MSSM Higgs Mass Limit Studies Using the DELPHI Detector

Jan Ludvig Vinningland

Thesis submitted to the Cand. Scient. degree
Department of Physics
University of Oslo

August 2000

Abstract

This thesis presents multidimensional likelihood methods for calculating confidences and exclusion limits on the CP-even, neutral Higgs scalar h^0 of the MSSM Higgs sector using the $h^0A^0 \to b\bar{b}b\bar{b}$ channel at 189 GeV. The limits are calculated using the ALRMC implementation of a maximum likelihood ratio test.

The likelihood methods prepare signal, background, and observed candidate distributions using various binning algorithms and different discriminating variables. A total of six methods are presented, using mass estimators of the best and second best jet pairings as discriminating variables, in addition to a variable obtained from a neural network-like analysis. The three dimensional and two dimensional methods use mass estimators of the best pairing, while a five dimensional method uses mass estimators of the two best pairings.

The m_h limit for all methods at a 95 % confidence level are compared, and conclusions drawn on which methods are preferred.

Acknowledgement

Firstly I would like to thank my supervisor professor Lars Bugge for his eminent support and comments during the work with these interesting problems. I would also like to thank professor Alex L. Read for his most helpful assistance and patience. Ph.D. student Jørgen Hansen deserves special thanks for providing the 189 GeV data, and not giving up on my repetitive questions. Special thanks also to Ph.D. student Trond Myklebust for all his assistance whenever a computer problem occurred. I am also thankful to the other graduate and Ph.D. students for making the Experimental Particle Physics Group an excellent place to work.

Jan Ludvig Vinningland
Blindern, Oslo
June 15, 2000

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Particle physics is the science of the fundamental particles that constitute matter, and the interactions these particles participate in. The theory known as the Standard Model (SM) summarizes the present knowledge in particle physics. Several extensions of the SM exist, but no experiment has yet rendered the SM inadequate.

1.1 Fundamental Particles

The SM classifies the fundamental particles in two main groups: *Fermions* and *bosons*. Fermions are matter constituents, while bosons act as force carriers between the fermions.

Fermions are further grouped into *leptons* and *quarks*, and they exist in three different generations, or families. The three lepton families are the electron (e), the muon (μ), the tau (τ), and their neutrino partners (ν_e, ν_μ, ν_τ). The quarks come in six *flavors*: up (u), down (d), charm (c), strange (s), top (t) and bottom (b); two flavors in each family. Each fermion also has an antiparticle partner of the same mass but with opposite charge. In addition to the six flavors, quarks exist in three different *color* states. Thus, matter consists of 6 leptons and 18 quarks, while antimatter consists of 6 antileptons and 18 antiquarks. Figure 1.1 gives a list of the fermion properties.

The force carriers, or mediators, of the fundamental forces are known as gauge bosons. The strong force is mediated by eight gluons (g), the electromagnetic force by the photon (γ), and the weak force is mediated by W^+, W^- and Z^0 bosons. Gluons couple to quarks, photons couple to electric charge, and W^+, W^- and Z^0 couple to left-handed quarks and leptons. Left-handed means that the particle spin direction is opposite to its momentum. Figure 1.2 gives a list of the boson properties.

1.2 Fundamental Forces

Four fundamental forces exist in nature: The strong force, the weak force, the electromagnetic force, and the gravitational force. The strong force keeps the quarks inside the nucleus, and the strength of the force increases as the

FERMIONS matter constituents spin = 1/2, 3/2, 5/2,...

Leptons spin = 1/2			Quarks spin = 1/2		
Flavor	Mass GeV/c²	Electric charge	Flavor	Approx. Mass GeV/c²	Electric charge
ν_e electron neutrino	$<7\times10^{-9}$	0	u up	0.005	2/3
e electron	0.000511	−1	d down	0.01	−1/3
ν_μ muon neutrino	<0.0003	0	c charm	1.5	2/3
μ muon	0.106	−1	s strange	0.2	−1/3
ν_τ tau neutrino	<0.03	0	t top (initial evidence)	170	2/3
τ tau	1.7771	−1	b bottom	4.7	−1/3

Figure 1.1: Fermion properties in the Standard Model [1].

quarks separate. It is in fact impossible to separate a quark-antiquark pair, because the energy needed is infinite. An example of this is seen in e^+e^- annihilations at high energies, performed at several accelerators around the world including LEP (discussed in Chapter 3), in which a quark-antiquark ($q\bar{q}$) pair is produced. The high energy of the collision forces the $q\bar{q}$ pair to separate, but because quarks do not exist as single particles, the quarks instead appear as *jets* of hadrons. A hadron is a particle interacting via the strong force. The jet momentum along the original quark direction is large compared to the transverse jet momentum, and the jets are collimated around the q and $\bar{q}$ axes. The transverse momentum is used as a measure to determine which jet a track belongs to. Chapter 5 presents the clustering of the individual tracks into jets in more detail, in addition to a discussion on kinematical fits.

The weak and the electromagnetic force are different manifestations of a single force named the *electroweak* force. The weak force makes the stars shine, and is responsible for natural radioactivity, for example β-decay of a neutron in a nucleus. The electromagnetic force keeps the electrons in orbit around the nucleus and thus holds the atoms together. It also makes the atoms group together to form molecules and solid objects. Without the electromagnetic force all matter would exist as ionized plasma.

The gravitational force is responsible for the motions of the solar system and the universe. It is the weakest of all the fundamental forces, and when dealing with elementary particles, it can be neglected.[1]

The SM describes strong interactions (QCD), electromagnetic interactions (QED), and weak interactions. The gravitational force is not included in the SM, but extensions of the SM try to explain all four fundamental forces within one single theory. Glashow, Weinberg and Salam concluded that the weak and electromagnetic force can be described as one single electroweak force. Future extensions of the SM, Grand Unification Theories (GUT), may show that all the forces are in fact only different manifestations of a single force.

[1]The Planck mass ($\sim 10^{19}$ GeV) is the mass scale at which the gravitational coupling approaches unity, and at this mass scale the gravitational force may no longer be neglected.

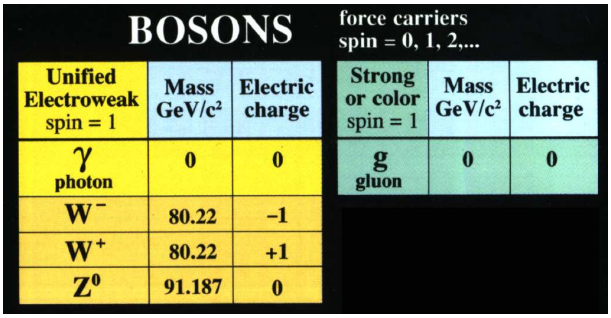

Figure 1.2: Gauge boson properties in the Standard Model [1].

1.3 Standard Model Problems

It is widely believed that the SM is not a final answer in particle physics. One of the most important arguments for a more fundamental theory is the great number of independent parameters needed as input to the SM: Eighteen parameters must be given by experiment (even more if the neutrinos are not massless) for which no SM explanations exist.

The missing understanding of the hierarchy of mass scales needed to explain the fundamental interactions, commonly known as the hierarchy problem, is another serious problem. Thus, physicists need to understand nature at a deeper level in order to give answers to these fundamental questions.

Among the extensions of the SM, supersymmetric (SUSY) extensions are particularly interesting because SUSY evades the hierarchy problem. The Minimal Supersymmetric extension of the Standard Model (MSSM) more than doubles the number of particles in the SM, because each SM particle adopts a SUSY partner.

1.4 The Higgs Mechanism

In order to explain how the fundamental particles acquire mass, the theory of the Higgs mechanism has been developed. This theory suggests that a certain symmetry of nature is broken through spontaneous symmetry breaking and, as a consequence of that, a Higgs particle is produced. In the SM only one Higgs particle exists, but the MSSM contains five Higgs particles. The Higgs particles are not easy to discover experimentally because they couple to fermions proportional to the mass of the fermion, and the most available fermions at particle accelerators are the light ones. Thus, the chance of discovering a Higgs particle is proportional to the amount of energy put into the accelerator. There has not yet been any experimental data supporting the existence of the Higgs particle. Therefore extensive Higgs searches are being carried out at particle physics laboratories and universities all over the world in order to discover the Higgs particle and measure its mass, or come up with an alternative theory.

Chapter 2 contains the theory of the SM, MSSM and Higgs mechanism.

1.5 The Analysis

This thesis presents the analysis performed to test different likelihood methods by comparing the obtained mass limits of the neutral MSSM Higgs scalar h^0. The search channel is the $h^0A^0 \to b\bar{b}b\bar{b}$ channel at 189 Gev. The backgrounds, signals and data candidates of the search are presented in Chapter 4.

The statistical treatment of the data, using a maximum likelihood test, are carried out by the ALRMC package which calculates confidence levels for the different signal hypotheses. Chapter 6 gives a brief discussion on the statistics of hypothesis testing in general, together with the specifics of ALRMC.

Chapters 7 and 8 present the different likelihood methods and algorithms developed to treat the signal, background and candidate distributions prior to ALRMC calculations.

The candidates used are the 189 GeV data collected by the DELPHI detector at CERN. The DELPHI detector and its subdetectors are outlined in Chapter 3.

Finally, the tests performed on the different likelihood methods are explained in Chapter 9, and the conclusions presented in Chapter 10.

Natural units ($\hbar c = 1$) are used throughout this thesis except where explicitly specified.

Chapter 2

The Theory

This chapter presents the Standard Model (SM), the Minimal Supersymmetric extension of the Standard Model (MSSM), the Higgs mechanism, and the Higgs spectrum of both the SM and the MSSM.

The SM Higgs theory is a one-doublet model which describes only one Higgs particle, ϕ^0. The SM theory and the SM Higgs mechanism are presented in Sec. 2.1. More complicated Higgs structures produce a larger number of Higgs particles, and thus additional possibilities for discovery. The simplest extensions of the one-doublet Higgs are models with two-doublets. No experiment has yet found evidence which favors the one-doublet model over more complicated structures [2]. Among the two-doublet models, the supersymmetric (SUSY) extension is particularly interesting because it solves some of the SM problems. The MSSM Higgs model produce two charged Higgs bosons, $H^{\pm}$, and three neutral Higgs bosons, h^0, H^0, and A^0. The motivations of SUSY and MSSM, together with the MSSM Higgs mechanism, are presented in Sec. 2.2.2.

The presentation of the SM theory and the SM Higgs sector are strongly inspired by the books "Dynamics of the Standard Model" by Donoghue, Golowich and Holstein [3], and "Quarks & Leptons" by Halzen and Martin [4]. The discussion on SUSY and the MSSM motivations are mainly influenced by the article "A 'theory of everything'?" by G.G.Ross [5], and the thesis "Search for Non-Minimal Higgs Bosons in Z^0 Decays with the L3 Detector at LEP" by André Sopczak [2]. The presentation of the MSSM Higgs sector is inspired by "The Higgs Hunter's Guide", by John F. Gunion, Howard E. Haber, Gordon Kane and Sally Dawson [6].

2.1 The Standard Model

The SM is a gauge theory, and within a gauge theory only two kinds of particles exist: Those who carry charge, and thus gives rise to currents, and those who mediate interactions between the currents by coupling directly to the charge. Fermions carry charge and produce currents, gauge bosons couples to the currents. If a gauge boson itself carries charge, the boson undergoes self-interactions. Charges and currents are treated in Sec. 2.1.3.

Quarks and leptons are both members of the fermion group. All leptons,

apart from the neutrinos, carry electric charge, while quarks in addition carry color charge. Neutrinos only participate in weak interactions, while the other leptons also experience electromagnetic interactions. Quarks participate in weak, electromagnetic and strong interactions.

Abelian[1] gauge bosons couple to currents with coupling constants determined by the charge of the particle producing the current (the coupling of the photon to the electron is $-e$). Since the non-Abelian gauge bosons are both charge carriers and mediators they undergo self-interactions that give rise to loop corrections complicating the renormalization of the theory.

A theory that is invariant under certain gauge transformations is said to be *gauge invariant*. Gauge transformations are classified as group transformations, and the SM is constructed to be gauge invariant under the $SU(3) \otimes SU(2)_L \otimes U(1)_Y$ gauge group transformation. The $SU(3)$ invariance is an exact symmetry and causes gluons to be massless, while the spontaneous breaking of the $SU(2)_L \otimes U(1)_Y$ invariance, through the Higgs mechanism, causes the $W^{\pm}$ and the Z^0 to be massive. How this comes about is discussed in Sec 2.1.5.

2.1.1 Fermions

Quarks and leptons are inputs to the SM in the sense that the model does not explain the variety and number of quarks and leptons, nor their properties. That is for experiments to decide.

Quarks and leptons exhibit certain empirical conservation laws that is not yet observed broken: Firstly each of the leptons belonging to a certain family has its own lepton number of value 1, which is conserved within the fermion family in every known interaction. Secondly quark flavor is conserved in all known strong interactions, but not in weak interactions.

Mixing occur between quarks, but not between leptons. The quark charged weak currents are responsible for the mixing. The quark states participating in transitions of the charged weak current are linear combinations of quark mass eigenstates. The quark-mixing matrix V contains the information about the linear combinations, and is named the *Cabbibo* matrix in the two quark family case, and the *Cabbibo-Kobayashi-Maskawa* matrix in the three family case after their inventors. The elements of the matrix are not predicted by theory, but have to be inferred from experiments.

In the three family case the matrix is a 3×3 matrix

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad , \tag{2.1}$$

where the subscripts indicate which quarks participate in the transitions. Because the V_{CKM} matrix contains a complex phase factor, the SM is not CP invariant, which is confirmed by experiments.

[1] Abelian groups is groups with a commutative group multiplication, in contrast to non-Abelian groups.

The diagonal elements of the V_{CKM} matrix are all very close to one. This reflects the experimental fact that transitions between quarks belonging to the same family are dominant.

The V_{CKM} matrix is unitary and, since quarks come in doublets, the electroweak theory fit experimental results showing that flavor changing neutral currents (FCNC) are highly suppressed. For example, the rate of neutral- to charged-current rates in kaon decay is [7]

$$\frac{K^+ \rightarrow \pi^+ \nu\overline{\nu}}{K^+ \rightarrow \pi^0 \mu^+ \nu_\mu} < 10^{-5} \quad . \tag{2.2}$$

2.1.2 Gauge Bosons

The SM describes strong interactions, electromagnetic interactions and weak interactions. The strong interactions are described by the $SU(3)$ group, and the electromagnetic and weak interactions are described by the $SU(2)_L \otimes U(1)_Y$ group as electroweak interactions.

The $SU(3)$ group is represented by eight generators. If the theory describing strong interactions is to be $SU(3)$ gauge invariant, eight gauge fields must be introduced, each field associated with a gauge boson. Thus, the eight gauge bosons are a consequence of the imposed $SU(3)$ gauge invariance, and the number of bosons corresponds to the number of generators of the group. The eight gauge bosons of the strong interaction are the massless, color charged gluons. Gluons are massless because the $SU(3)$ gauge invariance is an exact invariance.

Weak interactions are mediated by three massive gauge bosons, while electromagnetic interactions are mediated by one massless gauge boson. Weinberg and Salam showed that weak and electromagnetic interactions both arise from the electroweak force. The imposed $SU(2)_L \otimes U(1)_Y$ gauge invariance is not an exact invariance, because three of the gauge bosons are massive, and is broken via the Higgs mechanism to produce four gauge fields $W^1_\mu, W^2_\mu, W^3_\mu$ and B_μ. Two charged fields are defined as linear combinations of the W^1_μ and W^2_μ fields, and the charged fields correspond to two massive, charged gauge bosons, W^+ and W^-. Two neutral fields, corresponding to a massive Z^0 and a massless photon, are constructed by rotating the W^3_μ and B_μ fields by an angle θ_W (see (2.26)). The electroweak theory contains three massive bosons because the $SU(2)_L$ group is represented by three generators, and one massless boson because the $U(1)_Y$ group is represented by one single generator. The Higgs mechanism uses spontaneous symmetry breaking to produce massive bosons within a gauge invariant theory.

Local gauge invariance, spontaneous symmetry breaking, and transformations will be further discussed in sections 2.1.4 and 2.1.5.

2.1.3 Charges, Currents and Groups

Three different kinds of charge appear in the SM: color charge, weak isospin and weak hypercharge.

Color Charge

Each quark exists in one of three color states, red (R), blue (B) and green (G). These three color charges represent the fundamental representation of the $SU(3)$ symmetry group. The generators of this group are a set of eight 3×3 matrices denoted $\lambda_i, i = 1, ..., 8$, where the color states R,G,B are the base states of the group.

Weak Isospin

The three weak currents $J^i_\mu, i = 1, 2, 3$, one neutral and two charged, form an isospin triplet of weak currents whose corresponding charge operator is denoted $T^i, i = 1, 2, 3$. These three operators generate an $SU(2)_L$ algebra

$$[T^i, T^j] = i\varepsilon_{ijk}T^k \ . \tag{2.3}$$

The term 'isospin' is introduced because the isospin operators obey the same algebra as regular spin operators. The subscript L on $SU(2)$ is a reminder of the fact that weak currents only couple to left-handed fermions.

The symmetry of $SU(2)_L$ is broken by the observed weak neutral current because the neutral current has a right-handed component. However, the electromagnetic current comes to rescue because it is a neutral current with both right- and left-handed components. Neither the weak neutral current nor the electromagnetic current respects the $SU(2)_L$ symmetry alone, but two orthogonal combinations of the two currents diagonalize the mass matrix. One combination, J^3_μ, completes the weak current isospin triplet, while the other are the weak hypercharge current, J^Y_μ.

Weak Hypercharge

Weak hypercharge Y is defined by

$$Q = T^3 + \frac{Y}{2} \ , \tag{2.4}$$

with T^3 representing the isospin charge operator of J^3_μ, and Q being the electric charge operator of the electromagnetic current. The hypercharge operator Y generates a symmetry group $U(1)_Y$, and the electromagnetic interactions have been incorporated in the electroweak theory. By enlarging the symmetry group to $SU(2)_L \otimes U(1)_Y$ the electromagnetic and the weak interactions have been unified, but each group still has its own coupling strength. Because the gauge bosons of the weak interactions are massive, the $SU(2)_L \otimes U(1)_Y$ gauge group is spontaneously broken, induced by the Higgs mechanism, and an exact $U(1)_{em}$ gauge symmetry is left giving the photon zero mass.

The different groups presented in this section play a very important role in the SM. If, in the case of the $SU(3)$ group, the free Lagrangian of the quark color fields is constructed to be invariant under a $SU(3)$ transformation, the field theory of strong interactions (QCD) follows as a consequence. This kind of invariance is called *gauge invariance*.

2.1.4 Gauge Invariance

The SM is a quantum field theory. A quantum field theory is constructed by first quantizing the classical Lagrangian in order to produce fields that create and annihilate particles. The fields themselves are also described by Lagrangians, and the field equations are obtained by substituting the Lagrangian of each field into the Euler-Lagrange equation.

If a field $\psi(x_\mu)$ is given and a spacetime transformation $x_\mu \rightarrow x_\mu + a_\mu$ leaves the physics of the field unchanged, the field is said to be *invariant* under spacetime transformations; the field $\psi(x_\mu)$ has a spacetime symmetry.

Other symmetries exist apart from the spacetime symmetry, and Noether's Theorem states that for any invariance[2] there exists a classical, time-independent charge Q associated with a conserved current, $\partial_\mu J^\mu = 0$.

Spacetime symmetry is an external symmetry and leads to the conservation of energy and momentum. Internal symmetries are more interesting, and in particular phase transformations written as

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

where α is a real constant. Since α is independent of x_μ, (2.5) is called a *global* phase transformation. Phase transformations of the kind $U(\alpha) \equiv e^{i\alpha}$ forms a unitary Abelian group; the $U(1)$ group.

A generalization of (2.5) gives the transformation

$$\psi(x_\mu) \rightarrow e^{i\alpha(x^\mu)}\psi(x_\mu) \ , \tag{2.6}$$

where $\alpha(x^\mu)$ now depends on time and space coordinates. The transformation (2.6) is known as a *local* gauge transformation; it is possible to define $\alpha(x^\mu)$ locally in space and time, not just globally as in (2.5).

U(1) Gauge Invariance

As an example, the Lagrangian of the Dirac field equation is given by

$$\mathcal{L} = i\overline{\psi}\gamma_\mu\partial_\mu\psi - m\overline{\psi}\psi \ . \tag{2.7}$$

If (2.7) is supposed to be $U(1)$ gauge invariant, a covariant derivative

$$D_\mu \equiv \partial_\mu - ieA_\mu \tag{2.8}$$

must be constructed to replace ∂_μ. A_μ is the gauge field, and if (2.7) is to be gauge invariant it must transform as

$$A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu\alpha \tag{2.9}$$

where α is the same α as in (2.5). The gauge field A_μ is a consequence of the imposed gauge invariance on (2.7), and A_μ gives rise to a gauge boson, the photon. Thus, by demanding gauge invariance of the Dirac Lagrangian (2.7),

[2]Invariance of the action under a continuous transformation.

plus adding a term corresponding to the kinetic energy of the photon field[3], the QED Lagrangian follows:

$$\mathcal{L}_{QED} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi + e\overline{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.10}$$

A mass term like $\frac{1}{2}m^2A_\mu A^\mu$ is prohibited in (2.10) by gauge invariance, and the photon is massless. If massive bosons are to derive from a gauge theory, the Higgs mechanism is needed.

SU(3) Gauge Invariance

If the $SU(3)$ gauge group of phase transformations is imposed on the quark color fields, the Lagrangian of QCD is derived. The derivation is a bit more complicated than in the QED case because $SU(3)$ is a non-Abelian group. Gauge invariance of the quark color field leads to eight vector gluon fields G^a_μ, $a = 1, ..., 8$, that correspond to eight massless gluons interacting with quarks and gluons as discussed in Sec. 2.1.3. The final gauge invariant QCD Lagrangian, for the simplified case of one quark flavor, is given by

$$\mathcal{L} = \bar{q}(i\gamma^\mu\partial_\mu - m)q - g_3(\bar{q}\gamma^\mu T_a q)G^a_\mu - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a\ , \tag{2.11}$$

where q is the quark color field, T_a, $a = 1, ..., 8$, are the eight generators of the $SU(3)$ group, and g_3 is the coupling constant of $SU(3)$.

2.1.5 The Higgs Mechanism

Because gauge invariance prohibits mass terms of the gauge fields in the Lagrangian, alternative techniques must be used if a gauge theory of massive fermions and bosons is to be constructed. The Higgs mechanism produces massive fermions and gauge bosons by spontaneously breaking a local gauge symmetry, and introduces a new massive particle; the Higgs particle. Thus, observation of the Higgs particle is a very good signature of the Higgs mechanism. This section presents a more elaborate discussion on the Higgs mechanism and try to explain how elementary particles acquire mass.

The Higgs mechanism has not yet been confirmed by experimental data. Future experiments may reveal that the Higgs theory is not the correct one after all, but the general opinion among particle physicists is that there must be something left to discover beyond the SM. Whatever that may be, it could provide answers to the problems presently encountered in the SM.

On Symmetries

A given symmetry of the Lagrangian can behave in a number of different ways.

- The symmetry remains exact. $U(1)$ symmetry of QED and $SU(3)$ symmetry of QCD are examples, and the reason why photons and gluons are massless.

[3] Because the kinetic energy term must be gauge invariant, it can only involve the invariant field strenght tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

- The apparent symmetry may have an anomaly, and is not a true symmetry.
- The symmetry may be explicitly broken by terms in the Lagrangian which are not invariant under the symmetry. Isospin symmetry, broken by electromagnetism, and light-quark (up and down) mass difference, are examples.
- The symmetry may be hidden, meaning that it is an invariance of the Lagrangian but not of the ground state. This may come about if one or more scalar fields acquire vacuum expectation values, and thus symmetry is spontaneously broken.

If a theory of massive fermions and gauge bosons is to be renormalizable, spontaneous symmetry breaking is required.

Spontaneous Symmetry Breaking

The $SU(2)$ gauge invariant Lagrangian density, containing the Higgs-gauge boson coupling, is given by

$$\mathcal{L}_{HG} = (D^\mu \phi)^* D_\mu \phi - V(\phi) \quad , \tag{2.12}$$

where ϕ is an $SU(2)$ doublet of complex scalar fields

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad , \tag{2.13}$$

the covariant derivative ensuring $SU(2)$ invariance is

$$D_\mu = \partial_\mu + ig\frac{\tau_a}{2} W_\mu^a \quad , \tag{2.14}$$

and the potential

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \quad , \tag{2.15}$$

with $\mu^2 > 0$ and $\lambda > 0$. The global $SU(2)$ phase transformation is given by

$$\phi \to e^{i\alpha_a \tau_a/2}\, \phi \quad . \tag{2.16}$$

Three gaugefields are represented in (2.14) by $W_\mu^a(x^\mu)$ with $a = 1, 2, 3$, and τ_a denote the three generators of the group. The number of generators and the number of fields are always the same in a gauge theory. If (2.14) is compared to (2.8) it is easy to see that the $U(1)$ group has only one generator, while the $SU(2)$ group has three generators.

The minima of (2.15) is given by

$$\frac{\partial V}{\partial \phi} = \phi^\dagger(-\mu^2 + 2\lambda\phi^\dagger\phi) = 0 \tag{2.17}$$

and the non-trivial solution of these minima is

$$\langle \phi^\dagger \phi \rangle_0 = \frac{v^2}{2} \ , \tag{2.18}$$

with

$$v \equiv \sqrt{\frac{\mu^2}{\lambda}} \ . \tag{2.19}$$

The original symmetry is now broken in the ground state by expanding ϕ about a particular minimum

$$\phi_1 = \phi_2 = \phi_4 = 0, \quad \phi_3^2 = v^2 \ , \tag{2.20}$$

giving the vacuum expectation value

$$\langle \phi \rangle_0 = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \ . \tag{2.21}$$

The expansion looks like

$$\phi(x^\mu) = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x^\mu) \end{pmatrix} \ , \tag{2.22}$$

where $h(x^\mu)$ is the expanded Higgs field around the vacuum expectation value v. The four degrees of freedom of the Higgs doublet are transformed into massive $W^\pm$ and Z^0 bosons, and a neutral, physical Higgs field. The Higgs boson mass is given as the curvature of the Higgs field at the minimum of the potential (2.15) [2]

$$m_H^2 = \frac{\partial V^2}{\partial^2 \phi}|_{\phi=v} = 2v^2\lambda \ . \tag{2.23}$$

The mass of the Higgs boson is *not* predicted by the theory, since the Higgs self-coupling λ is a free parameter.

Gauge Boson Masses

To get the masses of the gauge bosons, substitute (2.21) and a definition of the charged Higgs fields into the Lagrangian (2.12). It leads to a mass term for the charged gauge bosons

$$m_W = \frac{v}{2} g \ , \tag{2.24}$$

where g is the coupling constant of $SU(2)_L$. A mixing occur in the neutral gauge bosons, but the fields Z_μ and A_μ diagonalize the mass matrix, as discussed in Sec. 2.1.3 and Sec. 2.1.2, and lead to

$$m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} \ \text{ and } \ m_\gamma = 0 \ , \tag{2.25}$$

where g' is the coupling constant of $U(1)_Y$. The ratio between the two coupling constants defines the mixing angle θ_W, $\tan\theta_W \equiv g/g'$. The Higgs mechanism thus leads to the important mass relation

$$\frac{m_W}{m_Z} = \cos\theta_W \quad . \tag{2.26}$$

Any extension of the one-doublet Higgs model must keep this fundamental relation in agreement with experiment, and a ρ-parameter is defined by

$$\rho = \frac{m_W}{m_Z \cos\theta_W} \quad . \tag{2.27}$$

The experimental world average is [2]

$$\rho = 0.995 \pm 0.013 \quad , \tag{2.28}$$

and thus in perfect agreement with the model.

Fermion Masses

The complete Higgs Lagrangian $\mathcal{L}_H$ also contains a Higgs-fermion coupling term $\mathcal{L}_{HF}$

$$\mathcal{L}_H = \mathcal{L}_{HG} + \mathcal{L}_{HF} \quad , \tag{2.29}$$

but for simplicity, only the Higgs-electron Lagrangian is given here

$$\mathcal{L}_{He} = -g_e \left[(\overline{\nu}_e \, e)_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \overline{e}_R (\phi^- \phi^{0^*}) \begin{pmatrix} \nu_e \\ \overline{e} \end{pmatrix}_L \right] \quad . \tag{2.30}$$

Analogous terms exist for the other fermions. If the expanded Higgs field vacuum expectation value (2.21) is substituted into (2.30), the electron mass is obtained. The masses of the other fermions follow equivalently, and are given by

$$m_f = \frac{v}{\sqrt{2}} g_f \quad , \tag{2.31}$$

with g_f being the Higgs-fermion coupling constant. Thus, even if v is known, the masses of the fermions still remain unpredicted. Equation (2.31) also imply that the Higgs boson tend to decay into the most massive kinematically accessible fermion pair, and thus makes Higgs observations depend on the accelerator-energy.

2.1.6 Higgs Production

The main SM Higgs production mechanism in e^+e^- collisions at LEP2 energies is

$$\text{Higgs-strahlung} \; : \quad e^+e^- \to \phi^0 Z^0 \, . \tag{2.32}$$

The fusion process

$$W^+W^- \text{ fusion} \quad : \quad e^+e^- \to \bar{\nu_e}\nu_e\phi^0 \tag{2.33}$$

has a considerably smaller cross section at LEP2 energies. The Feynman diagrams for the processes (2.32) and (2.33) are given in Fig. 2.1.

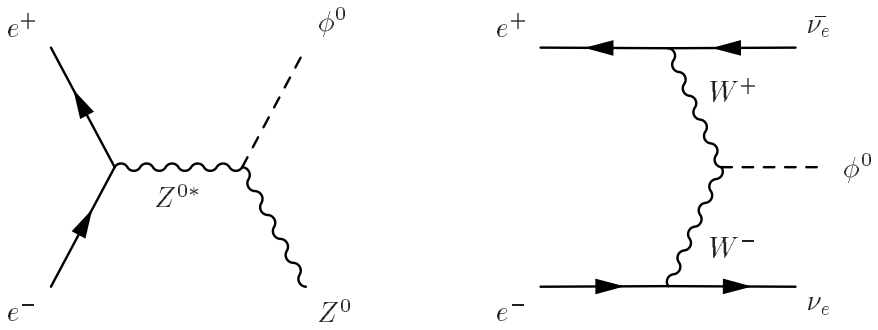

Figure 2.1: Fusion Feynman diagrams of Higgs production in the Standard Model. The left diagram is the Higgs-strahlung, and right diagram is the W^+W^--fusion

2.2 Supersymmetry

The SM is in perfect agreement with all existing data. Why is it then necessary to construct theories that go beyond the SM? Even if the SM explains all particle physics observations, many questions still remain unanswered if physicists are to understand nature at a deeper level. Such a theory is often referred to as a 'theory of everything'.

One of the problems suggesting that the SM is not final is the large number of unpredicted parameters in the model which have to be determined by experiment. These parameters include the three gauge couplings of the three gauge groups ($SU(3)$, $SU(2)_L$ and $U(1)_Y$), the six quark and the three lepton masses (as seen in Eq.(2.31)), the three relative mixing angles between the quark families which specify how the $W^{\pm}$ boson couples to quarks, the phase which determines the magnitude of CP violation, the mass of the $W^{\pm}$ boson, and finally the mass of the Higgs scalar [5]. A total of eighteen parameters. A fundamental theory has to explain the origin of these parameters from first principles.

A second problem is the origin of the structure assumed when constructing the SM: Why do fermions only interact with the weak current in a left-handed manner? Why are the charges of the quarks quantized in fractions one-third the charges of the leptons? Why are there three families of fermions, each having the same coupling to the gauge bosons but greatly different mass?

Further, the strengths of the interactions in the SM, the coupling constants, are not related by the model and the SM does not explain why they should be so different when measured. The Higgs mechanism spontaneously breaks the weak gauge symmetry and thus gives mass to the weak gauge bosons. But no reason is given for why the weak gauge bosons are massive, and not the electromagnetic nor strong gauge bosons.

Another flaw of the SM is that it only describes three of the four fundamental forces. Gravity is not included in the SM. A theory of everything would have to describe gravitational interactions as a quantum theory. At present the most

Figure 2.2: Some quadratically divergent Higgs self-energies.

promising candidate of such a theory is the String Theory, which will not be reviewed here.

However, the perhaps strongest argument for an extension of the SM is the hierarchy problem: What generates the hierarchy of mass scales needed to describe the fundamental interactions?

2.2.1 The Hierarchy Problem

The hierarchy of masses goes as follows: The Planck mass m_{Planck}, governing the gravitational coupling, is approximately 10^{19} GeV. The $W^{\pm}$ boson mass m_W, governing the weak coupling, is some 17 orders of magnitude smaller than m_{Planck}, approximately 10^2 GeV. The fermion masses range from $\frac{1}{2}10^{-3}$ GeV for the electron to $\frac{7}{4}10^2$ GeV for the top quark, while the neutrinos are close to being massless. The central question is what causes such a mass hierarchy? To complicate the problem even further, radiative corrections to the electroweak breaking scale estimated in the SM seem to drive the breaking scale up to the Planck scale.

The hierarchy problem is connected to the point-like nature of elementary particles. Virtual processes, in which a particle splits into more than one elementary state, lead to large radiative corrections to the particle mass. A one-loop electromagnetic correction to the electron mass, m_e, is quite modest. A cutoff Λ_e is imposed on the momentum flowing through the loop, and even if Λ_e is the entire mass of the observable universe the correction only amounts to $m_e \simeq 1.7 m_{e,0}$ [3]. The radiative correction to the Higgs mass does not behave quite as nice. The first diagram in Fig. 2.2 shows a Higgs loop with quadratic self-coupling. The shift of the Higgs mass is given by [3]

$$m_H^2 = m_{H,0}^2 + \frac{3}{16}\frac{\lambda}{\pi^2}\Lambda_H^2 \quad . \tag{2.34}$$

If Λ_H is as large as the Planck mass, the parameter $m_{H,0}^2$ must be negative and fine-tuned up to 30 decimal places, which is highly unnatural, if a renormalized mass governed by the electroweak breaking scale is to be obtained. Because of this unnatural tuning, the hierarchy problem is sometimes also referred to as the naturalness problem.

Even if these radiative processes are classically forbidden, they can proceed provided, in accordance with the uncertainty principle, the virtual states only exist for a short period of time. These radiative corrections push the mass of

the Higgs towards the Planck mass, which in turn leads to $W^{\pm}$ and Z^0 masses of the same order as the Higgs. Suggestions have therefore been made that the Higgs is not an elementary particle, but is composite and has structure on a certain mass scale, making virtual processes impossible above this mass scale. Thus, radiative corrections do not blow up the Higgs mass. Technicolour is one of these composite theories in which the Higgs boson is a fermion-antifermion state with a new class of fermions called techniquarks.

But what if all the SM states, including the Higgs, really are elementary? How do we then evade the hierarchy problem? How do we avoid radiative corrections driving the Higgs mass to the Planck scale? If there exists a symmetry in nature under which, if the symmetry is exact, the electroweak group is unbroken, then the radiative corrections are absent. The $W^{\pm}$ and Z^0 masses would be naturally small if the breaking of the symmetry is small. Only one symmetry consistent with relativity can play this role: Supersymmetry (SUSY).

2.2.2 MSSM

The Minimal Supersymmetric extension of the Standard Model (MSSM) is the top candidate of a non-composite theory without the hierarchy or naturalness problem. In MSSM, every elementary particle adopts a partner with a half integer spin difference compared to the SM particles. Thus, every SM fermion state is related to a SUSY boson state, and every SM boson state is related to a SUSY fermion state. In contrast, the local gauge symmetry of the SM only relates states of the same spin.

The partners of spin-$\frac{1}{2}$ quarks and leptons are spin-0 squarks and sleptons, respectively. The gluon, $W^{\pm}$, Z^0, and photon gauge bosons, adopt spin-$\frac{1}{2}$ partners named gluino, Wino, Zino and photino, respectively. These new states are considerably heavier than the SM states, and thus explains why no direct evidence of the existence of the SUSY particles has yet been found in experiments. The fact that the SUSY particles are heavier than their SM partners imply that SUSY must be broken. Hence, SUSY is not an exact symmetry.

SUSY particles contribute as virtual states to radiative corrections of the masses, even though they are too heavy to be observed on-shell. The crucial difference now is that the radiative corrections of the new states is of the same magnitude as the radiative corrections of the SM states but with *opposite* signs. If SUSY were an exact symmetry, there would be *no* corrections to the Higgs mass. But because SUSY is broken on a scale m_{SUSY}, the new SUSY states acquire masses of order $O(m_{SUSY})$. Cancellation is spoilt. However, no need for panic. If the resultant contribution of order $O(m_{SUSY})$ is not too large (i.e. $\leq O(1\,\text{TeV})$), it is consistent with the observed electroweak breaking. The hierarchy problem is solved. But the price paid is more than twice as many states as in the SM, and no explanations of the MSSM multiplet structure and couplings are provided.

Another important aspect of SUSY is encountered in connection with GUTs (Grand Unified Theories). If the three gauge couplings of the SM are to be unified as one single coupling constant, SUSY is needed. Without SUSY, the three couplings never meet at the same point [8].

2.2.3 Constraints

Extensions of the SM Higgs sector must satisfy three general constraints arising from the ρ-parameter measurements, absence of flavor changing neutral currents, and unitary requirements [2]. The two first are experimental constraints, while the last is a theoretical constraint.

ρ-parameter

Extensions of the Higgs sector must not spoil the successful predictions of the $W^{\pm}$ and Z^0 mass relation, that is to deviate from $\rho = 1$ (see (2.27)). If measurements indicate $\rho \neq 1$, new physics beyond the SM are required. It is therefore of great importance to determine ρ as accurately as possible. Higgs models consisting of any number of singlets and doublets satisfy $\rho = 1$. Other ways of satisfying the ρ measurements are restricted to models satisfying [6]

$$(2T+1)^2 - 3Y^2 = 1 \ , \tag{2.35}$$

with T and Y representing the total $SU(2)_L$ isospin and hypercharge, respectively. The two-doublet model with $T = \frac{1}{2}$ and $Y = \pm 1$ satisfy (2.35). Possibilities beyond $T = \frac{1}{2}$, $Y = \pm 1$ are usually dismissed because of complicated representations.

Absence of FCNC

As mentioned in section 2.1.1, flavor changing neutral currents (FCNC) are highly suppressed (see (2.2)). The absence of FCNC strongly restricts possible SM extensions. A Higgs mass of order 1 TeV would sufficiently suppress tree-level FCNC mediated by Higgs exchange. A more elegant possibility is restricted to models with more than one Higgs doublet, and is based on a theorem stating that tree-level FCNC are absent if fermions of a given electric charge only couples to one Higgs doublet. The last possibility is favored over an unnaturally high Higgs boson mass [2].

The coupling of the Higgs bosons to fermions in a two-doublet model generally proceed through one of two scenarios: The first possibility is that one doublet couples to up-type fermions and the other doublet couples to down-type fermions. MSSM requires this structure because the two Higgs doublets have opposite hypercharge, and the $Y = -1$ doublet only couples to down-type fermions while the $Y = 1$ doublet only couples to up-type fermions. This choice is also required if both up- and down-type fermions are to be massive in a supersymmetric theory.

Another two-doublet model avoiding FCNC is a model in which one doublet couples to both type of fermions, while the other doublet couples to none of the fermions. This kind of a two-doublet model is not useful if the doublets have opposite hypercharge.

To sum up: A two-doublet model is favored over a high Higgs boson mass, and MSSM requires the two doublets to couple to up- and down-type fermions separated.

Unitarity Condition

The unitarity condition is not only useful as a MSSM constraint: In the Fermi theory of β-decay the four fermions involved are assumed to have a pointlike interaction specified by the Fermi constant G_F. This assumption is correct at low energies, but at high energies the scattering amplitude $\mathcal{M}$ is highly divergent and thus violates the unitarity condition

$$|\mathcal{M}(s)| \leq 1 \ , \tag{2.36}$$

where s represents the available energy. The introduction of a heavy intermediating particle, the $W^{\pm}$ boson, prevents the divergence, and the unitarity condition is satisfied. A scattering amplitude greater than one implicates a probability greater than one, which clearly shows the importance of the unitarity condition.

Another example of the usefulness of the unitarity condition provide an alternative argument for the SM Higgs: The unitarity condition must be satisfied in $V_L V_L \to V_L V_L$ and $f_+ \bar{f}_+ \to V_L V_L$ amplitudes by any electroweak breaking model. (V_L is a left-handed $W^{\pm}$ or Z^0 and f_+ is a fermion with positive helicity.) If a gauge theory is to be renormalizable, non-trivial cancellations among Feynman diagrams of a given process are required. As an example, the scattering amplitude of the tree level $W^+W^- \to W^+W^-$ interaction is given by

$$\mathcal{M}_{WWWW} = -g^2 \frac{s}{4m_W^2} + O(1) \ . \tag{2.37}$$

The divergence of (2.37) is cured by introducing a diagram involving the exchange of a (neutral) spin-0 particle, i.e. by introducing a new interaction of the vector field W with a scalar field. The new particle is the SM Higgs boson ϕ^0, and the tree-level relation $g_{\phi^0 WW} = g m_W$, where g is the gauge coupling, guarantees cancellation of the growing energy terms. The Feynman diagram of the $W^+W^- \to W^+W^-$ interaction is shown in Fig. 2.3. The Higgs particle is often explained as the origin of massive fermions and bosons, but the unitarity condition alone provides an evidence for a new scalar particle if the SM is to be renormalizable at a certain energy level.

The MSSM two-doublet model do not depend on a single scalar boson to cure these unitarity problems alone. It suffices to obey the following sum rules for the scalar boson VV and $f\bar{f}$ couplings

$$\sum_i g^2_{h_i^0 VV} = g^2_{\phi^0 VV} \ , \tag{2.38}$$

$$\sum_i g_{h_i^0 VV} g_{h_i^0 f\bar{f}} = g_{\phi^0 VV} g_{\phi^0 f\bar{f}} \ . \tag{2.39}$$

These sum rules only apply if the Higgs fields are made of doublets and singlets.

If all these constraints, the ρ-parameter, the absence of FCNC and the unitarity condition are to be satisfied, the MSSM Higgs theory must be a two-doublet theory.

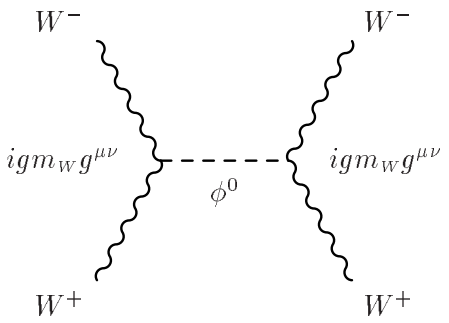

Figure 2.3: The lowest order Feynman diagram for $W^+W^- \to W^+W^-$ involving a neutral, scalar particle; the Standard Model Higgs ϕ^0.

2.2.4 The MSSM Higgs Model

In section 2.1.5 the spontaneous symmetry breaking of a one-doublet Higgs model is discussed. The spontaneous breaking of a two-doublet model follows the same structure as in the one-doublet case, although the Higgs potential is more complex, and the Higgs spectrum is richer: The two-doublet model has eight degrees of freedom, in contrast to the four degrees of freedom in the one-doublet case. It is important to point out that the two doublet model is not particular for SUSY models. Two-doublet models may exist within the SM as well, but MSSM requires a two-doublet, as discussed in Sec. 2.2.3

Two complex $|Y| = 1$, $SU(2)_L$ doublet scalar fields are given by

$$\phi_1 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix} , \qquad \phi_2 = \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix} . \tag{2.40}$$

A gauge invariant Higgs potential that avoids FCNC[4] has the form

$$\begin{aligned} V(\phi_1,\phi_2) \;=\; & \lambda_1(\phi_1^\dagger\phi_1 - v_1^2)^2 + \lambda_2(\phi_2^\dagger\phi_2 - v_2^2)^2 \\ & + \lambda_3\left[(\phi_1^\dagger\phi_1 - v_1^2)(\phi_2^\dagger\phi_2 - v_2^2)\right]^2 \\ & + \lambda_4\left[(\phi_1^\dagger\phi_1)(\phi_2^\dagger\phi_2) - (\phi_1^\dagger\phi_2)(\phi_2^\dagger\phi_1)\right] \\ & + \lambda_5\left[\mathrm{Re}\,(\phi_1^\dagger\phi_2) - v_1 v_2\cos\xi\right]^2 \\ & + \lambda_6\left[\mathrm{Im}\,(\phi_1^\dagger\phi_2) - v_1 v_2\sin\xi\right]^2 . \end{aligned} \tag{2.41}$$

The vacuum expectation values that minimize this potential for arbitrary real and positive parameters λ_i, $i = 1, ..., 6$, are

$$\langle\phi_1\rangle_0 = \begin{pmatrix} 0 \\ v_1 \end{pmatrix} , \qquad \langle\phi_2\rangle_0 = \begin{pmatrix} 0 \\ v_2 e^{i\xi} \end{pmatrix} . \tag{2.42}$$

If $\sin\xi \neq 0$, it leads to large CP violations in contradiction with measurements, and the phase factor ξ is therefore set to zero. Thus, the Higgs sector in MSSM is CP invariant even though the MSSM as a whole is not CP invariant. There is no argument for why the CP violation should occur in the Higgs sector.

[4]FCNC are avoided by making the Higgs potential respect the discrete symmetry $\phi_1 \leftrightarrow -\phi_1$.

Higgs Spectrum

A spontaneous breaking of the symmetry has occurred, and the Higgs spectrum is obtained by expanding the Higgs fields around their vacuum. Three Goldstone bosons are identified by their derivative couplings to the three W_μ fields appearing in the covariant derivative (2.14). The resulting gauge boson masses are given by

$$m_Z^2 = (v_1^2 + v_2^2)\frac{g^2}{2\cos\theta_W}, \quad m_\gamma = 0, \quad m_W^2 = (v_1^2 + v_2^2)\frac{g^2}{2} . \tag{2.43}$$

If (2.43) is compared to (2.24) and (2.25) it follows that the quadratically summed v_1 and v_2 must be equal to the square of the SM vacuum expectation value v. A key parameter in MSSM is defined as the ratio between the two vacuum expectation values:

$$\tan\beta \equiv v_2/v_1 . \tag{2.44}$$

The three Goldstone bosons must be removed if the physical Higgs states are to be determined. The Higgs states are orthogonal to the Goldstone bosons. In the charged sector the physical states are given by

$$H^\pm = -\phi_1^\pm \sin\beta + \phi_2^\pm \cos\beta . \tag{2.45}$$

Because of the assumed CP-invariance the imaginary and real parts of the neutral scalar fields decouple. The neutral Goldstone orthogonal Higgs state belongs to the imaginary (CP-odd) sector and is given by

$$A^0 = \sqrt{2}(-\mathrm{Im}\phi_1^0 \sin\beta + \mathrm{Im}\phi_2^0 \cos\beta) . \tag{2.46}$$

In the real (CP-even) sector the mass matrix $\mathcal{M}$ that mix the two neutral higgs bosons is given by

$$\mathcal{M} = \begin{pmatrix} 4v_1^2(\lambda_1+\lambda_3) + v_2^2\lambda_5 & (4\lambda_3+\lambda_5)v_1v_2 \\ (4\lambda_3+\lambda_5)v_1v_2 & 4v_2^2(\lambda_1+\lambda_3) + v_1^2\lambda_5 \end{pmatrix} , \tag{2.47}$$

and the mass eigenstates are

$$\begin{aligned} H^0 &= \sqrt{2}\left[(\mathrm{Re}\phi_1^0 - v_1)\cos\alpha + (\mathrm{Re}\phi_2^0 - v_2)\sin\alpha\right] , \\ h^0 &= \sqrt{2}\left[(-\mathrm{Re}\phi_1^0 - v_1)\sin\alpha + (\mathrm{Re}\phi_2^0 - v_2)\cos\alpha\right] . \end{aligned} \tag{2.48}$$

The masses of the Higgs bosons are given by

$$\begin{aligned} m_{H^\pm}^2 &= \lambda_4(v_1^2 + v_2^2) , \\ m_{A^0}^2 &= \lambda_6(v_1^2 + v_2^2) , \\ m_{h^0,H^0}^2 &= \frac{1}{2}\left[\mathcal{M}_{11} + \mathcal{M}_{22} \pm \sqrt{(\mathcal{M}_{11} - \mathcal{M}_{22})^2 + 4\mathcal{M}_{12}^2}\right] . \end{aligned} \tag{2.49}$$

At tree-level the Higgs masses simplify to

$$\begin{aligned} m_{H^\pm}^2 &= m_{A^0}^2 + m_{W^\pm}^2 , \\ m_{h^0,H^0}^2 &= \frac{1}{2}\left[m_{A^0}^2 + m_{Z^0}^2 \pm \sqrt{(m_{A^0}^2 + m_{Z^0}^2)^2 - 4m_{Z^0}^2 m_{A^0}^2 \cos^2 2\beta}\right] . \end{aligned} \tag{2.50}$$

The mixing angle α that diagonalizes the mass matrix $\mathcal{M}$ can in the tree-level case be expressed as

$$\cos^2(\beta-\alpha) = \frac{m_{h^0}^2(m_Z^2 - m_{h^0}^2)}{(m_{H^0}^2 - m_{h^0}^2)(m_{H^0}^2 + m_{h^0}^2 - m_Z^2)}. \tag{2.51}$$

The mass of the lightest MSSM neutral Higgs particle h^0 is bound to be smaller than the Z^0 mass at tree-level, but radiative corrections raise the upper mass limit of h^0 to about 150 Gev [9].

To summarize, the following Higgs spectrum has been obtained:

- One neutral CP-odd scalar A^0, often called a pseudoscalar.
- Two neutral CP-even scalars H^0 and h^0.
- Two charged scalars $H^\pm$.

The initial eight degrees of freedom (six λ parameters and two vacuum expectation values) have been turned into five Higgs boson mass parameters $m_{H^\pm}$, m_{A^0}, m_{H^0}, m_{h^0} and $\tan\beta$, while the remaining three have been absorbed in order to give mass to the three gauge bosons Z^0, W^+, W^-. The two-doublet model has six free parameters: Four Higgs masses, $\tan\beta$ and the mixing angle α. At tree-level the Higgs spectrum is determined by the weak gauge boson masses, the CP-odd Higgs boson mass, m_{A^0}, and $\tan\beta$. The other parameters of the model only enter through radiative corrections to the Higgs masses.

2.2.5 Higgs couplings

The Higgs couplings control production and decay of the Higgs bosons. The J^{PC} quantum numbers for the Higgs bosons determine which processes are allowed and which are not. If fermions are ignored, every boson of the MSSM theory is assigned a unique J^{PC} quantum number. If fermions are included P and C are no longer separately conserved, although CP still remains a good quantum number.

The J^{PC} quantum numbers of A^0, Z^0 and $W^\pm$ is 0^{+-}, 1^{--} and 1^-, respectively, and explain why there are no tree-level $A^0W^+W^-$ and $A^0Z^0Z^0$ couplings. Another argument is that in a CP conserving theory[5] the spontaneous symmetry breaking mechanism does not generate a coupling for the CP-odd A^0. The coupling only occur at the one-loop level through fermion loops. Because of this, there can be no A^0 bremsstrahlung emission of Z^0 or $W^\pm$ at tree-level.

The coupling of the Z^0 to a pair of identical Higgs bosons, $Z^0h^0h^0$ or $Z^0A^0A^0$, is forbidden by Bose symmetry. In the case of a nonidentical Higgs pair, the coupling is only present if the two Higgses have opposite CP quantum numbers, i.e. pair production.

[5]Recall that the ξ factor in (2.41) were set to zero to avoid large CP violations

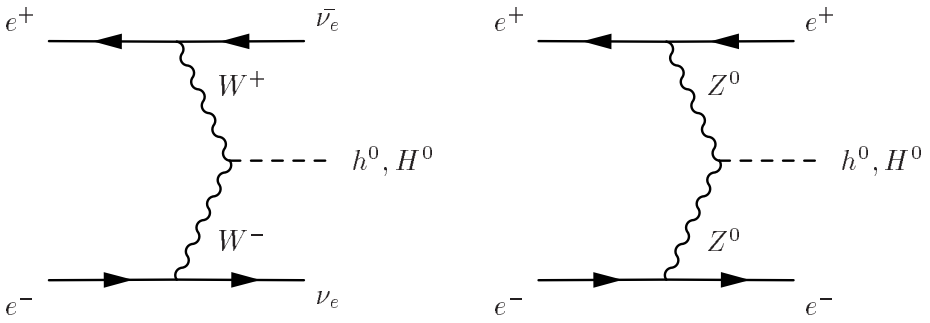

Figure 2.4: Fusion Feynman diagrams of Higgs production in the MSSM model. Left diagram is W^+W^--fusion, and right diagram is Z^0Z^0-fusion.

The couplings of H^0 and h^0 to W^+W^- and Z^0Z^0 are suppressed compared to the SM Higgs couplings according to the following sum rule derived from the more general rule (2.38)

$$g^2_{h^0VV} + g^2_{H^0VV} = g^2_{\phi^0VV} \ . \tag{2.52}$$

The SM and MSSM Higgs coupling relations expressed in terms of the angles α and β are given by

$$\begin{aligned} \frac{g_{h^0VV}}{g_{\phi^0VV}} &= \sin(\beta-\alpha) \\ \frac{g_{H^0VV}}{g_{\phi^0VV}} &= \cos(\beta-\alpha) \ , \end{aligned} \tag{2.53}$$

which satisfy (2.52). The Feynman diagrams of W^+W^- and Z^0Z^0 fusion processes are given in Fig. 2.4.

The remaining interactions for Higgs production near the Z^0 resonance are:

- The Bjorken process $Z^{0*} \to Z^0h^0$ or $Z^{0*} \to Z^0H^0$.
- Neutral pair production $Z^{0*} \to h^0A^0$ or $Z^{0*} \to H^0A^0$.
- Charged pair production $Z^{0*} \to H^+H^-$.

The corresponding Feynman diagrams are given in Fig. 2.5.

The Higgs couplings g_{ZhZ} and g_{ZhA} are complementary functions of the mixing angle α and $\tan\beta$:

$$g_{ZhZ} \propto \sin(\beta-\alpha) \ , \qquad g_{ZhA} \propto \cos(\beta-\alpha) \ . \tag{2.54}$$

Chapter 4 will give more details on h^0A^0 and h^0Z^0 interactions.

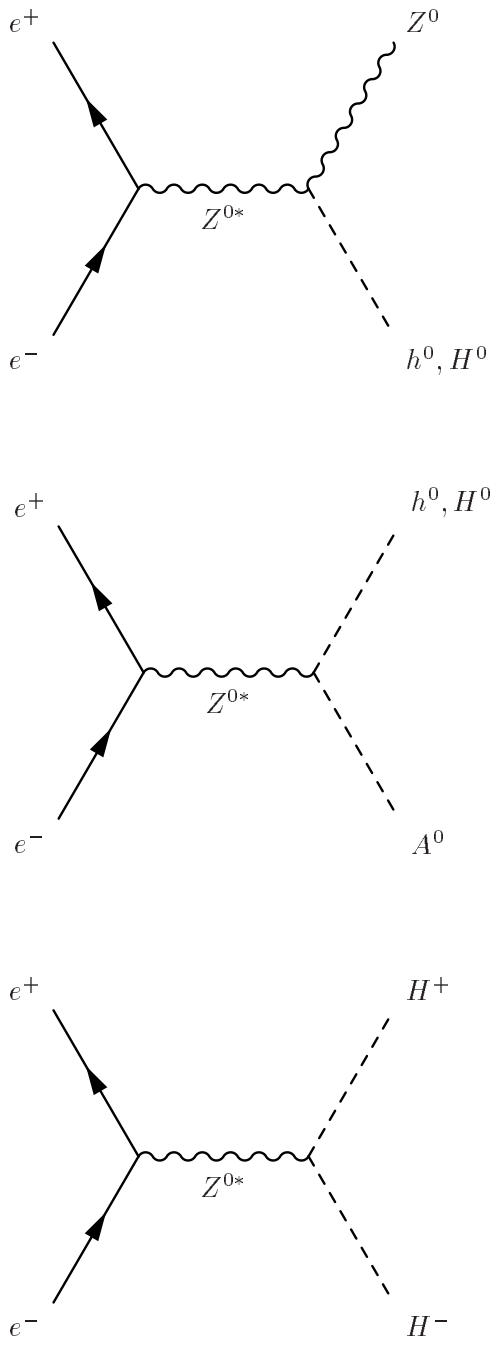

Figure 2.5: Feynman diagrams of Higgs production in the MSSM model. The top diagram is Higgs bremsstrahlung, the middle diagram is neutral Higgs pair-production, and the bottom diagram is charged Higgs pair-production.

Chapter 3

The Detector

CERN is a laboratory dedicated to the experimental and theoretical study of particle physics. The name CERN is an acronym for Centre Européen pour la Recherche Nucléaire.

CERN is located at the border between France and Switzerland with the Jura mountains to the west and Geneva to the east. CERN was established in 1953 to make new discoveries both in experimental and theoretical particle physics, in addition to help foster European cooperation after the war. At present CERN has twenty member countries. One of the most recent successes of CERN came in 1983 when Carlo Rubbia and Simon van der Meer were awarded the Nobel Prize in physics for leading the UA1 experiment which made the first discovery of the Z^0 and $W^{\pm}$ particles.

LEP is an acronym for Large Electron Positron collider, and it is the biggest of the accelerators at CERN. The LEP collider is situated inside a circular tunnel 50 to 100 m below ground surface. The tunnel is 3.5 m in diameter with a circumference of 26.7 km, and is one of the biggest accelerators yet built. The planning of the collider and the subsequent detectors started in 1976, and the construction work began in 1982. The first collisions were performed August 13, 1989 at a center-of-mass energy of 87 Gev, and the LEP accelerator has been running with great success up to now, reaching an energy of 208.8 GeV. The ability to probe the constituents of matter at a small scale and to produce heavy particles are closely related to the available accelerator energy via the famous Einstein equation $E = mc^2$.

LEP accelerates counter-rotating beams of bunches of e^+ and e^- until the maximum energy is reached. The beams are then brought into head-on collisions at four interaction points in the center of the four LEP detectors. The e^+ and e^- annihilate to produce a on- or off-shell Z^0, depending on the energy of the colliding beams. The Z^0 bosons decay with an average lifetime of $(2.65 \pm 0.01) \times 10^{-25}$ s [2], and the outmoving decay products are studied.

Year 2000 is the last year of data-taking with the LEP collider, and the dismantling of LEP is planned to start in the autumn 2000. LEP will then be replaced by the LHC (Large Hadron Collider) accelerator which is planned to start running at a center-of-mass energy of 14 TeV in 2005.

The main contribution of the LEP collider has been precision measurements

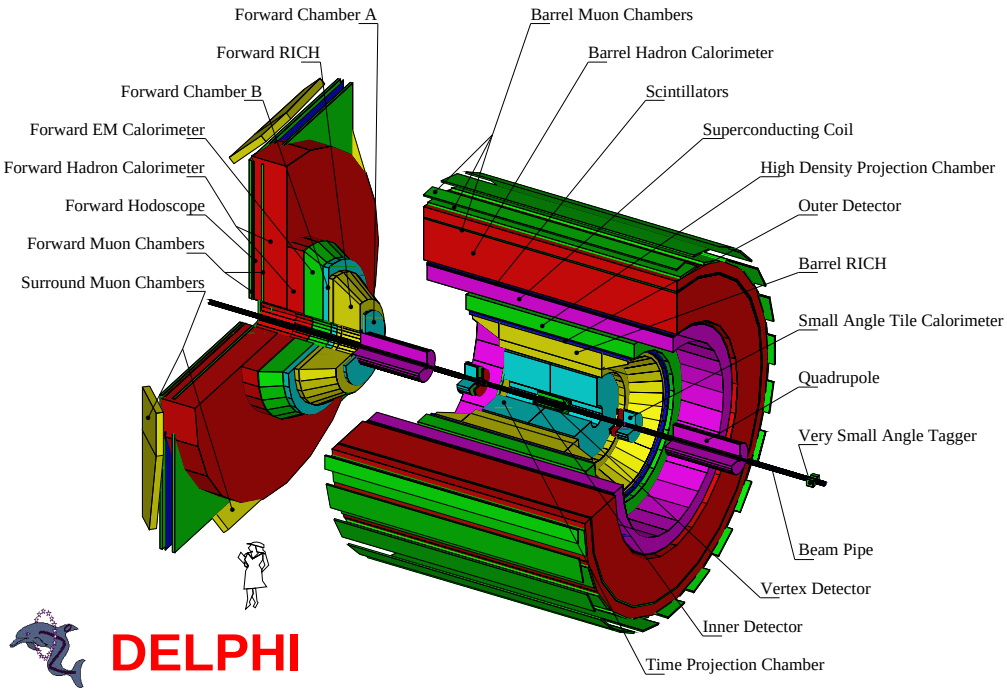

Figure 3.1: The DELPHI detector with all its subdetectors.

of the Z^0 and $W^{\pm}$ masses, and other tests of the SM, both in the electroweak and QCD sectors. The particle physics groups of the Oslo and Bergen universities made large contributions to the project measuring the number of fermion families using the now replaced SAT-detector of the DELPHI experiment.

Information on the DELPHI detector presented in this chapter is collected from DELPHI www-sites [10].

3.1 DELPHI

DELPHI is an acronym for DEtector with Lepton, Photon and Hadron Identification. DELPHI is one of the four detectors at LEP, the others are L3, OPAL and ALEPH. Fig. 3.1 gives a layout of the DELPHI detector and all its subdetectors.

The main part of the detector, the central cylindrical section, is often referred to as the barrel. The barrel are closed with to endcaps, referred to as the forward section. These endcaps can be removed during maintenance and service. The overall length and diameter of the detector are over 10 m, and the total weight is 3500 tons.

A huge superconducting solenoid, the magenta cylinder in Fig. 3.1, produces a magnetic field of 1.23 Tesla that bends the trajectory of the charged particles. The radius of the trajectory is proportional to the momentum of the particle.

The detector is made up of minor subdetectors specialized to extract as much information as possible about the particle properties and trajectories in

the detector. The detector has a layer structure of subdetectors similar to the layers of an onion. The subdetectors are classified according to what kind of information they provide:

- **Tracking Detectors**
  These detectors are placed close to the interaction point, and measures the particle path.

  - **Vertex Detector (VD)** Placed closest to the interaction point. It is a silicon detector, and the barrel part is composed of three layers of silicon strip plaquettes located at 66 mm, 92 mm, and 106 mm from the interaction point. The VD provides precise tracking information in order to detect very short lived particles, and covers a polar angle of about 20 degrees. The tracks are extrapolated backwards to the interaction point.

  - **Inner Detector (ID)** Located between the vertex detector and the time projection chamber. Provides intermediate positions of high precision in addition to trigger information. It consists of the JET chamber and the Trigger Layers (TL). A JET chamber is a driftchamber divided into sectors filled with sense wires that measures the drifttime. The trigger layers consist of cylindrical layers of straw tubes. Angular coverage down to 15 degrees. The resolution of the wires in the JET chamber is 90 μm, and after calibration the track extrapolation (TE) resolution is 40 μm in $r\phi$ and about 1.2 mrad in ϕ. The $r\phi$ measurement is used in the TL to provide an $r\phi$ TE resolution of 150 μm.

  - **Time Projection Chamber (TPC)** The principal tracking device of DELPHI. It also assists in identifying charged particles by measuring dE/dX (energy loss per unit length). Some trigger information are also given. It is a 2 x 1.3 m cylinder situated between the radii 0.29 m and 1.22 m. A charged particle passing through the TPC ionizes the gas, and an electric field causes the electrons of the ionized gas to drift toward one of the proportional chambers inside. The electric field originates from a plate separating the two drift volumes. The TPC provides information on the particle trajectory at radii from 40 to 110 cm between polar angles from 39 to 141 degrees. Precision of $r\phi$ is 250 μm per point, and the z precision is 900 μm per point.

  - **Outer Detector (OD)** A narrow cylinder placed outside the Barrel Ring Imaging Cherenkov detector (Barrel RICH). It is made of five layers of drift tubes located between the radii 197 and 206 cm. The active length corresponds to polar angles from 42 to 138 degrees. It provides final precise measurements of the momenta of charged

particles after the Barrel RICH. Precision of $r\phi$ is 100 μm per track, and the z precision is 4.4 cm.

- **Forward Chamber A (FCA)** Covering polar angles from 11 to 32 and 148 to 169 degrees. It is situated 160 cm from the interaction point and covers the ends of the TPC.

- **Forward Chamber B (FCB)** It is a drift chamber consisting of two independent modules at each endcap at an average distance of 275 cm from the interaction point. The sensitive area corresponds to polar angles from 11 to 36 and 148 to 169 degrees. Each hit is measured twice, first with an accurancy of 2ns, and second with an accurancy of 8 ns.

- **Very Forward Tracker (VFT)** Located on both sides of the vertex detector. It covers polar angles from 19 to 25 and 155 to 170 degrees. The VFT is the forward part of the Silicon Tracker (Vertex Detector).

- **Muon Chambers (MUC)** The DELPHI detector contains three muon chambers: Barrel Muon Chambers (MUB), Forward Muon Chambers (MUF) and Surrounding Muon Chambers (SMC). They are located furthest away from the interaction point because muons are the only charged particles that traverse the lead and iron of both calorimeters essentially unaffected: Most muons of momenta above 2 GeV are expected to reach the muon chambers, whereas the other charged particles are stopped at an earlier point in their trajectory. The muon identification is achieved by comparing extrapolations of reconstructed tracks, provided by subdetectors closer to the interaction point, and hits in the Barrel and Forward muon drift chambers. The SMC was installed in 1994 to fill the gap between the MUB and the MUF. Design accuracy for the MUB is 1 mm in $r\phi$ and 10 mm in z, and the accuracy when space points are associated to extrapolated tracks is 2 mm in $r\phi$ and 80 mm in z. The MUF accuracy on the x and y coordinates is about 5 mm. The muon chambers are the outer green layers in Fig. 3.1.

Having passed the tracking detectors, the particle has now traversed 5 m of the detector.

- **Electromagnetic Calorimeters and Scintillator Counters**
Electron and photon identification are provided primarily by the electromagnetic calorimetry system. The system is composed of a barrel calorimeter (HPC), a forward calorimeter (FEMC), and two very forward calorimeters; the Small angle TIle Calorimeter (STIC), and the Very Small Angle Tagger (VSAT). The latter two are mainly used for luminosity measurements. The STIC replaced the Small Angle Tagger (SAT) in 1994.

- **High-density Projection Chamber (HPC)** The barrel electromagnetic calorimeter. Installed as a cylindrical layer outside the OD. It is mounted on the inside of the solenoid, and consists of 144 independent modules, arranged in 6 rings of 24 modules each. Each module is a trapezoidal box filled with 41 layers of lead separated by gas gaps. An electromagnetic particle produce a shower in the lead and ionizes the gas. The electrons drift to one end of the box where they are collected by a proportional chamber, as in the TPC. The HPC is a cylinder of 2 x 254 cm situated between the radii 208 and 260 cm. The polar angle coverage is 43 to 137 degrees. Granularity is 1 degree in ϕ, 4 mm in z and 9 samples in r. The HPC in Fig. 3.1 is the green cylinder inside the superconducting coil.

- **Forward ElectroMagnetic Calorimeter (FEMC)** Electromagnetic calorimeter in the forward region of the detector. It consists of two disks with diameter of 5 m, and is made of lead-glass. The front faces are placed at a distance of 284 cm from the interaction point, covering polar angles from 8 to 35 and 145 to 172 degrees. For neutral showers of energy above 2 GeV the average precision on the reconstructed hit position in x and y, projected to $|z| = 284$ cm, is about 0.5 cm. Fig. 3.1 shows the FEMC as the green disc outside the forward HACL.

- **Scintillators** In order to achieve maximal coverage for high energy photons under all emission angles with minimal leakage, a so-called hermeticity counter, scintillators have been installed between the barrel and endcap, and in the HPC. In addition, the Time Of Flight (TOF) is situated in the barrel, and the HOrizontal Flight (HOF) tagger in the forward section. The scintillators are also used as fast triggers. The efficiency of the HOF is about 80 %, the rate is 0.1-0.4 Hz, and the response-time is less than 50 ns.

- **Hadron Calorimeter (HACL)**
  Measures the kinetic energy of neutral and charged strong interacting particles (hadrons). Situated between the superconducting coil and the two outermost muon chambers in the barrel, and between the forward EM calorimeter and the outermost muon chamber in the forward section. Both the barrel and forward calorimeter contain a muon chamber. The hadron calorimeters are indicated by red in Fig. 3.1. It is a sampling gas detector incorporated in the magnet yoke consisting mainly of iron. The barrel part covers polar angles from 42.6 to 137.4 degrees, and the two endcaps from 11.2 to 48.5 and 131.5 to 168.8 degrees.

- **Charged Hadron Identification**
  Particle identification of strongly interacting particles. In DELPHI it

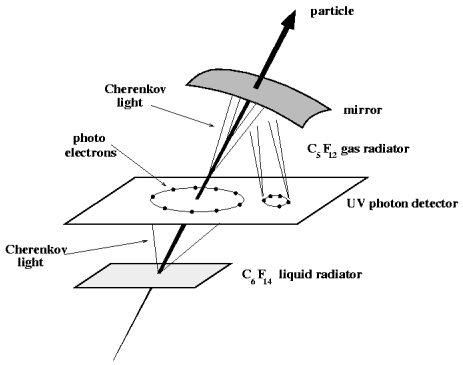

Figure 3.2: Figure showing the structure of the RICH detectors of DELPHI.

relies on the dE/dX measurement of the TPC, and on the velocity measurement of the RICH detectors. The RICH technique is based on the detection of Cherenkov light emitted by the moving particle. The DELPHI RICH contains two radiators: A liquid radiator operating in the momentum range from 0.7 to 9 GeV, and a gas radiator used from 2.5 to 25 GeV. Full solid angular coverage is provided by a Forward RICH and a Barrel RICH, both indicated by yellow in Fig. 3.1. Fig. 3.2 shows the structure of the RICH detectors of DELPHI. The radius of the ring gives the velocity of the moving particle. The photon detector is a gas drift chamber with multi wire proportional chambers (MWPC) to detect the drift electrons.

- **Barrel RICH** Located between the TPC and the OD. It is a 350 cm long cylinder with inner radius 123 cm and outer radius 197 cm. It covers polar angles between 40 and 140 degrees.

- **Forward RICH** Covers polar angles between 15 and 35 degrees. It is about 1 m thick, and is located 1.7 m away from the interaction point. The most important difference between the Barrel and Forward RICH is the presence of the magnetic field of 1.23 Tesla in the Forward RICH.

- **Luminosity Measurement**
  The luminosity at e^+e^- detectors is measured by counting the number of events with a clear experimental signature. Events with high statistics and a cross section that can be theoretically calculated to high precision are used. The process chosen is Bhabha scattering ($e^-e^+ \rightarrow e^-e^+$) at small angles, which proceeds almost entirely through the exchange of a photon in the t-channel. In DELPHI the absolute luminosity is measured using the Small angle TIle Calorimeter (STIC) and the Very Small Angle Tagger (VSAT).

  - **Small Angle Tile Calorimeter (STIC)** It is a sampling lead-

scintillator calorimeter formed by two cylindrical detectors placed on either side of the DELPHI interaction region at a distance of 220 cm, and covers polar angles between 29 and 185 mrad (from 6.5 to 42 cm in radius). The lead-scintillator calorimeter are made up of 47 lead-scintillator layers, each with 1600 holes to let the wavelength shifter fibers go through. The energy resolution at 45 GeV is 3 %, and the spatial resolution of the calorimeter alone is 1.5 degrees in ϕ and 300 μm to 1 mm in radius. The resolution of the silicon is 400 μm in radius, and the expected systematic error on luminosities is 0.2 %.

- **Very Small Angle Tagger (VSAT)** Consists of 4 calorimeter modules, each made of 11 silicon diodes separated by tungsten alloy absorber. Three silicon strip planes, placed at 5, 7 and 9 radiation lengths into the modules, are used for (x, y) shower position measurements. A lead block of 10 radiation lengths are placed at the back of each module to decrease the number of parasite events (photons or off-momentum leptons) from the interaction point. The VSAT detects electrons and positrons coming from Bhabha scattering, and photons between 5 and 7 mrad. The energy resolution at 45 GeV is 5 %, at 95 Gev around 4 %. The resolution of the silicon strips is about 170 μm of the x and y coordinates. Expected systematic error is 1 %. The energy resolution at 45 GeV is 5 %, and at 95 GeV around 4 %. The resolution of the silicon strips is about 170 μm.

Chapter 4

The h^0A^0 Analysis

The Higgs analysis presented is performed on the $Z^0 \to h^0A^0 \to b\bar{b}\,b\bar{b}$ channel of the MSSM Higgs sector. The Feynman diagram of the h^0A^0 channel is given in Fig. 4.1.

$Z^0 \to h^0A^0$ dominates over $Z^0 \to h^0Z^0$ at high $\tan\beta$, and $h^0A^0 \to b\bar{b}b\bar{b}$ is the dominant decay channel. A full scan of the $(m_A, \tan\beta)$ plane is not performed, only five m_A and two $\tan\beta$ values have been used in the limit calculations. This limited scan is performed because the objective of this analysis is to compare likelihood methods using different binning procedures and discriminating variables, rather than determine an excluded region of the $(m_A, \tan\beta)$ plane. The likelihood methods are presented in Chapter 7 and Chapter 8.

The decay modes of the h^0A^0 channel are presented in Sec. 4.2, the relevant background processes in Sec. 4.3, and the signals in Sec. 4.4.

4.1 Higgs Production

At LEP2 the neutral MSSM Higgs bosons h^0 and A^0 are produced mainly through the following two processes:

$$\begin{array}{ll} \text{Higgs} - \text{strahlung} & : e^+e^- \to h^0Z^0 \\ \text{Associated pair production} & : e^+e^- \to h^0A^0 \end{array} \qquad (4.1)$$

The Feynman diagrams of these processes are given in Fig. 2.5. The fusion processes (Fig. 2.4) play a minor role at the kinematical limit of Higgs-strahlung production of h^0. A^0 is only produced through pair production (to leading order).

The cross sections of the h^0Z^0 and h^0A^0 processes in (4.1), expressed in terms of the cross section σ_{SM} for Higgs-strahlung in the SM (Fig. 2.1), are given by

$$\sigma(e^+e^- \to h^0Z^0) = \sin^2(\beta-\alpha)\sigma_{SM} \qquad (4.2)$$

$$\sigma(e^+e^- \to h^0A^0) = \cos^2(\beta-\alpha)\bar{\lambda}\,\sigma_{SM}\,, \qquad (4.3)$$

where $\bar{\lambda}$ is the momentum factor of the two particle phase space. The cross sections for Higgs-strahlung, h^0Z^0, and pair production, h^0A^0, for $\tan\beta = 2.0$

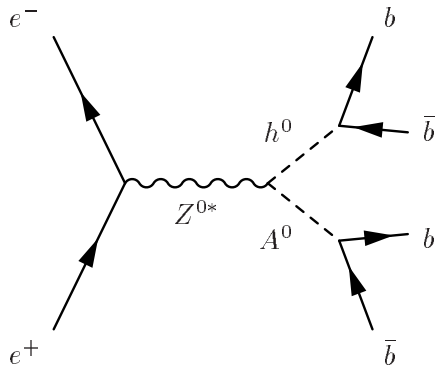

Figure 4.1: Feynman diagram of the MSSM process $e^+e^- \to Z^{0*} \to h^0A^0 \to b\bar{b}b\bar{b}$.

and $\tan\beta = 20.0$, are given in Fig. 4.2 [11], illustrating that pair production dominates over Higgs-strahlung at high $\tan\beta$.

4.2 Decay Modes

The CP-even scalar Higgs boson h^0 decays almost exclusively into a fermion-antifermion pair if $\tan\beta > 1$ and $m_h < 100$ GeV, which is the case for all the signal hypotheses of this analysis. Fermion decays are also the dominant decay mode of the CP-odd pseudoscalar Higgs boson A^0. The partial decay widths Γ of all the neutral Higgs bosons Φ into fermions f in the limit $m_\Phi^2 \gg m_f^2$ are given by [9]

$$\Gamma(\Phi \to f\bar{f}) = N_c \frac{G_F m_f^2}{4\sqrt{2\pi}} g_{\Phi ff}^2 m_\Phi \left[1 + \frac{17}{3}\frac{\alpha_s}{\pi}\right] , \tag{4.4}$$

where G_F (the Fermi constant) is the weak coupling constant, $g_{\Phi ff}$ are the Higgs-fermion couplings, and α_s is the QCD running coupling constant.

Fig. 4.3 [11] shows the branching ratios of h^0 and A^0 into SM particles, and it is evident that the dominant decay mode at both $\tan\beta = 2$ and $\tan\beta = 20$ is $b\bar{b}$.

4.3 Background

The three most important background processes to a Higgs search in the h^0A^0 channel are: $Z^{0*} \to W^+W^- \to q\bar{q}q\bar{q}$, $Z^0\gamma \to q\bar{q}\gamma$ and $Z^0Z^0 \to q\bar{q}q\bar{q}(q\bar{q}l\bar{l})$. The dominating process is the $Z^0\gamma$ process, followed by the W^+W^- process, and a small contribution from the Z^0Z^0 process. Fig. 4.5 [12] shows the background level vs. the efficiency for the different Monte Carlo generated backgrounds, together with the data of the 189 GeV run.

Channel		Events after cut	σ [pb]	Efficiency [%]	Events
$q\bar{q}(\gamma)$		2040	99.0	$(13.9 \pm 0.31) \cdot 10^{-2}$	13.4 ± 0.83
$q\bar{q}q\bar{q}$		1144	17.733	$(18.3 \pm 0.54) \cdot 10^{-2}$	9.45 ± 0.35
$l\bar{l}q\bar{q},\ l = \mu, \tau$		146	0.3688	$(69.4 \pm 5.7) \cdot 10^{-2}$	0.417 ± 0.036
Background		3330	—	—	23.2 ± 0.90
Data		33	—	—	33
Signal, $\tan\beta = 2$					
m_A [GeV]	70	1404	0.08329	70.2 ± 1.0	9.24 ± 0.16
	75	3735	0.06482	72.0 ± 0.62	7.37 ± 0.098
	80	3753	0.04890	75.2 ± 0.61	5.81 ± 0.075
	85	3900	0.03549	75.0 ± 0.60	4.21 ± 0.054
	90	2083	0.02457	74.4 ± 0.82	2.89 ± 0.043
Signal, $\tan\beta = 20$					
m_A [GeV]	70	1508	0.15892	75.4 ± 0.96	18.9 ± 0.31
	75	3779	0.11352	75.6 ± 0.60	13.6 ± 0.17
	80	3808	0.07287	76.2 ± 0.60	8.77 ± 0.11
	85	3551	0.03844	74.0 ± 0.63	4.50 ± 0.59
	90	2094	0.01270	69.8 ± 0.84	1.40 ± 0.22

Table 4.1: The final cut number of events, cross-sections, efficiencies and expected number of events for the different Monte Carlo generated backgrounds and signals, together with the 189 GeV data.

4.4 Signal

The h^0A^0 search presented is restricted to limit calculations of only ten points in the $(m_A, \tan\beta)$ plane: Five m_A values (70, 75, 80, 85, 90 GeV) at two $\tan\beta$ values (2, 20). The small number of $(m_A, \tan\beta)$ points is allowed because the objective of this search is to compare likelihood methods using different binning algorithms and discriminating variables, rather than exclude a region of the $(m_A, \tan\beta)$ plane. No interpolation between the Monte Carlo generated signal distributions is performed (see Sec. 6.3 and Fig. 6.2).

The signal, background and 189 GeV data distributions are provided by Jørgen Hansen, using the DELPHI TEAM C standard track [13, 14] and 4jet selection [15, 14]. The final cut is set using a neural network-like variable (presented in ref. [14]), and the cut corresponds to 33 observed candidates. Figures 4.6, 4.7, 4.8, 4.9 and 4.10 shows the the $m_A = 80$ GeV, $\tan\beta = 20$ signal, together with the different Monte Carlo backgrounds and the observed candidates at 189 GeV.

Table 4.1 gives the number of events after the final cut, cross sections, efficiencies, and expected number of events of the Monte Carlo generated backgrounds and signals, together with the 189 GeV data. The quoted uncertainties include a common 1 % relative error on both signals and backgrounds, to account for uncertainties in center-of-mass energy, luminosity, and cross section,

in addition to a systematic error on the difference between different generators of the various background channels and the modeling of the successive cuts, estimated to be 5.7 % for the PYTHIA background, and 2 % for the EXCALIBUR backgrounds [14, 16].

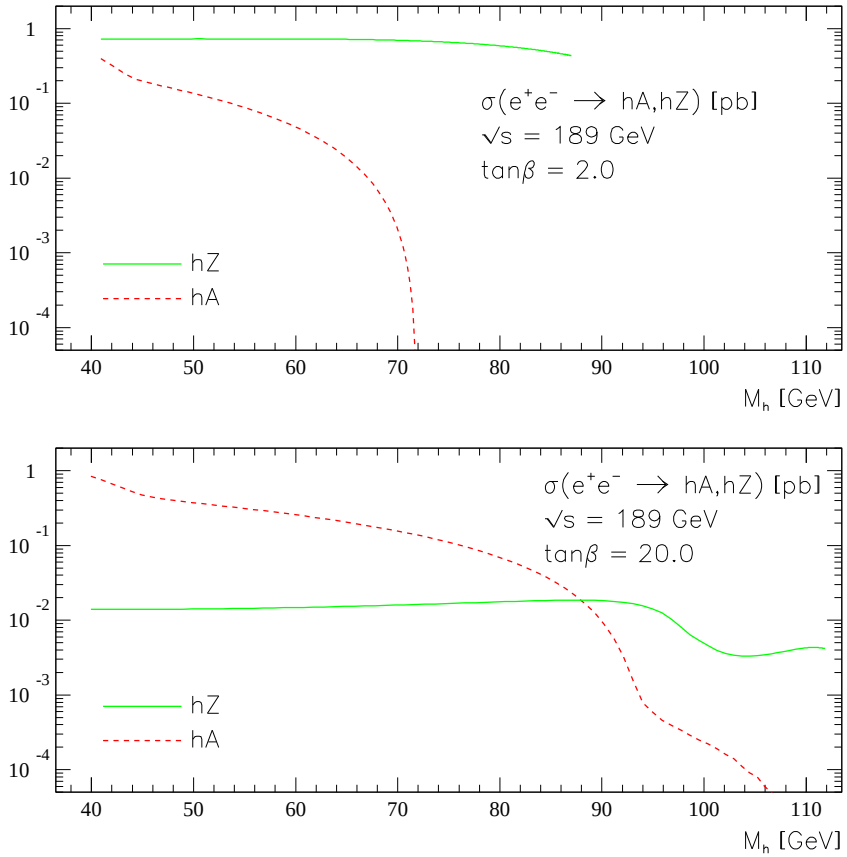

Figure 4.2: Cross sections of MSSM Higgs-strahlung, h^0Z^0, and pair production, h^0A^0, for $\tan\beta = 2.0$ and $\tan\beta = 20.0$ [11].

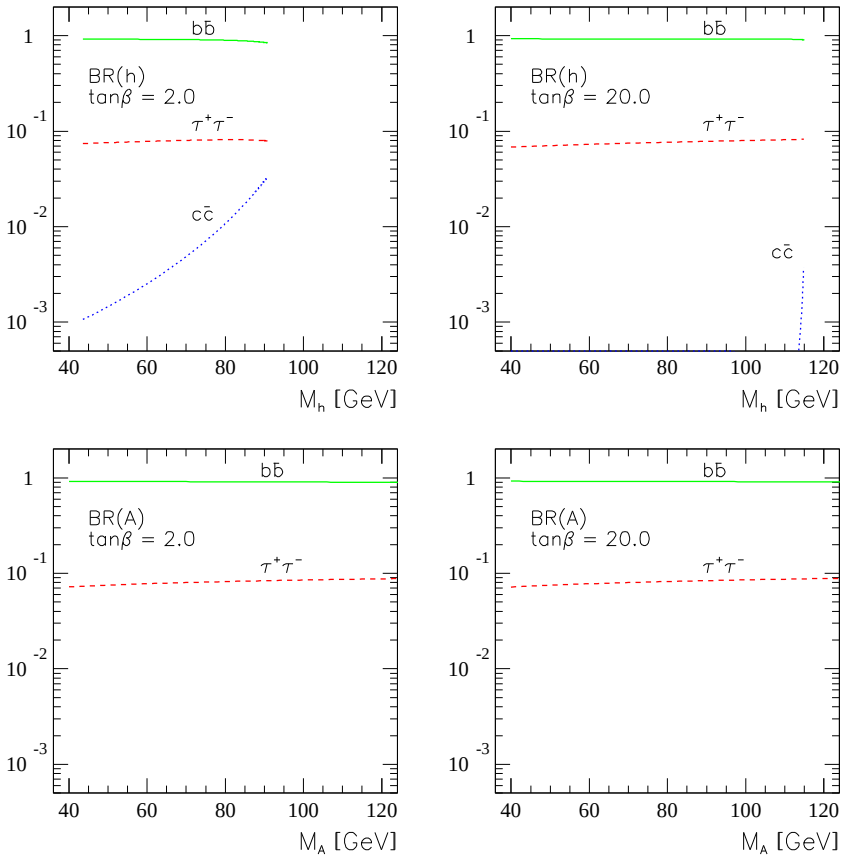

Figure 4.3: Branching ratios of the MSSM Higgs bosons h^0 and A^0 into Standard Model particles [11].

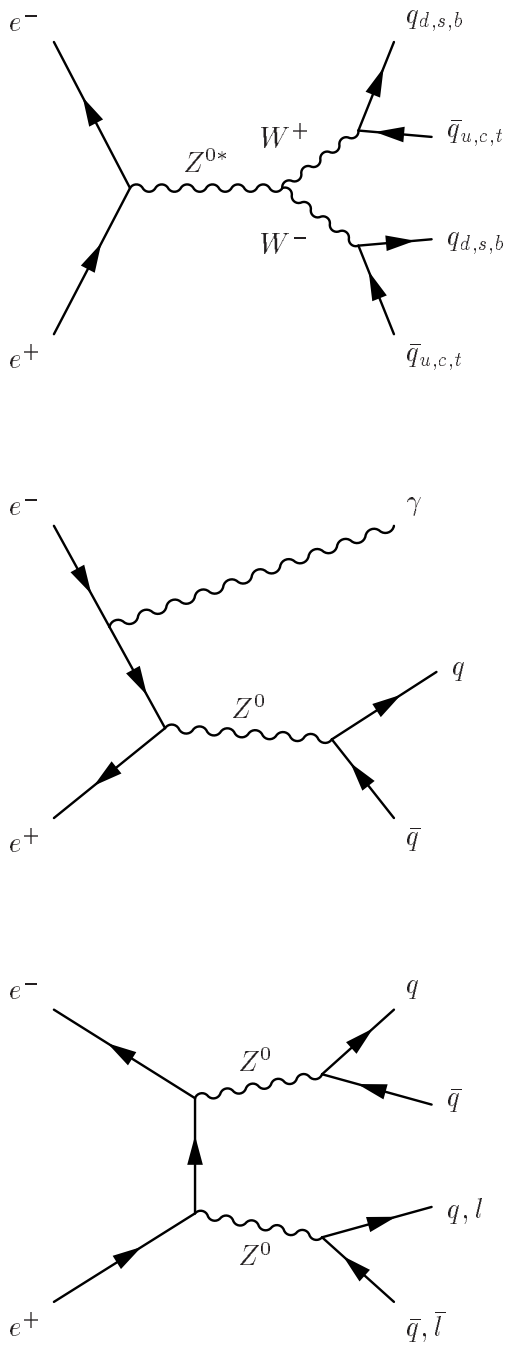

Figure 4.4: Feynman diagrams of the three most important background processes to a search in the h^0A^0 channel: $Z^{0*} \to W^+W^- \to q\bar{q}q\bar{q}$, $Z^{0*}\gamma \to q\bar{q}\gamma$ and $Z^0Z^0 \to q\bar{q}q\bar{q}(q\bar{q}l\bar{l})$.

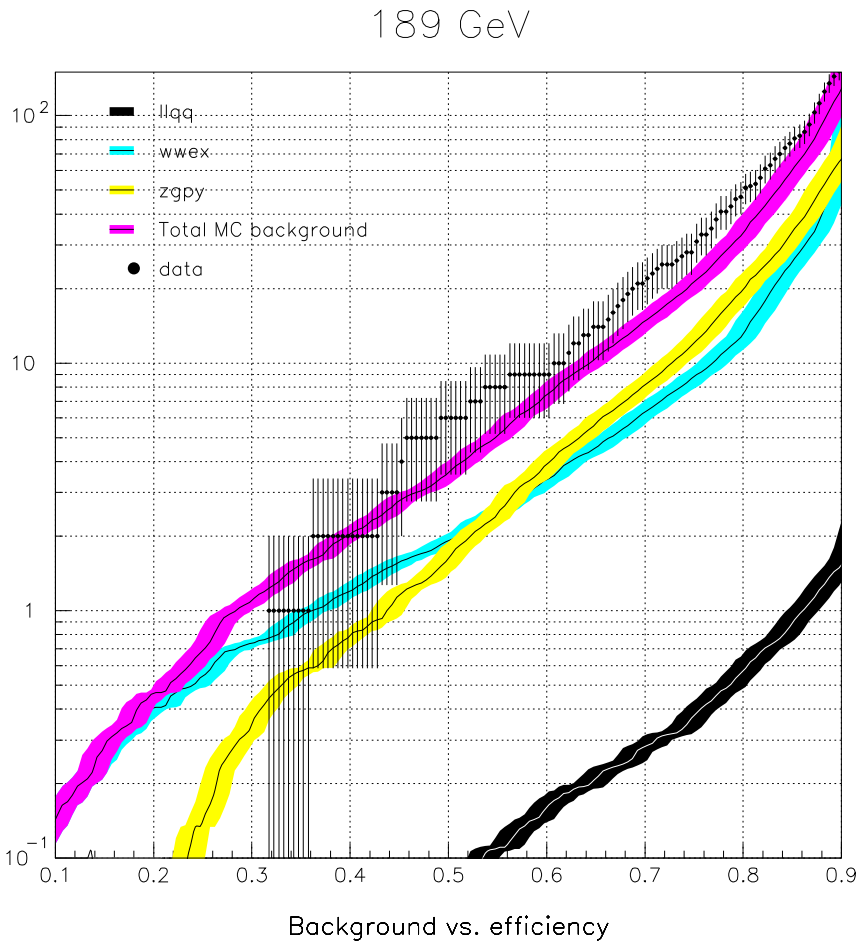

Figure 4.5: Figure showing the background level vs. the efficiency for the different Monte Carlo generated background samples, together with the data of the 189 GeV run [12].

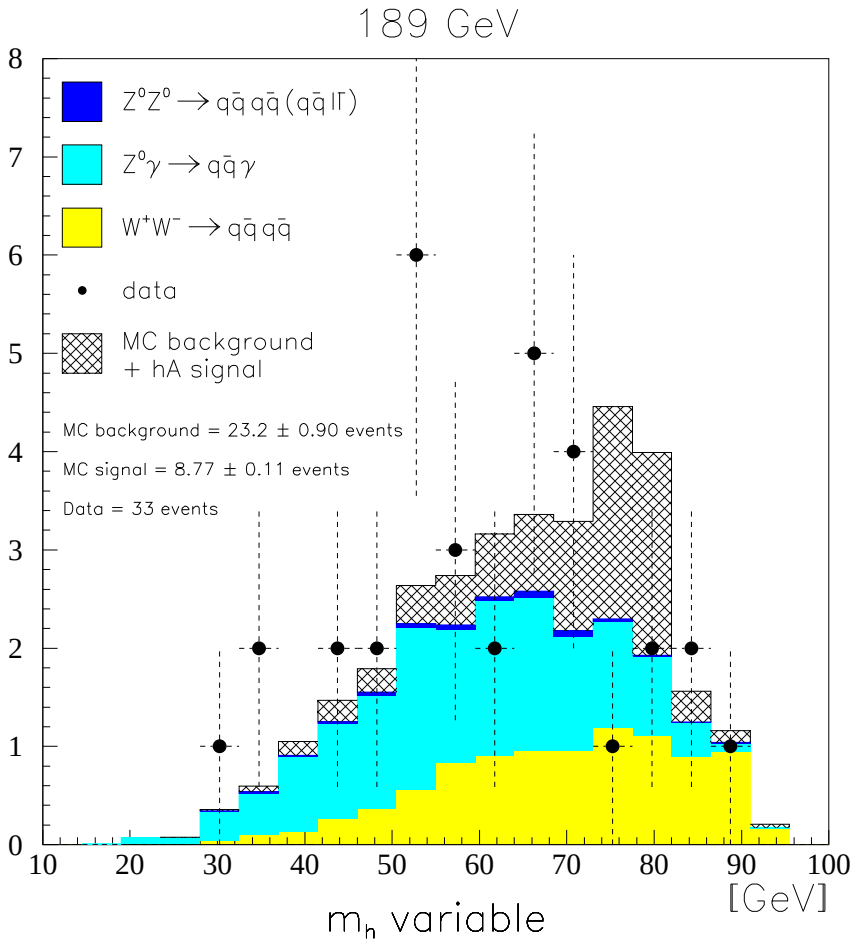

Figure 4.6: Figure showing the lightest of the paired jet masses, m_h. The signal sample used is the Monte Carlo generated signal for $m_A = 80$ GeV, $\tan\beta = 20$.

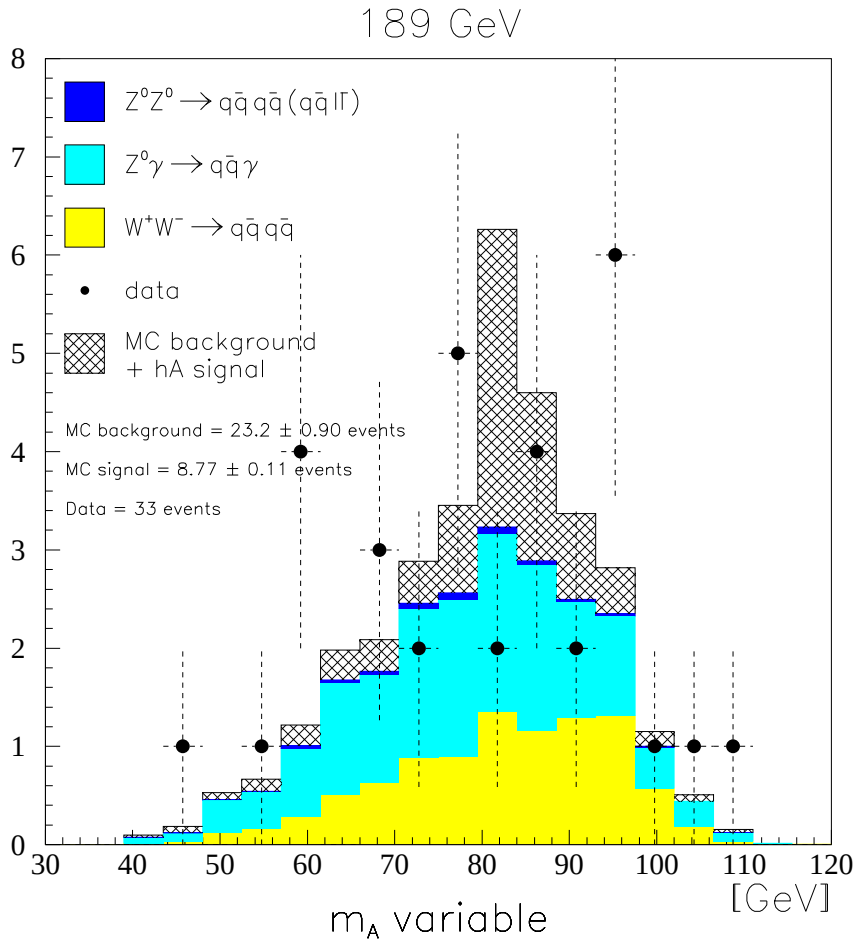

Figure 4.7: Figure showing the heaviest of the paired jet masses, m_A. The signal sample used is the Monte Carlo generated signal for $m_A = 80$ GeV, $\tan\beta = 20$.

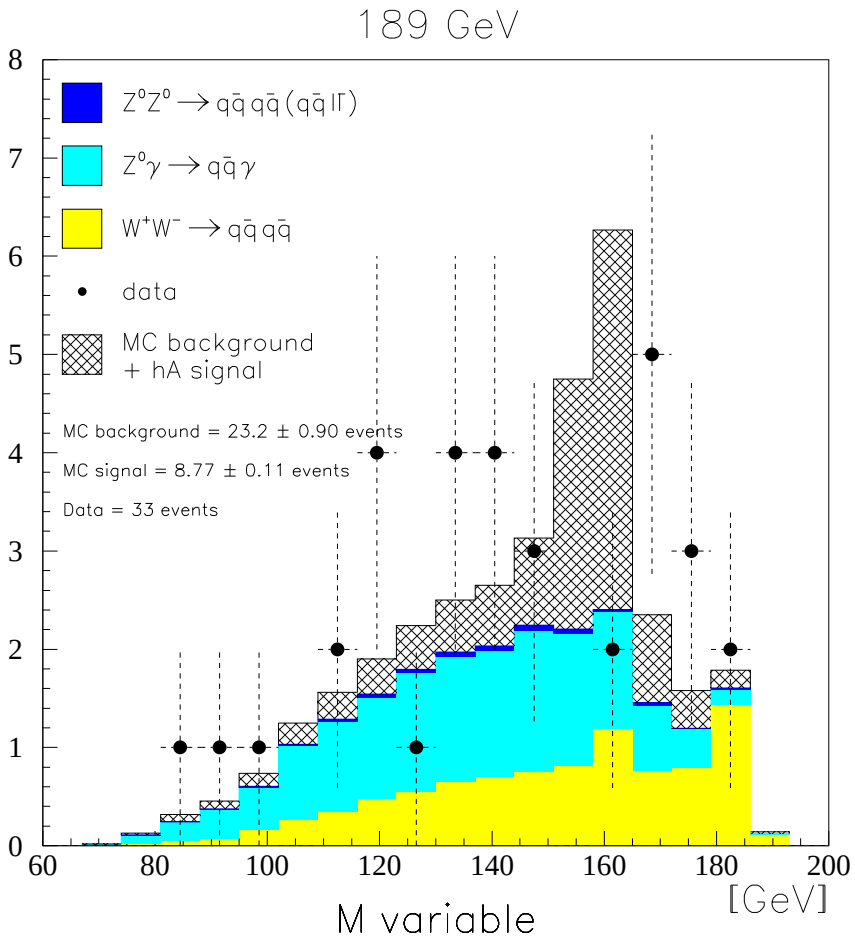

Figure 4.8: Figure showing the sum of the paired jet masses, $M = m_h + m_A$. The signal sample used is the Monte Carlo generated signal for $m_A = 80$ GeV, $\tan\beta = 20$.

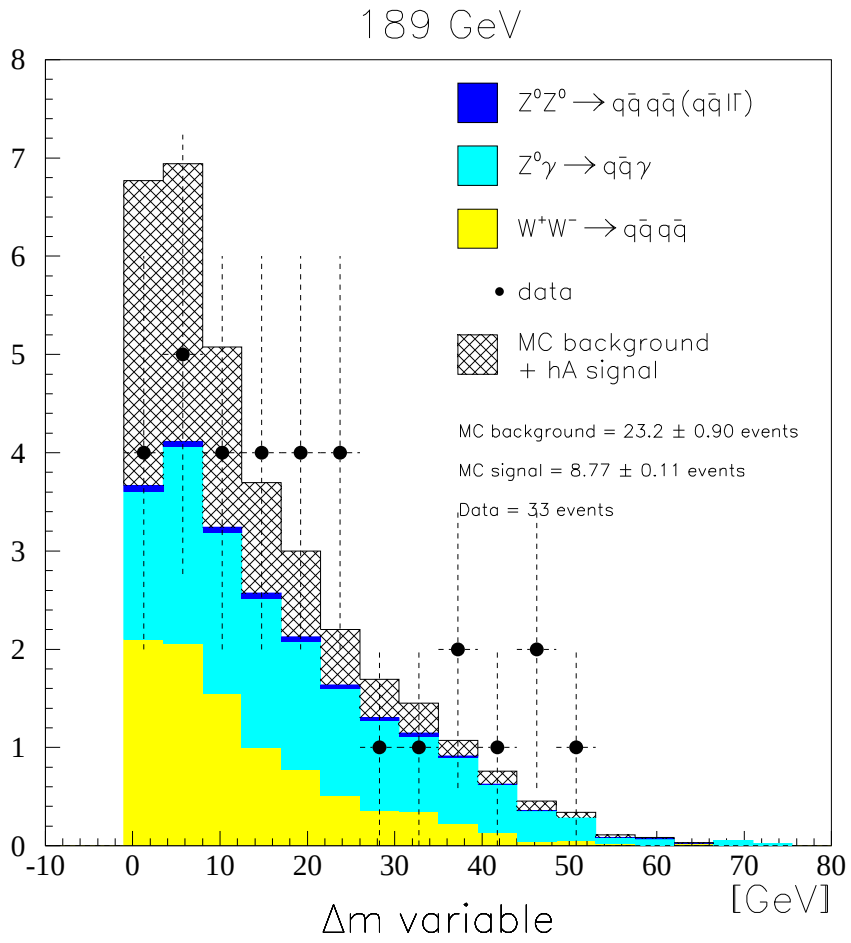

Figure 4.9: Figure showing the difference of the paired jet masses, $\Delta m = m_A - m_h$. The signal sample used is the Monte Carlo generated signal for $m_A = 80$ GeV, $\tan\beta = 20$.

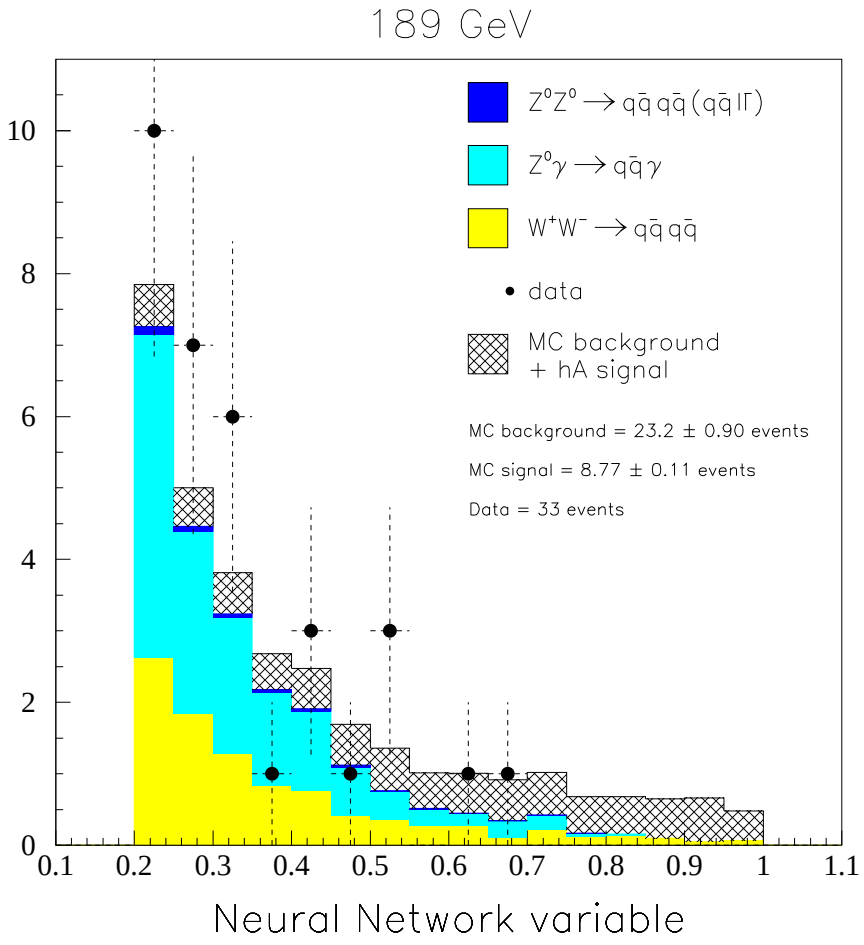

Figure 4.10: Figure showing the neural network-like variable presented in ref. [14] after the final cut. Strictly speaking the variable of ref. [14] is not a neural network variable, but because of many common features it is referred to as such (see Sec. 7.1). The signal sample used is the Monte Carlo generated signal for $m_A = 80$ GeV, $\tan\beta = 20$.

Chapter 5

Jet Clustering and Kinematical Fits

The single final-state particles of a high energy interaction event have to be grouped together in clusters in order to reconstruct the jet topology of the event, and calculate the jet momentum and energy. The jet clustering of the signals, backgrounds and data candidates used in the analysis of the likelihood methods presented in Chapter 7 and Chapter 8, is the standard DELPHI Durham [17] algorithm programmed by Jørgen Hansen. The presentation of the Durham jet clustering is based on the PYTHIA 5.7 manual by T. Sjöstrand [18].

To correct the event data for detector inefficiencies, a kinematical fit on the single particle tracks of the jets is performed. The effect is sharper mass distributions, as is seen in Fig. 5.1. Sec. 5.2 discuss kinematical fits in more detail. The discussion on kinematical fits are influenced by the book " Statistics for Nuclear and Particle Physicists" by Louis Lyons [19], and private communications with Jørgen Hansen.

5.1 Jet Clustering

The single hadronic tracks of an event have to be combined into clusters of tracks in order to calculate the momentum and energy of the jets.

To determine the individual jet axes in events of more than three jets, a technique known as *cluster finding* is developed: Each final-state particle is initially considered to be a cluster. The two nearest clusters are found using some distance measure. If the distance is smaller than a cut-off value, the clusters are joined to form a new cluster. This routine is repeated until the distance between any two clusters are above the cut-off value. The cut-off value controls the final jet picture, and it is possible to have the cluster algorithm find a predetermined number of jets, which is the case for the 4 jet clustering of $h^0A^0 \to b\bar{b}b\bar{b}$ events.

A jet is defined as a collection of particles which have a limited transverse momentum with respect to a common jet axis, and hence also with respect to each other [18]. The distance measure used should only depend on the transverse momentum, and the distance measure of the Durham algorithm is

written as [18]

$$\tilde{y}_{ij} = \frac{2\min(E_i^2, E_j^2)(1-\cos\theta_{ij})}{E_{cm}^2} \ , \tag{5.1}$$

where E_i and E_j are the energy of the two clusters, θ_{ij} is the relative angle between the clusters, and E_{cm} is the center-of-mass energy of the interaction. The physical interpretation of the $\tilde{y}_{ij}$ distance measure is the transverse momentum of the softer particle with respect to the direction of the harder one. In contrast, the `LUCLUS` routine [18] uses a distance measure with the physical interpretation as the transverse momentum of either particle with respect to the common direction given by the momentum vector sum.

Initially, each particle is considered to be a cluster, and the two clusters with smallest relative distance $\tilde{y}_{ij}$ are found and joined to one if $\tilde{y}_{ij} < \tilde{y}_{join}$, with $\tilde{y}_{join}$ some predetermined distance. The momentum of the new cluster is the vector sum of the momenta of the joined clusters. This procedure is repeated until the distance between any two clusters are greater than $\tilde{y}_{join}$. If the number of final clusters do not match a predetermined number of jets, the value of $\tilde{y}_{join}$ is modified, and the clustering algorithm repeated until the final number of clusters match the predetermined jet number.

The main difference between the Durham algorithm and the `LUCLUS` routine, is that Durham does not allow reassignments. Reassignment is performed in the `LUCLUS` routine after each joining because the particles of a new cluster may be closer to another cluster.

5.2 4C Fit

A kinematical fit is performed to make the measured quantities of an observed interaction satisfy a set of kinematical constraints. This is done by considering all configurations of the four-momentum vectors of the outgoing particles that satisfies these constraints. From this infinity of four-momentum vector sets, the set that has the least χ^2 value are used.

A set of n measurements $x_1^m, x_2^m, ..., x_n^m$ with errors $\sigma_1, \sigma_2, ..., \sigma_n$ are provided by the detector. These measurements are subjected to a number of constraints, $C_j(x_1, x_2, ..., x_n)$, $j = 1, ..., n_c$, satisfied by the numbers $\bar{x_1}, \bar{x_2}, ..., \bar{x_n}$. The quadratic sum $\mathcal{S}^2$ is defined as

$$\mathcal{S}^2 = \sum_{i=1}^{n} \left(\frac{x_i^m - \bar{x_i}}{\sigma_i} \right)^2 \ , \tag{5.2}$$

which gives a measurement of how much the x_i set of measurements have to be moved in order to fit the $\bar{x_i}$ set of values.[1] Further, χ^2 is defined as the

[1] The $\mathcal{S}^2$ definition (5.2) corresponds to the χ^2 definition [20]

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right)^2 \ ,$$

provided the measured values x_i^m are all collected from a normal distribution $N(\mu_i{=}\bar{x_i}, \sigma_i^2)$,

minimum of $\mathcal{S}^2$ when varying $\bar{x_i}$

$$\chi^2 = \min(\mathcal{S}^2) \ . \tag{5.3}$$

The problem now is to find the set $\bar{x_1}, \bar{x_2}, ..., \bar{x_n}$, among the infinite number of sets, that minimize $\mathcal{S}^2$.

In the case of a $h^0A^0 \to b\bar{b}b\bar{b}$ interaction the measured values are the momentum and energy of the four jets: $(p_x^i, p_y^i, p_z^i, E^i)$, $i = 1, ..., 4$.

5.2.1 The Constraints

The four constraints of a 4C fit are the constraints leading to conservation of momentum and energy and, in the case of a colliding beam experiment, the constraints are given as:

- $C_1 = \sum_{tracks} p_x = 0$
- $C_2 = \sum_{tracks} p_y = 0$
- $C_3 = \sum_{tracks} p_z = 0$
- $C_4 = \sum_{tracks} E = E_{CM} = \sqrt{s}$

These constraints are not hypothesis dependent; they are well founded physical laws. In contrast, a 5C fit imposes a fifth constraint: The masses of the two pairs of jets, resulting from a pairing of the four clustered jets two by two, are equal. This constraint is hypothesis dependent, and a 5C fit is rarely used in the $h^0A^0 \to b\bar{b}b\bar{b}$ channel.

After the kinematical fit is performed, the problem of which two jets belong to which two initial heavy objects must be addressed. If a 4C fit is used, the pairing of least mass difference is considered to be the best pairing. If a 5C fit is used, the pairing of least χ^2 is considered to be the best pairing. The h^0A^0 events used in the analysis presented in this thesis are subjected to a 4C fit, using the PUFITC package programmed by N. J. Kjaer of CERN and M. Mulders of NIKHEF/DELPHI. The three possible pairings are classified according to their jet mass difference. Chapter 8 provides a more elaborate discussion on the pairing problem.

Fig. 5.1 shows the invariant mass of $h^0A^0 \to b\bar{b}b\bar{b}$ events before and after a 4C fit performed by PUFITC [12]. The effect of the fit is a sharper mass distribution.

---

where μ_i is the mean. However, there is of course no reason to expect the x_i^m to be normal distributed with $\mu_i = \bar{x_i}$ prior to the kinematical fit. Hence, the definition (5.2) is referred to as a quadratic sum, $\mathcal{S}^2$, not as χ^2.

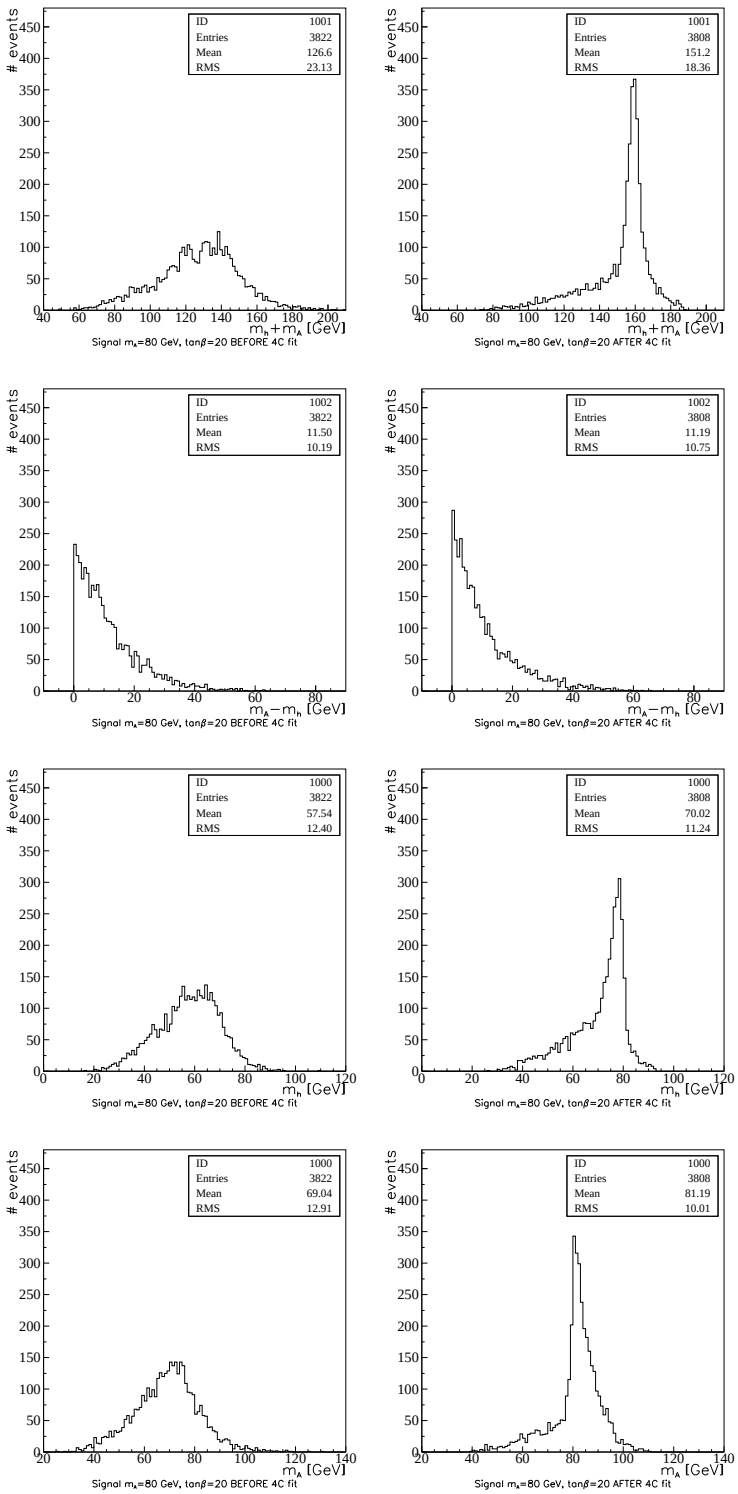

Figure 5.1: Plots of mass estimators of the $h^0A^0 \to b\bar{b}b\bar{b}$ channel before and after a 4C fit performed by the PUFITC package [12]. The plots to the left are the distributions of the mass estimator before the 4C fit, and the plots to the right are the distributions of the mass estimators after the 4C fit. As is evident from these plots, the 4C fit sharpens the mass distributions by reducing the RMS. The signal sample used is $m_A = 80$ GeV, $\tan\beta = 20$ ($m_h = 79.53$ GeV).

Chapter 6

The ALRMC Program

The ALRMC program [21] offers an optimal method for setting exclusion limits on the Higgs mass for multichannel searches using a likelihood ratio technique. The program is used for calculating confidence levels based on signal and background distributions prepared by the likelihood methods presented in Chapter 7 and Chapter 8. The limits are presented in Table 10.1.

A brief summary of hypothesis testing in general, together with the ALRMC specifics, are presented in Sec. 6.1. Sec. 6.2 presents the inputs of ALRMC, and Sec. 6.3 explains the customizations performed to make ALRMC fit the requirements of the analysis.

6.1 Statistics Theory

The discussion presented in this section is inspired by the book "Probability and Statistics in Particle Physics" by Frodesen, Skjeggestad and Tøfte [20], and the Delphi-note "Optimal Statistical Analysis of Search Results based on the Likelihood Ratio and its Application to the Search for the MSM Higgs Boson at $\sqrt{s} = 161$ and 172 GeV" by Alex L. Read [21].

6.1.1 Hypothesis Testing

The true value of a parameter θ is wanted. A number of n random samples, or *observables*, $x_1, x_2, ..., x_n$ are collected from a population described by a probability density function $F(x|\theta)$ which depends on θ. A *test-statistic* $X(x_1, x_2, ..., x_n)$, which does not depend on any unknown parameters, is constructed from the n samples. It is essential that X has some correspondence to θ; X is an *estimator* of θ. A null hypothesis is formulated stating a one- or two-sided limit on the true value of θ. A *confidence level* is interpreted as the probability of finding the true value of θ in the region stated in the null hypothesis. The confidence level is the integral of the probability distribution function of the test-statistic, $P(X)$, between the limits stated in the null hypothesis. A *confidence limit* is the value of a population parameter (such as a particle mass or production rate) which is excluded at a specific confidence level[21]. A confidence level expresses the confidence associated with a hypothesis.

Constructing a search analysis follows three general steps: Define the observables, define a test-statistic, and define rules for exclusion and discovery. The rules lead to ranges of values of the test-statistic, and these ranges are obtained by calculating the integration limits of the probability density function integral that gives the specified confidence level. Typical observables may be the number of candidates satisfying a set of criteria, the reconstructed invariant mass of the candidates, b-quark tagging probabilities, or a discriminant variable constructed from a neural network. The test-statistic is constructed to rank the experiments from the least to most signal-like, and the ALRMC program utilize the *likelihood ratio* test-statistic for parameter estimation.

Two kinds of probability density functions enter a search: The signal distribution of the mass hypothesis, and the background distribution of the known background. A number of candidates, satisfying certain criteria, are identified from the reconstructed data provided by the detector. To correct the candidates for backgrounds, the background rates are subtracted from the candidates. In case of small or absent signal rates, the result of this procedure may lead to unphysical rates. The way to deal with this is to normalize the confidence level observed for the signal+background together hypothesis, CL_{s+b}, to the confidence level observed for the background only hypothesis CL_b;

$$CL_s \equiv \frac{CL_{s+b}}{CL_b} \ . \tag{6.1}$$

The test-statistic X depends on the observables and the population parameters of the known background distribution and the signal hypothesis distribution, and it is constructed to increase monotonically for increasingly signal-like experiments. Hence, the confidence in the signal+background hypothesis is given as the probability that the test-statistic is less than or equal to the observed experimental value, X_{obs}:

$$CL_{s+b} = P_{s+b}(X \leq X_{obs}) \ , \tag{6.2}$$

with

$$P_{s+b}(X \leq X_{obs}) = \int_0^{X_{obs}} \frac{dP_{s+b}}{dX} dX \ , \tag{6.3}$$

and where dP_{s+b}/dX is the probability distribution function of the test-statistic X for the signal+background hypothesis. The confidence in the background only hypothesis is similarly given as

$$CL_b = P_b(X \leq X_{obs}) \ , \tag{6.4}$$

with

$$P_b(X \leq X_{obs}) = \int_0^{X_{obs}} \frac{dP_b}{dX} dX \ , \tag{6.5}$$

and where dP_b/dX is the probability distribution function of the test-statistic X for the background only hypothesis.

The signal hypothesis will be considered excluded at the confidence level CL when

$$CL_s \leq 1 - CL \ . \tag{6.6}$$

6.1.2 Maximum Likelihood Method

If a population has a probability distribution given by $F(x|\theta)$, the *likelihood* of the observations $x_1, x_2, ..., x_n$ for a specific θ is given by

$$\mathcal{L}(x_1, x_2, ..., x_n|\theta) = \prod_{i=1}^{n} F(x_i|\theta) \ , \tag{6.7}$$

and it expresses the joint conditional probability for obtaining the measurements, given θ. The *likelihood-ratio* λ is generally given as

$$\lambda \equiv \frac{\mathcal{L}(\hat{\omega})}{\mathcal{L}(\hat{\Omega})} \ , \tag{6.8}$$

where $\hat{\Omega}$ is the parameter space of θ, and $\hat{\omega}$ is a subspace of $\hat{\Omega}$, hence $0 < \lambda < 1$. $F(x|\theta)$ is now considered a function of $\theta = \{\theta_1, \theta_2, ..., \theta_k\} \in \hat{\Omega}$. If the null hypothesis is true, the parameters belong to the subgroup $\hat{\omega}$.

In search for new particles the approximate likelihood ratio Q is given by [21]

$$Q = \frac{\mathcal{L}(s+b)}{\mathcal{L}(b)} \ , \tag{6.9}$$

i.e. the ratio of products of probability densities for the signal+background hypothesis, to the products of probability densities for the background only (signal-free) hypothesis. The likelihood ratio, as a consequence of the Neyman-Pearson theorem [22, 21], maximizes the probability of rejecting a false hypothesis at a given confidence level, and conversely minimizes the probability of making a false discovery at a given discovery confidence level.

When searching for small signals in the presence of small backgrounds in several distinct channels, and where more than just the number of candidates will enter the likelihood ratio, it is in general not possible to obtain analytic expressions for the likelihood ratio probability distribution functions (p.d.f). The p.d.f's may instead be obtained by Monte Carlo generations of experiments according to the relevant hypothesis, and confidence levels, computed as the fraction of Monte Carlo experiments satisfying $Q \leq Q_{obs}$ [21] (see (6.2) and (6.4)).

Fig. 6.1 shows five different mass hypothesis distributions, normalized to one, together with the observed candidates, and gives a simple illustration of how the maximum likelihood method works in a search where the discriminating variable is the reconstructed mass of the candidate. Five Monte Carlo generated signal distributions are given for m_A values ranging from 70 to 90 GeV in steps of 5 GeV in the $\tan\beta = 20.0$ sector at $\sqrt{s} = 189$. The vertical bars on the top of the plot are the experimentally measured values, i.e. the reconstructed invariant mass of each observed candidate. As is seen in the plot, the mean of the signal distributions moves towards higher masses, and the width of the distributions increases as m_A increases. The likelihood is in this simple case given as

$$\mathcal{L} = \prod_{i=1}^{n} S_{m_i}(m_i) \ , \tag{6.10}$$

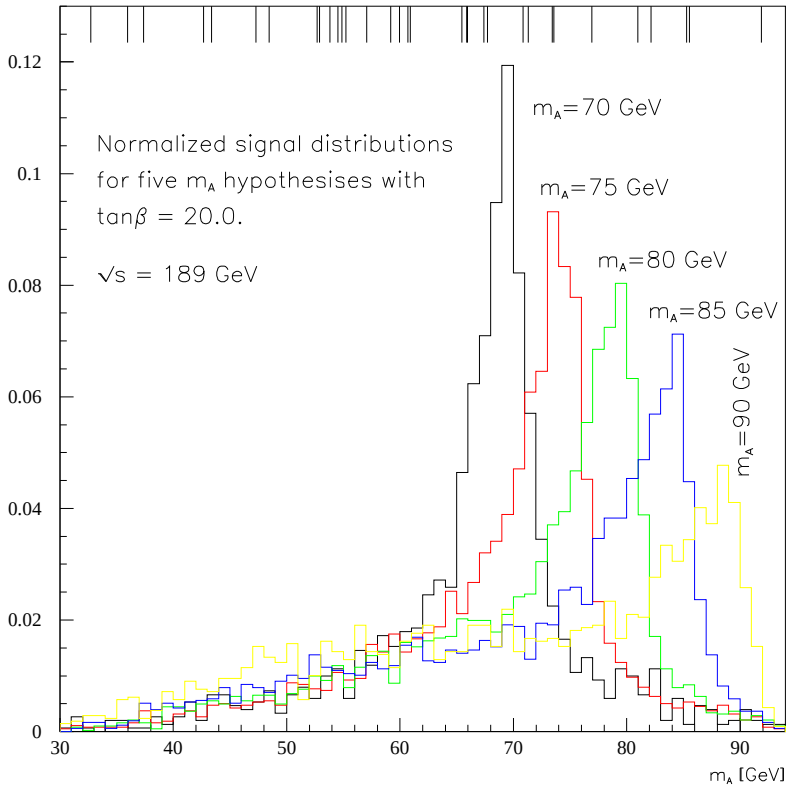

Figure 6.1: Five signal hypothesises with increasing mass in the $\tan\beta = 20$ sector together with the observed candidates of the 189 GeV data. The candidates are indicated by a vertical bar at the top of the plot.

where n is the number of observed candidates, $S_{m_i}(m_i)$ is the signal distribution for a given mass hypothesis evaluated for the i'th candidate, and m_i is the discriminating variable value of the i'th observed candidate. This likelihood is calculated for all the signal distributions, and the signal distribution that maximizes the likelihood $\mathcal{L}$ is the most likely m_A hypothesis.

In contrast to the single m_j used here, the methods of Chapter 7 use three discriminating variables, and the methods of Chapter 8 use five discriminating variables to separate signal from background.

Generic Likelihood Ratio

The likelihood ratio Q for experiments with N_{chan} independent search channels, with measurements of a single discriminating variable x for each candidate, and

with known absolute signal and background rates, can be written as [21]:

$$Q(m_H) = \frac{\prod_{i=1}^{N_{chan}} \frac{e^{-(s_i+b_i)}(s_i+b_i)^{n_i}}{n_i!}}{\prod_{i=1}^{N_{chan}} \frac{e^{-b_i} b_i^{n_i}}{n_i!}} \frac{\prod_{j=1}^{n_i} \frac{s_i S_i(x_{ij}) + b_i B_i(x_{ij})}{s_i+b_i}}{\prod_{j=1}^{n_i} B_i(x_{ij})} , \qquad (6.11)$$

where n_i is the number of observed candidates in each channel, x_{ij} is the value of the discriminating variable measured for each of the candidates, s_i and b_i are the integrated signal and background rates per channel, and $S_i(x_{ij})$ and $B_i(x_{ij})$ are the probability distribution functions of the discriminating variable for the signal and background, respectively.

6.2 The Inputs

The user defined inputs required by ALRMC are probability distributions for the signal hypotheses and the known background, the discriminating variable values of the observed candidates, and the efficiencies for each channel. Predefined inputs are tables of cross-sections and branching ratios used in the limit calculations.

The background distribution together with the signal distributions for all channels and mass hypotheses to be used in the ALRMC calculation, are provided by a file named `distributions.dat`, created from PAW-histograms of the distributions. It is of utmost importance to use same upper and lower edges for all the histograms of `distributions.dat`.

ALRMC only handles one or two discriminating variables. The values of the discriminating variable(s) for each observed candidate is given as `DATA` statements in the subroutine `d_xx_channels` of `d_higgs.f`, where `xx` is the name of the search channel[1]. `d_higgs.f` is a collection of Fortran subroutines that control which channels and efficiencies are used in the limit calculations, read the correct files, interpolate[2] the signal distributions, and return the SM expected signal, SM cross-sections and SM branching fractions.

The efficiencies[3] are not given as the efficiencies obtained at each simulated $(m_A, \tan\beta)$ point. Instead, a fit of the efficiency vs. mass is given as Fortran Real Functions. The efficiency-fits used in the limit calculations presented in this thesis are produced using the program `fiteff.kumac`[23] which takes ntuples of efficiency and mass as inputs. The efficiency Real Functions to be used in the calculations are concatenated into `efficiency.f`.

The predefined tables of cross-sections and branching fractions are made using the HZHA generator, and the tables are found in ref. [11].The `README` file of ref. [11] contains more details on the HZHA generations.

[1] `xx` is either `ha`, `hz` or `hinv`

[2] The interpolation procedure is not used in the analysis presented

[3] Efficiency used in this context means the experimentally obtained efficiency of each channel

6.3 ALRMC Customization

In a full limit calculation several channels at different center-of-mass energies are used. But the likelihood methods presented in this thesis are limited to one single channel at one single center-of-mass energy: $h^0A^0 \rightarrow b\bar{b}b\bar{b}$ at 189 Gev. The other channels are inhibited in the subroutine `fill_array` of `mssm.f` by setting the boolean array `QCHAN(I)`, where `I` is the channel number, false for all channels except for the h^0A^0 channel.

The Higgs mass hypothesis is characterized by three parameters in the h^0A^0 (or h^0Z^0) channel: The ratio $\tan\beta$ between the two vacuum expectation values in the MSSM Higgs model, the mass m_A of the neutral CP-odd scalar, and the mixing parameter α that diagonalizes the mass matrix. Section 2.2.4 gives more details on the parameters and mass spectrum of the MSSM Higgs model.

The h^0A^0 analysis presented is confined to calculate limits for only ten points in the $(m_A, \tan\beta)$ plane, because the objective of the analysis is to test different binning procedures and discriminating variables. Files of the general form `ecms_mixing.dat` [11] are tables of cross sections and branching fractions calculated for given m_A, $\tan\beta$ and m_h values. The center-of-mass energy is given by `ecms`, and `mixing` equals `no`, `max` or `typ` corresponding to the mixing applied in the HZHA generations. In order to customize ALRMC to only calculate limits for certain (m_A,$\tan\beta$) points, `ecms_mixing.dat` is modified to contain only the m_A and $\tan\beta$ values given by the signal hypotheses.

Further, the subroutine `excl_fast` of `mssm.f` is modified to only read the m_A and $\tan\beta$ values listed in the modified `ecms_mixing.dat`.

ALRMC contains an interpolation routine which interpolates between two Monte Carlo generated signal distributions in order to calculate limits for intermediate m_h values. However, the distributions produced by the likelihood methods presented in Chapter 7 and 8, are too complicated for the ALRMC interpolation procedure. The effect of the ALRMC interpolation routine is shown in Fig. 6.2 which is plot of the expected CL_s calculated with and without the interpolation procedure. From this plot it is evident that the intermediate CL_s (indicated by +), based on interpolated signal distributions, are not correct. Hence, the interpolation routine is skipped, using a `GOTO` statement in subroutine `d_sigdis_interpol` of `mssm.f`.

Finally the title-card `mssm.tit` must be set up correctly. The `LUMINOSITY` flag is set to the lowest m_A and $\tan\beta$ value, and the `OBSERVED` flag is set to a value that specifies which calculations ALRMC should perform.

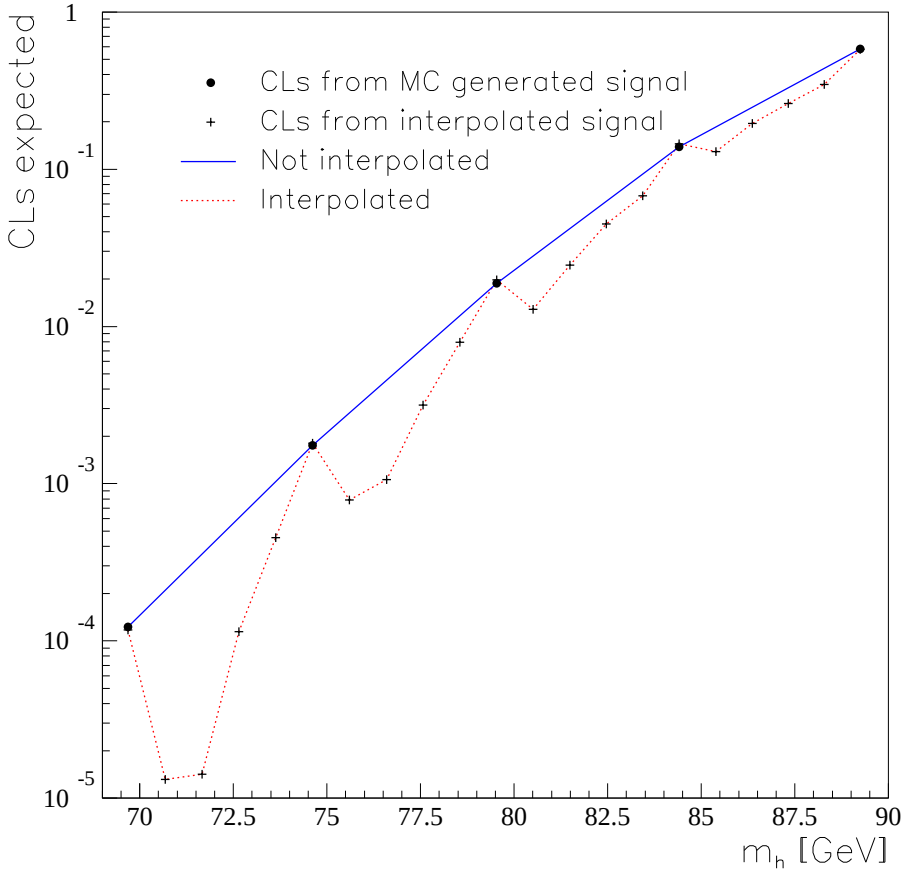

Figure 6.2: Plot of the expected CL_s calculated with and without the interpolation routine. The likelihood method used to calculate the CL_s of this plot is the variable/fixed binning procedure presented in Chapter 7.

Chapter 7

3D Likelihood Methods

In a search for a particular signal in a number of observed candidates, collected from a distribution of both background and signal, the signal is distinguished from the background using a variable with different distributions for signal and background. Such a variable is called a discriminating variable. The better the discriminating variable distinguishes the signal from background, the lower signal confidence levels (CL_s) are calculated, in the absence of a true signal, and the excluded region increases.

If three discriminating variables provide more information about the signal than two discriminating variables, it is likely to expect an improvement of the exclusion limits, in the absence of a true signal, if the three variables are used. This is the reasoning behind the development of the three dimensional (3D) likelihood methods.

ALRMC is implemented to handle a maximum of two discriminating variables. The 3D likelihood methods use three discriminating variables despite the ALRMC limit of two variables. This is possible because the three discriminating variable distributions are mapped into a one dimensional (1D) discriminating variable space. The mapping algorithm is presented in Sec. 7.5.

The three discriminating variable distributions are *binned* before the one dimensional discriminating variable is constructed. *Fixed* binning is the basic binning procedure, and *variable* binning is designed to enhance the resolution of the distributions, compared to the fixed binning procedure, by applying the available bins where the distributions change rapidly. The binning procedures are presented in Sec. 7.2 and Sec. 7.3, respectively.

There is a danger of overtraining the calculated confidence levels of the signal and background distributions if the bins are too small. Overtraining is further discussed in Sec. 7.4.

The observed candidates are given a treatment slightly different from the treatment of the signal and background distributions. The candidate binning and mapping are presented in Sec. 7.6.

Finally the PAW macros calling the binning and mapping routines, and initializing the appropriate 3D and 1D vectors of the 3D likelihood methods are presented in Sec. 7.7.

The term '3D likelihood method' is in this thesis used to denote a method

that takes a total of three discriminating variables as input, and gives one final discriminating variable, to be used with ALRMC, as output.

7.1 Three Discriminating Variable Distributions

Two sets of three discriminating variables are used as input to the 3D likelihood methods. The first set consists of the invariant masses m_h and m_A of the neutral MSSM Higgs scalars h^0 and A^0 (see Sec. 2.2.4) of the $h^0A^0 \to b\bar{b}b\bar{b}$ channel, and a discriminating variable which is the result of an analysis resembling a neural network analysis[1] presented in ref. [14]. Because the third discriminating variable resembles a neural network (NN) variable, it is for simplicity referred to as the *NN* variable in the following. Fig. 7.1 gives histograms of Monte Carlo generated m_h, m_A, and *NN* distributions for the $m_A = 80$ GeV, $\tan\beta = 20.0$ signal hypothesis, one of a total of ten signal hypotheses used with the 3D likelihood methods.

The second set of discriminating variables is only used with the fixed binning 3D likelihood method. This set consists of the sum of the invariant masses of the first set, $M = m_h + m_A$, the difference of the masses, $\Delta m = m_A - m_h$, and the same *NN* variable as in the first set. The implementation of the fixed binning method that uses the m_h, m_A set is identical to the implementation of the fixed binning method that uses the $M, \Delta m$ set. Only the m_h, m_A implementation is treated in detail in this chapter. Chapter 8 presents a 5D extension of the fixed binning 3D method, and the $M, \Delta m$ set is discussed in more detail in Chapter 8.

The three distributions of Fig. 7.1 are combined to form the 3D discriminating variable signal distribution of Fig. 7.2: The x-axis is the distribution of m_A, the y-axis is the distribution of m_h, and the z-axis is the distribution of the *NN* variable. The 3D background distribution is also shown in Fig. 7.2. The next step is to construct a 3D vector containing the 3D distribution by *binning* the distribution. The 3D vector is further treated in PAW using algorithms to produce a final 1D vector containing the discriminating variable distribution to be used as input to ALRMC. It is essential that exactly the same binning procedure is applied to all signal, background and candidate distributions used in the limit calculations.

7.2 Fixed Binning

The binning is performed by separating the 3D space of the 3D distribution into blocks, or *bins*, of a given size containing a small part of the whole distribution. An *event* of the distribution is a unique (m_A, m_h, NN) point in the 3D space spanned by the three variables, and a *bin* is a 3D sub-space that contains the number of events confined between the bin walls. A Fortran selection routine allocates each event to its corresponding bin. The routine loops over all the

[1] The analysis contain many features of a neural network, but the analysis is strictly speaking not a neural network. The variable is nonetheless referred to as a neural network variable for simplicity.

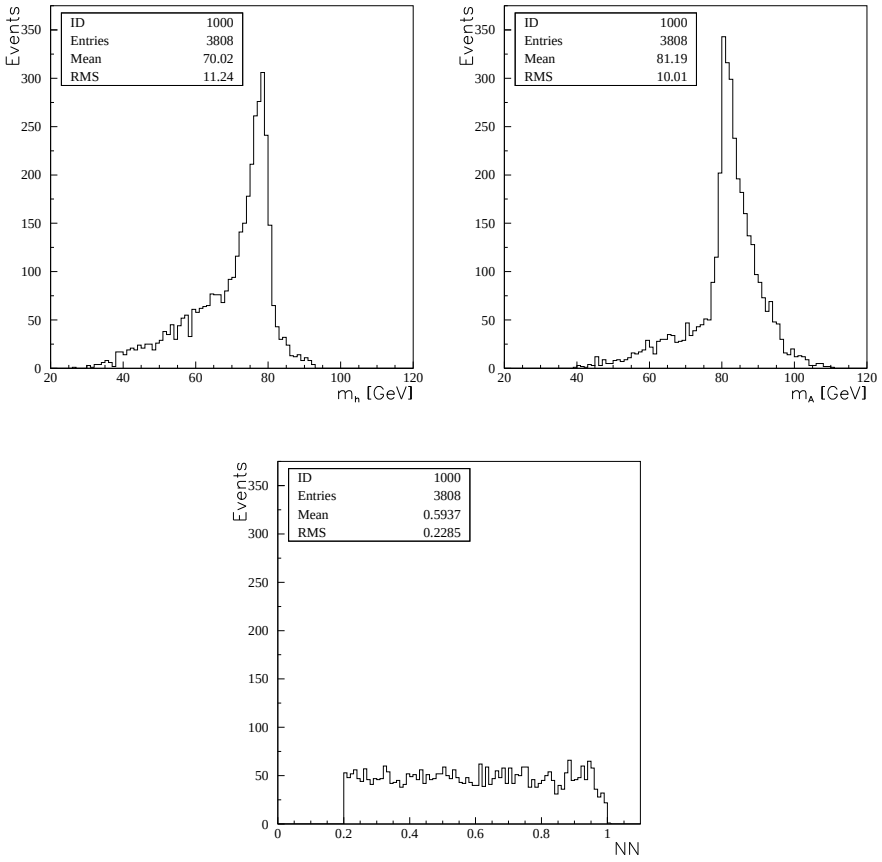

Figure 7.1: Monte Carlo generated distributions for the $m_A = 80$ GeV, $\tan\beta = 20.0$ signal hypothesis in the $h^0A^0 \to b\bar{b}b\bar{b}$ channel at 189 Gev. The top plots are the invariant mass distributions of m_h and m_A, respectively, and the bottom plot is the distribution of the neural network-like variable presented in ref. [14].

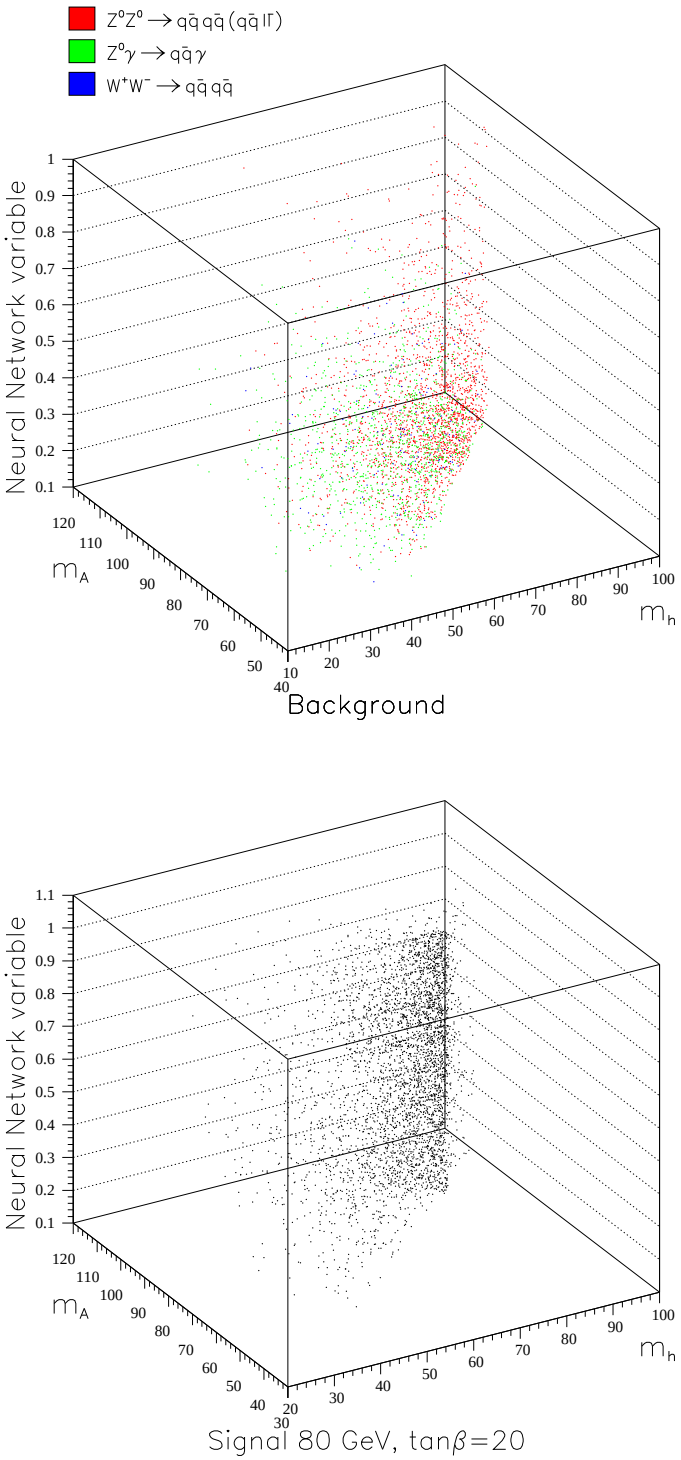

Figure 7.2: The three discriminating variable distributions of m_A, m_h, and NN plotted as 3D distributions. The top plot is the background 3D distribution, and the bottom plot is the $m_A = 80$ GeV, $\tan\beta = 20.0$ signal 3D distribution. The xy-plane is spanned by the two invariant masses, and the z-axis is the distribution of the neural network-like variable presented in ref. [14].

events of the distribution, and three bin index numbers, I_θ, are calculated for each discriminating variable θ using

$$I_\theta = \frac{\theta(j) - \theta_{min}}{\Delta_\theta} + 1 \ , \tag{7.1}$$

where θ is m_h, m_A, or NN, j is the event number, θ_{min} is the minimum value of the θ distribution, and Δ_θ is the binwidth of the bins of the θ distribution. The content of the bin is stored in a 3D vector of index $(I_{m_h}, I_{m_A}, I_{NN})$, and a control routine gives an error message if an event is not placed in any bin. The procedure given here is referred to as the *fixed binning* procedure.

7.2.1 Resolution

The total number of bins, N_{3D}, used in a 3D binning procedure, i.e. the size of the 3D vector, is given by

$$N_{3D} = N_{m_A} N_{m_h} N_{NN} \ , \tag{7.2}$$

and the number of bins N_θ used to bin each variable θ is given by

$$N_\theta = \frac{\theta_{max} - \theta_{min}}{\Delta_\theta} \ , \tag{7.3}$$

where θ_{max} and θ_{min} are the maximum and minimum value of the θ distribution, respectively. The only way to improve the resolution of a given distribution of fixed bins is to reduce Δ_θ. Hence, in the case of fixed binning, the size of the distribution vector is very sensitive to the resolution. In an attempt to increase the resolution, but not the size of the vector, a procedure referred to as *variable binning* is developed. It is also interesting to vary the binning without increasing the total size of the vector because ALRMC sets an upper limit on the number of bins the vector might contain[2].

7.3 Variable Binning

As is seen in Fig. 7.2 the 3D distribution of the signal is concentrated to the furthest corner of the box, leaving a lot of empty space in the box. Thus, when fixed binning is applied to the distribution, a few bins of the 3D distribution contain large samples of the distribution while a lot of the bins are almost empty. The information on the distribution inside the bin is lost after the binning has been performed. Hence, the larger portion of a distribution a single bin contains, the more information is lost, and the effect is a lower resolution of the distribution. The resolution therefore strongly depends on the size of the bins. If the resolution is to be enhanced without changing the total number of bins, the use of the bins must be optimized by applying the available bins to the parts of the distribution where the important information is. Thus, better

[2]The ALRMC maximum number of bin limit is not final, and may be changed in the ALRMC setup. However, a higher limit will result in using more CPU time and hence slow the calculations down.

resolution is achieved by making the size of the bin depend on the number of events inside the bin. This procedure is commonly known as a Fuzzy Box procedure. In the variable binning procedure presented here, the size of the bin is made to depend on the position of the bin rather than number of events, thus trying to simplify the implementation of the variable binning procedure. Because the number of events inside the bin depends on the position of the bin, the binsize of the variable binning procedure only indirectly depend on the number of events. Hence, variable binning is only an approximate Fuzzy Box procedure. The different binwidth regions are determined by inspection of the distributions.

The first step of variable binning is to bin the whole distribution using the binwidths required for the sensitive region of the signal or background distribution. The result is a large vector with many still empty bins, but the number of events in the filled bins is lower than in the fixed binning case because several bins have now been used to bin a part of the distribution that one single bin covered earlier.

The next step is to reduce the size of the vector by combining the bins that do not contain any important information about the distribution; a *rebinning* of the vector must be performed. The rebinning is done in a separate Fortran routine that takes the large vector containing the distribution in small, fixed bins as input, and gives a rebinned vector with a reduced number of entries as output, containing the distribution in bins of a size that depends on the position of the bin. The rebinning procedure adds the contents of the bins that are to be combined, and stores the sum in the correct bin of the output vector. The bins that are not to be combined, i.e. the bins that contain important distribution information, are just transferred to the correct bin of the output vector with their original bin content. A control routine is called after the rebinning is performed to check if the total number of events of the distribution has changed during rebinning. Fig. 7.3 gives the bins in the first and last step of the variable binning procedure to the left and right, respectively.

Two variable binning procedures are implemented. The first procedure applies variable binning to m_h and m_A, while the bins of the NN variable is unchanged. This procedure only takes the sensitive regions of the signal distributions into account. Sec. 7.3.1 provides more details.

The second procedure applies variable binning to all three variables. This procedure takes the sensitive regions of both the signal and background distributions into account. Sec. 7.3.2 provides more details.

7.3.1 Rebinning of m_h and m_A

In the case of the m_h and m_A distributions, the region where the smallest bins are used is the region of about two standard deviations centered around the mean, as seen in Fig. 7.3. The signal distribution of the NN variable in Fig. 7.1 is close to flat, i.e. no region of divergent event density exists, and fine binning is not needed to increase the resolution of the signal NN distribution.

The same binning procedure and resolution must be applied to all the distributions that enter the search. Hence, the appropriate region for fine binning

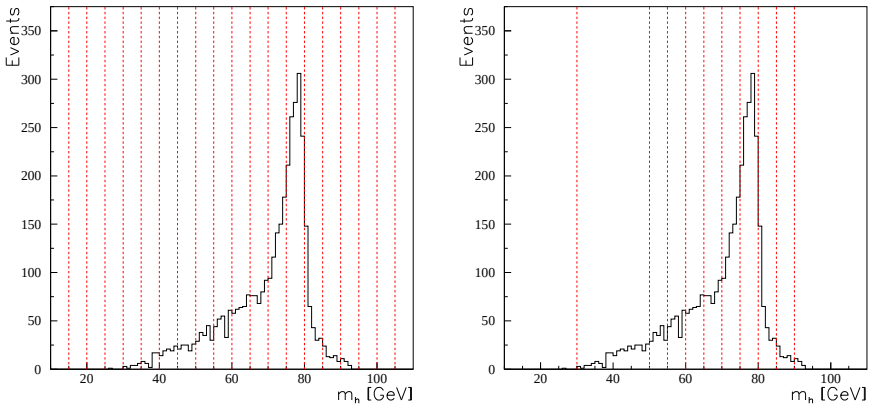

Figure 7.3: Plot of the m_h distribution where the bins are indicated by dotted lines. The plot to the left shows the bins in the first step of the variable binning procedure, and the plot to the right shows the bins after the rebinning procedure has been applied.

is the region that covers the peak of all the distributions that are to be rebinned. Fig. 7.6 and 7.7 show scatter plots of the m_h distribution versus the m_A distribution for the ten signal hypotheses, and the background distribution, presented in Chapter 4. As seen in these plots, the region of fine binning always covers the most event dense, or peaked, area of the plot.

7.3.2 Rebinning of m_h, m_A and NN

Fig. 7.4 shows the background *NN* distribution, and the distribution is not flat, as for the signal *NN* distribution, but monotonically decreasing. Thus, a higher resolution of the *NN* variable is of interest for only one distribution; the background distribution of the *NN* variable. A better resolution of the background distribution, might as well result in a lower CL_s as a better resolution of the signal distributions.

7.4 Overtraining

The smaller the bins are, the more information is transferred from the original distribution to the binned distribution. Because the calculated limit depends on the resolution of the input distribution, it follows that the limit is a function of the size of the bins. There is, however, a statistical limit on how small the bins might be, depending on the number of Monte Carlo (MC) generated events. If the bins are too small, an *overtraining* of the calculated limits occur. The result is a limit not based on the observed data, and a falsely excluded mass region. However, a plot of the confidence level as a function of binwidth, $CL_s(\Delta_\theta)$, exposes the overtraining region. Three functions are needed in order to make such a plot: CL_s of the first half MC events, CL_s of the second half

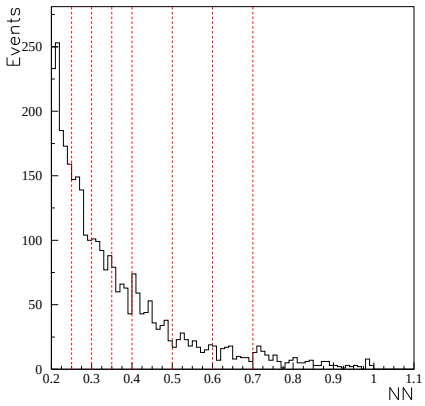

Figure 7.4: The NN variable distributions for the background hypothesis in the $h^0A^0 \to b\bar{b}b\bar{b}$ channel at 189 Gev. The bins are indicated by dotted lines

MC events, and CL_s of all MC events. A plot of three CL_s functions is given in Fig. 7.5 for the fixed binning 3D likelihood method. The area to the left of the vertical line in Fig. 7.5 is the region where there is an overtraining of the limits, because in this area the CL_s of all MC events (green line) is no longer the mean value of the CL_s of the first (red line) and second half (blue line) MC events. Overtraining plots of all the presented likelihood methods are given in Fig. 10.1.

7.5 Mapping from 3D Space to 1D Space

Because ALRMC does not handle three discriminating variables, the three variables are reduced to one variable by mapping a bin in 3D space to the corresponding bin in 1D space. The mapping is performed by a Fortran subroutine named `dim3tool.f` that takes the 3D vector of 3D bins as input, and assigns a 1D bin to each of the 3D bins. The subroutine loops over all the 1D bins of the 1D vector, and the index number I of the 1D bin that the content of the 3D bin of index $(I_{m_h}, I_{m_A}, I_{NN})$ is transferred to is given by

$$I = (I_{m_A} - 1)N_{m_h} + (I_{NN} - 1)N_{m_h}N_{m_A} + I_{m_h} \ , \tag{7.4}$$

where N_{m_h} and N_{m_A} are the total number of bins of the m_h and m_A distributions, respectively.

7.6 The Candidates

The observations of a search experiment are often referred to as *candidates*, and the candidates are identified by their position in the generated discriminating variable distributions. The likelihood $\mathcal{L}$ of the observed candidates is the product of the discriminating variable distribution evaluated at each candidate,

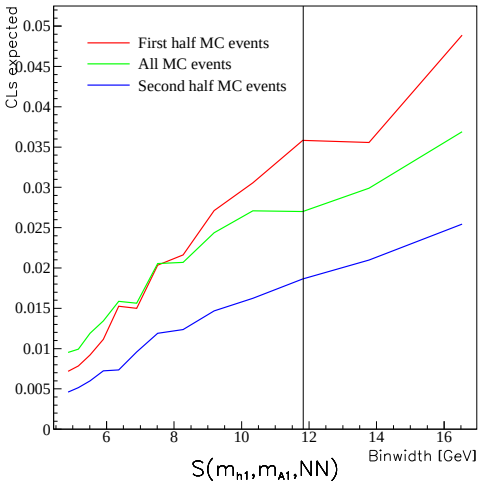

Figure 7.5: Plot of confidence level as a function of binwidth. This kind of plot is used to determine the optimal binwidth that does not lead to any overtraining of the obtained results.

see (6.10). Fig. 6.1 shows different signal distributions for one single discriminating variable, together with the observed candidates. If three discriminating variables are used in the limit calculations, three values must be given for each candidate, corresponding to the position of the candidate in each of the three distributions.

No information is stored during the binning processes described in Sec. 7.2 or Sec. 7.3 about in which 3D bin the single events of the 3D distribution ends up. A bin of a signal or background distribution contains the sum of events inside the bin. In contrast, a bin of a candidate contains a number that identify which candidate(s) belongs to the bin.

7.6.1 Fixed Binning

Only the bins containing candidates have to be mapped into 1D space, and the binning and mapping of the candidates are performed by one single routine. Which bin the candidate belongs to in 3D space is calculated using (7.1) for each of the three variables, and the 3D bin is mapped into 1D space using (7.4). The three variables given for each candidate is reduced to a single integer number that is the position of the candidate in the 1D space of the final 1D discriminating variable.

7.6.2 Variable Binning

If variable binning is applied to the distributions, the binning and mapping of the candidates are more complicated than in the fixed binning case. When several small bins are combined into one single large bin by the rebinning routine of Sec. 7.3, it is necessary to know which small bins contain which candidates,

if the candidate number, identifying the candidates of the resulting large bin, is to be correct. Two 3D vectors of equal size are needed in order to keep record of which bin contains which candidate(s): **A** denotes a 3D vector that contain a number which represent the candidate(s) of the bin, while **B** denotes a 3D vector which is a counter of how many candidates belong to the bin.

One bin may contain more than one candidate, and a simple algorithm that adds a new candidate to a bin without erasing the old candidate is given by

$$\mathbf{A}_{xyz} = \sum_{i=1}^{\mathbf{B}_{xyz}} C_i 10^{2(i-1)} \ , \tag{7.5}$$

where $\mathbf{A}_{xyz}$ is a number representing the candidates in bin (I_x, I_y, I_z), $\mathbf{B}_{xyz}$ is the total number of candidates in the same bin, and $C_i \in [1, 33]$ is the candidate number of the i'th candidate of the bin.[3] If the candidates numbered 3, 12, 27 all end up in the bin of index (1,1,2), the content of $\mathbf{A}_{112}$ is 271203, and the content of $\mathbf{B}_{112}$ is 3.

The candidate bins have to be rebinned in the same way as the distribution bins. The rebinning of the candidates is however more complicated than the rebinning of the distributions, because the contents of both **A** and **B** must be rebinned simultaneously. There is also an upper limit on how many candidates the **A** vector can keep track of. If more than four candidates end up in the same bin, the overflow candidate(s) is placed in an overflow $\mathbf{A}_+$ vector. The limit on four candidates is purely technical. A PAW vector can only contain integer numbers less than 10^8, and $\mathbf{A}_{xyz} > 10^8$ if $\mathbf{B}_{xyz} > 4$ and $C > 1$.

The contents of the bins of the rebinned **A** vector, $\mathbf{A}'$, are calculated using

$$\mathbf{A}'_{xyz} = \sum_{i=1}^{n} \mathbf{A}_{x^i y^i z^i} 10^{2(\mathbf{B}_{x^i y^i z^i} - 1)} \ , \tag{7.6}$$

where $\mathbf{A}'_{xyz}$ is a number representing the candidates in bin (I'_x, I'_y, I'_z), n is the number of small bins contained in the large bin, $\mathbf{A}_{x^i y^i z^i}$ is the number representing the candidate(s) of the i'th small bin, $\mathbf{B}_{x^i y^i z^i}$ is the number of candidates in the i'th small bin, and $\mathbf{A}'_{xyz}$ is the number representing the candidate(s) of the n small bins.

The last step of the candidate preparation is to map the rebinned candidate bins from 3D space to 1D space while keeping track of which candidate ends up in which 1D bin. The mapping procedure is implemented in a separate routine. A candidate number C is read from $\mathbf{A}'$ using

$$C = \frac{\mathbf{A}'_{xyz}}{10^{2\mathbf{B}'_{xyz}}} \ , \tag{7.7}$$

where $\mathbf{B}'_{xyz}$ is the number of candidates in $\mathbf{A}'_{xyz}$. Because C in declared to be an integer in the Fortran routine, the result of (7.7) is the integer part of the expression. The contents of $\mathbf{A}'_{xyz}$ after candidate C is read is the modulo of

[3] The analysis presented use the data of the 189 GeV run, and a total of 33 observed candidates are identified in the 189 GeV data (see Table 4.1).

$\mathbf{A}'_{xyz}$ and $10^{2\mathbf{B}'_{xyz}}$. Thus, if $\mathbf{A}_{112}$ contain 271203, the value of C is 27, and the new value of $\mathbf{A}_{112}$ is 1203. The value of $\mathbf{B}'_{xyz}$ is reduced by one each time a candidate number C is read, and (7.7) is repeatedly used until $\mathbf{B}'_{xyz} = 0$.

The index numbers I of the 1D bins that contain all 33 candidates are calculated using (7.4) with the bin index numbers (I'_x, I'_y, I'_z) of $\mathbf{A}'_{xyz}$ as input. The resulting candidate list, containing the 33 index numbers of the candidates, is given as `DATA`-statements in subroutine `d_ha_channels` of `h_higgs.f`.

7.7 The 3D Input Macros

All the different algorithms and Fortran routines described in this chapter are called from three different PAW macros. Fig. 7.8 gives the structure of the three PAW macros as one single flow chart. Each of the PAW macros are used with the signals, backgrounds, and candidates presented in Chapter 4.

Fig. 7.9 gives outputs of the 3D likelihood methods presented in the case of one signal hypothesis and the background hypothesis. The plots in the left column are background 1D distributions, and the plots in the right column are signal 1D distributions. As seen in the plots, the resolution of the background distribution is best when all three variables are binned with different binsizes, but the signal distribution is best when only the mass variables are binned with different binsizes. The resolution of both the background and signal distribution is low in the case of fixed binning of all variables. However, what really matters is which final discriminating variable provides the most *different* signal and background distribution.

Chapter 10 presents the limits calculated using the 3D methods, and Sec. 10.2.1 presents the conclusions on the 3D methods presented in this chapter.

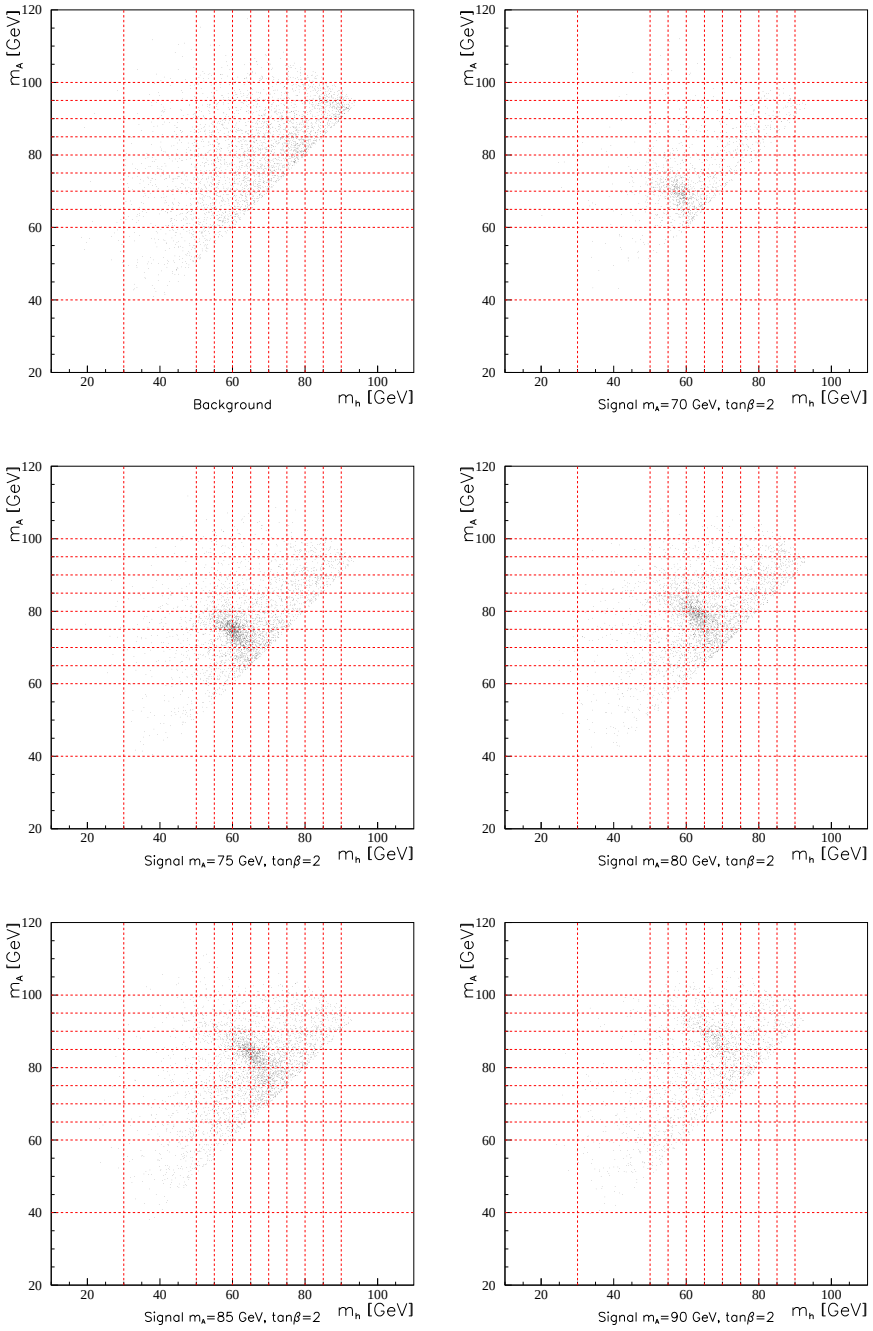

Figure 7.6: Scatter plots of m_h and m_A distributions of the background hypothesis and the five signal hypotheses $m_A = 70, 75, 80, 85, 90$ GeV at $\tan\beta = 2$ in the $h^0A^0 \to b\bar{b}b\bar{b}$ channel at 189 GeV . The dotted lines indicate the bins used in the variable binning of the distributions.

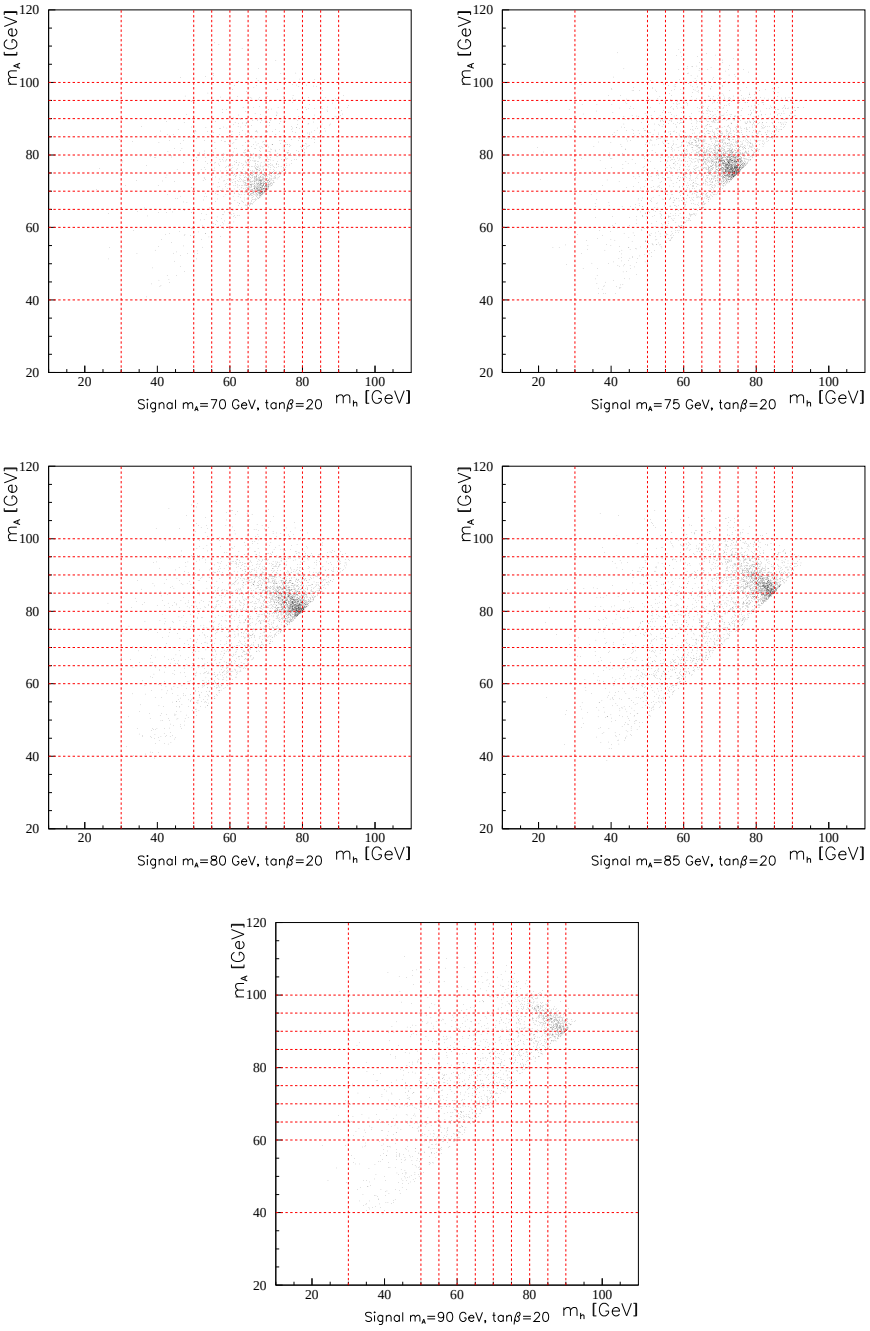

Figure 7.7: Scatter plots of $m_h$and m_A distributions of the background hypothesis and the five signal hypotheses $m_A = 70$, 75, 80, 85, 90 GeV at $\tan\beta = 20$ in the $h^0A^0 \to b\bar{b}b\bar{b}$ channel at 189 GeV . The dotted lines indicate the bins used in the variable binning of the distributions.

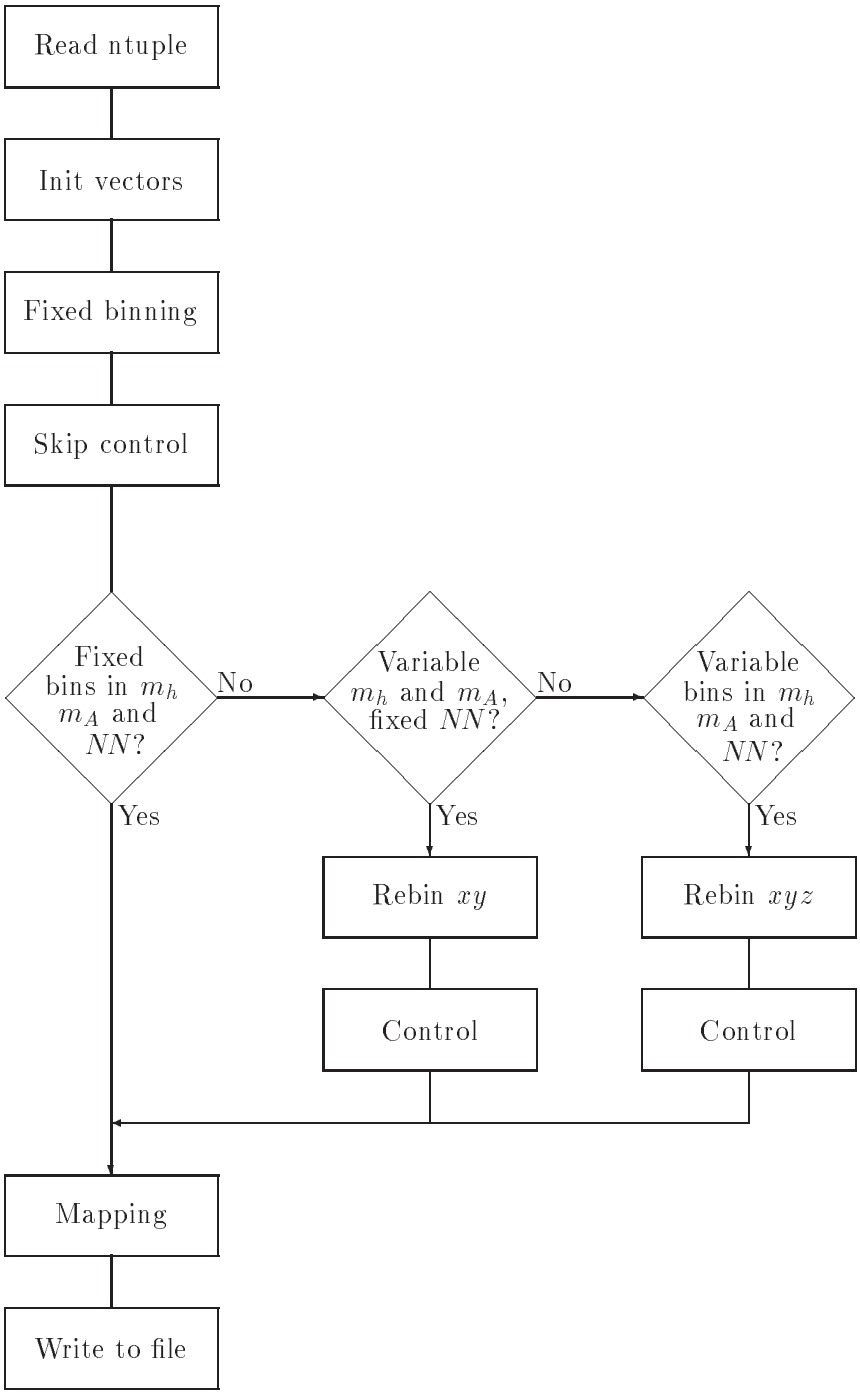

Figure 7.8: Flow chart of the three PAW macros preparing 3D inputs.

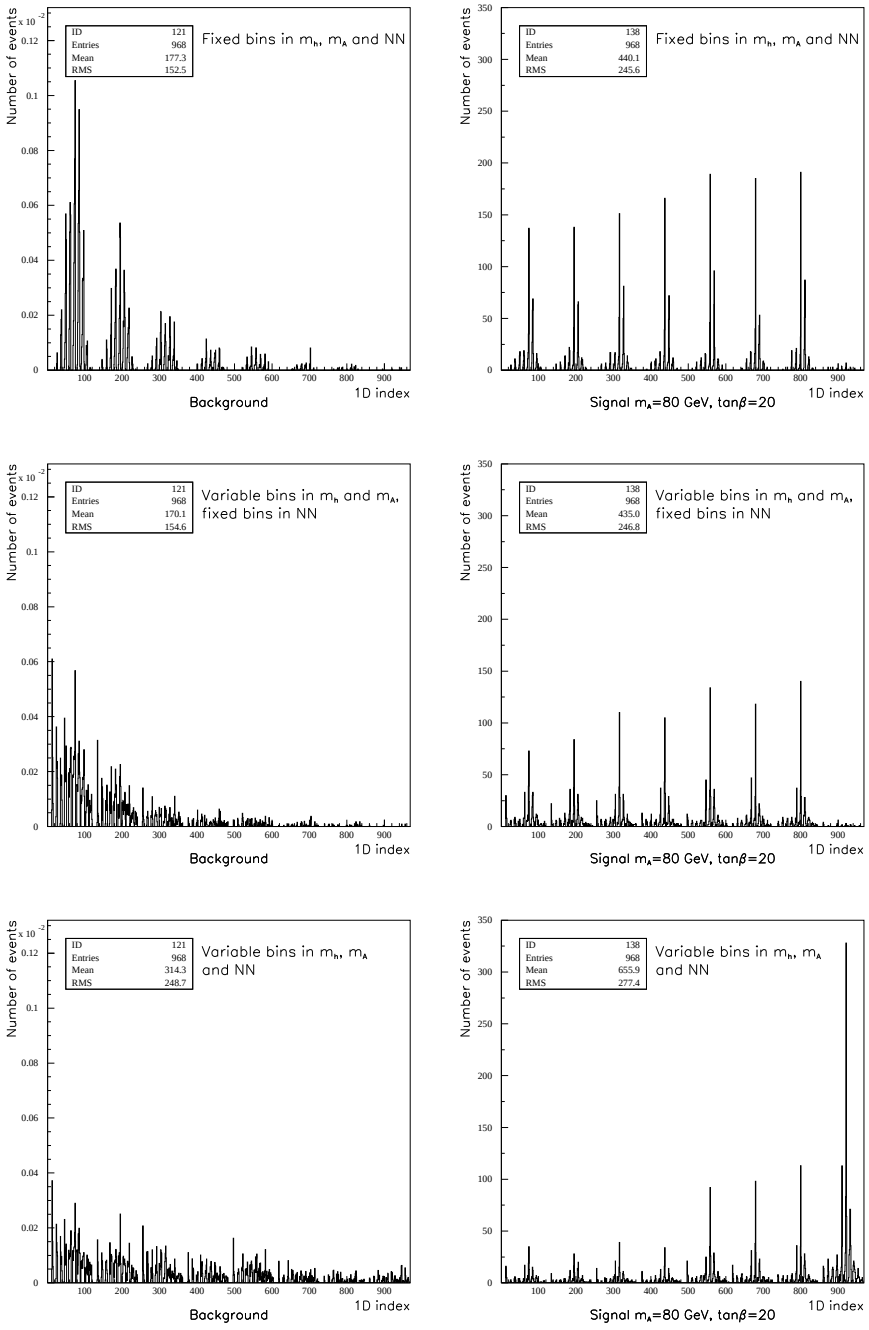

Figure 7.9: Plots of the final 1D discriminating variable distribution. The plots in the left column are background distributions, and the plots in the right column are signal distributions. The top row plots show distributions in the fixed binning case, the middle row plots show distributions in the case of variable binning of the masses, and the bottom row plots show distributions in the case of different binsizes in all three variables.

Chapter 8

5D Likelihood Methods

The jet clustering of the decay products of a $h^0A^0 \to b\bar{b}b\bar{b}$ event results in a four-jet event where each jet originates from one of the b-quarks. The problem of which two jets originate from h^0, and which two jets originate from A^0 is the problem of *pairing* the jets.

The mass of h^0 and A^0 are calculated as the invariant mass of the sum of the four-momenta of two of the four jets (see (8.1). The three possible combinations of grouping four jets into two groups of two jets each, are ordered according to a certain criterion. The criterion used for the 5D likelihood methods is the difference in mass, Δm, between the two groups of paired jets, i.e. $\Delta m = m_A - m_h$. Calculation and selection of the two best pairing h^0 and A^0 masses are further discussed in Sec. 8.1. However, the *best* pairing of the jets, according to the criterion, is not always the *correct* pairing of the jets. It is therefore of interest to look at a pairing extension of the 3D likelihood method of Chapter 7 which only included the best pairing.

The most intuitive approach to such an extension is to use the two invariant mass distributions (m_h and m_A, or M and Δm) of the second and third best pairing in addition to the three discriminating variable distributions used by the 3D likelihood method. The result is a 7D discriminating variable distribution. The binning and mapping of a 7D distribution is very complicated, and in addition the size of the 7D distribution rapidly increases beyond the limits imposed by ALRMC: If four bins are used to bin each of the seven distributions, the total number of bins is $4^7 = 16384$. Therefore, the extension of the 3D likelihood method is limited to include only the second best pairing, and a 5D discriminating variable distribution is constructed. The size of a 5D distribution with four bins in each variable is $4^5 = 1024$.

Sec. 8.2 present two sets of five discriminating variables used by the 5D likelihood methods. The two sets are two-pairing extensions of the two one-pairing sets used with the 3D likelihood methods. The binning and mapping of the distributions and the candidates are discussed in Sec. 8.3 and Sec. 8.4, respectively.

The PAW macros of the two 5D methods, each using one of the two sets of discriminating variables, are briefly discussed in Sec. 8.5

The term '5D likelihood method' is in this thesis used to denote a method

that takes a total of five discriminating variable distributions as input, and gives one final discriminating variable, to be used with ALRMC, as output.

8.1 Calculation of Two Best Pairings

Each event of the distributions used with the 5D likelihood methods are identified by the following variables: The two invariant masses m_{h1} and m_{A1} of the best pairing, the NN variable, the weight of the event, and the four-momentum p^{μ}_{α} of each of the four jets, where $\alpha = 1, ..., 4$ is the jet number. The invariant mass $m_{\alpha\beta}$ of the $\alpha\beta$ pair is calculated using

$$m^2_{\alpha\beta} = g_{\mu\nu}(p^{\nu}_{\alpha} + p^{\nu}_{\beta})(p^{\mu}_{\alpha} + p^{\mu}_{\beta}) \ , \tag{8.1}$$

where α and β are the numbers of the two combined jets, $\alpha \neq \beta$, and $m_{\alpha\beta} = m_{\beta\alpha}$. If $\tan\beta > 1$, the lightest mass of a $m_{\alpha\beta}$ and $m_{\gamma\delta}$ pair, with $\alpha \neq \beta \neq \gamma \neq \delta$, is m_h, while the other mass is m_A.

The three $m_{\alpha\beta}, m_{\gamma\delta}$ pairs are ordered according to the difference Δm between the masses of each pair. The pair with the lowest Δm is the best pair, and the pair with the second lowest Δm is the second best pair. The masses m_{h1} and m_{A1} are provided by the ntuple, together with the four-momenta, prior to the calculations. These masses are used to make a consistency check of the calculated best pairing masses. An error message is given if m_{h1} and m_{A1} do not match the calculated best pairing masses.

One ntuple is filled with the masses of the best pairing, and another ntuple is filled with the masses of the second best pairing. The mass calculations and the filling of the new ntuples are performed by a Fortran routine that is called for each of the ten signal distributions, the background distribution, and the observed candidates.

8.2 The Discriminating Variables

Two different sets of discriminating variables are used in the 5D likelihood method. The first set is a 5D extension of the first set of the 3D discriminating variables of Chapter 7: The two invariant masses m_{h1} and m_{A1} of the best pairing, the two invariant masses m_{h2} and m_{A2} of the second best pairing, and the NN variable.

The second set of discriminating variables is the first set replaced by the sum $M = m_A + m_h$, and the difference $\Delta m = m_A - m_h$. The variables of the second set are: M_1 and Δm_1 of the best pairing, M_2 and Δm_2 of the second best pairing, and the NN variable.

Fig. 8.1 shows distributions of the masses of both sets of discriminating variables for one particular signal hypothesis. The plots clearly show that the masses of the second best pair are less peaked than the masses of the best pair, and no variable binning of the distributions is performed. The reason is that no region of the second best pair distribution is sufficiently peaked.

Where the first bin of a distribution $S_{\theta}(\theta)$ starts and the last bin ends, is determined by θ_{min} and θ_{max}. Because exactly the same binning is applied

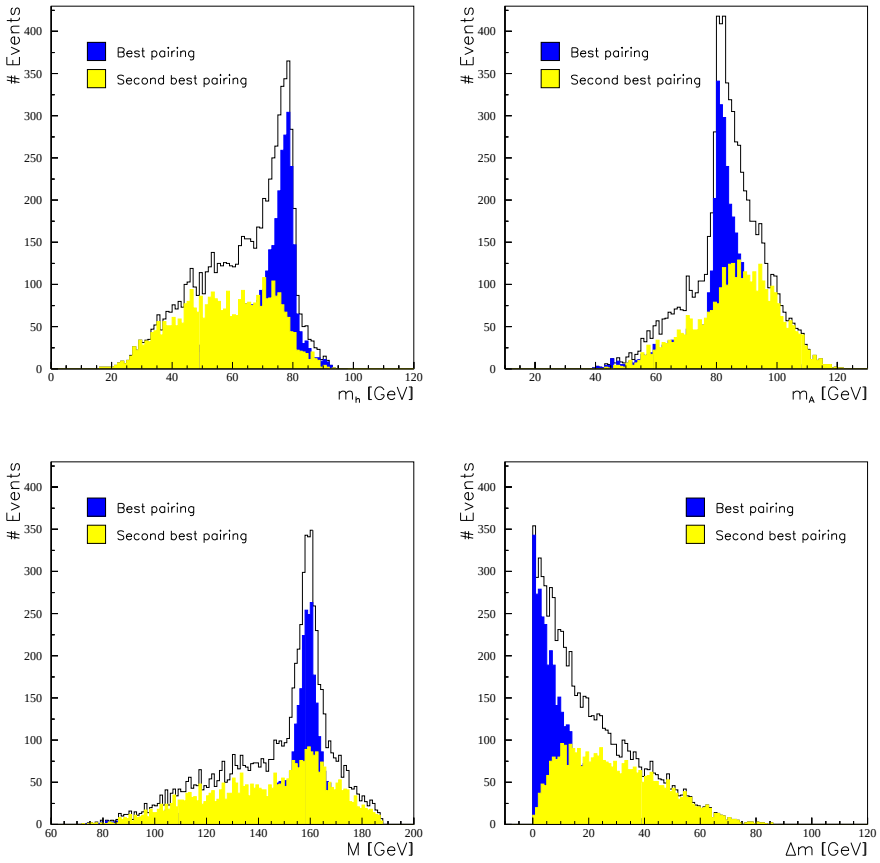

Figure 8.1: Plots of the mass discriminating distributions of the two best pairings used by the 5D likelihood methods. The top left plot shows m_{h1} and m_{h2}, and the top right plot shows m_{A1} and m_{A2}. The bottom left plot shows $M_1 = m_{h1} + m_{A1}$ and $M_2 = m_{h2} + m_{A2}$, and the bottom right plot shows $\Delta m_1 = m_{A1} - m_{h1}$ and $\Delta m_2 = m_{A2} - m_{h2}$. All four distributions are from the signal hypothesis $m_A = 80$ GeV, $\tan\beta = 20$, which corresponds to $m_h = 79.53$ GeV.

to all the distributions, θ_{min} and θ_{max} are calculated by looping over all the distributions that enter the search. The routine that calculates the masses of the different pairings also provides the θ_{min} and θ_{max} values used in the binning of the distributions.

The treatment of the two different sets of discriminating variables is identical, and only the $M, \Delta m$ variable set is discussed in this chapter.

8.3 Binning and Mapping of the Distributions

The binning of the 5D discriminating variable distribution can not proceed by allocating each $(M_1, \Delta m_1, M_2, \Delta m_2, NN)$ event to its corresponding bin in 5D space directly, as is the case with the 3D distribution, because PAW only supports vectors of maximum three dimensions. Hence, the 5D distribution must somehow be separated into less dimensioned distributions that is multiplied to construct an approximate 5D distribution. The distribution function S of the five discriminating variables is approximated as

$$S(M_1, \Delta m_1, M_2, \Delta m_2, NN) \approx S_1(M_1, \Delta m_1)\, S_2(M_2, \Delta m_2)\, S_{NN}(NN)\ , \quad (8.2)$$

which does not include correlations between masses of different pairings, nor between the masses and the NN variable. If all correlations are to be included, each event of the discriminating variable distribution must be treated as single $(M_1, \Delta m_1, M_2, \Delta m_2, NN)$ points and not as three separate $(M_1, \Delta m_1)$, $(M_2, \Delta m_2)$, and (NN) points.

The 5D method could more correctly be referred to as a 2D+2D+1D method because the 5D distribution of the five discriminating variables is reduced to two 2D and one 1D distribution. The term '5D' is instead used as a shorthand notation.

The 2D mass distributions $S_{p,1}(M_p, \Delta m_p)$, where $p = 1$ or 2, are binned using a binning procedure similar to the fixed binning procedure described in Sec. 7.2. The only difference is that this is a 2D binning instead of a 3D binning. Two bin index numbers I_{M_p} and $I_{\Delta m_p}$ are calculated using (7.1) twice for each $S_{p,1}(M_p, \Delta m_p)$ distribution.

The 2D bins of the two mass distributions are mapped from 2D space to 1D space using a 2D version of (7.4):

$$I = (I_{\Delta m_p} - 1) N_{M_p} + I_{M_p}\ , \quad (8.3)$$

where I is the index of the 1D vector, and N_{M_p} is the total number of bins of the $S_{M_p}(M_p)$ distribution.

The distribution $S_{NN}(NN)$ is binned using (7.1) once to calculate the 1D bin index number I_{NN}. No mapping is required.

All three 1D distributions are normalized to one, and a simple Fortran routine calculates the 2D distribution of the four mass variables as the product of $S_1^{1D}(M_1, \Delta m_1)$ and $S_2^{1D}(M_2, \Delta m_2)$. The resulting 2D distribution of the four masses is mapped from 2D to 1D space using (8.3). A 2D distribution of all five variables is calculated as the product of the 1D distributions of the four

masses and $S_{NN}(NN)$. The final 1D version of the discriminating variable is calculated using (8.3) once again.

Fig. 8.2 gives the signal distributions $S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)S_{NN}(NN)$ and $S_3(m_{h1}, m_{A1})S_4(m_{h2}, m_{A2})S_{NN}(NN)$ to the top and bottom right, respectively, as well as the background distributions $B_1(M_1, \Delta m_1)B_2(M_2, \Delta m_2)$ $B_{NN}(NN)$ and $B_3(m_{h1}, m_{A1})B_4(m_{h2}, m_{A2})B_{NN}(NN)$ to the top and bottom left, respectively. From these plots it looks as the peaks of the $S_1(M_1, \Delta m_1)$ $S_2(M_2, \Delta m_2)S_{NN}(NN)$ distribution is wider than the peaks of the $S_3(m_{h1}, m_{A1})$ $S_4(m_{h2}, m_{A2})S_{NN}(NN)$ distribution, which might indicate a better resolution in the former distribution. This is also the case for the two background distributions. The explanation might be that the events of the $M_1, \Delta m_1, M_2, \Delta m_2, NN$ set are more evenly distributed among the fixed bins than the events of the $m_{h1}, m_{A1}, m_{h2}, m_{A2}, NN$ set, and thus a better resolution is obtained for the $S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)S_{NN}(NN)$ and $B_1(M_1, \Delta m_1)B_2(M_2, \Delta m_2)B_{NN}(NN)$ distribution.

8.4 Binning and Mapping of the Candidates

The binning and mapping of the five discriminating variable distributions given for each signal or background hypothesis, is performed in order to produce the number of events inside each bin of the final 1D distribution. In contrast, the binning and mapping of the five discriminating values given for each observed *candidate*, is performed to produce the index number of the bin of the final 1D distribution that contains the five values given for the candidate; i.e. each candidate must be given a map of where to go in the 1D distribution. No information about the distribution of candidates is needed.

Before the address of a candidate bin of the final 1D distribution is calculated, five bin index numbers, corresponding to each of the initial five distributions, are calculated using (7.1) five times. The resulting bin index numbers, I_{M_1}, $I_{\Delta m_1}$, I_{M_2}, $I_{\Delta m_2}$, and I_{NN} are reduced to a single 1D bin index number I by applying (8.3) four times.

Fig. 8.3 gives a visualization of how the bin index number in the 1D space of the final distribution is calculated based on the bin index numbers of the five discriminating variable distributions. The 2D mapping algorithm (8.3) is applied at each vertex of Fig. 8.3.

8.5 The 5D Input Macros

Two different PAW macros prepare the 5D inputs: One macro handles the $S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)S_{NN}(NN)$ distribution, while the other macro handles the $S_3(m_{h1}, m_{A1})S_4(m_{h2}, m_{A2})S_{NN}(NN)$ distribution. The structures of the macros are identical. The only differences between the two macros are the input distributions, and hence the θ_{min} and θ_{max} values. Fig. 8.4 shows the flow chart of the macros.

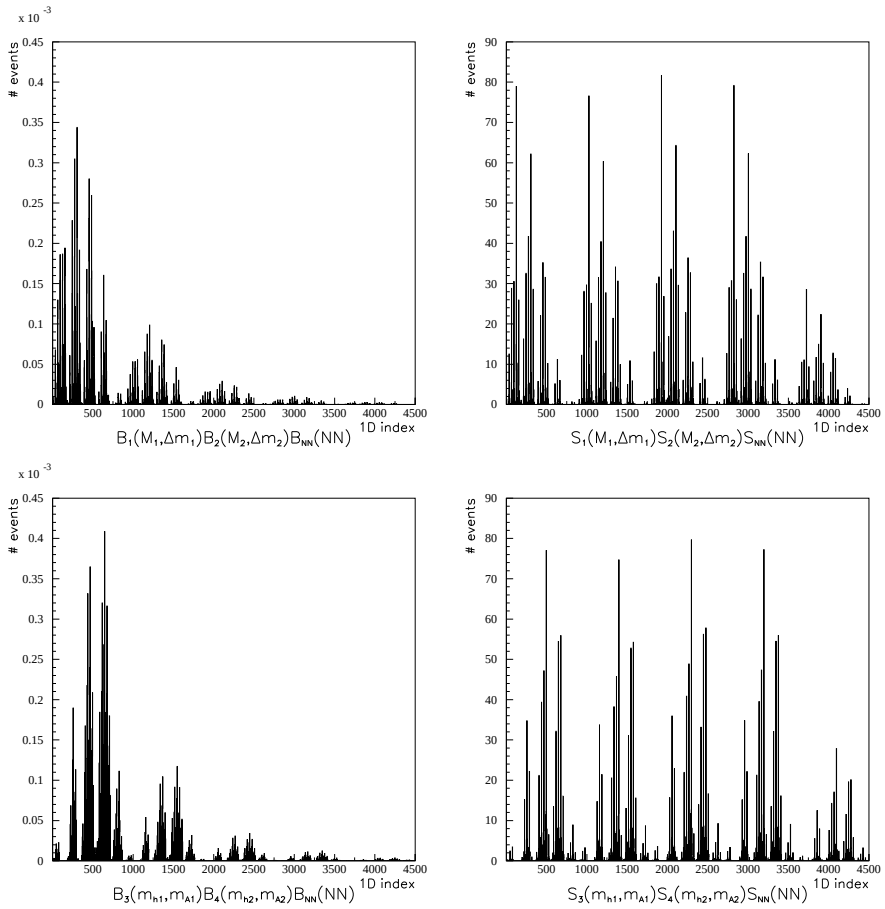

Figure 8.2: Plots of the final discriminating variable distribution of the 5D methods for the background distribution to the left, and the $m_A = 80$, $\tan\beta = 20$ signal distribution to the right. The top plots are the background and signal distributions of the first set of discriminating variables, M_1, Δm_1, M_2, Δm_2, and NN. The bottom plots are the background and signal distributions of the second set of discriminating variables m_{h1}, m_{A1}, m_{h2}, m_{A2}, and NN.

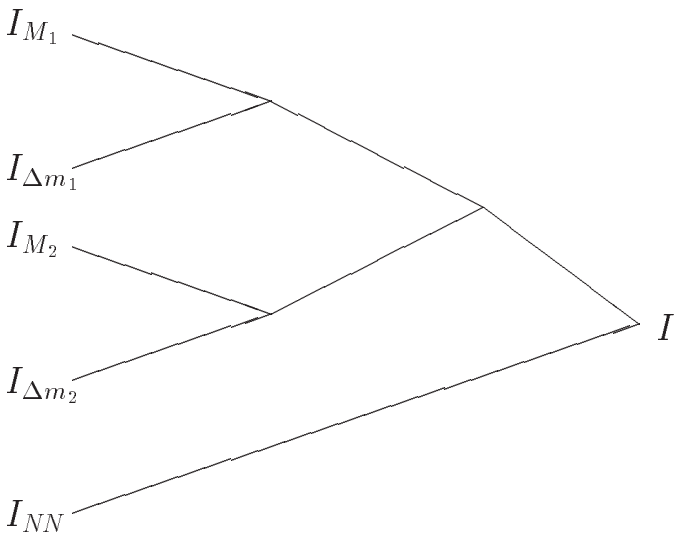

Figure 8.3: Reduction of bin index numbers in the 5D method. I_{M_1}, $I_{\Delta m_1}$, I_{M_2}, $I_{\Delta m_2}$, and I_{NN} are the bin index numbers of each of the five discriminating variables, and I is the bin index number of the final 1D distribution of the discriminating variables.

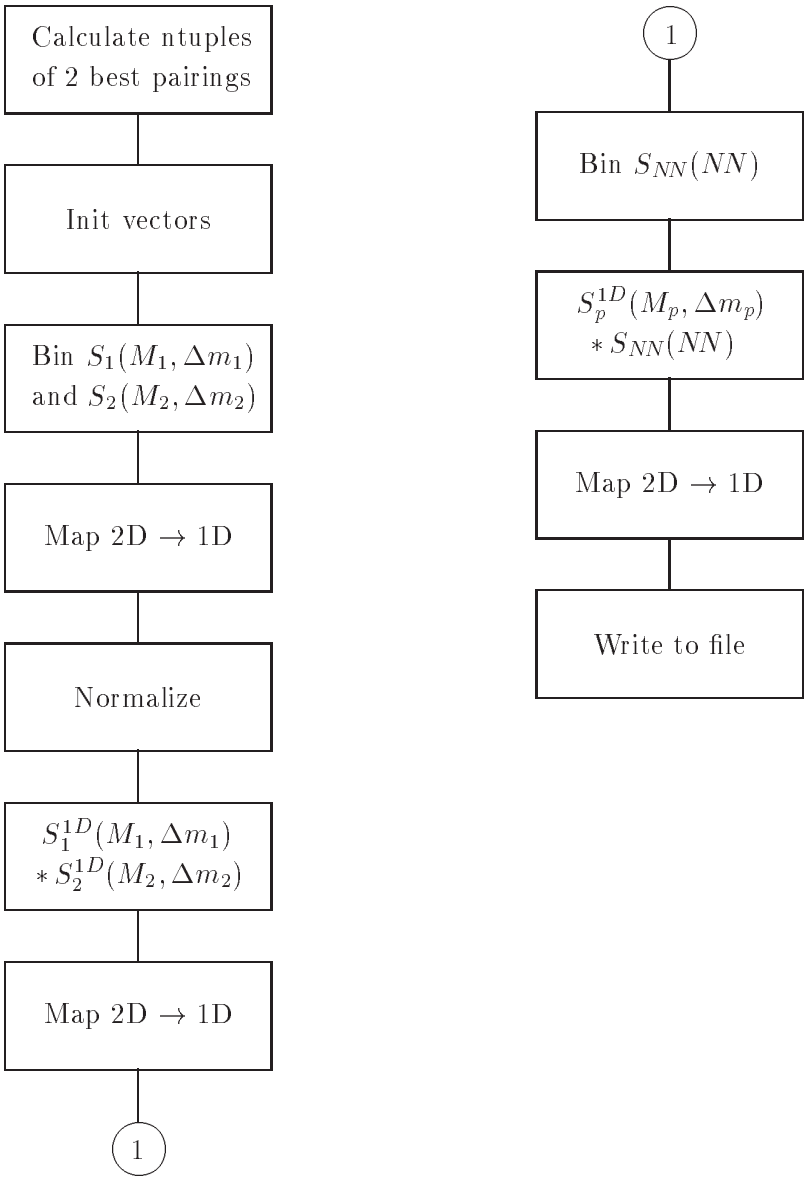

Figure 8.4: Flow chart of the PAW macros preparing the $S_1(M_1, \Delta m_1)$ $S_2(M_2, \Delta m_2) S_{NN}(NN)$ inputs. The structure of the $S_3(m_{h1}, m_{A1})$ $S_4(m_{h2}, m_{A2}) S_{NN}(NN)$ macro is identical.

Chapter 9

The Tests

A total of six likelihood methods are presented in Chapter 7 and Chapter 8: Four 3D methods, and two 5D methods. The 3D methods differ in the choice of binning procedures: Fixed binning of all three distributions, variable binning of two distributions and fixed binning of the third, or variable binning of all three distributions. Only a fixed binning 5D method is implemented.

The fixed binning 3D and 5D methods both use two sets of discriminating variables: The 3D method uses the distribution $S_{1,NN}(\Delta m_1, M_1, NN)$ and $S_{3,NN}(m_{h1}, m_{A1}, NN)$, and the 5D methods use the products of distributions $S_1(\Delta m_1, M_1)S_2(\Delta m_2, M_2)S_{NN}(NN)$ and $S_3(m_{h1}, m_{A1})S_4(m_{h2}, m_{A2})S_{NN}(NN)$. The only difference between the two sets is the choice of mass estimators, $M, \Delta m$ or m_h, m_A.

Two new likelihood methods are presented in this chapter: A 2D method and a 2D+1D method. The 2D method uses the distribution $S_{M_1,NN}(M_1, NN)$ as the discriminating variable, and the 2D+1D method uses the product of distributions $S_1(M_1, \Delta m_1)S_{NN}(NN)$ as the discriminating variable. The 2D+1D method is a simplification of the fixed binning 3D method, and the 2D method is again a simplification of the 2D+1D method.

The 2D and 2D+1D methods are used to investigate features of the 3D and 5D method that a direct comparison do not reveal. The 2D method is compared to the 3D method to investigate the effect of a third discriminating variable, and the 2D+1D method is compared to the 5D method to investigate the effects of using two pairings instead of one pairing. Because the 3D method use the distribution $S_{1,NN}(\Delta m_1, M_1, NN)$ and the 2D+1D method use the product of distributions $S_1(\Delta m_1, M_1)S_{NN}(NN)$, a comparison of the two methods reveals the effects of the correlation between the NN variable and the mass variables. The 5D method[1] does not take these correlations into account either, and hence the 5D method is compared to the 2D+1D method instead of the 3D method to investigate the effects of the second best pairing mass variables. The 2D+1D method is presented in Sec. 9.1, and the 2D method is presented in Sec. 9.2.

All the tests are performed by comparing expected m_h limits calculated using the distributions prepared by the different likelihood methods. Because

[1]The 5D method is actually a 2D+2D+1D method because it uses the product $S_1(\Delta m_1, M_1)S_2(\Delta m_2, M_2)S_{NN}(NN)$ as discriminating variable.

there is a danger of overtraining the confidence levels if the binwidths get too small an overtraining test, as explained in Sec. 7.4, is performed for the fixed binning methods. The variable binning methods are not subjected to an overtraining test due to complexity: If an overtraining test is to be performed, the method under investigation must be executed repeatedly with increasing number of bins if the three $CL_s(\Delta_\theta)$ functions, where Δ_θ is the width of the bin, are to be obtained. The calculation of $CL_s(\Delta_\theta)$ is considerably more complicated with variable binning than with fixed binning. Overtraining tests are discussed in further detail in Sec. 9.3.

The final tests of the total of nine combinations of methods and discriminating variables should answer three questions :

- What type of binning provides the optimal limit?
- What kind of mass estimator(s) provides the optimal limit?
- What number of jet pairings provides the optimal limit?

The tests focused on these questions are presented in Sec. 9.4, and the conclusions of the tests are presented in Chapter 10.

9.1 A 2D+1D Likelihood Method

The fixed binning 3D method, using a 3D distribution of $(M_1, \Delta m_1, NN)$ events as the discriminating variable distribution, is transformed into a fixed binning 2D+1D method by separating the 3D distribution into a 2D distribution of $(M_1, \Delta m_1)$ events and a 1D distribution of (NN) events.

The binning of the 2D and the 1D distribution are performed using (7.1), just as in the 3D case. The 2D vector containing the two mass distributions is mapped into 1D space using (8.3). The resulting 1D vector is normalized to one before it is multiplied with the 1D vector containing the normalized NN distribution to give the product $S_1(M_1, \Delta m_1)S_{NN}(NN)$ stored in a 2D vector. The 2D bin index numbers of the 2D vector are mapped into the 1D space of the final variable using (8.3) once again for each 2D bin. Because the $S_1(M_1, \Delta m_1)$ and $S_{NN}(NN)$ distributions are multiplied, the correlation between the NN variable and the mass variables is not taken into account.

The candidates are mapped into the final 1D variable space by applying (8.3) in the same manner as for the distributions. The only difference is that the final 1D bin do not contain the number of candidates, as is the case for the distributions, but instead a number that identifies which candidates are confined to the bin.

9.2 A 2D Likelihood Method

The 2D likelihood method is a simplification of the 2D+1D likelihood method: The $S_1(M_1, \Delta m_1)S_{NN}(NN)$ distribution product of the 2D+1D method are replaced by a $S_{M_1,NN}(M_1, NN)$ distribution which is binned using (7.1) twice for each (M_1, NN) event, and mapped using (8.3) for each 2D bin.

To map the candidates it suffices to apply (8.3) once for each candidate.

9.3 Overtraining Tests

As explained in Sec. 7.4 an overtraining test gives the limit on how small the bins used in the binning of a distribution can get before an overtraining of the calculated confidence levels occur. In other words the plot of the three $CL_s(\Delta_\theta)$ functions at different numbers of MC events expresses the potential of the method: How low signal confidence levels (and accordingly high m_h limits) the method is able to produce before overtraining effects influence the limits.

Overtraining tests are performed for the 5D method, the fixed binning 3D method, the 2D+1D method, and the 2D method. The tests monitor the development of the CL_s as a function of decreasing Δ_θ, and Δ_θ is determined by the width of the distribution and the number of bins in the following manner:

$$\Delta_\theta = \frac{\theta_{max} - \theta_{min}}{N_\theta} \ , \tag{9.1}$$

where θ_{max} and θ_{min} are the maximum and minimum value of θ, respectively, and N_θ is the number of bins in the θ distribution.

To give the mass variables equal weight the number of bins is selected to keep the binwidth of all the mass distributions equal. The width of the m_h and m_A distribution are the same, hence $N_{m_h} = N_{m_A}$. However, the difference between the widths of the M and Δm distributions is approximately 30 GeV, as is seen in Fig. 8.1, hence $N_M > N_{\Delta m}$.

The overtraining plots of the different methods, with cuts on the binwidths, are given in Chapter 10. The CL_s at the Δ_θ cut is the minimum CL_s (maximum m_h) not influenced by overtraining effects.

9.4 The Final Tests

A plot of expected signal confidence level versus m_h hypothesis, $CL_s(m_h)$, is the final test of the likelihood methods. A significance level of 95 % is required, and a cut is set at $CL_s = 0.05$ (see (6.6)). The m_h limit corresponding to the cut is the upper limit of the mass region where the existence of the h^0 scalar.

Three series of tests are performed in order to answer the questions regarding type of binning, type of mass estimator, and number of pairings:

Type of Binning The three 3D methods presented in Chapter 7, all using $S_{3,NN}(m_{h1}, m_{A1}, NN)$ as the discriminating variable distribution, are compared to conclude on which of the methods provides the most effective binning. No overtraining test of the three methods are performed because the implementation of such a test using variable binning methods is very complex. The results of the binning tests are presented in Sec. 10.2.1.

Type of Mass Estimator To investigate the use of mass estimators in detail, three mass estimator tests are performed: Firstly the fixed binning 3D method using $S_{3,NN}(m_{h1}, m_{A1}, NN)$ as discriminating variable

is compared to the 2D method using $S_{M_1,NN}(M_1, NN)$ as discriminating variable. Secondly the 3D method using $S_{3,NN}(m_{h1}, m_{A1}, NN)$ is again compared to the 3D method using $S_{1,NN}(M_1, \Delta m_1, NN)$ as the discriminating variable. The third comparison is performed between the 5D method using $S_3(m_{h1}, m_{A1})S_4(m_{h2}, m_{A2})S_{NN}(NN)$ and the 5D method using $S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)S_{NN}(NN)$ as the discriminating variable. All three comparisons are performed in order to determine if the $M, \Delta m$ set or the m_h, m_A set of discriminating variables is the best. All methods are fixed binning methods. Overtraining tests is used to set a limit on the binwidths. The results of the overtraining tests are presented in Sec. 10.1, and the results of the mass estimator tests are presented in Sec. 10.2.2.

Number of Pairings Two tests are performed: The 5D method using the product $S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)S_{NN}(NN)$ as discriminating variable is compared to the 2D+1D method using the product $S_1(M_1, \Delta m_1)S_{NN}(NN)$ as discriminating variable, to conclude on the effects of using mass estimators of the two best pairings as discriminating variables. The 5D method does not take the correlations between the masses of different pairings nor between the mass variables and the NN variable into account. A comparison of the fixed binning 3D method using $S_{1,NN}(M_1, \Delta m_1, NN)$ and the 2D+1D method using $S_1(M_1, \Delta m_1)S_{NN}(NN)$ as discriminating variable is performed to investigate the importance of the NN correlations. The correlations between the mass estimators of different pairings are not investigated any further, but there is no reason why such correlations should not be present.

Chapter 10

Conclusions

Signal and background distributions are created using the likelihood methods presented in Chapter 7, Chapter 8 and Chapter 9. Signal confidence levels (CL_s) based on these distributions are calculated, using the ALRMC implementation of a maximum likelihood ratio test, and the m_h limit at a 95 % confidence level (CL) is stated ($CL_s = 0.05$).

The conclusions are based on *expected* CL_s instead of observed CL_s because the observed CL_s are sensitive to the observed candidates of the search, i.e. the position of a candidate in the final distribution might have a major effect on the limits. Such an effect is not present if the expected CL_s are used.

10.1 The Overtraining Plots

Fig. 10.1 shows overtraining plots for the 2D, 2D+1D, 3D and 5D methods, all with fixed binning. The 3D and 5D methods are used with both sets of discriminating variables, $M, \Delta m$ and m_h, m_A.

The vertical line in the plots of Fig. 10.1 is the cut on the binwidth that sets an upper limit on number of bins allowed in the distributions. The tick marks on the $CL_s(\Delta_\theta)$ function of all Monte Carlo (MC) events (the green line) is the CL_s computed for a given number of bins for the $m_A = 80$ GeV, $\tan\beta = 20$ signal hypothesis. The binwidth Δ_θ is the average binwidth of the bins used.

No cut is applied to the 5D method in neither set of discriminating variables because the ALRMC limit on maximum number of allowed bins in the final 1D distribution is reached before the overtraining occurs.

10.2 The Final Tests

Expected CL_s are calculated at the binsize cut provided by the overtraining plots in Fig.10.1. The nine different combinations of methods, binning, and discriminating variables are summarized in Table 10.1 together with the m_h limits at $\tan\beta = 20$ and $\tan\beta = 2$. The nine combinations are further compared two by two, and conclusions drawn on which method offers the best limit regarding type of binning, type of mass estimators, and number of pairings. All limits presented correspond to a 95 % CL.

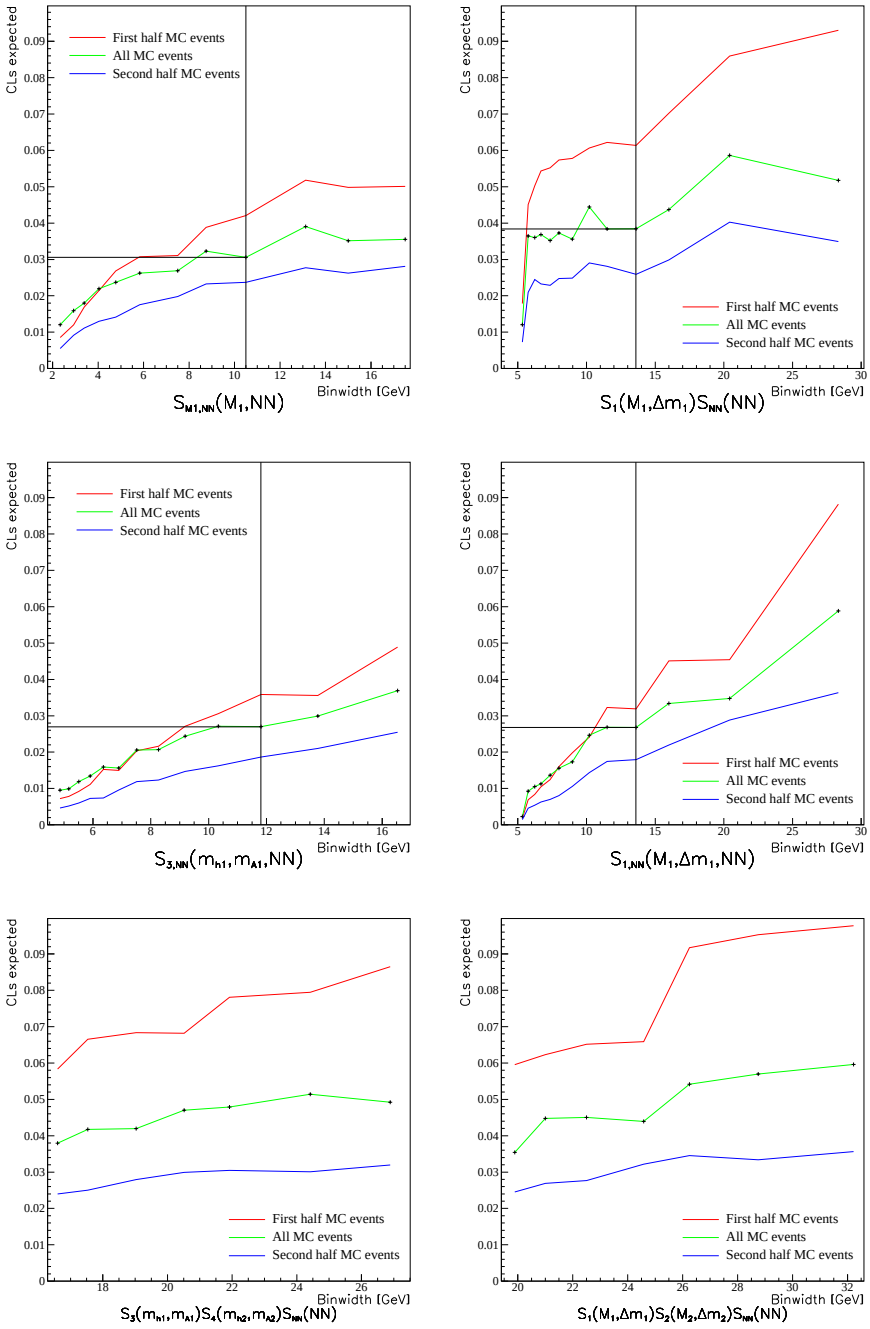

Figure 10.1: Overtraining plots for all fixed binning methods. The CL_s of all MC events (green line) is situated in the middle between the CL_s of the first and second half MC events if no overtraining is present. The vertical line in the plot is a cut on the binsize.

Method	Binning	Discriminating variables	m_h [GeV]	
			tg$\beta = 20$	tg$\beta = 2$
2D	Fixed	$S_{M_1,NN}(M_1, NN)$	80.85	57.59
2D+1D	Fixed	$S_1(M_1, \Delta m_1)S_{NN}(NN)$	80.25	57.43
3D	Fixed	$S_{3,NN}(m_{h1}, m_{A1}, NN)$	81.10	57.57
		$S_{1,NN}(M_1, \Delta m_1, NN)$	81.23	57.85
		$S_{3,NN}(m_{h1}, m_{A1}, NN)$	$81.57^{\dagger}$	$58.37^{\dagger}$
	Fix/Var	$S_{3,NN}(m_{h1}, m_{A1}, NN)$	$81.85^{\dagger}$	$58.70^{\dagger}$
	Variable	$S_{3,NN}(m_{h1}, m_{A1}, NN)$	$81.73^{\dagger}$	$58.40^{\dagger}$
5D	Fixed	$S_3(m_{h1}, m_{A1})S_4(m_{h2}, m_{A2})S_{NN}(NN)$	80.37	57.37
		$S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)S_{NN}(NN)$	80.46	57.54

† Not subjected to overtraining test, and not comparable to the other limits

Table 10.1: Summary table of the m_h limits of all nine combinations of method, binning procedure and discriminating variables.

10.2.1 Binning

Fig. 10.2 shows two plots of $CL_s(m_h)$ functions at $\tan\beta = 20$:

- Fixed binning of m_{h1}, m_{A1}, NN versus fixed binning of NN and variable binning of m_{h1}, m_{A1}.
- Fixed binning of NN and variable binning of m_{h1}, m_{A1} versus variable binning of m_{h1}, m_{A1}, NN.

Fig. 10.3 shows the same plots of $CL_s(m_h)$ functions at $\tan\beta = 2$, and Table 10.2 gives the measured differences of the m_h limits.

The conclusion drawn is that the method using both variable and fixed bins improves the fixed binning only method limit by 0.28 GeV (0.33 GeV), and the variable binning only method limit by 0.12 GeV (0.30 GeV) at $\tan\beta = 20$ ($\tan\beta = 2$). Hence, the method using both variable and fixed bins offers the most optimal binning of the $S_{3,NN}(m_{h1}, m_{A1}, NN)$ distribution.

Since the limits of these three methods have not been subjected to an overtraining test, the limits are not comparable to the other limits.

10.2.2 Mass Estimator

Fig. 10.4 shows three plots of $CL_s(m_h)$ functions at $\tan\beta = 20$:

- $S_{M_1,NN}(M_1, NN)$ versus $S_{3,NN}(m_{h1}, m_{A1}, NN)$.
- $S_{3,NN}(m_{h1}, m_{A1}, NN)$ versus $S_{1,NN}(M_1, \Delta m_1, NN)$.
- $S_3(m_{h1}, m_{A1})S_4(m_{h2}, m_{A2})S_{NN}(NN)$ versus $S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)S_{NN}(NN)$.

Fig. 10.5 shows the same plots of $CL_s(m_h)$ functions at $\tan\beta = 2$. The difference between the limits of these plots are given in Table 10.3.

Two conclusions are drawn from these tests: Firstly the $M, \Delta m$ set of discriminating variables improves the limit obtained using the m_h, m_A sets by 0.13 GeV (0.28 GeV) in the 3D case, and by 0.11 GeV (0.17 GeV) in the 5D case at $\tan\beta = 20$ ($\tan\beta = 2$). This is probably due to the fact that the events of the $M, \Delta m$ distributions are less clustered than the events of the m_h, m_A distribution, hence the $M, \Delta m$ set is more convenient to use with fixed bins.

The second conclusion is that $S_{1,NN}(M_1, \Delta m_1, NN)$ improves the limit obtained using $S_{M_1,NN}(M_1, NN)$ by 0.38 GeV (0.30 GeV), and in addition $S_{3,NN}(m_{h1}, m_{A1}, NN)$ also improves the limit obtained using $S_{M_1,NN}(M_1, NN)$ by 0.25 GeV (0.02 GeV) at $\tan\beta = 20$ ($\tan\beta = 2$). Hence, a 3D discriminating variable gives a better limit than a 2D discriminating variable. The improvement is however not impressive, and it is necessary to ask if not alternative extensions of a 2D likelihood method could result in larger improvements.

10.2.3 Pairing

Fig. 10.6 shows two plots of $CL_s(m_h)$ functions at $\tan\beta = 20$:

- $S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)S_{NN}(NN)$ versus $S_1(M_1, \Delta m_1)S_{NN}(NN)$.
- $S_1(M_1, \Delta m_1)S_{NN}(NN)$ versus $S_{1,NN}(M_1, \Delta m_1, NN)$.

Fig. 10.7 shows the same plots of $CL_s(m_h)$ functions at $\tan\beta = 2$. The difference between the m_h limits of the tests are given in Table 10.4.

Two conclusions are drawn from the tests: Firstly the two pairing variable $S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)S_{NN}(NN)$ improves the limit obtained using the $S_1(M_1, \Delta m_1)S_{NN}(NN)$ variable by 0.21 GeV (0.11 GeV) at $\tan\beta = 20$ ($\tan\beta = 2$). Hence, the two best pairing variable gives a slightly better m_h limit than the best pairing variable. The effect is not so clear at $\tan\beta = 2$. However the most interesting sector for the h^0A^0 channel is the $\tan\beta = 20$ sector because the h^0Z^0 cross section is much greater than the h^0A^0 cross section at $\tan\beta = 2$ (see Fig.4.2).

The second conclusion is that the $S_{1,NN}(M_1, \Delta m_1, NN)$ variable improves the limit obtained using the $S_1(M_1, \Delta m_1)S_{NN}(NN)$ variable by 0.98 Gev at $\tan\beta = 20$ and 0.36 Gev at $\tan\beta = 2$. Hence, the NN correlation is very significant to the m_h limit, and is at least some of the explanation to why the 3D method produce a better limit than the 5D method (see Table 10.1). This is one of the most important results obtained. If the $S_1(M_1, \Delta m_1)S_2(M_2, \Delta m_2)$ $S_{NN}(NN)$ variable is replaced by $S(M_1, \Delta m_1, M_2, \Delta m_2, NN)$ the m_h limit probably increases beyond the limit computed using $S_{1,NN}(M_1, \Delta m_1, NN)$.

10.3 Final Conclusion

Only minor improvements are obtained by extending the likelihood method from two discriminating variables to three and five discriminating variables, and the extension of the methods were probably not worth the effort.

The most important results seen are the the positive effect of using variable bins instead of fixed bins when the distributions are binned (see Fig. 10.2, 10.3 and Tab. 10.2), and especially the clear positive effect of including correlations between the discriminating variables (see Fig. 10.6, 10.7 and Tab. 10.4).

Test	Δm_h[GeV]	
	tg$\beta = 20$	tg$\beta = 2$
$m_h(S_{3,NN}^{fix/var}) - m_h(S_{3,NN}^{fix})$	0.28	0.33
$m_h(S_{3,NN}^{fix/var}) - m_h(S_{3,NN}^{var})$	0.12	0.30

Table 10.2: Difference in m_h limit between the three 3D methods using $S_{3,NN}(m_{h1}, m_{A1}, NN)$ as discriminating variable.

Test	Δm_h[GeV]	
	tg$\beta = 20$	tg$\beta = 2$
$m_h(S_{3,NN}) - m_h(S_{M_1,NN})$	0.25	0.02
$m_h(S_{1,NN}) - m_h(S_{3,NN})$	0.13	0.28
$m_h(S_1, S_2, S_{NN}) - m_h(S_3, S_4, S_{NN})$	0.11	0.17

Table 10.3: Difference in m_h limits between the 3D, 2D and 5D methods using the $M, \Delta m$ and m_h, m_A sets of discriminating variables.

Test	Δm_h[GeV]	
	tg$\beta = 20$	tg$\beta = 2$
$m_h(S_1, S_2, S_{NN}) - m_h(S_1, S_{NN})$	0.21	0.11
$m_h(S_{M_1,NN}) - m_h(S_1, S_{NN})$	0.98	0.36

Table 10.4: Difference in m_h limits between the 5D, 3D and 2D+1D method using the $M, \Delta m$ sets of discriminating variables.

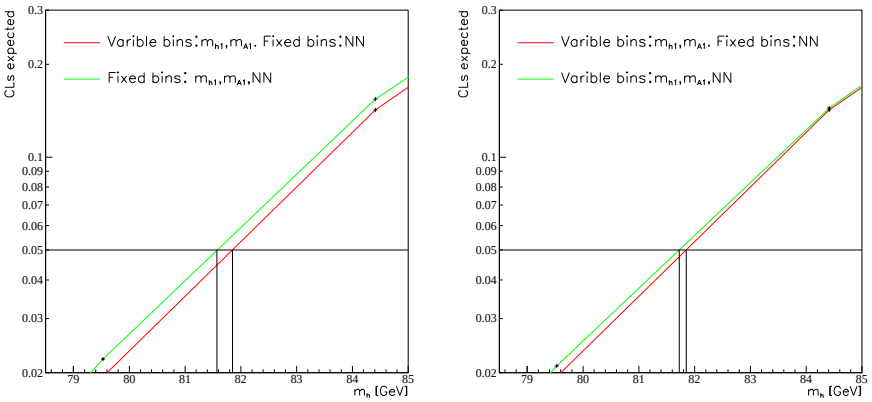

Figure 10.2: Comparisons of the two variable binning 3D methods and the fixed binning 3D method for the $\tan\beta = 20$ signal hypotheses. From these plots it is concluded that the variable/fixed binning 3D method improves the limits obtained using the other 3D methods.

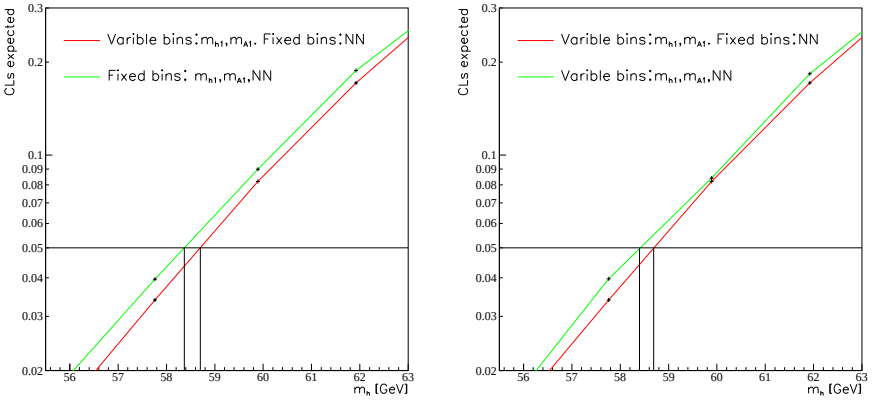

Figure 10.3: Comparisons of the two variable binning 3D methods and the fixed binning 3D method for the $\tan\beta = 2$ signal hypotheses. These plots support the conclusion from the $\tan\beta = 20$ plots: The variable/fixed binning 3D method improves the limits obtained using the other 3D methods. The difference between the variable/fixed method is slightly larger in the $\tan\beta = 2$ case than in the $\tan\beta = 20$ case.

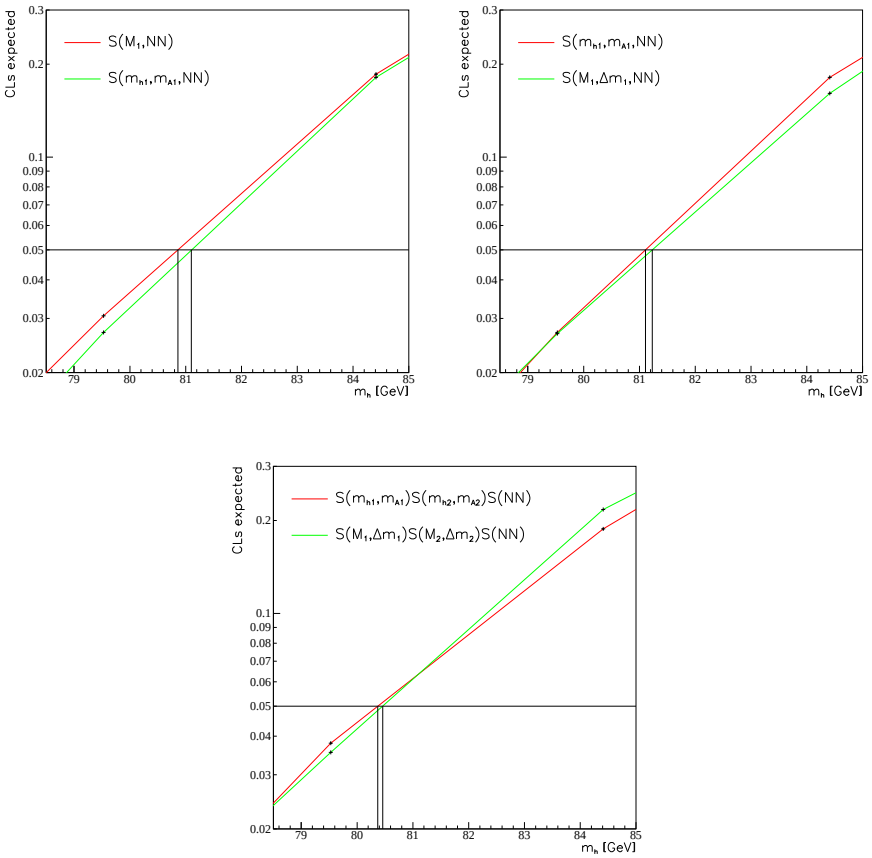

Figure 10.4: Comparisons of the 2D, 3D and 5D method in the $\tan\beta = 20$ sector using the $M, \Delta m$ and m_h, m_A sets of discriminating variables. The first conclusion is that the $M, \Delta m$ set is preferred when fixed binning is applied. The second conclusion is that the 3D method is preferred to both the 2D and 5D method.

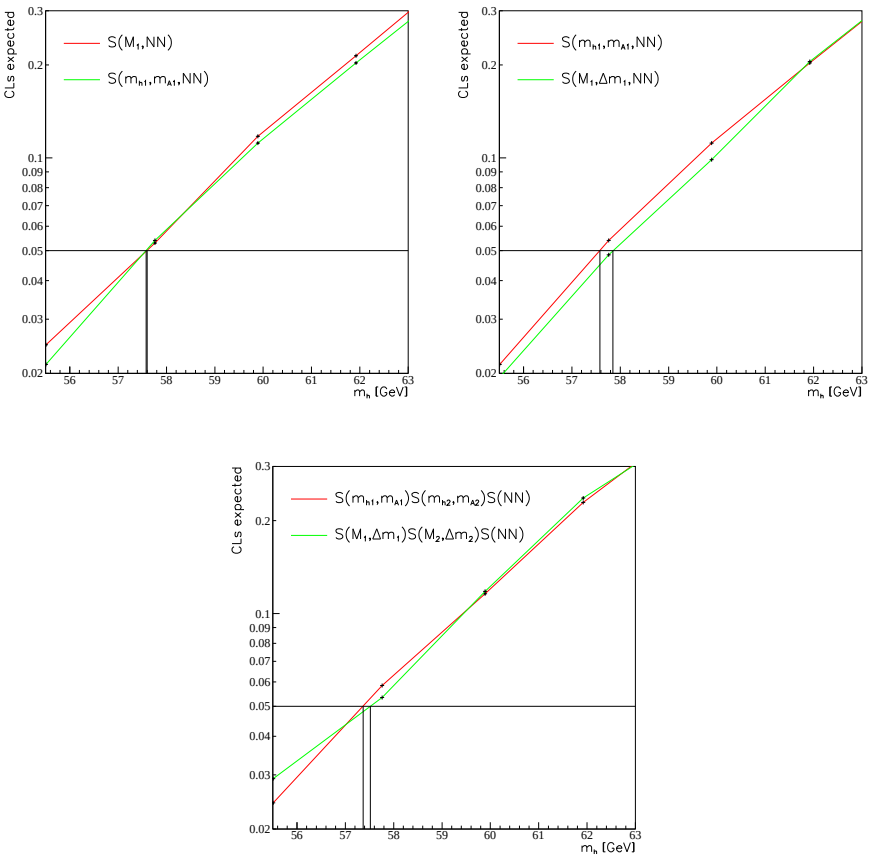

Figure 10.5: Comparisons of the 2D, 3D and 5D method in the $\tan\beta = 2$ sector using the $M, \Delta m$ and m_h, m_A sets of discriminating variables. The same conclusions is drawn as in the $\tan\beta = 20$ case: The $M, \Delta m$ set is preferred when fixed binning is applied, and the 3D method is preferred to both the 2D and 5D method. The limits provided by the m_h, m_A sets are lower in the $\tan\beta = 2$ case, but the limits provided by the $M, \Delta m$ sets are better in the $\tan\beta = 2$ case. The result is that the difference between the limits of $S_{1,NN}(M_1, \Delta m_1, NN)$ and $S_{M_1,NN}(M_1, NN)$ is almost the same in both $\tan\beta$ sectors.

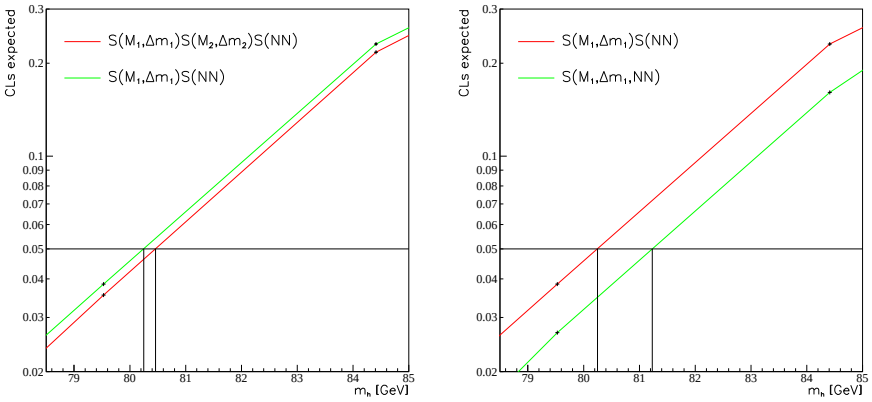

Figure 10.6: Plot of the 2D+1D versus the 5D method, both using the $M, \Delta m$ set of discriminating variables, for the $\tan\beta = 20$ signal hypotheses. The plot shows that the two pairing method gives the best m_h limit. The plot to the right shows the 3D versus the 2D+1D method and is performed to determine the significance of the correlation between the NN variable and the mass variables for the $\tan\beta = 20$ signal hypotheses. The plot shows that the limit is reduced by nearly 1 GeV if the NN correlations are not included in the discriminating variable.

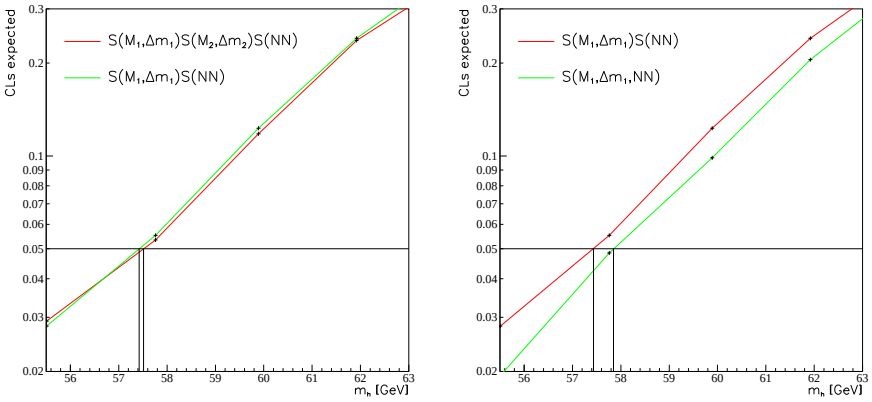

Figure 10.7: Plot of the 2D+1D versus the 5D method to the left, both using the $M, \Delta m$ set of discriminating variables, for the $\tan\beta = 2$ signal hypotheses. The plot shows that the two pairing method gives a better m_h limit than the one pairing method. The plot to the right shows the 3D versus the 2D+1D method and is performed to determine the significance of the correlation between the NN variable and the mass variables for the $\tan\beta = 2$ signal hypotheses. The plot shows that the limit is reduced by about 0.4 GeV if the NN correlations are not accounted for. In contrast, the reduction in the $\tan\beta = 20$ m_h limit is nearly 1 GeV.

Bibliography

[1] Contemporary Physics Education Project.
Standard Model of Fundamental Particles and Interactions.
Chart of tables, texts and pictures, 1995.

[2] André Sopczak.
Search for Non-Minimal Higgs Bosons in Z^0 Decays with the L3 Detector at LEP.
1992.

[3] John F. Donoghue, Eugene Golowich, and Barry R. Holstein.
Dynamics of the Standard Model.
Cambridge University Press, 1992.

[4] Francis Halzen and Alan D. Martin.
Quarks & Leptons. An introductory course in modern particle physics.
John Wiley & Sons, Inc., 1984.

[5] G. G. Ross.
A 'theory of everything'?
Contemporary Physics, 34(2):79–88, 1993.

[6] John F. Gunion, Howard E. Haber, Gordon Kane, and Sally Dawson.
The Higgs Hunter's Guide.
Addison-Wesley Publishing Company, Inc., 1990.

[7] Donald H. Perkins.
Introduction to High Energy Physics.
Addison-Wesley Publishing Company, Inc., 1987.

[8] U. Amaldi, W. Deboer, and H. Fürstenau.
Comparison of Grand Unified Theories with Electroweak and Strong Coupling-Constants Measured at LEP.
Physics Letters **B260**, (3-4):447–455, May 1991.

[9] G. Altarelli, T. Sjöstrand, and F. Zwirner, editors.
Physics at LEP2, Vol.1.
CERN 96-01.

[10] Philippe Charpentier.
The DELPHI detector at CERN's LEP collider.
`http://delphiwww.cern.ch/~offline/physics/delphi-detector.html`.

[11] Celso Martinez-Rivero.
Tables of h^0A^0 and h^0Z^0 cross-sections and branching ratios.
`/afs/cern.ch/delphi/tasks/lep200/searches/limits/higgs/MSSM/newbench175/189_no.dat`.

[12] Jørgen Hansen.
Private communications.
Physics Department, University of Oslo.

[13] P. Lutz.
DELPHI TEAM C standard track selection.
`http://delphiwww.cern.ch/~lutzp/delwww/tasks/lep200/searches/teamc/www/teamc_dq_particle_selection.html`.

[14] Jørgen Hansen.
Search for Pair Produced Neutral MSSM Higgs Bosons in the 4b Channel Using a Repeated 2D Likelihood Method.
DELPHI 00-xx PHYS xxx *in preparation.*

[15] P. Lutz.
DELPHI TEAM C 4jet selection.
`http://delphiwww.cern.ch/~lutzp/delwww/tasks/lep200/searches/teamc/www/news/4jsel_proposal97.html`.

[16] P. Lutz.
Preselection Performances for Processing D2.
`http://delphiwww.cern.ch/~lutzp/delwww/presel/presel_189D.html`.

[17] S. Catani, Yu. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber.
New Clustering-Algorithm for Multijet Cross-Sections in e^+e^- Annihilation.
Physics Letters **B269**, pages 432–438, 1991.

[18] Torbjörn Sjöstrand.
High-Energy-Physics Event Generation with PYTHIA 5.7 and JETSET 7.4.
Computer Physics Communications Journal, **82** (1994) 74.

[19] Louis Lyons.
Statistics for Nuclear and Particle Physicists.
Cambridge University Press, 1992.

[20] A. G. Frodesen, O. Skjeggestad, and H. Tøfte.
Probability and Statistics in Particle Physics.
Universitetsforlaget, 1979.

[21] Alex L. Read.
Optimal Statistical Analysis of Search Results Based on the Likelihood Ratio and its Application to the Search for the MSM Higgs Boson at $\sqrt{s} = 161$ *and* 172 *Gev.*
DELPHI 97-158 PHYS 737.

[22] A. G. Frodesen, O. Skjeggestad, and H. Tøfte.
Probability and Statistics in Particle Physics, pages 384–385.
Universitetsforlaget, 1979.

[23] V. Ruhlmann-Kleider.
Code for PAW macro `fiteff.kumac`.
`/afs/cern.ch/user/r/ruhlmann/public/limits/higgs/fitsig/fiteff.kumac`, 1997.