chi^2$'s for hypotheses 1,2 and 3 respectively. Events in which no beam background candidate object is found are defaulted to $\chi^2_{Ratio} = 0$.

With this definition it is possible to place a cut at a set $\chi^2_{Ratio}$ value which selects the beam related objects and yet leaves the standard physics events unaltered. Higher values of this ratio are more compatible with the beam background hypothesis. Figure 7.10 shows the distributions of $\chi^2_{Ratio}$ for background events in both Clean and Corrected MC. The figure is from the pre-selection level of the analysis. Any event with a $\chi^2_{Ratio}$ exceeding 2 will be treated as possessing a beam background object and the hypothesised object will be removed from the event.

## 7.6 Cleaning algorithm application

When determining the effect of the cleaning algorithm it is important to consider its application in two specific instances. Firstly it is important that the cleaning

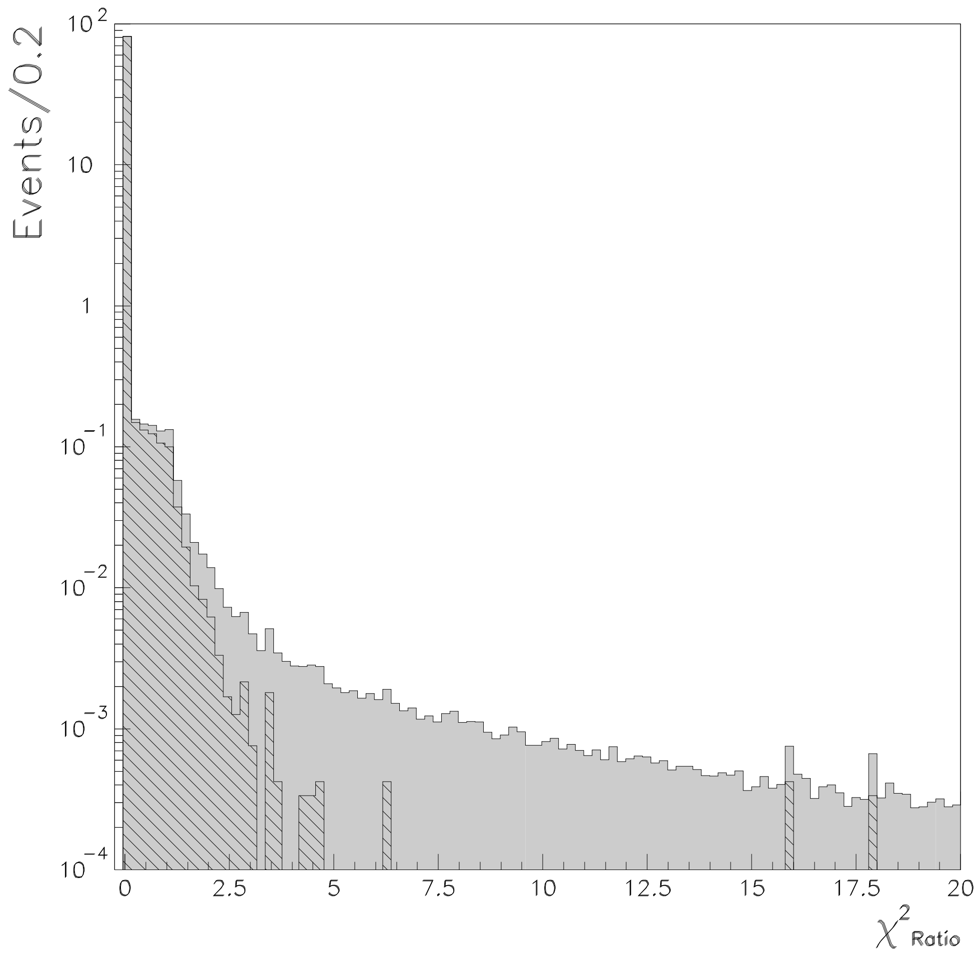


**Figure 7.10:** The distribution of the $\chi^2_{Ratio}$ for the clean (hashed) and corrected (solid) MC at the pre-selection level of the analysis.

algorithm does not degrade the performance of the analysis. It must also be assured that the cleaning algorithm accurately treats the events in which any beam related background is present. To test the cleaning algorithm it is applied to both the Clean MC and Dirty MC. When the cleaning algorithm is applied to the Clean MC it is found that only one simulated event is affected, where the definition of an affected event is as described previously.

The $e^+e^- \rightarrow q\bar{q}$ event which is affected is lost from the hZ stand-alone branch of the analysis. The details of this event both before and after the application of the cleaning algorithm are shown in table 7.6

| Variable | Before cleaning | After cleaning |
|---|---|---|
| $E_{vis}$ | 196.51 | 190.98 |
| $y_{34}$ | 0.0084 | 0.00697 |
| $\Theta$ | 360.27 | 360.21 |
| $\gamma$ | -1.930 | -1.977 |
| $m_{12}$ | 135.58 | 125.56 |
| $m_{34}$ | 64.44 | 75.67 |
| $m_{Reco}$ | 108.82 | 110.03 |
| $\min(\eta_3, \eta_4)$ | 0.99 | 0.99 |
| $(1-\eta_3)(1-\eta_4)$ | 9x$10^{-6}$ | 2x$10^{-6}$ |
| $\sum \eta_i$ | 2.145 | 2.150 |
| $X_\gamma$ | 0.19 | 0.21 |

**Table 7.6:** The details of the $q\bar{q}$ candidate which is rejected by the beam background cleaning algorithm both before and after the application of the algorithm.

The event fails the hZ selection criteria when the cleaning algorithm is applied because a 5 GeV object is found within the event and subsequently removed. The removal of this object from the event causes the $y_{34}$ value for the event to drop from 0.0084 to 0.00738. The $y_{34}$ cut for the hZ analysis is set at 0.008, therefore this event narrowly passes the selection before the object is removed and narrowly fails once the object is removed.

Therefore the effect of the cleaning algorithm on the Clean MC is negligible. When the Dirty MC is considered with the cleaning algorithm applied a much more substantial effect is observed, as expected.

Tables 7.7 and 7.8 present the number of affected events and common events with mass shifts when the Clean and Dirty MC samples are compared after the cleaning algorithm has been applied to the Dirty MC. These tables may be directly compared to tables 7.4 and 7.5, which present the same two quantities before the cleaning algorithm is applied. The affected number of events and the number of events experiencing a mass shift both fall once the cleaning algorithm is applied.

To further illustrate the impact of the cleaning algorithm table 7.9 shows the number of events which have been correctly treated by the cleaning algorithm and have been returned to their original mass.

Tables 7.7-7.9 show that the cleaning algorithm has a marked effect on the

Dirty MC and a sizable fraction of the contaminated events are indeed corrected by the algorithm.

The performance of the cleaning algorithm is also inferred by considering the number of candidates events in which the introduced RT object is correctly identified. Tables 7.10 and 7.11 show the number of introduced RT objects compared to the number of objects selected by the cleaning algorithm. The tables give the selection efficiency and the purity of the selected events for all RT object above 3 GeV and 10 GeV respectively. In both cases it is found that the selection purity is very high thus indicating the high degree of accuracy in the selection of a beam related object. The selection efficiency for objects with energy above 3 GeV is ≈28% compared to ≈50% for objects with energy above 10 GeV. This indicates that the cleaning algorithm is better at identifying beam related objects which are relatively high in energy.

This is displayed in figure 7.13 which shows the distribution of the introduced and selected RT objects. It is clear that the cleaning algorithm is better at identifying the higher energy RT objects. The loss in selection efficiency for low energy RT objects is also clearly visible in the region <10 GeV.

The difference between the $m_{\mathrm{Reco}}$ distributions for the Clean MC and the Corrected MC when the cleaning algorithm is applied are shown in figure 7.11 while the relative changes in the selection efficiency are shown in table 7.12. These two considerations show that there is a negligible shift in both the pdf's and event selection efficiencies and these need not be re-calculated.

The purpose of the cleaning algorithm is to return the Dirty MC to the original state as Clean MC. The demonstration that the algorithm does indeed perform this task may be seen when comparing figures 7.9 and 7.12 which show the comparisons of the Clean and Dirty MC $m_{\mathrm{Reco}}$ distributions before and after the application of the cleaning algorithm respectively. The figures show a marked shift towards the original Clean MC distributions once the cleaning algorithm is applied.

The cleaning algorithm has essentially no effect on events with no beam background and yet it is effective at cleaning a sizeable fraction of events in which the beam induced background is present.

| MC Type | Original Selected number | Affected events Dirty MC | Fraction affected Corrected MC (%) |
|---|---|---|---|
| hZ($m_{\mathrm{h}}$ = 114 GeV/$c^2$) | 2839 | 600 | 0.19 |
| ZZ | 1015 | 450 | 0.39 |
| WW | 194 | 152 | 0.69 |
| $\mathrm{q\bar{q}}$ | 153 | 87 | 0.50 |
| Total background | - | - | 0.46 |

**Table 7.7:** The number of affected hZ stand-alone events when the Dirty MC is compared to the Clean MC after the cleaning algorithm has been applied. The correctly weighted fraction of events is also given.

| Mass shift ( GeV/$c^2$) | hZ events | ZZ events | WW events | $\mathrm{q\bar{q}}$ events |
|---|---|---|---|---|
| >1 | 327 | 254 | 35 | 25 |
| >3 | 117 | 71 | 8 | 11 |
| >5 | 76 | 36 | 5 | 6 |
| >10 | 46 | 18 | 3 | 5 |
| >15 | 40 | 10 | 3 | 5 |
| >20 | 34 | 6 | 3 | 3 |

**Table 7.8:** The number of hZ stand-alone candidate events common to both the Clean and Dirty MC final selections that have undergone a shift in the value of the reconstructed Higgs mass equal to or greater than the value indicated after the cleaning has been applied.

| MC Type | Original Selected number | Corrected events Dirty MC |
|---|---|---|
| hZ($m_{\mathrm{h}}$ = 114 GeV/$c^2$) | 2839 | 1344 |
| ZZ | 1015 | 431 |
| WW | 194 | 81 |
| $\mathrm{q\bar{q}}$ | 153 | 59 |

**Table 7.9:** The number of events in the Dirty MC which have been returned to their original reconstructed Higgs mass and are selected in common with the hZ exclusive branch of the analysis when the cleaning algorithm has been applied.

## 7.7 Application of cleaning algorithm to 2000 data

The final stage in the application of the cleaning algorithm is to consider its impact on the 2000 data set. Figure 7.14 shows the distribution of the $\chi^2_{Ratio}$ for both Clean and Corrected MC, and the 2000 data in the hZ stand-alone branch. The

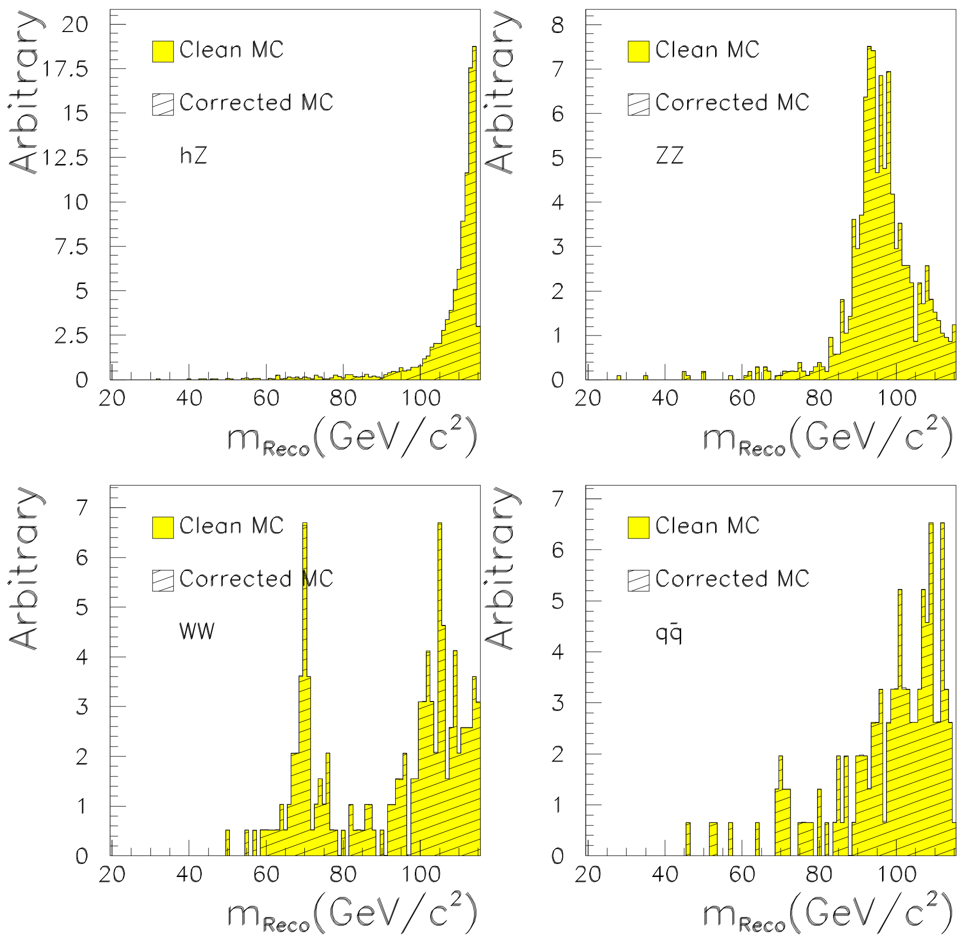


**Figure 7.11:** Comparisons of the $m_{\mathrm{Reco}}$ distributions, which form the discriminating variable, in signal and background cases for Clean MC and Corrected MC with the cleaning algorithm applied.

Clean MC is clearly peaked towards zero whilst the Corrected MC has a broad tail which extends out to higher values of $\chi^2_{Ratio}$. The data points, 33 in total, are concentrated in the lowest $\chi^2_{Ratio}$ bin with 32 of the candidates having no candidate beam background objects identified. This shows that, as expected, the majority of events are indeed clean and have no beam background present.

The remaining data candidate has a $\chi^2_{Ratio}$ value of 5.061 and is thus in a region which is consistent with the hypothesis that it contains a beam induced background object. Details of this candidate and the beam background object therein can be seen in table 7.13. The event possesses many properties which are con-

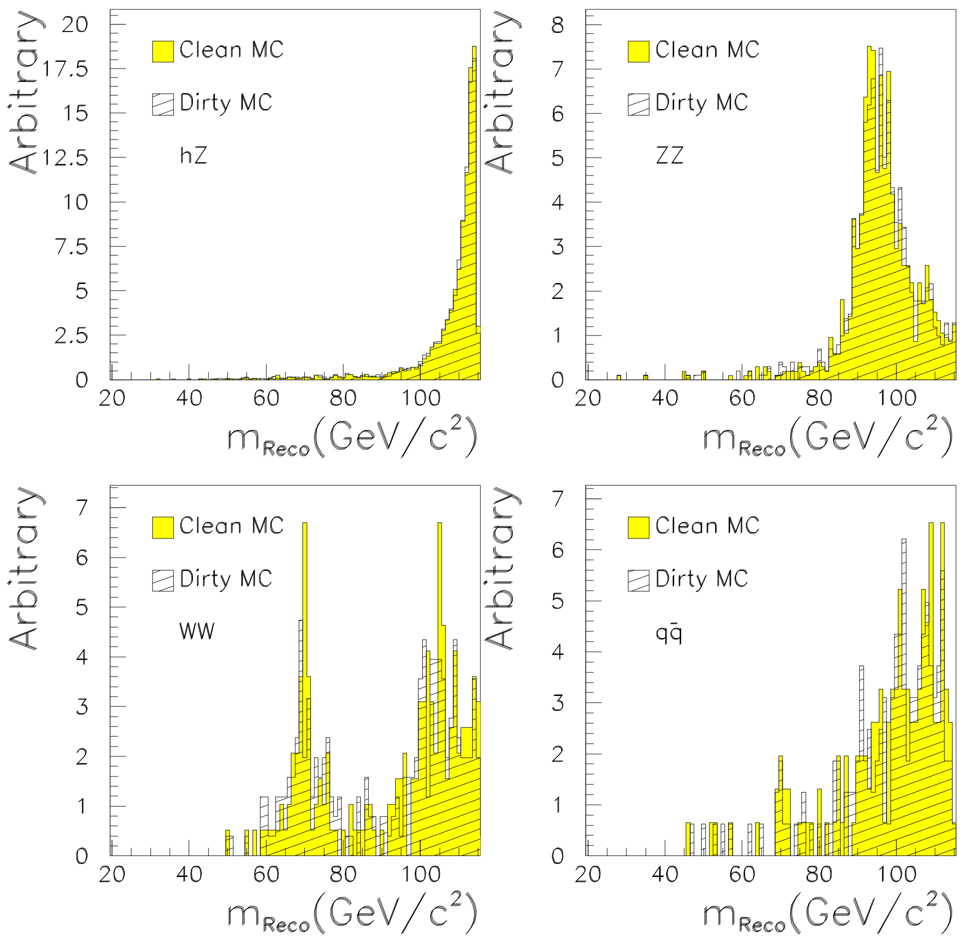


**Figure 7.12:** Comparisons of the $m_{\text{Reco}}$ distributions, which form the discriminating variable, in signal and background cases for Clean MC and Dirty MC with the cleaning algorithm applied.

sistent with events that have beam background present. The candidate beam background object lies within the angular region which is consistent with beam background. The object is also well isolated as is expected for a beam background object. The visible energy of the event (253 GeV) is also very high.

The details of this event, both before and after cleaning, as it is treated by the 4 jets cuts analysis are shown in table 7.14. The major impact of removing the object from the event is a change in the di-jet masses. The angular variables such as $\Theta$ and $\gamma$ show only a small deviation from their original values. Two direct impacts of the change in the di-jet masses are a change in the value of

| MC Type | Number Introduced | Number Selected | Correctly Selected | Selection Efficiency (%) | Selection Purity (%) |
|---|---|---|---|---|---|
| hZ | 7489 | 3211 | 3211 | 42.88 | 100 |
| ZZ | 29666 | 9357 | 9356 | 31.54 | 99.99 |
| WW | 206588 | 55384 | 55366 | 26.80 | 99.97 |
| $q\bar{q}$ | 202568 | 56695 | 56698 | 27.97 | 99.95 |
| Tot bgd | - | - | - | 27.75 | 99.96 |

**Table 7.10:** The performance of the beam background algorithm when selecting candidate beam background tracks for objects above 3 GeV. The number of introduced RT objects, selected events and correctly selected events is used to determine the selection efficiency and purity

| MC Type | Number Introduced | Number Selected | Correctly Selected | Selection Efficiency (%) | Selection Purity (%) |
|---|---|---|---|---|---|
| hZ | 3387 | 2935 | 2935 | 86.65 | 100 |
| ZZ | 13366 | 8703 | 8702 | 65.11 | 99.99 |
| WW | 98463 | 52150 | 52142 | 52.96 | 99.98 |
| $q\bar{q}$ | 107023 | 51886 | 51867 | 48.46 | 99.96 |
| Tot bgd | - | - | - | 50.21 | 99.97 |

**Table 7.11:** The performance of the beam background algorithm when selecting candidate beam background tracks for objects above 10 GeV. The number of introduced RT objects, selected events and correctly selected events is used to determine the selection efficiency and purity

| MC Type | hZ-SA |
|---|---|
| hZ | -0.002 |
| ZZ | -0.029 |
| WW | +0.269 |
| $q\bar{q}$ | +0.052 |
| Tot bgd | +0.032 |

**Table 7.12:** The relative % change in the selection efficiency for the Corrected MC when treated with the cleaning algorithm compared to the Clean MC

the discriminating variable and the transposition of the pairing choice within the event. The reconstructed Higgs mass changes from 112.8 GeV/$c^2$ to 114.4 GeV/$c^2$ when the beam related object is removed by the event cleaning procedure. The di-jet pairing is transposed such that the di-jets which are allocated to the h and Z bosons are swapped. The nature of the discriminating variable within the cuts analysis is such that this change of the di-jet pairings has no effect on the event as

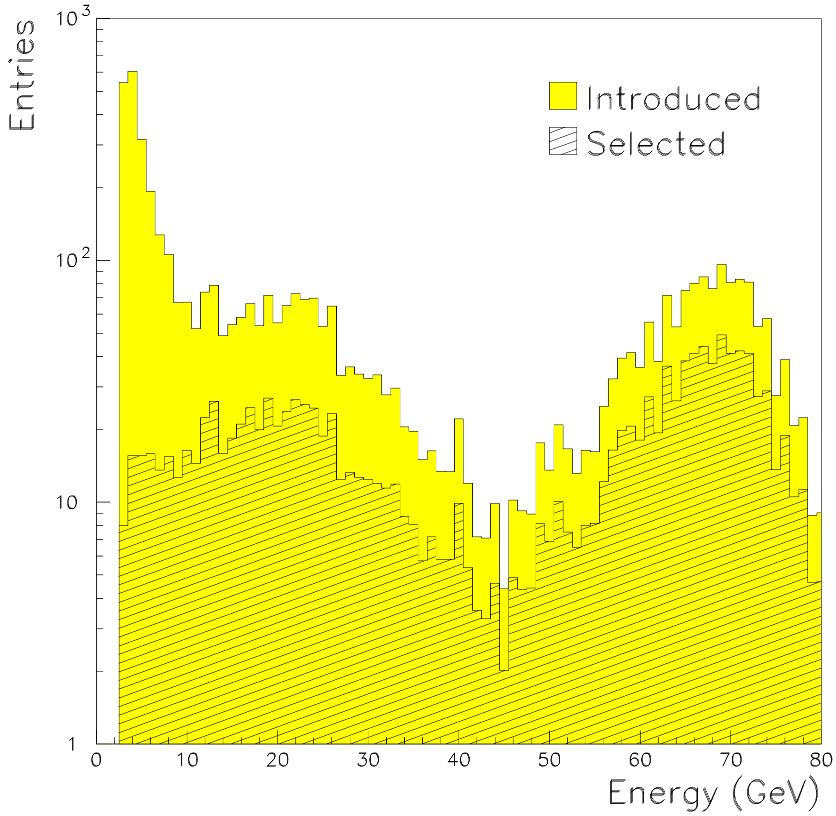


**Figure 7.13:** The energy distributions of the Introduced and Selected RT objects.

| Variable | Value |
|---|---|
| Run number | 56065 |
| Evt number | 3253 |
| QELEP | 206.65 GeV |
| $E_{vis}$ | 253.71 GeV |
| $\cos\theta$ | -0.999 |
| $\cos\phi$ | 0.965 |
| $E_{12}$ | 22.98 GeV |
| Object Energy | 21.93 GeV |
| $\cos\theta_{Iso}$ | 0.737 |
| $C_{ang}$ | 0.911 |
| $\chi^2_{Ratio}$ | 5.061 |

**Table 7.13:** Various variables associated with the presence of beam background for the candidate event which is selected by the cleaning analysis

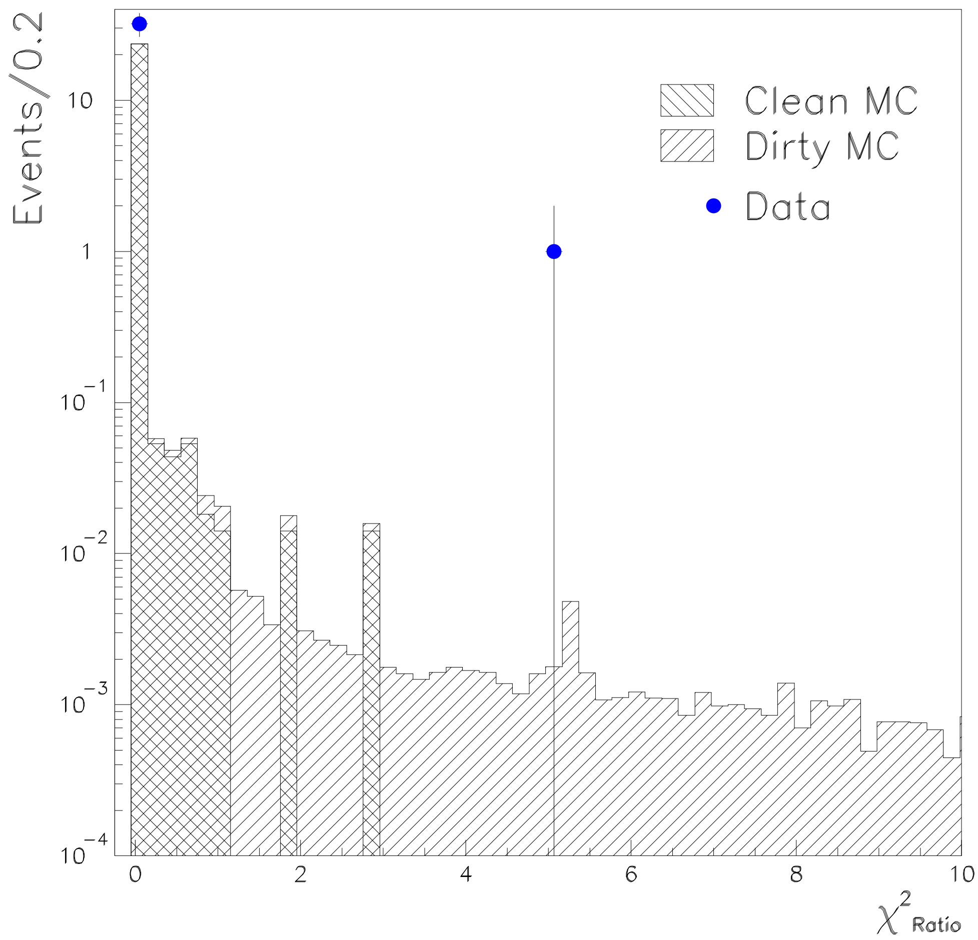


**Figure 7.14:** The distribution of the $\chi^2$ ratio in the hZ exclusive branch for the clean and corrected MC weighted to the final luminosity with the data candidates added

the transposed pairings are effectively correct.

## 7.8 Conclusion

The presence of the beam induced background has been studied in detail with its effect on the 4 jets cuts based analysis determined and a method of treating the background has been devised. The cleaning algorithm, as detailed above, provides a method of cleaning events in which an energetic beam related object is found whilst leaving events without this background unaltered. The application

| Variable | Before | After |
|---|---|---|
| $m_{\text{Reco}}$ | 112.80 | 114.43 |
| $\gamma$ | -1.905 | -1.963 |
| $\Theta$ | 359.70 | 359.99 |
| $y_{34}$ | 0.0134 | 0.0160 |
| $\min(\eta_3, \eta_4)$ | 0.99 | 0.66 |
| $(1-\eta_3)(1-\eta_4)$ | 0.347x$10^{-5}$ | 0.151x$10^{-2}$ |
| $E_{\text{vis}}$ | 253.72 | 231.78 |
| $X_{\gamma}$ | 0.27 | 0.21 |
| $\sum \eta_i$ | 3.651 | 3.655 |

**Table 7.14:** The details of the data candidate selected by the cleaning analysis both before and after the application of the beam background cleaning algorithm.

of this event cleaning to the analysis requires no modification of the event selection efficiency or the pdf shapes which are used for event weighting. One data event is affected by the cleaning algorithm, shown in figure 7.15. In this event a 22 GeV object is removed. The removal of this object from the event results in an increase in the reconstructed Higgs mass of 1.5 GeV/$c^2$ from 112.8 GeV/$c^2$ to 114.4 GeV/$c^2$ [75].

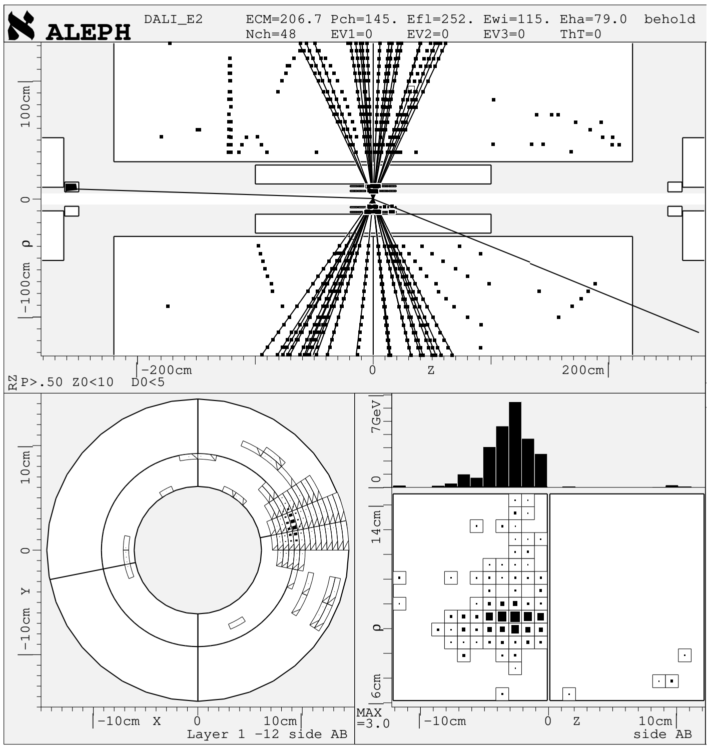


**Figure 7.15:** Event display of the data candidate selected by the beam background cleaning algorithm. The deposit associated with the beam background hypothesis is clearly visible.

# Chapter 8

# Summary, outlook and Conclusion

## 8.1 Introduction

The combined search channels used within the ALEPH Higgs search leads to an observed excess at $m_h \approx 115\,\text{GeV}/c^2$ which is incompatible with a Standard Model only hypothesis at the level of three standard deviations. The observed ALEPH result contributes to an excess in the combined LEP Higgs search which is consistent with the ALEPH result and also is of great interest for future Higgs boson searches[1].

## 8.2 Combined results from LEP Higgs searches

Data from searches at each of the four LEP experiments are statistically combined by the LEP Higgs Working Group to increase the sensitivity to a possible Higgs signal. The use of the likelihood ratio as a test-statistic aids this combination, with each sample being treated as a statistically independent data set.

Although the combination of the final results from all four collaborations has yet to be performed a preliminary combination was undertaken using the final results from the L3 experiment [76] and preliminary results from ALEPH [69], DELPHI [78] and OPAL [79]. The combination result [77] corresponds to a total integrated luminosity of 2465 $\text{pb}^{-1}$ collected by the four collaborations at centre of

[1] In the light of the publication of the final L3 search results [76], the LEP Higgs working group has produced a new (still preliminary) combined result [77]. The final combined result is expected by the end of 2001 after the publication of final results from ALEPH, DELPHI and OPAL.

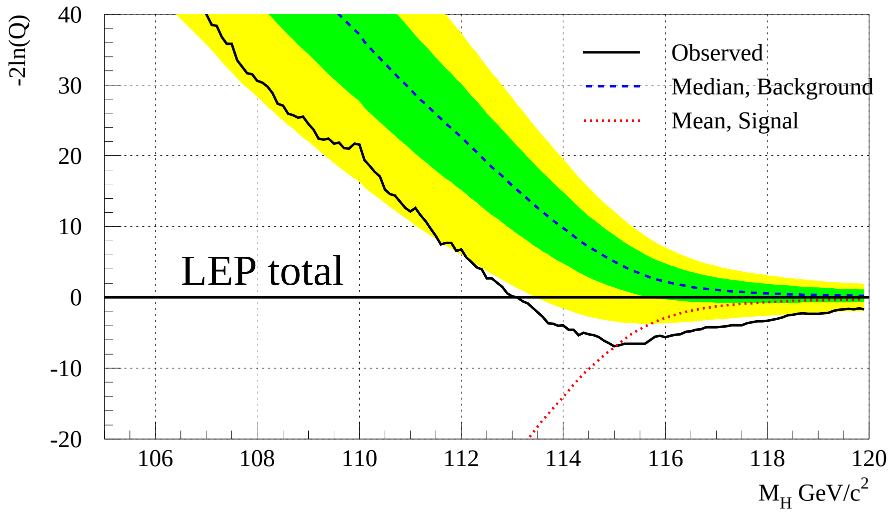


**Figure 8.1:** The distribution of the -2ln$Q$ test statistic as a function of the Higgs boson mass for the expected and observed data from the combination of the search results of the four LEP experiments. The shaded regions around the background only expectation represent the 1$\sigma$ and 2$\sigma$ bands respectively.

mass energies in the range 189-209 GeV with 542 pb$^{-1}$ collected above 206 GeV. The test-statistic, -2 ln$Q$, for the combined LEP experiments expressed as a function of the hypothesised Higgs mass is shown in figure 8.1. The minimum of the test statistic observed at 115.6 GeV/$c^2$ is indicative of a deviation from the SM background only hypothesis while it coincides with the results expected from a SM signal+background hypothesis.

The compatibility of the observed data with the background only hypothesis, as described by (1-CL$_b$), is shown in figure 8.2. For a hypothesised Higgs boson of mass 115.6 GeV/$c^2$, corresponding to the minimum of the test statistic, the value obtained for (1-CL$_b$) is 0.034. Thus the probability for the background to generate the observed effect is 3.4%. This corresponds to a deviation from the expected SM background-only hypothesis of 2.1 $\sigma$[2]. The combined experiments define a lower limit of 114.1 GeV/$c^2$ at the 95% confidence level on the Higgs boson mass via interpretation with the alternate CL$_s$ method [80].

[2]The conversion of 1-CL$_b$ into standard deviations uses a single sided Gaussian approximation.

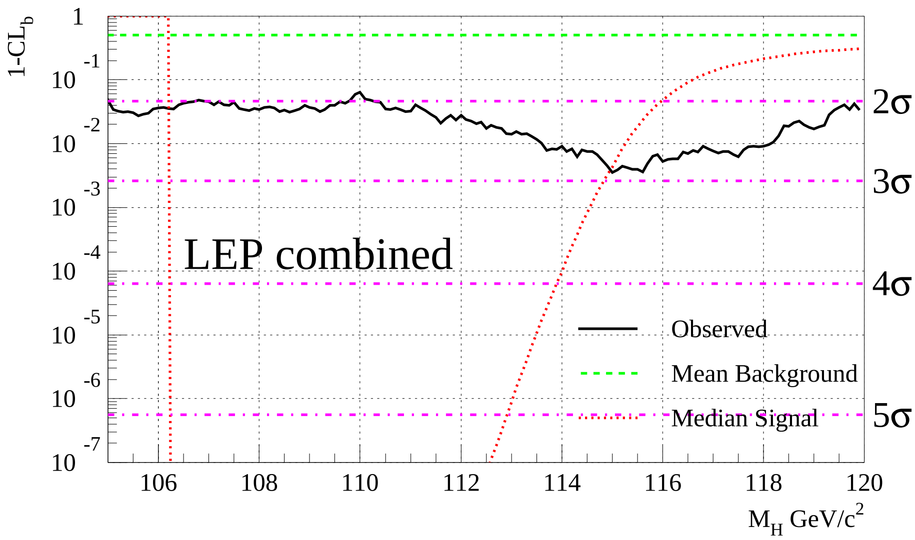


**Figure 8.2:** The distribution of $1 - \mathrm{CL_b}$ for the expected background and signal + background hypothesis and the observed data for the combined LEP experiments.

## 8.3 Future Higgs boson searches

The hints of a Higgs boson with a mass in the region of 115 GeV/$c^2$ are very exciting for future Higgs searches which will take place at the TeVatron and the Large Hadron Collider (LHC) high energy hadron colliders. Although it is now inevitable that special attention will be paid to searches in the region favoured by the LEP data the potential for Higgs searches at these colliders is much greater.

### 8.3.1 Higgs boson searches at the TeVatron

The recently upgraded TeVatron($p\bar{p}$) accelerator at Fermilab presents the best opportunity for a Higgs discovery within the short term future. The addition of a new Main Injector and Antiproton Recycler provide the TeVatron with an increased centre of mass energy from 1.8 TeV to 2.0 TeV, with physics cross sections typically increasing by 30-40%.

The initial goal of the TeVatron Run2, referred to as Run2a, is to deliver a total integrated luminosity of 2 fb$^{-1}$ with instantaneous luminosities up to 2x$10^{32}cm^{-2}s^{-1}$.

Subsequent TeVatron upgrades for the Run2b phase are intended to again raise the available luminosity into the range of 5x$10^{32}cm^{-2}s^{-1}$. The TeVatron

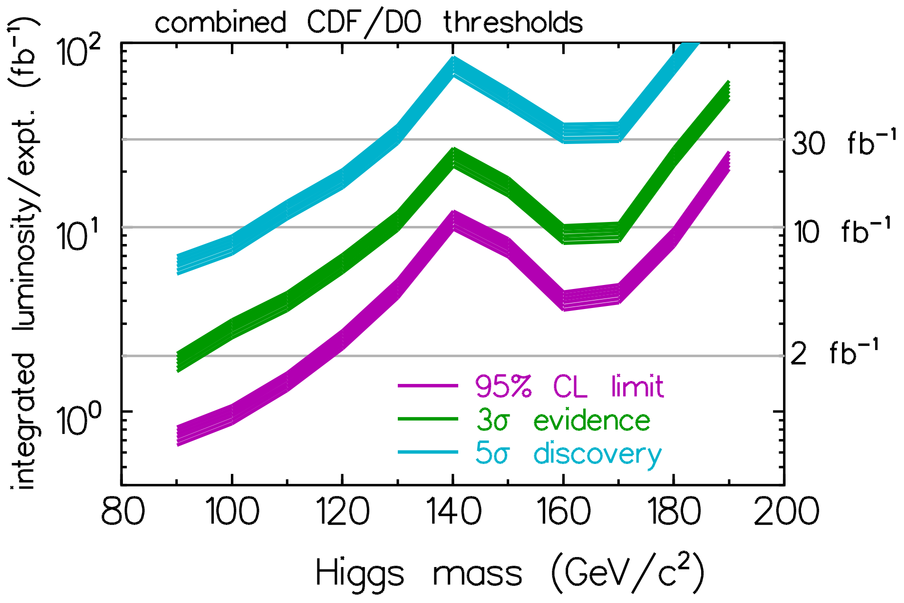


**Figure 8.3:** The discovery potential of the combined Higgs searches at the D0 and CDF detectors at Fermilab [81]. The 95% confidence level exclusion limit, $3\sigma$ observation and $5\sigma$ discovery curves are expressed as a function of the integrated luminosity and hypothesised Higgs boson mass.

thus aims to provide an integrated luminosity of 15 $\text{fb}^{-1}$ per detector by the end of 2007.

In the mass range of current interest, $m_{\text{h}} \approx 115$ GeV$/c^2$, the dominant production mechanism for the Higgs boson at the TeVatron is $gg$ fusion. However the associated decay channel $gg \to h \to \text{b}\bar{\text{b}}$ is not viable since the background cross section for QCD production of $\text{b}\bar{\text{b}}$ di-jet events is orders of magnitude larger than that for Higgs production [81].

The most promising channel in this mass range appears to be the production of a Higgs with an associated vector boson via $\text{p}\bar{\text{p}} \to \text{hW}$, $\text{p}\bar{\text{p}} \to \text{hZ}$. The production of a Higgs boson via this process leads to a number of promising final states, namely: $\ell\nu\text{b}\bar{\text{b}}$, $\nu\bar{\nu}\text{b}\bar{\text{b}}$, $\ell^+\ell^-\text{b}\bar{\text{b}}$ which possess distinct signatures with high $p_t$ leptons and/or missing $E_t$ which aid background rejection.

In the higher mass range with $m_{\text{h}}$ up to 180 GeV$/c^2$ the dominant decay mechanism is $\text{h} \to \text{W}^+\text{W}^-$. The subsequent decay of the W boson leads $\ell\nu\ell\nu$ and $\ell\nu\text{q}\bar{\text{q}}'$ final states which are more easily distinguished from the QCD jet background because of their lepton component.

Figure 8.3 shows the expected sensitivity of the combined CDF and D0 searches for SM Higgs boson production. In the absence of a signal, an integrated luminosity of 2 $fb^{-1}$ should allow the TeVatron experiments to exclude, at 95% confidence level, Higgs boson masses up to 115 $GeV/c^2$.

Higher integrated luminosities of 15 $fb^{-1}$ are expected to provide the potential of a $5\sigma$ discovery for Higgs boson masses in the 115 $GeV/c^2$ mass region while also allowing all masses up to 180 $GeV/c^2$ to be excluded in the absence of a Higgs boson signal. Integrated luminosities of 25 $fb^{-1}$ and above are required to provide discovery evidence of the order of $3\sigma$, in the mass range up to 180 $GeV/c^2$.

There exists a window of opportunity for the TeVatron collider to discover the Higgs boson in the current region of interest ($m_{\mathrm{h}} \approx 115$ $GeV/c^2$) before the commissioning of the LHC in 2006/2007. The discovery of a Higgs boson with a mass exceeding 180 $GeV/c^2$ however becomes increasingly difficult for the TeVatron and it is in this higher mass range that the LHC will excel.

### 8.3.2 Higgs boson searches at the LHC

The LHC at CERN will produce pp collisions with a centre of mass energy of 14 TeV and an instantaneous luminosity between $10^{33}$ and $10^{34} cm^{-2} s^{-1}$. The combination of the high centre of mass energy and luminosity allow two all purpose detectors, ATLAS and CMS, to search for the Higgs boson in the mass range from 100 $GeV/c^2$ to 1 $TeV/c^2$.

In the mass region $m_{\mathrm{h}} \approx 130$ $GeV/c^2$ the detection of the Higgs boson via its dominant decay mechanism, $\mathrm{h} \rightarrow \mathrm{b}\bar{\mathrm{b}}$, is problematic due to the overwhelmingly high QCD jet background expected at the LHC [82] [83].

In this mass range the LHC benefits from its high luminosity which means that even rare processes may make a significant contribution to the Higgs searches. The most promising channel in this mass range is $\mathrm{h} \rightarrow \gamma\gamma$, in which the Higgs boson is initially produced via gluon or vector boson fusion. This channel provides a clear signature and has good discovery potential for a Higgs boson of mass $\approx 115$ $GeV/c^2$ [82] [83].

The search for Higgs bosons in the higher mass region 140-800 GeV/$c^2$ is somewhat different. In this region the dominant decay of the Higgs boson to vector bosons $W^+W^-, ZZ^{(*)}$ provides the best opportunity for discovery. Although the decay to $W^+W^-$ is the dominant mode at these masses the presence of higher backgrounds make this channel less powerful than the so called "Gold Plated" $ZZ^{(*)}$ decay. The "Gold Plated" $ZZ^{(*)}$ decay mode refers to the $h \to ZZ \to \ell^+\ell^-\ell^+\ell^-$ process which has a low background and indeed provides the majority of the total discovery potential over a large mass range at the LHC.

At very high Higgs boson masses 800 GeV/$c^2$ $\to$ 1 TeV/$c^2$ the increasing decay width of the Higgs boson requires the addition of other channels to the "Golden Plated" channel to allow discovery.

The LHC is due to be commissioned in 2006 with a physics run providing 10 fb$^{-1}$ in 2007. The combined discovery potential of the ATLAS and CMS detectors is show in figure 8.4. The collection of an integrated luminosity of 10 fb$^{-1}$ is expected to lead to the discovery of a Higgs boson if it exists in the mass range 115 GeV/$c^2$ to 1 TeV/$c^2$.

Failure to observe the Higgs boson at the LHC would require the re-evaluation of the Higgs mechanism for spontaneous symmetry breaking in the electroweak sector with the possibility of its replacement with an alternative theory.

## 8.4 Conclusion

The observation of a significant deviation from the SM background expectation in the Higgs searches performed by the ALEPH collaboration and subsequently the LEP Higgs working Group have generated a great deal of excitement and are extremely promising for future Higgs searches.

The determination of the true nature of this deviation, whether due to a Higgs boson signature or statistical background fluctuation, requires more data which the TeVatron and LHC colliders will provide.

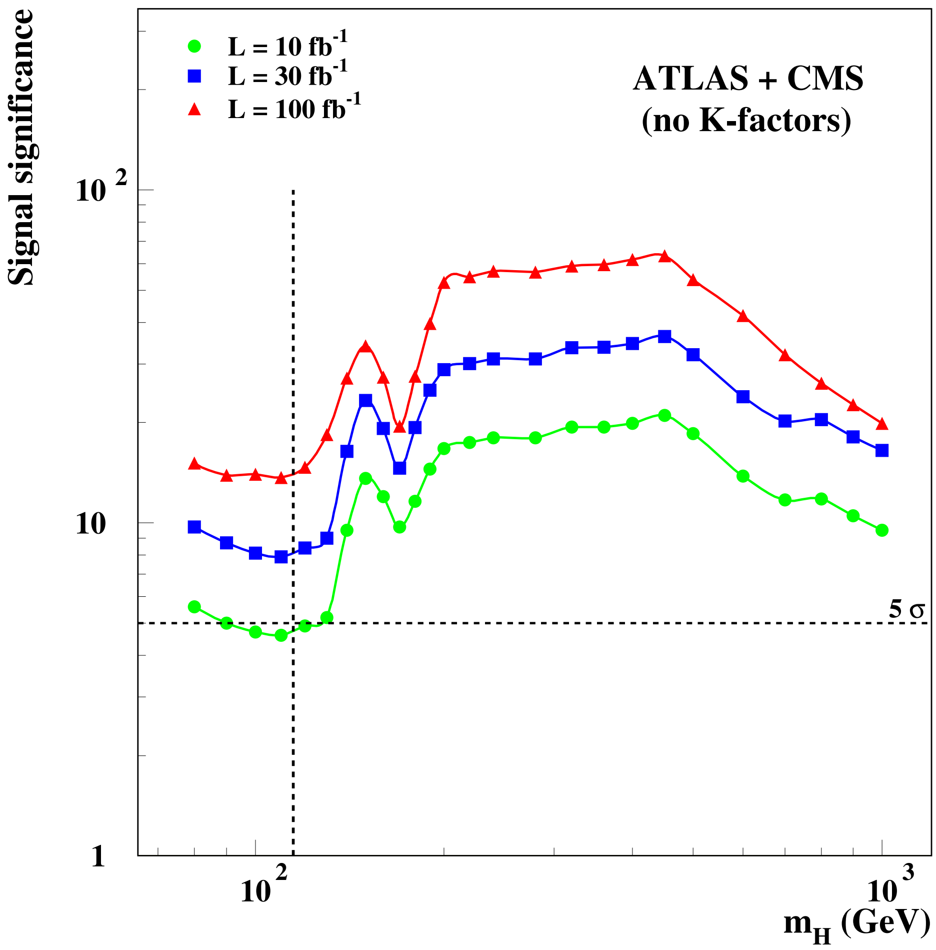


**Figure 8.4:** The combined discovery significance for the Higgs boson with the ATLAS and CMS detectors at the LHC. The significance is expressed as a function of the hypothesised Higgs boson mass for various integrated luminosities collected by each experiment.Taken from [84].

# Appendix A
# Jet pairing studies

## A.1 Introduction

All events which pass through the hZ four jets selection with more than one selected combination need a pairing choice to be implemented in order to select which combination is used in the construction of the discriminating variable, $m_{\text{Reco}}$. This chapter presents a brief overview of jet pairing studies performed in an attempt to find an improved pairing method.

## A.2 Pairing methods

Several possible pairing methods were studied, as detailed below:

- **DEC** - the current pairing method based on the decay angles likelihood method, as detailed in section 4.3.4.
- **COZ** - the pairing in which the di-jets are most back to back for both the Higgs and the Z bosons, formed from the sum of the cosines of the opening angles of the boson candidates.
- **MZ** - the pairing which has the Z boson candidate di-jet invariant mass closest to $m_Z$.
- **BTG** - the pairing in which the Higgs candidate di-jet possesses the maximum sum of b-tagging values.

- **ME** - the pairing with the maximum hZ matrix element value.
- **MTP** - the likelihood approach to the hZ matrix element pairing.
- **NID** - the pairing defined according to the product of the MTP and DEC likelihoods.

Initially the pairing studies focused on the comparison of the current DEC method, the MZ method, as used prior to the introduction of the DEC method, and the new COZ method which was proposed as a possible improvement. Later studies were performed using a wider range of variables with the hZ matrix element being introduced and possible 2D combinations of the pairing variables considered by forming likelihoods and linear discriminants.

All studies were performed using a simulated hZ sample with $m_{\mathrm{h}} = 107\ \mathrm{GeV}/c^2$ at $\sqrt{s} = 200$ GeV. The pairing methods were initially compared by considering the efficiency for a given method to correctly pair the jets in the simulated signal sample[1]. Events are said to be correctly paired if the two jets originating from the decay of the Higgs boson are correctly identified as such. The discriminating variable,

$$m_{\mathrm{Reco}} = \mathrm{m}_{12} + \mathrm{m}_{34} - 91.2\ \mathrm{GeV}/c^2, \tag{A.1}$$

is constructed such that it is invariant under the transposition of the two boson masses $\mathrm{m}_{12}$ and $\mathrm{m}_{34}$. Following from this the transposed assignment of the di-jets to the h and Z bosons, although reversing the boson masses themselves, leads to a value of the discriminating variable which is effectively correct. This effect is taken into consideration when we assign a pairing efficiency to each of the considered methods.

Table A.1 presents the correct and correct + transposed pairing efficiencies for each method. The table clearly shows that the NID method is the most performant at selecting the correct pairing while both the COZ and DEC pairing methods show a reasonable level of performance. When considering the correct

[1] All quoted efficiencies refer to events in which more than one possible di-jet combination is selected; as such they require a pairing choice to be made.

+ transposed pairing efficiency then we find that the COZ method is now the most performant followed by the DEC and NID methods. The nature of the COZ variable dictates that the same value of COZ is returned for both correct and transposed pairings and thus it is intuitive that this method will gain from the inclusion of the transposed pairs in the efficiency calculation.

| pairing variable | correct pairing efficiency (%) | correct + transposed pairing efficiency (%) |
|---|---|---|
| DEC | 51.6 | 89.2 |
| COZ | 60.7 | 95.6 |
| MZ | 43.9 | 55.4 |
| BTG | 37.9 | 49.7 |
| ME | 40.6 | 51.6 |
| MTP | 39.9 | 52.2 |
| NID | 65.4 | 85.2 |

**Table A.1:** The efficiencies of the different jet pairing variables for correct pairing and correct + transposed pairings.

The performance of the ME pairing appears to be quite poor compared to what was originally expected. To test the performance of the ME pairing the hZ matrix element was calculated using the MC truth level information for the final state quarks produced in the decay of the h and Z bosons and the ME pairing efficiency was re-evaluated using this information. The resulting efficiency for the correct pairing was found to be 66%. The decrease in performance when using the reconstructed di-jets is expected to be due to hadronisation and detector reconstruction effects. The MTP performance is found to be very close to that of the ME as would be expected and is introduced only to allow the combination of the MTP and DEC variables to form the NID variable.

The stability of the pairing methods as a function of Higgs boson mass was evaluated by applying the different pairing methods to MC simulation samples with various values of $m_{\mathrm{h}}$. Figure A.1 shows the pairing efficiency of the various variables as a function of Higgs boson mass. The mass dependence of the variables is immediately visible with the COZ method showing a strong mass bias towards the high Higgs boson mass region. The DEC and NID methods show similar performance to each other across the considered mass range, with their performance also peaked in the high Higgs mass range but with a less pronounced dependency than the COZ method. The MZ and MTP methods are found to be peaked in efficiency at $m_{\mathrm{h}} = 90\,\mathrm{GeV}/c^2$ as would be expected considering their dependence on $m_Z$ and the transposition effect. Lastly the BTG method is very stable across the considered mass range but has a consistently poor performance. The nature of the Higgs boson search would prefer the choice of a pairing method which shows improved pairing efficiency in the high Higgs mass range where the sensitivity of the search is most important.

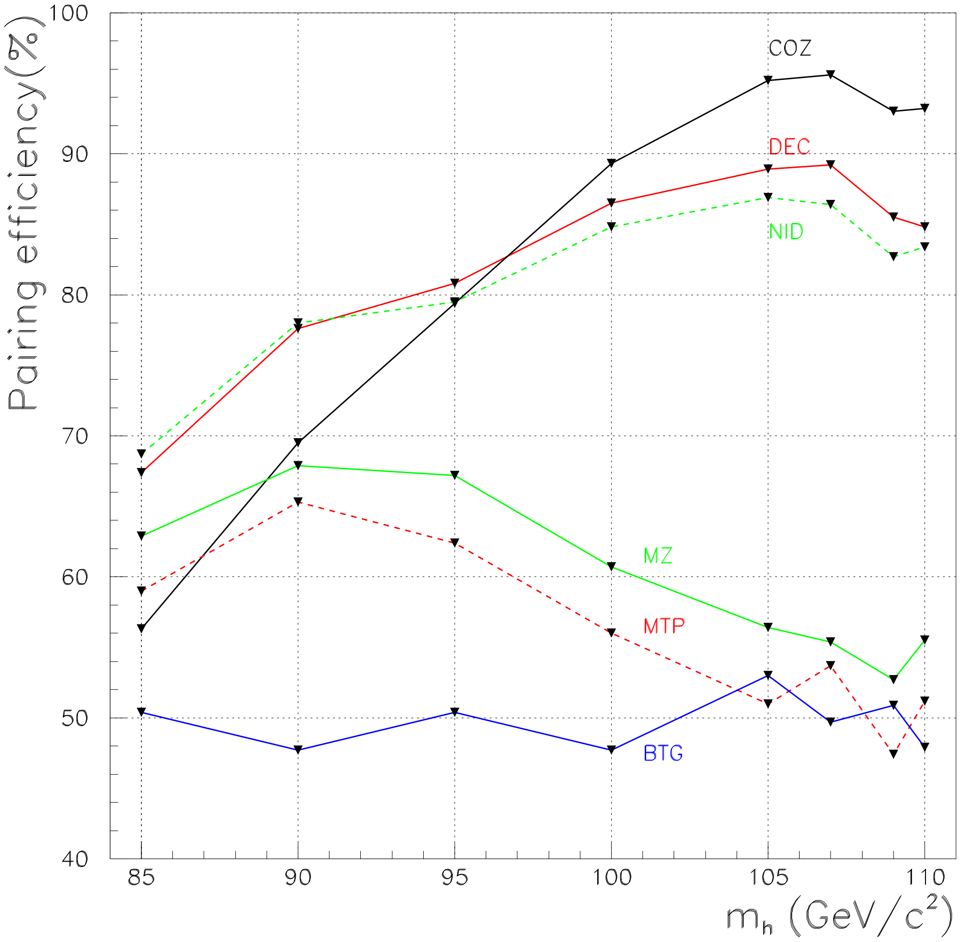


**Figure A.1:** The jet pairing efficiency as a function of Higgs boson mass for the various pairing variables

The pairing efficiency for the simulated hZ samples, although providing some indication about the performance of a given pairing variable, is not the only factor to be considered where the performance of the analysis is concerned. The performance of the hZ four jets analysis is ultimately dependent on the reconstructed mass distributions of both the hZ signal and full background samples which are used to set confidence levels (see appendix B). Following from this the distributions of $m_{\mathrm{Reco}}$ for both the hZ signal and full background are compared for the DEC vs COZ pairing methods in figures A.2a and A.2b and for the DEC vs MZ methods in figures A.3a and A.3b respectively. It can be seen that the COZ pairing method indeed improves the significance of the hZ signal and yet it also leads to higher background concentrations at high reconstructed Higgs mass. The MZ method on the other hand reduces the background concentration at high reconstructed mass but also results in reduced significance for the signal.

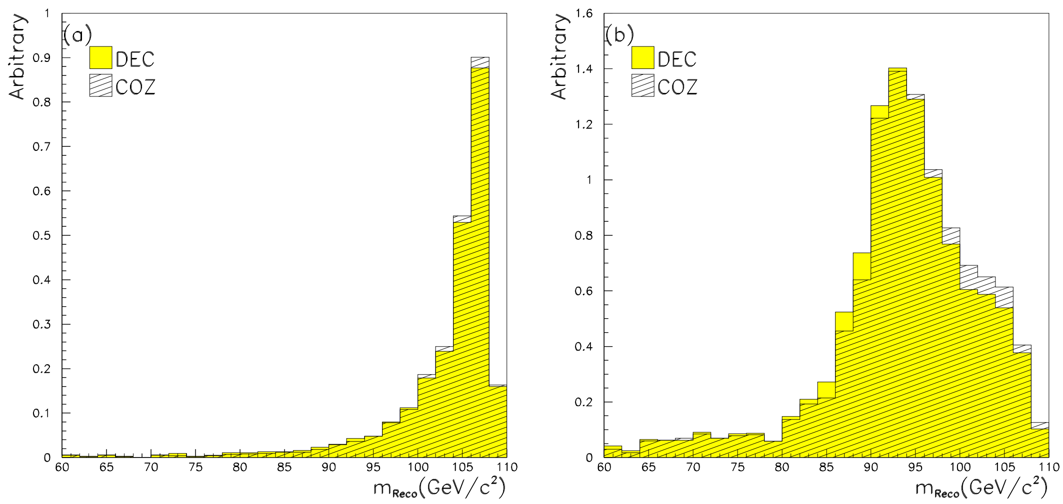


**Figure A.2:** A comparison of the distributions of the reconstructed Higgs boson mass, $m_{\mathrm{Reco}}$, for (a) hZ signal with $m_{\mathrm{h}} = 107\,\mathrm{GeV}/c^2$ and (b) the full background for the DEC versus the COZ pairing methods.

The calculation of the expected confidence level for the exclusion of a Higgs boson signal, $\mathrm{CL_{se}}$, was performed using the DEC, COZ and MZ methods and it is based on the results of this calculation that the analysis performance is assessed [85]. Figure A.4a shows the expected $\mathrm{CL_{se}}$ distributions for the three possible pairing methods over a large $m_{\mathrm{h}}$ range while figure A.4b shows the $\mathrm{CL_{se}}$ distributions in the range of the 95% confidence level. The dashed line at $\mathrm{CL_{se}}$=0.05

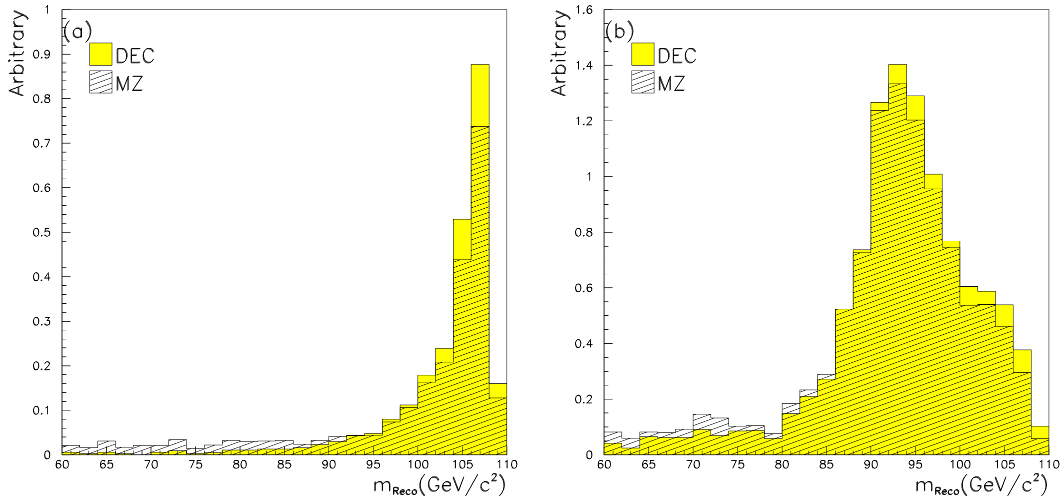


**Figure A.3:** A comparison of the distributions of the reconstructed Higgs boson mass, $m_{\rm Reco}$, for (a) hZ signal with $m_{\rm h} = 107\,{\rm GeV}/c^2$ and (b) the full background for the DEC versus the MZ pairing methods.

shows the 95% confidence line and the expected exclusion limit is set at the point where the curve crosses this line. From the figures it is clear that the MZ pairing is consistently worse than the DEC and COZ methods and the zoomed view shows that the DEC pairing is slightly more performant than the COZ method. Based on this comparison the chosen pairing method is the DEC method.

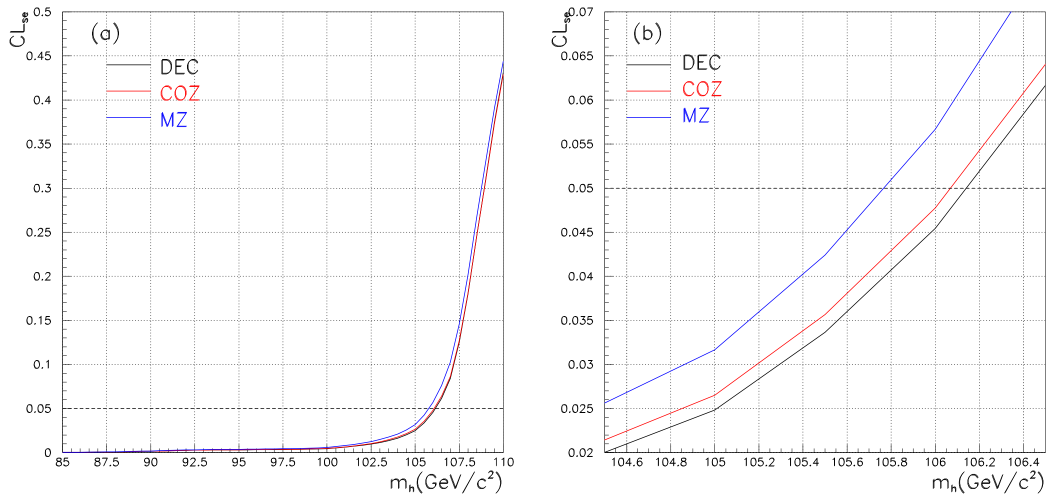


**Figure A.4:** A comparison the the $\rm CL_{se}$ curves confidence level curves for the DEC, COZ and MZ pairing methods (a) over a large Higgs boson mass range and (b) in the zoomed region where the expected 95% confidence level is set.

Figures A.5, A.6 and A.7 show the hZ signal and background $m_{\text{Reco}}$ distributions for the BTG, ME and NID pairing methods respectively. The performance of the BTG and ME methods is similar to that of the MZ pairing method and studies of these methods are taken no further since they are expected to result in a diminished performance of the analysis. The NID method shows more promise but seems to follow the DEC method very closely. It is considered that any resulting change in performance due to the implementation of the NID method would be small and as such studies of this method are also taken no further.

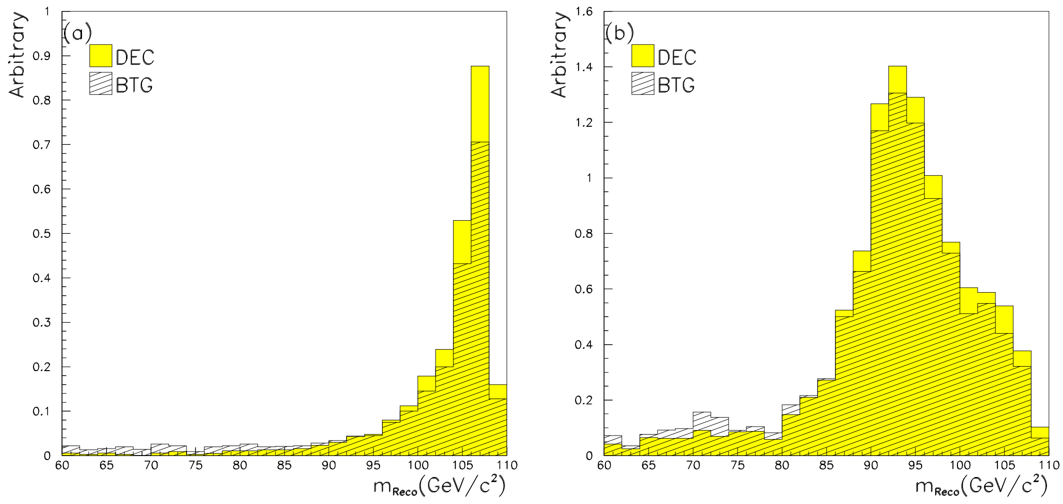


**Figure A.5:** A comparison of the distributions of the reconstructed Higgs boson mass, $m_{\text{Reco}}$, for (a) hZ signal with $m_{\text{h}} = 107\,\text{GeV}/c^2$ and (b) the full background for the DEC versus the BTG pairing methods.

## A.3 Conclusion

Numerous pairing methods were studied with none of the proposed methods showing improved performance compared to the DEC method which is currently utilised within the four jets analysis. The use of the decay angles (DEC) pairing method within the four jets cuts analysis is thus proven to be optimal and is continued [86].

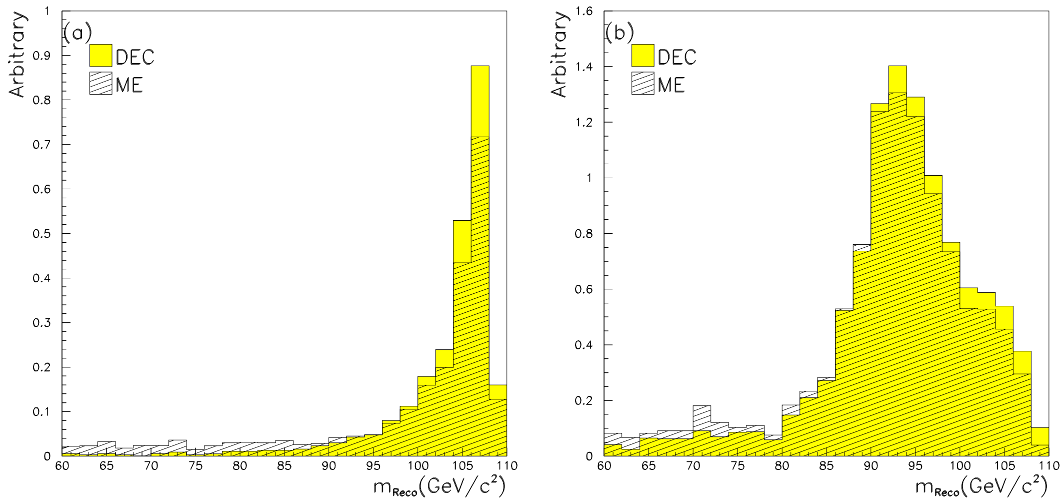


**Figure A.6:** A comparison of the distributions of the reconstructed Higgs boson mass, $m_{\mathrm{Reco}}$, for (a) hZ signal with $m_{\mathrm{h}} = 107\,\mathrm{GeV}/c^2$ and (b) the full background for the DEC versus the ME pairing methods.

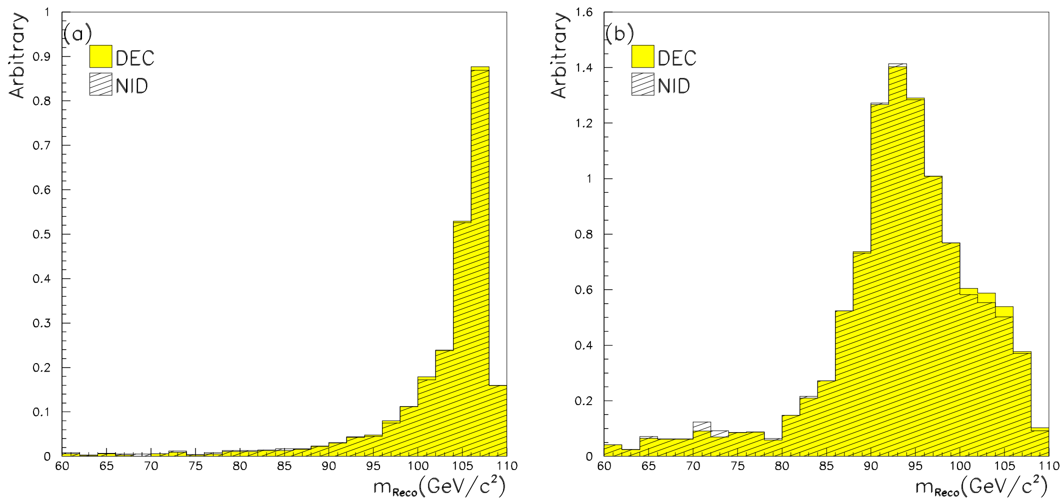


**Figure A.7:** A comparison of the distributions of the reconstructed Higgs boson mass, $m_{\mathrm{Reco}}$, for (a) hZ signal with $m_{\mathrm{h}} = 107\,\mathrm{GeV}/c^2$ and (b) the full background for the DEC versus the NID pairing methods.

# Appendix B

# Statistical Interpretation of search results

Scientists have calculated that the chance of anything so patently absurd actually existing are millions to one. But magicians have calculated that million-to-one chances crop up nine times out of ten.
- Mort, Terry Pratchett

## B.1 Introduction

The aim of the Higgs searches within the ALEPH collaboration is to determine the extent to which a possible Higgs boson signal may be excluded or conversely to determine the most compatible Higgs boson mass hypothesis in the presence of any signal like behaviour.

## B.2 The Statistical Estimator

To aid with the interpretation of the results from the experiment a test-statistic is defined. The test-statistic, or estimator, as we shall refer to it here, is a function of the experimental observables and population parameters which may be used to rank experiments from least to most signal-like (most to least background-like). The estimator is designed to increase monotonically for increasingly signal-like (decreasingly background-like) experiments. Following this the confidence in the signal+background hypothesis is given by the probability that the estimator is less than or equal to the value observed in the experiment.

In the case of the Higgs searches at ALEPH the chosen statistical estimator is the likelihood ratio, denoted as Q. The likelihood ratio is defined as

$$\mathrm{Q}(m_\mathrm{h}) = \frac{\mathcal{L}_{\mathrm{s+b}}(m_\mathrm{h})}{\mathcal{L}_\mathrm{b}(m_\mathrm{h})} \tag{B.1}$$

i.e., the ratio of the likelihoods of the signal+background hypothesis to the background only hypothesis.

The likelihood ratio [87] is a common test-statistic since it maximises the probability of rejecting a false hypothesis at a given confidence level whilst minimising the probability of making a false discovery at a given confidence level [87].

The two likelihoods used in the likelihood ratio are given by

$$\mathcal{L}_{\mathrm{s+b}} = \prod_{i=1}^{\mathrm{N}} \frac{\exp[-(\mathrm{s}_i(m_\mathrm{h}) + \mathrm{b}_i)](\mathrm{s}_i(m_\mathrm{h}) + \mathrm{b}_i)^{\mathrm{n}_i}}{\mathrm{n}_i!} \times \prod_{j=1}^{\mathrm{n}_i} \frac{\mathrm{s}_i(m_\mathrm{h})\mathrm{S}_i(m_\mathrm{h}, m_{ij}) + \mathrm{b}_i\mathrm{B}_i(m_{ij})}{\mathrm{s}_i(m_\mathrm{h}) + \mathrm{b}_i} \tag{B.2}$$

$$\mathcal{L}_\mathrm{b} = \mathcal{L}_{\mathrm{s+b}}\Big|_{\mathrm{s}=0} \tag{B.3}$$

where N is the number of independent search channels and $\mathrm{s}_i$, $\mathrm{b}_i$ and $\mathrm{n}_i$ are the expected number of signal, background and the observed number of events in channel $i$ respectively. The two functions $\mathrm{S}_i$ and $\mathrm{B}_i$ are the probability densities that a signal or background event, respectively, will be found in a given final state identified by a set of discriminating variables, here represented generically by $m_{ij}$.

The form of the likelihood defined in equation B.2 is seen to be comprised of two independent parts. Firstly there is a simple Poisson term corresponding to the ratio of the Poisson probabilities to observe $n_{obs}$ events for the signal + background and background only hypotheses.

$$\mathcal{P} = \frac{\exp[-(\mathrm{s}_i(m_\mathrm{h}) + \mathrm{b}_i)](\mathrm{s}_i(m_\mathrm{h}) + \mathrm{b}_i)^{\mathrm{n}_i}}{\mathrm{n}_i!} \tag{B.4}$$

The second term

$$\rho(m_{ij}) = \frac{\mathrm{s}_i(m_\mathrm{h})\mathrm{S}_i(m_\mathrm{h}, m_{ij}) + \mathrm{b}_i\mathrm{B}_i(m_{ij})}{\mathrm{s}_i(m_\mathrm{h}) + \mathrm{b}_i} \tag{B.5}$$

is the ratio of probability density for obtaining a candidate with a given set of discriminating variables in channel $i$ for the two hypotheses noted above.

The estimator may be written more simply as

$$\mathrm{Q}(m_\mathrm{h}) = \prod_{i=1}^{\mathrm{N}} \mathrm{e}^{-\mathrm{s}_i} \prod_{j=1}^{\mathrm{n}_i} \left[1 + \frac{\mathrm{s}_i\mathrm{S}_i(m_\mathrm{h}, m_{ij})}{\mathrm{b}_i\mathrm{B}_i(m_{ij})}\right] \tag{B.6}$$

and is often expressed in the form -2 ln $Q$ where

$$-2\ln(\mathrm{Q}) = 2\mathrm{s} \; - \; 2\sum_{i=1}^{\mathrm{N}}\sum_{j=1}^{\mathrm{n}_i} \ln\left[1 + \frac{\mathrm{s}_i\mathrm{S}_i(m_\mathrm{h}, m_{ij})}{\mathrm{b}_i\mathrm{B}_i(m_{ij})}\right] \tag{B.7}$$

In this form it can be readily seen that the individual candidates contribute to the combined result with a "weight" given by

$$\ln\left[1 + \frac{\mathrm{s}_i\mathrm{S}_i(m_\mathrm{h}, m_{ij})}{\mathrm{b}_i\mathrm{B}_i(m_{ij})}\right] \tag{B.8}$$

A high event weight implies that a candidate is more likely to have originated from the signal+background hypothesis than from the background only hypothesis.

## B.3 Deriving confidence levels from the estimator

The statistical estimator as described above provides a means of comparing the compatibility of an experimental result with the presence of a hypothesised Higgs boson of a given mass. The observed result may then be interpreted as signal-like or background-like based on the outcome of the estimator calculation. The value of the estimator itself however lacks any ability to provide a measure confidence in the observed result. Indeed in experiments where very few signal events are expected with a large underlying background, it may be quite possible for a fluctuation in the number of background events to produce an estimator value which could be interpreted as signal-like. To correctly interpret the estimator result and to obtain a measure of confidence in the experimental result the distribution of estimator values for all possible background and signal+background hypotheses is required. To find the estimator distributions many simulated experiments for background only and signal+background hypothesis are performed. Each of these experiments, known as Toy Monte Carlo (Toy MC), is generated by considering the Poisson distribution of the number of expected signal plus background events for the given hypothesis. The experiments are also generated according to the relevant distribution of the signal and background pdf's for the discriminating variable used within the relevant channel to correctly account for the event weights. The value of the estimator for each generated Toy MC experiment may then be calculated. Once this procedure has been performed for a large number of Toy MC experiments the corresponding distributions of the estimator may be found. Figure B.1 shows the distributions of the estimator values for the background only and signal + background hypotheses for three different possible Higgs signals. The distributions of the estimator are from now on referred to as $\rho_b$ and $\rho_{s+b}$ for the background only hypothesis and signal+background hypothesis respectively. It is clear that the distributions of the estimator are well separated for a light Higgs mass where many signal events may be expected but as the mass of the hypothesised Higgs boson increases and the expected number of signal events decreases the two distributions begin to converge until they become equal when the number of expected signal events is zero.

The estimator distributions, once defined, may be used to interpret the value of the estimator from the real experiment, in the form of a confidence level. Two confidence levels may be defined, $\mathrm{CL_b}$ and $\mathrm{CL_{s+b}}$, so that the confidence in the background only and signal+background hypotheses is given by the probability that the estimator is less signal like or equal to the value obtained in the real data experiment $\mathrm{x}_0(m_\mathrm{h})$, for a given signal hypothesis experiment. That is

$$\mathrm{CL_b}(m_\mathrm{h}) = \int_{\mathrm{x}_0(m_\mathrm{h})}^{\infty} \rho_b(m_\mathrm{h})\,\mathrm{d}x \tag{B.9}$$

$$\mathrm{CL_{s+b}}(m_\mathrm{h}) = \int_{\mathrm{x}_0(m_\mathrm{h})}^{\infty} \rho_{s+b}(m_\mathrm{h})\,\mathrm{d}x \tag{B.10}$$

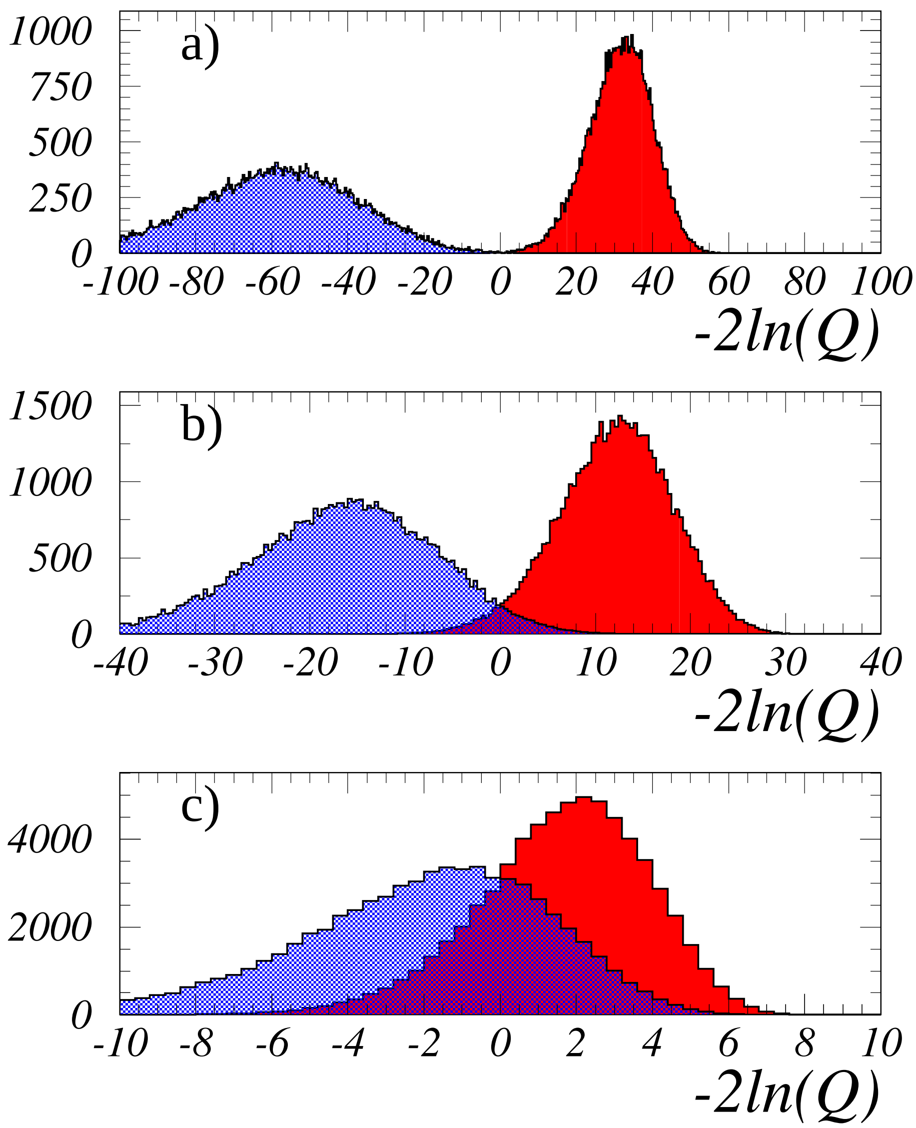


**Figure B.1:** Examples of distributions of -2ln$Q$ for signal+background(blue) and background only(red) hypotheses: a) for a light mass Higgs with a large cross section, b) for a moderate mass Higgs with a moderate cross section and c) for a heavy Higgs with a small cross section.Taken from [80]

where $x$ = -2ln$Q$.

The integrals of the two confidence levels for an example Toy MC experiment are shown in figure B.2.

The quantity $\mathrm{CL_b}$ defines the fraction of $\rho_b$ which is more background like than the observed value of the estimator from the experiment. $\mathrm{CL_{s+b}}$ similarly defines the fraction of $\rho_{s+b}$ which is more background-like than the result obtained from the experiment. To help with the interpretation of these results and to in the determination of the search performance, two quantities $\langle \mathrm{CL_b} \rangle$ and $\langle \mathrm{CL_{s+b}} \rangle$ are defined, where

$$\langle \mathrm{CL_b} \rangle = \overline{\mathrm{CL_b}} \quad \Big|_{\text{background only}} \tag{B.11}$$

$$\langle \mathrm{CL_{s+b}} \rangle = \overline{\mathrm{CL_{s+b}}} \Big|_{\text{background only}} \tag{B.12}$$

These two quantities are the median expected values of $\mathrm{CL_b}$ and $\mathrm{CL_{s+b}}$ respectively for experiments in which only background is present.

An analytic method based on fast Fourier transformation is used to calculate confidence levels within the ALEPH Higgs searches [88]. The high precision and fast computation associated with this analytic method means it is preferred to the slower Toy MC approach as described above.

## B.4 Interpretation for the exclusion of signals

In the absence of any significant deviation from the expected SM background the final interpretation of the search is to produce a lower limit on the Higgs boson mass. To provide such a lower limit we require a means of expressing a confidence in the signal hypothesis of the experiment. To determine a confidence for signal only is impossible in a real experiment since for each candidate it is impossible to define if it came from a signal or background process. In accordance with this we calculate $\mathrm{CL_{s+b}}$ which is a measure of the confidence in the signal + background hypothesis. This quantity however encounters problems when interpreting signal like behaviour and producing a mass limit. For cases in which the expected number of signal events is very small $\mathrm{CL_{s+b}}$ approaches $\mathrm{CL_b}$. In these cases any downwards fluctuation in the number of background events would result in an artificially low value of $\mathrm{CL_{s+b}}$ and, as such, an exaggerated (high) mass limit would be set. To overcome this problem the confidence in the signal hypothesis may be estimated in the following two ways.

$$\mathrm{CL_s} = \frac{\mathrm{CL_{s+b}}}{\mathrm{CL_b}} \tag{B.13}$$

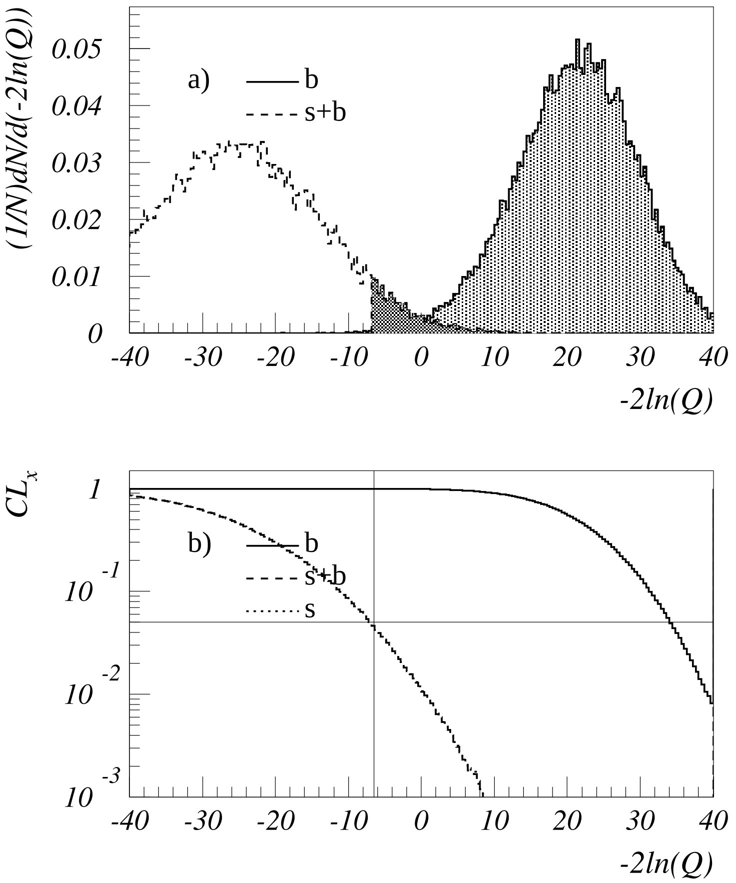


**Figure B.2:** Figure a) shows the distributions of the estimator -2ln$Q$ obtained for a given $m_\mathrm{h}$ from a large number of Toy MC experiments. The solid(dashed) line shows the estimator distribution for background(signal+background). The two distributions describe the probability density functions $\rho_b$ and $\rho_{s+b}$ respectively while the shaded regions represent the integrations corresponding to $\mathrm{CL_b}$ and $\mathrm{CL_{s+b}}$. Figure b) shows the the integrated values of $\mathrm{CL_b}$ and $\mathrm{CL_{s+b}}$ at any chosen value of -2ln$Q$ for $\mathrm{CL_b}$ , $\mathrm{CL_{s+b}}$ and $\mathrm{CL_s}$ [89].

$$\mathrm{CL_{se}} = \mathrm{CL_{s+b}} + (1 - \mathrm{CL_b})e^{-s} \tag{B.14}$$

The interpretation of the quantities $\mathrm{CL_s}$ and $\mathrm{CL_{se}}$ as confidence levels may then be used to determine the extent to which any signal is excluded. Both $\mathrm{CL_s}$ and $\mathrm{CL_{se}}$ are normalised in some way by $\mathrm{CL_b}$. This allows us to overcome the problems associated with background fluctuations and also reduces the probability of a signal mass being falsely excluded. Within the ALEPH collaboration the quantity $\mathrm{CL_{se}}$, called the confidence level "signal estimator" method, is used to interpret search results in the absence of any signal. The signal esti-

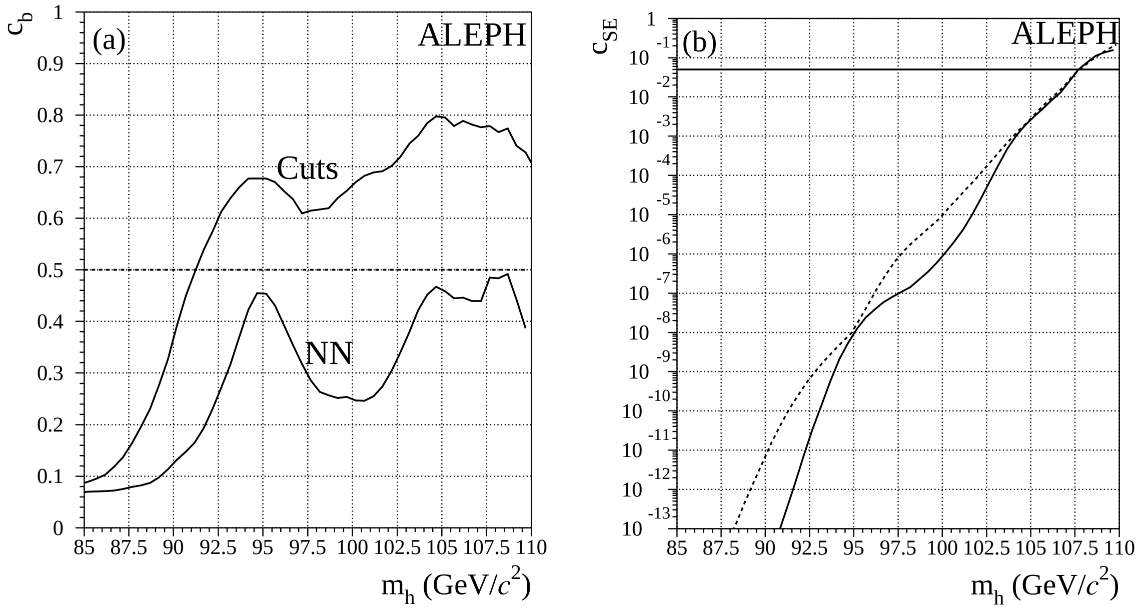


**Figure B.3:** Observed(solid) and expected(dashed) CL curves for (a) the background only hypothesis for both the NN and cuts streams and (b) the signal hypothesis for the NN stream, as a function of hypothesised Higgs boson mass $m_h$ taken from the 1999 ALEPH Higgs search [56].

mator method has been shown to yield stronger limits than the alternative $\mathrm{CL_s}$ method [90], whilst still returning the deontologically correct correct result, (ie, $\mathrm{CL_{se}} = 1 - \exp^{(-s)}$), when zero events are observed. A lower limit on the Higgs mass at the 95% confidence level is set at the lowest mass ($m_\mathrm{h}$) for which $\mathrm{CL_{se}}$ exceeds 5%. Figure B.3b shows an example of the $\mathrm{CL_{se}}$ distribution taken from the ALEPH Higgs search results in 1999 [56]. The solid line at 0.05 indicates the 95% confidence level line with the intersection of the curve and this line defining the 95% CL lower limit on the Higgs mass, in this case the limit was set at 107.7 GeV/$c^2$.

## B.5 Interpretation of results in the presence of a signal

In the presence of a deviation from the expected SM background and thus the possible observation of a signal the quantity $\mathrm{CL_b}$ is used to quantify the size of the deviation.

The quantity $CL_b$ is used to define the confidence in the background only hypothesis. The median value of $CL_b$ is expected to be 0.5 in the presence of background only and provides a means by which the experiment can be quantified as being more or less background compatible. Any significant deviations from this value indicate an excess or deficit of data in comparison to the expected behaviour of the SM. If a significant excess is observed then the value of $CL_b$ ($1-CL_b$) is expected to approach 1 (0) and a possible discovery may have been observed. In the presence of any such deviation from the expected SM background the deviation on $CL_b$ may be transformed into a measure of Gaussian standard deviations for interpretation. Table B.1 shows the conversion from $CL_b$ values to Gaussian standard deviations based on a single sided Gaussian integration. Figure B.3a shows an example of a $CL_b$ distribution for the data recorded and analysed by the ALEPH Higgs search in 1999 [56]. The median expected value of 0.5 is marked while the analyses for the cuts based stream showed an indication of an excess and the NN stream showed indication of a slight deficit in data.

| Standard Deviations | Confidence Level |
|---|---|
| +2.45 sigma | 0.993 |
| +2.41 sigma | 0.992 |
| +2.36 sigma | 0.991 |
| +2.33 sigma | 0.990 |
| +2.05 sigma | 0.980 |
| +1.88 sigma | 0.970 |
| +1.75 sigma | 0.960 |
| +1.64 sigma | 0.950 |
| +1.28 sigma | 0.900 |
| +0.84 sigma | 0.800 |
| +0.52 sigma | 0.700 |
| +0.25 sigma | 0.600 |
| 0 sigma | 0.500 |

| Standard Deviations | Confidence Level |
|---|---|
| +6 sigma | 0.999999999 |
| +5 sigma | 0.99999971 |
| +4 sigma | 0.9999685 |
| +3 sigma | 0.99865 |
| +2 sigma | 0.97725 |
| +1 sigma | 0.84135 |
| 0 sigma | 0.50000 |
| -1 sigma | 0.15865 |
| -2 sigma | 0.02275 |
| -3 sigma | 0.00135 |
| -4 sigma | 0.0000315 |
| -5 sigma | 0.000000285 |
| -6 sigma | 0.000000001 |

**Table B.1:** Conversion of a one sided confidence level to a Gaussian standard deviation [9].

# Appendix C

# Event selections in final states other than hZ four jets

## C.1 Introduction

Although the four jet final state is the most powerful channel in the search for the SM Higgs boson the presence of other distinct final states allows a significant increase in sensitivity to the signal hypothesis. Indeed the combination of the complementary final states to the four jets is found to provide approximately the same amount of statistical power as the four jets analysis itself. Each final state analysis is designed to search for the hZ decay signature from a distinct decay topology. The different topologies are defined by the nature of the hZ decay system. The dominant decay modes of the h and Z boson are summarised in table C.1 and it is from combinations of these h and Z decays that the search channels are formed.

| h decay | Z decay |
|---|---|
| $b\bar{b}$ (75%) | $q\bar{q}$ (70%) |
| $\tau^+\tau^-$ (7.4%) | $\nu\bar{\nu}$ (20%) |
| | $\ell^+\ell^-$ (6.6%) |
| | $\tau^+\tau^-$ (3.3%) |

**Table C.1:** The decays of the h ($m_{\mathrm{h}} = 114\,\mathrm{GeV}/c^2$) and Z bosons from the Higgsstrahlung process. The branching fractions are shown in brackets.

The Higgs boson itself decays predominantly to a $b\bar{b}$ pair with the majority of the remaining fraction decaying to $\tau^+\tau^-$. In the following analysis it is generally assumed that the Higgs boson decays to $b\bar{b}$ and the decay of the Z boson is used to define the specific topology for the search channel. However some decay analysis are indeed sensitive to the $h \rightarrow \tau^+\tau^-$ decay and this leads to an increase in their sensitivity.

The analyses which are combined to form the ALEPH Higgs search are

- The four jet final state $(hq\bar{q})^{\dagger}$ ($\approx$ 52.5%)
- The missing energy final state $(h\nu\bar{\nu})^{\dagger}$ ($\approx$ 15%)
- The leptonic final state $(h\ell^+\ell^-)$ where $\ell = e, \mu$ ($\approx$ 5%)
- The tau lepton final states ($h\tau\tau$ and $h \rightarrow \tau\tau$, $Z \rightarrow q\bar{q}$) ($\approx$ 8%)

where the approximate branching ratios from the Higgsstrahlung system are noted.

The two streams, cuts based and Neural Network based, are formed from the above four final state analyses where final states marked with a $^{\dagger}$ have alternate analyses based on NN's and cuts and the searches are identical in both streams for the leptonic and tau final states.

In this chapter the analysis used to search for the Higgs boson in channels other than four jets are briefly described. The described analyses are those which are combined to form the cuts based search. In addition the analysis for the search $hA \rightarrow b\bar{b}b\bar{b}$ within the MSSM is also described due to its overlap with the SM four jets search and the subsequent treatment of this overlap.

## C.2 The $h\nu\bar{\nu}$ final state

The $h\nu\bar{\nu}$ final state [70] [91], known as the missing energy final state, searches for Higgs boson production via Higgsstrahlung in which the Z boson subsequently decays invisibly to two neutrinos. The $h\nu\bar{\nu}$ final state is characterised by a large amount of missing mass $\not{M}$ and the presence of two b-tagged jets from the decay of the Higgs boson. The production of the Higgs boson via WW fusion produces the same final state and its positive interference enhances the number of expected signal events within this channel. Although the production cross section for WW fusion is much smaller than that for Higgsstrahlung its relative importance increases with Higgs boson mass and so is more important for searches at the edge of our sensitivity.

### C.2.1 Pre-Selection

The $h\nu\overline{\nu}$ pre-selection is designed to reduce the $\gamma\gamma$ background and remove events with very different topologies to that of the $h\nu\overline{\nu}$ signal.
Initially the pre-selection is applied to select hadronic events consistent with the signal hypothesis. Five or more good charged tracks are required with the total energy from all charged particles within the event being greater than $0.1\sqrt{s}$. The event is then divided into two hemispheres by a plane which is constructed perpendicular to the thrust axis of the event. Selected events are required to have energy deposited in each of the two hemispheres. The contamination from $\gamma\gamma$ events is reduced by requiring events to have $E_{30}$ greater than $0.25\sqrt{s}$ or $\not{P}t$ greater than $0.05\sqrt{s}$, where $E_{30}$ is the energy deposited more than $30^{\circ}$ away from the beam axis and $\not{P}t$ is the transverse component of the missing momentum. To reject $q\overline{q}(\gamma)$ events in which an energetic ISR photon is radiated, undetected, along the beam pipe the longitudinal component of the missing momentum $\not{P}z$ is required to be small $|\not{P}z| < 50$ GeV/c. To conform with the hypothesis of two missing neutrinos from the Z boson decay the missing mass, $\not{M}$, is required to be large $\not{M} > 50$ GeV/$c^2$. Upon the application of all pre-selection cuts the corresponding selection efficiency for a 95 GeV/$c^2$ Higgs boson is of the order of 85% with the dominant backgrounds formed from $e^+e^- \rightarrow W^+W^-$ and $e^+e^- \rightarrow q\bar{q}$.

### C.2.2 Selection

The cuts based $h\nu\overline{\nu}$ event selection is designed to further reduce the background associated with $q\bar{q}$, WW, We$\nu$ and Zee events. The event selection method applies cuts to reduce each of these background sources whilst attempting to maintain efficiency for signal selection. The application of a final b-tagging cut is undertaken to ensure that the events are consistent with the $h \rightarrow b\bar{b}$ decay associated with this analysis.

A large fraction of the $q\bar{q}$ events passing through pre-selection contain one or more undetected photons from initial state radiation. Much of this remaining background is removed by requiring that the missing momentum within the event should point away from the beam axis $\theta_{\not{p}_t} > 35^\circ$. In contrast to $q\bar{q}$ background events a large fraction of $h\nu\overline{\nu}$ signal events have acoplanar jets associated with them. To utilise this the modified acoplanarity is defined as

$$\widetilde{A} = (\hat{j}_1 \times \hat{j}_2) \cdot \hat{z} \tag{C.1}$$

where $\hat{j}_{1,2}$ are the unit vectors representing the normalised total momentum of each hemisphere and $\hat{z}$ is the unit vector along the beam direction. Selected events are required to have $\tilde{A} < 0.08$. Figure C.1a shows the distribution of the event acoplanarity.

The WW background remaining after the pre-selection consists predominantly of semileptonic decays in which one W decays into hadrons whilst the other decays into a tau lepton and the associated neutrino. These events may be rejected if the tau is well isolated and either decays leptonically or is sufficiently isolated and energetic. Two specific cuts are defined to reject this background, $E_{iso} < 8$GeV and $\alpha_{iso} < 25^\circ$. Here $E_{iso}$ is the sum of the energy within $30^\circ$ of the most energetic identified lepton and the isolation angle $\alpha_{iso}$ is the angle from the most energetic lepton to its nearest neighbour.

Events originating from $We\nu$ and Zee in which an energetic lepton is deflected at low angle into the detector are rejected by requiring the energy deposited within $12^\circ$ of the beam axis to be small, $E_{12^\circ} < 0.012\sqrt{s}$. Finally the cut on the missing mass as applied in the pre-selection is tightened to $\not{M} > 70$ GeV$/c^2$.

The selected events are then clustered into two or more jets using the DURHAM jet clustering algorithm with a $y_{cut}$ of 0.015. The b-tag associated with each jet is then found using the 6VNN (section 3.6). The two jets with the largest b-tag are then taken to be the products of the $h \to b\bar{b}$ decay and a cut is applied to ensure that the decay is sufficiently b-like $\sum_{i=1}^{2} \eta_i > 1.3$.

After the events selection is applied the most significant background remaining is from $ZZ \to b\bar{b}\nu\overline{\nu}$ and accounts for approximately 50% of the observed

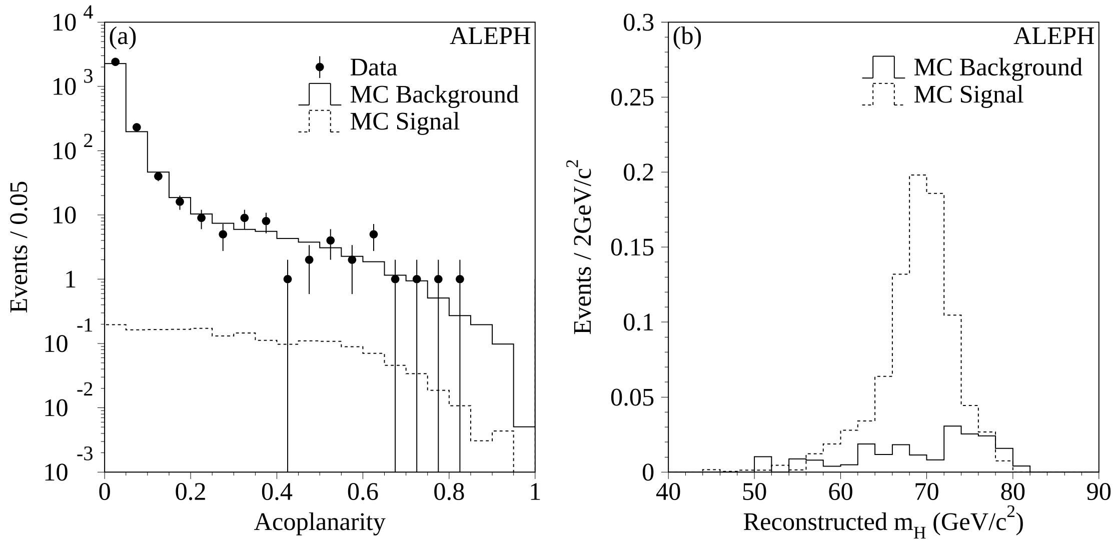


**Figure C.1:** The distributions of the event acoplanarity (a) for data(points), background(solid) and signal(dashed) and the reconstructed Higgs boson mass (b) for background(solid) and signal(dashed). All events selection criteria are applied except the visible mass cut.

background. Typical selection efficiencies are of the order of 35% for a Higgs boson of mass 95 GeV$/c^2$. Figure C.1b shows the distribution of the reconstructed Higgs boson mass.

To improve the rejection of three jet events from $\mathrm{q\overline{q}}g(\gamma)$, $\mathrm{q\bar{q}}\gamma\gamma$ a jet clustering algorithm is applied to form three jets [92]. Events from $\mathrm{q\bar{q}}g\gamma$ are removed by cuts on the minimum angle and minimum distance between two jets. Three jet events originating from the $\mathrm{q\overline{q}}\gamma(\gamma)$ process with a photon observed within the detector are removed if any of the three jets is predominantly electromagnetic.

## C.3 The $\mathrm{h}\ell^+\ell^-$ final state

The $\mathrm{h}\ell^+\ell^-$ final state [70] [93] concerns the decay of the hZ system in which the Z boson decays leptonically to either a $\mathrm{e^+e^-}$ or $\mu^+\mu^-$ pair. The same final state is produced by the ZZ fusion process which interferes negatively with the s-channel Higgsstrahlung process. The contribution of ZZ fusion is however small compared to the hZ process. The small branching ratio for this final state is countered by the fact that the experimental signature is very clear and the Higgs boson mass can be reconstructed with very good resolution due to the clean and well measured lepton candidates. Some sensitivity is also gained by including the $\tau^+\tau^-\ell^+\ell^-$ final state in which the Higgs boson decays to a pair of tau leptons.

The event selection proceeds by attempting to find pairs of leptons[1] from which hypothetical Z boson are reconstructed. The recoil mass is then calculated given the known centre of mass energy and the measured energy and momentum of the two lepton candidates. To achieve the good mass resolution and high efficiency which enhance the power of this channel much attention is paid to the correct identification of the two leptons, accounting for bremsstrahlung radiation for electrons and the identification of final state radiation (FSR) photons, from the Z boson decay, which are observed within the detector. The decay of the Higgs boson itself is only used to provide discriminating information such as the identification of tau decays and b jets from b-tagging. The mass of the Higgs boson decay system is not in itself directly measured.

## C.3.1 Selection

Events are initially required to have at least four good charged tracks with $|cos\theta|$ $< 0.95$. The total energy of the charged tracks is required to be larger than $0.1\sqrt{s}$. All possible combinations of oppositely charged lepton pairs within the event are then found and accepted pairs are then compared to the $m_Z$ hypothesis. To be accepted the lepton pair must have at least one identified lepton, (e or $\mu$), and may have one unidentified charged track. These unidentified charged tracks must have an isolation angle larger than $10^\circ$ where the isolation is defined as the half angle of the cone around the track that contains 5% of the energy of all other particles within the event. Furthermore, no mixed e - $\mu$ pairs are considered.

To improve the reconstruction of the $m_Z$ system several possibilities are considered. Firstly, to account for any possible bremsstrahlung photons from the lepton, neutral tracks within $2^\circ$ of a lepton candidate are excluded from the isolation calculation. In the case of an isolated electron candidate with a significant

[1]Within the $h\ell^+\ell^-$ selection the term leptons represents either the electrons or muons.

amount of detected bremsstrahlung the neutral objects within 2° of the electron are combined with the electron. Secondly a search for FSR photons from the decay of the Z boson is performed. Energy flow photons are considered as FSR candidates if their energy is larger than 2 GeV and their isolation is larger than 10° with $|cos\theta| < 0.95$. In the case of FSR photons the calculation of the isolation excludes the two lepton candidates and any bremsstrahlung associated with these. This consideration ensures that the FSR photon is considered as isolated only from the hadronic jets within the event but not necessarily from the two lepton candidates. Figure C.2 shows the distribution of the reconstructed Z boson mass with and without the applied Bremsstrahlung and FSR corrections.

Additionally, to reject events with an energetic photon from a radiative return to the Z, the most energetic isolated photon must have an energy which is less than 75% of the value of the most probable energy of a radiative return photon. The most probable ISR photon energy is given by

$$\gamma_{peak} = \frac{\sqrt{s}}{2} - \frac{m_Z^2}{2\sqrt{s}} \tag{C.2}$$

with the most probable energy for an ISR photon at a centre of mass energy of 206 GeV being 82.8 GeV.

The $\ell^+\ell^-(\gamma)$ system with an invariant mass closest to $m_Z$ is chosen as the Z boson decay system. The reconstructed Z boson mass is required to be larger than 77.5 GeV/$c^2$ or the combined masses of the hZ system $m_Z + m_h > 77.5$ + Threshold, where Threshold = ($\sqrt{s}$ - 91.2 GeV/$c^2$) [94].

After the selection of the lepton pair the remaining particles within the event are clustered into two jets using the DURHAM algorithm. To reject $e^+e^- \rightarrow q\bar{q}$ events, where the leptons are present within the jets themselves, the sum of the transverse momenta of the leptons with respect to their nearest jet is required to be greater than 20 GeV/c. In this case the nearest jet is defined as that which combines with the lepton to form the smallest invariant mass. To reject events such as $e^+e^-Z^*$ and $ZZ^{(*)}$ with a low mass $Z^*$ but high mass reconstructed Z boson, the invariant mass of the two hadronic jets must be larger than 15 GeV/$c^2$.

Events which have exactly four charged tracks are taken as candidates for $\tau^+\tau^-\ell^+\ell^-$ in which $h \rightarrow \tau^+\tau^-$. Signal events from this hypothesis are expected

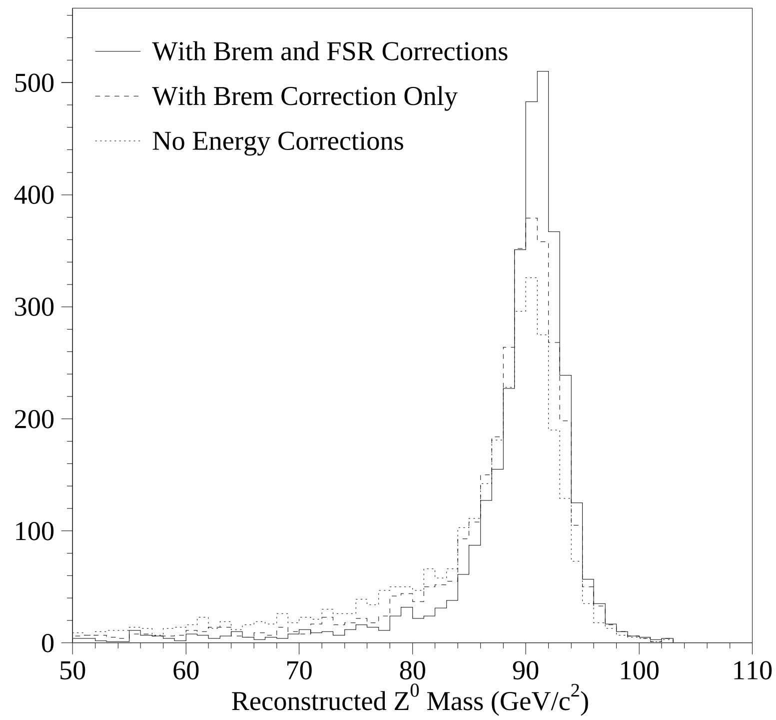


**Figure C.2:** The reconstructed Z mass for signal Monte Carlo events showing the effect of the Bremsstrahlung and final state radiation corrections [93].

to have a non-negligible amount of missing energy due to the loss of the neutrinos produced in the tau decays. Consequently a cut is placed that requires the missing energy should be at least $0.1\sqrt{s}$.

The background from semileptonic W pair decays ($WW \rightarrow \mathrm{q\bar{q}}'\ell\nu$) is rejected by explicitly reconstructing the two W bosons when one of the leptons is identified but the other is an unidentified but isolated charged particle. In this hypothesis the identified lepton and the missing four momentum (which is assigned to the neutrino) are assigned to the leptonic W decay. The unidentified charged track together with the remaining energy flow objects is assigned to the hadronic W decay. These events are then rejected if the two reconstructed W decays are compatible with the WW hypothesis. Events are thus rejected if $m_W^{hadronic} + m_W^{leptonic} < 150\,\mathrm{GeV}/c^2$ and $|m_W^{hadronic} - m_W^{leptonic}| < 20\,\mathrm{GeV}/c^2$. Background events originating from $\mathrm{Z}\gamma^{(*)}$ (where the low mass gamma decays to a pair of leptons) are eliminated by finding the leptons which originate from the

$\gamma^*$ and requiring that the sum of the lepton pair invariant mass and its recoil mass be greater than 115 GeV/$c^2$. Background from $\ell^+\ell^-\gamma\gamma$ events is removed by requiring both jets of the recoil system to contain at least one charged particle.

## C.4 Final states with $\tau^+\tau^-$ pairs

The $\tau^+\tau^-$Z analysis [70] [95] is designed to detect final states with tau leptons. In the majority of cases this final state will contain two tau leptons and two hadronic jets from either of the h $\rightarrow$ b$\bar{\text{b}}$,Z $\rightarrow \tau^+\tau^-$ or h $\rightarrow \tau^+\tau^-$,Z $\rightarrow$ q$\bar{\text{q}}$ decay processes. Some sensitivity also exists to the four tau final state arising form hZ $\rightarrow \tau^+\tau^-\tau^+\tau^-$. The final state with tau leptons is common to both SM hZ decays and also MSSM hA decays, as hA can decay to $\tau^+\tau^-$b$\bar{\text{b}}$. The analysis proceeds with a common pre-selection followed by two specific event selection procedures to search for hA and hZ decays.

### C.4.1 Pre-Selection

The events are initially required to be hadronic in origin. Events are required to have at least eight good charged tracks with a total charged energy exceeding $0.2\sqrt{s}$. Backgrounds from WW and ZZ are reduced by rejecting events which have an identified lepton with energy greater than $0.25\sqrt{s}$. Radiative returns to the Z, which characteristically have high missing energy, $\not{E}$, and high missing longitudinal momentum, $\not{P}$z, are rejected by requiring $|\not{P}\mathrm{z}| + \not{E} < 1.8\gamma_{peak}$ and $|\not{P}\mathrm{z}| < 0.60\gamma_{peak}$. Here $\gamma_{peak}$ refers to the most likely energy of an emitted ISR photon (equation C.2). Typically $\gamma_{peak}$ is of the order of 80 GeV for a centre of mass energy in the region of 200 GeV. Events are then clustered into mini-jets using the DURHAM algorithm, with each mini-jet constrained to have an invariant mass smaller than 2.7 GeV/$c^2$ to remain consistent with the tau hypothesis. The tau candidates are selected from the mini-jets using a series of quality cuts based on charged track multiplicity, isolation and momentum. Fully leptonic tau candidates are also included in the selection. Events in which two or more tau candidates are found are then treated further. The two tau candidates are paired together and the remainder of the event is clustered into two jets again using the

DURHAM algorithm. Events in which more than two tau candidates are found will indeed have several different clustering hypotheses, all of which are considered. All four jets within the event are rescaled using a kinematic consistency fit with the jet directions fixed and the tau candidates fixed to $m_\tau$. A $\chi^2$ fit estimator is then calculated based on energy-momentum conservation, hadronic jet resolutions and the compatibility of the di-jet invariant masses with the associated final state. The fit estimator calculation is dependent on the final state hypothesis in that two states exist for hZ , namely $\mathrm{h}\tau^+\tau^-$ and $\tau\tau Z$, and one further for hA. In the case of the hA hypothesis an extra term is added to the $\chi^2$ estimator comparing the di-jet masses of the $\tau^+\tau^-$ and hadronic systems assuming $m_\mathrm{h} = m_\mathrm{A}$ for this hypothesis. In the case of $\mathrm{h}\tau^+\tau^-$ the di-jet mass of the $\tau^+\tau^-$ system is compared to the Z mass $m_\mathrm{Z}$ and similarly in the case of $\tau^+\tau^-\mathrm{q\bar{q}}$ the hadronic system invariant mass is compared to $m_\mathrm{Z}$. Events failing the kinematic fit are rejected. From the several possible combinations of the tau mini-jet candidates the combination with the smallest $\chi^2$, thus the most compatible with the event hypothesis, is selected. Figure C.3a shows the $\chi^2$ fit variable for the $\mathrm{h} \to \tau^+\tau^-, \mathrm{Z} \to \mathrm{q\bar{q}}$ channel at the pre-selection level.

### C.4.2 Selection

Two Neural Networks (NN) are used to discriminate between $\mathrm{h}\tau^+\tau^-$ events, $\tau^+\tau^-\mathrm{q\bar{q}}$ events and background events. Each NN uses the following four variables as inputs. The kinematic fit estimator $\chi^2$, the event transverse momentum, the sum of the two tau mini-jet isolation angles and the sum of the fitted transverse momenta $P_t^{jet}$ of the tau mini-jets with respect to the nearest hadronic jet. The isolation angle of each tau mini-jet is defined as the half angle of the largest cone around the mini-jet direction which contains no more than 5% of the total energy outside the cone. To take advantage of the large probability for the Higgs boson to decay to a $\mathrm{b\bar{b}}$ system the $\mathrm{h}\tau^+\tau^-$ NN uses the sum of the NN b-tagging output of the two hadronic jets as a fifth input variable.

The NN associated with the hA event selection is formed from the same five variables as the $\mathrm{h}\tau^+\tau^-$ NN for the hZ analysis. Events in which the NN output exceeds 0.826 are selected as hA candidates.

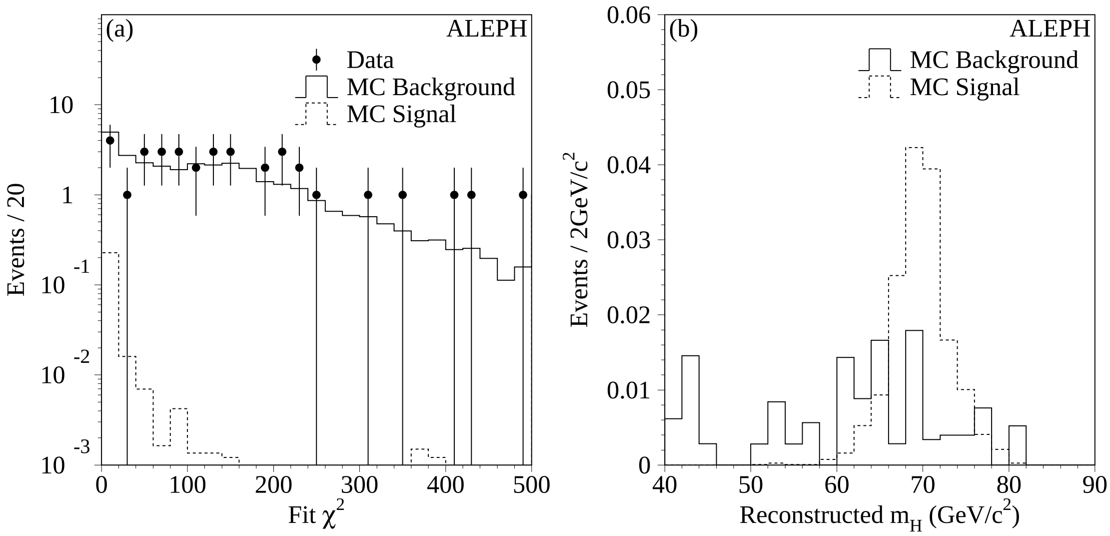


**Figure C.3:** The distributions of (a) the $\chi^2$ fit variable for the $h \to \tau^+\tau^-, Z \to q\bar{q}$ channel at the pre-selection level with data(points), background(solid) and signal(dashed) and (b) the fitted Higgs boson mass for both channels at the final selection level. With relaxed cuts on the Higgs boson mass.

Some hZ events may be selected by more than one NN. In this case the hypothesis with the largest NN output is taken as the correct interpretation. However if the sum of the two NN outputs is greater than 1.8 the difference between the two is assumed to be insignificant. In this case the kinematic fit estimator $\chi^2$ is used to determine which interpretation is taken. An offset is introduced in the $\chi^2$ cut to favour the classification as $h\tau^+\tau^-$. This offset is introduced since a high NN output in the $h\tau^+\tau^-$ hypothesis requires a high b-tagging content within the event. This additional information in the $h\tau^+\tau^-$ NN is taken to give the event a more signal like quality if the NN outputs are similar. Once the event hypothesis is defined a candidate is selected if the NN value exceeds 0.965. Figure C.3b shows the distribution of the reconstructed Higgs boson mass for both channels at the final selection level of the analysis.

## C.5 The $hA \to b\bar{b}b\bar{b}$ final state

The $hA \to b\bar{b}b\bar{b}$ final state [96] is very similar in topology to the $hZ \to b\bar{b}q\bar{q}$ final state, indeed part of the hZ four jets analysis is designed specifically to select

$b\bar{b}b\bar{b}$ final states. The main difference between the four jet final states from hZ and hA is indeed the enhanced b quark content of the hA decay and also the equal mass constraint $m_{\mathrm{h}} = m_{\mathrm{A}}$ associated with the MSSM hA hypothesis. The analysis proceeds with a loose pre-selection designed to select hadronic events and suppress $q\bar{q}\gamma$ events and then a final event selection based on an optimised linear discriminant.

### C.5.1 Pre-Selection

To select hadronic events at least eight good charged tracks are required with a total charged energy greater than $0.1\sqrt{s}$. Events originating from $q\bar{q}\gamma$ in which an ISR photon is observed within the detector are rejected by cutting on the fraction of the jet energy contained within a $1^\circ$ cone around any object within that jet. Events are rejected if in any of the four jets an object is found which has more than 80% of the jet energy within this $1^\circ$ cone. Events in which a photon escapes undetected down the beam pipe are rejected by requiring $\not{P}z$ < 1.5 ($m_{vis}$ - 90), where $m_{vis}$ is the total visible mass in the event and $\not{P}z$ is the missing longitudinal momentum. The events are then clustered into four jets using the DURHAM algorithm and events with $y_{34}$ <0.001 are rejected. The event thrust must be less than 0.9 and the smallest inter-jet angle $\theta_{ij}^{min}$ measured over all jets $i$ and $j$ must be larger than $20^\circ$.

### C.5.2 Selection

The final event selection combines information about the event topology as well as the b quark content of the four jets. A four variable NN b-tagger is used [97] with three of the four inputs being in common with the 6VNN (section 3.6) used in the SM four jets search and one new variable. The three common variables are: $\mathcal{P}_{jet}$ the confidence level of a jet being a light quark jet based upon the impact parameters of the tracks in the jet, the $\chi^2$ difference, $\Delta\chi^2$, between the fit assuming that all tracks originate form the primary vertex (IP) and a fit which assumes a secondary vertex exists and the transverse momentum $p_{\mathrm{t}}$ of the identified leptons with respect to the jet axis. The fourth and new variable is the scaled inclusive $X_e$,

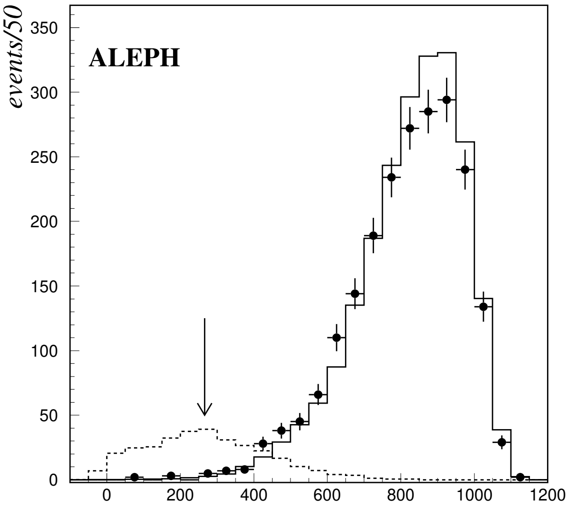


**Figure C.4:** Distribution of $\mathcal{F}$ for data (points), background (solid) and signal (dashed).

where $X_e$ is defined as the fraction of the jet energy carried by the most energetic particles which have a total invariant mass smaller than 2.1 GeV/$c^2$.

The main variable used within the hA four jet analysis is a linear discriminant defined as

$$\mathcal{F} = 300 \times (4 - \sum_{j=1}^{4} \eta_j) - \theta_{ij}^{min} \tag{C.3}$$

Here $\eta_j$ is the output of the 4VNN for the $j^{\text{th}}$ jet and $\theta_{ij}^{min}$ is the minimum inter-jet angle in degrees. Figure C.4 shows the distribution of $\mathcal{F}$ for background, data and a possible signal [70].

An additional cut is applied to reduce the number of events originating from the $q\overline{q}(g)$ process in which one jet is found recoiling against two more energetic jets. The variable $\Delta\theta_3$ is designed to reduce the background from these events. For each of the four possible three jet pairings within an event the variable $\Delta_{ijk}$ is calculated where

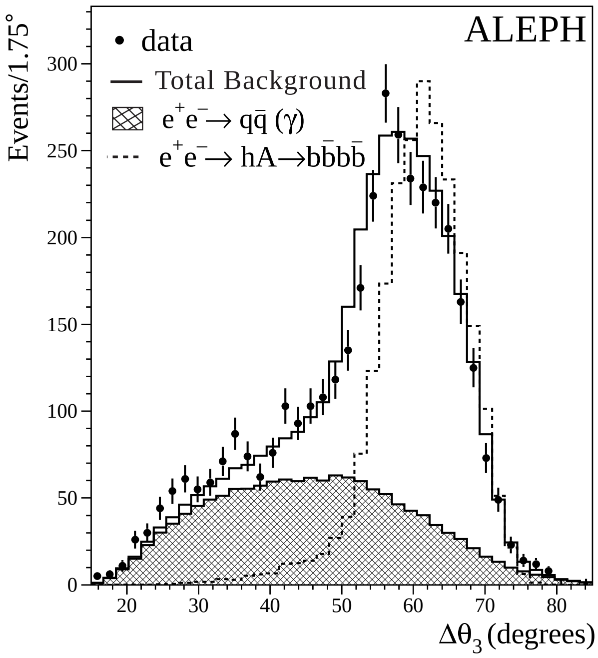


**Figure C.5:** Distribution of the minimum value of the energy-weighted angular dispersion, $\Delta\theta_3$, for data(points), background(solid) with the $q\bar{q}(\gamma)$ component of the background highlighted. The signal(dashed) is displayed with an arbitrary normalisation.

$$\Delta_{ijk} \equiv \sqrt{\frac{\sum_{l \in jet_{i,j,k}} \theta^2_{l,\vec{u}} E_l}{\sum_{l \in jet_{i,j,k}} E_l}} \tag{C.4}$$

$$\Delta\theta_3 \equiv \min(\Delta_{ijk}) \tag{C.5}$$

Here $ijk$ refers to the three jets, $\theta^2_{l,\vec{u}}$ is the angle between the momentum of the object $l$ contained in the jet $i, j, k$ and the sum of the three jet momenta $\vec{u}$ and $E_l$ is the energy of the object $l$. The variable $\Delta\theta_3$ is taken as the minimum value of $\Delta_{ijk}$ for all possible combinations. Figure C.5 shows the distribution of $\Delta\theta_3$, for background, data and a possible signal [56].

Events in which $\Delta\theta_3 > 50°$ and $\mathcal{F} < 351$ are selected as hA four jet candidates.

# Appendix D
# Analysis cross checks

## D.1 Introduction

The accurate modelling of observed data with Monte Carlo simulation is of great importance within the presented analysis. Inaccuracies in the simulations can lead to biases in the final interpretation of the experimental results. Pre-selection cuts are in general defined to eliminate any unmodeled backgrounds and provide a manageable data set for comparison with simulation. This chapter provides a comparison of data and MC simulation for various event selection variables at different selection levels from the cuts based four jet selection. Firstly the standard pre-selection level is covered and additional cuts are then alternatively applied on kinematic and b-tagging variables.

For each level of selection the number of expected and observed events is summarised and the applied selection is detailed.

The data/MC comparisons, for both variable distributions and event numbers, presented within this thesis use a subset of the available MC samples which is described in table D.1. The subset provides a very close approximation of the results which would be obtained with the use of the total available MC sample.

The $\chi^2/d.o.f$ presented for each of the data/MC comparisons is calculated using all histogram bins in which the recorded number of data events is greater than 0 and the expected number of MC events is also greater than 0.

| | ZZ | $q\bar{q}$ | WW | hZ SIGNAL (7.5k) |
|---|---|---|---|---|
| 204 | 50k | 100k | 250k | 114 GeV/$c^2$ |
| 205 | 50k | | 250k | |
| 206 | 50k | 100k | 250k | 114 GeV/$c^2$ |
| 207 | 50k | | 250k | |
| 208 | 50k | 100k | 250k | 114 GeV/$c^2$ |
| 209 | 50k | | 250k | |
| 210 | | | | 114 GeV/$c^2$ |

**Table D.1:** The MC simulation samples used during the data/MC comparisons within presented within this thesis.

## D.2 Standard pre-selection

This section presents the comparison of the observed data and MC simulation with the standard pre-selection cuts applied. Table D.2 shows the selected number of events from the data stream and also the expected number of events from MC simulation for background and a hypothesised ($m_{\rm h}$=114 GeV/$c^2$) Higgs boson signal while figures D.1, D.2, D.3 and D.4 show the data to MC comparisons for the b-tagging($\min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1-\eta_3)(1-\eta_4)$), kinematic ($\Theta, \gamma, \mathrm{m}_{12}, \mathrm{m}_{34}, \mathrm{y}_{34}, \mathrm{X}_\gamma$), linear discriminant($9.5\mathrm{y}_{34} + \sum \eta_i$) and jet pairing variables ($|\cos\theta_{h-decay}|,\ \ |\cos\theta_{z-decay}|$) respectively[1].

No significant deviations from the MC expectation are found in the number of expected events or the distributions of the considered variables.

| Event Type | Number of Events |
|---|---|
| hZ ($m_{\rm h} = 114$ GeV/$c^2$) | 7.103 |
| $q\bar{q}$ | 548.654 |
| WW | 1606.005 |
| ZZ | 133.259 |
| Total bgd | 2287.918 |
| Data | 2283 |

**Table D.2:** The number of expected signal ($m_{\rm h} = 114$ GeV/$c^2$) and background events from MC simulation and the number of observed data events for the pre-selection level of the analysis.

[1]Throughout this chapter the distributions are shown with the decay angles pairing applied.

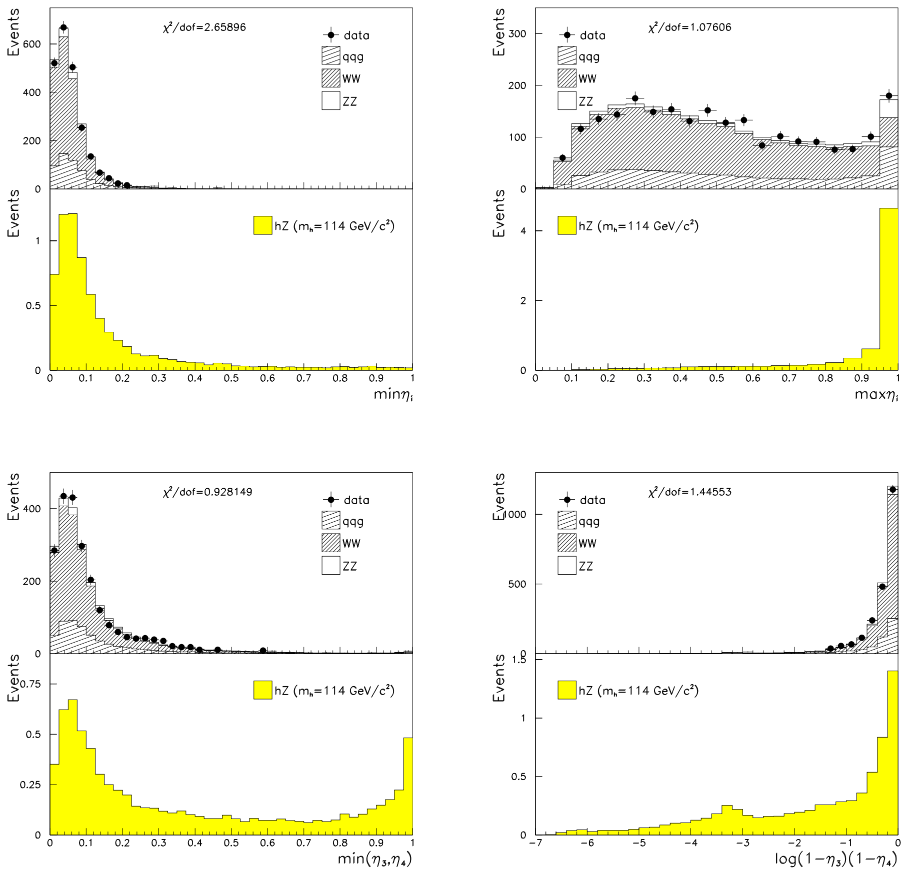


**Figure D.1:** Distributions of the b-tagging variables $(\min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1-\eta_3)(1-\eta_4))$ at the pre-selection level of the analysis for data, simulated background and a simulated Higgs signal with $m_{\mathrm{h}} = 114\ \mathrm{GeV}/c^2$.

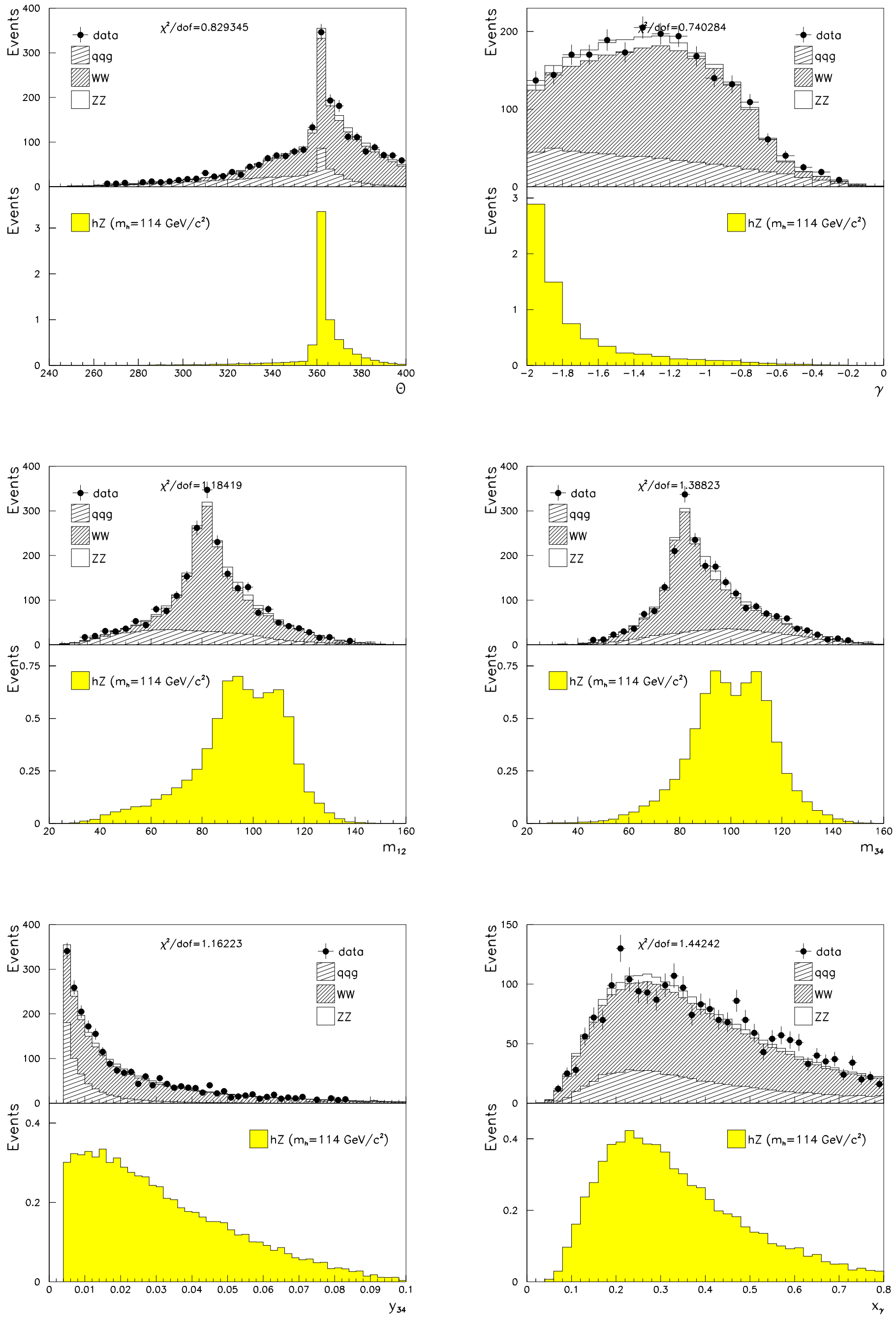


**Figure D.2:** The distributions of the kinematic variables $(\Theta, \gamma, \mathrm{m}_{12}, \mathrm{m}_{34}, \mathrm{y}_{34}, \mathrm{X}_\gamma)$ at the pre-selection level of the analysis for data, simulated background and a simulated Higgs signal with $m_\mathrm{h} = 114\ \mathrm{GeV}/c^2$.

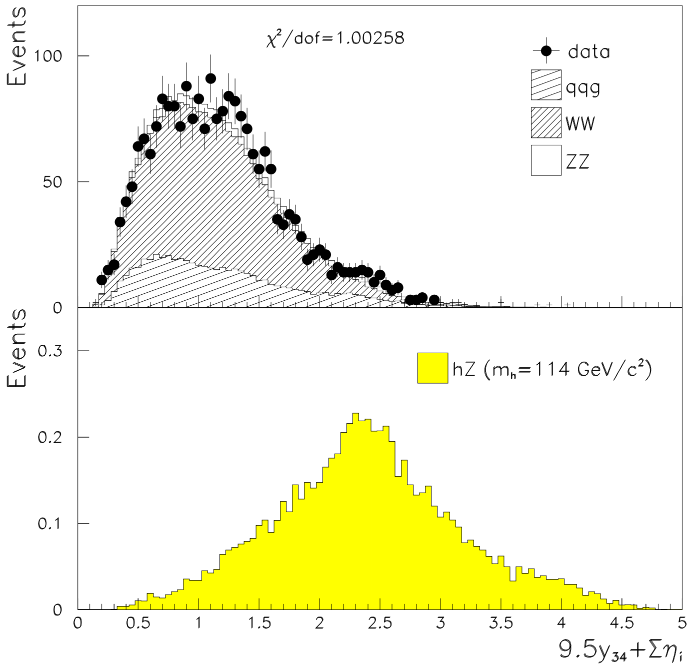


**Figure D.3:** The distributions of the linear discriminant variable ($9.5y_{34} + \sum \eta_i$), used in the 4b event selection, at the pre-selection level of the analysis for data, simulated background and a simulated Higgs signal with $m_h = 114\,\mathrm{GeV}/c^2$.

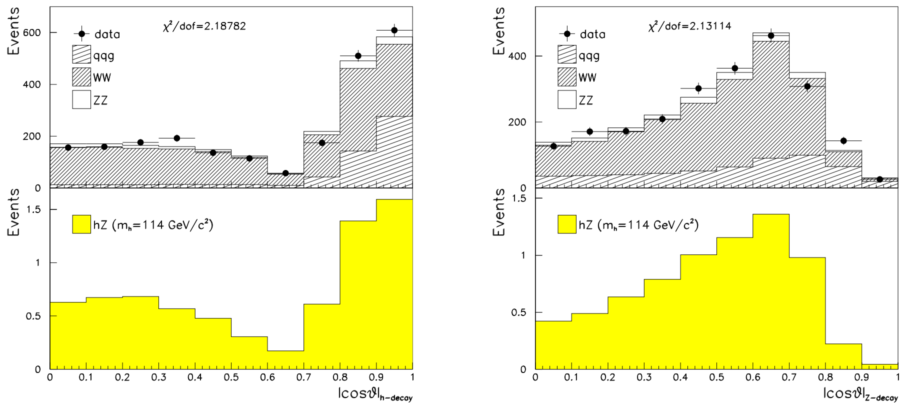


**Figure D.4:** The distributions of the h and Z decay angles ($|\cos\theta_{h-decay}|$, $|\cos\theta_{z-decay}|$) at the pre-selection level of the analysis for data, simulated background and a simulated Higgs signal with $m_h = 114\,\mathrm{GeV}/c^2$.

## D.3 Kinematics

This section presents the comparison of the observed data and MC simulation with standard pre-selection as well as with the following non-b-tagging cuts applied.

- $\gamma$ < -1.3
- $\Theta > 350^\circ$
- $m_{12} > 77$ GeV/$c^2$
- $m_{34} > 55$ GeV/$c^2$

Table D.3 shows the selected number of events from the data stream and also the expected number of events from MC simulation for background and a hypothesised 114 GeV/$c^2$ Higgs boson signal while figures D.5, D.6, D.7 and D.8 show the data to MC comparisons for various variables used within the four jets cuts analysis ($\Theta, \gamma, m_{12}, m_{34}, y_{34}, X_\gamma,\ \min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1-\eta_3)(1-\eta_4)$).

Good agreement is found between the number of expected and observed events as well as in the distributions of the various variables under consideration. This agreement shows that no bias is found in the application of the non-b-tagging cuts and that no observable inaccuracy in the modelling is present at this level of the event selection.

| Event Type | Number of Events |
|---|---|
| hZ ($m_h = 114$ GeV/$c^2$) | 6.202 |
| $q\bar{q}$ | 222.652 |
| WW | 716.199 |
| ZZ | 80.228 |
| Total bgd | 1019.079 |
| Data | 999 |

**Table D.3:** The number of expected signal ($m_h = 114$ GeV/$c^2$) and background events from MC simulation and the number of observed data events for the pre-selection level of the analysis with non-b-tagging cuts also applied.

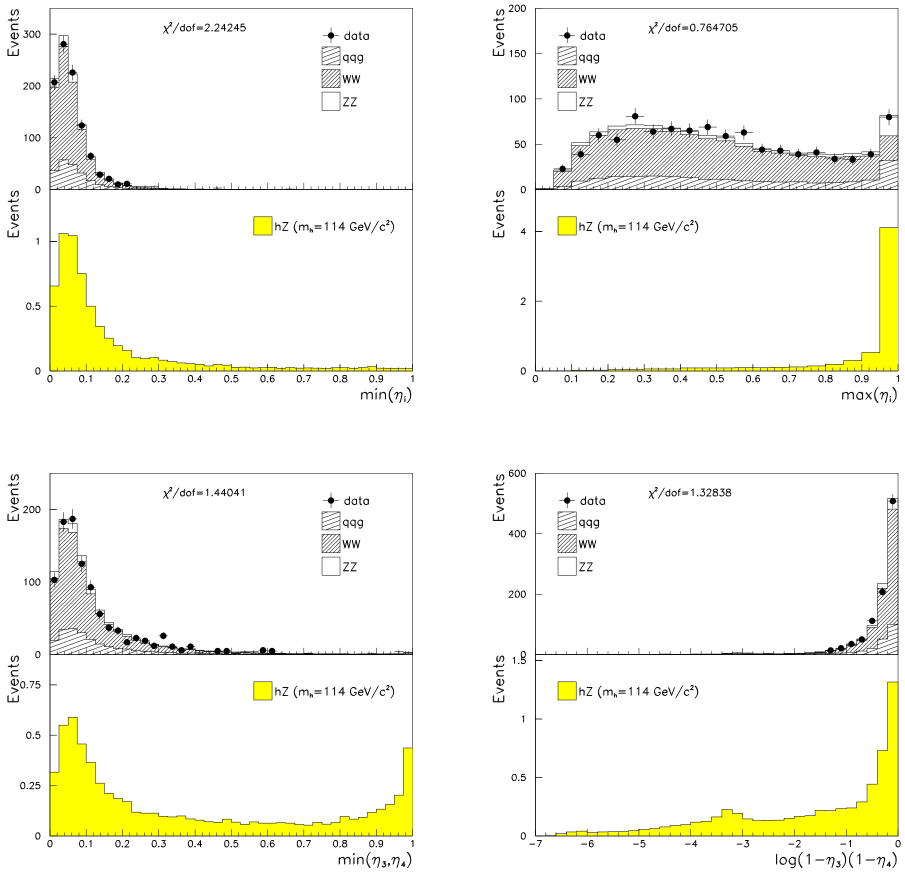


**Figure D.5:** Distributions of the b-tagging variables $(\min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1-\eta_3)(1-\eta_4))$ at the pre-selection level of the analysis with additional non-b-tagging cuts applied for data, simulated background and a simulated Higgs signal with $m_\mathrm{h} = 114\ \mathrm{GeV}/c^2$.

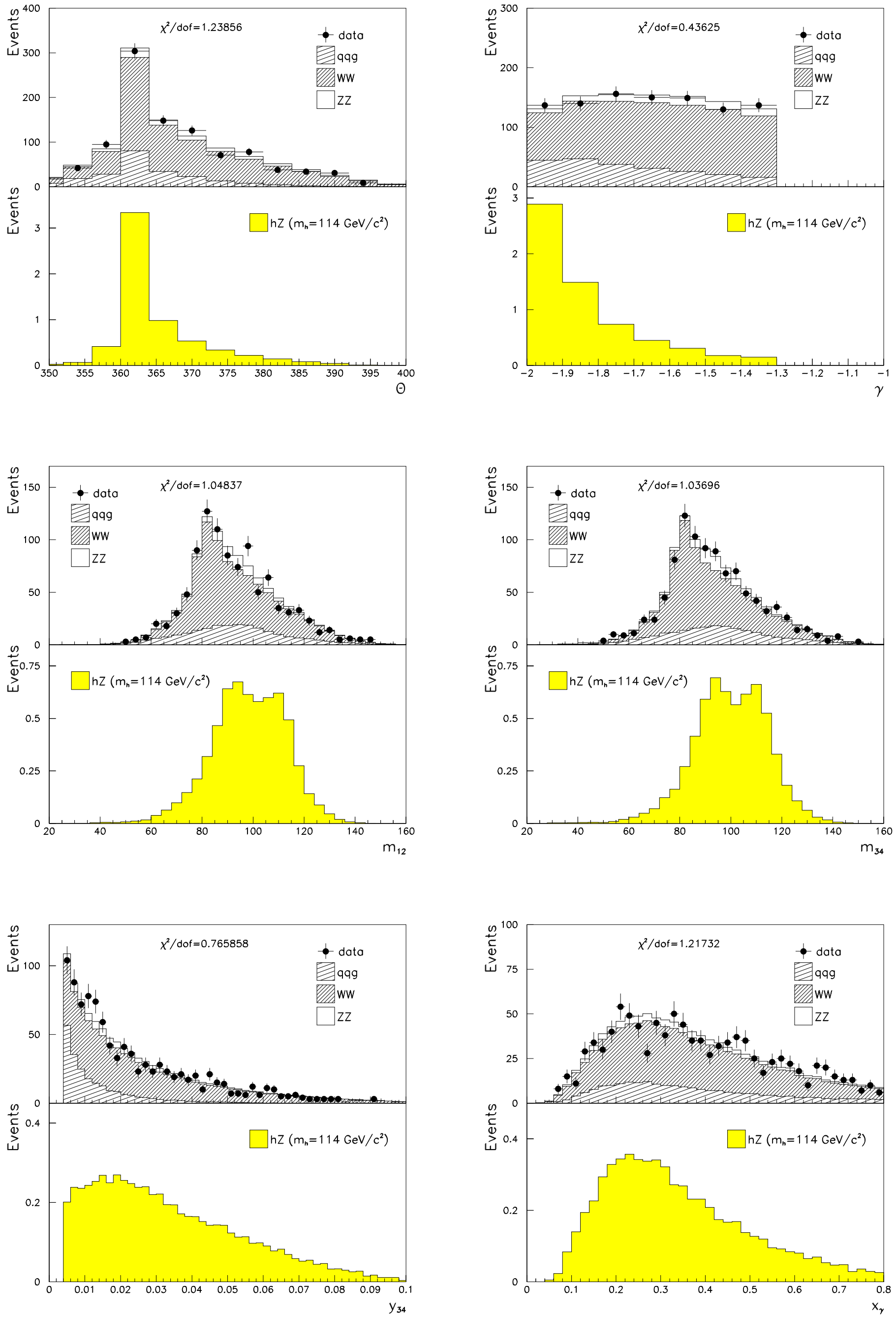


**Figure D.6:** The distributions of the non-b-tagging variables $(\Theta, \gamma, m_{12}, m_{34}, y_{34}, X_\gamma)$ at the pre-selection level of the analysis with additional non-b-tagging cuts applied for data, simulated background and a simulated Higgs signal with $m_h = 114\,\mathrm{GeV}/c^2$.

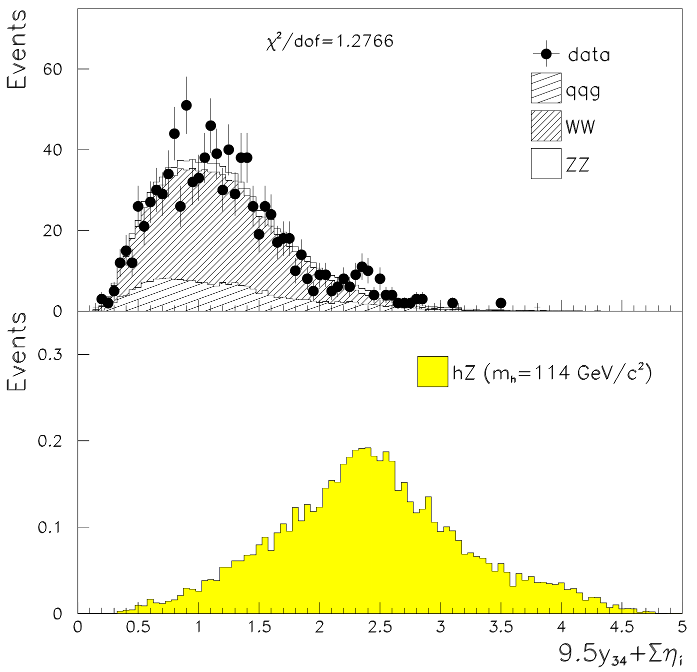


**Figure D.7:** The distributions of the linear discriminant variable ($9.5y_{34} + \sum\eta_i$) at the pre-selection level of the analysis with additional non-b-tagging cuts applied for data, simulated background and a simulated Higgs signal with $m_\mathrm{h} = 114\ \mathrm{GeV}/c^2$.

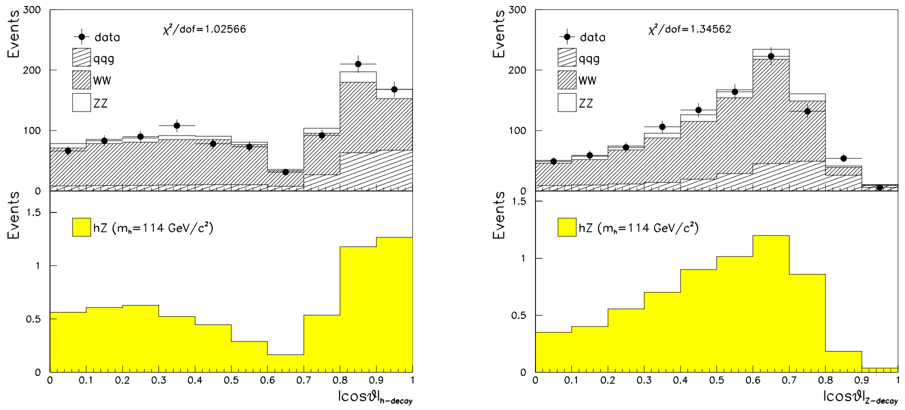


**Figure D.8:** The distributions of the h and Z decay angles ($|\cos\theta_{h-decay}|$, $|\cos\theta_{z-decay}|$) at the pre-selection level of the analysis with additional non-b-tagging cuts applied for data, simulated background and a simulated Higgs signal with $m_\mathrm{h} = 114\ \mathrm{GeV}/c^2$.

## D.4 Kinematics + anti-btag

In this section we presents the comparison of the observed data and MC simulation with standard pre-selection applied as well as the following non-b-tagging and anti-b-tagging cuts.

- $\gamma$ < -1.3
- $\Theta > 350^\circ$
- $m_{12} > 77$ GeV/$c^2$
- $m_{34} > 55$ GeV/$c^2$
- $\max(\eta_i) < 0.9$

Table D.4 shows the selected number of events from the data stream and also the expected number of events from MC simulation for background and a hypothesised 114 GeV/$c^2$ Higgs boson signal while figures D.9, D.10, D.11 and D.12 show the data to MC comparisons of the distributions of the considered variables ($\Theta, \gamma, m_{12}, m_{34}, y_{34}, X_\gamma,\ \min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1-\eta_3)(1-\eta_4)$).

The sample of events passing the anti-btag cut has no overlap with the sample of events that pass the full selection. The vast majority of events which pass the non-b-tagging cuts are rejected when the b-tagging cuts are applied. A large fraction of these events are selected by the anti-btag cut. Therefore the anti-btag sample effectively forms an orthogonal sample to the full selection level sample.

Good agreement between the observed data and the simulated MC in this orthogonal sample provides increased confidence in alternate b-tagged sample.

| Event Type | Number of Events |
|---|---|
| hZ ($m_{\mathrm{h}} = 114\ \mathrm{GeV}/c^2$) | 1.563 |
| $\mathrm{q\bar{q}}$ | 179.572 |
| WW | 662.859 |
| ZZ | 53.480 |
| Total bgd | 895.911 |
| Data | 880 |

**Table D.4:** The number of expected signal ($m_{\mathrm{h}} = 114\ \mathrm{GeV}/c^2$) and background events from MC simulation and the number of observed data events for the pre-selection level of the analysis with non-b-tagging and anti b-tagging cuts also applied.

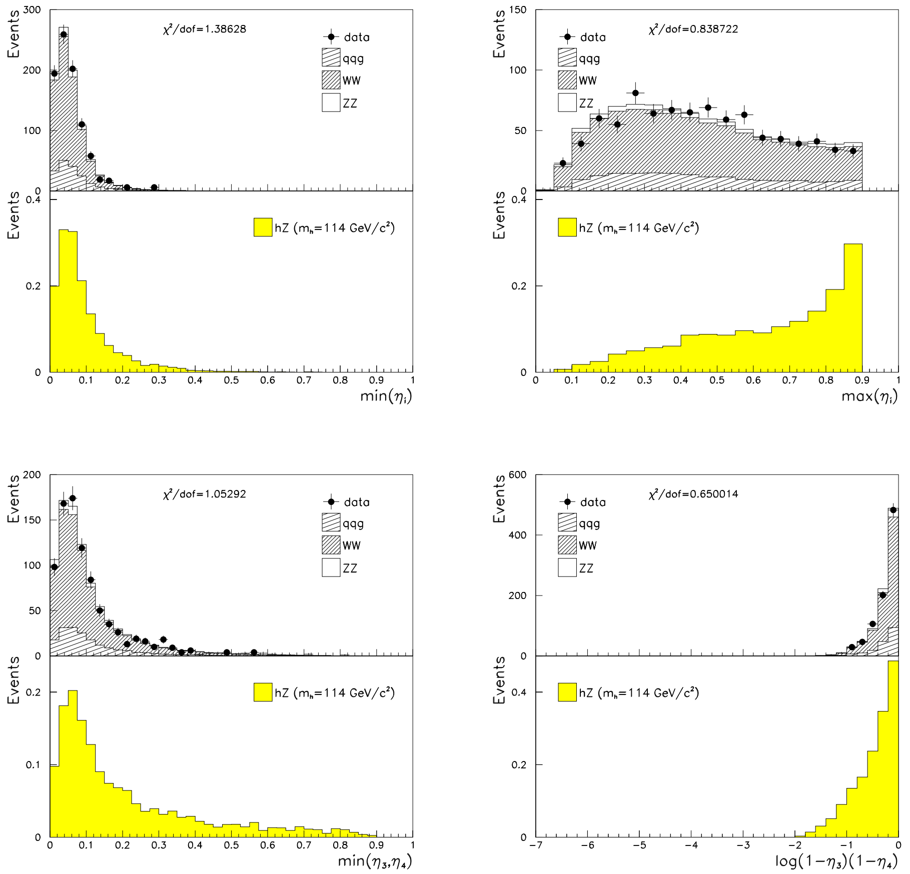


**Figure D.9:** The distributions of the b-tagging ($\min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1 - \eta_3)(1 - \eta_4)$) variables at the pre-selection level of the analysis with non-b-tagging and anti b-tagging cuts also applied for data, simulated background and a simulated Higgs signal with $m_{\mathrm{h}} = 114\ \mathrm{GeV}/c^2$.

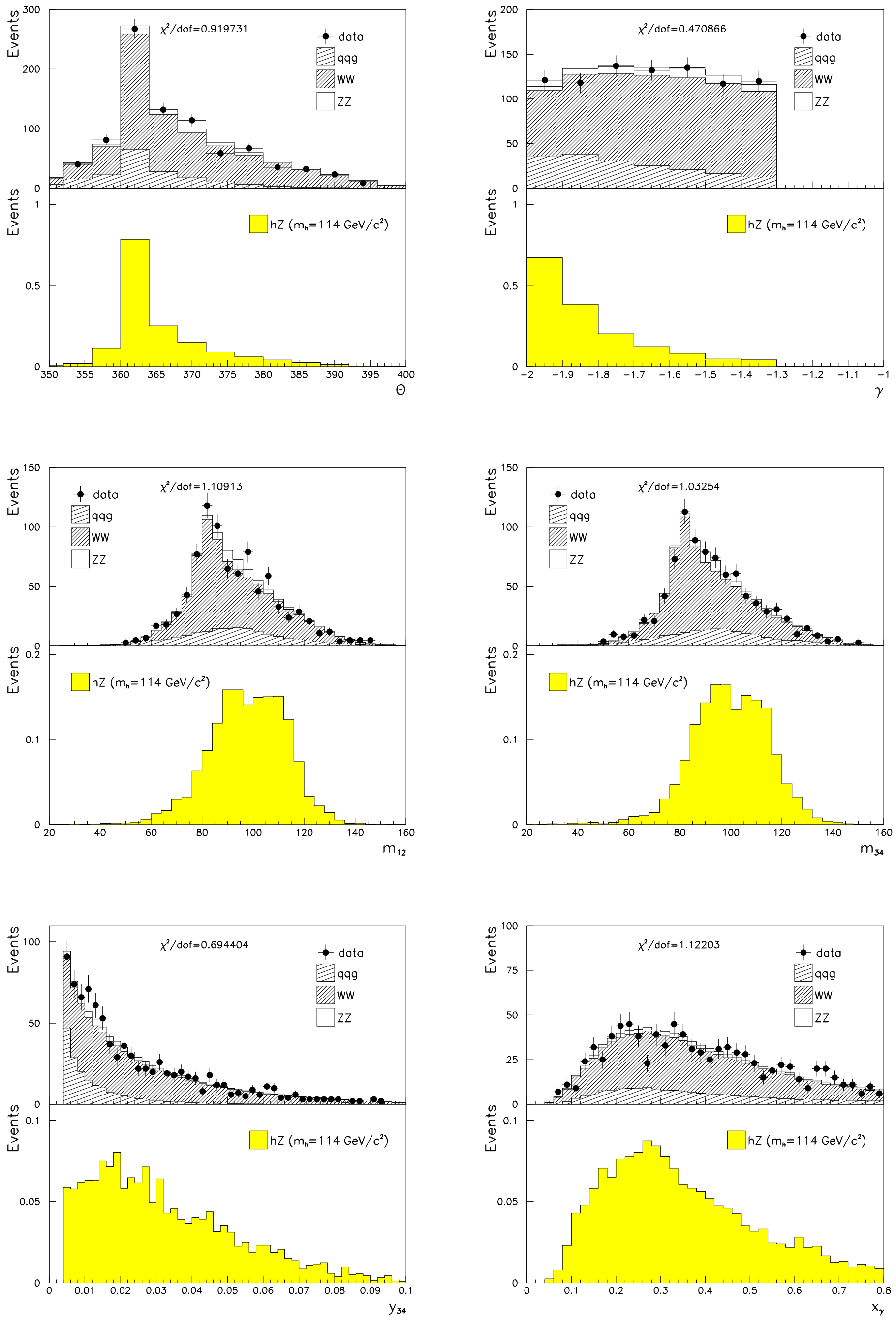


**Figure D.10:** The distributions of the non-b-tagging variables $(\Theta, \gamma, \mathrm{m}_{12}, \mathrm{m}_{34}, \mathrm{y}_{34}, \mathrm{X}_{\gamma})$ at the pre-selection level of the analysis with non-b-tagging and anti b-tagging cuts also applied for data, simulated background and a simulated Higgs signal with $m_{\mathrm{h}} = 114\ \mathrm{GeV}/c^2$.

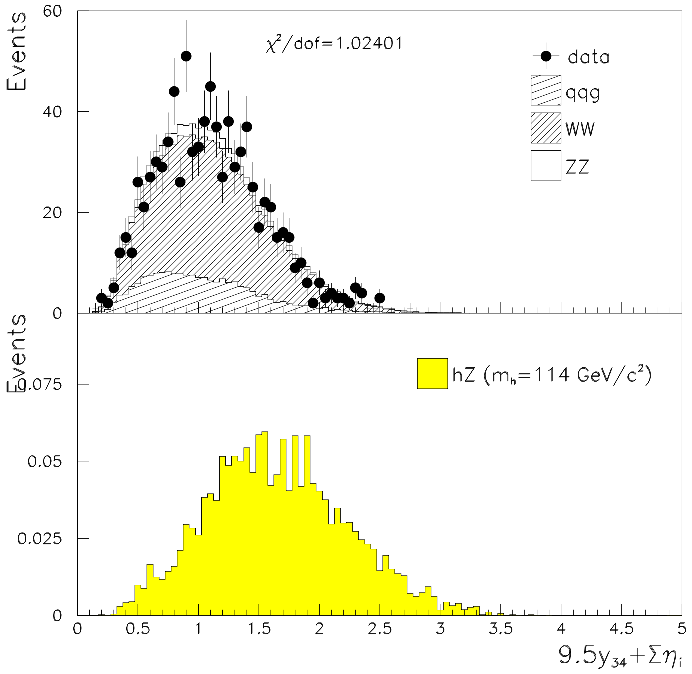


**Figure D.11:** The distributions of the linear discriminant variable ($9.5y_{34} + \sum \eta_i$) at the pre-selection level of the analysis with non-b-tagging and anti b-tagging cuts also applied for data, simulated background and a simulated Higgs signal with $m_{\rm h} = 114\ {\rm GeV}/c^2$.

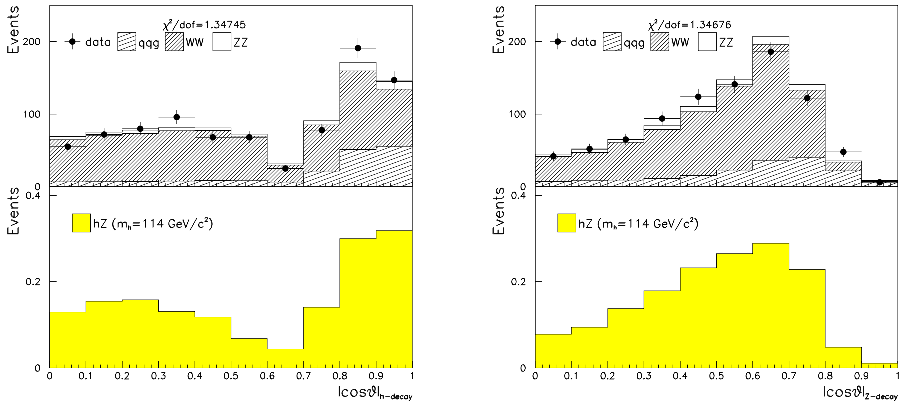


**Figure D.12:** The distributions of the h and Z decay angles ($|\cos\theta_{h-decay}|$, $|\cos\theta_{z-decay}|$) at the pre-selection level of the analysis with non-b-tagging and anti b-tagging cuts also applied for data, simulated background and a simulated Higgs signal with $m_{\rm h} = 114\ {\rm GeV}/c^2$.

# D.5 B-tagging

## D.5.1 Soft b-tag

This section presents the comparison of the observed data and MC simulation with standard pre-selection cuts applied as well as soft b-tagging cut applied to the Higgs candidate jets.

The applied cut is:

- $\min(\eta_3, \eta_4) > 0.35$

Table D.5 shows the selected number of events from the data stream and also the expected number of events from MC simulation for background and a hypothesised 114 GeV/$c^2$ Higgs boson signal while figures D.13, D.14, D.15 and D.16 show the data to MC comparisons for the considered variables.

The comparison provides a reasonably high statistics sample to test for any problems in the description of the b-tagging efficiencies and the modelling of the distributions. No significant discrepancies are found.

| Event Type | Number of Events |
|---|---|
| hZ ($m_\mathrm{h} = 114$ GeV/$c^2$) | 5.367 |
| $\mathrm{q\bar{q}}$ | 91.196 |
| WW | 268.409 |
| ZZ | 47.603 |
| Total bgd | 407.208 |
| Data | 425 |

**Table D.5:** The number of expected signal and background events from MC simulation and the number of observed data events for the pre-selection level of the analysis with soft b-tagging cuts also applied.

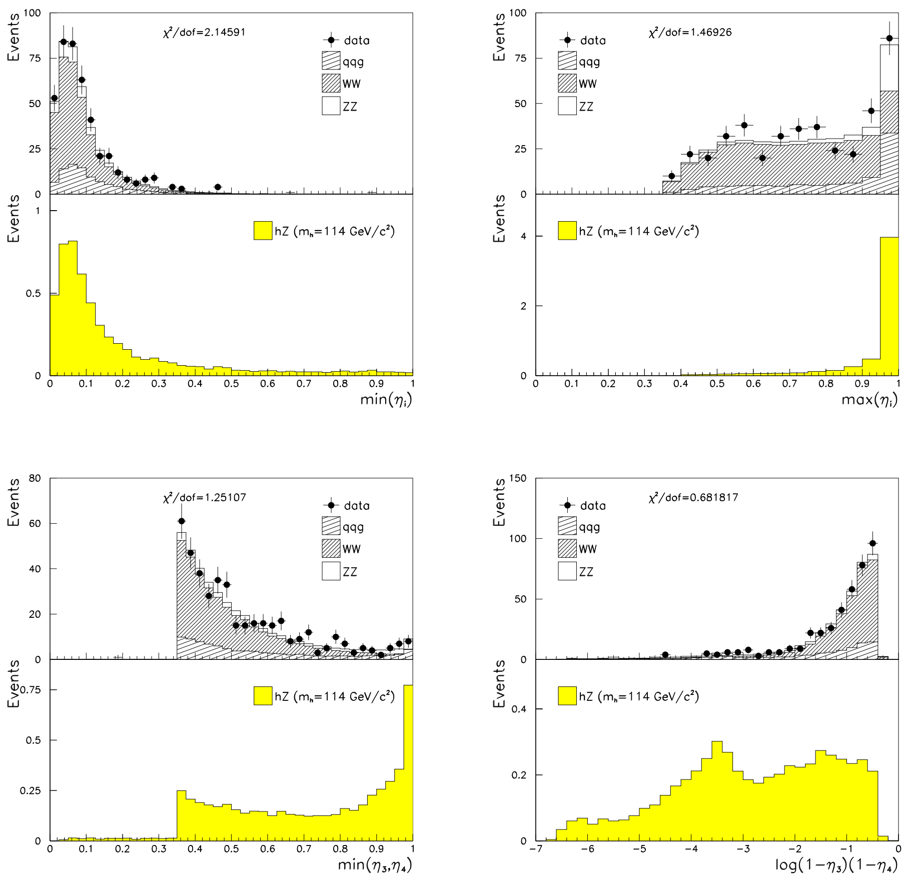


**Figure D.13:** The distributions of the b-tagging variables $(\min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1-\eta_3)(1-\eta_4))$ at the pre-selection level of the analysis with soft b-tagging cuts also applied for data, simulated background and a simulated Higgs signal with $m_{\mathrm{h}} = 114\ \mathrm{GeV}/c^2$.

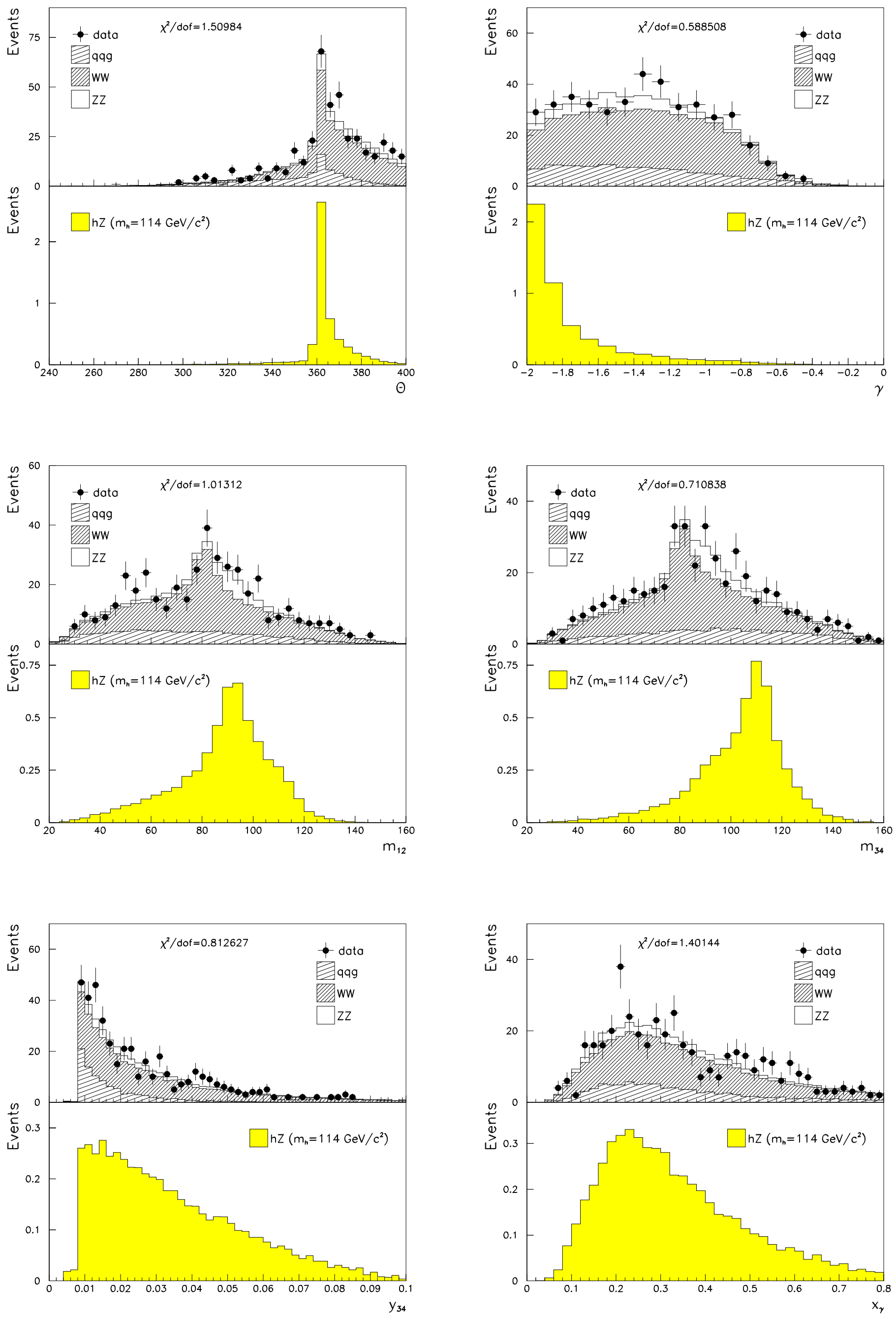


**Figure D.14:** The distributions of the non-b-tagging variables $(\Theta, \gamma, \mathrm{m}_{12}, \mathrm{m}_{34}, \mathrm{y}_{34}, \mathrm{X}_{\gamma})$ at the pre-selection level of the analysis with soft b-tagging cuts also applied for data, simulated background and a simulated Higgs signal with $m_{\mathrm{h}} = 114\,\mathrm{GeV}/c^2$.

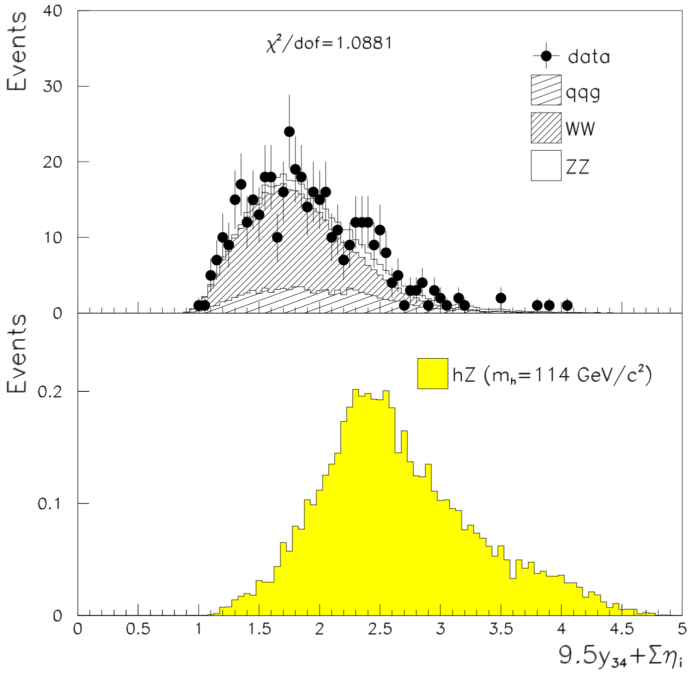


**Figure D.15:** The distributions of the linear discriminant variable($9.5y_{34} + \sum \eta_i$) at the pre-selection level of the analysis with soft b-tagging cuts also applied for data, simulated background and a simulated Higgs signal with $m_\mathrm{h} = 114\,\mathrm{GeV}/c^2$.

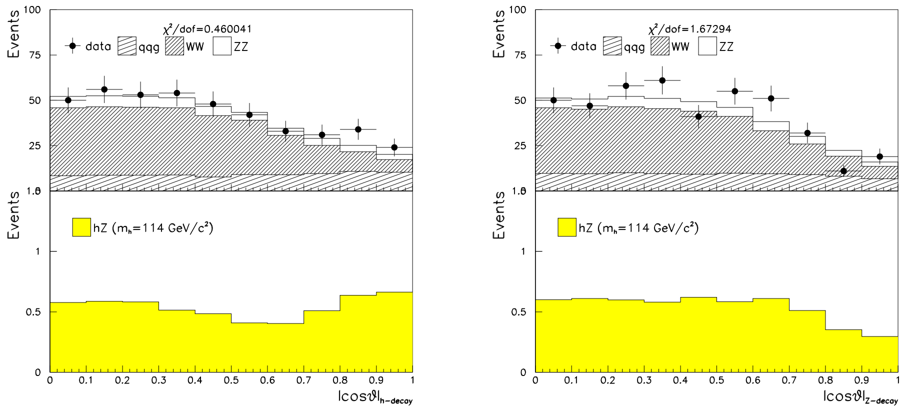


**Figure D.16:** The distributions of the h and Z decay angles ($|\cos\theta_{h-decay}|$, $|\cos\theta_{z-decay}|$) at the pre-selection level of the analysis with soft b-tagging cuts also applied for data, simulated background and a simulated Higgs signal with $m_\mathrm{h} = 114\,\mathrm{GeV}/c^2$.

### D.5.2 Full b-tag

This section presents the comparison of the observed data and MC simulation with standard pre-selection cuts applied as well as the full b-tagging cuts in the 2b event selection.

The applied cuts are:

- $\min(\eta_3, \eta_4) > 0.35$
- $(1 - \eta_3)(1 - \eta_3) < 4.8\text{x}10^{-3}$

Table D.6 shows the selected number of events from the data stream and also the expected number of events from MC simulation for background and a hypothesised 114 GeV/$c^2$ Higgs boson signal while figures D.17, D.18, D.19 and D.20 show the data to MC comparisons for the considered set of variables ($\Theta, \gamma, \mathrm{m}_{12}, \mathrm{m}_{34}, \mathrm{y}_{34}, \mathrm{X}_\gamma,\ \ \min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1 - \eta_3)(1 - \eta_4)$).

The comparison provides a test of the tight b-tagging cuts which are applied within the 2b event selection and enables us to identify any biases in the high b-tag region that is selected by the full event selection level of the analysis.

| Event Type | Number of Events |
|---|---|
| hZ ($m_\mathrm{h} = 114$ GeV/$c^2$) | 3.638 |
| $\mathrm{q\bar{q}}$ | 27.968 |
| WW | 9.428 |
| ZZ | 22.510 |
| Total bgd | 59.960 |
| Data | 65 |

**Table D.6:** The number of expected signal and background events from MC simulation and the number of observed data events for the pre-selection level of the analysis with b-tagging cuts also applied.

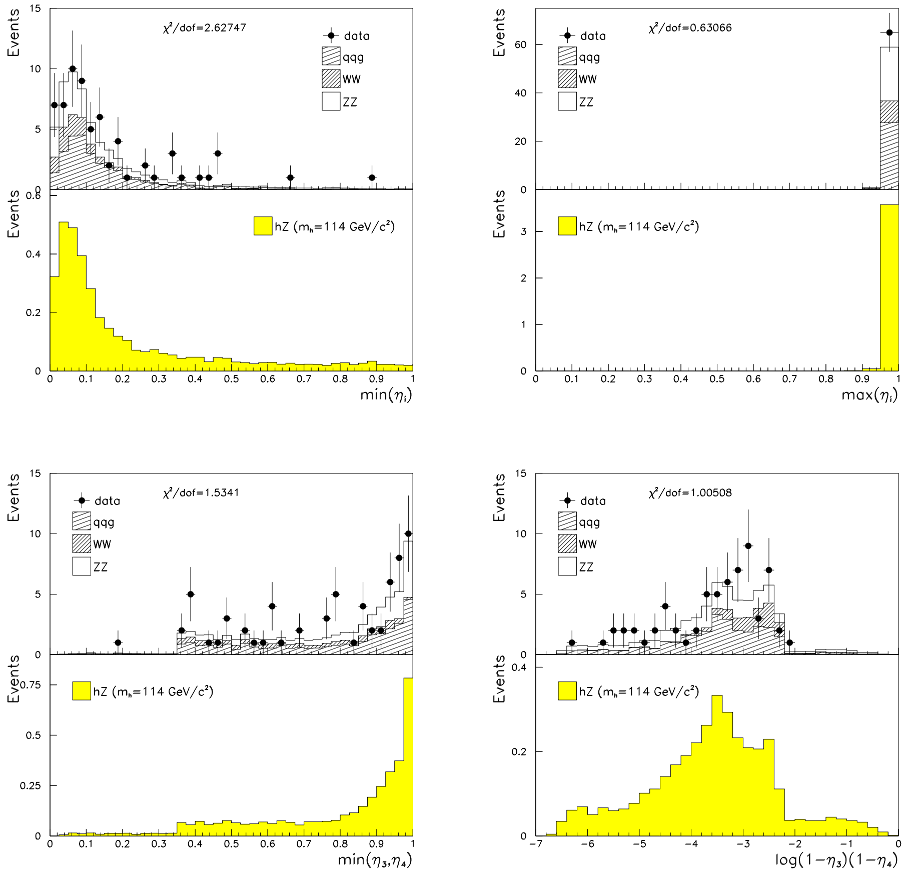


**Figure D.17:** The distributions of the b-tagging variables $(\min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1-\eta_3)(1-\eta_4))$ at the pre-selection level of the analysis with b-tagging cuts also applied.

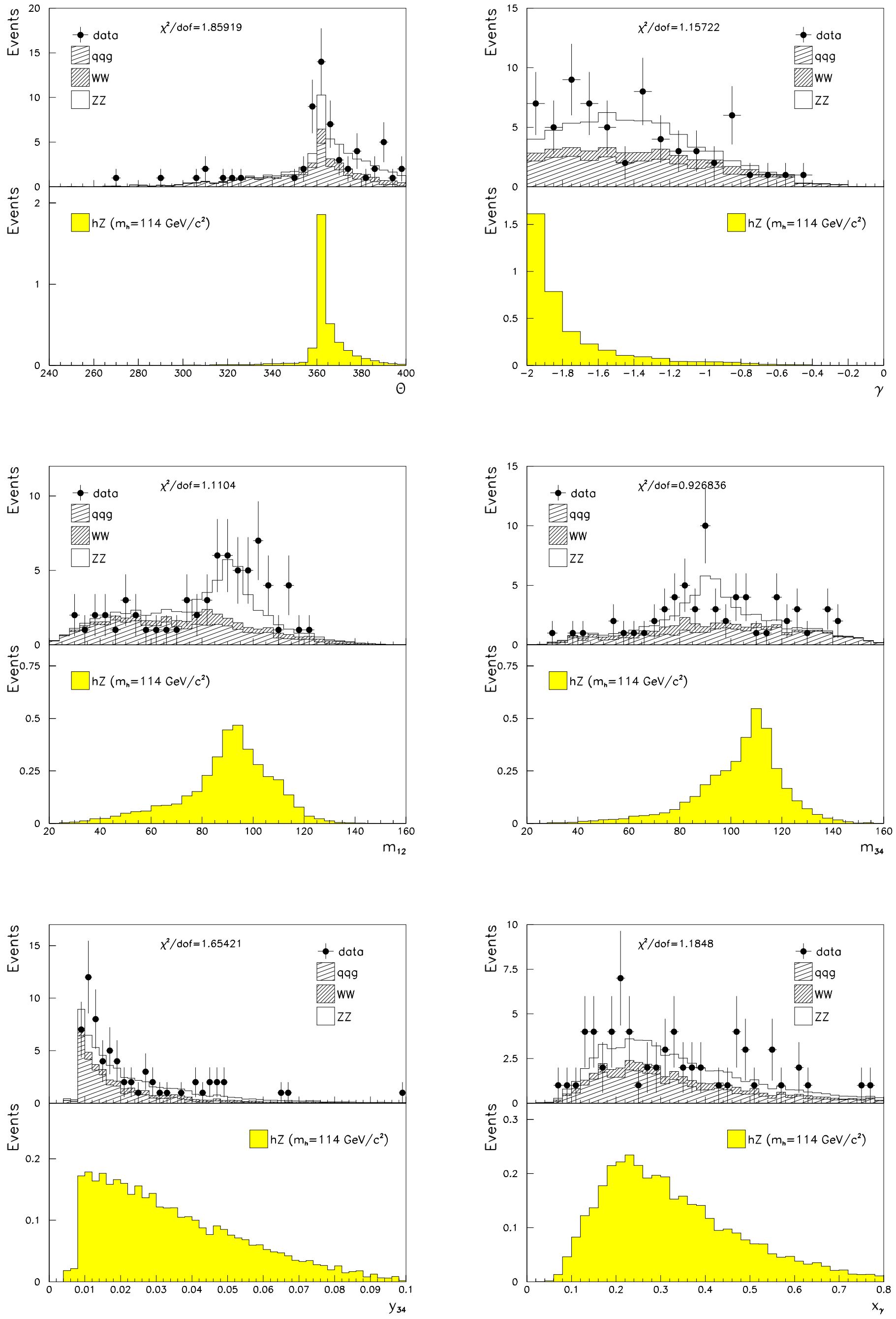


**Figure D.18:** The distributions of the non-b-tagging variables ($\Theta, \gamma, m_{12}, m_{34}, y_{34}, X_{\gamma}$) at the pre-selection level of the analysis with b-tagging cuts also applied.

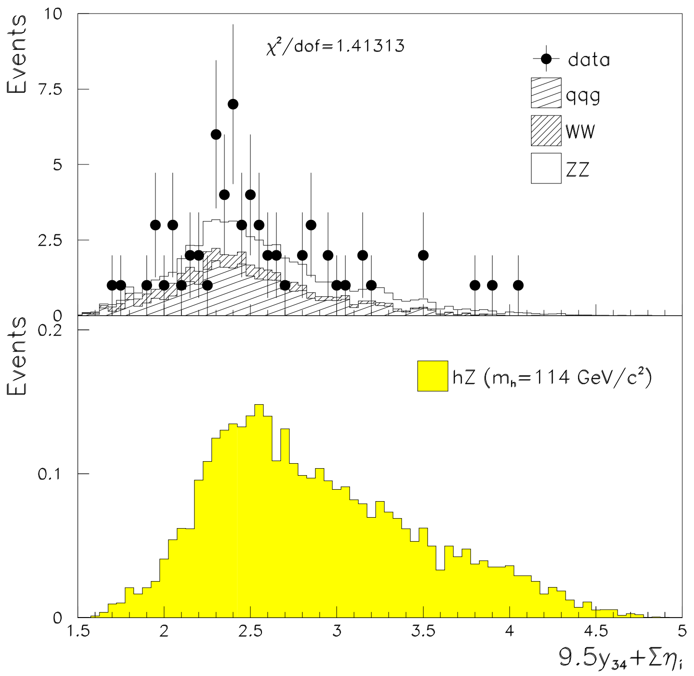


**Figure D.19:** The distributions of the linear discriminant variable ($9.5y_{34} + \sum \eta_i$) at the pre-selection level of the analysis with b-tagging cuts also applied.

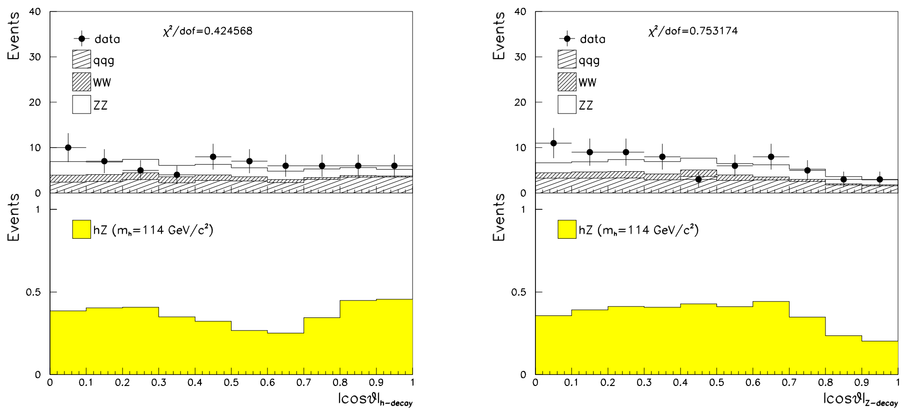


**Figure D.20:** The distributions of the h and Z decay angles ($|\cos\theta_{h-decay}|$, $|\cos\theta_{z-decay}|$) at the pre-selection level of the analysis with b-tagging cuts also applied.

## D.6 $\sqrt{s}$ dependence

The majority of data candidates observed during 2000 were recorded during the High Energy running, where High Energy running here refers to $\sqrt{s} > 206$ GeV. Figure D.21 shows the reconstructed mass and centre of mass energy for each data candidate. The kinematic limit for Higgs boson production via the Higgsstrahlung process is shown as well as the distribution of the recorded luminosity. The recorded luminosity is split into two regions above and below 206 GeV, referred to as the "Low Energy" and "High Energy" regions which have 89.55 pb$^{-1}$ and 127.63 pb$^{-1}$ of integrated luminosity respectively. The majority of collected data candidates are found in the High Energy region and in fact all the candidates with mass $\approx$ 114 GeV/$c^2$ are observed in this region. Figures D.22(a) and (b) show the distribution of the reconstructed Higgs boson mass for the expected background and hypothetical Higgs boson signal with $m_\mathrm{h} = 114$ GeV/$c^2$ and the observed data candidates at the selection level of the analysis. Table D.7 gives the number of expected background and signal events and the observed data events for the two luminosity regions at the selection level of the analysis. The possibility of a Higgs boson signal is compatible with the observation of more data candidates in the High Energy region due to threshold effects[2] and this is shown in the figures and tables describing the expected events. The possibility of poor modelling in the High energy MC is addressed by splitting the data/MC comparison into the two luminosity regions and comparing these.

Table D.8 shows the pre-selected number of events from the data stream and also the expected number of events from MC simulation for background and a hypothesised 114 GeV/$c^2$ Higgs boson signal while figures D.23, D.24, D.25 and D.26 show the data to MC comparisons for the considered analysis variables ($\Theta, \gamma, \mathrm{m}_{12}, \mathrm{m}_{34}, \mathrm{y}_{34}, \mathrm{X}_\gamma,\ \min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1-\eta_3)(1-\eta_4)$).

At the pre-selection level, presented here, no discrepancies are observed that would indicate a $\sqrt{s}$ bias.

[2]The kinematic threshold for the production of a hZ signal with $m_\mathrm{h} = 114$ GeV/$c^2$ is $\sqrt{s} =$ 205.2 GeV.

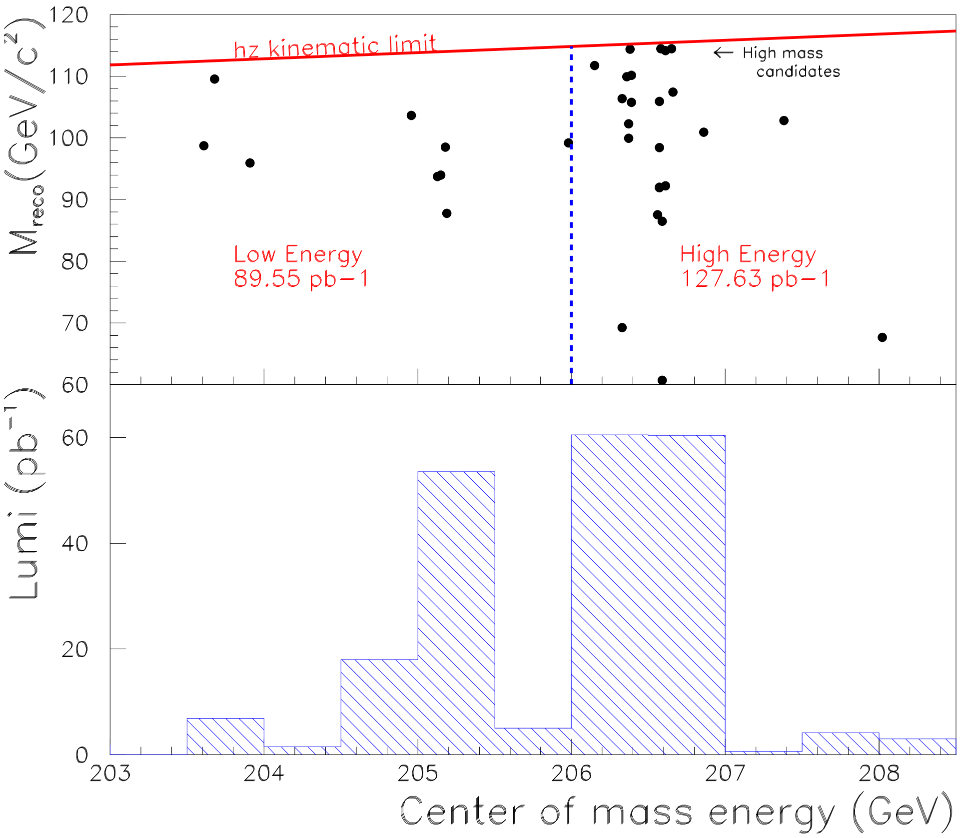


**Figure D.21:** The distribution of the data candidates in a $m_{\text{Reco}}$ versus $\sqrt{s}$ frame (upper) with the kinematic limit for each given $\sqrt{s}$ marked. The distribution of the collected integrated luminosity as a function of $\sqrt{s}$ is shown (lower) to display how the collected data candidates correspond to the collected luminosity.

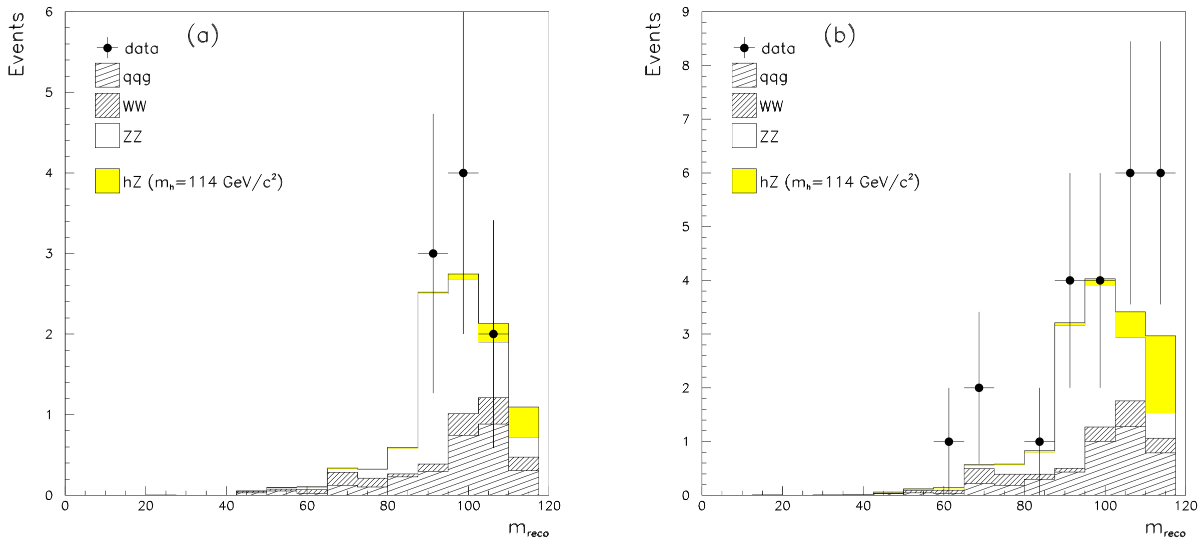


**Figure D.22:** The distributions of the reconstructed Higgs boson mass for the Low (a) and High (b) Energy data sets.

| Event Type | Number of Events |
|---|---|
| hZ | 0.772 |
| $q\bar{q}$ | 2.784 |
| WW | 1.258 |
| ZZ | 5.256 |
| Tot bgd | 9.298 |
| Data | 9 |

| Event Type | Number of Events |
|---|---|
| hZ | 2.185 |
| $q\bar{q}$ | 4.316 |
| WW | 1.779 |
| ZZ | 7.674 |
| Tot bgd | 13.769 |
| Data | 24 |

**Table D.7:** The number of expected signal and background events from MC simulation and the number of observed data events for the selection level of the analysis for the Low and High Energy regions.

| Event Type | Number of Events |
|---|---|
| hZ | 1.884 |
| $q\bar{q}$ | 223.216 |
| WW | 640.166 |
| ZZ | 52.395 |
| Tot bgd | 915.777 |
| Data | 934 |

| Event Type | Number of Events |
|---|---|
| hZ | 5.219 |
| $q\bar{q}$ | 325.439 |
| WW | 965.839 |
| ZZ | 80.864 |
| Tot bgd | 1372.142 |
| Data | 1349 |

**Table D.8:** The number of expected signal and background events from MC simulation and the number of observed data events for the pre-selection level of the analysis for the Low and High Energy regions.

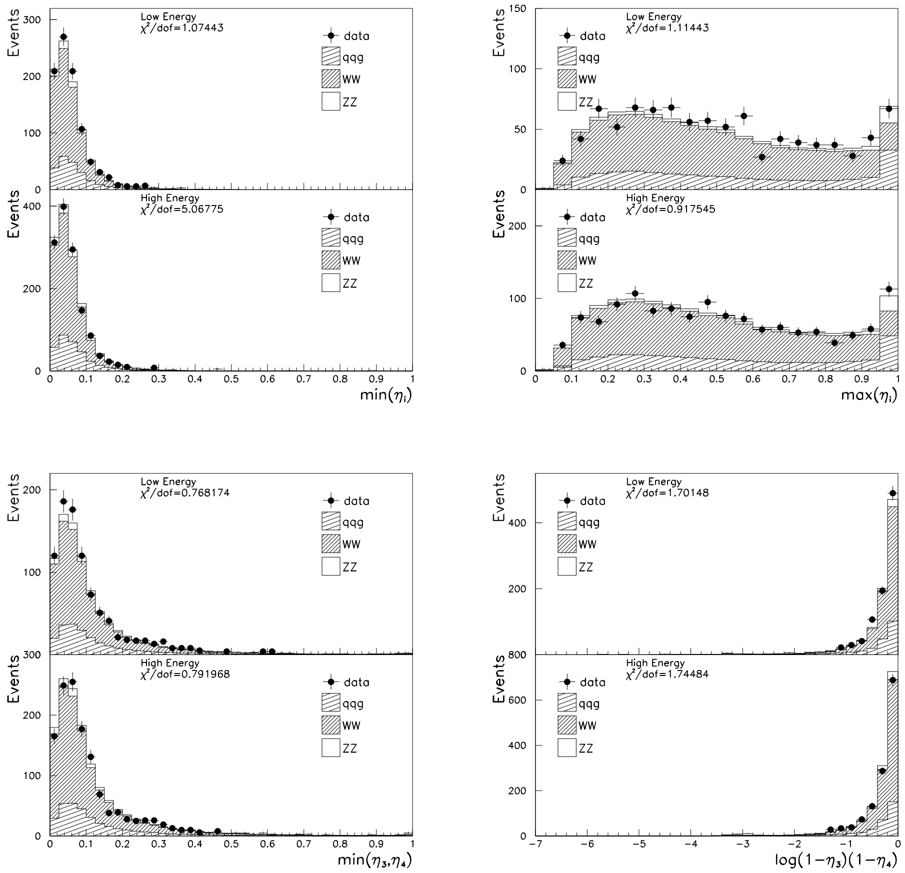


**Figure D.23:** The distributions of the b-tagging variables $(\min(\eta_i), \max(\eta_i), \min(\eta_3, \eta_4), (1-\eta_3)(1-\eta_4))$ at the pre-selection level of the analysis, for the "Low Energy" and "High Energy" event samples, as defined in the text.

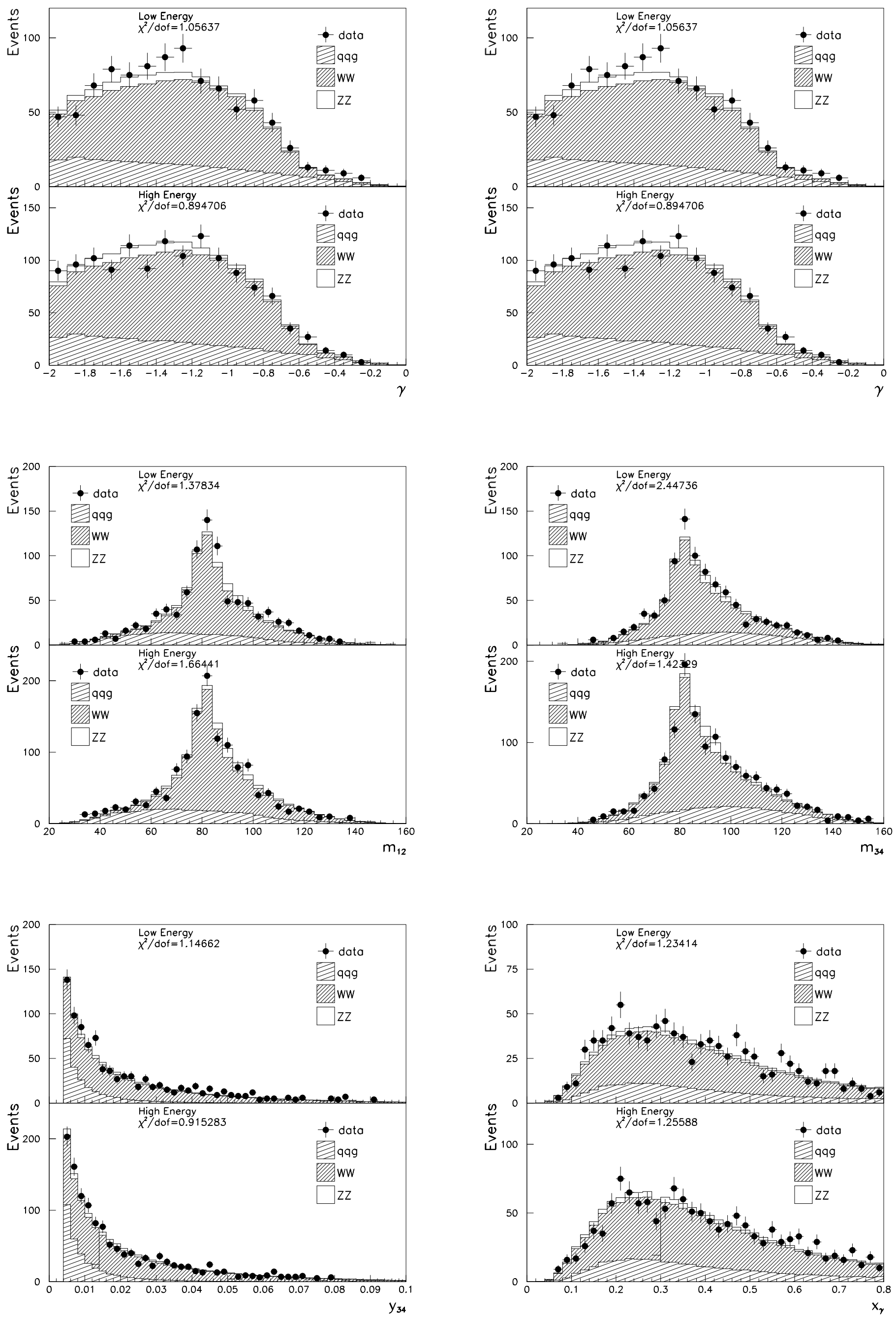


**Figure D.24:** The distributions of the non-b-tagging variables $(\Theta, \gamma, m_{12}, m_{34}, y_{34}, X_\gamma)$ at the pre-selection level of the analysis, for the "Low Energy" and "High Energy" event samples, as defined in the text.

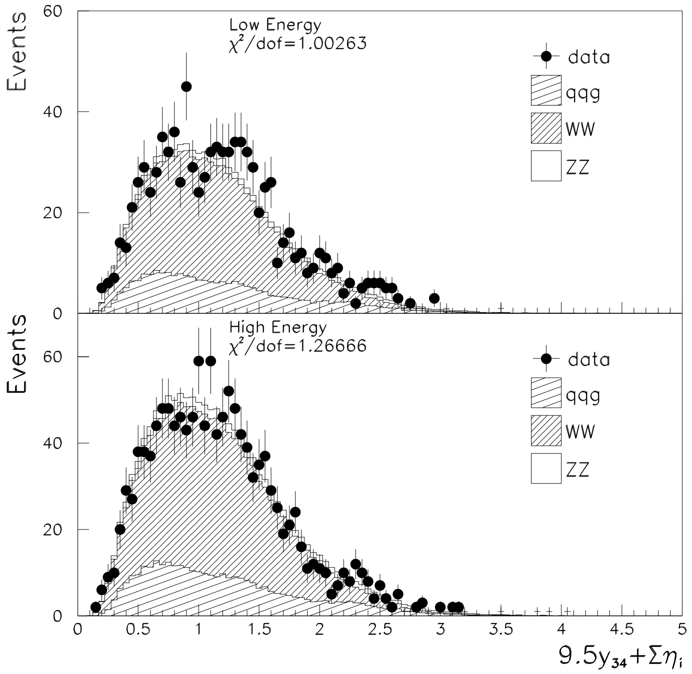


**Figure D.25:** The distributions of the linear discriminant variable ($9.5y_{34} + \sum \eta_i$) at the pre-selection level of the analysis, for the "Low Energy" and "High Energy" event samples, as defined in the text.

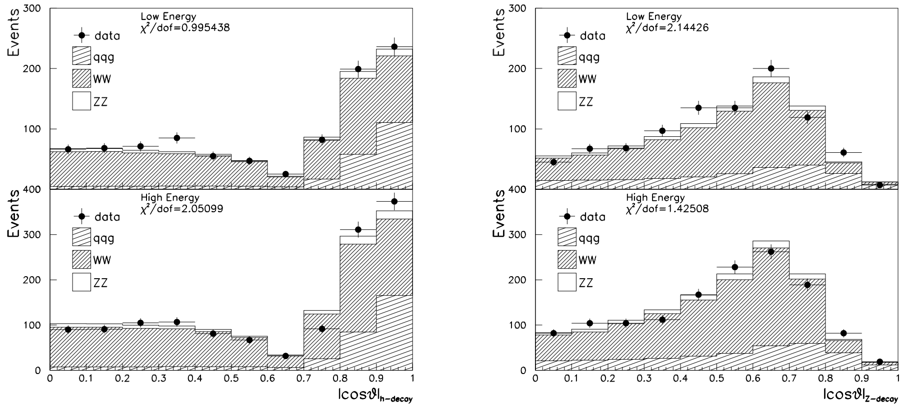


**Figure D.26:** The distributions of the h and Z decay angles ($|\cos\theta_{h-decay}|$, $|\cos\theta_{z-decay}|$) at the pre-selection level of the analysis, for the "Low Energy" and "High Energy" event samples, as defined in the text.

# Appendix E

# Candidate Details

This chapter presents the details of all the selected candidates. The candidates are broken down into their separate selection branches with table E.1 containing the hZ-exclusive candidates, table E.2 containing the hZ/hA overlap candidates and table E.3 containing the hA-exclusive candidates. Each table contains information about run and event numbers as well as the kinematic, mass and b-tagging variables for the selected events. Tables E.4-E.9 present the pairing dependent variables for the high mass candidates which have a large impact on the event analysis.

| Run | Event | $\sqrt{s}$ | $y_{34}$ | $\Theta$ | $\gamma$ | $m_{12}$ | $m_{34}$ | $m_{\mathrm{Reco}}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 53275 | 4089 | 203.91 | 0.0373 | 366.36 | -1.619 | 88.85 | 98.30 | 95.9547 | 0.039 | 0.048 | 0.967 | 0.999 |
| 53358 | 7581 | 203.68 | 0.0128 | 363.29 | -1.927 | 117.52 | 83.23 | 109.5427 | 0.707 | 0.267 | 0.999 | 0.611 |
| 53536 | 7071 | 203.61 | 0.0446 | 379.04 | -1.549 | 101.17 | 88.75 | 98.7136 | 0.171 | 0.383 | 0.999 | 0.898 |
| 53871 | 2333 | 204.96 | 0.0092 | 362.27 | -1.677 | 90.94 | 103.88 | 103.6235 | 0.038 | 0.065 | 0.972 | 0.999 |
| 54698 | 4881 | 206.61 | 0.0152 | 360.93 | -1.980 | 101.06 | 104.29 | 114.1493 | 0.124 | 0.012 | 0.998 | 0.999 |
| 54767 | 1934 | 206.59 | 0.0205 | 360.42 | -1.819 | 100.21 | 77.50 | 86.5080 | 0.283 | 0.197 | 0.934 | 0.981 |
| 54857 | 3550 | 205.15 | 0.0241 | 359.76 | -1.932 | 84.94 | 100.22 | 93.9593 | 0.024 | 0.113 | 0.972 | 0.999 |
| 54923 | 3660 | 206.57 | 0.0107 | 358.04 | -1.732 | 112.08 | 77.51 | 98.3973 | 0.333 | 0.148 | 0.998 | 0.882 |
| 54984 | 4852 | 206.57 | 0.0116 | 366.47 | -1.689 | 106.34 | 90.80 | 105.9396 | 0.825 | 0.128 | 0.999 | 0.353 |
| 55431 | 5026 | 206.57 | 0.0082 | 375.68 | -1.737 | 99.27 | 83.93 | 91.9980 | 0.214 | 0.044 | 0.994 | 0.954 |
| 55785 | 2061 | 208.02 | 0.0190 | 378.80 | -1.462 | 81.87 | 77.02 | 67.6902 | 0.194 | 0.358 | 0.492 | 0.992 |
| 55982 | 6125 | 206.58 | 0.0087 | 362.92 | -1.963 | 80.00 | 125.62 | 114.4212 | 0.071 | 0.306 | 0.449 | 0.998 |
| 56202 | 9802 | 206.39 | 0.0272 | 367.16 | -1.703 | 102.32 | 94.65 | 105.7764 | 0.920 | 0.052 | 0.593 | 0.998 |
| 56334 | 18059 | 205.19 | 0.0225 | 362.75 | -1.927 | 104.26 | 74.73 | 87.7827 | 0.354 | 0.369 | 0.971 | 0.970 |
| 56366 | 955 | 206.38 | 0.0122 | 360.69 | -1.971 | 78.55 | 127.03 | 114.3772 | 0.201 | 0.051 | 0.998 | 0.956 |
| 56669 | 3496 | 205.98 | 0.0274 | 361.97 | -1.511 | 115.52 | 74.84 | 99.1660 | 0.120 | 0.092 | 0.999 | 0.930 |
| 56875 | 987 | 206.56 | 0.0142 | 379.48 | -1.322 | 85.62 | 93.15 | 87.5760 | 0.057 | 0.172 | 0.398 | 0.999 |
| 57033 | 3973 | 205.18 | 0.0263 | 360.13 | -1.694 | 101.33 | 88.40 | 98.5374 | 0.068 | 0.015 | 0.773 | 0.999 |
| 57360 | 8956 | 206.66 | 0.0123 | 366.01 | -1.720 | 107.75 | 90.89 | 107.4426 | 0.682 | 0.141 | 0.601 | 0.990 |
| 57423 | 6845 | 206.37 | 0.0292 | 370.44 | -1.642 | 85.98 | 107.52 | 102.3006 | 0.079 | 0.151 | 0.997 | 0.479 |
| 57585 | 749 | 206.59 | 0.0645 | 379.91 | -1.500 | 84.58 | 67.29 | 60.6764 | 0.032 | 0.070 | 0.999 | 0.686 |
| 57747 | 8577 | 206.61 | 0.0425 | 383.99 | -1.362 | 112.74 | 70.71 | 92.2532 | 0.213 | 0.017 | 0.781 | 0.985 |
| 58023 | 117 | 206.37 | 0.0327 | 391.51 | -1.335 | 120.23 | 70.91 | 99.9407 | 0.186 | 0.296 | 0.969 | 0.872 |
| 58121 | 5044 | 206.39 | 0.0182 | 370.08 | -1.874 | 110.60 | 90.71 | 110.1098 | 0.078 | 0.018 | 0.995 | 0.775 |
| 58201 | 6835 | 206.15 | 0.0172 | 360.22 | -1.899 | 80.66 | 122.30 | 111.7545 | 0.096 | 0.277 | 0.965 | 0.870 |

**Table E.1:** Information about all selected candidate events within the hZ-exclusive branch of the four jets analysis.

| Run | Event | $\sqrt{s}$ | $y_{34}$ | $\Theta$ | $\gamma$ | $m_{12}$ | $m_{34}$ | $m_{\text{Reco}}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 53688 | 2208 | 206.33 | 0.0138 | 362.68 | -1.725 | 105.07 | 55.39 | 69.2597 | 0.097 | 0.581 | 0.979 | 0.935 |
| 53957 | 6617 | 206.33 | 0.0179 | 364.16 | -1.814 | 114.01 | 83.53 | 106.3389 | 1.000 | 0.271 | 0.995 | 0.387 |
| 54794 | 6706 | 205.13 | 0.0492 | 391.62 | -1.326 | 98.89 | 86.07 | 93.7666 | 0.410 | 0.617 | 0.997 | 0.998 |
| 55484 | 18148 | 206.57 | 0.0087 | 359.03 | -1.423 | 89.19 | 93.94 | 91.9312 | 0.526 | 0.466 | 0.997 | 0.771 |
| 56065 | 3253 | 206.65 | 0.0160 | 359.99 | -1.963 | 110.61 | 95.02 | 114.4278 | 1.000 | 0.996 | 0.996 | 0.663 |
| 56667 | 1617 | 207.38 | 0.0167 | 359.75 | -1.725 | 102.48 | 91.53 | 102.8025 | 0.686 | 0.347 | 0.983 | 0.962 |
| 56698 | 7455 | 206.36 | 0.0475 | 360.00 | -1.860 | 95.76 | 105.37 | 109.9299 | 0.999 | 0.831 | 0.999 | 0.197 |
| 58050 | 6532 | 206.86 | 0.0101 | 356.33 | -1.771 | 92.59 | 99.55 | 100.9465 | 0.527 | 0.457 | 0.996 | 0.915 |

**Table E.2:** Information about all selected candidate events within the hZ/hA overlap branch of the four jets analysis.

| Run | Event | $\sqrt{s}$ | $y_{34}$ | $\Theta$ | $\gamma$ | $m_{12}$ | $m_{34}$ | $m_{\text{Reco}}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 53715 | 6224 | 206.37 | 0.0097 | 386.64 | -1.181 | 78.95 | 89.75 | 77.5000 | 0.603 | 0.449 | 1.000 | 0.999 |
| 53790 | 2022 | 206.43 | 0.0033 | 360.07 | -1.759 | 94.81 | 89.70 | 93.3114 | 0.999 | 0.958 | 0.724 | 0.417 |
| 54662 | 9417 | 206.62 | 0.0163 | 319.92 | -1.289 | 46.51 | 66.95 | 22.2645 | 0.995 | 0.997 | 0.998 | 0.895 |
| 55688 | 6899 | 204.76 | 0.0133 | 350.04 | -1.241 | 59.48 | 59.63 | 27.9102 | 0.829 | 0.947 | 0.905 | 0.159 |
| 57100 | 7446 | 206.57 | 0.0049 | 361.75 | -1.604 | 107.75 | 72.58 | 89.1275 | 0.995 | 0.999 | 0.998 | 0.078 |
| 57183 | 279 | 206.57 | 0.0039 | 363.34 | -1.754 | 93.16 | 94.93 | 96.8913 | 0.940 | 0.716 | 0.512 | 0.914 |
| 57333 | 891 | 205.16 | 0.0441 | 403.72 | -0.834 | 84.01 | 81.21 | 74.0221 | 0.448 | 0.277 | 0.948 | 0.977 |
| 57660 | 780 | 206.61 | 0.0412 | 423.84 | -0.895 | 89.02 | 89.60 | 87.4129 | 0.766 | 0.971 | 0.611 | 0.466 |
| 57779 | 17001 | 206.59 | 0.0184 | 395.38 | -0.664 | 89.88 | 73.25 | 71.9391 | 0.726 | 0.978 | 0.911 | 0.084 |
| 58209 | 13823 | 206.17 | 0.0985 | 416.52 | -1.006 | 88.81 | 92.25 | 89.8578 | 0.981 | 0.998 | 0.343 | 0.617 |
| 58404 | 6132 | 206.36 | 0.0063 | 360.99 | -1.771 | 92.86 | 89.80 | 91.4638 | 0.458 | 0.869 | 0.766 | 0.967 |

**Table E.3:** Information about all selected candidate events within the hA-exclusive branch of the four jets analysis.

The following tables present the details of the six high mass candidates, as introduced in table 6.3 which have a large impact on the recorded significance. The pairing dependent variables for all possible combinations of a given event are shown. The pairings which pass the event selection are marked (y) while those which fail are marked (n) and the pairing which is final selected by the analysis is highlighted with a *

**Run: 56065 Event: 3253**

| Pair | $m_{12}$ | $m_{34}$ | $m_{\text{Reco}}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ | $pdf_z$ | $pdf_h$ | Sel |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 110.61 | 95.02 | 114.4278 | 1.000 | 0.996 | 0.996 | 0.663 | 0.1722 | 0.0604 | y* |
| 2 | 93.56 | 98.38 | 100.7382 | 1.000 | 0.996 | 0.996 | 0.663 | 0.0609 | 0.0846 | y |
| 3 | 35.34 | 51.47 | -4.3855 | 1.000 | 0.663 | 0.996 | 0.996 | 0.0554 | 0.0650 | y |
| 4 | 98.38 | 93.56 | 100.7382 | 0.996 | 0.663 | 1.000 | 0.996 | 0.0918 | 0.0625 | y |
| 5 | 51.47 | 35.34 | -4.3855 | 0.996 | 0.996 | 1.000 | 0.663 | 0.0652 | 0.0597 | y |
| 6 | 95.02 | 110.61 | 114.4278 | 0.996 | 0.663 | 1.000 | 0.996 | 0.0567 | 0.1669 | y |

**Table E.4:** Pairing dependent variables for high mass candidate run: 56065 event: 3253 .

**Run: 56698 Event: 7455**

| Pair | $m_{12}$ | $m_{34}$ | $m_{\text{Reco}}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ | $pdf_z$ | $pdf_h$ | Sel |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 93.22 | 72.00 | 74.0235 | 0.999 | 0.999 | 0.831 | 0.197 | 0.0974 | 0.0576 | y |
| 2 | 89.69 | 53.22 | 51.7081 | 0.999 | 0.831 | 0.999 | 0.197 | 0.1012 | 0.0604 | y |
| 3 | 105.37 | 95.76 | 109.9299 | 0.999 | 0.197 | 0.999 | 0.831 | 0.2780 | 0.0597 | y |
| 4 | 53.22 | 89.69 | 51.7081 | 0.999 | 0.197 | 0.999 | 0.831 | 0.0567 | 0.0929 | y |
| 5 | 95.76 | 105.37 | 109.9299 | 0.999 | 0.831 | 0.999 | 0.197 | 0.0554 | 0.2977 | y* |
| 6 | 72.00 | 93.22 | 74.0235 | 0.831 | 0.197 | 0.999 | 0.999 | 0.0503 | 0.0895 | y |

**Table E.5:** Pairing dependent variables for high mass candidate run: 56698 event: 7455 .

### Run: 54698 Event: 4881

| Pair | $m_{12}$ | $m_{34}$ | $m_{\text{Reco}}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ | $pdf_z$ | $pdf_h$ | Sel |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 108.02 | 86.48 | 103.3033 | 0.998 | 0.124 | 0.012 | 0.999 | -1.0000 | -1.0000 | n |
| 2 | 49.23 | 36.68 | -5.2966 | 0.998 | 0.012 | 0.124 | 0.999 | -1.0000 | -1.0000 | n |
| 3 | 104.29 | 101.06 | 114.1493 | 0.998 | 0.999 | 0.124 | 0.012 | -1.0000 | -1.0000 | n |
| 4 | 36.68 | 49.23 | -5.2966 | 0.124 | 0.999 | 0.998 | 0.012 | -1.0000 | -1.0000 | n |
| 5 | 101.06 | 104.29 | 114.1493 | 0.124 | 0.012 | 0.998 | 0.999 | 0.1670 | 0.2512 | y* |
| 6 | 86.48 | 108.02 | 103.3033 | 0.012 | 0.999 | 0.998 | 0.124 | -1.0000 | -1.0000 | n |

**Table E.6:** Pairing dependent variables for high mass candidate run: 54698 event: 4881 .

### Run: 55982 Event: 6125

| Pair | $m_{12}$ | $m_{34}$ | $m_{\text{Reco}}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ | $pdf_z$ | $pdf_h$ | Sel |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 125.62 | 80.00 | 114.4212 | 0.449 | 0.998 | 0.071 | 0.306 | -1.0000 | -1.0000 | n |
| 2 | 98.73 | 92.71 | 100.2402 | 0.449 | 0.071 | 0.998 | 0.306 | -1.0000 | -1.0000 | n |
| 3 | 47.12 | 36.27 | -7.8074 | 0.449 | 0.306 | 0.998 | 0.071 | -1.0000 | -1.0000 | n |
| 4 | 92.71 | 98.73 | 100.2402 | 0.998 | 0.306 | 0.449 | 0.071 | -1.0000 | -1.0000 | n |
| 5 | 36.27 | 47.12 | -7.8074 | 0.998 | 0.071 | 0.449 | 0.306 | -1.0000 | -1.0000 | n |
| 6 | 80.00 | 125.62 | 114.4212 | 0.071 | 0.306 | 0.449 | 0.998 | 0.1262 | 0.1286 | y* |

**Table E.7:** Pairing dependent variables for high mass candidate run: 55982 event: 6125 .

### Run: 56366 Event: 955

| Pair | $m_{12}$ | $m_{34}$ | $m_{\text{Reco}}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ | $pdf_z$ | $pdf_h$ | Sel |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 127.03 | 78.55 | 114.3772 | 0.998 | 0.956 | 0.201 | 0.051 | -1.0000 | -1.0000 | n |
| 2 | 69.13 | 57.36 | 35.2948 | 0.998 | 0.201 | 0.956 | 0.051 | -1.0000 | -1.0000 | n |
| 3 | 87.03 | 81.24 | 77.0692 | 0.998 | 0.051 | 0.956 | 0.201 | -1.0000 | -1.0000 | n |
| 4 | 57.36 | 69.13 | 35.2948 | 0.956 | 0.051 | 0.998 | 0.201 | -1.0000 | -1.0000 | n |
| 5 | 81.24 | 87.03 | 77.0692 | 0.956 | 0.201 | 0.998 | 0.051 | -1.0000 | -1.0000 | n |
| 6 | 78.55 | 127.03 | 114.3772 | 0.201 | 0.051 | 0.998 | 0.956 | 0.1175 | 0.2026 | y* |

**Table E.8:** Pairing dependent variables for high mass candidate run: 56366 event: 955 .

**Run: 58201 Event: 6835**

| Pair | $m_{12}$ | $m_{34}$ | $m_{\text{Reco}}$ | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ | $pdf_z$ | $pdf_h$ | Sel |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 122.30 | 80.66 | 111.7545 | 0.965 | 0.870 | 0.096 | 0.277 | -1.0000 | -1.0000 | n |
| 2 | 102.78 | 69.98 | 81.5621 | 0.965 | 0.096 | 0.870 | 0.277 | -1.0000 | -1.0000 | n |
| 3 | 63.38 | 55.93 | 28.1035 | 0.965 | 0.277 | 0.870 | 0.096 | -1.0000 | -1.0000 | n |
| 4 | 69.98 | 102.78 | 81.5621 | 0.870 | 0.277 | 0.965 | 0.096 | -1.0000 | -1.0000 | n |
| 5 | 55.93 | 63.38 | 28.1035 | 0.870 | 0.096 | 0.965 | 0.277 | -1.0000 | -1.0000 | n |
| 6 | 80.66 | 122.30 | 111.7545 | 0.096 | 0.277 | 0.965 | 0.870 | 0.2233 | 0.2242 | y* |

**Table E.9:** Pairing dependent variables for high mass candidate run: 58201 event: 6835 .

# Bibliography

[1] Lucretius - On the Nature of the Universe. Penguin books ltd, 1951.

[2] F. Halzen and A. D. Martin. Quarks and Leptons. John Wiley & Sons, New York USA, 1984.

[3] D. Griffiths. Introduction to Elementary Particles. John Wiley & Sons, New York USA, 1987.

[4] F. Mandl and G. Shaw. Quantum Field Theory. John Wiley & Sons, New York USA, revised addition edition, 1993.

[5] G.L. Kane J.F. Gunion, H.E. Haber and S. Dawson. The Higgs Hunter's Guide. Perseus Publishing, Cambridge, Massachusetts, 2000.

[6] S. Glashow. Partial symmetries of weak interactions. Nucl. Phys, 22:579, 1961.

[7] S. Weinberg. A model of leptons. Phys. Rev. Lett., 19:1264, 1967.

[8] A. Salam. Elementary particle theory. page 367, Stockholm, 1968. Almquist and Forlag.

[9] D.E.Groom et al. Review of particle physics. Eur. Phys. J., C(15):1, 2000.

[10] E. Norther. Nachr. Kgl. Ges. Wiss. Gottingen, page 235, 1918.

[11] P.W. Higgs. Broken symmetries, massless particles and gauge fields. Phys. Lett., 12:132, 1964.

[12] P.W. Higgs. Spontaneous symmetry breakdown without massless bosons. Phys. Rev., 145:1156, 1966.

[13] C.R. Hagen G.S. Guralnik and T.W.B. Kibble. Phys. Rev. Lett, 13:585, 1964.

[14] F. Englert and R. Brout. Phys. Rev. Lett, 13:1156, 1964.

[15] J. Goldstone. Nuovo Cimennto, 19:154, 1961.

[16] Y. Nambu. Phys. Rev. Lett, 4:380, 1960.

[17] J. Ellis. Beyond the standard model for hillwalkers. In 1998 European School on High-Energy Physics, page 133. CERN Yellow Report 1999-04, 1998.

[18] C. Quigg B.W. Lee and H.B. Thacker. Phys. Rev., D16:1519, 1977.

[19] E. Tournefier for the LEP Electroweak Working Group. Electroweak fits and the Higgs likelihood. In XXXVIme Recontres de Moriond, Moriond,France, March 2001.

[20] B.Pietrzyk. The global fit to electroweak data. In International Conference of High Energy Physics, Osaka, Japan, July 27 - August 2 2000.

[21] L3 ALEPH, DELPHI and OPAL Collaborations (The LEP working group for Higgs boson searches). Search for the Higgs boson: Preliminary combined results using LEP data collected at energies up to 202 GeV. CERN-EP/2000-055, 2000.

[22] M. Kramer W. Kilian and P.M. Zerwas. Higgs-strahlung and ww fusion in $e^+e^-$ collisions. Phys. Lett., B(373):135, 1996.

[23] G. Altarelli, T. Sjöstrand, and F. Zwirner Eds. Physics at LEP2. CERN 96-01, CERN, 1996. Volume 1.

[24] P. Janot. Hzha03. http://ALEPHwww.cern.ch/j̃anot/Generators.html.

[25] S. Myers and E. Picasso. The design construction and commissioning of the cern large electron positron collider. Contemporary Physics, 31:387, 1990.

[26] The ALEPH Coll. D. Decamp et al. ALEPH: A detector for electron-positron annihilations at LEP. Nucl. Instrum. and Methods, A(294):121, 1990.

[27] The ALEPH Coll. D. Buskulic et al. Performance of the ALEPH detector at LEP. Nucl. Instrum. and Methods, A(360):481, 1995.

[28] The ALEPH Coll. The ALEPH handbook. Vol. 1 and 2 are available from the ALEPH secretariat at CERN, 1995.

[29] O. Schneider D. Schlatter, P. Perrodo and A. Wagner. ALEPH in numbers. 1996.

[30] The DELPH Coll. P. Aarnio et al. The DELPHI detector at LEP. Nucl. Instrum and Methods, A303:233, 1991.

[31] The L3 Collaboration. The construction of the L3 experiment. Nucl. Instrum and Methods, A289:35, 1990.

[32] The OPAL Coll. K. Ahmet et al. The OPAL detector at LEP. Nucl. Instrum and Methods, A305:275, 1991.

[33] J. Lynch. The ALEPH operations group. Private communication.

[34] SPS and LEP performance workshop. How should we organise the Higgs Safari?, Chamonix, January 1999.

[35] D. Creanza et al. The new ALEPH vertex detector. Nucl. Instrum. and Methods, A(409):157, 1998.

[36] G.J. Barber et al. Performance of the three-dimensional readout of the ALEPH inner tracking chamber. Nucl. Inst and Methods, A279:212, 1989.

[37] S.J. Dorris. A Measurement of the Colour Factors of Quantum Chromodynamics from Four-Jet Events at LEP. PhD thesis, The University of Glasgow, 1997.

[38] E.J. Wolin and L.L. Ho. Covariant matrices for track fitting with the kalman filter. Nucl. Instrum. and Methods, A(329):493, 1993.

[39] F. Ranjard. GALEPH - monte carlo program for ALEPH. ALEPH Internal Note 88-119, 1998.

[40] M. Maire A. McPherson R. Brun, F. Bruyant and P. Zanarini. GEANT3. CERN-DD/EE/84-1.

[41] J. Knobloch. Julia users and programmers guide. ALEPH Internal Note 90-11, 1990.

[42] T. Sjostrand. High-energy physics event generation with pythia 5.7 and jetset 7.4. Comput. Phys. Commun, 82:74, 1994.

[43] B.F. Ward S. Jadach and Z. Was. Monte carlo program koralz version 4.0, for LEPton or quark pair production at LEP/slc energies. Comput. Phys. Commun, 79:503, 1994.

[44] W. Placzek M. Skrzypek, S. Jadach and Z. Was. Monte carlo program koralw-1.02 for w pair production at LEP2/nlc energies with yennie-frautschi-suura exponentiation. Comput. Phys. Commun, 94:216, 1996.

[45] C. Peterson and T. Rognvaldsson. An introduction to artificial neural networks. In C. Verkerk, editor, CERN School of Computing, 1991.

[46] S. Armstrong et al. Identification of b-jets in Higgs searches using a neural network. ALEPH Internal Note 96-100, 1996.

[47] D. Brown and M. Frank. Tagging b-hadrons using impact parameters. ALEPH Internal Note 92-135, 1992.

[48] T. Mattison. Qvsrch a tool for inclusive secondary vertex finding. ALEPH Internal Note 92-173, 1992.

[49] C. Paterson L. Lonnbald and T. Rognvaldsson. Jetnet 3.0: A versatile artificial neural network package. CERN-TH 7135/94, 1994.

[50] D. H. Smith. The search for Higgs bosons in the four jet channel at ALEPH. PhD thesis, The University of Glasgow, In Preparation.

[51] S.R. Armstrong. A Search for the Standard Model Higgs Boson in Four-Jet Final States at Centre-of-Mass Energies Near 183 GeV with the ALEPH Detector at LEP. PhD thesis, University of Wisconsin-Madison, 1998.

[52] S. Catani et al. New clustering algorithm for multi-jet cross-sections in $e^+e^-$ annihilation. Phys. Lett, B(269):432, 1991.

[53] The ALEPH Coll. R. Barate et al. Measurement of the w mass by direct reconstruction in $e^+e^-$ collisions at 172 GeV. Phys. Lett, B(422):384, 1998.

[54] S.L. Wu. Z. Phys., C9:329, 1981.

[55] T.L.Barklow. A Search for the Production of Charged Higgs and Technipions with Large Hadronic Branching Ratios in $e^+e^-$ Annihilation at 34.5 GeV Enter-of-Mass Energy. PhD thesis, University of Wisconsi-Madison, 1983.

[56] The ALEPH Coll. R. Barate et al. Searches for the neutral Higgs boson in $e^+e^-$ collisions at centre-of-mass energies from 192 to 202 GeV. Phys. Lett., B(299):53, 2001.

[57] H. Dreverman et al. Event display: Can we see what we want to see? CERN School of Computing, 1995.

[58] G. Cowan. Statistical Data Analysis. Oxford Science Publications, Oxford, England, 1998.

[59] Anders Waananen. 24 April 1999. Presentation to the ALEPH Higgs Task Force.

[60] David Smith. 26 January 2000. Presentation to the ALEPH Higgs Task Force.

[61] David Smith. 9 December 1999. Presentation to the ALEPH Higgs Task Force.

[62] K.S. Cranmer. Kernel estimation in high energy physics. Comput. Phys. Commun, (136):198, 2001.

[63] J. Allison. Multiquadratic radial basis functions for representing multidimensional high energy physics data. Manchester Particle Physics MAN/HEP/93/4, 1994. Submitted to Computer Physics Communications.

[64] John Kennedy. 24 January 2001. Presentation to the ALEPH Higgs Task Force.

[65] The ALEPH Coll. R. Barate et al. A measurement of the gluon splitting rate into $b\bar{b}$ pairs in hadronic z decays. Phys. Lett., B(434):437, 1998.

[66] I. Schmitt C. Peterson, D. Schlatter and P.M. Zerwas. Scaling violations in inclusive $e^+e^-$ annihilation spectra. Phys. Rev., D(27):105, 1983.

[67] The ALEPH Coll. Study of the fragmentation of b quarks into b mesons at the Z peak. CERN-EP/2001-039, 2001.

[68] The ALEPH Coll. R. Barate et al. Study of charm production in z decays. Euro. Phys. Journal, C(16):597, 2000.

[69] The ALEPH Coll. R. Barate et al. Observation of an excess in the search for the standard model Higgs boson at ALEPH. Phys. Lett., B(495):1, 2001.

[70] The ALEPH Coll. R. Barate et al. Search for the neutral Higgs boson of the standard model and the mssm in $e^+e^-$ collisions at $\sqrt{s} = 189$ GeV. E.Phys.J, C(17):223, 2000.

[71] S. Armstrong et al. Studies of Higgs candidates with masses around 114 GeV/$c^2$ in four-jet events at $\sqrt{s} > 205$ GeV using neural network techniques. ALEPH Internal Note 2000-079, 2000.

[72] J. Nowell. A Measurement of the W Boson Mass with the ALEPH Detector. PhD thesis, Imperial College, The University of London, 2000.

[73] O. Buchmuller. $e^+e^- \rightarrow hZ \rightarrow b\bar{b}q\bar{q}$ a kinematic point of view. ALEPH Internal Note 2000-087, 2000.

[74] Oliver Buchmuller. 22 August 2000. Presentation to the ALEPH Higgs Task Force.

[75] Tom Greening for John Kennedy. 9 May 2001. Presentation to the ALEPH Higgs Task Force.

[76] The L3 Coll. M. Acciarri et al. L3 Internal Note 2699, 2001.

[77] L3 ALEPH, DELPHI and OPAL Collaborations (The LEP working group for Higgs boson searches). Searches for the standard model Higgs boson at LEP. CERN-EP/2001-055, 2001.

[78] The DELPHI Coll. P. Abreu et al. DELPHI Internal Note 2000-113, 2001.

[79] The OPAL Coll. G. Abbiendi et al. OPAL Internal Note PN479, 2001.

[80] A. Read. Modified frequentist analysis of search results (the cls method). In Y. Perrin F. James, L. Lyons, editor, Workshop on Confidence Limits, page 81. CERN Yellow Report 2000-005, 2000.

[81] M. Carena et al. Report of the tevatron Higgs working group. FERMILAB conf 00/279, hep-ph/0010338, 2000.

[82] The ATLAS Collaboration. Atlas detector and physics performance design report. LHCC 99-14/15, 1999.

[83] The CMS Collaboration. CMS technical proposal. CERN/LHCC 94-38, 1994.

[84] J.A. Nielsen. Observation of an Excess in the Search for the Standard Model Higgs Boson at ALEPH. PhD thesis, University of Wisconsin-Madison.

[85] John Kennedy. 26 January 2000. Presentation to the ALEPH Higgs Task Force.

[86] John Kennedy. 12 April 2000. Presentation to the ALEPH Higgs Task Force.

[87] A.L. Read. Optimal statistical analysis of search results based on the likelihood ratio and its application to the search for the msm Higgs boson at $\sqrt{s} = 161$ GeV and $\sqrt{s} = 172$ GeV. DELPHI Internal Note 97-158, 1997.

[88] H. Hu and J. Nielsen. Analytic confidence level calculations using the likelihood ratio and fourier transform. page 109. CERN Yellow Report 2000-005, 2000.

[89] Pete McNamara. How does one interpret the background confidence level? http://www-wisconsin.cern.ch/m̃cnamara/whatisclb.html.

[90] S. Jin and P. McNamara. The signal estimator limit setting method. Nucl. Instrum. and Methods, A(462):561, 2001.

[91] C. Loomis et al. Search for standard model Higgs boson in $e^+e^-$ collisions at $\sqrt{s} = 183$ GeV using the hZ→ h$\nu\overline{\nu}$ channel. ALEPH Internal Note 98-009, 1998.

[92] Richard White. 12 April 2000. Presentation to the ALEPH Higgs Task Force.

[93] T. Greening. Search for the standard model Higgs boson in the electron and muon pair channels at $\sqrt{s} = 183$ GeV. ALEPH Internal Note 90-003, 1998.

[94] Tom Greening. 16 March 2000. Presentation to the ALEPH Higgs Task Force.

[95] The ALEPH Higgs Taskforce. Search for neutral Higgs bosons in the $\tau^+\tau^-q\bar{q}$ final states at 183 GeV. ALEPH Internal Note 98-010, 1998.

[96] The ALEPH Higgs Taskforce. Search for neutral Higgs bosons of the mssm in the channel hA→ $b\bar{b}b\bar{b}$ at $\sqrt{s} = 183$ GeV. ALEPH Internal Note 98-002.

[97] B. Tuchming. A four-variable neural network for b-tagging. application to the search for hA→ $b\bar{b}b\bar{b}$ at $\sqrt{s} = 189$ GeV. ALEPH Internal Note 99-040, 1999.