A Search for Neutral Higgs Bosons in the four Jet Channel using the DELPHI Detector at LEP

Doctor Scient. Thesis in Experimental Particle Physics

Jørgen Hansen

Department of Physics
University of Oslo
Norway

April 18, 2002

Abstract

A search for neutral Higgs bosons, both in the Standard Model and the Minimal Supersymmetric Standard Model, using the DELPHI detector at LEP with the data from the LEP runs in 1998, 1999, and 2000, is described. No signal was found, and 95% CL lower limits on the mass of the Higgs boson was found to be 114.3 GeV/c^2 in the SM case. For the MSSM case, limits are dependent on parameter choice and the specific MSSM benchmark, with the most conservative limits being 86.9 GeV/c^2 for m_{h^0}, 89.3 GeV/c^2 for m_{A^0}, and the minimal excluded range for the $\tan\beta$ parameter being 0.91–2.36.

Acknowledgements

I would like to thank first of all my supervisor, Prof. Lars Bugge, for continuous support and encouragement, for interesting and helpful discussions, for crucial insight into physics problems of importance for the analyses, and for cheerful enthusiasm. These are all valuable contributions to the work, and have helped greatly in making the analysis work a pleasant experience. I am also very grateful to Prof. Alexander L. Read for all the help and good advice he has given me, particularly concerning all aspects of statistics. I owe much to Vanina Ruhlmann-Kleider for instructing me on setting up the program for the final limit calculations, a very complex and difficult task that would have been very difficult, maybe impossible, without the guidance of an expert in the field. Among my fellow Ph. D. students, especially warm thanks go to Trond Myklebust, who, in his capacity as computer guru and Linux demigod, has helped me out of many technical difficulties. I would also like to express my gratitude to the rest of the staff and students at the Experimental Particle Physics group here at the University of Oslo, for creating a friendly environment full of challenges in the field of physics. Last, but not least, among my fellow physicists, I would like to thank all my collaborators in the DELPHI experiment for providing a highly competent and skilled environment in which challenging problems and interesting discussions are always in abundance.

My former room-mates, Petter Bøckman and Lene Aasland Kittelsen, also deserve thanks for having survived several years living together with me. Our co-existence has been intellectually stimulating, socially friendly and personally gratifying (not to mention financially *extremely* lucrative), and I would like to think that all our lives have been enriched by our five years together.

My parents, Ingrid Stensrud Hansen and Vidar Hansen, have contributed with continual support, both morally and financially (when needed), and with frequent questions and inquiries as to what, in layman's terms, I have been doing these years, something which has forced me to consider this very question carefully myself. This is, I feel, very important to all people working in a scientific environment, or indeed in any highly specialiced field, and one should make a serious effort to be able to explain at least some basic facts and principles about one's activities in a way understandable to most people willing to learn something about the subject. In this regard, my parents have played an essential part, in addition to important encouragement on a personal level, for which I thank them cordially.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Standard Model (SM) of particle physics correctly describes all current experimental results, and is a very successful theory of physics at the energies at which today's experiments operate. Ever since the top quark was observed at the Tevatron in 1995, the last missing component of the SM has been the Higgs boson. The particle content of the SM is shown in table 1.1, where matter is constructed from the fermions (hadrons, such as protons or neutrons, are made up from quarks, whereas the charged leptons can play the part of either individual particles, or by orbiting the protons and neutrons in stable matter structures known as atoms), whereas the vector bosons are the force carriers between the different particles subject to the force in question. The one scalar present in the SM, the Higgs boson, is intimately connected to the generation of what is known as the electro-weak symmetry breaking and the masses of the different particles of the theory. All the particles of table 1.1 have been experimentally seen, with the lone exception of the Higgs boson, which so far has proved to escape detection, most probably due to its large mass (or non-existence, in which case a different scheme for the electro-weak symmetry breaking and particle mass generation must be found).

Although not strictly necessary from the observations, the Higgs boson is an integral part of the model, as the Higgs mechanism is the most widely accepted way of addressing several theoretical issues associated with the SM. The Higgs boson is a direct consequence of the Higgs mechanism, a beautiful, but so far purely theoretical, construction, which allows for the assignment of masses to the fermions of the theory, as well as describing the electro-weak symmetry breaking, which gives masses to the Z^0, W^+ and W^- bosons. The mass of the Higgs boson is, however, a completely free parameter at tree-level, and the search for the Higgs boson has thus been conducted over a wide range of experiments over the years. Since 1995, LEP, the Large Electron-Positron collider at CERN, has been operating at continuously higher energies, as

Fermions			**Bosons**	
1st gen.	*Lepton*	*Quark*	*Vector bosons*	
up-type	electron-neutrino (ν_e)	up (u)	Force	Particle
down-type	electron (e)	down (d)	strong	gluon (g)
2nd gen.	*Lepton*	*Quark*		photon (γ)
up-type	muon-neutrino (ν_μ)	charm (c)	electro-weak	W^+W^-
down-type	muon (μ)	strange (s)		Z^0
3rd gen.	*Lepton*	*Quark*	*Scalar bosons*	
up-type	tau-neutrino (ν_τ)	top (t)	Higgs boson (H^0)	
down-type	tau (τ)	bottom (b)		

Table 1.1: The particle content of the SM. The force-carriers are all vector bosons (*i.e.* spin equals 1), whereas the Higgs boson is a scalar (*i.e.* spin equals 0).

the accelerator has undergone several upgrades. The resulting accelerator has usually been referred to as LEP-*II*, and has enabled the four LEP experiments, ALEPH, DELPHI, L3 and OPAL, to explore energy regions not earlier accessible to electron-positron colliders.

Even though the Higgs boson mass itself is a free parameter, both theoretical and experimental bounds do exist. Theoretical arguments give both upper and lower limits on the Higgs boson mass (such as the triviality and vacuum stability arguments, respectively, both treated more thoroughly in the next chapter, together with several other theoretical constraints), and thus reduces the possible values of the Higgs boson mass to a parameter range of which current experiments are able to probe a significant part. For the simplest case of the SM Higgs boson scenario, the one-doublet model, the upper limit on the Higgs boson mass is somewhere around 700 GeV/c^2, but for other Higgs scenarios, such as the two-doublet structure in the Minimal Supersymmetric Standard Model, significantly lower bounds, as low as $\sim$150 GeV/c^2, can be found [2, page 76]. Also, even though the Higgs boson mass is a free parameter, its value affects other measurable quantities through loop-diagrams. Thus, it is possible to make an indirect measurement of the Higgs boson mass, albeit with a fairly large error, as the aforementioned effect from loop diagrams are logarithmical in the mass parameter. The results of one such measurement, the fit of the Higgs boson mass to the electro-weak precision data, is shown in figure 1.1, which gives the SM Higgs boson mass as m_{H^0}=62^{+53}_{-39} GeV/c^2, with a 95% CL upper limit of 170 GeV/c^2 [1].

On this background, the searches for new physics beyond the SM, together with the investigations of experimentally missing or unclear parts of the

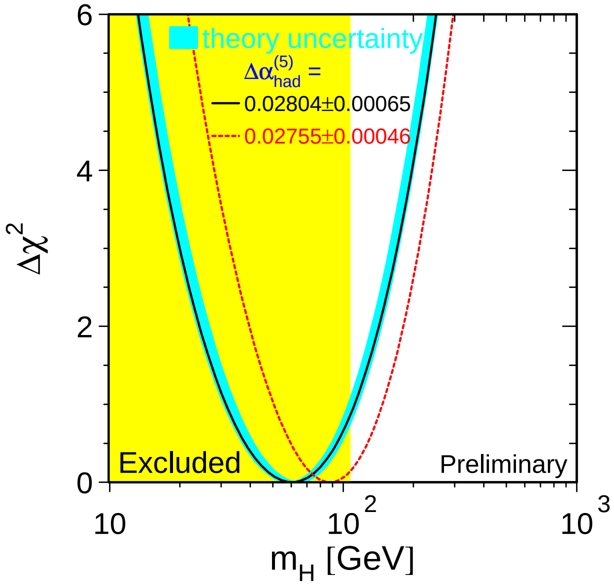


Figure 1.1: Plot showing the χ^2 of the fit of the Higgs boson mass to the electro-weak precision data. Figure taken from [1].

SM itself, have been some of the main activities at LEP-*II*. The search for the Higgs boson is one of the activities which has attracted most interest. Also, among the theories extending the SM, the Minimal Supersymmetric Standard Model (MSSM) has by many been considered the favourite model for a more fundamental theory. The MSSM is a supersymmetric theory (indeed, the simplest supersymmetric extension of the SM) and therefore introduces supersymmetric partners to all the particles in the SM. However, the Higgs sector of the theory must, due to the new supersymmetric structure, be composed of two complex scalar Higgs fields, instead of the one which is present in the SM. The result of this is an extended Higgs sector, which contains 5 Higgs bosons instead of the single scalar found in the SM.

In this work, the search for the Higgs boson was performed using the collected data from the DELPHI experiment during the runs of 1998, 1999 and 2000, totalling an amount of collected data of slightly less than 400 pb^{-1}. The search was performed in the fully hadronic channel, where the two heavy objects in the event both decay into a $q\overline{q}$ pair. The results are interpreted as mass limits on the Higgs boson(s) both in the SM and MSSM models.

The analysis presented here, developed by the author of this thesis, represents an independent search analysis method with respect to the standard DELPHI searches. Chapter 2 concerns the theoretical aspects of the Higgs boson; spontaneous symmetry breaking and the Higgs mechanism, radiative corrections to the Higgs mass and models beyond the simple one-doublet case, specifically the two-doublet model, which leads to the Minimal Supersymmetric Standard Model (MSSM). Chapter 3 deals with the Large Electron Positron collider (LEP) and the Detector with Lepton, Photon and Hadron Identification (DELPHI) detector. Chapter 4 describes the different backgrounds and signals, and discusses their most prominent features. The specific event selection cuts at the preselection level is presented in chapter 5 together with the studies of different track and event properties of the preselected events. Chapter 6 defines the analysis method, with the presentation of the repeated 2-dimensional likelihood network. This analysis method has been developed by the author, and is a central part of work. Here, the structure and underlying logic of the likelihood network is explained, as well as the task of 2-dimensional approximation of distributions. The actual implementation of the repeated 2-dimensional likelihood network in the presented analysis is detailed in chapter 7, with a description of the input classes and varibles, and their discriminating power with respect to different signals and backgrounds. Also, the different methods of Higgs mass estimation through the utilization of kinematic fits are outlined. The final results are presented in chapter 8, where the search results are interpreted in terms of excluded MSSM parameter regions, and the final conclusions are drawn.

Chapter 2

Theory

In this chapter, the theoretical motivation for the search for the Higgs boson(s) will be briefly described. In order to present an adequate picture of the Higgs mechanism, which leads to the existence of the Higgs boson, and its importance, a quick presentation of the Standard Model for the strong and electro-weak forces of nature will be given. The Standard Model consists of two distinct parts: QCD (Quantum Chromo Dynamics), which describes the strong forces acting between quarks and gluons, and the Glashow-Weinberg-Salam (GWS) theory of electro-weak interactions [3], linking together the weak force acting in nuclear decay and the electromagnetic force. These theories are formulated as quantum field theories with local gauge invariance, and the description of the SM therefore starts with a presentation of the principle of gauge invariance and its implications on the possible quantum field theories.

2.1 Gauge theories

Gauge theories are quantum field theories which, as well as being Lorentz-invariant, have an additional symmetry known as a gauge symmetry. This symmetry is built on representations of continuous symmetry groups (see for instance [4]), and requires the Lagrangian of the theory to be invariant under both local and global symmetry operations of the elements of the symmetry group chosen to be the gauge group of the theory. The symmetry groups are required to be simple (*i.e.* the group can not be subdivided into two mutually commuting sets of generators), compact (*i.e.* the algebra has finite-dimensional Hermitian representations and a finite number of generators) Lie groups. Such groups can be divided into three classes, with an additional five special groups. These are:

The $SU(N)$ groups: These groups are known as the unitary transformations of N-dimensional vectors, and preserve the inner product $\alpha_a^*\beta_a$ of two N-dimensional vectors α and β. The number of generators is N^2-1.

The $SO(N)$ groups: These groups are known as the orthogonal transformations of N-dimensional vectors, and preserve the symmetric inner product $\alpha_a\beta_a$ of two N-dimensional vectors α and β. The number of generators is $\frac{N(N-1)}{2}$.

The $Sp(N)$ groups: These groups are known as the symplectic transformations of N-dimensional vectors, and preserve the antisymmetric inner product $\epsilon_{\alpha\beta}\alpha_a\beta_b$ of two N-dimensional vectors α and β, where $\epsilon_{\alpha\beta}$ is the totally antisymmetric tensor. The number of generators is $\frac{N(N+1)}{2}$.

In addition to these three classes, there are five special groups denoted G_2, F_4, E_6, E_7 and E_8.

2.1.1 The simple $U(1)$ gauge group

When constructing quantum field theories to describe processes and interactions in particle physics, the principle of local gauge invariance has been found to be a very important guide to theories realized by nature. A classic example of this is the theory for electromagnetic interactions at particle level: consider the free electron field Lagrangian

$$\mathcal{L}_0 = \overline{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x). \tag{2.1}$$

This Lagrangian is invariant under a global $U(1)$ phase transformation; under such a phase transformation

$$\begin{array}{rcl} \psi(x) & \rightarrow & \psi'(x) = e^{i\alpha}\psi(x) \qquad \overline{\psi}(x) \;\; \rightarrow \;\; \overline{\psi}'(x) = e^{-i\alpha}\overline{\psi}(x) \\ \partial_\mu\psi(x) & \rightarrow & \partial_\mu\psi'(x) = e^{i\alpha}\partial_\mu\psi(x) \end{array} \tag{2.2}$$

and the Lagrangian in equation 2.1 is unchanged. However, when moving from global to local gauge invariance, *i.e.* when requiring that the parameter α is a function of the point x and not just a global number, the transformations in 2.2 change to

$$\begin{array}{rcl} \psi(x) & \rightarrow & \psi'(x) = e^{i\alpha(x)}\psi(x) \qquad \overline{\psi}(x) \;\; \rightarrow \;\; \overline{\psi}'(x) = e^{-i\alpha(x)}\overline{\psi}(x) \\ \partial_\mu\psi(x) & \rightarrow & \partial_\mu\psi'(x) = e^{i\alpha(x)}\partial_\mu\psi(x) + ie^{i\alpha(x)}\psi(x)\partial_\mu\alpha(x). \end{array} \tag{2.3}$$

As can be seen from equation 2.1, the Lagrangian is no longer invariant under these transformations, but rather

$$\mathcal{L}_0{}' = \mathcal{L}_0 + i\overline{\psi}(x)\gamma^\mu\psi(x)\partial_\mu\alpha(x). \tag{2.4}$$

Also, the expression $\partial_\mu\psi(x)$ is not strictly meaningful under a local gauge transformation, since the derivation process involves the subtraction of $\psi(x)$ at the point $x = x_0 + \epsilon$ from $\psi(x)$ at the point $x = x_0$, and the $U(1)$ phase transformation may well be different at these two points, so that the quantity $\partial_\mu\psi(x)$ has a priori no simple transformation law.

The simplest solution to this problem is to introduce a scalar quantity $U(y,x)$ which connects $\psi(x)$ at different points x and y. If this quantity is given the transformation

$$U(y,x) \to U'(y,x) = e^{i\alpha(y)}U(y,x)e^{-i\alpha(x)} \qquad U(y,y) = 1 \tag{2.5}$$

simultaneous with (2.3), the object $U(y,x)\psi(x)$ has the same transformation law as $\psi(y)$, given by (2.3), and the derivation $\partial_\mu\psi(x)$ can be given a meaningful definition by the *covariant derivative*

$$n^\mu D_\mu\psi(x) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\,\left[\psi(x+\epsilon n) - U(x+\epsilon n, x)\psi(x)\right] \tag{2.6}$$

where n^μ gives the direction in which the derivative of the field $\psi(x)$ is taken.

The object $U(y,x)$, which in general can be required to be a pure phase, can now be expanded in the separation of the two points in question:

$$U(x+\epsilon n, x) = 1 - ie\epsilon n^\mu A_\mu(x) + \mathcal{O}(\epsilon^2) \tag{2.7}$$

where the factor e is arbitrarily extracted. With this expansion, equation 2.6 can be written as

$$D_\mu\psi(x) = \partial_\mu\psi(x) + ieA_\mu(x)\psi(x) \qquad \Rightarrow \qquad D_\mu = \partial_\mu + ieA_\mu \tag{2.8}$$

which is the familiar expression for the covariant derivative in QED. To find how the field $A_\mu(x)$ transforms under this local gauge transformation, the equation 2.7 is inserted into equation 2.5, which gives

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x) \tag{2.9}$$

This definition of the covariant derivative ensures that it transforms under $U(1)$ gauge transformations in the same way that the field itself does, that is

$$D_\mu\psi(x) \to \left[\partial_\mu + ie(A_\mu(x) - \frac{1}{\epsilon}\partial_\mu\alpha(x))\right] e^{i\alpha(x)}\psi(x) = e^{i\alpha(x)}D_\mu\psi(x). \tag{2.10}$$

With the insertion of the covariant derivative D_μ instead of the ordinary derivative ∂_μ, the Lagrangian in equation 2.1 takes the form

$$\mathcal{L}_0 = \overline{\psi}(x)(i\gamma^\mu D_\mu - m)\psi(x) = \overline{\psi}(x)(i\gamma^\mu(\partial_\mu + ieA_\mu(x)) - m)\psi(x). \quad (2.11)$$

In order to write down the most general Lagrangian allowed by the local $U(1)$ gauge symmetry, a term including only the field $A_\mu(x)$ and its derivatives must be included. By considering the expansion in 2.7 a further order in perturbation, and in addition taking the relation $U(y,x)^\dagger = U(x,y)$ into account, the expansion can be written as

$$U(x+\epsilon n, x) = \exp\left[-ie\epsilon n^\mu A_\mu(x+\frac{\epsilon}{2}n) + \mathcal{O}(\epsilon^3)\right]. \quad (2.12)$$

Defining the quantity $\mathbf{U}(x)$ as the product of four $U(y,x)$s at different points, chosen as

$$\begin{aligned}\mathbf{U}(x) \equiv U(x, x+\epsilon\hat{2}) \times U(x+\epsilon\hat{2}, x+\epsilon\hat{1}+\epsilon\hat{2}) \\ \times U(x+\epsilon\hat{1}+\epsilon\hat{2}, x+\epsilon\hat{1}) \times U(x+\epsilon\hat{1}, x)\end{aligned} \quad (2.13)$$

where $\hat{1}(\hat{2})$ is the unit vector in the 1(2)-direction, ensures, by the transformation law 2.5, the local invariance of $\mathbf{U}(x)$. Expanding this equation by the use of 2.12 gives the expression

$$\begin{aligned}\mathbf{U}(x) = \exp\Big\{-i\epsilon e\big[-A_2(x+\frac{\epsilon}{2}\hat{2}) - A_1(x+\frac{\epsilon}{2}\hat{1}+\epsilon\hat{2}) \\ + A_2(x+\epsilon\hat{1}+\frac{\epsilon}{2}\hat{2}) + A_1(x+\frac{\epsilon}{2}\hat{1})\big] + \mathcal{O}(\epsilon^3)\Big\}\end{aligned} \quad (2.14)$$

which, when expanded in powers of ϵ, gives the following expression for $\mathbf{U}(x)$:

$$\mathbf{U}(x) = 1 - i\epsilon^2 e\left[\partial_1 A_2(x) - \partial_2 A_1(x)\right] + \mathcal{O}(\epsilon^3). \quad (2.15)$$

Since the definition of $\mathbf{U}(x)$ requires it to be locally invariant, the structure $\partial_1 A_2(x) - \partial_2 A_1(x)$ must also be locally invariant. When generalizing this argument to a general comparison of phases in an arbitrarily chosen round trip, and not just a specific one defined by the unit vectors $\hat{1}$ and $\hat{2}$, one finds that the locally invariant structure has the form

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x). \quad (2.16)$$

Another way of proving the invariance of $F_{\mu\nu}(x)$, is to consider the commutator of two different covariant derivatives. Since the covariant derivative has the same transformation as the field itself, the second covariant derivative also transforms in the same way, which implies

$$[D_\mu, D_\nu]\psi(x) \to e^{i\alpha(x)}[D_\mu, D_\nu]\psi(x). \tag{2.17}$$

On the other hand, when considering the expressions for the covariant derivative itself, the result is

$$\begin{aligned}
[D_\mu, D_\nu]\psi(x) &= [\partial_\mu, \partial_\nu]\psi(x) + ie([\partial_\mu, A_\nu(x)] - \\
&\qquad [\partial_\nu, A_\mu(x)])\psi(x) - e^2[A_\mu(x), A_\nu(x)]\psi(x) \\
&= ie(\partial_\mu A_\nu(x) - \partial_\nu A_\mu(x))\psi(x) \\
&= ieF_{\mu\nu}(x)\psi(x).
\end{aligned} \tag{2.18}$$

Thus, when comparing with equation 2.17, the construction $F_{\mu\nu}(x)$ is seen to be invariant under this local gauge transformation.

If we now want to construct the most general $U(1)$ locally (and globally) invariant Lagrangian, terms depending on $A_\mu(x)$ only through $F_{\mu\nu}(x)$ and its derivatives is invariant, and should be included in the Lagrangian. Up to operators of order 4 (and operators of higher orders prove to be nonrenormalizable), the only possible terms in the Lagrangian are therefore

$$\mathcal{L}_4 = \overline{\psi}(x)(i\gamma^\mu D_\mu)\psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x) - ic\epsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}(x)F_{\mu\nu}(x) - m\overline{\psi}(x)\psi, \tag{2.19}$$

where c is an arbitrary constant. The third term in this Lagrangian violates the discrete symmetries P (parity transformation) and T (time reversal), and should be excluded in a theory which postulates these symmetries. If doing so, the remaining Lagrangian is the familiar QED Lagrangian

$$\mathcal{L} = \overline{\psi}(x)(i\gamma^\mu D_\mu - m)\psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x). \tag{2.20}$$

Thus, by requiring local $U(1)$ phase invariance, the existence of the field $A_\mu(x)$, identified with the photon field, has been required in order to keep the free electron Lagrangian invariant, and has indeed led to the well-known expression for the QED Lagrangian. The principle of local gauge invariance has been shown to be very successful in the task of describing the theory of electromagnetic forces, and would therefore be expected to play an important role in the description of other interactions between particles. However, so far only the simple phase rotations of the $U(1)$ group has been considered, and in order to describe the more complex structures of the weak and strong interactions, other, and larger, groups have to be considered.

2.1.2 Yang-Mills gauge theories

When generalizing the arguments of the previous section to more complex gauge groups, specifically all continuous symmetry groups, the resulting theories are called Yang-Mills theories, after Yang and Mills, who were the first to propose this idea. However, when introducing more complex groups, there are several new features and problems arising when constructing a lagrangina which is locally invariant under such groups. Consider first a set of $n \times n$ unitary matrices as a representation of a general symmetry group. Let the basic field be an n-plet $\psi(x)$ which transforms under the local symmetry as

$$\psi(x) \to \psi'(x) = V(x)\psi(x). \tag{2.21}$$

$V(x)$ can now be expanded by the generators of the group as

$$V(x) = 1 + i\alpha^a(x)t^a + \mathcal{O}(\alpha^2) \tag{2.22}$$

where the group generators, represented by Hermitian matrices, are denoted t^a. These matrices are related by the structure constants f^{abc} of the group through the relations

$$[t^a, t^b] = if^{abc}t^c. \tag{2.23}$$

This result is one of the most important new features of Yang-Mills theories as opposed to QED. It shows that, due to the fact that there are several different generators of the group in the general case, and not just the single phase of the $U(1)$ transformation, there is a possibility of non-commuting group generators. This will become apparent when presenting the general expressions for the field tensor and the Lagrangian.

The covariant derivative associated with the general gauge transformation 2.21 is now

$$D_\mu = \partial_\mu - igA^a_\mu(x)t^a \tag{2.24}$$

which introduces one new vector field for each generator of the group, and ensures that the expression $D_\mu\psi(x)$ transforms as the field $\psi(x)$ itself. The corresponding infinitesimal transformation laws for the basic field $\psi(x)$ and the vector field $A^a_\mu(x)$ are:

$$\begin{aligned} \psi(x) &\to \psi'(x) = (1 + i\alpha^a(x)t^a)\psi \\ A^a_\mu(x) &\to A^{a\prime}_\mu(x) = A^a_\mu(x) + \tfrac{1}{g}\partial_\mu\alpha^a(x) + f^{abc}A^b_\mu(x)\alpha^c(x). \end{aligned} \tag{2.25}$$

The analogoue to the QED field tensor $F_{\mu\nu}(x)$ is now found by the relation

$$[D_\mu, D_\nu] = -igF^a_{\mu\nu}(x)t^a \tag{2.26}$$

which corresponds to the more explicit formula

$$F^a_{\mu\nu}(x) = \partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) + g f^{abc} A^b_\mu(x) A^c_\nu(x). \tag{2.27}$$

When comparing with the corresponding expression for the QED field tensor in equation 2.16, there is a new term containing the square of the $A^a_\mu(x)$ field and the structure constants of the group. This term is a direct consequence of the noncommuting nature of a general gauge group, as seen in equation 2.23.

The general Lagrangian is now, as in the case of the $U(1)$ gauge group, generated from the locally invariant terms found, and, if restricting the Lagrangian to contain only renormalizable terms which preserve the P and T symmetries, the result is very close in appearance to the $U(1)$ gauge group expression:

$$\mathcal{L} = \overline{\psi}(x)(i\gamma^\mu D_\mu - m)\psi(x) - \frac{1}{4} F^i_{\mu\nu}(x) F^{\mu\nu i}(x). \tag{2.28}$$

However, the nature of this Lagrangian is considerably more complex than the one presented in equation 2.20. Due to the additional term in the field strength, this Lagrangian contains, through the term $\frac{1}{4} F^i_{\mu\nu}(x) F^{\mu\nu i}(x)$, the possibility to have interactions between the vector $A^a_\mu(x)$ and itself. This feature of the Yang-Mills theories is a completely new development which has no analogy in QED, and is indeed present in several of the parts making up the Standard Model, most notably the self-interaction between the gluons of QCD.

The foregoing analysis accounts for the structure of a massive field $\psi(x)$ interacting with a massless vector boson field $A^a_\mu(x)$. Thus, the method of local gauge invariance seems, and has indeed been demonstrated, to describe theories with massless vector bosons, such as QED or QCD, very well. But, as can be seen from all the Lagrangians in equations 2.19, 2.20 and 2.28, the vector bosons are required to be massless, since an explicit mass term would break the local gauge symmetry. In order to have the opportunity to include such terms, as well as fermion mass terms, the mechanism and technique of spontaneous symmetry breaking is required.

2.2 Spontaneous symmetry breaking

Another important ingredient in the Standard Model of electro-weak interactions, is the concept of spontaneous symmetry breaking, which will provide a way of generating masses for the gauge bosons. This is crucial, since experiments show (see [5, 6]) that, although the gauge bosons of the electromagnetic force (the photon, γ) and the strong nuclear force (the gluons, g) are massless, the gauge bosons responsible for transmitting the weak interactions (the W^+, W^- and Z^0) are massive. Indeed, it is the large masses of these bosons that ensures the weakness and short range of the weak interactions.

As an illustration, consider the two cases in figure 2.1. Both show the function

$$f(\phi_1, \phi_2) = \frac{\mu^2}{2}\left(\phi_1^2 + \phi_2^2\right) + \frac{\lambda}{4}\left(\phi_1^2 + \phi_2^2\right)^2$$

the left part (a) with the parameter μ^2 positive, the right part (b) with μ^2 negative, λ being positive in both cases. For μ^2 positive, the global minimum is at $\phi_1 = \phi_2 = 0$, which is a point symmetrically placed with respect to the function $f(\phi_1, \phi_2)$. However, for μ^2 negative, the global minimum is not at a single point, but rather at a circle in the $\phi_1\phi_2$ plane, obeying the equation

$$\phi_{1,0}^2 + \phi_{2,0}^2 = v^2 = -\frac{\mu^2}{\lambda} \quad \Rightarrow \quad f(\phi_{1,0}, \phi_{2,0}) = -\frac{\mu^4}{4\lambda}. \tag{2.29}$$

If we now select one single point in this global minimum, *i.e.* any one pair of $\phi_{1,0}, \phi_{2,0}$ obeying equation 2.29, this point is not symmetrically placed with respect to the function $f(\phi_1, \phi_2)$, as is the case for the global minimum in the positive μ^2 case. Therefore, the point of global minimum of the function in part (a) of figure 2.1 exhibits a symmetry which is not present for any one point of the global minimum of the function in part (b); the symmetry has been broken.

When applying this idea to quantum field theories, consider first, as an illustration and a simple example, a complex scalar field $\phi(x) = \frac{1}{\sqrt{2}}(\phi_1(x) + i\phi_2(x))$. The Lagrangian of the system is

$$\begin{aligned} \mathcal{L} &= (\partial_\mu \phi)^*(\partial_\mu \phi) - \mu^2 \phi^* \phi - \lambda(\phi^* \phi)^2 \\ &= \tfrac{1}{2}(\partial_\mu \phi_1)^2 + \tfrac{1}{2}(\partial_\mu \phi_2)^2 - \tfrac{\mu^2}{2}(\phi_1^2 + \phi_2^2) - \tfrac{\lambda}{4}\left(\phi_1^2 + \phi_2^2\right)^2 \end{aligned} \tag{2.30}$$

and $f(\phi_1, \phi_2)$ is seen to appear as the potential of the system. As shown in 2.29, the minimum of the potential is a circle in the $\phi_1\phi_2$ plane. In order to expand the field ϕ around a minimum, we arbitrarily choose this minimum to be at the point

$$\phi_{1,0} = v \quad \phi_{2,0} = 0 \quad \Rightarrow \quad \phi_0 = \frac{v}{\sqrt{2}} = \sqrt{\frac{-\mu^2}{2\lambda}} \tag{2.31}$$

making the ground state ϕ_0 entirely real-valued, and make a change of variables:

$$\eta(x) = \phi_1(x) - v \quad \xi(x) = \phi_2(x) \quad \Rightarrow \quad \phi(x) = \frac{1}{\sqrt{2}}(v + \eta(x) + i\xi(x)) \tag{2.32}$$

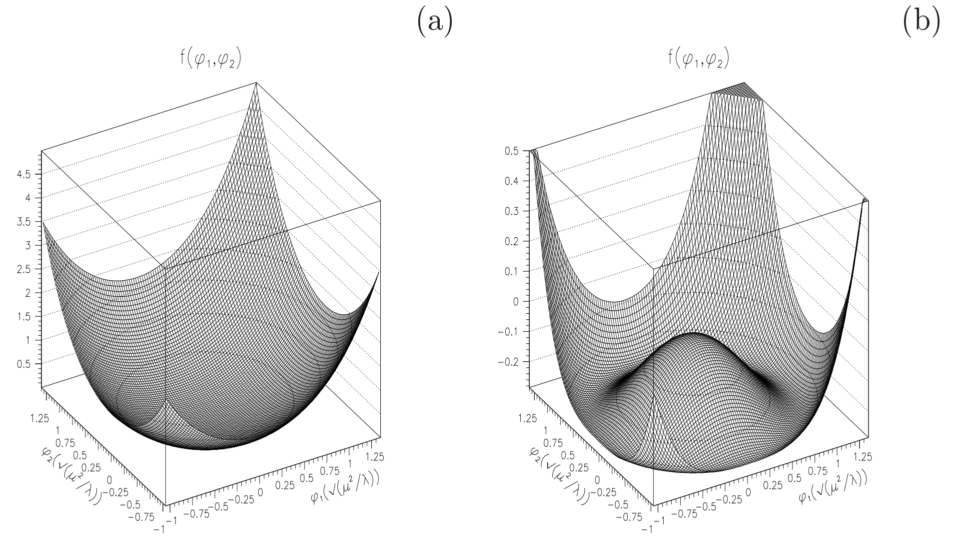


Figure 2.1: An example of a spontaneously broken symmetry: The function $f(\phi_1,\phi_2) = \frac{\mu^2}{2}\left(\phi_1^2+\phi_2^2\right)+\frac{\lambda}{4}\left(\phi_1^2+\phi_2^2\right)^2$. In (a), corresponding to an unbroken symmetry, the value of μ^2 is positive, whereas (b), corresponding to a broken symmetry, shows a case with negative value of μ^2.

With the substitutions in 2.32, and by making use of the relation 2.29 to eliminate λ from the expressions, the Lagrangian of the system can now be written as

$$\begin{aligned}\mathcal{L} = \tfrac{1}{2}(\partial_\mu\eta)^2 + \tfrac{1}{2}(\partial_\mu\xi)^2 + \mu^2\eta^2 - \tfrac{\mu^2 v^2}{4} \\ +\tfrac{\mu^2}{v}\eta^3 + \tfrac{\mu^2}{v}\eta\xi^2 + \tfrac{\mu^2}{4v^2}\eta^4 + \tfrac{\mu^2}{2v^2}\eta^2\xi^2 + \tfrac{\mu^2}{4v^2}\xi^4.\end{aligned} \tag{2.33}$$

This Lagrangian is now no longer symmetric in the variables η and ξ, as should be expected from the choice of ground state in 2.31. The third term in the Lagrangian is now an explicit mass term for the field $\eta(x)$ $(-\frac{1}{2}m_\eta^2\eta^2 = \mu^2\eta^2 \Rightarrow m_\eta = \sqrt{-2\mu^2})$, whereas no such term exists for the $\xi(x)$ field. There are also several terms describing interactions between the two $\eta(x)$ and $\xi(x)$ fields.

At first glance, this seems somewhat unreasonable; the physics of the system has apparently been completely changed by a simple change of variables. This is not the case, though. The Lagrangian 2.30 and 2.33 describe the same physical system, and if used in an exact calculation, they would yield the same physical quantities. However, when doing quantum field theory, one is forced to apply perturbation theory, and therefore perform the calculations in the vicinity of a local minimum. Therefore, the choice of the variables $\eta(x)$ and $\xi(x)$ are appropriate in the case of $\mu^2 < 0$, since the point ϕ_0 is such a local minimum in this scenario.

On this basis, it is also easy to understand why the $\eta(x)$ field has a mass term in the Lagrangian, whereas the $\xi(x)$ field is massless. The field $\eta(x)$ is directed in the purely real direction of the plane ($\eta(x) = \sqrt{2} \cdot \Re\{\phi(x)\}$), while the field $\xi(x)$ is directed in the purely imaginary direction of the plane ($\xi(x) = \sqrt{2} \cdot \Im\{\phi(x)\}$). At the point ϕ_0, the potential in the imaginary direction, *i.e.* in the direction of $\xi(x)$, is non-changing and equal to the minimum value of the potential. Thus, there is no resistance in the direction of the $\xi(x)$ field, and the result is a massless mode. The other part of the $\phi(x)$ field, $\eta(x)$, points in a direction of the $\phi_1\phi_2$ plane which is a potential well, and therefore describes a massive mode.

The existence of a massless particle in connection with a broken symmetry is not by any means a coincidence. In fact, this is guaranteed by Goldstone's theorem [7], which states that for every spontaneously broken continuous symmetry, the theory contains a massless particle. This might seem not to be very useful in attaining the original goal of spontaneous symmetry breaking, the generation of massive gauge bosons. But, as shall shortly be seen, when combining the two concepts of local gauge invariance and spontaneous symmetry breaking, the goal of massive gauge bosons will be achieved. In fact, the massless modes which accompany the broken symmetries, is an integral part of this whole mechanism, which is called the Higgs mechanism.

2.3 The Higgs mechanism

To illustrate the technique known as the Higgs mechanism [8], consider first a simple model containing a complex scalar field coupled both to itself and to an electromagnetic, or gauge, field. The Lagrangian has the form

$$\mathcal{L} = (D_\mu\phi(x))^*(D_\mu\phi(x)) - \frac{1}{4}(F_{\mu\nu}(x)^2) - V(\phi(x)), \tag{2.34}$$

where the field $\phi(x)$ can again be decomposed as $\phi(x) = \frac{1}{\sqrt{2}}(\phi_1(x) + i\phi_2(x))$. Letting the local gauge group be the simple $U(1)$ group, gives the standard transformations and covariant derivatives as seen earlier in equations 2.3, 2.9 and 2.8; $\psi(x) \to e^{i\alpha(x)}\psi(x)$, $A_\mu(x) \to A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x)$ and $D_\mu = \partial_\mu + ieA_\mu$. Also, let the form of $V(x)$ be the same as in 2.30, $V(\phi) = \mu^2\phi^*\phi + \lambda(\phi^*\phi)^2$ with μ^2 negative, and make the change of variables as in equation 2.32, $\phi(x) = \frac{1}{\sqrt{2}}(v + \eta(x) + i\xi(x))$. This gives the same Lagrangian for the $\eta(x)$ and $\xi(x)$ fields as in 2.33, and so far, no new development has taken place.

However, the Lagrangian no longer contains only the simple derivative ∂_μ, but rather the more complex covariant derivative D_μ. When explicitly computing the kinetic energy term, the result is

$$\begin{aligned}(D_\mu\phi)^*(D_\mu\phi) = \tfrac{1}{2}(\partial_\mu\eta)^2 + \tfrac{1}{2}(\partial_\mu\xi)^2 + \tfrac{1}{2}e^2v^2A_\mu A^\mu + evA_\mu\partial^\mu\xi \\ +e\eta A_\mu\partial^\mu\xi - e\xi A_\mu\partial^\mu\eta + e^2v\eta A_\mu A^\mu + \tfrac{1}{2}e^2\left(\eta^2+\xi^2\right)A_\mu A^\mu.\end{aligned} \tag{2.35}$$

The first two terms in this expression are the kinetic terms for the $\eta(x)$ and $\xi(x)$ field, and are also found in the Lagrangian 2.33. The third term, however, is a new construction not present in 2.33, and clearly describes a mass for the vector field $A_\mu(x)$ ($\frac{1}{2}m_\gamma^2 = e^2v^2 \Rightarrow m_\gamma = \sqrt{2}ev$), *i.e.* an explicit photon mass.

With these expressions for the different parts of the Lagrangian, and keeping in mind the expression for the field tensor $F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$, the form of the Lagrangian is

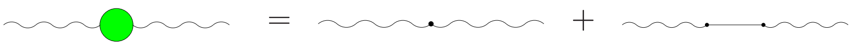
$$\mathcal{L} = \mathcal{L}_k + \mathcal{L}_m + \mathcal{L}_{\text{int1}} + \mathcal{L}_{\text{int2}}$$

where

$$\begin{aligned}\mathcal{L}_k &= -\tfrac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 + \tfrac{1}{2}(\partial_\mu\eta)^2 + \tfrac{1}{2}(\partial_\mu\xi)^2 - \tfrac{\mu^2v^2}{4} \\ \mathcal{L}_m &= \mu^2\eta^2 + \tfrac{1}{2}e^2v^2A_\mu A^\mu \\ \mathcal{L}_{\text{int1}} &= \tfrac{\mu^2}{v}\eta^3 + \tfrac{\mu^2}{v}\eta\xi^2 + \tfrac{\mu^2}{4v^2}\eta^4 + \tfrac{\mu^2}{2v^2}\eta^2\xi^2 + \tfrac{\mu^2}{4v^2}\xi^4 \\ \mathcal{L}_{\text{int2}} &= evA_\mu\partial^\mu\xi + e\eta A_\mu\partial^\mu\xi - e\xi A_\mu\partial^\mu\eta \\ &\quad +e^2v\eta A_\mu A^\mu + \tfrac{1}{2}e^2\left(\eta^2+\xi^2\right)A_\mu A^\mu.\end{aligned} \tag{2.36}$$

In this Lagrangian, $\mathcal{L}_k$ contains the kinetic terms for the $\eta(x)$, $\xi(x)$ and $A_\mu(x)$ fields, $\mathcal{L}_m$ contains the mass terms for the $\eta(x)$ and $A_\mu(x)$ fields, and $\mathcal{L}_{\text{int1}}$ and $\mathcal{L}_{\text{int2}}$ contains the interactions of the $\phi(x)$ field, and between the $A_\mu(x)$ and any part of the $\phi(x)$ field, respectively.

One may wonder about the role of the Goldstone bosons in this analysis, represented by the field $\xi(x)$. It can be shown [9, page 691] that a gauge boson cannot obtain a mass unless this mass term is associated with a pole in the vacuum polarization amplitude, which in four dimensions can be created only by a massless scalar particle. This role is filled perfectly by the Goldstone boson, and can even be seen directly in the Lagrangian. The first term of $\mathcal{L}_{\text{int2}}$ is a coupling between the gauge boson field $A_\mu(x)$ and the Goldstone field $\xi(x)$. If one performs perturbation theory, and treats the second term of $\mathcal{L}_m$, the gauge boson mass term, as a vertex, a leading-order calculation of the vacuum polarization amplitude yields the result



$$
\begin{aligned}
&= i m_\gamma^2 g^{\mu\nu} \quad + \quad (m_\gamma k^\mu) \tfrac{i}{k^2} (-m_\gamma k^\nu) \\
&= i m_\gamma^2 \left(g^{\mu\nu} - \tfrac{k^\mu k^\nu}{k^2} \right)
\end{aligned}
\qquad (2.37)
$$

and the result is seen to be properly transverse.

Even though the Goldstone boson is seen to play an essential role in this mechanism, the Goldstone boson itself does not appear as an independent physical particle. This can be seen by choosing a particular gauge called the unitary gauge, where the $U(1)$ parameter $\alpha(x)$ is chosen in such a way that $\phi(x)$ becomes real-valued, making the field $\phi_2(x)$ disappear from the Lagrangian 2.34, but retaining all the other characteristics of the theory. Thus, the role of the Goldstone boson has been reduced to supplying the means necessary for producing a massive gauge boson. Also, such a massive gauge boson has three physical polarization degrees of freedom, whereas a massless boson only has two. This makes it tempting to adopt the viewpoint that the gauge boson acquires its mass, and thereby its additional degree of freedom, by eating the Goldstone boson.

2.4 The Standard Model (SM)

The standard model of electro-weak interactions, also called the GSW model after its inventors Glashow, Weinberg and Salam, is a spontaneously broken gauge theory. The model has been shown to be very successful, and describes all experimental data well. The theory utilizes all the mechanisms and techniques of local gauge invariance, spontaneous symmetry breaking and the Higgs mechanism.

2.4.1 Gauge boson masses, and the Higgs boson

Experiments show [5, 6] that while there are three massive gauge bosons (the weak interaction vector bosons W^+, W^- and Z^0), one gauge boson remains massless (the electromagnetic gauge boson, the photon, γ). The gauge group is therefore chosen to be $SU(2)$ with an additional $U(1)$ group, such that the field $\phi(x)$, which is a scalar field in the spinor representation, transforms as

$$\phi(x) \to \phi'(x) = e^{i\alpha^a \tau^a} e^{i\beta/2} \phi(x) \qquad \text{with} \qquad \tau^a = \sigma^a/2 \qquad (2.38)$$

where the matrices σ^a can be represented by the Pauli matrices. This gives the covariant derivative for the field $\phi(x)$ as

$$D_\mu \phi(x) = \left(\partial_\mu - igA_\mu^a(x)\tau^a - i\frac{g'}{2}B_\mu(x) \right) \phi(x) \tag{2.39}$$

where the $A_\mu^a(x)$ and $B_\mu(x)$ fields are the $SU(2)$ and $U(1)$ gauge fields, respectively.

The spontaneous symmetry breaking is now achieved by the field $\phi(x)$ acquiring a vacuum expectation value in the presence of a potential $V(\phi)$ similar to the one in equation 2.30,

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \quad \Rightarrow \quad \langle\phi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \; v = \sqrt{\frac{-\mu^2}{\lambda}}, \tag{2.40}$$

with $\mu^2 < 0$ and $\lambda > 0$. The field $\phi(x)$ is now written as a perturbation around this value:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.41}$$

When using the expression 2.41 for the field $\phi(x)$, and the standard expressions for the Pauli matrices $\sigma^a = 2\tau^a$, the Lagrangian of the system corresponding to the interaction of the scalar Higgs field $\phi(x)$ with the gauge fields $A_\mu^a(x)$ and $B_\mu(x)$ is

$$\begin{aligned} \mathcal{L} &= (D^\mu\phi)^\dagger(D_\mu\phi) - V(\phi) \\ &= \left(\partial^\mu\phi^\dagger + ig\phi^\dagger A^{a\mu}\tau^a + i\tfrac{g'}{2}\phi^\dagger B^\mu\right)\left(\partial_\mu\phi - igA_\mu^b\tau^b\phi - i\tfrac{g'}{2}B_\mu\phi\right) \\ &\quad -\mu^2\phi^\dagger\phi - \lambda\left(\phi^\dagger\phi\right)^2 \\ &= \mathcal{L}_{AB} + \mathcal{L}_{AB\phi} + \mathcal{L}_\phi \end{aligned}$$

where

$$\begin{aligned} \mathcal{L}_{AB} &= \phi^\dagger \left\{ g^2 A^{a\mu}\tau^a A_\mu^b\tau^b + \tfrac{g'^2}{4}B^\mu B_\mu + \tfrac{gg'}{2}\left(A^{a\mu}\tau^a B_\mu + B^\mu A_\mu^a\tau^a\right) \right\} \phi \\ &= \tfrac{1}{2}\tfrac{v^2}{4}\left(g^2\left((A_\mu^1)^2 + (A_\mu^2)^2\right) + (g'B_\mu - gA_\mu^3)^2\right) \\ \mathcal{L}_{AB\phi} &= ig\left(\phi^\dagger A^{a\mu}\tau^a\partial_\mu\phi - \partial^\mu\phi^\dagger A_\mu^b\tau^b\phi\right) + i\tfrac{g'}{2}\left(\phi^\dagger B^\mu\partial_\mu\phi - \partial^\mu\phi^\dagger B_\mu\phi\right) \\ &= \tfrac{1}{2}\tfrac{2vh+h^2}{4}\left(g^2\left((A_\mu^1)^2 + (A_\mu^2)^2\right) + (gA_\mu^3 - g'B_\mu)^2\right) \\ \mathcal{L}_\phi &= (\partial^\mu\phi^\dagger)(\partial_\mu\phi) - \mu^2\phi^\dagger\phi - \lambda\left(\phi^\dagger\phi\right)^2 \\ &= \tfrac{1}{2}(\partial_\mu h)^2 - \tfrac{\mu^2 v^2}{4} + \mu^2 h^2 + \tfrac{\mu^2}{v}h^3 + \tfrac{\mu^2}{4v^2}h^4 \end{aligned} \tag{2.42}$$

From $\mathcal{L}_{AB}$ one can now see that there are three massive and one massless gauge bosons fields:

$$\begin{array}{rclcrcl} W^{\pm}_{\mu} & = & \frac{1}{\sqrt{2}}(A^1_{\mu} \mp iA^2_{\mu}) & \text{mass} & m_W & = & \frac{gv}{2} \\ Z^0_{\mu} & = & \frac{1}{\sqrt{g^2+g'^2}}(gA^3_{\mu} - g'B_{\mu}) & \text{mass} & m_Z & = & \frac{\sqrt{g^2+g'^2}v}{2} \\ A_{\mu} & = & \frac{1}{\sqrt{g^2+g'^2}}(g'A^3_{\mu} + gB_{\mu}) & \text{mass} & m_{\gamma} & = & 0 \end{array} \tag{2.43}$$

With these new gauge boson fields, the final form of the Lagrangian concerning the mass terms for the gauge bosons is

$$\mathcal{L}_{AB} = \frac{1}{2}m^2_W W^+_{\mu} W^{+\mu} + \frac{1}{2}m^2_W W^-_{\mu} W^{-\mu} + \frac{1}{2}m^2_Z Z^0_{\mu} Z^{0\mu} \tag{2.44}$$

It is now convenient to rewrite the covariant derivative in equation 2.39 in terms of these gauge boson mass eigenstate fields. Also, it will become necessary to have the expression for the covariant derivative for a general $SU(2)$ and $U(1)$ representation, and not just the spinor notation given by the Pauli matrices. This general form of the covariant derivative can now be expressed as

$$\begin{array}{rcl} D_{\mu} & = & \partial_{\mu} - igA^a_{\mu}(x)T^a - ig'YB_{\mu}(x) \\ & = & \partial_{\mu} - i\frac{g}{\sqrt{2}}\left(W^+_{\mu}T^+ + W^-_{\mu}T^-\right) - i\frac{1}{\sqrt{g^2+g'^2}}Z^0_{\mu}\left(g^2T^3 - g'^2Y\right) \\ & & -i\frac{gg'}{\sqrt{g^2+g'^2}}A_{\mu}\left(T^3 + Y\right), \end{array} \tag{2.45}$$

where T^a is the general representation of the $SU(2)$ generators, Y is the $U(1)$ charge, and $T^{\pm} = T^1 \pm iT^2$.[1] In equation 2.45, the field $A_{\mu}(x)$ is associated with the photon, and the coefficients of the last term of equation 2.45 should therefore be associated with the electromagnetic interaction, *i.e.* the electron charge e and the electric charge quantum number are given by the relations

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}} \qquad \text{and} \qquad Q = T^3 + Y. \tag{2.46}$$

It is also customary to introduce the weak mixing angle θ_W, also called the Weinberg angle, by rewriting the mixing between the gauge fields $A^3_{\mu}(x)$ and $B_{\mu}(x)$ resulting in the physical fields $Z^0_{\mu}(x)$ and $A_{\mu}(x)$ in the following way:

$$\begin{pmatrix} Z^0_{\mu}(x) \\ A_{\mu}(x) \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} A^3_{\mu}(x) \\ B_{\mu}(x) \end{pmatrix} \tag{2.47}$$

---

[1] In the spinor representation used so far, $T^a = \tau^a = \frac{1}{2}\sigma^a$.

where θ_W is given by the equation

$$\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}} = \frac{m_Z}{m_W} \quad \Rightarrow \quad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}. \tag{2.48}$$

By inserting equations 2.48 and 2.46 into 2.45, the final version of the covariant derivative takes the form

$$\begin{aligned} D_\mu &= \partial_\mu - ie\Big(\tfrac{1}{\sqrt{2}\sin\theta_W}\left(W_\mu^+T^+ + W_\mu^-T^-\right) \\ &\qquad + \tfrac{1}{\cos\theta_W\sin\theta_W}Z_\mu^0\left(T^3 - \sin^2\theta_W Q\right) + A_\mu Q\Big) \end{aligned} \tag{2.49}$$

Thus, the couplings of the electro-weak bosons are described by only two parameters, the gauge field coupling constants g and g', or, equivalently, the electric charge e and the weak mixing angle θ_W. It is also worth mentioning that the value of v is measurable from Fermi's effective theory of weak interactions, and gives

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} = \frac{g^2}{8(\frac{g^2v^2}{4})^2} \quad \Rightarrow \quad v = \left(\sqrt{2}G_F\right)^{-\frac{1}{2}} = 246 \text{ GeV} \tag{2.50}$$

The partial Lagrangian $\mathcal{L}_\phi$ of equation 2.42 shows that the theory contains a boson described by the field $h(x)$ with a mass given by

$$\mu^2h^2 = -\frac{1}{2}m_h^2h^2 \qquad \Rightarrow \qquad m_h = \sqrt{-2\mu^2} = \sqrt{2\lambda}v. \tag{2.51}$$

This is the Higgs boson, which is a scalar particle with a mass given by a combination of the parameters v (which is also found in the expressions for m_W and m_Z) and a new coupling constant λ.

2.4.2 Fermion couplings and masses

Fermion couplings

It is now time to consider the fermionic mass fields that make up the physical world. Note first that experiments show that charged weak currents violate parity (see [10]). This makes it natural to make the straightforward decomposition of the fermionic fields into their left- and right-handed components:

$$\begin{gathered} \psi = \psi_L + \psi_R \quad \Rightarrow \\ \overline{\psi}\gamma^\mu\psi = \overline{\psi}_L\gamma^\mu\psi_L + \overline{\psi}_R\gamma^\mu\psi_R, \qquad \overline{\psi}\gamma^5\psi = \overline{\psi}_L\gamma^5\psi_R + \overline{\psi}_R\gamma^5\psi_L \end{gathered} \tag{2.52}$$

Also, the fact that the W^+ and W^- couple differently to right-handed and left-handed components, makes it necessary to assign these components to different representations of $SU(2)$, and with different $U(1)$ charges Y. The left-handed fields are assigned to doublets of $SU(2)$, whereas the right-handed fields are singlets, and the structure of the fermionic fields is therefore[2]

$$E_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \; Q_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \; e_R, \; \nu_R, \; u_R, \; d_R, \tag{2.53}$$

where the subscripts L and R refer to the left- and right-handed projections respectively, $\psi_L = \frac{1}{2}(1-\gamma^5)\psi$, $\psi_R = \frac{1}{2}(1+\gamma^5)\psi$. This pattern of fields is repeated for the second and third generation of leptons and quarks.

The key to determining the different $U(1)$ charges of these fields lies in equation 2.46. The right-handed fields, which do not interact with the $SU(2)$ sector, will have $T^3 = 0$, so for these fields the $U(1)$ charge Y will simply equal the electric charge of the particle. For the left-handed fields, the $SU(2)$ factor T^3 gives values of $\pm\frac{1}{2}$ for the upper and lower component of the $SU(2)$ doublet, respectively, and in order to make the electric charge match this, a value of $Y = -\frac{1}{2}$ for the field E_L and $Y = \frac{1}{6}$ for the field Q_L is assigned. Thus, the kinetic term of the Lagrangian for the fermion fields has the form

$$\begin{aligned} \mathcal{L}_k \; = \; & \overline{E}_L(i\gamma^\mu D_\mu)E_L + \overline{e}_R(i\gamma^\mu D_\mu)e_R \\ & + \overline{Q}_L(i\gamma^\mu D_\mu)Q_L + \overline{u}_R(i\gamma^\mu D_\mu)u_R + \overline{d}_R(i\gamma^\mu D_\mu)d_R \end{aligned} \tag{2.54}$$

where the covariant derivative D_μ is given by the expression in equation 2.45 and the $U(1)$ charge Y is the one just assigned to each field. If we write this Lagrangian in terms of the mass eigenstate fields 2.43 and use the expression 2.49 for the covariant derivative, the result is

$$\begin{aligned} \mathcal{L}_k \; = \; & \overline{E}_L(i\gamma^\mu \partial_\mu)E_L + \overline{e}_R(i\gamma^\mu \partial_\mu)e_R \\ & + \overline{Q}_L(i\gamma^\mu \partial_\mu)Q_L + \overline{u}_R(i\gamma^\mu \partial_\mu)u_R + \overline{d}_R(i\gamma^\mu \partial_\mu)d_R \\ & + \tfrac{e}{\sin\theta_W}\left(W^+_\mu J^{+\mu}_W + W^-_\mu J^{-\mu}_W + Z^0_\mu J^\mu_Z\right) + eA_\mu J^\mu_{EM} \end{aligned} \tag{2.55}$$

---

[2] In the Minimal Standard Model, the neutrinos are assumed massless, and have therefore no right-handed component. This is however not in accordance with the latest neutrino oscillation experiments [11], and the right-handed neutrino fields are therefore kept in the formalism.

where

$$\begin{aligned}
J_W^{+\mu} &= \tfrac{1}{\sqrt{2}}(\overline{\nu}_L\gamma^\mu e_L + \overline{u}_L\gamma^\mu d_L) \\
J_W^{-\mu} &= \tfrac{1}{\sqrt{2}}(\overline{e}_L\gamma^\mu \nu_L + \overline{d}_L\gamma^\mu u_L) \\
J_Z^{\mu} &= \tfrac{1}{\cos\theta_W}\left[\overline{\nu}_L\gamma^\mu\left(\tfrac{1}{2}\right)\nu_L + \overline{e}_L\gamma^\mu\left(-\tfrac{1}{2}+\sin^2\theta_W\right)e_L + \overline{e}_R\gamma^\mu\left(\sin^2\theta_W\right)e_R\right. \\
&\qquad +\overline{u}_L\gamma^\mu\left(\tfrac{1}{2}-\tfrac{2}{3}\sin^2\theta_W\right)u_L + \overline{u}_R\gamma^\mu\left(-\tfrac{2}{3}\sin^2\theta_W\right)u_R \\
&\qquad \left.+\overline{d}_L\gamma^\mu\left(-\tfrac{1}{2}+\tfrac{1}{3}\sin^2\theta_W\right)d_L + \overline{d}_R\gamma^\mu\left(\tfrac{1}{3}\sin^2\theta_W\right)d_R\right] \\
J_{EM}^{\mu} &= \overline{e}\gamma^\mu(-1)\,e + \overline{u}\gamma^\mu\left(+\tfrac{2}{3}\right)u + \overline{d}\gamma^\mu\left(-\tfrac{1}{3}\right)d,
\end{aligned}$$

where the field e is a short-hand notation for $e_L + e_R$. As can be seen directly, the electromagnetic current J_{EM}^{μ} is regained with the correct factors corresponding to the electric charge of the electron, up- and down-type quarks.

Fermion mass terms

The task of writing down fermion mass terms seems at first glance to be an impossible one. Since the left- and right-handed fields necessarily reside in different $SU(2)$ representation, due to the chiral nature of the weak interactions, gauge invariance does not permit direct mass terms of the form $-m_f(\overline{f_L}f_R + \overline{f_R}f_L)$. This is because the left-handed and right-handed fields have different $U(1)$ charge Y, and the mass term therefore violates global gauge invariance.

The solution to this problem lies, once more, in the spontaneous symmetry breaking of the field $\phi(x)$; two gauge-invariant term, with $U(1)$ charges Y summing to zero, can be written for the quark fields Q_L, u_R and d_R as

$$\Delta\mathcal{L}_q = -\lambda_d\{(\overline{Q}_L\cdot\phi)d_R + \overline{d}_R(\phi^\dagger\cdot Q_L)\} - \lambda_u\epsilon^{ab}\{\overline{Q}_{La}\phi_b^\dagger u_R + \overline{u}_R\phi_b Q_{La}\} \quad (2.56)$$

which, if we insert the expansion 2.41 into this expression, takes the form

$$\begin{aligned}
\Delta\mathcal{L}_q &= -\tfrac{\lambda_d v}{\sqrt{2}}\left(\overline{d}_L d_R + \overline{d}_R d_L\right) - \tfrac{\lambda_u v}{\sqrt{2}}\left(\overline{u}_L u_R + \overline{u}_R u_L\right) \\
&\quad -\tfrac{\lambda_d}{\sqrt{2}}\left(\overline{d}_L d_R + \overline{d}_R d_L\right)h(x) - \tfrac{\lambda_u}{\sqrt{2}}\left(\overline{u}_L u_R + \overline{u}_R u_L\right)h(x)
\end{aligned} \quad (2.57)$$

The first two terms of this partial Lagrangian are mass term for the d and u quarks, with the quark masses being $m_u = \frac{1}{\sqrt{2}}\lambda_u v$ and $m_d = \frac{1}{\sqrt{2}}\lambda_d v$. Thus, the quark masses are given by two new coupling constant to the $\phi(x)$ field. If we introduce the quark masses into the expression, and the total fields as the sum of the right-handed and left-handed components, the partial Lagrangian takes the form

$$\Delta\mathcal{L}_q = -m_d\overline{d}d - m_u\overline{u}u - \frac{m_d}{v}\overline{d}dh(x) - \frac{m_u}{v}\overline{u}uh(x) \tag{2.58}$$

The last two terms in this equation are couplings between the Higgs boson field $h(x)$ and a pair of quark/anti-quark fields, and the strength of the coupling is proportional to the quark mass. Thus, the Higgs field couples stronger to fermion fields the more massive the fermion field is.

The same mechanism is used for generating masses for the leptons, *i.e.* the Lagrangian contains a term

$$\begin{aligned}\Delta\mathcal{L}_e &= -\lambda_e\{(\overline{E}_L\cdot\phi)e_R + \overline{e}_R\cdot(\phi^\dagger)E_L)\} - \lambda_\nu\epsilon^{ab}\{\overline{E}_{La}\phi_b^\dagger\nu_R + \overline{\nu}_R\phi_bE_{La}\} \\ &= -m_e\overline{e}e - m_\nu\overline{\nu}\nu - \tfrac{m_e}{v}\overline{e}eh(x) - \tfrac{m_\nu}{v}\overline{\nu}\nu h(x).^3\end{aligned} \tag{2.59}$$

When introducing more than one generation of quarks and leptons, couplings mixing different generations of quarks can be introduced, of the type

$$\Delta\mathcal{L} = -\lambda_{dij}(\overline{Q}_L^i\cdot\phi)d_R^j - \lambda_{uij}\epsilon^{ab}\overline{Q}_{La}^i\phi_b^\dagger u_R^j \tag{2.60}$$

where the indices i and j denote the different quark generations. In order to diagonalize the Higgs couplings to avoid these mixings, and thus produce the physical mass matrix of the quarks and leptons, it is always possible to choose bases $u_L^i = (u_L, c_L, t_L)$, $d_L^i = (d_L, s_L, b_L)$ (the original basis) and $u_L^{'i} = (u_L^{'}, c_L^{'}, t_L^{'})$, $d_L^{'i} = (d_L^{'}, s_L^{'}, b_L^{'})$ (the basis that diagonalizes the Higgs couplings) connected by the relations

$$u_L'^{\,i} = U_u^{ij}u_L^j, \qquad d_L'^{\,i} = U_d^{ij}d_L^j \tag{2.61}$$

where the matrices U_u and U_d are unitary, and the indices i and j denote matrix indices. Using the transformations 2.61, the expression for the quark part of the W^+W^- boson currents takes the form

$$\begin{aligned}J_W^{+\mu} &= \tfrac{1}{\sqrt{2}}\overline{u}_L^i\gamma^\mu d_L^i &= \tfrac{1}{\sqrt{2}}\overline{u}_L'^{\,i}\gamma^\mu(U_u^\dagger U_d)^{ij}d_L'^{\,j} &= \tfrac{1}{\sqrt{2}}\overline{u}_L'^{\,i}\gamma^\mu(V_{\text{CKM}})^{ij}d_L'^{\,j} \\ J_W^{-\mu} &= \tfrac{1}{\sqrt{2}}\overline{d}_L\gamma^\mu u_L &= \tfrac{1}{\sqrt{2}}\overline{d}_L'^{\,i}\gamma^\mu(U_d^\dagger U_u)^{ij}u_L'^{\,j} &= \tfrac{1}{\sqrt{2}}\overline{d}_L'^{\,i}\gamma^\mu(V_{\text{CKM}}^\dagger)^{ij}u_L'^{\,j}\end{aligned} \tag{2.62}$$

---

[3] In the Minimal Standard Model, the neutrino mass is assumed to be zero, making the second and last terms in this equation disappear. However, the latest neutrino oscillation searches favour, as already mentioned, a non-zero mass difference between neutrinos of different generations. [11]

The unitary matrix V_{CKM} is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix [12], and induces mixing between different quark generations in the weak charged currents[4].

An alternative, but completely equivalent, viewpoint to this approach, is to retain the original basis u_L and d_L, *i.e.* the original weak eigenstate bases. In this case, there is no mixing introduced between the states, as this is incorporated in the general quark mass matrix of the weak eigenstates. Thus, these two alternatives can be illustrated as follows. (The mass matrix shown is the one in the up-type quark sector, corresponding to the second term of equation 2.58, whereas the one in the down quark sector is analogous and therefore omitted. The same applies to the CKM matrix in the leptonic sector.)

$$\begin{array}{cc}
\text{Weak eigenstate basis} & \text{Physical eigenstate basis} \\
\underset{\text{mass}}{\begin{pmatrix} M_{uu} & M_{uc} & M_{ut} \\ M_{cu} & M_{cc} & M_{ct} \\ M_{tu} & M_{tc} & M_{tt} \end{pmatrix}} \quad \underset{\text{mixing}}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} & \underset{\text{mass}}{\begin{pmatrix} M_{uu} & 0 & 0 \\ 0 & M_{cc} & 0 \\ 0 & 0 & M_{tt} \end{pmatrix}} \quad \underset{\text{mixing}}{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}
\end{array}$$

In the first case, all of the parameters are contained in the mass sector as couplings to the Higgs field, and the mixing sector is the trivial unity operator. This approach can in many respects be regarded as the most fundamental one, as it encompasses the weak eigenstates alone, which are the ones connected with the fundamental symmetries of the theory. However, the physically observable particles are the ones found when diagonalizing the mass matrix; also, the parameters of the CKM matrix are related to and measurable from weak hadronic interactions. Therefore, the last case is the most useful one from a practical and experimental point of view, and is therefore the most widely used in the physics community. However, when presenting the aspects and importance of the Higgs mechanism, the first approach is worth notice, as the mixing sector is reduced to triviality and the origin of both the physical particle masses and the CKM mixing as parameters in the Higgs sector is fully revealed.

---

[4] In the Minima Standard Model, this method of generating mixing between the mass eigenstates and the weak eigenstates will not work in the leptonic sector, due to the non-existence of the right-handed massless neutrino fiels. If, on the other hand, the neutrinos have mass, the right-handed fields will exist, and the generation of a matrix analogous to V_{CKM} is straightforward.

2.4.3 Radiative corrections and theoretical Higgs mass limits

So far all the considerations regarding the Higgs mass have been performed on the level of the Lagrangian, *i.e.* at the tree-level. At this level, the mass of the Higgs boson is a completely undetermined parameter, which can only be determined by direct measurement on the Higgs boson itself. When including radiative corrections, bounds on the Higgs mass can be achieved. Also, fundamental theoretical considerations and arguments can be used to set limits on the mass of the Higgs.

Vacuum stability: the Linde-Weinberg limit

The Linde-Weinberg lower bound on the Higgs mass [13] comes from the requirement of a stable electro-weak symmetry breaking vacuum. When considering the renormalization group equation for the Higgs self-coupling λ of equation 2.40, together with the one-loop effective potential, one finds that, in order to have a symmetry-breaking vacuum (*i.e.* $V_{\text{eff}}(\frac{v}{\sqrt{2}}) < 0$) the Higgs mass has a lower limit given by

$$m_h^2 > \frac{3}{16\pi^2 v^2}\left[2m_W^4 + m_Z^4 - 4m_t^4\right] \tag{2.63}$$

where m_t is the top mass, and the lighter fermion masses approximated by zero. Unfortunately, with the top mass being as high as 175 GeV/c^2 [14], the right-hand side of equation 2.63 becomes negative, rendering this approach to a lower limit on the Higgs mass useless.

The Coleman-Weinberg potential

The Coleman-Weinberg mechanism [15] is based on the approach that the symmetry breaking in the Higgs sector is due solely to radiative corrections, *i.e.* that the parameter μ^2 in equation 2.40 is set to zero. Although there is no compelling reason from any symmetry to set this parameter to zero, the result is a Higgs boson mass term induced by the radiative corrections (due to an induced minimum in the potential) given, at the one-loop correction level, by

$$m_{\text{CW}}^2 = 8v^2 B, \qquad B = \frac{1}{64\pi^2 v^4}\sum_i C_i(2J_i+1)(-1)^{2J_i} m_i^4, \tag{2.64}$$

where the sum is taken over all vector bosons and fermions, of mass m_i, spin J_i and counting factor C_i (which counts electric and colour charge of the particles; $C_i = 1, 2, 2, 6$ for the Z^0, W^+W^-, charged lepton and quark, respectively). This number provides a lower limit on the Higgs mass. Unfortunately, for top quark masses above ~78 GeV/c^2 (which is a value well below the observed value [14]), the value for m^2_{CW} becomes negative. However, this approach also gives a Higgs mass limit in the two-doublet case, where the value indeed gives a non-zero lower limit.

Triviality

When considering the one-loop renormalization group equation for the Higgs self-coupling parameter λ of equation 2.40, the leading term in λ is given by [16, page 65]

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \beta(t) = \frac{12}{16\pi^2}\lambda^2(t) + \mathcal{O}(\lambda) \tag{2.65}$$

where $t = \ln(Q^2/Q_0^2)$, Q_0 and Q being the experimental and the cutoff energy scale, respectively. This equation has the solution

$$\frac{1}{\lambda(v)} - \frac{1}{\lambda(Q)} = \frac{3}{4\pi^2}\ln\left(\frac{Q^2}{v^2}\right) \quad \Rightarrow \quad \lambda(v) = \frac{\lambda(Q)}{1 + \frac{3\lambda(Q)}{2\pi^2}\ln\left(\frac{Q}{v}\right)} \tag{2.66}$$

when evaluating from the energy scale Q down to the scale $Q_0 = v$. The term "triviality" is explained by noting that the coupling $\lambda(v)$ vanishes as the cutoff Q is taken to infinity. This signifies than the theory, with an elementary Higgs-Lagrangian, is meaningful only up to a certain energy scale, where new physics must be assumed to enter. By rewriting the solution:

$$\lambda(Q) = \frac{\lambda(v)}{1 - \frac{3\lambda(v)}{2\pi^2}\ln\left(\frac{Q}{v}\right)} \tag{2.67}$$

one can see that as Q increases, so does the self-coupling, and will eventualy blow up at a certain large value of Q, called the Landau pole. This is under the assumption that the β function of equation 2.65 remains an adequate description of the evaluation of λ, *i.e.* that the theory remains in the perturbative regime. By letting $\lambda(Q)$ go to infinity (which is clearly outside the perturbative domain, and therefore requires new physics or a strongly interacting Higgs sector), one finds an upper limit on the Higgs mass which, from equation 2.67, is given by

$$\frac{1}{\lambda(v)} \geq \frac{3}{2\pi^2}\ln\left(\frac{Q}{v}\right) \quad \Rightarrow \quad \lambda(v) \leq \frac{2\pi^2}{3\ln\left(\frac{Q}{v}\right)} \quad \Rightarrow \quad m^2_{H^0} \leq \frac{4v^2\pi^2}{3\ln\left(\frac{Q}{v}\right)}. \tag{2.68}$$

The choice of $\frac{Q}{v} = 10$ (which is not crucial, since this factor only contributes logarithmically) gives an upper Higgs boson mass limit of approximately 600 GeV/c^2.

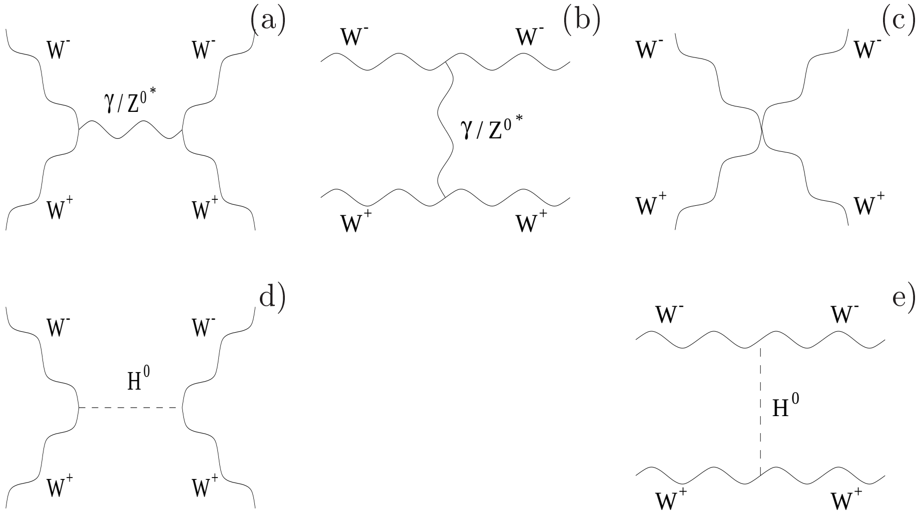


Figure 2.2: Tree-level Feynman diagrams contributing to the W^+W^- scattering process $W^+W^- \to W^+W^-$. In figures a, b and c are shown the SM processes without the Higgs boson, whereas figures d and e show the processes where the Higgs boson contributes.

Unitarity constraints: longitudinally polarized $\boldsymbol{W^+W^-}$ scattering

If one removes the Higgs boson from the standard model, the resulting theory is no longer renormalizable. This is seen when computing the longitudinal W^+W^- scattering amplitude (given at tree-level by the Feynman diagrams in parts a, b and c of figure 2.2), and the amplitude grows linearly with the centre-of-mass energy s. However, removing the Higgs boson from the model can effectively be equally well achieved by letting the Higgs boson mass go to infinity, thus making its influence on the theory negligible. In a total tree-level computation of the longitudinal W^+W^- scattering amplitude, *i.e.* including all the Feynman diagrams, of figure 2.2, the result in the limit $s, m_{H^0}^2 \gg m_W^2,\ m_Z^2$ is [16, page 82]

$$\mathcal{A}(W_L^+W_L^- \to W_L^+W_L^-) = \sqrt{2}G_F\left[\frac{s}{1-\frac{s}{m_{H^0}}} + \frac{t}{1-\frac{t}{m_{H^0}}}\right]. \tag{2.69}$$

Equation 2.69 shows that if one lets $m_{H^0} \gg s$, the scattering amplitude increases linearly with s, breaking unitarity at a specific point. A detailed analysis of this problem [17] shows that the $J = 0$ partial-wave gives the strictest bound on the Higgs boson mass:

$$m_{H^0}^2 \leq \frac{4\pi\sqrt{2}}{3G_F} \approx \left(700\,\mathrm{GeV/c^2}\right)^2. \tag{2.70}$$

Since this calculation is done at tree-level, one might ask what effect radiative corrections have on the Higgs boson mass bound of equation 2.70. Attempts at including one-loop corrections to this number have been made [18], but they do not help in restoring unitarity. The limit found in equation 2.70 is however not strictly a limit on the Higgs boson mass, but must be interpreted as the point at which normal perturbation theory breaks down. This may lead to a non-perturbative regime, for instance in a strongly interacting W^+W^- sector, or to the appearance of new physics.

2.4.4 Problems with the SM, and further outlook

Despite the success of the SM in describing the present situation in particle physics, theorists generally believe that the SM is not the final theory of particle interactions, but rather the low-energy effective theory of some more fundamental theory. This argument is based on several features of the SM which are somewhat undesirable, and which one would not expect a final physical theory to have.

First, the SM contains a rather large number of free parameters: all fermion masses of all three generations (12 parameters), the coupling constants g_s (the coupling constant for the QCD sector, also known as the strong coupling constant), g and g' (3 parameters), the three angles and one phase of the CKM matrix, plus the equivalent matrix in the leptonic sector (if one assumes neutrino masses different from zero, which has already been indicated when including them in the parameters for the fermion masses) (8 parameters) and the Higgs boson mass (1 parameter), in total 24 free parameters. In addition to this, the gauge structure of $SU(3)_C \times SU(2)_L \times U(1)_Y$ is a somewhat arbitrary choice, and is not motivated by any fundamental principle. This large degree of arbitrariness is generally considered very undesirable for a fundamental theory, and is an important reason why the SM is not generally thought of as fundamental.

There are two other arguments against the fundamentalness of the SM which are closely connected to the Higgs sector and the electro-weak spontaneous symmetry breaking. These are known as the hierarchy problem and the fine-tuning problem, and are briefly described below.

The hierarchy problem

When searching for a more fundamental theory of particle physics, of which the SM is only the low-energy effective theory, two fundamental scales are usually considered:

The Planck scale, M_P: This is the scale at which gravitational effects become important at the elementary particle level, and is therefore assumed to be the energy scale at which the gravitational force is unified with the three other forces of nature (QCD and the electro-weak GWS theory). This energy scale is $\sim 10^{19}$ GeV.

The GUT scale, M_{GUT}: This is the scale at which the electro-weak force (GWS) and the strong force (QCD) is assumed to be united. The theoretical motivation for such an assumption comes from studying the renormalization group equations for the three coupling constants of these theories. This analysis shows that whereas the strong coupling constant g_s, which is by far the larger one at the currently accessible energy scales, decreases with increasing energy scale, the other two coupling constants, g and g', both increase with increasing energy scale. Thus, one might hope that the coupling constants evolve to a common value, at which the three forces unite. This is the GUT scale, which, in most scenarios, have values of about $10^{15} - 10^{16}$ GeV.

In addition to these two energy scales, there is a third energy scale present in particle physics, which is the scale at which the electro-weak force breaks down to the two forces acting in low-energy physics: the electromagnetic and the weak force. This scale, however, is described by the symmetry-breaking vacuum expectation value given in equation 2.40, and is of the order of 10^2 GeV, which is extremely small compared to the other two scales M_P and M_{GUT}. Such a large gap in energy between fundamental scales of the theory seems unnatural for a fundamental theory, and this problem is therefore often labeled the hierarchy problem, as there seems to be a clear hierarchy in the fundamental energy scales of the theory.

The fine-tuning problem

When studying radiative corrections to the Higgs boson mass, one discovers that there are large problems connected to these. The running Higgs mass at a given scale μ_2, when evolving down from a higher scale μ_1, can generally be expressed as [19, page 81]

$$m_{H^0}^2(\mu_2) = m_{H^0}^2(\mu_1) + Cg^2 \int_{\mu_2^2}^{\mu_1^2} \mathrm{d}k^2 + Rg^2 + \mathcal{O}(g^4) \tag{2.71}$$

where g is a coupling constant and C and R dimensionless. R grows almost logarithmically with μ_1, and C diverges quadratically as $\mu_1 \to \infty$.

If now μ_1 is taken to be the fundamental scale of either the Planck or the GUT scale, both being much higher than the upper limits already obtained for the Higgs boson mass, a problem arises in keeping the Higgs boson mass small compared to the scale μ_1. The second term of equation 2.71 is of the order μ_1^2, which necessitates an extreme cancellation with the first term in order to keep the Higgs mass far below the scale μ_1. It would appear that the "natural" value of the Higgs boson mass should be of the order of μ_1, and this fine-tuning problem has therefore also been referred to as the naturalness problem.

With these shortcomings of the SM as a fundamental theory of nature, the search for theories of a more fundamental nature has been pursued extensively by theorists. Several attempts have been made in different directions, and one of the most extensively explored and by many considered promising, is the supersymmetric theories.

2.5 Beyond the SM: Supersymmetry

2.5.1 Motivations for supersymmetry

The principle of symmetry has historically been an important guide in constructing models for physics applications. The symmetry of the gauge groups is at the heart of the gauge theories described in section 2.1. However, as was shown by Coleman and Mandula [20, 21], there are strong restrictions on what type of symmetry groups can be included in a non-trivial theory. Indeed, at first glance the standard gauge theories studied so far appear to contain the maximum allowed symmetries, as the Coleman-Mandula theorem states that [20, page 28] "Under certain conditions, any physically interesting theory can contain only the Poincaré symmetries, internal symmetries and discrete symmetries". Thus, the opportunities for extending the symmetries present in the SM seem very restricted.

However, there is one more symmetry to consider. The Coleman-Mandula theorem is based on bosonic group generators obeying commutator relations, which thereby make up a Lie algebra. In addition to the bosonic generators, one can introduce fermionic generators obeying anticommutator relations, which is a situation not covered by the original assumptions made in the Coleman-Mandula theorem. By introducing these fermionic generators, and commutator relations between them and the bosonic generators, one can construct a so-called graded Lie algebra. This construction can now be used to define a theory which “sidesteps” the Coleman-Mandula theorem, and uses this graded Lie algebra instead of the Poincaré algebra, thus defining a non-trivial theory. Such theories are called supersymmetric theories, and supersymmetry can thus be considered the final symmetry possible under the restrictions set by the Coleman-Mandula theorem.

One consequence of supersymmetry is that the number of fermionic degrees of freedom equals the number of bosonic degrees of freedom. In other words, each fermion field will have a supersymmetric bosonic partner, and vice versa. Thus, supersymmetry can be seen as a symmetry between the fermionic fields, corresponding to the particles making up matter, and the bosonic fields, making up the force carriers. In this way, supersymmetric theories introduce a new symmetry between forces and ordinary matter, which can be seen as a step towards unification of the different components of the physical world, and a more fundamental theory.

Supersymmetry requires the bosonic and fermionic partner field to have the same masses and couplings. This is obviously not realized in the physical world in which we live, as there is yet no experimental evidence for supersymmetric partners of the SM fields, and the supersymmetry must therefore be broken at some scale above the energy range that has been experimentally searched up until now. A popular theoretical assumption is that the supersymmetry is broken in such a way that all the supersymmetric partner fields acquire masses substantially above their SM counterparts, and therefore hitherto have been too massive to be experimentally detected.

The supersymmetry between bosonic and fermionic degrees of freedom has another desirable consequence in the Higgs sector. It turns out that having bosonic and fermionic partner fields of the same mass and coupling (as is required by supersymmetry) removes all quadratic divergencies of the radiative corrections to the Higgs mass, *i.e.* the factor C of equation 2.71 is guaranteed to be zero. In this way, supersymmetry provides a natural solution to the fine-tuning problem, as long as the supersymmetric scale is manifest not too far above the electro-weak symmetry breaking scale. However, any solution to the fine-tuning problem must fulfill this requirement, and supersymmetry is therefore a perfectly adequate solution to this problem.

The Minimal Supersymmetric Standard Model is the simplest possible consistent realization of a supersymmetric theory in which the SM can be embedded, and has therefore been subject to much theoretical interest. A supersymmetric bosonic partner field is assigned to each SM fermion field (and denoted by the original fermion field with a preceding s-), and every SM boson field is assigned a sypersymmetric fermion field (denoted by the original boson field with a trailing -ino). It is also customary to write all supersymmetric partners with a tilde above, in order to distinguish them

SM Fermions		**SM Bosons**			
SM field	SUSY field	SM field	SUSY field		particle
Leptons	*Sleptons*	$H^{\pm}$	Higgsino	⇒	chargino
electron (e)	selectron ($\tilde{e}$)	$W^{\pm}$	Wino		$\tilde{\chi}_1^{\pm}$–$\tilde{\chi}_2^{\pm}$
muon (μ)	smuon ($\tilde{\mu}$)	H^0	Higgsino		
tau (τ)	stau ($\tilde{\tau}$)	h^0	Higgsino	⇒	neutralino
neutrino (ν)	sneutrino ($\tilde{\nu}$)	Z^0	Zino		$\tilde{\chi}_1^0$–$\tilde{\chi}_4^0$
Quarks	*Squarks*	photon (γ)	photino		
bottom (b)	sbottom ($\tilde{b}$)	gluon (g)			gluino ($\tilde{g}$)
top (t)	stop ($\tilde{t}$)	graviton (G)			gravitino ($\tilde{G}$)

Table 2.1: The different SM particle fields and their supersymmetric partners. The necessety of two Higgs doublets, and thereby the enlarged Higgs sector, will be treated in the following sections. The charged supersymmetric partners of the $W^{\pm}$ and $H^{\pm}$ mix to form a total of four charginos, whereas the neutral supersymmetric partners of the Z^0, h^0, H^0, and A^0 mix to form a total of four neutralinos. Also, only the sbottom and stop squarks are listed, as the other squarks are rarely referred to explicitly by their supersymmetric names.

from ordinary particles. The structure and nomenclature of the new supersymmetric particles is given in table 2.1.

In the Standard Model, only a single Higgs doublet is needed in order to assign masses to the gauge bosons and fermions in the theory. This is no longer possible within the supersymmetric framework, because a mechanism like that of the last term of equation 2.56 in order to generate masses for the up-type quarks would violate gauge-symmetry. The Higgs sector therefore contains two complex $SU(2)$ doublets in the spinor representation.[5] This makes it necessary to study a general two-doublet model.

2.5.2 The two-doublet Higgs model

Any extension of the simple one-doublet Higgs structure must retain several important results from the SM, both experimentally and theoretically. Such results imply important restrictions on the structure of the new model, the

---

[5] There is no fundamental reason for only having a single Higgs doublet in the SM, but it is sufficient. The theoretical development of a two-doublet Higgs sector in the SM follows from the general analysis of a two-doublet model, treated in the next section, and is therefore omitted.

most important of which are:

The ρ-parameter: The parameter ρ is defined as the ratio $\rho = \frac{m_W}{m_Z \cos\theta_W}$. In the SM, this parameter is by definition, as a result of the Higgs structure, equal to 1. Measurements of this parameter give a world average value of $\rho = 0.995 \pm 0.013$ [22], in perfect agreement with the SM prediction. For a Higgs structure consisting of only singlets and doublets, the value of $\rho = 1$ follows automatically [23]. For more complicated Higgs structures, the value of ρ is dependent on parameter values (such as the isospin T and hypercharge Y, together with the vacuum expectation value) of each Higgs representation. This imposes constraints on the parameters of the Higgs representations (possibly by a custodial $SU(2)$ symmetry in the Higgs sector) in order to avoid a fine-tuning between them to insure a value of ρ close to 1. Such models are therefore generally considered more "unnatural" than models consisting of only Higgs singlets and doublets.

Flavor-changing neutral currents (FCNCs): The observed absence of flavour-changing neutral currents [24], which is guaranteed at tree-level by the SM Higgs structure because of the mass diagonalizing operation which simultaneously diagonalize the Higgs-fermion couplings, strongly constrains the possible Higgs structures. As more general Higgs structures than the simple one-doublet model in the SM do not prohibit FCNCs, a method is needed to control them. One way of doing this, is to require relatively large Higgs masses (of the order of 1 TeV), which will supress such FCNCs. The other method, which is often favoured to the somewhat artificially large values of Higgs masses, is due to a theorem by Glashow and Weinberg [25], which states that if all fermions of a given electric charge couple to no more than one Higgs doublet, FCNCs mediated by Higgs bosons is absent at tree-level. Such a requirement constrains the Higgs-fermion couplings, but not uniquely. In realised models with more than one Higgs doublet, this theorem is usually utilized in one of two ways:

1. One doublet couples to up-type fermions, and the other doublet couples to down-type fermions.
2. One doublet couples to all fermions, both up-type and down-type, while the other doublet does not couple to fermions at all.

As will be shown below, supersymmetry requires the first of these two schemes.

Unitarity: As already discussed in section 2.4.3, a scalar field is required in order to keep the longitudinally polarized W^+W^- scattering amplitude from violating unitarity; indeed, this requirement was used to set an upper limit on the Higgs mass. Essential in this mechanism, is the fact that the Higgs coupling to vector bosons is given by the relation

$$g_{HVV} = g\, m_V \tag{2.72}$$

where V denotes either Z^0 or $W^\pm$, and g is the gauge coupling. In more complex models (admittedly restricted to models containig only Higgs singlets and doublets), the relation 2.72, which ensures the correct cancellation in the W^+W^- scattering amplitude, is replaced by

$$\sum_i (g_{H_iVV})^2 = (g_{HVV})^2, \quad \sum_i (g_{H_iVV})(g_{H_iff}) = (g_{HVV})(g_{Hff}) \tag{2.73}$$

where the sum is carried out over the different Higgs singlets and doublets, g_{H_iVV} and g_{H_iff} are the Higgs couplings of singlet or doublet i to vector bosons and fermions, respectively, and the right-hand side of the equations corresponds to the SM values of the couplings. As can be seen from equation 2.73, an extended Higgs sector containing more than a single Higgs doublet implies that the Higgs couplings to both vector bosons and fermions are reduced with respect to the SM couplings.

These arguments show that, although there is still room for more complex Higgs structures, the models containing only Higgs doublets and singlets have several theoretical advantages. Furthermore, the simplest extension of the one-doublet model (in the sense of adding the smallest number of new free parameters) is the two-doublet model, something that makes the two-doublet model particularly interesting.

The symmetry-breaking potential $V(\phi_1, \phi_2)$ of two complex Higgs doublets ϕ_1 and ϕ_2 is given in its most general form as [26]

$$\begin{aligned}
& \phi_1 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix}, \qquad \phi_2 = \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix} \\
V(\phi_1, \phi_2) \ = \ & \lambda_1 \left(\phi_1^\dagger \phi_1 - v_1^2\right)^2 + \lambda_2 \left(\phi_2^\dagger \phi_2 - v_2^2\right)^2 \\
& + \lambda_3 \left[\left(\phi_1^\dagger \phi_1 - v_1^2\right) + \left(\phi_2^\dagger \phi_2 - v_2^2\right)\right]^2 \\
& + \lambda_4 \left[\left(\phi_1^\dagger \phi_1\right)\left(\phi_2^\dagger \phi_2\right) - \left(\phi_1^\dagger \phi_2\right)\left(\phi_2^\dagger \phi_1\right)\right]^2 \\
& + \lambda_5 \left[\Re\left(\phi_1^\dagger \phi_2\right) - v_1 v_2 \cos\xi\right]^2 \\
& + \lambda_6 \left[\Im\left(\phi_1^\dagger \phi_2\right) - v_1 v_2 \sin\xi\right]^2 .
\end{aligned} \tag{2.74}$$

This potential guarantees the correct symmetry breaking of $SU(2) \times U(1)$, and is the most general one which respects gauge symmetry and the symmetry $\phi_1 \to -\phi_1$, which is necessary in order to ensure that there are no large FCNCs. The vacuum expectation values of the Higgs fields are

$$\langle \phi_1 \rangle = \begin{pmatrix} 0 \\ v_1 \end{pmatrix}, \qquad \langle \phi_2 \rangle = \begin{pmatrix} 0 \\ v_2 e^{i\xi} \end{pmatrix} \tag{2.75}$$

which is analogous to equation 2.40 in the SM case. The parameter ξ introduces CP violation in the Higgs sector (for a summary of such models, see [27]); although this is allowed in the general two-doublet model case, supersymmetry explicitly requires the two parameters λ_5 and λ_6 of equation 2.74 to be equal, making it possible to rotate the parameter ξ out of the Higgs sector, and thereby making sure there is no CP violation. Therefore, the parameter ξ is set to zero for simplicity in the following analysis.

The resulting gauge boson masses are given by

$$m_W^2 = \frac{g^2 \left(v_1^2 + v_2^2\right)}{2} \quad m_Z^2 = \frac{\left(g^2 + g'^2\right)\left(v_1^2 + v_2^2\right)}{2} \quad m_\gamma^2 = 0. \tag{2.76}$$

When comparing this result to the one found in 2.43, the only difference, apart from a conventional factor of $\sqrt{2}$ in the vacuum expactation values 2.40 and 2.75, is the replacement of v^2 by $v_1^2 + v_2^2$.

One key parameter can be introduced already at this level; the ratio of the vacuum expectation values of the two Higgs doublets,

$$\tan\beta = \frac{v_2}{v_1}. \tag{2.77}$$

The two complex Higgs doublets contain a total of eight parameters. Three of these are eaten up by the Goldstone bosons, which leaves five physical Higgs bosons. These are two charged Higgs bosons, H^+ and H^-, and three neutral Higgs bosons, the CP-odd A^0, and the two CP-even H^0 and h^0. These are given by:

$$\begin{aligned} H^{\pm} &= -\phi_1^{\pm}\sin\beta + \phi_2^{\pm}\cos\beta & \text{mass} \quad m_{H^{\pm}}^2 &= \lambda_4\left(v_1^2+v_2^2\right) \\ A^0 &= \sqrt{2}\left(-\Im\left(\phi_1^0\right)\sin\beta + \Im\left(\phi_2^0\right)\cos\beta\right) & \text{mass} \quad m_{A^0}^2 &= \lambda_6\left(v_1^2+v_2^2\right). \end{aligned} \tag{2.78}$$

The two CP-even Higgs bosons mix through their mass-squared matrix

$$\mathcal{M} = \begin{pmatrix} 4v_1^2(\lambda_1+\lambda_3)+v_2^2\lambda_5 & (4\lambda_3+\lambda_5)v_1v_2 \\ (4\lambda_3+\lambda_5)v_1v_2 & 4v_2^2(\lambda_2+\lambda_3)+v_1^2\lambda_5 \end{pmatrix}, \tag{2.79}$$

with the resulting Higgs boson fields and masses being

$$\begin{aligned} H^0 &= \sqrt{2}\left[\left(\Re(\phi_1^0)-v_1\right)\cos\alpha + \left(\Re(\phi_2^0)-v_2\right)\sin\alpha\right] \\ h^0 &= \sqrt{2}\left[\left(-\Re(\phi_1^0)-v_1\right)\sin\alpha + \left(\Re(\phi_2^0)-v_2\right)\cos\alpha\right] \\ m_{H,h} &= \tfrac{1}{2}\left[\mathcal{M}_{11}+\mathcal{M}_{22} \pm \sqrt{\left(\mathcal{M}_{11}-\mathcal{M}_{22}\right)^2+4\mathcal{M}_{12}^2}\right]. \end{aligned} \tag{2.80}$$

A second parameter of the model, the Higgs mixing angle α, has been introduced in equation 2.80, which is given by the following expression:

$$\sin 2\alpha = \frac{2\mathcal{M}_{12}}{\sqrt{\left(\mathcal{M}_{11}-\mathcal{M}_{22}\right)^2+4\mathcal{M}_{12}^2}}. \tag{2.81}$$

Thus, the Higgs sector of the two-doublet model contains six free parameters: four Higgs boson masses, one mixing angle α and the ratio of vacuum expectations $\tan\beta$, as opposed to the one free parameter, the Higgs boson mass, in the one-doublet model.

In order to discuss the couplings of the different Higgs bosons to the vector bosons, it is necessary to study their C, P and J quantum numbers. The assignment of such quantum numbers is justified by noting that the SM, in the absence of quarks and leptons (*i.e.* with only the bosons) separately conserves C and P [16, page 197]. Thus, these are good quantum numbers for the fundamental bosons of the theory, given the aforementioned conditions of a fermion-free theroy, and their values are given in table 2.2 [28].

The coupling of Z^0 to a pair of identical Higgs bosons (H^0H^0, h^0h^0 or A^0A^0) is forbidden by Bose symmetry, since the Z^0 wave function is antisymmetric, whereas the two identical Higgs bosons have a symmetric wave

	Higgs bosons				Vector bosons		
	h^0	H^0	A^0	$H^\pm$	γ	Z^0	$W^\pm$
J^{PC} (J^P)	0^{++}	0^{++}	0^{+-}	0^+	1^{--}	1^{--}	1^-

Table 2.2: The C, P and J quantum numbers of the different Higgs bosons of the two-doublet model, as well as the vector bosons, when disregarding the fermions of the theory, *i.e.* when C and P are both conserved. For the charged $H^\pm$ and $W^\pm$, only the P and J quantum numbers are given, as the C quantum number is only well-defined for neutral particles.

function, as required by Bose statistics. This leaves the coupling of Z^0 to a pair of non-identical Higgs bosons to be considered, something which is only allowed if the two bosons have opposite CP quantum numbers (*i.e.* h^0A^0 and H^0A^0). Furthermore, the coupling of two vector bosons and one Higgs boson (*i.e.* the Higgs-strahlung process) requires the Higgs boson to be CP-even, thus eliminating the couplings of $Z^0Z^0A^0$ and $W^+W^-A^0$. In addition to this, coupling involving only neutral Higgs bosons and one or two photons or gluons vanish, due to the massless nature of these gauge bosons. Also, the couplings $H^+W^-\gamma$ and $H^+W^-Z^0$ are prohibited, the first from conservation of the electromagnetic current, and the last as a consequence of a Higgs structure consisting of only doublets and singlets [29]. In summary, the only three allowed types of couplings between Higgs bosons and vector bosons are:

- Higgs-strahlung: $Z^0 \to h^0Z^0$ or H^0Z^0, (known as the Bjorken process) and $W^\pm \to h^0W^\pm$ or $H^0W^\pm$.
- Neutral pair production: $Z^0 \to h^0A^0$ or H^0A^0.
- Charged pair production: $Z^0 \to H^+H^-$.

2.5.3 The Minimal Supersymmetric Standard Model (MSSM)

As already mentioned, supersymmetry requires an additional Higgs doublet in order to assign masses to the up-type quarks in a gauge-invariant way. This can be seen from the superpotential describing interactions between Higgs bosons and fermions,

$$W_F = \epsilon_{ij}\left[\lambda_l \hat{H}_1^i \hat{L}^j \hat{R} + \lambda_d \hat{H}_1^i \hat{Q}^j \hat{D} + \lambda_u \hat{H}_2^j \hat{Q}^i \hat{U}\right] \tag{2.82}$$

where $\hat{H}_1$ and $\hat{H}_2$ are the Higgs superfields, $\hat{Q}$ and $\hat{L}$ are the $SU(2)$ weak-doublet quark and lepton superfields, $\hat{U}$ and $\hat{D}$ are the $SU(2)$ singlet up- and down-quark superfields, respectively, and $\hat{R}$ is an $SU(2)$ weak-singlet charged lepton superfield. Since supersymmetry forbids the appearance of $\hat{H}_1^*$, which is used in the second term of equation 2.56, and a term like $\hat{H}_1\hat{Q}\hat{U}$ would violate gauge symmetry, the last term of equation 2.82 is the only gauge- and supersymmetry-invariant way of assigning mass to the up-type quarks.

Another way of seeing the necessity for a second Higgs doublet, is to require the disappearance of the anomalies in the theory. This corresponds to requiring all fermionic charges to sum to zero. Since the fermion charges in the quark and lepton sector sum to zero separately, this puts the same requirement on the Higgs sector itself. Thus, the fermionic partners of one Higgs doublet, $(\tilde{H}_1^0, \tilde{H}_1^-)$, must be complemented with those of a second Higgs doublet, $(\tilde{H}_2^+, \tilde{H}_2^0)$.

Since supersymmetry is broken at all energy scales currently accessible to experimental study, one must introduce supersymmetry breaking terms in the Lagrangian. The usual way of accomplishing this, is by breaking supersymmetry dynamically (see, for instance, [30] for a review) at a high energy scale (constrained between about 10^4 GeV and M_P [31]), and evolving the supersymmetry breaking parameters down to the weak scale through their renormalization group equations. These supersymmetry breaking terms must, however, not spoil the cancellation of the quadratically divergent radiative corrections to the Higgs boson mass mentioned in the previous paragraph. Terms fulfilling these requirements are known as soft supersymmetry breaking (SSB) terms, and enable the total Lagrangian to be written in the form [2, page 11]

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{soft}} \tag{2.83}$$

where $\mathcal{L}_{\text{SUSY}}$ contains the supersymmetrically invariant terms, and $\mathcal{L}_{\text{soft}}$ violates supersymmetry but contains only mass terms and couplings with *positive* mass dimention. Such terms are limited to gaugino masses, sfermion masses, and bilinear and trilinear scalar couplings. These terms combine, if they are allowed to be treated as free parameters, into a rather overwhelming degree of arbitrariness, as there in the Minimal Supersymmetric Standard Model are no less than 105 masses, phases and mixing angles which cannot be rotated away [32]. Many of these parameters introduce physical effects, such as flavor mixing and CP violation, at a level excluded by experiments, and some guiding organizing principle for these SSB terms must therefore be applied, in order to make a phenomenologically viable model. This is often done by making assumptions about the origin of the SSB terms, something which in most models greatly decreases the number of free parameters. In other words, the dynamical supersymmetry breaking at high energy scale is parametrized by some specific scheme or scenario, usually in such a way that this supersymmetry breaking is communicated through a "hidden" sector of particles with no or very small couplings to the "visible" sector of the SM particles and their superpartners. This field has seen major developments over the last few years, and scenarios for the dynamical high-energy breaking of supersymmetry now includes the following [31] (the term "supersymmetry breaking" is usually atttached to all scenarios listed):

Gavity-mediated (SUGRA): In this scenario, the hidden and visible sectors communicate through gravity-based interactions, with the energy scale of the supersymmetry breaking interactions of the hidden sector at an energy of $\gtrsim 10^{10}$ GeV.

Gauge-mediated (GMSB): Here, the supersymmetry breaking is accomplished by interactions of the ordinary electroweak and QCD gauge type, mediated by messenger particles which couple to a field with supersymmetry breaking vacuum expectation value. In this scenario, the scale of supersymmetry breaking can be as low as $\sim 10^4$ GeV.

Anomaly-mediated (AMSB): This scenario has no supergravity couplings, and the supersymmetry breaking is induced by loop effects. These contributions also exist in the SUGRA and GMSB scenarios, but are there negligible.

Gaugino-dominated: This scenario is based on the *brane world scenario*, where our world, with SM particles and supersymmetric partners, exists on a brane separated from the one which is responsible for the supersymmetry breaking.

The MSSM employs the scheme of gravity-mediated supersymmetry breaking, and in addition assumes a "minimal" form for the normalization of kinetic terms and gauge interactions in the full, nonrenormalizable supergravity lagrangian [2]. The remaining free parameters are then:

The universal gaugino mass term $m_{1/2}$: The parameters giving mass terms for the $U(1)$, $SU(2)$, and $SU(3)$ gauginos, M_1, M_2, and M_3, respectively, gets a common value $m_{1/2}$.

The universal sfermion mass term M_S: The common mass term for all scalar supersymmetric particles above the supersymmetry breaking scale (this parameter is also known as m_0).

The universal squark trilinear coupling A: The interaction parameter between three squarks; usually, only the third, *i.e.* heaviest, generation is considered.

In addition to these, parameters describing the Higgs sector itself, as outlined in section 2.5.2, are needed.

Supersymmetry places several restrictions on the two-doublet structure which are not present in the general case studied in section 2.5.2. In order to keep the cancellation of the quadratic divergencies of the Higgs boson mass, the dimension-four terms of the Higgs potential must respect the supersymmetry. This relates the gauge boson couplings to the Higgs couplings, and gives definite relations between Higgs and gauge boson masses. Furthermore, the MSSM, which is the *minimal* supersymmetric extension of the SM, poses further constraints on the theory. In a general supersymmetric model, a gauge singlet scalar field denoted by N is included and must be taken into account when studying the Higgs sector, but this is not present in the MSSM. Therefore, the most general superpotential that respects baryon and lepton number conservation is

$$W = \mu \epsilon_{ij} \hat{H}_1^i \hat{H}_2^j + W_F \tag{2.84}$$

where W_F is given by equation 2.82.

The scalar field potential V of the MSSM is now, under the aforementioned conditions, given as

$$\begin{aligned} V \ = \ & (m_1^2 + |\mu|^2 + \text{h.c.})\, H_1^{i*} H_1^i + (m_2^2 + |\mu|^2 + \text{h.c.})\, H_2^{i*} H_2^i \\ & - (m_{12}^2 + \text{h.c.})\, \epsilon_{ij} H_1^i H_2^j + \tfrac{1}{2} g^2 |H_1^{i*} H_2^i|^2 \\ & + \tfrac{1}{8}(g^2 + g'^2) \left(H_1^{i*} H_1^i - H_2^{i*} H_2^i \right)^2 \end{aligned} \tag{2.85}$$

where each field (without a hat) denotes the scalar field component of the superfield (with a hat). The free parameters m_1, m_2, m_{12} and μ all have dimensions of mass, and the Higgs doublet notation used corresponds to the earlier, general two-doublet model of equation 2.74, in the following way:

$$\begin{aligned} H_1 &= \begin{pmatrix} H_1^1 \\ H_1^2 \end{pmatrix} = \begin{pmatrix} \phi_1^{0*} \\ -\phi_1^- \end{pmatrix} \Rightarrow \langle H_1 \rangle = \begin{pmatrix} v_1 \\ 0 \end{pmatrix} \\ H_2 &= \begin{pmatrix} H_2^1 \\ H_2^2 \end{pmatrix} = \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix} \Rightarrow \langle H_2 \rangle = \begin{pmatrix} 0 \\ v_2 \end{pmatrix} \end{aligned} \tag{2.86}$$

If one now works out the minimization constraints that guarantees v_1 and

v_2 to be non-zero, various relations between parameters can be found [33]:

$$\begin{aligned} \lambda_2 &= \lambda_1 \\ \lambda_3 &= \tfrac{1}{8}(g^2 + g'^2) - \lambda_1 \\ \lambda_4 &= 2\lambda_1 - \tfrac{1}{2}g'^2 \\ \lambda_5 &= 2\lambda_1 - \tfrac{1}{2}(g^2 + g'^2) \\ \lambda_6 &= \lambda_5. \end{aligned} \tag{2.87}$$

The last of these relations shows that it is possible to rotate away the phase ξ of equation 2.74 by a redefinition of one of the fields without affecting the other terms in the potential. Thus, the phase ξ is omitted, *i.e.* no CP-violation occurs, which has already been assumed in the treatment of the general two-doublet model.

It is also possible to obtain relations between the mass parameters of equation 2.85 and other fundamental constants of the theory:

$$\begin{aligned} m_1^2 &= -|\mu|^2 + 2\lambda_1 v_2^2 - \tfrac{1}{2}m_Z^2 \\ m_2^2 &= -|\mu|^2 + 2\lambda_1 v_1^2 - \tfrac{1}{2}m_Z^2 \\ m_{12}^2 &= -\tfrac{1}{2}v_1 v_2 (g^2 + g'^2 - 4\lambda_1). \end{aligned} \tag{2.88}$$

These relations, together with the ones in 2.78 and 2.80, give predictions for several of the different Higgs boson tree-level masses, in terms of other parameters of the model. Several different choices for independent parameters of the Higgs sector of the MSSM can be found [28]; in particular, $\tan\beta$ and any one of the two Higgs masses m_{A^0} and $m_{H^\pm}$ predict, at tree-level, the other Higgs boson masses through the relations

$$\begin{aligned} m_{A^0}^2 &= m_{12}^2(\tan\beta + \cot\beta) \\ m_{H^\pm}^2 &= m_{A^0}^2 + m_W^2 \\ m_{H^0,h^0}^2 &= \tfrac{1}{2}\left[m_{A^0}^2 + m_Z^2 \pm \sqrt{\left(m_{A^0}^2 + m_Z^2\right)^2 - 4m_Z^2 m_{A^0}^2 \cos^2 2\beta}\right]. \end{aligned} \tag{2.89}$$

From the expressions in 2.89 several important relations between masses of the Higgs and gauge bosons of the theory can be found:

$$\begin{gathered} m_{H^\pm} \geq m_W, \quad m_{H^0} \geq m_Z, \quad m_{A^0} \geq m_{h^0}, \\ m_{h^0} \leq \min\{m_Z, m_{A^0}\}|\cos(2\beta)| \leq m_Z \end{gathered} \tag{2.90}$$

It is worth emphasizing that these relations are based on tree-level calculations, and that several of them can be violated when introducing loop corrections.

2.5.4 Radiative corrections in the two-doublet model

So far, all results concerning the Higgs boson masses and couplings in the two-doublet model have been based on tree-level calculations. As in the case of one Higgs boublet, radiative corrections induce new effects, however, the analyses of these are generally more complicated in the two-doublet case.

It is important to notice that the mass of the pseudoscalar A^0 can become very light, as no Linde-Weinberg type of lower limits exists for this particle. The reason for this is the following: in the case of m_{A^0} being zero at tree-level, there will be an extra $U(1)$ global symmetry in the Higgs potential which is spontaneously broken, with the A^0 being identified as the Goldstone boson corresponding to this broken $U(1)$ symmetry. Therefore, the analysis focuses on the scalar Higgs sector.

Also, it is convenient to redefine the scalar fields ϕ_1^0 and ϕ_2^0 so that only one field has a vacuum expectation value:

$$\xi = \cos\beta\ \Re\{\phi_1^0\} + \sin\beta\ \Re\{\phi_2^0\}, \qquad \eta = -\sin\beta\ \Re\{\phi_1^0\} + \cos\beta\ \Re\{\phi_2^0\} \tag{2.91}$$

which gives corresponding vacuum expectation values

$$\langle\xi\rangle = \frac{v}{\sqrt{2}} = \sqrt{v_1^2 + v_2^2}, \qquad \langle\eta\rangle = 0. \tag{2.92}$$

The Coleman-Weinberg potential

As already mentioned, even though the Coleman-Weinberg lower Higgs boson mass limit does not yield a physically interesting result in the one-doublet model, quite the opposite can be true in the two-doublet model, depending on specific parameters of the model. If the masses M_i of the different particles in the theory which couple to the field ξ are assumed to depend on v according to the formula

$$M_i^2(\xi)|_{\xi=v/\sqrt{2}} = \mu_i^2 + \lambda_i v^2 \tag{2.93}$$

(the simplest possibility), the Coleman-Weinberg mass limit now takes the form

$$m_{\mathrm{CW}}^2 = \frac{v^2}{8\pi^2}\mathrm{Str}\left\{\lambda_i^2\left(1 - \frac{\mu_i^2}{\lambda_i v^2}\log\left(\frac{\mu_i^2 + \lambda_i v^2}{\mu_i^2}\right)\right)\right\}, \tag{2.94}$$

where the notation

$$\mathrm{Str}\{\cdots\} = \sum_i C_i(2J_i + 1)(-1)^{2J_i}\{\cdots\}$$

is introduced, as in equation 2.64. When dividing the particles i of the theory into two classes j and k such that $\mu_j = 0$ and $\mu_k \gg v$, the result in 2.94

takes a form very reminiscent of the one found in 2.64:

$$m^2_{\mathrm{CW}} = \frac{1}{8\pi^2 v^2}\mathrm{Str}M_j^4 + \frac{v^4}{16\pi^2}\mathrm{Str}\frac{\lambda_k^3}{\mu_k^2} \tag{2.95}$$

This is now the Higgs boson mass generated from the assumption of a vanishing quadratic term in the Higgs potential. In terms of the fields η and ξ, m^2_{CW} corresponds to the diagonal matrix element of the Higgs mass-squared matrix in this basis.

The Linde-Weinberg limit

As in the one-doublet case, requiring a stable vacuum, *i.e.* that the symmetry breaking vacuum is a global minimum, can be used to set a lower limit on the mass of the Higgs boson. The new expression for the Linde-Weinberg mass is now

$$m^2_{\mathrm{LW}} = \frac{v^2}{16\pi^2}\mathrm{Str}\left\{\lambda_i^2\left(1 - \frac{2\mu_i^2}{\lambda_i v^2}\left[1 - \frac{\mu_i^2}{\lambda_i v^2}\log\left(\frac{\mu_i^2 + \lambda_i v^2}{\mu_i^2}\right)\right]\right)\right\}, \tag{2.96}$$

which, when, as for the Coleman-Weinberg case, one assumes the particles divided into two classes with $\mu_j = 0$ and $\mu_k \gg v$, takes the form

$$m^2_{\mathrm{LW}} = \frac{1}{16\pi^2 v^2}\mathrm{Str}M_j^4 + \frac{v^4}{24\pi^2}\mathrm{Str}\frac{\lambda_k^3}{\mu_k^2}. \tag{2.97}$$

This mass is however a limit in the basis of η and ξ, which is a mixing of the physical states h^0 and H^0. Thus, the Linde-Weinberg mass bound is of the form

$$m^2_{H^0}\cos^2(\beta - \alpha) + m^2_{h^0}\cos^2(\beta - \alpha) > m^2_{\mathrm{LW}} \tag{2.98}$$

where m_{h^0} and m_{H^0} are the radiatively corrected physical Higgs masses of the two neutral scalar Higgs bosons.

2.6 Alternatives to the Higgs mechanism

Although the structure of one or more fundamental scalar Higgs fields provides a theoretical framework for describing both the electro-weak symmetry breaking and the generation of mass terms for the fermions and weak vector bosons, alternative theoretical approaches exist. The emergence of such theories are in many cases motivated by the problems connected to the Higgs sector mentioned in section 2.4.4, and in particular the fine-tuning problem.

The introduction of a new symmetry (*i.e.* supersymmetry) in order to ensure the cancellation of quadratically divergent terms in the Higgs mass (the factor C of equation 2.71) is one way to cure the fine-tuning problem. Another solution is the possibility of leaving elementary scalar bosons out of the theory entirely, and describing the equivalent of the Higgs bosons as composite particles, where the non-elementary nature of the "Higgs boson" would become revealed at energies where the fine-tuning effects set in. A short summary of such models, with references to more detailed literature, can be found in [34]. Of these models, one of the most popular is the technicolor approach.

2.6.1 Technicolor

Technicolor models [35, 36, 37] are based on the theoretical framework and understanding of QCD, where the gauge coupling becomes strong at an energy of approximately 200 MeV (often called Λ_{QCD}), breaking the chiral symmetry $SU(6)_L \times SU(6)_R$ down to the diagonal (vectorial) $SU(6)$ subgroup, giving rise to pseudo-Goldstone bosons, *i.e.* the pions as quark-condensates. Technicolor introduces a new set of fermions, the so-called technifermions, which are subject to all the usual interactions of the $SU(3)_C \times SU(2)_L \times U(1)_Y$ structure of the SM, but in addition carry their own technicolor charges. If this new techniforce becomes strongly interacting at an energy of several hundred GeV (named Λ_{TC}, in reflection of Λ_{QCD}), the resulting pseudo-Goldstone bosons in the technicolor sector (called technipions, analogous to the ordinary pions of QCD) can play the role of the scalar Higgs boson and give masses to the weak vector bosons through the absorption of three such technipions for the three longitudinal components of W^+W^- and Z^0.

This method of generating the electro-weak symmetry breaking has several theoretically appealing features. First, as technicolor, like QCD, is asymptotically free, the fine-tuning problem, as well as the hierarchy and triviality problems, are not present in the theory. The fine-tuning described by equation 2.71 simply does not exist, as the scale μ_1 must be replaced by the much lower scale Λ_{TC}, and the scale of the Higgs boson mass becomes naturally this new energy scale. If the technicolor gauge symmetry is embedded in a larger symmetry at a very large energy Λ (such as the GUT or Planck scale) with a relatively weak coupling, the scale Λ_{TC} is naturally exponentially smaller than Λ, explaining the large difference in scale between the GUT/Planck scale and the electro-weak scale (the hierarchy problem). Finally, asymptotically free theories are non-trivial, something which can be seen from a plus sign instead of a minus sign in the analog of equation 2.67.

On the other hand, there are generally more technipions left after the generation of the longitudinal components of the W^+W^- and Z^0, and these are, in most models, usually detectable and sufficiently light for experimental production. Also, the generation of masses for the ordinary SM fermions of the theory (quarks and leptons) is not explained by this mechanism, and technicolor models thus need to be further extended. Such extensions are known as extended technicolor.

Extended Technicolor

In extended technicolor models [38, 39], masses for quarks and leptons are generated through additional interactions that couple these fermions to the techniquarks. The gauge structure of the SM is unified with the technicolor gauge structure into a larger Extended Technicolor (ETC) gauge group, called G_{ETC}. The specific structure of such models differ, but they are all plagued with several difficulties. First, FCNCs generated by the ETC mechanisms tend to be much larger than experimentally allowed, and require careful avoidance in the construction of the models. Second, the generation of the top quark mass, being very high, requires fine-tuning in the ETC mechanism, and brings the mass of the required ETC boson down to $\sim$1 TeV, where it would interfere with the technicolor dynamics which generate the electro-weak symmetry breaking. And finally, electro-weak parameters tend to be modified by the ETC interactions at a level unacceptable to precision electro-weak measurements.

There are two ways in which technicolor models are usually modified to meet these requirements: walking technicolor and topcolor-assisted technicolor.

Walking Technicolor

In the discussions so far, technicolor has been assumed to be simply a scaled-up version of QCD. If one relaxes this assumption, and allows effects in the ETC gauge dynamics not present in QCD, the problems of the previous section can be addressed. In QCD, the evolving gauge coupling α_s goes rather quickly to the weakly coupled regime, due to asymptotic freedom. If, on the other hand, the ETC coupling evolves rather slowly (*i.e.* "walking" coupling constant, as opposed to "running" coupling constant), the FCNC effects can be suppressed to an acceptable level, and the electro-weak precision measurements can be accomodated. The resulting theories are known as walking technicolor [40]. Whereas such theories can correct the disagreement with experiment with regard to the FCNCs and electro-weak precision quantities, it does not explain the large top quark mass. Additional mechanisms have been proposed to solve this problem, of which topcolor-assisted technicolor is a particularly promising one.

Topcolor-Assisted Technicolor

The original idea, developed in the early 1990s, assumed a new, special interaction known as topcolor [41], for the third quark generation. This involves a large top-quark condensate $\langle \bar{t}t \rangle$ responsible for the symmetry breaking which generates the top quark mass. However, the simplest approach is an unnatural one, and there are also problems connected with the large mass difference between the top and bottom quarks. The two concepts of topcolor and technicolor were later added into what is known as topcolor-assisted technicolor [42], where the electro-weak symmetry breaking is driven mainly by technicolor interactions strong near 1 TeV, light fermion (and technipion) masses are generated by ETC, and the topcolor interactions, also at scales near 1 TeV, generate the $\langle \bar{t}t \rangle$ condensate and the large top quark mass. This is a rather young field in constant development, and new discoveries concerning these theories should be expected. Whether this can provide a definite cure for the problems connected with the technicolor approach, remains to be seen.

2.6.2 Extra dimensions

A completely different explanation to the fine-tuning problem is offered by the development of multi-dimensional theories, usually in connection with string theory [43]. Here, as is the case for the technicolor approach, the fine-tuning effect present in equation 2.71 vanishes because the scale μ_1 is no longer a scale much larger than μ_2, but some smaller scale of roughly the same order as the Higgs mass itself. For string theory, this new scale appears due to large extra dimensions, and is therefore related to the compactification scale of these; see reference [44] for details.

At first, the Higgs mechanism (with either a fundamental or a composite scalar playing the role of the Higgs boson) seems unavoidable; the only known renormalizable theories of massive vector bosons in 4 dimensions, are gauge theories with spontaneous symmetry breaking mediated by the Higgs mechanism [45, page 1]. However, for theories formulated in $4+N$ dimensions, where the extra N dimensions compactify, there are other possibilities. Firstly, extra-dimensional components of gauge fields can, through various mechanisms, generate scalar fields which play the role of the Higgs field and generate massive vector bosons and masses for the fermions of the theory (see for instance refs. [46] and [47] for two such examples). It is even possible, as shown in ref. [45], to construct models where a scalar field is entirely absent from the theory, but where both massive vector bosons and massive fermions remain. This would then provide a theoretical framework where the Higgs boson is entirely superfluous, and thus a viable alternative to the Higgs mechanism. By starting with a pure gauge theory in more than 4 dimensions, and making use of the localization of a vector field on a lower-dimensional defect, the procedure, as taken from [45], is as follows.

In 5 dimensions (the generalization to higher dimensions is straightforward), the action can be expressed as

$$S = -\frac{1}{4}\int \mathrm{d}^4x\, \mathrm{d}z \Delta(z) F_{AB}F^{AB}, \tag{2.99}$$

where the normal 4 dimensions are described by x^ν, and the fifth being z. F^{AB} is the ordinary field strength, and $\Delta(z) > 0$ is the weight function, depending, in general, on the fifth coordinate.

In the case of a purely Abelian gauge field $A_B(x_\nu, z)$, and expanding this field in a Fourier-type series along the coordinate z,

$$A_B(x_\nu, z) = \sum_n A^n_B(x_\nu)\psi_n(z), \tag{2.100}$$

the equations of motion for the gauge field gives the following equations for the $\psi_n(z)$ fields:

$$-\frac{1}{\Delta(z)}\frac{\partial}{\partial z}\left(\Delta(z)\frac{\partial}{\partial z}\psi_n(z)\right) = m_n^2\psi_n(z) \tag{2.101}$$

with the orthogonality and completeness conditions

$$\int \mathrm{d}z \Delta(z)\psi_n(z)\psi_m(z) = \delta_{mn}, \quad \sum_n \psi_n(z)\psi_n(z') = \frac{1}{\Delta(z)}\delta(z - z'). \tag{2.102}$$

The fields $A^n_\mu(x_\nu)$ now describe vector fields in the ordinary 4-dimensional time-space, whereas $A^n_z(x_\nu)$ are scalars in the extra fifth dimension.

The m_n of equation 2.101 give the masses of the fields, all non-negative. It now turns out that, for a wide class of weight function $\Delta(z)$ ($\Delta(z)$ being an even function which decreases at low z, reaches a minimum and then grows sufficiantly rapidly with increasing z), m_0 is non-zero with a gap to the rest of the eigenvalues m_n. This is a situation which will give the desired spectrum of a 4-dimensional gauge theory of massive vector bosons, without any fundamental scalars.

It should be noted that neither of these extra dimensional theories are to be considered realistic models which yield the SM as an effective low-energy theory. Indeed, this is neither the intention nor the expectation of such models, as the string theorists themselves freely admit. To quote one such theorist: “In this paper our intention is not to recover the standard model of particle physics from a 10-dimensional theory. In fact we believe that within our present understanding this is not possible.” [47, page 7]. However, string theories offer the only currently available serious candidate for building a Theory of Everything (TOE), *i.e.* a theory which encompasses all the forces of nature: gravity, currently described by Einsteins theory of general relativity, the strong force, currently described by QCD, and the electro-weak force, currently described by the GWS theory, would all be united once one enters the domain where the string interactions become the important physical processes. What role, if any, the Higgs mechanism and the Higgs boson plays in such a scheme is yet to be seen, but a TOE must under any circumstances address and explain the questions of electro-weak symmetry breaking and massive fermions that the Higgs boson currently provides the most widely accepted answers to.

With the theoretical framework presented in this chapter, and from that the upper limits on the Higgs mass and consequently the energy scale at which Higgs-like (or alternative) phenomena must appear, there is an exciting experimental outlook on the present and relatively near future for experimentalists in high energy physics world-wide. Combined, the experiments at LEP-*II*, the Tevatron, a proton/anti-proton collider operating at centre-of-mass energies up to 2 TeV at Fermilab, the Large Hadron Collider (LHC), a proton/proton collider which is to be built in the LEP tunnel, operating at centre-of-mass energies of up to 14 TeV when finished in 2006, and possibly one or more of the proposed Next Linear Colliders (NLC) to be built in the next 10–20 years, where electrons and positrons are collided at centre-of-mass energies of 500 GeV or more, should cover most, and possibly all of the consistent parameter space for all models of Higgs or Higgs-like mechanisms. Therefore, one expects to discover either the Higgs boson itself, or new physics playing the role of the Higgs boson, within the next $\sim$10 years. The first step in this process was taken in the autumn of 1995, when LEP-*II* started running, at energies significantly above the Z^0 mass. The next chapter in this thesis describes the LEP accelerator and the experimental tool used in the present analyses of the LEP data, the DELPHI detector. LEP and its pre-accelerators and injectors are briefly described, before the DELPHI general layout, different sub-detectors and trigger system are presented.

Chapter 3

LEP and the DELPHI detector

3.1 LEP

The Large Electron Positron collider (LEP) at CERN started operating on August 13, 1989, colliding electrons and positrons at high energies. Up until 1995 the accelerator was operated at centre-of-mass energies around the Z^0 mass of 91.19 GeV [22], whereas the end of the 1995 run saw the first of several upgrades enabling continually increased centre-of-mass energies to be reached. This process culminated during the 2000 run, in which energies up to 209 GeV were achieved. The last electron-positron collisions in LEP occurred on November 2, 2000, and the accelerator has since been dismantled to prepare for the next large particle accelerator at CERN, the LHC (Large Hadron Collider), which will be the next generation accelerator at CERN, colliding protons against protons at centre-of-mass energies of up to 14 TeV (in addition to an extended heavy ion program).

The LEP injection system consists of a number of smaller accelerators and injectors, as shown in figure 3.1. An electron/positron which eventually circulates in the LEP ring, goes through the following chain of accelerators:

LEP Injection Linacs (LIL): The first linear accelerator brings electrons, produced by an electron gun, to energies of 200 MeV before colliding them against a tungsten target. This produces hard gamma radiation, which in turn converts to electron-positron pairs. The second linac accelerates these up to energies of 600 MeV.

Electron Positron Accumulator ring (EPA): The EPA stores the electrons and positrons in bunches and serves as a buffer for the synchrotrons which are next in the chain.

Proton Synchrotron (PS): Originally built in 1959 as a proton-proton collider, the PS is an integral part of the LEP accelerator system, and accelerates the electrons and positrons up to 3.5 GeV.

Super Proton Synchrotron (SPS): The final pre-accelerator, originally built as the successor of the PS with which the two experiments UA1 and UA2 discovered the intermediate vector bosons W^+, W^- and Z^0 [5, 6], accelerates the particles to 20 GeV before injecting then into the LEP ring.

The LEP ring itself consistes of eight circular segments 2840 m in length, and eight straight sections of length 490 m, making the total circumference of LEP 26.7 km. Electron and positron beams consisting of an even number of bunches, of which four bunches has been the scheme most frequently used, are accelerated in the LEP storage ring and collided at four interaction points where the LEP experiments (ALEPH, DELPHI, L3 and OPAL) are located. These are located in the straight sections, and two of the remaining straight sections house the radio-frequency cavities (RF), which accelerate the beams from the injection energy of 20 GeV up to the beam energy.

3.2 The DELPHI detector

The DELPHI detector [49, 50, 51] (DEtector with Lepton, Photon and Hadron Identification) is one of four detectors at LEP. It is designed as a general purpose detector with emphasis on particle identification, which is accomplished by ring imaging Cherenkov counters, three-dimensional information with high granularity and good vertex determination.

3.2.1 General layout

The DELPHI detector is situated in a cavern 100 meter below ground level at Interaction Point 8 (IP8; see figure 3.1) in the LEP ring, with the main computer and control centre in a surface building. The general layout of the detector is shown in figure 3.2, where the individual sub-detectors are also indicated. The coordinate system adopted in DELPHI has the z-axis along the beam pipe, with positive z in the direction travelled by the electrons. The x-axis points towards the centre of the LEP ring, and the y-axis points upwards. An alternative set of coordinates is a polar angle θ to the z-axis (with $\theta = 0$ along positive z), an azimuthal angle ϕ around the z-axis and a radial coordinate R given by $R = \sqrt{x^2 + y^2}$. The detector is divided into a barrel part, covering polar angles of about 40° to 140°, and two forward endcap parts, covering the remaining polar angles. The plane at $z = 0$

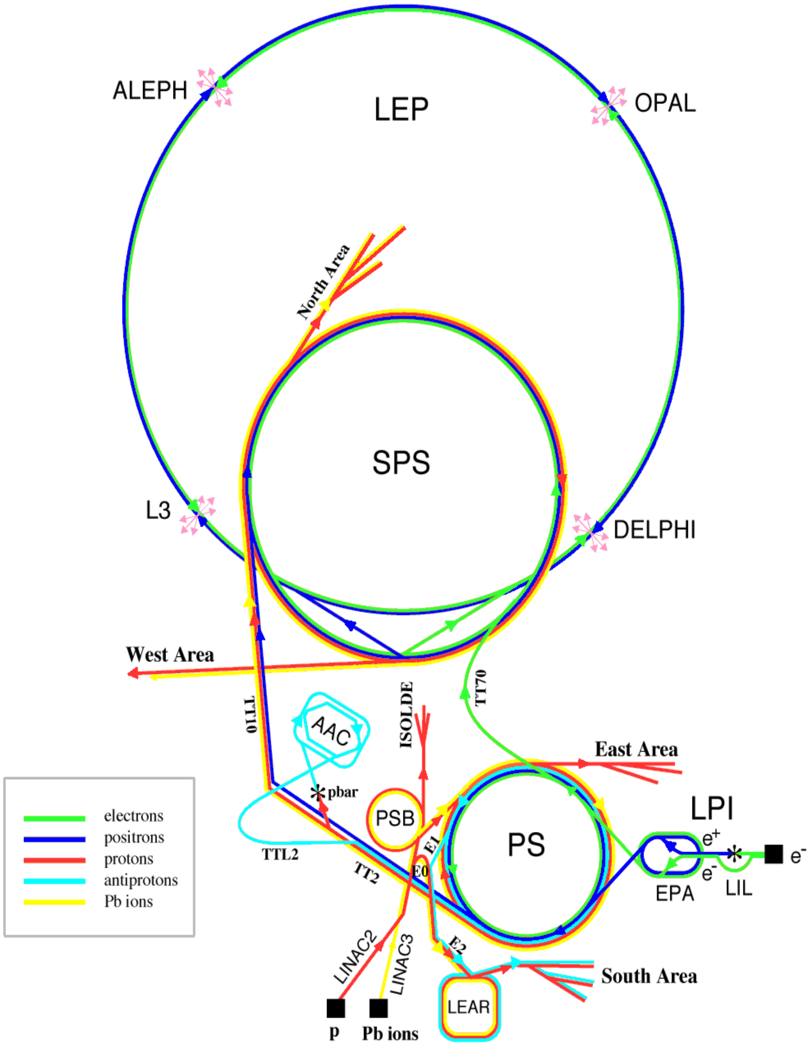


Figure 3.1: Schematic figure showing the LEP accelerator complex and the different pre-accelerators and injectors involved in producing the high-energy electron and positron beams in the LEP ring. Figure taken from [48]

divides DELPHI into two hemibarrels, denoted A and C, which are also used to denote the two different endcaps.

3.2.2 Tracking

The tracking of charged particles through the detector is achieved in DELPHI through the use of several sub-detectors, each with its own special features and purposes.

Microvertex detector (VD)

The DELPHI Microvertex Detector [52, 53] (MVD, more commonly known as the VD) is a silicon strip detector placed as close as possible to the beam pipe. Its primary objective is to give good $R\phi$ resolution for charged tracks, and it is the most important tool for heavy flavour tagging. This sub-detector has gone through a series of upgrades throughout the history of DELPHI, and the performance has been substantially improved, both in terms of precision and angular coverage. The original VD was a two-layer barrel silicon strip detector with concentric layers (known as the Outer and Inner layers) of length 24 cm at average radii 9 and 11 cm from the centre of the beam pipe. Each layer consists of 24 modules, each of 4 silicon detectors with strips along the beam direction. There is a $\sim$10% overlap in ϕ between modules. Each silicon detector has a width of 285 μm, a diode pitch of 25 μm and a readout pitch of 50 μm. There are 512 readout channels in the Inner layer, and 640 readout channels for the outer layer, giving a total of 54254 silicon strips. The impact parameter resolution in $R\phi$, measured on di-muon events, for this setup is well described by the expression

$$\sigma_{\mathrm{IP}} = \sqrt{(80)^2 + \left(\frac{120}{p_t}\right)^2}\ \mu\mathrm{m},$$

where p_t is the transverse momentum measured in GeV/c.

Before the 1991 run, a new beam pipe was installed in DELPHI, reducing the outer radius of the beam pipe from $\sim$7.9 cm to $\sim$5.4 cm. This made it possible to include a third layer in the Microvertex Detector, the Closer layer, at an average radius of 6.3 cm from the centre of the beam pipe and a length of 22 cm. The number of readout channels for this layer is 384, which brings the total number of strips to 73728. The $R\phi$ resolution of the VD was now measured to be

$$\sigma_{\mathrm{IP}} = \sqrt{(24)^2 + \left(\frac{69}{p_t}\right)^2}\ \mu\mathrm{m},$$

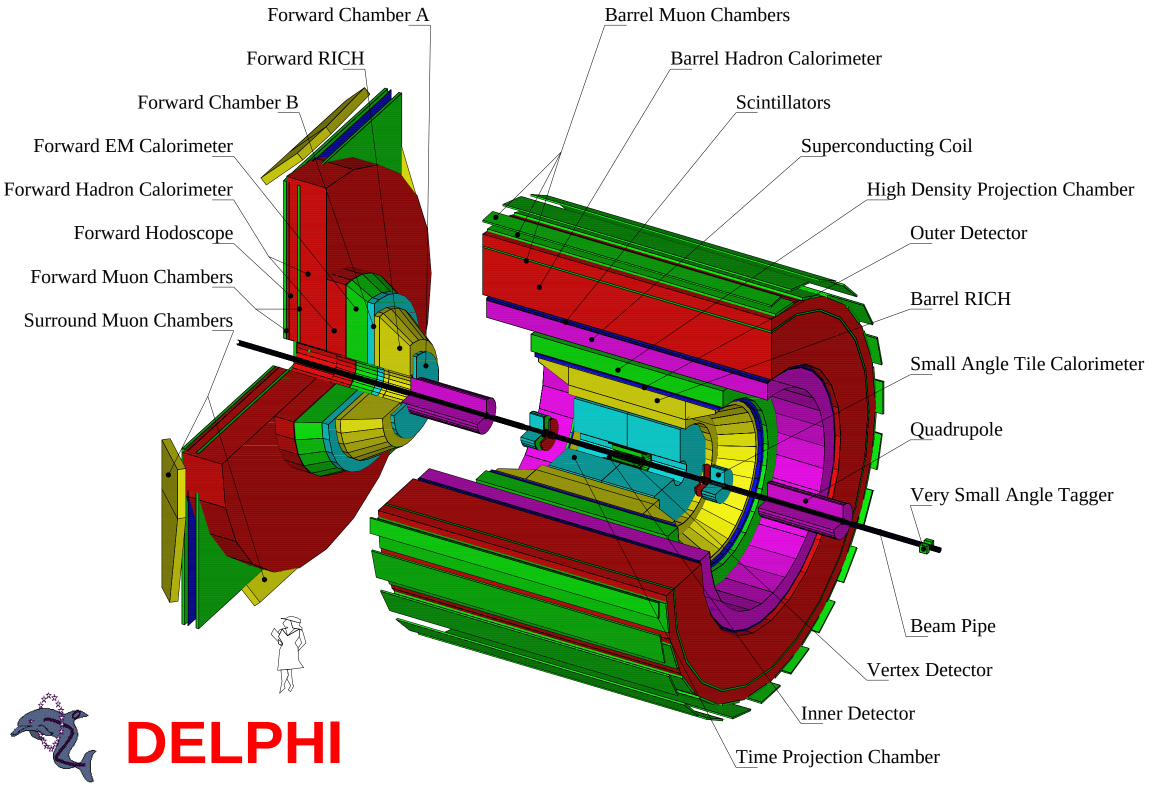


Figure 3.2: Figure giving a schematic view of the DELPHI detector with the sub-detectors indicated. The barrel part and one of the two endcaps are shown. Figure taken from [50].

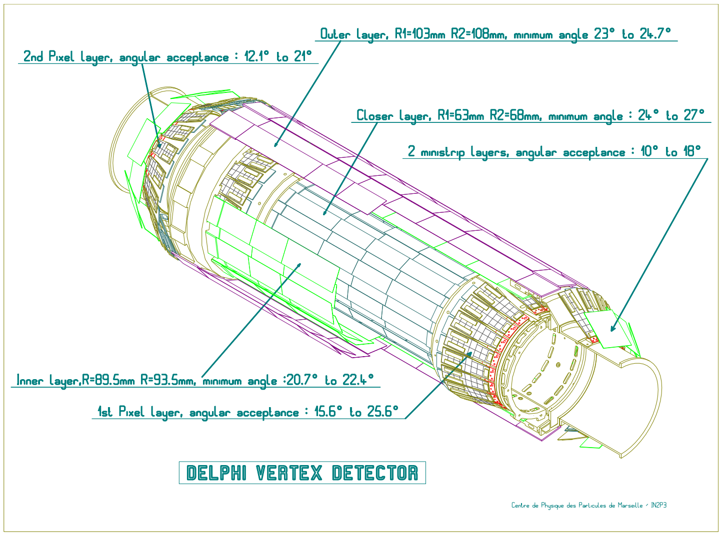


Figure 3.3: Figure showing the general layout of the Microvertex Detector, including the pixels and ministrips of the Very Forward Tracker. Figure taken from [50].

a significant improvement from the earlier, two-layer configuration.

The next change of the VD was performed in the spring of 1994, and saw a major upgrade of the detector. The single sided detector modules in the Closer and Outer layers were replaced by double sided silicon strip detectors with orthogonally oriented strips, enabling the layers to provide three coordinates reconstruction. The Inner layer retained its use of single sided modules, but by reusing detector modules from the old Inner and Outer layers, the overlap between neighbouring modules was increased to ∼20%. Also, the Closer layer was extended in the z-direction to cover polar angles down to 25°. The number of readout strips now totals 125952.

The impact parameter resolution in the $R\phi$ direction for this new VD was now measured to be

$$\sigma_{\mathrm{IP}} = \sqrt{(20)^2 + \left(\frac{65}{p_t \sin^{\frac{3}{2}}\theta}\right)^2}\ \mu\mathrm{m},$$

where θ is the polar angle of the track. This is not a large improvement with respect to the previous value, but the main gain is achieved in the Rz plane. The impact parameter resolution in the Rz plane is strongly dependent on the polar angle and momentum of the track, and ranges from 46 μm for tracks with momentum above 8 GeV/c and polar angle less than 10° away from the vertical (90°), and 203 μm for tracks with momentum around 1 GeV/c and polar angle between 45° and 55°. This represents a general improvement of almost a factor 20 with respect to the Rz plane impact parameter resolution one obtains from a fit with only $R\phi$ information.

The final upgrade of the VD was performed in 1996, and consisted of two major steps. Firstly, the Outer and Inner layers were upgraded, doubling their lengths from 24 cm to 48 cm, and the Inner layer was equipped with double sided detector modules. The Closer layer was left unchanged. In addition, the detector was improved in the forward region at polar angles between 10° (170°) and 25° (155°) by the inclusion of the Very Forward Tracker (VFT). The VFT consists of two parts, one in each hemibarrel, mounted on the end of the barrel VD. Each part consists of two planes of ministrips detectors [54], and two planes of pixel detectors [55]. Each plane of the ministrips consists of two half rings with 6 detector modules each surrounding the beam pipe. Each module consists of two single-sided strip detectors glued back to back orthogonally oriented. Each detector has a read out pitch of 200 μm. The ministrips total 25376 readout channels, and has a spacial resolution on track elements of 10 to 30 μm, depending on the track inclination. For the pixel detectors, each part consists of 38 modules, each of 8064 square pixels of 330 μm pitch. The total number of channels in the pixel detectors amount to 1225728 detector elements (pixels), of which 5/8 were installed in 1996, and the remaining ones installed in 1997. Thus, the final version of the VD has a general layout as shown in figure 3.3.

Inner detector (ID)

The Inner Detector (ID) is situated just outside the VD, covering radii from 11.8 to 28 cm. The sub-detector is made up of two parts, the inner part, the jet chamber, is a drift chamber of 24 azimuthal sectors at radial range up to $\sim$23 cm, giving 24 $R\phi$ points. The angular coverage is 15° to 165° for tracks with hits in the 10 innermost wires. Outside the jet chamber is a cylindrical structure of five layers of straw tubes, with a total of 192 tubes in each layer. The tubes have a width of $\sim$8 mm, and are staggered by half the width of a tube in subsequent layers. Information from these straw tube layers provide up to 5 $R\phi$ points, and also play an important role in the trigger.

Both the jet chamber and the straw tubes measure $R\phi$ coordinates for tracks. The single wire resolution of the jet chamber is of the order of 90 μm, giving a total track element resolution of $\sim$40 μm in $R\phi$ and about 1.2 mrad in ϕ. The $R\phi$ resolution of the straw tubes is approximately 150 μm, which therefore resolves the left/right ambiguity of the drift chamber.

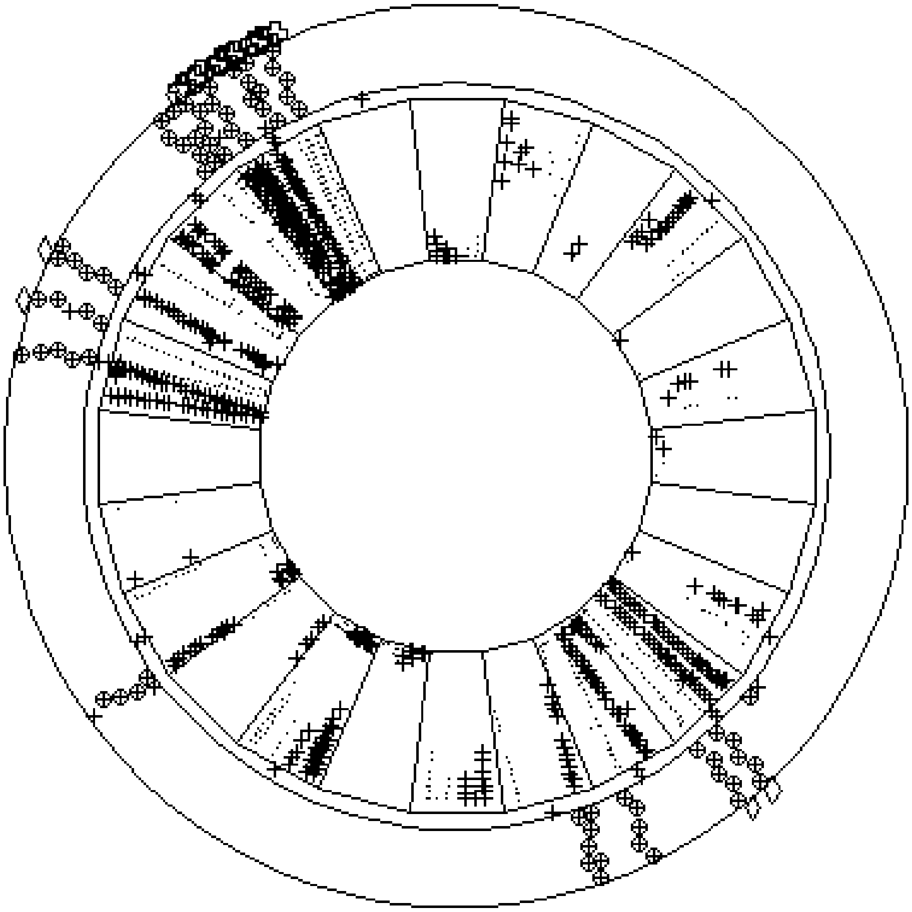


Figure 3.4: Figure showing a transverse view of the Inner Detector. Track points in the jet chamber (a maximum of 24 points per track) are shown as crosses, whereas track points in the straw tubes (a maximum of 5 points per track) are shown as crosses in circles. Figure taken from [50].

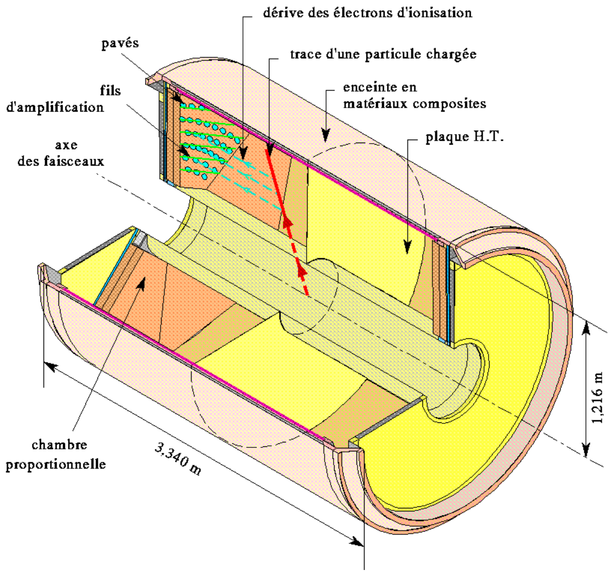


Figure 3.5: The general layout of the Time Projection Chamber (TPC), the main tracking device of DELPHI. Figure taken from [50].

Time projection chamber (TPC)

The Time Projection Chamber (TPC) is the main tracking device of DELPHI, occupying the barrel region outside the ID up to radii of 122 cm (radial acceptance from 35 cm to 111 cm), and with a total length of 334 cm in z. At $z = 0$, a high voltage plane, creating a drift field of 187 V/cm, divides the detector into two drift volumes. Each endplate of the TPC is divided into six azimuthal sectors, each with 192 sense wires, and 16 circular pad rows with constant radial spacing, providing up to 16 space points for each charged particle traversing the TPC volume. The sense wires, of diameter 20 μm, have a spacing of 4 mm, and a voltage of 1435 V, with the drift velocity in the gas volume (80% Ar, 20% CH_4) being 7 cm/μs. This velocity is monitored constantly by six miniature nitrogen lasers, one for each sector, at each endcap, giving a relative drift velocity measurement of better than 2×10^{-4}. The general layout of the TPC is shown in figure 3.5.

The single point precision of the TPC, measured on dimuons, is found to be 250 μm in $R\phi$ and 880 μm in Rz, with the two-point resolution being $\sim$1 cm in both directions. Distortions in the $R\phi$ and Rz planes limit the precision of the track elements to $\sim$150 μm in $R\phi$ and $\sim$600 μm in z.

Outer detector (OD)

The Outer Detector (OD) is situated outside the Barrel-RICH (see page 61), covering radii of 198 cm to 206 cm, and $|z| < 232$ cm. Its primary goal is to provide fast trigger information with full coverage in ϕ, and to improve the momentum resolution, improving on the constraints the Barrel-RICH puts on the geometry of the TPC. The detector consists of 24 modules, each of 145 drift tubes in five layers, running the length of the detector (4.7 m) and operating in the limited streamer mode. Subsequent layers are staggered, and detector modules overlap in ϕ, providing full azimuthal coverage. All layers give points in $R\phi$, and three layers give in addition z information, by timing the signals at the end of the anode wires. The single point precision of the OD is measured at 110 μm in $R\phi$ and 3.5 cm in z.

Forward chambers A and B (FCA/FCB)

The forward chambers A and B play more or less the role in the forward direction that the OD plays for the barrel. The forward chamber A (FCA) consists of two halves mounted on the end of each side of the TPC. It covers polar angle regions of 11° to 32° (169° to 148°) in a plane of constant z (155 cm to 165 cm). One side consists of three chambers, each with two staggered layers and split into half-discs with an outer radius of 103 cm, running in the limited streamer mode. The wires of the three modules are rotated by 120° with respect to each other. See part (a) of figure 3.6 for details of the geometry of the staggered drift tubes. Test beam measurements show single wire average root mean square residuals of 190 μm, but with deterioration near the sense wires and in the corners of the drift tubes. Under normal operational conditions, where the particle direction is not known, this transforms to a track element precision of 290 μm in x and 240 μm in y, and 8.5 mrad in polar angle θ and 24 mrad in ϕ (averaged over θ).

The forward chamber B (FCB) is situated further from the interaction point than the FCA, in two modules at z positions of ± 267 cm to ± 283 cm, covering polar angle regions of 11° to 36° (169° to 144°). Each module is made up of two chambers each forming a half-disc, with a complete disc being a regular dodecagon of inner radius R=48 cm and outer radius R=211 cm. There are 12 read-out planes in each module, with the wire direction rotated by 120° (an internal coordinate system parallell to the one in FCA), giving a total of four space points. Precision on single track elements are 150 μm in x and y, 3.5 mrad in θ, and $(4.0/\sin\theta)$ mrad in ϕ.

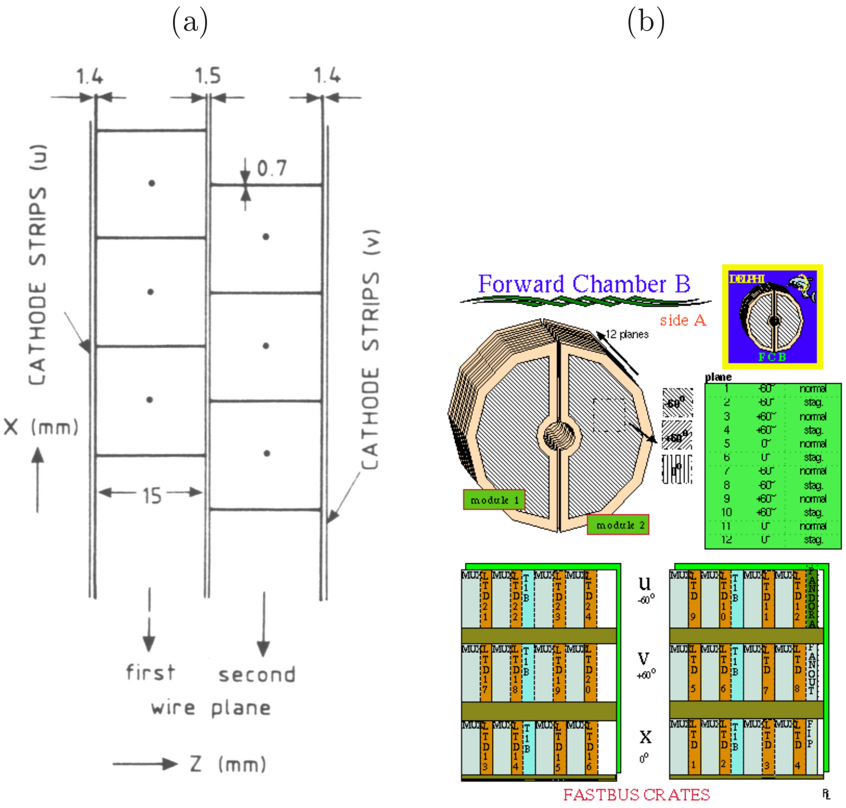


Figure 3.6: Figure showing details of the forward chambers A and B. Part (a) shows detail of the staggered double layers of FCA, whereas the general structure of FCB is shown in part (b). Figure (a) is taken from [49], whereas figure (b) is taken from [50].

Muon chambers (MUB/MUF/SMC)

The muon chambers are, with the exception of the luminosity monitor at very small angles, the VSAT (see page 66), the sub-detectors which are situated furthest from the centre (*i.e.* the interaction point) of DELPHI. Their primary objective is the detection and momentum measurement of minimum ionizing particles, which at LEP energies are muons. The sub-detector consists of a barrel part (MUB), covering a radial range of ∼445 cm to ∼485 cm and a range along the beam pipe of $|z| <\sim 365$ cm, and two endplates (MUF), situated at $z=\pm(463\text{–}500)$ cm and covering radial ranges of 70 cm to 460 cm. In addition to this, a new set of surround muon chambers (SMC), in all eight modules, were installed in 1994 to cover the intermediate region between the barrel and forward part of DELPHI.

The barrel muon chambers (MUB) is made up of three modules, each consisting of 24 sectors, with an additional 2 sectors between the legs of the detector. The inner module is a structure of 2×24 planks inserted into the return yoke after 90 cm of iron, where half of the planks makes up a shell covering ranges in z of 0 cm up to 365 cm, with the other half making up the corresponding shell on the negative z side. Each plank consists of 3 staggered layers of drift chamber, two of which contain 5 chambers and the third containing 4 chambers in a 5–4–5 arrangement. Two of the layers are read out, the third layer being regarded as a spare. The outer and peripheral modules of the MUB are situated outside of the return yoke behind a further 20 cm of iron, and both consist of 2 layers of 4 and 3 staggered chambers, respectively (see figure 3.7). The layers are placed so that the ones of the peripheral module cover the holes left by the structure of the outer module. The majority of the drift chambers, which are operated in the proportional mode, have an active length of 365 cm. Single hit resolutions of ∼1 mm in $R\phi$ and ∼10 mm in z transform to ∼2 mm in $R\phi$ and ∼80 mm in z when associating to extrapolated tracks (dimuons).

The forward muon chambers are made up of two halves, one in each endcap, each containing two detection planes. The first plane is embedded in the return yoke behind $\geq$85 cm of iron, the second plane behind a further ∼20cm of iron and the forward scintillators. Each plane consists of 4 quadrants of dimension 450 cm×450 cm×8 cm, each quadrant containing 2 layers of 22 drift chambers, staggered by 90° and operating in the limited streamer mode. Two space points are measured, x and y, with an accuracy of ∼5 mm.

The surround muon chambers (SMC) consist of 8 parts mounted on the side, top and bottom of both endcaps, covering the holes between the muon chamber coverage in the barrel and forward (see figure 3.8). Each part contains two modules, which again are composed of two detector planes each.

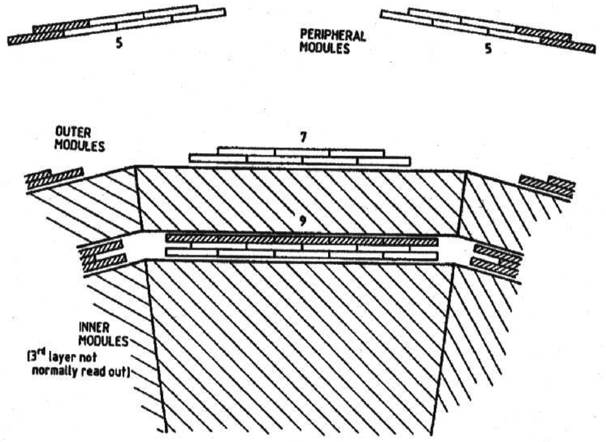


Figure 3.7: The general layout of one sector of the barrel muon chambers (MUB), showing the structure of the three modules and the configuration of the chambers. Figure taken from [56].

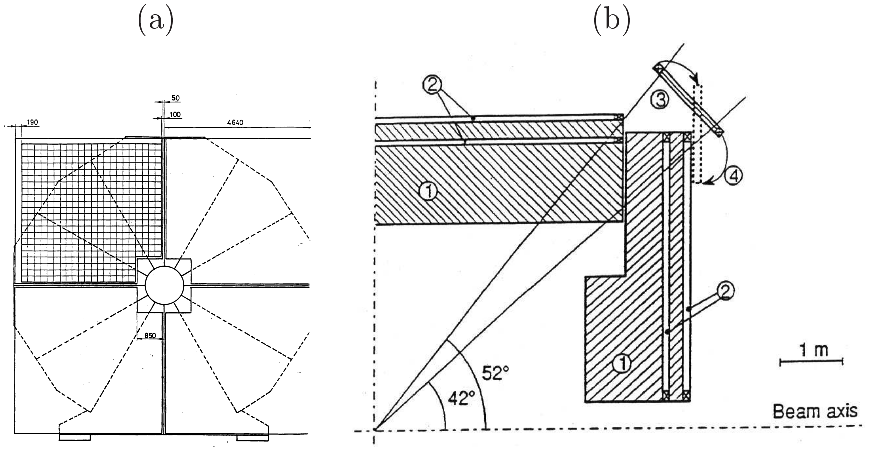


Figure 3.8: The forward and surround muon chambers. Part (a) shows the general layout of the four quadrant structure of the MUF in one endcap, whereas part (b) shows the coverage of the SMC. Figures taken from [57] (a) and [58] (b).

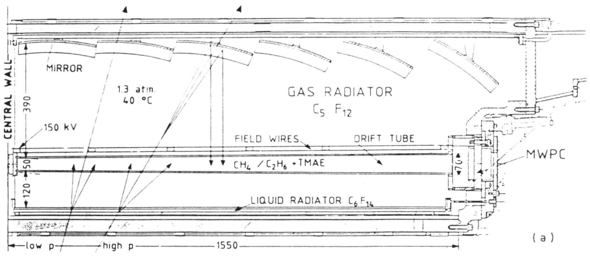


Figure 3.9: The general layout of the barrel-RICH detector, showing the liquid radiator, the drift volume and the gas radiator, with the principle of production and detection of Cherenkov light indicated. Figure taken from [49].

3.2.3 Ring imaging Cherenkov counters (RICH)

The ring imaging Cherenkov counters (RICH) in DELPHI are designed to provide particle identification based on the principle of Cherenkov radiation. This is achieved both by measuring the Cherenkov angle of the emitted Cherenkov photons, and as a veto counter for particles not emitting such photons. The detector is divided into two sub-detectors. In the barrel, the RICH has the structure of a cylindrical shell, and is located between the TPC and the OD at radial distances between 123 cm and 197 cm, and a length along the beam of 3.5 m. The forward RICH consists of two parts, one in each endcap, each occupying a truncated conical area at distances of $|z|$ between 172 cm and 266 cm, and radial range from 125 cm (end nearest the centre of DELPHI) to 180 cm (end furthest from the centre of DELPHI).

Barrel-RICH

The barrel-RICH is divided in two halves by a central plane at z=0. Each half is constructed with boxes of liquid radiators of thickness 1 cm and refractive index 1.278 near the inner radius, outside of which are drift tubes where the photons are detected. Outside of this again is a gas volume of thickness 40 cm and refractive index 1.00174, and Cherenkov photons produced here are reflected by parabloic mirrors back into the same drift tubes, at the end of which there are multi wire proportional chambers (MWPCs). On average, a charged track emits 12 Cherenkov photons in the liquid RICH, and 8 in the gas. See figure 3.9 for a detailed view.

Space points in three dimensions from the photon conversions are measured in the MWPC chambers at the end of the drift tubes from anode and cathode readout and time information. The detector provides 4.2σ separation of pions and kaons for particle momenta up to 18 GeV/c, and up to 33 GeV/c for separation of kaons and protons.

Forward-RICH

The forward-RICH consists of two endcap parts which are divided into two half-cones and 12 modular sectors. Each sector contains one driftbox, two MWPCs, three liquid-radiator containers and five mirrors. The MWPCs are mounted radially on the two borders of each 30° sector, and is equipped with two crossed layers of photon screens due to the crossed $E - B$ field configuration. On average, a charged particle produces ∼20 photoelectrons in the liquid, and ∼10 in the gas for each particle ring.

3.2.4 Scintillator counters

The scintillator counters are used mostly for fast triggering of beam events and cosmics. The scintillator counters originally installed in DELPHI are the time-of-flight counters in the barrel, and the forward hodoscope in the forward. In 1995, additional counters were installed to cover the gap between the barrel and the endcap, and the small gaps between the HPC modules not covered by other scintillators. These are referred to as the hermeticity taggers.

Time-of-flight counters (TOF)

The time-of-flight counters (TOF) is situated just outside the solenoid and mounted on the inside of the return yoke. The detector consists of a single layer of 192 counters with dimension 354 cm×20 cm×2 cm. The polar angle acceptance is from 41° to 139° with small acceptance holes near the plane at z=0 and at the support legs of the cryostat. Each counter is supplied with a Photo Multiplier Tube (PM) at each end, and the detector is sub-divided into 24 sectors of 4 counters (8 PMs).

Each particle traversing the scintillator counters generates a light pulse, which is generated to an electronic pulse by the PMs at both ends. Both the arrival time and charge is measured. Time resolution is measured on cosmics to 1.2 ns, which corresponds to a resolution in z of 20 cm. The detection efficiency for minimum ionizing particles is ∼99.9%.

Forward hodoscope (HOF)

The forward hodoscope (HOF) consists of two parts, one in each endcap, and is situated in the 45 mm gap between the end of the endcap and the second

MUF layer. Each endcap half consists of four quadrants, each containing 28 scintillator counters which gives a total of 224 counters. The counters are 1 cm thich, 20 cm wide and have lengths from 1.5 m up to more than 4 m. Each counter is supplied with one PM on the far side with respect to the beam pipe. Time resolution has been measured to 5 ns

The hermeticity taggers (TAG)

The hermeticity taggers are lead-scintillator counters installed in DELPHI to provide track detection for small areas where other detectors have cracks and/or dead regions. These regions can be structured in three groups:

The 90° polar angle region: The plane at z=0 (polar angle 90°) is a major division plane between two halves of many barrel sub-detectors. Therefore, the detector coverage in this region is poor, and the taggers are installed to enable particle detection here. There are 24 channels.

Phi cracks: Due to the modular structure of the high density projection chamber (HPC), there are some cracks between these modules that point straight to the interaction region, making it possible for particles to escape undetected. Between the HPC and the cryostat there is room for some taggers, and 36 such have been installed around the feet of DELPHI.

The 40° polar angle region: In the region between the barrel and the forward part of the detector, there are holes in the active region of a few detectors. Therefore, three concentric rings of scintillators have been installed in the gap between for forward and the barrel part of the RICH, giving a total of 46 readout channels.

3.2.5 Calorimetry

The calorimeters measure energy of particles passing through the detector. There are two types: hadron calorimeters, and electromagnetic calorimeters, both with separate modules in the barrel and forward. A special kind of electromagnetec calorimeters are the luminosity monitors, specifically designed to detect the energy of electrons at low polar angles, used to measure the luminosity of the accelerator.

Hadron calorimeter (HCAL)

The hadron calorimeter (HCAL) is the largest of the DELPHI sub-detectors, and covers almost the full solid angle, at polar angles from θ=11° to θ=169°.

The detector is installed into the return yoke of the solenoid, and consists 19032 limited streamer mode tubes installed in the 18 mm wide slots between the 50 mm thick iron plates. The detectors are wire chambers consisting of a plastic cathode forming 8 cells of 9 mm×9 mm with one anode wire of 80 μm in each.

The HCAL is divided into two parts, one barrel covering radii of 320 cm to 479 cm and direction along the beam of $|z|$ <380 cm, and two endcaps located at z=±(340 cm to 489 cm) between radii of 65 cm and 460 cm. The barrel HCAL is comprised of 24 modules with a depth of 20 detector layers, whereas the endcaps both have 12 modules with a depth of 19 detector layers. The readout is performed by pads covering fixed angular regions: 3.75° in ϕ, and 2.96° in θ for the barrel or 2.62° in θ in the endcaps. For the readout, several pads in the radial direction are read out in the same channel, which is called a tower. In the barrel, a tower is made up of 5 pads, whereas 4 or 7 pads, depending on the geometry, make up a tower in the forward. From the start of the 1994 run, a system which reads out the cathodes of the individual streamer tubes was implemented. This is a system which is independent of the tower readout, and improves the granularity in ϕ by a factor of 3, and in R by a factor of 5. The energy resolution in the barrel is found to be $\sigma(E)/E = 0.21 \oplus (1.12/\sqrt{E})$ (E being measured in GeV).

High density projection chamber (HPC)

The high density projection chamber (HPC) is the electromagnetic calorimeter for the barrel part of DELPHI. It is situated between the barrel-RICH and the supercondicting coil, at radial distances of 208 cm to 260 cm, and a distance along the beam of $|z|$ ≤254 cm. The detector uses the time-projection principle to measure 3-dimensional charge distribution with very high granularity (1° in ϕ, 4 mm in z and 9 samplings in R). The detector consists of 144 independent modules arranged in 6 circular shells of 24 modules each. Each module is a trapezoidal box with a length at small radius of 52 cm and 64 cm at large radius, a height of 465 mm, and a length of 90 cm. Each module is divided into nine radial rows, with each row being subdivided further into pads (see figure 3.10 for details). Each module has 128 pads which are read out, giving a total of 18432 channels.

Each module is filled with 41 layers of lead separated by gas gaps. Electromagnetic particles traversing the modules shower in the lead and ionize the gas. The ions travel to one end of the box, and the signal is read out by the pad readout. In the 10th sampling layer (∼4.5 radiation length, or the maximum length of showers), the gas is replaced by a scintillator counter for fast triggering purposes. An energy resolution of $\sigma_E/E = 0.043 \oplus (0.32/\sqrt{E})$

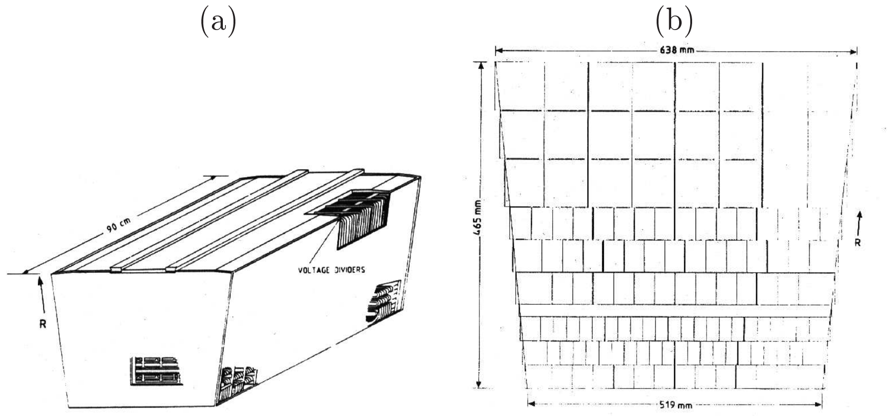


Figure 3.10: Figure showing the general layout of a HPC module. Part (a) shows a geometrical view of the entire module, whereas part (b) is a view from the endplate, showing the structure of nine rows and 128 pads. Both figures taken from [49].

(E being measured in GeV) has been measured, together with angular precisions of 1.7 mrad in ϕ and 1.0 mrad in θ.

Forward electromagnetic calorimeter (FEMC)

The forward electromagnetic calorimeter (FEMC) comprises two halves, one in each endcap, with each half being made up of 4532 Cherenkov lead glass blocks in an array of diameter ~5 m. The detector covers radial ranges from 46 cm to 240 cm, with the front faces at $|z|$=284 cm. Each glass block is a truncated pyramid with inner face dimensions 5.0 cm×5.0 cm, outer face dimensions 5.6 cm×5.6 cm, and depth 40 cm. The blocks are mounted so as to point almost to the interaction region, but tilted about 1° in order to avoid particles escaping in the dead zones between blocks. The block structure allows for good granularity, about 1° in both θ and ϕ.

The Cherenkov signal induced from a traversing particle is read out by a single stage photomultiplier, coupled to a low noise preamplifier. Calibration is done on Bhabha electrons (*i.e.* electrons form the process $e^+e^- \to e^+e^-(\gamma)$), with the exception of the region $\theta > 32°$ ($<148°$), where muons are used, due to the electron energy degradation from interactions in the TPC. Bhabhas are measured with an energy resolution of 4.8%, and the relative precision on the measured energy can be parametrized as $\sigma(E)/E = 0.03 \oplus (0.12/\sqrt{E}) \oplus (0.11/E)$ (E being measured in GeV).

3.2.6 The luminosity monitors

The luminosity at LEP is determined by measuring the number of events (preferably large, in order to minimize statistical errors) of a specific process which has a theoretically well known cross section. Such a process is the Bhabha electron scattering, with a production cross section which rises sharply at small angles, where the t-channel photon exchange diagram dominates. Therefore, the luminosity monitors at DELPHI consist of electromagnetic calorimeters at small radii, specifically optimized for detecting electrons.

Small angle tile calorimeter (STIC)

The small angle tile calorimeter [59, 60] (STIC) is the principal luminosity monitor of DELPHI. It consists of one cylinder in each endcap, placed at distances along z of 220 cm from the interaction point, and covering radii from 6.5 cm to 42.0 cm. Each cylinder is divided into two halves, giving a total of four modules. The calorimeter is a lead/scintillator sampling detector with 49 layers of 3.4 mm steel laminated lead plates and 3 mm thick scintillator tiles. Perpendicular to and through these planes run wavelength shifting fibres of 1 mm diameter with density $\sim$1 fibre/cm^2 for a total of 1600 fibres. Each module is arranged in eight azimuthal sectors of 22.5° and ten radial sectors of 3 cm, giving a total of 320 towers. Planes 8 and 15 are replaced with silicon microstrip detectors, for purposes of tracking through the detector. In addition, a tungsten mask, machined with a precision of 10 μm, is placed in front of each cylinder to provide an accurate definition of the acceptance. See figure 3.11 for details.

The energy resolution of electrons at 45 GeV is 3%, and the spatial resolution of the calorimeter alone is 1.5° in ϕ and ranges from 0.3 mm to 1 mm in R. In total, this enables the STIC to measure the luminosity at an expected systematic error of 0.2%.

Very small angle tagger (VSAT)

The VSAT is the sub-detector of DELPHI which is situated furthest from the interaction point, at a z distance of $\pm$770 cm. It is aimed at measuring Bhabha electrons at very small angles, thus achieving very large statistics, used both for fast luminosity measurements and machine operation monitoring. The sub-detector is comprised of 4 modules, two in each arm, with one arm consisting of two modules on either side of the beam pipe ($\pm$(6-8) cm in x). Each module is a rectangular W-Si calorimeter stack of dimensions 5 cm (height), 3 cm (width) and 10 cm (length), composed of 11 W-plates

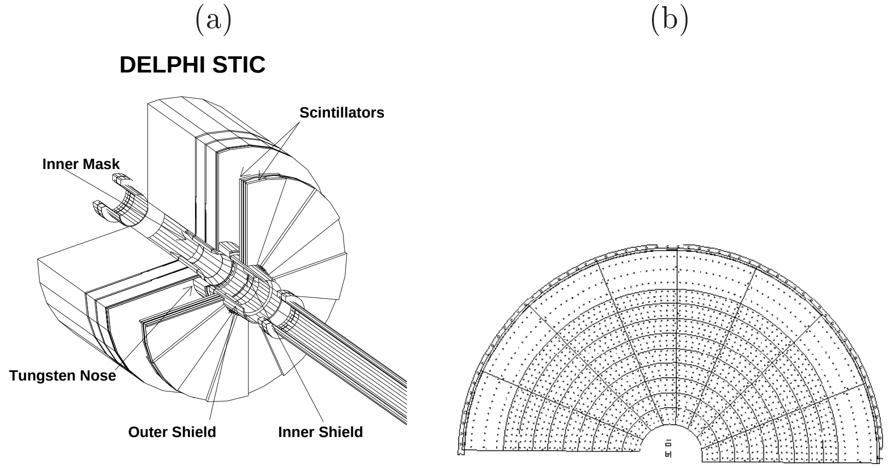


Figure 3.11: Figure showing details of the small angle tile calorimeter (STIC), the main luminosity monitor of DELPHI. Part (a) shows a general layout of one cylinder, and part (b) shows the structure of one scintillator plane of a module, with the 1600 waveshift fibres. Both figures from [50].

interleaved with silicon diodes of thichness 0.3 mm and separated by tungsten alloy absorbers. Silicon strip planes are inserted behind planes 5, 7, and 9, the ones behind layers 5 and 9 having vertical strips (32 strips of width 1 mm and height 50 mm), whereas the one behind plane 7 has horizontal strips (48 strips of height 1 mm and width 50 mm).

The resolution of the silicon strips is $\sim$170 μm over x and y coordinates. The energy resolution is 5% at 45 GeV and $\sim$4% at 95 GeV, which follows the parametrization of $\sigma(E)/E = 0.35/\sqrt{E}$ (E being measured in GeV). Expected systematic error on the relative luminosity is 1% in offline processing.

3.2.7 Trigger

The DELPHI trigger [61] is a structure of four successive layers of increasing selectivity, named T1, T2, T3, and T4. Of these, the first two are hardware triggers synchronous with the beam cross over signal (BCO), whereas the last two are software filters. T1 and T2 have been active since the beginning of DELPHI, while T3 and T4 were introduced in 1992 and 1993, respectively. The following gives a brief description of the four trigger layers:

T1: The trigger decision of T1 is taken 3.5 μs after the BCO, and works only as a loose pre-trigger. Requirements are restricted to simple patterns in track chambers, scintillator hits, or low energy single clusters in the calorimeters. The detectors which contribute are the fast tracking detectors ID, OD, FCA, and FCB, the scintillator arrays TOF and HOF, the scintillators embedded in the HPC, the FEMC, and the MUB. No correlation between signals from these sub-detectors are introduced. The T1 trigger rate for a normal luminosity ($\sim 1.5 \times 10^{31}$ cm^{-2}s^{-1}) is $\sim$700 Hz.

T2: This trigger decision is taken 39 μs after the BCO, and complements the T1 by adding information from the TPC, HPC, and MUF, which, due to the longer drift times of these detectors, was not available at T1. Combination of sub-detector signals are introduced. T2 is organized in majorities in order to maintain redundancy and efficiency for physics events. These majorities correspond to polar angle ranges: the barrel, the endcaps and the intermediate region between them. Individual sub-detectors with low counting rates produce their own triggers, while more noisy sub-detectors are grouped in majorities, where two sub-detector signals, excluding certain signal combinations, are required to trigger in coincidence for the majority trigger to fire. The T2 trigger rate for a normal luminosity is $\sim$4.5 Hz, of which around 20% are real physics events.

T3: This is a software trigger which uses the same logic as T2 and introduces no new signals from sub-detectors. But, being a software trigger, T3 can make use of calibration constants, thereby enabling tighter cuts on the pointing of tracks towards the primary vertex and shower energy thresholds, thus reducing the T2 trigger rate by a factor $\sim$2 and enriching the physics content of the events written to tape.

T4: This trigger was originally introduced in order to tag, in real time, all Z^0 decays, and in particular those with interesting topologies as predicted by models for new physics. The T3 trigger rate is reduced by a factor $\sim$2.

The main background for T1 is random noise from the sub-detectors. This is greatly reduced in T2, when sub-detector correlations are introduced, and from this trigger level onwards the main background sources are beam-gas interactions, synchrotron radiation and cosmic ray events. The trigger efficiency depends on the process in question, but is generally very high for actual physics events. The global trigger efficiency for electron and muon pairs is consistent with 1 to the level of 1×10^{-4} for polar angles between 20° to 160°, and is hardly distinguishable from 1 for hadronic events over nearly the full solid angle.

After this description of the experimental apparatus, the next chapter goes on to describe the most typical characteristics of the signals and the most important backgrounds. The difference between backgrounds and signal is stressed, but the difference between signals of different type and parameters (*i.e.* the Higgs mass and/or $\tan\beta$ value) is also pointed out. This difference is mainly described in terms of two distinct classes of variables: event shape variables and b-tag variables. The concluding section of the chapter describes the different Monte Carlo generated samples used to estimate the expected backgrounds and signals for each individual analysis.

Chapter 4

Signature of signal and background

Since the cross sections for the Higgs production typically lie far below the cross sections for the relevant SM backgrounds, as can be seen in tables 4.2–4.7 and summarized in the plots in figure 4.6, (several orders of magnitude in the region where the limits are obtained), there is a need for strong discrimination between the Higgs signal and the backgrounds. Fortunately, the Higgs signal events contain rather distinctive features which makes it possible to obtain a good separation between signal and background events. Many of these features are common to the H^0Z^0 and h^0A^0 channels, and are therefore treated as general classes of separation variables.

As was shown in section 2.4.2, the Higgs boson coupling to fermions is proportional to the fermion mass. This makes it favourable for the Higgs boson to decay to the most massive kinematically available particle/anti-particle pair, which at LEP is the $b\bar{b}$ quark pair. In the work presented here, the search channels are restricted to this decay mode for the Higgs boson(s), and in addition the Z^0 is assumed to decay to a quark/anti-quark pair (which is the largest branching ratio of the Z^0 at ∼70 %). Therefore, the signal is a four jet structure originating from four quarks, of which, in the h^0Z^0 (h^0A^0) channel, two (all four) of the quarks are b-quarks. This is briefly sketched in table 4.1.

4.1 General signal event characteristics

The discrimination between signal and background is, in the discriminating variable part of the analysis, obtained by studying two different aspects of the events: event shape and b-tag. Of these two, the information from the b-

Channel	Production process	Decay	Decay
H^0Z^0	$e^+e^- \to Z^{0*} \to H^0Z^0$	$H^0 \to b\bar{b}$	$Z^0 \to q\bar{q}$ $(q = u, d, s, c, b)$
h^0A^0	$e^+e^- \to Z^{0*} \to h^0A^0$	$h^0 \to b\bar{b}$	$A^0 \to b\bar{b}$

Table 4.1: A short schematic description of the two different search channels analysed in this work.

tag is the most important one, the high content of b-hadrons in the h^0Z^0 and particularly the h^0A^0 signal events being a very good discriminator between signal and background, whereas the event shape of the hadronic four jet structure has comparatively large counterparts in non-Higgs SM background processes.

In addition to these two characteristics, which will be combined into a single discriminating variable in the different analyses, the mass of the Higgs boson(s) will be used to discriminate between the signal and the background hypothesis. This aspect of the analysis makes use of the method of constrained fits, and is specific to the search channel in question, and to a lesser degree to the specific analysis. The method of mass reconstruction, as well as the choice of mass estimator, will be treated further in section 7.2.

4.1.1 Event shape

Since the Higgs boson signal events in both the H^0Z^0 and h^0A^0 case consist of two heavy bosons, both decaying to a quark/anti-quark pair, the event has a structure of four hadronic jets, with (ideally) no missing energy or momentum. The characteristics of such events can be summarized in a few points:

- Many charged tracks in the event.
- Large visible energy.
- Four (ideally) relatively clearly separated hadronic jets in the event.
- Topology of the distribution of tracks in the event corresponding to a more isotropic structure than would be the case for more back-to-back events.
- No high-energy photons in the event, as opposed to a photon radiated from the initial-state electron or positron (initial state radiation, ISR) or from the decay products of the process in question (final state radiation, FSR), as is the case for the majority of the $q\bar{q}(\gamma)$ events.

The specific event shape variables used in the analysis will be treated in section 7.1.1.

4.1.2 B-tag

As already mentioned, the content of b-hadrons in the decay products of the Higgs boson(s) is an important trait in the signal events. The b-tag is a very effective way of distinguishing the signal from the different backgrounds, most notably the W^+W^- background, and is therefore an important ingredient in the analyses. The aspect of the b-hadrons used for tagging purposes are:

- The B meson lifetime.
- The effective mass of the secondary B meson vertex.
- Rapidity of tracks in the secondary B meson vertex.
- Charged jet energy fraction of the secondary B meson vertex.
- High transverse momentum leptons.

The specific b-tag variables used in the analysis will be treated in section 7.1.2.

4.2 The H^0Z^0 signal

The production mode for the H^0Z^0 signal at tree-level is the s-channel e^+e^- Higgs-strahlung process $e^+e^- \to Z^{0*} \to H^0Z^0$, also known as the Bjorken process (see figure 4.1). This channel is present both in the SM and MSSM, where the Higgs boson can be either the (light) h^0 or the (heavy) H^0. Since, in the MSSM, the h^0 is the lighter, and therefore kinematically easier accessible, the analysis will be aimed at this channel. In order to remain consistent with the SM description, both the SM and the MSSM Higgs-strahlung channel will be labeled H^0Z^0.

The cross section for the SM Higgs-strahlung process is given at tree-level by the expression [62, page 361]

$$\begin{aligned} \sigma_{\mathrm{SM}}^{\mathrm{HZ}} &= \sigma(e^+e^- \to H^0Z^0) = \tfrac{G_F^2 m_Z^4}{96\,\pi s}(v_e^2 + a_e^2)\lambda^{\frac{1}{2}}\tfrac{\lambda+12z}{(1-z)^2} \qquad \text{where} \\ z &= \tfrac{m_Z^2}{s}, \qquad \lambda = \tfrac{1}{s^2}\left\{(s - m_H^2 - m_Z^2)^2 - 4m_H^2 m_Z^2\right\}, \end{aligned} \tag{4.1}$$

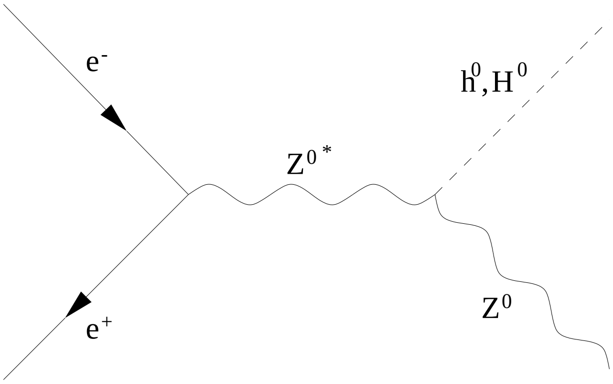


Figure 4.1: The tree-level Feynman diagram of the production channel for the H^0Z^0 final state, the s-channel e^+e^- Higgs-strahlung, also known as the Bjorken process.

$\sqrt{s}$ denotes the centre-of-mass energy, and $a_e = -1,\ v_e = -1 + 4\sin^2\theta_W$ are the Z charges of the electron. Radiative corrections to the tree-level production rate are relatively small, and the main correction to this expression is the inclusion of photon radiation [63].

In the MSSM, the production cross section for the Higgs-strahlung process is given as a simple correction to the SM expression of equation 4.1 as [62, page 400]

$$\sigma^{\mathrm{HZ}}_{\mathrm{MSSM}} = \sigma(e^+e^- \to h^0Z^0) = \sin^2(\beta-\alpha)\sigma^{\mathrm{HZ}}_{\mathrm{SM}} \tag{4.2}$$

where the parameter $\sin(\beta-\alpha)$ refers to the parameters $\tan\beta$ (equation 2.77) and α (equation 2.81) of the general two-doublet model.

Since the Higgs boson predominantly decays to a $b\bar{b}$ quark pair, the signature of the Higgs-strahlung events in the four jet channel is a four jet hadronic system with at least two b-tagged jets and the dijet invariant mass of the opposite jet pair being close to the Z^0 mass. This signature has several non-Higgs SM process backgrounds, as will be discussed in section 4.4.

4.3 The h^0A^0 signal

The production mode for the h^0A^0 signal at tree-level is the s-channel e^+e^- pair production $e^+e^- \to Z^{0*} \to h^0A^0$ (see figure 4.2).

The production cross section for the MSSM pair production process is given in terms of the SM Higgs-strahlung cross section of equation 4.1 as [62,

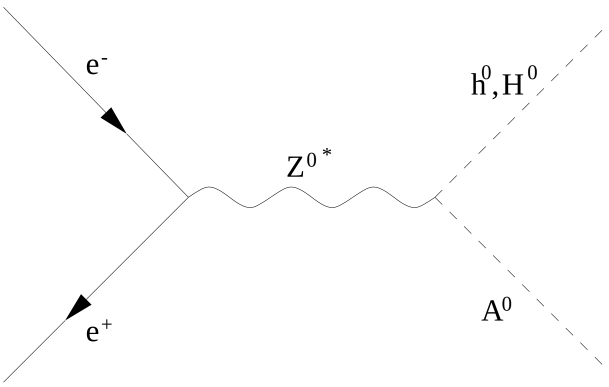


Figure 4.2: The leading-order Feynman diagram for the production channel for the h^0A^0 final state, the s-channel e^+e^- pair production channel.

page 400]

$$\begin{gathered}\sigma^{\mathrm{hA}}_{\mathrm{MSSM}} = \sigma(e^+e^- \to h^0A^0) = \cos^2(\beta-\alpha)\overline{\lambda}\sigma^{\mathrm{HZ}}_{\mathrm{SM}} \qquad \text{where} \\ \overline{\lambda} = \lambda^{\frac{3}{2}}_{hA}\left\{\lambda^{\frac{1}{2}}_{hZ}\left[12z+\lambda_{hZ}\right]\right\},\ \lambda_{ij} = \left(1-\tfrac{(m_i+m_j)^2}{s}\right)\left(1-\tfrac{(m_i-m_j)^2}{s}\right)\end{gathered} \tag{4.3}$$

In the h^0A^0 channel both the h^0 and A^0 decay predominantly to a $b\overline{b}$ quark pair, and the signature of the pair production events in the four jet channel is therefore a four jet hadronic system with all four jets being b-tagged. The masses of the two heavy objects in the event (the h^0 and A^0) are both unknown, and since both the h^0 and A^0 decay to a $b\overline{b}$ quark system, there is in the general case no kinematic information in the event which can be used to determine which of the jets originates from which of the original heavy objects. This point will be further elaborated in section 7.2.

4.4 Backgrounds

The different backgrounds important to the four jet channel are categorized in three classes: the $q\overline{q}(\gamma)$ channel, the W^+W^- channel and the Z^0Z^0 channel. Notably, the largest background channel at LEP-*II*, the t-channel gamma-exchange $e^+e^-\to e^+e^-$+hadrons through multiperipheral diagrams (also known as the two-photon or $\gamma\gamma$ channel), is missing. This is due to the requirement of a large number of charged tracks, large visible energy and maximum photon energy in the event, which cuts away very close to all $\gamma\gamma$ events at the preselection level; see section 5.2.

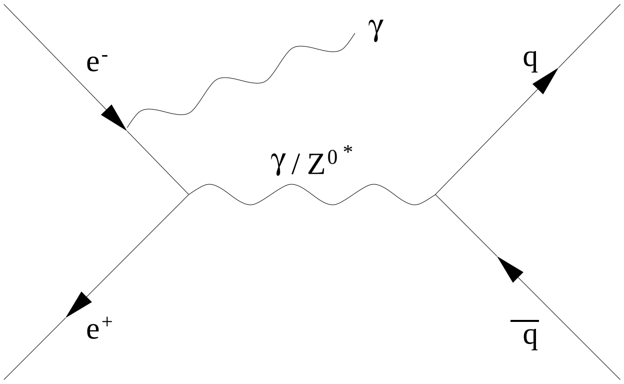


Figure 4.3: The leading-order Feynman diagram for the $q\overline{q}(\gamma)$ background.

4.4.1 The $e^+e^- \to q\overline{q}(\gamma)$ background

The tree-level production mode for the $q\overline{q}(\gamma)$ channel is the s-channel e^+e^- Z^0/γ annihilation, with the Z^0/γ decaying to a quark pair $q\overline{q}$ (see figure 4.3). In about 75% of such events, there is an ISR photon present which gives the event an effective centre-of-mass energy considerably lower than the sum of the beam energies.

This channel is not a true hadronic four jet channel, but rather a hadronic system with two jets, which, in about 75% of the cases, is accompanied by a high-energy photon. However, hadronic events generally have a less clean structure than leptonic events, which can cause a two jet event to have a jet which is recognized as two different hadronic jets (jet splitting). Also, gluon radiation decaying to quark-pairs can cause secondary hadronic jets in the event, which may be very difficult to distinguish from the quark jets originating from the heavy bosons in the event. In addition to this, the high-energy photon can, if present in the event, give secondary reactions in the detector which resembles hadronic structures, and can, together with the aforementioned effects, be recognized as hadronic jets. For this background, it is also possible for all jets in the event to originate from b-quarks, which can give such events very high b-tag values.

The cross section of the $q\overline{q}(\gamma)$ channel has been subject to much study prior to the LEP-*II* runs [62, page 210, ff], and a simple, fairly accurate, cross section is not easily given. The $q\overline{q}(\gamma)$ cross section at energies of interest for the analyses presented here, generally lies at values of approximately 100 pb. This is substantially larger than the cross section for the other two background channels, which compensates for the lower preselection efficiency of this channel, due to the requirement on the $q\overline{q}(\gamma)$ events having special

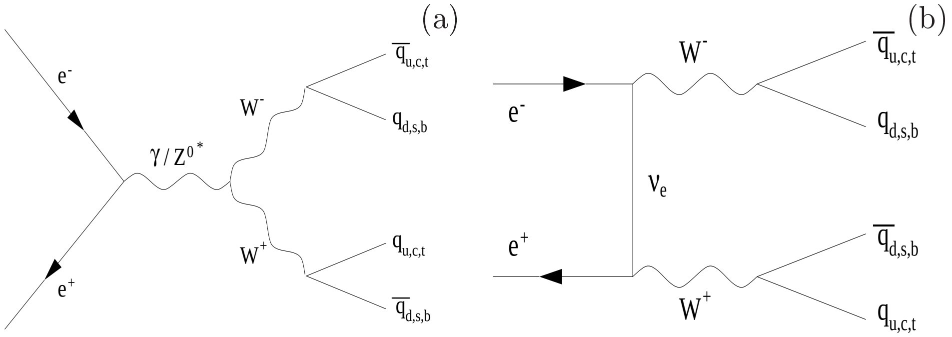


Figure 4.4: The leading-order Feynman diagrams for the W^+W^- background. Part (a) shows the s-channel annihilation diagrams, whereas part (b) shows the conversion diagram.

features in order to appear as four jet events. Therefore, the expected number of events for this background channel is still comparable to the other two.

4.4.2 The $e^+e^- \to W^+W^-$ background

The tree-level production mode for the W^+W^- channel contains three diagrams: first the two s-channel e^+e^- Z^0 or γ, non-Abelian annihilation diagrams, with the $Z^0(\gamma)$ decaying to a pair of $W^\pm$ bosons, shown in figure 4.4 (a), and last the t-channel e^+e^- ν_e conversion diagram, shown in figure 4.4 (b).

This channel is, as opposed to the $q\overline{q}(\gamma)$ channel, a true four jet hadronic background, and is, in terms of event shape and kinematics, almost indistinguishable from a Higgs signal event of comparable Higgs masses (the Z^0 mass in the H^0Z^0 signal events giving the only difference). However, the $W^\pm$ boson decays to one up-type quark and one down-type quark (one of which being an anti-quark). This, coupled with the fact that the t-quark is kinemtically unavailable, means that the $W^\pm$ boson can only decay to a b-quark and an up-type quark through a Cabbibo-suppressed decay, either the bc or the even more strongly suppressed bu quark channel. Using the latest CKM values [22, page 94], the probability that a $W^\pm$ pair decays to two b-quarks (which is the maximum number) is as low as $\sim 2 \times 10^{-3}$, as compared to the $\gtrsim$85% branching ratio (for most MSSM points of interest) for H^0 and A^0 into the $b\bar{b}$ topology. In all, this means that events from the W^+W^- channel do not have very high b-tag values, which is a feature that makes such events fairly simple to reject.

The tree-level cross section of the W^+W^- channel is given in the Born approximation as [62, page 89]

$$\sigma_{\mathrm{WW,Born}} \approx \frac{\pi\alpha^2\beta}{s\sin^4\theta_W} \qquad \text{where} \qquad \beta = \sqrt{1-\frac{4m_W^2}{s}}. \tag{4.4}$$

Radiative corrections to this expression exist, and bring the total theoretical uncertainty on the cross section down to around 0.5% [62, page 90, ff]. This cross section is generally much smaller than that of the $q\overline{q}(\gamma)$ background (a factor 4–6 for the centre-of-mass energy values of interest in the analyses presented here), but the feature of the W^+W^- being a true four jet background makes it the numerically dominant channel at the preselection level. Also, the W^+W^- cross section is rising with increasing centre-of-mass energy, as opposed to the $q\overline{q}(\gamma)$ cross section, which is falling, and this makes the relative importance of the W^+W^- background compared to the $q\overline{q}(\gamma)$ background larger as the centre-of-mass energy increases.

4.4.3 The $e^+e^- \rightarrow Z^0Z^0$ background

The tree-level production mode for the Z^0Z^0 channel consists of one diagram, the t-channel e^+e^-conversion diagram through neutral gauge bosons, Z^0 or γ (see figure 4.5). This channel is, like the W^+W^- channel, a true four jet hadronic background, with an even more kinematically indistinguishable structure with respect to the true Higgs signal events, as there are now two Z^0's in the event. (This will in the case of degenerate Higgs and Z^0 masses, give a signature exactly like the Higgs signal when disregarding the difference in the production polar angle distributions, which is only measurable on a relatively large sample, and not on single events.) But, contrary to the W^+W^- channel, the jets in a Z^0Z^0 channel event can all originate from b-quarks, making it possible for events of the Z^0Z^0 background to have very high b-tag values, and the Z^0Z^0 channel is therefore the only background which contains irreducible events with respect to the h^0A^0 signal channel.

The tree-level cross section of the Z^0Z^0 channel in the narrow-width Z^0 decay approximation is given as [62, page 234]

$$\begin{aligned} \sigma_{\mathrm{ZZ}} &= \tfrac{\alpha^2\pi}{s} C_D \left(A_D \log\left(\tfrac{\alpha_D+\beta_D}{\alpha_D-\beta_D}\right) - 3\alpha_D\beta_D \right) \qquad \text{where} \\ \alpha_D &= 1-2x_Z, \qquad \beta_D = \sqrt{1-4x_Z}, \qquad x_Z = \tfrac{m_Z^2}{s} \\ C_D &= \tfrac{38\sin^8\theta_W - 32\sin^6\theta_W + 24\sin^4\theta_W - \sin^2\theta_W + 1}{16\sin^4\theta_W\cos^4\theta_W(1-2x_Z)}, \qquad A_D = 1+4x_Z^2. \end{aligned} \tag{4.5}$$

Due to the larger mass of the Z^0 boson, and thereby the larger energy requirement in order to produce the Z^0Z^0 events, the cross section of the Z^0Z^0 channel lies well below that of the W^+W^- channel.

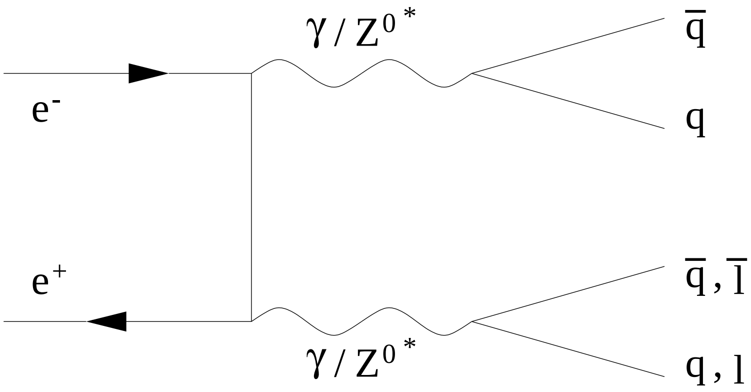


Figure 4.5: The leading-order Feynman diagram for the Z^0Z^0 background.

4.5 Monte Carlo samples

In order to estimate the various expected backgrounds and signals, a large number of Monte Carlo events have been generated by various physics processes generators, to be described later in this chapter. The generated events were then passed through the DELPHI detector simulation program DELSIM [64], which ensures the correct detector setup. The generated events have been divided into different samples, each with a corresponding cross section, according to what background or signal channel this sample is estimating. However, this separation is not the same as the one described for the background channels in the earlier sections, and therefore needs explanation.

4.5.1 Background samples

The $q\bar{q}(\gamma)$ background has been estimated using the PYTHIA [65] generator. This has been the DELPHI standard choice of generator during the LEP-*I* and LEP-*II* runs, and is well documented and tested.

For the W^+W^- and Z^0Z^0 channels, the EXCALIBUR [66] generator has been used. This is a 4-fermion generator which operates on the basis of final states, and not on Feynman diagrams, and the different samples generated from EXCALIBUR therefore do not necessarily correspond to the background channels described in the previous sections. The two different EXCALIBUR samples are:

WW-like sample (WWEX): This sample consists of all possible generic (*i.e.* $SU(2)$ weak isospin blind) final states from a W^+W^- event: $q\bar{q}q\bar{q}$ (four quarks), $q\bar{q}l\nu$ (two quarks, a lepton and a neutrino) or $l\nu l\nu$ (two leptons and two neutrinos). However, since the EXCALIBUR generator is $SU(2)$ weak isospin flavor (and generation) blind, the quarks are only guaranteed to come in quark/anti-quark pairs, in addition to preserving the total charge of the system. This means that in EXCALIBUR $q\overline{q}q\overline{q}$ events, the four quarks can all be up- or down-type quarks, whereas $q\overline{q}q\overline{q}$ events from true W^+W^- events necessarily must consist of an up-type quark and an anti down-type quark (from the W^+) in addition to an anti up-type quark and a down-type quark (from the W^-). This means that the important irreducible background of four b-quarks is contained in this sample, even though this can only originate from Z^0Z^0 events.

$q\overline{q}l\overline{l}$ sample (QQLL): This sample contains the 4-fermion processes resulting in two quarks (one of which is an anti-quark) and two leptons (one of which is an anti-lepton). Such events cannot originate from W^+W^- events, and they are therefore not part of the WW-like channel. The sample is further subdivided into the three different parts $q\overline{q}\tau\overline{\tau}$, $q\overline{q}\mu\overline{\mu}$ and $q\overline{q}e\overline{e}$.

The statistics for the generated Monte Carlo background samples is presented in table 4.2 and 4.3. Here, the tables are subdivided into ten parts, which corresponds to the average energy of the ten different centre-of-mass energy windows which the collected data have been divided into. These are:

- For the 1998 data: one centre-of-mass energy window with average energy 188.6 GeV.

- For the 1999 data: four centre-of-mass energy windows with average energies 191.6 GeV, 195.5 GeV, 199.5 GeV and 201.6 GeV.

- For the 2000 data: four centre-of-mass energy windows with average energies 203.6 GeV, 205.2 GeV, 206.7 GeV and 208.2 GeV. In addition to this, there occurred on September 6th the permanent trip of sector 6 of one half of the TPC, making it necessary to produce new Monte Carlo for the remainder of the 2000 run. The data from this period was collected in a separate centre-of-mass energy window, with average energy 206.3 GeV.

The centre-of-mass energy distribution for the preselected events can be found in figure 5.1 on page 94. Also, the cross sections listed in table 4.2 and 4.3 are summarized in the lower plot of figure 4.6.

Channel	X-section (pb)	Generated events	$\mathcal{L}_{\mathrm{MC}}$ (pb^{-1})	$\frac{\mathcal{L}_{\mathrm{MC}}}{\mathcal{L}_{\mathrm{Data}}}$
E_{CM}=188.6 GeV, $\mathcal{L}_{\mathrm{Data}}$=158.0 pb^{-1}				
$q\bar{q}(\gamma)$ (ZGPY)	99.0	1665299	$\sim$ 16800	$\sim$ 106
$q\bar{q}q\bar{q}$, $q\bar{q}l\nu$, $l\nu l\nu$ (WWEX)	17.733	686214	$\sim$ 38700	$\sim$ 245
$q\bar{q}\tau\bar{\tau}$ (QQLL)	0.1058	14996	$\sim$ 142000	$\sim$ 897
$q\bar{q}\mu\bar{\mu}$ (QQLL)	0.263	14995	$\sim$ 63500	$\sim$ 402
$q\bar{q}e\bar{e}$ (QQLL)	0.468	14694	$\sim$ 31400	$\sim$ 199
E_{CM}=191.6 GeV, $\mathcal{L}_{\mathrm{Data}}$=25.89 pb^{-1}				
$q\bar{q}(\gamma)$ (ZGPY)	96.04	151370	$\sim$ 1580	$\sim$ 61
$q\bar{q}q\bar{q}$, $q\bar{q}l\nu$, $l\nu l\nu$ (WWEX)	18.127	236920	$\sim$ 13100	$\sim$ 505
$q\bar{q}\tau\bar{\tau}$ (QQLL)	0.1095	14994	$\sim$ 137000	$\sim$ 5290
$q\bar{q}\mu\bar{\mu}$ (QQLL)	0.2614	15000	$\sim$ 57400	$\sim$ 2220
$q\bar{q}e\bar{e}$ (QQLL)	0.4247	14994	$\sim$ 35300	$\sim$ 1360
E_{CM}=195.5 GeV, $\mathcal{L}_{\mathrm{Data}}$=76.90 pb^{-1}				
$q\bar{q}(\gamma)$ (ZGPY)	90.04	464307	$\sim$ 4490	$\sim$ 58
$q\bar{q}q\bar{q}$, $q\bar{q}l\nu$, $l\nu l\nu$ (WWEX)	18.482	477146	$\sim$ 25800	$\sim$ 336
$q\bar{q}\tau\bar{\tau}$ (QQLL)	0.1125	14693	$\sim$ 131000	$\sim$ 1700
$q\bar{q}\mu\bar{\mu}$ (QQLL)	0.2578	14999	$\sim$ 58200	$\sim$ 757
$q\bar{q}e\bar{e}$ (QQLL)	0.4148	14991	$\sim$ 36100	$\sim$ 470
E_{CM}=199.5 GeV, $\mathcal{L}_{\mathrm{Data}}$=84.28 pb^{-1}				
$q\bar{q}(\gamma)$ (ZGPY)	86.11	538948	$\sim$ 6260	$\sim$ 74
$q\bar{q}q\bar{q}$, $q\bar{q}l\nu$, $l\nu l\nu$ (WWEX)	18.720	608798	$\sim$ 32500	$\sim$ 386
$q\bar{q}\tau\bar{\tau}$ (QQLL)	0.1140	14993	$\sim$ 132000	$\sim$ 1560
$q\bar{q}\mu\bar{\mu}$ (QQLL)	0.2530	14698	$\sim$ 58100	$\sim$ 689
$q\bar{q}e\bar{e}$ (QQLL)	0.4112	14991	$\sim$ 36500	$\sim$ 433
E_{CM}=201.6 GeV, $\mathcal{L}_{\mathrm{Data}}$=41.11 pb^{-1}				
$q\bar{q}(\gamma)$ (ZGPY)	83.27	505875	$\sim$ 6080	$\sim$ 148
$q\bar{q}q\bar{q}$, $q\bar{q}l\nu$, $l\nu l\nu$ (WWEX)	18.792	237144	$\sim$ 12700	$\sim$ 307
$q\bar{q}\tau\bar{\tau}$ (QQLL)	0.1142	14992	$\sim$ 131000	$\sim$ 3190
$q\bar{q}\mu\bar{\mu}$ (QQLL)	0.2504	15000	$\sim$ 59900	$\sim$ 1460
$q\bar{q}e\bar{e}$ (QQLL)	0.4076	14992	$\sim$ 36800	$\sim$ 895

Table 4.2: The Monte Carlo generated production cross section, number of generated events and equivalent luminosity for the different backgrounds. The statistics shown corresponds to the 1998 and 1999 samples.

Channel	X-section (pb)	Generated events	$\mathcal{L}_{\text{MC}}$ (pb^{-1})	$\frac{\mathcal{L}_{\text{MC}}}{\mathcal{L}_{\text{Data}}}$
E_{CM}=203.6 GeV, $\mathcal{L}_{\text{Data}}$=8.77 pb^{-1}				
$q\overline{q}(\gamma)$ (ZGPY)	82.23	989322	~ 12000	~ 1370
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	18.857	165537	~ 8780	~ 1000
$q\overline{q}\tau\overline{\tau}$ (QQLL)	0.1142	14993	~ 131000	~ 15000
$q\overline{q}\mu\overline{\mu}$ (QQLL)	0.2482	14998	~ 60400	~ 6890
$q\overline{q}e\overline{e}$ (QQLL)	0.4045	14388	~ 35600	~ 4060
E_{CM}=205.2 GeV, $\mathcal{L}_{\text{Data}}$=63.18 pb^{-1}				
$q\overline{q}(\gamma)$ (ZGPY)	80.77	2806314	~ 34700	~ 550
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	18.897	652490	~ 34500	~ 547
$q\overline{q}\tau\overline{\tau}$ (QQLL)	0.1139	30115	~ 264000	~ 4180
$q\overline{q}\mu\overline{\mu}$ (QQLL)	0.2460	29815	~ 120000	~ 1900
$q\overline{q}e\overline{e}$ (QQLL)	0.4013	30099	~ 75000	~ 1190
E_{CM}=206.7 GeV, $\mathcal{L}_{\text{Data}}$=77.66 pb^{-1}				
$q\overline{q}(\gamma)$ (ZGPY)	79.32	2669386	~ 33700	~ 433
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	18.935	638675	~ 33700	~ 434
$q\overline{q}\tau\overline{\tau}$ (QQLL)	0.1134	79468	~ 700000	~ 9020
$q\overline{q}\mu\overline{\mu}$ (QQLL)	0.2435	96285	~ 39500	~ 5090
$q\overline{q}e\overline{e}$ (QQLL)	0.3983	92942	~ 233000	~ 3000
E_{CM}=208.2 GeV, $\mathcal{L}_{\text{Data}}$=7.08 pb^{-1}				
$q\overline{q}(\gamma)$ (ZGPY)	77.94	1820584	~ 23400	~ 3300
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	18.972	367524	~ 19400	~ 2740
$q\overline{q}\tau\overline{\tau}$ (QQLL)	0.1130	30297	~ 268000	~ 37900
$q\overline{q}\mu\overline{\mu}$ (QQLL)	0.2413	30305	~ 126000	~ 17700
$q\overline{q}e\overline{e}$ (QQLL)	0.3954	52471	~ 133000	~ 18700
E_{CM}=206.3 GeV, $\mathcal{L}_{\text{Data}}$=59.88 pb^{-1}				
$q\overline{q}(\gamma)$ (ZGPY)	79.73	995549	~ 12500	~ 209
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	18.924	283691	~ 15000	~ 250
$q\overline{q}\tau\overline{\tau}$ (QQLL)	0.1135	15291	~ 135000	~ 2250
$q\overline{q}\mu\overline{\mu}$ (QQLL)	0.2441	16797	~ 68800	~ 1150
$q\overline{q}e\overline{e}$ (QQLL)	0.3991	16790	~ 42100	~ 703

Table 4.3: The Monte Carlo generated production cross section, number of generated events and equivalent luminosity for the different backgrounds. The statistics shown corresponds to the 2000 data sample.

4.5.2 Signal samples

The signal samples were all created using the HZHA [67] generator. The samples used in the analyses presented here are generated in two different classes:

The HZQQ samples: These samples are h^0Z^0 signal events generated in the four jet channel, *i.e.* the h^0 decaying to a $b\bar{b}$ quark/anti-quark pair, and the Z^0 decaying to a $q\bar{q}$ quark/anti-quark pair. Several different samples have been generated, corresponding to different values of the Higgs mass m_{H^0}, ranging from 70 GeV/c^2 to 120 GeV/c^2.

The HZHA samples: These samples are h^0A^0 signal events generated in the four jet channel, *i.e.* both the h^0 and the A^0 decaying to a $b\bar{b}$ quark/anti-quark pair. The samples have been generated at several different values of the MSSM parameter $\tan\beta$, $\tan\beta$=2,20,50, all with several different values of the Higgs mass m_{A^0}, ranging from 70 GeV/c^2 to 100 GeV/c^2 in steps of 5 GeV/c^2.

The statistics for the generated Monte Carlo signal samples are presented in table 4.4 and 4.5 (for the h^0Z^0 signal) and table 4.6 and 4.7 (for the h^0A^0 signal). The cross sections for a few different signal hypothesis is shown on the upper left of figure 4.6 for the H^0Z^0 signal, and on the upper right for the h^0A^0 signal.

Each generated event sample was generated at a specific centre-of-mass energy. This, however, does not represent the actual experimental situation, particularly for the year 2000, very well, as data was taken over a range of centre-of-mass energies (albeit at rather narrow and peaked ranges for the 1998 and 1999 data). Therefore, each generated event was reboosted to a new centre-of-mass energy according to the distribution in data for the centre-of-mass range in question. This was done according to the following procedure: When boosting an event from the old centre-of-mass energy E_{old} to the new value E_{new}, events were divided into two classes:

- Events with two heavy bosons in the original state (*i.e.* the W^+W^-, Z^0Z^0, H^0Z^0 and h^0A^0 channels) were treated in the following way: For each particle in the event, an attempt at finding a link to the original simulated heavy boson through the use of the various simulation banks was made. With all the particles where such a link was found, a re-boosting routine [68] was applied, which reboosts the original heavy bosons in the event to a new centre-of-mass energy and propagates the results to the final state particles. The remaining particles in the event, where the search for a simulation link to the original heavy bosons was

m_{H^0} (GeV/c^2)	X-section (pb)	Generated events	m_{H^0} (GeV/c^2)	X-section (pb)	Generated events
E_{CM}=188.6 GeV					
70.0	0.8001	2000	75.0	0.6961	1995
80.0	0.5862	1799	85.0	0.4667	4996
90.0	0.3319	4999	92.5	0.2545	2996
95.0	0.1652	25682	97.5	0.0652	3000
100.0	0.0251	4996			
E_{CM}=191.6 GeV					
85.0	0.4928	2049	90.0	0.3788	2050
95.0	0.2469	2050	100.0	0.0758	4917
105.0	0.0145	2048			
E_{CM}=195.5 GeV					
85.0	0.5123	2049	90.0	0.4177	2050
95.0	0.3131	2050	100.0	0.1900	4917
105.0	0.0396	2048			
E_{CM}=199.5 GeV					
85.0	0.5207	2049	90.0	0.4405	2050
95.0	0.3543	2048	100.0	0.2586	2047
105.0	0.1436	5122	110.0	0.0237	2048
115.0	0.0079	1022			
E_{CM}=201.6 GeV					
85.0	0.5217	1999	90.0	0.4476	1998
95.0	0.3684	1999	100.0	0.2829	2000
105.0	0.1840	1816	110.0	0.0564	2021
115.0	0.0107	1997			

Table 4.4: The Monte Carlo generated production cross section and number of generated events for the different H^0Z^0 signals. The cross sections refer to the SM cross sections, whereas the MSSM numbers are found by multiplying with the correction factor $\sin^2(\beta-\alpha)$. The statistics shown corresponds to the 1998 and 1999 samples.

m_{H^0} (GeV/c^2)	X-section (pb)	Generated events	m_{H^0} (GeV/c^2)	X-section (pb)	Generated events
E_{CM}=203.6 GeV					
90.0	0.4521	2012	95.0	0.3794	2050
100.0	0.3011	2050	105.0	0.2135	1930
110.0	0.1059	2050	115.0	0.0160	2049
120.0	0.0061	1998			
E_{CM}=205.2 GeV					
85.0	0.5194	2000	90.0	0.4543	1999
95.0	0.3859	2000	100.0	0.3128	1999
105.0	0.2324	2000	108.0	0.1780	4998
110.0	0.1375	6993	112.0	0.0915	4996
114.0	0.0419	4999	115.0	0.0262	6998
120.0	0.0075	1998			
E_{CM}=206.7 GeV					
90.0	0.4554	2000	95.0	0.3907	1999
100.0	0.3215	2000	105.0	0.2472	2000
108.0	0.1976	9999	110.0	0.1615	11995
112.0	0.1215	9794	114.0	0.0757	9195
115.0	0.0512	11993	120.0	0.0095	1999
E_{CM}=208.2 GeV					
90.0	0.4557	3998	95.0	0.3943	3996
100.0	0.3295	3997	105.0	0.2597	3998
110.0	0.1812	3998	115.0	0.0842	3998
120.0	0.0127	3999			
E_{CM}=206.3 GeV					
90.0	0.4552	2000	95.0	0.3895	1999
100.0	0.3198	2000	105.0	0.2435	2000
108.0	0.1927	9999	110.0	0.1555	11995
112.0	0.1141	9794	114.0	0.0665	9195
115.0	0.0427	11993	120.0	0.0089	1999

Table 4.5: The Monte Carlo generated production cross section and number of generated events for the different H^0Z^0 signals. The cross sections refer to the SM cross sections, whereas the MSSM numbers are found by multiplying with the correction factor $\sin^2(\beta - \alpha)$. The statistics shown corresponds to the 2000 samples.

m_{A^0} (GeV/c^2)	Generated events	m_{A^0} (GeV/c^2)	Generated events	m_{A^0} (GeV/c^2)	Generated events
E_{CM}=188.6 GeV					
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
70.0	1199	70.0	2000	70.0	1999
75.0	5191	75.0	4997	75.0	1999
80.0	4994	80.0	4997	80.0	1998
85.0	5197	85.0	4797	85.0	3998
90.0	2001	90.0	2999	90.0	3999
E_{CM}=191.6 GeV					
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
80.0	1999	80.0	2000	—	—
85.0	4398	85.0	1821	—	—
90.0	1998	90.0	2799	—	—
95.0	1999	95.0	2000	—	—
100.0	2095	—	—	—	—
E_{CM}=195.5 GeV					
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
80.0	1999	80.0	2000	—	—
85.0	4398	85.0	1821	—	—
90.0	1998	90.0	2799	—	—
95.0	1999	95.0	2000	—	—
100.0	2095	—	—	—	—
E_{CM}=199.5 GeV					
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
80.0	1998	80.0	1997	80.0	1997
85.0	1998	85.0	1999	85.0	2000
90.0	2000	90.0	1999	90.0	2000
95.0	2000	95.0	1999	95.0	2000
E_{CM}=201.6 GeV					
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
80.0	1972	80.0	1800	—	—
85.0	1998	85.0	1999	—	—
90.0	1999	90.0	2000	—	—
95.0	2000	95.0	2000	—	—

Table 4.6: The number of events for the different Monte Carlo generated h^0A^0 signals. The statistics shown corresponds to the 1998 and 1999 samples.

m_{A^0} (GeV/c^2)	Generated events	m_{A^0} (GeV/c^2)	Generated events	m_{A^0} (GeV/c^2)	Generated events
		E_{CM}=203.6 GeV			
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
80.0	1999	80.0	1998	—	—
85.0	2000	85.0	1999	—	—
90.0	1981	90.0	1998	—	—
95.0	1999	95.0	1999	—	—
		E_{CM}=205.2 GeV			
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
80.0	1999	80.0	1999	80.0	1996
85.0	2000	85.0	1999	85.0	2000
90.0	1999	90.0	1999	90.0	1999
95.0	2000	95.0	1999	95.0	2000
		E_{CM}=206.7 GeV			
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
80.0	3997	80.0	3995	80.0	1999
85.0	3995	85.0	4000	85.0	1999
90.0	3999	90.0	3999	90.0	1999
95.0	3995	95.0	3998	95.0	1999
		E_{CM}=208.2 GeV			
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
80.0	1999	80.0	1999	80.0	1998
85.0	1996	85.0	1997	85.0	2000
90.0	1999	90.0	1999	90.0	1957
95.0	1999	95.0	1999	95.0	1995
		E_{CM}=206.3 GeV			
$\tan\beta = 2$		$\tan\beta = 20$		$\tan\beta = 50$	
80.0	5998	80.0	5969	80.0	1998
85.0	1997	85.0	1998	85.0	3995
90.0	1999	90.0	1998	90.0	2000
95.0	1997	95.0	2000	95.0	2800
100.0	1999	100.0	1999	100.0	1996

Table 4.7: The number of events for the different Monte Carlo generated h^0A^0 signals. The statistics shown corresponds to the 2000 sample.

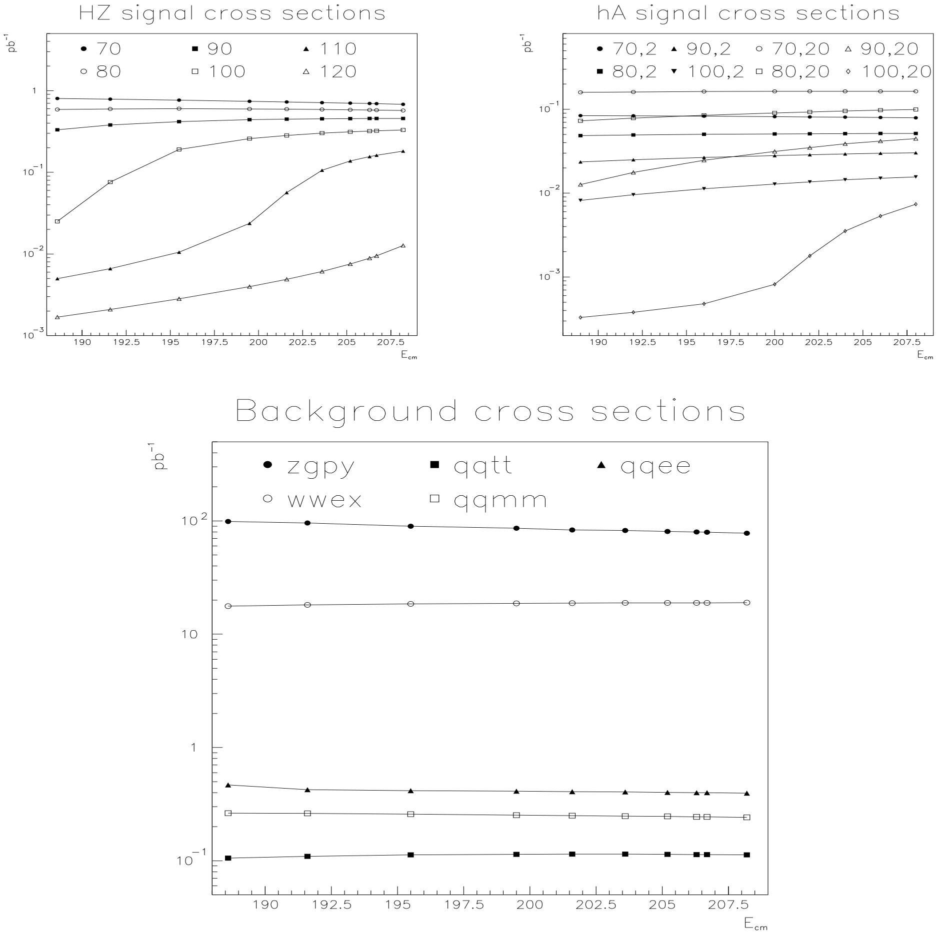


Figure 4.6: Cross sections of the relevant backgrounds and a few of the signals, as a function of the centre-of-mass energy. On the top left is shown the cross sections for the H^0Z^0 Higgs-strahlung process for signals of different Higgs masses, the numbers in the legends being the Higgs mass in GeV/c^2. The h^0A^0 pair production cross section is shown on the top right for different MSSM signals, all in the no mixing hypothesis (see section 8.1). The legends indicate the mass of the A^0 boson in GeV/c^2 followed by the value of the parameter $\tan\beta$. In the lower plot is shown the cross section for different relevant backgrounds described in the text.

unsuccessful, the particle energies and momenta were rescaled with the factor $\frac{E_{\text{new}}}{E_{\text{old}}}$.

- For the remaining channels (*i.e.* the $q\overline{q}(\gamma)$ channel), all particle energies and momenta were rescaled with the factor $\frac{E_{\text{new}}}{E_{\text{old}}}$.

After this short presentation of the most important backgrounds and signals, the next chapter goes more into detail. The cuts of the four jet preselection are presented, and their effect on data and Monte Carlo generated backgrounds and signals is shown. Two different track selections are presented and compared, and a few properties from the four jet selected events using both track selections are shown. These properties are then used to distinguish between the two track selections, and a choice is made as to what is to be the track selection used for the different analyses presented later in the work.

Chapter 5

Selection of events

Due to the special topology of the Higgs signal events, both in the MSSM and the SM scenarios, a rather tight preselection of events, aimed at reducing the expected non-Higgs SM background by a large factor, could be employed without loosing large fractions of the signal. Also, the signal event feature of a large number of tracks and, in principle, no missing energy, makes a rather loose track selection possible. This is useful in order to include as many of the original tracks from the event as possible, even though the number of false tracks introduced by such a selection will be larger than for a tight selection. This, however, will not be as critical as for search topologies with few tracks, since the four jet topology is a purely hadronic system with each hadronic jet being built up of several tracks.

5.1 Track selection

Early in the history of DELPHI, the different analysis teams each employed their own set of track selection criteria, according to the needs of the individual analyses. However, as analysis work progressed and the understanding of the detector increased, in conjunction with several upgrades of the detector itself, a desire to have a standardized track selection grew within the experiment. This was achieved in 1998 [69], and the new standard track selection has gradually taken over for the old, individualized track selections of the different analysis teams. This has been achieved in parallell with the emergence and widespread use of the standard analysis tool SKELANA [70]. In the following sections the two sets of track selection cuts are presented, and some central variables are shown.

5.1.1 Old hadronic search team track selection

Because of the large number of tracks, and the fact that there is nominally no missing energy in the event, the purely hadronic four jet analyses have traditionally relied upon a rather loose track selection. This track selection has been used by all the purely hadronic topology searches in DELPHI, and consists of the following set of cuts: [71]

- Charged particles
  - Minimum track momentum 100 MeV/c
  - Maximum impact parameter in the transverse direction: 4 cm
  - Maximum impact parameter in the z-direction: 10 cm

  Charged particles with energy above the beam energy, are rescaled to half the beam energy.

- Neutral particles
  - For the Electromagnetic clusters: Minimum energy 200 MeV
  - For the Hadronic clusters: Minimum energy 500 MeV

5.1.2 New standard track selection

The new, standard track selection [69], intended to be used by all of the DELPHI analyses, is made more flexible by the introduction of several adjustable parameters, in order to accomodate to the specific needs of different analyses. This track selection consists of the following cuts: [72]

- Charged particles
  - Minimum track momentum 100 MeV, maximum $1.5 \times E_{beam}$
  - Maximum fractional track momentum error $\Delta p/p$: 1
  - Maximum impact parameter in the transverse direction: 4 cm
  - Maximum impact parameter in the z-direction: $\sin(\theta)\times 4$ cm, θ being the polar angle of the track.
- Neutral particles
  - Electromagnetic clusters:
    - HPC: Minimum energy 300 MeV
    - FEMC: Minimum energy 400 MeV

    - ∗ STIC: Minimum energy 300 MeV
      In addition, neutrals of over 2 GeV energy produced from only one STIC tower, are rejected. Also, off-momentum electrons are rejected by rejecting STIC showers below 3 degrees in polar angle.
  - Hadronic clusters:
    Here, no minimum energy cut is applied. Instead, a noise reduction routine is used, aimed at giving a better data/simulation agreement at the expense of removing some genuine low energy showers.

Also, all ID-VD tracks (*i.e.* tracks seen only by the Inner Detector and the Vertex Detector) without z information, and all VD only tracks (*i.e.* tracks seen only by the Vertex Detector) without z information, are rejected.

In addition to these cuts, a method to recover charged tracks rejected by the original track rejection has been implemented. Due to the aforementioned properties of the purely hadronic events, the options chosen with this track selection is aimed at rejecting as few of the tracks as possible. Therefore, the recovery procedures are set to recover as many of the rejected tracks as possible. This procecdure consists of three steps:

MAMMOTH Recovery: [1] For tracks passing the impact parameter cuts, but rejected by another cut in the charged track selection, a recovery routine aimed at correcting for kinks and hadronic interactions is applied.

Unphysical high momentum particles: Charged particles which are rejected by the maximum momentum cut is re-fitted using a primary vertex constraint. If the track parameters and the re-fit probability are acceptable, the track is accepted with the new re-fitted track parameters.

Neutral Energy Recovery: If a rejected charged track has calorimeter energy of at least 5 GeV associated to it, the track is accepted as a neutral, which again is subject to the neutral cluster selection routine.

Comparisons of track and event properties of events with these two different track selections are presented in section 5.3.

---

[1] The term MAMMOTH is in DELPHI used about the program which attempts to improve the event reconstruction and aspects of the current tracking through the use of tracks that have been discarded due to specific conditions, such as certain detector combination or certain behaviour in the detector.

5.2 Preselection of four jet events

When performing the preselection, the event is forced into a four jet configuration by the DURHAM clustering algorithm [65, page 277]. A hadronic identification algorithm [73] is applied to charged tracks, which, when successful, assigns masses to the tracks, whereas unidentified charged tracks are assigned the pion mass (135 GeV/c^2). Each neutral cluster is assigned a mass of zero. The jet mass is then calculated by the relations

$$\begin{gathered} M_{\text{jet}} = \tfrac{1}{c^2}\sqrt{E_{\text{jet}}^2 - c^2\vec{P}_{\text{jet}}^2}, \quad \text{where} \\ \vec{P}_{\text{jet}} = \textstyle\sum_{i=1}^{n_{\text{tracks}}} \vec{P}_i, \quad E_{\text{jet}} = \sum_{i=1}^{n_{\text{tracks}}} E_i = \sum_{i=1}^{n_{\text{tracks}}} c\sqrt{\vec{P}_i^2 + M_i^2 c^2} \end{gathered} \tag{5.1}$$

where n_{tracks} is the number of tracks, both charged and neutral, belonging to the jet, $\vec{P}_i$, E_i and M_i are the momentum, energy and mass of track i, and $\vec{P}_{\text{jet}}$, $E_{\text{jet}}$and M_{jet} are the corresponding quantities for the jet.

The four jet event selection consists of the following cuts: [74]

- Minimum 18 charged tracks
- Visible energy larger than 60% of $\sqrt{s}$
- Neutral energy less than 50% of $\sqrt{s}$
- No neutrals with electromagnetic energy above 30 GeV
- The energy of an invisible photon as calculated by SPRIME+ [75] less than 30 GeV
- No electromagnetic calorimeter shower energy above 30 GeV
- At least 1 charged particle per jet, and all jet masses at least 1.5 GeV/c^2
- The Fox-Wolfram moments (see page 129) $H_2 + H_4$, normalized to H_0, less than 1.1
- Event thrust (see page 129) less than 0.92

After this preselection, the only remaining backgrounds are the ones treated in section 4.4; specifically, the $\gamma\gamma$ background is completely removed (see figure 5.2). Each of the variables in the four jet selection is presented for the remaining expected backgrounds together with the data in figures 5.3 to 5.8. The plots are shown for the new standard track selection, with the signal being a h^0A^0 signal of m_{A^0}=85 GeV/c^2, $\tan\beta$=20. The corresponding four jet selection statistics for the different Monte Carlo generated backgrounds, as well as for the data, are shown in tables 5.1 to 5.3. The systematic errors in the four jet selection have been extensively studied by DELPHI, and the errors included in the numbers found in the tables, contain contribution from the following sources:

- A contribution from the purely statistical uncertainty in the selection efficiency for the different Monte Carlo samples, *i.e.*
  $\sigma_{\text{eff}} = \sqrt{\epsilon(1-\epsilon)/N}$, where
  ϵ is the selection efficiency, and
  N is the size of the Monte Carlo sample.
  This contribution is typically relatively small compared to the other contributions mentioned below, the exception being for the different signal samples and some of the QQLL samples, which are generally of smaller size (see tables 4.2 to 4.7).

- A common 1% contribution from uncertainty in the luminosity and cross sections [76].

- A common 4% systematic contribution estimated from differences between different Monte Carlo generators and differences between data and Monte Carlo [74].

Figure 5.1 shows the centre-of-mass energy distribution of the four jet selected events for all three years of data taking.

At this selection level, the signal efficiencies lie for the h^0A^0 channel mostly above 90%, with no Monte Carlo generated signal sample having an efficiency below 86%. The efficiencies for the H^0Z^0 samples lie slightly below this level, but more than 84% of the signal is still retained for all Monte Carlo generated signal samples. In order to compare different signal hypotheses to each other, plots corresponding to the ones shown in figures 5.3 and 5.4 are shown for four different signal hypothesis in figures 5.9 and 5.10.

The numbers in tables 5.1 to 5.3 show generally good agreement between the expected and the observed number of events. The only two centre-of-mass energy windows for which the expectation is more than one standard deviation away from the observation are both in the 2000 data; the 203.6 GeV point, where the expectation lies below the data, and the 206.3 GeV point, where the expectation lies above the data. In terms of total events in the 2000 data sample, these two effects counteract each other, indicating statistical fluctuations rather than a systematic effect.

The distributions in figures 5.3 to 5.8 generally show a fairly good agreement between data and expected background. Although som distributions show features looking somewhat unnatural (the neutral energy and Fox-Wolfram moments of the 1998 data, in figures 5.3 and 5.4, and the total

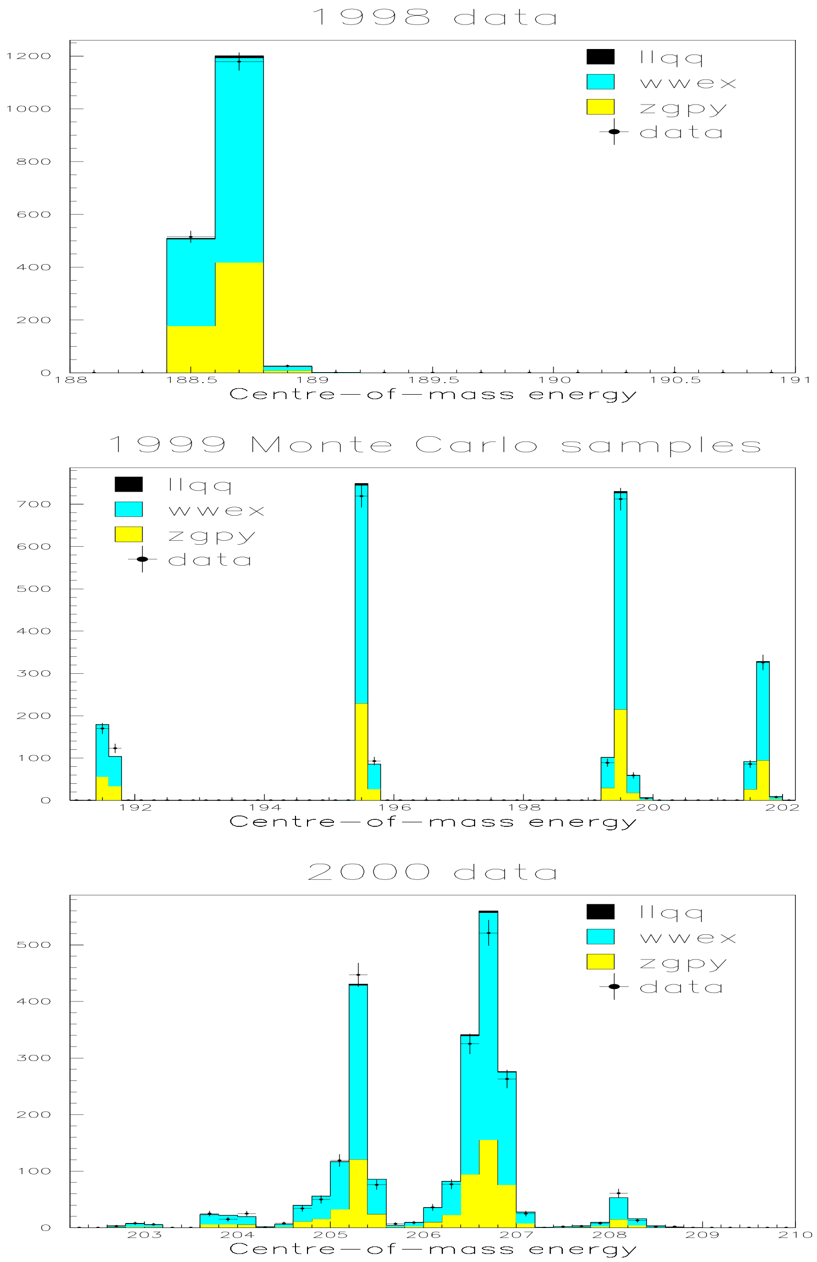


Figure 5.1: Centre-of-mass energy distributions for the four jet selected events for all collected data as well as excpected Monte Carlo background. On top is shown the 1998 data set, in the middle is the 1999 data set, and the 2000 data set is shown on the bottom.

Channel	Four jet eff.(%)	Expected events
E_{CM}=188.6 GeV		
$q\overline{q}(\gamma)$ (ZGPY)	3.87 ± 0.16	605 ± 25
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	39.9 ± 1.6	1119 ± 46
$q\overline{q}\tau\overline{\tau}$ (QQLL)	20.71 ± 0.89	3.46 ± 0.15
$q\overline{q}\mu\overline{\mu}$ (QQLL)	12.61 ± 0.57	5.24 ± 0.24
$q\overline{q}e\overline{e}$ (QQLL)	2.28 ± 0.15	1.69 ± 0.12
Total expected background	—	1734 ± 53
Data	—	1721
E_{CM}=191.6 GeV		
$q\overline{q}(\gamma)$ (ZGPY)	3.63 ± 0.15	90.2 ± 3.9
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	40.9 ± 1.6	192.0 ± 7.9
$q\overline{q}\tau\overline{\tau}$ (QQLL)	21.48 ± 0.92	0.609 ± 0.027
$q\overline{q}\mu\overline{\mu}$ (QQLL)	12.79 ± 0.58	0.865 ± 0.040
$q\overline{q}e\overline{e}$ (QQLL)	1.79 ± 0.13	0.197 ± 0.014
Total expected background	—	283.9 ± 8.8
Data	—	293
E_{CM}=195.5 GeV		
$q\overline{q}(\gamma)$ (ZGPY)	3.70 ± 0.15	256 ± 11
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	40.3 ± 1.6	573 ± 24
$q\overline{q}\tau\overline{\tau}$ (QQLL)	19.77 ± 0.86	1.710 ± 0.076
$q\overline{q}\mu\overline{\mu}$ (QQLL)	12.2 ± 0.56	2.42 ± 0.11
$q\overline{q}e\overline{e}$ (QQLL)	1.65 ± 0.12	0.528 ± 0.040
Total expected background	—	834 ± 26
Data	—	812
E_{CM}=199.5 GeV		
$q\overline{q}(\gamma)$ (ZGPY)	3.65 ± 0.15	265 ± 11
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	39.8 ± 1.6	628 ± 26
$q\overline{q}\tau\overline{\tau}$ (QQLL)	19.29 ± 0.84	1.853 ± 0.082
$q\overline{q}\mu\overline{\mu}$ (QQLL)	11.85 ± 0.54	2.53 ± 0.12
$q\overline{q}e\overline{e}$ (QQLL)	1.37 ± 0.11	0.474 ± 0.038
Total expected background	—	898 ± 28
Data	—	865

Table 5.1: The four jet selection efficiency and number of events for expected background and data. The errors contain contributions from the sources mentioned on page 93.

Channel	Four jet eff.(%)	Expected events
E_{CM}=201.6 GeV		
$q\bar{q}(\gamma)$ (ZGPY)	3.60 ± 0.15	123.3 ± 5.2
$q\bar{q}q\bar{q}$, $q\bar{q}l\nu$, $l\nu l\nu$ (WWEX)	39.5 ± 1.6	305 ± 13
$q\bar{q}\tau\bar{\tau}$ (QQLL)	18.90 ± 0.82	0.887 ± 0.040
$q\bar{q}\mu\bar{\mu}$ (QQLL)	11.71 ± 0.54	1.206 ± 0.057
$q\bar{q}e\bar{e}$ (QQLL)	1.52 ± 0.12	0.255 ± 0.020
Total expected background	—	431 ± 14
Data	—	420
E_{CM}=203.6 GeV		
$q\bar{q}(\gamma)$ (ZGPY)	3.55 ± 0.14	25.6 ± 1.1
$q\bar{q}q\bar{q}$, $q\bar{q}l\nu$, $l\nu l\nu$ (WWEX)	39.3 ± 1.6	64.9 ± 2.7
$q\bar{q}\tau\bar{\tau}$ (QQLL)	18.12 ± 0.79	0.181 ± 0.008
$q\bar{q}\mu\bar{\mu}$ (QQLL)	11.68 ± 0.54	0.254 ± 0.012
$q\bar{q}e\bar{e}$ (QQLL)	1.54 ± 0.12	0.054 ± 0.004
Total expected background	—	91.0 ± 2.9
Data	—	111
E_{CM}=205.2 GeV		
$q\bar{q}(\gamma)$ (ZGPY)	3.58 ± 0.14	182.7 ± 7.6
$q\bar{q}q\bar{q}$, $q\bar{q}l\nu$, $l\nu l\nu$ (WWEX)	39.0 ± 1.6	466 ± 19
$q\bar{q}\tau\bar{\tau}$ (QQLL)	19.54 ± 0.81	1.406 ± 0.060
$q\bar{q}\mu\bar{\mu}$ (QQLL)	11.63 ± 0.50	1.823 ± 0.081
$q\bar{q}e\bar{e}$ (QQLL)	1.452 ± 0.090	0.368 ± 0.023
Total expected background	—	652 ± 21
Data	—	656

Table 5.2: The four jet selection efficiency and number of events for expected background and data. The errors contain contributions from the sources mentioned on page 93.

Channel	Four jet eff.(%)	Expected events
E_{CM}=206.7 GeV		
$q\overline{q}(\gamma)$ (ZGPY)	3.54 ± 0.14	218.2 ± 9.0
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	38.8 ± 1.6	570 ± 24
$q\overline{q}\tau\overline{\tau}$ (QQLL)	18.65 ± 0.76	1.643 ± 0.069
$q\overline{q}\mu\overline{\mu}$ (QQLL)	11.78 ± 0.48	2.227 ± 0.094
$q\overline{q}e\overline{e}$ (QQLL)	1.347 ± 0.066	0.417 ± 0.021
Total expected background	—	792 ± 25
Data	—	789
E_{CM}=208.2 GeV		
$q\overline{q}(\gamma)$ (ZGPY)	3.53 ± 0.14	19.49 ± 0.81
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	38.5 ± 1.5	51.7 ± 2.1
$q\overline{q}\tau\overline{\tau}$ (QQLL)	18.69 ± 0.78	0.149 ± 0.006
$q\overline{q}\mu\overline{\mu}$ (QQLL)	11.86 ± 0.51	0.203 ± 0.009
$q\overline{q}e\overline{e}$ (QQLL)	1.277 ± 0.071	0.036 ± 0.002
Total expected background	—	71.5 ± 2.3
Data	—	71
E_{CM}=206.3 GeV		
$q\overline{q}(\gamma)$ (ZGPY)	3.55 ± 0.14	169.4 ± 7.0
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)	38.4 ± 1.5	435 ± 18
$q\overline{q}\tau\overline{\tau}$ (QQLL)	18.08 ± 0.79	1.229 ± 0.055
$q\overline{q}\mu\overline{\mu}$ (QQLL)	11.94 ± 0.54	1.745 ± 0.081
$q\overline{q}e\overline{e}$ (QQLL)	1.37 ± 0.11	0.327 ± 0.025
Total expected background	—	607 ± 19
Data	—	555

Table 5.3: The four jet selection efficiency and number of events for expected background and data. The errors contain contributions from the sources mentioned on page 93.

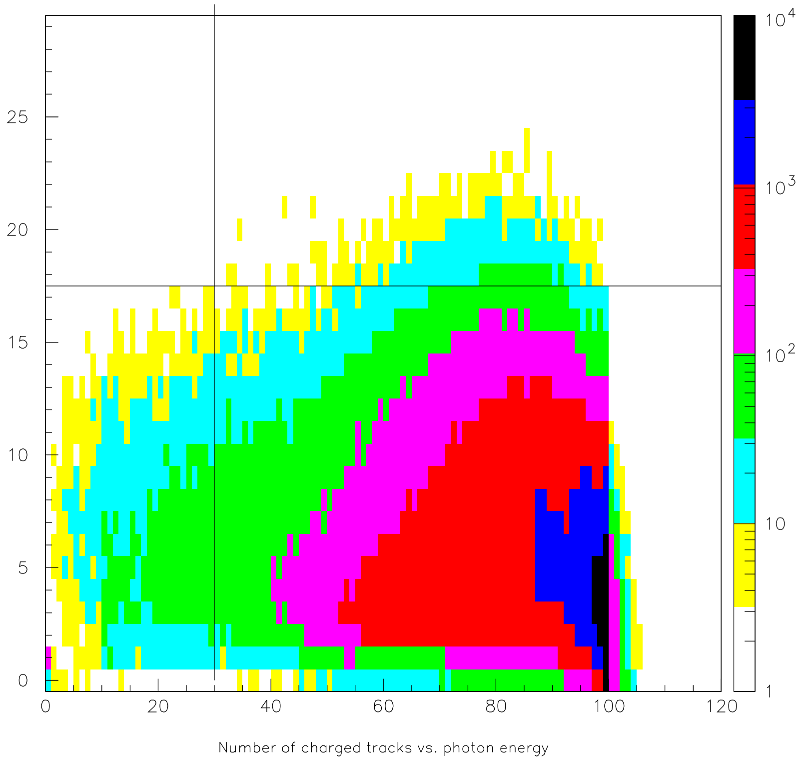


Figure 5.2: Plot showing the number of charged tracks in the event versus the maximum of the three photon energy variables (used in cuts number 4, 5, and 6 in the four jet preselection on page 92) for a typical $\gamma\gamma$ generated sample. The cuts are shown as lines, with the events passing the cuts in the upper left-hand quadrant.

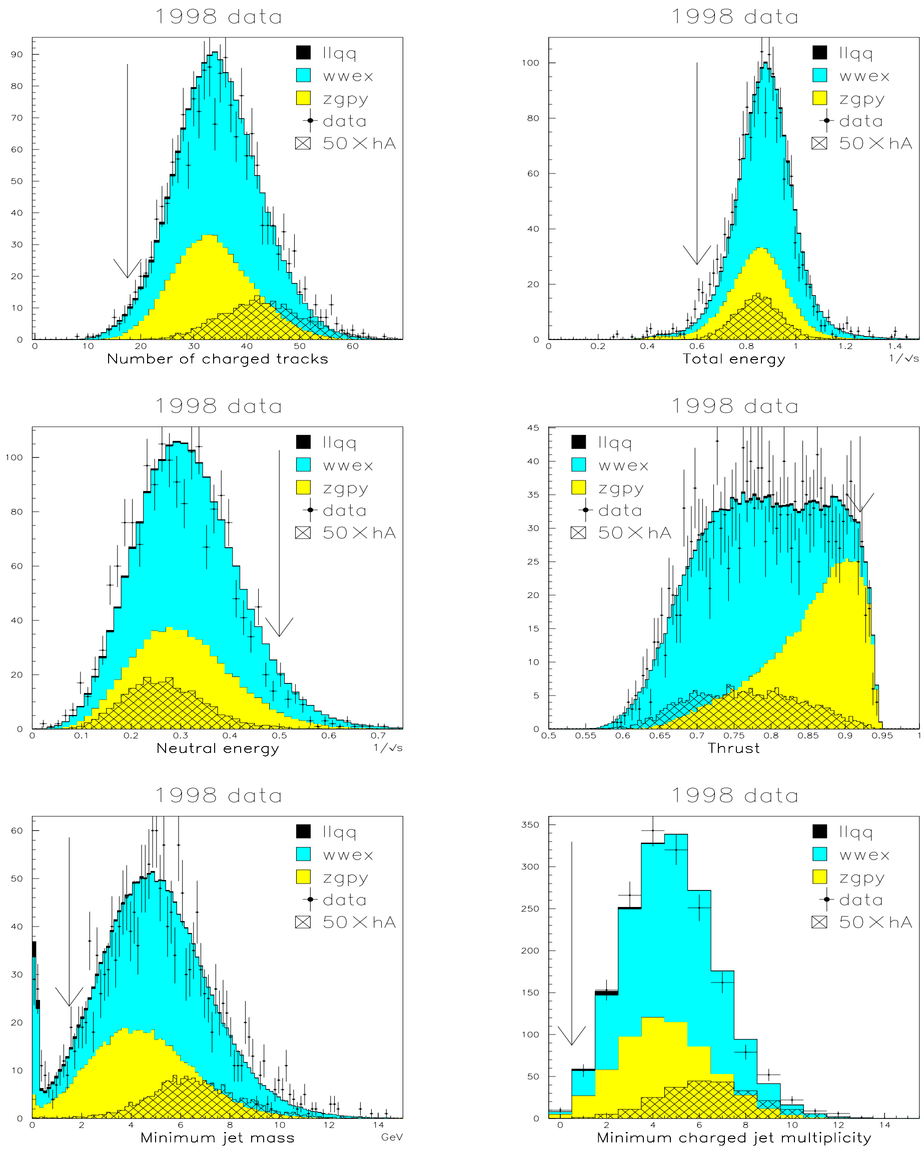


Figure 5.3: Figure showing plots of six different four jet selection variables, when all four jet selection cuts have been performed except for the variable shown in the plot in question. The Monte Carlo generated signal corresponds to the m_{A^0}=85 GeV/c^2, $\tan\beta$=20 signal. The plots are for the 1998 data sample.

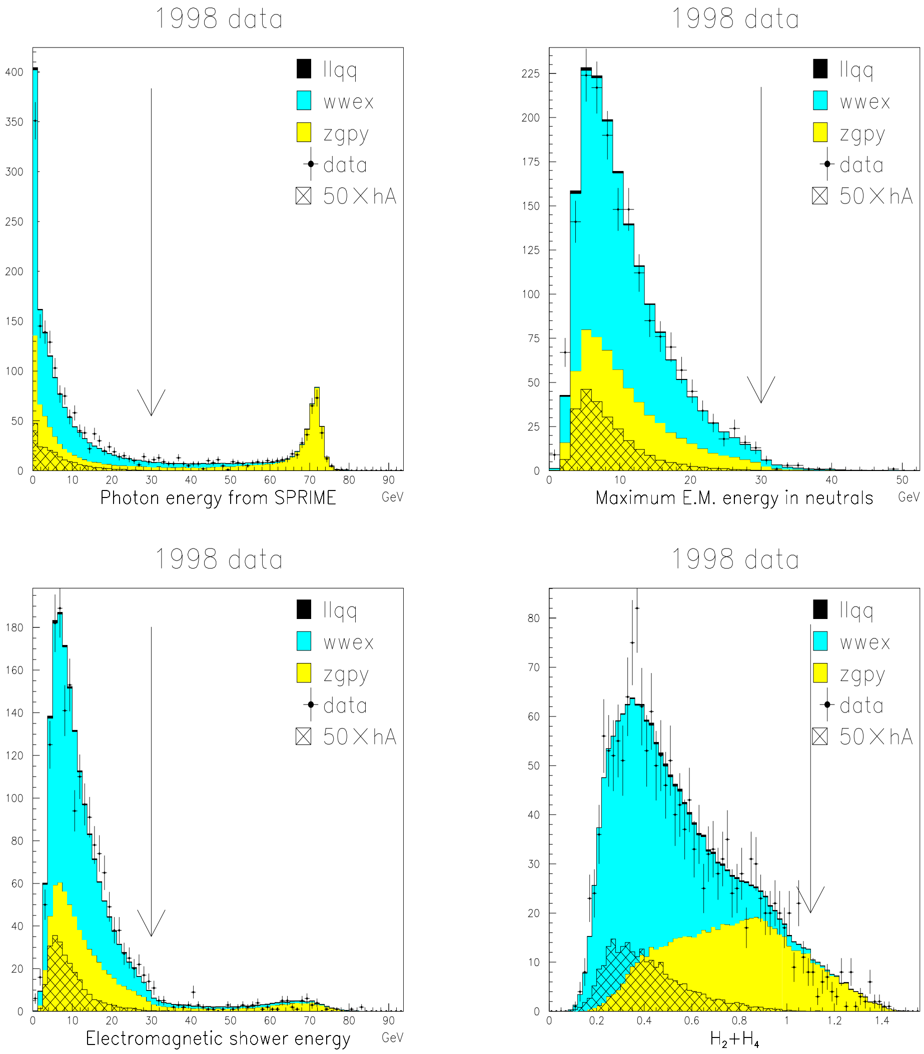


Figure 5.4: Figure showing plots of four different four jet selection variables, when all four jet selection cuts have been performed except for the variable shown in the plot in question. The Monte Carlo generated signal corresponds to the m_{A^0}=85 GeV/c^2, $\tan\beta$=20 signal. The plots are for the 1998 data sample.

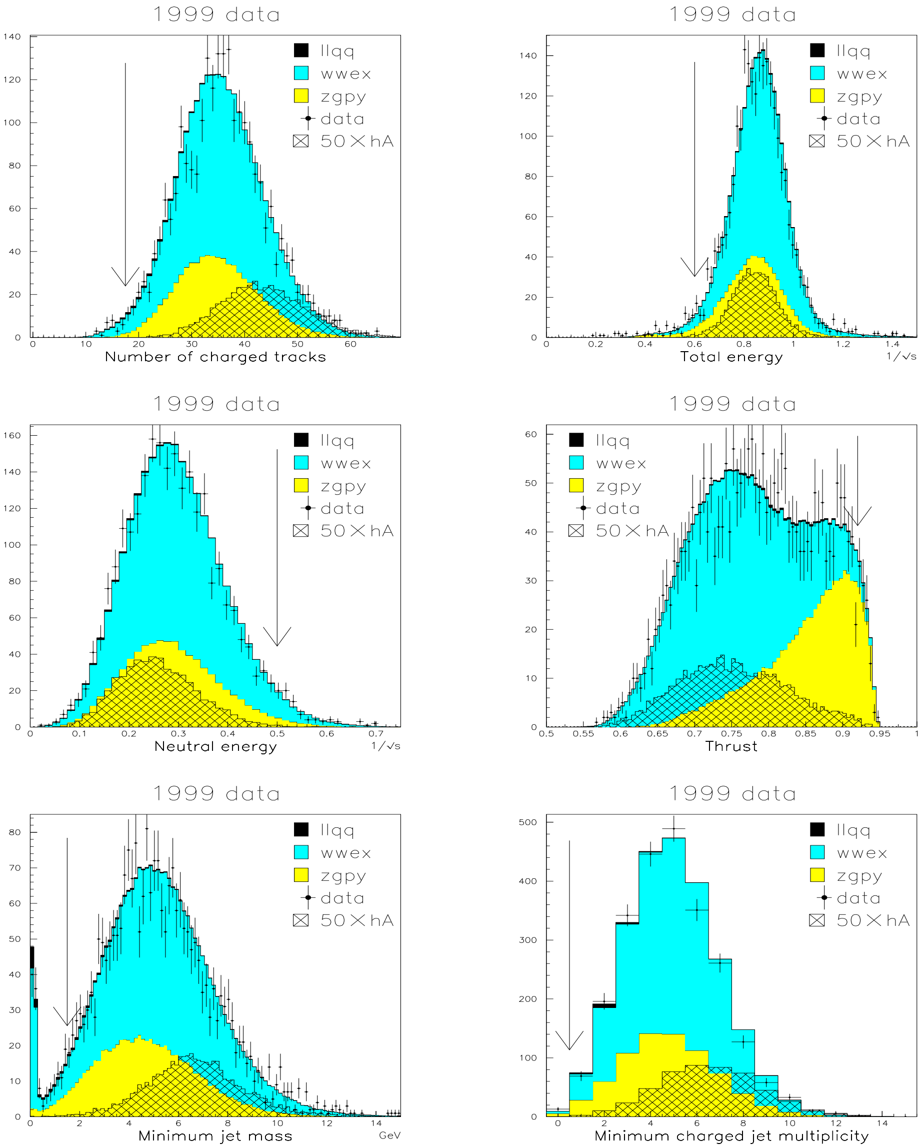


Figure 5.5: Figure showing the same plots as figure 5.3 for the 1999 data sample.

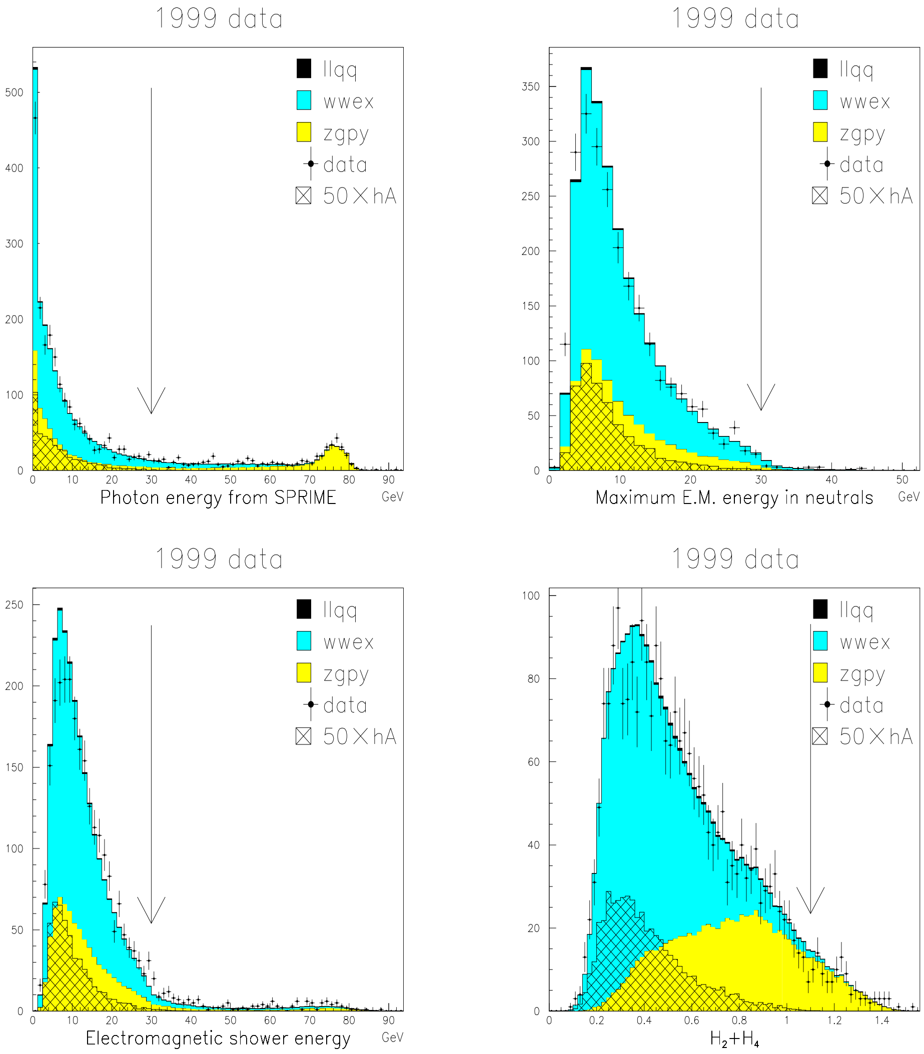


Figure 5.6: Figure showing the same plots as figure 5.4 for the 1999 data sample.

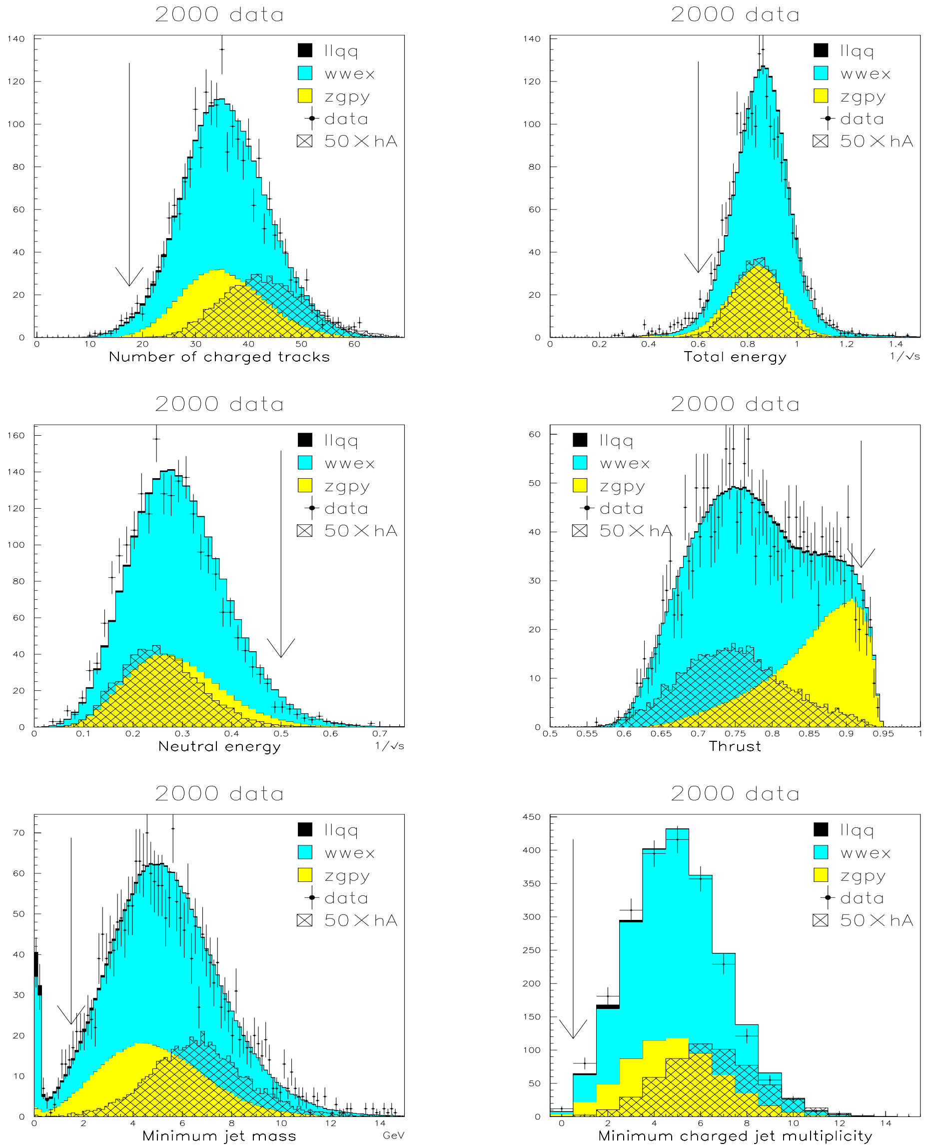


Figure 5.7: Figure showing the same plots as figure 5.3 for the 2000 data sample.

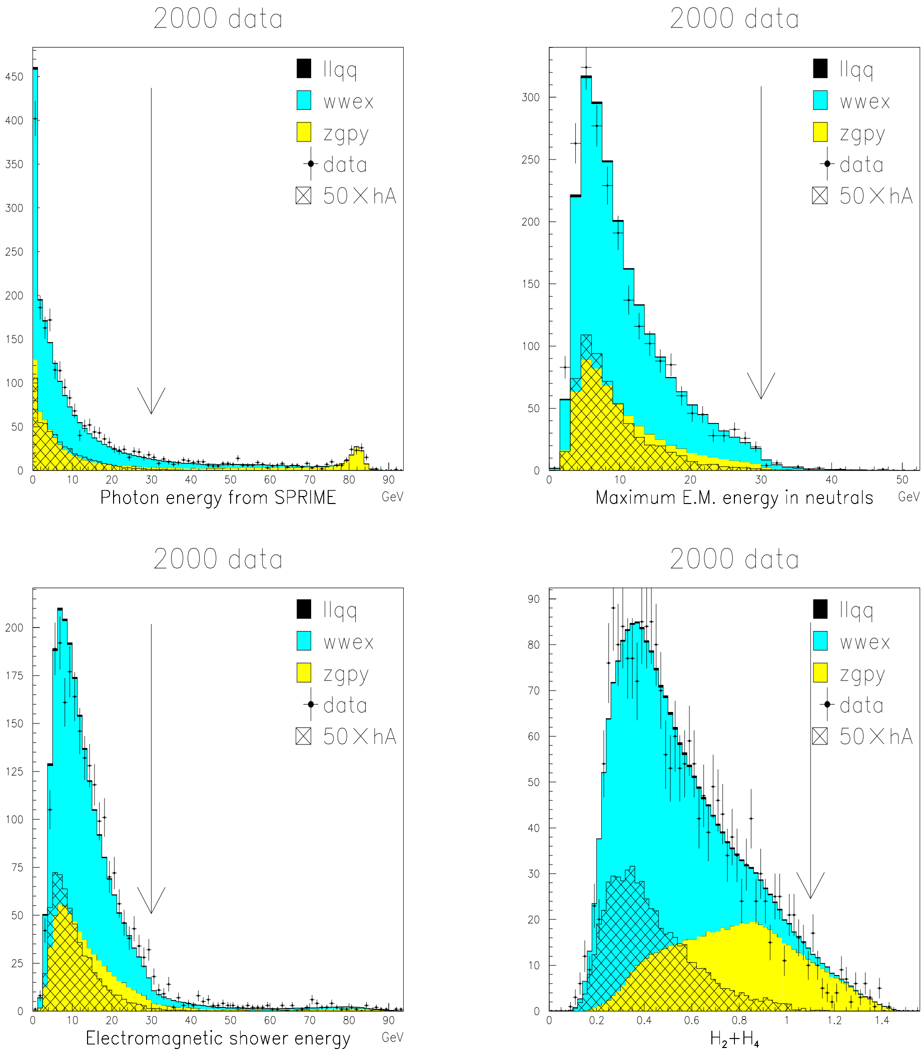


Figure 5.8: Figure showing the same plots as figure 5.4 for the 2000 data sample.

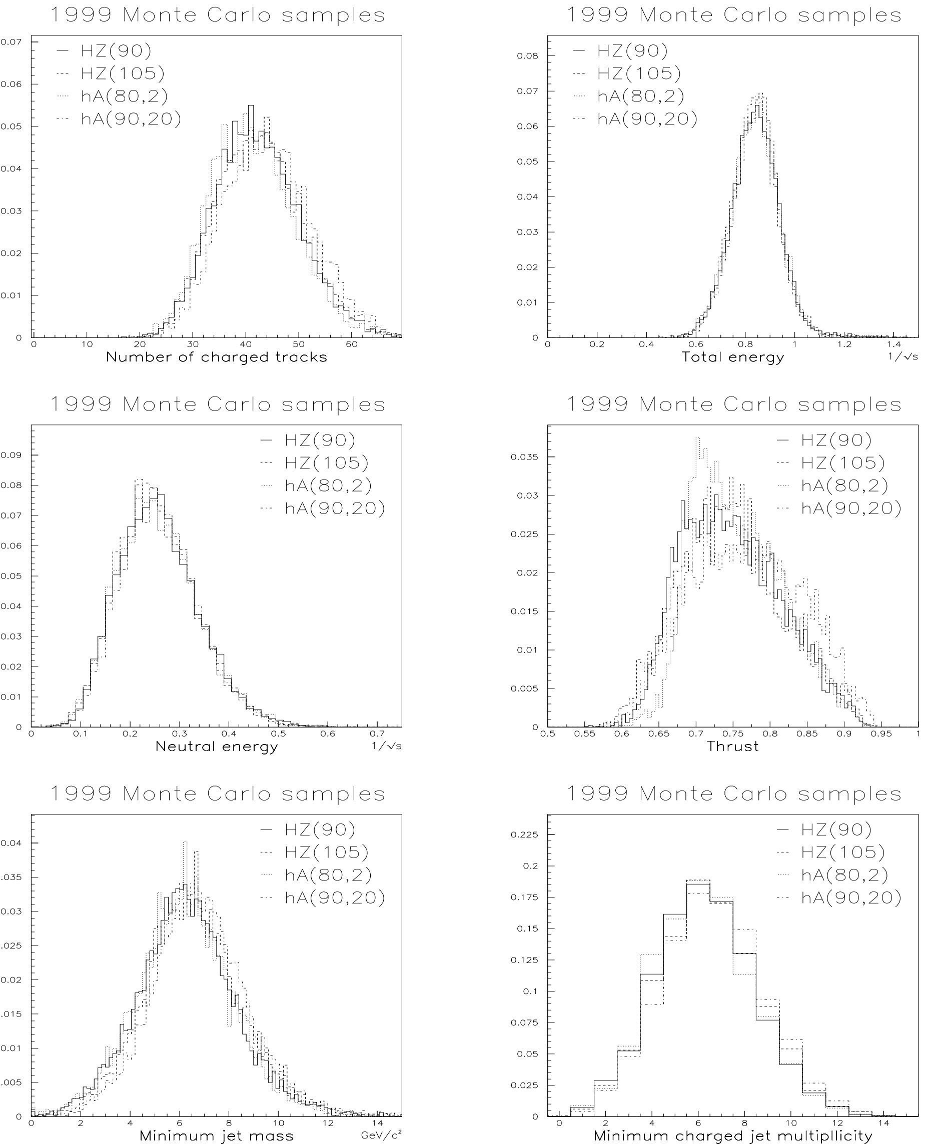


Figure 5.9: Figure showing the same plots as figure 5.3 for different Monte Carlo generated signal samples for the sum of centre-of-mass energy points of the 1999 data.

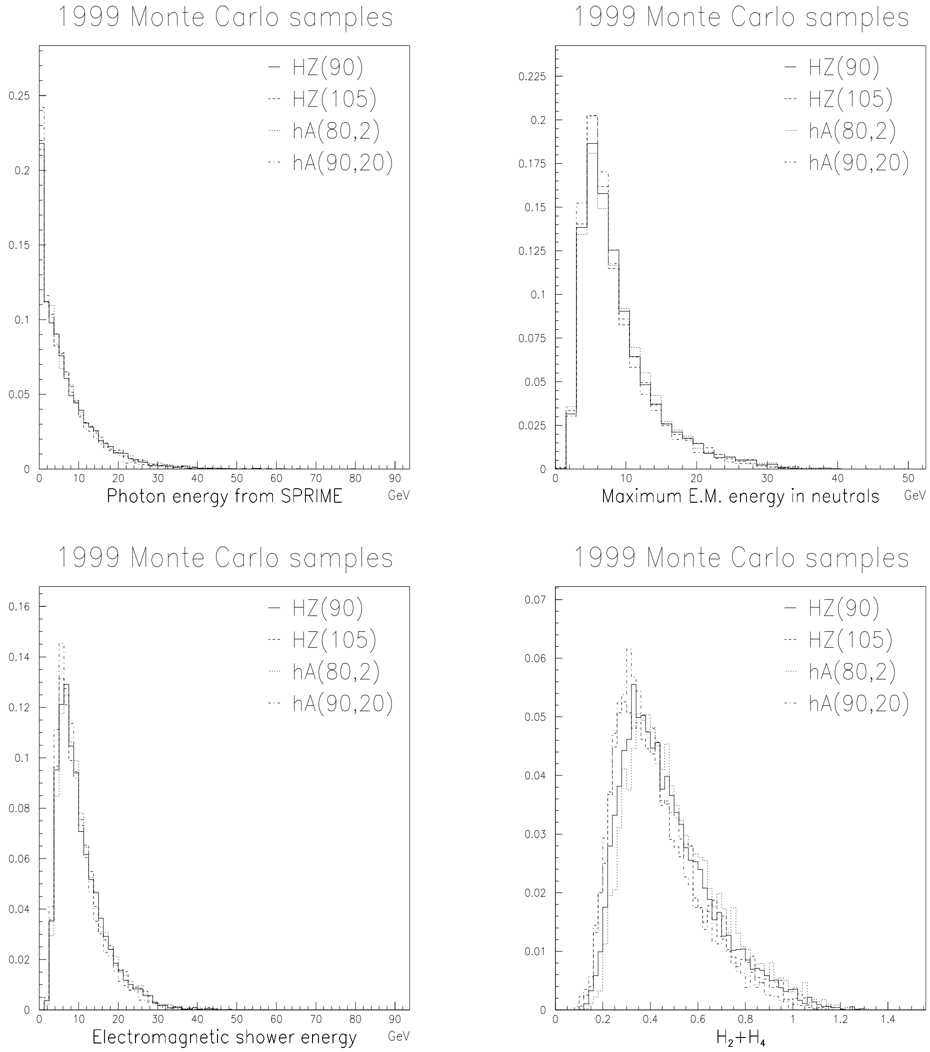


Figure 5.10: Figure showing the same plots as figure 5.4 for different Monte Carlo generated signal samples for the sum of centre-of-mass energy points of the 1999 data.

energy of the 2000 data in figure 5.7, to name some), these are often not persistent over the three year. In a few cases one can see a small shift between data and expected background for all the three year (the neutral energy and the minimum charged jet multiplicity), but in these cases the selection cut is situated far from these discrepancies.

5.3 Track and event property studies on four jet events

In this section a comparison is made between the two different track selections mentioned earlier in this chapter. The agreement of data with expected background from Monte Carlo is emphasized, as this has been an important consideration in the work with the standard track selection. The comparisons are all made at the four jet preselection level. For the plots in this section, the left-hand column shows data and expected background for the old hadronic search team track selection, whereas the right-hand column shows the same for the new standard track selection.

The plots in figure 5.11 show the momentum, polar angle and total energy of charged tracks for the two track selections for the sum over all centre-of-mass energy windows, whereas the same plots are given for neutral particles in figure 5.12.

For the charged particle plots in figure 5.11, the difference between the two track selections is not particularly large. Both the old hadronic search team selection (left column) and the new standard selection (right column) show a reasonably good agreement between expected background and data. The charged tracks momentum distribution in the old track selection shows some slight deterioration in the agreement between data and expected background for the very largest track momenta, whereas the new track selection shows good agreement for the entire range of track momenta. The charged tracks which in the old track selection are rescaled to half of the beam energy can be seen as a bump around 50 GeV/c. Also, the plots of charged tracks polar angle shows a slightly less good agreement between data and expected background for polar angles far from 90 degrees for the old track selection. In the same plots, the small dip in the distribution at polar angle 90° is due to the crack between the two hemibarrels of DELPHI.

The plots concerning the neutral particles in figure 5.12, however, show larger differences, both between the expected background and data, and between the different track selections. The two plots of the neutral particle energy both show a deficit of data with respect to expected background in

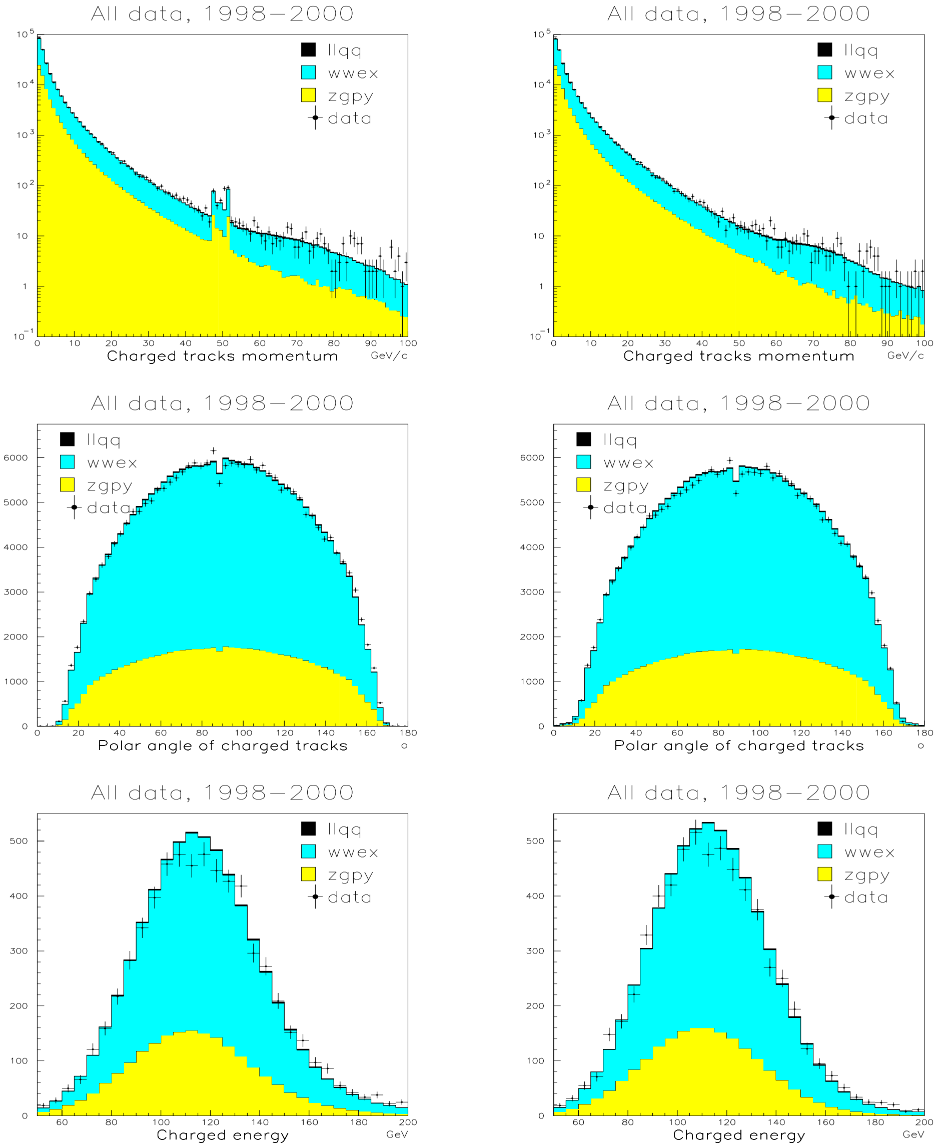


Figure 5.11: Figure showing plots of different charged tracks quantities; the charged track momentum on top, the charged track polar angle in the middle and the number of charged tracks in the event at the bottom. The left-hand column shows the old hadronic search team track selection, and the new standard track selection is shown on the right. The plots are for the sum over all centre-of-mass energy windows.

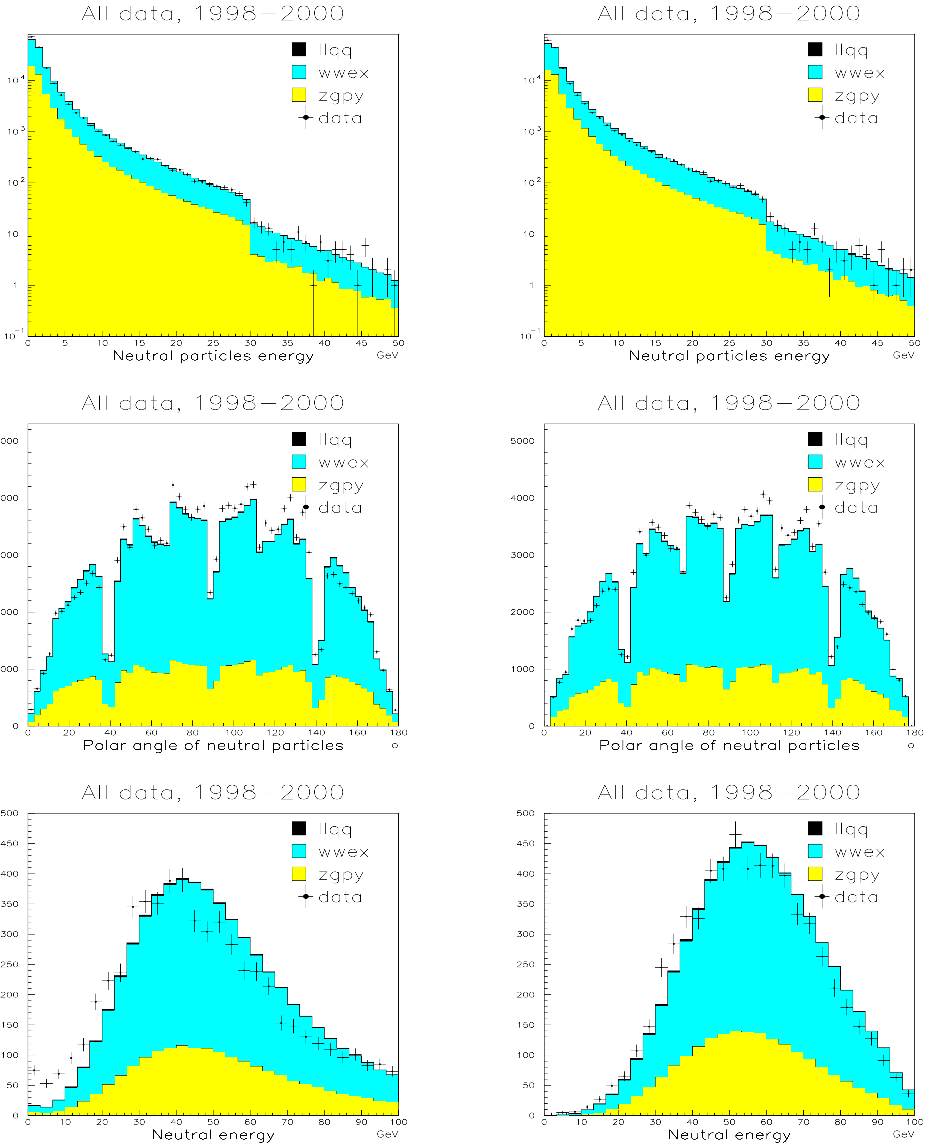


Figure 5.12: Figure showing plots of different neutral particles quantities; the neutral particle energy on top, the polar angle of the neutrals in the middle and the number of neutral particles in the event at the bottom. The left-hand column shows the old hadronic search team track selection, and the new standard track selection is shown on the right. The plots are for the sum over all centre-of-mass energy windows.

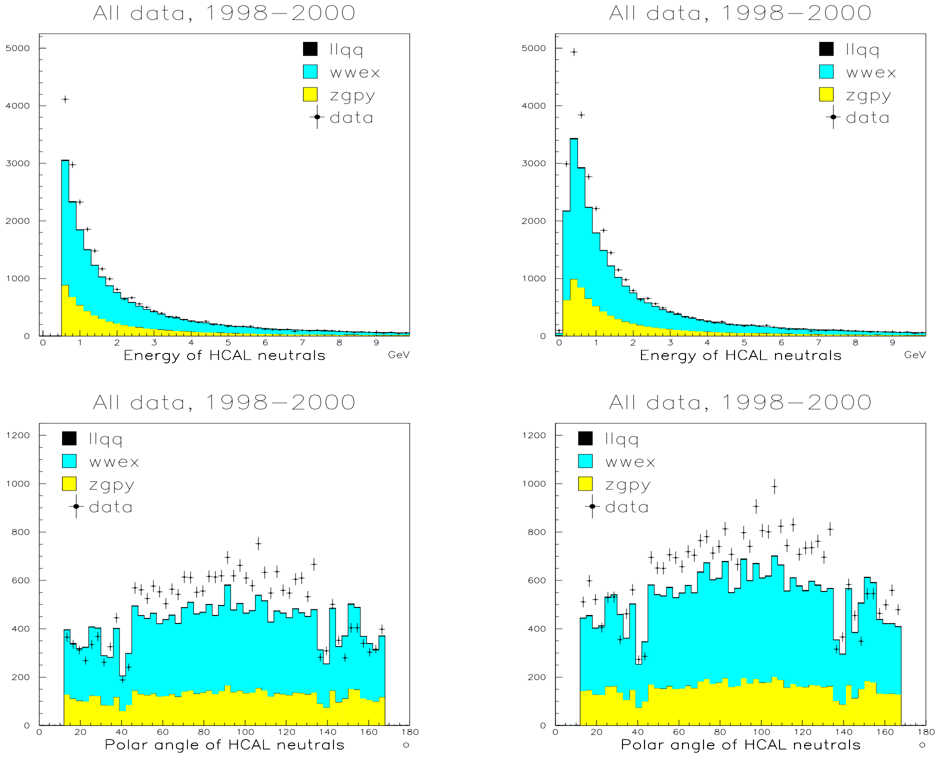


Figure 5.13: Figure showing plots of different particle quantities for neutral particles in the HCAL, and comparison between the old hadronic search team track selection (left column), and the new standard track selection (right column). In the top row is shown the energy of neutral particles in the HCAL, whereas the polar angle of such particles is shown in the lower row. The plots are for all centre-of-mass energy windows.

the energy range up to about 15 GeV, with the exception of the first bin, where there is an excess of data (the edge at 30 GeV reflects the fact that no events with a neutral particle of electromagnetic energy above 30 GeV are excepted as four jet events). Thus, one would expect the total neutral energy of the event to be too small in data with respect to the expected background, which can indeed be seen to be the case in the bottom row of figure 5.12. However, this effect is much more pronounced in the old track selection than in the new. Also, the polar angle distribution of neutral particles, where one can clearly see the cracks between different calorimeters, shows that the excess of data is mainly concentrated in the barrel.

When examining these results more closely, it is natural to look at the response of individual sub-detectors. Concentrating first on the different

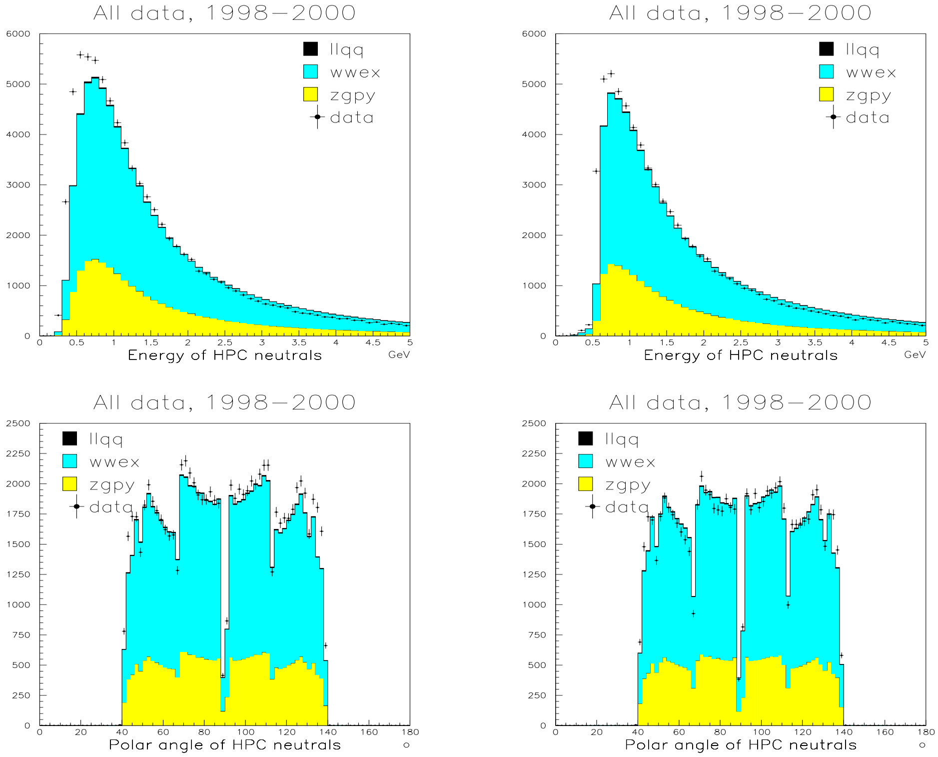


Figure 5.14: Figure showing plots of different particle quantities for neutral particles in the HPC, and comparison between the old hadronic search team track selection (left column), and the new standard track selection (right column). In the top row is shown the energy of neutral particles in the HPC, whereas the polar angle of such particles is shown in the lower row. The plots are for all centre-of-mass energy windows.

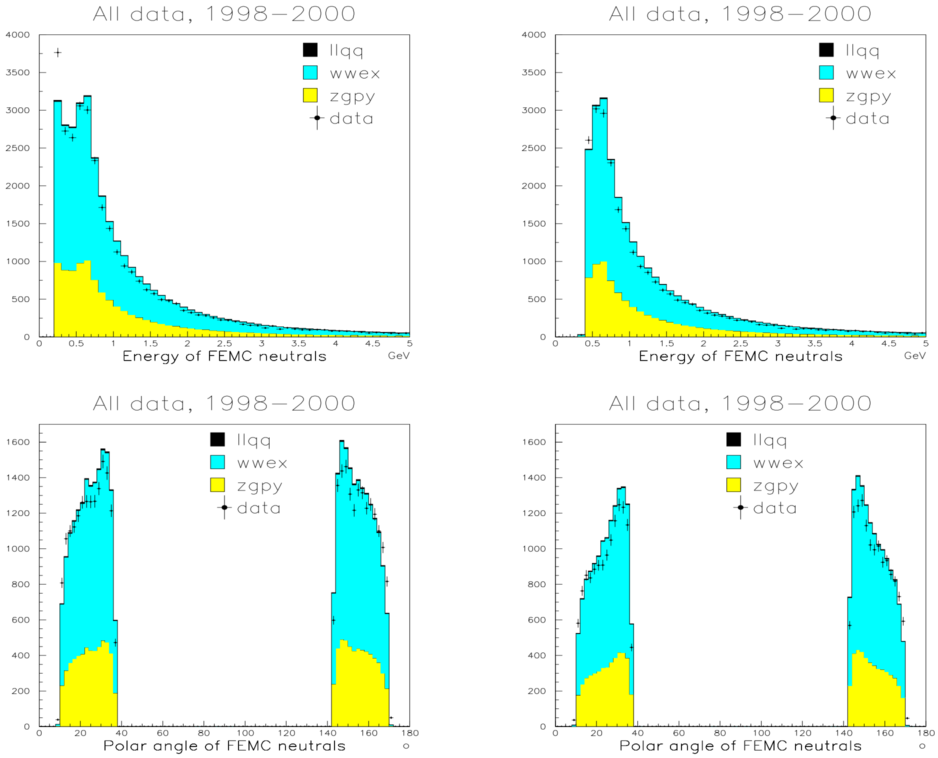


Figure 5.15: Figure showing plots of different particle quantities for neutral particles in the FEMC, and comparison between the old hadronic search team track selection (left column), and the new standard track selection (right column). In the top row is shown the energy of neutral particles in the FEMC, whereas the polar angle of such particles is shown in the lower row. The plots are for all centre-of-mass energy windows.

calorimeters, the calorimeter energy for neutral particles and the polar angle distribution of these are shown for the three most important calorimeters, the HCAL, HPC and FEMC, in figures 5.13 to 5.15. The plots of the calorimetric energies are all concentrated at the lower edge of the energy range, as the discrepancies in figure 5.12 are mainly located here.

The plots for the HCAL energies in figure 5.13 show fairly large discrepancies between data and expected background for both track selections. However, this discrepancy is at its largest for slightly higher energies for the old track selection (fourth bin) than for the new track selection (third bin). Also, the new track selection allows for the recovery of HCAL neutral particles below 500 MeV, albeit at the price of a somewhat large discrepancy between data and expected background.

For the plots in figure 5.14 regarding neutral particles in the HPC, the main part of the data excess for the energy plot is located below 1 GeV, and this discrepancy is larger for the old track selection than for the new. Also, the polar angle plot shows a slightly better overall agreement between data and expected background for the new track selection, as there is a small excess of data for the old track selection case, mainly at polar angles above 90 degrees.

The plots in figure 5.15 for the neutral particles in the EMF show a general deficit of data with respect to the expected background; the polar angle plots show that this effect is distributed over most polar angles covered by the EMF, with the exception of the polar angles furthest away from 90 degrees (smallest angles with respect to the beam pipe), where this deficit is reversed to an excess. An excess can also be seen in the first significantly populated bin of the energy distribution (the third bin for the old track selection, and the fifth bin for the new track selection), which is much more pronounced in the old track selection than in the new.

These comparisons show that the agreement between Monte Carlo expected and observed data is slightly improved with the new standard track selection, even though there is still some discrepancy present (an excess of neutral particles in data with respect to the Monte Carlo expectation). As the new standard track selection has been promoted as useable for most, possibly all, analyses within DELPHI, this is in itself an argument in favour of using this in the analyses. Also, since the tests and comparisons performed in this section show the new track selection to be superior, if only slightly, to the old in terms of agreement between expectation and observation, there is no good reason to select anything other than the standard choice of track selection, which is the new one. Therefore, the rest of the analysis will be performed using the new standard track selection.

After now having decided on and presented the basic starting point for the analyses, the next chapter describes the analysis philosophy and method, the repeated 2D likelihood network. First, some motivation points are given to justify the specific choice of analysis method. Then, a few different methods for approximating 2-dimensional distributions are shown, and tested on some realistic distributions in order to choose the one most ideally suited to the current use in the repeated 2D likelihood network.

Chapter 6

Analysis method

Most analyses aimed at discriminating a given background from a signal (possibly parameter dependent), falls into one of two classes. These can be summarized as follows:

Sequential cuts: This method is based on a series of cuts in different variables which distinguishes between signal and background. After these cuts are performed, one is left with a certain number of expected events of background and signal, and the final stastistical treatment of the search result is then performed on this, often relatively small, number of events. The advantage of this method is that it is simple, both in implementation and in philosophy, and therefore transparent, making comparison between the expected background and signal with the observed result easy. The disadvantage is that correlations between different variables are not taken into account, making the analysis suboptimal.

Multivariable discriminating methods: The goal of such methods is to take the correlation between different variables into account, thus making the analysis more optimal, *i.e.* improve the separation between signal and background. The implementation is often such that all the different variables are fed into a discrimination mechanism (likelihood method, neural net, or some similar mechanism), which then returns a small number of, often only one, discriminating variable(s). The advantage of such methods is that they take a larger part of the available discriminating information, that is the correlation between different variables, into account, making the analysis more efficient. The disadvantage is that the discriminating mechanism can often be a rather closed system, and thus have many of the characteristics of a "black box", making it difficult to control how the analysis behaves with respect to consistency checks in such areas as the agreement between the expected background and the observation.

The analysis presented here is an attempt at keeping the favourable features from both these classes, and minimizing the disadvantages, *i.e.* taking the information from the different correlations between variables into account, while at the same time keeping the analysis transparent and convenient with respect to controling the agreement between simulation and real data.

6.1 The repeated 2D likelihood network

The analysis procedure consists of a repeated technique comprising the following steps:

- Construct the 2-dimensional distribution of two variables.
- Make an approximation of this distribution (see section 6.2), and construct the $\frac{s}{b}$(signal divided by background)-distribution.
- Transform the $\frac{s}{b}$-distribution into one which has a flat signal distribution (see section 6.3).
- Use this transformed $\frac{s}{b}$-distribution as the output variable of the combination step.

This procedure is then repeated until a small number, possibly only a single, variable remains.

Thus, the analysis combines a number of input variables to at most a few output variables, taking the most important correlations into account. In this respect, the analysis behaves much like a multivariable discriminating method as described earlier, with the important difference that each new combination step yields another input variable, which then in turn may be combined further. The advantage with this approach, is that one can check the development of both signal and background distributions throughout the analysis, since the result from each combination step is available as a single variable. This makes it easier to check the analysis for unexpected features or unnatural behaviour. Also, it becomes significantly easier to monitor the agreement between the Monte Carlo generated background and the data.

6.2 Approximation of 2D distributions

An essential part of the analysis procedure is the approximation of the different 2-dimensional distributions in the combination step. This is done by smoothing the original distributions, which is necessary due to the limited amount of Monte Carlo generated signal and background samples, and therefore spikes and discontinuities in the distributions. There is also a strong incentive from the analysis method itself to have smooth distributions which have non-zero values in the entire range of possible values for the different search variables, since the fraction $\frac{s}{b}$, which the whole analysis technique is based on, will become infinite in regions of zero background and non-zero signal, signifying a discovery in the case of observing only the smallest amount of data (*i.e.* a single event) in such an area. Even though background-free experiments, in which such a scenario might be feasible, could by constructed, this is certainly not the case for the analysis presented here, as there is irreducible background (specifically, the Z^0Z^0 background) in the entire range of search variables.

Another reason for wanting to approximate the distributions by smoothing them out, is to avoid overtraining. Due to the limited amount of Monte Carlo simulated data, one would like to use all available generated samples in the analysis, both to construct the repeated 2D likelihood network and to estimate the expected background, in order to minimize the statistical error in the signal and background estimates. This solution does however introduce the problem of overtraining, as statistical fluctuations in the Monte Carlo generated samples will propagate into the 2D likelihood network, where such fluctuations might be utilized to distinguish between signal and background, when these differences have no basis in physical realities. Thus, an important goal for a smoothing procedure is to correct for statistical fluctuations in the 2-dimensional distribution. The overtraining problem in the specific implementation of the h^0A^0 and H^0Z^0 search channels presented here, is further studied in section 7.6.

Several different smoothing algorithms and methods were tried, most of which were found to have different strengths and weaknesses. But the smoothing procedures all have a common goal, which can be summarized in the following points:

- The smoothed distribution should represent, as closely as possible, the underlying, exact distribution approximated by the original histogram.

- The smoothed distribution should not contain traces of statistical fluctuations in the original histogram, leading to features with no basis in physical realities in the true, underlying distribution. Such features would lead to overtraining and a non-optimal distribution optimization.

- The smoothed distribution should not introduce new features present neither in the underlying exact distribution nor in the original histogram. This is particularly important in the case of spikes, steep edges and other sharply varying features in the distribution.

The following subsections present a short review of the different smoothing algorithms and approaches which have been studied. The original, unsmoothed histogram, of a number n_x bins in the x-direction and n_y bins in the y-direction, is denoted by $H_{i,j}$, which is the number of events in histogram bin (i, j), and the total number of events in the histogram is $H_{\text{tot}} = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} H_{i,j}$. Also, the histogram covers a certain rectangular region in the 2-dimensional space of the two variables making up the histogram; this is given by the four parameters x_{min} (lower edge in the x-direction), x_{max} (upper edge in the x-direction), y_{min} (lower edge in the y-direction) and y_{max} (upper edge in the y-direction).

6.2.1 Gaussian smearing approximation

The basic idea behind the gaussian smearing procedure is that the content of each bin is to be considered a delta function in that bin, and is thereafter subject to a gaussian smearing. That is, if one considers the content $H_{i,j}$ of a single bin, the bin edges being the two points $(x^i_{\text{min}}, y^j_{\text{min}})$ (lower left-hand corner) and $(x^i_{\text{max}}, y^j_{\text{max}})$ (upper right-hand corner), the content of this bin is smeared out in the bin (i', j') with bin edges $(x^{i'}_{\text{min}}, y^{j'}_{\text{min}})$ (lower left-hand corner) and $(x^{i'}_{\text{max}}, y^{j'}_{\text{max}})$ (upper right-hand corner) according to the gaussian integral

$$H^{i',j'}_{i,j} = \frac{H_{i,j}}{2\pi\sqrt{\sigma_x \sigma_y}} \int_{y^{j'}_{\text{min}}}^{y^{j'}_{\text{max}}} \int_{x^{i'}_{\text{min}}}^{x^{i'}_{\text{max}}} \exp\left(\frac{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2}{-2} \right) \mathrm{d}x\, \mathrm{d}y \qquad (6.1)$$

where (μ_x, μ_y) is the middle point of the bin with content $H_{i,j}$, *i.e.* $\mu_x = \frac{1}{2}\left(x^i_{\text{min}} + x^i_{\text{max}}\right)$ and $\mu_y = \frac{1}{2}\left(y^j_{\text{min}} + y^j_{\text{max}}\right)$. This smearing guarantees that the total content $H_{i,j}$ of bin (i, j) is preserved in the histogram.[1]

---

[1] This is only strictly true if the histogram covers the entire xy-plane, *i.e.* from $-\infty$ to ∞, which of course is not the case. However, the histogram is usually large enough compared to the smearing constants σ_x and σ_y for this effect to be very small for all bins except the ones near the histogram edges. In any case, the difference between the original

Equation 6.1 shows that the smearing, for reasons of simplicity, is assumed to be uncorrelated in the two spatial dimensions. Also, the amount of smearing, given by the parameters σ_x and σ_y, has not yet been determined. One might imagine determining these parameters from knowledge of the uncertainty in the two variables making up the histogram, but since this method is intended to be completely general, such knowledge can not be a priori assumed. However, some guidelines can be established by considering the following two points:

- In order to get a handle on the scale of the smearing, one should take into account the total width of the histogram, that is the span of the histogram in the two variables x and y. It is therefore natural to let the parameters σ_x and σ_y scale with the numbers $x_{\max} - x_{\min}$ and $y_{\max} - y_{\min}$, respectively.
- As the smearing is intended to correct for effects due to insufficiently accurate description of the 2-dimensional distribution, it is of interest to study the uncertainty within a single bin. If one has a sample of n events all with the same expected population mean in the x- and y-variables (which in this context can be viewed as a good approximation in the case of events within one single bin), the standard deviation of the sample mean is equal to $\sigma/\sqrt{n}$, where σ is the expected population standard deviation. That is, the uncertainty scales with the inverse of the square root of the number of events, which makes it natural to scale the parameters σ_x and σ_y with the number ${H_{i,j}}^{-\frac{1}{2}}$.

Thus, the parameters σ_x and σ_y should be a function of which bin is being smeared, and therefore take the form $\sigma_x^{i,j}$ and $\sigma_y^{i,j}$. In addition, a "cutoff" factor B is introduced for the scaling of the parameters σ_x and σ_y with the factor ${H_{i,j}}^{-\frac{1}{2}}$ in order to make sure that the amount of smearing does not fall below a certain minimum. This is done to make sure that imperfections in the histograms originating from other sources than pure statistics will also be smeared. Also, there is no a priori reason to distinguish between the two spatial directions of the histogram, due to the generality of the method, and the final form of the smearing parameters becomes

$$\sigma_x^{i,j} = A(x_{\max}-x_{\min})\min(H_{i,j}, B)^{-\frac{1}{2}} \qquad \sigma_y^{i,j} = A(y_{\max}-y_{\min})\min(H_{i,j}, B)^{-\frac{1}{2}} \tag{6.2}$$

---

bin content $H_{i,j}$ and the actuall amount contained within the histogram is corrected for by rescaling all the smeared contributions in all the bins of the histogram so as to sum up exactly to the value $H_{i,j}$.

where the only free parameters are the constants A and B, which should be determined in such a way as to ensure enough smearing in order to avoid overtraining.

The total smeared histogram $H'_{i',j'}$ is achieved by summing the smeared contributions from each bin, so that the final form of the contents of the smeared histogram in bin (i', j') is given by the following formula:

$$
\begin{gathered}
H'_{i',j'} = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} H_{i,j}^{i',j'} \qquad \text{where} \\
H_{i,j}^{i',j'} = \frac{H_{i,j}}{2\pi\sqrt{\sigma_x^{i,j}\sigma_y^{i,j}}} \int_{y_{\mathrm{min}}^{j'}}^{y_{\mathrm{max}}^{j'}} \int_{x_{\mathrm{min}}^{i'}}^{x_{\mathrm{max}}^{i'}} \exp\left(\frac{\left(\frac{x-\mu_x}{\sigma_x^{i,j}}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y^{i,j}}\right)^2}{-2}\right) \mathrm{d}x\,\mathrm{d}y \\
\mu_x = \tfrac{1}{2}\left(x_{\mathrm{min}}^i + x_{\mathrm{max}}^i\right) \quad \mu_y = \tfrac{1}{2}\left(y_{\mathrm{min}}^j + y_{\mathrm{max}}^j\right) \quad X = \tfrac{A}{\min(H_{i,j},B)^{\frac{1}{2}}} \\
\sigma_x^{i,j} = X\left(x_{\mathrm{max}} - x_{\mathrm{min}}\right) \qquad \sigma_y^{i,j} = X\left(y_{\mathrm{max}} - y_{\mathrm{min}}\right)
\end{gathered}
\tag{6.3}
$$

As already mentioned, the free parameters A and B should be determined by giving it as small a value as possible while still avoiding overtraining problems. After testing several different values for the parameter, the lowest value before overtraining effects start to be visible was found to be

$$A = 0.300, \qquad B = 300 \tag{6.4}$$

which therefore are the selected values for these parameters.

6.2.2 Spline approximation

In this section, smoothing of the 2-dimensional distributions by applying the method of B-splines [77] is discussed. The definition of the B-spline interpolation and its parameters can be found in [78, page 149], and is summarized in appendix F. The resulting 2-dimensional B-spline is, for the implementation used in this context, described by two parameters, the number of spline knots in the x- and y-direction, m_x and m_y. When determining these two parameters to be used in the smoothing algorithm, a loop was performed over both m_x and m_y over a limited range of values as not to introduce overtraining. In the 1-dimensional case, a suggested value for the number of spline knots is given by $4N_P + 6$, where N_P is the number of statistically significant peaks in the distribution [79, page 132]. With this as a guideline, and assuming at most two statistically significant peaks in the 2-dimensional distribution, the range of values for m_x and m_y was chosen to be from 6 to 14, and the corresponding spline approximation was performed for each pair of m_x and m_y values. An estimate of the difference with respect to the original histogram, a χ^2 value given by

$$\chi^2 = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \left(\frac{H'_{i,j} - H_{i,j}}{\max\left(1, \sqrt{H_{i,j}}\right)} \right)^2 , \tag{6.5}$$

where $H'_{i,j}$ is the histogram bin content of bin (i, j) for the histogram resulting from the spline procedure, was found for each histogram corresponding to a set of spline parameters. After this χ^2 value had been calculated for all the histograms resulting from values of m_x and m_y in the loop, the pair of m_x and m_y giving the smallest χ^2 was chosen as the one used in the smoothing of the original histogram.

6.2.3 Neural Net approximation

Neural nets [80] have become very popular as tools in experimental particle physics due to their generality and flexibility, their ability to provide highly optimized solutions to complex problems and to utilize correlations in a multi-dimensional environment. Neural nets can be adapted to a multitude of problems, such as final search and measurements analysis of experimental results, particle identification, track search and fitting in complex environments and triggering of events in noisy and highly track rich experimental situations, to name a few. In this context, however, a neural network will be used to approximate 2-dimensional histograms, and thus give a description of the true, underlying 2-dimensional distribution from which the histogram has been constructed, hopefully conforming to the points given on page 117. The specific implementation of the neural network is in the form of a Multi-Layer Perceptron (MLP) [81].

However, when used for the 2-dimensional histograms in the context of this analysis method, several problems appear. Firstly, as explained in section 6.3, the distributions, although by definition flat for the signal, will peak towards small values for the background. The analysis is usually performed on events passing a fairly loose preselection, in order to obtain as high a signal efficiency as possible, and the histograms forming the combination steps in the analysis will therefore for the background become highly peaked towards lower values. Such histograms can in many cases be difficult to approximate by neural nets, and may lead to the output of the network having very little resemblance with the original histogram. A few such cases are shown in figures 6.1 to 6.3. This problem can be remedied by using larger and more complex networks, but this in turn leads to other problems. One fundamental problem is that such complex neural nets are likely to introduce overtraining problems, particularly for the histograms which are less peaked than the ones which require such nets to be introduced. This is a severe problem which in itself could be reason enough for disregarding this method of smoothing. In addition to this, there is also a more practical problem with such complex networks, which is that they require a rather large number of training cycles in the optimization process, something that quickly becomes very CPU consuming, making the neural network approach to the smoothing of the 2-dimensional distributions a very time consuming one, and therefore rather inconvenient.

When all these considerations are taken into account, the smoothing procedure which gives the overall best results in terms of stability, flexibility, time consumption and non-overtraining is the gaussian smearing approximation; see figures 6.1 to 6.3. Therefore, this has been chosen as the method used for smearing the different 2-dimensional distributions in the analysis.

6.3 Construction of the discriminating variable(s)

The final variable(s) is(are) constructed from the previous combination steps. If the original variables are divided into different classes, where the correlation between variables belonging to the same class is expected to be large, whereas the correlation between variables in different classes is expected to be small, and the combination is then performed at the earlier stages within the classes, one might hope to include most of the important correlation in the analysis tree. In this way, the analysis may be schematically described as in figure 6.4.

The transformation of the $\frac{s}{b}$-distribution is done in the following way: Assume the two input variables on which the combination is to be performed are named x and y, and define the function

$$z(x,y) = \frac{s(x,y)}{b(x,y)}$$

where $s(x,y)$ ($b(x,y)$) is the distribution of the signal (background) in the 2-dimensional space spanned by x and y, normalized to the number of expected events. Now let $P_s(z)$ be the probability distribution of this variable for the signal. The output variable from the combination procedure of x and y is now defined as:

$$var_{xy} = \int_0^z P_s(z)\,\mathrm{d}z \text{ (See figure 6.5)}$$

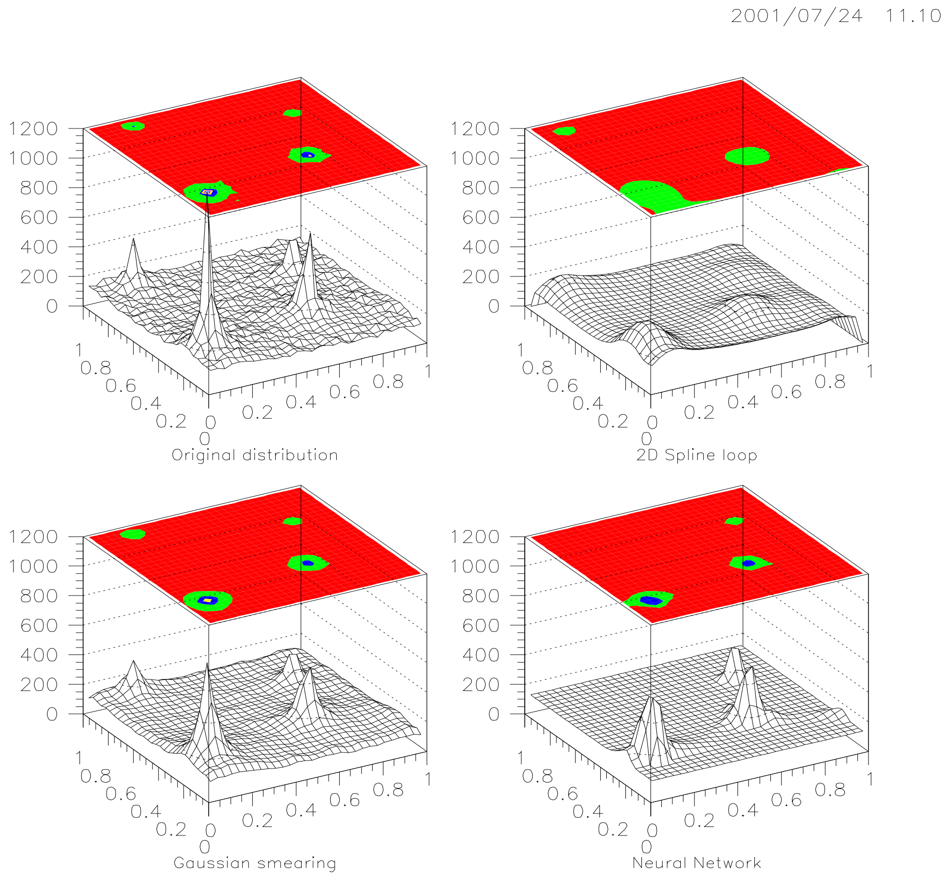


Figure 6.1: Figure showing the three different smoothing algorithms tested on a distribution with several bumps/spikes. On the upper left is shown the original distribution, the SPLINE approximation is shown on the upper right, the lower left shows the gaussian smearing approximation whereas the neural network approximation is shown on the bottom right. The neural network in question is one with 50 nodes in the first hidden layer and 10 nodes in the second hidden layer, and the network is trained for 2000 epochs with the learning method of conjugate gradients with Polak-Ribiere updating formula.

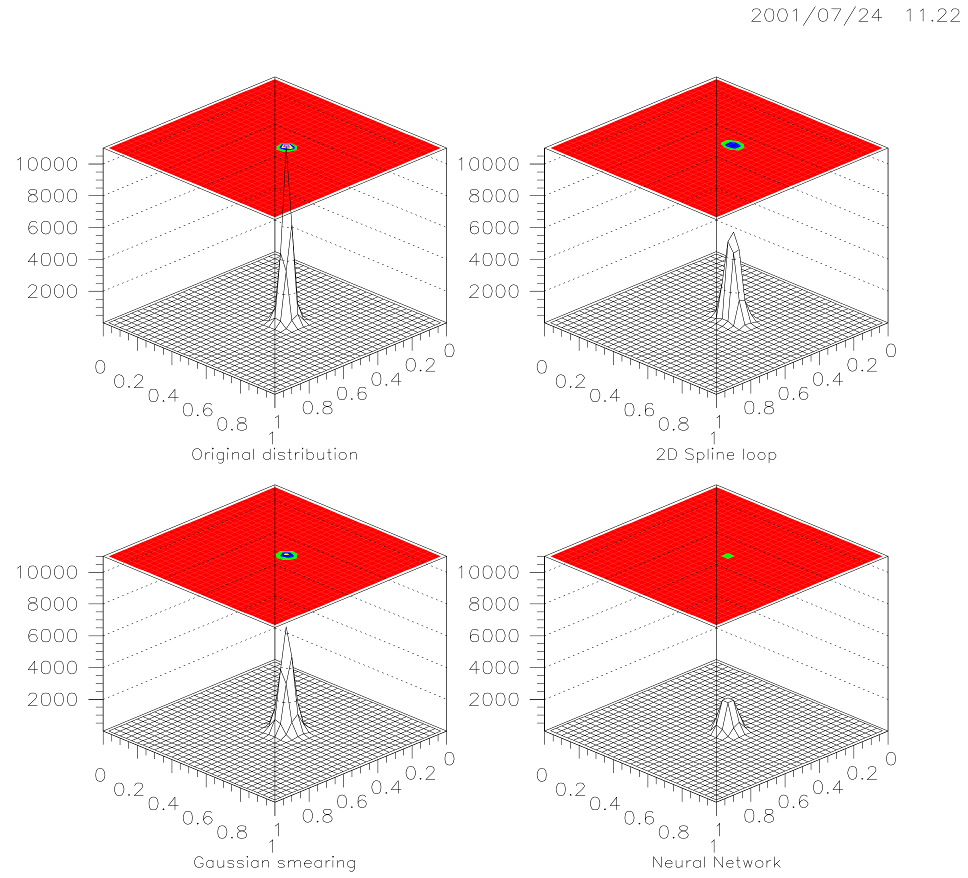


Figure 6.2: Figure showing the three different smoothing algorithms tested on a distribution with one central spike. On the upper left is shown the original distribution, the SPLINE approximation is shown on the upper right, the lower left shows the gaussian smearing approximation whereas the neural network approximation is shown on the bottom right. The neural network is the same as used in figure 6.1.

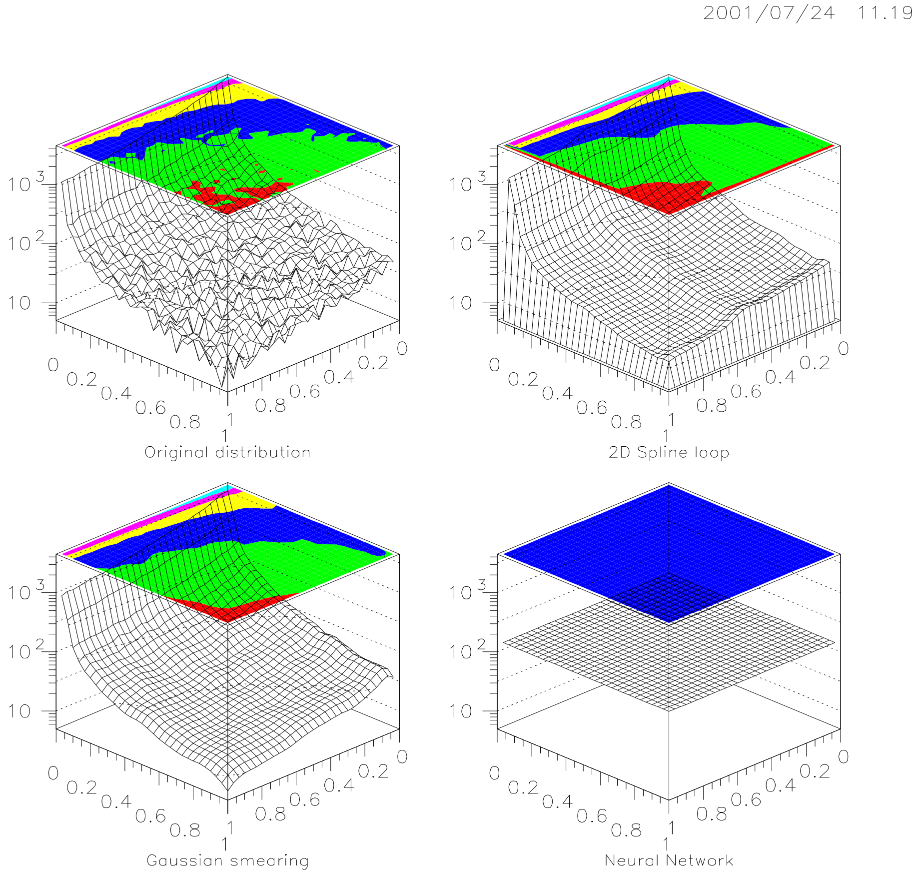


Figure 6.3: Figure showing the three different smoothing algorithms tested on a distribution peaked towards small values in both dimensions. On the upper left is shown the original distribution, the SPLINE approximation is shown on the upper right, the lower left shows the gaussian smearing approximation whereas the neural network approximation is shown on the bottom right. The neural network is the same as used in figure 6.1.

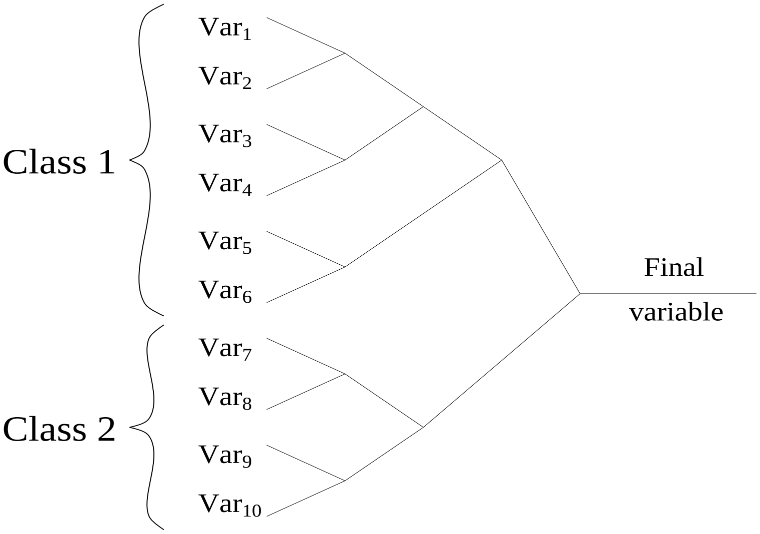


Figure 6.4: The general structure of the analysis, for the generic case of ten variables in two classes.

This procedure gives an output variable distributed between 0 and 1 for both the signal and the background, the signal being a flat distribution, whereas the background is concentrated at low values.

As the analysis method and underlying philosophy has now been described, the next chapter moves on to present the specific implementation of the repeated 2D likelihood network. The different variables and classes for both the H^0Z^0 and h^0A^0 analyses are discussed and motivated, and the specific shapes of the different parts of the network (*i.e.* the way and order in which the different variables are connected) are shown. Emphasis is put on the choice of the Higgs mass estimator for the two search channels, and different methods of constrained fits are discussed. The input variables to the analysis are shown for expected background, observed data, and a range of different signal hypotheses, and the analysis tree is presented for both the H^0Z^0 and h^0A^0 analysis. Studies of possible overtraining effects and systematic errors are presented, as well as the final stages of the analyses, *i.e.* the eventshape, b-tag and final discriminating variable of the two analyses.

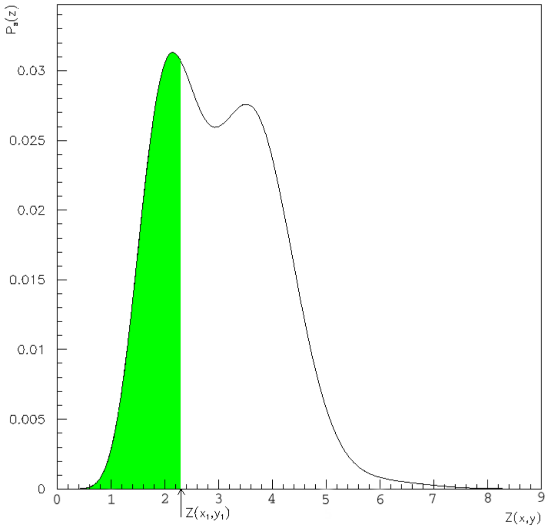


Figure 6.5: Transformation of the $\frac{s}{b}$-distribution: The output variable for a value of z equal to the $\frac{s}{b}$ -value of x_1 and y_1 equals the shaded area.

Chapter 7

Analysis

7.1 Discriminating variables and classes

The analyses presented here have 13 input variables to the 2D likelihood network, divided into two different classes, with an additional variable in an extra class for the 5C fit analysis in the h^0A^0 and H^0Z^0 channels (see section 7.2.3). These two classes are designed to reject different backgrounds, and thus emphasize different aspects of the expected signal events. The correlation between variables of different classes can therefore be assumed to be relatively small, and thus the variables fall naturally into the different classes as described in the repeated 2D likelihood network analysis.

7.1.1 Event shape variables

The event shape variables are mostly used for rejection of the backgrounds which are not true four jet hadronic events, *i.e.* the $q\overline{q}(\gamma)$ background. Thus, these variables are independent of the b-tag information, and are therefore gathered in one analysis class.

The $q\overline{q}(\gamma)$ background is originally a two jet configuration, and $q\overline{q}(\gamma)$ events which have been identified as four jet events are therefore events which, due to some perturbation, such as jet splitting, gluon radiation, an ISR photon giving a jet-like detector response or something similar, looks more or less like a four jet hadronic event. Since the four jet structure in this case is created due to some perturbation, one would expect that at least one of the jets would have a less clear hadronic structure, and the three first event shape variables are therefore intended to identify such jets, with less energy, number of tracks and being less clearly separated from other jets.

$\mathbf{Tr_{min}}$: the minimum number of tracks of a jet in a forced four jet configuration.

$E_{\mathbf{min}}$: the minimum energy of a jet in a forced four jet configuration.

$\alpha_{\mathbf{min}}$: the minimum angle between two jets in a forced four jet configuration.

The $q\bar{q}(\gamma)$ background is also topologically different from the structure of the true hadronic four jets, being a more back-to-back topology than the more isotropic four jet topology. Therefore, the next four event shape variables are chosen as variables describing overall features of the distribution of tracks in the event.

Fox-Wolfram moments: The Fox-Wolfram moments $H_l,\ l = 0, 1, 2, 3, \ldots$ [65, page 258] are defined by the relation

$$H_l = \sum_{i,j} \frac{|\vec{p_i}| \cdot |\vec{p_j}|}{E_{vis}^2} P_l(\cos\theta_{ij})$$

where the sum goes over the different particles in the event, $p_i(p_j)$ is the momentum of particle $i(j)$, $\theta_{i,j}$ is the opening angle between particles i and j and E_{vis} is the total visible energy of the event. The $P_l(x)$ is the Legendre polynomial of order l. Normally, the different Fox-Wolfram moments are normalized to H_0, which will only differ from 1 when there are non-negligible masses among the particles in the event. The ones used in the analyses are:

H_2: the second Fox-Wolfram moment, normalized to H_0.

H_4: the fourth Fox-Wolfram moment, normalized to H_0.

M_L: the light jet mass [65, page 258]. When dividing the tracks of an event into two different objects, each object will have an invariant mass given by the energy and momenta of the tracks belonging to that object. When the sum of squares of these two masses is minimized, *i.e.*

$$\overset{\min}{\text{(combination of tracks)}}\ M_1^2 + M_2^2 \overset{\text{def}}{=} M_L^2 + M_H^2$$

the two masses in question are called the light (for the smallest) and heavy (for the largest) jet masses, M_L and M_H respectively.

Event thrust: [65, page 256] The thrust is defined by the relation

$$T = \max_{|\vec{n}|=1} \frac{\sum |\vec{n} \cdot \vec{p_i}|}{\sum |\vec{p_i}|}$$

where the sum goes over the different particles in the event, and p_i is the momentum of track number i. The thrust axis is given by the $\vec{n}$ which maximizes T, which has values ranging from 0.5 to 1.

The characteristics of the Fox-Wolfram moments are such that for two jet events, the H_l tends towards 1 for l even, and towards 0 for l odd. The light jet mass is nominally smaller in a back-to-back event, due to the fact that the particles here generally have momenta in the same direction, making the vector sum of the particle momenta relatively closer to the scalar sum of the particle energies than would be the case in a more isotropic event, and for particle masses small compared to the momenta. The event thrust has values close to the maximum value of 1 for back-to-back events, and close to the minimum value of 0.5 for isotropic events, whereas the value for three jet events is between $\frac{2}{3}$ and 1.

Lastly, a variable is introduced on the jet clustering procedure. This variable is aimed at selecting events where one jet has been split in two, so that the separation between these two jets are less clear than for the other jets. The variable chosen is therefore:

Y_{34}: the y-cut transition value between 3 and 4 jets in the JADE clustering algorithm [65, page 276].

This variable will typically have smaller values for the $q\overline{q}(\gamma)$ background than for the true hadronic four jet events, as it gives a measure of the amount of separation between the least clearly separated jet and the closest of the other jets.

The event shape properties of the signal events are very similar for the H^0Z^0 and the h^0A^0 events, the main difference coming from the fact that the masses of the heavy objects in the event are generally larger in the H^0Z^0 channel than in the h^0A^0 channel. However, the analyses are not intended to be optimized at a specific Higgs boson mass point, but rather cover a relatively large area of the model parameter space, at least within what is reachable at LEP-*II*. Therefore, one would expect that, due to the similar structure of the event shape of the signal in the H^0Z^0 and h^0A^0 channels, the event shape part of the 2D likelihood network should be very close between the two analyses, and the network structure chosen is indeed the same for the two analyses.

The variables in the event shape part of the analysis have been combined to form a final event shape output variable from the repeated 2D likelihood network according to figure 7.1

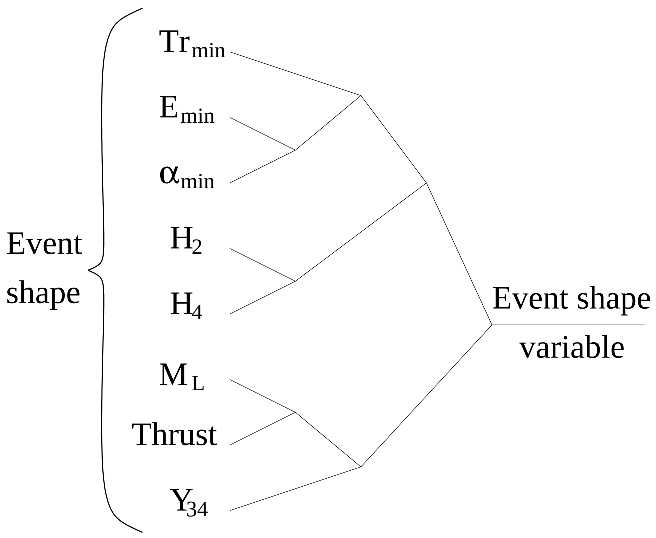


Figure 7.1: The repeated 2D likelihood network combination of the variables in the event shape class. The structure is the same for both the H^0Z^0 and the h^0A^0 search.

7.1.2 B-tag variables

The b-tag is a discriminating variable distinguishing between jets containing b-hadrons and other hadronic jets, and consists of several components.

Jet lifetime probability: [82] Due to the relatively long lifetime of the b-hadrons resulting from b and $\overline{b}$ quarks (typically around 1.5 ps), the b-hadrons will travel a short distance before decaying. The decay products will therefore produce tracks with significantly positive impact parameters, up to ∼1–2 mm.

In ∼50% of the jets from b-hadrons, a secondary hadronic vertex can be identified inside the jet. This corresponds to the place where the b-hadron has decayed, and the information from the tracks originating from this secondary vertex can be used for tagging purposes:

Effective mass of secondary vertex: [83] If an invariant mass is constructed from the track originating from the secondary vertex, this should be equal to the mass of the b-hadron, and thus have higher values than for the lighter quarks.

Rapidity of tracks in secondary vertex: [83] Due to the higher mass and multiplicity of the b-hadrons with respect to the c-hadrons originating from c-quark decays, the rapidity of tracks from secondary vertices originating from b-quarks is on average less than for c-quarks. As for light quarks, the secondary vertices originate in this case mainly from wrong measurements, induced by multiple scattering, interactions in the material etc. Thus, tracks from secondary vertices originating from light quarks are usually soft, shifting the rapidity distribution to lower values.

Charged jet energy fraction in secondary vertex: [83] This is a variable which is shifted towards larger values for b-quarks, than is the case for c and light quarks. The overlap between distributions is, however, rather large, making the discriminating power of this variable rather weak.

High p_t leptons: [84] In semi-leptonic b-hadron decays, the leptons coming from the b-hadron will typically have larger transverse momenta with respect to the jet axis, p_t. This can be used for tagging the b-hadron, provided the decay is semi-leptonic.

The different components of the b-tagging are combined to yield a total b-tagging value. This can be calculated on groups of tracks, jets, and of the event as a whole.

The b-tag variables used in this analysis are the b-tag values for each of the four forced jets in the event, and in addition a total event b-tag:

B_1: the most b-like value for a jet in a forced four jet configuration.

B_2: the second most b-like value for a jet in a forced four jet configuration.

B_3: the third most b-like value for a jet in a forced four jet configuration.

B_4: the fourth most b-like value for a jet in a forced four jet configuration.

B_{HA}: the total event b-tag variable, defined as the sum of $B_i,\ i = 1, \ldots, 4$.

Due to the fact that the analysis method does not take all possible correlations into account (only the ones between the two variables to be combined into one in a combination step), including variables that are simple combinations of other variables in the analysis might contribute discriminating information. This has been seen to be the case for the B_{HA} variable.

7.1.3 χ^2 from the 5C fit

In the cases where the analysis uses a 5C kinematic fit in order to construct the Higgs mass estimator (see section 7.2.3), the χ^2 from the 5C fit

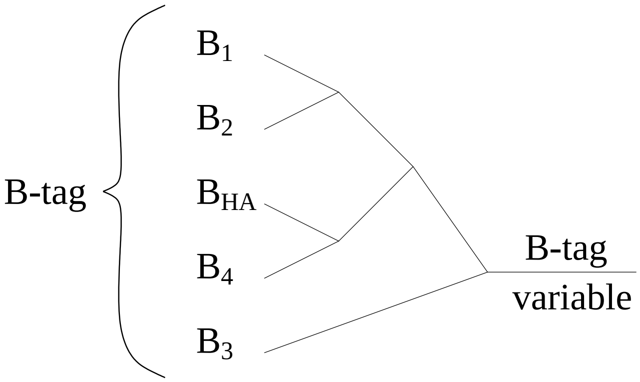


Figure 7.2: The repeated 2D likelihood network combination of the variables in the b-tag class. The structure is the same for both the H^0Z^0 and the h^0A^0 search.

can be used as a variable to discriminate between background and signal. For these analyses, the fifth constraint is assigning a specific mass for the original bosons of the event, either by forcing one of the two bosons to a specific mass (as is the case for the H^0Z^0 analysis), or by assigning a specific mass difference to the bosons (as is the case for the h^0A^0 analysis). In both cases, the χ^2 will be useful as a discriminating variable against backgrounds without two heavy bosons in the original state (*i.e.* the $q\overline{q}(\gamma)$ background) and backgrounds with bosons of masses significantly different from the value compelled by the kinematic fit. This point will be treated more thoroughly in the next section.

7.2 Estimator of the Higgs mass

The analysis has so far been concentrated at constructing a variable where the discrimination between the background and the signal is as good as possible, but without separating different signal hypotheses from each other. Indeed, it has hitherto been an advantage that the discriminating variable has similar properties in all signal hypotheses of interest, since this has enabled the signal statistics to be increased, by adding several different Monte Carlo generated samples of different signal parameter choices together when constructing the different 2-dimensional distributions.

However, when one wants to measure a specific Higgs mass, or, in the absence of a signal, set a lower limit, there are essentially two different characteristics to be taken into account:

1. Measure the cross section of the Higgs production, which is a function of the Higgs mass, and from this measurement determine a value for the Higgs mass, or, in the case of no signal, setting a lower limit. This will be an indirect measurement of the Higgs mass. Due to the relatively low expected cross section where one might expect to discover the Higgs boson, the statistical uncertainty in the cross section measurement will make this aspect rather inconclusive from the point of view of a discovery, but it is an important ingredient in setting a limit in the abscence of a signal.
2. Determine the Higgs mass from direct reconstruction of the invariant mass from each individual event. This will be a direct measurement of the Higgs mass, and will also supply added information to a simple cross section measurement.

When discovering the Higgs boson, it will be an important check of the theoretical prediction that these two characteristics are in accordance with expectations. Also, when setting a lower limit, one should make use of both these characteristics in order to extract maximum information from the available data. With this in mind, it is obvious that there is a strong incentive to have as good a Higgs mass measure from each event as possible. This is achieved by using the method of constrained fits.

7.2.1 Constrained fits, general remarks

A constrained fit consists of the following elements and relations:

- A set of measured quantities, $x_i^m, i = 1, \ldots, n_x$ with corresponding uncertainties σ_{x_i}
- A set of functions on these quantities, $C_j(x_1, x_2, \ldots, x_{n_x}),\ j = 1, \ldots, n_C$
- A different set of quantities $\tilde{x_i}, i = 1, \ldots, n_x$ fulfilling the relations $C_j(\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_{n_x}}) = 0,\ i = 1, \ldots, n_x,\ j = 1, \ldots, n_C$

The goal of the constrained fit is now to find the one out of the (usually infinitely) many possible sets $\tilde{x_i}, i = 1, \ldots, n_x$ which minimizes the function

$$\chi^2 = \sum_{i=1}^{n_x} \left(\frac{\tilde{x_i} - x_i^m}{\sigma_{x_i}} \right)^2$$

Thus, one can view this procedure as the constrained fit seeking to change the measured quantities in such a way as to fulfill the constraints, but changing the quantities as little as possible, relative to their errors, as measured by the χ^2.

In the application of this analysis, the measured quantities are the four-vectors of the four jets in the event ($p_x^i, p_y^i, p_z^i, E^i, i = 1, \ldots, 4$), making the number of measured quantities (*i.e.* the number of degrees of freedom) equal to 16. What the constraints will be, is dependent on the signal and approach to the analysis method, and will be treated in the sections following this one.

Another issue to be taken into account is the so-called pairing problem. This consists of the following: when reconstructing a four jet event originating from two heavy objects each decaying to two quarks, one can cluster the event into four hadronic jets with energies and momentua indicative of the original energies and momenta of the four original quarks. However, one cannot know with certainty which of the two heavy objects a certain hadronic jet belongs to. This means that although one may have a good approximation of the quark impulses, this does not guarantee a good approximation of the invariant masses of the two heavy objects, which is the ultimate goal. However, there are methods, both in the 4C fit and the 5C fit case, to deal with this problem.

7.2.2 4C fit

The 4C fit is a completely model independent, safe and simple way of implementing the constrained fits method to the analysis. The constraints take into account the original, underlying physics process, where an electron of mometum p_e and energy E_{beam} collides head-on with a positron of mometum p_e and energy E_{beam}. Thus, the initial conditions are a system with no total momentum, and energy $2 \cdot E_{beam} = E_{CM}$, which is reflected in the four constraints.

The 4C fit is only used in the h^0A^0 search, and not in the H^0Z^0 search, since there is more information in this search scenario which can be taken into account in a 5C fit (see section 7.2.3).

Constraints

The four constraints are designed to preserve the total momentum and energy in the event, and are defined as follows:

- $\sum_{i=1}^{4} \tilde{p}_x^i = 0$
- $\sum_{i=1}^{4} \tilde{p}_y^i = 0$
- $\sum_{i=1}^{4} \tilde{p}_z^i = 0$

- $\sum_{i=1}^{4} \tilde{E}^i = 2 \cdot E_{beam} = E_{CM}$

Pairing

The pairing is chosen by looking at all three possibilities (jet 1 coupled to jet 2, jet 3 or jet 4, and the other coupling being trivially given by the two remaining jets) and choosing the one with the smallest absolute difference between the invariant masses of the two objects. The mass estimator is now chosen as the sum of the invariant masses of the two heavy objects; this because the sum is better determined than any of the two individual values, due to the fact that if one of the invariant masses is too low (or high), the other will be likely to be too high (or low), making the sum a more correct measure.

Advantages

- The method is simple and intuitive, and only aimed at correcting detector imperfections.

- It introduces no hypothesis dependence, apart from the conservation of energy and momentum, which is a very non-controversial assumption and holds for all signals under study.

Disadvantages

- The method is not optimal, as there is information present in the signal hypothesis which is not taken into account (see section 7.2.3).

- For certain signal hypotheses where Δm ($\Delta m = m_{A^0} - m_{h^0}$) is large, the chosen pairing will not be the correct one, since the two heavy objects in the event have mass difference significantly different from zero.

7.2.3 5C fit

The 5C fit attempts to take more information from the expected signal hypothesis into account, and thus improve the resolution of the invariant mass of the dijet corresponding to the supposed Higgs boson of the event. The fit uses the four constraints already described in the 4C case, and in addition a constraint aimed at the invariant mass of one or both of the dijet objects in the event.

The 5C fit is the standard choice in the H^0Z^0 channel, and can also be used in the h^0A^0 channel to improve the resolution of the reconstructed Higgs mass, albeit by the use of a slightly controversial method, as described below.

Constraints in the H^0Z^0 case

The fifth constraint in the H^0Z^0 case is that the invariant mass of a specified one of the two dijet objects is to be equal to the Z^0 mass:

- $M_{inv}^{1,2}$=91.19 GeV/c^2

Constraints in the h^0A^0 case

The fifth constraint in the h^0A^0 case is that the invariant masses of the two dijet objects are to have a definite difference, given by the MSSM parameters for the specific signal hypothesis in question:

- $M_{inv}^{1,2} - M_{inv}^{3,4}$=$\Delta m$ (given by the MSSM signal hypothesis)

For simplicity, and in order to be able to implement this scheme in the analysis, the value of Δm is rounded off to its nearest integer, with a maximal value of 30 GeV/c^2.

Pairing

The pairing is chosen by looking at all six possibilities (jet 1 coupled to jet 2, jet 3 or jet 4, and the other coupling being trivially given by the two remaining jets, with all three possibilities having a second choice as of which dijet object is to be assigned to the Z^0 in the H^0Z^0 case, or the heaviest of the h^0 and A^0, which in most areas of the MSSM parameter space will be the A^0, in the h^0A^0 case), which will give six different values of the χ^2 from the 5C fit.

Advantages

- The method gives an improved mass resolution compared to that of the 4C fit method, since more information from the expected signal events is taken into account.

- The choice of pairing comes directly from the fit (in the SM case coupled with b-tag information, as described in section 7.3), and gives the nominally correct choice of pairing for all signal hypotheses, as opposed to the 4C fit method, which only gives a correct choice of pairing for signal hypotheses where the two heavy objects in the event are close in mass.

- The χ^2 from the 5C fit can be used as a discriminating variable, mostly against the $q\overline{q}(\gamma)$ background, since this is the background which does not contain two heavy objects in the event, and which therefore would be expected to have a higher value of the χ^2.

Disadvantages

The disadvantages of the 5C fit method appear only in the h^0A^0 channel, and all originate from the fact that the method is no longer independent of signal hypothesis, *i.e.* that the treatment of the candidate events depend on parameters of the MSSM, more specifically the difference between the mass of the A^0 and the h^0. Thus, the list below applies only to the h^0A^0 analysis, and no specific problems are envisioned in the H^0Z^0 channel.

- The method introduces signal hypothesis dependencies, which is a somewhat undesirable feature.

- Background distributions differ between different signal hypotheses, and a large system of "book-keeping" is required to keep track of the different distributions for both the background and the different signal hypotheses.

- Candidates (and their choices of pairing) in data become signal hypothesis dependent, and it is therefore difficult to assess what signal hypothesis best fits specific data candidates. Even though this is not a fundamental problem, but one of a more aesthetic nature, it is still problematic when presenting the actual data results.

- In the method used for extracting the lower limit on the Higgs boson masses [85], the likelihood ratio of $\frac{\mathcal{L}(s+b)}{\mathcal{L}(b)}$ is used. However, the denominator of this expression is to be evaluated on a signal-free background, which becomes somewhat fundamentally problematic for a method where the background is signal hypothesis dependent. This problem is not of immediate worry in the case of setting a limit on the Higgs boson mass, but in the context of a discovery this will immediately create problems.

7.2.4 Choice of mass estimation method

The discussion in the preceding sections has given a basis for choosing the method for constructing the Higgs mass estimator best suited to the two different analyses.

In the H^0Z^0 channel, the 5C fit with the choice of pairing being determined from the χ^2 of the fit together with the b-tag information as described in section 7.3 is used.

In the h^0A^0 channel, both the standard 4C fit, as described in section 7.2.2, and the new method of the 5C fit, as described in section 7.2.3, are presented.

7.3 The H^0Z^0 analysis

The total analysis tree for the discriminating variable in the H^0Z^0 channel is given in figure 7.3. This analysis is very similar to the h^0A^0 analysis, using the same input variables and network structure, the difference being the expected distributions from the signal. This is, as already mentioned, mainly concentrated in the b-tag part of the analysis, as there are now only two guaranteed b-quarks in the signal events.

In the H^0Z^0 channel, the χ^2 is modified with the probability that the two jets opposite of the ones assigned to the Z^0 have b-tag values compatible with coming from the H^0. This is done because the jets with high b-tag values are more likely to originate from the H^0 than from the Z^0, since the branching ratio of H^0 into the $b\bar{b}$ topology is much larger (*i.e.* $\gtrsim$85% for most MSSM points of interest) than that of the Z^0 into $b\bar{b}$ ($\sim$20% of the $Z^0 \to q\bar{q}$ channel). The final expression for the pairing selection function is the following:

$$\mathcal{P}_b^{j_1}\mathcal{P}_b^{j_2}\left((1-R_b-R_c)\,\mathcal{P}_q^{j_3}\mathcal{P}_q^{j_4}+R_b\mathcal{P}_b^{j_3}\mathcal{P}_b^{j_4}+R_c\mathcal{P}_c^{j_3}\mathcal{P}_c^{j_4}\right)\mathcal{P}^{5C}_{j_3,j_4} \tag{7.1}$$

where $\mathcal{P}_b$, $\mathcal{P}_c$ and $\mathcal{P}_q$ are the probability distribution functions for the b-tag value of a jet originating from b, c and light quarks, respectively, R_b, R_c and R_q are the branching fractions of the Z^0 going to a $b\bar{b}$, $c\bar{c}$ and light quark/anti-quark pair, respectively, and $\mathcal{P}^{5C}_{j_3,j_4}$ is the χ^2 probability of the 5C fit, forcing jets j_3 and j_4 to the Z^0 mass. Thus, the pairing maximizing this function is chosen, and the Higgs mass estimator is calculated as the invariant mass of the two jets j_1 and j_2.

For Higgs masses approaching the kinematical limit (*i.e.* $E_{CM}-m_{H^0}-m_Z$ is small), the most probable value of the invariant mass of the quark/anti-quark pair originating from the Z^0 will no longer be at the nominal Z^0 mass

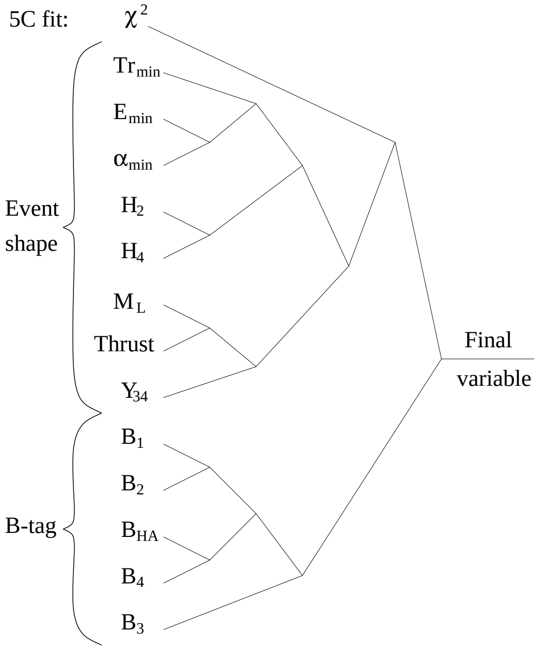


Figure 7.3: The repeated 2D likelihood network combination of the variables in the H^0Z^0 search.

of 91.19 GeV/c^2, due to the reduction in the available kinematical phase space. Thus, for searches where the final result is close to the kinematical edge, as is the case for the SM search, forcing the two jets opposite the Higgs boson to the nominal Z^0 mass is not entirely optimal. However, for a search intended to cover large areas of MSSM parameters, which is the overall goal of the work presented here, the limits for large ranges of MSSM parameter choices will be sufficiently far from the kinematical edge that the invariant mass of the quark/anti-quark system is centred at the nominal Z^0 mass. Therefore, the procedure described in the previous paragraph should be followed in order to use the same procedure for all signal hypotheses, albeit at the expense of a slightly suboptimal SM Higgs mass limit.

A plot showing the H^0Z^0 Higgs mass estimator at the four jet selection level for two different Higgs mass hypotheses is shown in figure 7.4.

7.4 The h^0A^0 analysis

As has already been mentioned, the h^0A^0 channel is presented in two different scenarios, with either a 4C fit or a 5C fit. Both the discriminating variable

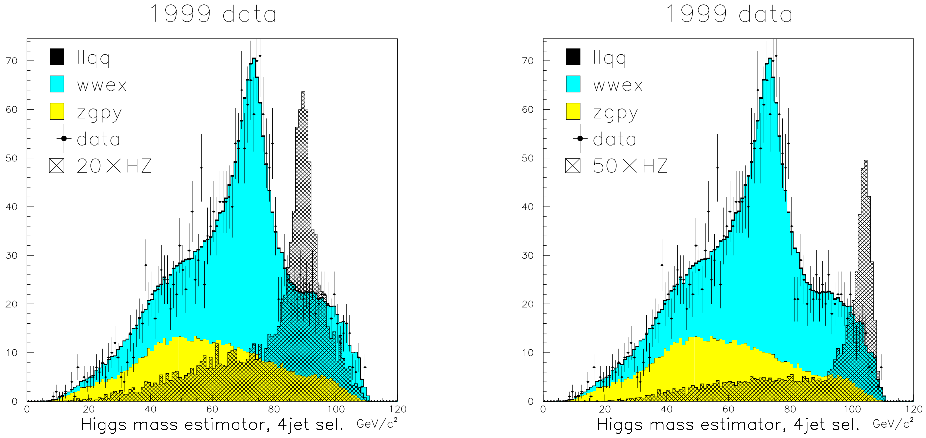


Figure 7.4: Plots showing the Higgs mass estimator (the invariant mass of the dijet opposite the one forced to the Z^0 mass) at the four jet selection level for the H^0Z^0 analysis. Two different Monte Carlo generated signals are plotted: on the left is shown the m_{H^0}=90 GeV/c^2 signal, and on the right the m_{H^0}=105 GeV/c^2 signal is shown. Both plots show the sum of the four centre-of-mass energy points for the 1999 data.

analysis part and the mass estimator are different in the two variants. In the h^0A^0 channel, the h^0 and A^0 have very close to equal branching ratios into the $b\bar{b}$ topology, at least in the MSSM parameter region of interest for this search, and there is therefore no information from the b-tag values as to which jets belong to the h^0 and A^0, respectively. For this reason, one must use certain assumptions about the signal in order to choose the pairing.

7.4.1 4C fit analysis

The total analysis tree for the discriminating variable in the h^0A^0 channel in the 4C fit mode is given in figure 7.5.

For most parts of the MSSM parameter region of interest for this search, the h^0 and A^0 are fairly close in mass. This is even more true when one considers the parts of the MSSM regions where the h^0A^0 production dominates over the H^0Z^0 production (*i.e.* large values of $\tan\beta$), and the pairing is therefore chosen as the one which minimizes the difference of the invariant mass of the two dijet-objects. The mass estimator is then chosen as the sum of the invariant masses of these two dijet objects. A plot showing the h^0A^0 Higgs mass estimator at the four jet selection level for two different m_{A^0}, $\tan\beta$ hypotheses in the 4C fit analysis case is shown in figure 7.6.

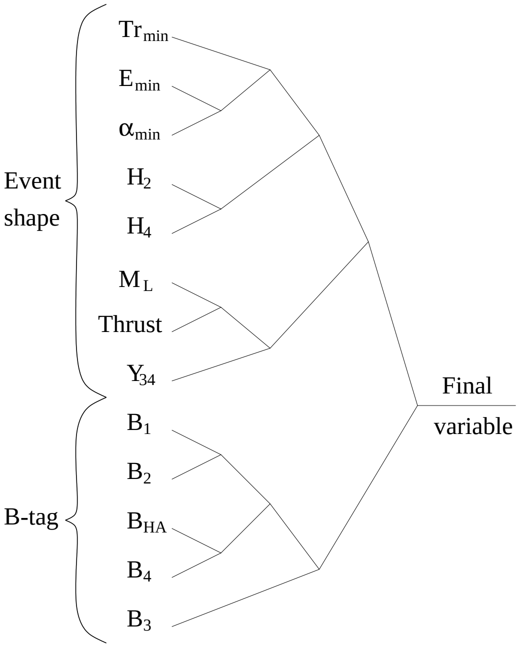


Figure 7.5: The repeated 2D likelihood network combination of the variables in the h^0A^0 search in the 4C fit case.

7.4.2 5C fit analysis

The total analysis tree for the discriminating variable in the h^0A^0 channel in the 5C fit mode is given in figure 7.7.

In the 5C fit method, the pairing is taken into account by forcing the mass difference between the two dijet objects to a certain value, determined by the MSSM parameter assumption. Thus, the agreement of a particular pairing with the underlying MSSM parameter assumption is expressed directly in the χ^2 value from the 5C fit. Therefore, the pairing is chosen in the simple way of selecting the one pairing out of the six posibilities with the minimal χ^2 from the fit. The mass estimator is still chosen as the sum of the invariant masses of the two resulting dijet objects. A plot showing the h^0A^0 Higgs mass estimator at the four jet selection level in the 5C fit analysis case for the same two m_{A^0}, $\tan\beta$ hypotheses as in figure 7.6 is shown in figure 7.8.

When comparing the methods of 4C fit and 5C fit, the plots in figures 7.6 and 7.8 shows much of the effect. By comparing the right plots in the two figures (corresponding to the m_{A^0}=85 GeV/c^2, $\tan\beta$=20 signal) it is apparent that the difference between the two methods is not very large. The signal is slightly more peaked for the 5C fit method and the the signal tail down

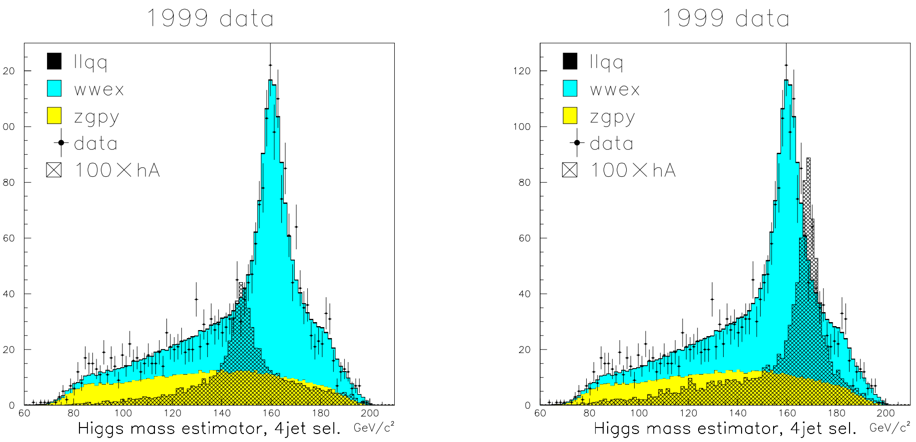


Figure 7.6: Plots showing the Higgs mass estimator (the sum of the invariant masses of the dijets, with the pairing chosen to minimize the dijet mass difference) at the four jet selection level for the h^0A^0 analysis with the 4C fit mass estimator method. Two different Monte Carlo generated signals are plotted: on the left is shown the m_{A^0}=85 GeV/c^2, $\tan\beta$=2 signal (corresponding to $m_{h^0}\sim$65 GeV/c^2) , whereas the m_{A^0}=85 GeV/c^2, $\tan\beta$=20 signal (corresponding to $m_{h^0}\sim$85 GeV/c^2) is shown on the right. Both plots show the sum of the four centre-of-mass energy points for the 1999 data.

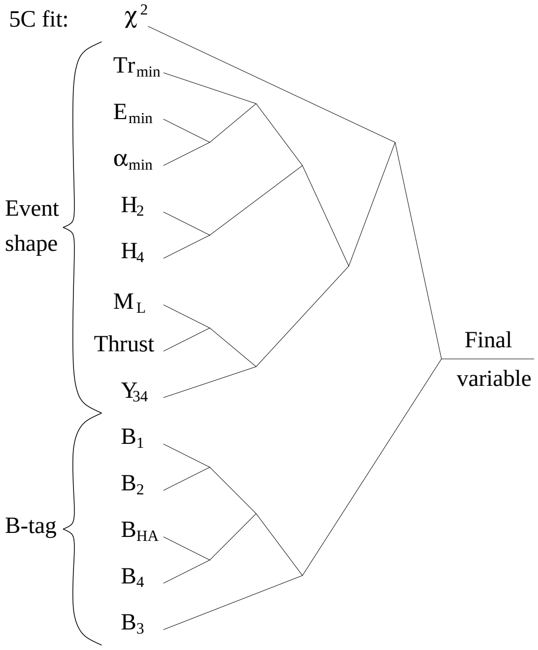


Figure 7.7: The repeated 2D likelihood network combination of the variables in the h^0A^0 search in the 5C fit case.

to lower masses is reduced with respect to the 4C fit method, but the tail at larger masses above the signal peak itself is slightly enhanced. It is also noteworthy that the Z^0Z^0 peak at 180 GeV/c^2 in the expected background is clearly visible for in 5C fit case, whereas this is obscured by the other 4-fermion background in the WWEX sample for the 4C fit method.

The left plots in the two figures show larger differences. The signal is peaked significantly more in the 5C fit case in comparison to the 4C fit case, and the tail at lower masses is correspondingly reduced. But the large difference is represented by the background distribution, which is significantly less peaked for the 5C fit case. This is explained by the fact that Δm in the 5C fit now is 20 GeV/c^2, and therefore does not correspond to the situation in the 4-fermion backgrounds, where the dijet-objects (W^+W^- or Z^0Z^0) are usually of equal mass. Therefore, the 5C fit will frequently force the masses of the W^+W^- and Z^0Z^0 events to incorrect values or pick the wrong pairing. At this point it is worth noting that this large difference between the 4C fit and 5C fit methods appears at low values of $\tan\beta$, where the H^0Z^0 process dominates over the h^0A^0 process, and the importance of this difference is therefore decreased in terms of the final Higgs mass limits. Therefore, even though the 5C fit method has obvious advantages over the 4C fit method, it

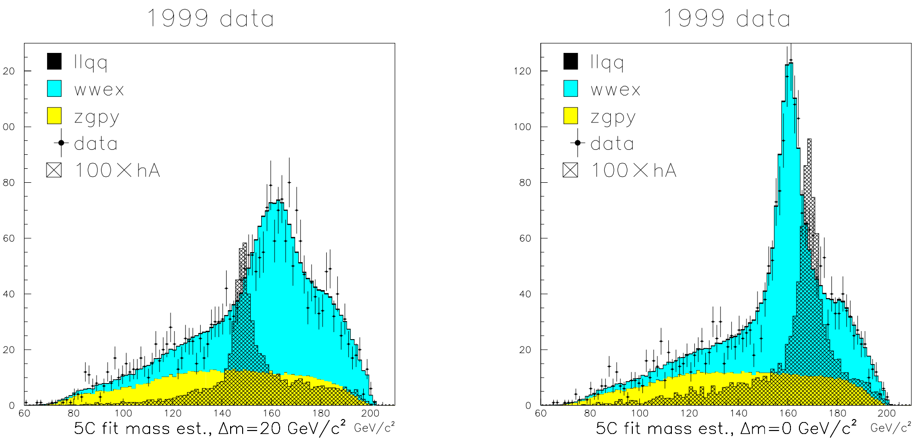


Figure 7.8: Plots showing the Higgs mass estimator (the sum of the invariant masses of the dijets, with the pairing chosen from the χ^2 of the 5C fit) at the four jet selection level for the h^0A^0 analysis with the 5C fit mass estimator method. Two different Monte Carlo generated signals are plotted: on the left is shown the m_{A^0}=85 GeV/c^2, $\tan\beta$=2 signal (corresponding to $m_{h^0}\sim$65 GeV/c^2, *i.e.* $\Delta m = 20$ GeV/c^2 in the fit), whereas the m_{A^0}=85 GeV/c^2, $\tan\beta$=20 signal (corresponding to $m_{h^0}\sim$85 GeV/c^2, *i.e.* $\Delta m = 0$ GeV/c^2 in the fit), is shown on the right. Both plots show the sum of the four centre-of-mass energy points for the 1999 data.

will be shown in section 8.4 that this advantage does not transform into a measurable effect on the final MSSM parameter limits.

7.5 Input variables for the analyses

The input variables for the analyses, with the exception of the thrust, which has already been shown in the four jet selection, are shown in figures 7.9 to 7.14, divided into the three years of data taking and showing the expected background and signal compared to the observed data. In addition to this, figures 7.15 and 7.16 show the same information, in arbitrary normalization, for four different Monte Carlo generated signals. The variables are shown for events passing the four jet selection, which is the level at which the construction of the discriminating variable is performed.

These plots show generally good agreement between expected background and data. The most prominent feature in terms of disagreement is an excess of data with respect to the Monte Carlo background at moderately large values of b-tag variables for the 1998 data (figure 7.10). This discrepancy, which does not appear at the highest b-tag levels, has been extensively studied by DELPHI, and is believed to be understood. In any case, as it does not appear in the most signal-like region, the effect is not crucial to the analyses.

In order to check that each input variable contributes information into the final analysis steps, a systematic procedure was adopted for the h^0A^0 analysis in the 4C fit case (the assumption being that the H^0Z^0 signal is rather similar, and that variables relevant for the h^0A^0 analysis apply also to the h^0Z^0 analysis). In turn, each of the 13 input variables in the analysis was removed and the analysis performed, resulting in each case in a final b-tag variable or event shape variable (depending on whether the removed variable is part of the b-tag or event shape class of variables) slightly different from the original. As the signal distribution is by definition flat (see section 6.3), such comparisons are fairly straight-forward, as all the information on the quality of the analysis lies in the background distribution (the more peaked towards lower values, the better the analysis). The relevant final variable distributions from these 13 different analyses are shown in figures 7.17 to 7.19.

The difference between the distribution from the original analysis and the one with one variable removed can in most cases be seen, although some variables contribute only a rather small improvement. In the few cases where removing the variable does not measurably worsen the background distribution, the variables are included, even though they contribute no overall analysis gain, in order to remain consistent with other DELPHI analyses, where these variables are used. An example of such a variable is the second

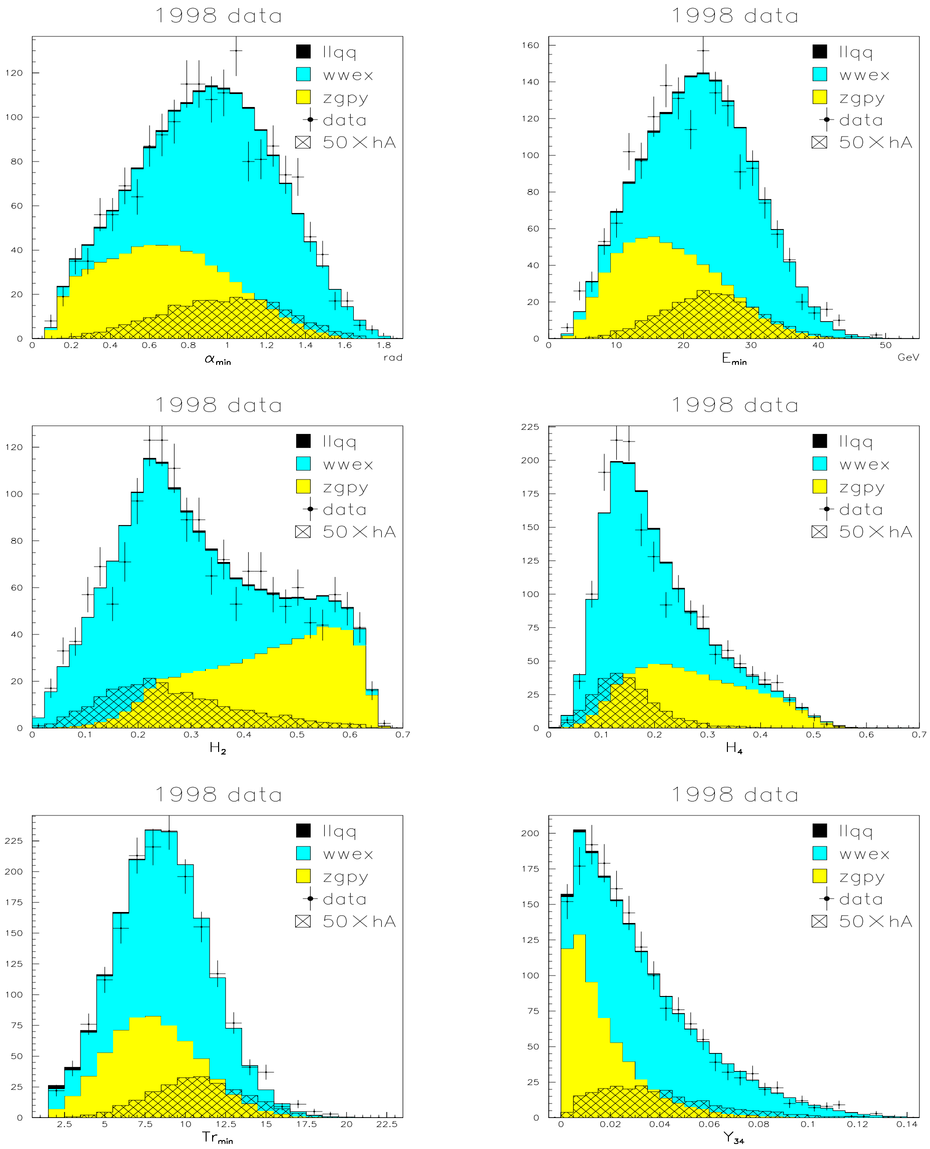


Figure 7.9: Plots of the first six analysis variables for the 1998 data set. These are the minimum inter-jet angle, the minimum jet energy, the second Fox-Wolfram moment, the fourth Fox-Wolfram moment, minimum jet multiplicity and the y-cut transition value from three to four jets. The Monte Carlo generated signal corresponds to the m_{A^0}=85 GeV/c^2, tan β=20 signal.

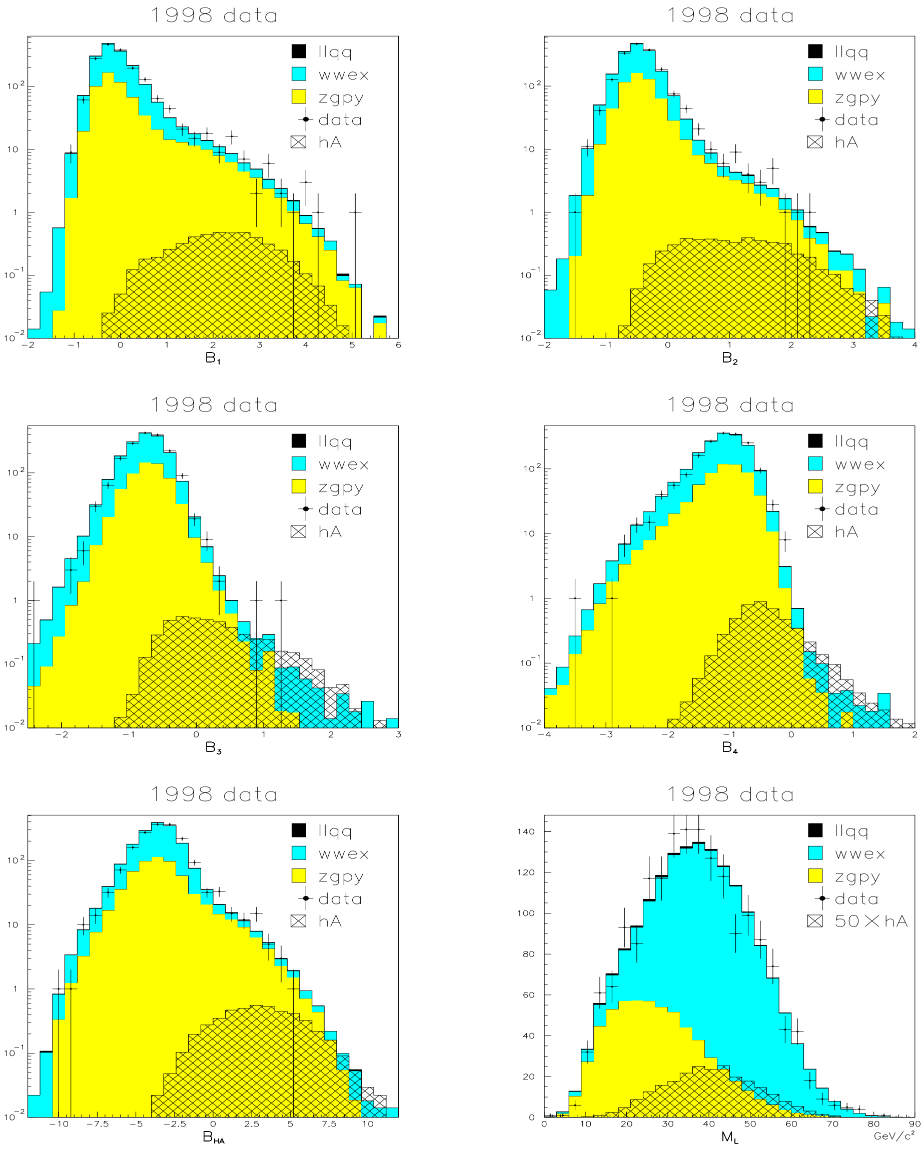


Figure 7.10: Plots of the last six analysis variables for the 1998 data set. These are the best, second best, third best and fourth best jet b-tag values, the total b-tag value and the light jet mass. The Monte Carlo generated signal corresponds to the m_{A^0}=85 GeV/c^2, $\tan\beta$=20 signal.

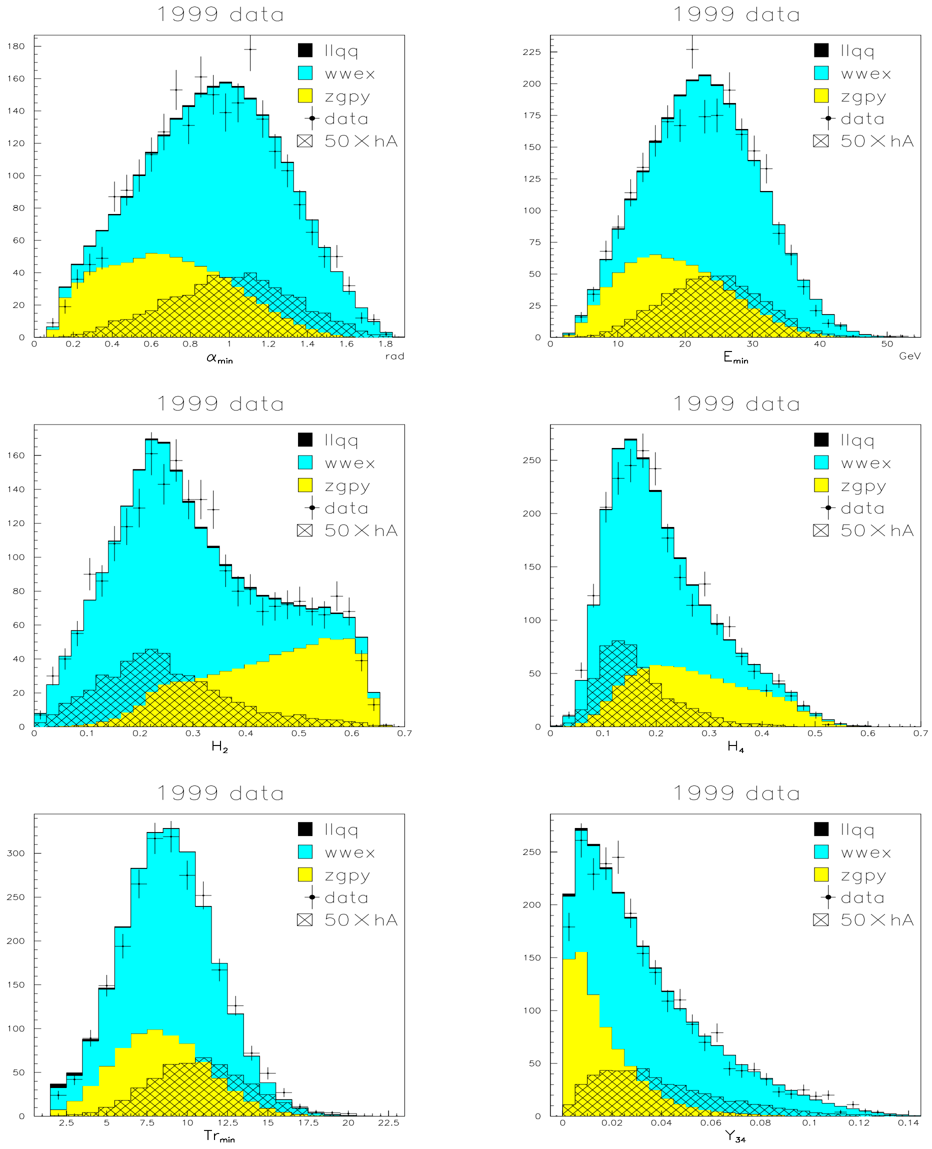


Figure 7.11: Plots of the first six analysis variables for the 1999 data set. The variables are the same as in figure 7.9

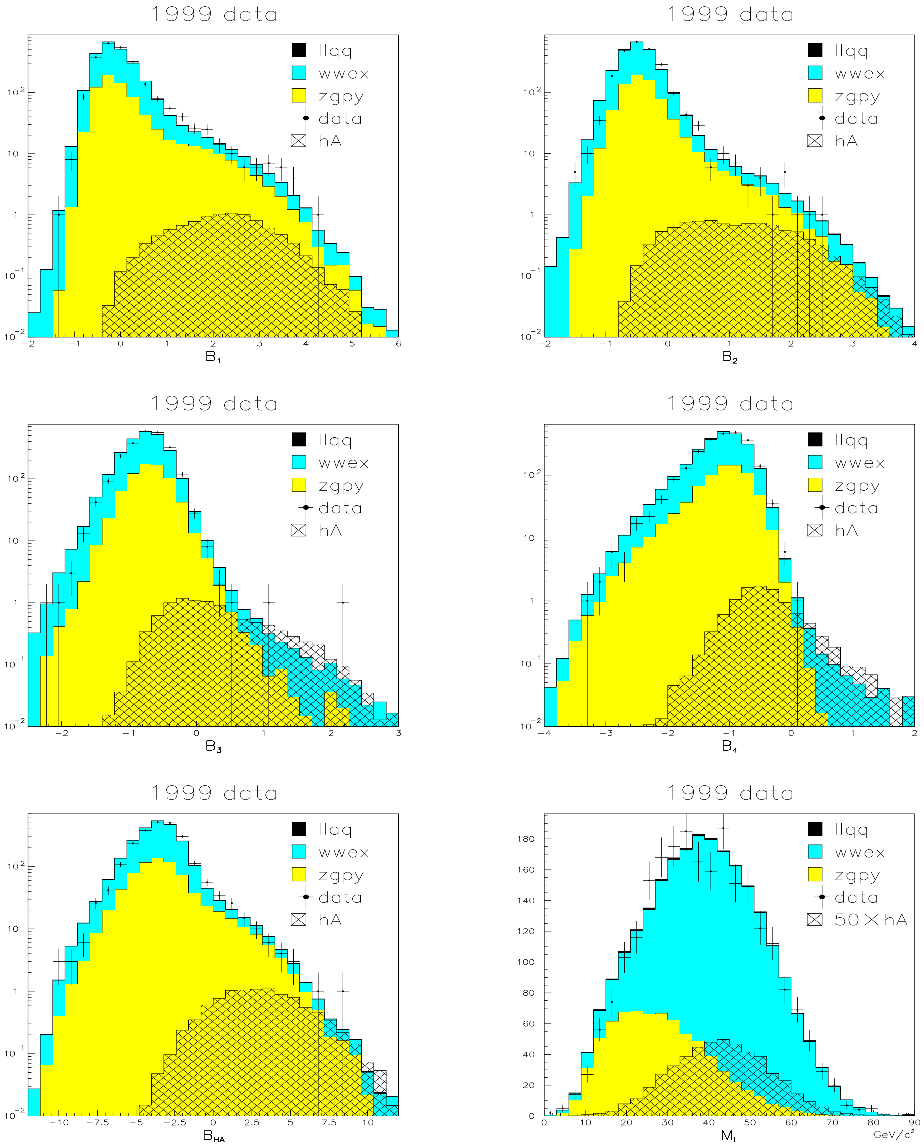


Figure 7.12: Plots of the last six analysis variables for the 1999 data set. The variables are the same as in figure 7.10

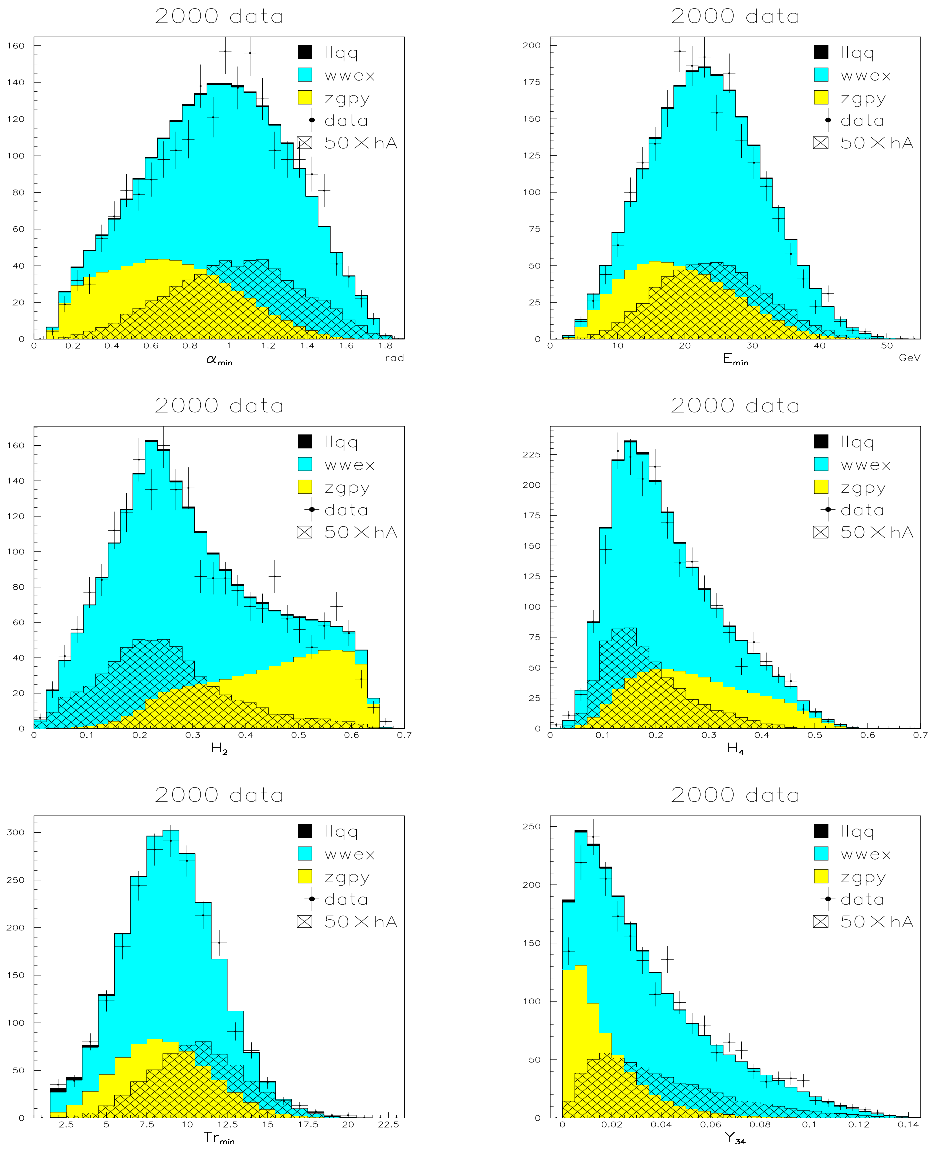


Figure 7.13: Plots of the first six analysis variables for the 2000 data set. The variables are the same as in figure 7.9

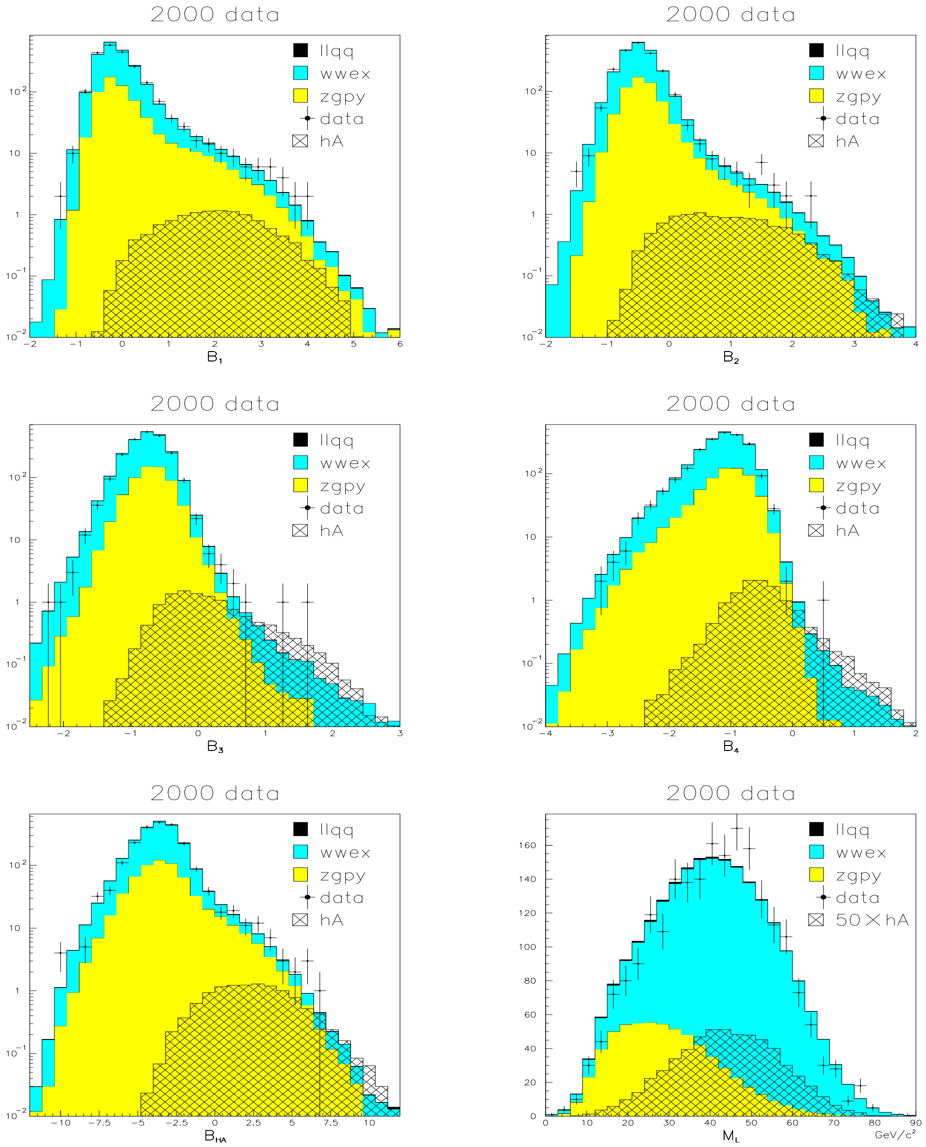


Figure 7.14: Plots of the last six analysis variables for the 2000 data set. The variables are the same as in figure 7.10

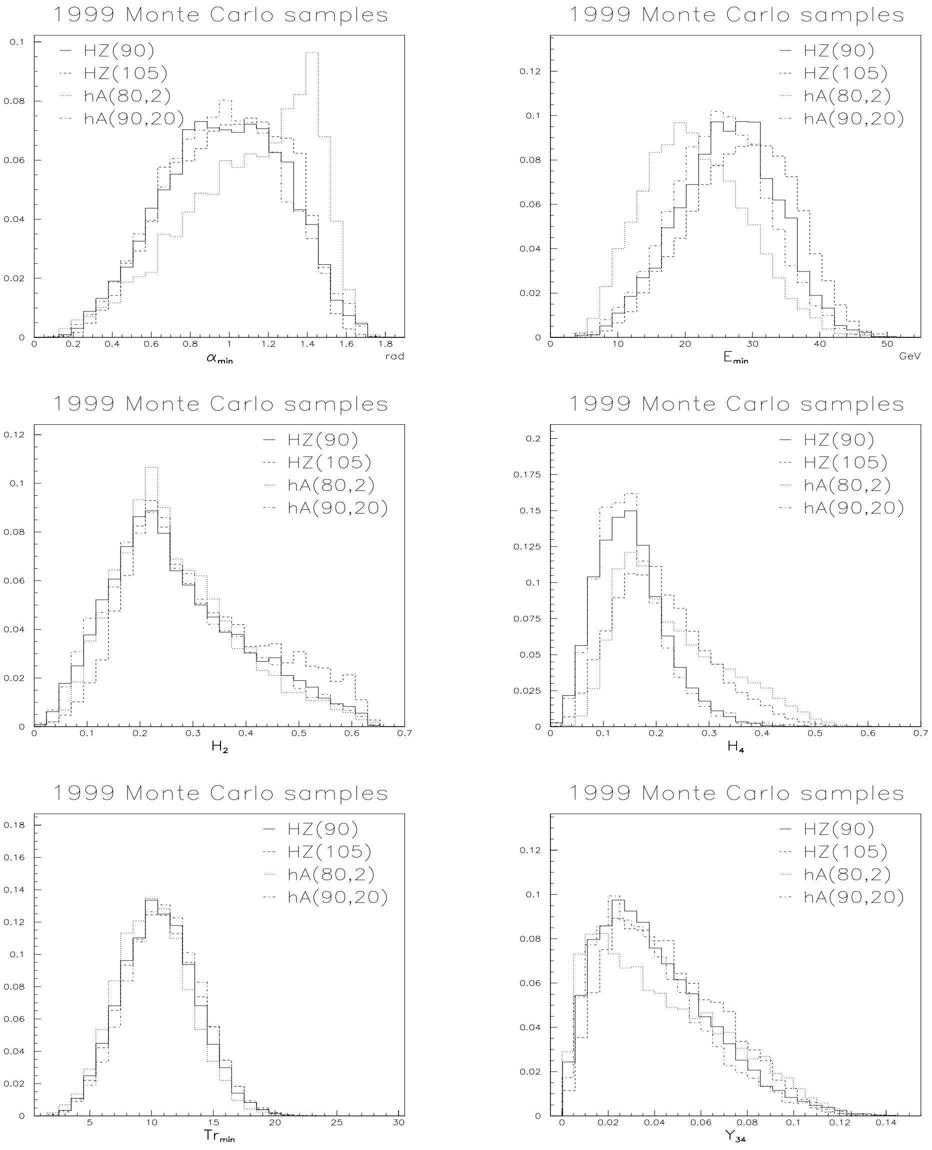


Figure 7.15: Plots of the last six analysis variables for different Monte Carlo generated signals. These are for the h^0A^0 case the m_{A^0}=80 GeV/c^2, $\tan\beta$=2 and m_{A^0}=90 GeV/c^2, $\tan\beta$=20, and for the H^0Z^0 case the m_{H^0}=90 GeV/c^2 and m_{H^0}=105 GeV/c^2. The plots correspond to the sum of all generated signal Monte Carlo for these four signal hypotheses at the four different centre-of-mass points for the 1999 data set. The variables are the same as in figure 7.9

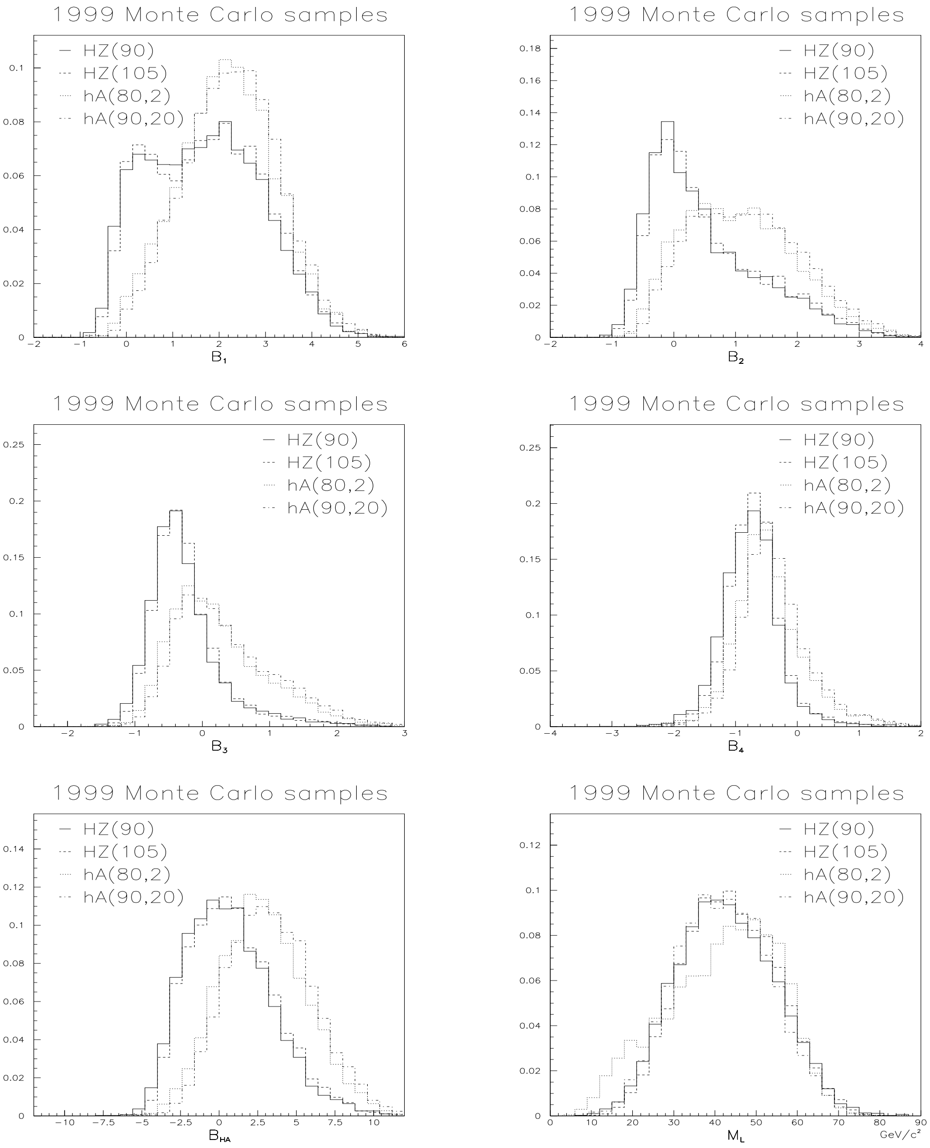


Figure 7.16: Plots of the last six analysis variables for different Monte Carlo generated signals. These are for the h^0A^0 case the m_{A^0}=80 GeV/c^2, $\tan\beta$=2 and m_{A^0}=90 GeV/c^2, $\tan\beta$=20, and for the H^0Z^0 case the m_{H^0}=90 GeV/c^2 and m_{H^0}=105 GeV/c^2. The plots correspond to the sum of all generated signal Monte Carlo for these four signal hypotheses at the four different centre-of-mass points for the 1999 data set. The variables are the same as in figure 7.10

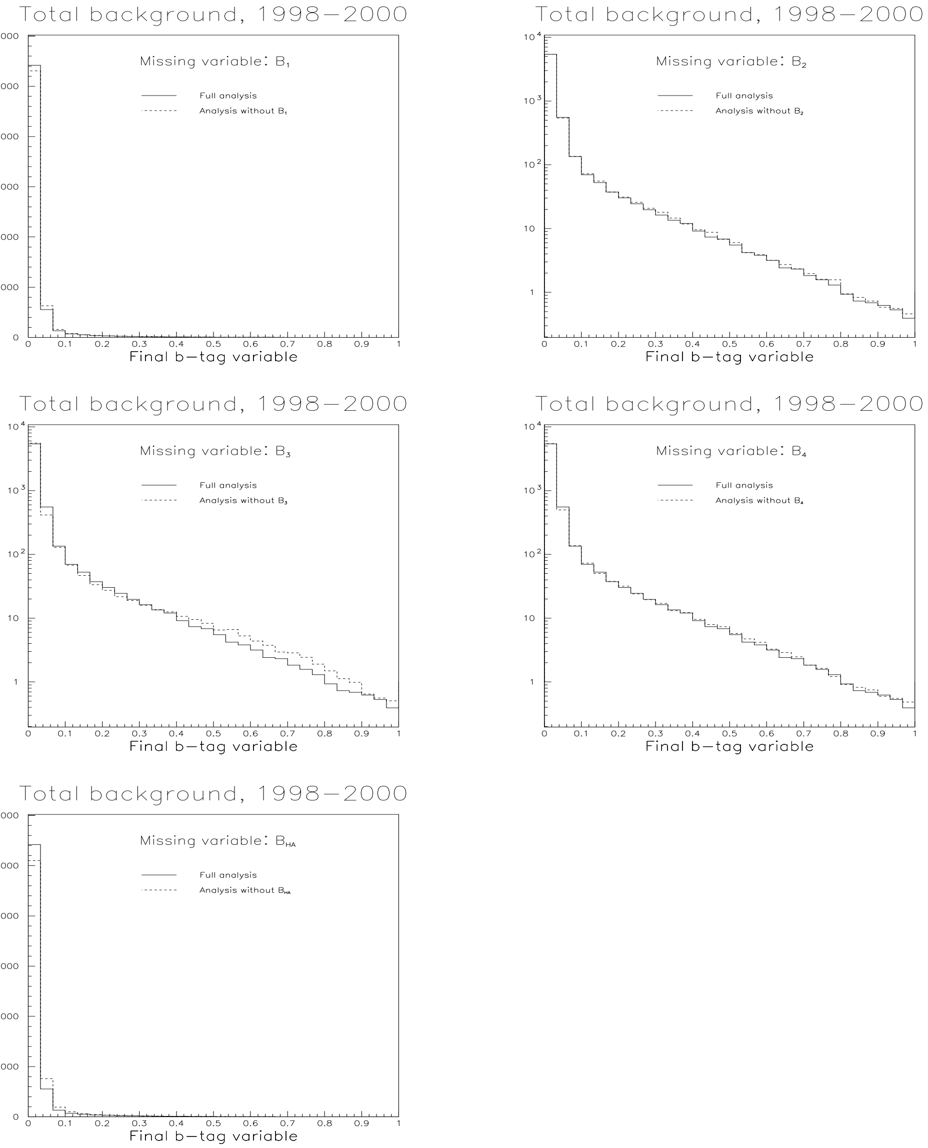


Figure 7.17: Plots of the final b-tag variable for the five h^0A^0 analyses with each of the b-tag variables removed, and comparisons with the final b-tag variable for the original h^0A^0 analysis. The distributions shown correspond to the total background for all centre-of-mass energy points. The original background is shown in the solid histogram, whereas the distribution with the variable in question removed from the analysis is shown in dashed. The choice of linear or logarithmic scale is made in order to best highlight the difference between the two distributions.

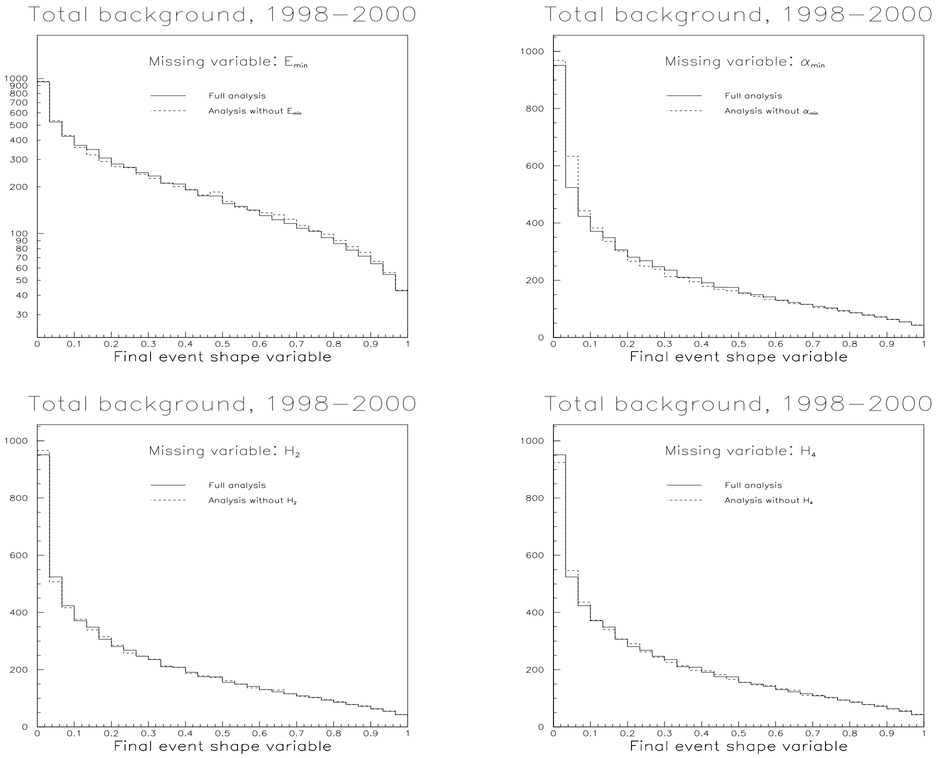


Figure 7.18: Plots of the final event shape variable for the four h^0A^0 analyses with each of the first four event shape variables removed, and comparisons with the final event shape variable for the original h^0A^0 analysis. The distributions shown correspond to the total background for all centre-of-mass energy points. The original background is shown in the solid histogram, whereas the distribution with the variable in question removed from the analysis is shown in dashed. The choice of linear or logarithmic scale is made in order to best highlight the difference between the two distributions.

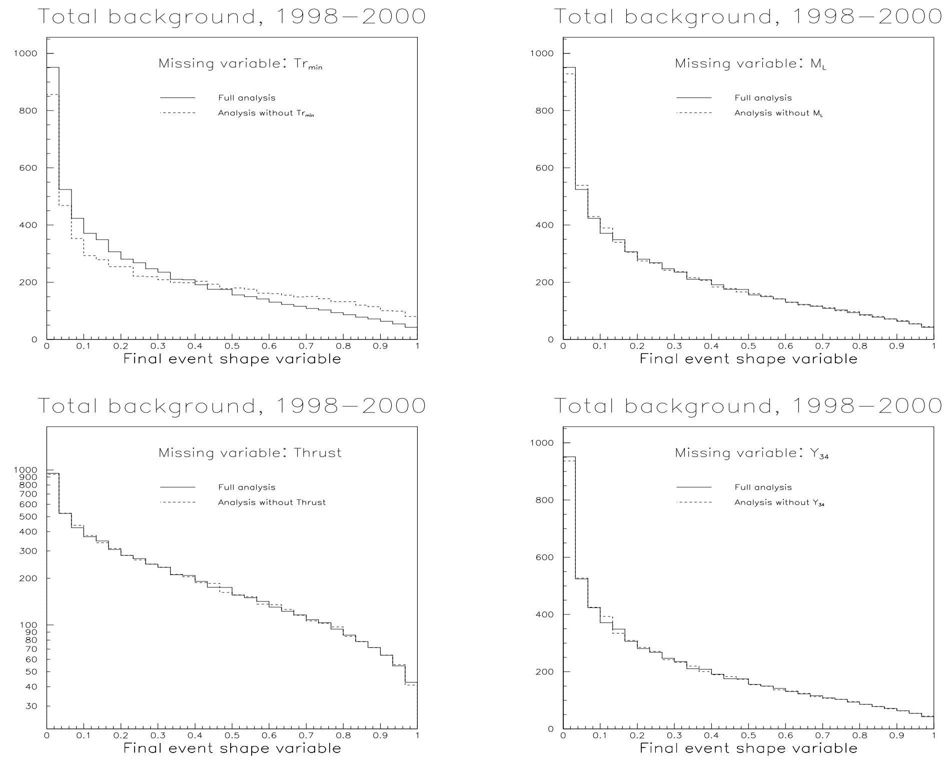


Figure 7.19: Plots of the final event shape variable for the four h^0A^0 analyses with each of the last four event shape variables removed, and comparisons with the final event shape variable for the original h^0A^0 analysis. The distributions shown correspond to the total background for all centre-of-mass energy points. The original background is shown in the solid histogram, whereas the distribution with the variable in question removed from the analysis is shown in dashed. The choice of linear or logarithmic scale is made in order to best highlight the difference between the two distributions.

Fox-Wolfram moment, H_2, which does not contribute discriminating information beyond that which is contributed by the fourth Fox-Wolfram moment, H_4. In other DELPHI analyses, the sum of these two variables is often used, whereas the analysis technique presented in this work provides a method for checking whether this is the optimal way of combining the two variables, and each of the two variables were therefore used on its own.

Another aspect is that even though the inclusion of a new variable does not on average improve the analysis, it might still contain information. Given the crucial problem of overtraining, and the consequent necessity of smoothing the different distributions before using them in the analysis, one must expect a little discriminating power to be lost in each combination step in the analysis, something which will be counter-acted by the discriminating information in the variable to be added. If these two effects are almost equally large, the net result will of course be that the analysis does not on average improve. However, there is still the possibility that single events which have extreme values in the added variable will have a significantly different final discriminating variable value in two analyses whose only difference is the inclusion of the variable in question for one of the analyses. Such effects are most important for data candidates, where one wants to assure oneself of the validity of each candidate as genuinely Higgs-like. Thus, the inclusion of variables contributing only marginal new discriminating information can still be used as a way of increasing the probability that only very signal-like data candidates are recognized as such by the analysis.

An illustration of this is provided by table 7.1, showing three actual data candidates from two different h^0A^0 analysis with 4C fit, the first with B_1 in the analysis tree (*i.e.* the original analysis) and the second without B_1 in the analysis tree (*i.e.* the variable B_1, which is a variable contributing only very slightly to the overall analysis performance, is removed from the analysis tree, whereas the rest of the analysis remains unchanged). The final discriminating variable for the first two of these events move significantly (a factor >2 and a factor >10, respectively) when including B_1 in the analysis, but both at a value below the final cut value of 0.15 (see section 7.8). The last event, on the other hand, moves from a value below the cut value and well into the region of events which survive to the final statistical treatment, thus affecting the final result. This shows that the inclusion of B_1 in the analysis does have a measurable effect, even though this will on average, for the expected result, be very small.

	Analysis with B_1		Analysis without B_1	
B_1	Final btag var.	Final disc. var.	Final btag var.	Final disc. var.
4.28	0.091	0.103	0.044	0.048
3.00	0.058	0.037	0.001	0.003
4.07	0.152	0.198	0.123	0.146

Table 7.1: Table showing the value of B_1 and the corresponding final btag and discriminating variable for both the original h^0A^0 analysis with 4C fit, and for the analysis with B_1 removed from the analysis tree. Three actual data events are shown, where the inclusion of B_1 changes the final btag and discriminating variable significantly, despite the fact that the overall effect of including B_1 in the analysis tree is very small if not negligible.

7.6 Overtraining studies

In order to study possible overtraining problems originating from inaccurate approximation of the 2-dimensional distributions, the Monte Carlo generated background and signal samples were divided in two halves. The analysis network, *i.e.* the total transformation, defined by the 2-dimensional histograms and the method described in chapter 6, of the input variables into the final discriminating variable, was constructed using only the first half. The analysis was then performed on both halves, giving a final discriminating variable distribution for background and signal for both halves of the sample. In the presence of overtraining, one would expect the performance of the first half to be better than that of the second half.

When performing a cut at a specific value in the final discriminating variable, one finds a certain background level corresponding to a certain signal efficiency. If such cuts are performed subsequently tighter and tighter in the final variable distribution, a set of corresponding levels of background and signal efficiency is amassed, and by plotting these as points on a curve, the total performance of the analysis can be envisaged. For the overtraining studies, the difference in the signal efficiency between the first and the second half of the Monte Carlo generated signal samples can be plotted as a function of the background level, which in the case of overtraining would make up a curve lying significantly above the abscissa. Such plots are shown for the H^0Z^0 analysis at all ten centre-of-mass energy windows in figure 7.20, and for the h^0A^0 analysis with 4C fit Higgs mass estimator for all ten centre-of-mass energy windows in figure 7.21. The h^0A^0 analysis with 5C fit Higgs mass estimator is assumed to behave very similarly to the h^0A^0 analysis with 4C fit mass estimator, due to the very small differences between them, and the conclusions regarding overtraining for the latter is assumed to hold for the former.

The plots in figure 7.20 and 7.21 show signs of overall overtraining only for a few centre-of-mass energy windows: the 192 GeV, and to some extent the 205 GeV and 208 GeV energy windows in the H^0Z^0 analysis, and the 200 GeV and 207 GeV, and to some extent the 192 GeV energy windows in the h^0A^0 analysis. On the other hand, the 196 GeV centre-of-mass energy windows in both the H^0Z^0 and the h^0A^0 analysis, and to some extent the 192 GeV centre-of-mass energy window in the h^0A^0 analysis, show the opposite effect, *i.e.* the analysis on the second half of the Monte Carlo generated signal and background samples is superior to the one on the first half. This lends support to the assumption that the differences between the first and second half of the samples are due to fluctuations rather than systematic overtraining. The signal efficiency difference is seen to stay within an absolute value of $\pm 3\%$ for close to the entire range of the analyses, with the exceptions occuring only at fairly small background expectations (below $\sim$1 event), where the number of generated Monte Carlo events is small, and the analyses therefore more vulnerable to fluctuations. A common 3% absolute error on the signal efficiency is therefore included for all signals, both in the H^0Z^0 and the h^0A^0 signals.

7.7 Final steps in the H^0Z^0 search

The final analysis steps and output variables for the H^0Z^0 search are shown in figures 7.22 to 7.24 for the three years of data taking. Also, the background versus efficiency curve is shown, with both the statistical and the systematic errors included.

In order to prepare the analysis results for the final statistical treatment, in which the limits on the Higgs boson masses are set, the number of events needs to be reduced, in order to keep the procedure managable. The final statistical treatment of the search results (see section 8.2), uses a likelihood technique for the limit determination, where the likelihood contains the variable in which the cut is made; a 2-dimensional parameter space of discriminating variable vs. Higgs mass estimator. For such a technique, one would like to cut away as few events as possible, since the added information contained in the events which are otherwise cut away can only improve the expected limit, as the likelihood weighs these accordingly. (The proof of this can be found in ref. [85, appendix A]. The one exception is if the error in the expected estimate of the included events is significantly larger than the one already considered in the likelihood, but this is not a problem for the work

189 GeV Monte Carlo sample

192 GeV Monte Carlo sample

196 GeV Monte Carlo sample

200 GeV Monte Carlo sample

202 GeV Monte Carlo sample

204 GeV Monte Carlo sample

205 GeV Monte Carlo sample

207 GeV Monte Carlo sample

208 GeV Monte Carlo sample

206 GeV Monte Carlo sample

Figure 7.20: Plots showing the signal efficiency difference between the first half (on which the analysis is constructed) and the second half of the Monte Carlo generated signal, as a function of the logarithm of the expected background for the H^0Z^0 analysis. All ten centre-of-mass energy windows are shown, and the error bands originate from the statistical error in the Monte Carlo samples only.

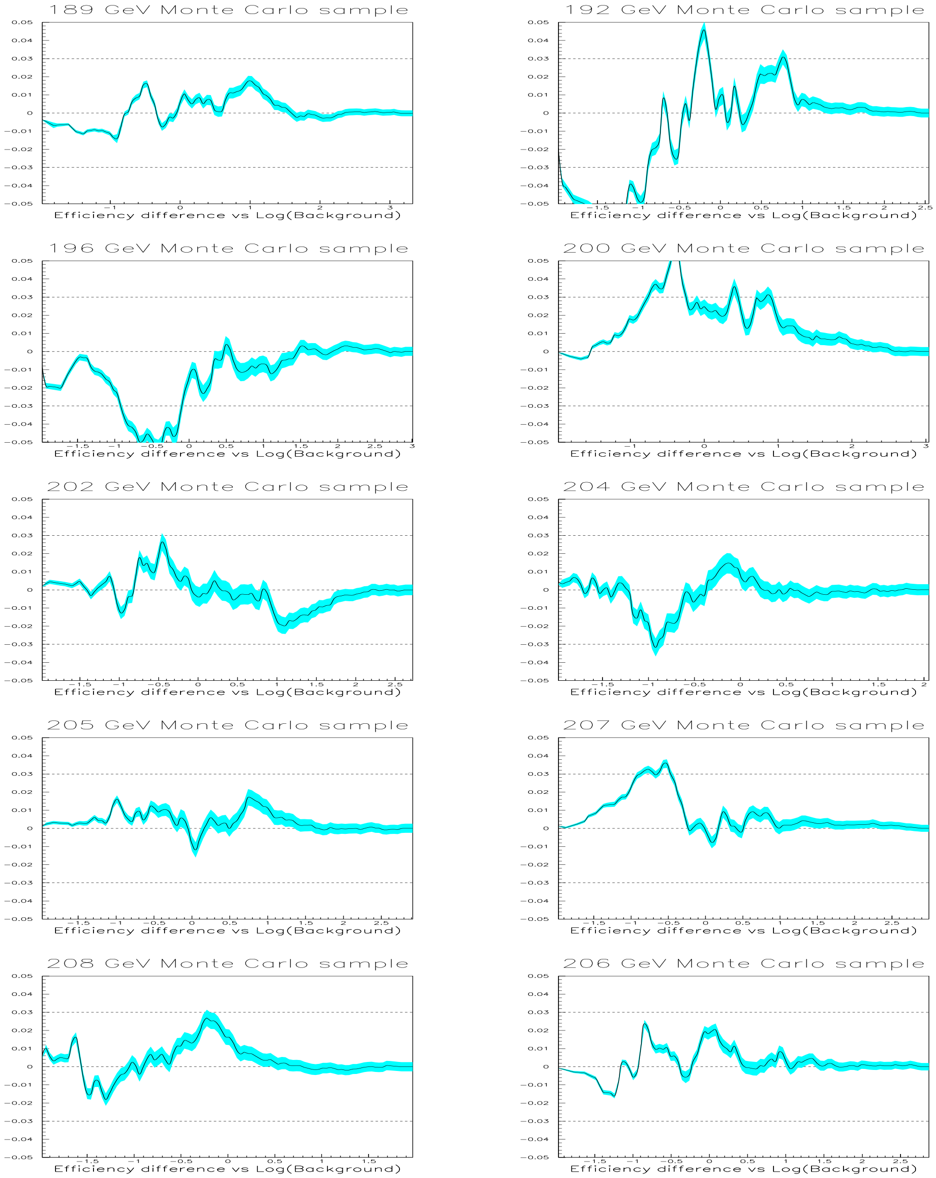


Figure 7.21: Plots showing the signal efficiency difference between the first half (on which the analysis is constructed) and the second half of the Monte Carlo generated signal, as a function of the logarithm of the expected background for the h^0A^0 analysis. All ten centre-of-mass energy windows are shown, and the error bands originate from the statistical error in the Monte Carlo samples only.

at hand.) For this reason, one should in principle include as much as possible of both the signal and background for the final statistical treatment of the search results. This is, however, a practical problem of both manageability and computer resource consumption. Therefore, a cut in the final variable at a point where the difference in expected limit is expected to be very small when compared to what would be achieved if all the data were included, is performed. For the h^0Z^0 analysis, this value is chosen to be 0.5, a number which results in approximately the same number of data candidates passing the cut as is the case for the official DELPHI analysis, so as to make comparisons easy.

Table 7.2 summarizes, on the left, the number of expected background and observed data events above the final cut for all centre-of-mass energy windows. The corresponding signal efficiencies are shown underneath. The plot of the Higgs mass estimator for the remaining events is shown on the left in figure 7.28

7.8 Final steps in the h^0A^0 search

For the 4C fit Higgs mass estimator case, the final steps of the analysis are shown in figures 7.25 to 7.27, where the final event shape variable, the final b-tag variable and the final discriminating variable are shown, together with the background versus signal efficiency plot.

Due to the h^0A^0 signal being a cleaner signal with respect to the background than what is the case for the H^0Z^0 signal, the cut for the final statistical treatment can be set at a lower value for the h^0A^0 channel, retaining a larger part of the signal. The cut value is chosen at 0.15, and the remaining expected background and observed data events, as well as the signal efficiency, is shown on the right of table 7.2. The Higgs mass estimator, which is the sum of the invariant masses of the two dijet objects, at this selection level is shown on the right in figure 7.28.

For the 5C fit mass estimator case, the discriminating variable is very similar to the one in the 4C fit case, the chief difference being in the Higgs mass estimator sector of the analysis. The final cut in the discriminating variable is also here set at 0.15, and the plot of the Higgs mass estimator for the resulting events for two different values of the mass difference Δm used in the kinematic fit (see section 7.2.3) is shown in figure 7.29.

E_{CM}	H^0Z^0		h^0A^0		
	Exp. bck.	Data	Exp. bck.		Data
188.6	23.1 ± 1.2	31	49.0 ± 2.1		61
191.6	4.4 ± 0.3	5	7.4 ± 0.4		11
195.5	11.9 ± 0.7	15	20.6 ± 0.9		20
199.5	13.5 ± 0.7	13	27.3 ± 1.2		28
201.6	7.2 ± 0.4	9	11.7 ± 0.5		10
203.6	1.5 ± 0.1	0	2.6 ± 0.1		2
205.2	10.2 ± 0.5	12	18.1 ± 0.7		19
206.7	10.8 ± 0.6	12	22.5 ± 0.9		24
208.2	1.3 ± 0.1	2	2.0 ± 0.1		4
206.3	7.9 ± 0.4	7	18.8 ± 0.8		20
Total	91.8 ± 4.8	106	180.0 ± 7.6		199
	m_{H^0}(GeV/c^2)	eff(%)	m_{A^0}(GeV/c^2)	$\tan\beta$	eff. (%)
	80	36.5 ± 4.1	80	2	75.7 ± 5.2
	85	38.6 ± 4.6	85	2	77.1 ± 5.3
	90	40.3 ± 4.2	90	2	79.2 ± 5.4
	92.5	47.4 ± 4.5	95	2	79.0 ± 5.4
	95	44.0 ± 4.4	100	2	81.0 ± 5.5
	97.5	46.0 ± 4.4	80	20	79.2 ± 5.4
	100	44.7 ± 4.4	85	20	81.5 ± 5.5
	105	46.2 ± 4.5	90	20	83.3 ± 5.5
	108	47.0 ± 4.4	95	20	83.4 ± 5.6
	110	47.4 ± 4.5	100	20	81.4 ± 5.5
	112	45.8 ± 4.4	80	50	76.0 ± 5.2
	114	44.8 ± 4.3	85	50	77.6 ± 5.3
	115	42.2 ± 4.3	90	50	78.7 ± 5.3
	120	35.9 ± 4.0	95	50	76.6 ± 5.3
	—	—	100	50	76.5 ± 5.3

Table 7.2: The numbers of expected background and observed data events for all centre-of-mass energy windows, as well as in total, are shown in the upper half, whereas the signal efficiencies for different signals are shown in the lower half. The H^0Z^0 analysis is shown on the left, and the h^0A^0 analysis with 4C fit is shown on the right. The signal efficiency corresponds to the luminosity-weighted total over the sum of the centre-of-mass energy windows, as described in appendix D.

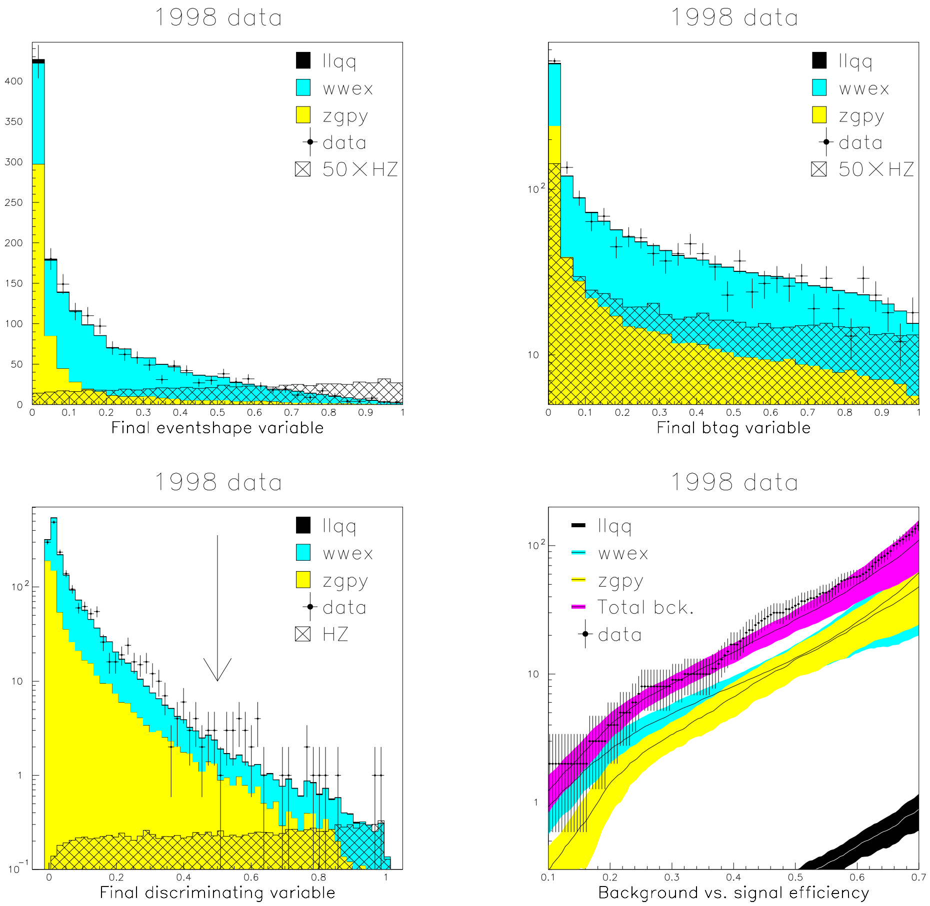


Figure 7.22: Plots of the final event shape variable, the final b-tag variable and the final discriminating variable, as well as the background versus efficiency curve for the H^0Z^0 analysis. The signal plotted in the plots of the variables, as well as the signal used for the background versus efficiency curve, is the m_{H^0}=95 GeV/c^2 signal. The 1998 data are shown.

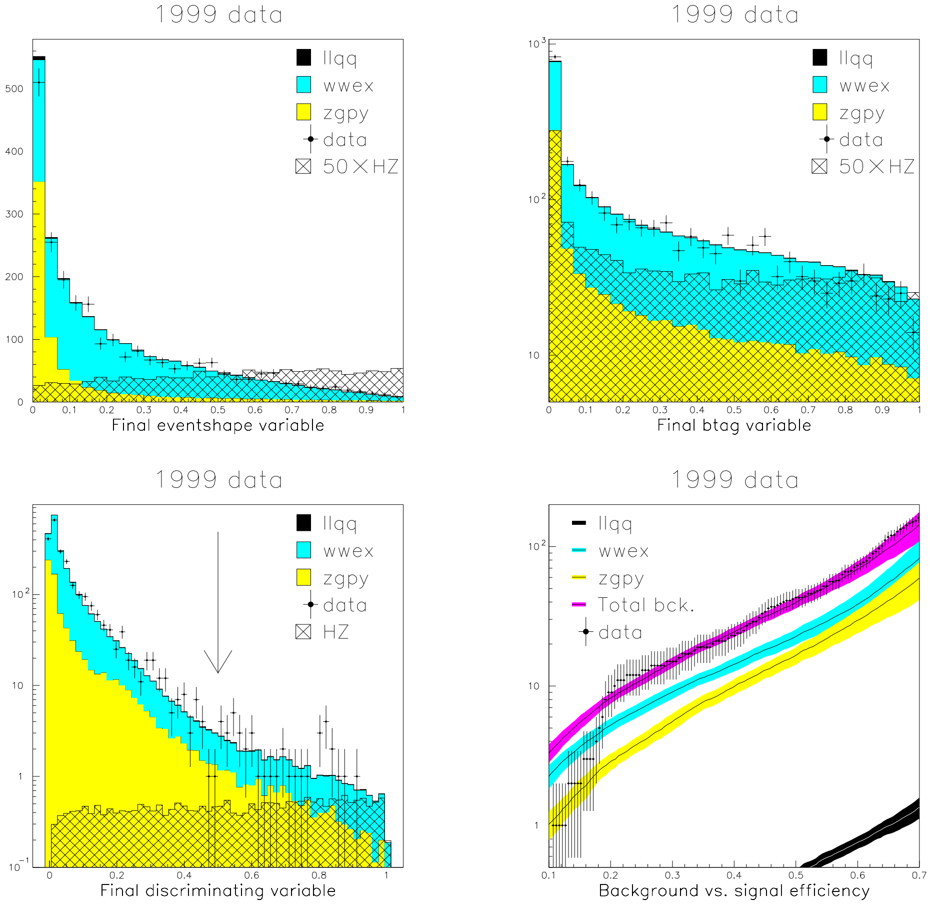


Figure 7.23: Plots of the final event shape variable, the final b-tag variable and the final discriminating variable, as well as the background versus efficiency curve for the H^0Z^0 analysis. The signal plotted in the plots of the variables, as well as the signal used for the background versus efficiency curve, is the m_{H^0}=100 GeV/c^2 signal. The 1999 data are shown.

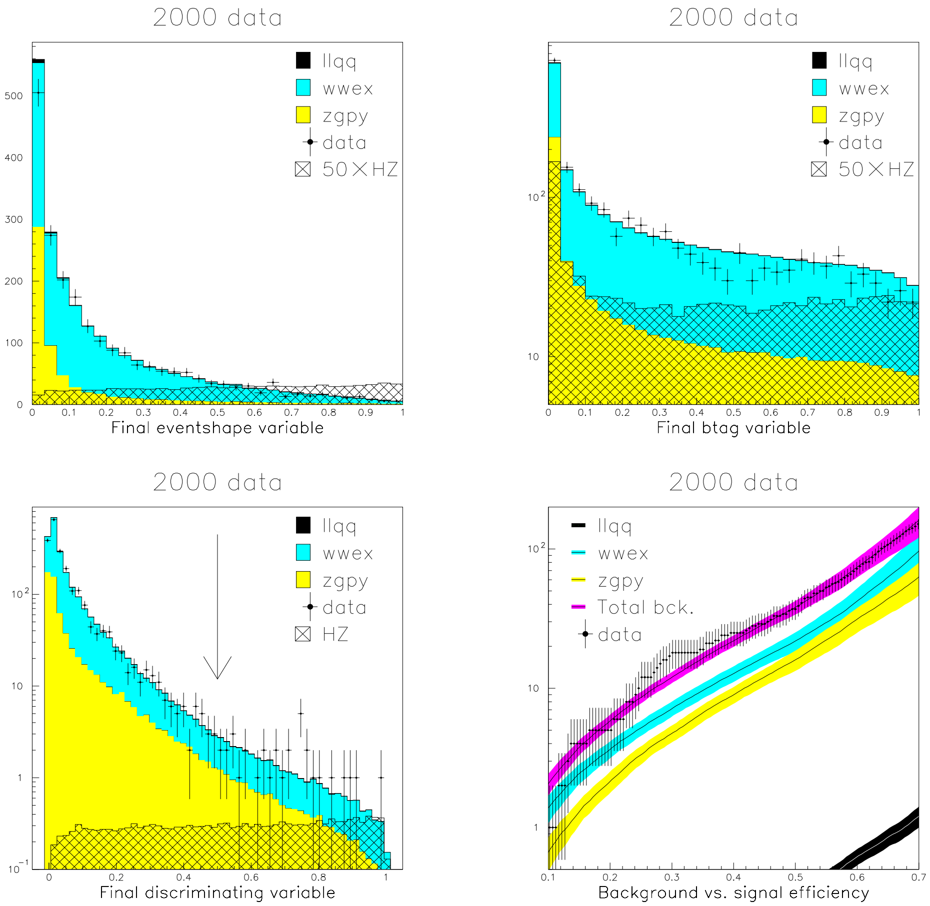


Figure 7.24: Plots of the final event shape variable, the final b-tag variable and the final discriminating variable, as well as the background versus efficiency curve for the H^0Z^0 analysis. The signal plotted in the plots of the variables, as well as the signal used for the background versus efficiency curve, is the m_{H^0}=110 GeV/c^2 signal. The 2000 data are shown.

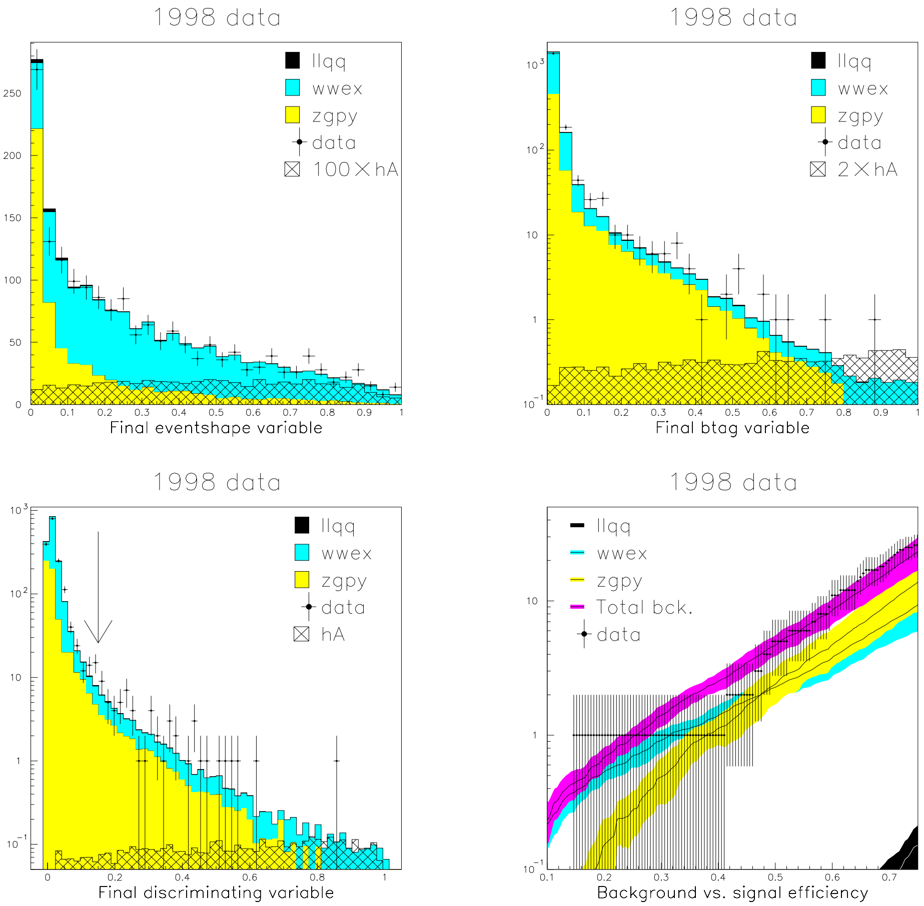


Figure 7.25: Plots of the final event shape variable, the final b-tag variable and the final discriminating variable, as well as the background versus efficiency curve for the h^0A^0, 4C fit analysis. The signal shown in the plots of the variables, as well as the signal used for the background versus efficiency curve, is the m_{A^0}=85 GeV/c^2, $\tan\beta$=20 signal. The 1998 data are shown.

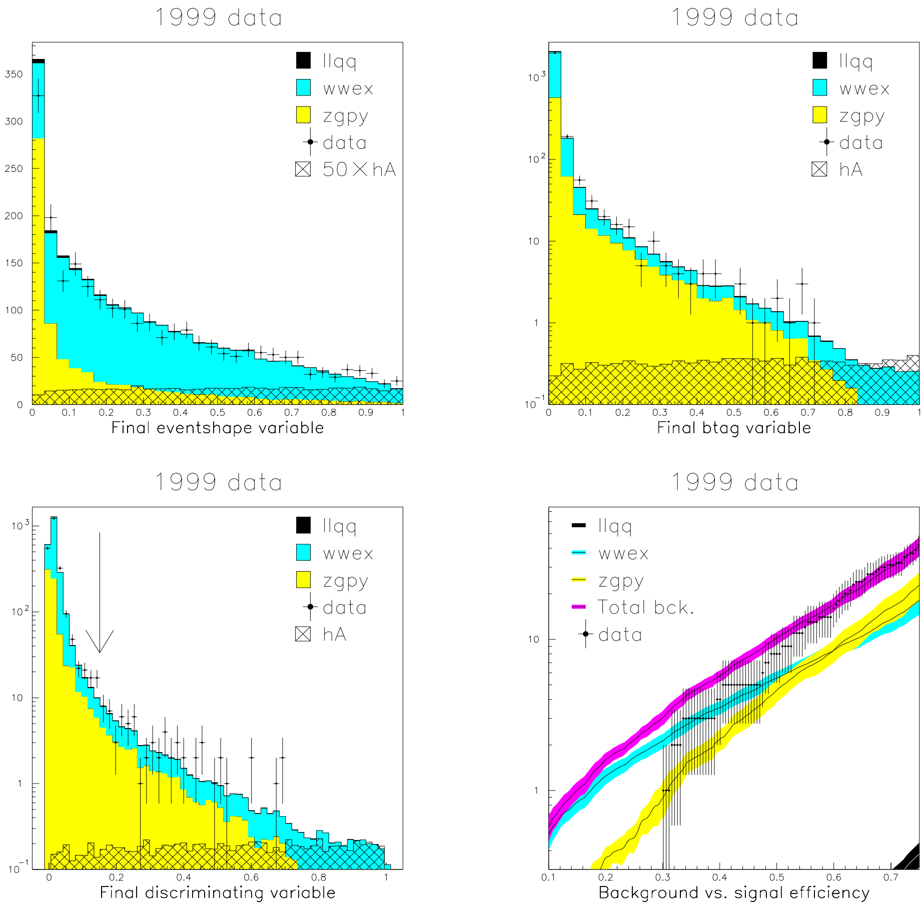


Figure 7.26: Plots of the final event shape variable, the final b-tag variable and the final discriminating variable, as well as the background versus efficiency curve for the h^0A^0, 4C fit analysis. The signal shown in the plots of the variables, as well as the signal used for the background versus efficiency curve, is the m_{A^0}=85 GeV/c^2, $\tan\beta$=20 signal. The 1999 data are shown.

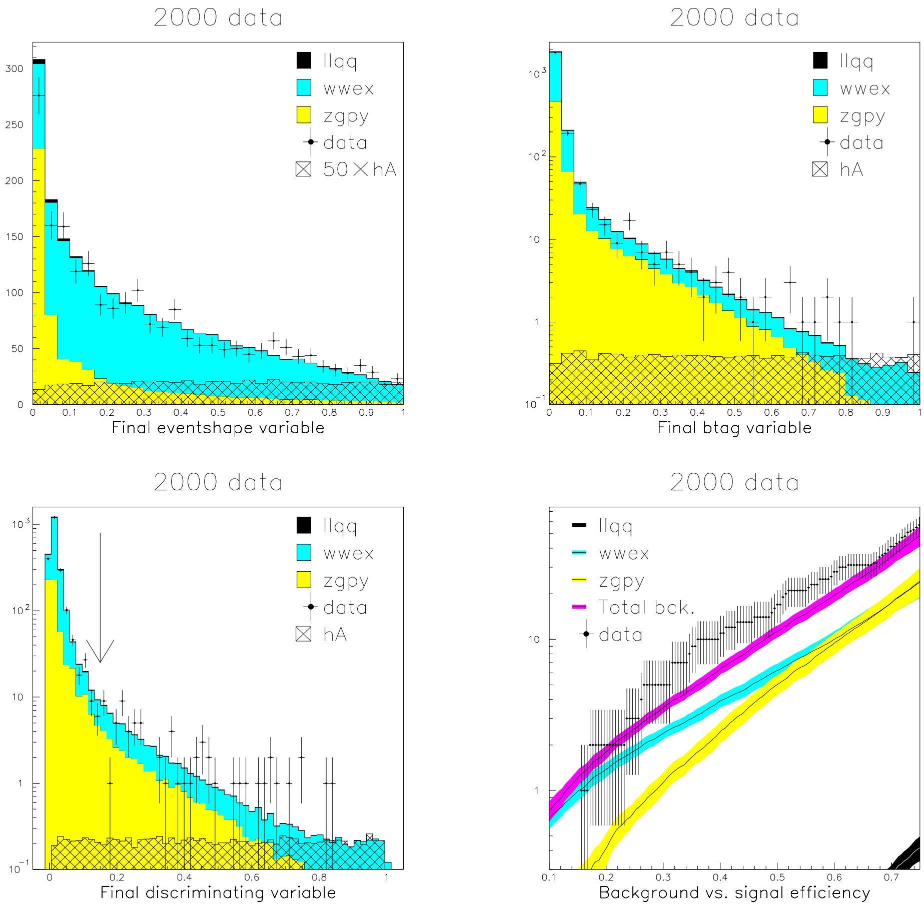


Figure 7.27: Plots of the final event shape variable, the final b-tag variable and the final discriminating variable, as well as the background versus efficiency curve for the h^0A^0, 4C fit analysis. The signal shown in the plots of the variables, as well as the signal used for the background versus efficiency curve, is the m_{A^0}=85 GeV/c^2, $\tan\beta$=20 signal. The 2000 data are shown.

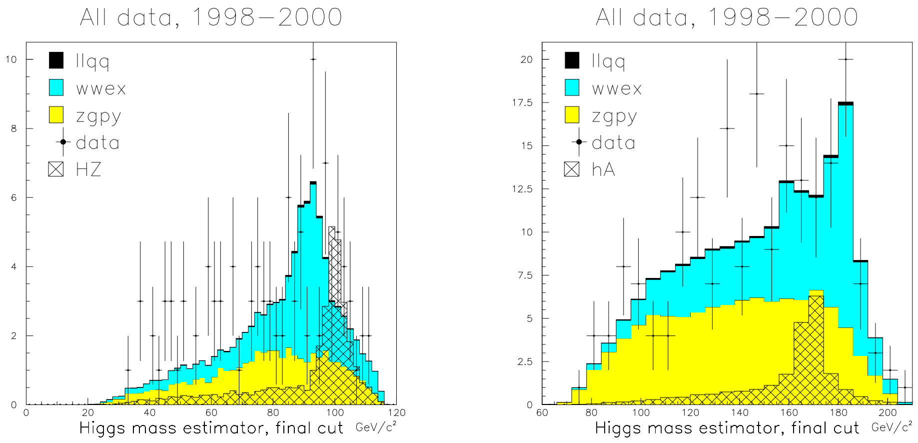


Figure 7.28: Plots showing the Higgs mass estimator at the final cut selection level for both the H^0Z^0 (the invariant mass of the dijet opposite the one forced to the Z^0 mass) and h^0A^0 with 4C fit mass estimator (the sum of the invariant masses of the dijets, with the pairing chosen to minimize the dijet mass difference) analyses. All centre-of-mass windows are added, and the signal corresponds to m_{H^0}=100 GeV/c^2 for the H^0Z^0 case, and m_{A^0}=85 GeV/c^2, $\tan\beta$=20 for the h^0A^0 case.

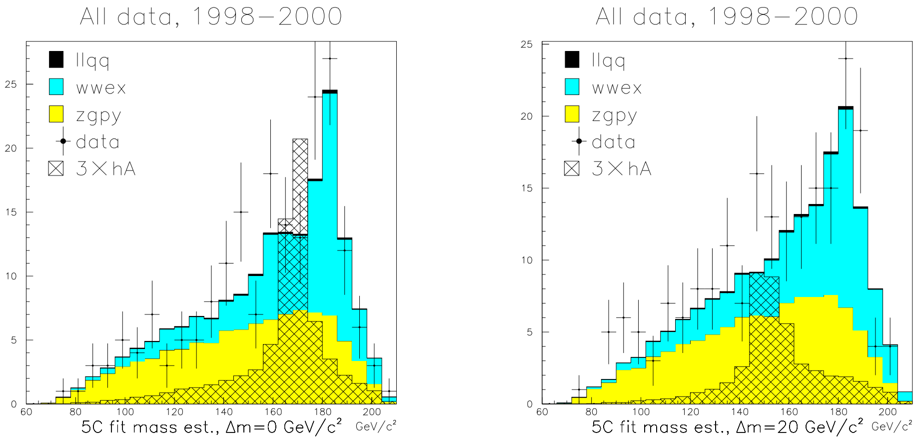


Figure 7.29: Plots showing the Higgs mass estimator at the final cut selection level for the h^0A^0 analysis with 5C fit mass estimator. All centre-of-mass energy windows are added, and the signal corresponds to m_{A^0}=85 GeV/c^2, $\tan\beta$=20 with Δm=0 GeV/c^2 on the left, and m_{A^0}=85 GeV/c^2, $\tan\beta$=2 with Δm=20 GeV/c^2 on the right.

7.9 Systematic errors

In order to estimate the systematic errors at the final selection level, *i.e.* after the cut in the final discriminating variable at a value of 0.5 for the H^0Z^0 analysis, and 0.15 for the h^0A^0 analysis, a systematic procedure was used. This procedure consists of studying each input variable separately, varying the value of the variable by approximately one standard deviation, and observing the change in the final discriminating variable at the final selection level. After having gone through this procedure for every variable, the total systematic error was estimated by adding the individual contributions for each variable in quadrature. Again, the addition of the χ^2 from the fifth constraint in the 5C fit analysis is assumed to have only a negligible effect on the systematic error estimation, and the contribution from this variable is therefore omitted. Also, the choice of track selection is not assumed to contribute to the total systematic error, as it tends to affect the results only very slightly. One possible effect from the different track selections would be that events just barely meeting the four jet preselection criteria with one track selection, could fail this preselection with another track selection. However, such events tend not to be very Higgs-like in any case, as the preselection is designed to retain a large portion of the Higgs signal, making it fairly unimportant whether such events are rejected by the preselection or given a very low weight in the final analysis. The b-tag package, which is the most important analysis tool for the four jet channel, uses its own track selection algorithm, resulting in b-tag variables which are independent of the track selection used for the rest of the analysis. The remaining effect of the different track selections is mostly the difference in track momenta and energies left in the event; this difference is for a large part "smoothed over" by the use of kinematic fits (see section 7.2), which are indeed used to correct for and improve on imperfections in the detector.

Table 7.3 shows the resulting change in the total background from varying each individual variable, and finally the total, as the sum of squares of the 13 individual contributions. The total systematic error at the final selection level in the background estimate amounts to 5.5% in the H^0Z^0 channel, and 3.8% in the h^0A^0 channel.

For the final systematic error used in the final statistical treatment of the search results, the contribution from the four jet selection (see page 93) is added in quadrature with the systematic errors for the final variable selection cut. The resulting total systematic error is the one which is quoted in table 7.2, and corresponds to the error used in the final statistical treatment in the ALRMC program (see section 8.2).

Variable	H^0Z^0 (orig. bck.: 92.364)		h^0A^0 (orig. bck.: 181.151)	
	After shift	Shift (%)	After shift	Shift (%)
B_1	91.500	-0.94	180.186	-0.53
B_2	91.387	-1.06	179.068	-1.15
B_3	91.544	-0.89	177.934	-1.78
B_4	92.364	-0.58	179.575	-0.87
$B_{\rm HA}$	91.621	-0.80	178.737	-1.33
$\rm Tr_{min}$	94.213	2.00	184.583	1.89
$E_{\rm min}$	91.709	-0.71	181.016	-0.07
$\alpha_{\rm min}$	92.473	0.12	182.141	0.55
H_2	91.176	-1.29	181.510	0.20
H_4	93.775	1.53	181.707	0.31
M_L	89.181	-3.45	180.951	-0.11
Thrust	90.073	-2.48	184.349	1.77
Y_{34}	91.732	0.68	180.812	-0.19
	Total error	5.5	Total error	3.8

Table 7.3: The results from the systematic error estimation procedure. The numbers for each individual variable reflect the change in the expected background for the sum of all centre-of-mass energy windows.

Chapter 8

Results

In this chapter, the final results from the analyses is presented. The three different MSSM benchmarks are first introduced, and the statistical method of search analysis results based on the likelihood ratio, the ALRMC method [85], is briefly recalled. The extraction of the excluded parameter ranges is then performed using only the Higgs search channels presented in this work, both for the analysis presented here and for the official DELPHI search channels, in order to compare their performances. Finally, the Higgs search channels from this work is combined with the official DELPHI Higgs search channels in order to extract the final excluded MSSM parameter regions, which also leads to absolute lower limits on the Higgs masses m_{h^0} and m_{A^0}.

8.1 MSSM benchmark scenarios

As mentioned in section 2.5.3, the Higgs sector of the MSSM can be described at tree-level by two parameters, usually chosen as $\tan\beta$ and one of the Higgs boson masses. Thus, at tree-level, confidence limits on the parameters describing the Higgs sector can be set in a plane of two of the parameters mentioned in section 2.5.3, and the limits on the remaining parameters can be calculated from these. However, radiative corrections introduce dependencies on the other parameters of the theory. In DELPHI, differences originating from these radiative corrections are taken into account by choosing different benchmarks, *i.e.* different specific values for the MSSM parameters. The framework in which this is done, is the universality assumption, where the remaining free parameters beyond tree-level are:

- The Higgs mixing parameter, μ.
- The universal sfermion mass term, M_S.

- The universal gaugino mass term, M_2.
- The gluino mass, $m_{\tilde{g}}$.
- The universal squark trilinear coupling, A.

The benchmarks are:

No mixing benchmark: This scenario differs from the $m_h^{\max}$ scenario only in the choice of the parameter $X_t = A - \mu \cot\beta$, which controls the mixing in the stop sector. This scenario sets the parameter X_t to zero (thus, the name no mixing). The values for the parameters are:

- M_S=1 TeV/c^2
- M_2=-μ=200 GeV/c^2
- $m_{\tilde{g}}$=800 GeV/c^2
- X_t=0

$m_h^{\max}$ benchmark: This scenario is designed to maximize the largest value of m_{h^0} as a function of $\tan\beta$. As this gives the kinematically most unfavourable situation at a specific choice of the other MSSM parameters, this scenario gives the most conservative limits on the m_{h^0} parameter, given that the h^0 is always present in the signal channels under study. At small values of $\tan\beta$, this scenario reduces to the SM. The values of the parameters are:

- M_S=1 TeV/c^2
- M_2=-μ=200 GeV/c^2
- $m_{\tilde{g}}$=800 GeV/c^2
- X_t=2M_S

Large μ benchmark: This scenario predicts at least one scalar Higgs boson with a mass within kinematical reach of LEP-*II* in every theoretically allowed point of the MSSM parameter space, even though there are some regions where the Higgs boson would be undetectable due to $b\bar{b}$ branching fractions being strongly suppressed by radiative corrections. The values for the parameters are:

- M_S=M_2=400 GeV/c^2
- μ=1 TeV/c^2
- $m_{\tilde{g}}$=200 GeV/c^2
- X_t=-300 GeV/c^2

8.2 The ALRMC statistical procedure

In order to obtain the excluded ranges of MSSM parameters, the ALRMC [85] method was used. The test-statistic X used is given by $X = \ln(Q)$, where Q is the likelihood ratio for N_{chan} independent search channels,

$$Q = e^{-s_{\text{tot}}} \prod_{i=1}^{N_{\text{chan}}} \prod_{j=1}^{n_i} \left(1 + \frac{s_i S_i(\vec{x_{ij}})}{b_i B_i(\vec{x_{ij}})}, \right) \tag{8.1}$$

where n_i is the observed number of candidates in each channel, $\vec{x_{ij}}$ is the value of the discriminating variables measured for each of the candidates (in the DELPHI Higgs searches, these are the discriminating variable information and the Higgs mass estimator), s_i and b_i are the integrated signal and background rates per channel, s_{tot} is the total signal rate for all channels, $s_{\text{tot}} = \sum_{i=1}^{N_{\text{chan}}} s_i$, and $S_i(\vec{x})$ and $B_i(\vec{x})$ are the probability distribution functions (p.d.f.'s) of the discriminating variable for the signal and background respectively. These p.d.f.'s are typically not available in analytic form, as they are constructed from simulated signal and background events passing through first the detector simulation and then the specific search channel analysis, and distribution of Q must therefore be constructed from Monte Carlo simulations. When doing such simulations, each Monte Carlo event is assigned values of $\vec{x_{ij}}$ according to the appropriate p.d.f., and the number of expected signal and background events for the channel in question, s_i and b_i, are varied according to a gaussian function with the total systematic error (see section 7.9) as the standard deviation.

The different confidences are now calculated by comparing the observed value of the test statistic, $X_{\text{obs}} = \ln(Q_{\text{obs}})$, to the distributions obtained from the Monte Carlo simulations. The confidence in the signal+background hypothesis is given by

$$\text{CL}_{s+b} = P_{s+b}(X \leq X_{\text{obs}}) \tag{8.2}$$

where

$$P_{s+b}(X \leq X_{\text{obs}}) = \int_0^{X_{\text{obs}}} \frac{dP_{s+b}}{dX} dX \tag{8.3}$$

and where dP_{s+b}/dX is the p.d.f. of the test-statistic for signal+background experiments. In the same manner, the confidence in the background is given by

$$\text{CL}_b = P_b(X \leq X_{\text{obs}}) \tag{8.4}$$

where

$$P_b(X \leq X_{\text{obs}}) = \int_0^{X_{\text{obs}}} \frac{dP_b}{dX} dX \tag{8.5}$$

and where dP_b/dX is the p.d.f. of the test-statistic for background-only experiments. The background hypothesis will be considered excluded at a specific confidence level if $1-\mathrm{CL}_b \leq \mathrm{CL}$, where CL gives the confidence level. At the 5σ confidence level, CL is equal to 5.7×10^{-7}, and if the confidence in the background hypothesis reaches a level closer to unity than this, the background hypothesis is rejected, signifying a 5σ signal discovery.

The modified frequentist procedure, as the ALRMC method is based on, now defines the signal confidence as

$$\mathrm{CL}_s \equiv \mathrm{CL}_{s+b}/\mathrm{CL}_b. \tag{8.6}$$

Strictly speaking, this makes CL_s not a true confidence, but rather a ratio of confidences. The significance of this definition, is to make CL_s approximate the confidence in the signal hypothesis one might have obtained if the experiment had been performed in the absence of background, *i.e.* if a precise correction to the data could be made in order to compensate for the presence of background. As with CL_b, the signal hypothesis is rejected at a specific confidence level CL if $1-\mathrm{CL}_s \leq \mathrm{CL}$. For the exclusion in the MSSM parameter space, a 2σ confidence level is used, giving a CL of 0.05.

In the extraction of excluded MSSM parameter ranges, the MSSM parameter space is scanned, calculating CL_s for a number of specific points in the MSSM parameter space. Points which are excluded at 95% confidence level or more, *i.e.* points where 1-CL_s falls below 0.05, are then mapped out in the 2-dimensional planes made up of two of the three parameters on which limits are set, thus creating the excluded regions in three possible projections ($\tan\beta$ versus m_{h^0}, $\tan\beta$ versus m_{A^0} and m_{A^0} versus m_{h^0}). This procedure is performed covering a range in $\tan\beta$ from 0.4–50 for the no mixing and m_h^{max} benchmarks, and 0.7–50 for the large μ benchmark.

8.3 Comparison between this work and official DELPHI search channels

In order to assess the discriminating power of the analysis presented in this work, the excluded MSSM parameter ranges were extracted using only the search channels presented in this work, *i.e.* the channels described in table 4.1 which are marked by *, and at centre-of-mass energy points from 189 GeV and upward. This was done both for the analyses presented in this work and for the official DELPHI search channels. By comparing these two different exclusions, one can form an opinion about how the analysis method presented in this work performs when compared to more traditional search methods.

Production channel $e^+e^- \to Z^{0*} \to H^0Z^0$		
Channel name	H^0 Decay	Z^0 Decay
hadronic*	$b\bar{b}$	$q\bar{q}(q = u, d, s, c, b)$
neutrino or invisible	$b\bar{b}$	$\nu_l\overline{\nu_l}(l = e, \mu, \tau)$
electron	$b\bar{b}$	e^+e^-
muon	$b\bar{b}$	$\mu^+\mu^-$
tau	$b\bar{b}$	$\tau^+\tau^-$
	$\tau^+\tau^-$	$q\bar{q}(q = u, d, s, c, b)$

Production channel $e^+e^- \to Z^{0*} \to h^0A^0$		
Channel name	h^0 Decay	A^0 Decay
four b*	$b\bar{b}$	$b\bar{b}$
tau	$b\bar{b}$	$\tau^+\tau^-$
	$\tau^+\tau^-$	$b\bar{b}$

Table 8.1: A short schematic view of the different neutral Higgs search channels in DELPHI. The channels marked by * are channels analysed in this work.

The exclusion plots for the no mixing and $m_h^{\max}$ benchmarks are shown in figure 8.1, both for the search analysis described in this work, and for the corresponding official DELPHI search channels. In both cases, only the projection $\tan\beta$ vs. m_{h^0} is shown, as the excluded regions in the other projections tend to look rather ragged, due to the use of only a subset of the total search channels, and therefore offer more confusion than illumination. The h^0A^0 analysis used is the one with 4C fit mass estimator, the assumption being that the difference between the two different h^0A^0 analysis cases is sufficiently small that comparing one of them with the official DELPHI search results gives an adequate opinion about their performance. Indeed, as will be shown in section 8.4, the difference in the expected final results between the two methods is negligible. For the large μ benchmark, at some points of the MSSM parameter space the branching ratio of h^0 to $b\bar{b}$ falls close to zero, which of course renders the four jet channels useless. Therefore, the exclusion in the MSSM parameter space with only the four jet channels becomes rather difficult for this benchmark, and the comparison is therefore only carried out in the no mixing and $m_h^{\max}$ benchmarks..

The first thing to note about the plots in figure 8.1, is that the excluded regions of the no mixing benchmark contains a large hole for $\tan\beta$ values from ∼1 to ∼2, which is not present in the final exclusion plots comprised of all the different DELPHI Higgs search channels, shown in figures 8.3 to 8.8.

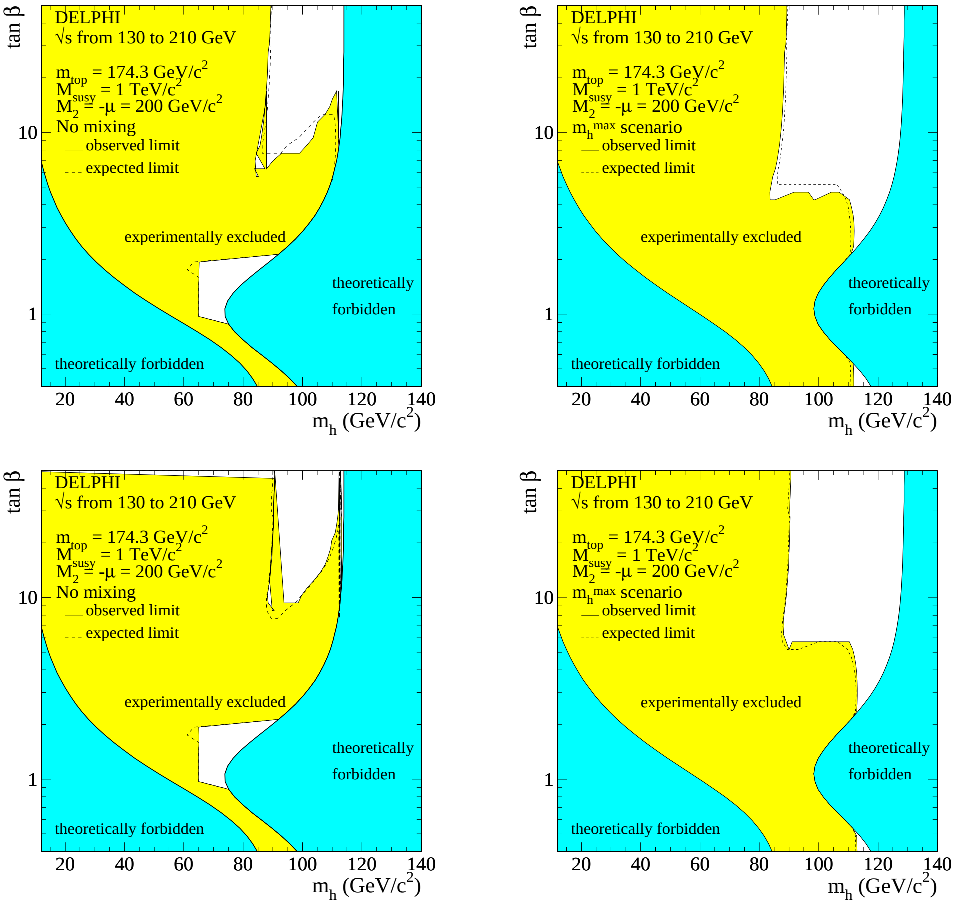


Figure 8.1: Exclusion plots for the no mixing and $m_h^{\max}$ benchmarks both for search results using the search channels from this work and for search results using the corresponding official DELPHI search channels. All plots show the projection of $\tan\beta$ vs. m_{h^0}. On the top left is shown the exclusion plot of the no mixing benchmark, and on the top right is shown the $m_h^{\max}$ benchmark, both for the analysis presented in this work. The bottom row shows the same two plots when using the official DELPHI search channels.

The reason for this, is that the fact that one restricts oneself to the four jet channels, means that new unexcluded regions of the MSSM parameter space appear, specifically at regions where other decay modes than $b\bar{b}$ becomes important for the h^0. Such holes tend to appear at specific places in the MSSM parameter space, with little influence from the search channels themselves, making a comparison of different methods rather difficult at such MSSM parameter regions. Therefore, the absolute limits on the MSSM parameters might be rather different in this test case than what would be the case for the exclusion with all search channels present, and in addition might not reflect very well the difference between the two analyses.

Thus, the best way to assess the difference between the analysis presented in this work and the official DELPHI analysis, is to compare the lowest unexcluded value of m_{h^0} for the two search results at each value of $\tan\beta$, and see how this behaves at different values of $\tan\beta$. For most values of $\tan\beta$ this will give a meaningful measure of the difference in analysis power. For all these points, the official DELPHI analysis performs slightly better than the analysis presented here. For the no mixing benchmark, this procedure shows no difference between the two search results at low values of $\tan\beta$, as the entire MSSM parameter range allowed by the theory is excluded here. For $\tan\beta$ values between $\sim$1 and $\sim$2, there is the earlier mentioned unexcluded hole whose lower edge in m_{h^0} lies within a difference of 0.2 GeV/c^2 between the two searches. At larger values of $\tan\beta$, the difference becomes larger, as the "transition region" where the h^0A^0 analysis becomes the more important one occurs at slightly different values of $\tan\beta$ for the two cases, but quickly settles at a level of <1 GeV/c^2 above $\tan\beta \sim 9$. For the m_h^{max} benchmark, the same procedure shows a relatively larger difference at small values of $\tan\beta$, where the difference between the two analyses is $\sim$1.7 GeV/c^2 in terms of excluded values of m_{h^0}. In this region, the H^0Z^0 analysis dominates over the h^0A^0 and the difference therefore illustrates the fact the the H^0Z^0 analysis in this work is not optimized at Higgs boson masses close to the kinematical edge. After the $\tan\beta$ range of $\sim$0.6–$\sim$2, which are excluded by both analyses, the difference between them are at about the same level as at very low values of $\tan\beta$, increasing slightly for values of $\tan\beta$ up to $\sim$4, where the same difference between where the "transition region" is situated as in the no mixing benchmark leads to the same relatively large difference between the two search results. However, for values of $\tan\beta$ above $\sim$6, where the h^0A^0 channel begins to be the dominant one, the difference between the two analyses in terms of excluded values of m_{h^0} quickly settles to a level of less than 1 GeV/c^2, reminiscent of the situation for the no mixing benchmark.

This comparison is summarized in figure 8.2, where the difference in expected lower excluded value of m_{h^0} is shown as a function of $\tan\beta$ for the

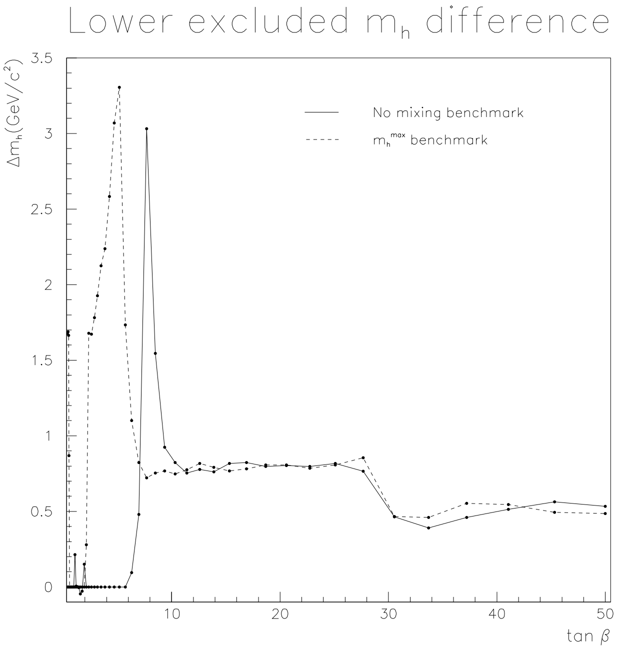


Figure 8.2: Plot showing the difference in expected lower excluded value of m_{h^0} for the search results obtained using the search channels from this work and for search results using the corresponding official DELPHI search channels. The no mixing benchmark is shown in solid, while the $m_h^{\max}$ benchmark is shown in dashed. The dots indicate the values in $\tan\beta$ at which the exclusion was calculated.

two different search results.

8.4 Final exclusion in the MSSM parameter space

When calculating the final limits, the results from the H^0Z^0 analysis were used together with the two different h^0A^0 analyses (4C fit and 5C fit Higgs mass estimator) in two different runs. In addition, the official DELPHI analysis results were used for the Higgs search channels not covered by this work. These channels are shown in table 8.1.

The three exclusion plots from the three different MSSM benchmarks for both runs are shown in figures 8.3 to 8.8. From these plots, one can find extreme values of the three parameters m_{h^0}, m_{A^0} and $\tan\beta$ (for m_{h^0} and

m_{A^0} only lower values, but for $\tan\beta$ both upper and lower values) which are not excluded when scanning over all points in a given exclusion plot. This procedure gives the overall limits of these parameters for each benchmark, and they are summarized in table 8.2. Also, the SM results extracted from these runs are shown in this table (see appendix A for the extraction of the SM m_{H^0} parameter from the MSSM search).

A few points should be mentioned concerning these limits. For the large μ benchmark, only two unexcluded regions remain, around m_{h^0}=60 GeV/c^2 at values of $\tan\beta$ below ~0.9, and at values of $\tan\beta$ above ~7 for m_{h^0} from ~90 GeV/c^2 to ~105 GeV/c^2. The first of these appears at Higgs mass values far below the "characteristic" excluded Higgs mass, and would continue further down at lower values of $\tan\beta$ than 0.7. Therefore, the actual limits in the Higgs mass ranges do not very well represent the excluded regions, as these limits would be much lower than the "characteristic" excluded Higgs masses (recall that the large μ benchmark predicts at least one scalar Higgs boson with a mass within kinematical reach of LEP-*II* in every theoretically allowed MSSM parameter point). For this reason, it is not customary to quote limits on the parameters m_{h^0} and m_{A^0} for this benchmark, but rather to refer to the figures showing the excluded regions. For the same reason, it is also customary to disregard the small unexcluded hole around m_{h^0}~85 GeV/c^2, m_{A^0}~35 GeV/c^2 for the no mixing benchmark, but, as this hole is very small, nevertheless quote Higgs mass limits.

A last point concerning the limits, is that the uncertainty in the Higgs mass results due to Monte Carlo statistics from the ALRMC run is of the order of ~50 MeV/c^2, which means that differences in the limits of 0.1 GeV/c^2 or less should not be considered significant.

For comparisons, the official search results including all LEP-*II* data for the four different LEP experiments are shown in table 8.3. These numbers show that for two of the LEP experiments, L3 and OPAL, the observed excluded mass ranges for the m_{h^0} and m_{A^0} parameters lie well below the expectation ($\gtrsim$5 GeV/c^2 difference). One of the LEP experiments, ALEPH, shows a smaller deviation of the same tendency, whereas the remaining experiment, DELPHI, shows the opposite trend, *i.e.* a stronger observed exclusion than expected. For the SM results, the situation is somewhat similar, as two experiments, ALEPH and OPAL, show an observed result well below the expectation (2.7 GeV/c^2 and 2.8 GeV/c^2, respectively), one observed result, from L3, lies just below the expectation, and the reamining observed result, from DELPHI, lies slightly above the expectation. A review of the status of Higgs boson searches at different experiments and for different theoretical models, with emphasis on future outlook, can be found in ref. [86], and an analysis of the general two-doublet CP-conserving Higgs model (not nec-

Benchmark	m_{h^0}(GeV/c^2)	m_{A^0}(GeV/c^2)	$\tan\beta$
Results from this work in H^0Z^0 and h^0A^0 channels			
	4C fit h^0A^0 analysis		
No mixing	87.5 (88.5)	89.3 (89.2)	0.44–7.69 (0.40–7.69)
$m_h^{\max}$	86.9 (88.2)	89.4 (89.4)	0.49–2.36 (0.54–2.14)
Large μ	— (—)	— (—)	0.91–6.74 (0.91–7.36)
	5C fit h^0A^0 analysis		
No mixing	87.5 (88.5)	89.2 (89.3)	0.44–7.69 (0.40–7.69)
$m_h^{\max}$	86.8 (88.2)	89.2 (89.4)	0.49–2.36 (0.54–2.14)
Large μ	— (—)	— (—)	0.91–6.74 (0.91–7.36)
SM	114.3 (113.1)	—	—

Table 8.2: The 95% CL excluded regions (for $\tan\beta$) and 95% CL lower limits (for m_{A^0} and m_{h^0}) of the three MSSM paramters m_{h^0}, m_{A^0} and $\tan\beta$ for all three MSSM benchmarks. The numbers show results obtained for the analyses presented in this work together with the other official DELPHI search channels, for both choices of the h^0A^0 analysis, the 4C fit case being in the upper half and the 5C fit case in the lower half. The SM results are also shown, in which case the difference between using the 4C fit and 5C fit method for the h^0A^0 analysis is non-measurable. Expected results are shown in parantheses after observed results.

essarilly supersymmetric) of type II (i.e. the second of the two choices on page 32) is given in ref. [87] with emphasis on light Higgs bosons. Ref. [88] gives explanations of the 2.6σ SM deviation of the measured value of the muon anomalous magnetic moment from the BNL E821 experiment in terms of a light CP-even Higgs boson of the two-doublet model, whereas ref. [89] argues in favour of a supersymmetric explanation for the possible hints of a Higgs signal with a mass at about 115 GeV/c^2.

When choosing between the two different analyses in the h^0A^0 channel (4C fit or 5C fit Higgs mass estimator), there are several considerations to take into account, some of which have been presented in sections 7.2.2 and 7.2.3. The added complexity, and hence the need for a better systematized implementation, of the 5C fit method with respect to the 4C fit method may be a point in favour of the 4C fit method when starting an analysis, but is certainly not so when choosing between final results. The problem of data events changing significance and Higgs mass estimator value with the changing signal hypotheses is not a fundamental one, but rather one of presenting the analysis before the final results (the mass limits), *i.e.* at the level of showing the final selected events, and should therefore not be assigned too

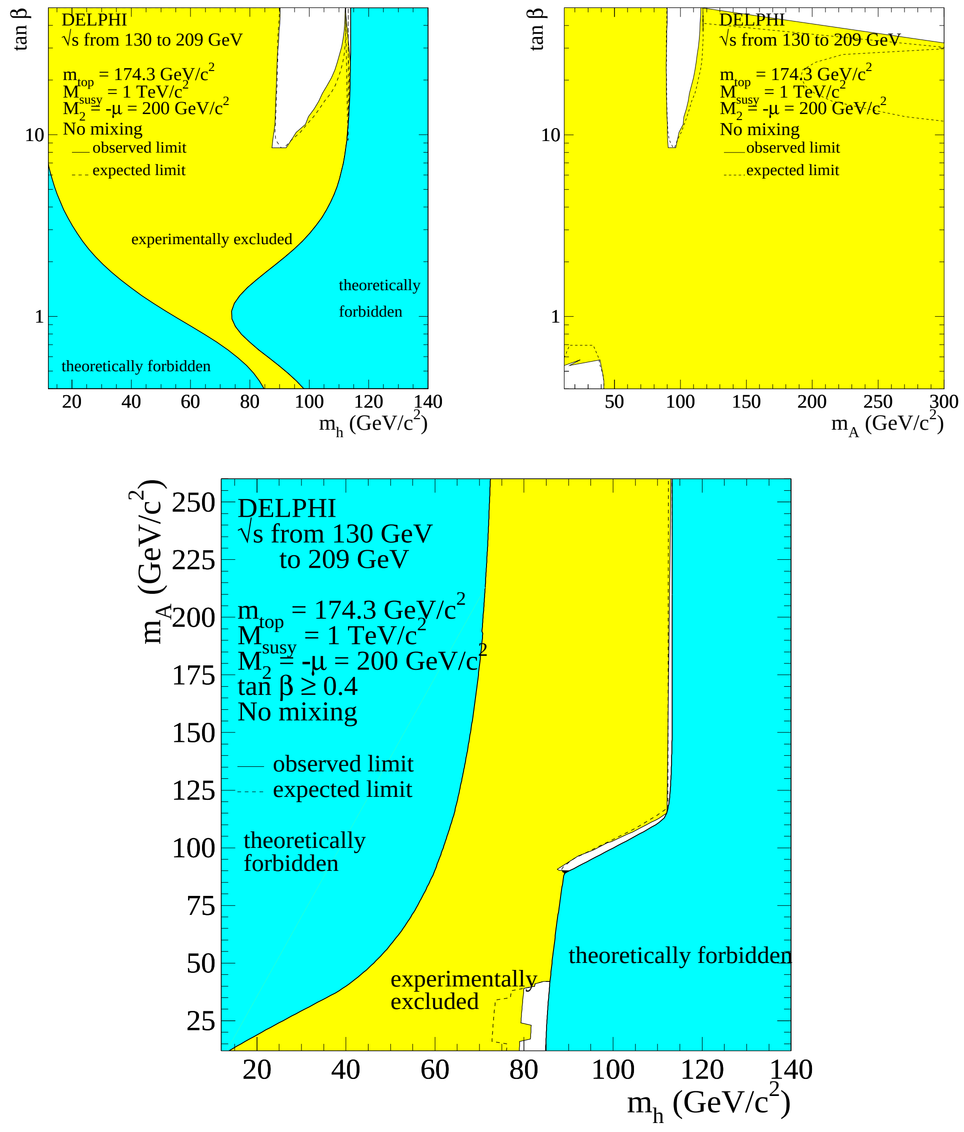


Figure 8.3: Exclusion plots for the no mixing benchmark: On the top left is shown the $\tan\beta$ versus m_{h^0} plot, the $\tan\beta$ versus m_{A^0} plot is shown on the top right, whereas the bottom plot shows the m_{A^0} versus m_{h^0} plot. The h^0A^0 analysis used is the 4C fit Higgs mass estimator case.

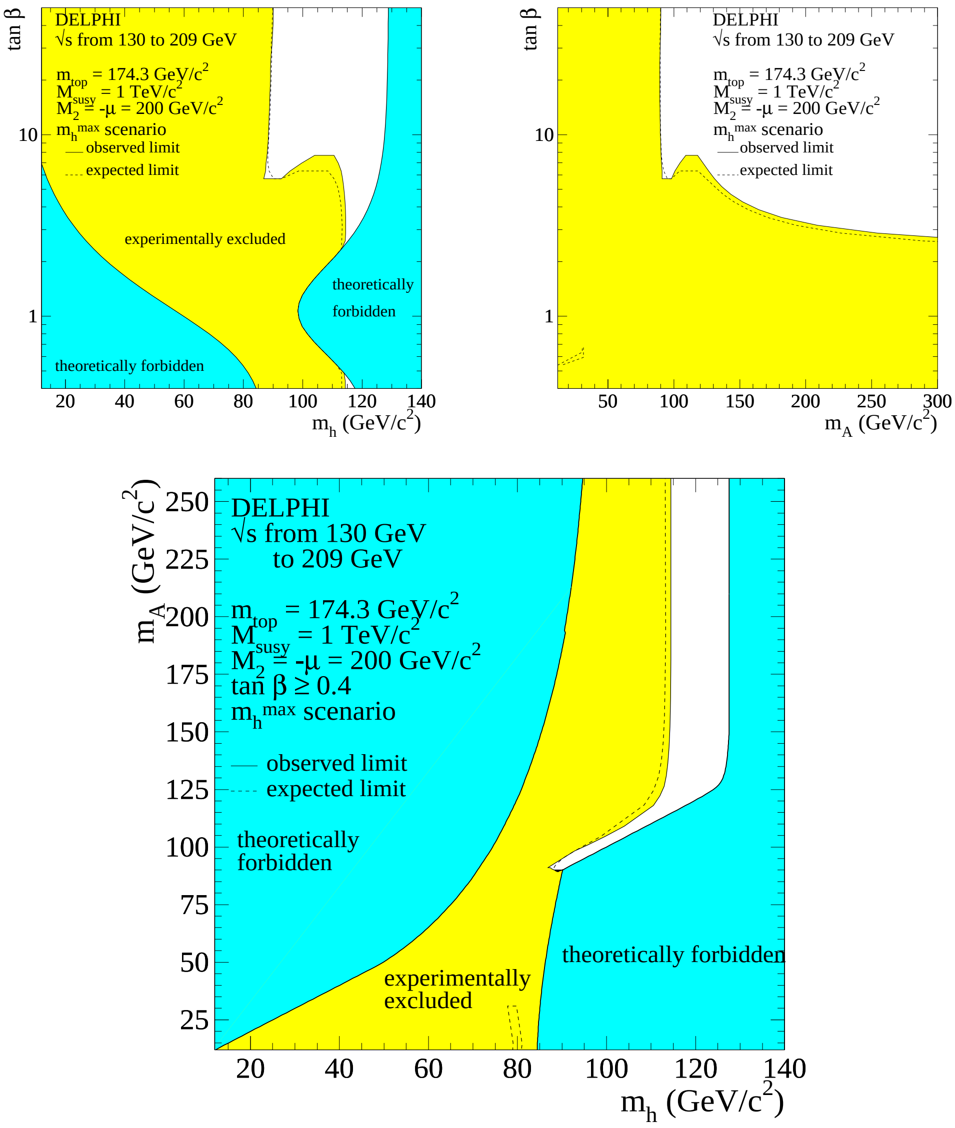


Figure 8.4: Exclusion plots for the $m_h^{\max}$ benchmark: On the top left is shown the $\tan\beta$ versus m_{h^0} plot, the $\tan\beta$ versus m_{A^0} plot is shown on the top right, whereas the bottom plot shows the m_{A^0} versus m_{h^0} plot. The h^0A^0 analysis used is the 4C fit Higgs mass estimator case.

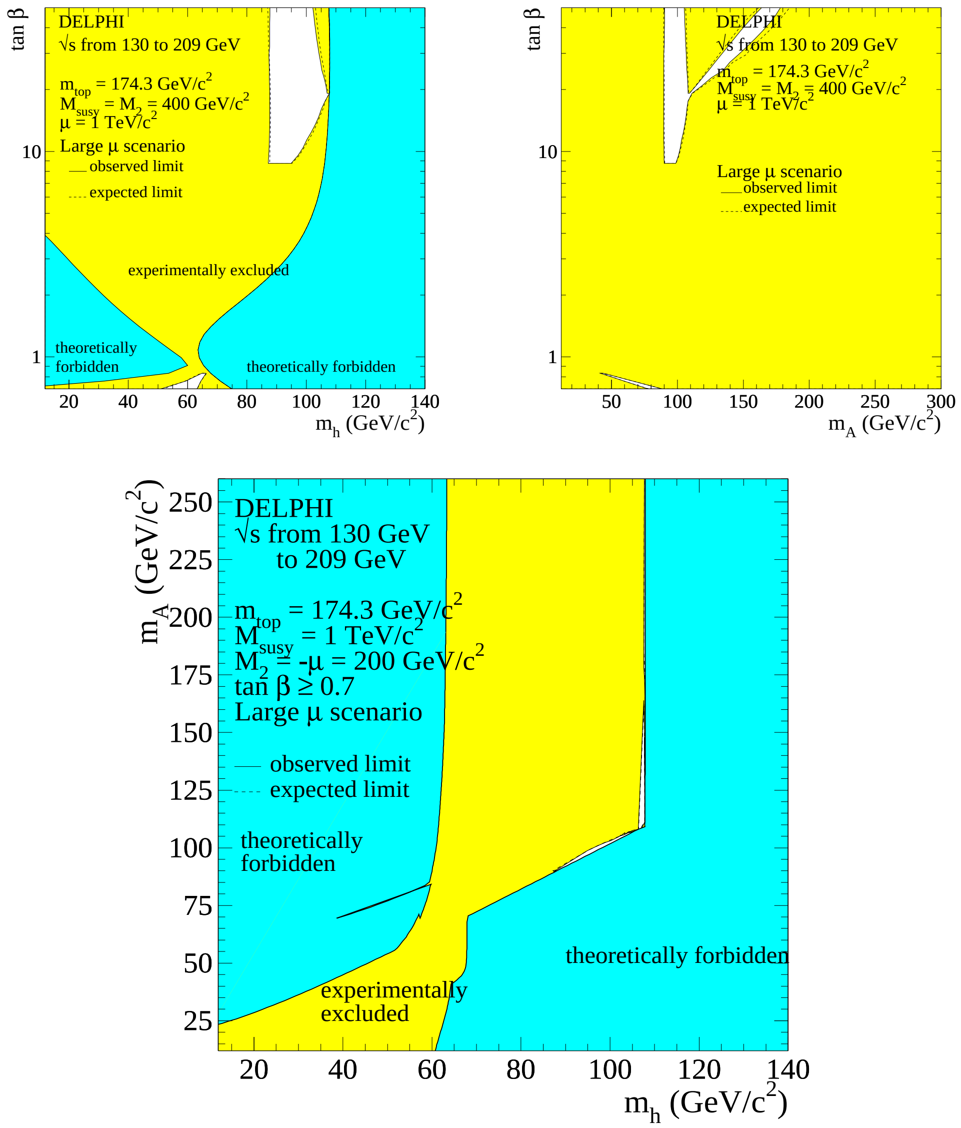


Figure 8.5: Exclusion plots for the large μ benchmark: On the top left is shown the $\tan\beta$ versus m_{h^0} plot, the $\tan\beta$ versus m_{A^0} plot is shown on the top right, whereas the bottom plot shows the m_{A^0} versus m_{h^0} plot. The h^0A^0 analysis used is the 4C fit Higgs mass estimator case.

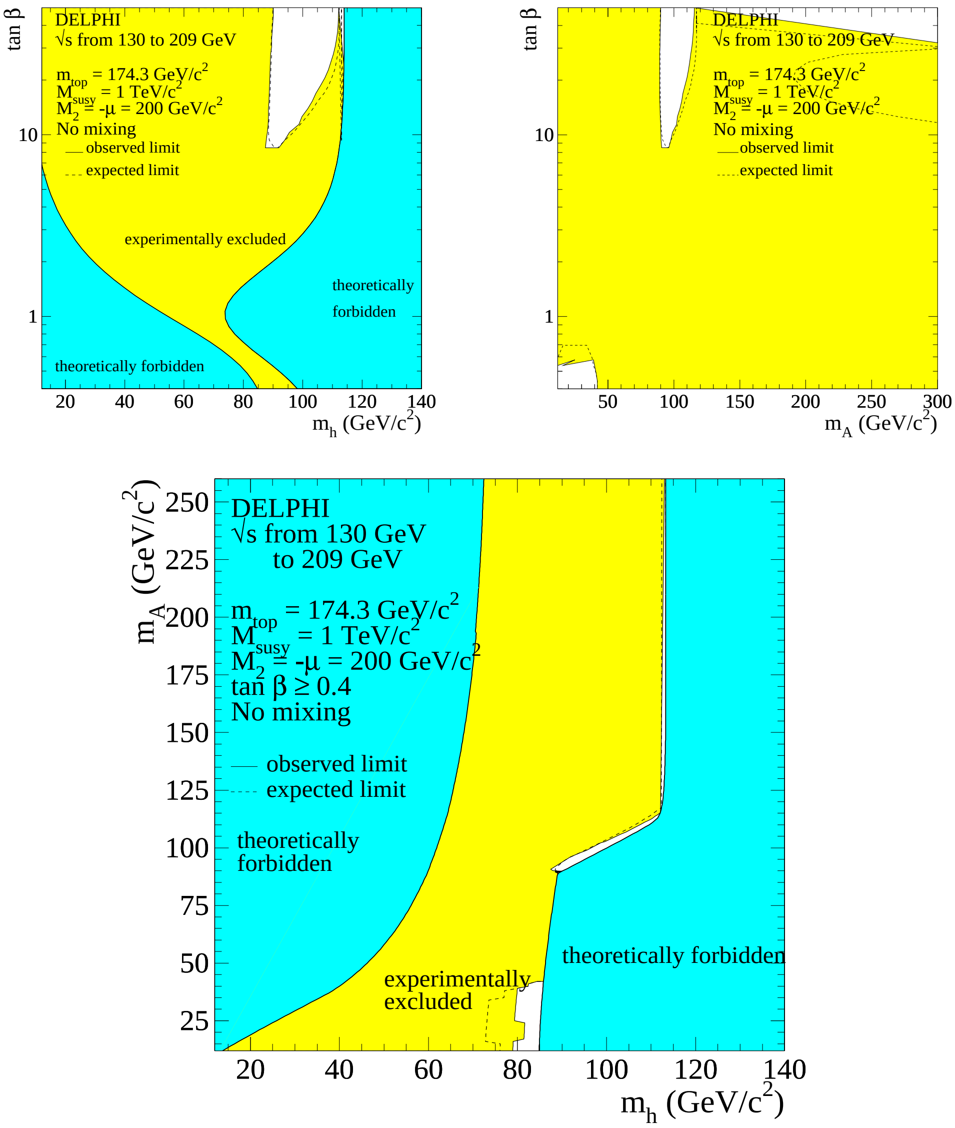


Figure 8.6: Exclusion plots for the no mixing benchmark: On the top left is shown the $\tan\beta$ versus m_{h^0} plot, the $\tan\beta$ versus m_{A^0} plot is shown on the top right, whereas the bottom plot shows the m_{A^0} versus m_{h^0} plot. The h^0A^0 analysis used is the 5C fit Higgs mass estimator case.

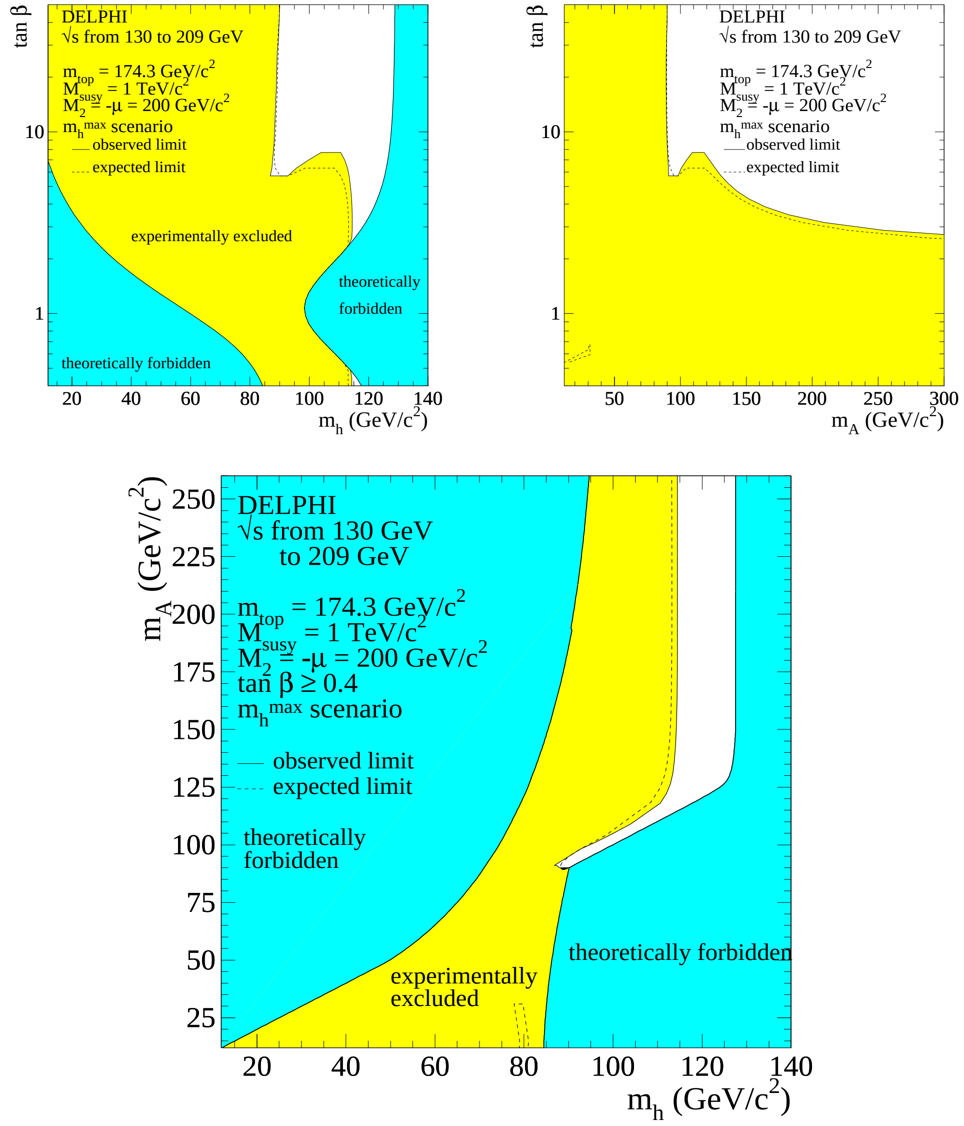


Figure 8.7: Exclusion plots for the $m_h^{\max}$ benchmark: On the top left is shown the $\tan\beta$ versus m_{h^0} plot, the $\tan\beta$ versus m_{A^0} plot is shown on the top right, whereas the bottom plot shows the m_{A^0} versus m_{h^0} plot. The h^0A^0 analysis used is the 5C fit Higgs mass estimator case.

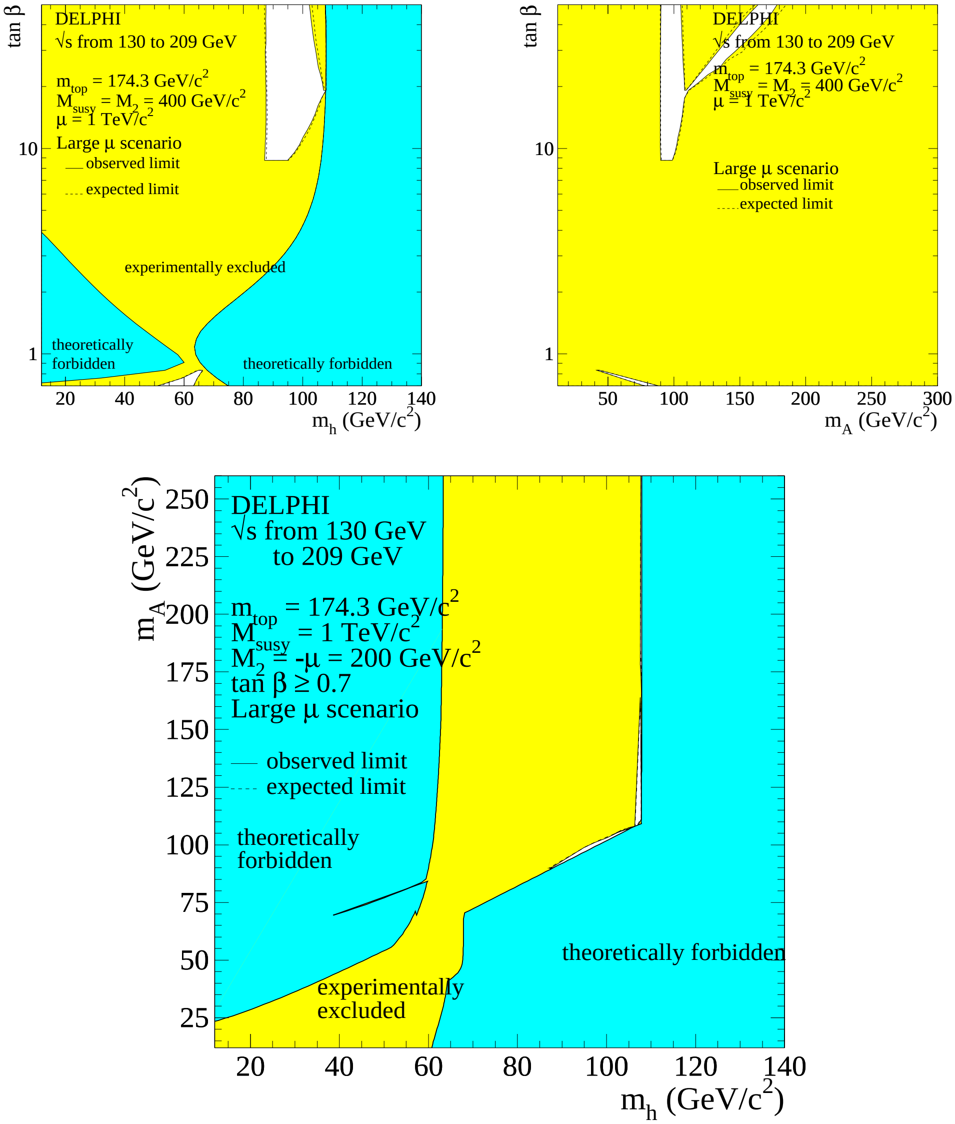


Figure 8.8: Exclusion plots for the large μ benchmark: On the top left is shown the $\tan\beta$ versus m_{h^0} plot, the $\tan\beta$ versus m_{A^0} plot is shown on the top right, whereas the bottom plot shows the m_{A^0} versus m_{h^0} plot. The h^0A^0 analysis used is the 5C fit Higgs mass estimator case.

Benchmark	m_{h^0}(GeV/c^2)	m_{A^0}(GeV/c^2)	$\tan\beta$
Official ALEPH search results [90]			
No mixing	89.8 (91.3)	90.1 (91.6)	0.5–6.2 (———)
$m_h^{\max}$	89.8 (91.3)	90.1 (91.6)	0.7–2.3 (———)
Large μ	— (—)	— (—)	——— (———)
SM	111.5 (114.2)	—	—
Official DELPHI search results [91] [1]			
No mixing	89.7 (88.8)	90.7 (89.7)	0.49–9.36 (0.49–8.49)
$m_h^{\max}$	89.7 (88.8)	90.7 (89.7)	0.49–2.36 (0.54–2.36)
Large μ	— (—)	— (—)	——— (———)
SM	114.3 (113.5)	—	—
Official L3 search results [92] [2]			
No mixing	83.2 (88.1)	83.9 (88.3)	——— (———)
$m_h^{\max}$	83.2 (88.1)	83.9 (88.3)	——— (———)
Large μ	— (—)	— (—)	——— (———)
SM	112.0 (112.4)	—	—
Official OPAL search results [93] [3]			
No mixing	80.9 (85.6)	82.3 (86.9)	1.2–3.8 (1.2–4.5)
$m_h^{\max}$	79.3 (85.1)	80.6 (86.9)	0.9–1.7 (1.0–1.9)
Large μ	79.8 (84.9)	82.4 (88.0)	0.7–4.8 (0.7–6.2)
SM	109.7 (112.5)	—	—

Table 8.3: Official search results for the four LEP experiments ALEPH, DELPHI, L3 and OPAL. The 95% CL excluded regions (for $\tan\beta$) and 95% CL lower limits (for m_{A^0} and m_{h^0}) of the three MSSM paramters m_{h^0}, m_{A^0} and $\tan\beta$ for all three MSSM benchmarks are shown, as well as the lower limit on the SM Higgs boson mass. Expected results are shown in parentheses after observed results. — indicates unavailable or inapplicable numbers.

---

[1] The MSSM neutral Higgs mass limits quoted are the limits for either choice of X_t and for $\tan\beta$ values above 0.49.

[2] The quoted MSSM neutral Higgs mass limits are obtained for either choice of X_t and for $\tan\beta$ values above 0.8.

[3] The MSSM neutral Higgs mass limits are obtained with the additional constraint of $\tan\beta > 1.2$.

much weight. The more fundamental problem connected to the definition of the likelihood ratio $\frac{s+b}{b}$ used in the final limits extraction (see section 7.2.3), is the most serious argument against the 5C fit method. This objection is not necessarily of paramount importance, as it mostly concerns the question of a discovery, and is not so important for an exclusion, but given that the results are indistinguishable between the two methods, the author feels that the arguments in favour of the 5C fit method have largely been rendered void, and that it therefore is better to use the more conventional method of the 4C fit. It should be noted that this could not be known a priori, as the 5C fit method does indeed have advantages over the 4C fit method, but, as was argued in section 7.4, this advantage appears in a part of the MSSM parameter space where the h^0A^0 channel is inferior to other channels, and the 5C fit method is therefore not optimal for this search channel. Nevertheless, for reference and to allow comparisons, the results from the 5C fit method are retained in the lower part of table 8.2.

8.5 Conclusions

A search for the MSSM Higgs bosons h^0 and A^0 in the four jet channel using the DELPHI detector and data of the runs of 1998, 1999 and 2000 has been performed. No signal was found in either channel, and 95% CL excluded regions in the MSSM parameter space of m_{A^0}, m_{h^0}, and $\tan\beta$ were mapped out for three different mixing benchmarks. Also, 95% CL absolute limits were set on the aforementioned parameters in the three mixing benchmarks.

- No mizing benchmark: The excluded regions are shown in figure 8.3, with the excluded ranges of the parameters being
  - $m_{A^0} < 89.3$ GeV/c^2 (expected: 89.2 GeV/c^2)
  - $m_{h^0} < 87.5$ GeV/c^2 (expected: 88.5 GeV/c^2)
  - $0.44 < \tan\beta < 7.69$ (expected: $0.40 < \tan\beta < 7.69$)
- $m_h^{\max}$ benchmark: The excluded regions are shown in figure 8.4, with the excluded ranges of the parameters being
  - $m_{A^0} < 89.4$ GeV/c^2 (expected: 89.4 GeV/c^2)
  - $m_{h^0} < 86.9$ GeV/c^2 (expected: 88.2 GeV/c^2)
  - $0.49 < \tan\beta < 2.36$ (expected: $0.54 < \tan\beta < 2.14$)
- Large μ benchmark: The excluded regions are shown in figure 8.5. As explained on page 182, it is customary to only quote a limit on the

$\tan\beta$ parameter for this benchmark, and the excluded region for this parameter is

  - $0.91 < \tan\beta < 6.74$ (expected: $0.91 < \tan\beta < 7.36$)

When interpreted in terms of the SM, a lower limit was set on the Higgs boson mass m_{H^0} at 114.3 GeV/c^2 (expected: 113.1 GeV/c^2).

Appendix A

Extracting the SM Higgs mass limit from MSSM exclusion

Given the difference between the one-doublet Higgs structure found in the minimal SM and the two-doublet structure in the MSSM, one might expect that the connection between the different parameters describing the two sectors is rather complex. Indeed, in the general MSSM case, deducing the SM Higgs boson limit is far from easy. However, there is a special case in which the MSSM Higgs sector reduces to the SM, making the SM Higgs boson mass limit determination a trivial task given the general MSSM exclusion plot. This special case occurs when the mass of the pseudo-scalar A^0, m_{A^0}, goes to infinity, in which case, at tree level, masses of both the charged Higgs bosons $H^{\pm}$ and the heaviest neutral Higgs boson, H^0, also go to infinity, effectively decoupling all these from the theory. In addition, the correction factor to the MSSM h^0Z^0 production factor of $\sin^2(\beta - \alpha)$ (see equation 4.2 on page 73) goes to unity, giving the h^0 the same couplings as the SM H^0, and thus all Feynman rules of the Higgs sector of the theory reduce to the ones in the SM [16, page 239 ff, 356]. If also, as is assumed in the benchmarks, the supersymmetric partners of the ordinary particles of the SM are too heavy to be kinematically available, the decay of the MSSM h^0 can only go through SM particles, and the whole Higgs search sector is reduced to the SM case.

The requirement for the Higgs sector of the MSSM to reduce to the SM, is thus that m_{A^0} goes to infinity. In an actual simulation of the MSSM Higgs sector, this is of course not a realizable scenario, as the computations require finite parameters. The important question then becomes where in the scanned MSSM parameter space the parameter m_{A^0} reaches large enough values so as to give results indistinguishable from the SM case. This happens in the $m_h^{\max}$ benchmark at low values of $\tan\beta$. Therefore, the limit on the

SM Higgs boson mass can be found as the limit on m_{h^0} at the low end of the range for $\tan\beta$ in the plot of $\tan\beta$ versus m_{h^0} in the $m_h^{\max}$ benchmark (figures 8.4 and 8.7). At these values of $\tan\beta$, the H^0Z^0 signal dominates over the h^0A^0, and the choice of 4C or 5C fit Higgs mass estimator in the h^0A^0 channel becomes negligible. The expected limit is in both cases 113.1 GeV/c^2, with the observed limit at 114.3 GeV/c^2. The SM results are also shown in table 8.2.

Appendix B

DELPHI internal note 2000-156 PHYS 872

This appendix includes a DELPHI internal note by the author describing the analysis technique used in the analysis presented in this work, known as the repeated 2D likelihood network.

DELPHI Collaboration

DELPHI 2000-156 PHYS 872
14 April, 2000

Search for pair produced neutral MSSM Higgs bosons in the 4b channel using a repeated 2D likelihood method.

J. Hansen
University of Oslo

Abstract

A search for pair produced neutral MSSM Higgs bosons decaying into 4 b-quarks using the DELPHI 1999 data is presented. The analysis method used is a repeated likelihood technique, based on the step by step combination of pairs of variables by the use of 2-dimensional likelihood functions. No evidence of an MSSM signal was found, and the analysis yields a lower limit on the mass of the pseudo-scalar Higgs boson A^0 at a value of $\tan\beta$=10 and in a maximal mixing scenario of 86.1 GeV/c^2.

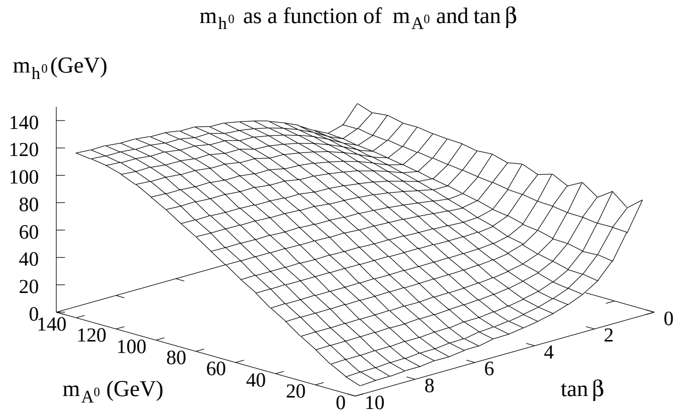


Figure 1: The relation between the model parameters $\tan\beta$ and m_A, and the mass of the lightest of the scalar Higgs particles, m_h

1 Introduction

The Higgs sector of the Minimal Supersymmetric Standard Model (MSSM) [1] contains 5 Higgs bosons, 3 neutral (the scalars h^0 and H^0, and the pseudo-scalar A^0) and 2 charged ($H^{\pm}$). At tree level there are only two relavant parameters (whereas radiative corrections introduce several others), which are usually chosen to be $\tan\beta$ (the ratio of the vacuum expectation values of the two Higgs doublets, $\tan\beta = \frac{v_2}{v_1}$) and m_A (the mass of the pseudo-scalar). At tree level, there are specific relations between these parameters and the masses of the different Higgs bosons:

$$m_A^2 = m_{H^\pm}^2 - m_W^2 \qquad m_{H,h}^2 = \frac{1}{2}\left[m_A^2 + m_Z^2 \pm \sqrt{(m_A^2+m_Z^2)^2 - 4m_Z^2 m_A^2 \cos^2 2\beta}\right] .$$

When including radiative corrections, these relations no longer hold, and the relation between the two main parameters $\tan\beta$ and m_A, and the mass of the lightest of the scalar Higgs particles, m_h, for the specific choice of a maximal mixing scenario (the common sfermion mass term M_S=1000 GeV/c^2, the mixing in the stop sector A_t=$\sqrt{6}M_S$ and the Higgs mixing parameter μ=-100 GeV/c^2), is shown in figure 1.

As can be seen in figure 1, the masses of the A^0 and h^0 are close to equal for relatively high values of $\tan\beta$ (above ∼5) and for m_A up to ∼120 GeV/c^2, which is above the region of interest for this analysis. At lower values of $\tan\beta$, down to values of ∼1, the mass of the h^0 lies below m_A in the m_A range in interest for this analysis, but the two masses are still fairly close.

An important production mode, especially at high values of $\tan\beta$, is the pair production of h^0A^0 through the s-channel $e^+e^- \to Z^{0*} \to h^0A^0$. Since the Higgs bosons have fermionic couplings proportional to the fermion mass, the decay process at LEP2 results in a final state which in most cases, i.e. above 80 % for a wide range of MSSM parameter choices, is a system of 4 b-quarks. Thus, the experimental signature for these events is a purely hadronic 4-jet system with a high b-quark content.

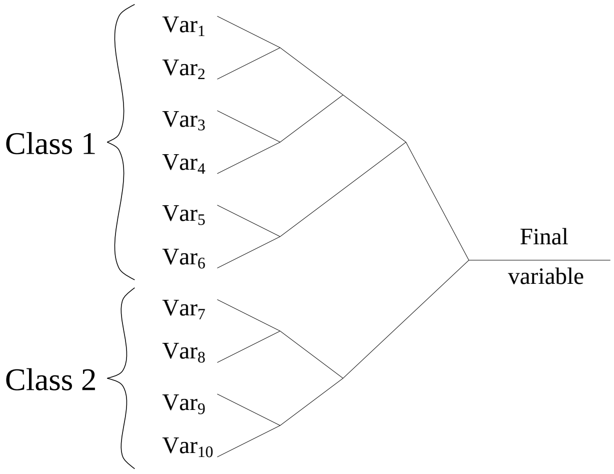


Figure 2: The general structure of the analysis method, for a generic case of 10 variables in 2 classes.

2 Method

The method implemented to separate the signal from the background is based on a repeated 2-dimensional likelihood approach. The method has, as multivariable discriminant methods in general, the goal of taking advantage of the discriminant power of the different variables while reducing the number of variables involved in the final steps of the analysis. Also, the analysis method makes optimization procedures and checks of the agreement between data and Monte Carlo simulation particularly easy, and in addition tries to use the correlations between variables, taking advantage of their discriminant power. By choosing variables where the correlation is expected to be large as belonging to the same class, and then doing the combination of variables within these classes at the earlier stages, one may hope to include most of the important correlations, and the method can thus be schematically described as in figure 2.

The combination of two variables is done by constructing the ratio of the likelihood functions in the signal+background and in the background only hypotheses ($\frac{s+b}{b}$) in the 2-dimensional space spanned by the two variables in question. This is, from a likelihood ratio point of view, the optimal way of combining them [2]. The $\frac{s+b}{b}$ distributions are then transformed in such a way that the distribution for the signal is a uniform distribution between 0 and 1, while the background will peak at lower values for this new distribution, which then becomes the output variable from the combination.

Due to limited statistics in the Monte Carlo simulations (leading to discontinuities in the 2-dimensional distributions), a method of smoothing the distributions has been adopted. This is done by applying the method of B-splines [3]. In addition to this, the tails of the distributions have been smoothed, while keeping the overall normalization, in order to take into account the low statistics in the less populated regions of the histogram, and to avoid overtraining. In fact, it is an inherent feature of the method that to avoid infinities in the $\frac{s+b}{b}$ distribution, a non-zero (albeit possibly very small) background has to be considered in the entire range of all the variables. This smoothing is done by choosing significant peaks in the distribution, and for all bins in the 2-dimensional histogram (after the SPLINE procedure has been applied) where the content is below the level corresponding to 1 simulated event, calculating a total distance measure: let L_i be the distance from the bin to peak number i, and let l_i be the distance from peak number i to the edge where the distribution reaches the level of 1 simulated event, in the direction from the peak to the bin in question. The total distance measure over a total of n different peaks is now defined as

$$l_{tot} = \frac{\sum_{i=1}^{n} \frac{l_i}{L_i^2}}{\sum_{i=1}^{n} \frac{1}{L_i^2}} \ .$$

This distance measure is then a weighted average over all peaks of the distance from the peak to the level corresponding to 1 simulated event in the direction of the bin in question, making the closest peak the most significant one. The bin content value is now decided by letting the distribution fall off as r^{-p}, where r is the shortest distance from the bin to the edge where the distribution reaches the level of 1 simulated event, and p is a function of l_{tot} which is increasing with decreasing l_{tot} ($p = 3 \cdot {l_{tot}}^{-0.25}$).

3 Analysis

3.1 Monte Carlo simulations and Data samples

For the simulation of the background, the PYTHIA [4] generator was used for the QCD samples ($e^+e^- \to q\overline{q}\ (\gamma)$), whereas EXCALIBUR [5] was used for the four-fermion background ($e^+e^- \to q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$, and $l\overline{l}q\overline{q}$, $l = e, \mu$ or τ, the $e^+e^-q\overline{q}$ and $\nu\overline{\nu}q\overline{q}$ final states being discarded due to their small cross-sections and 4-jet selection efficiencies). Due to the requirements of a large number of charged tracks and a large visible energy (see section 3.2), the $e^+e^- \to \gamma\gamma \to$ hadrons background samples need not be considered. The signal was generated using the HZHA [6] generator, using different values of $\tan\beta$ and m_A to cover a range of possible MSSM parameter choices. The number of generated events, together with the equivalent luminosity for each sample, is summarized in table 1. The data collected during the 1999 run amounts to 228.18 pb^{-1}, divided into 25.89 pb^{-1} at 192 GeV, 76.90 pb^{-1} at 196 GeV, 84.28 pb^{-1} at 200 GeV and 41.11 pb^{-1} at 202 GeV. Due to a limited amount of Monte Carlo simulated signal samples at 192 and 202 GeV, the signal samples used here were created using a boosting routine [7], boosting the signal samples from 196 GeV down to 192 GeV, and the samples from 200 GeV up to 202 GeV. A conservative value of 100 % for the statistical correlation between the boosted sample and the original was assumed.

3.2 Event selection

The analysis starts with the standard track selection [8] and 4-jet event selection [9] applied in the DELPHI Higgs searches in 4-jet channels.

The track selection consists of the following cuts:

- Charged particles
  - Track momentum larger than 100 MeV

Channel		Generator	X-section (pb)	Nr. of events	Equiv. lumi. (pb^{-1})
$q\overline{q}$ (γ) (ZGPY)		PYTHIA	88.1	1926560	$\sim$ 22200
$q\overline{q}q\overline{q}$, $q\overline{q}l\nu$, $l\nu l\nu$ (WWEX)		EXCALIBUR	18.6	816873	$\sim$ 44000
$l\overline{l}q\overline{q}$, $l = \mu$ or τ (LLQQ)		EXCALIBUR	0.368	86423	$\sim$ 235000
Signal, tan β=2					
m_A (GeV/c^2)	80	HZHA	0.04181	7994	$\sim$ 191000
	85	HZHA	0.03160	12792	$\sim$ 410000
	90	HZHA	0.02309	7996	$\sim$ 350000
	95	HZHA	0.01618	7998	$\sim$ 503000
Signal, tan β=20					
m_A (GeV/c^2)	80	HZHA	0.07366	7994	$\sim$ 110000
	85	HZHA	0.04668	7640	$\sim$ 169000
	90	HZHA	0.02418	9596	$\sim$ 455000
	95	HZHA	0.00795	7998	$\sim$ 1710000

Table 1: The Monte Carlo generated number of events for the different background and signal channels, with their corresponding simulated luminosity. The numbers are a sum over all four centre-of-mass energies, the cross-sections (which for the signal are multiplied by the branching ratio $h^0A^0 \rightarrow 4b$) being weighted with luminosity and 4-jet selection efficiency for each centre-of-mass point.

  - Impact parameter less than 4 cm in the transverse direction
  - Impact parameter less than 10 cm in the z direction

- Neutral particles
  - For the Electromagnetic clusters: Energy larger than 200 MeV
  - For the Hadronic clusters: Energy larger than 500 MeV

The 4-jet event selection consists of the following cuts:

- Minimum 18 charged tracks
- Visible energy larger than 60 % of $\sqrt{s}$
- Neutral energy less than 50 % of $\sqrt{s}$
- No visible γ with an energy above 30 GeV
- The energy of an invisible γ returned from SPRIME+ [11] less than 30 GeV
- At least 1 charged particle per jet, and a jet mass of at least 1.5 GeV/c^2 in a forced 4-jet configuration
- The Fox-Wolfram momenta $H_2 + H_4$, normalized to H_0, less than 1.1
- Event thrust less than 0.92

The remaining efficiencies after these cuts for the different Monte Carlo simulated backgrounds and signals are presented in the left half of table 2. The remaining number of data events is 2440. The quoted uncertainties on the Monte Carlo simulated backgrounds include a common 1 % relative error accounting for the uncertainties in centre-of-mass energy, luminosity and cross-sections, and in addition a systematic error on the difference between different generators of the various background channels and the modelling of the successive cuts, estimated to be 5.7 % for the PYTHIA background, and 2 % for the EXCALIBUR background [10]. The uncertainties on the Monte Carlo simulated signals have a 1 % relative contribution from the uncertainties in the centre-of-mass energy, luminosity and cross-sections, and an additional absolute uncertainty in the efficiency of 1.5 % coming from possible small biases introduced by the method itself, estimated on different half samples; see subsection 3.3.

3.3 The construction of the discriminating variable in the 2D likelihood net

The variables chosen, in two classes, for distinguishing the signal from the background, are listed below. The algorithm used for forcing a 4-jet configuration, is the Durham [12] algorithm.

- Event shape variables
  - Tr_{min}, the minimum number of tracks in a jet in a forced 4-jet configuration
  - E_{min}, the minimum energy of a jet in a forced 4-jet configuration
  - α_{min}, the minimum angle between two jets in a forced 4-jet configuration
  - H_2, the second Fox-Wolfram moment, normalized to H_0
  - H_4, the fourth Fox-Wolfram moment, normalized to H_0
  - M_L, the light jet mass [4, page 276]
  - Event thrust
  - Y_{34}, the y-cut transition value between 3 and 4 jets
- B tagging variables
  - B_1, the most b-like value for a jet in a forced 4-jet configuration
  - B_2, the second most b-like value for a jet in a forced 4-jet configuration
  - B_3, the third most b-like value for a jet in a forced 4-jet configuration
  - B_4, the fourth most b-like value for a jet in a forced 4-jet configuration
  - B_{HA}, the total event B tagging variable, defined as the sum of B_i, $i = 1, \ldots, 4$[1]

---

[1] Due to the fact that the analysis method does not take all possible correlations into account, only the ones between the two variables to be combined into one in a combination step, including variables that are simple combinations of other variables in the analysis might contribute discriminating information. This has been seen to be the case for the B_{HA} variable.

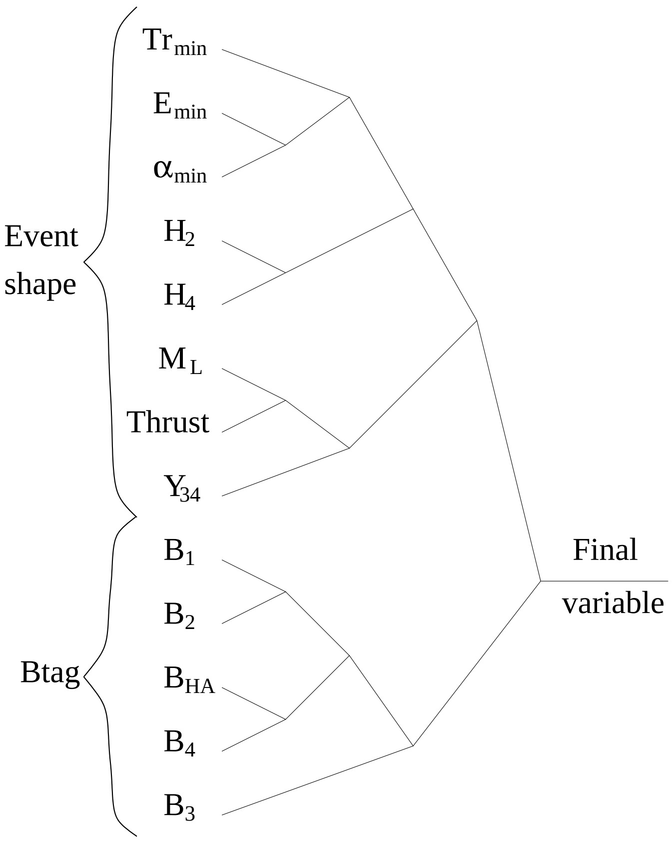


Figure 3: The combination of the 13 input variables in this analysis

The distributions of these variables for the simulated Monte Carlo background and signal, together with the data, are shown in figures 4 and 5; the distribution for the B_{HA} variable being omitted, since this is simply the sum of the four other B tagging variables.

The dependence of both the event shape and B tagging variables on the two parameters of the model (m_A and $\tan\beta$) is small in the parameter region of interest, with the high values of m_A, i.e. values of m_A approaching the beam energy, giving a signal most clearly separated from the QCD background. In fact, table 2 shows that the obtained efficiencies depend only weakly on the parameter values, at least for relatively high values of m_A. The signal used to construct the $\frac{s+b}{b}$ distributions is therefore chosen to be a mix of the Monte Carlo simulated signal for m_A=80, 85, 90, 95 GeV/c^2 and $\tan\beta$=2,20, in order to increase the signal statistics. As seen from figure 1, the variation of m_h folows closely the one of m_A within this parameter range. At $\tan\beta$=20, the two masses are close to equal, whereas at $\tan\beta$=2, m_h increases slower than m_A.

The variables and classes are combined to form the final discriminating variable according to the repeated 2D likelihood method as shown in figure 3. Due to the risk of overtraining, the Monte Carlo simulated signal and background samples were initially divided into two equally large sub-samples, with the $\frac{s+b}{b}$ distributions for the repeated 2D likelihood net being built up on one half, and the analysis performed on both, in order to

		4-jet selection level		Final cut level	
Channel		Efficiency (%)	Events	Efficiency (%)	Events
$q\bar{q}$ (γ)		3.74 ± 0.12	752 ± 24	$(3.62 \pm 0.18)\cdot 10^{-2}$	7.25 ± 0.37
$q\bar{q}q\bar{q}$, $q\bar{q}l\nu$, $l\nu l\nu$		40.32 ± 0.44	1710 ± 21	$(17.47 \pm 0.52)\cdot 10^{-2}$	7.41 ± 0.22
$l\bar{l}q\bar{q}$, $l = \mu$ or τ		15.47 ± 0.21	12.99 ± 0.19	$(16.5 \pm 1.5)\cdot 10^{-2}$	0.139 ± 0.013
Total background		—	2475 ± 32	—	14.80 ± 0.43
Data		—	2440	—	15
Signal, $\tan\beta$=2					
m_A (GeV/c^2)	80	89.7 ± 1.2	8.56 ± 0.12	60.0 ± 1.3	5.73 ± 0.13
	85	91.4 ± 1.1	6.59 ± 0.09	60.3 ± 1.3	4.35 ± 0.10
	90	92.9 ± 1.1	4.90 ± 0.07	60.5 ± 1.3	3.19 ± 0.07
	95	93.6 ± 1.1	3.45 ± 0.05	60.0 ± 1.3	2.21 ± 0.05
Signal, $\tan\beta$=20					
m_A (GeV/c^2)	80	94.1 ± 1.1	15.82 ± 0.21	61.7 ± 1.3	10.37 ± 0.23
	85	94.5 ± 1.1	10.07 ± 0.13	62.3 ± 1.3	6.63 ± 0.15
	90	94.2 ± 1.1	5.20 ± 0.07	62.2 ± 1.3	3.44 ± 0.08
	95	92.1 ± 1.1	1.67 ± 0.03	60.4 ± 1.3	1.09 ± 0.03

Table 2: The 4-jet and final cut selection efficiencies and number of events for the different Monte Carlo simulated backgrounds and signals, together with the data. The numbers are a sum over the four centre-of-mass energies, the efficiencies being weighted with the luminosity for each centre-of-mass point.

see the difference in the two samples. The result is presented in figure 6. The difference is seen to be small, with the only evidence of overtraining appearing at extremely high purity for the signal, and correspondingly low level of background. For this reason, and in order to minimize the statistical error and increase the accuracy, the method adopted was to do both the initial build-up of the distributions and the analysis on the entire Monte Carlo samples. An absolute systematic error in the signal efficiency of 1.5 %, which is a value that the difference band in figure 6 is seen to stay within for all but the highest purity region, was included.

The distribution of the final variable is shown in figure 7 for the different Monte Carlo simulated backgrounds and signal, together with the data. Cuts at different values for this variable give corresponding levels of background and efficiency, thus generating a curve of background level vs. efficiency; this is shown in figure 8 for a standard choice of m_A and $\tan\beta$. The figure shows good agreement between the observed data and the expected background, the two lying within one standard deviation of each other in all regions except for a small area around an efficiency of 60 %.

A Higgs boson mass estimator for each event is needed in order to distinguish between different signal hypothesis. This is constructed by forcing a 4-jet configuration, and then performing a 4C fit, the constraints being conservation of energy and momentum in the event. The jet pairing is chosen by looking at the difference in the invariant mass of the two dijet objects, and choosing the pairing with the smallest difference. This is a natural choice for signal hypothesis at large values of $\tan\beta$, since the masses of A^0 and h^0 are close to equal in this area. At lower values of $\tan\beta$, the mass difference between A^0 and

h^0 starts to increase, but is still sufficiently small that the choice of pairing is reasonable. The mass estimator is now chosen as the sum of the invariant masses of the two dijet objects. The final discriminating variable is then used, together with this Higgs mass estimator, as the 2D input to the final statistical analysis. The distributions are shown in figure 9.

In order to show the most signal-like region of the analysis, a cut in the final discriminating variable at a value of 0.35 was performed, leaving the number of background and signal events cited in the right half of table 2. The distribution of the sum of the dijet masses of events above this cut in the discriminating variable, together with the expected background and signal, is shown in figure 10.

4 Results

The extraction of the 95 % CL lower limits on the different Higgs mass hypotheses is done using the ALRMC [2] program, and the resulting signal confidence level curves are shown in figure 11. These are compared to the ones derived using the standard DELPHI analysis in the 4b channel, and the expected results in the $m_A/\tan\beta$ plane are seen to be very similar, with the difference in the excluded mass ranges between the two analyses being at most a few hundred MeV in the entire $m_A/\tan\beta$ plane shown in figure 11.

The resulting observed 95 % CL lower limits for the two $\tan\beta$ cases can be seen from figure 11 to be 75.9 GeV/c^2 for $\tan\beta$=2, and 86.1 GeV/c^2 for $\tan\beta$=20. The expected values are 80.9 GeV/c^2 for $\tan\beta$=2, and 85.2 GeV/c^2 for $\tan\beta$=20. Also, the confidence limits are seen to be approximately constant in a range from $\tan\beta$=5 to 20, and the limit in the middle of that range, at $\tan\beta$=10, is therefore taken to be the value at which the comparison with the standard DELPHI analysis is made. The analysis presented in this note yields an expected limit at 85.2 GeV/c^2, whereas the observed limit is at 86.1 GeV/c^2. The expected limit for the standard DELPHI analysis is 85.3 GeV/c^2, which is 0.1 GeV/c^2 above the results from the analysis presented in this note. The observed result for the standard DELPHI analysis is seen to lie somewhat below this, at 82.6 GeV/c^2.

5 Conclusion

In this note, a repeated 2D likelihood analysis method has been presented, and applied to the search for the $hA \rightarrow 4b$ topology with data from the LEP running of 1999. No evidence of an MSSM signal was found, and the resulting 95 % CL lower limits on the mass of the pseudo-scalar A^0 was found to be 75.9 GeV/c^2 for $\tan\beta$=2, and 86.1 GeV/c^2 for $\tan\beta$=20.

References

[1] J. F. Gunion, H. E. Haber, G. L. Kane, S. Dawson *The Higgs Hunter's Guide*, Addison-Wesley, (1990).
G. Altarelli, T. Sjöstrand and F. Zwirner *Cern Yellow Reports: Physics at LEP2*, CERN96-01 (1996).

[2] A. L. Read *Optimal Statistical Analysis of Search Results based on the Likelihood Ratio and its Application to the Search for the MSSM Higgs Boson at* $\sqrt{s} = 161$ *and* 172 *GeV*, DELPHI 97-158 PHYS 737 (1997).

[3] B. Schorr *Spline estimation of Distributions and Density Functions* Technical Report DD/75/13, CERN (1975)

[4] T. Sjöstrand , Comp. Phys. Comm. 39 (1986) 347
Pythia 5.7 and Jetset 7.4, Physics and Manual, CERN-TH 7112/93 (1993)

[5] F. A. Berends, R. Kleiss, R. Pittan, Nucl. Phys. B426 (1994) 344; Nucl. Phys. (Proc. Suppl.) 37B (1994) 163-168.

[6] P. Janot in CERN Report 96-01, Vol. 2, 309. Comp. Phys. Comm. 39 (1986) 347.

[7] P. Roudeau *Subroutine boostmc* /afs/cern.ch/delphi/tasks/lep200/searches/teamc/boostmc.car

[8] P. Lutz *DELPHI TEAM C standard track selection*,

`http://delphiwww.cern.ch/~pubxx/delwww/tasks/lep200/searches/teamc/www/teamc_dq_particle_selection.html`

[9] P. Lutz *DELPHI TEAM C 4-jet selection*,

`http://delphiwww.cern.ch/~pubxx/delwww/tasks/lep200/searches/teamc/www/news/4jsel_proposal97.html`

[10] P. Lutz *Preselection performances for processing D2*,

`http://delphiwww.cern.ch/~lutzp/delwww/presel/presel_189D.html`

[11] P. Abreu, D. Fassouliotis, A. Grefrath, R. P. Henriques, L. Vitale *SPRIME A Package for Estimating the Effective* $\sqrt{s'}$ *Center of Mass Energy in* $q\overline{q}\gamma$ *Events*, DELPHI 96-124 PHYS 632 (1996).

[12] S. Catani, Yu. L. Dokshitzer, M. Olsson, G. Turnoch, B. R. Webber, Phys. Lett B269 (1991), 432

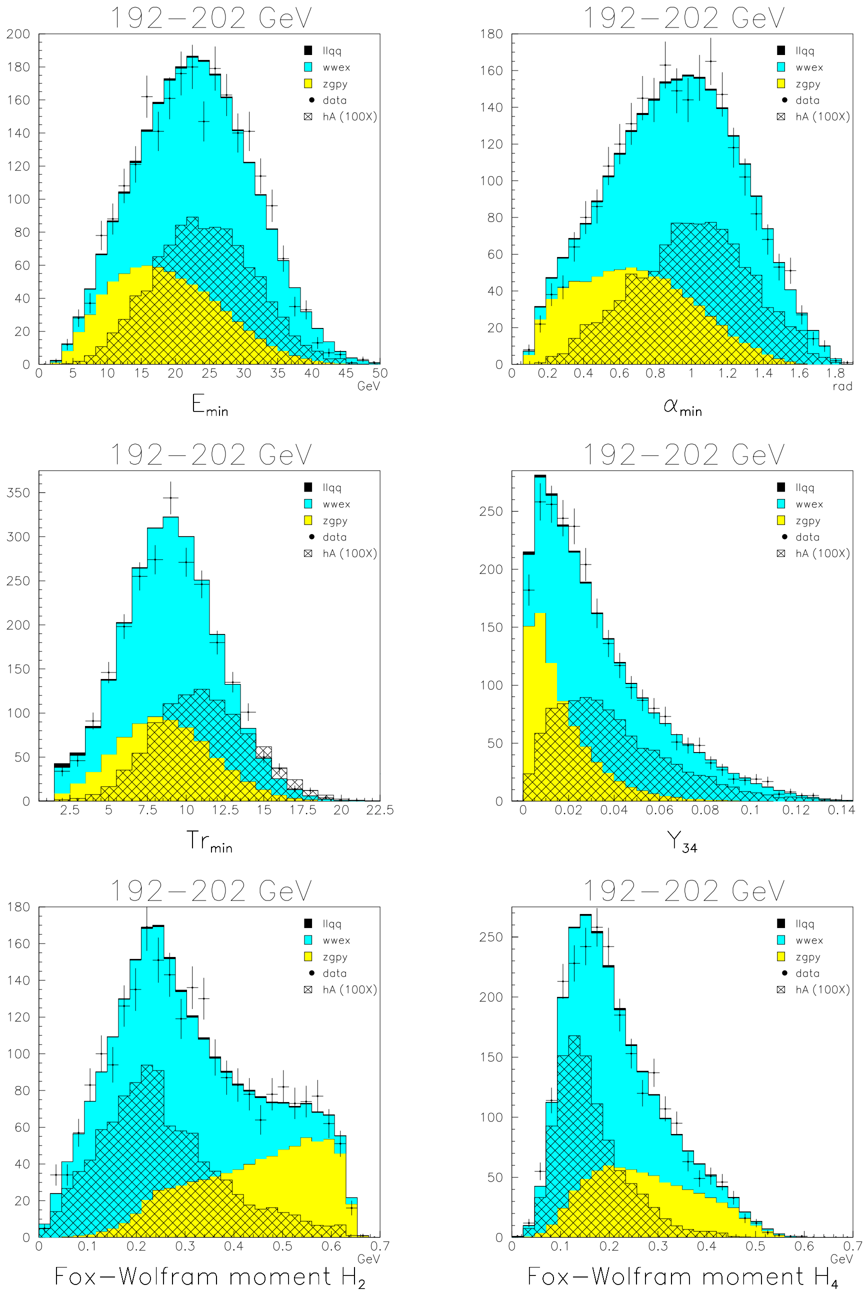


Figure 4: The first six input variables for the 2D likelihood net. The hA signal refers to the m_A=85 GeV/c^2, $\tan\beta$=20 Monte Carlo simulated signal sample. The signal has been scaled up by a factor of 100 where indicated.

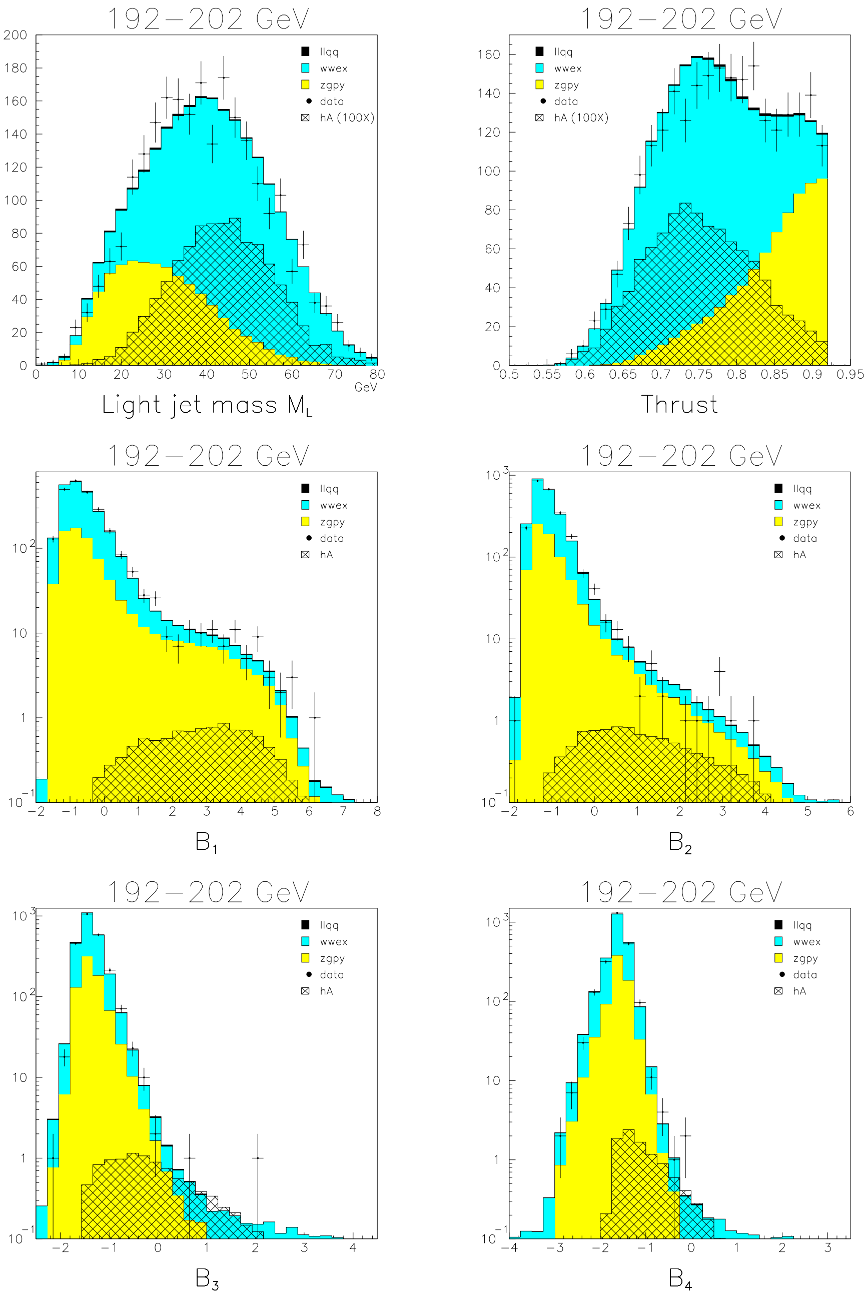


Figure 5: The last six input variables for the 2D likelihood net. The hA signal refers to the m_A=85 GeV/c^2, tan β=20 Monte Carlo simulated signal sample. The signal has been scaled up by a factor of 100 where indicated.

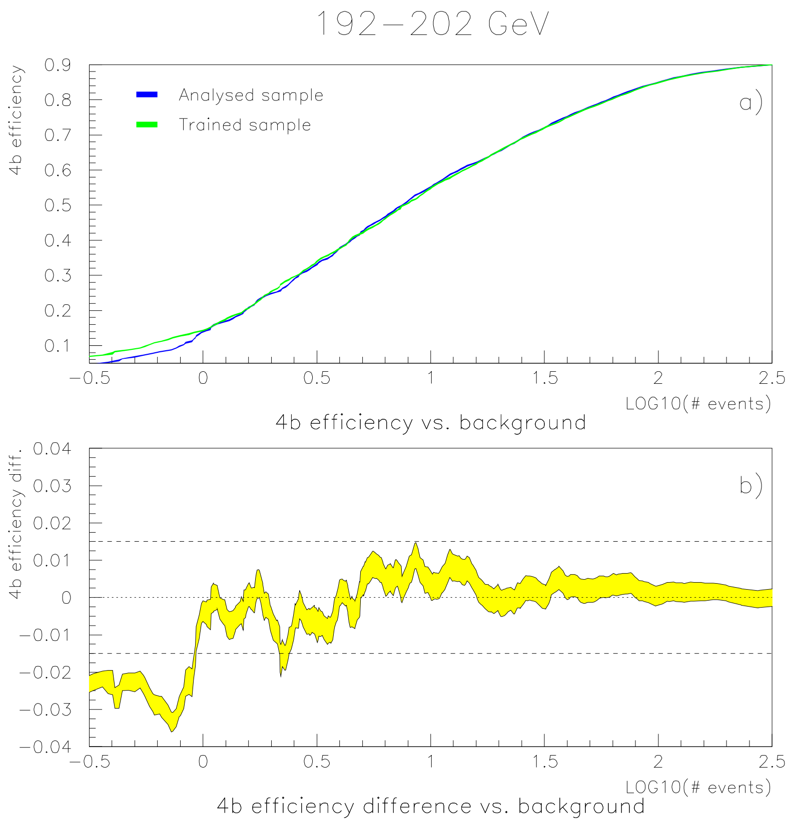


Figure 6: Figure showing the difference between the analysis result on the trained vs the untrained sample. a) shows the efficiency vs the logarithm of the total background for the two cases (trained and untrained), whereas b) shows the difference (untrained minus trained). In both cases the error band indicates the statistical error in the Monte Carlo simulated signal and background samples, as well as the systematics in the cross-sections, luminosities and background Monte Carlo simulated samples. The signal sample used is the sum of the Monte Carlo simulated signal samples for m_A=75, 80, 85, 90 GeV/c^2 and $\tan\beta$=2,20.

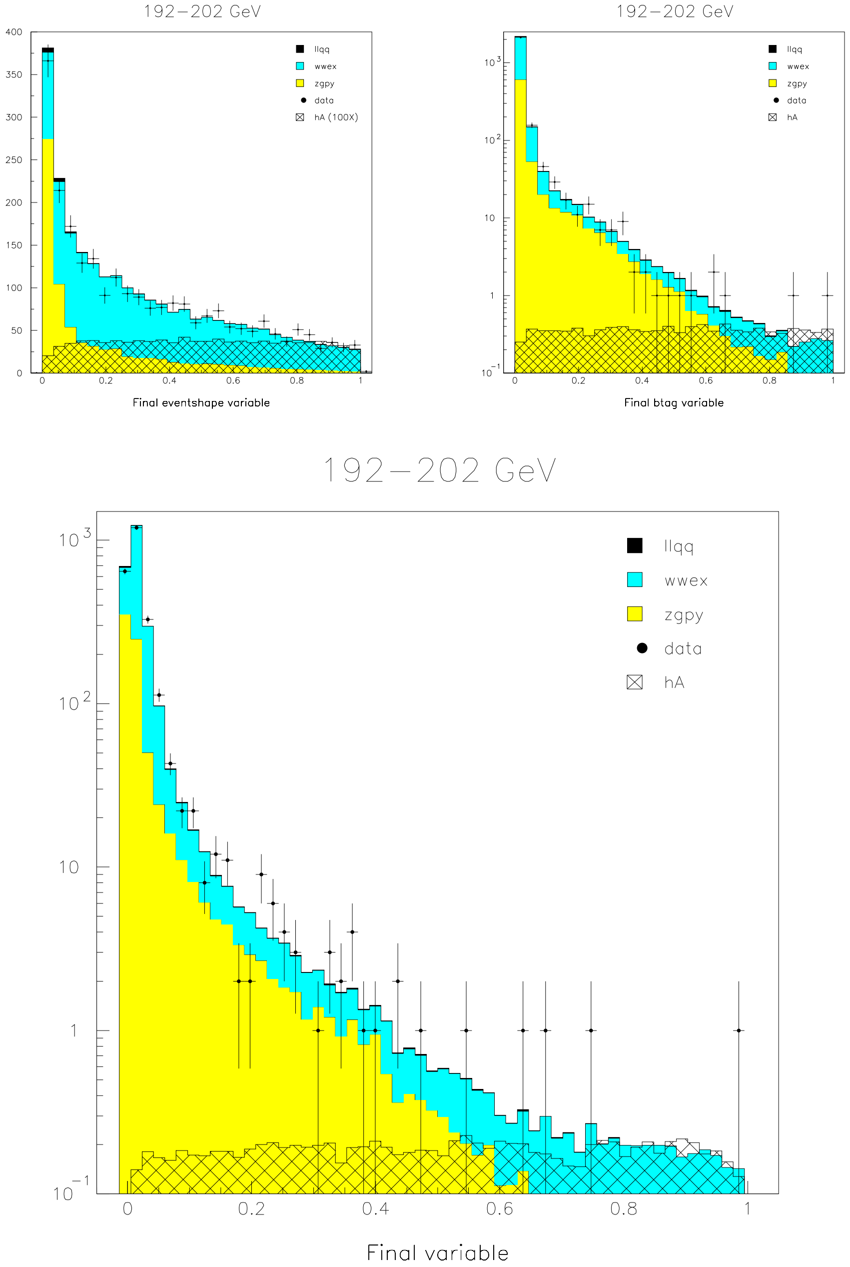


Figure 7: Figure showing the variables in the final step. On top, the inputs to the final variable (the final event shape and the final B tagging variables) are shown, whereas the final discriminating variable is shown in the bottom plot. The signal sample used is the Monte Carlo simulated signal for m_A=85 GeV/c^2, tan β=20.

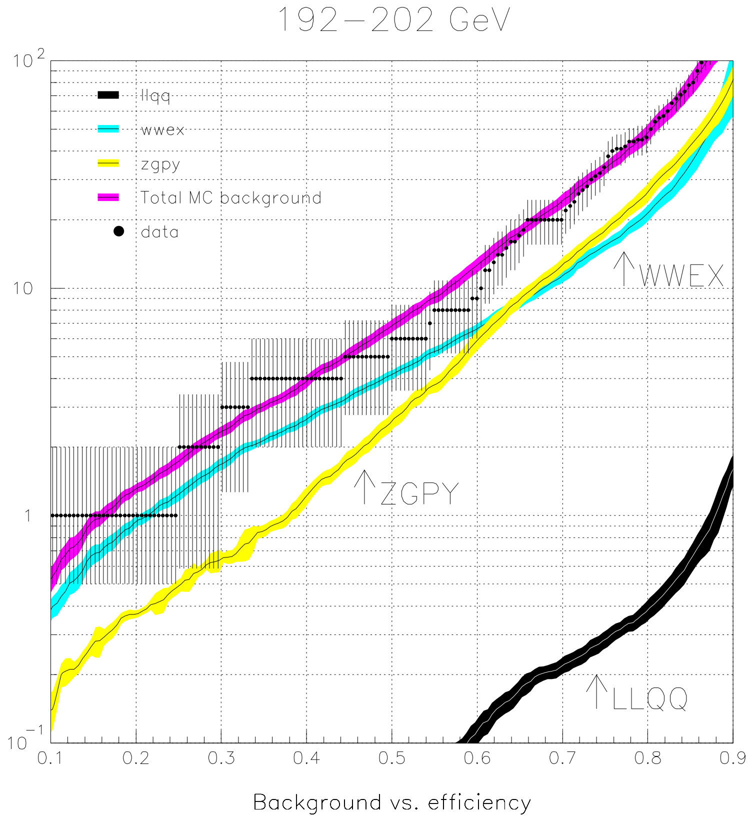


Figure 8: Figure showing the background level vs. the 4b efficiency for the different Monte Carlo simulated background samples, together with the data. The plot is made by adding the background levels for all four centre-of-mass points with a cut in the final variable (not necessarilly at exactly the same value for all four centre-of-mass points) giving the same signal efficiency for all centre-of-mass points. The error bands include systematic errors. The signal sample used is the Monte Carlo simulated signal for m_A=85 GeV/c^2, tan β=20.

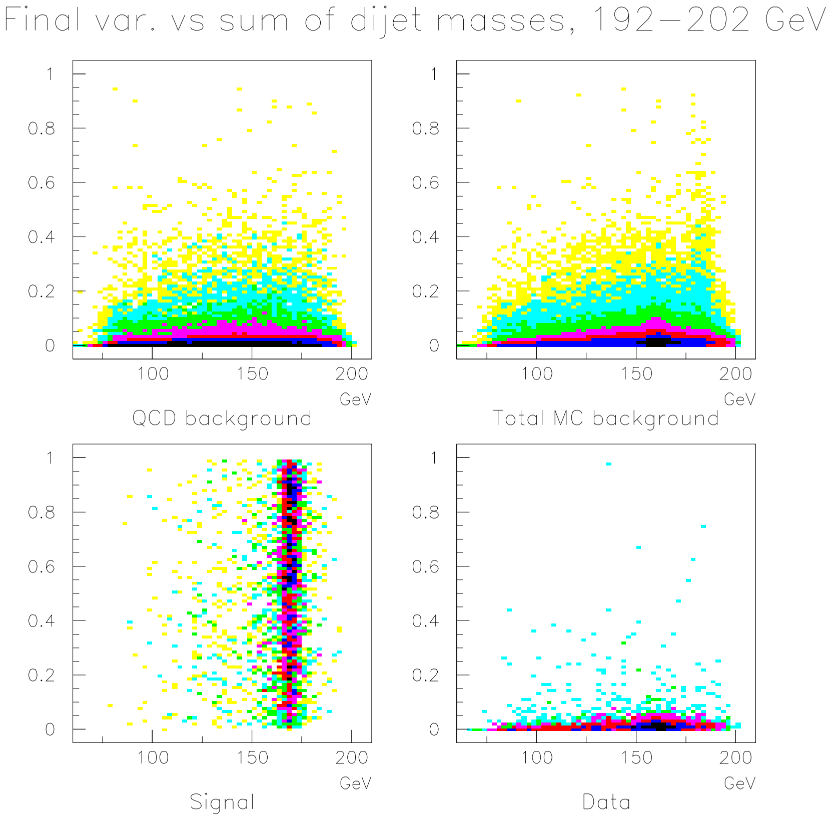


Figure 9: Figure showing the final variable vs the sum of the dijet masses. The signal sample used is the Monte Carlo simulated signal for m_A=85 GeV/c^2, $\tan\beta$=20.

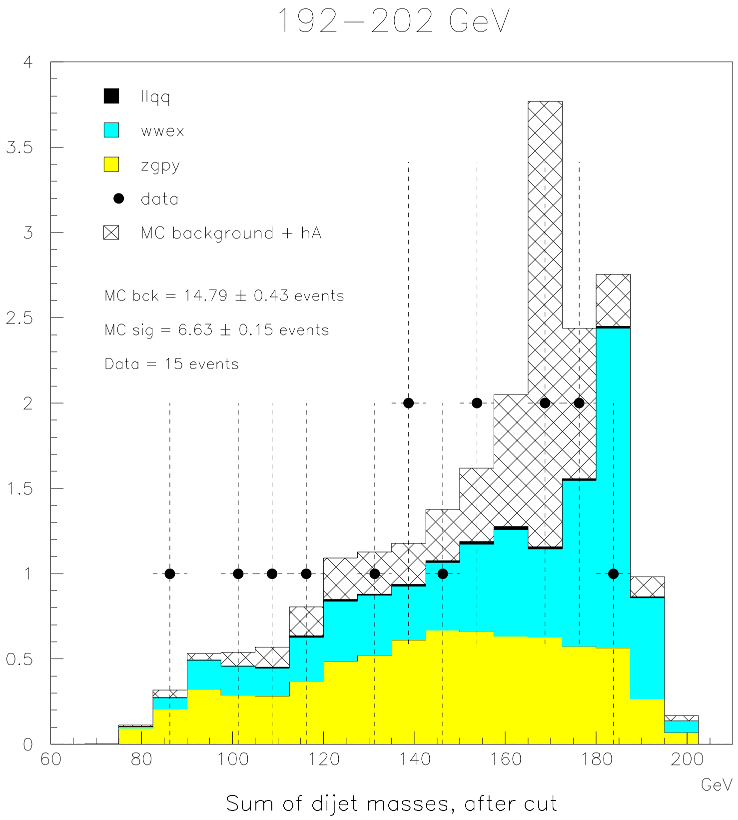


Figure 10: Figure showing the sum of the dijet masses after a cut in the final variable at a value of 0.35. The signal sample used is the Monte Carlo simulated signal for m_A=85 GeV/c^2, $\tan\beta$=20.

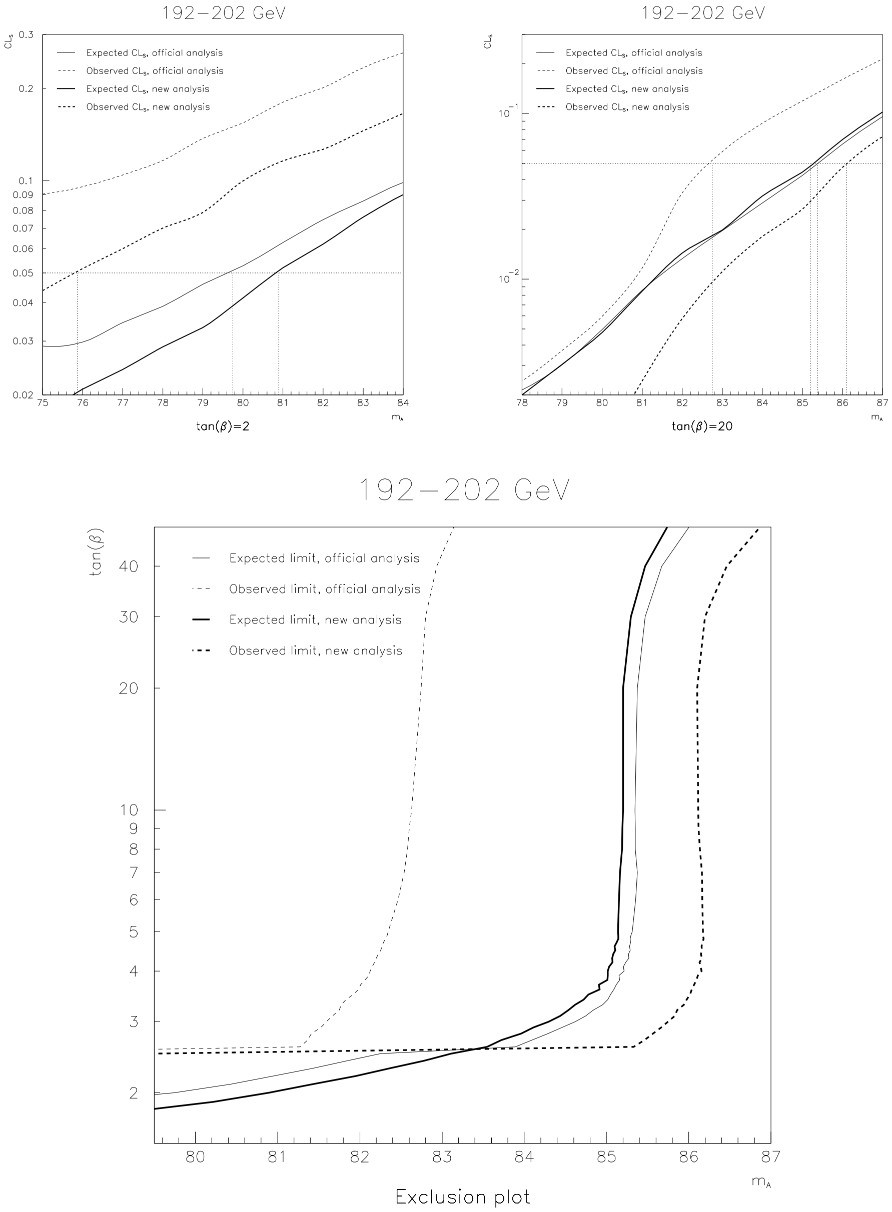


Figure 11: Figure showing different signal confidence levels and limits. On the top left is shown the expected and observed signal confidence levels for the $\tan\beta$=2 case, whereas the same for the $\tan\beta$=20 case is shown on the top right. The exclusion curve in the $m_A/\tan\beta$ plane, i.e. the points where the signal confidence reaches the value of 0.05, is shown on the bottom. The excluded regions lie to the left of the curves.

Appendix C

Contribution to the Lake Louise Winter Institute 2001, “Fundamental Interactions” Conference

This appendix includes the abstract from a talk on the Higgs searches in DELPHI in the year 2000, given by the author. The abstract is to be included in the 2001 proceedings from the conference, to be published by World Scientific.

HIGGS SEARCHES WITH THE DELPHI DETECTOR

JØRGEN HANSEN
Department of Physics, University of Oslo, P.O. Box 1048 Blindern,
0316 Oslo, Norway
E-mail: jorgen.hansen@fys.uio.no

Preliminary results for the search for the Higgs boson(s) are presented in the framework of the Standard Model (SM) and the Minimal Supersymmetric Standard Model (MSSM). The results from the year 2000 running are presented, and show no sign of a Higgs signal. Lower limits on the Higgs boson mass(es) are obtained using the data from the entire LEP-*II* running period (presented in full in reference [1]), which in a few cases are updates of what was shown at the conference.

1 Introduction

In the Standard Model, a single complex scalar Higgs doublet is sufficient in order to accomodate the Electroweak symmetry breaking and the assignment of fermionic masses. After the Higgs mechanism has absorbed three of the four parameters of the complex field, a single free parameter, the Higgs boson mass m_{H^0}, remains. The tree level production of the SM Higgs boson is at LEP-*II* the s-channel Z^0 exchange Higgs strahlung process $e^+e^-{\to}H^0Z^0$.

In a two doublet Higgs model, which is needed for a supersymmetric theory, five Higgs bosons remain: two neutral scalars h^0 and H^0(by definition, $m_{h^0}{<}m_{H^0}$), one neutral pseudoscalar A^0 and a charged particle/anti-particle pair H^+H^-. In addition to the Higgs strahlung process, Higgs bosons can now be produced through s-channel Z^0 exchange Higgs pair production, either as a pair of charged Higgs bosons $e^+e^-{\to}H^+H^-$, or as a pair of neutrals, one of which must be the A^0, $e^+e^-{\to}h^0A^0$ or H^0A^0.

During the running of LEP in the year 2000, DELPHI has collected a total of 224.3 pb^{-1} at an efficiency of 95.4%, distributed at centre-of-mass energies ranging from 200 GeV to 209 GeV. Details of the simulation of background, and an overview of the DELPHI detector in the year 2000, can be found in the first paper listed in this reference [1].

2 The SM search

2.1 The leptonic channels

The leptonic channels, where the Z^0 decays to a pair of leptons (e, μ or τ) and the H^0 decays to a $b\bar{b}$ quark pair, each make up about 2.5% of the total decay width of the H^0Z^0 channel (for a Higgs signal of around m_{h^0}=115 GeV/c^2). The $H^0e^+e^-$ and $H^0\mu^+\mu^-$ channels are both sequential cut analyses using the b-tagging (i.e. the sum of the b-tag for the two hadronic jets) as the final discriminant variable, and the Higgs mass estimator being the invariant dijet mass of the two hadronic jets after a 5C fit, imposing energy and momentum conservation, and in addition forcing the invariant mass of the e^+e^- or $\mu^+\mu^-$ system to the mass of the Z^0.

The $\tau^+\tau^-q\overline{q}$ analysis covers three different channels, two in the SM H^0Z^0 (either the Z^0 going to a $q\overline{q}$ quark pair and the H^0 decaying into $\tau^+\tau^-$, or the Z^0 decaying into $\tau^+\tau^-$ and the H^0 going to a $b\overline{b}$ quark pair) and one in the MSSM h^0A^0 (as the analysis has no way of distinguishing the h^0 from the A^0, the signal consists of a $b\overline{b}$ quark pair recoiling against a $\tau^+\tau^-$ system). In all three cases, the analysis starts with a search for τ candidates in one- or three-prong decays. After rescaling the energies and masses of the τ jets, imposing energy and momentum conservation to give a better estimate of the masses of both di-jets ($\tau^+\tau^-$ and $q\overline{q}$), the final discriminant variable is built from these rescaling factors and the b-tag information. The Higgs mass estimator is constructed by calculating the sum of the invariant masses of the $\tau^+\tau^-$ and the $q\overline{q}$ system, and in the two SM channels subtracting the nominal mass of the Z^0.

2.2 The neutrino channel

The neutrino channel, where $Z^0 \to \nu\overline{\nu}$ and $H^0 \to b\overline{b}$, makes up about 15% of the total decay width of the H^0Z^0 channel. The analysis consists of a series of stringent cuts [2] before a final discriminant variable is constructed, taking b-tagging information and various event shape caracteristics into account. The Higgs mass estimator is calculated as the visible mass after a 1C fit is performed, the constraint being that the invisible recoil system is an on-shell Z^0 boson.

2.3 The hadronic four-jet channel

The hadronic four-jet channel, where $Z^0 \to q\overline{q}$ and $H^0 \to b\overline{b}$ (about 52% of the H^0Z^0 final states) starts with a common four-jet preselection [3] and forces the event into a four-jet topology using the DURHAM algorithm. The final discriminant variable is defined as the output of an artificial neural network combining 13 variables: the global b-tag of the event, four variables which test the compatibility of the event kinematics with the hypotheses of W^+W^- and Z^0Z^0 production to either four or five jets, and eight variables intended to distinguish the $q\overline{q}(\gamma)$ background from the true four-jet signal.

Table 1. Expected background and signal together with the observed data for the different SM Higgs boson channels at two different cut levels of the final discriminant variable of the analyses. The signal efficiency quoted corresponds to a signal of m_{H^0}=114 GeV/c^2.

Channel	Data	Background	$q\bar{q}(\gamma)$	4-fermion	Other	Signal Eff.
Candidates selection level						
$H^0e^+e^-$	7	11.6	0.5	10.4	0.7	57%
$H^0\mu^+\mu^-$	7	10.6	0.2	10.4	—	67%
$\tau^+\tau^-q\bar{q}$	5	6.0	0.4	5.6	—	22%
$H^0\nu\bar{\nu}$	90	99.7	50.4	49.3	—	60%
$H^0q\bar{q}$	398	423.7	154.9	268.8	—	79%
Tight selection						
$H^0e^+e^-$	1	3.5	0.1	3.2	0.2	49%
$H^0\mu^+\mu^-$	2	3.6	0.1	3.5	—	56%
$\tau^+\tau^-q\bar{q}$	2	4.1	0.1	4.0	—	19%
$H^0\nu\bar{\nu}$	3	4.9	1.4	3.5	—	30%
$H^0q\bar{q}$	8	7.4	2.8	4.6	—	36%
Total	16	23.5	4.5	18.8	0.2	27%

The Higgs mass estimator is determined from the likelihood pairing function,

$$\mathcal{P}_b^{j_1}\cdot\mathcal{P}_b^{j_2}\cdot\Big(\big(1-R_b^Z-R_c^Z\big)\cdot\mathcal{P}_q^{j_3}\cdot\mathcal{P}_q^{j_4}+R_b^Z\cdot\mathcal{P}_b^{j_3}\cdot\mathcal{P}_b^{j_4}+R_c^Z\cdot\mathcal{P}_c^{j_3}\cdot\mathcal{P}_c^{j_4}\Big)\cdot P_{j_3,j_4}^{5C}$$

which is calculated for each of the six possibilities to combine the jets j_1, j_2, j_3 and j_4. $\mathcal{P}_b^{j_i}$, $\mathcal{P}_c^{j_i}$, $\mathcal{P}_q^{j_i}$ are the probability densities of getting the observed b-tagging value for the jet j_i when originating from b, c or light (u, d or s) quarks, respectively, R_b^Z and R_c^Z are the hadronic branching fractions of the Z^0 into b or c quarks, respectively, and P_{j_3,j_4}^{5C} is the probability corresponding to the kinematical 5C-fit with the jets j_3 and j_4 assigned to the Z^0. The pairing which minimizes this function is now chosen, and the Higgs mass estimator is set to the invariant mass of the jet pair j_1, j_2.

2.4 Results and limits for the SM search

The agreement between data and expected background at different cut levels for the different SM Higgs search channels is shown in table 1. The plot of the Higgs mass estimator at the tight selection level for the sum of the SM Higgs search channels is shown in figure 1.

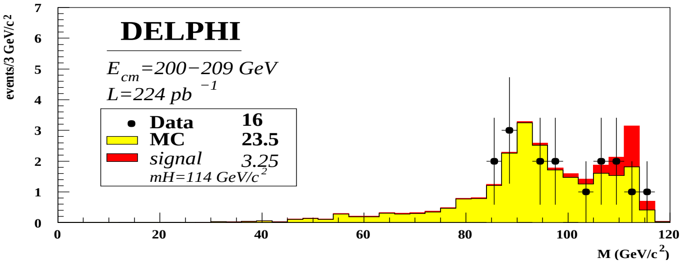


Figure 1. Distribution of the Higgs mass estimator at the tight selection level for the sum of all SM search channels in the year 2000. Data (dots) are compared to the SM background expectation and the expectation from a Higgs boson of mass 114 GeV/c^2.

In the SM Higgs boson search, the only free parameter of the model is the Higgs boson mass, and the final statistical analysis is therefore concerned with the confidence of a specific Higgs boson mass hypothesis. A likelihood method [4] based on the ratio of the confidence level for the signal plus background (CL$_{s+b}$) to the background (CL$_b$) is used to extract the 95% CL on the SM Higgs boson mass. This result for the SM search is shown in figure 2.

3 The MSSM search

3.1 The h^0A^0 channel

The h^0A^0 channel is rather similar to the SM $H^0q\overline{q}$ channel, the main difference being a higher content of B-hadrons in the h^0A^0 signal events, leading to a better discrimination of signal and background due to the b-tagging. After the common four-jet preselection (as in the $H^0q\overline{q}$ channel), a further preselection is performed before the final discriminant variable is constructed by a likelihood method using eight input variables accounting for the b-tag and event kinematics properties. At the final cut selection level, 6.0 events, of which the majority (5.6 events) originates from the 4-fermion processes, is expected in background, whereas the observed number of events is 5. At this level, the signal efficiency is 25% for a Higgs boson signal of m_{A^0}=90 GeV/c^2, $\tan\beta$=20.

As the h^0A^0 production mode dominates over the h^0Z^0 production at large values of $\tan\beta$, where the mass difference between the h^0 and A^0 is small at LEP-*II* energies, the pairing defining the Higgs boson candidates is chosen as that which minimizes the mass difference between the two di-jets after a 4C fit is performed. The Higgs mass estimator is now chosen to be the sum of the invariant masses of these two di-jet objects.

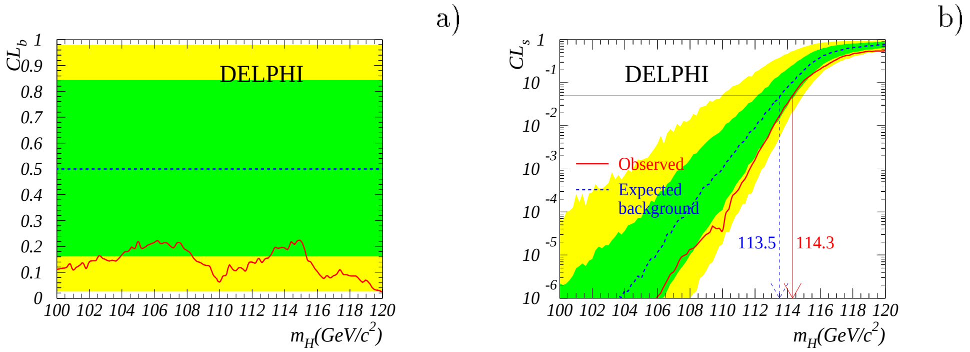


Figure 2. The results from the statistical treatment of the SM search results. In part a, the CL for the background is shown, where the observed result can be seen to lie about one sigma away from the expectation on the non-signal like side, whereas part b shows the CL for the signal, which gives the lower limits for the Higgs boson mass.

3.2 The H^+H^- channel

The decays considered in the H^+H^- analysis include one hadronic and one leptonic decay mode for the charged Higgs bosons. Thus, there are three search channels for the H^+H^- analysis: the leptonic channel ($\tau^-\overline{\nu}_\tau\tau^+\nu_\tau$), the hadronic channel ($\overline{c}s c\overline{s}$), and the semileptonic channel ($\tau^-\overline{\nu}_\tau c\overline{s}$ or $\tau^+\nu_\tau\overline{c}s$). After a channel-specific preselection, the main background of all three channels comes from W^+W^- events. Therefore, all channels construct a channel-specific anti-WW likelihood, containing information from several of the following characteristics and variables: the boson polar angle, τ polarization and identification, di-jet momentum polar angle, hadronic jet flavour tagging information and variables containing information on pairing. Also, the hadronic and semileptonic channels use an anti-QCD likelihood based on event kinematics information.

After the final likelihoods, the number of events remaining in the three channels amounts to 585 for an expected background of 627.6. The same statistical procedure as in the SM search [4] was used to extract 95% CL lower limits on the charged Higgs mass. This was done keeping the branching ratio of $H \rightarrow \tau\nu$ as a free parameter, and the results are shown in part a of figure 3, which gives an overall lower limit on the charged Higgs mass of 73.8 GeV/c^2.

3.3 The invisible Higgs channel

For this channel, where the Higgs boson decays into stable, non-interacting particles, rendering it invisible, the analysis covers four different channels,

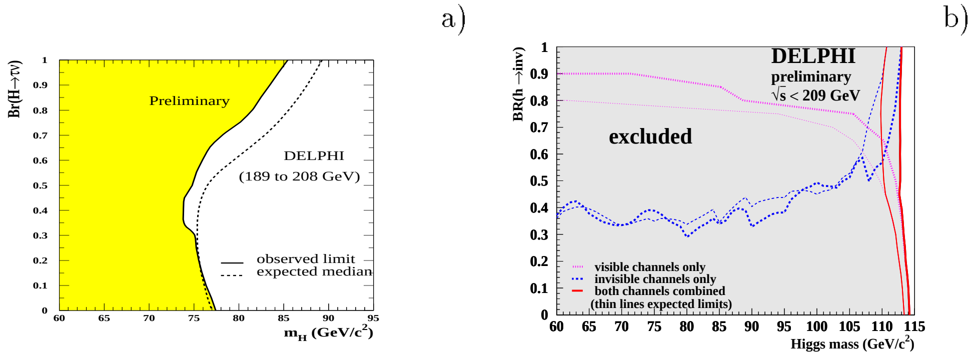


Figure 3. Excluded Higgs mass regions for the H^+H^- and invisible Higgs search. Part a shows, for the charged Higgs search, the regions excluded at 95% CL in the plane of the branching ratio of $H \to \tau\nu$ (the branching ratio of $H \to cs$ is assumed to be 1-BR($H \to \tau\nu$)) vs. the charged Higgs mass. Part b shows, for the invisible Higgs search, the regions excluded at 95% CL in the plane of the branching ratio of h →invisible vs. the Higgs mass.

according to the decay of the Z^0: one hadronic ($Z^0 \to q\bar{q}$) and three leptonic ($Z^0 \to e^+e^-$, $\mu^+\mu^-$ or $\tau^+\tau^-$). After a preselection, the hadronic channel uses an Iterative Discriminant Analysis Program (IDA) to combine information on event energy and momentum deposits as well as event shape in a discriminant variable. The leptonic channels apply cuts on lepton identification and event kinematics to reduce the background. When adding the remaining candidates, the number of observed events is 41 for an expected background of 58.4 events. 95% CL lower limits on the Higgs mass was found as a function of the branching ratio into invisible decays BR$_{inv}$, assuming visible branchnig ratios BR$_{vis}$=1-BR$_{inv}$ and SM production cross sections. Part b of figure 3 shows the result, which yields a global lower Higgs boson mass limit of 112.7 GeV/c^2.

3.4 Results and limits for the MSSM search

In the MSSM, there are at tree level only two free parameters of the Higgs sector, usually chosen as two of the three parameters $\tan\beta$, m_{A^0} and m_{h^0}. Radiative corrections introduce additional parameters [2]. The search translate into regions in the MSSM parameter space excluded at the 95% CL or more, and are given for two different benchmarks, called the m_h^{max} scenario (which gives the maximal value of m_{h^0} as a function of $\tan\beta$) and the no mixing scenario [2]. The results are shown in figure 4.

Finally, the excluded regions in the different Higgs boson masses and, for the MSSM, $\tan\beta$ values are summarized in table 2. [5]

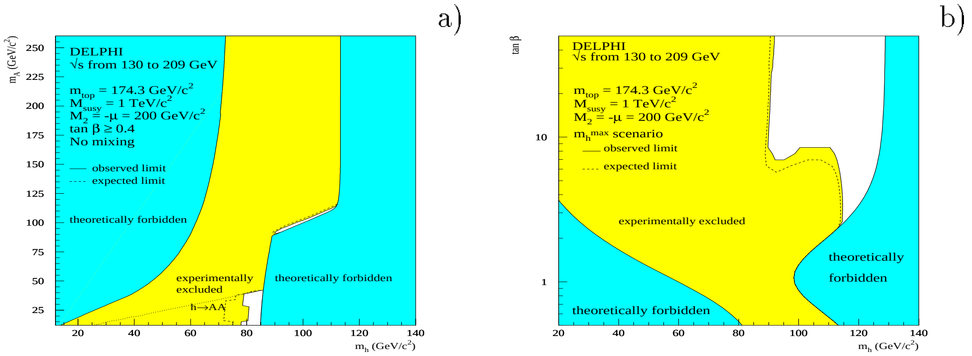


Figure 4. MSSM regions in the (m_{h^0}, m_{A^0}) plane for the no mixing scenario (part a) and in the (m_{h^0}, $\tan\beta$) plane for the m_h^{max} scenario (part b) excluded at 95% CL. All SM and MSSM channels up to $\sqrt{s}$=209 GeV are included.

Table 2. The search results, given in terms of excluded ranges of Higgs boson masses and, for the MSSM case, values of $\tan\beta$. For MSSM exclusion of Higgs boson masses, the assumptions m_{A^0}>12 GeV/c^2 and $\tan\beta$<0.6 is made.

Framework	Observed limit	Median expected limit	Scenario
SM	m_{H^0}<114.3 GeV/c^2	m_{H^0}<113.5 GeV/c^2	—
MSSM	m_{A^0}<90.9 GeV/c^2	m_{A^0}<90.1 GeV/c^2	m_h^{max}
	m_{h^0}<89.8 GeV/c^2	m_{h^0}<89.0 GeV/c^2	
	$\tan\beta \in$ [0.49,2.36]	$\tan\beta \in$ [0.54,2.36]	
MSSM	m_{A^0}<90.8 GeV/c^2	m_{A^0}<90.0 GeV/c^2	No mixing
	m_{h^0}<90.0 GeV/c^2	m_{h^0}<89.1 GeV/c^2	
	$\tan\beta \in$ [0.59,9.36]	$\tan\beta \in$ [0.72,8.49]	

References

1. DELPHI Collaboration, *Phys. Lett.* B **499**, 23 (2001);
   DELPHI Higgs Research Line Teams, DELPHI 2001-017 CONF 458;
   M. Battaglia, M. Ellert, T. Ekelof, G. Gómez-Ceballos, A. Kiiskinen, P. Lutz and F. Matorras, DELPHI 2001-030 CONF 471;
   A. Sopczak and M. Stanitzki, DELPHI 2001-025 CONF 466;
2. DELPHI Collaboration, DELPHI 2000-077 CONF 376.
3. DELPHI Collaboration, *Eur. Phus. J.* C **17**, 187 (2000).
4. A. L. Read, in CERN Report 2000-005, p. 81 (2000).
5. http://delphiwww.cern.ch/teams/searches/delwww/higgs/inputs/results/results.html#010227

Appendix D

Statistical errors for Monte Carlo samples

When estimating the number of expected events for a signal or background channel, the number of expected events for channel i is found by the equation

$$N_i = \epsilon_i \mathcal{L}_i \sigma_i \tag{D.1}$$

where ϵ_i is the efficiency of the channel, $\mathcal{L}_i$ is the luminosity and σ_i is the cross section of the channel. If there are more channels which are to be summed together, the total number of events is now given by

$$N_{\text{tot}} = \sum_{i=1}^{n_i} N_i = \sum_{i=1}^{n_i} \epsilon_i \mathcal{L}_i \sigma_i . \tag{D.2}$$

The goal is now to express equation D.2 in terms of total quantities instead of summations, and to find the total error in these quantities. Two typical cases, both applicable to the situation in the analysis presented in this work, will be considered, and are treated in the following two sections. In both cases, the relation for the total variance of a function $f(x_1, x_2, \ldots x_n)$,

$$\begin{aligned} \text{Var}(f) &= \textstyle\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2 \text{Var}(x_i) + 2 \sum_{i<j}^{n} \left(\frac{\partial f}{\partial x_i}\right)\left(\frac{\partial f}{\partial x_j}\right) \text{Cov}(x_i, x_j) \\ &= \textstyle\sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)\left(\frac{\partial f}{\partial x_j}\right) \text{Cov}(x_i, x_j) \end{aligned} \tag{D.3}$$

will be used extensively. In the special case that $f(x_1, x_2, \ldots x_n)$ is a simple multiplicative function of the powers of the variables (which is the case for our application of this relation), $f(x_1, x_2, \ldots x_n) = x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n}$, the relation D.3

simplifies to

$$\frac{\mathrm{Var}(f)}{f^2} = \sum_{i=1}^{n} p_i^2 \frac{\mathrm{Var}(x_i)}{x_i^2} + 2\sum_{i<j}^{n} p_i p_j \frac{\mathrm{Cov}(x_i, x_j)}{x_i x_j} = \sum_{i=1}^{n}\sum_{j=1}^{n} p_i p_j \frac{\mathrm{Cov}(x_i, x_j)}{x_i x_j}. \tag{D.4}$$

For literature regarding the fundamental relations of variance and other statistical concepts, see for instance [94].

In what follows, it is assumed that variables of different types are uncorrelated, *i.e.*

$$\mathrm{Cov}(\epsilon_i, \mathcal{L}_j) = \mathrm{Cov}(\mathcal{L}_i, \sigma_j) = \mathrm{Cov}(\sigma_i, \epsilon_j) = 0. \tag{D.5}$$

D.1 Sum over channels of different centre-of-mass energy points

In this case, the total number of events N_{tot} of equation D.2 is the result of a summation over (in general) both different channels and different centre-of-mass energy points. Thus, equation D.2 can be expressed as

$$N_{\text{tot}} = \sum_{i=1}^{n_i} N_i = \sum_{i=1}^{n_i} \epsilon_i \mathcal{L}_i \sigma_i \overset{\text{def}}{=} \epsilon_{\text{tot}} \mathcal{L}_{\text{tot}} \sigma_{\text{tot}} \tag{D.6}$$

This equation does not uniquely define the total quantities ϵ_{tot}, $\mathcal{L}_{\text{tot}}$ and σ_{tot}, which necessitates further assumptions. The luminosity $\mathcal{L}_i$ is usually considered an additive quantity,[1] and the total efficiency ϵ_{tot} should be confined to values between zero and one. These considerations make the choice

$$\mathcal{L}_{\text{tot}} = \sum_{i=1}^{n} \mathcal{L}_i, \quad \epsilon_{\text{tot}} = \frac{\sum_{i=1}^{n} \mathcal{L}_i \epsilon_i}{\sum_{i=1}^{n} \mathcal{L}_i} \overset{\text{def}}{=} \frac{A}{\mathcal{L}_{\text{tot}}}, \quad \sigma_{\text{tot}} = \frac{\sum_{i=1}^{n} \mathcal{L}_i \epsilon_i \sigma_i}{\sum_{i=1}^{n} \mathcal{L}_i \epsilon_i} = \frac{N_{\text{tot}}}{A} \tag{D.7}$$

a natural, albeit not the only, one.

By using the assumptions in D.5, together with equation D.4 (keeping in mind that the covariance of a variable with itself is simply the variance,

---

[1] This is the case for luminosities at different centre-of-mass energy points, whereas the luminosity at channels of the same centre-of-mass energy point is a common factor, making the correlation between these luminosities 100%. In order to have the total luminosity $\mathcal{L}_{\text{tot}}$ reflect this, factors of $\frac{1}{n_{ch}}$ and n_{ch} (where $n_{ch}$is the number of different channels within the same centre-of-mass energy point) can be introduced into the definitions of the total quantities $\mathcal{L}_{\text{tot}}$ and σ_{tot}, respectively. However, since these factors are only multiplicative constants, the basic mathematical relations between the total quantities are only modified by the same constants.

$\mathrm{Cov}(x, x) = \mathrm{Var}(x)$), the variance of the different total quantities (including the intermediate variable A) is now:

$$
\begin{array}{rcl}
\mathrm{Var}(A) & = & \sum_{i,j=1}^{n} \mathrm{Cov}(\mathcal{L}_i\epsilon_i, \mathcal{L}_j\epsilon_j) \\
 & = & \sum_{i,j=1}^{n} \mathcal{L}_i\mathcal{L}_j \mathrm{Cov}(\epsilon_i, \epsilon_j) + \epsilon_i\epsilon_j \mathrm{Cov}(\mathcal{L}_i, \mathcal{L}_j) \\
 & & \\
\mathrm{Var}(\mathcal{L}_{\mathrm{tot}}) & = & \sum_{i,j=1}^{n} \mathrm{Cov}(\mathcal{L}_i, \mathcal{L}_j) \\
 & & \\
\mathrm{Var}(\epsilon_{\mathrm{tot}}) & = & \left(\frac{A}{\mathcal{L}_{\mathrm{tot}}}\right)^2 \left(\frac{\mathrm{Var}(A)}{A^2} + \frac{\mathrm{Var}(\mathcal{L}_{\mathrm{tot}})}{\mathcal{L}_{\mathrm{tot}}^2} - 2\frac{\mathrm{Cov}(A, \mathcal{L}_{\mathrm{tot}})}{A\mathcal{L}_{\mathrm{tot}}}\right) \\
 & = & \left(\frac{1}{\mathcal{L}_{\mathrm{tot}}}\right)^4 \sum_{i,j=1}^{n} \{\mathcal{L}_{\mathrm{tot}}^2 \left[\mathcal{L}_i\mathcal{L}_j \mathrm{Cov}(\epsilon_i, \epsilon_j)\right. \\
 & & \quad \left.+\epsilon_i\epsilon_j \mathrm{Cov}(\mathcal{L}_i, \mathcal{L}_j)\right] + (A^2 - 2A\mathcal{L}_{\mathrm{tot}}\epsilon_i)\,\mathrm{Cov}(\mathcal{L}_i, \mathcal{L}_j)\} \\
 & & \\
\mathrm{Var}(\sigma_{\mathrm{tot}}) & = & \left(\frac{N_{\mathrm{tot}}}{A}\right)^2 \left(\frac{\mathrm{Var}(N_{\mathrm{tot}})}{N_{\mathrm{tot}}^2} + \frac{\mathrm{Var}(A)}{A^2} - 2\frac{\mathrm{Cov}(N_{\mathrm{tot}}, A)}{N_{\mathrm{tot}}A}\right) \\
 & = & \left(\frac{1}{A}\right)^4 \sum_{i,j=1}^{n} \{A^2 \left[\epsilon_i\epsilon_j\sigma_i\sigma_j \mathrm{Cov}(\mathcal{L}_i, \mathcal{L}_j)\right. \\
 & & \qquad \left.+\sigma_i\sigma_j\mathcal{L}_i\mathcal{L}_j \mathrm{Cov}(\epsilon_i, \epsilon_j) + \mathcal{L}_i\mathcal{L}_j\epsilon_i\epsilon_j \mathrm{Cov}(\sigma_i, \sigma_j)\right] \\
 & & \quad + (N_{\mathrm{tot}}^2 - 2AN_{\mathrm{tot}}\sigma_i)\left[\mathcal{L}_i\mathcal{L}_j \mathrm{Cov}(\epsilon_i, \epsilon_j) + \epsilon_i\epsilon_j \mathrm{Cov}(\mathcal{L}_i, \mathcal{L}_j)\right]\} \\
 & & \\
\mathrm{Var}(N_{\mathrm{tot}}) & = & \sum_{i,j=1}^{n} \{\epsilon_i\epsilon_j\mathcal{L}_i\mathcal{L}_j \mathrm{Cov}(\sigma_i, \sigma_j) + \mathcal{L}_i\mathcal{L}_j\sigma_i\sigma_j \mathrm{Cov}(\epsilon_i, \epsilon_j) \\
 & & \qquad +\sigma_i\sigma_j\epsilon_i\epsilon_j \mathrm{Cov}(\mathcal{L}_i, \mathcal{L}_j)\}
\end{array}
\tag{D.8}
$$

D.2 Sum over channels within one centre-of-mass energy point

In this case, which in some sense is a special version of the one treated in the previous section, the channels are all assumed to be within the same centre-of-mass energy point. This means that the luminosity for each channel, $\mathcal{L}_i$, is a common factor, and the expression for the total number of events, equation D.2 takes the form

$$
N_{\mathrm{tot}} = \sum_{i=1}^{n_i} N_i = \sum_{i=1}^{n_i} \epsilon_i\mathcal{L}_i\sigma_i = \mathcal{L}\sum_{i=1}^{n_i} \epsilon_i\sigma_i \overset{\mathrm{def}}{=} \epsilon_{\mathrm{tot}}\sigma_{\mathrm{tot}}\mathcal{L} \tag{D.9}
$$

where the channels subscript is removed for the luminosity, $\mathcal{L} = \mathcal{L}_i$, since this is, by assumption, equal for all channels. Again, this equation does not uniquely define the total quantities ϵ_{tot} and σ_{tot}, but the requirement of confining the total efficiency ϵ_{tot} to values between zero and one, makes the definitions

$$\sigma_{\text{tot}} = \sum_{i=1}^{n} \sigma_i, \quad \epsilon_{\text{tot}} = \frac{\sum_{i=1}^{n} \sigma_i \epsilon_i}{\sum_{i=1}^{n} \sigma_i} \overset{\text{def}}{=} \frac{B}{\sigma_{\text{tot}}} \tag{D.10}$$

a natural choice.

Making use of this definition, together with the general assumptions made in D.5, the variance of these quantities are now:

$$\begin{aligned}
\text{Var}(B) &= \textstyle\sum_{i,j=1}^{n} \{\sigma_i \sigma_j \text{Cov}(\epsilon_i, \epsilon_j) + \epsilon_i \epsilon_j \text{Cov}(\sigma_i, \sigma_j)\} \\
\text{Var}(\sigma_{\text{tot}}) &= \textstyle\sum_{i,j=1}^{n} \{\text{Cov}(\sigma_i, \sigma_j)\} \\
\text{Var}(\epsilon_{\text{tot}}) &= \left(\tfrac{1}{\sigma_{\text{tot}}}\right)^4 \textstyle\sum_{i,j=1}^{n} \{\sigma_{\text{tot}}^2 \left[\sigma_i \sigma_j \text{Cov}(\epsilon_i, \epsilon_j) + \epsilon_i \epsilon_j \text{Cov}(\sigma_i, \sigma_j)\right] \\
&\qquad + (B^2 - 2B\sigma_{\text{tot}}\epsilon_i)\text{Cov}(\sigma_i, \sigma_j)\} \\
\text{Var}(N_{\text{tot}}) &= \textstyle\sum_{i,j=1}^{n} \{\epsilon_i \epsilon_j \mathcal{L}^2 \text{Cov}(\sigma_i, \sigma_j) + \mathcal{L}^2 \sigma_i \sigma_j \text{Cov}(\epsilon_i, \epsilon_j) \\
&\qquad + \sigma_i \sigma_j \epsilon_i \epsilon_j \text{Var}(\mathcal{L})\}
\end{aligned} \tag{D.11}$$

Appendix E

Error band for curves with errors in both dimensions

When studying how different cuts and selection criteria affect different parts of an analysis, one is frequently in the situation of looking at the evolution of two quantities as a function of this cut or selection criteria. This is often presented as a curve; one of these quantities as a function of the other (for instance, the expected background as a function of signal efficiency). However, these two quantities often have errors associated to them, and the curve is therefore enlarged to a band. This appendix describes how this band is extracted from the original curve, and the standard deviations in the two quantities.

Consider two variables x and y, $y = f(x)$. The variables have standard deviations σ_x and σ_y, respectively, which are assumed to be uncorrelated, but not necessarily constant. The error ellipse around each point (x_0, y_0) is therefore an ellipse obeying the equation

$$\left(\frac{x - x_0}{\sigma_x}\right)^2 + \left(\frac{y - y_0}{\sigma_y}\right)^2 = 1 \quad \Rightarrow \quad y = \pm\sigma_y\sqrt{1 - \left(\frac{x - x_0}{\sigma_x}\right)^2} + y_0 \tag{E.1}$$

For simplicity, consider first the special case of the curve being a straight line, and the standard deviations σ_x and σ_y being constants. In this case, the error ellipse will trace out a band around the curve $y = f(x)$ according to the angle α of the line with the x-aksis, in such a way that the ellipse always stays within this band (see figure E.1). The error band is now found by determining the point on the error ellipse where the derivative of the ellipsis function (equation E.1) is equal to the slope of the tangent to the curve. Thus, if we look at the upper half of the ellipse, $y > y_0$, the equation for the

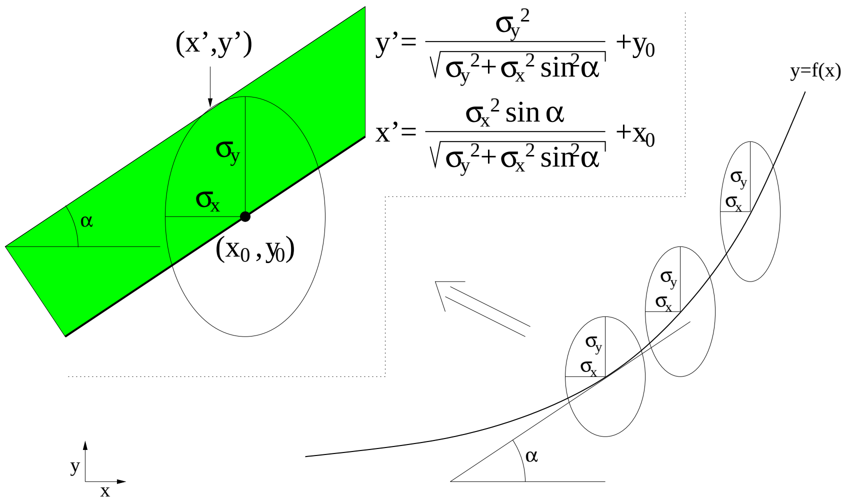


Figure E.1: Figure showing the construction of error bands around the function $y = f(x)$ in the case of errors in both variables x and y. The errors are assumed uncorrelated, and the function $f(x)$ is assumed differentiable for all points of interest.

derivative is

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\sigma_y(x-x_0)}{\sigma_x^2\sqrt{1-\left(\frac{x-x_0}{\sigma_x}\right)^2}} = \sin\alpha \overset{\text{def}}{=} a \tag{E.2}$$

with the lower half of the ellipse having an additional minus sign. This equation has two solutions, depending on the sign of $\sin\alpha$, giving the points (x', y') as solutions to equation E.2:

$$\begin{aligned} a < 0 \;&\Rightarrow\; x' = \tfrac{-a\sigma_x}{\sqrt{\sigma_y^2+a^2\sigma_x^2}} + x_0 \;\Rightarrow\; y' = \tfrac{\sigma_y}{\sqrt{\sigma_y^2+a^2\sigma_x^2}} + y_0 \\ a > 0 \;&\Rightarrow\; x' = \tfrac{a\sigma_x}{\sqrt{\sigma_y^2+a^2\sigma_x^2}} + x_0 \;\Rightarrow\; y' = \tfrac{\sigma_y}{\sqrt{\sigma_y^2+a^2\sigma_x^2}} + y_0. \end{aligned} \tag{E.3}$$

For the lower half $y < y_0$, the only change compared to these results is an additional minus sign, such that in this case, the points are given by

$$\begin{aligned} a < 0 \;&\Rightarrow\; x' = \tfrac{a\sigma_x}{\sqrt{\sigma_y^2+a^2\sigma_x^2}} + x_0 \;\Rightarrow\; y' = \tfrac{-\sigma_y}{\sqrt{\sigma_y^2+a^2\sigma_x^2}} + y_0 \\ a > 0 \;&\Rightarrow\; x' = \tfrac{-a\sigma_x}{\sqrt{\sigma_y^2+a^2\sigma_x^2}} + x_0 \;\Rightarrow\; y' = \tfrac{-\sigma_y}{\sqrt{\sigma_y^2+a^2\sigma_x^2}} + y_0, \end{aligned} \tag{E.4}$$

so that the point (x', y') is simply reflected through the point (x_0, y_0), as one would expect from figure E.1. For this reason, only the upper half $y > y_0$ will be considered from this point.

When considering the general case of non-constant standard deviations σ_x and σ_y, the expression for the derivative $\frac{\mathrm{d}y}{\mathrm{d}x}$ takes a more complex form than the simple relation found in equation E.2:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \sqrt{1 - \left(\frac{x - x_0}{\sigma_x}\right)^2} \left\{ \frac{\mathrm{d}\sigma_y}{\mathrm{d}x} - \frac{\sigma_y (x - x_0)}{\sigma_x \left(1 - \left(\frac{x - x_0}{\sigma_x}\right)^2\right)} \left[\frac{1}{\sigma_x} - \frac{\mathrm{d}\sigma_x}{\mathrm{d}x}(x - x_0)\right] \right\} \tag{E.5}$$

The factor in square brackets reduces to $\frac{1}{\sigma_x}$ when assuming constant standard deviations (*i.e.* $\frac{\mathrm{d}\sigma_y}{\mathrm{d}x} = \frac{\mathrm{d}\sigma_x}{\mathrm{d}x} = 0$, reducing equation E.5 to E.2), and can be considered a perturbation of this value. In order to estimate how this factor behaves, assume that it is a perturbation of the original value, that is

$$\frac{1}{\sigma_x} - \frac{\mathrm{d}\sigma_x}{\mathrm{d}x}(x - x_0) = \frac{1}{\sigma_x}(1 - \epsilon) \quad \Rightarrow \quad \frac{\mathrm{d}\sigma_x}{\mathrm{d}x} = \frac{\epsilon}{\sigma_x (x - x_0)}. \tag{E.6}$$

This is a simple differential equation for σ_x, which has the solution

$$\sigma_x = \sqrt{2\epsilon \ln(|x - x_0|) + C}, \tag{E.7}$$

where the factor C is an arbitrary constant of integration. In other words, if σ_x changes no faster with the variable x than the function given in equation E.7, the correction to the factor in square brackets in equation E.5 changes less than a factor $(1 - \epsilon)$.

In order to tackle the total factor in braces in equation E.5, the assumption is made that σ_x conforms to the requirements mentioned in order to ensure that the term in the square brackets remains close to the original value, that is

$$\frac{\mathrm{d}\sigma_y}{\mathrm{d}x} - \frac{\sigma_y (x - x_0)}{\sigma_x \left(1 - \left(\frac{x - x_0}{\sigma_x}\right)^2\right)} \left[\frac{1}{\sigma_x} - \frac{\mathrm{d}\sigma_x}{\mathrm{d}x}(x - x_0)\right] = \frac{\mathrm{d}\sigma_y}{\mathrm{d}x} - \frac{\sigma_y (x - x_0)(1 - \epsilon)}{\sigma_x^2 \left(1 - \left(\frac{x - x_0}{\sigma_x}\right)^2\right)}. \tag{E.8}$$

The factor σ_x is now considered a constant, under the assumption that the inclusion of the derivative $\frac{\mathrm{d}\sigma_x}{\mathrm{d}x}$ only changes the expression by a factor $(1 - \epsilon)$. Thus, if we now consider the right-hand side of equation E.8 as a perturbation of the original value E.2, the corresponding equation to E.6 (with the factor

ϵ' instead of ϵ) for the total factor in braces is now

$$\frac{\mathrm{d}\sigma_y}{\mathrm{d}x} - \frac{\sigma_y(x-x_0)(1-\epsilon)}{\sigma_x^2\left(1-\left(\frac{x-x_0}{\sigma_x}\right)^2\right)} = \frac{\sigma_y(x-x_0)(1-\epsilon)(1-\epsilon')}{\sigma_x^2\left(1-\left(\frac{x-x_0}{\sigma_x}\right)^2\right)} \quad \Rightarrow$$

$$\frac{\mathrm{d}\sigma_y}{\mathrm{d}x} = \frac{(1-\epsilon)\epsilon'\sigma_y}{\sigma_x^2}\frac{(x-x_0)}{\left(1-\left(\frac{x-x_0}{\sigma_x}\right)^2\right)} \approx \frac{\epsilon'\sigma_y}{\sigma_x^2}\frac{(x-x_0)}{\left(1-\left(\frac{x-x_0}{\sigma_x}\right)^2\right)} \qquad \text{(E.9)}$$

when only keeping the first-order terms in ϵ and ϵ'. This is again a simple differential equation for σ_y, with the solution

$$\sigma_y = C'\left(1-\left(\frac{x-x_0}{\sigma_x}\right)^2\right)^{-\frac{\epsilon'}{2}}. \qquad \text{(E.10)}$$

This means, again, that as long as σ_y changes less rapidly than the function given in E.10, the change with respect to E.2 is less than the factor $(1-\epsilon')$.

Under these assumptions about the change of the standard deviations σ_x and σ_y, the points (x', y') given in equations E.3 and E.4 are correct to within an accuracy of $(1-\epsilon)(1-\epsilon')$.

Appendix F

Description of polynomial splines

In the 1-dimensional case, a polynomial spline is defined by the following parameters and coefficients [78, page 149]:

k	Degree (order-1) of the B-spline ($0 \leq k \leq 25$).
m	Number of spline-knots ($m \geq 2k+2$).
i	Index of the B-spline($1 \leq i \leq m-k-1$).
τ	Set of m spline-knots $\tau = \{t_1, t_2, \ldots, t_m\}$, in non-decreasing order, with multiplicity $\leq k+1$, (*i.e.* no more than $k+1$ knots coincide).
$[a,b]$	Interval, defined by $a = t_{k+1}$, $b = t_{m-k}$.
$B_i(x)$	Normalized B-spline of degree k over τ with index i. The value of B_i is identically zero outside the interval $t_i \leq x \leq t_{i+k+1}$, and the normalization of $B_i(x)$ is such that $\displaystyle\int_{-\infty}^{+\infty} B_i(x)\mathrm{d}x = \frac{t_{i+k+1} - t_i}{k+1} \quad (i = 1, \ldots, m-k-1). \qquad \text{(F.1)}$
$s(x)$	Polynomial spline at $x \in [a,b]$ in B-spline representation $\displaystyle y = s(x) = \sum_{i=1}^{m-k-1} c_i B_i(x). \qquad \text{(F.2)}$

When using this B-spline for approximating a data set $\{x_l, y_l\}_{l=1,\ldots,n}$ the coefficients $\{c_i\}_{i=1,\ldots,m-k-1}$ are determined by minimizing the function

$$\phi(c_1, \ldots, c_{m-k-1}) = \sum_{l=1}^{n} (y_l - s(x_l))^2 . \tag{F.3}$$

This is easily generalized to the case of 2-dimensional B-splines, where the number of the aforementioned parameters and coefficients is doubled, with one set acting in each spatial dimension. Thus, the new parameters and coefficients introduced are k_x, k_y, m_x, m_y, i, j, τ_x, τ_y, $[a_x, b_x]$, $[a_y, b_y]$, $B_i(x)$ and $B_j(y)$, with the total B-spline given by

$$B_{i,j}(x, y) = B_i(x)B_j(y). \tag{F.4}$$

The polynomial spline is given by

$$z = s(x, y) = \sum_{i=1}^{m_x-k_x-1} \sum_{j=1}^{m_y-k_y-1} c_{i,j} B_{i,j}(x, y), \tag{F.5}$$

and the function which is minimized is now

$$\phi(c_{1,1}, \ldots, c_{m_x-k_x-1,m_y-k_y-1}) = \sum_{l_x=1}^{n_x} \sum_{l_y=1}^{n_y} \left(z_{l_x,l_y} - s(x_{l_x}, y_{l_y})\right)^2 , \tag{F.6}$$

where the data set is given by $\{x_{l_x}, y_{l_y}, z_{l_x,l_y}\}_{l_x=1,\ldots,n_x; l_y=1,\ldots,n_y}$.[1]

For the implementation of B-splines for the purpose of smoothing 2-dimensional histograms used in this work [79, page 132], the maximum degree of the B-spline (k_x and k_y) is 3 in both spatial dimensions. For reasons of simplicity, and in order to keep the smoothing routine as flexible and general as possible, the choice $k_x = k_y = 3$ was made. Also, the sets of spline knots τ_x and τ_y are pre-defined, in such a way that, in the x-directon (y-direction), the first k_x+1 (k_y+1) knots coincide with a_x (a_y), and the last k_x+1 (k_y+1) knots coincide with b_x (b_y), leaving the remaining m_x-2k_x-2 (m_y-2k_y-2) knots equidistantly distributed in the interval $[a_x, b_x]$ ($[a_y, b_y]$).

---

[1] In the notation of the histogram bin content, this has the following correspondence: x_{l_x}=x-value of bin (i, j), y_{l_y}=y-value of bin (i, j), $z_{l_x,l_y} = H_{i,j}$, the bin content of bin (i, j), with the sum in equation F.6 being over i and j instead of l_x and l_y.

Bibliography

[1] T. Junk, *Searches at LEP*, talk given at the RADCOR 2000 conference, September 12, 2000, World Wide Web,

`http://radcor2000.slac.stanford.edu/talks/pdf/junk.pdf`

[2] G. L. Kane (ed.), *Perspectives on supersymmetry*, World Scientific Publishing Company, ISBN 981–02–3553–4 (1998).

[3] S. L. Glashow, Nucl. Phys. **22** (1961), 579;
S. Weinberg, Phys. Rev. Lett. **19** (1967), 1264;
A. Salam, *Elementary Particle Theory*, ed. N. Svartholm (Stockholm 1968), 361.

[4] J. P. Elliott and P. G. Dawber, *Symmetry in physics, Volume 1: principles and simple applications*, MACMILLAN PRESS LTD, ISBN 0–333–38270–8 (1979).

[5] UA1 Collaboration, G. Arnison *et al.*, Phys. Lett. **B122** (1983), 103;
UA2 Collaboration, M. Banner *et al.*, Phys. Lett. **B122** (1983), 476.

[6] UA1 Collaboration, G. Arnison *et al.*, Phys. Lett. **B126** (1983), 398;
UA2 Collaboration, P. Bagnaia *et al.*, Phys. Lett. **B129** (1983), 130.

[7] J. Goldstone, Nuovo Cim.**19** (1961),154.

[8] P. W. Higgs, Phys. Lett. **12** (1964), 132;
P. W. Higgs, Phys. Rev. Lett. **13** (1964), 508;
P. W. Higgs, Phys. Rev. **145** (1966), 1156;
F. Englert and R. Brout, Phys. Rev. Lett. **13** (1964), 321;
G. S. Guralnik, C. S. Hagen and T. W. B. Kibble, Phys. Rev. Lett. **13** (1964), 585;
T. W. B. Kibble, Phys. Rev. **155** (1967), 1554.

[9] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley Publishing Company, ISBN 0–201–50397–2 (1995).

[10] R. T. Cox, C. G. McIlwraith, and B. Kurrelmeyer, Nat. Acad. Sci. Proc. **14** (1928), 544. C. S. Wu *et al.*, Phys. Rev. **105** (1957), 1413.

[11] Y. Fukuda *et al.*, Phys. Rev. Lett. **81** (1998), 1562;
Y. Fukuda *et al.*, Phys. Lett. **B433** (1998), 9; Phys. Lett. **B436** (1998), 33;
Y. Fukuda *et al.*, Phys. Rev. Lett. **82** (1999), 2664; Phys. Lett. **B467** (1999), 185.

[12] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49** (1975), 652.

[13] A. D. Linde, JETP Lett. **23** (1976), 64; Phys. Lett. **B62** (1976), 435;
S. Weinberg, Phys. Rev. Lett. **36** (1976), 294.

[14] S. Abachi *et al.*, Phys. Rev. Lett. **79** (1997), 1197;
F. Abe *et al.*, Phys. Rev. Lett. **80** (1998), 2767.

[15] S. Coleman and E. Weinberg, Phys. Rev. **D7** (1973), 1888.

[16] J. F. Gunion, H. E. Haber, G. L. Kane and S. Dawson, *The Higgs Hunter's Guide*, Addison-Wesley Publishing Company, ISBN 0–201–50935–0 (1990).

[17] B. W. Lee, C. Quigg and G. B. Thacker, Phys. Rev. Lett. **38** (1977), 883;
Phys. Rev. **D16** (1977), 1519.

[18] S. Dawson and S. Willenbrock, Phys. Rev. Lett. **62** (1989), 1232;
Phys. Rev. **D40** (1989), 2880.

[19] H. E. Haber and G. L. Kane, *The search for supersymmetry: probing physics beyond the Standard Model*, Phys. Rep. **117** (1985), 75.

[20] H. Elvang, *Symmetri, Coleman-Mandulas Teorem og Supersymmetri*, Gamma Γ, Tidsskrift for fysik **118** (2000), 26.

[21] S. Coleman and J. Mandula, *All possible symmetries of the S-matrix*, Phys. Rev. **159** (1967), 1251.

[22] S. Lowell, D. Brown and D. L. Nordstrom, *Review of Particle Properties*, Phys. Rev. **D50** (1994), 1173.

[23] D. A. Ross and M. Veltman, Nucl. Phys. **B95** (1975), 135.

[24] M. Savage, Phys. Lett. **B266** (1991), 135.

[25] S. G. Glashow and S. Weinberg, Phys. Rev. **D15** (1977), 1958.

[26] A. Sopczak, *Search for Non-Minimal Higgs Bosons in Z^0 Decays with the L3 Detector at LEP*, Ph.D. dissertation, University of California, San Diego (1992).

[27] W. Grimus, Fortschr. Phys. **36** (1988), 201.

[28] J. F. Gunion and H. E. Haber, Nucl. Phys. **278** (1986), 449.

[29] J. A. Grifols and A. Mendez, Phys. Rev. **D22** (1980), 1725.

[30] I. Affleck, M. Dine and N. Seiberg, Nucl. Phys. **B256** (1985), 557.

[31] ACF Linear Colider Working Group, *SUSY breaking scenarios*, World Wide Web,

`http://acfahep.kek.jp/acfareport/node66.html`

[32] S. Dimopoulos and D. Sutter, Nucl. Phys. **B452** (1995), 496.

[33] J. F. Gunion and H. E. Haber, Nucl. Phys. **272** (1986), 1.

[34] T. L. Barklow *et al.*, *Strong Coupling Electroweak Symmetry Breaking*, Working Group Summary Report from the 1996 DPF-DPB Summer Study, Snowmass, Colorado, June 25-July 12, 1996, SLAC-PUB-7397, CLNS 97/1473 (1997).

[35] S. Weinberg, Phys. Rev. **D19** (1979), 1277; newline L. Susskind, Phys. Rev. **D20** (1979), 2619.

[36] R. K. Kaul, Rev. Mod. Phys. **55** (1983), 449.

[37] K. Lane, *Technicolor 2000*, preprint BUHEP–00–15, e-Print arXiv: hep-ph/0007304 (2000).

[38] S. Dimopoulos and L. Susskind, Nucl. Phys. **B155** (1979), 237.

[39] E. Eichten and K. D. Lane, Phys. Lett. **B90** (1980), 125.

[40] B. Holdom, Phys. Lett. **B150** (1985), 301.

[41] C. T. Hill, Phys. Lett. **B266** (1991), 419.

[42] C. T. Hill, Phys. Lett. **B345** (1995), 483.

[43] M. B. Green, J. H. Schwarz, E. Witten, *Superstring Theory, two volumes*, Cambridge University Press, ISBN 0–521–35752–7 and ISBN 0–521–35753–5 (1988).
J. Polchinski, *String Theory (aka "Joe's Big Book of Strings")*, Cambridge University Press, ISBN 0–521–63312–5 (1998);

[44] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, Phys. Lett. **B429** (1998), 263; Phys. Rev. **D59** (1999), 086004;
I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali, Phys. Lett. **B436** (1998), 257.

[45] M. Shaposhnikov and P. Tinyakov, Phys. Lett. **B515** (2001), 442.

[46] Y. Hositani, Phys. Lett. **B126** (1983), 309.

[47] G. Dvali, S. Randjbar-Daemi and R. Tabbash, e-Print arXiv: hep-ph/0102307 (2001).

[48] M. Schiøtz, *The use of semianalytic confidence limit calculations in searches for new particles at LEP200*, Thesis submitted to the Cand. Scient. Degree (1999).

[49] DELPHI Collaboration, *The DELPHI detector at LEP*, Nucl. Instr. Meth. **A303** (1991), 233.

[50] K. Cieslik and P. Charpentier, *The DELPHI detector at CERN's LEP collider*, World Wide Web,

`http://delphiwww.cern.ch/$\sim$offline/physics/`
`delphi-detector.html`

and links within.

[51] DELPHI Collaboration, *Performance of the DELPHIA detector*, Nucl. Instr. Meth. **A378** (1996), 57.

[52] N. Bingefors *et al.*, *The DELPHI Microvertex detector*, Nucl. Instr. Meth. **A328** (1993), 447.

[53] V. Chabaud *et al.*, *The DELPHI silicon strip microvertex detector with double sided readout*, Nucl. Instr. Meth. **A368** (1996), 314.

[54] W. Adan *et al.*, *The Status of the DELPHI Very Forward Ministrip Detector*, contribution to the 6th International Conference for Experiments at e^+e^--Colliders, Novosibirsk, Russia, Feb. 1996.

[55] K. H. Becks *et al.*, *Progress in the construction of the DELPHI pixel detector*, DELPHI 96–56 MVX 14 (2. May 1996).

[56] K. Yip, *Extending the Muon Identification Tools Available for DELPHI Analysis*, unpublished

[57] J. Buytaert *et al.*, *The Forward Muon Detector of the DELPHI Experiment at LEP*, DELPHI 91–61 TRACK 65 (22. Jul 1991).

[58] The DELPHI Collaboration, *Proposal for the DELPHI Surround Muon Chambers*, DELPHI 92–139 TRACK 71 (21. Oct 1991).

[59] A. Maio *et al.*, *STIC, The New DELPHI Luminosity Monitor*, contribution to the Elba conference, 1993.

[60] V. Obrastsov *et al.*, *Status of the DELPHI Small angle TIle Calorimeter project*, contribution to the Marseille conference, 1993.

[61] DELPHI trigger group, Nucl. Instr. Meth. **A362** (1995), 361.

[62] G. Altarelli, T. Sjöstrand and F. Zwirner (ed.), *Cern Yellow Reports: Physics at LEP2*, CERN96–01 (1996).

[63] F. A. Berends, W. L. van Neerven and G. J. Burgers, Nucl. Phys. **B297** (1988), 429; erratum ibid. **B304** (1988), 921.

[64] DELPHI Collaboration, *DELSIM: DELPHI Event Generation and Detector Simulation*, Reference Manual, DELPHI 89–68 PROG 143 (1989), unpublished.

[65] T. Sjöstrand , Comp. Phys. Comm. **39** (1986), 347; *Pythia 5.7 and Jetset 7.4, Physics and Manual*, CERN-TH 7112/93 (1993).

[66] F. A. Berends, R. Kleiss and R. Pittan, Nucl. Phys. **B426** (1994), 344; Nucl. Phys. (Proc. Suppl.) **37B** (1994), 163-168.

[67] P. Janot in CERN Report 96-01, Vol. 2, 309. Comp. Phys. Comm. **39** (1986), 347.

[68] P. Roudeau, *Subroutine boostmc*,

`/afs/cern.ch/delphi/tasks/lep200/searches/teamc/boostmc.car`

[69] C. Parkes, *Standard Track Selection for Hadronic Analyses - short summary*, 26 June, 1998, unpublished.

[70] T. Spassoff, N. Smirnov, *SKELANA skeleton analysis program, Version 1.05, User's guide*, unpublished.

[71] P. Lutz, *DELPHI TEAM C standard track selection*, World Wide Web,

`http://www.fys.uio.no/~jorgenha/physics/teamc_partsel.html`

[72] F. Cossutti, M. Elsing, M. Espirito-Santo, C. Parkes, T. Spassoff, A. Tonazzo, C. Weiser, K. Beloous, V. Hedberg, F. Mazzucato, *Improvements to SKELANA for Version 2.0*, DELPHI 99-175 PROG 239 (1999).

[73] V. Ruhlmann-Kleider, *Subroutine ASSIGM*, private communication.

[74] DELPHI Collaboration, P. Abreu *et al.*, Eur. Phys. J. **C10** (1999), 563.

[75] P. Abreu, D. Fassouliotis, A. Grefrath, R. P. Henriques, L. Vitale, *SPRIME A Package for Estimating the Effective $\sqrt{s'}$ Center of Mass Energy in $q\bar{q}\gamma$ Events*, DELPHI 96-124 PHYS 632 (1996).

[76] P. Lutz, *Preselection performances for processing D2*, World Wide Web,

`http://www.fys.uio.no/~jorgenha/physics/presel_189D.html`

[77] B. Schorr *Spline estimation of Distributions and Density Functions*, Technical Report DD/75/13, CERN (1975).

[78] J. Shiers and M. Goossens, *CERNLIB – CERN Program Library Short Writeups*, CERN Computing and Networks Division (1995).

[79] J. Bunn and M. Goossens, *HBOOK Reference Manual*, CERN Computing and Networks Division (1995).

[80] R. P. Lippmann, *An Introduction to Computing with Neural Nets*, IEEE ASSP Magasine, Vol. 3, No. 4 (1987), 4.

[81] J. Schwindling, B. Mansoulié and O. Couet, *Using Multi-Layer Perceptrons in PAW*, World Wide Web,

`http://wwwinfo.cern.ch/asd/paw/mlpfit/pawmlp.html`

J. Schwindling, *MLPfit: a tool for Multi-Layer Perceptrons*, World Wide Web,

`http://schwind.home.cern.ch/schwind/MLPfit`

[82] G. V. Borisov *Lifetime Tag of events* $Z^0 \to b\bar{b}$ *with the DELPHI detector, AABTAG program*, DELPHI 94-125 PROG 208 (1994).

[83] G. Borisov *Combined b-tagging*, DELPHI 97-94 PHYS 716 (1997).

[84] M. Boonekamp *b-tagging with high* p_T *leptons*, DELPHI 98-54 PHYS 779 (1998).

[85] A. L. Read, *Optimal Statistical Analysis of Search Results based on the Likelihood Ratio and its Application to the Search for the MSSM Higgs Boson at* $\sqrt{s} = 161$ *and* 172 *GeV*, DELPHI 97–158 PHYS 737 (1997).

[86] M. Krawczyk, *Where is the Higgs boson?*, Acta Phys. Polon. **B29** (1998), 3543.

[87] P. H. Chankowski, M. Krawczyk and J. Żochowski, Eur. Phys. J. **C11** (1999), 661.

[88] M. Krawczyk, e-Print arXiv: hep-ph/0103223 (2001);
A. Dedes and H. E. Haber, e-Print arXiv: hep-ph/0105014 (2001).

[89] J. Ellis and D. Ross, e-Print arXiv: hep-ph/0012067 (2000).

[90] ALEPH Collaboration, e-Print arXiv: hep-ex/0201014 (2002), CERN-EP/2001-095 (2001).

[91] DELPHI Collaboration, Phys. Lett. **B499** (2001), 23; DELPHI 2001-078 CONF 506 (2001);
e-Print arXiv: hep-ex/0201022 (2002).

[92] L3 Collaboration, Phys. Lett. **B517** (2001), 319; L3 Note 2692 (2001);
Phys. Lett. **B503** (2001), 21.

[93] OPAL Collaboration, G. Abbiendi *et al.*, Phys. Lett. **B499** (2001), 38;
OPAL Physics Note PN472 (2001);
e-Print arXiv: hep-ex/9908002 (1999), CERN-EP/99-096 (1999).

[94] G. K. Bhattacharyya and R. A. Johnson, *Statistical Concepts and Methods*, John Wiley & Sons, Inc., ISBN 0–471–03532–7 (1977).