Complete kinematics reconstruction in the golden cascade

by

Jørgen Eriksson Midtbø

Thesis

for the degree of

Master of Science

Faculty of Mathematics and Natural Sciences
University of Oslo

June 2015

Samandrag

Vi undersøker ein metode, opphavleg foreslått i [1], for å bestemme massane til resonansar og usynlege sluttilstandar i kaskadehenfall, i teoriar utanfor Standardmodellen med ein paritetssymmetri. Vi gjer undersøkinga vår innanfor rammeverket til den Minimale Supersymmetriske Standardmodellen med bevaring av R-paritet. Metoden formulerer massebestemminga som eit optimeringsproblem som kan løysast med numeriske verktøy. I [1] vart det demonstrert med Monte Carlo-simuleringar at metoden har gode utsikter for nøyaktig massebestemming. Vi finn at det er visse problem med metoden og analysen, som leiar til underestimering av feilen. I lys av dette presenterer vi reviderte feilestimat. Vi held fram med å undersøke alternative formuleringar av metoden som kanskje kan løyse problema, og presenterer masseestimat og usikkerheiter for Monte Carloanalysar av desse alternative formuleringane. Vi undersøker også effektane av jet-rekonstruksjon, og viser at dette er ei viktig kjelde til feil for metoden.

Abstract

We investigate a method, originally suggested in [1], for determination of the masses of resonances and invisible final states in cascade decays, in Beyond-Standard-Model theories with a parity symmetry. We do our study within the framework of the Minimal Supersymmetric Standard Model with conservation of R-parity. The method formulates the mass determination as an optimization problem for which numerical techniques can be applied. The method was demonstrated in [1] by Monte Carlo simulations to have good prospects for accurate mass determination. We discover that there are certain problems with the method and the analysis, which lead to an under-estimation of the errors. In light of this, we present revised error estimates. We proceed to investigate alternative formulations of the method which might amend the issues, and present mass estimates and uncertainties for Monte Carlo analyses of these alternative formulations. We also investigate the effects of jet reconstruction, and show that this is an important source of error for the method.

Til mamma

Takk

Takk til Are Raklev for god rettleiing. Takk til Lars Andreas Dal og Anders Kvellestad for uvurderleg hjelp, og til Anders Lauvland og Anders Hafreager for gode diskusjonar. Takk til Janne-Kristin Svarstad Nygård for at eg får dele kvardagane med deg. Takk til pappa for at du alltid stiller opp. Takk også til Mona Semb, Bryan Webber, Tor Gjerrestad, Arild Gaasdal, Bjørn Hallvard Samset, Fysikkforeningen, Fysisk fagutval og miljøet på Lillefy, venner og familie.

Contents

List of Figures

List of Tables

Introduction

Supersymmetry is a proposed extension of the Standard Model of particle physics which adds heavier partner particles, called superpartners, to the Standard Model particles. Supersymmetry is appealing because it can provide solutions to many of the unsolved problems of high-energy physics, problems which are either not adressed by the Standard Model or are inherent in the Standard Model itself. These problems include the Dark Matter which is observed to permeate the universe and the apparent fine-tuning needed to avoid large corrections to the Higgs boson mass.

This thesis begins with an introduction to the principles that underlie the Standard Model, and outlines the construction of the model in Chapter 1. Then, the extension of the Standard Model by supersymmetry is introduced and elaborated, culminating in the definition of the Minimal Supersymmetric Standard Model (MSSM), which is the minimal (in terms of field content) supersymmetric extension that contains the Standard Model. The MSSM has over 100 free parameters, so it is conventional to introduce some constraining assumptions. We present one popular restricted model, called the Constrained MSSM (CMSSM), and discuss its features and review its experimental status in light of LHC data. All of this takes place in Chapter 2.

In a scenario where supersymmetry (SUSY) exists with conservation of R-parity, there should be a neutral lightest supersymmetric particle (LSP) which is absolutely stable. Additionally, all SUSY-particles will be produced in pairs in collisions. Any supersymmetric particle will decay down to the LSP, possibly in multiple steps. At the Large Hadron Collider (LHC) at CERN, one typically expects squarks or gluinos, the supersymmetric partners of the quarks and gluons, which subsequently decay in a cascade down to the LSP, emitting multiple Standard Model particles in the process, see *e.g.* [2]. The LSP escapes detection, but the energy and momentum of all the SM particles can be measured (assuming no neutrinos are produced).

Should hints of SUSY be seen at the LHC in the future, it will be important to determine the masses of the SUSY particles, and many methods have been investigated for this purpose, *e.g.* see [3] for a review. In [1] a method for determining the masses by formulating the kinematics as a linear algebra problem is presented and discussed. The article specifically considers a pair of SUSY parti-

cles, each decaying to the LSP in a three-step cascade, emitting three Standard Model particles each. Determination of the masses then involve inverting an eight by eight matrix containing the measured Standard Model particle momenta, using it to determine the four-momenta, p^{μ}_{inv}, of the two invisible LSPs. Having calculated the invisible four-momenta, one can make a hypothesis for the eight unknown SUSY masses involved in the chains, and minimize $p^2_{\text{inv}} - M^2_{\text{hyp}}$ in order to obtain a best fit hypothesis (the minimized quantity should be zero for an on-shell particle). The article further demonstrates by a Monte Carlo simulation that a good fit can be obtained using only a few tens of events. This is a very desirable feature, since only a very limited number of SUSY candidate events are expected at the LHC, given current limits on SUSY particle masses. Being able to reconstruct the four-momenta of the invisibles is also useful for measuring spins.

In Chapter 3 and 4, the method of [1] is presented and discussed. Chapter 3 formulates the type of process we are studying and the problems faced, and defines the method. We also discuss some minor issues with the original formulation and suggest ways to amend these. The subsequent chapters deal with an investigation of, and improvements on the method. We begin in Chapter 4 by simulating events using the Monte Carlo generator `Herwig++` and attempt to reproduce the results presented in [1]. In the course of this, we discover what appears to be problematic issues with the analysis in [1], concerned with the numerical function minimization. We are able to reproduce our results using the same code as in the original article [4], and this confirms the issues.

In light of this, we present revised estimates for the accuracy we can expect from the method. In Chapter 5, we investigate ways to amend the problems by modifying the method in various ways. Although we are able to construct modified versions of the method which is free of the technical problems in the original formulation, we are unable to recover the accuracy promised by the original paper.

We finalize our analysis in Chapter 5 by taking the effects of jet reconstruction into account, something which was not considered in the original paper. We find that the uncertainties introduced by the jet reconstruction increases the errors significantly. Finally, we make our conclusions.

Chapter 1

The Standard Model of Particle Physics

The Standard Model of particle physics has been hugely successful in explaining what our universe consists of at the smallest length scales, and how these constituents interact with each other. It recieved a final, spectacular confirmation in 2012, when a Higgs boson consistent with the predictions of the Standard Model was discovered by the CMS and ATLAS experiments at CERN [5, 6]. It is well known, however, that the Standard Model is incomplete as a description of our universe, for instance since it gives no explanation for Dark Matter. There are also more technical problems with the Standard Model, such as the hierarchy problem of the Higgs boson mass loop corrections and the arbitrariness of the model parameters.

The present chapter gives a brief introduction to the principles that underlie the Standard Model, and outlines the construction of the model.

1.1 Symmetries and conservation laws

Symmetries are manifest in most physical systems. For instance, the special theory of relativity is symmetric under boosts and rotations, as well as translations in space and time. There is a deep relationship between symmetries and the conservation of physical quantities. This result is known as Noether's theorem, and was proven by Emmy Noether in 1915 [7]. It states that *every differentiable symmetry of the action of a physical system has a corresponding conservation law*. In the example of special relativity, the symmetries under translations in time and space correspond to conservation of energy and momentum.

1.1.1 Description by groups

It is often convenient to describe the symmetries of physical systems in the language of group theory. A group G is a set of objects together with a binary operation $\cdot$ which satisfies the following properties:

$$g_1 \cdot g_2 \in G \quad \forall g_1, g_2 \in G \qquad \text{(closure)}, \quad (1.1)$$

$$\exists\, e \in G \quad \text{such that} \quad g \cdot e = e \cdot g = g \quad \forall\, g \in G \qquad \text{(identity)}, \quad (1.2)$$

$$\forall\, g \in G \quad \exists\, g^{-1} \in G \quad \text{such that} \quad g \cdot g^{-1} = g^{-1} \cdot g = e \qquad \text{(inverse)}, \quad (1.3)$$

$$(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3) \quad \forall\, g_1, g_2, g_3 \in G \qquad \text{(associativity)}. \quad (1.4)$$

The set of all Lorentz boosts and rotations in special relativity form a group, called the Lorentz group, and together with all spatial translations they form the Poincaré group.

The experimental fact that there exist a number of conserved quantities in particle physical systems — examples include energy and momentum, but also electrical and colour charge, among others — can be used to construct a theory of particle interactions, by finding the symmetries, and the symmetry groups, that correspond to these quantities and demanding that the theory be symmetric under their action.

An important type of group is the $SU(n)$ group. In the defining representation, this group consists of all complex unitary $n \times n$ matrices M with $\det M = 1$. The $SU(n)$ groups are *Lie groups*, which means that they are continuous — *i.e.* that it is possible to find group elements that are arbitrarily close to the identity. Also, any transformation in the group may be constructed by successive application of such infinitesimal transformations.

The group elements, and the objects on which the group acts, may be given in several *representations*. In the case of matrix groups this means matrices and vectors of different dimension. For an $SU(n)$ group, the two most important representations are the defining, or *fundamental*, representation, where the vectors have dimension n, and the *adjoint* representation, where the vectors have dimension $n^2 - 1$.

The infinitesimal behaviour of a Lie group is given by its corresponding *Lie algebra*, which is written[1]

$$[T_a, T_b] = i f_{abc} T_c. \quad (1.5)$$

The objects T_a are called the *generators* of the Lie group, and the f_{abc} are called the *structure coefficients*. The structure coefficients uniquely determine the algebra. For $SU(n)$, there are $n^2 - 1$ generators, so $a, b, c = 1, ..., n^2 - 1$. For a general Lie group, the binary operation $[\cdot, \cdot]$, called the *Lie bracket*, must be specified,

[1] Here, and in the following, repeated indices are summed over, unless otherwise stated.

but for $SU(n)$ is is just the commutator

$$[T_a, T_b] = T_a T_b - T_b T_a. \tag{1.6}$$

An element G of an $SU(n)$ group may generally be written as

$$G = e^{i\alpha_a O_a}, \tag{1.7}$$

in an open neighbourhood of the identity element in $SU(n)$. Here, $O_a = T_a$ in the fundamental representation, and $O_a = (f_{ij})_a$ in the adjoint representation. For $SU(2)$, the fundamental representation of the generators are proportional to the Pauli matrices σ_i, and for $SU(3)$ they are proportional to the Gell-Mann matrices λ_i. Formally, G is the *exponential map* from a Lie algebra to its corresponding Lie group.

The exponential map may be extended from a *global* transformation to a *local* transformation with respect to some underlying manifold X by letting $\alpha_a = \alpha_a(x)$, where $x \in X$. For the purposes of the Standard Model, X is space-time.

1.2 Ingredients of the Standard Model

The Standard Model consists of 12 fermions with corresponding antifermions, a number of vector gauge bosons and one scalar boson, whose existence have all been verified experimentally. The gauge bosons mediate interactions between the particles. There are three fundamental interactions in the Standard Model: The electromagnetic interaction, the weak interaction and the strong interaction. The combination of electromagnetic and weak interactions are often referred to as the *electroweak* theory, or the *Glashow-Weinberg-Salam model*, while the theory describing the strong interaction is known as *Quantum Chromodynamics (QCD)*. Not all particles couple to each other with all of the interactions.

The fermions are divided into two groups, the quarks and leptons. There are six different *flavours* of quarks, called up, down, strange, charm, bottom and top, in order of increasing mass. They are subdivided into three generations of pairs, up/down, charm/strange and top/bottom. The up, charm and top quarks carry quanta of $+\frac{2}{3}$ of the fundamental electrical charge e, while the down, strange and bottom quarks carry $-\frac{1}{3}e$. There are also six leptons, of which three are charged. They are called the electron, the muon and the tau. They each belong in a generation of their own, together with their neutral counterparts, the electron neutrino, muon neutrino and tau neutrino, respectively.

The vector gauge bosons consist of the photon, the Z and W bosons and the gluon. The photon is the mediator of electromagnetic interactions, the Z and W mediate the weak interaction and the gluon mediates the strong interaction. The photon, Z boson and gluon are all neutral, and they are their own antiparticles. The photon and gluon are massless, while the W and Z bosons are quite heavy.

The W carries one elementary unit of electric charge, and is thus distinct from its antiparticle, with a difference in sign for the particle and antiparticle states. The scalar boson of the Standard Model is the Higgs boson, which is responsible for giving particles their observed mass through the Higgs mechanism. It is electrically neutral and very massive.

Among the fermions, only the quarks couple to the strong interaction. All the known fermions couple with the weak interaction, while only the electrically charged fermions couple electromagnetically — *i.e.* all except the neutrinos. They couple to the Higgs field proportionally to their mass, so that for instance the top quark, which is the heaviest Standard Model particle, couples the strongest. A schematic overview of the particles in the Standard Model is shown in Fig. 1.1.

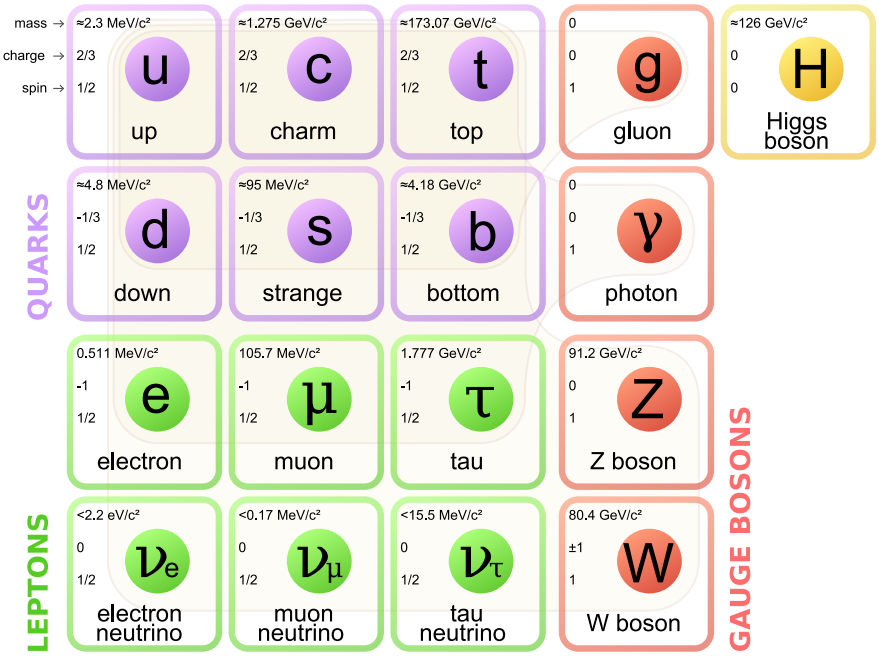


Figure 1.1: An overview of the particles of the Standard Model and their interactions, from [8].

The flavours of the quarks and leptons are conserved in the electromagnetic and strong interactions. For instance, a top quark cannot change into a charm or up quark by emission of a photon or gluon. The weak interaction enables the top quark to change into a bottom quark, or a tau lepton to change into a tau neutrino, through the emission of a charged W boson. This would still seem to conserve the *generation* of quark or lepton, but breaking of generation is also made possible through the mechanism of *generation mixing*, quantified by the Cabibbo-Kobayashi-Maskawa (CKM) matrix for the case of quarks and the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix for the leptons. The PMNS

mixing also explains the observed phenomenon of *neutrino oscillations*.[2]

1.3 Constructing a gauge theory

The Standard Model is a quantum field theoretic model, and may be stated in terms of a Lagrangian density function $\mathcal{L}$. The guiding principle for constructing the Lagrangian is *gauge invariance*. Gauge degrees of freedom are physical degrees of freedom which are "superflous", in the sense that they do not have any observable consequences. An example is Maxwell's theory of electromagnetism, where the electromagnetic vector potential A^μ is undetermined up to the addition of a total derivative term $\partial^\mu \phi$. The gauge freedom is exploited by requiring that the Lagrangian, which determines the physical dynamics, does not change when the gauge degrees of freedom are varied, *i.e.* that it is gauge invariant. This invariance is related to conservation of physical quantities by Noether's theorem.

The Standard Model is based on gauge invariance under three Lie groups, the famous $U(1)_Y \times SU(2)_L \times SU(3)_C$. The different particle types transform in different representations of the groups. In the Standard Model, the fermions transform in the fundamental representation, while the gauge bosons transform in the adjoint representation.

The particle content of the Standard Model is input into the Lagrangian by inserting fermionic fields, *i.e.* Dirac spinor fields, and imposing the desired gauge invariance on these fields. The basic Dirac term, called the Dirac bilinear, for some spinor field ψ, is[3]

$$\mathcal{L}_D = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi = \bar{\psi}(i\partial\!\!\!/ - m)\psi, \tag{1.8}$$

where γ_μ are the Dirac matrices, m is the mass of the spinor field, and $\bar{\psi} \equiv \psi^\dagger \gamma_0$. The Dirac bilinear results in the Dirac equation for the field when the equations of motion are applied. The Dirac equation and the Dirac matrices are derived in Appendix A.

Next, we impose gauge invariance. The local group transformation of an $SU(n)$ group may be written in the fundamental representation as

$$G(x) = e^{ig\alpha_a(x)T_a}, \tag{1.9}$$

where $\alpha_a(x)$ are n arbitrary real differentiable functions, g is the charge of the field, and T_a are the generators of $SU(n)$ in the fundamental representation. We assume that the Lagrangian consists of n Dirac bilinear terms with fields ψ_i, and that they are put into an n-dimensional multiplet $\Psi = (\psi_1, \psi_2, ..., \psi_n)^T$ such that

[2] If the neutrinos are assumed to be massive, which they are technically not in the Standard Model.

[3] We will, for what follows, set $\hbar = c = 1$.

the basic Dirac Lagrangian reads

$$\mathcal{L}_0 = \bar{\Psi}(i\partial\!\!\!/ - m)\Psi, \tag{1.10}$$

where we assume that all fields have the same mass m.[4] The group transformations of the multiplet and its adjoint are then

$$\begin{aligned} \Psi(x) &\xrightarrow{G} G(x)\Psi(x), \\ \bar{\Psi}(x) &\xrightarrow{G} \bar{\Psi}(x)G^\dagger(x). \end{aligned} \tag{1.11}$$

If we apply these transformations to the basic Lagrangian, it becomes

$$\begin{aligned} \mathcal{L}_0 =& \bar{\Psi}(x)(i\partial\!\!\!/ - m)\Psi(x) \\ \xrightarrow{G}& \bar{\Psi}(x)G^\dagger(x)(i\partial\!\!\!/ - m)G(x)\Psi(x) \\ =& \bar{\Psi}(x)(i\partial\!\!\!/ - m)\Psi(x) + i\bar{\Psi}(x)G^\dagger(x)(\partial\!\!\!/ G(x))\Psi(x). \end{aligned} \tag{1.12}$$

Thus, the basic Dirac Lagrangian is not gauge invariant, since we have picked up an additional term. Gauge invariance may be achieved by adding a term of the form

$$g\bar{\Psi}(x)\gamma^\mu A_{a,\mu}(x)T_a\Psi(x), \tag{1.13}$$

to the Lagrangian, where $A_{a,\mu}(x)$ is some new field, which we require to transform under G as

$$A_{a,\mu}(x)T_a \xrightarrow{G} G(x)\left(A_{a,\mu}(x)T_a + \frac{1}{g}\partial_\mu\right)G^\dagger(x). \tag{1.14}$$

If we apply G to the sum of the Dirac bilinear with this new term, it is invariant:

$$\begin{aligned} \mathcal{L}_0 =& \bar{\Psi}(x)(i\partial\!\!\!/ - m)\Psi(x) + g\bar{\Psi}(x)\gamma^\mu A_{a,\mu}(x)T_a\Psi(x) \\ \xrightarrow{G}& \bar{\Psi}(x)(i\partial\!\!\!/ - m)\Psi(x) + i\bar{\Psi}(x)G^\dagger(x)(\partial\!\!\!/ G(x))\Psi(x) \\ &+ g\bar{\Psi}(x)\gamma^\mu A_{a,\mu}(x)T_a\Psi(x) - i\bar{\Psi}(x)G^\dagger(x)(\partial\!\!\!/ G(x)) \\ =& \bar{\Psi}(x)(i\partial\!\!\!/ - m)\Psi(x) + g\bar{\Psi}(x)\gamma^\mu A_{a,\mu}(x)T_a\Psi(x). \end{aligned} \tag{1.15}$$

The term from Eq. (1.13) is usually included by replacing ∂_μ with the *covariant derivative*

$$D_\mu = \partial_\mu - igA_{a,\mu}T_a. \tag{1.16}$$

The fields $A_{a,\mu}$ are called gauge boson fields, and are responsible for mediating interactions between the Dirac fermion fields. The gauge boson fields must also

[4] This assumption is often wrong in the case of the Standard Model, but finds its solution in the Higgs mechanism.

have their own free-field term in the Lagrangian, called the *field strength*, which is given from the Proca Lagrangian for spin-1 fields as

$$-\frac{1}{4}F_{a,\mu\nu}F_a^{\mu\nu}, \tag{1.17}$$

where $F_a^{\mu\nu}$ is given as

$$F_a^{\mu\nu} = \partial^\mu A_a^\nu - \partial^\nu A_a^\mu + g f_{abc} A_b^\mu A_b^\nu, \tag{1.18}$$

where, again, f_{abc} are the structure coefficients of $SU(n)$. Note that with this definition, the field strength (1.17) is gauge invariant under Eq. 1.14.

With this, the total gauge invariant Lagrangian consists of n fermion fields and $n^2 - 1$ gauge boson fields, and reads

$$\mathcal{L} = \bar{\Psi}(i\not{D} - m)\Psi - \frac{1}{4}F_{a,\mu\nu}F_a^{\mu\nu}. \tag{1.19}$$

The covariant derivative gives rise to terms coupling the fermion and gauge fields. In the case of $n = 1$, the gauge group is the $U(1)$ group, which describes the theory of quantum electrodynamics, the simplest gauge theory. For $U(1)$, the structure coefficients vanish, since there is only a single gauge field,[5] making the Lagrangian particularily simple. In QED, there are no gauge boson self-interactions. For $n > 1$, the structure coefficients do not vanish, and this gives rise to cross terms in the field strength term $-\frac{1}{4}F_{a,\mu\nu}F_a^{\mu\nu}$ coupling the gauge bosons among themselves. These couplings are of great importance in the theories of weak and strong interactions.

1.4 Singlets, doublets and triplets

As mentioned previously, not all the fermion fields are subject to all the different interactions. If a field couples through a certain interaction, it is said to be *charged* under the transformations corresponding to that interaction. A specific amount g of charge is assigned to every field, and enters into the group transformations in Eq. (1.9). Thus, for $g = 0$, the transformation is the identity and has no effect. In the electromagnetic $U(1)$ case, this charge is the electrical charge, $g = q$. Analogous charges are associated with the $U(1)_Y$, $SU(2)_L$ and $SU(3)_C$ groups. They are called hypercharge, weak isospin and colour charge, respectively. These charges are the conserved quantitites associated with the gauge symmetries, as implied by Noether's theorem. Since the terms in the Lagrangian coupling fermions to gauge bosons are proportional to the charges, they are often called *coupling constants* — but the term coupling constant may also be used in

[5] This contradicts the claim that there are $n^2 - 1$ gauge fields — for $U(n)$ there are n^2. The reason is that $U(1)$ is not an $SU(n)$ group, but the above derivation works for $U(1)$ as well.

a broader sense to mean the constant of proportionality in any Lagrangian term which gives rise to a coupling vertex.

In the case of $SU(2)$ and $SU(3)$, the fields have to be put into vectors in order to be acted upon by the transformations, as was done in the previous section. Since the fermionic fields transform in the fundamental representation, the dimensions of the vectors are two and three, respectively. These types of vectors are referred to as $SU(2)$ *doublets* and $SU(3)$ *triplets*.

A Dirac field can be written as the sum of a left-chiral and a right-chiral part, defined by the projection operators

$$P_{R/L} = \frac{1}{2}\left(1 \pm \gamma^5\right), \tag{1.20}$$

where $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$. Given a Dirac field ψ, we may write

$$\psi = P_R\psi + P_L\psi \equiv \psi_R + \psi_L. \tag{1.21}$$

In the case of $SU(2)_L$, only the left chiral part of the fields are charged under the symmetry. For instance, the left-chiral parts of the quark fields are put in doublets, *e.g.*

$$q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \tag{1.22}$$

for the up- and down-quarks, while the right-handed parts are put in two separate singlets u_R and d_R, upon which the $SU(2)_L$ transformation has no effect. This has the consequence that the $SU(2)_L$ interaction is left-chiral — it only couples to left-handed parts of fields. Due to the spontaneous symmetry breaking of the $U(1)_Y \times SU(2)_L$ symmetry, the chirality is not exact in the resulting weak interactions, but it is still an important feature of the Standard Model.

The $SU(3)_C$ symmetry is the symmetry of the strong force, and among the fermions, only the quarks are charged under it. The quarks transform under $SU(3)$ in triplets — one for each quark flavour — where the components of the triplet are discriminated by differing *colour*, denoted red, green or blue.

While the fermions transform under the groups in the fundamental representation, which has dimension n for $SU(n)$, the gauge vector boson fields transform in the adjoint representation, which has dimension $n^2 - 1$. This number then determines the number of different gauge bosons for each group: $U(1)_Y$ has a single gauge boson field labeled B^μ, $SU(2)_L$ has three, labeled $W^\mu_{1,2,3}$, and $SU(3)_C$ has eight different gauge boson fields, labeled A^μ_a for $a = 1, ..., 8$. The $SU(3)_C$ bosons are called *gluons*. The $U(1)_Y$ and $SU(2)_L$ bosons are not the ones that we observe — the physical gauge boson eigenstates are linear combinations of them, mixed together by the spontaneous symmetry breaking of the Higgs mechanism.

1.5 The Higgs mechanism

Gauge invariance forbids the inclusion of terms of the form $m^2 A^\mu A_\mu$ into the Lagrangian, which are required in order to give vector bosons such as the Z their observed mass. To include the terms, we must also include a new complex scalar field doublet $\Phi = (\phi_a, \phi_b)^T$. This mechanism was first suggested by Anderson [9], and subsequently generalized to a relativistic field theory independently by three groups: Guralnik, Hagen and Kibble [10], Brout and Englert [11], and Higgs [12]. It is commonly referred to as the Higgs mechanism. The mechanism introduces the following terms into the Lagrangian:

$$\mathcal{L} \ni |D_\mu \Phi(x)|^2 - \mu^2 |\Phi(x)|^2 - \lambda |\Phi(x)|^4. \tag{1.23}$$

The last two terms comprise the Higgs *potential*. If μ^2 is assumed to be negative and λ positive, then the potential assumes the shape of a "mexican hat" as a function of $|\Phi|$. Figure 1.2 shows an analogous potential in the case of only a single complex scalar field ϕ. This potential has a minimum and a circle of degenerate

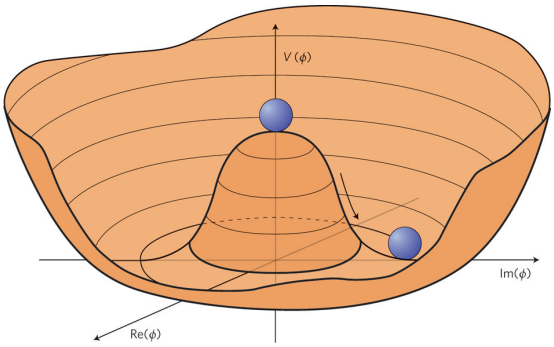


Figure 1.2: The shape of the Higgs potential, from [13].

energy at the field value $|\Phi_0|^2 = |\phi_a^0|^2 + |\phi_b^0|^2 = -\mu^2/2\lambda$. The mechanism of *spontaneous symmetry breaking* occurs when, as the energy decreases, the Higgs field falls to the bottom of the degenerate circle and is forced to *choose* a particular point on the circle for its lowest energy state, the vacuum. This causes Φ to obtain a non-zero *vacuum expectation value* (vev). Without loss of generality, we may write Φ_0 as

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{1.24}$$

in the vacuum state, where $v \equiv \sqrt{-\mu^2/\lambda}$. We may write Φ generally as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \eta_1(x) + i\eta_2(x) \\ v + H(x) + i\eta_3(x) \end{pmatrix}. \tag{1.25}$$

One can find parameters ξ_j, $j = 1, ..., 4$, such that, for a gauge transformation $G \in U(1)_Y \times SU(2)_L$, if

$$G^{-1} = e^{\frac{i}{v}\left(\xi_j \frac{1}{2}\sigma_j\right)} e^{\frac{i}{v}\xi_4}, \tag{1.26}$$

then

$$\Phi = \frac{1}{\sqrt{2}} G^{-1} \begin{pmatrix} 0 \\ v + H \end{pmatrix}. \tag{1.27}$$

Then, by gauge transforming the Lagrangian $\mathcal{L}$ using G, one obtains

$$\Phi \xrightarrow{G} \Phi' = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \tag{1.28}$$

and simultaneously, the vector gauge fields transform as

$$\begin{aligned} W_i^\mu &\xrightarrow{G} W_i'^\mu, \\ B^\mu &\xrightarrow{G} B'^\mu. \end{aligned} \tag{1.29}$$

In this gauge, the three degrees of freedom represented by the real scalar fields $\eta_i(x)$ are not present. The interpretation is that they are absorbed into the three bosons $W^\pm$ and Z, providing the longitudinal polarization degrees of freedom required for massive vector bosons. The remaining real scalar field $H(x)$ is the Higgs field.

This gauge choice also makes apparent that the gauge fields W_i^μ and B^μ mix together into the physical mass eigenstates $W^{\pm\mu}$ and Z^μ. This can be seen from the covariant derivative term in Eq. (1.23). In this gauge, the covariant derivative of Φ is

$$D'^\mu \Phi' = \left[\partial^\mu + ig\frac{\sigma_j}{2} W_j'^\mu + i\frac{1}{2} g' B'^\mu\right] \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \tag{1.30}$$

where g and g' are the coupling constants of $SU(2)_L$ and $U(1)_Y$, respectively, and σ_j are the Pauli matrices. By multiplying out, this becomes

$$D'^\mu \Phi' = \begin{bmatrix} \frac{ig}{2}(W_1'^\mu - iW_2'^\mu)\frac{v+H(x)}{\sqrt{2}} \\ \left(\partial^\mu - \frac{ig}{2} W_3'^\mu + \frac{ig'}{2} B'^\mu\right)\frac{v+H(x)}{\sqrt{2}} \end{bmatrix}.$$

By making the following definitions,

$$\tan\theta_W = \frac{g'}{g}, \tag{1.31}$$

$$W_3'^\mu = \cos\theta_W Z^\mu + \sin\theta_W A^\mu, \tag{1.32}$$

$$B'^\mu = -\sin\theta_W Z^\mu + \cos\theta_W A^\mu, \tag{1.33}$$

$$W^{\pm\mu} = \frac{1}{\sqrt{2}}\left(W'^\mu_1 \mp iW'^\mu_2\right), \tag{1.34}$$

where θ_W is called the *weak mixing angle*, Eq. (1.31) becomes

$$D'^{\mu}\Phi' = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{ig}{\sqrt{2}}W^{+\mu}\,(v+H) \\ \partial^{\mu}H(x) - \frac{ig}{2\cos\theta_W}Z^{\mu}\,(v+H) \end{pmatrix}. \tag{1.35}$$

Thus, the covariant derivative term $|D^{\mu}\Phi|^2$ is

$$\begin{aligned} |D^{\mu}\Phi|^2 = \frac{1}{2}\Bigg(& \frac{g^2v^2}{2}W^{+\mu}W^{-}_{\mu} + g^2vHW^{+\mu}W^{-}_{\mu} \\ & + \frac{g^2}{2}W^{+\mu}W^{-}_{\mu}H^2 + \partial^{\mu}H\partial_{\mu}H + \frac{g^2v^2}{4\cos^2\theta_W}Z^{\mu}Z_{\mu} \\ & + \frac{g^2}{2\cos^2\theta_W}vZ^{\mu}Z_{\mu}H + \frac{g^2}{4\cos^2\theta_W}Z^{\mu}Z_{\mu}H^2\Bigg). \end{aligned} \tag{1.36}$$

This expression contains the terms $\frac{g^2v^2}{8\cos^2\theta_W}Z^{\mu}Z_{\mu}$ and $\frac{g^2v^2}{4}W^{+\mu}W^{-}_{\mu}$, which are mass terms for the Z and $W^{\pm}$ bosons, respectively, fixing their masses to

$$m_Z = \frac{1}{2\cos\theta_W}vg, \quad m_W = \frac{1}{2}vg. \tag{1.37}$$

Note that this means that at the Lagrangian level

$$m_Z = \frac{m_W}{\cos\theta_W}. \tag{1.38}$$

Note also that there is no mass term for the photon field A^{μ} — it remains massless after symmetry breaking. This allows for the remaining $U(1)_{\text{em}}$ symmetry, where A^{μ} is the corresponding gauge boson and the gauge charge is the electrical charge, which by Noether's theorem must be conserved in all processes of the Standard Model.

The Higgs mechanism also provides for the possibility of fermionic terms of the form

$$\bar{\psi}_i\psi_j y_{ij}(v+H) = y_{ij}\bar{\psi}_i\psi_j H + vy_{ij}\bar{\psi}_i\psi_j. \tag{1.39}$$

The terms of the first type couple the fermions to the Higgs field. For $i = j$, the terms of the second type are mass terms of the form written in the Dirac bilinear, Eq. (1.8), and for $i \neq j$ they give rise to off-diagonal terms in the CKM and PMNS matrices. The coupling constants of these terms are called *Yukawa couplings*.

1.6 The Feynman calculus and loop corrections

Very few problems in the framework of the Standard Model can be solved exactly. Instead, calculations are done using perturbation theory to series expand the solution as an infinite sum of increasingly complicated, but decreasingly important,

contributions in terms of powers of some small parameter. Feynman invented a technique for visualizing these expansions using diagrams, known as Feynman diagrams [14]. For instance, the problem of electron-positron scattering has as one of its leading contributions the diagram shown in Fig. 1.3a. The next-to-leading order in the fine-structure constant $\alpha = e^2/4\pi$ includes the diagrams in figs. 1.3 (b), (c) and (d). The Feynman calculus associates each diagram with a

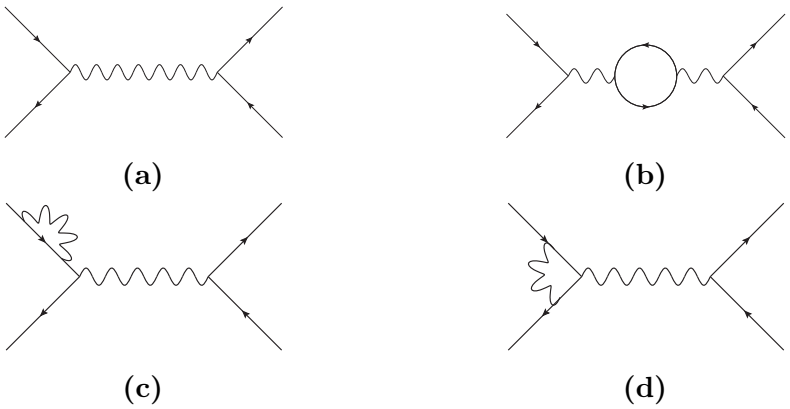


Figure 1.3: Feynman diagrams of contributions to e^+e^- scattering. Made using `JaxoDraw` [15].

specific mathematical expression called the Feynman amplitude $\mathcal{M}$ for that diagram. When several diagrams are included, the total amplitude for the process to the given order is the sum of the amplitudes from each diagram. The physical quantities of interest, *e.g.* cross sections and decay widths, are obtained by integrating the amplitude, or rather its absolute square, over all spin and momentum configurations of the system.

1.7 Renormalization

The subleading diagrams in a perturbative expansion, such as those shown in Fig. 1.3 (b–d), contain closed loops. These loops introduce extra momentum integrals into the calculations. Often, the integrals are divergent — which is unacceptable from a physical viewpoint. The divergences can be understood and dealt with by using the techniques of *regularization* and *renormalization*.

1.7.1 Regularization

Regularization is a means for parametrizing the divergence in terms of some small parameter ϵ which is zero in the physical limit. The most modern way to regularize a momentum integral is by using *dimensional regularization*: The

original loop integral is an integral over four space-time dimensions. Dimensional regularization makes the substitution $4 \to d = 4 - \epsilon$, so the integral becomes

$$\int d^4\, k \to \int d^d\, k. \tag{1.40}$$

This integral is mathematically well-defined, and allows the divergences to be parametrized in terms of ϵ.

An alternative to the dimensional regularization technique is to use a *momentum cutoff* Λ, which is infinite in the physical limit. One may even argue that new physics should come into play at higher energy scales, and therefore set the cutoff to a finite value, since the theory is valid only up to this energy. This will be applied to the calculation of Higgs mass loop corrections in Section 1.8.

1.7.2 Renormalization

When the divergence has been isolated and parametrized, it needs to be explained physically. This is done by the process of renormalizing the theory. For instance, in the case of the photon propagator in quantum electrodynamics, illustrated in Fig. 1.3b, the dimensionally regularized expression for the leading-order loop correction to the propagator is proportional to

$$\int_0^1 dx\, x(1-x) \left(\frac{2}{\epsilon} - \log\left(m^2 - x(1-x)q^2\right) + \text{constant terms} \right), \tag{1.41}$$

which blows up as $\epsilon \to 0$. Here, q is the momentum of the propagating photon and m is the mass of the fermions in the loop. Renormalization is the claim that this infinity is a part of the *bare* physical constants which are present in the Lagrangian, in this case the electrical charge, whose bare value is denoted e_0. These bare parameters are not observable quantities, only parameters in the Lagrangian. What is observed is the *renormalized* charge $e = e_0 + \delta e$, where δe is the infinite shift that cancels Eq. (1.41).

All the coupling constants of the Standard Model are renormalized. The renormalization introduces an *energy dependence* into the coupling constants, since the shift comes from loop corrections which depend on the energy of the process. For instance, the effective value of the electron charge in quantum electrodynamics, at some momentum q, is at one-loop order given as

$$e^2(q) = \frac{e_r^2}{1 - (e_r^2/6\pi^2)\log(q/M)}, \tag{1.42}$$

where e_r is some reference value for the charge, defined at the energy scale $q_r = M$. The fact that the coupling constants are not constant is referred to as the *running of the coupling constants*.

1.7.3 The Callan-Symanzik equation

The Callan-Symanzik, or *renormalization group* (RG), equation, is the equation which describes the running of the coupling constants in a systematic way for any interaction in a quantum field theory. It is obtained by requiring that the Greens function for the interaction, G, *i.e.* the propagator or coupling vertex, varies with the renormalization scale M in such a way that the bare parameters of the Lagrangian are unchanged. For the example of massless QED, the Callan-Symanzik equation for a Greens function $G^{(n,m)}$ with n electron fields and m photon fields is [16]

$$\left[M\frac{\partial}{\partial M} + \beta(e)\frac{\partial}{\partial e} + n\gamma_2(e) + m\gamma_3(e)\right] G^{(n,m)}(x_1, ..., x_n; M, e) = 0. \tag{1.43}$$

The functions beta and gamma are defined as

$$\beta \equiv M\frac{\partial e}{\partial M}, \quad \gamma_i \equiv -M\frac{\partial \eta_i}{\partial M}, \tag{1.44}$$

where $\delta\eta_i$ are the field-strength renormalization terms, shifting the field values of the electron and photon fields,

$$\psi \to (1 + \delta\eta_2)\psi \quad \text{and} \quad A_\mu \to (1 + \delta\eta_3)A_\mu, \tag{1.45}$$

respectively. The Callan-Symanzik equation states that the combined effect of all the shifts in parameters induced by the renormalization should exactly weigh up for the shift in the Green's function itself, which is given by

$$G^{(n,m)} \to (1 + n\delta\eta_2 + m\delta\eta_3)G^{(n,m)}. \tag{1.46}$$

This is what is stated in Eq. (1.43). The Callan-Symanzik equation for other interactions, such as the $SU(3)$ quantum chromodynamics, may be derived similarily, but its complexity grows with the complexity of the interaction.

The primary quantities of interest from a phenomenological viewpoint are the beta functions. They describe the change in the coupling constant and other parameters as a function of renormalization scale, and in the case of QED they may be used to derive the formula (1.42) for the running of the electromagnetic coupling constant e. Equation (1.42) shows that the electromagnetic coupling constant increases as a function of the energy q. The same turns out to be true for the weak coupling constant, while the strong coupling constant of QCD decreases with increasing energy. This last fact is called *asymptotic freedom*, and means that the quarks and gluons are unbound by strong forces in the limit of high energy, or equivalently, short distances.

1.8 Motivations for extending the Standard Model

The Standard Model provides a very accurate description of the matter we have around us, as well as three of the fundamental forces that are responsible for matter interactions. As a "theory of everything", it comes short since it does not describe the fourth fundamental force, gravity. Attempts at unifying the Standard Model with gravity have so far been unsuccessful. There are, however, also other motivations besides gravity for extending the Standard Model.

Since the Standard Model is widely believed to be a low-energy effective model of some more fundamental high-energy theory, it is speculated that the three interactions of the Standard Model unite at a higher energy and act as a single interaction under some larger gauge group, called a Grand Unified Theory (GUT). However, when the three couplings are evolved to high energies using the Callan-Symanzik equations, they do not meet at a single point. This is seen by many as a flaw of the Standard Model. In the theory of supersymmetry, the evolution of the couplings is altered, and they may meet at a single point. This effect is shown in Fig. 1.4. Supersymmetry is discussed in more detail in the next chapter.

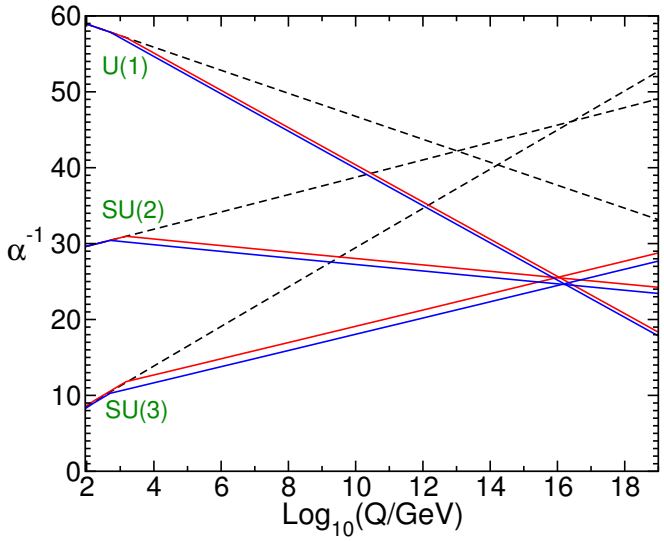


Figure 1.4: Evolution of the inverse coupling constants $\alpha_i^{-1} = 4\pi/g_i^2$, for the cases of the Standard Model (dashed lines) and models with supersymmetry (solid lines). From [2].

Another issue with the Standard Model is that is has no candidate for particle Dark Matter. Observations over the last century have given strong evidence for the existence of some as yet unkown form of matter which is distributed in large quantites all over the universe — in fact four times as much as our ordinary

matter. It is widely believed that this Dark Matter is some form of particle. Dark matter interacts primarily, or possibly even solely, via gravitation, so the particle has to be colourless and electrically neutral, because the strength of these interactions would otherwise have lead to the particle having been observed by now. It also has to be long-lived in order to explain the abundance of Dark Matter that we observe in the universe today. These restrictions rule out most of the Standard Model particles, with the exception of neutrinos. But neutrinos are known to be very light, almost massless, and calculations of early-universe dynamics show that they are too light to be candidates for Dark Matter.

There is also a more technical problem with the Standard Model, related to the scalar Higgs field. As discussed in Section 1.6, the calculations of parameters in a quantum field theory are subject to loop corrections. The Higgs mass parameter recieves corrections from loops containing all massive fermions, with the largest contribution coming from the top quark. The leading-order fermion contribution is shown in Fig. 1.5a and is divergent. When it is evaluated using a

Figure 1.5: Loop corrections to the Higgs mass. **(a)** shows the leading Standard Model contributions, and **(b)** shows the corresponding supersymmetric contributions which cancel them.

cutoff regulator Λ, the diagram turns out to be quadratic in the regulator. This is shown in Appendix B. By arguing that the Standard Model only is valid up to the energy scale of grand unification, one may set $\Lambda = M_{\text{GUT}} \sim 10^{16}$ GeV. This means that the Higgs mass corrections are many orders of magnitude larger than the observed Higgs mass of 126 GeV, implying the need very nearly exact cancellations among the correction terms with different signs.

The fermions and vector bosons of the Standard Model also recieve corrections, but because of chiral and gauge symmetries, these can be shown to be at most logarithmically divergent in terms of Λ. For the scalar Higgs particle, there is no such "protective symmetry". Neither are there any other symmetries in the Standard Model which says that a cancellation should occur, so it appears to be an "accident" of nature. Such accidents are seen as unnatural, and this explanation is thus unsatisfactory from a theoretical viewpoint. This is referred to as the *hierarchy problem* of the Higgs mass.

In supersymmetry, new scalar degrees of freedom enter into the loop corrections as illustrated in Fig. 1.5b for scalar "sfermions" $\tilde{f}$. This diagram is also calculated in Appendix B. The leading-order loop correction contributions from

fermions and sfermions combined are in unbroken supersymmetry

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda^2 + \frac{\lambda_{\tilde{f}}}{16\pi^2}\Lambda^2 + \cdots \tag{1.47}$$

to leading order in Λ, where $\lambda_{f/\tilde{f}}$ are the coupling strengths to fermions/sfermions. In supersymmetry, the coupling for a fermion f is related to its sfermion partner by $|\lambda_f|^2 = \lambda_{\tilde{f}}$, and there are exactly two sfermions for each fermion. Thus, in supersymmetry, the corrections cancel each other in a natural way.

Chapter 2

Supersymmetry

The theory of supersymmetry (SUSY) is a proposed extension of the Standard Model which increases the number of degrees of freedom by introducing a symmetry between fermions and bosons, called a supersymmetry. The construction of supersymmetry is in some sense a two-step process, where one first derives the Lagrangian of a theory with complete symmetry between fermions and bosons, meaning that every bosonic degree of freedom gets a corresponding 'supersymmetric' fermionic degree of freedom, and *vice versa*. These fields only differ in spin. But since, for example, scalar, colour charged particles with the same mass as the quarks are not observed in experiments, the symmetry cannot be exact. To make the theory physically viable, the supersymmetric partners must in most cases be significantly heavier than their Standard Model counterparts. This means that the supersymmetry must be a broken symmetry, and this breaking is in practice put into the theory by hand.

In this chapter we will outline the construction of a supersymmetric theory. First, we introduce the group theoretic framework of the symmetries. We define the concept of superfields, fields transforming under representations of the supersymmetry group. We go on to construct a fully supersymmetric Lagrangian in the framework of the Minimal Supersymmetric Standard Model (MSSM). Then the breaking of supersymmetry is achieved by manually inserting so-called "soft" supersymmetry-breaking terms. Also, the concept of R-parity is introduced in order to ensure the stability of the proton. R-parity will also make the lightest supersymmetric particle a good Dark Matter candidate. From the broken supersymmetry Lagrangian, we extract the particle content — identifying the familiar fields of the Standard Model as well as their supersymmetric counterparts. We then introduce a popular phenomenological model used to constrain and study the parameter space of the MSSM, and discuss its implications for the hierarchy of supersymmetric masses. This constrained model is subsequently adopted for the study of particular supersymmetric cascade decays, which is the topic for the remainder of the thesis. We will also review the current experimental status of supersymmetry.

2.1 Extending the Poincaré symmetry

In the beginning of Chapter 1, the Poincaré group was introduced. It is the group of all Lorentz boosts and rotations, as well as all translations in spacetime. It is defined by its Lie algebra, called the *Poincaré algebra*,

$$[M_{\mu\nu}, M_{\rho\sigma}] = -i(g_{\mu\rho}M_{\nu\sigma} - g_{\mu\sigma}M_{\nu\rho} - g_{\nu\rho}M_{\mu\sigma} + g_{\nu\sigma}M_{\mu\rho}), \tag{2.1}$$
$$[P_\mu, P_\nu] = 0, \tag{2.2}$$
$$[M_{\mu\nu}, P_\rho] = -i(g_{\mu\rho}P_\nu - g_{\nu\rho}P_\mu), \tag{2.3}$$

where $M_{\mu\nu}$ are the generators of Lorentz boosts and rotations and P_μ are the momentum operators, the generators of translations. Any physical theory obeying Special Relativity must be invariant under the Poincaré group. It was shown in 1967 by Coleman and Mandula [17], during attempts to unify Special Relativity with the observed global hadron flavour symmetry groups in a larger symmetry group structure, that there exists no Lie-algebra based extension of the Poincaré symmetry which includes the gauge groups of the Standard Model in a non-trivial way, *i.e.* a way by which the extended group cannot be written as a direct product such that the groups do not couple to each other.

This prompted Haag, Łopuszański and Sohnius [18] to introduce the concept of a *superalgebra*. A superalgebra, or *graded* Lie algebra, L, is a direct sum of two Lie algebras L_0 and L_1, $L = L_0 \oplus L_1$, with a special binary operation called a *grading*. For $x_i \in L_i$, the grading operation is given by

$$x_i \cdot x_j = x_k \in L_{i+j \bmod 2}, \tag{2.4}$$

which means that $x_0 \cdot x_0 \in L_0$, $x_1 \cdot x_1 \in L_0$ and $x_0 \cdot x_1 \in L_1$.

Haag *et. al.* constructed a superalgebra by combining the Poincaré algebra with an algebra spanned by four operators called *Majorana spinor charges*, represented by a two-component Weyl spinor Q_A (to be defined shortly) and its hermitian conjugate $\bar{Q}_{\dot{A}}$. The resulting superalgebra is given by the (anti)commutation relations

$$[Q_A, P_\mu] = [\bar{Q}_{\dot{A}}, P_\mu] = 0, \tag{2.5}$$
$$[Q_A, M_{\mu\nu}] = \sigma^B_{\mu\nu,A} Q_B, \tag{2.6}$$
$$\{Q_A, Q_B\} = \{\bar{Q}_{\dot{A}}, \bar{Q}_{\dot{B}}\} = 0, \tag{2.7}$$
$$\{Q_A, \bar{Q}_{\dot{B}}\} = 2\sigma^\mu_{A\dot{B}} P_\mu, \tag{2.8}$$

where $\sigma_\mu = (1_{2\times 2}, \sigma_i)$, with σ_i the Pauli matrices and $\sigma_{\mu\nu} = \frac{i}{4}(\sigma_\mu\bar{\sigma}_\nu - \sigma_\nu\bar{\sigma}_\mu)$. It is possible to extend the superalgebra further by introducing more Majorana spinor charges, labeled Q^α_A for $\alpha = 1, ..., N$. For general N, this extension can be shown to be the largest possible extension of the Poincaré group. The extension

defined by Eqs. (2.5)–(2.8) is called $N = 1$ supersymmetry. The transformations corresponding to the superalgebra are called *supersymmetry transformations*.

In the usual spinor representation of the Poincaré group, the fermion fields are represented as four-component Dirac spinors. It can be shown that the Poincaré group is isomorphic to $SL(2,\mathbb{C}) \times SL(2,\mathbb{C})$, so it is possible to define the theory using the representations of this group instead. The $SL(2,\mathbb{C})$ group has two inequivalent fundamental representations by two-component spinors which are called *left- and right-handed Weyl spinors* and written as ψ_A and $\bar{\psi}_{\dot{A}}$, respectively.

2.2 Superfields

The objects transforming under supersymmetry transformations can be represented by *superfields*, which are functions defined on the *superspace* spanned by the spacetime coordinates x^μ and four anti-commuting Grassman numbers θ_A and $\bar{\theta}_{\dot{A}}$. There are two important types of superfields, called chiral and vector superfields. Because of the anticommutativity, which means that any Grassman number squared vanishes, a function of a Grassman number, $f(\theta_A)$, has an all-order expansion given by

$$f(\theta_A) = a + b\theta_A. \tag{2.9}$$

Using this fact, a superfield Φ may generally be written as

$$\begin{aligned}\Phi\left(x^\mu, \theta_A, \bar{\theta}_{\dot{A}}\right) =& f(x) + \theta^A\phi_A(x) + \bar{\theta}_{\dot{A}}\bar{\chi}^{\dot{A}}(x) + \theta\theta m(x) \\ &+ \bar{\theta}\bar{\theta}n(x) + \theta\sigma^\mu\bar{\theta}V_\mu(x) + \theta\theta\bar{\theta}_{\dot{A}}\bar{\lambda}^{\dot{A}}(x) + \bar{\theta}\bar{\theta}\theta^A\psi_A(x) + \theta\theta\bar{\theta}\bar{\theta}d(x).\end{aligned} \tag{2.10}$$

The different field components have the following properties: $f(x)$, $m(x)$ and $n(x)$ are complex (pseudo)scalars, $\psi_A(x)$ and $\phi_A(x)$ are left-handed Weyl spinors, $\bar{\chi}^{\dot{A}}(x)$ and $\bar{\lambda}^{\dot{A}}(x)$ are right-handed Weyl spinors, $V_\mu(x)$ is a complex Lorentz four-vector and $d(x)$ is a complex scalar.

A supersymmetry transformation on a superfield $\Phi(x^\mu, \theta, \bar{\theta})$ in terms of infinitesimal parameters ξ_A and $\bar{\xi}^{\dot{A}}$ may be written [2]

$$\begin{aligned}\delta_\xi\Phi &= \left(\xi^A\frac{\partial}{\partial\theta^A} + \bar{\xi}_{\dot{A}}\frac{\partial}{\partial\bar{\theta}_{\dot{A}}} + i\left[\xi\sigma^\mu\bar{\theta} + \bar{\xi}\bar{\sigma}^\mu\theta\right]\partial_\mu\right)\Phi \\ &= \Phi(x^\mu + i\xi\sigma^\mu\bar{\theta} + i\bar{\xi}\bar{\sigma}^\mu\theta, \theta + \xi, \bar{\theta} + \bar{\xi}) - \Phi(x^\mu, \theta, \bar{\theta}).\end{aligned} \tag{2.11}$$

A set of covariant derivatives are defined by

$$D_A = \frac{\partial}{\partial\theta^A} - i\sigma^\mu_{A\dot{A}}\bar{\theta}^{\dot{A}}\partial_\mu, \quad \bar{D}^{\dot{A}} = -\frac{\partial}{\partial\bar{\theta}_{\dot{A}}} + i\bar{\sigma}^{\mu,A\dot{A}}\theta_A\partial_\mu, \tag{2.12}$$

which can be shown to satisfy

$$\delta_\xi(D_A\Phi) = D_A(\delta_\xi\Phi), \quad \delta_\xi(\bar{D}^{\dot{A}}\Phi) = \bar{D}^{\dot{A}}(\delta_\xi\Phi), \tag{2.13}$$

so that they are indeed supersymmetrically covariant. In terms of these, a *left chiral* superfield Φ is defined by the condition

$$\bar{D}^{\dot{A}}\Phi = 0. \tag{2.14}$$

By substituting $y^\mu = x^\mu - i\theta\sigma^\mu\bar{\theta}$, the covariant derivative $\bar{D}^{\dot{A}}$ is given as

$$\bar{D}^{\dot{A}} = -\frac{\partial}{\partial\bar{\theta}_{\dot{A}}}. \tag{2.15}$$

This shows that a left-chiral superfield must be independent of $\bar{\theta}$ in these coordinates, so it may generally be written as

$$\Phi(y,\theta) = A(y) + \sqrt{2}\theta\psi(y) + \theta\theta F(y), \tag{2.16}$$

thus containing two complex scalar fields and a left-handed Weyl spinor. Under the infinitesimal supersymmetry transformation defined in Eq. (2.11), the component field F can be shown to transform into a total derivative. It will thus not contribute to the action, since all fields must vanish on the boundary at infinity. For this reason it is called an auxillary field. Thus we see that a left-chiral superfield contains two bosonic (scalar) degrees of freedom and two fermionic degrees of freedom contained in a left-handed Weyl spinor. Similar arguments may be applied to define a right-chiral superfield by the condition

$$D_A\Phi^\dagger = 0, \tag{2.17}$$

and to show that it contains two auxillary and two proper scalar degrees of freedom, as well as a right-handed Weyl spinor.

A *vector superfield* V is defined by the condition

$$V^\dagger = V. \tag{2.18}$$

This condition allows the field content

$$V = f(x) + \theta^A\phi_A(x) + \bar{\theta}_{\dot{A}}\bar{\chi}^{\dot{A}}(x) + \theta\theta m(x) + \bar{\theta}\bar{\theta}m^*(x) \tag{2.19}$$

$$+ \theta\sigma^\mu\bar{\theta}V_\mu(x) + \theta\theta\bar{\theta}_{\dot{A}}\bar{\lambda}^{\dot{A}}(x) + \bar{\theta}\bar{\theta}\theta^A\lambda_A(x) + \theta\theta\bar{\theta}\bar{\theta}d(x). \tag{2.20}$$

Here, the scalar fields $f(x)$ and $d(x)$, as well as the four-vector $V_\mu(x)$, are required to be real fields, thus halving their amount of degrees of freedom. There are auxillary degrees of freedom which may be removed by a gauge transformation. One gauge which removes all auxillary fields from a vector superfield is the *Wess-Zumino* gauge. A vector superfield may in this gauge be written as

$$V_{\mathrm{WZ}} = (\theta\sigma^\mu\bar{\theta})\left[V_\mu(x) + i\partial_\mu(A(x) - A^*(x))\right] + \theta\theta\bar{\theta}_{\dot{A}}\bar{\lambda}^{\dot{A}}(x) + \bar{\theta}\bar{\theta}\theta_A\lambda^A(x) + \theta\theta\bar{\theta}\bar{\theta}d(x), \tag{2.21}$$

where $A(x)$ is a complex scalar field obeying $A(x) + A^*(x) = 2\Re A(x) = -f(x)$. In this gauge, the vector superfield contains one real scalar field degree of freedom (d.o.f.) from $d(x)$, three gauge field d.o.f.'s from $[V_\mu(x) + i\partial_\mu(A(x) - A^*(x))]$ and four fermion d.o.f.'s from the Weyl spinors $\lambda(x)$ and $\bar{\lambda}(x)$.[1]

2.3 The unbroken supersymmetric Lagrangian

To obtain a theory which is supersymmetric, the action, given by

$$S = \int d^4x \mathcal{L}, \tag{2.22}$$

needs to be invariant under supersymmetry transformations. As mentioned in the previous section, a total derivative has this property because its integral is determined by the boundary conditions, where it has to vanish. It can be shown that the highest-order component fields in θ and $\bar{\theta}$, *i.e.* the term proportional to $\theta\theta\bar{\theta}\bar{\theta}$, always has this property for both chiral and vector superfields and products thereof. Thus the invariance of the action may be ensured by redefining the Lagrangian using superfields such that

$$S = \int d^4x \int d^4\theta \mathcal{L}, \tag{2.23}$$

where the last integral is over the four Grassman variables. This will project out only the desired terms, because of how the Grassman integral is defined. Thus the supersymmetric Lagrangian may be constructed from superfields and their products. It can be written generically as

$$\mathcal{L} = \mathcal{L}_{\theta\theta\bar{\theta}\bar{\theta}} + \theta\theta\mathcal{L}_{\bar{\theta}\bar{\theta}} + \bar{\theta}\bar{\theta}\mathcal{L}_{\theta\theta}, \tag{2.24}$$

where the indices indicate the highest power of θ in each term. Note that this alters the dimensions of $\mathcal{L}$, since the Grassman variables θ have mass dimension $-1/2$.

The *superpotential* W is defined as a product of left-chiral superfields,

$$W(\Phi) = L^i\Phi_i + \frac{1}{2}m^{ij}\Phi_i\Phi_j + \frac{1}{3}\lambda^{ijk}\Phi_i\Phi_j\Phi_k. \tag{2.25}$$

The inclusion of higher-order field terms is ruled out by the condition of renormalizability, which forbids terms where the combined mass dimension of the factors (including the leading Grassman coordinates in Eq. 2.24 and the differential $d^4\theta$ in Eq. 2.23) are larger than four. Scalar, fermionic and auxillary fields have mass

[1] $V_\mu(x)$ has four d.o.f.'s, but one d.o.f. is removed by the gauge freedom of the imaginary part of $A(x)$.

dimension 1, 3/2 and 2, respectively, and the Grassman coordinates θ have dimension $-1/2$. The most general Lagrangian that can be written in terms of chiral superfields is

$$\mathcal{L} = \Phi_i^\dagger \Phi_i + \bar{\theta}\bar{\theta} W(\Phi) + \theta\theta W(\Phi^\dagger), \tag{2.26}$$

where the first term is called the kinetic term.

The Lagrangian has to be gauge invariant. The general gauge transformation of a chiral superfield under a group G is defined by

$$\Phi \xrightarrow{G} e^{-iq\Lambda^a T_a}\Phi, \tag{2.27}$$

where T_a are the group generators, q is the charge of Φ under G and the gauge parameters Λ_a can be shown to be left-chiral superfields. The equivalent transformation for a right-chiral superfield $\Phi^\dagger$ involves a right-chiral superfield gauge parameter $\Lambda_a^\dagger$.

Analogously to the Standard Model, the supersymmetric gauge interactions are introduced as compensating terms to the gauge transformation of the chiral superfields. The analogue to the gauge boson fields are the vector superfields V^a, which are introduced into the kinetic terms of the Lagrangian by writing them as

$$\Phi_i^\dagger e^{q_i V^a T_a}\Phi_i, \tag{2.28}$$

so that the kinetic term transforms as

$$\Phi_i^\dagger e^{q_i V^a T_a}\Phi_i \xrightarrow{G} \Phi_i^\dagger e^{iq_i(\Lambda^a)^\dagger T_a} e^{q_i V'^a T_a} e^{-iq_i\Lambda^a T_a}\Phi_i, \tag{2.29}$$

which is invariant given that the vector superfields transform as

$$e^{qV'^a T_a} = e^{-iq(\Lambda^a)^\dagger T_a} e^{qV^a T_a} e^{iq\Lambda^a T_a}. \tag{2.30}$$

For infinitesimal Λ, this is to leading order

$$V'^a = V^a + i(\Lambda^a - (\Lambda^a)^\dagger) - \frac{1}{2} q f^a_{bc} V^b((\Lambda^c)^\dagger + \Lambda^c). \tag{2.31}$$

This gives for the vector component fields of the vector superfields, V_μ^a,

$$V_\mu^a \xrightarrow{G} V'^a_\mu = V^a_\mu + i\partial_\mu(A^a - (A^a)^* - q f^{abc} V^b_\mu (A^c + (A^c)^*). \tag{2.32}$$

With these definitions, it can be shown that the Standard Model couplings of fermions with bosons are recovered by defining the covariant derivative

$$D^i_\mu = \partial_\mu - \frac{i}{2} q_i V_\mu. \tag{2.33}$$

The terms of the supersymmetric Lagrangian containing the field strengths of the gauge fields are written as

$$\mathrm{Tr}[W^A W_A], \tag{2.34}$$

where W_A and $\bar{W}_{\dot{A}}$ are left- and right-handed chiral superfields, respectively, given by

$$W_A \equiv -\frac{1}{4}\bar{D}\bar{D}e^{-qV^aT_a}D_A e^{qV^aT_a}, \tag{2.35}$$

$$\bar{W}_{\dot{A}} \equiv -\frac{1}{4}DDe^{-qV^aT_a}\bar{D}_{\dot{A}}e^{qV^aT_a}. \tag{2.36}$$

The most general form of the supersymmetric Lagrangian is then

$$\mathcal{L} = \Phi^\dagger e^{qV^aT_a}\Phi + \bar{\theta}\bar{\theta}W(\Phi) + \theta\theta W(\Phi^\dagger) + \frac{1}{2T(R)}\bar{\theta}\bar{\theta}\mathrm{Tr}(W^A W_A), \tag{2.37}$$

where $T(R)$, the *Dynkin index* of the representation of the gauge group, is a normalization constant.

2.4 Supersymmetry breaking

Supersymmetry has to be a broken theory, at least in the low-energy limit, since supersymmetric particles with Standard Model masses are not observed. The breaking can be inserted into the supersymmetry Lagrangian by hand, by explicitly adding terms that break supersymmetry and allow for mass splitting. The rationale for these terms is that the SUSY Lagrangian is only an effective Lagrangian where some heavy fields have been integrated out, and that the breaking of SUSY occurs through these fields at a higher scale through a spontaneous symmetry breaking mechanism similar to the Higgs mechanism in the Standard Model.

There are several alternatives for the mechanisms of SUSY breaking, some of which are Planck-scale mediated SUSY breaking, gauge mediated SUSY breaking and anomaly mediated SUSY breaking. Whichever of the mechanisms is chosen, there are only a finite set of terms that may be added to the Lagrangian without reintroducing the hierarchy problems of the Higgs mass loop corrections. They are called *soft* SUSY breaking terms, and are required to have couplings of mass dimension one or higher. They may in the most general form be written

$$\begin{aligned}\mathcal{L}_{\mathrm{soft}} = &-\frac{1}{4T(R)}M\theta\theta\bar{\theta}\bar{\theta}\mathrm{Tr}[W^A W_A] - \frac{1}{6}a_{ijk}\theta\theta\bar{\theta}\bar{\theta}\Phi_i\Phi_j\Phi_k \\ &-\frac{1}{2}b_{ij}\theta\theta\bar{\theta}\bar{\theta}\Phi_i\Phi_j - t_i\theta\theta\bar{\theta}\bar{\theta}\Phi_i + \mathrm{h.c.} \\ &-m_{ij}\theta\theta\bar{\theta}\bar{\theta}\Phi_i^\dagger\Phi_j.\end{aligned} \tag{2.38}$$

In terms of the component fields of the superfields, the soft Lagrangian may be written

$$\mathcal{L}_{\text{soft}} = -\frac{1}{2}M\lambda^A\lambda_A - \left(\frac{1}{6}a_{ijk}A_iA_jA_k + \frac{1}{2}b_{ij}A_iA_j + t_iA_i + \frac{1}{2}c_{ijk}A_i^*A_jA_k + \text{c.c.}\right) - m_{ij}^2A_i^*A_j. \tag{2.39}$$

Since this contains both Weyl spinor fields λ_A and scalar fields A_i, the soft terms may be used to modify masses and couplings of the superpartner scalar and fermionic fields which will appear below.

2.5 The Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model (MSSM) is the minimal supersymmetric theory which contains the field content of the Standard Model. It is constructed by choosing superfields in accordance with the requirements deduced in the previous sections. To construct a Dirac fermion, we use one left-chiral and one right-chiral superfield together. This gives the four fermionic degrees of freedom that a Dirac fermion and its antiparticle require. Since each chiral superfield also contains two scalar degrees of freedom (after removing the auxillary fields), this introduces two scalar particle-antiparticle pairs, which are called the supersymmetric partners, or *superpartners*, of the Dirac fermion. An important point is that all superfield components must have the same charge under all gauge groups, due to the way the gauge transformation was defined. This also means that the scalar fields generally will be charged, and that the supersymmetric partner fields have the same couplings as their Standard Model counterparts.

The superfields for the charged leptons are denoted l_i and $\bar{E}_i$ for the left- and right-chiral superfields, respectively, and the left-handed neutrino superfields are denoted ν_i. Here, $i = 1, 2, 3$ is a generation index. The $SU(2)_L$ doublet of the Standard Model is recovered by setting $L_i = (\nu_i, l_i)$. The quark superfields are denoted u_i, $\bar{U}_i$, d_i and $\bar{D}_i$, where $Q_i = (u_i, d_i)$ makes the $SU(2)_L$ doublet.

The gauge boson fields come from the vector superfields, each of which also contains two Weyl-spinor fields of opposite handedness. To obey gauge invariance, $n^2 - 1$ vector superfields are required for each of the $SU(n)$ groups just as in the Standard Model, *i.e.* gauge invariance under $U(1)_Y \times SU(2)_L \times SU(3)_C$ requires 1+3+8 vector superfields.[2] These are denoted B^0, W^a and C^a, respectively. The Weyl spinor fields, the superpartners of the gauge fields, are written as $\tilde{B}^0$, $\tilde{W}^a$ and $\tilde{g}$, respectively. In the literature, these are referred to as *bino*, *wino* and *gluino*.

[2] Again, n^2 rather than $n^2 - 1$ for the $U(1)$ group.

The MSSM requires two Higgs superfield $SU(2)_L$ doublets to be able to give mass to both up- and down-type quarks. In the Standard Model, the same Higgs doublet can be used for both types by rotating the components using the $SU(2)_L$ generators, but this is not possible in SUSY because it would mix left- and right-handed superfields in the superpotential. The Higgs doublets in the MSSM are

$$H_u = \begin{pmatrix} H_u^+ \\ H_u^0 \end{pmatrix}, \quad H_d = \begin{pmatrix} H_d^0 \\ H_d^- \end{pmatrix}, \tag{2.40}$$

where the sign indicates electric charge. This introduces several additional Higgs scalars into the model, as well as their fermionic superpartner fields.

Using the fields listed above, the MSSM Lagrangian $\mathcal{L}_{\text{MSSM}}$ may be constructed, subject to the rules for a general gauge invariant SUSY Lagrangian that were derived in sections 2.3 and 2.4. This gives rise to kinetic terms, superpotential terms, supersymmetric field strength terms and soft SUSY breaking terms. The total MSSM Lagrangian is

$$\mathcal{L} = \mathcal{L}_{\text{kin}} + W + \mathcal{L}_V + \mathcal{L}_{\text{soft}}. \tag{2.41}$$

The kinetic terms are given by

$$\begin{aligned} \mathcal{L}_{\text{kin}} =& L_i^\dagger e^{\frac{1}{2}g\sigma W - \frac{1}{2}g'B} L_i + Q_i^\dagger e^{\frac{1}{2}g_s\lambda C + \frac{1}{2}g\sigma W + \frac{1}{3}\cdot\frac{1}{2}g'B} Q_i \\ &+ \bar{U}_i^\dagger e^{\frac{1}{2}g_s\lambda C - \frac{4}{3}\cdot\frac{1}{2}g'B} \bar{U}_i + \bar{D}_i^\dagger e^{\frac{1}{2}g_s\lambda C + \frac{2}{3}\cdot\frac{1}{2}g'B} \bar{D}_i \\ &+ \bar{E}_i^\dagger e^{2\frac{1}{2}g'B} \bar{E}_i + H_u^\dagger e^{\frac{1}{2}g\sigma W + \frac{1}{2}g'B} H_u + H_d^\dagger e^{\frac{1}{2}g\sigma W - \frac{1}{2}g'B} H_d, \end{aligned} \tag{2.42}$$

where g', g and g_s are the coupling constants of $U(1)_Y$, $SU(2)_L$ and $SU(3)_C$, respectively. The hypercharge under $U(1)_Y$ is assigned in units of $\frac{1}{2}g'$ as a convention. Also, factors of $\frac{1}{2}$ are used in the transformations of $SU(2)_L$ and $SU(3)_C$ to avoid accumulation of numerical factors because of how the generators are defined. With these conventions, the electroweak relationship between electric charge Q, hypercharge Y and weak isospin T_3 is

$$Q = \frac{Y}{2} + T_3. \tag{2.43}$$

The field strength terms are

$$\mathcal{L}_V = \frac{1}{2}\text{Tr}\left[W^A W_A\right]\bar{\theta}\bar{\theta} + \frac{1}{2}\text{Tr}\left[C^A C_A\right]\bar{\theta}\bar{\theta} + \frac{1}{4}B^A B_A \bar{\theta}\bar{\theta} + \text{h.c.}, \tag{2.44}$$

where the field strengths are given as

$$W_A = -\frac{1}{4}\bar{D}\bar{D}e^{-W} D_A e^W, \quad W = \frac{1}{2}g\sigma^a W^a, \tag{2.45}$$

$$C_A = -\frac{1}{4}\bar{D}\bar{D}e^{-C} D_A e^C, \quad C = \frac{1}{2}g_s\lambda^a C^a, \tag{2.46}$$

$$B_A = -\frac{1}{4}\bar{D}\bar{D}D_A B, \quad B = \frac{1}{2}g'B^0. \tag{2.47}$$

The possible gauge invariant terms of the MSSM superpotential W are

$$\begin{aligned} W =& \mu H_u H_d + \mu'_i L_i H_u + y^e_{ij} L_i H_d E_j y^u_{ij} Q_i H_u \bar{U}_j + y^d_{ij} Q_i H_d \bar{D}_j \\ &+ \lambda_{ijk} L_i L_j \bar{E}_k + \lambda'_{ijk} L_i Q_j \bar{D}_k + \lambda''_{ijk} \bar{U}_i \bar{D}_j \bar{D}_k, \end{aligned} \tag{2.48}$$

where $H_u H_d$ is shorthand for $H_u^T i\sigma_2 H_d$, to make the term $SU(2)_L$ invariant.

2.5.1 R-parity

The superpotential in Eq. (2.48) contains terms which break lepton and baryon number conservation, namely LH_u, LLE and $LQ\bar{D}$. If the couplings for these terms are large, they allow for proton decay, which is experimentally very heavily constrained, with a lifetime of $\tau_{\text{proton}} > 10^{33}$ yr [19]. To avoid this, it is conventional to introduce the concept of *R-parity*, which gives an explanation for why these couplings are zero.

R-parity is a multiplicative quantum number which is assumed to be conserved in all SUSY interactions. Formally, a particle has R-parity given by

$$R = (-1)^{2s+3B+L}, \tag{2.49}$$

where s is the particle's spin, B its baryon number and L its lepton number. The important point is that all Standard Model particles have $R = +1$ while all superpartners have $R = -1$. This leads to the very important prediction that superpartner particles can only be produced and annihilate in pairs. In particular, it means that the lightest superpartner (LSP) must be stable against decay. This makes the LSP very attractive as a candidate for Dark Matter, if it is electrically neutral and has no colour charge. In a SUSY scenario with conservation of R-parity, the supersymmetric particles will typically decay in *cascades* down to the LSP, emitting multiple Standard Model particles along the way.

With R-parity, the possible supersymmetry breaking soft terms of the MSSM are, in component fields,

$$\begin{aligned} \mathcal{L}_{\text{soft}} =& \left(-\frac{1}{2} M_1 \tilde{B}\tilde{B} - \frac{1}{2} M_2 \tilde{W}^{i,A} \tilde{W}^i_A - \frac{1}{2} M_3 \tilde{g}^{a,A} \tilde{g}^a_A + \text{c.c.} \right) \\ &+ \left(-a^e_{ij} \tilde{L}_i H_d \tilde{e}^*_{iR} - a^u_{ij} \tilde{Q}_i H_u \tilde{u}^*_{iR} - a^d_{ij} \tilde{Q}_i H_d \tilde{d}^*_{jR} + \text{c.c.} \right) \\ &- (m^L_{ij})^2 \tilde{L}^\dagger_i \tilde{L}_j - (m^e_{ij})^2 \tilde{e}^*_{iR} \tilde{e}_{jR} - (m^Q_{ij})^2 \tilde{Q}^\dagger_i \tilde{Q}_j \\ &- (m^u_{ij})^2 \tilde{u}^*_{iR} \tilde{u}_{jR} - (m^d_{ij})^2 \tilde{d}^*_{iR} \tilde{d}_{jR} - m^2_{H_u} H^\dagger_u H_u - m^2_{H_d} H^\dagger_d H_d. \end{aligned} \tag{2.50}$$

The soft terms constitute the main contributions to the superpartner masses.

2.6 Radiative electroweak symmetry breaking

The scalar potential of the MSSM Higgs fields is, in terms of component fields,

$$\begin{aligned} V(H_u, H_d) =& |\mu|^2 \left(|H_u^0|^2 + |H_u^+|^2 + |H_d^0|^2 + |H_d^-|^2\right) \\ &+ \frac{1}{8}\left(g^2 + g'^2\right)\left(|H_u^0|^2 + |H_u^+|^2 - |H_d^0|^2 - |H_d^-|^2\right) \\ &+ \frac{1}{2} g^2 |H_u^+ H_d^{0*} + H_u^0 H_d^{-*}|^2 \\ &+ m_{H_u}^2 \left(|H_u^0|^2 + |H_u^+|^2\right) + m_{H_d}^2 \left(|H_d^0|^2 + |H_d^+|^2\right) \\ &+ \left[b\left(H_u^+ H_d^- - H_u^0 H_d^0\right) + \text{c.c.}\right]. \end{aligned} \tag{2.51}$$

It can be shown that a necessary and sufficient set of conditions for achieving spontaneous electroweak symmetry breaking in this potential is [2]

$$b^2 > \left(|\mu|^2 + m_{H_u}^2\right)\left(|\mu|^2 + m_{H_d}^2\right), \tag{2.52}$$

$$2b < 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2. \tag{2.53}$$

Like any quantum field theory, the MSSM is subject to renormalization, which induces the running of the coupling constants and masses of the model as discussed in Chapter 1. In particular, the mass parameters $m_{H_{u/d}}$ for the Higgs doublets, which come from soft breaking terms in the Lagrangian, run with energy. To break the electroweak symmetry, *i.e.* satisfy Eqs. (2.52) and (2.53), one can assume that these are equal at some high scale, and run down. It can be shown that the beta functions which determine the running to lowest order are given by

$$16\pi^2 \beta_{m_{H_u}^2} = 6|y_t|^2 \left(m_{H_u}^2 + (m_{33}^Q)^2 + (m_{33}^u)^2\right), \tag{2.54}$$

$$16\pi^2 \beta_{m_{H_d}^2} = 6|y_b|^2 \left(m_{H_d}^2 + (m_{33}^Q)^2 + (m_{33}^d)^2\right), \tag{2.55}$$

where y_t and y_b are the top and bottom Yukawa couplings, respectively. Because the top Yukawa coupling is much larger than the bottom one, the m_{H_u} parameter runs down much faster with energy and can become negative, facilitating the symmetry breaking given by Eqs. (2.52) and (2.53). This is called *radiative electroweak symmetry breaking*, and gives an explanation for the Higgs mechanism.

As a result of the symmetry breaking, the neutral components of both the Higgs doublets acquire a non-vanishing vacuum expectation value, $v_u = \langle H_u^0 \rangle$ and $v_d = \langle H_d^0 \rangle$, respectively. These must relate to the vector boson masses of the Standard Model as

$$v_u^2 + v_d^2 = \frac{2m_Z^2}{g^2 + g'^2} \approx (174 \text{ GeV})^2. \tag{2.56}$$

The remaining free parameter is conventionally parametrized as

$$\tan\beta \equiv \frac{v_u}{v_d}. \tag{2.57}$$

2.7 Particle phenomenology of the MSSM

Standard model particles[3]	Spin	Superpartners	Spin
e, μ, τ	1/2	$\tilde{e}_L, \tilde{e}_R, \tilde{\mu}_L, \tilde{\mu}_R, \tilde{\tau}_1, \tilde{\tau}_2$	0
ν_e, ν_μ, ν_τ	1/2	$\tilde{\nu}_{e,L}, \tilde{\nu}_{e,R}, \tilde{\nu}_{\mu,L}, \tilde{\nu}_{\mu,R}, \tilde{\nu}_{\tau,L}, \tilde{\nu}_{\tau,R}$	0
u, d, s, c	1/2	$\tilde{u}_L, \tilde{u}_R, \tilde{d}_L, \tilde{d}_R, \tilde{s}_L, \tilde{s}_R, \tilde{c}_L, \tilde{c}_R$	0
b, t	1/2	$\tilde{b}_1, \tilde{b}_2, \tilde{t}_1, \tilde{t}_2$	0
$h, H, A, H^\pm$	0	$\chi_1^0, \chi_2^0, \chi_3^0, \chi_4^0$	1/2
$B^0, W^{1,2,3}$	1	$\chi_1^\pm, \chi_2^\pm$	1/2
g	1	$\tilde{g}$	1/2

Table 2.1: Summary of MSSM particle and sparticle content.

The total particle content of the MSSM is as follows: The Standard Model particles are present: electrons, muons, taus and their corresponding neutrinos, the up, down, strange, charm, bottom and top quarks, the photon, Z boson, W bosons, gluons, and the Higgs boson. In addition to the Standard Model Higgs h, there are four other scalar Higgs particles with positive R-parity, labeled H, $H^\pm$ and A. H is identical to h except for its larger mass, and is therefore termed "heavy Higgs", in contrast to the "light Higgs" h of the Standard Model. The other neutral field A is a pseudo-scalar.

All the Standard Model particles have superpartners, often termed *sparticles*. For the gluons, they are called gluinos and labeled $\tilde{g}$. The partners of the B^0 and W^a fields, which in the Standard Model make up the photon, Z and $W^\pm$, mix with the superpartner Higgs fields to form four neutral Majorana fermions called neutralinos, labeled $\tilde{\chi}_i^0$, $i = 1, ..., 4$, and two charged fermion-antifermion pairs called charginos, $\tilde{\chi}_i^\pm$, $i = 1, 2$. Each of the Standard Model fermions get two corresponding scalar particles with the same gauge charges. For the first two generations, the mass matrix is to good approximation diagonal in the basis spanned by the superpartner fields to the chiral Weyl fermions, so *e.g.* the superpartners of the up quark u are labeled $\tilde{u}_R$ and $\tilde{u}_L$. For the third-generation fermions, the chiral approximation is cruder, so these mass eigenstates are just numbered, *e.g.* $\tilde{b}_1$ and $\tilde{b}_2$ for the b quark. A summary of the MSSM particle content is shown in Table 2.1.

The fact that superpartner fields inherit the couplings of their Standard Model counterparts is very useful when drawing Feynman diagrams for processes, since

[3]Plus the extended Higgs sector.

intuition from the Standard Model can be applied through *supersymmetrization* of the diagrams. For instance, the fermion-fermion-gauge boson vertices (Fig. 2.1a) of the Standard Model have supersymmetrized fermion-sfermion-gaugino and sfermion-sfermion-gauge boson versions (figs. 2.1 b and c). However, sfermion-sfermion-gaugino does not exist, because of R-parity.

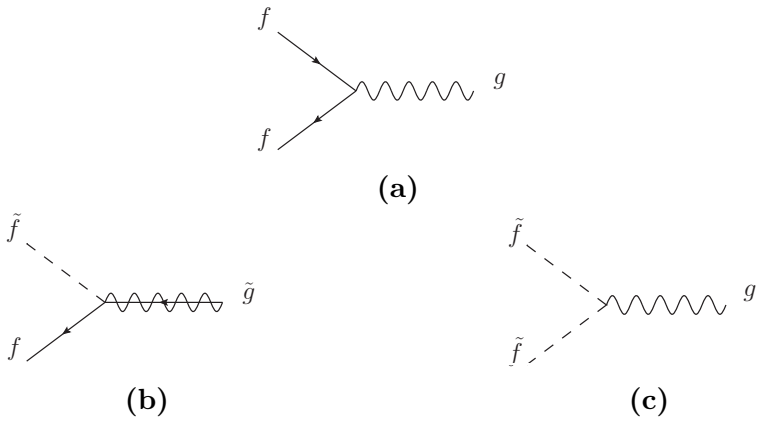


Figure 2.1: Feynman diagrams of a *fermion-fermion-gauge boson* vertex **(a)** and the supersymmetrized *fermion-sfermion-gaugino* **(b)** and *sfermion-sfermion-gauge boson* **(c)** vertices.

An important consequence of the coupling inheritance is that the gaugino fields $\tilde{\chi}_i^{0/\pm}$ will couple differently depending on the mass mixing matrix for these states. Since the wino field inherits the left-chiral $SU(2)_L$ coupling, the chirality of the couplings of the four neutralinos may differ considerably if the mass eigenstates are close to the gauge eigenstates. In the parameter point used for the analysis in the coming chapters, the second-generation neutralino consists of a large wino part, and this means that it has a small coupling to right-chiral quarks.

2.7.1 Sparticle masses

The main contributions to the sparticle masses naturally come from the soft terms, since these are responsible for the symmetry breaking which would otherwise give Standard Model masses to sparticles. The mass parameters from these terms are the gaugino mass terms $M_{1,2,3}$, the sfermion mass terms m_{ij} and the Higgs terms $m_{H_{u/d}}$. Additionally, the parameter μ from the superpotential term coupling the Higgs doublets together in the unbroken SUSY Lagrangian contributes [2].

- At tree level, the different Higgs masses (which have positive R-parity) are

$$m_A^2 = 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2, \tag{2.58}$$

$$m_{h,H}^2 = \frac{1}{2}\left(m_A^2 + m_Z^2 \pm \sqrt{(m_A^2 - m_Z^2)^2 + 4m_Z^2 m_A^2 \sin^2 2\beta}\right), \tag{2.59}$$

$$m_{H^\pm}^2 = m_A^2 + m_W^2. \tag{2.60}$$

- The gluino mass $m_{\tilde{g}}$ is given as

$$m_{\tilde{g}} = M_3\left[1 + \frac{\alpha_s}{4\pi}\left(15 + 6\ln\frac{\mu}{M_3} + \sum_{\text{all } \tilde{q}} A_{\tilde{q}}\right)\right], \tag{2.61}$$

  where $A_{\tilde{q}}$ are the significant squark one-loop contributions given by

$$A_{\tilde{q}} = \int_0^1 dx\, x \ln\left(x\frac{m_{\tilde{q}}^2}{M_3^2} + (1-x)\frac{m_q^2}{M_3^2} - x(1-x) - i\epsilon\right). \tag{2.62}$$

- The neutralinos $\tilde{\chi}_i^0$ are the mass eigenstates of the bino, wino and higgsino fields. In the gauge eigenstate basis, where $\tilde{\chi}^0 = (\tilde{B}^0, \tilde{W}^0, \tilde{H}_d^0, \tilde{H}_u^0)^T$, the mass matrix may be written at tree level as

$$M_{\tilde{\chi}^0} = \begin{pmatrix} M_1 & 0 & -c_\beta s_{\theta_W} m_Z & s_\beta s_{\theta_W} m_Z \\ 0 & M_2 & c_\beta c_{\theta_W} m_Z & -s_\beta c_{\theta_W} m_Z \\ -c_\beta s_{\theta_W} m_Z & s_\beta s_{\theta_W} m_Z & 0 & -\mu \\ c_\beta c_{\theta_W} m_Z & -s_\beta c_{\theta_W} m_Z & -\mu & 0 \end{pmatrix}, \tag{2.63}$$

  where $c_x = \cos x$ and $s_x = \sin x$. For a given parameter choice, this matrix must be diagonalized to find the neutralino masses and field content.

- The charginos $\tilde{\chi}_i^\pm$ have analogous structure. In the gauge eigenstate basis $\tilde{\chi}^\pm = (\tilde{W}^+, \tilde{H}_u^+, \tilde{W}^-, \tilde{H}_d^-)^T$, the mass matrix is

$$M_{\tilde{\chi}^\pm} = \begin{pmatrix} 0 & 0 & M_2 & \sqrt{2}c_\beta m_W \\ 0 & 0 & \sqrt{2}s_\beta m_W & \mu \\ M_2 & \sqrt{2}s_\beta m_W & 0 & 0 \\ \sqrt{2}s_\beta m_W & \mu & 0 & 0 \end{pmatrix}. \tag{2.64}$$

- The first two generations of *sfermions*, superpartners of the Standard Model fermions, get masses according to

$$m_F^2 = m_{F,\text{soft}}^2 + \Delta_F, \tag{2.65}$$

where $m^2_{F,\text{soft}}$ is the mass term from the soft term of the form $-m_F^2 F^\dagger F$ and Δ_F is given by

$$\Delta_F = (T_{3F} - Q_F \sin^2\theta_W)\cos 2\beta\, m_Z^2, \tag{2.66}$$

where T_{3F} and Q_F are the weak isospin and electric charge, respectively, of the left-handed supermultiplet to which the sfermion belongs. The masses are then *e.g.*

$$m^2_{\tilde{e}_L} = m^2_{L_1} + \Delta\tilde{e}_L, \tag{2.67}$$
$$m^2_{\tilde{e}_R} = m^2_{e_R} + \Delta\tilde{e}_R. \tag{2.68}$$

The mass-splitting between same-generation sleptons and squarks are universal and given by *e.g.*

$$m^2_{\tilde{e}_L} - m^2_{\tilde{\nu}_L} = m^2_{\tilde{d}_L} - m^2_{\tilde{u}_L} = -\cos 2\beta\, m_W^2. \tag{2.69}$$

- The third-generation sfermions have more complicated mass contributions, *e.g.* the stop squark mass matrix in the chiral gauge eigenstate basis $\tilde{t} = (\tilde{t}_L, \tilde{t}_R)^T$ is given by

$$m^2_{\tilde{t}} = \begin{pmatrix} m^2_{Q_3} + m_t^2 + \Delta\tilde{u}_L & v((a^u_{33})^* \sin\beta - \mu y_t \cos\beta) \\ v(a^u_{33}\sin\beta - \mu^* y_t\cos\beta) & m^2_{u3} + m_t^2 + \Delta\tilde{u}_R \end{pmatrix}, \tag{2.70}$$

which can be diagonalized to find the mass eigenstates. The reason that the third-generation sfermions have more complicated mass matrices is that the Yukawa and trilinear couplings are larger, giving non-negligible contributions to the masses.

2.8 Gauge unification and mass hierarchies

It was mentioned in Section 1.8 that one very appealing feature of supersymmetry is that the coupling constants of the electromagnetic, weak and strong interactions can be made to unite at a high energy scale. In the MSSM, this scale, called the "grand unification scale" (GUT), is $m_{\text{GUT}} \approx 2 \times 10^{16}$ GeV. If we assume that the grand unified gauge group is $SU(5)$ or $SO(10)$,[4], then we can take the three gauge couplings of the grand unified theory to be [20]

$$g_1 = \sqrt{\frac{5}{3}} g', \quad g_2 = g, \quad g_3 = g_s. \tag{2.71}$$

[4] $SO(n)$ is the *special orthogonal group* given in the defining matrix representation as the group of all orthogonal $n \times n$ matrices of determinant one.

It can be shown that the one-loop β-functions of the gauge couplings g_i and the gaugino soft mass parameters M_i are

$$\beta_{g_i}|_{1-\text{loop}} = \frac{1}{16\pi^2} b_i g_i^3, \tag{2.72}$$

$$\beta_{M_i}|_{1-\text{loop}} = \frac{1}{8\pi^2} g_i^2 M_i b_i. \tag{2.73}$$

The ratios M_i/g_i^2 for all three parameter pairs are scale independent at one-loop order. To see this, it is convenent to write the β-functions as

$$\beta_X \equiv \frac{d}{dt} X, \tag{2.74}$$

for a running variable X, where $t = \log \mu$ is the *scale parameter* for the renormalization scale μ. If we define

$$R = \frac{M_i}{g_i^2}, \tag{2.75}$$

then

$$\beta_R = \frac{dR}{dt} = \frac{(\frac{d}{dt} M_i) g_i^2 - M_i \frac{d}{dt} g_i^2}{g_i^4} = 0,$$

which proves the claim.

By assuming that the coupling constants g_i unite to the value g_u at the GUT scale, and that the gaugino masses M_i have the common value $m_{1/2}$ at the same scale, then it follows that

$$\frac{M_1}{g_1^2} = \frac{M_2}{g_2^2} = \frac{M_3}{g_3^2} = \frac{m_{1/2}}{g_u}. \tag{2.76}$$

In terms of the electromagnetic and strong fine-structure constants α and α_s, and the weak mixing angle θ_W, this gives

$$M_3 = \frac{\alpha_s}{\alpha} \sin^2 \theta_W M_2 = \frac{3}{5} \frac{\alpha_s}{\alpha} \cos^2 \theta_W M_1. \tag{2.77}$$

At a scale of 1 TeV, this evaluates to approximately

$$M_3 : M_2 : M_1 \approxeq 6 : 2 : 1. \tag{2.78}$$

This propagates into the mass formulae to predict the approximate mass relationships

$$m_{\tilde{g}} \approx 6 m_{\tilde{\chi}_1^0} \quad \text{and} \quad m_{\tilde{\chi}_2^0} \approx m_{\tilde{\chi}_1^\pm} \approx 2 m_{\tilde{\chi}_1^0}. \tag{2.79}$$

2.9 The Constrained MSSM

The available parameter space of the general MSSM is very large, of the order 100 free parameters. The reason for the vast number of parameters are the soft supersymmetry breaking terms, which contain coupling parameters for all the extra terms without any *a priori* way to relate them. It is therefore conventional to make some restricting assumptions. A much-studied restriction is the *Constrained MSSM* (CMSSM), also known as *minimal supergravity* (mSUGRA). This model is constructed by assuming that supersymmetry breaking is mediated by some mechanism of gravity at the Planck scale of $M_P = 2.4 \times 10^{18}$ GeV. By assuming a minimal form for the parameters at the GUT scale, to obtain gauge unification, the resulting model is parametrized in terms of five parameters,

$$m_{1/2},\ m_0,\ A_0,\ \tan\beta \text{ and } \mathrm{sign}(\mu). \tag{2.80}$$

The mass parameters $m_{1/2}$ and m_0 are the common soft masses of all gauginos and sfermions, respectively, at the GUT scale. The mass splitting between sparticles appears when the individual sparticle masses are evolved down to a lower scale. This is illustrated in Fig. 2.2 for one choice of $m_{1/2}$ and m_0. The figure also

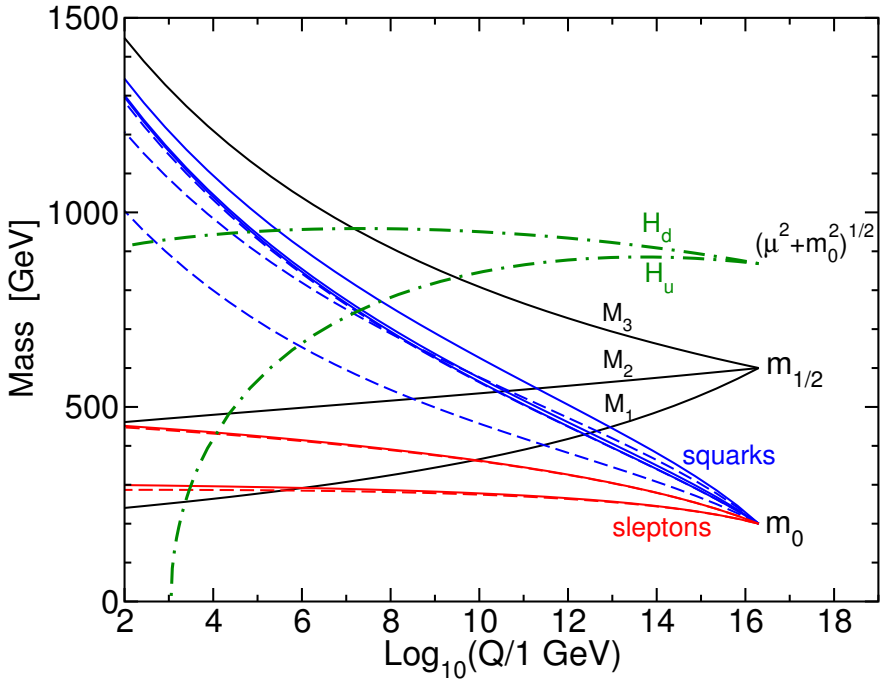


Figure 2.2: MSSM RG-running, from [2]. In this figure, $m_0 = 200$ GeV and $m_{1/2} = 600$ GeV.

illustrates the evolution of m_{H_u} down to a negative value, facilitating the radiative electroweak symmetry breaking discussed in Section 2.6.

Because the renormalization running is proportional to the Yukawa couplings of the Standard Model partners, the sfermion partners will often get an inverted mass hierarchy compared to the Standard Model. Especially in the colour sector, where stop and sbottom are often the lightest squarks. However, the effect may be compensated by mixing terms, and the first and second generation mass splittings are small or even degenerate, since the masses are dominated by the common soft mass terms. Since all the sfermion masses evolve from the common mass m_0, the squarks will typically be heavier than the sleptons in CMSSM because the running of the former is affected by QCD interactions. Also, because the $SU(3)$ coupling g_3 influences the RG-running of $m_{\tilde{q}}$, typically $m_{\tilde{q}} \approx m_{\tilde{g}}$.

This thesis studies a particular type of cascade decay chain, namely

$$\tilde{q} \to \tilde{\chi}_2^0 q \to \tilde{l} l q \to \tilde{\chi}_1^0 l l q. \tag{2.81}$$

which exhibits several features which enable determination of SUSY mass parameters. For this chain to be available to study in a CMSSM parameter point, the following conditions must be satisfied:

1. The mass hierarchy must be $m_{\tilde{q}} > m_{\tilde{\chi}_2^0} > m_{\tilde{l}} > m_{\tilde{\chi}_1^0}$.

2. The gluino must be as heavy or heavier than the squarks, to disable the $\tilde{q} \to \tilde{g} q$ channel, which otherwise dominates over the chain process.

3. All the intermediate particles must have sizeable branching fractions to the subsequent chain particles.

In [21] it is shown that these conditions are indeed satisfied in large parts of the CMSSM parameter space. Figure 2.3 shows plots of a selection of the CMSSM parameter space, indicating different mass hierarchies. The green areas have the correct hierarchy for the chain and the right LSP. The darker green area inside it additionally has a sizeable branching fraction into the chain, *i.e.* more than 10% of the branching at the SPS1a benchmark point defined in [22], which is marked and denoted α in Fig. 2.3a. Regions where a charged particle (the stau) is the LSP are excluded. The SPS1a point is the benchmark point that will be used in subsequent chapters.

2.9.1 The SPS1a benchmark point

The SPS1a benchmark point is defined in [22] as the CMSSM point with $m_0 = 100$ GeV, $m_{1/2} = 250$ GeV, $A_0 = -100$ GeV, $\tan\beta = 10$ and $\mu > 0$. The point is actually defined as the MSSM parameters obtained by using `ISAJET 7.58` [23] to RG-run down from the defining CMSSM parameters. In this thesis we will instead use the parameters obtained by using `SOFTSUSY 3.4.1` [24] for RG-running. This gives a slightly different mass spectrum, but the main features

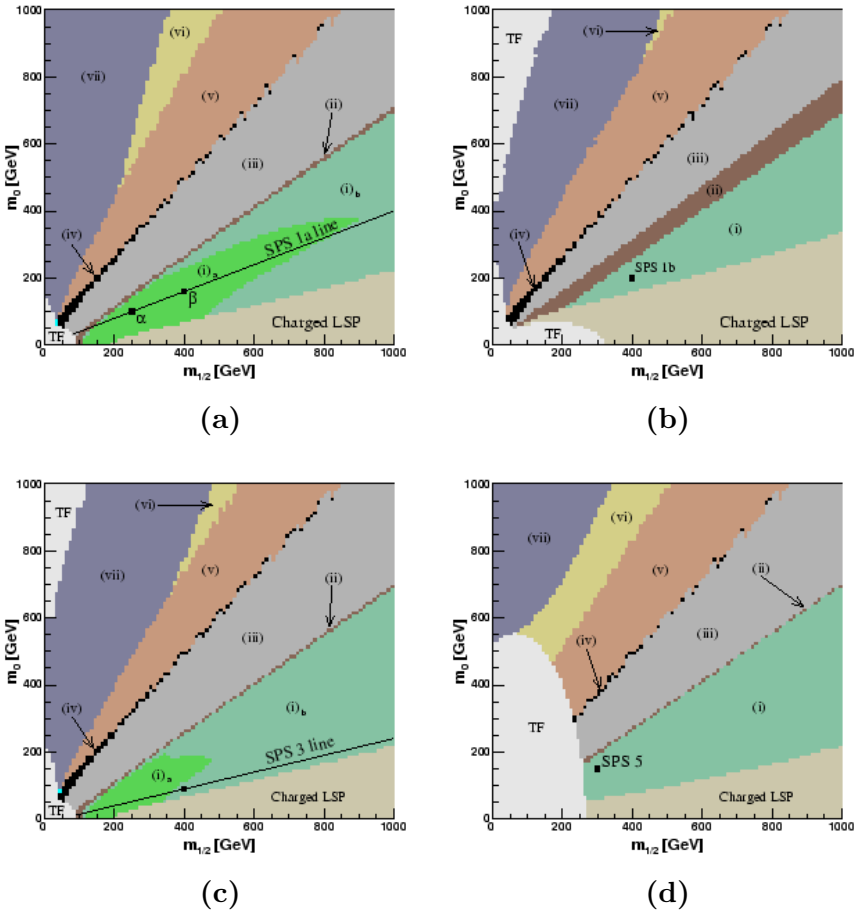


Figure 2.3: CMSSM mass hierarchies for different parameter choices: $A_0 = -m_0$, $\tan\beta = 10$ **(a)**, $A_0 = 0$, $\tan\beta = 30$ **(b)**, $A_0 = 0$, $\tan\beta = 10$ **(c)**, and $A_0 = -1000$ GeV, $\tan\beta = 5$ **(d)**. The dark and light green areas satisfy the conditions for the decay cascade (2.81). From [21].

are the same. Figure 2.4 shows the mass and decay spectrum of the SPS1a point in `SOFTSUSY`. The dashed lines indicate the decay channels, and the line width is proportional to the branching ratio. Decays with a branching ratio of less than 1% are not shown. Among the features to note is that the right-handed squarks have a much smaller branching ratio to $\tilde{\chi}_2^0$ than the left-handeds: 1% and 30%, respectively. This limits the amount of right-handed squarks initiating the chain, which is helpful for mass reconstruction since it limits the mass splitting between the squarks in the chain. Also, the left-handed sleptons are heavier than $\tilde{\chi}_2^0$, switching off that decay channel and limiting the slepton mass splitting.

The branching fractions of the χ_2^0 are divided between the the different $\tilde{l}l$ decay modes, with nearly 90% branching to $\tilde{\tau}\tau$ and the rest evenly divided between $\tilde{e}e$ and $\tilde{\mu}\mu$. Because the tau is short-lived and has a large branching fraction to

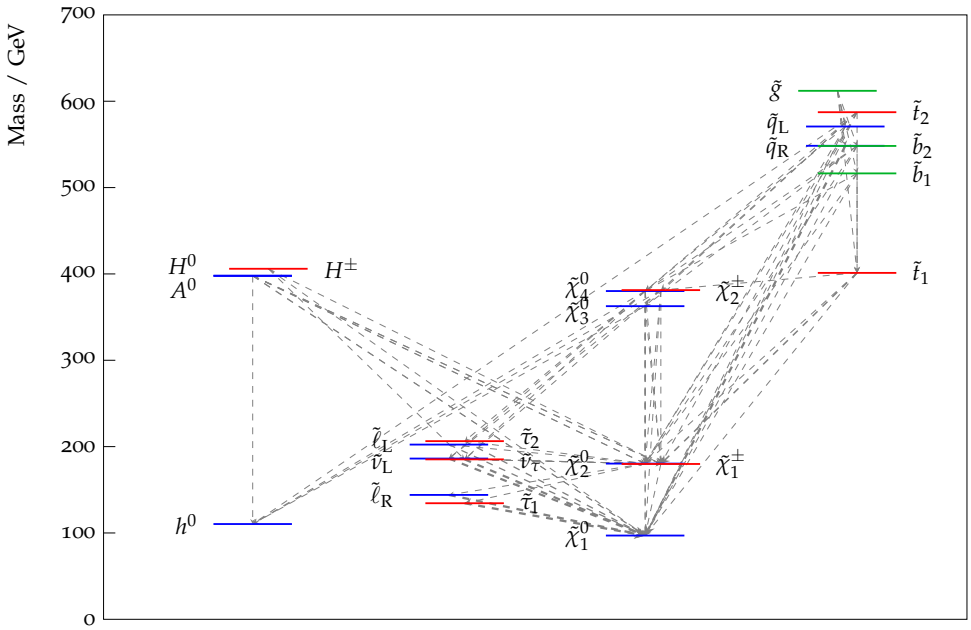


Figure 2.4: The mass spectrum of SPS1a from `SOFTSUSY`. The dashed lines show the decay channels. The figure is made using `PySLHA 3.0.2` [25]. See the text for details.

hadrons (65%) [19], it is difficult to use the cascade events which contain a tau for kinematical analysis, and this limits the number of useful chain events at SPS1a. For the analysis in the coming chapters, we will work with event samples of 25 events generated in the SPS1a parameter point at a centre-of-momentum energy of $\sqrt{s} = 14$ TeV, selected such that the leptons in the two chains are of different flavour. This corresponds to an integrated luminosity of 3 fb^{-1} [1].

2.10 The experimental status of supersymmetry

The search for supersymmetric particles is one of the main activities ongoing at CERN's Large Hadron Collider, particularly in the general purpose ATLAS and CMS experiments. Before the LHC was turned off for maintenance and upgrades at the end of 2012, each experiment had collected $\sim$ 20 fb^{-1} of data at a centre-of-momentum energy of $\sqrt{s} = 8$ TeV in what is called LHC Run I. No significant excesses of events that could indicate the existence of SUSY have so far been observed, with the exception of a few minor deviations in data that do not reach the level required for a discovery, the most significant being a 3σ excess in events consistent with a Z-boson, missing energy and jets [26]. At the time of writing, the LHC has begun circulating beams in preparation for Run

II, which will achieve a centre-of-momentum energy of $\sqrt{s} = 13$ TeV, and is expected to record ~ 300 fb^{-1} of data over the next ten years. In an even longer perspective, an upgrade to the *high luminosity LHC* is planned, which brings the total expected integrated luminosity up to 3000 fb^{-1} at $\sqrt{s} = 14$ TeV.

The discovery of a Higgs boson with [27]

$$m_h = 125.09 \pm 0.21\,(\text{stat.}) \pm 0.11\,(\text{syst.})\ \text{GeV}, \tag{2.82}$$

has important implications for supersymmetry. If we assume that it is the light SM-like Higgs of the MSSM, Eqs. (2.58) and (2.59) restrict combinations of the Higgs potential parameters $|\mu|$, $\tan\beta$ and $m^2_{H_{u/d}}$ at tree level. In corrections to its mass at loop-level, this Higgs impacts even more supersymmetric particles, in particular those with large couplings, such as the stop squark, as discussed in Section 1.8 and Appendix B.

Figure 2.5 shows the current reach of ATLAS in the CMSSM model which we discussed in Section 2.9. The plot shows the $(m_0, m_{1/2})$ plane for fixed values of $\tan\beta = 30$, $A_0 = -2m_0$ and $\mu > 0$. The choice of $\tan\beta$ and A_0 is made to accomodate a large Higgs mass consistent with (2.82), here shown as dashed–dotted blue lines. The various coloured (solid) lines indicate the lower bound for the parameter combination at the 95% confidence level in different searches, while the dashed lines are the expected exclusion limits for the given data assuming only the Standard Model. We see that for low values of m_0, where the decay chain in eq. (2.81) exists, searches for events with multiple jets and no leptons put the most constraining limits on the $m_{1/2}$ parameter.

Figure 2.6 shows the same kinds of exclusion limits, but now in the $(m_{\tilde{q}/\tilde{g}}, m_{\tilde{\chi}^0_1})$ mass plane for the MSSM. These searches assume simplified models with the mass hierarchy $m_{\tilde{q}/\tilde{g}} > m_{\tilde{\chi}^0_2} > m_{\tilde{l}} > m_{\tilde{\chi}^0_1}$ and all other sparticles decoupled.[5] The events considered in this analysis are required to have two (a) and four (b) jets, respectively, where none of the jets have been tagged as b-jets. Additionally, there must be two leptons of opposite sign and same flavour, which are not allowed to have an invariant mass between them that is compatible with a Z boson decaying. Also, there must be a significant amount (> 200 GeV) of missing transverse energy. We see that for a $\tilde{\chi}^0_1$ with mass below 200 GeV, squarks below 800 GeV are excluded, and below 300 GeV, gluinos below 1200 GeV are ruled out, thus putting considerable constraints on the decay chain we are interested in.

Referring back to Fig. 2.3, we see that the bright green region, where the chain hierarchy is present with a sizeable branching fraction, favours low values for m_0 and $m_{1/2}$ in the CMSSM. Comparing to the experimental constraints of figure 2.5, we see that much of these parts of the parameter space are ruled out. In particular we note that the SPS1a point is excluded. The top left region,

[5] Only the squarks of the first two generations are included. The third generation squarks are assumed to be decoupled.

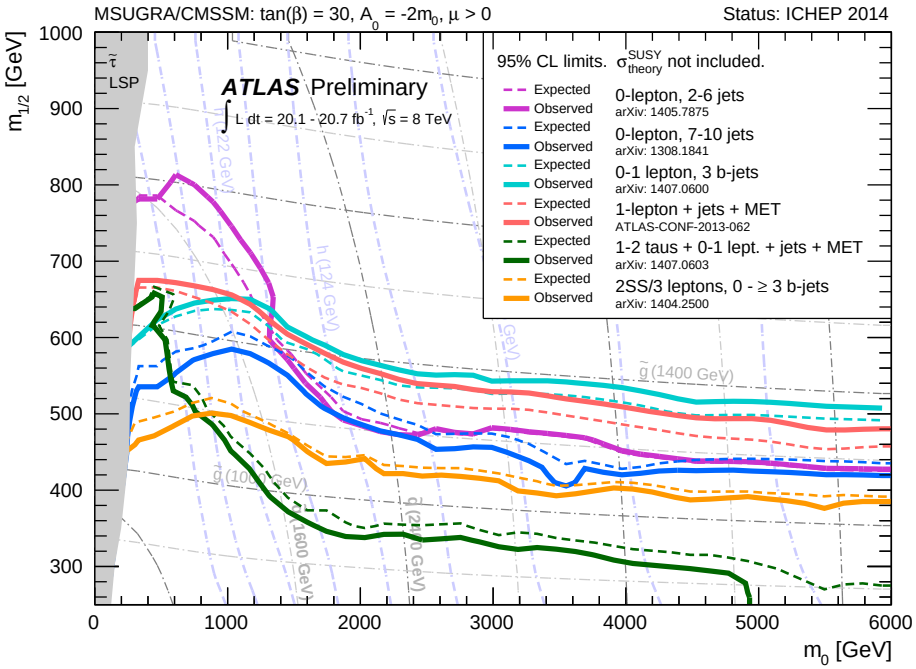
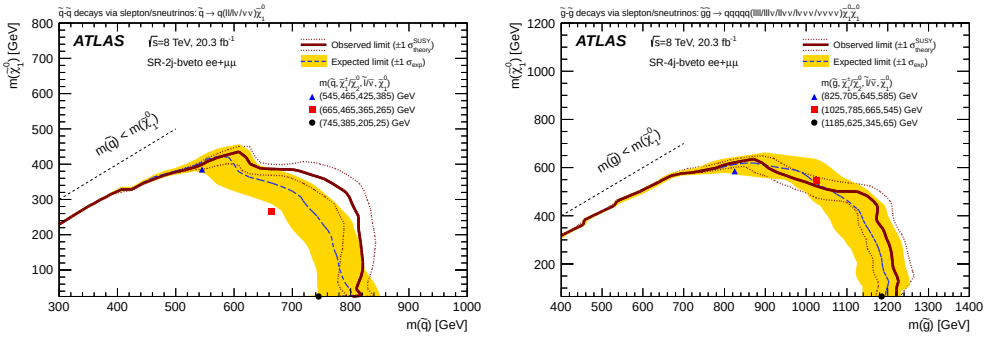


Figure 2.5: CMSSM limits from ATLAS following Run I [28–33].

where $m_{1/2} \gtrsim 800$ GeV and 1000 GeV $\gtrsim m_0 \gtrsim 400$ GeV, is compatible with the green areas of Fig. 2.3 and is not ruled out by the ATLAS limits themselves, but it corresponds to a Higgs mass $m_h \lesssim 123$ GeV, which is somewhat far from the measured value in (2.82). However, it should be noted that current state-of-the-art calculations of the Higgs mass in the MSSM (and CMSSM) come with theoretical errors of the order of 2 GeV [34], thus this region is not *per se* ruled out. There is also a question of what values are chosen for A_0 and $\tan\beta$, since the searches in Fig. 2.5 assume one specific choice. Tuning these parameters may bring closer agreement with the Higgs mass, but this has not been investigated here since it would require a full re-evaluation of the ATLAS searches.

As a benchmark for the event yield of the chain (2.81), we have calculated the expected number of events in which there are two such chains, for the parameter point $m_0 = 400$ GeV, $m_{1/2} = 800$ GeV, $A_0 = -800$ GeV, $\tan\beta = 30$, $\mu > 0$ for LHC Run II. This is the lowest $(m_0, m_{1/2})$ combination that is not excluded by Fig. 2.5 and at the same time provides the conditions needed for the chain. We calculate sparticle masses and branching fractions using `SOFTSUSY 3.4.1` [24] and squark and gluino cross sections for 13 TeV LHC using `NLL-fast 3.0` [35–41]. Assuming an integrated luminosity of 3000 fb^{-1} for a high-luminosity upgraded LHC,[6] we find an expectation of the order 10^{-5} events.

[6]The high-luminosity upgrade of the LHC will probably take data at $\sqrt{s} = 14$ TeV. There is, however, no complete implementation of `NLL-fast` for 14 TeV.



(a) **(b)**

Figure 2.6: Limits on the mass of $\tilde{q}$ **(a)** and $\tilde{g}$ **(b)** versus the $\tilde{\chi}_1^0$ mass in events with least two **(a)** or four **(b)** jets (with a b-jet veto), respectively, and a same-flavour dilepton pair. From ATLAS Run-I [26].

To match the measured Higgs mass and avoid the ATLAS direct searches, and at the same time keep the conditions needed for the chain as illustrated in Fig. 2.3, it appears that both the m_0 and $m_{1/2}$ parameters need to be pushed to or above 1000 GeV. Increasing the sparticle masses reduces the production cross section, which means that the expected number of events decreases even further. Thus, it seems that it will be difficult to measure the golden decay chain in the CMSSM at the LHC in a way that will be useful in the present context. However, by requiring only one chain rather than two, the expected event yield increases significantly. There are several measurements that can be done using only one chain [42]. For the aforementioned parameter point, the expected number of events with one chain is 0.10 — which is still very small, but it illustrates how much rarer the two-sided chain is.

By loosening some of the constraining assumptions of the CMSSM, such as the soft mass unification at a high scale, more general parametrizations of SUSY may be defined. The failure of the LHC to find signs of supersymmetry has prompted a large interest in studies of these more general models, in particular those termed "Natural SUSY", since the main guiding principle is to avoid the need for fine-tuning, such as the *little hierarchy problem* of the Higgs mass discussed in Appendix B. In these more general models, the squark and/or gluino masses may decouple, *e.g.* pushing the gluino and third generation squarks to high masses, while retaining light squarks of the two first generations. This may evade experimental constraints on the Higgs mass, and still give relatively large cross sections for squark production to initiate the chain. However, one would still be subject to the constraints shown in fig. 2.6 (a).

As a benchmark for this scenario, we have used `NLL-fast` to find the squark-pair production cross section at 14 TeV LHC, in the decoupling limit of large gluino mass. Assuming $m_{\tilde{q}} = 850$ GeV, we get $\sigma_{\tilde{q}\tilde{q}} = 0.258$ pb. If we assume,

as a best-case scenario, that the squarks have BR($\tilde{q}_L \to q\tilde{\chi}_2^0$) = 100%, BR($\tilde{q}_R \to q\tilde{\chi}_1^0$) = 100%,[7] and that BR($\tilde{\chi}_2^0 \to l\tilde{l}$) = 100% and BR($\tilde{l} \to l\tilde{\chi}_1^0$) = 100%, then the expected event yield with an integrated luminosity of 3000 fb^{-1} is about 200 000 events. This is an order of magnitude larger than for SPS1a.

[7]This corresponds to a scenario where $\tilde{\chi}_2^0 \sim \tilde{W}^0$, $\tilde{\chi}_1^0 \sim \tilde{B}^0$.

Chapter 3

Determination of SUSY particle masses from cascade decays

3.1 The problem

Consider an LHC event where two chains of the form

$$\tilde{q} \to \tilde{\chi}_2^0 q \to \tilde{l} l q \to \tilde{\chi}_1^0 l l q \tag{3.1}$$

are present. Combined, the measurable particles in the two chains are the two quarks and four leptons, where the lepton pairs are opposite-sign same-flavour. The LSPs escape detection, but the sum of their transverse momenta can be measured as the missing p_T. The quantities of interest, however, are the masses of the supersymmetric particles, $m_{\tilde{q}}, m_{\tilde{\chi}_2^0}, m_{\tilde{l}}$ and $m_{\tilde{\chi}_1^0}$. Potentially with several values for the squarks and sleptons if they differ in generation and/or chirality between the sides. These masses are not directly measurable, but the kinematics of the problem depend on them.

Many methods have been investigated for the purpose of measuring supersymmetric masses [3]. One well known example is the end-point method [43]. We measure *e.g.* the dilepton invariant mass in the process (3.1). The distribution of the dilepton invariant mass can be shown to form a right triangle where the maximal value is given by

$$(m_{ll}^{\max})^2 = \frac{\left(m_{\tilde{\chi}_2^0}^2 - m_{\tilde{l}}^2\right)\left(m_{\tilde{l}}^2 - m_{\tilde{\chi}_1^0}^2\right)}{m_{\tilde{l}}^2}, \tag{3.2}$$

thus constraining $m_{\tilde{\chi}_2^0}$, $m_{\tilde{l}}$ and $m_{\tilde{\chi}_1^0}$. Similar constraints may be obtained for the three other visible particle combinations, giving four equations with four unknowns. This method is very dependent on statistics, since each measured event only contributes one point to the distribution. A large number of events is required to get a reliable value. However, the number of events contributing to

the measurement is also much larger, since each side of the decay chain contribute individually, thus making use of events with leptons on one side only as well.

3.2 Webber's method

Webber [1] suggests a different method, where all available kinematical info from every event is used. Consider the general decay tree in fig. 3.1. We assume that we have an event with two such chains, but not necessarily with identical particles in the two. We will distinguish the two chains by referring to them as the "first" and "second" one, although the assignment is arbitrary. Assuming that

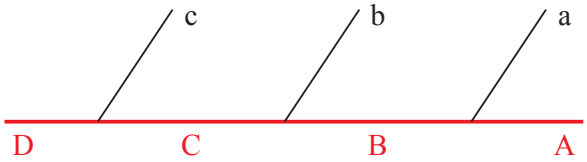


Figure 3.1: Decay topology, from [42].

the decaying particles are on-shell, the four-momenta in the first chain should satisfy

$$\begin{aligned}(p_c + p_b + p_a + p_A)^2 &= m_D^2,\\ (p_b + p_a + p_A)^2 &= m_C^2,\\ (p_a + p_A)^2 &= m_B^2,\\ p_A^2 &= m_A^2.\end{aligned} \tag{3.3}$$

The first three equations give three linear constraints on the invisible four-momentum p_A:

$$-2p_c \cdot p_A = m_C^2 - m_D^2 + 2p_c \cdot p_b + 2p_c \cdot p_a + m_c^2 \equiv S_1, \tag{3.4}$$
$$-2p_b \cdot p_A = m_B^2 - m_C^2 + 2p_b \cdot p_a + m_b^2 \equiv S_2, \tag{3.5}$$
$$-2p_a \cdot p_A = m_A^2 - m_B^2 + m_a^2 \equiv S_3. \tag{3.6}$$

Physically, these correspond to projections of the invisible particles on the measured ones. Equivalently the second chain (with primed indices) gets the constraints

$$-2p_{c'} \cdot p_{A'} = m_{C'}^2 - m_{D'}^2 + 2p_{c'} \cdot p_{b'} + 2p_{c'} \cdot p_{a'} + m_{c'}^2 \equiv S_5, \tag{3.7}$$
$$-2p_{b'} \cdot p_{A'} = m_{B'}^2 - m_{C'}^2 + 2p_{b'} \cdot p_{a'} + m_{b'}^2 \equiv S_6, \tag{3.8}$$
$$-2p_{a'} \cdot p_{A'} = m_{A'}^2 - m_{B'}^2 + m_{a'}^2 \equiv S_7. \tag{3.9}$$

In addition we have the transverse momentum constraints

$$\begin{aligned}p_A^x + p_{A'}^x &= p_{\text{miss}}^x \equiv S_4,\\ p_A^y + p_{A'}^y &= p_{\text{miss}}^y \equiv S_8.\end{aligned} \tag{3.10}$$

The vector $\mathbf{S} = (S_1, S_2, ...)$ thus depends on the eight unknown masses

$$\mathbf{M} = (m_D^2, m_C^2, m_B^2, m_{A'}^2, m_{D'}^2, m_{C'}^2, m_{B'}^2, m_{A'}^2) \tag{3.11}$$

and the visible momenta, which in principle are measurable. We define a vector containing the four-momenta of the invisible final state particles as

$$\mathbf{P} = (p_A^x, p_A^y, p_A^z, E_A, p_{A'}^x, p_{A'}^y, p_{A'}^z, E_{A'}). \tag{3.12}$$

We then have

$$\mathbf{AP} = \mathbf{S}, \tag{3.13}$$

where

$$\mathbf{A} = 2 \begin{pmatrix} p_c^x & p_c^y & p_c^z & -E_c & 0 & 0 & 0 & 0 \\ p_b^x & p_b^y & p_b^z & -E_b & 0 & 0 & 0 & 0 \\ p_a^x & p_a^y & p_a^z & -E_a & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{c'}^x & p_{c'}^y & p_{c'}^z & -E_{c'} \\ 0 & 0 & 0 & 0 & p_{b'}^x & p_{b'}^y & p_{b'}^z & -E_{b'} \\ 0 & 0 & 0 & 0 & p_{a'}^x & p_{a'}^y & p_{a'}^z & -E_{a'} \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \end{pmatrix}. \tag{3.14}$$

Furthermore, $\mathbf{S}$ may be written as

$$\mathbf{S} = \mathbf{BM} + \mathbf{C}, \tag{3.15}$$

where

$$\mathbf{B} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.16}$$

is a sparse matrix, and

$$\begin{aligned} \mathbf{C} = (&2p_c \cdot p_b + 2p_c \cdot p_a + m_c^2, 2p_2 \cdot p_3 + m_b^2, m_a^2, p_{\text{miss}}^x, \\ &2p_{c'} \cdot p_{b'} + 2p_{c'} \cdot p_{a'} + m_{c'}^2, 2p_{b'} \cdot p_{a'} + m_{b'}^2, m_{a'}^2, p_{\text{miss}}^y). \end{aligned} \tag{3.17}$$

With all of this, the solution for the invisible four-momenta, given the unknown masses, is

$$\mathbf{P} = \mathbf{A}^{-1}\mathbf{S} = \mathbf{DM} + \mathbf{E}, \tag{3.18}$$

where $\mathbf{D} = \mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{E} = \mathbf{A}^{-1}\mathbf{C}$.

The matrix $\mathbf{D}$ and vector $\mathbf{E}$ contain only measurable quantities, hence they only need to be calculated once for every event. For the true value of the unknown masses $\mathbf{M}$, the system should satisfy the on-shell conditions

$$\begin{aligned} p_A^2 &= P_4^2 - P_1^2 - P_2^2 - P_3^2 = m_A^2, \\ p_{A'}^2 &= P_8^2 - P_5^2 - P_6^2 - P_7^2 = m_{A'}^2. \end{aligned} \tag{3.19}$$

So by calculating $\mathbf{D}_n$ and $\mathbf{E}_n$ for each event n, and making a *hypothesis* $\mathbf{M}$ for the unknown masses, we can measure the goodness of fit for our hypothesis by the quantity

$$\xi^2(\mathbf{M}) = \sum_n \left[(p_A^2)_n - m_A^2\right]^2 + \left[(p_{A'}^2)_n - m_{A'}^2\right]^2. \tag{3.20}$$

Note that this quantity measures the goodness-of-fit of all the unknown masses equally, since it follows from the constraint equations (3.3) that for example

$$\begin{aligned} (p_B^2)_n - m_B^2 &= (p_a + p_A)_n^2 - m_B^2 \\ &= (p_a^2)_n + (p_A^2)_n + 2p_a \cdot p_A - m_B^2 \\ &= (p_a^2)_n + (p_A^2)_n - m_A^2 + m_B^2 - m_a^2 - m_B^2 \\ &= (p_A^2)_n - m_A^2. \end{aligned} \tag{3.21}$$

For each event there are eight constraining equations. There are eight unknowns from the masses plus six from the unknown LSP momenta (using the on-shell constraint on the LSPs). The system is thus underconstrained, with six free parameters. The point of the method is to minimize ξ^2 as a function of $\mathbf{M}$. This is generally an eight-dimensional minimization problem with a potentially very complicated behaviour, and thus not easy to solve. However, in the case of identical chains, it reduces to a much more handleable four-dimensional one which one could hope could be solved. In this case the number of free parameters reduces from six to two. The condition of identical chains can often be satisfied by a combination of vetoing (*e.g.* b-jets) and assuming small mass splittings between different generations, thus approximating their masses as equal. As was discussed in Section 2.9, this is a realistic assumption in many SUSY scenarios. Note that because the system is underconstrained, more than one event is needed to be able to solve for the masses.

Note that the method is applicable to a wider range of models than supersymmetry — it may be used in any model which adds Standard Model partner particles under a parity symmetry, for example in universal extra dimension (UED) with the *Kaluza-Klein* excitations cascade

$$q_1 \to qZ_1 \to qll_1 \to qll\gamma_1, \tag{3.22}$$

where γ_1 is the Kaluza-Klein excitation of the photon [44]. This is why we discuss the chain in the abstract notation used above. However, we will only consider supersymmetric applications in this thesis.

3.3 Two technical modifications

For the method to be applicable, the matrix $\mathbf{A}$ must be invertible. A matrix is invertible if its determinant is non-zero. However, the matrix $\mathbf{A}$ as it stands in (3.14) as defined by Webber [1], is technically ill-defined for inversion since not all rows have the same units. The rows 4 and 8, corresponding to the components 4 and 8 of the vector $\mathbf{S}$, have no dimension, while the other rows have dimension [mass]^1. This is reflected in the components of $\mathbf{S}$, where all, except 4 and 8, have dimension [mass]^2. This means both that the magnitude of the determinant is sensitive to the choice of mass scale, since some rows have non-zero dimension, and that it does not scale properly for numerical calculations, since not all rows have the same dimension. This is something that Webber does not comment on, and although it is more sensitive to numerical instability the method still in principle works, but we make some minor reformulations of the method in order to amend both problems.

For the first, we redefine S_4 and S_8 to be

$$\begin{aligned} S_4 &\equiv (p_A^x + p_{A'}^x)^2 = (p_{\text{miss}}^x)^2, \\ S_8 &\equiv (p_A^y + p_{A'}^y)^2 = (p_{\text{miss}}^y)^2. \end{aligned} \tag{3.23}$$

We do not wish to redefine $\mathbf{P}$ (3.12), so to keep the relationship $\mathbf{S} = \mathbf{AP}$ we modify rows 4 and 8 of $\mathbf{A}$ to

$$\begin{aligned} \mathbf{A}_4 &= (p_{\text{miss}}^x, 0, 0, 0, p_{\text{miss}}^x, 0, 0, 0), \\ \mathbf{A}_8 &= (0, p_{\text{miss}}^y, 0, 0, 0, p_{\text{miss}}^y, 0, 0), \end{aligned} \tag{3.24}$$

so that $\mathbf{A}$ now is

$$\mathbf{A} = 2 \begin{pmatrix} p_c^x & p_c^y & p_c^z & -E_c & 0 & 0 & 0 & 0 \\ p_b^x & p_b^y & p_b^z & -E_b & 0 & 0 & 0 & 0 \\ p_a^x & p_a^y & p_a^z & -E_a & 0 & 0 & 0 & 0 \\ p_{\text{miss}}^x/2 & 0 & 0 & 0 & p_{\text{miss}}^x/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{c'}^x & p_{c'}^y & p_{c'}^z & -E_{c'} \\ 0 & 0 & 0 & 0 & p_{b'}^x & p_{b'}^y & p_{b'}^z & -E_{b'} \\ 0 & 0 & 0 & 0 & p_{a'}^x & p_{a'}^y & p_{a'}^z & -E_{a'} \\ 0 & p_{\text{miss}}^y/2 & 0 & 0 & 0 & p_{\text{miss}}^y/2 & 0 & 0 \end{pmatrix}. \tag{3.25}$$

This redefinition does not alter the solvability of the problem, since the only information lost in $\mathbf{S}$ is the sign of p_{miss}^i which is kept in $\mathbf{A}$ instead. Also it keeps the essential feature that $\mathbf{A}$ only contains measured quantities, so that it can be inverted prior to making a mass hypothesis. The redefinition of $\mathbf{S}$ means we also have to modify $\mathbf{C}$ to keep the relationship $\mathbf{S} = \mathbf{BM} + \mathbf{C}$ from Eq. (3.15). We make the same redefinitions here, *i.e.*

$$\begin{aligned} C_4 &\equiv (p_{\text{miss}}^x)^2, \\ C_8 &\equiv (p_{\text{miss}}^y)^2. \end{aligned} \tag{3.26}$$

The other issue is to make the numerical problem dimensionless. All elements of $\mathbf{A}$ and $\mathbf{P}$ now have mass dimension 1, while all elements of $\mathbf{S}$, and thus $\mathbf{M}$ and $\mathbf{C}$, have dimension 2. We are free to multiply both sides of Eq. (3.13) by some normalization mass M_{norm} squared,

$$\frac{1}{M_{\text{norm}}^2}\mathbf{AP} = \frac{1}{M_{\text{norm}}^2}\mathbf{S}, \tag{3.27}$$

and we choose to take it into the matrix and vectors such that they all become dimensionless, *i.e.* we modify

$$\begin{aligned}\hat{\mathbf{A}} &= \frac{1}{M_{\text{norm}}}\mathbf{A},\\ \hat{\mathbf{P}} &= \frac{1}{M_{\text{norm}}}\mathbf{P},\\ \hat{\mathbf{S}} &= \frac{1}{M_{\text{norm}}^2}\mathbf{S},\end{aligned} \tag{3.28}$$

thus modifying $\mathbf{M}$ and $\mathbf{C}$ in the same way as $\mathbf{S}$ to comply with Eq. (3.15). We also modify the fitting function ξ^2 accordingly, so that it becomes

$$\xi^2(\mathbf{M}) = \sum_n \left[(\hat{p}_A^2)_n - \frac{m_A^2}{M_{\text{norm}}^2}\right]^2 + \left[(\hat{p}_{A'}^2)_n - \frac{m_{A'}^2}{M_{\text{norm}}^2}\right]^2. \tag{3.29}$$

To obtain numbers of order one, which is optimal for numerical purposes, we should pick a mass of the relevant scale for the problem. This is not something that is known *a priori*, since it depends on the supersymmetric masses that we are trying to determine. We might be tempted to use something based on the measured momenta, but this is a bad idea since it would mean weighting different events differently. We choose the normalization constant

$$M_{\text{norm}} = 100 \text{ GeV}, \tag{3.30}$$

the same order of magnitude as we expect for the supersymmetric masses (electroweak scale).

We have made thorough checks that this formulation and the original one produce identical results within numerical accuracy, so that indeed the formulations are equivalent. The normed matrix and vectors $\hat{\mathbf{A}}$, $\hat{\mathbf{P}}$ and $\hat{\mathbf{S}}$ are the ones we will use throughout the rest of the thesis, and we will refer to them without the hats, *i.e.* $\mathbf{A}$, $\mathbf{P}$ and $\mathbf{S}$.

3.4 Taking account of combinatorics

In a real detector event, the ordering of the quarks and leptons in and between chains is not known — all we have are the measured particle types and their

momenta. We must take this into account when applying the method to Monte Carlo simulated datasets. Webber does this by evaluating all possible combinations in each event at each mass point and selecting the combination which gives the lowest ξ^2 value, choosing to add this value to the sum in Eq. (3.29). The number of possible combinations are 8 or 16, depending on whether the lepton pairs in the two chains are the same flavour or not.

For two pairs of different-generation leptons, the possible orderings are (given some 'base ordering' which we permute from): Switching the ordering of the leptons in the first chain, switching the ordering of the leptons in the second chain, or switching the leptons in both chains.[1] For each of these permutations we have the option to switch the two quarks, so the total number of combinations is 8. In the case of identical leptons, we may additionally interchange leptons between the chains — but this only increases the total combinations by a factor of 2 because the same-chain leptons must have opposite charge.

Note that in order to switch the ordering of near and far leptons within the same chain, all we need to do is permute the rows of the matrix $\mathbf{A}$. The vector $\mathbf{C}$ is invariant as long as the same-chain leptons have the same mass. When the quarks are flipped, however, or when leptons are interchanged between chains, then we must redefine $\mathbf{A}$ and $\mathbf{C}$. Webber makes a point that this property can save computer memory, since one only has to store two or four versions of the matrix and vector for each event. We have not found this to be an issue.

3.5 Outline of our plan

In the following chapters we will investigate and develop this method further. The plan of attack is as follows:

1. We begin by attempting to reproduce Webber's parton level results using Monte Carlo simulations. As a benchmark we investigate the precision attainable for the CMSSM parameter point SPS1a [22], which was used by Webber [1]. We also apply a crude momentum smearing, in order to simulate that the measurement of the kinematics of final-state particles is not exact, as well as a quality cut on the best-fit value of the ξ^2 function to eliminate bad event samples.

2. We proceed to analyze the effects of changing the settings of the numerical routine used for minimizing the ξ^2 function. In light of this, we present revised estimates for the precision attainable for the method in the different cases with and without momentum smearing.

[1] We will in the following refer to the lepton which is nearest to the quark in the decay chain as the *near* lepton, and the other as the *far* lepton.

3. We then compare results with and without including the combinatorical issues from identifying the decay products. We also investigate the effects of using a more sophisticated parametrization of the momentum smearing, based on realistic detector response.

4. We proceed with investigating potential improvements to the method, such as using the value of the determinant of the matrix $\mathbf{A}$ to discard bad events, and schemes for reducing the combinatorical ambiguities.

5. The partons that emerge from the hadron showers will hadronize, forming a hadron jet before arriving in the detector. Measurement of the initial parton from reconstructing such jets is one of the big challenges of collider physics. We use the `FastJet` [45] program for jet reconstruction, with algorithms used in the LHC experiments, and study the effect of jet reconstruction on the mass measurement.

6. In an analysis of real data, one would have to use selection criteria such as cuts on transverse momentum and number of jets and leptons to discriminate between signal and background events. We will apply such cut schemes to our simulated events based on expectations for 14 TeV LHC, and present estimates for the precision we can expect from the method in this case.

Chapter 4

Investigating Webber's method by Monte Carlo simulations

Webber demonstrates the aptitude of the method on a Monte Carlo generated dataset. A natural starting point for our study is to try to reproduce his results, which is the main focus of the present chapter.

4.1 Reproducing Webber's results

There are several ways in which a pair of squarks can be produced in pp collisions. The three main categories are: direct pair production of two squarks or a squark-antisquark pair, squark plus gluino with subsequent gluino decay to a squark, and pair production of two gluinos which both subsequently decay to squarks. The mechanism of production affects how much momentum goes into the squark, and thus the momentum boost of the subsequent decay chain. However, the quantity which is reconstructed, p_A^2, is Lorentz invariant. Still, the numerical aspects of the reconstruction depend on the measured momenta. The three different categories also affect how many hard jets are present in the event, which affects the combinatorical aspects of reconstructing the chain. This problem we will return to in Chapter 5.

In [1], Webber studies the benchmark CMSSM parameter point SPS1a [22] which we discussed in Section 2.9.1. He uses `HERWIG 6.510` [46, 47] to produce events, selecting only events with two left-handed first- or second generation squarks (to limit the mass splitting), but irrespective of the hard production process. He further specifies to only include events where the lepton pairs of the two chains are of different flavour, and require them both to be of first or second generation. He uses the quark four-momenta before parton showering, which means that the quarks potentially are off-shell, and takes the quark masses to be the invariant mass of these four-momenta.

The analysis, *i.e.* minimization of the ξ^2, is performed on 100 samples of

25 events each, using the `Minuit` [48] `Simplex` routine for minimization. He models the effects of measurement errors in a real detector by applying a common momentum smearing on all particles according to a gaussian distribution, and he investigates a cut on the total ξ^2 obtained at the minimum to eliminate samples which give a bad result. His final results are summarized in Table 1 of [1], which we for convenience show in Table 4.1. The four columns labeled m_i show the mean values and r.m.s. errors of each of the fitted masses in the samples, with the true masses shown in parenthesis. The column $\delta p/p$ indicates the standard deviation of the gaussian smearing applied on the momenta, ξ^2_{max} indicates the cut value on the ξ^2, f_ξ is the fraction of samples surviving the cut and f_{corr} is the fraction of events where the chosen particle combination, selected as described in Section 3.4, is the correct one. Note that if a sample is rejected by the ξ^2 cut, it means that this sample would have failed experiment, since each sample is a separate simulation of a run of the LHC.

$\delta p/p$	ξ^2_{max}	f_ξ	f_{corr}	$m_{\tilde{q}}(540)$	$m_{\tilde{\chi}^0_2}(177)$	$m_{\tilde{l}}(143)$	$m_{\tilde{\chi}^0_1}(96)$
0	∞	100%	72%	538 ± 20	176 ± 12	143 ± 7	95 ± 10
0	100	80%	76%	539 ± 7	177 ± 1	144 ± 1	96 ± 2
5%	∞	100%	52%	534 ± 28	176 ± 11	143 ± 10	95 ± 13
5%	100	57%	55%	539 ± 9	178 ± 3	144 ± 2	96 ± 4
10%	∞	100%	40%	522 ± 37	171 ± 18	140 ± 17	88 ± 26
10%	200	42%	43%	530 ± 22	173 ± 12	140 ± 12	89 ± 20

Table 4.1: Webber's table of results, taken from Table 1 of [1].

To reproduce Webber's results, we have mainly generated events using `Herwig++ 2.7.1` [49]. As a control, and to enable us to interface the simulation with jet reconstruction software later, we have also used `Pythia 8.2` [50]. We have also had access to the `HERWIG 6.510` code used in the original paper, enabling us to make a quite close reconstruction of Webber's analysis [4]. The reproductions of Webber's results in the present chapter are done with the `HERWIG` code. In Section 4.5 we switch to a `Herwig++` generated dataset, which we also use through most of Chapter 5, before we employ `Pythia` to study jet reconstruction effects in Section 5.6. Even though the SPS1a parameter point is excluded by experimental data, as discussed in Section 2.9.1, we choose to use it for our study to be able to compare our results as closely as possible to Webber's.

To minimize the ξ^2 function, we have used the `Simplex` algorithm. For the reproduction of Webber's results in the present chapter we have used the `Minuit` implementation, which was also used in [1]. For the subsequent investigation of improvements, we have also used a custom implementation listed in Appendix D. In the course of our investigation, we have discovered that the mass fit is heavily dependent on the input parameters to the `Simplex` minimization, and this makes the fit more challenging. To facilitate the subsequent discussion, we

briefly introduce the Simplex method.

4.2 The Nelder-Mead Simplex algorithm

Simplex [51] is a heuristic minimization search method for minimizing a scalar function of N variables. It takes a starting parameter point as input from the user. From this parameter point it erects a *simplex*, an $N+1$-dimensional generalization of a triangle, by slightly increasing each of the coordinates of the parameter point in turn.[1] It then begins to evaluate the function in the vertices of this simplex. A new simplex is constructed by reflecting the vertex with the highest function value around the (hyper)line made out of the other N vertices. Hopefully this new simplex lies lower in the function terrain, and thus the algorithm iterates towards a local minimum. In case of trouble, it may also try contracting the simplex or distorting its shape in various ways to obtain points of lower function values.

Since the method is heuristic, so is the convergence criterion. Convergence is said to be obtained when the *estimated distance to minimum (EDM)* is smaller than some set tolerance value. Usually there is also a predefined maximal number of iterations before the method gives up, to avoid going on forever on non-converging problems. The EDM is commonly defined as

$$\mathrm{EDM}(f_{\min}, f_{\max}) = \frac{|f_{\max} - f_{\min}|}{|f_{\max}| + |f_{\min}|}, \tag{4.1}$$

where $f_{\min}$ and $f_{\max}$ are the function values at the lowest and highest point of the current simplex, respectively. This means that the convergence criterion really measures how "flat" the simplex is, and thus how steep the function is in the region. If the tolerance is too high, then we run the risk of obtaining convergence in a region where the gradient is not steep enough to be resolved by the set tolerance, but which may still be far from the minimum.

A pitfall of any minimization routine, also for Minuit Simplex, is that it has a default tolerance value which is used automatically unless the user specifically changes it. The default tolerance in Minuit Simplex is 0.1. This appears to be what Webber has used. We have confirmed that we obtain statistically consistent results when choosing that value. However, this does not always resolve our particular function well enough, because it tends to have almost flat directions in mass space for some sets of events. This is illustrated in Fig. 4.1, which shows contour plots of the ξ^2 for one sample of 25 events as a function of pairs of mass parameters. We see that in this sample, the $\tilde{\chi}_1^0$ direction is flat at low mass

[1] The simplex can be constructed in various ways. Another option is to scale the coordinates by a factor, rather than adding a constant value. The behaviour of the minimization turns out to be very similar in this case.

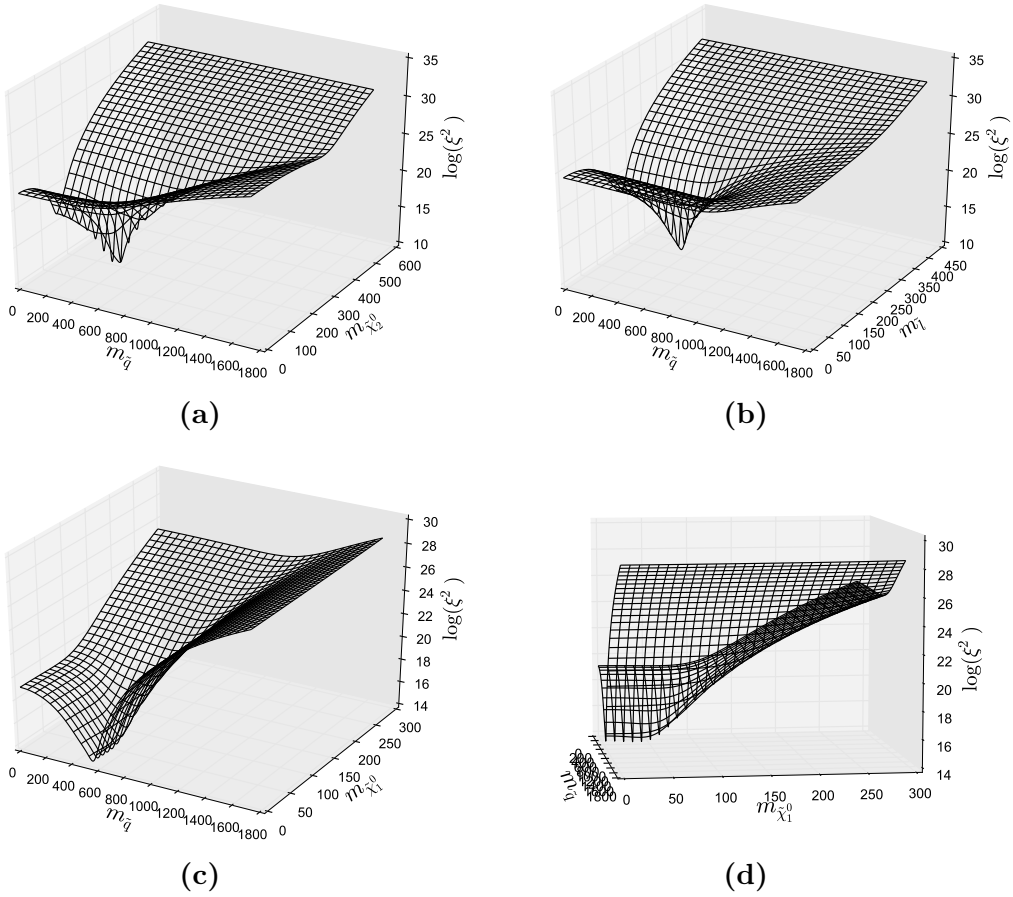


Figure 4.1: 3D contour plot of $\log(\xi^2)$ in $(m_{\tilde{q}}, m_i)$ plane around the point of true minimum, where $i = \tilde{\chi}_2^0$ for **(a)**, $i = \tilde{l}$ for **(b)** and $i = \tilde{\chi}_1^0$ for **(c)** and **(d)**. The other two masses are in each case fixed to their true value.

values, a feature which stretches down to $m_{\tilde{\chi}_1^0} = 0$ GeV. Note, however, that there may be additional higher-dimensional degeneracies in the four-mass space which will not show in a two-dimensional plot.

Setting the tolerance high may therefore lead to convergence at a non-minimal point. If, additionally, the search is started at or close to the masses used to generate the Monte Carlo, then the minimization may obtain "convergence" at points very close to the true value, but these points are not minimal points, just regions where the function is not very steep.

4.3 The effect of tolerance

Refer to the scatter plot in Fig. 2 of [1], displayed in Fig. 4.2 for convenience. It is a scatter plot of the best-fit mass points in the fit corresponding to the

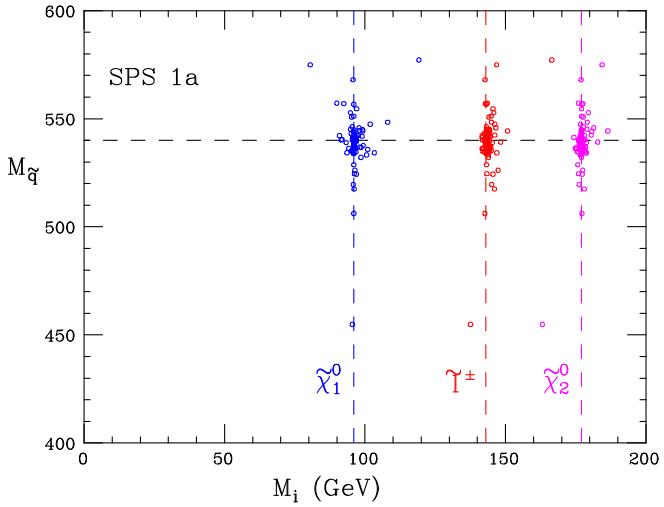


Figure 4.2: Figure 2 from [1], showing the best-fit points for each sample corresponding to the first row of Table 4.1, plotted as functions of pairs of masses.

first row of Table 4.1. We have reproduced this fit using Webber's code [4], but we have generated our SPS1a parameter point using `SOFTSUSY 3.4.1` [24] rather than `ISAJET 7.59`, as discussed in Section 2.9.1. We have converted the resulting SLHA [52] model file to the `ISAWIG` format readable by `HERWIG 6.510` using the package `PySLHA 3.0.2` [25].[2] The effects of using a different RG-runner is that the SUSY masses are not exactly the same. The most significant shift is the squarks, which in Webber's case have a mass of 540 GeV, compared to 565 GeV in our case. We observe similar results as Webber gives in his article when we run his code with the original settings for our SUSY parameter point. Our reproduction of Fig. 4.2 is shown in Fig. 4.3a,[3] and our reproduction of Table 4.1 is given in Table 4.2.

However, the tolerance setting in `Minuit` can be adjusted. When we rerun the code used to produce Fig. 4.3a with the tolerance set to 10^{-12},[4] we get the fit shown in 4.3b and listed in Table 4.3. The results are not dramatically altered, but there are some features to notice: There is a clear tendency to a linear correlation between the masses. This is a feature we should expect physically from the kinematics: By co-varying all the masses, the fit does not change much. This can be seen in the matrix dependence on the masses.[5] The fact that this

[2] We have had to make several modifications to the `PySLHA` code to make the `ISAWIG` output readable by `HERWIG 6.510`. These changes have been included in `PySLHA 3.0.3` by the author.

[3] All the plots in this thesis are in vector graphics format. When the thesis is viewed electronically, they are zoomable.

[4] This will be referred to as *low tolerance* below.

[5] This is part of the reason why these kinds of mass reconstruction methods very often reconstruct the squared mass *difference* rather than the masses themselves.

$\delta p/p$	$\xi^2_{\max}$	f_ξ	f_{corr}	$m_{\tilde{q}}(568)$	$m_{\tilde{\chi}^0_2}(180)$	$m_{\tilde{l}}(144)$	$m_{\tilde{\chi}^0_1}(97)$
0	∞	100%	65%	566 ± 9	180 ± 2	144 ± 2	97 ± 3
0	100	85%	67%	567 ± 6	180 ± 1	144 ± 1	97 ± 3
5%	∞	100%	43%	564 ± 26	181 ± 14	145 ± 10	94 ± 15
5%	100	52%	48%	566 ± 10	180 ± 2	145 ± 2	96 ± 4
10%	∞	100%	33%	551 ± 33	180 ± 15	144 ± 11	91 ± 24
10%	200	43%	36%	559 ± 17	177 ± 11	143 ± 11	91 ± 20

Table 4.2: Our reproduction of Table 4.1, using Webber's code [4] with original settings, except with the SPS1a masses from SOFTSUSY.

physically reasonable degenerate direction appears when the tolerance is reduced indicates that a such a reduction is necessary to achieve reliable results. We also see that the fitted masses now seem slightly biased toward lower values. Finally, we note that while the mean value and errors are still consistent with the true values, their accuracy is somewhat reduced. Particularly so for the LSP, where the fit is reduced from 99 ± 3 GeV to 83 ± 19 GeV, compared to the true value of 97 GeV.

These fit results, with the low tolerance setting, are still not bad. However, in Table 4.1, Webber also gives best-fit values where he has applied smearing to the four-momenta, as a crude approximation to the effects of limited detector resolution. He does this by smearing the spatial components according to a gaussian distribution of a given r.m.s. width $\delta p/p$, and then setting the energy component to the value which leaves the invariant mass of the four-vector unchanged. In Fig. 4.4 we show scatter plots of the fits to the dataset smeared with $\delta p/p = 5\%$, minimized with original and reduced tolerance, again using Webber's code for event generation and minimization. The fit with original tolerance is consistent with Fig. 3 of [1], as it should be. However, when the tolerance is reduced, the fit results are worsened considerably. Since each event is smeared individually, this appears to greatly affect the position of the minimum. Again we see that the LSP (yellow) recieves the roughest treatment, being pushed to much lower values than the true one in most cases. The results are even worse for the 10% smeared dataset.

However, Webber also investigates the effects of imposing a *cut* on the ξ^2 value obtained at the minimum. In his fits, this cut tends to remove the bad events, giving a better fit. Applying a cut also helps for the reduced-tolerance fit, although it does not recover Webber's original low error estimate. In Table 4.3 we reproduce Table 4.2 for the reduced-tolerance fit. We note that the fraction of samples passing the ξ^2 cut is drastically reduced compared to Table 4.2. The fraction of events where the best-fit combination is the true one is also reduced. The fact that the fraction of samples passing the ξ^2 cut decreases is counter-intuitive — one would expect that a lower tolerance ensures that the fit obtains

a lower ξ^2 value. However, the form of the ξ^2 is very complicated due to the combinatorical ambiguities, and it may be difficult for `Simplex` to handle.

$\delta p/p$	$\xi^2_{\max}$	f_ξ	f_{corr}	$m_{\tilde{q}}(568)$	$m_{\tilde{\chi}^0_2}(180)$	$m_{\tilde{l}}(144)$	$m_{\tilde{\chi}^0_1}(97)$
0	∞	100%	36%	563 ± 13	173 ± 10	136 ± 11	83 ± 19
0	100	35%	52%	565 ± 9	175 ± 8	138 ± 9	86 ± 16
5%	∞	100%	31%	557 ± 27	165 ± 17	125 ± 15	58 ± 27
5%	100	13%	43%	558 ± 14	164 ± 11	126 ± 12	65 ± 22
10%	∞	100%	29%	542 ± 35	158 ± 20	116 ± 17	36 ± 28
10%	200	15%	33%	549 ± 20	155 ± 12	116 ± 12	38 ± 25

Table 4.3: Reproduction of the fits in Table 4.2, but with reduced convergence tolerance.

4.4 Starting point dependence of the best-fit point

There is also another potential issue with Webber's analysis. It has to do with the fact that the best-fit search is started at the true mass values. In a real experiment, these parameters are the unknowns we wish to find. Starting the search here is in principle fine as long as we are sure that the algorithm always finds the true global minimum. So we must investigate what happens if we start our search in some other point. We have done this for a motivated set of starting points, and discover that this greatly affects the location of the best-fit points.

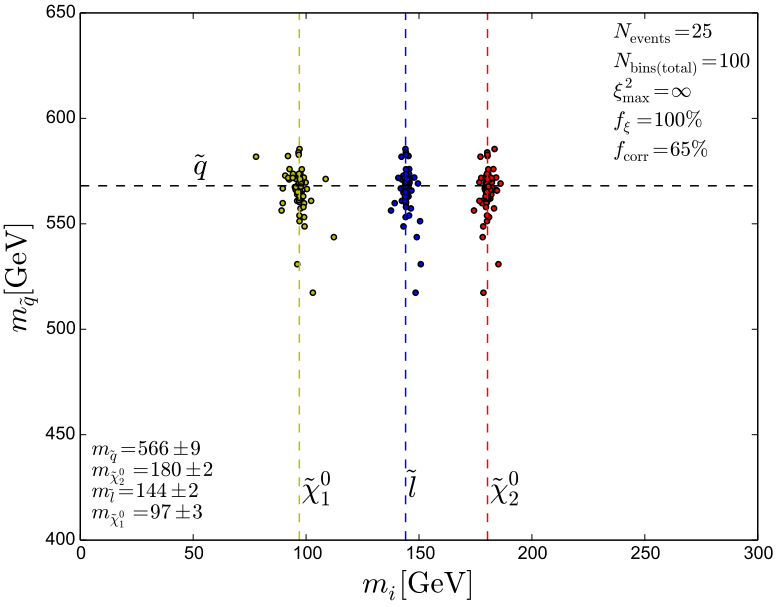
In Fig. 4.5 we show the best-fit points for four low-tolerance minimizations on the unsmeared `HERWIG 6.510` dataset. Figure 4.5a is the same as Fig. 4.3b, the minimization started from the true mass point (TMP). The other three are minimizations from starting points selected to illustrate other plausible mass spectra: one where both the masses and the mass differences are smaller (b), one where they are larger (c), and one where there is a large splitting between a heavy squark and three very light masses (d). It is obvious that the results of the fit is hugely dependent on where we start the search. This concerns both the location of the best-fit point and the number of samples where convergence, as defined in Section 4.2, is obtained within 500 iterations, indicated by the number N_{bins} in each plot. For instance, the mean value and standard errors of the squark masses for the samples range from 506 ± 98 GeV to 563 ± 13 GeV. We also note that the margins of error in the latter case, which is minimization from the TMP, exclude the mean values obtained from the other three starting points.

It might, however, be that the function has multiple local minima, giving rise to the behaviour in Fig. 4.5, but that the *global* minimum is the one we find by starting in the TMP. To investigate this, we have run minimizations (with low

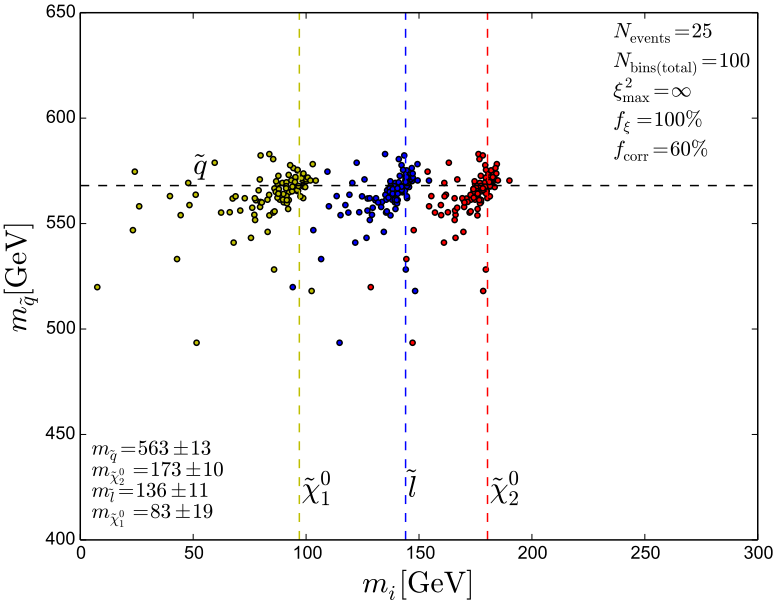
tolerance) where we perturb the starting point of each of the four parameters for each event sample away from the TMP by a gaussian distribution of width 10% or 20% of the TMP mass values. This is a minor perturbation relative to the masses. The minimization results are shown in Table 4.4 for perturbations of 10% (top) and 20% (bottom). The standard errors of the fit increase considerably compared with Table 4.3.

$\delta p/p$	$\xi^2_{\max}$	f_ξ	f_{corr}	$m_{\tilde{q}}(568)$	$m_{\tilde{\chi}^0_2}(180)$	$m_{\tilde{l}}(144)$	$m_{\tilde{\chi}^0_1}(97)$
10%:							
0	∞	100%	44%	553 ± 41	165 ± 23	125 ± 26	53 ± 43
0	100	58%	52%	562 ± 16	171 ± 16	133 ± 18	78 ± 28
5%	∞	100%	36%	547 ± 46	159 ± 21	117 ± 23	38 ± 39
5%	100	46%	47%	558 ± 18	165 ± 18	126 ± 20	62 ± 32
10%	∞	100%	30%	540 ± 46	151 ± 22	110 ± 24	12 ± 28
10%	200	39%	35%	545 ± 21	152 ± 18	113 ± 20	24 ± 34
20%:							
0	∞	100%	40%	540 ± 58	156 ± 30	116 ± 32	44 ± 42
0	100	50%	56%	562 ± 16	171 ± 15	133 ± 16	77 ± 28
5%	∞	100%	32%	548 ± 49	159 ± 24	115 ± 28	33 ± 39
5%	100	37%	42%	552 ± 19	160 ± 18	122 ± 18	59 ± 30
10%	∞	100%	27%	529 ± 51	150 ± 26	107 ± 22	11 ± 27
10%	200	38%	33%	543 ± 23	151 ± 20	113 ± 19	24 ± 33

Table 4.4: Reproduction of the fits in Table 4.3 with random perturbations of 10% and 20% of the TMP masses, respectively, on the starting points of the `Simplex` search.

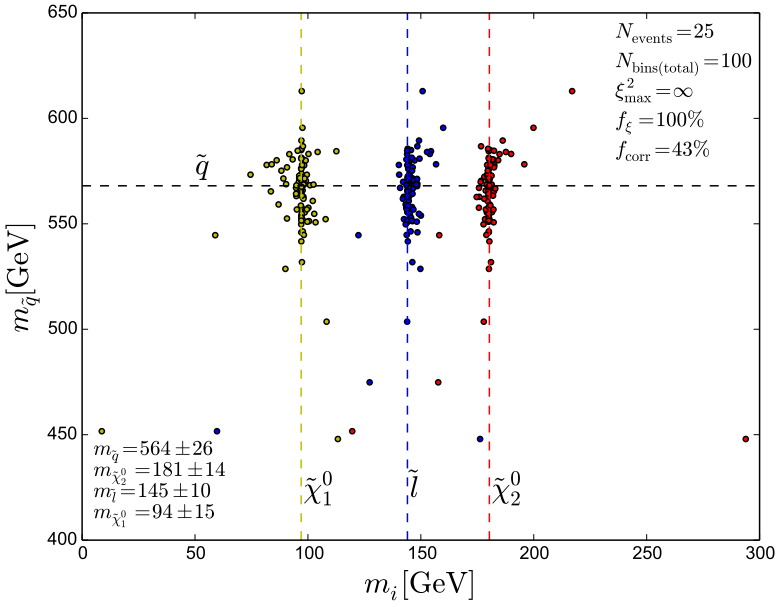


(a)

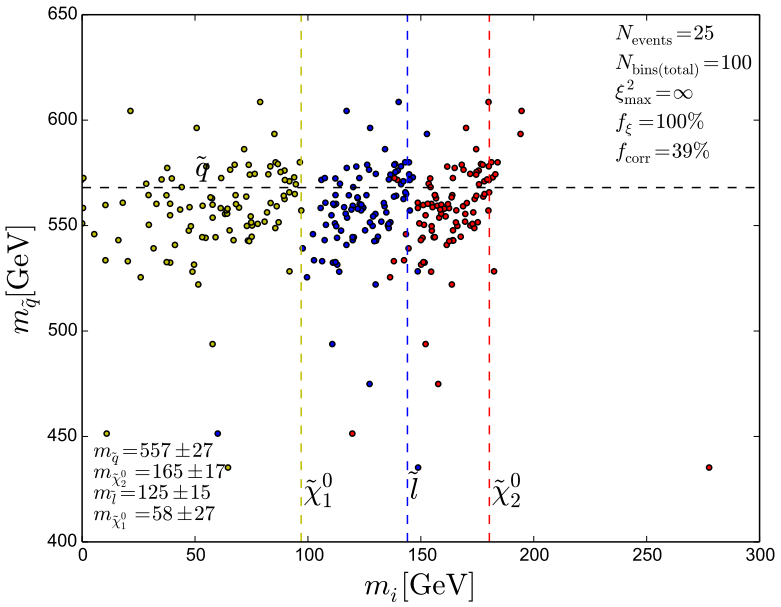


(b)

Figure 4.3: Reproduction of Webber's results corresponding to fig. 4.2 and the first row of Table 4.1 for **(a)** original convergence tolerance and **(b)** a lower tolerance criterion of 10^{-12}.



(a)



(b)

Figure 4.4: Reproduction of Webber's 5% momentum-smeared fit, corresponding to the third row of Table 4.1, for **(a)** original convergence tolerance and **(b)** the lower tolerance criterion.

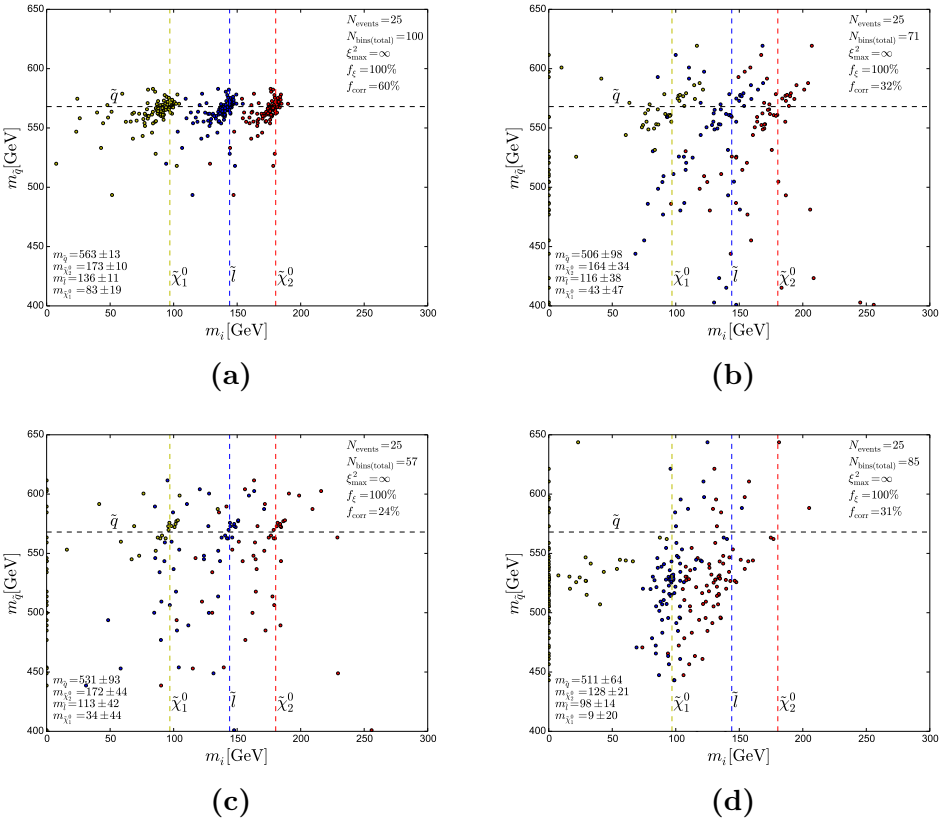


Figure 4.5: Minimization on the unsmeared `HERWIG 6.510` dataset for different starting points: $\vec{M} = (568, 180, 144, 97)$ GeV (the TMP) in **(a)**, $\vec{M} = (400, 300, 200, 100)$ GeV in **(b)**, $\vec{M} = (800, 500, 300, 50)$ GeV in **(c)** and $\vec{M} = (1000, 100, 80, 30)$ GeV in **(d)**.

4.5 Sans combinatorics

If we for a moment forget about the combinatorics, and evaluate only the true particle combination for each event, then the minimization gives consistent results irrespective of starting point. This is illustrated in Fig. 4.6, which shows

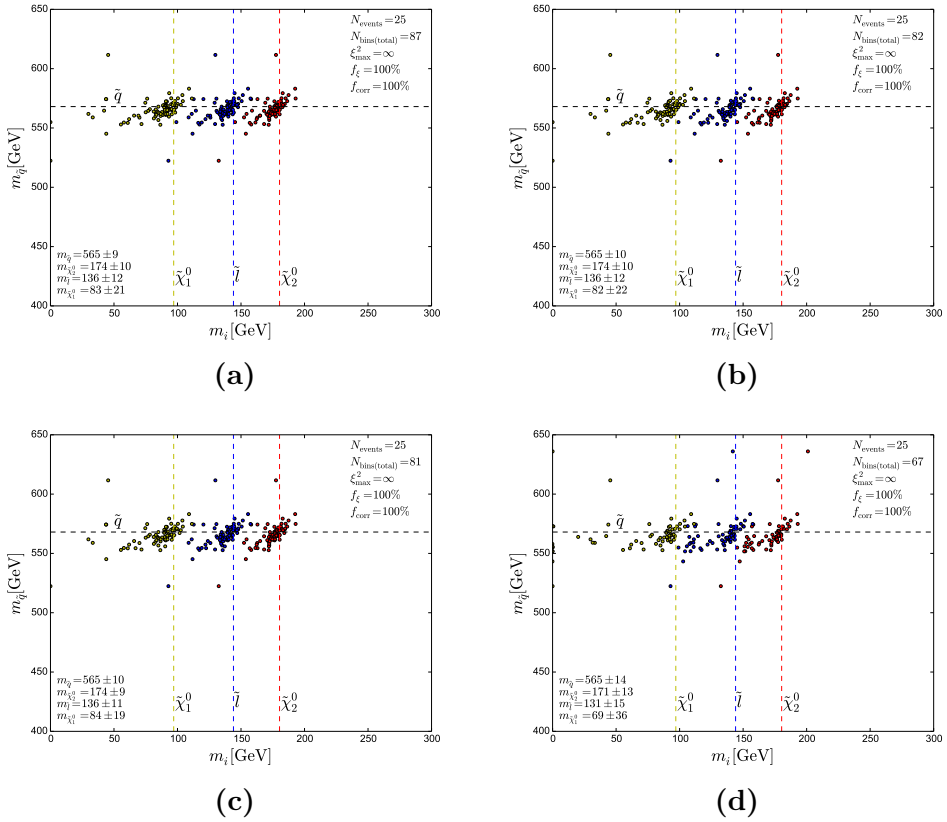


Figure 4.6: An equivalent fit to fig. 4.5 (on a `Herwig++` dataset), however, the ξ^2 contribution is only evaluated for the true particle combination in each event.

minimization of 100 samples of 25 events minimized with low tolerance, but only evaluated with the correct combination of chain particles.[6] There are differences between the fits, mainly between Fig. 4.6d and the others. In particular, while about 85 of the samples converge within the set limit of 500 iterations in each of the cases a, b and c (indicated by the value of $N_{\text{bins(total)}}$ in the top right of

[6] This dataset is generated with `Herwig++ 2.7.1` [49] and minimized using our own implementation of the `Simplex` algorithm in C++, included in Appendix D. We have checked that this dataset gives consistent results with the `HERWIG 6.510` program used in Chapter 4. The reason we switch is that the `HERWIG 6.510` code [4] does both Monte Carlo generation and `Simplex` minimization together, without storing event data, while we want to be able to rerun the minimization without regenerating the Monte Carlo.

each plot), this number is reduced to 67 in d. The starting point in case (d) is characterized by a much larger mass gap between the squark and the $\tilde{\chi}_2^0$ than in the TMP, which we might surmise would give rise to very different kinematics than we have in our events. We tentatively conclude that the combinatorics is the main culprit in making the method sensitive to the starting point.

We also keep in mind the effects of momentum smearing, and check the no-combinatorics minimization on the 5% smeared dataset for the same four starting points. The plots are shown in Fig. 4.7. We find that it also in this case gives consistent results irrespective of starting point, with mean best-fit values of $m_{\tilde{q}} = 556 \pm 33$ GeV, $m_{\tilde{\chi}_2^0} = 156 \pm 21$ GeV, $m_{\tilde{l}} = 112 \pm 20$ GeV and $m_{\tilde{\chi}_1^0} = 22 \pm 34$ GeV — but the LSP mass is fitted to zero in 63 of 94 samples. It appears that when the data are smeared, the method loses its sensitivity to the LSP mass. Or equivalently, we could say that it loses its sensitivity to the absolute mass scale of the problem.

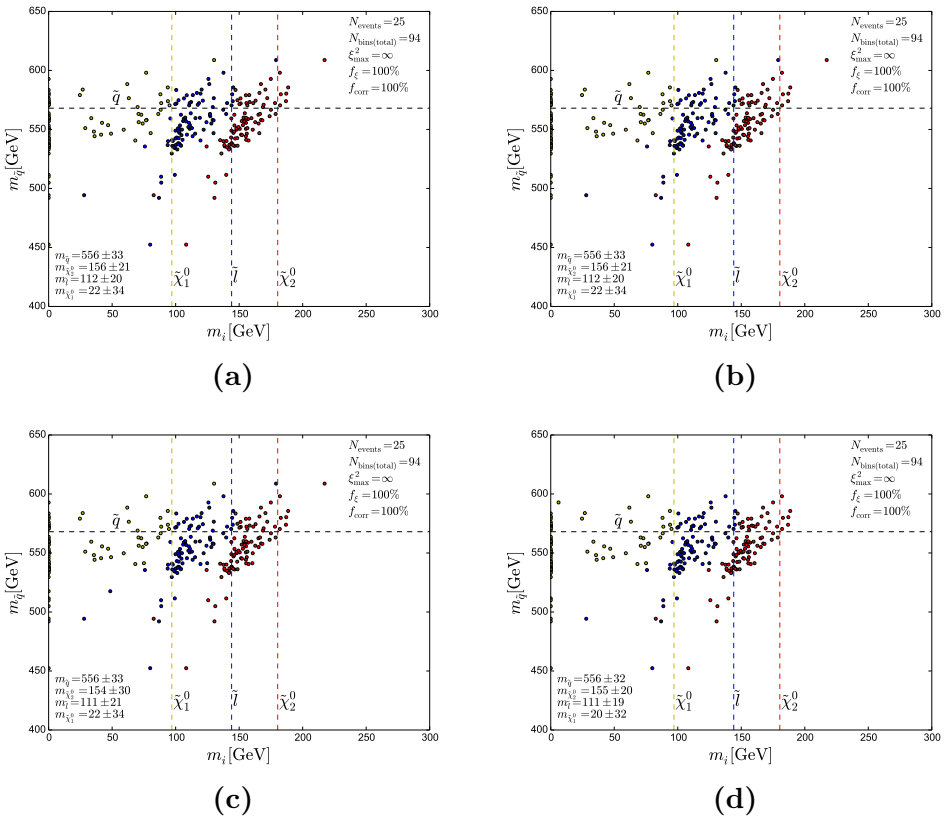
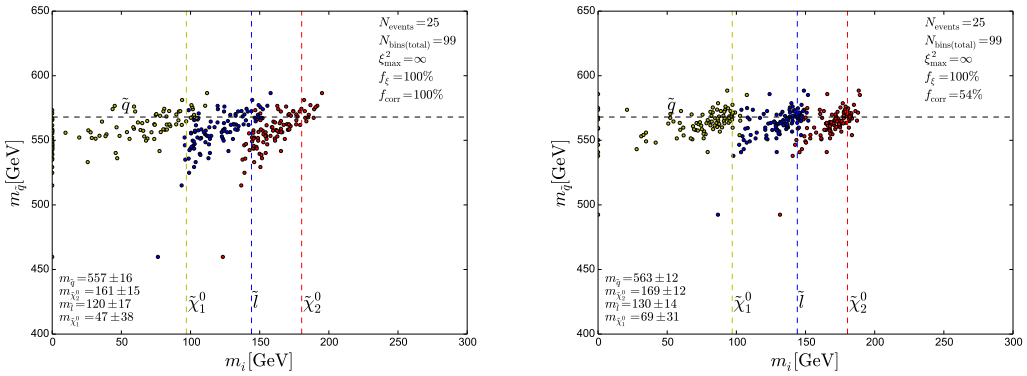


Figure 4.7: Again the same fit as in 4.5 and 4.6, here with a 5% smeared dataset and no combinatorics.

4.6 Effects of a more realistic momentum smearing

The momentum smearing that Webber uses is crude. More realistic smearing formulae can be obtained from the manual to `AcerDET 1.0` [53]. We apply these smearing formulae, updated to be consistent with the ATLAS Technical Paper [54], to the `Herwig++` dataset. In Fig. 4.8 we show minimizations with low tolerance on this dataset. In Fig. 4.8a combinatorics is not taken into account, and in Fig. 4.8b it is. In the former case, the LSP mass is fitted to zero in 32 of the samples, and in the latter case it is fitted to zero in 12 of the samples.

The best-fit masses in the no-combinatorics case of Fig. 4.8a are $m_{\tilde{q}} = 557 \pm 16$ GeV, $m_{\tilde{\chi}_2^0} = 161 \pm 15$ GeV, $m_{\tilde{l}} = 120 \pm 17$ GeV and $m_{\tilde{\chi}_1^0} = 47 \pm 38$ GeV, which have smaller errors and biases compared to the fit with 5% momentum smearing in the previous section. This indicates that the momentum smearing used by Webber overestimates the uncertainties, since we obtain a better fit with a more realistic smearing. The fit in the case with combinatorics, Fig. 4.8b, also has less spread compared to Fig. 4.4b. However, it has the same issues with dependence on the `Simplex` starting point.



(a) **(b)**

Figure 4.8: Minimization with low tolerance on the `Herwig++` dataset with momentum smearing according to the `AcerDET 1.0` manual [53]. Combinatorics is not taken into account in **(a)**.

Chapter 5

Investigating potential improvements

With the potentially significant systematic uncertainties inherent in Webber's original method, from the choice of starting point and the tolerance of the minimization, we will now turn to investigate possible improvements.

5.1 Fitting mass squared differences

We saw in the previous chapter that, even without taking combinatorical ambiguities into account, the method is insensitive to the absolute mass scale of the decay in many of the samples when the momentum resolution is smeared. In a later article [55], Webber *et al.* reformulate the method in terms of squared mass differences. We can borrow their idea and reformulate the problem as a mass-squared-difference fit. Such a fit may be combined with measurements of the dilepton invariant mass edge to find the LSP mass, using Eq. (3.2), which can be rewritten as

$$m^2_{\tilde{\chi}^0_1} = (m^2_{\tilde{l}} - m^2_{\tilde{\chi}^0_1}) \left[\frac{m^2_{\tilde{\chi}^0_2} - m^2_{\tilde{l}}}{(m^{\max}_{ll})^2} - 1 \right], \tag{5.1}$$

or in the more abstract notation of Fig. 3.1,

$$M^2_A = (M^2_B - M^2_A) \left[\frac{M^2_C - M^2_B}{(m^{\max}_{ab})^2} - 1 \right]. \tag{5.2}$$

Thus we see that the LSP mass can be found from knowing only the mass-squared differences plus the invariant mass edge. This formulation is inspired by [56].

Referring back to Chapter 3, and the way the reconstruction was formulated in terms of matrices, the only modifications we have to make in order to reformulate

the problem as a mass-squared-difference fit are the following: Define a vector $\mathbf{M}$ of mass-squared differences

$$\mathbf{M} = (M_1^2, M_2^2, M_3^2), \tag{5.3}$$

where

$$M_1^2 = m_D^2 - m_C^2, \quad M_2^2 = m_C^2 - m_B^2, \quad M_3^2 = m_B^2 - m_A^2, \tag{5.4}$$

and observe that the vector $\mathbf{S}$ may still be written as

$$\mathbf{S} = \mathbf{B}\mathbf{M} + \mathbf{C}, \tag{5.5}$$

provided we let

$$\mathbf{B} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.6}$$

Thus the reconstructed LSP momenta $\mathbf{P} = (p_A^x, p_A^y, p_A^z, E_A, p_{A'}^x, p_{A'}^y, p_{A'}^z, E_{A'})$ are still given as

$$\mathbf{P} = \mathbf{A}^{-1}\mathbf{B}\mathbf{M} + \mathbf{A}^{-1}\mathbf{C}, \tag{5.7}$$

where $\mathbf{M}$ and $\mathbf{B}$ are modified and $\mathbf{A}$ and $\mathbf{C}$ are as before.

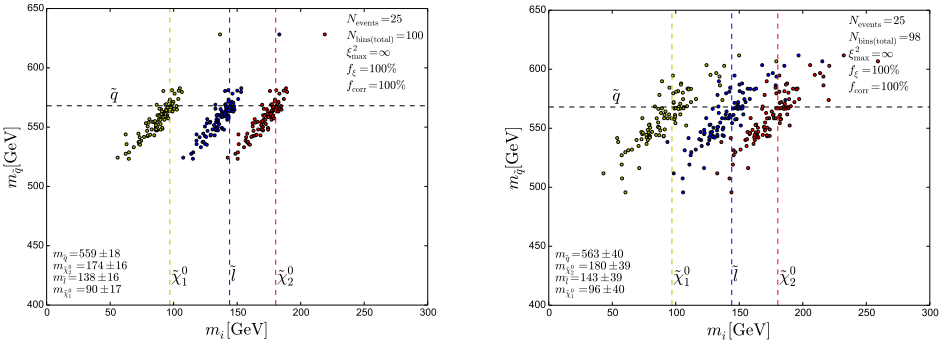
This means that we can reformulate our problem to fit $M_{1,2,3}^2$ instead. However, since we in this case do not fit the masses themselves, our ξ^2 function,

$$\xi^2(\mathbf{M}) = \sum_n \left[(\hat{p}_A^2)_n - \frac{m_A^2}{M_{\text{norm}}^2} \right]^2 + \left[(\hat{p}_{A'}^2)_n - \frac{m_{A'}^2}{M_{\text{norm}}^2} \right]^2, \tag{5.8}$$

has an unknown variable m_A. We choose to use the dilepton mass edge constraint, Eq. (5.2), to calculate the value of m_A^2 from the squared mass difference at each function evaluation. In terms of the mass-squared differences $M_{1,2,3}^2$, m_A^2 is given as

$$m_A^2 = M_3^2 \left[\frac{M_2^2}{(m_{ll}^{\max})^2} - 1 \right]. \tag{5.9}$$

We note that with these modifications, we have introduced another constraining equation into our problem, thus reducing the number of free parameters from



(a) **(b)**

Figure 5.1: MSD minimizations on the `Herwig++` dataset **(a)** without smearing and **(b)** with 5% momentum smearing, without combinatorics.

two to one, as discussed in Chapter 3. In addition, the minimization problem has been reduced from a four-dimensional one to a three-dimensional one.

A fit of the unsmeared dataset with this method, not considering combinatorics, is shown in Fig. 5.1a. We have used the theoretical value of $m_{ll}^{\max} = 80.1$ GeV for the SPS1a masses, calculated using Eq. (3.2). The experimental uncertainty on $m_{ll}^{\max}$ is not expected to be significant compared to other sources of uncertainty. We have checked that the fit also in this case is independent of where we start the search. We also show the same fit on the dataset with 5% momentum smearing in Fig. 5.1b. In the last chapter we saw that with momentum smearing, the LSP mass was estimated to zero in 63 of of the 94 convergent samples, Fig. 4.7, when we used the original formulation of the method. In this case all the samples obtain a nonzero LSP mass. For the remainder of the thesis we will use the mass-squared-difference (MSD) minimization.

We also want to investigate whether the modifications have affected the combinatorical problems we faced when using the original formulation, where we pick the lowest value of ξ^2 among all combinations for each event in each point. We saw in Section 4.4 that the results were dependent on where the minimization search was started.

In Fig. 4.5 we demonstrated the performance of the minimization of the unsmeared `HERWIG 6.510` dataset using four different starting points for the `Simplex` search. In Fig. 5.2 we show the same plot using the MSD technique on the `Herwig++` dataset, selecting between the different combinatorical choices by always choosing the lowest ξ^2 value. Although there are large biases and standard errors, the *differences* between the four scatter plots in Fig. 5.2 appear less significant than in Fig. 4.5. This is reflected in the mean values, which for *e.g.* the squark range from 507 to 519 GeV in the four plots, and in the standard errors, which all are about 60 GeV for the squark. There are similar consisten-

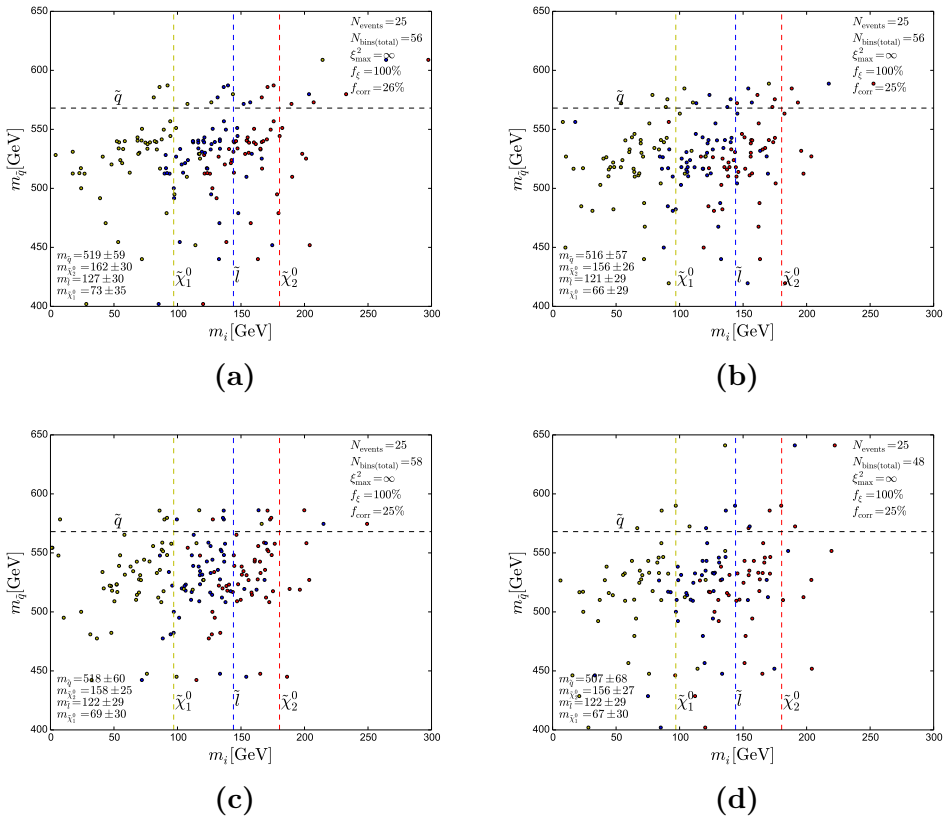


Figure 5.2: MSD minimization on the unsmeared `Herwig++` dataset with combinatorics done according to Webber, for the four different starting points introduced in Section 4.4.

cies between the plots for the other three masses. The four different starting points are also consistent with respect to how many samples obtain convergence in the minimization (about 50 of 100), and how many events obtain the correct combinatorical choice in the best-fit point (about 25% of the convergent samples).

Even though the four plots in Fig. 5.2 are statistically similar, the individual event samples are not consistently fitted to the same points, and the same samples do not obtain convergence in all cases. 32 of the 100 event samples obtain convergence from all four starting points, although each starting point has about 50 convergent samples. Only three of the samples obtain best-fit points less than 0.1 GeV apart in all four masses in all four cases.

5.2 Summing the combinations

We saw in Section 4.5 that when we minimize only the true combinatorical choice, the minimization is independent of starting point. If we include all combinations

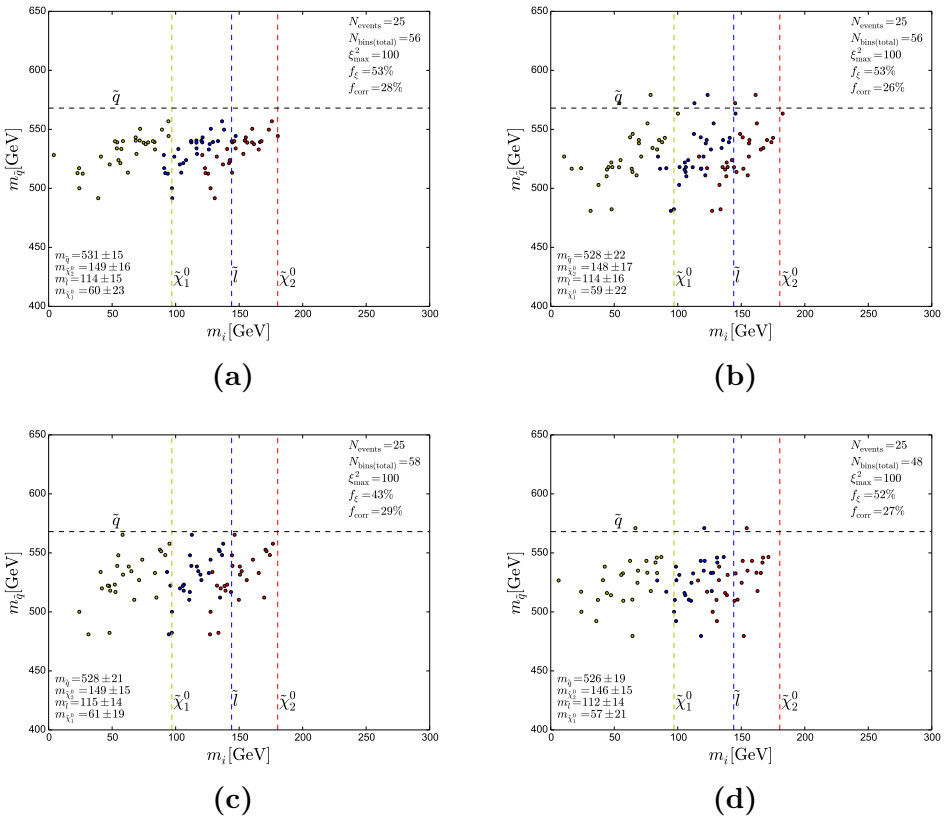


Figure 5.3: MSD minimization on the unsmeared `Herwig++` dataset with combinatorics done according to Webber, for the four different starting points introduced in Section 4.4, subject to a ξ^2 cut of 100.

by always selecting the lowest among the ξ^2 values in each mass point, a starting point dependence is introduced. The mathematical difference between these two problems is that in the former case, the ξ^2 is a smooth polynomial, while in the latter case it is not. We can make the ξ^2 surface smooth if we add the different combinatorical choices together instead of jumping between them. Webber mentions this option in his article, but discards it, saying *"[t]he surface would be smooth if one added the ξ^2 contributions of all combinations, but then the sensitivity to the correct solution is reduced and biases are introduced by the huge contributions of wrong combinations"* [1].

However, not all combinations will contribute equally if they are included. Reference [57] features a general discussion on momentum reconstruction problems with missing energy, and also a specific discussion of the golden MSSM decay chain. Although they discuss other mass reconstruction methods than our present minimization scheme, some of the discussion applies to our case as well. They argue that the permutation of the leptons within the same chains intro-

duces only a minor error into the reconstruction, which because of the kinematics should vanish in the limit of large $m_{\tilde{q}} - m_{\tilde{\chi}_1^0}$ mass gap.

For events where the dileptons differ in flavour between the chains, which is what we have used in our analysis thus far, there are eight combinatorical possibilities. These can be divided into two categories depending on which lepton pair is paired with which quark. For each of these pairings, there are four combinations of near and far leptons, as discussed in Section 3.4. The matrix **A** is invariant, up to a permutation of rows, for each of the two categories. If we include all four combinations of near and far leptons for each event by adding them to the ξ^2, then the problem reduces from an eight-fold ambiguity to a two-fold ambiguity, distinguished only by the permutation of quarks between chains.

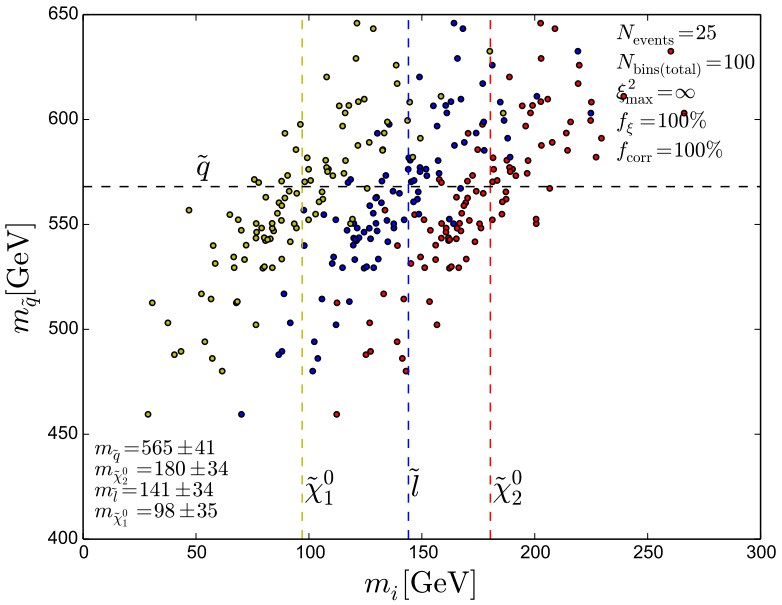


Figure 5.4: Minimization of the unsmeared `Herwig++` dataset where all orderings of the leptons within the same chains are included in the calculation of ξ^2, and only the true quark-lepton combinations are considered.

In Fig. 5.4 we show a fit of the unsmeared dataset where the ξ^2 has been constructed in this way, but where only the true pairing of quarks and leptons is considered. The mean values and standard errors for the mass fits in the 100 samples are $m_{\tilde{q}} = 565 \pm 41$ GeV, $m_{\tilde{\chi}_2^0} 180 \pm 34$ GeV, $m_{\tilde{l}} = 141 \pm 34$ GeV and $m_{\tilde{\chi}_1^0} = 98 \pm 35$ GeV. The standard error is quite large, about 40 GeV for all the four masses, but the mean values are very close to the true values in all four cases. The minimization is also completely robust against starting points: For the four different starting points defined earlier, all 100 samples obtain convergence in all cases, and each sample is fitted to the same point in all cases. In the case

where we jump between values, 50% of the event samples obtain convergence, see Fig. 5.2. Since a real LHC analysis only will have one sample of events, the convergence robustness is an important factor.

We have also tried including all eight combinations instead of only the four lepton permutations. Then the results worsen considerably. While the robustness against starting points is retained, only half of the samples obtain convergence, and the errors on the masses are about 90 GeV for the three lightest particles and 150 GeV for the squark. There is also a significant downward bias on the squark mass.

5.2.1 Error introduced when summing lepton permutations

The lepton permutations consist of permuting the leptons in one, both or none of the two chains. Consider one decay chain, shown schematically in fig. 5.5. For the true assignment of near and far lepton, we have the relationship $m_B^2 =$

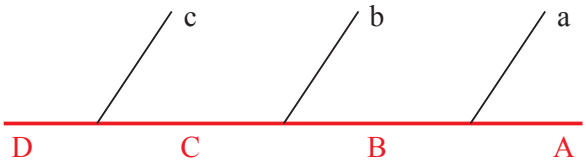


Figure 5.5: Decay topology, from [42].

$(p_A + p_a)^2 \equiv m_{Aa}^2$. If the lepton assigments are switched, it corresponds to reconstructing the mass $(p_A + p_b)^2 = m_{Ab}^2 \equiv m_X$ instead of the true mass m_B. The kinematics of this misidentification is investigated in Appendix C. It turns out that m_X has a probablity distribution limited by the values

$$m_{X\,\mathrm{high}}^2 = m_A^2 + m_C^2 - m_B^2, \qquad (5.10)$$
$$m_{X\,\mathrm{low}}^2 = \frac{m_C^2 m_A^2}{m_B^2}.$$

Note that m_X is contained in the interval (m_A, m_C) for any kinematically allowed choice of masses $m_{A,B,C}$. The probability distribution of m_X is shown in Fig. 5.6. For SPS1a, the limits evaluate to $m_{X\,\mathrm{low}} = 121$ GeV and $m_{X\,\mathrm{high}} = 145$ GeV. Thus, in this case, the high m_X limit is close to the true $m_B = 144$ GeV. This does not have to be the case generally.

In principle the measurement of the other masses should be unaffected by this, *e.g.* since $m_C^2 = (p_b + p_a + p_A)^2$ which is invariant under $b \leftrightarrow a$. However, there is a nontrivial relationship between the fitted masses in the method, *e.g.* due to the dilepton edge formula, so the misidentification may affect them indirectly.

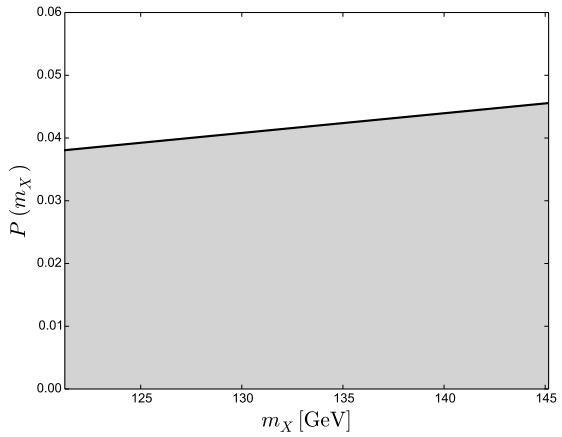


Figure 5.6: Probability distribution of m_X.

5.3 Cutting on the determinant

In this section we investigate what can be gained by applying a minimal value cut on the determinant of the matrix **A**. The idea behind this is that since a matrix with zero determinant is uninvertible, the magnitude of the determinant may be small for events that are "difficult" to invert in some sense, and these may contribute poorly to the fit.

This turns out to have some merit, as can be seen from Fig. 5.7: The plot (a) is exactly the same as in Fig. 5.4, except that the det**A** cut has been applied on an event-wise basis. The errors on the masses decrease by a factor two. They are further reduced by combining with a ξ^2 cut, but at the expense of losing a lot of samples. About 60% of events pass a $|\det\mathbf{A}| > 10$ cut. The events passing the cut are as expected quite evenly distributed between samples, with the worst sample retaining 10 of 25 events after cuts and the best retaining 24. The distribution of the number of events in each sample passing the cut have a mean value of 15 with a standard deviation of 2.5.

Note that while the ξ^2 cut used earlier will remove the whole sample, the det**A** cut only removes parts of it, meaning that all the samples still may yield an answer, albeit possibly with low statistics.

Figure 5.8 shows the distribution of det**A** for the true quark-lepton combinations and the misidentified combinations. We see that the distribution exhibits a smooth decrease from zero, and it is symmetric. There do not appear to be any jumps in det**A** values which could be used to set the cut value.

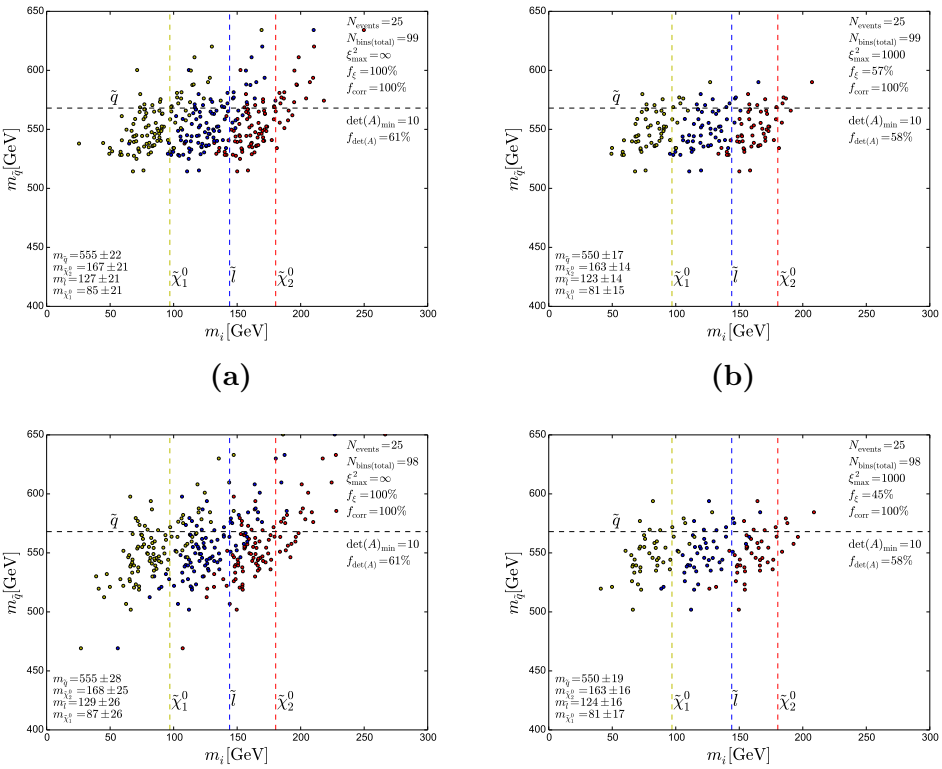


Figure 5.7: Minimization where a minimal value cut is applied on $|\det\mathbf{A}|$ for each event, in this case $|\det\mathbf{A}| > 10$. Minimization is done using the MSD technique, on the `Herwig++` dataset without smearing in **(a)** and **(b)** and with 5 percent smearing in **(c)** and **(d)**. ξ^2 cuts of 1000 are applied in **(b)** and **(d)**. Only the true quark-lepton combination for each event is considered, and the four lepton permutations are summed.

5.3.1 Sub-determinant cuts

Consider again the full form of the matrix $\mathbf{A}$,

$$\mathbf{A} = \frac{2}{M_{\text{norm}}} \begin{pmatrix} p_c^x & p_c^y & p_c^z & -E_c & 0 & 0 & 0 & 0 \\ p_b^x & p_b^y & p_b^z & -E_b & 0 & 0 & 0 & 0 \\ p_a^x & p_a^y & p_a^z & -E_a & 0 & 0 & 0 & 0 \\ p_{\text{miss}}^x/2 & 0 & 0 & 0 & p_{\text{miss}}^x/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{c'}^x & p_{c'}^y & p_{c'}^z & -E_{c'} \\ 0 & 0 & 0 & 0 & p_{b'}^x & p_{b'}^y & p_{b'}^z & -E_{b'} \\ 0 & 0 & 0 & 0 & p_{a'}^x & p_{a'}^y & p_{a'}^z & -E_{a'} \\ 0 & p_{\text{miss}}^y/2 & 0 & 0 & 0 & p_{\text{miss}}^y/2 & 0 & 0 \end{pmatrix}. \quad (5.11)$$

It comes from the constraints in Eqs. (3.4)–(3.9). These equations are projections of the unknown momenta $p_A/p_{A'}$ onto the measured momenta $p_{abc}/p_{a'b'c'}$. To be

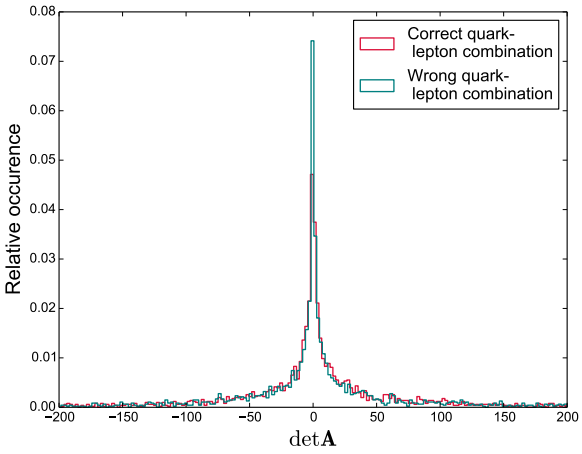


Figure 5.8: Distribution of det$\mathbf{A}$ for the correct and the wrong combination of quarks with lepton pairs. The tails of the distribution have been trimmed at ± 200 to remove outliers.

able to invert the equations and solve for $p_A/p_{A'}$, the three measured momenta for each chain must be linearly independent, *i.e.* they must form a basis for the subspace we are projecting $p_A/p_{A'}$ onto. Of course, three vectors cannot span a four-dimensional space — this is the reason the problem is underconstrained and we have to find a best-fit solution, since there are not enough measured momenta to fully constrain the invisible ones. But we can look at the three-dimensional spatial momentum subspace. For *e.g.* p_A to be reconstructible, we should have span $\{\mathbf{p}_a, \mathbf{p}_b, \mathbf{p}_c\} = \mathbb{R}^3$. This is equivalent to

$$|\mathbf{p}_a \cdot (\mathbf{p}_b \times \mathbf{p}_c)| \neq 0, \tag{5.12}$$

which is proportional to the sub-determinant of $\mathbf{A}$ given by

$$\text{subdet}(\mathbf{A}, 1, 3) = \frac{1}{M_{\text{norm}}^3} \begin{vmatrix} p_c^x & p_c^y & p_c^z \\ p_b^x & p_b^y & p_b^z \\ p_a^x & p_a^y & p_a^z \end{vmatrix}. \tag{5.13}$$

If the value of the left hand side of Eq. (5.12) is small, then the projections of the invisible momenta are numerically poor. Analogously to cutting on the full det$\mathbf{A}$, we can apply a simultaneous cut on the values of the two subdeterminants

$$\text{subdet}(\mathbf{A}, 1, 3) = \frac{1}{|\mathbf{p}_a||\mathbf{p}_b||\mathbf{p}_c|} \begin{vmatrix} p_c^x & p_c^y & p_c^z \\ p_b^x & p_b^y & p_b^z \\ p_a^x & p_a^y & p_a^z \end{vmatrix} = \frac{|\mathbf{p}_a \cdot (\mathbf{p}_b \times \mathbf{p}_c)|}{|\mathbf{p}_a||\mathbf{p}_b||\mathbf{p}_c|}, \tag{5.14}$$

$$\text{subdet}(\mathbf{A}, 5, 7) = \frac{1}{|\mathbf{p}_{a'}||\mathbf{p}_{b'}||\mathbf{p}_{c'}|} \begin{vmatrix} p_{c'}^x & p_{c'}^y & p_{c'}^z \\ p_{b'}^x & p_{b'}^y & p_{b'}^z \\ p_{a'}^x & p_{a'}^y & p_{a'}^z \end{vmatrix} = \frac{|\mathbf{p}_{a'} \cdot (\mathbf{p}_{b'} \times \mathbf{p}_{c'})|}{|\mathbf{p}_{a'}||\mathbf{p}_{b'}||\mathbf{p}_{c'}|}, \tag{5.15}$$

where we have chosen a different normalization to make the subdeterminant values lie in the interval $[-1, 1]$. The determinant with this normalization is what we will refer to as the subdeterminant in the following.

The distribution of the subdeterminant values for the unsmeared `Herwig++` dataset is shown in figure 5.9. The subdeterminants corresponding to the two chains are included independently. It is interesting to note that the correlation coefficient between the determinant and subdeterminants is of the order -0.01, and a correlation hypothesis test shows consistency with the assumption that they are uncorrelated, at 95% confidence level.

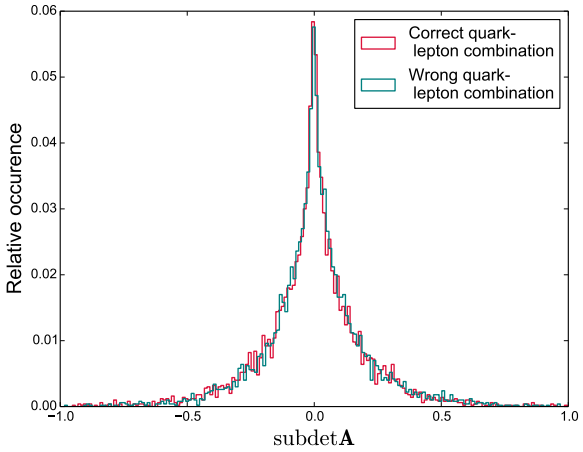


Figure 5.9: Distribution of subdet$(\mathbf{A}, 1, 3)$ and subdet$(\mathbf{A}, 5, 7)$ for the correct combination and the wrong combination of quarks with lepton pairs.

To test this approach, we apply cuts on the subdeterminants, by requiring that both subdeterminants are larger than some value, to the `Herwig++` dataset. We use the cut values 0.03 and 0.005, chosen so that about 60% and 90% of the events pass the cut, respectively. 60% is the same fraction as for the det$\mathbf{A}$ cut used in Fig. 5.7, to enable comparison. We use the MSD fitting technique with a sum over the four lepton combinations, considering only the true $\mathbf{A}$ matrices, as in Fig. 5.7. The resulting scatter plot is shown in Fig. 5.10. We see that the errors on the masses in both cases are of the same order as in Fig. 5.4, where no determinant cut was applied, and thus they are a factor two larger than the errors in the case of the det$\mathbf{A}$ cut in Fig. 5.7a. Thus, the subdeterminant cut does not seem to make the fit better.

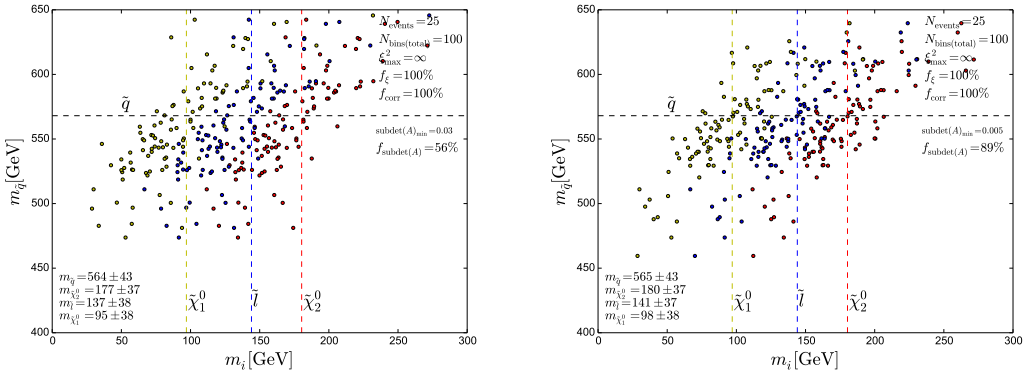


Figure 5.10: Minimization where a cut of $|\text{subdet}(\mathbf{A}, 1, 3)| > x$ and $|\text{subdet}(\mathbf{A}, 5, 7)| > x$ is applied for each event, with $x = 0.03$ in **(a)** and $x = 0.005$ in **(b)**. Minimization is done using the MSD technique, on the `Herwig++` dataset without smearing. Only the true quark-lepton combinations are considered, and the remaining four lepton permutations are summed.

5.4 Handling combinatorics in the summed-combination approach

We proceed with investigating the method of summing the four lepton combinations. In this formulation, the combinatorical ambiguity is reduced. When considering only events where the chains have opposite-flavour leptons (OFL), the ambiguity reduces from eight in the original formulation to two. For same-flavour leptons (SFL) it reduces from 16 to 4. Each of these combinations correspond to a different $\mathbf{A}$ matrix.

A complication is that, as mentioned in the beginning of Chapter 4, there may be more than two high-momentum quarks in an event. This is in fact quite likely, since one or both of the squarks in the chains may come from gluinos, originating from the processes $pp \to \tilde{g}\tilde{q}$ and $pp \to \tilde{g}\tilde{g}$ and decaying as $\tilde{g} \to \tilde{q}q$. In a realistic detector event, the quarks will fragment into hadron jets before reaching the detector. These jets must be reconstructed using specially constructed algorithms in order to determine, as well as possible, the four-momentum and invariant mass of the originating quark. This process potentially introduces large inaccuracies into the measurements of the four-momenta of the jets.

Since we do not *a priori* know which jets belong to the event, the possibility of additional jets quickly increases the combinatorical ambiguity — with three jets, there are six different $\mathbf{A}$ matrices for OFL events, and 12 for SFL events. With four jets, there are 12/24 different $\mathbf{A}$'s. It is clear that we need a way to handle the remaining ambiguity.

There are several options for doing this. One is to utilize the original method of "jumping" between values, always selecting the lowest among the combinations for each event at each mass point. When there are fewer different values to choose between, the amount of jumping and the resulting difficulties might be reduced.

We check this for the unsmeared `Herwig++` dataset with only OFL events, where there are only two combinations to jump between, by starting the minimization from the four different starting points. The resulting scatter plots are shown in Fig. 5.11. We find in this case for the unsmeared dataset that ~ 90 samples converge in each case, while 84 of the samples converge in all four cases. 62 of the samples are fitted consistently, meaning that all four minimizations agree on all four mass values within 0.1 GeV. We also find that about 80% of the events obtain the minimum with the correct combination, while the remaining 20% apparently obtain smaller values using the wrong combination of quarks and leptons.

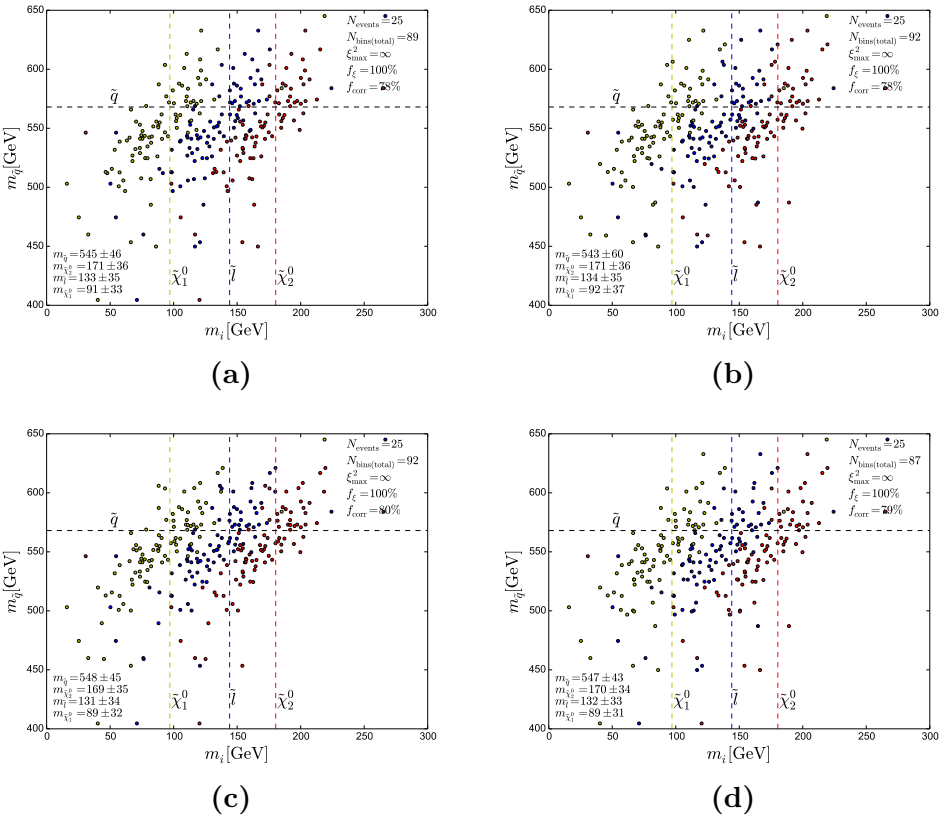


Figure 5.11: MSD minimization on the unsmeared `Herwig++` dataset with combinatorics handled by summing the four closest combinations and jumping between the two quark-lepton combinations, for the four different starting points introduced in Section 4.4.

5.4.1 An event-pair algorithm for handling combinatorics

Another possibility is to combine a few events at a time, and minimize the ξ^2, with the lepton-combinations summed, separately for each of the quark-lepton combinations between chains, *i.e.* each matrix $\mathbf{A}$, in each event. For instance, if we select two events, then there are four ways to combine the quark-lepton assignments in the events: the right combination in both events, one right and one wrong, or both wrong. If the lowest ξ^2 minimum value among these four tend to belong to the right quark-lepton combination, then this may be used to select the correct quark-lepton combinations, and thus the correct $mathbf A$ for the event, before all events are summed together for the total mass fit.

We have implemented this in the following way: We use the mass-squared difference fitting technique, and we sum the four leptonic combinations. For each sample of events, we run through all the individual events, indexing them by i. For each event, we run through all the other events, indexed by j. For each pair (i, j), we minimize the ξ^2 for the four possible combinations of $(\mathbf{A}_i^a, \mathbf{A}_j^b)$ matrices, where $a, b = 1, 2$ denotes the different $\mathbf{A}$ matrices for each event. We make a note of which choice a for event i gives the lowest minimum among the four. We allow the minima to take unphysical values for the fitted masses, either in the form of wrong mass hierarchies or even negative mass values, since we only are interested in the ξ^2 value.

For each event i, we count how many of the (i, j) pairs prefer the different choices of $\mathbf{A}_i^a$. If there is a clear preference for a specific choice a, then we select this as the "true" matrix. In the case of events where all four leptons are the same flavour, there will be four different $\mathbf{A}^a$ matrices instead of two, thus increasing the number of combinations to evaluate.

For the unsmeared `Herwig++` data set of 100 samples of 25 events, there seems to be a clear preference for the correct $\mathbf{A}$ matrix. The fraction of events where the correct $\mathbf{A}$ is chosen, for each event sample, is shown in Fig. 5.12. We see that for the samples with the lowest fraction of correct events, the fraction is 60%. Thus even the worst sample does better than a random guess, which has an expectance of 50%. No cut on det$\mathbf{A}$ is applied prior to the selection, since we find that we get a significantly clearer signal by including all events. We also apply a determinant cut to remove events after the selection. There is a slight shift towards a higher fraction of correct $\mathbf{A}$ after this cut. We obtain similar results for the dataset with 5% momentum smearing.

Figure 5.13 shows scatter plots of the best-fit points found by selecting $\mathbf{A}$ for each event according to this algorithm, and summing the remaining four combinations. Figure 5.13a has no cut on the determinant, while a $|\det\mathbf{A}| > 10$ cut is applied in Fig. 5.13b.[1]

[1]The number $f_{\det\mathbf{A}} = 96\,\%$ in the plot in Fig. 5.13a seems to indicate that not all events have passed the determinant cut, even though no cut has been applied. This is an artifact due to not all events passing the $\mathbf{A}$ matrix selection algorithm.

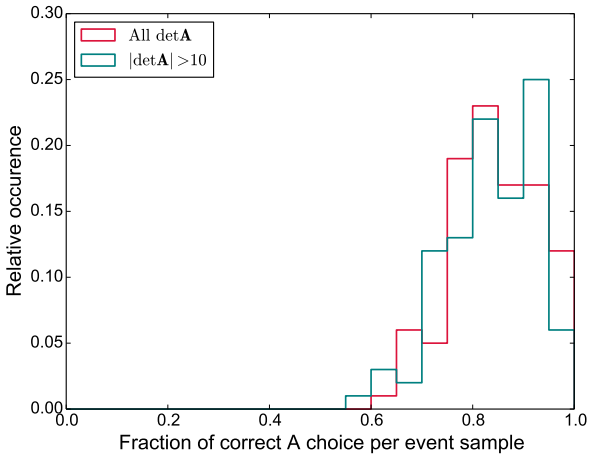
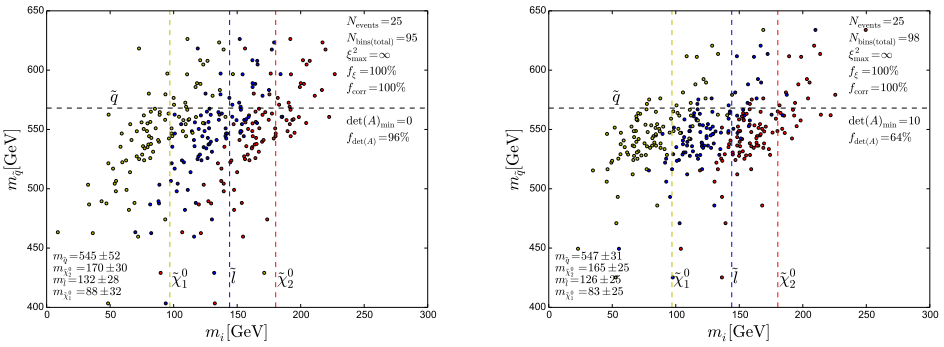


Figure 5.12: Distribution of the 25-event samples of `Herwig++` events, indicating, for each sample, the fraction of events where the event-pair selection algorithm prefers the true **A** matrix. A $|\det\mathbf{A} > 10$ cut is used in the blue distribution, discarding some events. See the text for details.

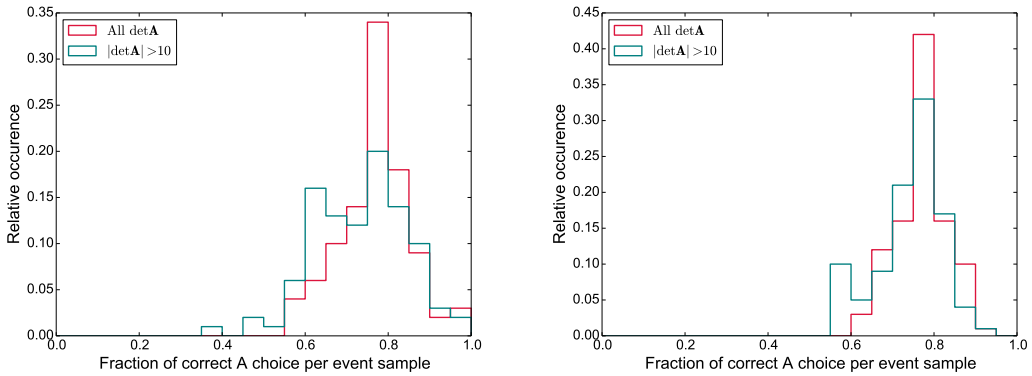


(a) **(b)**

Figure 5.13: MSD minimization on the unsmeared `Herwig++` dataset with combinatorics handled by selecting **A** matrices for each event using the event-pairing algorithm described in Section 5.4 and summing the remaining four combinations. A minimal-value determinant cut of 10 is applied in **(b)**.

5.4.2 Including same-flavour lepton events

In our analysis so far, we have discarded events where all the four leptons are of the same flavour (SFL events). This was done to limit the combinatorical ambiguity, but comes at the expense of halving the amount of events we can expect. In the language of our method, the inclusion of SFL events amounts to including two additional matrices **A** which come from permuting the leptons across the chains. It is interesting to see how the algorithm for choosing **A** tackles this case. We find that it handles it quite well. Figure 5.14a shows the fraction of correctly identified **A** matrices with and without a determinant cut, on the unsmeared `Herwig++` dataset with SFL events included. We see that almost all event samples obtain a fraction larger than 50%, which should be compared to the expectance of random chance which is 25% in this case. The SFL events are distributed as they come from the Monte Carlo, which means they constitute about half the events in each sample.



(a) **(b)**

Figure 5.14: Distribution of the fraction of correctly identified **A** matrices in samples when same-flavour lepton events are included. A determinant cut of 10 is applied in **(b)**.

In Fig. 5.15, we show the best-fit points obtained in this case. The errors are somewhat larger than when using only opposite-flavour leptons (OFL), and the downward bias is enhanced. The best-fit points with a determinant cut are $m_{\tilde{q}} = 547 \pm 31$ GeV, $m_{\tilde{\chi}_2^0} = 165 \pm 25$ GeV, $m_{\tilde{l}} = 126 \pm 25$ GeV and $m_{\tilde{\chi}_1^0} = 83 \pm 25$ GeV for the OFL only case (Fig. 5.13b), and $m_{\tilde{q}} = 534 \pm 35$ GeV, $m_{\tilde{\chi}_2^0} = 160 \pm 27$ GeV, $m_{\tilde{l}} = 121 \pm 27$ GeV and $m_{\tilde{\chi}_1^0} = 78 \pm 28$ GeV when SFL events are included as well (Fig. 5.15b).

Since using only OFL events amounts to discarding half the available events, when SFL events are included we should really double the sample size. Figure 5.16 shows scatter plots of this. The mean best-fit points with a determinant cut of 10 are $m_{\tilde{q}} = 539 \pm 32$ GeV, $m_{\tilde{\chi}_2^0} = 165 \pm 19$ GeV, $m_{\tilde{l}} = 125 \pm 19$ GeV

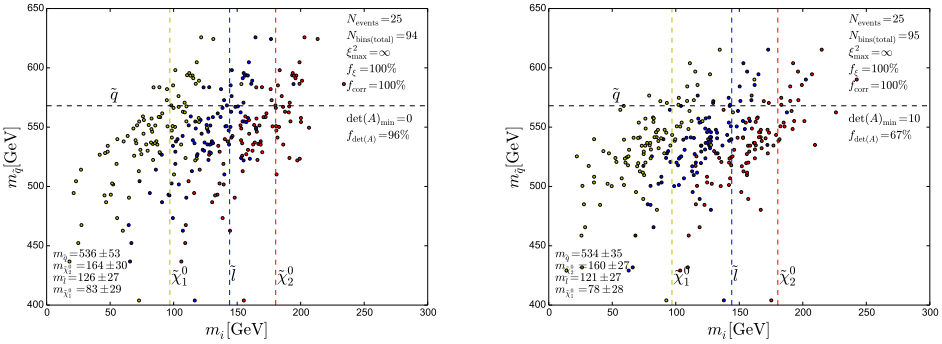


Figure 5.15: MSD minimization on the unsmeared `Herwig++` dataset including same-flavour lepton events, with combinatorics handled by selecting **A** matrices for each event using the event-pairing algorithm described in Section 5.4 and summing the remaining four combinations. A minimal-value determinant cut of 10 is applied in **(b)**.

and $m_{\tilde{\chi}^0_1} = 83 \pm 20$ GeV. The errors are somewhat reduced compared to the minimization with 25-event samples, and they are also slightly lower than the 25-event sample minimization with only OFL events, though there is a larger downward bias on the squark mass than in the only-OFL case.

5.4.3 Handling the remaining combinatorical ambiguity

The method of event-pair selection described in the previous section provides a way to select between the quark-lepton combinations. With this, we in principle have a working method which can be applied to a sample of events without knowing the true combinatorical configuration, and which is robust against starting points and convergence issues. The plot of minimization on the unsmeared `Herwig++` dataset with a determinant cut, shown in Fig. 5.13b, can be viewed as a best-case scenario for this technique. The mean best-fit values are $m_{\tilde{q}} = 547 \pm 31$ GeV, $m_{\tilde{\chi}^0_2} = 165 \pm 25$ GeV, $m_{\tilde{l}} = 126 \pm 25$ GeV and $m_{\tilde{\chi}^0_1} = 83 \pm 25$ GeV. Although the determinant cut has reduced the errors compared to Fig. 5.13a, the errors are still very large — as large as 30% for the LSP. There is also a significant downward bias, which is enhanced by the determiant cut. Part of this uncertainty is introduced by the summing of the four leptonic combinations. It would be desirable to find a way to discriminate between these combinations also.

One might try to revert to the method of "jumping" between combinations in the ξ^2 function, but jumping between only the four remaining combinations instead of all eight. We have tried this option, but it gives a very poor fit.

Another possibility is to do something analogous to the algorithm we applied

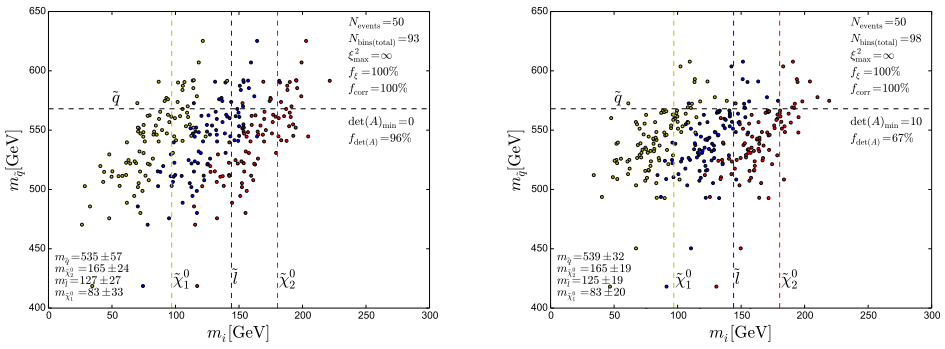


Figure 5.16: MSD minimization on 50-event samples of the unsmeared `Herwig++` dataset, including same-flavour lepton events, with combinatorics handled by selecting **A** matrices for each event using the event-pairing algorithm described in Section 5.4 and summing the remaining four combinations. A minimal-value determinant cut of 10 is applied in **(b)**.

to choose between the quark assignments. One could imagine pairing events and minimizing all combinations ($4^2 = 16$, in this case), selecting the combination which is minimal in most cases. However, this would only give two contributing terms to the ξ^2 for each minimization, instead of eight terms in the case where four permutations are summed. The ξ^2 turns out to be difficult to minimize in this case, and it often does not converge at all.

5.5 Comparison benchmark

We are now in a position to produce a table analogous to those of Chapter 4, for our modified implementation of the method. We apply the MSD fit with dilepton edge to the `Herwig++` dataset with only OFL events. We use the event-pair selection algorithm to choose between the two **A** matrices for each event, and sum the remaining four lepton permutations in the ξ^2. We apply a minimal value determinant cut of 10. We also test applying a ξ^2 cut as in Chapter 4, although the cut value has to be set higher to get a sizeable fraction of samples past the cut because of the changes made to the method. We have chosen a cut value which gives a similar pass-cut fraction as in Webber's original analysis, Table 4.1. The results are summarized in Table 5.1.

We see that the ξ^2 helps to reduce the errors as before. Note that the fraction of events passing the ξ^2 cut does not decrease as rapidly with increased momentum smearing as in the original analysis of Webber. Particularily, we have not doubled the ξ^2 cut value for the 10% smearing case, contrary to the original analysis. The mass errors from the fits in Table 5.1 are larger compared to those

$\delta p/p$	$\xi^2_{\max}$	$f_{\det \mathbf{A}}$	f_ξ	f_{corr}	$m_{\tilde{q}}(568)$	$m_{\tilde{\chi}^0_2}(180)$	$m_{\tilde{l}}(144)$	$m_{\tilde{\chi}^0_1}(97)$
0	∞	64%	100%	83%	547 ± 31	165 ± 25	126 ± 25	83 ± 25
0	5000	63%	81%	82%	546 ± 19	162 ± 18	122 ± 19	80 ± 19
5%	∞	65%	100%	82%	548 ± 33	167 ± 27	128 ± 27	85 ± 28
5%	5000	64%	75%	83%	547 ± 22	166 ± 18	126 ± 18	84 ± 19
10%	∞	64%	100%	81%	545 ± 40	161 ± 29	122 ± 29	78 ± 31
10%	5000	64%	65%	80%	543 ± 31	159 ± 25	120 ± 25	76 ± 26

Table 5.1: MSD minimization on the `Herwig++` dataset with and without momentum smearing and ξ^2 cuts. See the text for details.

of the original analysis [1], given in Table 4.1. For example, the 5% smeared fit with ξ^2 cut has about twice as large error on the squark mass, while the errors on the other three masses are an order of magnitude larger. By reducing the ξ^2 cut value to 2000, we bring f_{corr} down to 56%, the same as in Table 4.1. This brings the errors somewhat down, to give the best-fit values $m_{\tilde{q}} = 548 \pm 19$ GeV, $m_{\tilde{\chi}^0_2} = 163 \pm 16$ GeV, $m_{\tilde{l}} = 124 \pm 16$ GeV, $m_{\tilde{\chi}^0_1} = 81 \pm 16$ GeV.

Note that this analysis is done with a determinant cut of 10. In a future analysis one could experiment with different combinations of determinant and ξ^2 cuts to find an optimal combination.

5.6 Taking jet reconstruction effects into account

Finally, we take the effects of jet reconstruction into account. We have simulated jet reconstruction using `Pythia 8.2` [50] with the `FastJet` [45] jet reconstruction program. It utilizes the anti-k_t jet clustering algorithm [58], which does sequential jet recombination in the following way:

One defines a minimal transverse momentum p_T that a jet can have, and an angular radius squared $R^2 = \Delta\eta^2 + \Delta\phi^2$ in the pseudorapidity–azimuthal angle (η, ϕ) plane of the detector. This radius defines the maximal angular size that a jet is allowed to have.

For an object i with transverse momentum $k_{t,i}$, the distance between i and the beam axis B is defined as

$$d_{iB} = k_{t,i}^{-2}, \tag{5.16}$$

and the distance between two objects i and j is defined as

$$d_{ij} = \min\left\{k_{t,i}^{-2}, k_{t,j}^{-2}\right\} \frac{\Delta R_{ij}^2}{R^2}, \tag{5.17}$$

where

$$\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2, \tag{5.18}$$

with (η_i, ϕ_i) the angular coordinates of i.

The transverse momenta k_t of all particles in the detector are indexed by i. Then the algorithm does a sequential recombination by looping over all particles as follows:

1. For a particle i, find the smallest value among all d_{ij} and d_{iB}.
2. If one of the d_{ij} is smallest, then recombine the objects i and j by combining their four-momenta.
3. If d_{iB} is smallest, set i to be a jet and remove it from the list of particles.
4. Repeat from 1. until there are no particles left.
5. All reconstructed jets which have a p_T larger than the set threshold are kept.

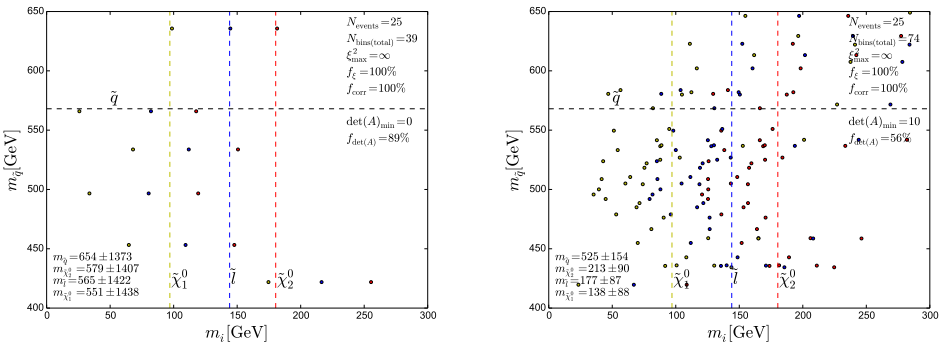
We use $R = 0.4$, $|\eta| < 4$ as parameters for the jet reconstruction. We adapt the cuts used in [56], which are:

1. There must be at least two jets with $p_T > 100$ GeV, $|\eta| < 2.5$.
2. All the four leptons must have $p_T > 10$ GeV, $|\eta| < 2.5$.
3. The event must have missing $p_T > 50$ GeV.

Since we use `Pythia` rather than `Herwig++` for the jet reconstruction, we have checked that the events are consistent with previous results when jet reconstruction is not taken into account.

In Fig. 5.17, we show scatter plots of the best-fit points obtained for the `Pythia` dataset with jet reconstruction, minimized on samples of 25 events with only opposite-flavour leptons. We take the quark masses to be the invariant masses of the reconstructed jets. We use the MSD minimization with a sum of lepton permutations and the event-pair selection algorithm for the quark-lepton permutation. As usual a $|\det\mathbf{A}| > 10$ cut is applied in Fig. 5.17b. The two hardest jets in each event have been taken as the signal jets. This is justified in SPS1a because of the small mass splitting between the gluino and the squarks, which limit the momentum attainable for the quarks from gluino decay in events containing $\tilde{g} \to q\tilde{q}$.

Another reason for selecting jets in this way is that the alternative would require discrimination between three or more jets. This quickly increases the combinatorical ambiguity, as discussed in Section 5.4. One could hope to reduce the ambiguity with additional selection criteria, such as constraints from the invariant mass edges discussed in the beginning of Chapter 3, but the prospects for selecting the correct combination are still dim.



(a) **(b)**

Figure 5.17: Minimization on the `Pythia` dataset with jet reconstruction and kinematical cuts, using 25-event samples with only OFL events. Most of the points in **(a)** lie outside the plotting region.

Note that when we use the jet reconstructed quark momenta, we lose information on which quark belongs to which chain in the Monte Carlo data. We thus no longer have the possibility to minimize only the correct quark-lepton combinations as a control, unless we match the jets to the quarks, which is a non-trivial procedure.

Also, since the numbering of jets is independent from the numbering of chains, we should expect the **A** selection algorithm to have equal preference for the two matrices which have correct lepton combinations. Therefore, the fraction of events where the first **A** is selected should be centered on 0.5. The distribution of this fraction, with a $|\det\mathbf{A}| > 10$ cut, is shown in Fig. 5.18 for 25-event samples of only OFL events and 50-event samples of both OFL and SFL events (abbreviated OSFL). We see that the distributions are centered close to 0.5 as we expect. The OSFL distribution has a slight preference for lower fractions, as we should expect since it chooses between four **A** matrices rather than two.[2]

In Fig. 5.19, we show the minimization on jet-reconstructed samples of 50 events each, including both OFL and SFL events.

The best-fit values are summarized in Table 5.2. We have also tried applying cuts to the ξ^2 values in the minimization with determinant cuts. We have found ξ^2 cut values for each of the OFL and OSFL cases which yield about 15%, 40% and 75% of samples past, respectively. We see that the two lowest ξ^2 cuts approximately halve the errors for both the OFL and the OSFL cases, while the loosest cut gives less of an error reduction. The OFL and OSFL cases obtain similar results — the OSFL case has somewhat smaller errors, but a larger downward bias.

[2]The two extra matrices are never the right ones, since they correspond to wrong lepton permutations between chains.

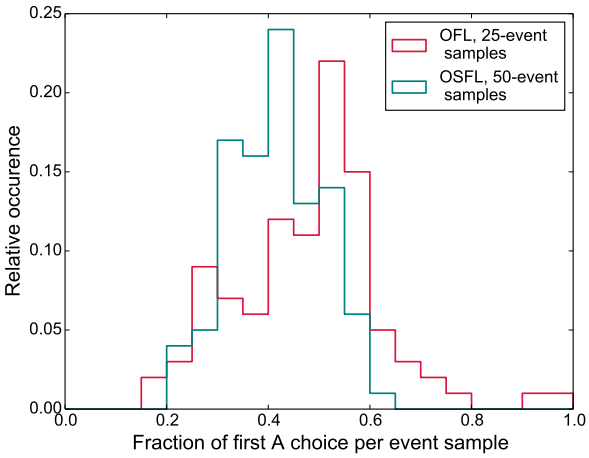


Figure 5.18: Distribution of the fraction of events where the **A** selection algorithm prefers the first **A**. The `Pythia` dataset with jet reconstruction is used, for samples of 25 events with only opposite flavour leptons, and samples of 50 events with both opposite and same flavour leptons.

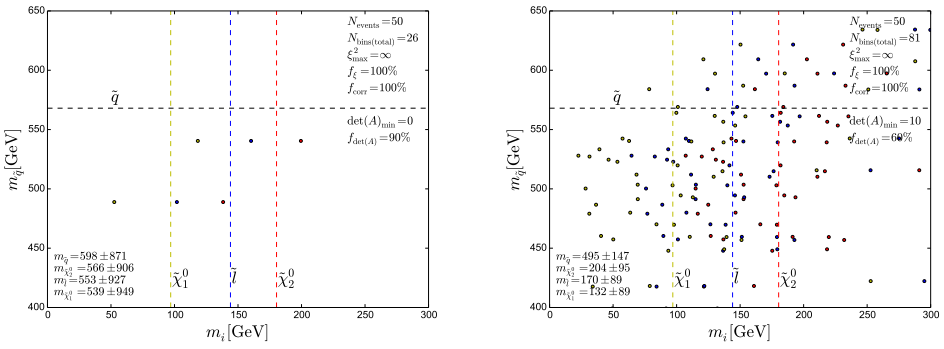


Figure 5.19: Minimization on the `Pythia` dataset with jet reconstruction and kinematical cuts, using 50-event samples with both OFL and SFL events. Most of the points in **(a)** lie outside the plotting region.

Type (N_{evt})	ξ^2_{max} (f_ξ)	$f_{\mathrm{det}\mathbf{A}}$	$m_{\tilde{q}}(568)$	$m_{\tilde{\chi}^0_2}(180)$	$m_{\tilde{l}}(144)$	$m_{\tilde{\chi}^0_1}(97)$
OFL (25)	∞	56%	525 ± 154	213 ± 90	177 ± 87	138 ± 88
OFL (25)	$5 \cdot 10^5$ (71%)	55%	551 ± 105	205 ± 82	166 ± 82	123 ± 84
OFL (25)	$5 \cdot 10^4$ (41%)	54%	536 ± 64	180 ± 56	141 ± 56	98 ± 57
OFL (25)	10^4 (18%)	50%	535 ± 67	178 ± 46	139 ± 45	96 ± 48
OSFL (50)	∞	60%	495 ± 147	204 ± 95	170 ± 89	132 ± 89
OSFL (50)	$3 \cdot 10^6$ (77%)	60%	494 ± 110	190 ± 80	155 ± 75	115 ± 76
OSFL (50)	$3 \cdot 10^5$ (39%)	60%	503 ± 52	167 ± 38	128 ± 37	83 ± 40
OSFL (50)	$5 \cdot 10^4$ (14%)	58%	504 ± 54	163 ± 45	125 ± 44	79 ± 47

Table 5.2: Summary of results obtained on the `Pythia` dataset with jet reconstruction and kinematical cuts, using the MSD technique with the event-pair selection algorithm for choosing between $\mathbf{A}^a$ matrices and summing over the four lepton permutations. A $|\mathrm{det}\mathbf{A}| > 10$ cut has been applied.

Conclusions

The method presented by Webber in [1], although promising, has been shown to have inherent technical problems which lead to an underestimation of the errors and biases in the original article, as presented in Table 4.1. The challenges stem in part from the combinatorical ambiguity inherent in the chain.

When the estimates are revised, taking into account the need for a lower `Simplex` tolerance in order to resolve the function minimum, the error is significantly increased, and biases are introduced, as can be seen from the results presented in Table 4.3. In addition, there is a dependence upon the minimization starting point which adds to the uncertainties, demonstrated in Table 4.4. Of particular importance is the feature that when experimental uncertainties are simulated in the form of momentum smearing, and the convergence tolerance is set sufficiently low, the mass of the LSP is fitted to zero in many of the samples, see Fig. 4.7. This is the case even without taking into account combinatorics.

Attempts to modify the method to avoid these problems only has limited success. Although we have devised an implementation that is robust against different starting points for the minimization algorithm, handles the combinatorical ambiguities and resolves the LSP mass, we have not found a way to reduce the errors to values that are comparable to the results pertained in the original article. Table 5.1 gives a summary of the capabilities of the alternative implementation of the method, and should be compared to the original estimates given in Table 4.1. For the three lightest particles, the errors in our analysis are an order of magnitude larger than the original estimate. We also observe a significant downward bias on all four masses.

Jet reconstruction effects were not investigated by Webber in the original article. Taking these into account should be expected to affect the precision attainable by the method, and we indeed find that it drastically increases the errors. Our results with jet reconstruction are summarized in Table 5.2. Even when we apply a quality cut on the ξ^2 values of the fits which remove about half the event samples, along with a minimal value cut on the determinant of the kinematics matrix for each event, we obtain errors of the order 50 GeV on all four masses. For the LSP mass this corresponds to an error of 50%. Doubling the event samples by including events where all four leptons have the same flavour reduces the errors somewhat, but at the expense of solidifying a significant downward

bias in mass estimate.

In summary, we have shown that the errors and biases were underestimated in the original article. We have revised the estimates to be more realistic. In addition, our analysis employs event samples which are large compared to the expected event yield given current limits on sparticle masses, at least in the CMSSM. In non-CMSSM models, the cross section for the chain can be larger. Reducing the sample size is expected to further increase the uncertainties.

Appendix A

The Dirac equation

Einstein's fundamental relationship between energy, momentum and mass in special relativity is

$$p^\mu p_\mu - m^2 = 0. \tag{A.1}$$

Following the prescriptions of quantum mechanics, energy and momentum are promoted to operators:

$$p_\mu \to i\hbar \frac{\partial}{\partial x^\mu} \equiv i\hbar \partial_\mu. \tag{A.2}$$

Taken as an eigenvalue equation for a free field ϕ, this gives rise to the relativistic *Klein-Gordon* equation

$$-\hbar^2 \partial^\mu \partial_\mu \phi - m^2 \phi = 0. \tag{A.3}$$

This equation differs from the Schrödinger equation of non-relativistic quantum mechanics in that it is second order in $\partial_0 = \partial_t$, while the Schrödinger equation is first order. To obtain a first-order differential equation in time, Dirac attempted [59] to factor the energy-momentum relation (A.1) as

$$p^\mu p_\mu - m^2 = \left(\gamma^\kappa p_\kappa + m\right)\left(\gamma^\lambda p_\lambda - m\right). \tag{A.4}$$

This leads to the requirement

$$p^\mu p_\mu = \gamma^\mu \gamma^\nu p_\mu p_\nu. \tag{A.5}$$

The right hand side can be written as

$$\gamma^\mu \gamma^\nu p_\mu p_\nu = \frac{1}{2}\left(\gamma^\mu \gamma^\nu p_\mu p_\nu + \gamma^\nu \gamma^\mu p_\nu p_\mu\right) \tag{A.6}$$

$$= \frac{1}{2}\left(\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu\right) p_\mu p_\nu, \tag{A.7}$$

where we have used the fact that μ and ν are dummy summation variables. This implies that

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \tag{A.8}$$

where $g^{\mu\nu}$ is the Minkowski metric tensor, and the curly brackets denote the anticommutator $\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu$. Dirac realized that this condition can be met provided we let the γ^μ be *matrices*. It turns out that the lowest possible dimension of these matrices is 4×4. They may be represented in several equivalent bases, but one standard convention is

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{A.9}$$

where σ^i are the Pauli matrices, 1 denotes the 2×2 identity matrix and 0 the 2×2 matrix of zeroes. With this factorization, the relativistic momentum-energy relation is satisfied by

$$\gamma^\mu p_\mu - m = 0. \tag{A.10}$$

The Dirac equation is obtained by substituting operators and letting them act on an object Ψ, giving

$$(i\hbar\gamma^\mu\partial_\mu - m)\,\Psi = 0, \tag{A.11}$$

where Ψ is required by dimensionality to be a four-column vector of fields,

$$\Psi = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix}, \tag{A.12}$$

called a Dirac *spinor*.

The *equations of motion* are the equations describing the field configurations which minimize the *action*,

$$S = \int d^4x\,\mathcal{L}. \tag{A.13}$$

Demanding stationarity of action for a spinor field Ψ and its adjoint $\bar{\Psi} \equiv \Psi^\dagger\gamma^0$,

$$\delta S = 0, \tag{A.14}$$

gives the *Euler-Lagrange* equations of motion

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\Psi\right)}\right) - \frac{\partial\mathcal{L}}{\partial\Psi} = 0, \tag{A.15}$$

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\bar{\Psi}\right)}\right) - \frac{\partial\mathcal{L}}{\partial\bar{\Psi}} = 0. \tag{A.16}$$

The Lagrangian $\mathcal{L}_D$ which yields the Dirac equation upon application of the equations of motion is

$$\mathcal{L}_D = \bar{\Psi}\left(i\hbar\partial\!\!\!/ - m\right)\Psi, \tag{A.17}$$

which is called the Dirac bilinear.

Appendix B

Higgs mass loop corrections

This appendix contains calculations of the one-loop contributions to the Higgs mass from fermionic and scalar particles.

B.1 Fermion loop

The diagram is given in Fig. B.1. The vertex factors are $-i\frac{y}{\sqrt{2}}$, where y is the

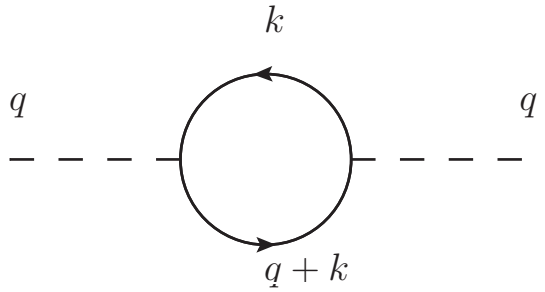


Figure B.1: Fermionic loop correction to the Higgs mass.

Yukawa coupling of the fermion, which has mass m. The amplitude is then given by

$$i\mathcal{M} = -\,(-i)^2\frac{y^2}{2}\mathrm{Tr}\left\{\int\frac{d^4k}{(2\pi)^4}\frac{\left[\not{q}+\not{k}+m\right]\left[\not{k}+m\right]}{\left[(q+k)^2-m^2\right]\left[k^2-m^2\right]}\right\}. \tag{B.1}$$

The denominators can be combined by a Feynman parametrization,

$$\begin{aligned}\frac{1}{\left[(q+k)^2-m^2\right]\left[k^2-m^2\right]} &= \int_0^1 dx\,\frac{1}{\left[\left((q+k)^2-m^2\right)x+\left(k^2-m^2\right)(1-x)\right]^2}\\ &= \int_0^1 dx\,\frac{1}{\left[(k+qx)^2-\Delta\right]^2},\end{aligned} \tag{B.2}$$

where $\Delta \equiv q^2 x(x-1) + m^2$. Defining $\ell \equiv k + qx$, we get

$$i\mathcal{M} = -\frac{y^2}{2}\int_0^1 dx \int \frac{d^4\ell}{(2\pi)^4} \mathrm{Tr}\left\{\frac{\not{\ell}^2 + \not{q}^2 x(x-1) + m^2}{\left[\ell^2 - \Delta\right]^2}\right\}, \tag{B.3}$$

where we have omitted terms linear in ℓ, since they vanish by the antisymmetry of the integral, and terms with an odd number of γ matrices since they trace to zero. The remaining trace is found by using that

$$\mathrm{Tr}\left\{\gamma^\mu \gamma^\nu\right\} = 4g^{\mu\nu}, \tag{B.4}$$

and remembering that there is really also a 4×4 identity matrix in the m^2 term. We then get

$$\begin{aligned} i\mathcal{M} &= -\frac{y^2}{2}\int_0^1 dx \int \frac{d^4\ell}{(2\pi)^4} 4 \frac{\ell^2 + q^2 x(x-1) + m^2}{\left[\ell^2 - \Delta\right]^2} \\ &= -2y^2 \int_0^1 dx \int \frac{d^4\ell}{(2\pi)^4} 4 \frac{\ell^2 + \Delta}{\left[\ell^2 - \Delta\right]^2}. \end{aligned} \tag{B.5}$$

We now make a Wick rotation by introducing the Euclidian variable ℓ_E, given by $\ell^0 = i\ell_E^0$, $\vec{\ell} = \vec{\ell}_E$, such that

$$\ell^2 = -\ell_E^2, \quad d^4\ell = i d^4\ell_E. \tag{B.6}$$

This gives

$$i\mathcal{M} = i8y^2 \int_0^1 dx \int \frac{d^4\ell_E}{(2\pi)^4} \frac{{\ell_E}^2 - \Delta}{\left[\ell_E^2 + \Delta\right]^2}, \tag{B.7}$$

which can be evaluated in spherical coordinates,

$$\int d^4\ell_E = \int d\Omega_4 \int d|\ell_E| \, |\ell_E|^3, \tag{B.8}$$

where Ω_4 is the four-dimensional solid angle, which by spherical symmetry just gives

$$\int d\Omega_4 = 2\pi^2. \tag{B.9}$$

By using a cutoff regularization parametrized by Λ, the whole amplitude is then

$$\begin{aligned} i\mathcal{M} =& -i\frac{4\pi^2}{(2\pi)^4}y^2\int_0^1 dx\,\left[\int_0^\Lambda d|\ell_E|\frac{|\ell_E|^5}{\left(|\ell_E|^2+\Delta\right)^2}-\Delta\int_0^\Lambda\frac{|\ell_E|^3}{\left(|\ell_E|^2+\Delta\right)^2}\right] \\ =& -iy^2\frac{1}{4\pi^2}\int_0^1 dx\,\left[\frac{-2\Delta^2+\Delta\Lambda^2+\Lambda^4-3\Delta\left(\Delta+\Lambda^2\right)\log\left(\Delta+\Lambda^2\right)}{2\left(\Delta+\Lambda^2\right)}\right. \\ &\left.-\frac{-2\Delta^2-4\Delta^2\log\Delta}{2\Delta}\right] \\ \overset{\Lambda^2\gg\Delta}{\to}& -iy^2\frac{1}{4\pi^2}\frac{\Lambda^4}{2\Lambda^2} \\ =& -i\frac{y^2}{8\pi^2}\Lambda^2. \end{aligned} \tag{B.10}$$

For $y=|\lambda_f|$, this gives the first term in Eq. (1.47).

B.2 Scalar loop

The diagram is given in Fig. B.2. The vertex factor is taken to be $i\lambda$, and m is

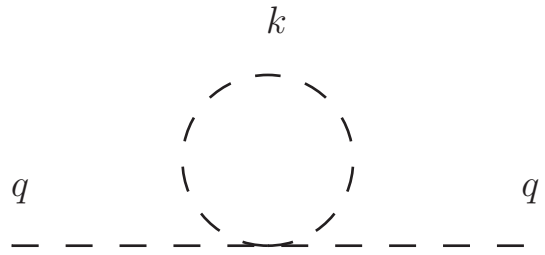


Figure B.2: Scalar loop correction to the Higgs mass.

the mass of the scalar in the loop. The amplitude is given by

$$i\mathcal{M}=i\lambda\int\frac{d^4k}{(2\pi)^4}\,\frac{i}{k^2-m^2}. \tag{B.11}$$

By Wick rotating $k\to k_E$ and switching to spherical coordinates as in the previous section, we get

$$\begin{aligned} i\mathcal{M} =& i\lambda\frac{1}{(2\pi)^4}\int d\Omega_4\int_0^\Lambda d|k_E|\,\frac{|k_E|^3}{|k_E|^2+m^2} \\ =& i\frac{\lambda}{16\pi^2}\left(\Lambda^2+m^2\log\left(\frac{m^2}{m^2+\Lambda^2}\right)\right) \\ \overset{\Lambda^2\gg m^2}{\to}& i\frac{\lambda}{16\pi^2}\Lambda^2. \end{aligned} \tag{B.12}$$

For $\lambda = \lambda_{\tilde{f}}$, this is the second term in Eq. (1.47).

Note that when the leading, quadratic terms cancel each other, the remainder in Eq. (B.12) is

$$m^2 \log\left(\frac{m^2}{\Lambda^2}\right), \tag{B.13}$$

which does *not* cancel with the fermion loop, since it depends on the scalar mass. If the scalar, *i.e.* the superpartner particle, is heavy, then this contribution may be significant. This puts restrictions on the sfermion masses in order not to reintroduce the fine-tuning, and is sometimes called the *little hierarchy problem*.

Appendix C

Invariant mass with lepton misidentification

The invariant mass of particle a and A in the decay topology of Fig. C.1 is given by $m_{aA}^2 = (p_a + p_A)^2 = m_B^2$. However, if the particle b is misidentified as a, what is the resulting invariant mass m_X^2? The calculation is most easily done in the rest frame of B. Let $p_a = (E_a, \mathbf{p}_a), p_A = (E_A, \mathbf{p}_A)$ in this frame. We

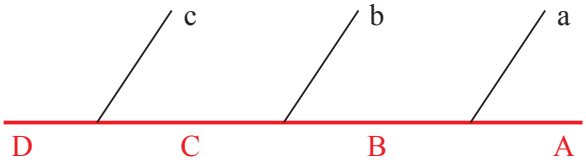


Figure C.1: Decay topology, from [42].

assume $m_b = m_a = 0$. The decay must go back-to-back, so $\mathbf{p}_A = -\mathbf{p}_a$. Then, the condition

$$m_B^2 = (p_A + p_a)^2 = p_A^2 + 2p_A \cdot p_a + p_a^2 \tag{C.1}$$

$$= m_A^2 + 2\left(m_A^2 + |\mathbf{p}_a|^2\right)^{1/2} |\mathbf{p}_a| + 2|\mathbf{p}_a|^2, \tag{C.2}$$

gives

$$|\mathbf{p}_a| = |\mathbf{p}_A| = \frac{m_B^2 - m_A^2}{2m_B}. \tag{C.3}$$

Thus, if $\hat{\mathbf{p}}$ is a unit vector in the direction of A in the rest frame of B, we have

$$p_A = \left(\sqrt{m_A^2 + \frac{(m_B^2 - m_A^2)^2}{4m_B^2}}, \frac{m_B^2 - m_A^2}{2m_B}\hat{\mathbf{p}}\right) = \left(\frac{m_B^2 + m_A^2}{2m_B}, \frac{m_B^2 - m_A^2}{2m_B}\hat{\mathbf{p}}\right), \tag{C.4}$$

$$p_a = \left(\frac{m_B^2 - m_A^2}{2m_B}, -\frac{m_B^2 - m_A^2}{2m_B}\hat{\mathbf{p}}\right).$$

To evaluate m_X^2, we must evaluate p_b and p_A in the same reference frame. We choose to transform p_b' to p_b in the rest frame of B. This is easiest by using the invariance of the Minkowski norm, which gives

$$\begin{aligned}(p_b + p_B)^2 &= (p_b' + p_B')^2 = m_C^2 \\ &= m_B^2 + 2m_B|\mathbf{p}_b| \\ \Rightarrow |\mathbf{p}_b| &= \frac{m_C^2 - m_B^2}{2m_B},\end{aligned} \tag{C.5}$$

where we have used that $\mathbf{p}_B = \mathbf{0}$ in the rest frame of B. Now, we may assume without loss of generality that the decay products of $C \to Bb$ are aligned along the x-axis. The decay $B \to Aa$ will make an angle θ with the x-axis. This is illustrated in figure C.2. The spherical distribution of the decay direction of A

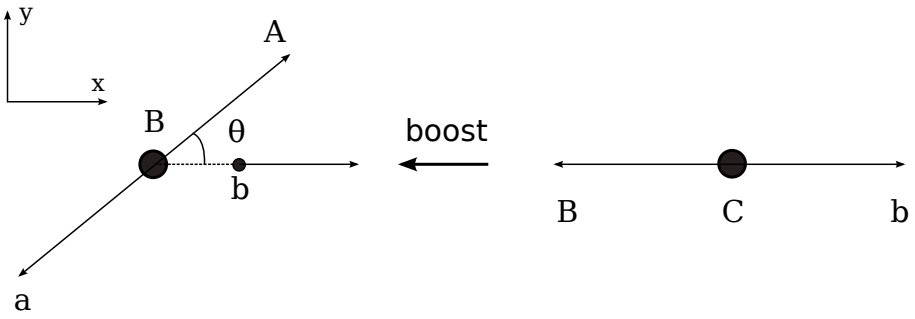


Figure C.2: Lorentz boost from rest frame of C to rest frame of B.

is isotropic in the rest frame of B if B is a scalar, which means that $\cos\theta$ is uniformly distributed. We then get

$$\begin{aligned}m_X^2 = (p_A + p_b)^2 &= m_A^2 + 2E_A|\mathbf{p}_b| - 2|\mathbf{p}_A||\mathbf{p}_b|\cos\theta \\ &= m_A^2 + 2\left(m_A^2 + \frac{(m_B^2 - m_A^2)^2}{4m_B^2}\right)^{1/2}\frac{m_C^2 - m_B^2}{2m_B} - 2\frac{m_B^2 - m_A^2}{2m_B}\frac{m_C^2 - m_B^2}{2m_B}\cos\theta \\ &= m_A^2 + \frac{m_C^2 - m_B^2}{2m_B^2}\left(m_B^2(1 - \cos\theta) + m_A^2(1 + \cos\theta)\right).\end{aligned} \tag{C.6}$$

The limiting cases are

$$\begin{aligned}\cos\theta = -1 &\Rightarrow m_{X\,\mathrm{high}}^2 = m_A^2 + m_C^2 - m_B^2, \\ \cos\theta = +1 &\Rightarrow m_{X\,\mathrm{low}}^2 = \frac{m_C^2 m_A^2}{m_B^2}.\end{aligned} \tag{C.7}$$

We see that m_X^2 is monotonically decreasing as a function of $\cos\theta$ between these limits, which means m_X is as well. We note that the interval $(m_{X\mathrm{low}}, m_{X\mathrm{high}})$ is always a subset of the interval (m_A, m_C) for any physical combination of $m_{A,B,C}$,

so m_X is always consistent with chain kinematics. For the example of SPS1a discussed in the text, m_A, m_B and m_C are 97 GeV, 144 GeV and 180 GeV, respectively. The limits then evaluate to $m_{X\,\mathrm{high}} = 145$ GeV and $m_{X\,\mathrm{low}} = 121$ GeV. The probability distribution for m_X in this case — assuming that B is a scalar, as indeed it is if it is a slepton — is shown in Fig. C.3.

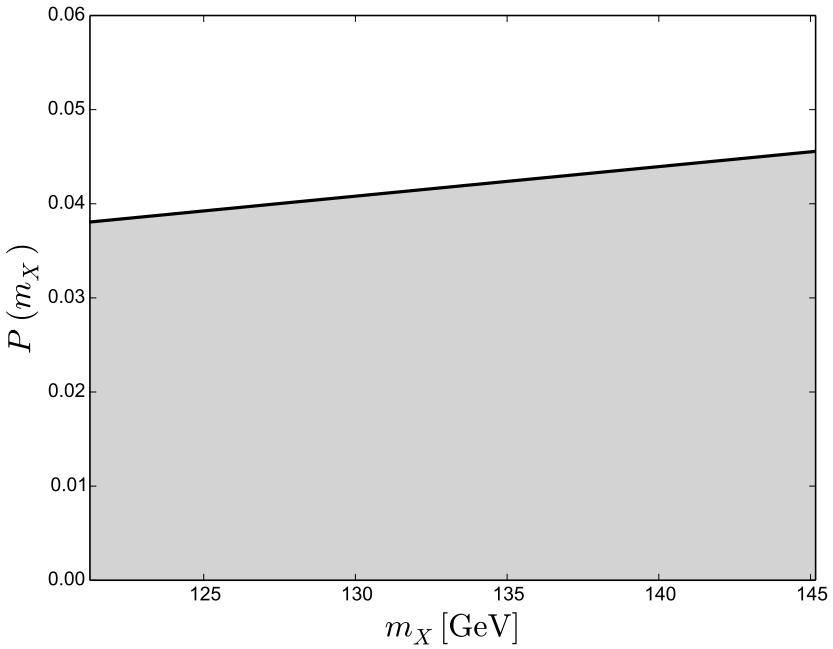


Figure C.3: Probability distribution of m_X.

Appendix D

A C++ implementation of the Nelder-Mead `Simplex` algorithm

Adapted from [60]. The function to call for minimization is `amoeba`. It returns true or false depending on whether convergence has been obtained within the set number of iterations. The `xisquared` function uses the `Armadillo` linear algebra library [61].

```
// Implementation of Nelder-Mead Simplex method:
double * alloc_vector(int cols)
{
  return (double *) malloc(sizeof(double) * cols);
}
void free_vector(double * vector , int cols)
{
  free(vector);
}
double ** alloc_matrix(int rows, int cols)
{
  int i;
  double ** matrix = (double **) malloc(sizeof(double *) *
     rows);
  for (i = 0; i < rows; i++)
    matrix[i] = alloc_vector(cols);
  return matrix;
}
void free_matrix(double ** matrix, int rows, int cols)
{
  int i;
  for (i =0; i < rows; i++)
    free_vector(matrix[i], cols);
  free(matrix);
}
double ** make_simplex(double * point, int dim)
{
```

```
  int i, j;
  double ** simplex = alloc_matrix(dim + 1, dim);
  for (i = 0; i < dim + 1; i++)
    for (j = 0; j < dim; j++)
      simplex[i][j] = point[j];
  for (i = 0; i < dim; i++)
    simplex[i][i] += 1.0;
  return simplex;
}
void evaluate_simplex(double ** simplex, int dim, double *
   fx, double (*func)(double *, int, int, double,
   vector<bool> &, vector<vector<mat>> &, vector<vector<vec>>
   &, vector<pair<int,int> >, int, int),
  int Nevents, int jBin, double Mnorm, vector<bool>
     &all_leptons_equal_list, vector<vector<mat>> &D_lists,
     vector<vector<vec>> &E_lists, vector<pair<int,int> >
     eventIndices, int minimization_method, int combosum)
{
  int i;
  for (i = 0; i < dim + 1; i++)
  {
    fx[i] = (*func)(simplex[i], Nevents, jBin, Mnorm,
       all_leptons_equal_list, D_lists, E_lists,
       eventIndices, minimization_method, combosum);
  }
}

void simplex_extremes(double *fx, int dim, int & ihi, int &
   ilo,int & inhi)
{
  int i;
  if (fx[0] > fx[1])
  { ihi = 0; ilo = inhi = 1; }
  else
  { ihi = 1; ilo = inhi = 0; }
  for (i = 2; i < dim + 1; i++)
    if (fx[i] <= fx[ilo])
      ilo = i;
    else if (fx[i] > fx[ihi])
      { inhi = ihi; ihi = i; }
    else if (fx[i] > fx[inhi])
      inhi = i;
}
void simplex_bearings(double ** simplex, int dim,double *
   midpoint, double * line, int ihi)
{
  int i, j;
  for (j = 0; j < dim; j++)
    midpoint[j] = 0.0;
  for (i = 0; i < dim + 1; i++)
    if (i != ihi)
```

```
      for (j = 0; j < dim; j++)
        midpoint[j] += simplex[i][j];

  for (j = 0; j < dim; j++)
  {
    midpoint[j] /= dim;
    line[j] = simplex[ihi][j] - midpoint[j];
  }
}
int update_simplex(double * point, int dim, double &
   fmax,double * midpoint, double * line, double scale,
   double (*func)(double *, int, int, double, vector<bool> &,
   vector<vector<mat>> &, vector<vector<vec>> &,
   vector<pair<int,int> >, int, int),
  int Nevents, int jBin, double Mnorm, vector<bool>
     &all_leptons_equal_list, vector<vector<mat>> &D_lists,
     vector<vector<vec>> &E_lists, vector<pair<int,int> >
     eventIndices, int minimization_method, int combosum)
{
  int i, update = 0;
  double * next = alloc_vector(dim), fx;
  for (i = 0; i < dim; i++)
    next[i] = midpoint[i] + scale * line[i];
  fx = (*func)(next, Nevents, jBin, Mnorm,
      all_leptons_equal_list, D_lists, E_lists, eventIndices,
      minimization_method, combosum);
  if (fx < fmax)
  {
    for (i = 0; i < dim; i++)
      point[i] = next[i];
    fmax = fx;
    update = 1;
  }
  free_vector(next, dim);
  return update;
}

void contract_simplex(double ** simplex, int dim, double *
   fx, int ilo, double (*func)(double *, int, int, double,
   vector<bool> &, vector<vector<mat>> &, vector<vector<vec>>
   &, vector<pair<int,int> >, int, int), int Nevents, int
   jBin, double Mnorm, vector<bool> &all_leptons_equal_list,
   vector<vector<mat>> &D_lists, vector<vector<vec>>
   &E_lists, vector<pair<int,int> > eventIndices, int
   minimization_method, int combosum)
{
  int i, j;
  for (i = 0; i < dim + 1; i++)
    if (i != ilo)
    {
      for (j = 0; j < dim; j++)
```

```
        simplex[i][j] = (simplex[ilo][j]+simplex[i][j])*0.5;
      fx[i] = (*func)(simplex[i], Nevents, jBin, Mnorm,
          all_leptons_equal_list, D_lists, E_lists,
          eventIndices, minimization_method, combosum);
    }
}


#define ZEPS 1e-30
int check_tol(double fmax, double fmin, double ftol)
{
double delta = fabs(fmax - fmin);
double accuracy = (fabs(fmax) + fabs(fmin)) * ftol;
// cout << delta << ", " << accuracy << ", " << ftol << endl;
return (delta < (accuracy + ZEPS));
}

bool amoeba(double *point, double &fmin, double
   (*func)(double *, int, int, double, vector<bool> &,
   vector<vector<mat>> &, vector<vector<vec>> &,
   vector<pair<int,int> >, int, int),
  int dim, double tol, int maxiter,
  int Nevents, int jBin, double Mnorm, vector<bool>
     &all_leptons_equal_list, vector<vector<mat> > &D_lists,
     vector<vector<vec> > &E_lists, vector<pair<int,int> >
     eventIndices, int minimization_method, int combosum)
{
  // Usage: Point is an allocated dim-dimensional array of
     doubles
  // to be filled with coordinates of the best-fit point,
  // func is the function to minimize.
  int ihi, ilo, inhi, j;
  double * fx = alloc_vector(dim + 1);
  double * midpoint = alloc_vector(dim);
  double * line = alloc_vector(dim);
  double ** simplex = make_simplex(point, dim);
  evaluate_simplex(simplex, dim, fx, func,
    Nevents, jBin, Mnorm, all_leptons_equal_list, D_lists,
       E_lists, eventIndices, minimization_method, combosum);

  int iter = 0;
  while (iter < maxiter)
  {
    simplex_extremes(fx, dim, ihi, ilo, inhi);
    simplex_bearings(simplex, dim, midpoint, line, ihi);
    if (check_tol(fx[ihi], fx[ilo], tol)) { /*cout << "below
       tol = " << tol << endl;*/ break; }
    update_simplex(simplex[ihi], dim, fx[ihi],
    midpoint, line, -1.0, func,
    Nevents, jBin, Mnorm, all_leptons_equal_list, D_lists,
       E_lists, eventIndices, minimization_method, combosum);
```

```
    if (fx[ihi] < fx[ilo])
      update_simplex(simplex[ihi], dim, fx[ihi], midpoint,
          line, -2.0, func,
        Nevents, jBin, Mnorm, all_leptons_equal_list,
            D_lists, E_lists, eventIndices,
            minimization_method, combosum);
    else if (fx[ihi] >= fx[inhi])
      if (!update_simplex(simplex[ihi], dim, fx[ihi],
          midpoint, line, 0.5, func, Nevents, jBin, Mnorm,
          all_leptons_equal_list, D_lists, E_lists,
          eventIndices, minimization_method, combosum))
        contract_simplex(simplex, dim, fx, ilo, func,
            Nevents, jBin, Mnorm, all_leptons_equal_list,
            D_lists, E_lists, eventIndices,
            minimization_method, combosum);
    iter += 1;
  }

  for (j = 0; j < dim; j++)
    point[j] = simplex[ilo][j];
  fmin = fx[ilo];
  free_vector(fx, dim);
  free_vector(midpoint, dim);
  free_vector(line, dim);
  free_matrix(simplex, dim + 1, dim);

  if (iter < maxiter)
  {
    return true;
  }
  else
    return false;
}
struct MomentumVector
{
  int id;
  arma::vec p;
};

double minkowskidot(vec a, vec b)
{
  return a[3]*b[3]-a[0]*b[0]-a[1]*b[1]-a[2]*b[2];
}

double xisquared(double *Masses, int Nevents, int j, double
    Mnorm, vector<bool> &all_leptons_equal_list,
    vector<vector<mat>> &D_lists, vector<vector<vec>>
    &E_lists, vector<pair<int,int> > eventIndices, int
    minimization_method, int combosum)
{
```

```
double xisquared = 0;

if (minimization_method == 1)
{
  // ORIGINAL FOUR-PARAMETER FIT
  vec M;
  M << Masses[0]*Masses[0] << Masses[1]*Masses[1] <<
      Masses[2]*Masses[2] << Masses[3]*Masses[3]
    << Masses[0]*Masses[0] << Masses[1]*Masses[1] <<
       Masses[2]*Masses[2] << Masses[3]*Masses[3];
  M = M/pow(Mnorm, 2);


  // Avoid regions of unphysical mass combinations by
     adding a huge contribution in a continuous way
  double hugefactor = 10000000.0;
  if (Masses[0] < 0) xisquared = xisquared +
     hugefactor*M[0]*M[0];
  if (Masses[1] < 0) xisquared = xisquared +
     hugefactor*M[1]*M[1];
  if (Masses[2] < 0) xisquared = xisquared +
     hugefactor*M[2]*M[2];
  if (Masses[3] < 0) xisquared = xisquared +
     hugefactor*M[3]*M[3];
  if (M[0] < M[1]) xisquared = xisquared +
     hugefactor*(M[0]-M[1])*(M[0]-M[1]);
  if (M[1] < M[2]) xisquared = xisquared +
     hugefactor*(M[1]-M[2])*(M[1]-M[2]);
  if (M[2] < M[3]) xisquared = xisquared +
     hugefactor*(M[2]-M[3])*(M[2]-M[3]);


  for (auto iteratorCurrentEvent:eventIndices)
  {
    int iEvent = iteratorCurrentEvent.first;
    int iCombo = iteratorCurrentEvent.second;
    // cout << "iteratorCurrentEvent = " << iEvent << ",
       "<< iCombo << endl;

    double xisquared_current;
    vec P;
    P = D_lists[iCombo][iEvent]*M + E_lists[0][iEvent];
    xisquared_current = pow(P[3]*P[3] - P[0]*P[0] -
       P[1]*P[1] -P[2]*P[2] - M[3], 2) + pow(P[7]*P[7] -
       P[4]*P[4] - P[5]*P[5] - P[6]*P[6] - M[3], 2);

    if (combosum == 2)
    {
      P = D_lists[1+iCombo][iEvent]*M + E_lists[1][iEvent];
      xisquared_current = xisquared_current + pow(P[3]*P[3]
         - P[0]*P[0] - P[1]*P[1] - P[2]*P[2] - M[3], 2 ) +
```

```
            pow(P[7]*P[7   ] - P[4]*P[4] - P[5]*P[5] -
            P[6]*P[6] - M[3], 2);
        P = D_lists[2+iCombo][iEvent]*M + E_lists[2][iEvent];
        xisquared_current = xisquared_current + pow(P[3]*P[3]
            - P[0]*P[0] - P[1]*P[1] - P[2]*P[2] - M[3], 2 ) +
            pow(P[7]*P[7   ] - P[4]*P[4] - P[5]*P[5] -
            P[6]*P[6] - M[3], 2);
        P = D_lists[3+iCombo][iEvent]*M + E_lists[3][iEvent];
        xisquared_current = xisquared_current + pow(P[3]*P[3]
            - P[0]*P[0] - P[1]*P[1] - P[2]*P[2] - M[3], 2 ) +
            pow(P[7]*P[7] - P[4]*P[4] - P[5]*P[5] - P[6]*P[6] -
             M[3], 2);
      }

      xisquared = xisquared + xisquared_current;

      // ADD IF TEST FOR COMBOSUM + REMEMBER TO CHECK SHIFT
         VALUE FOR COMBINATORICS DEPENDENT ON COMBOSUM
    }

  } // END IF minimization_method == 1
  else if (minimization_method == 2)
  {
    // MASS-DIFFERENCE FIT
    vec M;
    M << Masses[0]
      << Masses[1]
      << Masses[2];
    M = M/pow(Mnorm, 2);

    // Avoid regions of unphysical mass combinations by
       adding a huge contribution in a continuous way
    double hugefactor = 100000.0;
    if (Masses[0] < 0) xisquared = xisquared +
       hugefactor*M[0]*M[0];
    if (Masses[1] < 0) xisquared = xisquared +
       hugefactor*M[1]*M[1];
    if (Masses[2] < 0) xisquared = xisquared +
       hugefactor*M[2]*M[2];

    // Calculate current estimate for LSP mass from dilepton
       invariant mass edge
    double mllinv = 80.1; // Calculated from true masses
       using formula
    double MLSPsq = M[2]*(M[1]/(mllinv*mllinv) - 1.0);

    for (auto iteratorCurrentEvent : eventIndices)
    {
      int iEvent = iteratorCurrentEvent.first;
      int iCombo = iteratorCurrentEvent.second;
      // cout << "iteratorCurrentEvent = " << iEvent << ",
```

```
            "<< iCombo << endl;

        double xisquared_current;
        vec P;
        // cout << iCombo << endl;

        P = D_lists[0+4*iCombo][iEvent]*M +
            E_lists[0+4*iCombo][iEvent];
        xisquared_current = pow(P[3]*P[3] - P[0]*P[0] -
            P[1]*P[1] - P[2]*P[2] - MLSPsq, 2) + pow(P[7]*P[7] -
            P[4]*P[4] - P[5]*P[5] - P[6]*P[6] -MLSPsq, 2);
        // cout << xisquared_current << endl;

        if (combosum == 2)
        {
          // Add contributions from the three first wrong
              combos, where the leptons are flipped inside chains
          P = D_lists[1+4*iCombo][iEvent]*M +
              E_lists[1+4*iCombo][iEvent];
          xisquared_current = xisquared_current + pow(P[3]*P[3]
              - P[0]*P[0] - P[1]*P[1] - P[2]*P[2] - MLSPsq, 2 )
              + pow(P[7]*P[7] - P[4]*P[4] - P[5]*P[5] -
              P[6]*P[6] - MLSPsq, 2);
          P = D_lists[2+4*iCombo][iEvent]*M +
              E_lists[2+4*iCombo][iEvent];
          xisquared_current = xisquared_current + pow(P[3]*P[3]
              - P[0]*P[0] - P[1]*P[1] - P[2]*P[2] - MLSPsq, 2 )
              + pow(P[7]*P[7] - P[4]*P[4] - P[5]*P[5] -
              P[6]*P[6] - MLSPsq, 2);
          P = D_lists[3+4*iCombo][iEvent]*M +
              E_lists[3+4*iCombo][iEvent];
          xisquared_current = xisquared_current + pow(P[3]*P[3]
              - P[0]*P[0] - P[1]*P[1] - P[2]*P[2] - MLSPsq, 2 )
              + pow(P[7]*P[7] - P[4]*P[4] - P[5]*P[5] -
              P[6]*P[6] - MLSPsq, 2);
        }
        xisquared = xisquared + xisquared_current;
      }
    } // END IF minimization_method == 2


    return xisquared;

}
```

Bibliography

[1] B. Webber, «Mass determination in sequential particle decay chains», *JHEP*, vol. 0909, p. 124, 2009. DOI: 10.1088/1126-6708/2009/09/124. arXiv: 0907.5307 [hep-ph].

[2] S. P. Martin, «A Supersymmetry primer», *Adv. Ser. Direct. High Energy Phys.*, vol. 21, pp. 1–153, 2010. DOI: 10.1142/9789814307505_0001. arXiv: hep-ph/9709356 [hep-ph].

[3] A. J. Barr and C. G. Lester, «A Review of the Mass Measurement Techniques proposed for the Large Hadron Collider», *J. Phys.*, vol. G37, p. 123 001, 2010. DOI: 10.1088/0954-3899/37/12/123001. arXiv: 1004.2732 [hep-ph].

[4] B. Webber, Private communication.

[5] G. Aad *et al.*, «Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC», *Phys. Lett.*, vol. B716, pp. 1–29, 2012. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[6] S. Chatrchyan *et al.*, «Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC», *Phys. Lett.*, vol. B716, pp. 30–61, 2012. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[7] E. Noether, «Invariante Variationsprobleme», *Gott. Nachr.*, vol. 1918, pp. 235–257, 1918. DOI: 10.1080/00411457108231446. arXiv: physics/0503066 [physics].

[8] Wikimedia Commons, *Standard Model of Elementary Particles*, Licensed under the Creative Commons Attribution 3.0 Unported license. [Online]. Available: {http://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg}.

[9] P. W. Anderson, «Plasmons, Gauge Invariance, and Mass», *Physical Review*, vol. 130, pp. 439–442, Apr. 1963. DOI: 10.1103/PhysRev.130.439.

[10] G. S. Guralnik, C. R. Hagen, and T. W. Kibble, «Global Conservation Laws and Massless Particles», *Physical Review Letters*, vol. 13, pp. 585–587, Nov. 1964. DOI: 10.1103/PhysRevLett.13.585.

[11] F. Englert and R. Brout, «Broken Symmetry and the Mass of Gauge Vector Mesons», *Physical Review Letters*, vol. 13, pp. 321–323, Aug. 1964. DOI: 10.1103/PhysRevLett.13.321.

[12] P. W. Higgs, «Broken Symmetries and the Masses of Gauge Bosons», *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964. DOI: 10.1103/PhysRevLett.13.508.

[13] L. Alvarez-Gaume and J. Ellis, «Eyes on a prize particle», *Nat Phys*, vol. 7, pp. 2–3, 2011.

[14] D. Kaiser, «Physics and Feynman's Diagrams», *American Scientist*, vol. 93, no. 2, p. 156, 2005.

[15] D. Binosi and L. Theussl, «JaxoDraw: A Graphical user interface for drawing Feynman diagrams», *Comput. Phys. Commun.*, vol. 161, pp. 76–86, 2004. DOI: 10.1016/j.cpc.2004.05.001. arXiv: hep-ph/0309015 [hep-ph].

[16] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[17] S. Coleman and J. Mandula, «All possible symmetries of the S matrix», *Phys. Rev.*, vol. 159, pp. 1251–1256, 5 Jul. 1967. DOI: 10.1103/PhysRev.159.1251. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRev.159.1251.

[18] R. Haag, J. T. Łopuszański, and M. Sohnius, «All possible generators of supersymmetries of the S-matrix», *Nuclear Physics B*, vol. 88, no. 2, pp. 257–274, 1975, ISSN: 0550-3213. DOI: http://dx.doi.org/10.1016/0550-3213(75)90279-5. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0550321375902795.

[19] K. Olive *et al.*, «Review of Particle Physics», *Chin. Phys.*, vol. C38, p. 090 001, 2014. DOI: 10.1088/1674-1137/38/9/090001.

[20] P. Bätzing and A. Raklev, *Lecture notes for FYS5190/FYS9190 – Supersymmetry*, A course given at the University of Oslo, 2013.

[21] B. Gjelsten, D. Miller, and P. Osland, «Measurement of SUSY masses via cascade decays for SPS 1a», *JHEP*, vol. 0412, p. 003, 2004. DOI: 10.1088/1126-6708/2004/12/003. arXiv: hep-ph/0410303 [hep-ph].

[22] B. Allanach, M. Battaglia, G. Blair, M. S. Carena, A. De Roeck, *et al.*, «The Snowmass points and slopes: Benchmarks for SUSY searches», *Eur. Phys. J.*, vol. C25, pp. 113–123, 2002. DOI: 10.1007/s10052-002-0949-3. arXiv: hep-ph/0202233 [hep-ph].

[23] H. Baer, F. E. Paige, S. D. Protopopescu, and X. Tata, «ISAJET 7.48: A Monte Carlo event generator for p p, anti-p, p, and e+ e- reactions», 1999. arXiv: hep-ph/0001086 [hep-ph].

[24] B. Allanach, «SOFTSUSY: a program for calculating supersymmetric spectra», *Comput. Phys. Commun.*, vol. 143, pp. 305–331, 2002. DOI: 10.1016/S0010-4655(01)00460-X. arXiv: hep-ph/0104145 [hep-ph].

[25] A. Buckley, «PySLHA: a Pythonic interface to SUSY Les Houches Accord data», 2013. arXiv: 1305.4194 [hep-ph].

[26] ATLAS collaboration, «Search for supersymmetry in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector», 2015. arXiv: 1503.03290 [hep-ex].

[27] ATLAS and CMS collaborations, «Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments», *Phys. Rev. Lett.*, vol. 114, p. 191 803, 2015. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex].

[28] ATLAS collaboration, «Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using $\sqrt{s} = 8$ TeV proton–proton collision data», *JHEP*, vol. 1409, p. 176, 2014. DOI: 10.1007/JHEP09(2014)176. arXiv: 1405.7875 [hep-ex].

[29] ATLAS collaboration, «Search for new phenomena in final states with large jet multiplicities and missing transverse momentum at $\sqrt{s}$=8 TeV proton-proton collisions using the ATLAS experiment», *JHEP*, vol. 1310, p. 130, 2013. DOI: 10.1007/JHEP10(2013)130,10.1007/JHEP01(2014)109. arXiv: 1308.1841 [hep-ex].

[30] ATLAS collaboration, «Search for strong production of supersymmetric particles in final states with missing transverse momentum and at least three b-jets at $\sqrt{s}$= 8 TeV proton-proton collisions with the ATLAS detector», *JHEP*, vol. 1410, p. 24, 2014. DOI: 10.1007/JHEP10(2014)024. arXiv: 1407.0600 [hep-ex].

[31] ATLAS collaboration, «Search for squarks and gluinos in events with isolated leptons, jets and missing transverse momentum at $\sqrt{s} = 8$ TeV with the ATLAS detector», *JHEP*, vol. 1504, p. 116, 2015. DOI: 10.1007/JHEP04(2015)116. arXiv: 1501.03555 [hep-ex].

[32] ATLAS collaboration, «Search for supersymmetry in events with large missing transverse momentum, jets, and at least one tau lepton in 20 fb^{-1} of $\sqrt{s} = 8$ TeV proton-proton collision data with the ATLAS detector», *JHEP*, vol. 1409, p. 103, 2014. DOI: 10.1007/JHEP09(2014)103. arXiv: 1407.0603 [hep-ex].

[33] ATLAS collaboration, «Search for supersymmetry at $\sqrt{s}$=8 TeV in final states with jets and two same-sign leptons or three leptons with the ATLAS detector», *JHEP*, vol. 1406, p. 035, 2014. DOI: 10.1007/JHEP06(2014)035. arXiv: 1404.2500 [hep-ex].

[34] S. Heinemeyer, «The MSSM Higgs Sector at the LHC and Beyond», 2015. arXiv: 1504.07957 [hep-ph].

[35] W. Beenakker, R. Hopker, M. Spira, and P. Zerwas, «Squark and gluino production at hadron colliders», *Nucl. Phys.*, vol. B492, pp. 51–103, 1997. DOI: 10.1016/S0550-3213(97)80027-2. arXiv: hep-ph/9610490 [hep-ph].

[36] A. Kulesza and L. Motyka, «Threshold resummation for squark-antisquark and gluino-pair production at the LHC», *Phys. Rev. Lett.*, vol. 102, p. 111 802, 2009. DOI: 10.1103/PhysRevLett.102.111802. arXiv: 0807.2405 [hep-ph].

[37] A. Kulesza *et al.*, «Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC», *Phys. Rev.*, vol. D80, p. 095 004, 2009. DOI: 10.1103/PhysRevD.80.095004. arXiv: 0905.4749 [hep-ph].

[38] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, *et al.*, «Soft-gluon resummation for squark and gluino hadroproduction», *JHEP*, vol. 0912, p. 041, 2009. DOI: 10.1088/1126-6708/2009/12/041. arXiv: 0909.4418 [hep-ph].

[39] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, *et al.*, «Squark and Gluino Hadroproduction», *Int. J. Mod. Phys.*, vol. A26, pp. 2637–2664, 2011. DOI: 10.1142/S0217751X11053560. arXiv: 1105.1110 [hep-ph].

[40] W. Beenakker, M. Kramer, T. Plehn, M. Spira, and P. Zerwas, «Stop production at hadron colliders», *Nucl. Phys.*, vol. B515, pp. 3–14, 1998. DOI: 10.1016/S0550-3213(98)00014-5. arXiv: hep-ph/9710451 [hep-ph].

[41] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, *et al.*, «Supersymmetric top and bottom squark production at hadron colliders», *JHEP*, vol. 1008, p. 098, 2010. DOI: 10.1007/JHEP08(2010)098. arXiv: 1006.4771 [hep-ph].

[42] D. Miller, P. Osland, and A. Raklev, «Invariant mass distributions in cascade decays», *JHEP*, vol. 0603, p. 034, 2006. DOI: 10.1088/1126-6708/2006/03/034. arXiv: hep-ph/0510356 [hep-ph].

[43] B. C. Allanach, C. G. Lester, M. A. Parker, and B. R. Webber, «Measuring sparticle masses in non-universal string inspired models at the lhc», *Journal of High Energy Physics*, vol. 2000, no. 09, p. 004, 2000. [Online]. Available: http://stacks.iop.org/1126-6708/2000/i=09/a=004.

[44] H.-C. Cheng, K. T. Matchev, and M. Schmaltz, «Bosonic supersymmetry? Getting fooled at the CERN LHC», *Phys.Rev.*, vol. D66, p. 056 006, 2002. DOI: 10.1103/PhysRevD.66.056006. arXiv: hep-ph/0205314 [hep-ph].

[45] M. Cacciari, G. P. Salam, and G. Soyez, «FastJet User Manual», *Eur. Phys. J.*, vol. C72, p. 1896, 2012. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[46] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, *et al.*, «HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)», *JHEP*, vol. 0101, p. 010, 2001. DOI: 10.1088/1126-6708/2001/01/010. arXiv: hep-ph/0011363 [hep-ph].

[47] S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, «Implementation of supersymmetric processes in the HERWIG event generator», *JHEP*, vol. 0204, p. 028, 2002. DOI: 10.1088/1126-6708/2002/04/028. arXiv: hep-ph/0204123 [hep-ph].

[48] F. James and M. Roos, «Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations», *Comput. Phys. Commun.*, vol. 10, pp. 343–367, 1975. DOI: 10.1016/0010-4655(75)90039-9.

[49] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, *et al.*, «Herwig++ Physics and Manual», *Eur. Phys. J.*, vol. C58, pp. 639–707, 2008. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[50] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, *et al.*, «An Introduction to PYTHIA 8.2», 2014. arXiv: 1410.3012 [hep-ph].

[51] J. A. Nelder and R. Mead, «A simplex method for function minimization», *The computer journal*, vol. 7, no. 4, pp. 308–313, 1965.

[52] P. Z. Skands, B. Allanach, H. Baer, C. Balazs, G. Belanger, *et al.*, «SUSY Les Houches accord: Interfacing SUSY spectrum calculators, decay packages, and event generators», *JHEP*, vol. 0407, p. 036, 2004. DOI: 10.1088/1126-6708/2004/07/036. arXiv: hep-ph/0311123 [hep-ph].

[53] E. Richter-Was, «AcerDET: A Particle level fast simulation and reconstruction package for phenomenological studies on high p(T) physics at LHC», 2002. arXiv: hep-ph/0207355 [hep-ph].

[54] G. Aad *et al.*, «The ATLAS Experiment at the CERN Large Hadron Collider», *JINST*, vol. 3, S08003, 2008. DOI: 10.1088/1748-0221/3/08/S08003.

[55] M. M. Nojiri, K. Sakurai, and B. R. Webber, «Reconstructing particle masses from pairs of decay chains», *JHEP*, vol. 1006, p. 069, 2010. DOI: 10.1007/JHEP06(2010)069. arXiv: 1005.2532 [hep-ph].

[56] H.-C. Cheng, J. F. Gunion, Z. Han, and B. McElrath, «Accurate Mass Determinations in Decay Chains with Missing Energy. II», *Phys. Rev.*, vol. D80, p. 035 020, 2009. DOI: 10.1103/PhysRevD.80.035020. arXiv: 0905.1344 [hep-ph].

[57] B. Gripaios, K. Sakurai, and B. Webber, «Polynomials, Riemann surfaces, and reconstructing missing-energy events», *JHEP*, vol. 1109, p. 140, 2011. DOI: 10.1007/JHEP09(2011)140. arXiv: 1103.3438 [hep-ph].

[58] M. Cacciari, G. P. Salam, and G. Soyez, ≪The Anti-k(t) jet clustering algorithm≫, *JHEP*, vol. 0804, p. 063, 2008. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[59] D. Griffiths, *Introduction to Elementary Particles. Second, revised edition.* Wiley-VCH, 2008.

[60] G. Abecasis, *A C++ implementation of* `Simplex`, From the lecture notes to the course BIOSTAT 615/815 – Statistical Computing at the University of Michigan. [Online]. Available: {http://csg.sph.umich.edu/abecasis/class/2008/615.16.pdf}.

[61] C. Sanderson *et al.*, *Armadillo C++ linear algebra library*, `http://arma.sourceforge.net/`.