



PARTICLE PHYSICS

Master Thesis

---

**Enhancing detectability of  
tau-neutrino signatures  
using machine learning**

---

By: Daniel Alvestad

Supervisors: Jörn Kersten and Inga Strümke

June 1, 2018

---

## Abstract

In this thesis, the collider signatures of the scenario with a tau-sneutrino *next-to-lightest supersymmetric particle* (NLSP) at LHC are studied using machine learning. The parameter region of the *non-universal Higgs masses* model, where the tau-sneutrino is the NLSP, is studied to find a parameter point which satisfies constraints from recent experimental results. We look at the tri-lepton signature from two same sign hadronic taus and a muon ( $\tau_h^\pm \tau_h^\pm \mu^\mp$ ). This signature have its main contribution from the slepton and sneutrino pair production channel,  $pp \rightarrow \tilde{\ell} \tilde{\nu}_\ell$ . The aim is to enhance detectability of this signature by using a deep neural network trained on monte carlo simulated collider events. The best performing deep neural network is a multi class classifier, which is compared to other neural network architectures and a boosted decision tree. The required integrated luminosity for a  $5\sigma$  significance discovery using  $\sqrt{s} = 13$  TeV is found to be  $\mathcal{L}_{\text{int}}(5\sigma) = (3.4 \pm 0.7) \times 10^3 \text{ fb}^{-1}$ . We find that the multi class deep neural network performs better by a factor of 2.0 than the traditional optimized cuts.

## Acknowledgment

Firstly, I would like to express my gratitude to my supervisor Jörn Kersten and co-supervisor Inga Strümke for their guidance, motivation and inspirational discussions related to this thesis. A spacial thanks to Inga Strümke for inspiring me to pursue this masters project and for being available at any time. Thanks to Steffen Mæland and Nikolai Fomin for many helpful discussions concerning the cut analysis. I also thank my fellow physics students for five great years at the University of Bergen.

Last but not least, I am very grateful to my girlfriend, Stine, for proofreading the thesis and for all her love, support and encouragement.

# Contents

<a href="#">Abstract</a> . . . . .	i
<a href="#">Acknowledgment</a> . . . . .	iii
<a href="#">Acronyms</a> . . . . .	ix
<b><a href="#">1 Introduction</a></b>	<b>1</b>
<b><a href="#">2 Supersymmetry</a></b>	<b>5</b>
<a href="#">2.1 Standard model</a> . . . . .	5
<a href="#">2.1.1 Problems of the SM</a> . . . . .	7
<a href="#">2.2 Symmetries</a> . . . . .	8
<a href="#">2.2.1 Most general symmetry</a> . . . . .	9
<a href="#">2.2.2 Supercurrent</a> . . . . .	10
<a href="#">2.2.3 Weyl spinors</a> . . . . .	11
<a href="#">2.3 The Minimal Supersymmetric Standard Model</a> . . . . .	11
<a href="#">2.3.1 R-parity</a> . . . . .	13
<a href="#">2.4 Soft supersymmetry breaking in the MSSM</a> . . . . .	13
<a href="#">2.5 Electroweak symmetry breaking</a> . . . . .	14
<a href="#">2.5.1 Sfermion masses</a> . . . . .	16
<a href="#">2.5.2 Neutralinos and charginos</a> . . . . .	17
<a href="#">2.6 GUT scale models</a> . . . . .	17
<a href="#">2.6.1 Organization Principle</a> . . . . .	17
<a href="#">2.6.2 Underlying symmetry</a> . . . . .	18
<a href="#">2.6.3 The constrained MSSM</a> . . . . .	18
<b><a href="#">3 Supersymmetry breaking</a></b>	<b>21</b>
<a href="#">3.1 Spontaneous supersymmetry breaking</a> . . . . .	21
<a href="#">3.2 The goldstino</a> . . . . .	23
<a href="#">3.3 Supergravity</a> . . . . .	24
<a href="#">3.4 The Gravitino</a> . . . . .	25
<a href="#">3.4.1 The super-Higgs mechanism</a> . . . . .	26



3.4.2	The gravitino-goldstino equivalence	27
<b>4</b>	<b>The lightest supersymmetric particle</b>	<b>29</b>
4.1	The gravitino problem	29
4.2	The lightest observable supersymmetric particle	30
4.3	Sneutrino (NLSP) lifetime	30
4.3.1	Spin 3/2 formalism	31
4.3.2	Goldstino approximation	32
<b>5</b>	<b>NUHM parameter space</b>	<b>35</b>
5.1	Renormalization group equations	35
5.2	Parameter point	37
5.2.1	Spectrum generator	37
5.2.2	Reducing scanning parameters	38
5.2.3	Parameter scan	39
5.3	Mass spectrum	41
5.4	Tests against experiemtns	43
<b>6</b>	<b>Sneutrino signals in colliders</b>	<b>45</b>
6.1	Production cross-sections	45
6.2	Leptonic signatures	47
6.2.1	Sleptons	47
6.2.2	Chargino and neutralino	49
6.2.3	Squarks and gluino	49
6.3	Trilepton signal and backgrounds	50
6.3.1	Background	51
<b>7</b>	<b>Monte carlo simulation</b>	<b>53</b>
7.1	Structure of an event	53
7.1.1	Hard subprocess	54
7.1.2	Parton Shower	56
7.1.3	Hadronization	56
7.1.4	Jet algorithms	56
7.2	Event Generators	57
7.2.1	Validating Simulations	58
7.3	Analysis Scheme	59
7.3.1	Detector cuts	59
7.4	Signal selection efficiency	60

7.5 Simulations	61
7.6 Feature plots	62
7.6.1 Low-level features	62
7.6.2 High-level features	64
7.7 Detectable signal	64
7.8 Optimized cuts	67
<b>8 Machine Learning</b>	<b>71</b>
8.1 Neural Networks	72
8.2 Binary classifier setup	74
8.2.1 Input features	74
8.3 Pre-processing feature data	75
8.3.1 Training the network	75
8.4 Neural network as a cut	77
8.5 Class weights	77
8.6 Classifying individual signals	80
8.7 Performance	82
<b>9 A comparison of deep neural networks and optimized cuts</b>	<b>85</b>
9.1 Event simulation	85
9.2 Feature plots	86
9.3 Optimizing cuts	88
9.4 Neural network optimization	89
9.5 Machine learning vs. optimized cuts	90
<b>10 Summary and conclusion</b>	<b>93</b>
10.1 Overview of results	93
10.2 Outlook	95
<b>Appendices</b>	<b>97</b>
<b>A Notation and identities</b>	<b>99</b>
A.1 Notation and conventions	99
A.2 Identities	100
A.2.1 Gamma matrices identities	100
<b>B Lifetime Calculations</b>	<b>101</b>
B.1 2-body decay	101
B.2 Gravitino calculation	102

<b>B.3 Goldstino calculation</b> . . . . .	105
--	-----

# Acronyms

<b>ATLAS</b> A Toroidal LHC Experiment	<b>MSSM</b> Minimal Supersymmetric Standard Model
<b>BBN</b> Big bang nucleosynthesis	<b>mSUGRA</b> Minimal Supergravity
<b>BDT</b> Boosted decision tree	<b>NLO</b> Next-to-leading order
<b>CERN</b> European Organization of Nuclear Research	<b>NLSP</b> Next-to lightest supersymmetric particle
<b>CMSSM</b> Constrain Minimal Supersymmetric Standard Model	<b>NNLO</b> Next-to-next-to leading order
<b>CP</b> Charge conjugate and Parity	<b>NUHM</b> Non-Universal Higgs Masses
<b>DNN</b> Deep neural network	<b>PDF</b> Parton distribution function
<b>FCNC</b> Flavour changing neutral currents	<b>QCD</b> Quantum Chromodynamics
<b>GUT</b> Grand unification theory	<b>RGE</b> Renormalization Group Equations
<b>LHC</b> Large Hadron Collider	<b>SM</b> Standard Model
<b>LO</b> Leading order	<b>SUSY</b> Supersymmetry
<b>LSP</b> Lightest supersymmetric particle	<b>VEV</b> Vacuum expectation value



# Chapter 1

## Introduction

The Standard Model (SM) of particle physics is a theory that describes the fundamental constituencies of matter and their interactions. The model has proven extremely successful and is well-verified experimentally. With the discovery of the Higgs Boson, all the particles are discovered and no laboratory experiments have contradicted the theory as of this date. However, cosmological observations show that only 4% of the universe consists of "ordinary matter" made up by known particles described by the SM. This is one of the main reasons why the model is considered incomplete. The model also fails to incorporate quantum gravity and predict the cosmological constant. There are numerous proposed theories for physics beyond the Standard Model. One of the more popular is supersymmetry (SUSY).

Supersymmetry proposes a new set of particles, where each particle is a superpartner to one of the SM particles. The model resolves some of the problems related to the SM; most notably has a Dark Matter candidate and an explanation to the so called Hierarchy problem causing fine-tuning of the Higgs mass. The superpartners introduced in this theory have the same quantum numbers, except for the spin, which is  $\pm\frac{1}{2}$  of the SM particle. This introduces a problem, because these particles have not been observed in any experiments. SUSY must therefore be a broken symmetry to be consistent with current experimental observations.

The *Large Hadron Collider* (LHC) at CERN search for supersymmetry at the TeV scale. It operates with a center-of-mass energy of  $\sqrt{s} = 13$  TeV and has gathered in total about  $100 \text{ fb}^{-1}$  of integrated luminosity. Still without any hints of additional

structures. This is where the machine learning tools utilized in this thesis can make a contribution. In the last decades the improvements in computational power has increased and driven an innovation in data analysis. The applications of machine learning are from particle physics all the way to speak recognition and art projects. This thesis focuses on one application of deep neural network to the detection of supersymmetric particles.

Neural networks are good at recognizing patterns in data. This is ideal for data analysis tasks, such as separating signal and background events. The traditional way to do this is by optimized cuts. The cuts are applied to the observable feature data from each event such that they remove as much background as possible without losing too much signal. A neural network can explore the whole observable feature space at onces, whereas the process of constructing optimized cuts is to visualize the observable data in maximum three dimension and from this make cuts. Thus, it becomes hard to find efficient cuts if the number of dimensions in the feature space is large. Applying machine learning as a tool for enhancing detectability of signals, is studied in various scenarios, e.g. [1-4]. These studies show that the neural network can separate the signal from background, without using high-level observables like the invariant mass and angular separation between particles.

To interpret the data from collision experiments, we need theoretical models describing how the particles interact. Most of these studies assume that the lightest *neutralino* is the *lightest supersymmetric particle* (LSP). This is motivated by the neutralinos feasibility to be a *dark matter* candidate. Another less-explored region of the *minimal supersymmetric standard model* (MSSM) is studied in this thesis. In this case the tau-sneutrino ( $\tilde{\nu}_\tau$ ) is regarded as the lightest of the MSSM particles. Since the tau-sneutrino would have been discovered in direct detection [5, 6], it is not suitable as a dark matter candidate. However, the sneutrino can decay into a lighter *gravitino*, which is the superpartner of the graviton from quantum gravity. Since it has very weak couplings, the gravitino itself would not be seen directly and the tau-sneutrino appear as a stable particle in colliders because of the relative long lifetime [7].

Although the sneutrino behaves in the same way as the lightest neutralino, i.e. both escapes undetected through the detector and leaves a high missing energy, the mass spectrum, production rates and branching ratios are in general different. This leads to different characteristics and therefore a separate study is needed. The study of collider signatures from tau-sneutrino as the NLSP, within a *supergravity* model with

---

*non-universal Higgs masses* was performed by Figy et al. [8]. This is the starting-point of this thesis. We focus on the leptonic channels, in particular the  $\tau^\pm\tau^\pm\mu^\mp$  signature. Other similar studies within different supersymmetry breaking frameworks were performed by Covi and Kraml in [9] and Katz and Tweedie [10]. Since the above mentioned studies were conducted before the discovery of the Higgs boson at the LHC, the Higgs mass in their analysis is excluded.

This thesis begins with an introduction of supersymmetry and how the symmetry is broken, which leads to a short overview of the necessary concepts of supergravity used to calculate the sneutrino lifetime in chapter 4. Then the parameter space is scanned to obtain a parameter point matching current experimental limits followed by the relevant collider signatures in chapter 6. The parameter point is used to simulate collider events in chapter 7 which are used to train a deep neural network in classifying signal and background events in chapter 8. The deep neural network is then compared to the traditional cut and count method in chapter 9 before summarizing the obtained results.





# Chapter 2

## Supersymmetry

This chapter introduces supersymmetry and focuses on an extension of the Standard Model (SM). The chapter starts with a short description of the SM and why we need to extend this theory. There are numerous books and reviews written about the SM (see, for instance, [11, 12]) and therefore the many details are not included in this short overview. Then the supersymmetric extension of the standard model, the *Minimal Supersymmetric Standard Model* (MSSM), is introduced. Most of the supersymmetry theory in this chapter is taken from S.P. Martins ‘A supersymmetry Primer’ [13], which provides an excellent introduction to supersymmetry. For a summary of notation and convention used in this thesis, see chapter A.

### 2.1 Standard model

The *Standard Model* (SM) is a gauge theory describing the fundamental particles of the universe. It embodies our current understanding of the elementary particles and the interactions between them. The particles in the theory can be divided into fermions, gauge bosons and the Higgs particle. The fermions are the leptons and quarks, which all are spin- $\frac{1}{2}$  particles and have three generations of particles. The first generation of fermions are the particles that build up the matter around us; the electron, the up quark and the down quark. The second and third generations are irrelevant for the matter we see around us, but they are important in the study of the universe. These are the muon and tau for the fermions and charm, strange, top and bottom for the quarks. There are also some neutral leptons called neutrinos, which

are regarded as massless in the SM. Recent experiments show that these are in fact massive<sup>[14]</sup>, but this is usually not considered as a part of the SM<sup>[1]</sup>. The forces that make structures like atoms possible are the gauge bosons, also called force carriers. These are the W and Z bosons, the photon that mediate the *electroweak force* <sup>[16-18]</sup><sup>[2]</sup> and the gluons that mediate the *strong force* <sup>[19]</sup>. The gauge bosons are *vector bosons*, i.e. they have spin 1, whereas the Higgs boson is a scalar boson, i.e. it has spin 0. The Higgs field gives mass to the massive fermions and gauge bosons <sup>[20, 21]</sup>.

The system is described using the Lagrangian mechanism. The terms in the Lagrangian completely describes the particles and their motion, including the interactions. The SM is based on the gauge group  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ , which means that the Lagrangian must be invariant under this symmetry. Noether's theorem <sup>[22]</sup> relates these transformational symmetries to a conserved current and a corresponding conserved charge. The gauge groups are;

- $SU(3)_C$  is a non-abelian group that describes the strong interactions between the quarks and the gluons. There are eight conserved currents given by;

$$S_i^\mu = \frac{1}{2} \bar{\Psi}^f(x) \gamma^\mu \lambda_i \Psi^f(x), \quad (2.1.1)$$

where  $\Psi^f = (\psi_r^f \ \psi_g^f \ \psi_b^f)^T$ ,  $f = u, d, s, c, b, t$  (all quarks) and  $\lambda_i$  with  $i = 1, 2, \dots, 8$  is the eight generators of the group<sup>[3]</sup>. This means that any quark field has either red, green or blue as the color charge and the gluon fields has a color combination of these three colors e.g.  $r\bar{b}$  or  $b\bar{g}$  <sup>[19]</sup>. This theory is called *Quantum Chromodynamics*.

- $SU(2)_L$  is a non-abelian group describing the weak interactions between the left-handed fermions of the theory, ergo the subscript  $L$ . The conserved current is

$$J_i^\alpha(a) = \frac{1}{2} \bar{\Psi}_l^L(x) \gamma^\alpha \tau_i \Psi_l^L, \quad (2.1.2)$$

where  $\Psi_l^L(x) = (\psi_{\nu_e}^L \ \psi_l^L)^T$  and  $\tau_i$  is the group generator with  $i = 1, 2, 3$ . The spinor are all left-handed projections of the full Dirac spinor. The conserved charge is the weak isospin  $I_i^W$ .

<sup>1</sup>See <sup>[15]</sup> chapter 14 for an overview.

<sup>2</sup>This is the mass eigenstates, the gauge eigenstates are the W and B bosons, see section <sup>[2.5]</sup>

<sup>3</sup>This is the Gell-Mann matrices <sup>[11]</sup>.

- $U(1)_Y$  is the abelian group mediating by a single vector boson  $B^\mu$  field with the conserved hypercharge  $Y$ . The weak hypercharge current is

$$J_Y^\alpha(x) = -\frac{1}{2}\bar{\Psi}_l^L(x)\gamma^\alpha\Psi_l^L(x) - \bar{\psi}_l^R(x)\gamma^\alpha\psi_l^R(x), \quad (2.1.3)$$

which is the same ‘left-handed’ spinors as in eq. (2.1.2), the  $\psi_l^R$  is the ‘right-handed’ spinor fields. This group, together with the  $SU(2)_L$ , form the electroweak symmetry group;  $SU(2)_L \otimes U(1)_Y$ , and is called the *electroweak theory*.

The symmetry is broken spontaneously via the *Brout-Englert-Higgs mechanism* [20, 21] to the subgroup  $SU(3)_C \otimes U(1)_{\text{em}}$ . The  $U(1)_{\text{em}}$  group correspond to the theory of quantum electrodynamics (QED), which describes the interactions of charged particles with photons. The breaking of the symmetry gives rise to the masses of the massive particles in the theory. See [11] for a detailed review of the Brout-Englert-Higgs mechanism.

### 2.1.1 Problems of the SM

As stated in the introduction, the SM is considered as incomplete in the description of the universe. The main motivations for extending the SM are listed below;

- **Number of parameters:** The SM contains at least 19 free parameters that cannot be derived from any theory and have to be measured by experiments. This includes the masses of the particles, the weak mixing angle and the gauge coupling constants to name some.
- **Gravity:** It is unknown how to incorporate the gravitational force in the theory.
- **Dark matter:** There are no dark matter candidates in the SM.
- **Hierarchy problem** has been one of the main motivations since the beginning of the 1970. This is a technical problem concerning the loop corrections to the Higgs mass, which receives contributions from all particles that couples directly, or indirectly, to the Higgs field. The mass of the Higgs field squared can be expressed as

$$m_h^2 = (m_h^0)^2 + \Delta m_h^2, \quad (2.1.4)$$

where  $m_h^0$  is the tree-level mass of the Higgs and  $\Delta m_h^2$  represents the radiative corrections to the Higgs mass due to self-energy energy loop-diagrams. The

corrections from the fermions  $f$  that couples to the Higgs field by the Lagrangian term  $-\lambda_f H f \bar{f}$ , yields a correction to the Higgs mass by

$$\Delta m_h^2 = \frac{|\lambda_f|^2}{8\pi} \Lambda_{\text{UV}}^2 + \dots, \quad (2.1.5)$$

where  $\Lambda_{\text{UV}}$  is an ultraviolet momentum cutoff. This cutoff is regulating the loop integrals to avoid infinities in the theory and is interpreted as the scale at which new physics enters to regulate the high-energy behavior of the theory. If we assume new physics to enter at the Planck scale, where a theory of quantum gravity enters, the quantum corrections to  $m_h^2$  is roughly 30 orders of magnitude larger than the measured value at the  $\mathcal{O}(10^4)$  GeV scale. This means that the tree-level mass and the correction terms in eq. (2.1.4) needs to be very fine tuned in order to obtain the measured value. This is often referred to as the *fine-tuning problem*.

- **Why questions:** There are many unanswered questions in the SM, such as why there are three generations of fermions and why some of the particles mix to form mass eigenstates.
- **Grand unification:** The couplings of the SM appear to be changing depending on the energy. As we will see later in this chapter, these couplings never unify at a high scale [13]. It is presumed, however, that there is an energy where these couplings unify and that the SM is an effective field theory [23, 24]. This implies that the three interactions of the SM unite at a higher energy and act as a single interaction under a larger gauge group. These theories is referred to as *Grand Unified Theories* (GUT)<sup>4</sup>, where the couplings unite at  $\Lambda_{\text{GUT}} \sim 10^{16}$  GeV.

## 2.2 Symmetries

The new symmetry relates bosons and fermions. The supersymmetry Lagrangian is invariant under a transformation of the bosonic fields into fermionic fields and vice versa. This global transformation is schematically

$$\delta B = \bar{\epsilon} F, \quad \delta F = \epsilon \partial B, \quad (2.2.1)$$

---

<sup>4</sup>One of these GUTs is the  $SU(5)$  theory [25]

where  $B$  and  $F$  are bosonic and fermionic fields,  $\epsilon$  is the infinitesimal supersymmetry parameter carrying spinor indices and  $\partial$  stands for a space-time derivative. The SM (spin 1/2) fermions transform into (spin 0) scalars and the SM (spin 1) vector bosons transform into (spin 1/2) fermions. The Higgs scalar bosons (spin 0) also transform into (spin 1/2) fermions. The naming convention for the supersymmetric scalars is adding a prefix  $s-$  to the corresponding SM fermion name. For the supersymmetric fermions the suffix  $-ino$  is added to the SM boson name. For instance, the superpartner of the electron is the *selectron*, while the superpartner of the W boson is the *Wino*.

### 2.2.1 Most general symmetry

For a theory to obey special relativity, it must be invariant under the *Poincare group*. This is the group of all Lorentz boost and rotations, as well as translation. The generators of Lorentz boost and rotations,  $M_{\mu\nu}$ , and the generators of translation,  $P_\mu$ , satisfy the *Poincare algebra*:

$$[P_\mu, P_\nu] = 0, \quad (2.2.2)$$

$$[M_{\mu\nu}, P_\rho] = -i(g_{\mu\rho}P_\nu - g_{\nu\rho}P_\mu), \quad (2.2.3)$$

$$[M_{\mu\nu}, M_{\rho\sigma}] = -i(g_{\mu\rho}M_{\nu\sigma} - g_{\mu\sigma}M_{\nu\rho} - g_{\nu\rho}M_{\mu\sigma} + g_{\nu\sigma}M_{\mu\rho}), \quad (2.2.4)$$

which is a Lie algebra<sup>5</sup>. These generators commute with the generators of the internal gauge symmetries of the SM.

In the 1970's supersymmetry was introduced as the most general visible symmetry. This means that the most general symmetry is *Supersymmetry*  $\otimes$  *Poincare*. This was proven in two parts, first part by Coleman and Mandula and second part by Haag, Lopuszanski and Sohnius. First out was Coleman and Mandula [27], who asked; Is it possible to write down some physically meaningful theory where the charges  $T^a$  do not commute with the generators of the Poincare group ( $M_{\mu\nu}$  and  $P_\mu$ )? A physical meaningful theory, is a theory with the charges  $T^a$  obeying the algebra  $[T^a, T^b] = if^{abc}T^c$  [13], where  $f^{abc}$  is a structure constant. The answer to this question is *no*, and this seemed the end of the idea of building theories mixing space-time transformations and internal symmetries non-trivially. However, SUSY does circumvent the theorem in a very simple way; the supercharges ( $Q, Q^\dagger$ ) do not obey a set of commutation relations, but instead obey an algebra of anticommutation relations. This anticommutation relation

<sup>5</sup>See [26] for an overview of Lie Groups.

is schematically

$$\{Q, Q^\dagger\} \propto P^\mu, \{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0 \quad (2.2.5)$$

where  $P^\mu$  is the four-momentum (translation) generator of the Poincare group. If one extend the supersymmetry algebra to include  $N$  sets of supercharges, called extended supersymmetry, the *Haag-Lopuszanski-Sohnius Theorem* [28] states that this is the most general continuous symmetry extension of the Poincare group. Only the  $N=1$  version of supersymmetry is discussed in this thesis.

## 2.2.2 Supercurrent

The Noether's theorem [22] states that all continuous symmetry transformations imply the existence of a conserved current. This is also true for supersymmetry and it is called the *supercurrent*  $J^\mu$ , which is an anti-commuting four-vector carrying spinor indices. We can use *Noethers procedure* to find the *supercurrent* and its hermitian conjugate by

$$\bar{\epsilon} J^\mu(x) = \sum_X \delta X \frac{\delta \mathcal{L}}{\delta(\partial_\mu X)} - K^\mu, \quad (2.2.6)$$

where  $X$  can be any field in the theory and  $K^\mu$  is an object where the divergence follow the variation of the Lagrangian under supersymmetry transformation,  $\delta \mathcal{L} = \partial_\mu K^\mu$ . The general 4-component supercurrent for a global supersymmetry is given by [29]

$$\begin{aligned} J^\mu = & -\sqrt{2}[(D_\nu \phi_i)^* \gamma^\nu \gamma^\mu \psi_L^i + (D_\nu \phi_i) \gamma^\nu \gamma^\mu \psi_R^i - iW_i \gamma^\mu \psi_L^i - iW_i^* \gamma^\mu \psi_R^i] \\ & - \frac{1}{4} F_{\nu\rho}^a [\gamma^\nu, \gamma^\rho] \gamma^\mu \gamma_5 \lambda^a - ig(\phi_i^* T_{ij}^a \phi_j) \gamma^\mu \lambda^a, \end{aligned} \quad (2.2.7)$$

where  $\phi_i$  is the scalar fields,  $\psi_{L/R}^i$  is the left/right-handed fermion fields,  $\lambda^a$  is the gaugino fields, i.e. superpartners of the gauge fields,  $F_{\nu\rho}^a$  is the field strength tensor,  $D_\nu$  is the covariant derivative of  $\phi_i$  and  $T_{ij}^a$  is the corresponding group generator. The superpotential  $W$  is a holomorphic function that describes supersymmetric interactions and is defined in the most general form as

$$W = L^i \phi_i + \frac{1}{2} M^{ij} \phi_i \phi_j + \frac{1}{6} y^{ijk} \phi_i \phi_j \phi_k \quad (2.2.8)$$

where  $M^{ij}$  is the symmetric mass matrix for the fermion fields and  $y^{ijk}$  is the Yukawa coupling we know from the SM. The first term in eq. (2.2.8) requires a gauge singlet, which do not exist in the supersymmetric extension of the SM discussed in this thesis. The terms in the superpotential must contain a set of fields that transform in a representation that combine to form a singlet. This is to ensure renormalizability. In

terms of this superpotential, we have the definitions

$$W^i = \frac{\delta}{\delta\phi_i} W, \text{ and } W^{ij} = \frac{\delta^2}{\delta\phi_i \delta\phi_j} W. \quad (2.2.9)$$

### 2.2.3 Weyl spinors

Before we continue to the minimal extension of the SM, we need to develop some notation. When developing the supersymmetry theory the two-component Weyl-spinor notation is convenient. This is because each Dirac fermion has a left-handed and a right-handed part, which do not share all the same interactions. This means that in the Lagrangian, the four-component fermion needs to be projected to either a left-handed or right-handed fermion. The Dirac spinor is represented by two-component Weyl fermions as

$$\psi_D = \begin{pmatrix} \xi_\alpha \\ \chi^{\dagger\dot{\alpha}} \end{pmatrix}, \quad (2.2.10)$$

where  $\xi$  is a left-handed Weyl spinor and  $\chi^\dagger$  is a right-handed Weyl spinor. The indices  $\alpha = 1, 2$  and  $\dot{\alpha} = 1, 2$  are two distinct spinor indices. The hermitian conjugate of a left-handed Weyl fermion is a right-handed Weyl fermion;  $\xi_\alpha^\dagger \equiv (\xi_\alpha)^\dagger = (\xi^\dagger)_\alpha$ . The spinor indices of the Weyl-fermions are raised and lowered by the *antisymmetric* object  $\epsilon_{\alpha\beta} = (i\sigma_2)_{\alpha\beta}$ , i.e.

$$\xi^\alpha = \epsilon^{\alpha\beta} \xi_\beta, \quad \chi_{\dot{\alpha}} = \epsilon_{\dot{\alpha}\dot{\beta}} \chi^{\dagger\dot{\beta}}. \quad (2.2.11)$$

The two-component notation is used throughout this chapter.

## 2.3 The Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model (MSSM), is an extension of the SM and is minimal in the sense that it has the smallest field content consistent with the SM [13]. The particle content of the MSSM is listed in table 2.1. The tilde symbol ( $\tilde{\phantom{x}}$ ) denotes superpartners of the SM particles. Due to supersymmetry there is a supersymmetric particle for each SM particle. These pairs of fermions and bosons are represented in the irreducible representation of SUSY, called *supermultiplets* for short. One supermultiplet always consist of one fermion and one boson and the particles transform in identical gauge group representations. There are two types of supermultiplets in the MSSM; *chiral multiplets* that consist of a spin 0 scalar boson and a



spin 1/2 fermion and *vector multiplets* that consist of a spin 1/2 fermion and a spin 1 vector boson. The fermions, i.e. the lepton and quark families, resides in chiral supermultiplets with their spin 0 scalar boson partners. The Higgses and their spin 1/2 superpartners, the higgsinos, also resides in chiral supermultiplets. The particles building up the vector supermultiplets are the gauge bosons;  $W$  and  $B$  bosons, photon and the gluon with their superpartners called wino, bino, photino and gluino. The gauge singlets, in  $SU(2)_L$  and  $SU(3)_C$ , are denoted by  $\mathbf{1}$ , whereas in the case of  $U(1)_Y$  singlets carry hypercharge  $Y = 0$ . The normalization of hypercharges are such that the electric charge  $Q$  is given by  $Q = I_3^W + Y/2$ , where  $I_3^W$  denotes the weak isospin. The upper entry in a  $SU(2)$  doublet is  $I_3^W = +1/2$  and the lower is  $I_3^W = -1/2$ .

Super-multiplets	Bosons	Fermions	$SU(3)_C, SU(2)_L, U(1)_Y$
Gluon/gluino	$g$	$\tilde{g}$	$(\mathbf{8}, \mathbf{1}, 0)$
Gauge/gaugino	$W$	$\tilde{W}^\pm, \tilde{W}^0$	$(\mathbf{1}, \mathbf{3}, 0)$
	$B$	$\tilde{B}^0$	$(\mathbf{1}, \mathbf{1}, 0)$
Slepton/lepton	$L$	$(\tilde{\nu}, \tilde{e}_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	$\bar{e}$	$\tilde{e}_R^*$	$(\mathbf{1}, \mathbf{1}, 1)$
Squark/quark	$Q$	$(\tilde{u}_L, \tilde{d}_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	$\bar{u}$	$\tilde{u}_R^*$	$(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	$\bar{d}$	$\tilde{d}_R^*$	$(\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
Higgs/higgsinos	$H_u$	$(H_u^+, H_u^0)$	$(\mathbf{1}, \mathbf{2}, \frac{1}{2})$
	$H_d$	$(H_d^0, H_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$

Table 2.1: Supermultiplets in the MSSM and their irreducible representations of the  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ .

The superpotential for the MSSM is

$$W_{MSSM} = \tilde{u}^* \mathbf{y}_d \tilde{Q} \cdot H_u - \tilde{d}^* \mathbf{y}_d \tilde{Q} \cdot H_d - \tilde{e}^* \mathbf{y}_e \tilde{L} \cdot H_d + \mu H_u \cdot H_d. \quad (2.3.1)$$

The doublet structures of the superpotential, i.e. terms combining two  $SU(2)_L$  doublets, is tied together by  $\tilde{Q} \cdot H_u = \epsilon^{\alpha\beta} \tilde{Q}_\alpha H_{u\beta}$ . Furthermore,  $\tilde{u}^* \mathbf{y}_d \tilde{Q}$  is a matrix multiplication in family space, i.e.  $\tilde{u}^* \mathbf{y}_d \tilde{Q} = \tilde{u}_i^* \mathbf{y}_d^{ij} \tilde{Q}_{j\alpha}$ , where  $i, j = 1, 2, 3$  is the family index and  $a = 1, 2, 3$  is the color index. The first term is written out as  $\tilde{u}_i^* \mathbf{y}_d^{ij} \tilde{Q}_{j\alpha} H_{u\beta} \epsilon^{\alpha\beta}$ .

### 2.3.1 R-parity

In the Standard Model there are no renormalizable terms that violate baryon number ( $B$ ) and lepton number ( $L$ ). This is backed up by experiments, where no violations have been observed. Thus, we cannot allow any  $B$  or  $L$  violating terms in the MSSM. The superpotential, eq. (2.3.1), is the minimal in the sense that it is sufficient to produce a phenomenologically viable model. However, there are other gauge-invariant and holomorphic terms one can write that violate  $B$  or  $L$ . The additional terms  $\tilde{L} \cdot H_u$ ,  $\tilde{L} \cdot \tilde{L} \tilde{e}$  and  $\tilde{L} \cdot \tilde{Q} \tilde{d}$  violate the lepton number and the term  $\tilde{u} \tilde{d} \tilde{d}$  violate the baryon number. The most obvious constraint on these terms is related to proton decay, which would decay rapidly, i.e. fraction of a second, if these terms are allowed [13].

To solve this in the MSSM, without assuming  $B$  and  $L$  conservation, we introduce a new symmetry. This new symmetry is called ‘R-parity’ and is defined for each particle as

$$P_R = (-1)^{3(B-L)+2s} \quad (2.3.2)$$

where  $s$  is the spin of the particle. All SM particles and the additional Higgs bosons of the MSSM have even R-parity ( $P_R = +1$ ), while all the supersymmetric particles, i.e. squarks, sleptons, gauginos, and higgsinos, have odd R-parity ( $P_R = -1$ ). Any interactions must conserve R-parity and must therefore contain zero or an even number of supersymmetric particles. This has three important consequences

- The lightest supersymmetric particle (LSP) must be stable.
- All supersymmetric particles other than the LSP must eventually decay to the LSP.
- In colliders, supersymmetric particles can only be produced in even numbers.

The first two consequences mentioned above makes the LSP a good dark matter candidate.

## 2.4 Soft supersymmetry breaking in the MSSM

In each supermultiplet the mass of the two particles must be the same, which is a phenomenological problem because no supersymmetric particles have been observed.

This means that the symmetry is not exact. We expect supersymmetry to be broken spontaneously and there are many proposed models of how this spontaneous breaking occur<sup>6</sup>. What these models have in common are the resulting symmetry breaking terms, which should break the symmetry *soft*, i.e. naturally maintain the hierarchy between the electroweak scale and Planck mass scale. The *soft terms* have couplings with mass dimension one or higher. The possible soft breaking terms in the MSSM are

$$\begin{aligned}
\mathcal{L}_{soft}^{MSSM} = & -\frac{1}{2} \left( M_3 \widetilde{g}\widetilde{g} + M_2 \widetilde{W}\widetilde{W} + M_1 \widetilde{B}\widetilde{B} + h.c \right) \\
& - \left( \widetilde{u} \mathbf{a}_u \widetilde{Q} \cdot H_u - \widetilde{d} \mathbf{a}_d \widetilde{Q} \cdot H_d - \widetilde{e} \mathbf{a}_e \widetilde{L} \cdot H_d + h.c \right) \\
& - \widetilde{Q}^\dagger \mathbf{m}_Q^2 \widetilde{Q} - \widetilde{L}^\dagger \mathbf{m}_L^2 \widetilde{L} - \widetilde{u} \mathbf{m}_u^2 \widetilde{u}^\dagger - \widetilde{d} \mathbf{m}_d^2 \widetilde{d}^\dagger - \widetilde{e} \mathbf{m}_e^2 \widetilde{e}^\dagger \\
& - m_{H_u}^2 H_u^* H_u - m_{H_d}^2 H_d^* H_d - (b H_u \cdot H_d + b^\dagger H_d^\dagger \cdot H_u^\dagger),
\end{aligned} \tag{2.4.1}$$

where  $M_1, M_2, M_3$  are the bino, wino and gluino mass terms. The second line in eq. (2.4.1) contains the trilinear terms, coupling the sfermions to the Higgs with the trilinear couplings  $\mathbf{a}_u, \mathbf{a}_d$  and  $\mathbf{a}_e$ , which are complex  $3 \times 3$  matrices in family space. The third line in eq. (2.4.1) consists of squark and slepton mass terms, where each  $\mathbf{m}_Q^2, \mathbf{m}_L^2, \mathbf{m}_u^2, \mathbf{m}_d^2$  and  $\mathbf{m}_e^2$  represents a  $3 \times 3$  matrix in family space. The last line in eq. (2.4.1) describes the supersymmetry breaking contribution to the Higgs potential.

These soft supersymmetry breaking terms introduce new parameters and in the MSSM there are in total 105 new parameters. This is in addition to the SM particles and most of the new parameters are from the soft supersymmetry breaking Lagrangian.

## 2.5 Electroweak symmetry breaking

In the SM the vector bosons and the fermions acquired mass through the electroweak symmetry breaking, which break down to electromagnetism,

$$SU(2)_L \otimes U(1)_Y \rightarrow U(1)_{em}. \tag{2.5.1}$$

Since there are two complex Higgs doublets in the MSSM, the electroweak symmetry breaking is slightly more complicated. In the MSSM the electroweak symmetry breaking is achieved after supersymmetry breaking when one of the Higgs masses evolve

<sup>6</sup>This is discussed in more detail in the next chapter.

to a negative value. This breaks the symmetry because the neutral Higgs fields,  $H_u^0$  and  $H_d^0$ , acquire non-zero *vacuum expectation values* (VEV),  $\langle H_u^0 \rangle = v_d$  and  $\langle H_d^0 \rangle = v_u$ . These VEVs need to be compatible with experiments, which imply that they must follow

$$v_u^2 + v_d^2 = v^2 = 4m_Z^2/(g^2 + g'^2), \quad (2.5.2)$$

where  $m_Z$  is the Z-boson mass. The ratio of the VEVs is written as  $\tan \beta = v_u/v_d$ , indicating that they can be written in terms of  $v$  as  $v_u = v \sin \beta$  and  $v_d = v \cos \beta$ .

The two Higgs fields in the MSSM are two complex  $SU(2)_L$  doublets. This means that they have 8 degrees of freedom combined. As in the SM, three of these are the *Nambu-Goldstone* bosons [30],  $G^0$  and  $G^\pm$ , which become the longitudinal modes of the  $Z^0$  and  $W^\pm$  massive vector bosons. The mass eigenstates of the gauge bosons are given by

$$\begin{pmatrix} A_\mu \\ Z_\mu^0 \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^0 \end{pmatrix}, \quad (2.5.3)$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \pm W_\mu^2) \quad (2.5.4)$$

where  $\theta_W$  is the Weinberg mixing angle and  $B_\mu$ ,  $W_\mu^0$  and  $W_\mu^{(1,2)}$  are the gauge eigenstates.

The remaining five degrees of freedom consist of two CP-even neutral scalars  $h^0$  and  $H^0$ , one CP-odd neutral scalar  $A^0$  and a charged scalar  $H^+$  with its charged conjugate scalar  $H^-$ . The lightest Higgs is the  $h^0$  and represents the most likely candidate to be the discovered Higgs boson, with  $m_h = 125.09$  GeV [31]. The gauge-eigenstate Higgs fields can be expressed in terms of the mass eigenstates as

$$\begin{pmatrix} H_u^0 \\ H_d^0 \end{pmatrix} = \begin{pmatrix} v_u \\ v_d \end{pmatrix} + \frac{1}{\sqrt{2}}R_\alpha \begin{pmatrix} h^0 \\ H^0 \end{pmatrix} + \frac{i}{\sqrt{2}}R_{\beta_0} \begin{pmatrix} G^0 \\ A^0 \end{pmatrix}, \quad (2.5.5)$$

$$\begin{pmatrix} H_u^+ \\ H_d^{-*} \end{pmatrix} = R_{\beta_\pm} \begin{pmatrix} G^+ \\ h^+ \end{pmatrix}, \quad (2.5.6)$$

$$(2.5.7)$$

where the orthogonal rotational matrices  $R_\alpha$ ,  $R_{\beta_0}$  and  $R_{\beta_\pm}$  contain the mixing angles  $\alpha$ ,  $\beta_0$  and  $\beta_\pm$  respectively. At the minimum in the scalar potential, the mixing angles

are  $\beta_0 = \beta_{\pm} = \beta$  and the Higgs masses are given by

$$m_{A^0}^2 = 2b/\sin 2\beta = 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2, \quad (2.5.8)$$

$$m_{h^0, H^0}^2 = \frac{1}{2} \left( m_{A^0}^2 + m_Z^2 \mp \sqrt{(m_{A^0}^2 - m_Z^2)^2 \pm 4m_Z^2 m_{A^0}^2 \sin^2(2\beta)} \right), \quad (2.5.9)$$

$$m_H^{\pm} = m_{A^0}^2 + m_W^2 \quad (2.5.10)$$

relating them to the soft symmetry breaking parameters.

### 2.5.1 Sfermion masses

The contribution to the mass of the supersymmetry particles is mainly from the soft SUSY breaking Lagrangian terms, eq. (2.4.1). This is as expected since without the SUSY breaking, the masses of the particles within a supermultiplet would be the same. The first and second generation sleptons and squarks masses have mainly contribution from the  $\tilde{\phi}^\dagger \mathbf{m}_\phi^2 \tilde{\phi}$  terms in  $\mathcal{L}_{\text{soft}}^{\text{MSSM}}$  and the hyperfine splitting  $\Delta_\phi$ . The latter one is given by

$$\Delta_\phi = (T_{3\phi} - Q_\phi \sin \theta_W) \cos(2\beta) m_Z^2 \quad (2.5.11)$$

where  $T_{3\phi}$  and  $Q_\phi$  are the weak isospin and the electric charge, respectively.

The third generation leptons and squarks mix the right- and left-handed particles, because of the larger Yukawa and trilinear couplings. They have the same contribution as the first and second generation, in addition to contribution from the Yukawa and trilinear mixing terms. The gauge-eigenstates for the staus are  $\tilde{\tau}_L$  and  $\tilde{\tau}_R^*$ , and the mass terms are

$$\mathcal{L}_{\tilde{\tau} \text{ mass}} = \begin{pmatrix} \tilde{\tau}_L^* & \tilde{\tau}_R^* \end{pmatrix} \mathbf{m}_{\tilde{\tau}}^2 \begin{pmatrix} \tilde{\tau}_L \\ \tilde{\tau}_R \end{pmatrix} \quad (2.5.12)$$

where

$$\mathbf{m}_{\tilde{\tau}}^2 = \begin{pmatrix} m_{L_3}^2 + \Delta_{\tilde{e}_L} & v(a_\tau^* \cos \beta - \mu y_\tau \sin \beta) \\ v(a_\tau \cos \beta - \mu^* y_\tau \sin \beta) & m_{\tilde{e}_3}^2 + \Delta_{\tilde{e}_3} \end{pmatrix} \quad (2.5.13)$$

which after diagonalizing gives the mass-eigenstates;  $\tau_1$  and  $\tau_2$ . We do the same to get the mass-eigenstates of the the stops,  $\tilde{t}_1, \tilde{t}_2$ , and the sbottoms,  $\tilde{b}_1, \tilde{b}_2$ .

## 2.5.2 Neutralinos and charginos

A consequence of the electroweak symmetry breaking is that the higgsinos and electroweak gauginos mix with each other. This forms the mass eigenstates called *neutralinos* for the neutral and *charginos* for the charged. The neutralinos consist of the neutral higgsinos,  $\tilde{H}_u^0$  and  $\tilde{H}_d^0$ , and the neutral gauginos,  $\tilde{B}$  and  $\tilde{W}^0$ , combined to form the mass eigenstate  $\tilde{\chi}_i^0$ , where  $i = 1, 2, 3, 4$ . The neutralino mass terms in the Lagrangian, in terms of the gauge-eigenstate basis  $\psi^0 = (\tilde{B}, \tilde{W}^0, \tilde{H}_d^0, \tilde{H}_u^0)$ , are

$$\mathcal{L}_{\text{neutralino mass}} = -\frac{1}{2}(\psi^0)^T \mathbf{M}_{\tilde{\chi}^0} \psi^0 + \text{c.c.}, \quad (2.5.14)$$

where  $\mathbf{M}_{\tilde{\chi}^0}$  is the mass matrix. To get the mass eigenstates this matrix must be diagonalized. The mass eigenstates are given by

$$\tilde{\chi}_i^0 = \mathbf{N}_{ij} \psi_j^0, \quad (2.5.15)$$

where  $\mathbf{N}$  is a unitary matrix diagonalizing  $\mathbf{M}_{\tilde{\chi}^0}$ , such that

$$\mathbf{N}^* \mathbf{M}_{\tilde{\chi}^0} \mathbf{N}^{-1} = \text{diag}(m_{\tilde{\chi}_1^0}, m_{\tilde{\chi}_2^0}, m_{\tilde{\chi}_3^0}, m_{\tilde{\chi}_4^0}). \quad (2.5.16)$$

For the charginos we get the mass term  $-\frac{1}{2}(\psi^-)^T \mathbf{M}_{\tilde{\chi}^\pm} \psi^- + \text{c.c.}$ , where the gauge-eigenstate consists of  $\psi^+ = (\tilde{W}^+, \tilde{H}_u^+)$  and  $\psi^- = (\tilde{W}^-, \tilde{H}_d^-)$ . The mixing matrices are diagonalized by the  $2 \times 2$  unitary matrices  $\mathbf{U}$  and  $\mathbf{V}$ , to create the mass eigenstates

$$\begin{pmatrix} \tilde{\chi}_1^+ \\ \tilde{\chi}_2^+ \end{pmatrix} = \mathbf{V} \begin{pmatrix} \tilde{W}^+ \\ \tilde{H}_u^+ \end{pmatrix}, \quad \begin{pmatrix} \tilde{\chi}_1^- \\ \tilde{\chi}_2^- \end{pmatrix} = \mathbf{U} \begin{pmatrix} \tilde{W}^- \\ \tilde{H}_d^- \end{pmatrix}. \quad (2.5.17)$$

## 2.6 GUT scale models

### 2.6.1 Organization Principle

The 105 masses, phases and mixing angles that cannot be redefined by known constants in the MSSM, must be measured and is independent of each other. By using already existing experimental evidence from the Standard Model, the number of unknown parameters decrease [13]. It seems like some organization principles are governing the soft supersymmetry breaking Lagrangian eq. (2.4.1). This organisation principle is

necessary because without it there would be large CP-violation processes.

One example of flavor mixing of leptons can be shown by the process  $\mu \rightarrow e\gamma$ , which is suppressed by experimental evidence. The contributions to this process arise from terms in the soft supersymmetry breaking Lagrangian (eq. (2.4.1)), e.g.  $\tilde{e} \mathbf{m}_{\tilde{e}}^2 \tilde{e}^\dagger$  and  $\tilde{e} \mathbf{a}_e \tilde{L} \cdot H_d$ . In an idealized limit  $\mathbf{m}_{\tilde{e}}^2$  and  $\mathbf{a}_e$  are assumed to be diagonal. The process will not get any more contributions from eq. (2.4.1) and we have the first pattern of the organization principle;  $\mathbf{m}_{\tilde{e}}^2 = m_{\tilde{e}}^2 \mathbf{1}$ ,  $\mathbf{a}_e = A_{e0} \mathbf{y}_e$ , where  $\mathbf{y}_e$  is the Yukawa matrix for leptons. The reason why  $\mathbf{a}_e$  should be proportional to the Yukawa matrix, is to achieve equal organisation of all parameters and to ensure quark mixing as in the SM. There are similar examples with all the mass terms from eq. (2.4.1), which avoid flavor mixing and CP-violating processes.

## 2.6.2 Underlying symmetry

The organization principle could indicate an underlying simplicity, or symmetry of the Lagrangian at a high energy scale. At ordinary energies where the masses, cross-sections and decay-rates are measured, the Lagrangian calculation of these would involve large logarithms. These can be resummed away using the renormalization group equations (RGEs), as usual in quantum field theory. To make this work we treat all couplings and masses appearing in the Lagrangian as running parameters. In section 5.1 we discuss the full renormalization group equations.

The running of the coupling constants in supersymmetry is one of the main motivations for the theory. This is, as mentioned in the beginning, because the coupling constants  $g'$ ,  $g$  and  $g_s$  in the SM do not unify at a high energy scale. In supersymmetry, the extra loop corrections from the MSSM field content alters it to unify at a scale  $M_U \sim 1.5 \times 10^{16}$  GeV. This is shown in fig. 2.1, where we clearly see that the dashed SM lines do not unify at a higher scale as the solid MSSM lines do.

## 2.6.3 The constrained MSSM

Motivated by the wish for unification at  $M_U$ , we introduce the much studied *constrained MSSM* (CMSSM), also known as *minimal supergravity* (mSUGRA). This model assumes a minimal form of the parameters at the GUT scale. In terms of the the soft

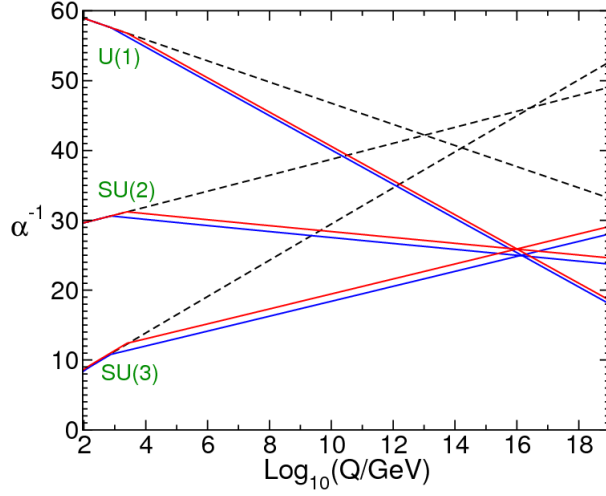


Figure 2.1: RGE evolution of the inverse coupling  $\alpha_i^{-1} = \frac{4\pi}{g_i^2}$  at two-loop order. The dashed line is the SM and the solid lines are the MSSM. The figure is taken from [13].

breaking Lagrangian parameters in eq. (2.4.1), they are

$$M_3 = M_2 = M_1 = m_{1/2}, \quad (2.6.1)$$

$$\mathbf{m}_Q^2 = \mathbf{m}_{\bar{u}}^2 = \mathbf{m}_d^2 = \mathbf{m}_L^2 = \mathbf{m}_e^2 = m_0^2 \mathbf{1}, \quad m_{H_u}^2 = m_{H_d}^2 = m_0^2, \quad (2.6.2)$$

$$\mathbf{a}_u = A_0 \mathbf{y}_u, \quad \mathbf{a}_d = A_0 \mathbf{y}_d, \quad \mathbf{a}_e = A_0 \mathbf{y}_e, \quad (2.6.3)$$

$$b = B_0 \mu. \quad (2.6.4)$$

The  $b$  parameter is replaced by  $\tan \beta = \frac{v_u}{v_d}$ , where  $v_u$  and  $v_d$  are the VEVs acquired by the two neutral Higgs fields when the electroweak symmetry is broken. The constraints from the electroweak symmetry breaking allow us to exchange both  $|\mu|$  and  $b$  for  $\tan \beta$  and  $\text{sign}(\mu)$ . Thus, the free parameters in the theory are

$$m_0, m_{1/2}, A_0, \tan \beta \text{ and } \text{sign}(\mu). \quad (2.6.5)$$

In chapter 5 the *non-universal Higgs masses* (NUHM) model is introduced, which is the model used for the analysis. This model have the same constraints as the CMSSM except for the soft Higgs mass parameters,  $m_{H_u}^2$  and  $m_{H_d}^2$ , i.e. the condition

$$m_{H_u}^2 = m_{H_d}^2 = m_0 \quad (2.6.6)$$

do not need to be satisfied. Thus, the free parameters is the same as in the CMSSM including the  $m_{H_u}^2$  and  $m_{H_d}^2$ .





# Chapter 3

## Supersymmetry breaking

We know from observations that supersymmetry cannot be an exact symmetry of nature. Otherwise, there would have to exist a spin zero particle with exactly the mass and charge of an electron and such a particle could not have evaded experimental detection. The only way out of this is if the supersymmetric particles are heavier than the already known particles. Thus, supersymmetry must be a broken symmetry.

There are two possible scenarios of how supersymmetry can be broken: explicitly or spontaneously. If supersymmetry is broken explicitly one adds terms to the supersymmetric Lagrangian that violate the symmetry, whereas the Lagrangian remains supersymmetric when supersymmetry is spontaneously broken. In the latter scenario the vacuum state of the theory is not supersymmetric [30] and is similar to spontaneous symmetry breaking in the *Brout, Englert, Higgs mechanism* [20, 21]. This second scenario is the preferred scenario since the theory is built to be invariant under supersymmetry transformation.

### 3.1 Spontaneous supersymmetry breaking

The condition for a spontaneously broken symmetry is that the vacuum state  $|0\rangle$  is not invariant under the symmetry transformation [30]. In the vacuum state, no real fields are present. If the supersymmetry generators  $Q_\alpha$  and  $Q_\alpha^\dagger$  operate on the vacuum state, it should be annihilated. It follows that supersymmetry is spontaneously broken

if

$$Q_\alpha |0\rangle \neq 0, \quad Q_\alpha^\dagger |0\rangle \neq 0. \quad (3.1.1)$$

In global supersymmetry, the Hamilton operator  $H$  is related to the supersymmetry generators through the algebra  $\{Q_\alpha, Q_\alpha^\dagger\} = -2\sigma_{\alpha\dot{\alpha}}P_\mu$ :

$$H = P^0 = \frac{1}{4}(Q_1 Q_1^\dagger + Q_1^\dagger Q_1 + Q_2 Q_2^\dagger + Q_2^\dagger Q_2). \quad (3.1.2)$$

This leads to the criterion for supersymmetry breaking  $\langle 0|H|0\rangle > 0$ . The only fields that may have nonzero VEV without breaking Lorentz invariance are scalar fields. This reduces the criterion to  $\langle 0|V|0\rangle > 0$ . The condition for spontaneous supersymmetry breaking, given a general global SUSY theory with the scalar potential  $V(\phi, \phi^*) = F_i^* F_i + \frac{1}{2} \sum_a D^a D^a$ , is that one of the auxiliary fields  $F_i$  or  $D^a$  obtains a nonzero VEV. The auxiliary fields  $F_i$  and  $D^a$  are a complex scalar field and a real bosonic field. The equation of motion of the auxiliary fields are  $F_i = W_i^*$  and  $D^a = -g(\phi^* T^a \phi)$ <sup>1</sup>.

There are no candidate gauge or scalar singlet to acquire a VEV in the MSSM and therefore the model need to be extended [13]. It is usually assumed that the supersymmetry breaking happens in a so-called *hidden sector* and then mediates to the *visible sector* of known fields by some mediator [32, 33]. Some of the more popular frameworks for supersymmetry breaking are *Planck-scale-mediated supersymmetry breaking* [34], *Gauge-mediated supersymmetry breaking* [35] and *Gaugino-mediated supersymmetry breaking* [36].

In Planck-scale-mediation the mediation of supersymmetry breaking from the hidden sector to the visible sector occurs only through gravitational interactions. The couplings in this scenario are inversely proportional to the Planck mass  $M_p = 2.435 \times 10^{18}$  GeV [34]. If supersymmetry is broken in the hidden sector by a VEV  $\langle F \rangle$ , the masses of the supersymmetric particles in the visible sector are derived to [13]

$$m_{soft} \sim \frac{\langle F \rangle}{M_p}. \quad (3.1.3)$$

The various supersymmetry breaking models produces a supersymmetry breaking Lagrangian, which includes terms on the same form as the soft breaking Lagrangian, eq. (2.4.1). Thus, the soft parameter can be related to the parameter of the model by comparing terms.

---

<sup>1</sup>For details, see [13].

The notation used in the rest of this chapter and the next is the four-component Dirac notation, where a small summary is presented in chapter [A](#)

## 3.2 The goldstino

The goldstino<sup>2</sup> is a massless Nambu-Goldstone mode that always appears in the spontaneous breaking of a continuous global symmetry. This is in complete analogy to the Goldstone bosons for the case of the spontaneous breakdown of  $SU(2)_L \times U(1)_Y$  in the SM [\[30\]](#). The generators of supersymmetry are fermionic and the goldstino must be a massless neutral fermion [\[37\]](#).

The goldstino can be connected to the spontaneous supersymmetry breaking criteria by looking at the fermionic mass matrix. This matrix transforms under a fermionic eigenfunction to a eigenvalue corresponding to the mass of the fermion mode. In the goldstino case, there exists a fermionic eigenfunction which annihilates the fermionic mass matrix. This mode is proportional to the VEV of the auxiliary fields and only appears when supersymmetry is broken, i.e.  $\langle F_i \rangle \neq 0$  or  $\langle D^a \rangle \neq 0$ .

To construct the goldstino Lagrangian, we need to revisit the supercurrent after the spontaneous supersymmetry breaking. The goldstino appears in the first four terms in the general supercurrent eq. [\(2.2.7\)](#). The new supercurrent including the goldstino is

$$J^\mu = i\sqrt{2}\langle F \rangle \gamma^\mu \tilde{G} + \sqrt{2}\gamma^\nu \gamma^\mu \tilde{G} D_\nu \phi^* + J_{\text{vis}}^\mu, \quad (3.2.1)$$

where  $\tilde{G}$  is the gravitino field and  $J_{\text{vis}}^\mu$  the supercurrent for all particles in the visible sector. The supercurrent conservation equation is given by

$$\partial_\mu J^\mu = i\sqrt{2}\langle F \rangle \gamma^\mu \partial_\mu \tilde{G} + \sqrt{2}\gamma^\nu \gamma^\mu \partial_\mu (\tilde{G} D_\nu \phi^*) + \partial_\mu J_{\text{vis}}^\mu = 0. \quad (3.2.2)$$

The equation of motion for the goldstino can be derived from eq. [\(3.2.2\)](#). This can be substituted into the kinetic goldstino term, which is the standard kinetic term for fermions. By splitting the term in one half part and substitute in the equation of motion, the Lagrangian can be expressed as

$$\mathcal{L}_{\tilde{G}} = \frac{i}{2} \tilde{G} \gamma^\mu \partial_\mu \tilde{G} - \frac{1}{2\langle F \rangle} \tilde{G} \gamma^\nu \gamma^\mu \partial_\mu (\tilde{G} D_\nu \phi^*) - \frac{1}{2\sqrt{2}\langle F \rangle} \tilde{G} \partial_\mu J_{\text{vis}}^\mu. \quad (3.2.3)$$

<sup>2</sup>Contrary to what the name suggests, the goldstino is not the superpartner of a goldstone boson.

The last term describes the interaction of the goldstino with all other fermion-boson pairs in the visible sector. It contains two derivatives, which turns out to always produce a kinematic factor proportional to the squared-mass difference of the superpartners when they are on-shell<sup>3</sup> [38]. For the goldstino-scalar-fermion interaction, the factor is the squared mass difference between the scalar and the fermion;  $m_\phi^2 - m_\psi^2$ .

The interaction Lagrangian of a single goldstino coupling to the visible sector particles becomes

$$\begin{aligned} \mathcal{L}_{\tilde{G},int} = & \frac{i(m_\phi^2 - m_\psi^2)}{\langle F \rangle} (\bar{\psi} P_L \tilde{G} \phi_L - \bar{\psi} P_R \tilde{G} \phi_R + \bar{\tilde{G}} P_R \psi \phi_R^* - \bar{\tilde{G}} P_L \psi \phi_L^*) \\ & + \frac{im_{\lambda^a}}{4\sqrt{2}\langle F \rangle} \bar{\tilde{G}} [\gamma^\mu, \gamma^\nu] \lambda^a F_{\mu\nu} - \frac{g_s m_{\lambda^a}}{\sqrt{2}\langle F \rangle} \bar{\tilde{G}} \lambda^a \phi_i^* T_{ij}^a \phi_j, \end{aligned} \quad (3.2.4)$$

where  $P_{L/R} = \frac{1}{2}(1 \pm \gamma_5)$  are the chiral projection operators,  $g_s$  the gauge couplings and  $T_{ij}^a$  the generators of the gauge group.

To this point we have looked at supersymmetry as a global symmetry, which means that the goldstino is a physical particle. As we will see in the next section, the goldstino is not a physical particle when we make the theory local.

### 3.3 Supergravity

The goldstino in the previous section is a would be Nambu Goldstone particle. Since the goldstino must have a massless supersymmetric partner, which is not observed, it cannot be a physical particle. The way we got rid of the Goldstone boson in the SM was to make the symmetry transformation of the broken symmetry local [39]. In such a case the supersymmetry transformation parameter is allowed to be space and time dependent. This means that the transformation parameter  $\epsilon_a$  is no longer a constant i.e.  $\epsilon_a \rightarrow \epsilon_a(x)$  (see eq. (2.2.1)). A local supersymmetry transformation is known as *supergravity* since it necessarily brings a massless spin 2 field, known as the Graviton in general relativity, into the theory [40, 41].

A gauge field has to be introduced to make a theory invariant under a local  $U(1)$  symmetry. This gauge field couples to the fields in the Lagrangian, and transform under the local symmetry in a way that restores the invariance. In local supersymmetry,

<sup>3</sup>On-shell means that the physical system satisfies the classical equation of motion.

this gauge field is called a gravitino  $\Psi_\mu$ . It is introduced through the Rarita-Schwinger equation for spin 3/2 particles [42] and the interaction with the supercurrent. The term that cancels the extra terms in the variation of the Lagrangian ( $\delta\mathcal{L} \neq 0$ ) is the gravitino interaction term. The interaction Lagrangian of the gravitino to the supercurrent is given by [43, 44]

$$\mathcal{L}_{3/2,int} = -\frac{1}{2M_p}\bar{\Psi}_\mu J^\mu. \quad (3.3.1)$$

This can be derived either from supersymmetry becoming a local theory (see [45]) or using the Rarita-Schwinger equation for a massless spin 3/2 particle [42]

$$\mathcal{L}_{RS} = -\frac{1}{2}e^{\mu\nu\rho\sigma}\bar{\Psi}_\mu\gamma_5\gamma_\nu\partial_\rho\Psi_\sigma, \quad (3.3.2)$$

where the gravitino field must be invariant under the gauge SUSY transformation with parameter  $\epsilon$ :  $\delta\Psi_\mu = -2M_p\partial_\mu\epsilon(x)$ . The variation of the matter action defines the Majorana supercurrent  $\delta J_{\text{matter}} = \int d^4x J^{\dagger\mu}\partial_\mu\epsilon(x)$  demanding the interaction Lagrangian up to order  $\mathcal{O}(M_p^{-2})$  to be eq. (3.3.1) [43].

Since all supersymmetric particles must be part of a supermultiplet, we must introduce one more particle. The gravitino is a supersymmetric particle, and must be the superpartner of a spin 2 field. This is the Graviton from general relativity, and together with the gravitino is the helicity 3/2 supermultiplet.

## 3.4 The Gravitino

The interactions of the gravitino are defined in eq. (3.3.1). This equation shows the fundamental features of the gravitino: 1) the gravitino have both vector and spinor indices, which makes the gravitino a spin 3/2 particle, 2) the dimension is  $[\Psi_\mu] = 3/2$  as for spin 1/2 fermions [43] and 3) the inverse planck mass factor making all interactions with the gravitino weak.

The interactions of the gravitino are defined in eq. (3.3.1). From this, we can see the fundamental features of the gravitino. First, we notice the gravitino have both vector and spinor indices, this makes the gravitino a spin 3/2 particle. The dimension is  $[\Psi_\mu] = 3/2$  as for spin 1/2 fermions [43]. Most notably is the inverse planck mass factor making all interactions with the gravitino weak.

### 3.4.1 The super-Higgs mechanism

In this section, we describe how the gravitino obtains its mass. The derivation follows the appendix of [44] closely. The gauge-invariant Lagrangian for the goldstino-gravitino system is [46, 47]

$$\mathcal{L} \supset -\frac{1}{2}\epsilon\bar{\Psi}_\mu\gamma_5\gamma_\nu\partial_\rho\Psi_\sigma + \frac{1}{2}\bar{G}i\gamma^\mu\partial_\mu\tilde{G} - m_{3/2}\left(\frac{1}{4}\bar{\Psi}_\mu[\gamma^\mu,\gamma^\nu]\Psi_\nu + \bar{G}\tilde{G} - \sqrt{\frac{3}{2}}\bar{\Psi}_\mu i\gamma^\mu\tilde{G}\right), \quad (3.4.1)$$

where the two first terms are the kinetic terms of the gravitino and the goldstino, then the mass terms with a mass mixing term. The gravitino-goldstino mass mixing term comes from the gravitino interaction Lagrangian eq. (3.3.1), where the relevant supercurrent term is  $i\sqrt{2}\langle F\rangle\gamma^\mu\tilde{G}$  (eq. (3.2.2)). This Lagrangian is invariant under the local transformations of the gravitino ( $\delta\Psi_\mu$ ) and goldstino ( $\delta\tilde{G}$ ) fields, provided that the gravitino mass  $m_{3/2}$  and the supersymmetry breaking VEV  $\langle F\rangle$  are related as [44, 46, 47]

$$m_{3/2} = \frac{\langle F\rangle}{\sqrt{3}M_p}. \quad (3.4.2)$$

To avoid the mass mixing term in eq. (3.4.1), we choose a unitary gauge

$$\Psi'_\mu = \Psi_\mu - \frac{i}{\sqrt{6}}\gamma_\mu\tilde{G} - \sqrt{\frac{2}{3}}\frac{1}{m_{3/2}}\partial_\mu\tilde{G}, \quad (3.4.3)$$

which describes a massive gravitino. The goldstino has been ‘absorbed’ in the gravitino, gaining the helicity  $\pm 1/2$  states. The Lagrangian of the massive gravitino becomes [47]

$$\mathcal{L} = -\frac{1}{2}e^{\mu\nu\rho\sigma}\bar{\Psi}'_\mu\gamma_5\gamma_\nu\partial_\rho\Psi'_\sigma - \frac{m_{3/2}}{4}\bar{\Psi}'_\mu[\gamma^\mu,\gamma^\nu]\Psi'_\nu + \frac{1}{2M_p}\bar{\Psi}'_\mu J_{vis}^\mu. \quad (3.4.4)$$

This Lagrangian is the Rarita-Schwinger Lagrangian for massive spin 3/2 particles. If we find the equation of motion for the Lagrangian, we obtain the Rarita-Schwinger equation, similar to the Dirac equation for spin 1/2 particles. This can be reduced to the following equations

$$\gamma^\mu\psi_\mu^\pm(\mathbf{p}, r) = 0, \quad (3.4.5a)$$

$$p^\mu\psi_\mu^\pm(\mathbf{p}, r) = 0, \quad (3.4.5b)$$

$$(\not{p} - m_{3/2})\psi_\mu^\pm(\mathbf{p}, r) = 0, \quad (3.4.5c)$$

where  $\psi_\mu(\mathbf{p})$  is the wave function of the gravitino in momentum space,  $\Psi_\mu \sim e^{ipx} \psi_\mu(\mathbf{p})$ . The polarization  $r$  is the four gravitino helicities,  $s = \pm 3/2, \pm 1/2$ . The spin sum for spin 3/2 particle, with momentum  $p$  is [48]

$$\begin{aligned} \Pi^{(\pm)}(p)_{\mu\nu} &= \sum_r \psi_\mu^{(\pm)}(\mathbf{p}; r) \bar{\psi}_\nu^{(\pm)}(\mathbf{p}; r) \\ &= -(\not{p} \pm m_{3/2}) \left\{ g_{\mu\nu} - \frac{1}{3} (\gamma_\mu \gamma_\nu + \gamma_\mu \frac{p_\nu}{m_{3/2}} - \frac{p_\mu}{m_{3/2}} \gamma_\nu + 2 \frac{p_\mu p_\nu}{m_{3/2}^2}) \right\} \end{aligned} \quad (3.4.6)$$

and obeys the following equations, corresponding to eq. (3.4.5a)

$$\gamma^\mu \Pi_{\mu\nu}^\pm(p) = \Pi_{\mu\nu}^\pm(p) \gamma^\mu = 0, \quad (3.4.7a)$$

$$p^\mu \Pi_{\mu\nu}^\pm(p) = \Pi_{\mu\nu}^\pm(p) p^\mu = 0, \quad (3.4.7b)$$

$$(\not{p} \mp m_{3/2}) \Pi_{\mu\nu}^\pm(p) = \Pi_{\mu\nu}^\pm(p) (\not{p} \mp m_{3/2}) = 0. \quad (3.4.7c)$$

### 3.4.2 The gravitino-goldstino equivalence

After the absorption of the goldstino, the gravitino inherits the interactions of the goldstino. The interactions of the gravitino and goldstino can be separated into two terms, making it possible for equivalence between the massive gravitino and goldstino [49, 50]. The relevant interaction term is the last term in eq. (3.4.4) and describes the interaction of the massive gravitino field to the visible sector particles. We put the redefined gravitino gauge (eq. (3.4.3)) back into this term

$$\begin{aligned} \mathcal{L}_{3/2,int} &\supset -\frac{1}{2M_p} \bar{\Psi}'_\mu J_{\text{vis}}^\mu \\ &= -\frac{1}{2M_p} \left( \bar{\Psi}_\mu J_{\text{vis}}^\mu - \frac{i}{\sqrt{6}} \gamma_\mu \bar{G} J_{\text{vis}}^\mu - \sqrt{\frac{2}{3}} \frac{1}{m_{3/2}} (\partial_\mu \bar{G}) J_{\text{vis}}^\mu \right) \\ &= -\frac{1}{2M_p} \left( \bar{\Psi}_\mu J_{\text{vis}}^\mu - \frac{i}{\sqrt{6}} \gamma_\mu \bar{G} J_{\text{vis}}^\mu + \sqrt{\frac{2}{3}} \frac{1}{m_{3/2}} \bar{G} \partial_\mu J_{\text{vis}}^\mu \right), \end{aligned} \quad (3.4.8)$$

where integrating by parts is used;  $(\partial_\mu \bar{G}) J_{\text{vis}}^\mu = -\bar{G} \partial_\mu J_{\text{vis}}^\mu + (\text{Higher order derivatives})$ . We know from section 3.2 that the divergence of the supercurrent of the visible sector are proportional to the mass difference of the particles in the supermultiplet, i.e.,  $m_{\tilde{f}} - m_f$  for fermion-scalar supermultiplets and  $m_{\tilde{g}} - m_g$  for gaugino-gauge supermultiplets. For light gravitinos, satisfying  $m_{3/2} \ll \sqrt{s}$ , we have  $m_{3/2} \ll m_{\tilde{f}} - m_f$ . In this case the last term will be the dominant interaction term and the gravitino can be replaced by



the last term in the gravitino gauge [50]

$$\psi_\mu \sim \sqrt{\frac{2}{3}} \frac{1}{m_{3/2}} \partial_\mu \tilde{G}. \quad (3.4.9)$$

After doing the substitution, we arrive at the goldstino Lagrangian with two derivatives (eq. (3.2.3)). By using that the term  $\tilde{G} \partial_\mu J^{\dagger\mu}$  is proportional to the mass difference of the fermion and scalar in a supermultiplet, we obtain the goldstino Lagrangian without derivatives (eq. (3.2.4)). This shows that the goldstino decouples from the gravitino in the supersymmetry limit ( $\langle F \rangle \rightarrow 0$ ) and that the interaction Lagrangian for the goldstino vanishes [13]. This is called the gravitino-goldstino equivalence.

By utilizing this substitution, decay rates and cross sections can be calculated by the use of a standard spin 1/2 fermion and not the spin 3/2 spinorial vector field. This makes the calculation simpler because the standard spin 1/2 spin sum can be used instead of the gravitino spin sum eq. (3.4.6). In the next section, this equivalence is used to calculate the tau-sneutrino lifetime.

# Chapter 4

## The lightest supersymmetric particle

In this thesis, the LSP is the gravitino, which is stable under the assumption of R-parity. This also means that all supersymmetric particles eventually decay into the gravitino. As we have seen in the previous chapter, the coupling of the gravitino to the visible sector is weak. This leads to the cosmological gravitino problem discussed below, and to long-lived supersymmetric particles. We calculate the lifetime of the next-to lightest supersymmetric particle, which decays into the gravitino.

### 4.1 The gravitino problem

If we assume the Neutralino to be the LSP, the gravitino is in this case unstable [13]. As we saw in the previous chapter the gravitino couples weakly to other particles, and hence has a relatively long lifetime of up to several years [51]. In the early universe, when the nuclear interaction becomes effective, the synthesis of light elements starts. The theory describing this is called the *Big Bang Nucleosynthesis* (BBN) and predicts the abundance of the light elements  $D$ ,  ${}^3H_e$ ,  ${}^4H_e$  and  ${}^7Li$  [52, 53]. The observed abundance of the primordial light elements should match predictions of the BBN. If the gravitino decays after these elements have been produced, its decay product destroys nuclei produced by BBN. This is called the gravitino problem [51, 54]. To match the observed data to predictions requires the gravitino mass to be  $m_{3/2} \gg 1\text{TeV}$  or reheating temperature after inflation to be  $T_R \lesssim 10^6\text{GeV}$ . The first constraint requires an unnatural mass spectrum in most scenarios of SUSY breaking, whereas the latter one prevents thermal leptogenesis without fine-tuning. This motivates scenarios where

the gravitino is the LSP. A relatively massive gravitino, order tens of GeV, can also be a viable cold dark matter candidate [55].

With the gravitino as the LSP, there are still a long-lived particle decaying into gravitino called the next-to lightest supersymmetric particle (NLSP). This can only decay into the gravitino because R-parity is assumed, and because of the weak coupling it is long-lived. This is still a problem for the BBN prediction. We can make the NLSP mass  $\lesssim 1$  TeV, this rules out the neutralinos, charginos and gluinos [56] as the NLSP because of the large hadronic branching ratios. A light charged slepton or stau is also constrained because they can bind together forming bound states with nuclei, altering the BBN reaction rates [57]. For a light NLSP ( $\lesssim 1$ TeV) the particle we are left with is the tau-sneutrino ( $\tilde{\nu}_\tau$ ).

## 4.2 The lightest observable supersymmetric particle

The present analysis depends on the NLSP being long-lived on detector scales, so that it can escape the detector. In the next section, the lifetime of the tau-sneutrino is calculated. The dominant decay channel for the sneutrino is

$$\tilde{\nu}_\tau \rightarrow \tilde{\Psi} + \nu_\tau, \quad (4.2.1)$$

where  $\tilde{\Psi}$  is the gravitino. There are also three body decays such as  $\tilde{\nu}_\tau \rightarrow \tilde{G} + \nu_\tau + \gamma$ , but these are negligible [7] for the present discussion.

Because of the long-lived properties of the sneutrinos, the dark matter relic density is related to the sneutrino density before it decays. The relation depends on the decay modes of the sneutrino, which in this case is the gravitino and neutrino. This means that the relation will be proportional to the factor  $m_{\tilde{G}}/m_{\tilde{\nu}_\tau}$ , leaving few cosmological constraints on the sneutrino mass for  $m_{\tilde{\nu}_\tau} \lesssim 1$  TeV [7].

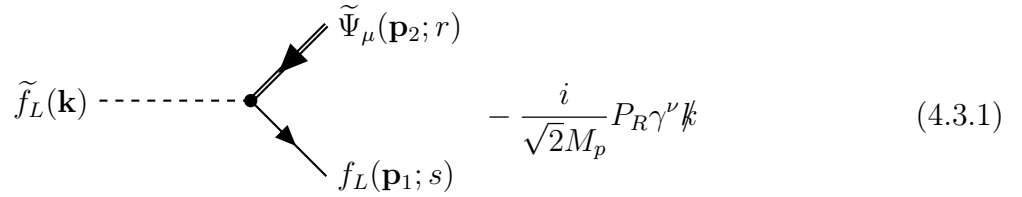
## 4.3 Sneutrino (NLSP) lifetime

From the previous chapter, we know that there are two different ways to calculate lifetimes concerning a single gravitino. One is to calculate it using the full spin 3/2 formalism to describe the gravitino, and the other is using the gravitino-goldstino

equivalence theorem to reduce the problem to spin 1/2 formalism, by considering only the goldstino component of the gravitino. This section contains both calculations, followed by a comparison of the two and a discussion about the validity of the goldstino approximation.

### 4.3.1 Spin 3/2 formalism

We start by calculating the lifetime using the gravitino field  $\tilde{\Psi}_\mu$ . The Feynman diagram for this decay with the corresponding vertex factor is [58]



$$\tilde{f}_L(\mathbf{k}) \text{ --- } \bullet \begin{cases} \nearrow \tilde{\Psi}_\mu(\mathbf{p}_2; r) \\ \searrow f_L(\mathbf{p}_1; s) \end{cases} \quad - \frac{i}{\sqrt{2}M_p} P_R \gamma^\nu \not{k} \quad (4.3.1)$$

The Feynman rules for the external lines are

- $\tilde{f}_L(\mathbf{k}) \rightarrow 1$  for the initial left-handed sfermion,
- $f_L(\mathbf{p}_1; s) \rightarrow u_s(\mathbf{p}_1)$  for the final left-handed fermion,
- $\tilde{\Psi}_\mu(\mathbf{p}_2; r) \rightarrow \psi_\mu^{(-)}(\mathbf{p}_2; r)$  for the final gravitino.

This gives the matrix element

$$\mathcal{M} = -\frac{i}{\sqrt{2}M_p} \bar{u}_s(\mathbf{p}_1) P_R \gamma^\mu \not{k} \psi_\mu^{(-)}(\mathbf{p}_2; r). \quad (4.3.2)$$

Squaring the matrix element, summing over the final state polarizations and using the spin sum for the fermions eqs. (B.2.5) and (3.4.6), yields

$$\begin{aligned} |\mathcal{M}|_{\text{total}}^2 &= \frac{1}{2M_p^2} \text{Tr}[\Pi_{\mu\nu}^{(-)}(p_2) \not{k} \gamma^\nu P_L (\not{p}_1 + m_f) P_R \gamma^\mu \not{k}] \\ &= \frac{1}{2M_p^2} \text{Tr}[\Pi_{\mu\nu}^{(-)}(p_2) \not{k} \gamma^\nu \not{p}_1 \gamma^\mu \not{k} P_R], \end{aligned} \quad (4.3.3)$$

where the Clifford algebra eq. (A.2.3), the spin sum relations  $\gamma^\mu \Pi_{\mu\nu}(p) = 0$  and  $\Pi_{\mu\nu} \gamma^\nu = 0$  for spin 3/2 particles are used. The matrix element squared becomes

(see section [B.2](#) for the full calculation)

$$|\mathcal{M}|_{\mathcal{P}\mathcal{O}}^2 = \frac{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)^3}{3M_p^2 m_{3/2}^2} \left( 1 - \frac{4m_{3/2}^2 m_f^2}{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)^2} \right)^3. \quad (4.3.4)$$

Since the final state fermion is a neutrino, its mass can be neglected ( $m_f = m_{\nu_\tau} \approx 0$ ) and the squared matrix element becomes

$$|\mathcal{M}|_{\mathcal{P}\mathcal{O}}^2 = \frac{m_{\tilde{\nu}_\tau}^6}{3M_p^2 m_{3/2}^2} \left( 1 - \frac{m_{3/2}^2}{m_{\tilde{\nu}_\tau}^2} \right)^3, \quad (4.3.5)$$

where the tau-sneutrino mass is substituted for the general sfermion mass  $m_{\tilde{f}} = m_{\tilde{\nu}_\tau}$ . The decay rate is

$$\Gamma_\Psi = \frac{m_{\tilde{\nu}_\tau}^5}{48M_p^2 m_{3/2}^2} \left( 1 - \frac{m_{3/2}^2}{m_{\tilde{\nu}_\tau}^2} \right)^4, \quad (4.3.6)$$

using the 2-body decay rate formulas in section [B.1](#). This is the decay rate calculated using the gravitino, which is also obtained in the same form in [7](#).

### 4.3.2 Goldstino approximation

From section [3.4.2](#) we saw that the calculation can be done using the gravitino-goldstino equivalence theorem. This means that the spin 3/2 gravitino can be substituted by a spin 1/2 goldstino (eq. [\(3.4.9\)](#)) which makes the spin sum the same as any other fermion. The corresponding Feynman diagram and the vertex factor is [59](#)

$$-i \frac{m_{\tilde{f}}^2 - m_f^2}{\sqrt{3} M_p m_{3/2}} \quad (4.3.7)$$

The Feynman rules for the external lines are

- $\tilde{f}_L(\mathbf{k}) \rightarrow 1$  for the initial left-handed sfermion,
- $f_L(\mathbf{p}_1; s) \rightarrow u_s(\mathbf{p}_1)$  for the final left-handed fermion,
- $\tilde{G}(\mathbf{p}_2; r) \rightarrow v_r(\mathbf{p}_2)$  for the final goldstino.

The matrix element is

$$\mathcal{M} = -i \frac{m_{\tilde{f}}^2 - m_f^2}{\sqrt{3} M_p m_{3/2}} \bar{u}_s(p_1) P_L v_r(p_2). \quad (4.3.8)$$

Summing over the polarizations and multiplying with the hermitian conjugate

$$\sum_{s,r} |\mathcal{M}|_{\cancel{p}\cancel{\sigma}}^2 = \frac{(m_{\tilde{f}}^2 - m_f^2)^2}{3M_p^2 m_{3/2}^2} \text{Tr} \left[ \sum_r v_r(p_2) \bar{v}_r(p_2) \sum_s u_s(p_1) \bar{u}_s(p_1) P_L \right]. \quad (4.3.9)$$

Using the spin sums for fermions eq. (B.2.5) and the trace identities eqs. (A.2.4) and (A.2.7), the matrix element yields

$$|\mathcal{M}|_{\cancel{p}\cancel{\sigma}}^2 = \frac{m_{\tilde{f}}^2 (m_{\tilde{f}}^2 - m_f^2)^2}{3M_p^2 m_{3/2}^2} \left( 1 - \frac{(m_{3/2} + m_f)^2}{m_{\tilde{f}}^2} \right). \quad (4.3.10)$$

The decay rate and the momentum is the same as in previous section, putting this and the matrix element into the decay rate and neglecting the neutrino mass

$$\Gamma_{\tilde{G}} = \frac{m_{\tilde{\nu}_\tau}^5}{48\pi M_p^2 m_{3/2}^2} \left( 1 - \frac{m_{3/2}^2}{m_{\tilde{\nu}_\tau}^2} \right)^2. \quad (4.3.11)$$

This is the decay rate for a tau-sneutrino  $m_{\tilde{\nu}_\tau}$  decaying to a goldstino/gravitino and a fermion.

There is a factor  $\frac{\Gamma_{\Psi}}{\Gamma_{\tilde{G}}} = \left( 1 - \frac{m_{3/2}^2}{m_{\tilde{\nu}_\tau}^2} \right)^2$  difference between the decay rate calculated using gravitino and goldstino. According to the assumption made in the gravitino-goldstino equivalence (section 3.4.2), that  $m_{3/2} \ll m_{\tilde{f}} - m_f$ , this factor is not relevant as long as the assumption holds. For gravitino masses close to the sneutrino mass, the assumption do not hold and the factor is relevant.

We can plot the lifetime of the tau-sneutrino to see the difference in two calculations. The lifetime is given by

$$\tau = \frac{1}{\sum_i \Gamma_i} \quad (4.3.12)$$

where the summation is over all decay modes of the particle. For the tau-sneutrino, this is the only possible 2-body decay, and the three body decays can be neglected. The lifetime is plotted in fig. 4.3 together with the matrix element squared. The red is the gravitino calculation, and the dashed green is the goldstino calculation, the grey area is where the decay and matrix element is non-physical, i.e. not possible for

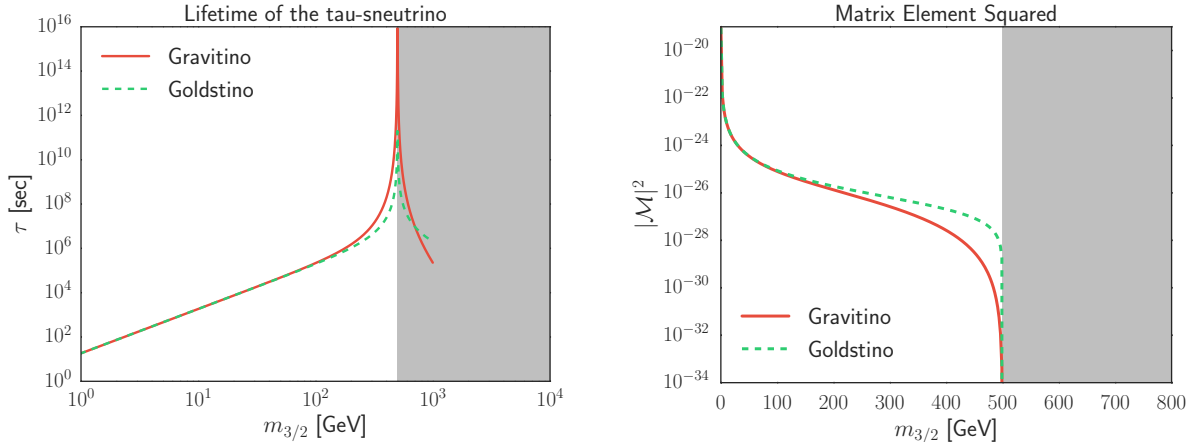


Figure 4.3: The lifetime (left) and the corresponding matrix element (right), of a tau-sneutrino decaying to a gravitino (red solid line) or a goldstino (green dashed line). For decays of  $m_{\tilde{\nu}_\tau} = 500\text{GeV}$ . The grey area is non-physical, since  $m_{3/2} > m_{\tilde{\nu}_\tau}$ .

the tau-sneutrino to decay to a heavier gravitino. The equivalence holds up to about  $m_{3/2} = 100\text{GeV}$  for a  $500\text{GeV}$  tau-sneutrino.

As we can see from the plot the lifetime of the sneutrino is much longer than the time it takes to escape the detector. This means that in collider physics the tau-sneutrino can be treated as a stable particle, and all supersymmetric decay chains will end with a tau-sneutrino.

# Chapter 5

## NUHM parameter space

In this thesis, the Non-Universal Higgs Masses (NUHM) model is utilized. This means that the soft supersymmetry breaking mass parameters for gaugino, squark and leptons are universal as in mSUGRA, see eq. (2.6.1). Additionally, the soft supersymmetry-breaking contributions to Higgs scalar masses are non-universal, i.e.  $m_{H_d}^2 = m_{H_u}^2 = m_0^2$  does not need to be satisfied. This is the NUHM2 model which we call NUHM from now on, the NUHM1 is when  $m_{H_d}^2 = m_{H_u}^2$  [60]. The reason for using the NUHM model is that in the CMSSM there are no parameter points with the tau-sneutrino as the NLSP [7]. This can be seen in the fig. 5.1 where the points  $m_{H_d}^2 = m_{H_u}^2 = m_0^2$ , are far from the region where the tau-sneutrino is the the NLSP. The reason why we make the Higgs masses non-universal, is because the Higgses does not introduce any *flavor changing neutral current* (FCNC) to the model<sup>1</sup>.

### 5.1 Renormalization group equations

To understand the behavior of the masses and parameters in the NUHM model, we first examine the renormalization group equations (RGE). These equations describe how parameters are evolved from the GUT scale down to a measurable scale. The RGEs for the NUHM model have additional terms beyond those appearing in the CMSSM. This leads to some novel feature differences between the two models which

---

<sup>1</sup>A non-universal gaugino masses (NUGM) model do not contribute to FCNC either.



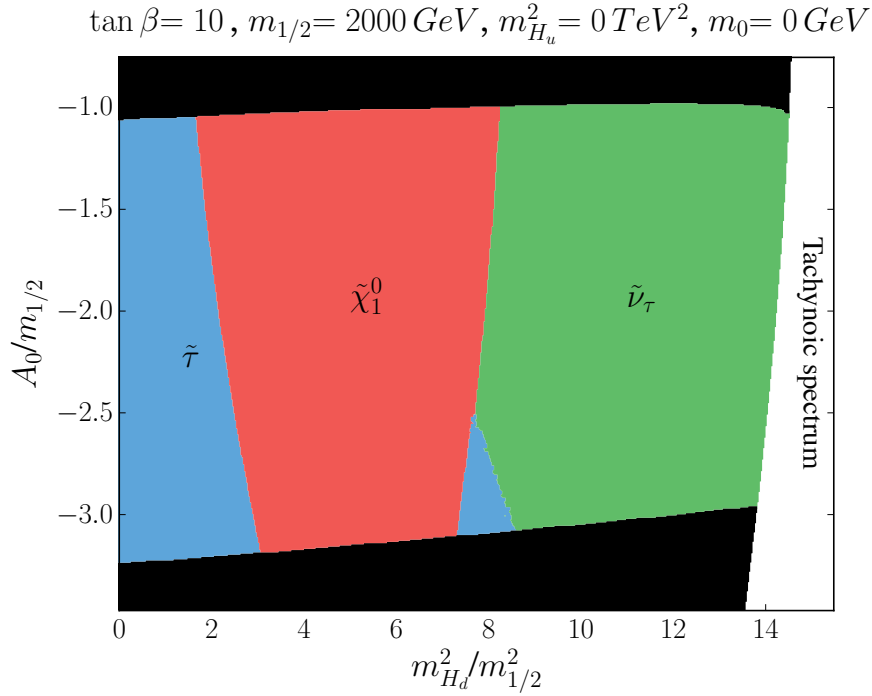


Figure 5.1: Regions corresponding to a tau-sneutrino (green), neutralino (red) or stau (blue) NLSP. The black areas corresponds to points where the Higgs mass is outside the range  $123 \text{ GeV} \leq m_{h^0} \leq 127 \text{ GeV}$ .

are discussed in this section. The new features come from a non-zero  $S$  term [61]

$$S \equiv \frac{g_1^2}{5} \left( m_{H_u}^2 - m_{H_d}^2 + 2(m_{\tilde{Q}}^2 - m_{\tilde{L}}^2 - 2m_{\tilde{u}_R}^2 + m_{\tilde{d}_R}^2 + m_{\tilde{e}_R}^2) \right. \\ \left. + (m_{\tilde{Q}_3}^2 - m_{\tilde{L}_3}^2 - 2m_{\tilde{t}_R}^2 + m_{\tilde{b}_R}^2 + m_{\tilde{\tau}_R}^2) \right), \quad (5.1.1)$$

which at the GUT scale is only obtained by NUHM. This is because the dominant terms are  $m_{H_u}^2 - m_{H_d}^2$  [13] which is zero in CMSSM.

The third generation slepton, squarks and the soft Higgs masses obtain a contribution from the Yukawa ( $y_{t,b,\tau}$ ) and soft ( $a_{t,b,\tau}$ ) couplings. These are neglected in the first and second generation squarks and sleptons. At one-loop order, these contributions appear in three combinations

$$X_t = 2|y_t|^2(m_{H_u}^2 + m_{\tilde{Q}_3}^2 + m_{\tilde{u}_R}^2) + 2|a_t|^2 \\ X_b = 2|y_b|^2(m_{H_d}^2 + m_{\tilde{Q}_3}^2 + m_{\tilde{d}_R}^2) + 2|a_b|^2 \\ X_\tau = 2|y_\tau|^2(m_{H_d}^2 + m_{\tilde{L}_3}^2 + m_{\tilde{\tau}_R}^2) + 2|a_\tau|^2, \quad (5.1.2)$$

where the trilinear term  $|a_x|^2 = |y_x|^2|A_x|^2$  for  $x = t, b$  and  $\tau$ . The RGEs for the two

soft Higgs mass parameters and the sleptons ( $L$  and  $e$  is the first and second generation and  $L_3$  and  $\tau$  is the third) is given by

$$\begin{aligned}
16\pi^2 \frac{d}{dt} m_{H_u}^2 &= 3X_t - 6g_2^2 |M_2|^2 - \frac{6}{5}g_1^2 |M_1|^2 + 3S \\
16\pi^2 \frac{d}{dt} m_{H_d}^2 &= 3X_b + X_\tau - 6g_2^2 |M_2|^2 - \frac{6}{5}g_1^2 |M_1|^2 - 3S \\
16\pi^2 \frac{d}{dt} m_{L_L}^2 &= -6g_2^2 |M_2|^2 - \frac{6}{5}g_1^2 |M_1|^2 - 3S \\
16\pi^2 \frac{d}{dt} m_{e_R}^2 &= -\frac{24}{5}g_1^2 |M_1|^2 + 6S \\
16\pi^2 \frac{d}{dt} m_{L_3}^2 &= X_\tau - 6g_2^2 |M_2|^2 - \frac{6}{5}g_1^2 |M_1|^2 - 3S \\
16\pi^2 \frac{d}{dt} m_{\tau_R}^2 &= 2X_\tau - \frac{24}{5}g_1^2 |M_1|^2 + 6S,
\end{aligned} \tag{5.1.3}$$

where  $M_1$ ,  $M_2$  and  $M_3$  are the soft gaugino mass parameters and  $t=\ln(Q/Q_0)$ .  $Q$  is the renormalization scale and  $Q_0$  is the reference mass<sup>2</sup>. One example of new features in the NUHM model is when  $S < 0$ , then the  $3S$  term contributes positively to the change in the left-chiral slepton soft mass. This opens the possibility of the left-handed sleptons to be lighter than the right-handed [62], making a tau-sneutrino NLSP possible. For  $S > 0$  the difference between the right-handed and the left-handed masses increase, with the right-handed as the lightest.

The Yukawa and trilinear couplings  $X_t$ ,  $X_b$  and  $X_\tau$  are in general positive. This means that the contribution to the masses is negative and that the third generation becomes the lightest generation. When  $X_t$ ,  $X_b$  and  $X_\tau$  are negative, the first generation is the NLSP.

## 5.2 Parameter point

### 5.2.1 Spectrum generator

The RGEs must be applied to obtain the masses of the supersymmetric particles. This is already implemented in spectrum generators, which take the supersymmetric parameters as input. These spectrum generators then utilize the parameters as boundary conditions to solve the MSSM RGEs. In this project SPheno 4.0.3 is used [63, 64] and here the calculation is done in two-loop order. SPheno also calculates other important

<sup>2</sup>Usually at  $m_Z$

quantities, such as the the decay width and branching ratios of all the supersymmetric particles and all the model parameters. The SM masses and gauge couplings utilized in the parameter scan is the default values from SPheno

$$G_F = 1.66379 \times 10^{-5} \text{ GeV} \quad \alpha_s(M_Z) = 1.184 \times 10^{-1} \quad m_Z(\text{pole}) = 91.1876 \text{ GeV}$$

$$m_b(m_b) = 4.18 \text{ GeV} \quad m_t(\text{pole}) = 173.1 \text{ GeV} \quad m_\tau(\text{pole}) = 1.77682 \text{ GeV}$$

Here,  $\alpha_s \equiv \frac{g_s^2}{4\pi}$ , where  $g_s$  is the QCD gauge coupling.  $G_F$  is the Fermi constant,  $m_Z$  is the pole mass of the Z-boson and  $m_t$ ,  $m_b$ ,  $m_\tau$  are the three third-generation fermion masses.

### 5.2.2 Reducing scanning parameters

In NUHM there are six free parameters. This results in a six-dimensional space of points, which is hard to visualize. Thus, for visualization purposes, it is necessary to reduce the number of parameters. The parameter point must satisfy the tau-sneutrino NLSP scenario and should not be ruled out by experiments. By considering parameters that are most affected by these constraints, it is possible to reduce the number of parameters to scan. We can use the knowledge from the previous section about the RGEs (eq. (5.1.3)) to find these parameters.

As seen from the RGEs, the difference in the Higgs soft parameters must be non-zero and negative ( $m_{H_d}^2 \gg m_{H_u}^2$ ) to obtain a tau-sneutrino NLSP. Additionally, because of the discovery of the Higgs boson at  $125.09 \pm 0.24 \text{ GeV}$  [31], which we assume to be the lightest CP-even Higgs in MSSM, we want the the lightest Higgs mass to be in the range  $123 \text{ GeV} \leq m_{h^0} \leq 127 \text{ GeV}$  [65, 66]. The uncertainty in this constraint is due to the theoretical calculations. The contribution of  $m_{H_u}^2$  and  $m_{H_d}^2$  to the lightest Higgs mass is small [67]. It is therefore possible to fix one of the soft Higgs masses, and then scan the other, to change the difference between them. A convenient choice is fixing  $m_{H_u}^2 = 0$  and then scan for  $m_{H_d} > 0$ .

A second parameter to constrain is the soft scalar mass parameter  $m_0$ . This contributes positively to all the scalar masses, and thereby raise the total spectrum. Since having a Higgs mass in the expected range usually comes with a heavy spectrum [67], many of these points are undetectable in our scenario. Another reason for choosing as low as possible spectrum is to get as close as possible to the spectrum in [8]. Their mass spectrum have  $m_{\tilde{\nu}_\tau} = 90.5 \text{ GeV}$ ,  $m_{\tilde{t}_1} = 723.6 \text{ GeV}$ ,  $m_{\tilde{\chi}_1^0} = 206.0$  and  $m_{h^0} = 115.9 \text{ GeV}$ . Because these masses are lighter then the corresponding one in a spectrum with the

Higgs mass of  $125 \pm 2$  GeV, the amount of signal is reduced in the correct Higgs mass scenario. We choose  $m_0 = 0$ , to obtain the scalar masses as low as possible because these are the primary production particles in this scenario<sup>3</sup>

The last parameter to fix is  $\tan \beta = 10$  as this gave the smallest sneutrino mass.  $\tan \beta$  is defined at the GUT scale together with  $m_{H_u}^2, m_{H_d}^2$  and the other parameters.

### 5.2.3 Parameter scan

The remaining free parameters are  $A_0$ ,  $m_{1/2}$  and  $m_{H_d}^2$ , which have to be plotted to visualize the mass spectrum change. Figure 5.2 shows a scan of the  $(m_{H_d}^2, A_0)$  and  $(m_{H_d}^2, m_{1/2})$  planes with a colormap of the sneutrino mass. If the sneutrino mass is below 200 GeV, the points lie on the yellow spots under the red line. The lines represent the sneutrino mass contours. We can see that the mass of the sneutrino is dependent on the soft Higgs masses difference and the soft gaugino mass parameter  $m_{1/2}$ . In regard to the soft Higgs masses, the sneutrino mass decrease as the difference increase. This is, as explained above, because of the  $S$  in the RGEs. Since the gaugino masses,  $M_1$  and  $M_2$ , contributes negative to the lepton eq. (5.1.3)s, the increase in  $m_{1/2}$  increases the sneutrino mass.

The limit of how light the sneutrino mass can become as  $|m_{H_u}^2 - m_{H_d}^2|$  increase is when the squared stau mass runs negative. This gives a tachyon error in *SPheno* and is represented in the white areas of the plots. The gray areas are where the Higgs mass is outside the  $123 \text{ GeV} \leq m_{h^0} \leq 127 \text{ GeV}$  range.

In fig. 5.3 the individual masses is plotted as a function of the soft Higgs mass parameter  $m_{H_d}^2$ . This plot shows how the sneutrino mass becomes lighter as the  $m_{H_d}^2$  increases, and overlaps the stau mass. Since there are some uncertainty in the theoretical calculation of the masses, the difference between the sneutrino mass and the stau mass should not be too small. This is to be sure the point does not have the stau as its NLSP for higher accuracy calculations. The difference is plotted in the right plot of fig. 5.3. They are part of the same  $SU(2)_L$  multiplet, and have therefore the same mass contributions except for the hyperfine splitting ( $\Delta_{\tilde{\nu}}$  and  $\Delta_{\tilde{\tau}}$ , see eq. (2.5.11)) and loop corrections.

The last thing we need to look at is the Higgs mass which must lie in the range  $123 \text{ GeV} < m_{h^0} < 127 \text{ GeV}$ . The range is chosen based on the uncertainty in the Higgs

<sup>3</sup>Setting  $m_0 = 0$  is required in gaugino mediated supersymmetry breaking [36, 67]

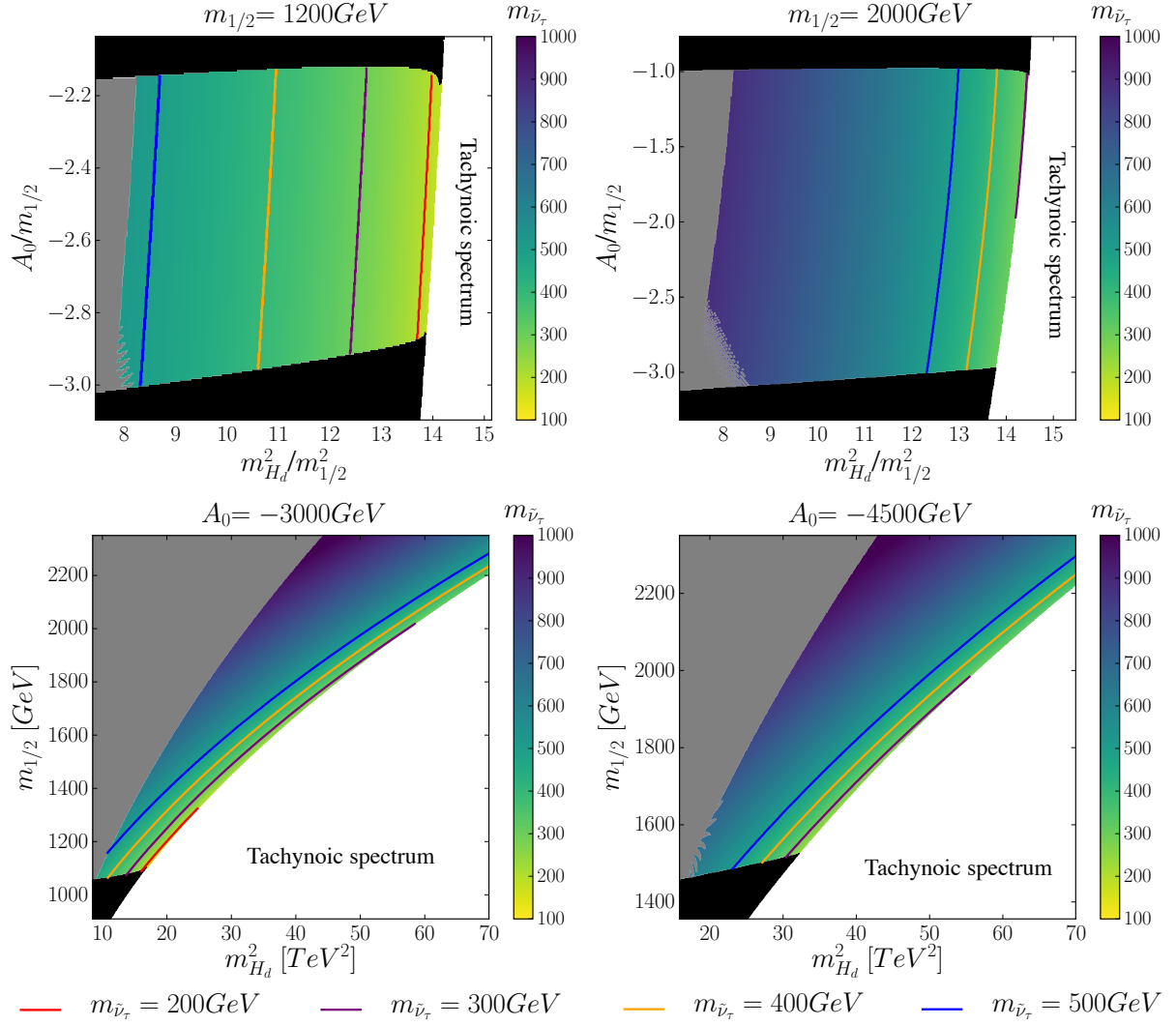


Figure 5.2: The variation of the sneutrino mass ( $m_{\tilde{\nu}_\tau}$ ), represented as a heatmap, in the  $(m_{H_d}^2, A_0)$  and  $(m_{H_d}^2, m_{1/2})$  plane using  $m_0 = 0$ ,  $\tan \beta = 10$  and  $m_{H_u}^2 = 0$ . The gray region is where the sneutrino is not the NLSP and the black region is where the Higgs mass is outside the range  $123 \text{ GeV} < m_{h^0} < 127 \text{ GeV}$ .

mass calculations from *SPheno* [65, 66]. The sneutrino mass plots, fig. 5.2, show the lightest sneutrino mass is at  $m_{1/2} = 1200 \text{ GeV}$ . The Higgs mass is plotted in fig. 5.4 for the same parameter planes as the sneutrino mass plots (fig. 5.2). The Higgs mass follows  $A_0$  and  $m_{1/2}$  parameters. The reason for the small contribution to the lightest Higgs mass from  $m_{H_d}^2$  is because of the loop corrections. These loop corrections are dominated by the stop mass [13]. Contributions to increase or decrease in the stop

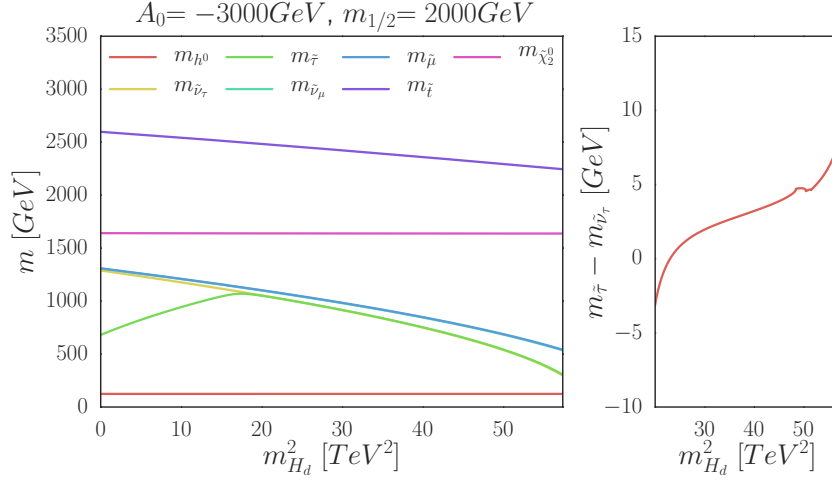


Figure 5.3: The change of masses due to change in  $m_{H_d}^2$  to the left, and the mass split between the stau ( $m_{\tilde{\tau}}$ ) and the tau-sneutrino ( $m_{\tilde{\nu}_\tau}$ ) to the right. The parameters  $m_0 = 0$ ,  $\tan \beta = 10$  and  $m_{H_u}^2 = 0$ , as in the parameter scans, are used.

mass can be seen from the third generation soft squark masses RGEs

$$\begin{aligned}
 16\pi^2 \frac{d}{dt} m_{\tilde{Q}_3}^2 &= X_t - X_b - \frac{32}{2} g_3^2 |M_3|^2 - 6g_2^2 |M_2|^2 - \frac{2}{15} g_1^2 |M_1|^2 + S \\
 16\pi^2 \frac{d}{dt} m_{\tilde{u}_3}^2 &= 2X_t - \frac{32}{2} g_3^2 |M_3|^2 - 6g_2^2 |M_2|^2 - \frac{2}{15} g_1^2 |M_1|^2 + 4S,
 \end{aligned} \tag{5.2.1}$$

where  $m_{1/2}$ , which is  $m_{1/2} = M_1 = M_2 = M_3$  at the GUT scale, and  $A_0$ , which is in  $X_t$  and  $X_b$  (eq. (5.1.2)), make the most contribution, considering  $m_0 = 0$ . Note that at the GUT scale the parameter  $\mathbf{a}_x = A_0 \mathbf{y}_x$ , for  $x = t, b, \tau$ .

The parameter point used in the rest of this thesis is:  $m_0 = 0$ ,  $m_{1/2} = 1200 \text{ GeV}$ ,  $A_0 = -3000 \text{ GeV}$ ,  $m_{H_d}^2 = 20 \text{ TeV}^2$ ,  $m_{H_u}^2 = 0$ . This point satisfies all the constraints. The Higgs mass ( $m_{h^0} = 123.3 \text{ GeV}$ ) is close to the boundary of the allowed range, but increasing the Higgs mass would decrease the amount of signal.

## 5.3 Mass spectrum

The mass spectrum of the point is given in table 5.1. The mass hierarchy follows for the most part

$$m_{\tilde{\ell}} < m_{\tilde{\chi}} < m_{\tilde{q}}. \tag{5.3.1}$$

We see there is a significant gap between the sleptons and the squarks. This makes the slepton productions dominant, as we will see in the next chapter. The slepton mass

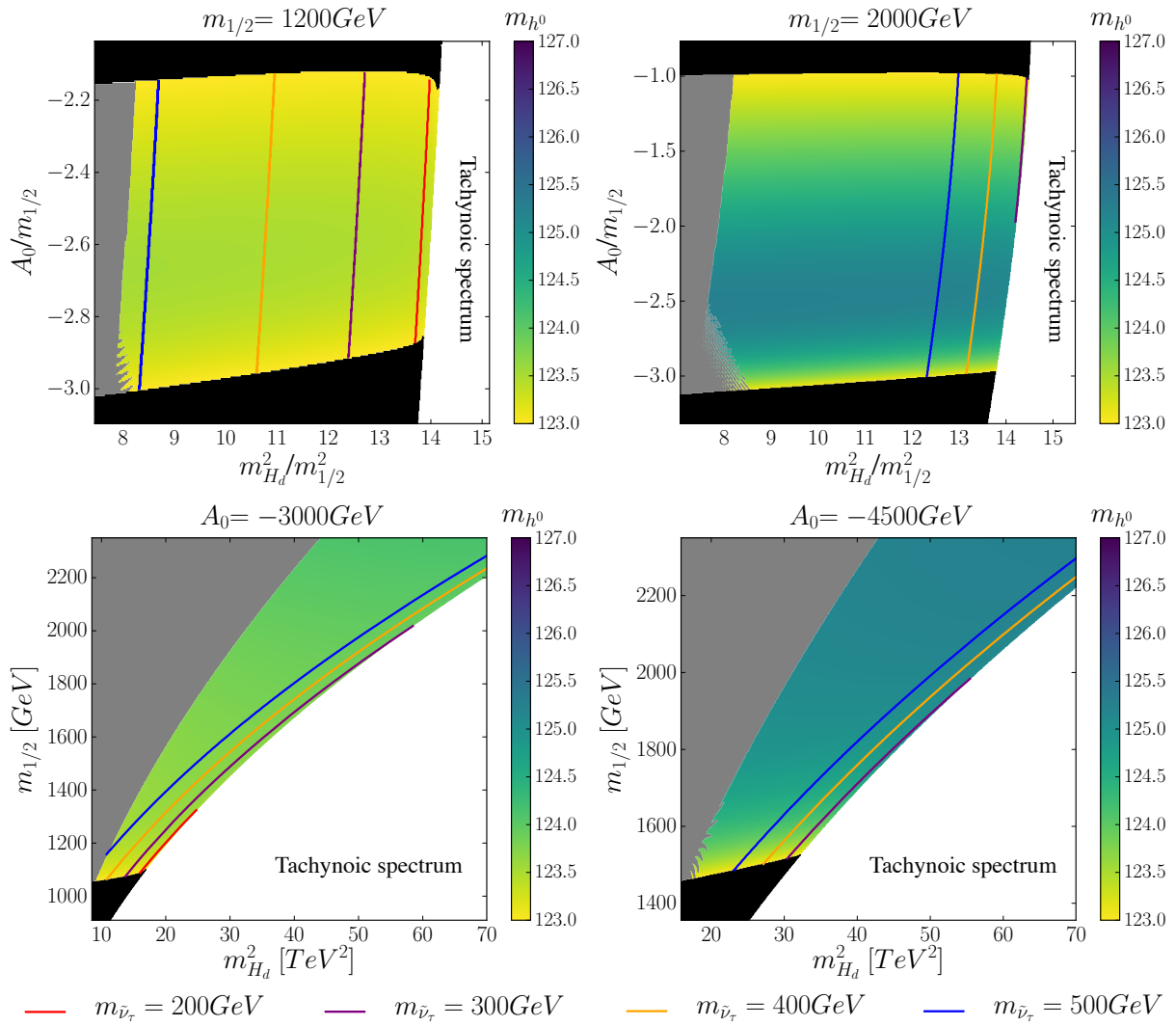


Figure 5.4: The variation of the Higgs mass ( $m_{\tilde{\nu}_\tau}$ ), represented as a heatmap, in the  $(m_{H_d}^2, A_0)$  and  $(m_{H_d}^2, m_{1/2})$  plane using  $m_0 = 0$ ,  $\tan\beta = 10$  and  $m_{H_u}^2 = 0$ . The gray region is where the sneutrino is not the NLSP and the black region is where the Higgs mass is outside the range  $123 \text{ GeV} < m_{h^0} < 127 \text{ GeV}$ .

hierarchy is as expected with the sneutrino as the NLSP;

$$m_{\tilde{\nu}_\tau} < m_{\tilde{\tau}_1} < m_{\tilde{\nu}_\mu} < m_{\tilde{\mu}_L}. \quad (5.3.2)$$

As discussed in the previous chapter, the tau-sneutrino is long-lived. The stau can also be a long-lived particle, and because of the charge of the stau, it is possible to detect it in the muon-chamber of a detector at LHC. If the stau is light and long-lived, the parameter point is ruled out [68, 69]. The lifetime of the stau is affected by the left-right mixing and mass difference to  $\tilde{\nu}_\tau$  [70]. The mixing of the  $\tilde{\tau}$ s at this parameter points gives  $\tilde{\tau}_1 \approx \tilde{\tau}_L$ . We can see this from the RGEs, eq. (5.1.3), that when

Mass [GeV]		Mass [GeV]	
		$m_{\tilde{\chi}_3^0}$	1976
$m_{\tilde{\nu}_\tau}$	195.7	$m_{\tilde{\chi}_4^0, \tilde{\chi}_2^\pm}$	1979
$m_{\tilde{\tau}_1}$	206.3	$m_{\tilde{\mu}_R, \tilde{e}_R}$	1084
$m_{\tilde{\nu}_\mu, \tilde{\nu}_e}$	352	$m_{\tilde{\tau}_2}$	1002
$m_{\tilde{\mu}_L, \tilde{e}_L}$	361	$m_{\tilde{t}_1}$	1122
$m_{\tilde{\chi}_1^0}$	515.9	$m_{\tilde{q}}$	$\sim 2000$
$m_{\tilde{\chi}_1^\pm}$	974.7	$m_{\tilde{g}}$	2589
$m_{\tilde{\chi}_2^0}$	974.5	$m_{h_0}$	123.3
		$m_{H^0, A, H^\pm}$	4882

Table 5.1: The mass spectrum for the parameter point  $m_0 = 0$ ,  $m_{1/2} = 1200$  GeV,  $A_0 = -3000$  GeV,  $m_{H_d}^2 = 20$  TeV<sup>2</sup>,  $m_{H_u}^2 = 0$ . All the squarks except  $m_{\tilde{t}_1}$  have approximately the same mass; noting them by  $\tilde{q} = \tilde{t}_2, \tilde{b}_1, \dots, \tilde{d}_R, \tilde{d}_L$ .

$m_{H_d}^2 \gg m_{H_u}^2$ , the left-chiral soft masses is lowered and the right-chiral soft masses is increased. This means that  $\tau_1$  is mostly left-handed. The width of the  $\tilde{\tau}$ -decay is  $\Gamma_{\tilde{\tau}_1} = 2.87 \times 10^{-7}$  GeV. This corresponds to a lifetime of  $\tau_{\tilde{\tau}_1} = 2.3 \times 10^{-18}$  sec, which is not long enough to reach any detectors.

## 5.4 Tests against experimtns

In this section, we test the parameter point using CheckMATE 2.0 [71-76]. To test if the parameter point is excluded from searches at LHC, CheckMATE simulated collision events using Pythia 8.2 [77], including detector simulations done by DELPHES 3 [78]. An analysis of each collision event is carried out based on the searches we want to compare to. The analysis of this parameter point is then compared to the experimental limits in the respective signal regions<sup>4</sup> of the analysis at 95% confidence interval [79]. The signal region, from all the analysis, which have the best sensitivity determines if the point is allowed or excluded.

In this thesis the parameter point is tested against all ATLAS analyses implemented in CheckMATE [80-124], which includes  $\sqrt{s} = 7$  GeV, 8 GeV and 13 GeV searches. The point was not excluded from these analyses.

<sup>4</sup>A signal region is a part of a signal after doing necessary cuts to remove background. This is discussed in the next chapter.





# Chapter 6

## Sneutrino signals in colliders

This section focuses on possible collider signatures at LHC for the parameter point found in the previous chapter;  $m_0 = 0$ ,  $m_{1/2} = 1200$  GeV,  $A_0 = -3000$  GeV,  $m_{H_d}^2 = 20$  TeV<sup>2</sup>,  $m_{H_u}^2 = 0$ . The starting point of this thesis is the cut-and-count analysis done in [8], where a trilepton signal ( $\tau_h^\pm \tau_h^\pm \ell^\mp$ ) is used ( $\tau_h$  is hadronic taus). Since we have a new parameter point, the collider features could be different from the analysis done in [8]. Thus, it is necessary to find out if the same signal can be utilized.

### 6.1 Production cross-sections

The first step is to calculate the production cross-sections of the supersymmetric particles at LHC. This is to find which production channels that dominate at LHC. The higher production cross-section, the more events can be detected. The current center of mass energy at LHC is 13TeV, which is used in this thesis. We define  $\ell = e, \mu$  and  $\ell' = e, \mu, \tau$ . When  $\ell$  is used in one of the tables below the value can be either the electron or muon version, for the squarks ( $\tilde{q}$ ) we sum over all but the stop flavor;  $\tilde{u} + \tilde{d} + \tilde{s} + \tilde{c} + \tilde{b}$ . The production cross-sections are presented in table 6.1<sup>1</sup>. The simulation is done in *Herwig 7.1.1* [125]<sup>2</sup>.

From [8] we know the sleptons are the dominant production channel, which is also the case for the new parameter point. This is as expected from the mass spectrum since

---

<sup>1</sup>The equivalent table in [8] can be found in on page 6, table 3

<sup>2</sup>Details on how the simulation is done is explained in the next chapter

the sleptons are the lightest supersymmetric particles. We notice a higher production of staus than the first two generations. This is due to the slepton mass hierarchy where the third generation is the lightest one.

The neutralino and chargino production cross-sections are small compared to the slepton production cross-sections. The dominant productions are  $\tilde{\chi}_2^0 \tilde{\chi}_1^\pm$  and  $\tilde{\chi}_1^- \tilde{\chi}_1^+$  because they are mostly wino<sup>3</sup>, which have a significant coupling to  $\gamma$ ,  $Z$  and  $W$ . The lightest neutralino,  $\tilde{\chi}_1^0$ , is mostly bino, which do not couple in pairs to  $\gamma$ ,  $Z$  and  $W$ . This makes the production channels including  $\tilde{\chi}_1^0$  small [13]. The heavier charginos and neutralinos,  $\tilde{\chi}_0^3$ ,  $\tilde{\chi}_0^4$  and  $\tilde{\chi}_2^\pm$ , are mostly higgsino. The production channels including these higgsino like charginos and neutralinos are suppressed because there are no  $\tilde{H}^+ \tilde{H}^- (W/Z)$  or  $\tilde{H}^0 \tilde{H}^0 (W/Z)$  vertex couplings [13]. They are also heavy, see table 5.1, which suppress the production further.

The squarks and gluinos have small production cross-sections because of their high mass. The dominant one is the stop pair production, which represents the dominant contribution to our signal from non-lepton productions. This is the reason why obtaining a higher Higgs mass by increasing the stop mass would lead to a weaker signal.

a/	$\tilde{\ell}_L^+ \tilde{\ell}_L^-$	$\tilde{\nu}_\ell \tilde{\nu}_\ell^*$	$\tilde{\ell}_L^+ \tilde{\nu}_\ell$	$\tilde{\ell}_L^- \tilde{\nu}_\ell^*$	$\tilde{\tau}_1^+ \tilde{\tau}_1^-$	$\tilde{\tau}_1^+ \tilde{\nu}_\tau$	$\tilde{\tau}_1^- \tilde{\nu}_\tau^*$	$\tilde{\nu}_\tau \tilde{\nu}_\tau^*$
$\sigma$ [fb]	1.52	1.596	3.818	1.702	13.68	33.8	17.31	15.80
b/	$\tilde{\chi}_1^0 \tilde{\chi}_1^0$	$\tilde{\chi}_1^0 \tilde{\chi}_1^-$	$\tilde{\chi}_1^0 \tilde{\chi}_1^+$	$\tilde{\chi}_2^0 \tilde{\chi}_1^-$	$\tilde{\chi}_2^0 \tilde{\chi}_1^+$	$\tilde{\chi}_2^0 \tilde{\chi}_2^0$	$\tilde{\chi}_1^- \tilde{\chi}_1^+$	
$\sigma$ [fb]	0.017	0.018	0.019	0.11	0.322	0.008	0.220	
c/	$\tilde{q} \tilde{q}$	$\tilde{t}_1 \tilde{t}_1^*$	$\tilde{g} \tilde{q}$	$\tilde{g} \tilde{g}$	$\tilde{\chi}_1^0 \tilde{q}$	$\tilde{\chi}_2^0 \tilde{q}$	$\tilde{\chi}_1^+ \tilde{q}$	$\tilde{\chi}_1^- \tilde{q}$
$\sigma$ [fb]	0.284	1.226	0.354	0.004	0.021	0.037	0.328	0.300

Table 6.1: The production cross-sections in fb using  $\sqrt{s} = 13$  TeV and the mass spectrum in table 5.1 for a/ slepton pairs, b/ chargino and neutralino pair, and c/ squarks and gluino production at the LHC. Calculations were done in *Herwig 7.1.1* [125]. Here  $\tilde{q}$  represents the sum over the squarks  $\tilde{u} + \tilde{d} + \tilde{s} + \tilde{c} + \tilde{b}$ , while  $\tilde{\ell}$  can be either  $\tilde{e}$  or  $\tilde{\mu}$ .

<sup>3</sup>This is found in the SLHA file from SPheno.

## 6.2 Leptonic signatures

In this tau-sneutrino NLSP scenario the squarks, neutralino and charginos are heavy. This makes the production cross-section of these particles small compared to the sleptons, which are the lightest particles. Sleptons do not produce high energy jets which is often used as a signature of supersymmetry [126]. Thus, we look at lepton signatures in this thesis.

Considering the production cross-sections above, we combine this with the decay modes and their branching ratios to find possible signatures. When choosing a signature, the amount of background from the SM must be taken into account. In this part we only consider background processes that have the same final state signal and not those that only have the same signal at detector level. By detector level, we mean the signal measured in the detector after the detector cuts, see section 7.3.1. At this level signals miss-identification of particles also have to be considered. This is discussed later in this chapter.

Before we proceed, we need to clarify what a detector can measure. When measuring a leptonic signal, the only detectable particles are electrons and muons. The neutrinos are not detectable and the taus decay before they reach the detectors and therefore the decay products have to be measured instead. The taus can decay to leptons or quark pairs, we call the taus decaying to lepton pairs for leptonic taus and taus decaying to quark pairs for hadronic taus. This is important when discussing the signature, because the signatures with hadronic taus have much less background than with leptonic taus [8]. The detection efficiency is smaller for hadronic taus with about 50% [127], while for muons it is about 95% [128] and 90% for electrons [129], depending on the transverse momentum cuts.

### 6.2.1 Sleptons

The slepton production cross-sections in table 6.1 show that the stau pairs have the dominating cross-section. The decay modes of the lightest stau are

$$\tilde{\tau}_1^- \rightarrow \tilde{\nu}_\tau^* q \bar{q}, \quad \tilde{\nu}_\tau \ell^- \nu_\ell, \quad \tilde{\nu}_\tau \tau^- \bar{\nu}_\tau. \quad (6.2.1)$$

The signal from quarks has a large background in many QCD productions. Thus, we look at the two-lepton signal from staus, i.e. a stau pair production channel where each

stau decays to leptons and a sneutrino. The SM background for this signal is coming from a single  $Z$  boson production,  $ZZ$  production and from  $W^+W^-$  production, which all have large cross-sections compared to the supersymmetry signal [130–132]. For instance, the  $Z$  boson can decay as

$$Z \rightarrow \ell^+\ell^-, \quad (6.2.2)$$

which leave the same signature in a detector as a stau pair decaying to  $\tilde{\nu}_\tau\ell\nu_\ell$  because the neutrinos and the sneutrinos are not detectable. Even if the two-lepton supersymmetry signals get a contribution from all the supersymmetry productions channels, the signal would be too small.

Since we cannot use the stau productions channel, the first and second generation slepton productions must be considered. The largest production cross-section of these is the  $\tilde{\ell}_L^\pm\tilde{\nu}_\ell$ . The possible signals are the one and two lepton signals, which both have too much background, and the trilepton signals. One of the trilepton signals comes from the decay

$$\tilde{\ell}^+ \rightarrow \tilde{\nu}_\tau\tau^+\nu_\ell \text{ and } \tilde{\nu}_\ell \rightarrow \tilde{\nu}_\tau\ell^-\tau^+. \quad (6.2.3)$$

This is the same sign version of the trilepton signal and has much less total branching ratio than the opposite sign version. The opposite sign version is from

$$\tilde{\ell}^- \rightarrow \tilde{\tau}_1^+\tau^-\ell^- \text{ and } \tilde{\nu}_\ell \rightarrow \tilde{\nu}_\tau^*\nu_\ell\nu_\tau. \quad (6.2.4)$$

In the latter case, the stau decay as  $\tilde{\tau}_1^+ \rightarrow \tilde{\nu}_\tau\ell^-\nu_\ell, \tilde{\nu}_\tau\tau^-\bar{\nu}_\tau$ . The detectable signals from these are  $\tau^\pm\tau^\pm\ell^\mp$  for same sign taus and  $\tau^\pm\tau^\mp\ell^\mp$  for opposite sign taus. The main background of these signals is the  $W^\pm Z$  production. The opposite sign background occurs when the  $Z$ -boson decay  $Z \rightarrow \tau^+\tau^-$  and the  $W$ -boson decay  $W^\pm \rightarrow \ell^\pm\nu_\ell$ . This branching ratio is larger than the same sign background where both the  $Z$ - and  $W$ -boson decay to taus and then the last of the three taus  $\tau^\pm\tau^\pm\tau^\mp$ , decays to a lepton. For the dominant slepton production, the same sign signal is the preferred one. The Feynman diagram of this signal is drawn in fig. 6.1

The  $\tilde{\ell}_L^+\tilde{\ell}_L^-$  and  $\tilde{\nu}_\ell\tilde{\nu}_\ell^*$  production channels also have cross-sections that are higher than the chargino and neutralino pairs, and the squarks-gluino pairs. Both contribute to the trilepton signal discussed above. The branching ratio for the decay mode  $\tilde{\ell}^- \rightarrow \tilde{\tau}_1^+\tau^-\ell^-$  is significantly higher for the parameter point utilized in this thesis compared to the one from [8]. The trilepton signal is  $\tilde{\ell}^+ \rightarrow \tilde{\tau}_1^+\tau^-\ell^+$  followed by  $\tilde{\tau}_1^+ \rightarrow \tilde{\nu}_\tau^*q\bar{q}$  and

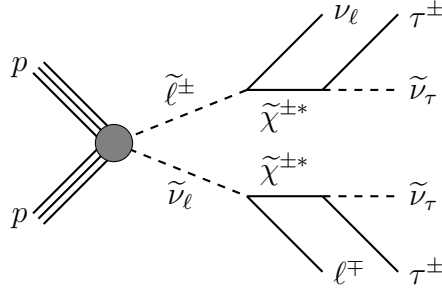


Figure 6.1: Example of a trilepton signal from a nslepton-sneutrino production

$\tilde{\ell}^- \rightarrow$	$\tilde{\nu}_\tau^* \tau^- \nu_\ell$	$\tilde{\tau}_1^+ \tau^- \ell^-$	$\tilde{\tau}_1^- \tau^+ \ell^-$	$\tilde{\nu}_\tau^* \ell^- \nu_\tau$	$\tilde{\nu}_\tau \ell^- \bar{\nu}_\tau$	$\tilde{\nu}_\ell q \bar{q}$	$\Gamma$ [GeV]
BR [%]	57.7	34.8	5.9	0.6	0.4	0.4	$1.7 \times 10^{-5}$

$\tilde{\nu}_\ell \rightarrow$	$\tilde{\nu}_\tau^* \nu_\ell \nu_\tau, \tilde{\nu}_\tau \nu_\ell \bar{\nu}_\tau$	$\tilde{\tau}_1^+ \ell^- \nu_\tau$	$\tilde{\nu}_\tau \ell^- \tau^+$	$\tilde{\tau}_1^\pm \tau^\mp \nu_\ell$	$\Gamma$ [GeV]
BR [%]	53.2	41.5	4.5	0.8	$1.3 \times 10^{-5}$

$\tilde{\tau}_1^- \rightarrow$	$\tilde{\nu}_\tau^* q \bar{q}$	$\tilde{\nu}_\tau \ell^- \nu_\ell$	$\tilde{\nu}_\tau \tau^- \bar{\nu}_\tau$	$\Gamma$ [GeV]
BR [%]	66.7	23.1	10.1	$2.87 \times 10^{-7}$

Table 6.2: Decays and total width of the sleptons, for the mass spectrum in section 5.3. Only decays with branching ratio  $\geq 0.1\%$  are included. When noting  $\pm$  as the charge of a particle in a decay mode it means the sum of the charged and charged conjugate version.  $\ell$  is either e or  $\mu$ .

$\tilde{\ell}^- \rightarrow \tilde{\nu}_\tau^* \tau^- \nu_\ell$ . The detectable signal from this would be  $\tau^- \tau^- \ell^+ + j$ , where  $j$  is a jet produced by the  $q\bar{q}$  pair.

## 6.2.2 Chargino and neutralino

The chargino and neutralino production pairs have a small cross-section and therefore a small contribution to the total signal. The trilepton signal discussed above can be produced by  $\tilde{\chi}_2^0 \tilde{\chi}_1^\pm$ , which is the dominant production pair. One of the possible decay chains of this production is shown in fig. 6.2. The two other decay chains are the chargino decaying  $\tilde{\chi}_1^\pm \rightarrow \tilde{\nu}_\tau \tau^\pm$  and the neutralino decaying either as (a)  $\tilde{\chi}_2^0 \rightarrow \tilde{\tau}_1^\pm \tau^\mp$  followed by  $\tilde{\tau}_1^\pm \rightarrow \tilde{\nu}_\tau + \nu_\ell + \ell'$  or (b)  $\tilde{\chi}_2^0 \rightarrow \tilde{\nu}_\ell \nu_\ell$  followed by  $\tilde{\nu}_\ell \rightarrow \tilde{\nu}_\tau + \tau^\pm + \ell^\mp$ .

## 6.2.3 Squarks and gluino

All productions, except for the  $\tilde{t}_1 \tilde{t}_1^*$  production, have low cross-sections. Thus, only the stop pair is considered when deciding on a signal. The dominant decay of the stop

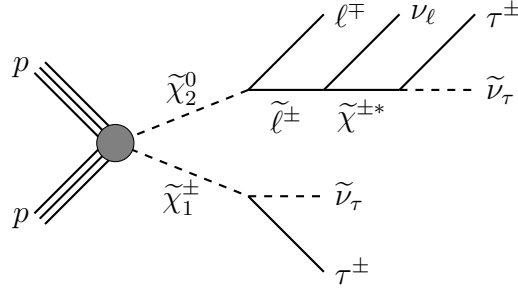


Figure 6.2: Example of a trilepton signal from a neutralino and chargino production

$\tilde{\chi}_{1,2}^0 \rightarrow$	$\tilde{\nu}_\tau \bar{\nu}_\tau + \text{c.c.}$	$\tilde{\tau}_1^\pm \tau^\mp$	$\tilde{\nu}_\ell \ell^\pm$	$\tilde{\ell}_L^\pm \ell^\mp$	$\Gamma$ [GeV]
BR ( $\tilde{\chi}_1^0$ ) [%]	29.1	27.8	22.6	20.4	1.72
BR ( $\tilde{\chi}_2^0$ ) [%]	19.0	19.0	31.1	30.9	1.72

$\tilde{\chi}_1^\pm \rightarrow$	$\tilde{\nu}_\tau \tau^\pm$	$\tilde{\tau}_1^\pm \nu_\tau$	$\tilde{\nu}_\ell \ell^\pm$	$\tilde{\ell}_L^\pm \nu_\ell$	$\Gamma$ [GeV]
BR [%]	19.1	18.8	31.3	30.7	19.0

Table 6.3: Decays and the total widths of the first and second neutralinos and the first chargino. Only decays with branching ratio  $\geq 0.1\%$  are included. When noting  $\pm$  as the charge of a particle in a decay mode it means the sum of the charged and charged conjugate version.  $\ell$  is either e or  $\mu$ .

is  $\tilde{t}_1 \rightarrow \tilde{\chi}_1^0 t$  with a branching ratio of 98%. This indicates that we are looking for a signal that is coming from  $\tilde{\chi}_1^0 \tilde{\chi}_1^0$ . The dominant of the trilepton signals from the two neutralinos is by combining  $\tilde{\chi} \rightarrow \tilde{\tau}_1^\pm \tau^\mp$  followed by  $\tilde{\tau}_1^\pm \rightarrow \tilde{\nu}_\tau q \bar{q}$  and  $\tilde{\chi}_1^0 \rightarrow \tilde{\ell}_L^\pm \ell^\mp$  followed by  $\tilde{\ell}_L^\pm \rightarrow \tilde{\nu}_\tau^* \tau^\pm \nu_\ell$ .

### 6.3 Trilepton signal and backgrounds

As shown in the previous section, the trilepton signature with a pair of same sign taus is particularly interesting. The signal is  $\tau^\pm \tau^\pm (e, \mu)^\mp + nj + \cancel{E}_T$  where  $n$  is the number of jets ( $n = 0, 1, 2, \dots$ ) and  $\cancel{E}_T$  is the missing transverse energy<sup>4</sup>. As discussed in section 6.2 the taus are not measured. When the taus decay, the branching ratio for hadrons is 64%. The rest is lepton to either a muon and muon-neutrino or electron and electron neutrino (including a tau-neutrino). This implies that the branching ratio for hadronic jets is higher than considering leptonic taus. We chose the hadronic tau

<sup>4</sup>Transverse energy is the energy deposited perpendicular to the collider beam axis, the missing transverse energy is the energy which we do not measure, but by conservation of momentum know is there.

signature in this thesis, because the background is smaller and the hadronic tau decay branching ratio is higher.

### 6.3.1 Background

The main backgrounds for the case of  $\tau_h^+ \tau_h^+ \ell^-$  signature are the  $W^+ Z \rightarrow \tau^+ \nu_\tau \tau_h^+ \tau_h^-$ , where the two same sign taus decay hadronically, while the opposite sign decays leptonically, and the  $W^+ W^- b\bar{b} \rightarrow \tau_h^+ \nu_\tau \ell^- \bar{\nu}_\ell b\bar{b}$  where the  $b\bar{b}$  decays to a hadronic tau plus extra jets. Note that the  $t\bar{t}$  and  $tW^+$  backgrounds are both covered by the  $W^+ W^- b\bar{b}$  background. The other backgrounds contributing to the signal are  $W^+ W^+ W^- \rightarrow \tau_h^+ \nu_\tau \tau_h^+ \nu_\tau \ell^- \bar{\nu}_\ell$  and, at detector level,  $ZZ$ ,  $t\bar{t}W^+$ ,  $t\bar{t}Z$ ,  $W^+ W^- Z$ . Processes which are backgrounds at detector level do not necessarily produce the signal we are looking for, but it mimics it if one or two particles are miss-identified or do not pass the detector cuts<sup>5</sup>. One example is the  $ZZ \rightarrow \tau^+ \tau^- \tau^+ \tau^-$  background, which is only a background if one of the taus, either as a hadronic tau or a lepton, do not pass the detector cuts or is not detected. Then the final state can be  $ZZ \rightarrow \tau_h^+ \cancel{\tau}^+ \tau_h^+ \tau^- [\rightarrow \ell^- \bar{\nu}_\ell]$ , where the slashed tau ( $\cancel{\tau}$ ) represents the miss-identified  $\tau$  or a  $\tau$  which did not pass the detector cuts.

There are also charged conjugate versions of all these backgrounds in the case of the  $\tau_h^- \tau_h^- \ell^+$  signature. The backgrounds are denoted in the next chapters without the charge of the  $W$  bosons, this means that  $W^+ Z$  and  $W^- Z$  are denoted as  $WZ$  where both processes are included.

---

<sup>5</sup>The detector cuts are explained in the next chapter





# Chapter 7

## Monte carlo simulation

### 7.1 Structure of an event

We start this chapter by explaining the key components in an event generator. The simulation of a proton-proton collision is separated into steps by which the event generator build up the structure of an event<sup>1</sup>. The following primary phases of the process that need to be simulated;

- a primary hard subprocess,
- parton (quark or gluon) showers associated with the incoming and outgoing particles in the subprocess,
- secondary interactions that give rise to the underlying event,
- hadronization of the colored particles,
- and the decay of unstable particles that do not escape the detector.

In all these steps there are also higher order corrections, which makes new steps necessary (e.g. combining the matrix element and the parton shower at higher orders [133]).

Before these steps are described in more detail, some notes about the detectors in the particles colliders are needed. The protons collide in the middle of the detectors

---

<sup>1</sup>For a more in dept description see [133].

which then measures as many particles as possible that comes out of the collision. The collider detectors are cylindrical about the beam axis. We define the transverse direction, to be perpendicular to the beam axis. The momentum of a particle in this direction is called the transverse momentum ( $p_T$ ). In this transverse plain we define the azimuthal angle  $\phi$ . It is useful to define the angle between the beam and the transverse axis in terms of the rapidity;

$$\eta \equiv -\ln\left(\tan\frac{\theta}{2}\right), \quad (7.1.1)$$

where  $\theta$  is the polar angle with respect to the beam axis. The detector coverage is defined in the  $\eta$  range, i.e. how large  $\eta$  the detector can measure.

### 7.1.1 Hard subprocess

The interesting events involve a large momentum transferer to produce heavy particles or jets with high momentum. Thus, the simulation of subprocesses with large invariant momentum transfer is the core of any simulation of collider events. Cross-section for scattering subprocess  $ab \rightarrow n$  at hadron colliders is [134]

$$\sigma = \sum_{a,b} \int_0^1 dx_a dx_b \int f_a^{h_1}(x_a, \mu_F) f_a^{h_2}(x_b, \mu_F) d\hat{\sigma}_{ab \rightarrow n}(\mu_F, \mu_R) \quad (7.1.2)$$

where  $f_a^h(x, \mu_F)$  is a *parton distribution function* (PDF). It depends on the momentum fraction  $x$  of particle  $a$  with respect to its parent hadron  $h$  and on the factorization scale  $\mu_F$ .  $\hat{\sigma}_{ab \rightarrow n}$  is the parton level cross-section of the final state  $n$  coming from the initial particles  $a$  and  $b$ , which depends on the factorization and renormalization scale  $\mu_F$  and  $\mu_R$ . We integrate over the momentum fractions  $x_a$  and  $x_b$  to get the total cross section. The part of the process where this equation applies is sketched in red in fig. [7.1].

This equation holds to all orders in perturbation theory [133]. However, when the cross-section from eq. (7.1.2) is computed beyond tree level, i.e. at *next-to leading order*(NLO) or *next-to-next-to leading order*(NNLO), two additional correcting terms are added to the calculation. The two new terms are a virtual- and a real-emission term, which are particles emitted from the hard process [135]. These extra particles make the combination with the parton shower more cumbersome.

To the lowest order, the function  $f_i(x, \mu_F)$  describes the probability of finding a parton

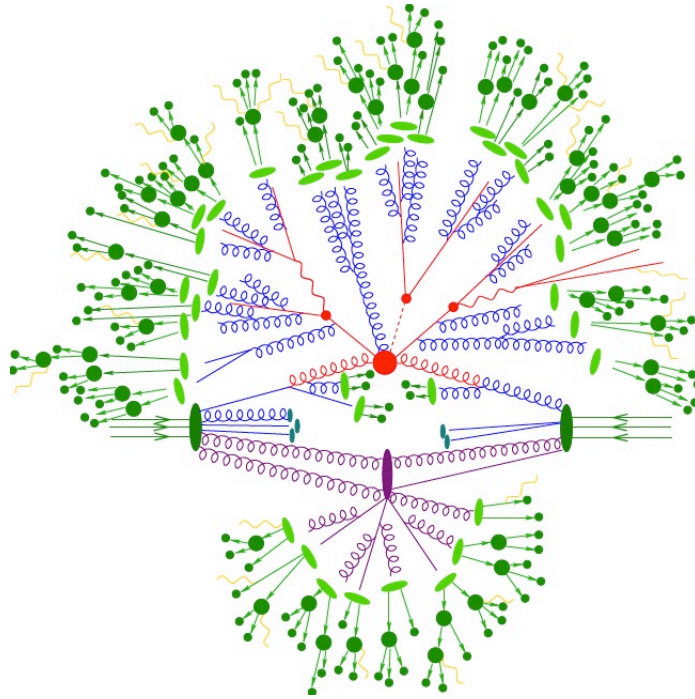


Figure 7.1: Sketch of a hadron-hadron collision as simulated by a Monte-Carlo event generator. The hard subprocess is shown in red, parton showers produced by Bremsstrahlung in blue, a secondary hard subprocess interaction in purple and the hadronization in green. Figure taken from [138].

(quark or gluon) of species  $i$  with the momentum fraction  $x$  when a proton is probed at the scale  $\mu_F$ . This function cannot be predicted from first principles because it depends on the proton wave function, which is non-perturbative, i.e. perturbative QCD cannot be used [136]. The factorization scale corresponds to the resolution in which the hadron is being probed [137].

The parton level cross-section  $\hat{\sigma}_{ab \rightarrow n}(\mu_F, \mu_R)$  is dependent on both the factorization scale and the renormalization scale. The renormalization scale is the same as the reference mass in the renormalization group equations. These scales are not physical parameters, but appear when calculating at a finite order in perturbation theory<sup>2</sup>. One usually defines the renormalization and factorization scales as  $\mu_F = \mu_R = Q^2$ . For s-channel resonances of mass  $M$ , the common choice of scale is;  $Q^2 = M^2$  and for the production of massless particles with transverse momentum  $p_T$  a common choice is;  $Q^2 = p_T^2$ .

<sup>2</sup>See [1] part 2, for more detail

### 7.1.2 Parton Shower

In all interactions, the scattered, annihilated or created partons can radiate gluons. These gluons are themselves colored, which give rise to further gluon radiation and can also produce quark-antiquark pairs. The quark-antiquark pairs and the radiating gluons generate showers of outgoing partons. Parton shower algorithms simulate this evolution from the scale of the hard subprocess to the infrared scale where non-perturbative confinement effects sets in and the hadronization process begins [138]. This part of the simulated process is sketched with blue gluon lines in fig. 7.1.

The event generator has to merge the parton shower with the hard subprocess. They cannot be blindly combined because the tree level matrix element is calculated to a fixed order in perturbation theory and the parton shower is approximated to all orders in perturbation theory [133]. There are two different ways to solve this; 1) matching the parton shower and the hard subprocess together by modifying one or both of them to fit the other [139] and 2) utilizing a merging algorithm [140, 141].

### 7.1.3 Hadronization

As the event is evolved downward in momentum scales, it ultimately reaches the scale where QCD becomes strongly interactive and perturbation theory breaks down. At this scale, the parton evolution models must be replaced by a non-perturbatively hadronization model (green lines in fig. 7.1). This describes the confinement of partons in colorless hadrons. The two most commonly used models are the string [142] and cluster model [143]. The hadronic and tau decays are simulated at the same stage in the simulation.

### 7.1.4 Jet algorithms

Jets can be defined as a collimated spray of particles arising from the fragmentation and hadronization of a quark or gluon after a collision. In a detector, only colorless hadrons are measured. Reconstruction algorithms are applied on these measured hadrons to combine them into jets. These jets can be used in the reconstruction of the event, in the same way as other measured particles. Two main classes of jet algorithms are in use, i.e. cone algorithms [144] and sequential algorithms [145, 146]. In this thesis, a sequential algorithm is utilized, namely the anti- $k_T$  algorithm [73]. The sequential clustering algorithm work by finding the minimum separation of particles in an area

using the formula

$$d_{ij} = \min(p_{T,i}^a, p_{T,j}^a) \times \frac{R_{ij}^2}{R} \quad (7.1.3)$$

where  $R_{ij}$  is the radial distance between particle  $i$  and  $j$ ,  $R$  is the cone radius parameter which determines the final size of the jet and  $a = -2$  in the anti- $k_T$  algorithm. The distance is calculated using  $R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$ , where  $\eta$  is the rapidity and  $\phi$  is the azimuthal angle of the particle. The two particles are then combined into one particle ( $ij$ ) using summation of four vectors and removing  $i$  and  $j$  from the list of particles. The algorithm does the same thing over and over again until the minimum of  $d_{ij}$  is greater than  $d_{iB} = p_{T,i}^a$ , where  $p_{T,i}^a$  is the transverse momentum of particle  $i$  which is now a final jet. This algorithm can be repeated until all hadrons are clustered into jets with cone radius  $R$ . The cone radius of a jet is usually set to  $0.4 - 0.7$ , i.e. the angular distance from the center of the cone to the boundary.

## 7.2 Event Generators

In this thesis the multi-purpose Monte Carlo event generators Sherpa 2.2.4 [147–151] and Herwig 7.1.1 [125, 152] have been utilized. They both use the principles described above and many more to make accurate predictions in colliders. Since simulating collision events is highly complex, they both rely on third-party tools created for specific parts of the event generation.

All simulations are done at  $\sqrt{s} = 13$  TeV, with the same factorization and renormalization scale. The supersymmetric processes simulated in Herwig uses the default settings with the MMHT2014 PDF [153] and a scale that is the invariant mass of the production pair. For the background processes the NNPDF [154] PDF is utilized and the scales are for most of them set to the invariant mass of the production particles. The hard sub-processes in Sherpa is calculated at *next-to leading order* (NLO) for all but the  $t\bar{t}Z$  and  $t\bar{t}W$  which are calculated at *leading order* (LO)<sup>3</sup>. The matrix element calculations are obtained from Amegic [149] for the tree level calculation and OpenLoops [155] to estimate the one-loop corrections. The matching of the hard sub-process to the parton shower is carried out by the matching algorithm from MC@NLO [156]. Additionally, the extra jets, calculated at NLO, are merged with the hard processes using the MEPS@NLO method at NLO [151]. The number of extra jets is set to the maximal number of jets

<sup>3</sup>The reason  $t\bar{t}Z$  and  $t\bar{t}W$  is calculated at LO is because this is the highest order available in Sherpa.

Process	Program	PDF	Order	Scale ( $\mu_R = \mu_F$ )
Signals	Herwig 7	MMHT2014	LO	$M_{\tilde{p}_1\tilde{p}_2}$
$W^\pm Z$	Sherpa	NNPDF3.0	NLO	$\frac{1}{2}(m_Z + m_W)$
$ZZ$	Sherpa	NNPDF3.0	NLO	$m_Z$
$WWbb$	Sherpa	NNPDF3.0	NLO	$M_{WW}$
$t\bar{t}W$	Sherpa	NNPDF3.0	LO	$M_{t\bar{t}W}$
$t\bar{t}Z$	Sherpa	NNPDF3.0	LO	$M_{t\bar{t}Z}$
$WWW$	Sherpa	NNPDF2.3	NLO	$\frac{3}{2}M_{W_1}$
$ZWW$	Sherpa	NNPDF3.0	NLO	$M_{ZWW}$

Table 7.1: Signal and Background processes with corresponding event generator, scale and PDF used. The scale of the signals is set to the invariant mass of the production sparticles;  $M_{\tilde{p}_1\tilde{p}_2}$ .

available for the given process (e.g. the  $WZ$  production can produce up to 3 extra jets). A summary of the settings for each process are presented in table [7.1](#), where in the scale column, the capital  $M$  means invariant mass of the two subscript particles, e.g.  $M_{WW}$  is the invariant mass of the two  $W$  bosons produced in  $WWt\bar{t}$ .

### 7.2.1 Validating Simulations

In this section we validate and normalize the background simulations to published results for the various background processes;  $WZ$  [\[157\]](#),  $ZZ$  [\[131\]](#),  $t\bar{t}Z$  [\[158\]](#),  $t\bar{t}W$  [\[159\]](#),  $WWW$  [\[160\]](#) and  $ZWW$  [\[161\]](#). The  $WWb\bar{b}$  is missing because there are no published results we can use for normalization<sup>4</sup>. These published results are all calculations done at the highest order for the specific process to this date, i.e. there are no calculation of the same process carried out at a higher order. They are summarized in table [7.2](#) where the total cross-section is presented for the published results and the cross-section obtained from our simulations. Note that there are no similar studies of the supersymmetry signal to compare with in this parameter space region. Since the published results are in general calculated at a higher order, we can normalize our result to the higher order calculation to obtain the same order. This normalization is done for all result cross-sections in the proceeding chapters. Note that the same CM energy as in the compared reference is used in table [7.2](#).

<sup>4</sup>Studies on the  $WWb\bar{b}$  process have been restricted to only parts of momentum space and at  $\sqrt{s} = 7$  TeV, such as [\[162\]](#). This means that it can not be used for normalizing the result cross-sections.

Process	$\sqrt{s}$	Order	$\sigma_{tot}$ ref. [pb]	Our $\sigma_{tot}$ result [pb]
$W^\pm Z$	13 TeV	NNLO	$51_{-1}^{+1}$	$47 \pm 2$
$ZZ$	13 TeV	NNLO	$16.91_{-0.5}^{+0.4}$	$14.8 \pm 0.3$
$t\bar{t}W$	13 TeV	NLO	$0.52_{-0.01}^{+0.03}$	$0.462 \pm 0.005$
$t\bar{t}Z$	13 TeV	NLO	$0.78_{-0.07}^{+0.06}$	$0.716 \pm 0.004$
$WWW$	14 TeV	NLO	$0.21_{-0.02}^{+0.02}$	$0.245 \pm 0.003$
$ZWW$	14 TeV	NLO	0.228	$0.261 \pm 0.004$

Table 7.2: Comparing the total production cross-section ( $\sigma_{tot}$ ) of the referenced published results, with the given order and CM-energy ( $\sqrt{s}$ ), to our  $\sigma_{tot}$  results. The calculation of our  $\sigma_{tot}$  is done with the same CM-energy, PDF, renormalization scale and factorization scale as the references; [131, 157-161].

## 7.3 Analysis Scheme

The event analysis tool `Rivet 2.5.4` is used when analyzing the simulated events [163]. The advantage of Rivet is that it is compatible with both Sherpa and Herwig (i.e. it is possible to use the same analyses script for all processes). The main part of the analysis is to check if the event has the correct signal ( $\tau_h^\pm \tau_h^\pm \mu^\mp$ ) and check if it passed the cuts. If the event is accepted, all relevant features of the particles, including the first three jets, are saved to a text file to be used in training and testing of the neural network (chapter 8). The analysis scheme is the same for all processes (both signal and background).

### 7.3.1 Detector cuts

For the signal to be detectable, the particles need to pass some criterions, called detector cuts. These cuts are made based on what is possible to detect and what gives the highest efficiency of detection. The cuts that are utilized in this analysis are based on cuts from [8], but slightly modified to match recent supersymmetry searches in ATLAS [102, 164-166].

The muon have a transverse momentum cut of  $p_T > 10$  GeV and a rapidity of  $|\eta| < 2.5$ . These are both detector specific cuts, the first based on the detection efficiency of the muon and the second on the placement of the detectors. For the hadronic taus the cuts are  $p_T > 15$  GeV and  $|\eta| < 2.5$ . The jets are reconstructed using the anti- $k_T$  clustering algorithm [73] in `FastJet 3.2.2` [72]. The radius parameter is  $R = 0.4$



with a rapidity of  $|\eta| < 4.5$ .

To resolve reconstruction ambiguities, an overlap removal algorithm is applied to remove overlapping muons or jets. Jet candidates overlapping muons by  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.2$  are removed. Muon candidates overlapping jets by  $\Delta R < 0.4$  are also removed. These two overlapping removals has to be made, in this order, because the jet could be created by the muon, which could be the case if the muon is detected in the middle of the jet. If the muon is at the edge of the jet, it is unlikely that it created the jet. In this latter case it is more likely that the jet produced the muon in some decay or deexcitation. Any jets overlapping hadronic taus candidates with  $\Delta R < 0.2$  are excluded.

Below is a summary of how an event is selected as signal, applied in the given order:

1. Jet reconstruction according to anti- $k_T$  algorithm with  $R = 0.4$ ,  $p_T^j > 20$  GeV and  $|\eta_j| < 4.5$ .
2. Muons are required to have  $p_T^\mu > 10$  GeV,  $|\eta_\mu| < 2.5$ .
3. Hadronic taus are required to have  $p_T^{\tau_h} > 15$  GeV,  $|\eta_{\tau_h}| < 2.5$ .
4. Remove jets with  $\Delta R_{\mu,j} < 0.2$ .
5. Remove muons with  $\Delta R_{\mu,j} < 0.4$ .
6. Remove jets with  $\Delta R_{\tau_h,j} < 0.2$ .
7. Require  $N_\mu = 1$  and  $N_{\tau_h} = 2$ .
8. Require same sign taus and opposite sign muon;  $Q_{\tau_1} = Q_{\tau_2}$ ,  $Q_\mu \neq Q_{\tau_1}$ .
9. Muons and taus are required to be isolated with  $\Delta R_{\ell,\tau_h} > 0.4$  and  $\Delta R_{\tau_h,\tau_h} > 0.4$ .

## 7.4 Signal selection efficiency

In an event generator, the number of accepted events to the total number of SUSY events follows a physical ratio. This can be used to calculate the efficiency  $\epsilon$  of accepted events [\[167\]](#)

$$\epsilon = \frac{n}{N}, \quad (7.4.1)$$

Process ( $\tau_h^\pm \tau_h^\pm \mu^\mp$ )	$N_{\text{Det}}$	$\sigma_{\text{Det}}$ [fb]	$\epsilon_{\text{Det}}$
SUSY EW	172260	$(8.18 \pm 0.02) \times 10^{-2}$	$0.0044 \pm 0.0001$
SUSY QCD	337134	$(2.747 \pm 0.001) \times 10^{-2}$	$0.0072 \pm 0.0001$
SUSY $\chi\chi$	241896	$(5.22 \pm 0.01) \times 10^{-3}$	$0.0135 \pm 0.0001$
$WZ$	393806	$3.09 \pm 0.15$	$0.0197 \pm 0.0003$
$ZZ$	613308	$0.508 \pm 0.013$	$0.0307 \pm 0.0004$
$WWb\bar{b}$	413868	$45.19 \pm 0.27$	$0.0023 \pm 0.0001$
$t\bar{t}Z$	38659	$0.418 \pm 0.006$	$0.00062 \pm 0.00002$
$t\bar{t}W$	210639	$0.181 \pm 0.001$	$0.0105 \pm 0.0002$
$WWW$	157833	$0.072 \pm 0.001$	$0.0079 \pm 0.0002$
$ZWW$	667933	$0.026 \pm 0.001$	$0.0334 \pm 0.0004$

Table 7.3: The number of accepted events after the detector cut, the detector cut cross sections and efficiency for the signal and background processes at  $\sqrt{s} = 13\text{TeV}$  for the  $\tau_h^\pm \tau_h^\pm \mu^\mp$  signature. See background process definition in section [6.3.1](#).

where  $n$  is the number of accepted events and  $N$  is the number of tried events. The uncertainty is [\[167\]](#)

$$\Delta\epsilon = \frac{\sqrt{n(1-n/N)}}{N}. \quad (7.4.2)$$

## 7.5 Simulations

The cross-section for the signal and background processes are presented in table [7.3](#), together with the efficiency of the detector cuts. Note that the detector cut selects the signal ( $\tau^\pm \tau^\pm \mu^\mp$ ) and cut on the candidate particle observables. This means that the event generator cross-section includes production of signatures that are irrelevant for this analysis. Since there are some differences in measuring muons and electrons, the detector cuts should be slightly different. The simulations below are for the  $\tau_h^\pm \tau_h^\pm \mu^\mp$  signal, which is used for the rest for this thesis, because the muons have a higher efficiency.

The signal process is divided into three groups; SUSY EW, SUSY  $\chi\chi$  and SUSY QCD. This is the same naming convention as in Figy et al. [\[8\]](#). The SUSY EW signal consists of the slepton production channels, i.e.  $\tilde{\ell}^+ \tilde{\ell}^-$ ,  $\tilde{\ell}^\pm \tilde{\nu}_\ell$  and  $\tilde{\nu}_\ell \tilde{\nu}_\ell^*$ . The SUSY  $\chi\chi$  signal is the neutralino and chargino production channels, i.e.  $\tilde{\chi}_0^2 \tilde{\chi}_1^\pm$  and  $\tilde{\chi}_1^- \tilde{\chi}_1^+$ . The SUSY QCD signal is the squark and gluino production channels, i.e.  $\tilde{q}\tilde{q}$ ,  $\tilde{t}_1 \tilde{t}_1^*$ ,  $\tilde{g}\tilde{q}$  and  $\tilde{\chi}_1^\pm \tilde{q}$ .

The total signal cross-section is  $\sigma_{Det,S} = (11.45 \pm 0.02) \times 10^{-2}$  fb, where  $S$  represents all signals, while the total background cross-section is  $\sigma_{Det,B} = (49.485 \pm 0.31)$  fb, where  $B$  represents all backgrounds. The dominating background process is the  $WWb\bar{b}$  process, which has about 450 times larger cross-section than the total signal. As expected, the electroweak signal is the dominant SUSY signal. The SUSY QCD is still stronger than the SUSY  $\chi\chi$ , which is subdominant in this parameter point.

The contributions to the cross-sections uncertainty are the event generation uncertainty, the uncertainty from the normalization of the cross-sections to higher orders in perturbation theory and the efficiency uncertainty (eq. (7.4.2)). To find the proper uncertainty in the event generator, one has to vary the the renormalization and factorization scale. This requires long computation time and is therefore not performed in this thesis. However, this is included in the normalization cross-sections from section 7.2.1

## 7.6 Feature plots

### 7.6.1 Low-level features

The observable features of the signature are plotted in fig. 7.2. The blue, green and red distributions are the SUSY EW, SUSY  $\chi\chi$  and SUSY QCD signals, while the black distribution represents background processes. The distributions are all normalized to one for the fraction of events, i.e. the area under each distribution is one. The reason for not normalizing to the relative cross-sections are to make the plots more readable. Since the background cross-section is about 480 times higher then the signal cross-section, a plot where the the signal and background distributions are normalized relative to the cross-section would be dominated by the background. This would make it hard to compare the signal and background distribution. The contributions to the background distributions from each process are proportional to the cross-section.

The feature data used to plot the distributions in fig. 7.2 is after the detector cuts are applied. This means that these plots show if it is possible to find differences in the background and signal observables. The first row in fig. 7.2 is the transverse momentum of the signature particles  $(\mu, \tau_1, \tau_2)$ , where  $\tau_1$  and  $\tau_2$  are the hardest and second hardest hadronic tau, i.e.  $p_T^{\tau_1} > p_T^{\tau_2}$ . In the rest of this thesis we will always mean hadronic tau ( $\tau_h$ ) when we refer to tau ( $\tau$ ), unless explicitly stated otherwise. Due to higher mass

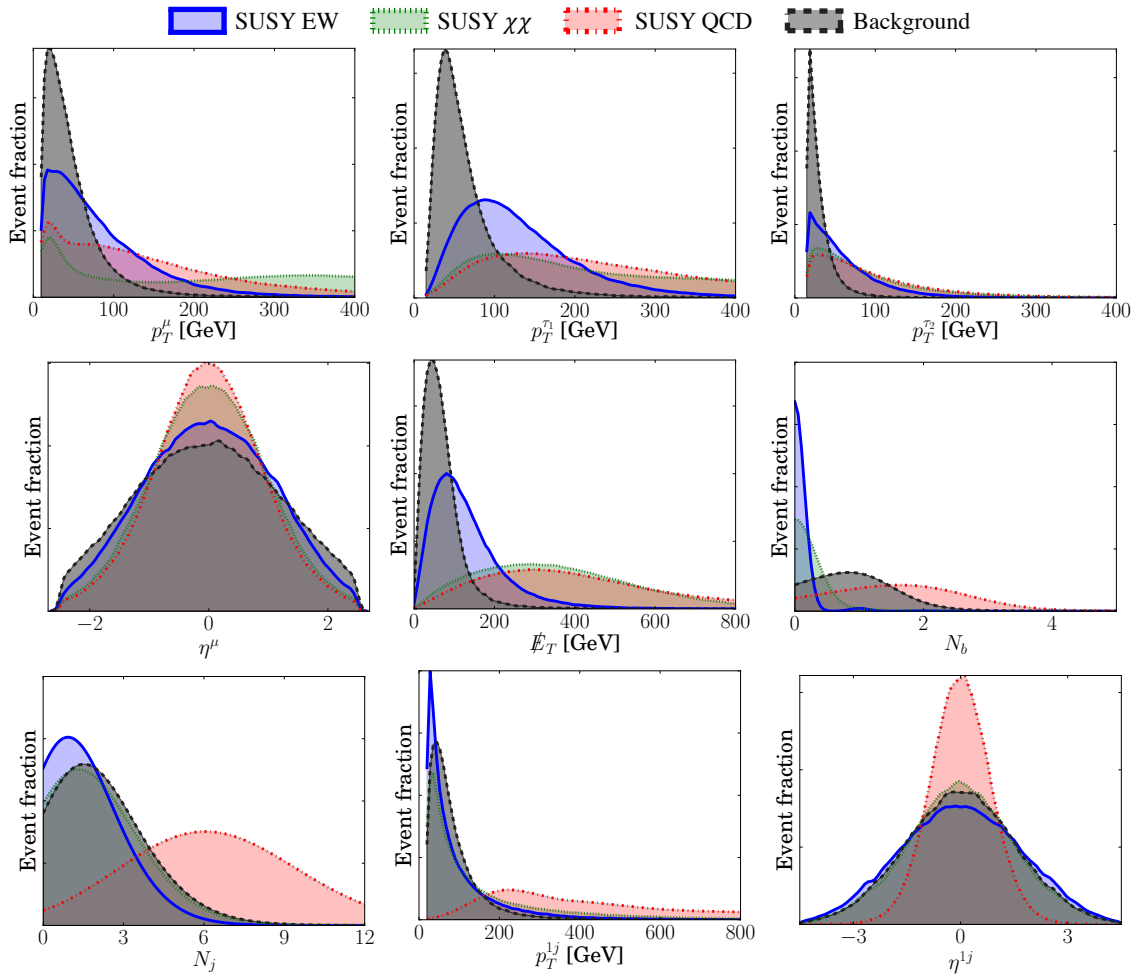


Figure 7.2: Selected low-level feature distributions using the parameter point found in section 5.3, for the signature particles ( $\mu$ ,  $\tau_1$ ,  $\tau_2$ ), the missing transverse energy  $\cancel{E}_T$  and the hardest jets  $1j$  ( $N_j$ : Number of jets,  $N_b$ : Number of b-tagged jets).

spectrum, the  $p_T$  distributions are higher for the supersymmetry signals. The next line in fig. 7.2 shows the rapidity distribution for the muon, the missing transverse momentum ( $\cancel{E}_T$ ) and the number of b-tagged jets. The missing transverse energy is often used in supersymmetry searches. This is because the LSP, or in this case NLSP, is not detectable. Additionally, the high mass of the tau-sneutrino ( $m_{\tilde{\nu}_\tau} = 195.7 \text{ GeV}$ ) results in a high  $\cancel{E}_T$ . The number of b-tagged jets is probably the distribution with the largest difference. The SUSY EW signal does for the most part not include b-tagged jets, whereas the SUSY QCD signal have two or more. The background has a peak at the number of b-tagged jets because of the  $WWb\bar{b}$ , which for the most part includes a b-tagged jet. The last line consists of the distribution for number of jets ( $N_j$ ) in addition the  $p_T$  and  $\eta$  of the hardest jet.

## 7.6.2 High-level features

In fig. [7.3](#) the invariant mass and the separation of the signature particles are plotted, i.e. some of the high-level features. The top row is the invariant mass of the signature particles. Notice in the  $M_{\tau_1\tau_2}$  there are no definite endpoint where the distribution decrease rapidly for the signals. This suggest that the taus originate from different decays, which is the case for the signal processes. For the  $M_{\tau_{1/2}\mu}$  there are a definite endpoint to the SUSY EW distribution. This is because the tau and muon comes from the same decay, i.e. one of the sleptons decay to one hadronic tau, one muon and one tau-sneutrino. The invariant mass peak for  $m_{\mu\tau_{1/2}}$  is at about 100 GeV, which do fit the mass spectrum with a sneutrino mass at  $m_{\tilde{\nu}} = 195.7$  GeV and the other slepton masses at about 350 GeV. The second row shows the radial difference between the signature particles, defined as in section [7.1.4](#). This could indicate how far away in the decay chain the two particles originate from. This does fit the SUSY EW in  $\Delta R^{\mu,\tau_{1/2}}$  because the muon and tau originate from the same 3-body decay, while in the SUSY  $\chi\chi$  the two does not originate from the same decay, see the Feynman diagrams in figs. [6.1](#) and [6.2](#).

To separate the signal and background, there must be some difference in feature distributions. This could be some unknown high-level feature that separates the distributions more than those shown in figs. [7.2](#) and [7.3](#). To find such high-level features, one has to be more creative in the combination of features. Since the machine learning algorithm devise its own high-level features, it can help simplify this process. In the next chapter, we will see how machine learning algorithms perform on the features presented in fig. [7.2](#).

## 7.7 Detectable signal

To evaluate if the signal is detectable in a collider, we calculate the integrated luminosity required for a  $5\sigma$  significance detection ( $\mathcal{L}_{\text{int}}(5\sigma)$ ). This number is also used to compare the performance of the machine learning algorithm compared to the traditional optimized cuts that are presented in the next section. The statistical significance  $Z$  is defined as [\[167\]](#)

$$Z = \sqrt{2 \left( (S + B) \ln \left( 1 + \frac{S}{B} \right) - S \right)}, \quad (7.7.1)$$

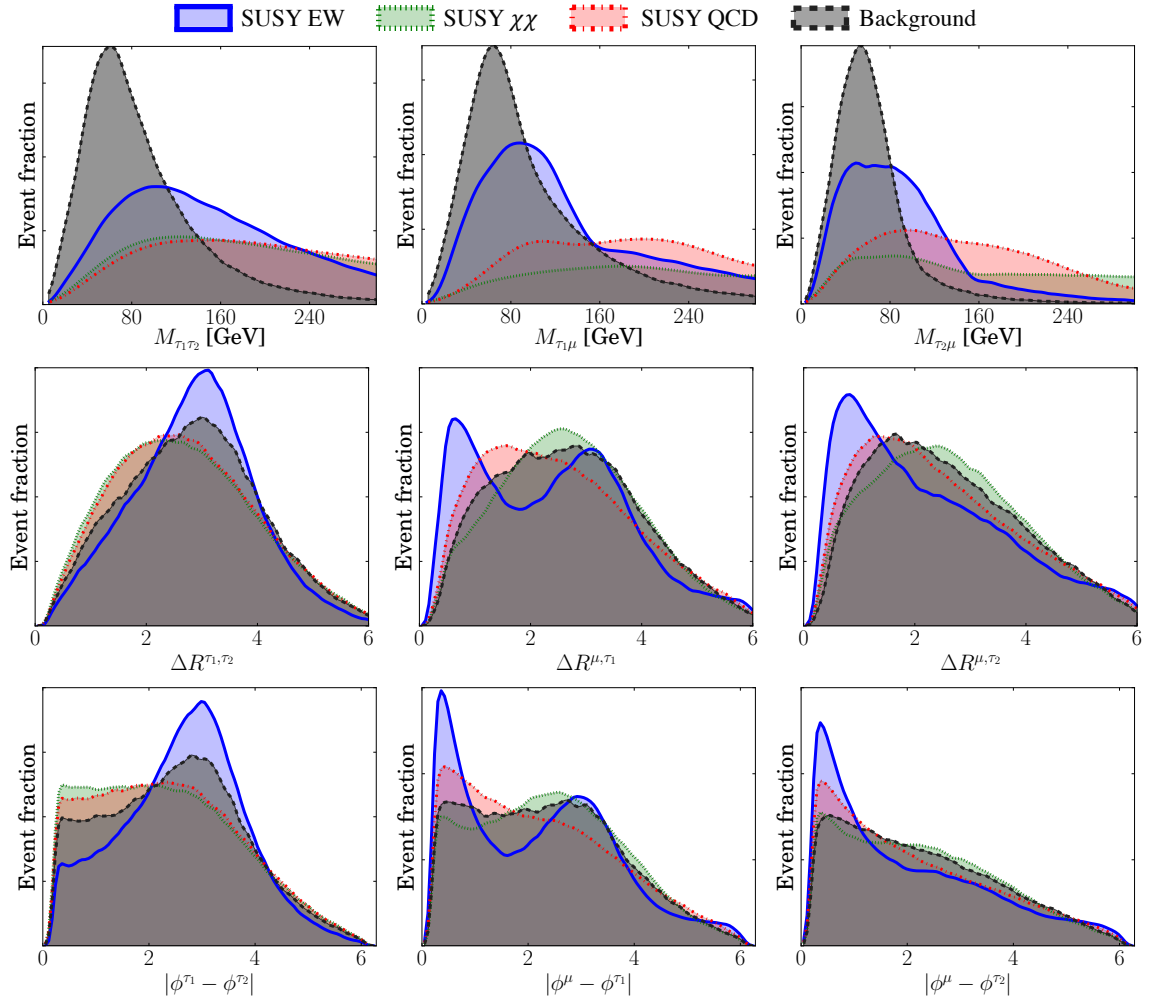


Figure 7.3: Selected high-level feature distributions (invariant mass, angular distance ( $\Delta R$ ) and azimuthal angle separation) using the parameter point found in section 5.3, for the signature particles ( $\mu$ ,  $\tau_1$ ,  $\tau_2$ ).

where  $S$  is the number of signal events and  $B$  is the number of background events. To make a discovery, the significance has to be at  $5\sigma$  ( $Z = 5$ ). This corresponds to a probability of the discovery to be a fluctuation in the background of  $5.7 \times 10^{-5}\%$  [167].

The number of signal events  $S$  can be translated into integrated luminosity  $\mathcal{L}_{\text{int}}$  and cross-section  $\sigma_S$  by  $S = \sigma_S \mathcal{L}_{\text{int}}$ . We also have to include the efficiency of detecting the signal  $S$  given it was produced. For the signal in this thesis ( $\tau_h^\pm \tau_h^\pm \mu^\mp$ ) the efficiency depends on the efficiency of simultaneously measuring two hadronic taus. This efficiency is taken from ATLAS [166] searches and is  $\epsilon_\tau = 0.5$  for each hadronic tau. Since there are two hadronic taus, the total efficiency is  $\epsilon_\tau^2$ . The number of measured signal events are

$$S = \sigma_S \mathcal{L}_{\text{int}} \epsilon_\tau^2. \quad (7.7.2)$$

The integrated luminosity can be translated into the number of collisions needed in the collider for the signal to be detectable. Using eqs. (7.7.1) and (7.7.2), we get a measure of the integrated luminosity needed for detection of the signal with signal cross-section  $\sigma_S$  and background cross-section  $\sigma_B$  at  $5\sigma$

$$\mathcal{L}_{\text{int}}(5\sigma) = \frac{5^2}{2((\sigma_S + \sigma_B) \ln(1 + \sigma_S/\sigma_B) - \sigma_S) \epsilon_\tau^2}. \quad (7.7.3)$$

For our analysis, with 3 signal processes and 7 background processes, the cross-section  $\sigma_{\text{Det},S}(\sigma_{\text{Det},B})$  is the sum of the signal (background) cross-sections. The calculation of cross-sections include multiplying by the cut efficiencies for the detector cut, i.e.

$$\sigma_{\text{Det},S} = \sum_i \sigma_{\text{gen},i} \epsilon_{\text{Det},i}, \quad (7.7.4)$$

where we sum over all signals,  $\sigma_{\text{gen},i}$  is the generation cross-section from the event generator and  $\epsilon_{\text{Det},i}$  is the detector cut efficiency for signal  $i$ . The efficiencies are calculated using eq. (7.4.1).

To evaluate the uncertainty in the integrated luminosity, the uncertainty from the event generators and efficiencies, eq. (7.4.2), are propagated through eq. (7.7.3). To propagate the uncertainty, the standard formula of propagation is used:

$$\Delta \mathcal{L}_{\text{int}} = \sqrt{\sum_X \left( \frac{\partial \mathcal{L}_{\text{int}}}{\partial X} \right)^2 (\sigma_X)^2}, \quad (7.7.5)$$

where  $X$  is one of the variables in  $\mathcal{L}$  with uncertainty  $\sigma_X$ .

## 7.8 Optimized cuts

To see if machine learning improves detection, we have to compare it with the traditional way of analysing events, i.e. utilize optimized cuts to reduce the background without removing too much signal. These optimized cuts are based on the difference in signal and background features. In this thesis we use the same optimized cuts as in Figy et al. [8]. The optimized cuts are separated in two versions; 1) favoring the electroweak supersymmetry signal and 2) favoring the supersymmetry QCD signal. The main difference between the two optimized cuts are how they cut on the jet observables. From fig. 7.2 we see that the electroweak signal generally has no high- $p_T$  jets, usually under approximately 200 GeV. For the supersymmetry QCD jets the  $p_T$  is higher. This makes it hard to construct cuts that separate both signals from the background. The background processes that include quark pairs in the production have many of the same features as the supersymmetry QCD signal. The same is true for the electroweak backgrounds and the supersymmetry electroweak signals.

It should be noted that since the two optimized cuts favors different supersymmetry signals, the information about the nature of the signal is different. How the slepton production channels behave represents the unique signature of a tau-sneutrino NLSP. If, for instance, the neutralino was the LSP the lighter production channels would not produce as many taus due to lepton flavour conservation. Thus, for the lighter production channels, e.g. in SUSY EW and SUSY  $\chi\chi$ , there would be a difference between tau-sneutrino NLSP and neutralino LSP. If we look at the SUSY QCD channel, there would not be much of a difference between the supersymmetry scenarios. This means that the first optimized cut, favoring SUSY EW, can only be used for detecting supersymmetry and determining what SUSY scenario it is. The second optimized cut, favoring SUSY QCD, can only be used to detect supersymmetry [8].

The optimized cut A (optA) is optimized for the electroweak signal and the cuts are as follows [8]:

1. Veto on b-jets and more then one jet, i.e.  $N_b = 0$  and  $N_j \leq 1$
2. Cut on the transverse momentum of the hardest jet in the event above 200 GeV, i.e. we require  $20 \text{ GeV} < p_T^{j_1} \leq 200 \text{ GeV}$



3. Require the minimum invariant mass<sup>5</sup> of a muon and tau to be less than 55 GeV, i.e.  $m_{\min}(\mu, \tau) = \min(M_{\mu, \tau_1}, M_{\mu, \tau_2}) \leq 55 \text{ GeV}$ .
4. Require the azimuthal angle between the muon and the missing transverse momentum to be greater than 1.5, i.e.  $|\phi_\mu - \phi_{\cancel{p}_T}| < 1.5$ .

Because of the new parameter point, which has a higher mass spectrum than [8], the optA cut is not optimized anymore. This can be solved by changing the limits of the cuts. However, this causes a problem with the invariant mass cut (nr. 3) because the invariant mass of the SUSY EW signal overlaps with the background, see fig. 7.3. Thus, the optA cut does not work for this parameter point. This optimized cut is compared to machine learning algorithms in chapter 9 by utilizing the parameter point from Figy et al. [8].

The optimized cut B (optB) is designed to promote the high- $p_T$  jets from the supersymmetry QCD signal. It cut on the number of jets and the sum of the transverse energy of the leptons and jets. If we compare cut number one, which require two jets or more, to the background feature distribution of number of jets ( $N_j$ ) in fig. 7.2, it does not match. The  $N_j$  background distribution has a peak at two jets and the cut is on  $N_j \geq 2$ . The change in number of jets can be caused by several factors; 1) evolution of event generators and higher order corrections, 2) the energy cut made in the MEPS@NLO merging, which could be stricter in [8] and 3) the use of  $R = 0.4$  instead of  $R = 0.35$  ( $D = 0.7$ ) in the anti- $k_T$  jet reconstruction algorithm.

Since the OptB cut should not be affected by the change in parameter point, it is used with our parameter point as a comparison. This cut is therefore optimized to the simulations conducted in this chapter. This is done by varying the required number of jets ( $N_j$ ) and the required hardest jet transverse momenta ( $p_T^{j_1}$ ) to find the optimal value, i.e. minimizing the integrated luminosity required for a detection significance of  $5\sigma$  ( $\mathcal{L}_{int}(5\sigma)$ ) eq. (7.7.3). This is shown in fig. 7.4, where the minimum is at  $N_j \geq 8$  and  $p_T^{j_1} \geq 250 \text{ GeV}$ . The optimized cut B (optB), after optimization, is as follows:

1. We require  $N_j \geq 8$  and  $p_T^{j_1} \geq 250 \text{ GeV}$ .
2. We require the sum of the transverse energy of the leptons ( $\mu$  and  $\tau_h$ ) and jets to be greater than 300 GeV, i.e.  $\sum_{i=\ell, j} E_T^i \geq 300 \text{ GeV}$ .

---

<sup>5</sup>The invariant mass from particle 1 and 2 with 4-momentum  $p_1$  and  $p_2$  is  $M_{12}^2 = (p_1 + p_2)^2 = (E_1 + E_2)^2 - |\mathbf{p}_1 - \mathbf{p}_2|^2$

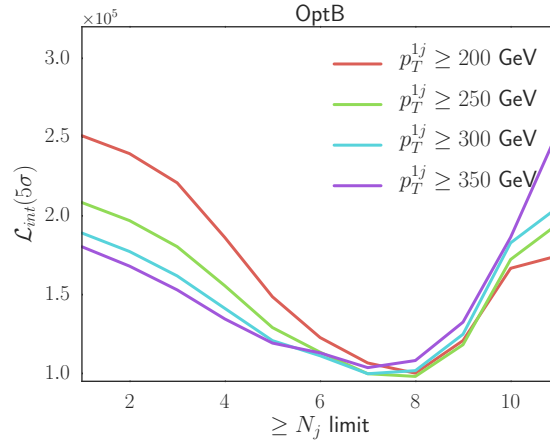


Figure 7.4: The optimization of OptB by varying the number of jets ( $N_j$ ) required and the required hardest jet transverse momentum ( $p_T^{1j}$ ).

One problem with this cut is that it focuses on the SUSY QCD signal, which in this parameter point is weaker than the dominant SUSY EW signal. The latter signal is suppressed because it overlaps with the background on the cut parameters. The efficiencies for the individual processes are shown in table 7.4, where the background is almost removed. However, the signal efficiency is not as high as in Figy et. al [8].

Process	$\sigma_{OptB} / 10^{-4}$ [fb]	$\epsilon_{OptB} / 10^{-2}$	$\epsilon_{OptB} / 10^{-2}$ (Figy et. al. [8])
SUSY EW	$0.06 \pm 0.02$	$0.8 \pm 0.2$	2.2
SUSY QCD	$60.2 \pm 0.2$	$21.9 \pm 0.07$	91.0
SUSY $\chi\chi$	$0.025 \pm 0.007$	$0.48 \pm 0.01$	20.9
$WZ$	$89.2 \pm 0.5$	$0.29 \pm 0.01$	3.5
$ZZ$	$0.79 \pm 0.09$	$0.013 \pm 0.001$	0.9
$WWb\bar{b}$	$237.7 \pm 0.2$	$0.05 \pm 0.04$	1.7
$t\bar{t}Z$	$11 \pm 1$	$0.25 \pm 0.03$	4.6
$t\bar{t}W$	$0.21 \pm 0.08$	$0.010 \pm 0.002$	5.3
$WWW$	$0.014 \pm 0.001$	$0.002 \pm 0.001$	5.9
$ZWW$	$0.36 \pm 0.01$	$0.136 \pm 0.005$	6.9

Table 7.4: The optB cut cross-section and efficiency, including comparison in efficiency with Figy et al. [8].

The integrated luminosity required for a  $5\sigma$  discovery at the LHC using the optB cut is  $(98 \pm 5) \times 10^3 \text{ fb}^{-1}$ . This is much higher than what future planned experiments will reach, which means that the signature is not detectable when using these optimized cuts. There could be better suited cuts for this parameter point, but a full cut analysis

is beyond the scope of this thesis. Therefore a comparison is done in chapter [9](#) using the same parameter point as Figy et al. [\[8\]](#).

# Chapter 8

## Machine Learning

In the past few decades the development of computer hardware has made significant progress, increasing the available computational power. This has made possible new machine learning techniques such as *deep neural networks* which requires large amounts of computational power to be efficient. The machine learning algorithms can be classified into three main categories according to their style of learning;

- Supervised learning

These algorithms is trained to find a general rule to map an input data set to an output data set. It is called supervised because the algorithm learns from a set of training data with known targets. More formally is that it learns a target function  $f$  that best maps a set of input variables  $\mathbf{x}$  to an output  $y$ . This is explained in more detail later in this section.

- Unsupervised learning

In some cases, there are data sets with no associated labels, and the machine learning algorithm has to classify the output based on the structure in the data. This is called unsupervised learning. An example is grouping of similar data into clusters.

- Reinforcement learning

This is different from the two above because it learns on constraint input from the outside world. It is basically a "learning by doing" algorithm

which has to make an action in response to each data point at runtime (e.g. sensor data from a self-driving car). The algorithm improves its strategy by receiving rewards or punishments for the decisions.

The algorithm used in this thesis is supervised learning. After the training on the labeled data is finished, the algorithm uses its 'experience' to predict estimated output values for unseen input data. A typical use-case outside physics is estimating real estate prices, with input features such as the size, number of rooms, and distance to the city center. The output value in this example is continuous and therefore in the regression subgroup. In this thesis, the classification subgroup is utilized, which learns on data belonging to separate classes. The algorithm then has to classify new data to one of the classes.

## 8.1 Neural Networks

The simplest unit in a neural networks is a single neuron, which is based on a McCulloch and Pitts' Neuron [168]. It consists of a set of input nodes  $x_1, x_2, \dots, x_m$ , that is multiplied by a weight and summed,

$$h = \sum_{i=0}^m w_i x_i. \quad (8.1.1)$$

Each input has a corresponding weight  $w_1, w_2, \dots, w_m$ . This sum is then used as input to an activation function, for instance the *sigmoid* function [169]

$$y = g(h) = \frac{1}{1 + \exp(-\beta h)}, \quad (8.1.2)$$

where  $\beta$  is some positive parameter. This is the output of the neuron and the predicted value for the simplest neural network [170]. For this neuron to be able to predict useful results, it needs to be trained. In supervised learning, this is done by feeding a set of data with known labels into the neuron with one input for each feature of the data set. The network is initialized by assigning random values to the weights before the first run. The output  $y$  is then compared to the target value  $t$ , i.e. the true value, which is either 0 or 1 in the single neuron case. This is done by calculating the error using an error, or loss, function. An example of such a loss function is the *binary cross entropy*

loss function [171]

$$L = -[t \ln(y) + (1 - t) \ln(1 - y)]. \quad (8.1.3)$$

The goal of the training process is to minimize this loss function, which is done after each iteration by adjusting the weights towards the minimum of the loss function. This direction is the derivative with respect to the weights. This is commonly referred to as the gradient descent algorithm. Hence, the weights are updated as

$$w_i \leftarrow w_i - \eta \frac{\partial L}{\partial w_i}, \quad (8.1.4)$$

where  $\eta$  is the learning rate, which is usually set to a number less than one to prevent unstable training of the network. By repeating these steps many times, the network learns to classify a set of input features that it has not seen before. Note that for this simple one neuron network, i.e. the inputs directly connected to a single output node, only linear relations in the data can be learned. Formally we can think of the mapping function  $f$  as a linear function. Thus, data which is not separable by a hyperplane can not be classified using this network.

If we add layers of neurons between the inputs and output, the network can learn to map a set of input features to an output by a non-linear function. These layers are referred to as hidden layers. The neurons, which we call nodes from now on, in the first hidden layer are connected to each node in the input layer. The hidden layer nodes is built up by the same components as the output node, i.e. each node has weighted inputs which are summed and passed through an activation function. One commonly used activation function for the hidden layer nodes is the *ReLU* function

$$g(h) = h^+ = \max(0, h), \quad (8.1.5)$$

which is 0 for  $h < 0$  and a linear function for  $h > 0$ . One of the advantages is that when  $h < 0$  the node becomes inactive, which reduces the computations required. A sketch of such a multi-layer neural network with two hidden layers can be seen in fig. [8.1], where the green nodes are the input nodes, the purple nodes are the two hidden layers and the red is the output node. A neural network with two or more hidden layers are commonly referred to as a *deep neural network* (DNN).

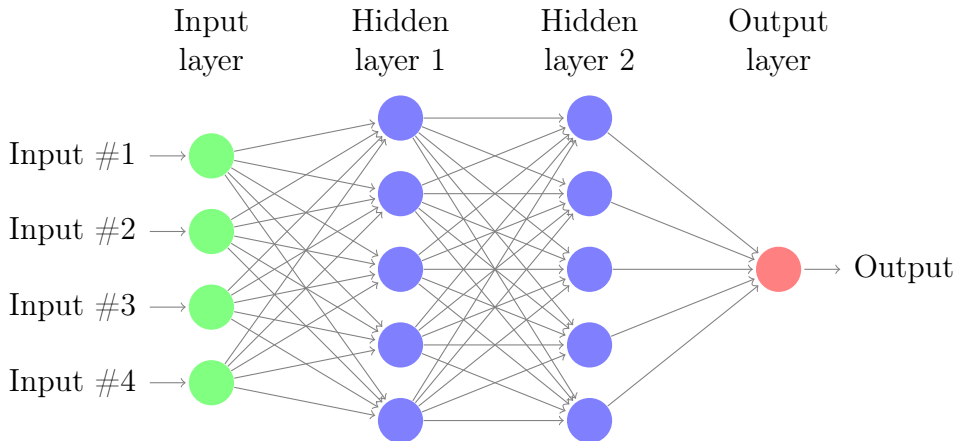


Figure 8.1: Example of a deep neural network. This has two hidden layers with 5 nodes represented in purple, there are 4 input nodes and one output.

## 8.2 Binary classifier setup

To set up the neural network algorithm, the *Keras 2.1.1* [172] API is utilized. This uses *TensorFlow* [173] as backend, which is a high-performance numerical computation library built for machine learning tasks. The neural network architecture is based on [3], where the number of hidden layers is four, with 250, 200, 100 and 50 nodes. The 'ReLU' activation function is used for the hidden layers [174]. This has in general better performance than other widely used activation functions [175], such as the sigmoid and 'tanh' functions for the hidden layer nodes. The output layer has a sigmoid activation function and the loss function is the *binary cross-entropy* function eq. (8.1.3). In the neural network introduction, section 8.1, we updated the weights by the gradient descent method. In *Keras* a more sophisticated algorithm called the *Adam optimizer* is implemented [176].

### 8.2.1 Input features

The case studied in this thesis has 26 input features, which includes the  $p_T$ ,  $\eta$  and  $\phi$  for the three signature particles ( $\tau_h$ ,  $\tau_h$  and  $\mu$ ) and the three hardest jets, the total number of jets, number of b-tagged jets, the invariant masses  $M_{\tau_1\mu}$ ,  $M_{\tau_2\mu}$  and  $M_{\tau_1\tau_2}$  and the missing transverse energy ( $\cancel{E}_T$ ) with its azimuthal angle ( $\phi^{\cancel{E}_T}$ ). The included input features are only observables visible for detectors, other features of the signal such as the neutrino  $p_T$  is not included as this is not possible to detect in experiments. Some of the input feature distributions are shown in fig. 7.2 and fig. 7.3. Other high-level features than the invariant mass are not included as the discriminating features are

learned by the DNN from the low-level features [3].

## 8.3 Pre-processing feature data

Before the data is passed into the neural network, it is preprocessed to enhance the performance of the neural network training process [177]. We import the simulated events, one dataset for each process where the distribution of training data reflects the relative cross-section, see table 7.3. Because of the high  $WWb\bar{b}$  cross-section, we need about 450 times more events from this process than the signal processes. To obtain a reasonable amount of signal events, i.e. above 10000 events, the number of  $WWb\bar{b}$  events has to be 4500000 for the distribution of training data to reflect the relative cross-section. To simulate 50000 events of this process, two weeks of computation is needed and thus simulating over a million events is not possible within the given time limit for completion of this work. The relative cross section of the  $WWb\bar{b}$  process is therefore reduced to  $0.1 * \sigma_{\text{Det}, WWb\bar{b}}$ . We can do this change in training data class distribution without losing useful information. This is because the feature distributions of signal and background, such as  $p_T^{\tau_1}$  and  $N_j$  in fig. 7.2, is separated enough to avoid problems regarding bias in this distribution of input data. For a recent example where the input data class distribution do not avoid this problem, see sec. 2.4 in [178] where an instructive example of this is presented.

The datasets are split into training and testing sets, with a ratio of 0.7 to 0.3 for the  $WWb\bar{b}$  process. Since this is the process which the distribution of training data is based on, the other processes need to split their data such that it match this training data distribution. Signal events are labeled 1 and background events 0. The data sets are then shuffled and centered with the mean at zero unit variance; this is called statistical standardization [171]. To be consistent with the scaling, the testing data is scaled to the same variance as the training data.

### 8.3.1 Training the network

The neural network is trained until it only memorize the training data, and do not learn any new general feature correlations. Passing this point in the training process is called overfitting. After each epoch, which is one iteration through the training dataset, the accuracy and loss are calculated with a set of validation data. This is



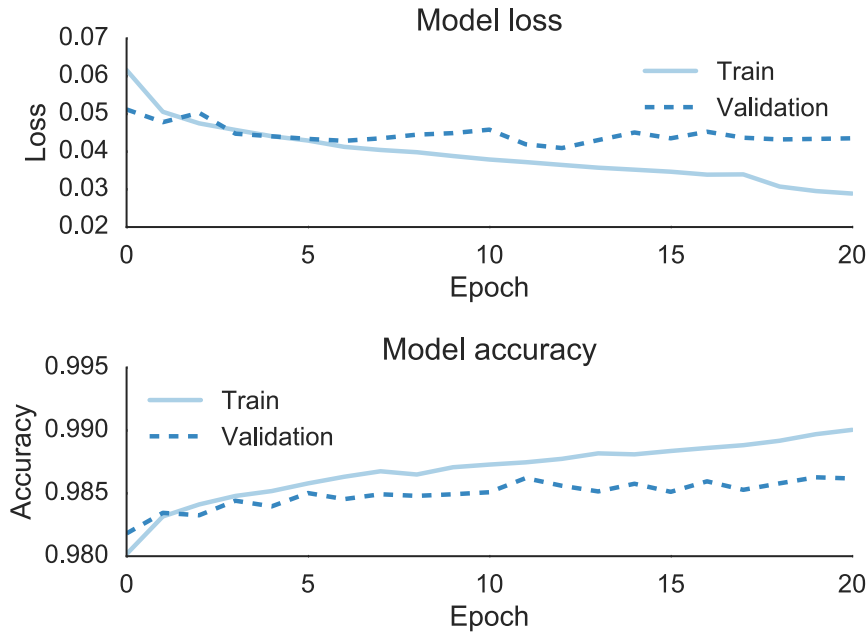


Figure 8.2: The evolution of model loss and accuracy while training the deep neural network, showing that the loss decreases and the accuracy increases as the network learns the feature correlations. The point at which the loss and accuracy for the validation set flattens out while the loss (accuracy) of the training set continue to decrease (increase) indicates when the network begins to memorize the training set.

used to find the point where overfitting starts. The accuracy is defined as

$$\text{Acc} = \frac{\#TP + \#TN}{N}, \quad (8.3.1)$$

where true positive (TP) is the number of signal events classified correctly, true negative (TN) is the number of background events classified correctly and  $N$  is the total number of events. Incorrectly classifying an event as signal (background) when it is indeed a background (signal) event, is referred to as false positive (FP) and false negative (FN). The neural network stops when there is no improvement in the loss after eight epochs or if it passes 50 epochs in total. In fig. 8.2 the loss and accuracy are plotted for the set of training and testing data at each epoch.

The accuracy in for the training and validation set in fig. 8.2 is misleading because of the distribution of training data relative to the cross-section, i.e. there are more background than signal events. The reason this accuracy is misleading is because if the classifier predicts background for all events, i.e.  $TP = 0$ , then the accuracy according to eq. (8.3.1) is  $\#TN/N$ . For the dataset used to produce fig. 8.2, the

$\#TN/N = 0.976$  which fits with the accuracy obtained after the first epoch. This means that a small increase in accuracy is a significant improvement for the network.

## 8.4 Neural network as a cut

To express the result in terms of a physical quantity, we use the deep neural network in the same way as the optimized cuts in the previous chapter. We calculate the efficiency using eq. (7.4.1), where signal efficiency can be expressed in terms of the machine learning terminology, as explained in the previous section, as

$$\epsilon_{\text{ML}} = \frac{\#TP}{N}. \quad (8.4.1)$$

and the corresponding uncertainty is calculated using eq. (7.4.2). Then the cross-sections after the machine learning cut ( $\sigma_{\text{ML}}$ ) are obtained by multiplying the efficiency with the  $\sigma_{\text{Det}}$  from table 7.3. From this, we calculate the integrated luminosity required for a detection significance of  $5\sigma$ , denoted  $\mathcal{L}_{\text{int}}(5\sigma)$ , as in section 7.7. This is the value we want to optimize, as we did in the previous chapter. The parameters of the neural network, often referred to as the hyperparameters, can be adjusted and optimized for the given dataset. The next section contains a discussion about how we optimized one of the hyperparameters. This turns out to be an important factor for the  $\mathcal{L}_{\text{int}}(5\sigma)$  result.

## 8.5 Class weights

In this section, the deep neural network is optimized by changing the loss function to punish false classification of the two classes (signal and background) differently. This is done by assigning different weights to the classes, referred to as *class weights*. The cross-entropy loss function, eq. (8.1.3), is changed to

$$L = -[\omega_1 t \ln(y) + \omega_0 (1 - t) \ln(1 - y)] \quad (8.5.1)$$

where  $\omega_1(\omega_0)$  is the signal(background) class weight. If we set  $\omega_1$  higher than  $\omega_0$ , we see from eq. (8.5.1) that an output of, for instance,  $|y - t| = 0.4$  for a signal event ( $t = 1$ ) yields a higher loss than  $|y - t| = 0.4$  for a background event ( $t = 0$ ).

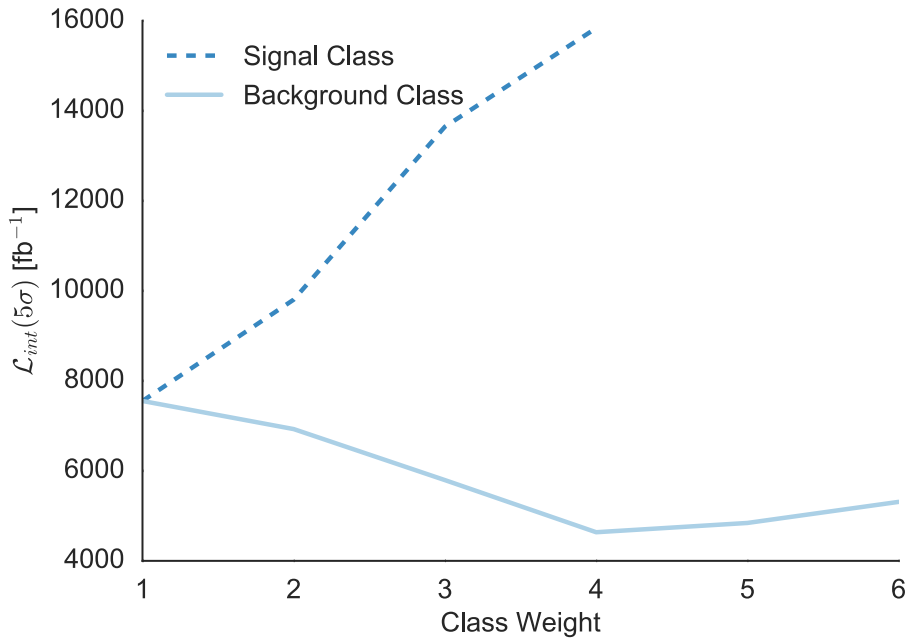


Figure 8.3: Change in the required integrated luminosity for a  $5\sigma$  significance of discovery as a function of class weights, represented by the dashed (solid) line.

The class weights are optimized by training the network several times and manually adjusting the class weights after each full training process. The result is plotted in fig. 8.3, for change in both the signal and background class weights. The blue dashed line shows how increasing the signal or background class weights changes  $\mathcal{L}_{\text{int}}(5\sigma)$ . We clearly see how increasing the background class weight lowers the required  $\mathcal{L}_{\text{int}}(5\sigma)$  up to a class weight of 4.

To explain why increasing the background class weight improves our result we need to look at the  $\mathcal{L}_{\text{int}}(5\sigma)$  function, eq. (7.7.3). We use the total signal and background cross-sections from table 7.3, i.e.  $\sigma_{\text{Det},S} = (11.45 \pm 0.02) \times 10^{-2}$  fb and  $\sigma_{\text{Det},B} = (49.485 \pm 0.31)$  fb. This is plotted for signal efficiency and the background rejection efficiency in fig. 8.4, with the  $\mathcal{L}_{\text{int}}(5\sigma)$  represented in a colormap. From this plot, we see that a change in the background rejection efficiency changes  $\mathcal{L}_{\text{int}}(5\sigma)$  more than a change in signal efficiency. This explains why increasing the background class weight lowered the required  $\mathcal{L}_{\text{int}}(5\sigma)$ . It also explains why the optimization of optB in section 7.8 favored a high background rejection. The plot marks the line where the signal could be detectable at a future high-luminosity LHC; this limit is set to an integrated luminosity of  $3000 \text{ fb}^{-1}$ . Some results from the background rejection and signal efficiency are shown as red points in the plot. These move closer to the detectable line for increased

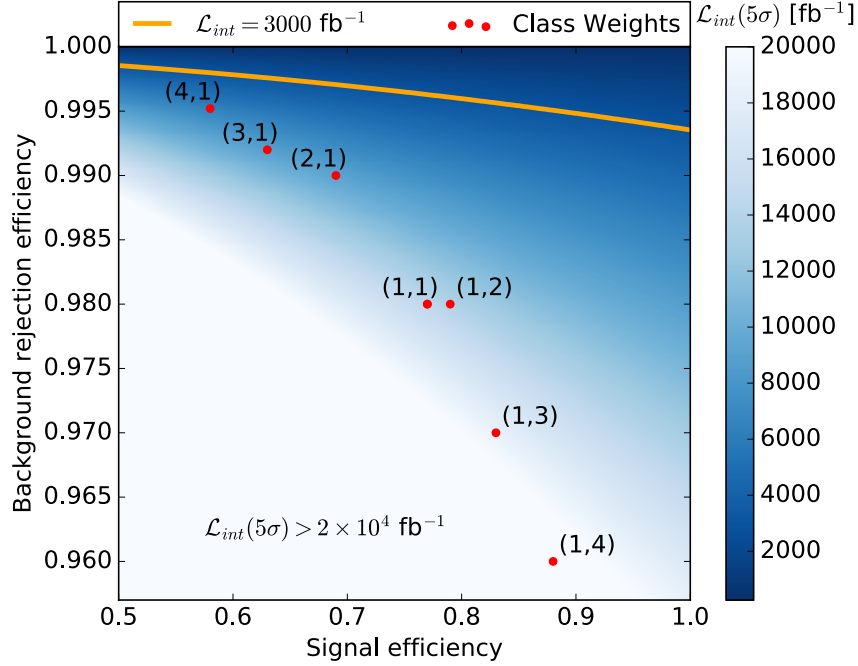


Figure 8.4: The change in the required integrated luminosity for a detection significance of  $5\sigma$  (represented as a colormap) for background rejection and signal efficiency, where the initial cross-section is  $\sigma_{\text{Det},S} = 0.115\text{fb}$  for the signal and  $\sigma_{\text{Det},B} = 49.485\text{fb}$  for the background. The red points indicate different class weights for the background ( $B$ ) and signal ( $S$ ) class ( $B, S$ ).

background class weight and correspondingly farther away with increased signal class weight. One thing to note about this plot is that it does not represent the true value of  $\mathcal{L}_{\text{int}}(5\sigma)$  because we use the same efficiency for all the signal and background processes and not the efficiency for each individual process as in section 7.7. This means that  $\sigma_S$  and  $\sigma_B$  in eq. (7.7.3) is

$$\sigma_S = \epsilon_{\text{ML},S} \sum_i^3 \sigma_{\text{Det},i} = \epsilon_{\text{ML},S} \sigma_{\text{Det},S}, \quad \sigma_B = \epsilon_{\text{ML},B} \sum_i^7 \sigma_{\text{Det},i} = \epsilon_{\text{ML},B} \sigma_{\text{Det},B}, \quad (8.5.2)$$

where the sum is over the individual signals and backgrounds correspondingly. The neural network cut efficiency is denoted as  $\epsilon_{\text{ML},S}$  ( $\epsilon_{\text{ML},B}$ ) for signal (background) in eq. (8.5.2). Using this simplification yields a slightly higher  $\mathcal{L}_{\text{int}}(5\sigma)$ , about 5% increase, but it is necessary to be able to plot in two dimensions.

## 8.6 Classifying individual signals

Until now, we have looked at the performance of a neural network which classifies an event as either signal or background. In this section, on the other hand, all signal and background processes are classified individually. This means that there are 10 outputs, one for each process. This is motivated by two reasons; 1) we want to see if there are certain background processes that mimics the signal more then other background processes, 2) test if this lowers  $\mathcal{L}_{\text{int}}(5\sigma)$ . The first can be used to achieve a lower  $\mathcal{L}_{\text{int}}(5\sigma)$  because we can target the background class weights individually. Another reason for a multi class classifier is that we can separate the three signals. In the single class classifier above, we had to tell the network that SUSY EW and SUSY QCD is the same. This is not ideal for the network, because these signals do not possess the same feature distributions. Thus, by separating the two into different classes makes it easier for the network to ‘specialize’ on each signal.

The architecture of the neural network changes to have 10 output nodes, while the rest of the architecture; input features, the number of nodes and hidden layers as well as the activation function for the hidden layer nodes, stay the same as previously described. The output of the neural network is now an array of 10 values between 0 and 1, where each indicates the prediction confidence of the corresponding class. All these output nodes, are all connected to each node in the last hidden layer. The output nodes use the *soft-max* activation function [171]

$$y_k = g(h_k) = \frac{e^{h_k}}{\sum_{j=1}^{10} e^{h_j}}, \quad (8.6.1)$$

where  $k$  is the output node index and  $h_k$  is the sum of the inputs no node  $k$ , as in eq. (8.1.1). After the activation function all the outputs sums up to 1. The cross-entropy function is used as the loss function, which for a multi class classifier is given by

$$L = - \sum_{k=1}^{10} \omega_k t_k \ln(y_k), \quad (8.6.2)$$

where the sums runs over all the output nodes. The class weight parameter is included so that each process can be weighted individually. Since the  $WWb\bar{b}$  background process is the dominant one, we divide the class weights into three groups; signal,  $WWb\bar{b}$  and the other backgrounds. The class weights used are 1 for signals, 3 for all but the  $WWb\bar{b}$  process which has a class weight of 10.

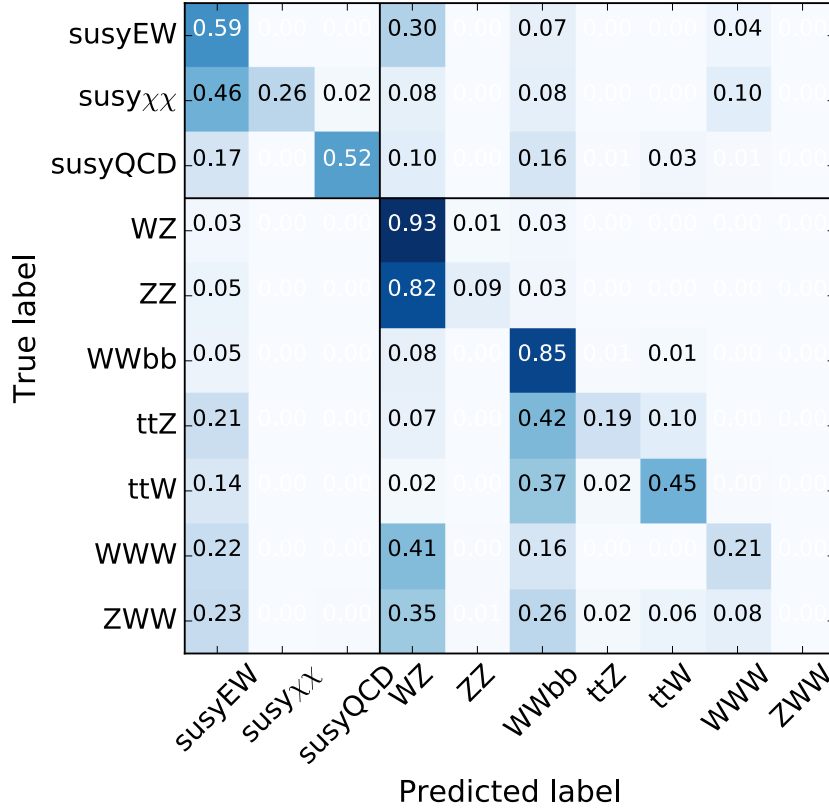


Figure 8.5: The confusion matrix for the multi class deep neural network, where each row shows the distribution of the predicted label for the corresponding process. The sum of each row is one and, entries with a value less than 0.01 is left blank. This is for the class weighted deep neural network; ( $WWb\bar{b}$ : 10, Background: 3, Signal: 1)

It is interesting to look at the confusion matrix for the multi-classifier, which is shown in fig. 8.5. The x-axis is the predicted class, and the y-axis is the target value. From this, we can see which processes the network is more prone to classify incorrectly and the processes it incorrectly classified it as. The three backgrounds which the classifier confuses the signal with, is  $WZ$ ,  $WWb\bar{b}$  and  $WWW$ . The reason why the signals are confused with these backgrounds are because they produce the signature  $(\tau_h^\pm \tau_h^\pm \mu^\mp + n_j)$  without any miss-identification of prompt particles. Another reason for this is the dominance of  $WZ$  and  $WWb\bar{b}$  in the distribution of the data. This cause the neural network to learn the characteristics of these processes better then the other processes. Some of these characteristics could be similar in the other processes too, but because of the smaller training data set, the neural network was not able to learn these characteristics in other processes.

The last background processes,  $t\bar{t}Z$ ,  $t\bar{t}W$ ,  $WWW$  and  $ZWW$ , is often classified incorrectly as SUSY EW. Take, for instance, the  $ZWW$  background, which is never classified correctly as itself. There could be many reasons for this. The first is because  $ZWW$  can decay to the signature in different ways;

$$ZW^+W^- \rightarrow \tau_h^+ \cancel{\tau}^- \tau_h^+ \mu^- \nu_\tau \bar{\nu}_\mu, \quad \tau_h^+ \tau^- [\rightarrow \mu^+ \nu_\mu] \tau_h^+ \cancel{\tau}, \quad (8.6.3)$$

where  $\cancel{\tau}$  means the particle is not identified as the correct particle or did not pass the detector cuts. The neutrinos escapes the detector, and is not measured. The latter decay chain from the  $ZWW$  background have the exact same signature as  $W^+Z \rightarrow \tau_h^+ \tau_h^+ \tau^- [\rightarrow \mu^+ \nu_\mu]$ , which means that these have very similar features. For the  $WWb\bar{b}$ , which decays as  $W^+[\rightarrow \tau^+ \nu_\tau] W^-[\rightarrow \mu^- \bar{\nu}_\mu] b\bar{b}$  and the two bottom quarks decay to an hadronic tau and jets, is similar to the first decay channel of  $ZWW$  in eq. (8.6.3). By using the same argument as when talking about miss-identified signal processes as background, the  $WZ$  decay directly to the signal without any extra miss-identified particles. This means again that since  $ZWW$  is similar to  $WZ$ , it is also similar to SUSY EW.

When calculating the cut efficiency, we select which events that are signal or background by summing the output values of the signal classes and background classes. Which sum is higher determines whether an event is interpreted as signal or background. The resulting  $\mathcal{L}_{\text{int}}(5\sigma)$  obtained by this method is better than classifying the event as signal or background based on the process with the highest confidence. The reason why the method where all signal and backgrounds are summed and then compared, performed better can be seen from the confusion matrix. The background processes are often confused to be other background processes, this yields small individual prediction confidences (outputs), which can be lower than the signal prediction confidence.

## 8.7 Performance

In this section we compare the performance of several classification methods; single class DNN (SC DNN), single class DNN with class weights (SC DNN (class weights)), multi class DNN (MC DNN) and a boosted decision tree (BDT). The BDT was created using the *AdaBoost* classifier [179] implementation in the scikit-learn library [180], with the max tree depth of two, 600 estimators and a learning rate of 1. The performances

of the classifiers are measured in the required  $\mathcal{L}_{\text{int}}(5\sigma)$  and presented in table [8.1](#).

Technique	$\mathcal{L}_{\text{int}}(5\sigma)$ [fb <sup>-1</sup> ]
SC DNN	7533 ± 875
SC DNN (Class Weights)	4664 ± 853
MC DNN	5449 ± 644
MC DNN (Class Weights)	3425 ± 748
BDT	10182 ± 1374

Table 8.1: Comparison of the machine learning algorithms performance for the integrated luminosity required for a detection significance of  $5\sigma$  ( $\mathcal{L}_{\text{int}}(5\sigma)$ ).

The BDT has poorer performance than the DNNs. This is because it is not able to learn the high-level features [3](#), such as the radial separation of the  $\mu$  and  $\tau_2$  ( $\Delta R^{\mu, \tau_2}$ ), see fig. [7.3](#). Setting class weights of the background class in the DNN lowered the required  $\mathcal{L}_{\text{int}}(5\sigma)$ . Further improvements were made by creating a multi class classifier. The cross-section and cut efficiencies for the multi class classifier are presented in table [8.2](#), and this is the network which is used when comparing the performance to the optimized cuts in the next section.

Signal	$\sigma_{ML} / 10^{-2}$ [fb]	$\epsilon_{ML} / 10^{-2}$
SUSY EW	3.11 ± 0.01	38.0 ± 0.1
SUSY QCD	1.460 ± 0.002	53.2 ± 0.1
SUSY $\chi\chi$	0.222 ± 0.001	42.5 ± 0.1
$WZ$	0.22 ± 0.02	0.072 ± 0.008
$ZZ$	0.017 ± 0.001	0.029 ± 0.003
$WWb\bar{b}$	5.8 ± 1.69	0.129 ± 0.046
$t\bar{t}Z$	0.04 ± 0.03	0.09 ± 0.01
$t\bar{t}W$	0.073 ± 0.003	0.37 ± 0.01
$WWW$	0.16 ± 0.01	2.36 ± 0.04
$ZWW$	0.055 ± 0.001	2.10 ± 0.02

Table 8.2: The cross-section and efficiency for the best performing neural network; the MC DNN with background class weight of 3 except for the  $WWb\bar{b}$  class which have 10.





# Chapter 9

## A comparison of deep neural networks and optimized cuts

We compare the machine learning performance to the optimized cuts from [8]. To compare the two methods, we use the same parameter point as [8], which have  $m_0 = 100$  GeV,  $m_{1/2} = 500$  GeV,  $\tan\beta = 10$ ,  $A_0 = 0$ ,  $m_{A^0} = 2000$  GeV and  $\mu = 600$  GeV. Note that the  $m_{A^0}$  is the pole mass of the CP-odd Higgs boson. Thus, representing this point using eq. (2.5.8) in the  $m_{H_d}^2$  and  $m_{H_u}^2$  parameter space is not trivial. The mass spectrum and the decay branching ratios is generated using *SPheno*, as in section 5.2.1. In this point the tau-sneutrino mass is  $m_{\tilde{\nu}_\tau} = 90.5$  GeV, the smuon and muon-sneutrino masses are  $m_{\tilde{e}_L} = 161.4$  GeV and  $m_{\tilde{\nu}_e} = 140.6$ , the second neutralino and first chargino are  $m_{\tilde{\chi}_2^0} \simeq m_{\tilde{\chi}_1^\pm} = 396.0$  GeV and the squark masses are at  $\sim 1$  TeV. Most notably is perhaps the lightest Higgs boson at  $m_{h^0} = 115.9$  GeV, which is ruled out by experiments [31]. This means that this chapter is solely a comparison between the two methods.

### 9.1 Event simulation

The same event generator settings and analysis script, as in chapter 7, are used in the event simulation. The background samples are the same and the only difference in the signal simulations is the parameter point passed to *Herwig*. The number of accepted events, the detector cut cross-sections and efficiency for the signal processes are shown in table 9.1 and for the background processes in table 7.3. In this point,

Process ( $\tau_h^\pm \tau_h^\pm \mu^\mp$ )	$N_{\text{Det}}$	$\sigma_{\text{Det}}$ [fb]	$\epsilon_{\text{Det}} / 10^{-2}$
SUSY EW	158666	$2.115 \pm 0.001$	$0.45 \pm 0.01$
SUSY QCD	136899	$3.726 \pm 0.001$	$0.68 \pm 0.02$
SUSY $\chi\chi$	145959	$0.137 \pm 0.001$	$0.49 \pm 0.01$

Table 9.1: The number of accepted events, cross sections and detector cut efficiency for the signal processes at  $\sqrt{s} = 13\text{TeV}$  for the  $\tau_h^\pm \tau_h^\pm \mu^\mp$  signature using the parameter point from Figy et al. [8].

the dominant signal is SUSY QCD. This is due to the squark masses, which are significantly lighter in this parameter point at about 1 TeV. The total signal cross-section is  $\sigma_{\text{Det},S} = (5.978 \pm 0.002)$  fb. Since the same background samples are used, the total background cross-section is  $\sigma_{\text{Det},B} = (49.485 \pm 0.31)$  fb. This means that the signal to background ratio is larger compared to the parameter point used in chapters 7 and 8.

## 9.2 Feature plots

Some of the feature distributions are shown in fig. 9.1. These are all normalized in the same way as in figs. 7.2 and 7.3, i.e. the distributions are normalized to one and not to the individual cross-sections because of the large background to signal ratio. The first row show the transverse momentum distributions for the signature particles. For this parameter point the SUSY EW signal overlap the background  $p_T$  distribution. In the next two rows, the jet feature distributions and the missing transverse energy ( $\cancel{E}_T$ ) are presented, which show the possibility of extracting the SUSY QCD signal from the background. For instance, the number of jets ( $N_j$ ) and the hardest jet transverse momentum ( $p_T^{1j}$ ) have both the majority of the distribution higher than the background.

High-level feature distributions are presented in the last plot in row three and the two last rows in fig. 9.1. These plots show, as in section 7.6, that it is possible to construct cuts that separate the SUSY EW and the background. This is for the most part due to the clean signature of the SUSY EW, i.e. the decay chain is short from the production particles to the tau-sneutrino. In the last row the invariant mass,  $M$ , of the signature particles is presented. The two last distributions in this row show how the invariant mass of the muon with the two taus have a lower energy peak than the background. This is the case for all the signals, but is most dominant in the SUSY EW signal.

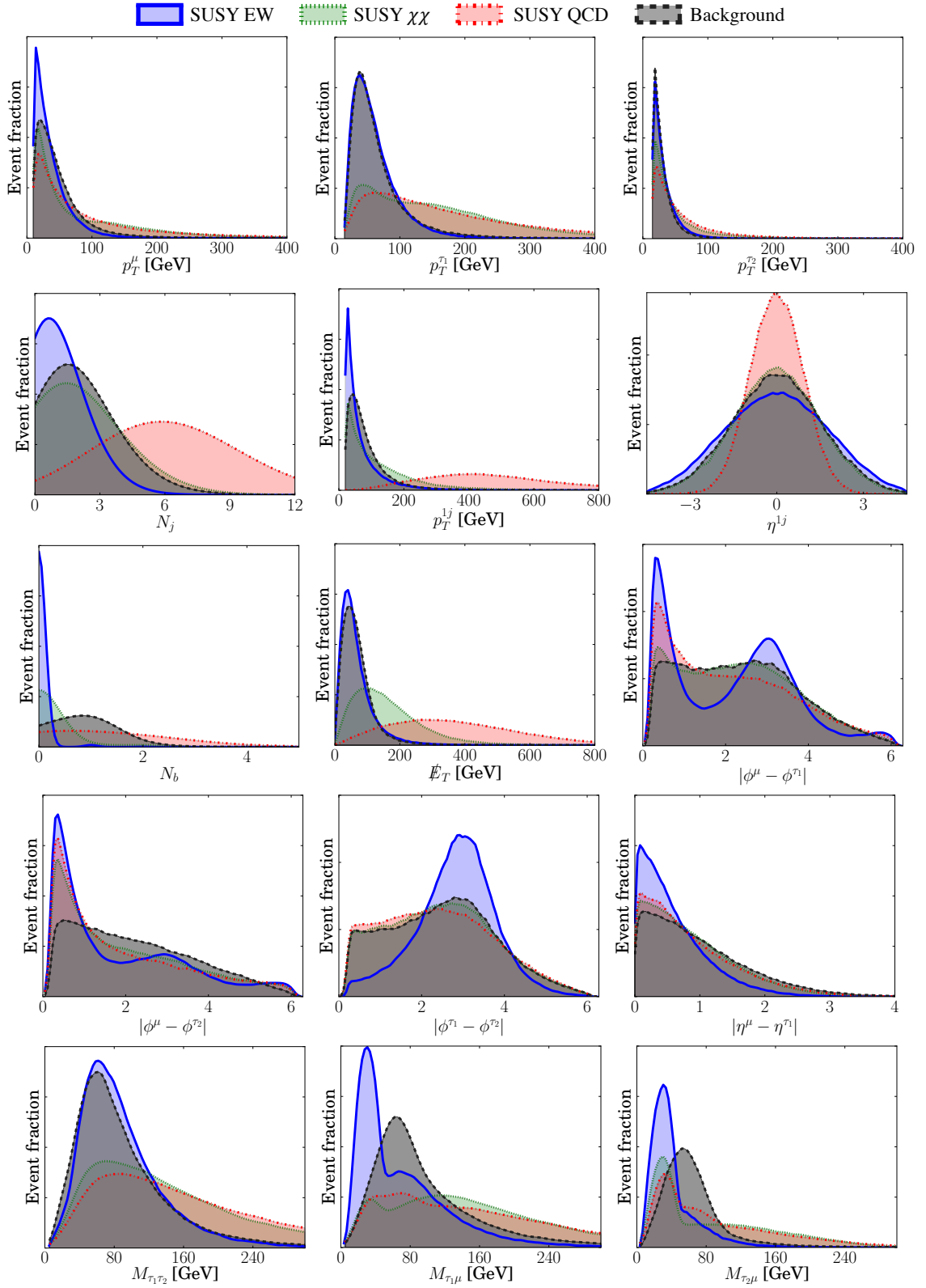


Figure 9.1: Selected low- and high-level feature distributions for the signature particles ( $\mu$ ,  $\tau_1$ ,  $\tau_2$ ), the missing transverse momentum ( $\cancel{E}_T$ ) and the hardest jet ( $1j$ ).

Thus, it is possible to cut on the minimum of these two invariant masses. We want the minimum because the muon could originate from the same decay as both the  $\tau_1$  and the  $\tau_2$ , which means that it is only one of the them that have an invariant mass with a energy peak below the background. The invariant mass from when the muon and the tau does not originate from the same decay can be seen as the tail behind the energy peak of the  $M_{\mu\tau_1}$  and  $M_{\mu\tau_2}$  distributions.

### 9.3 Optimizing cuts

In this section the optimized cuts from Figy et al. [8], presented in section 7.8, are again optimized for the parameter point used in this chapter. The optimization of the cuts used for the same parameter point in [8], must be redone because of the improvements in event generators and of the order in which the matrix element is calculated. In the rest of this chapter, all values are normalized to the following reference value;  $\mathcal{L}_{\text{int}}(5\sigma)/\mathcal{L}_{\text{ref.}}$ , where  $\mathcal{L}_{\text{ref.}}$  is some reference integrated luminosity. This is to make it easier to compare values. For the optimization in this section, all values are normalized to the required integrated luminosity using optimized cut as they are presented in Figy et al. [8].

The OptB cut first considers the required number of jets cut and the required  $p_T$  of the hardest jet cut. The results are normalized to the OptB cuts from [8], i.e.  $N_j \geq 2$ ,  $p_T^{1j} \geq 200$  GeV and  $\sum E_T^i \geq 300$  GeV. This yields an required integrated luminosity for a discovery of  $\mathcal{L}_{\text{ref. OptB}} = (24.07 \pm 0.30)$  fb<sup>-1</sup>. The normalization of the  $\mathcal{L}_{\text{int}}$  results obtained, is represented as  $R_{\text{OptB}} = \mathcal{L}_{\text{int}}/\mathcal{L}_{\text{ref. OptB}}$ . Note that we do not specify the  $5\sigma$  significance, which is due to the cancellation of the significance in the ratio, i.e. the ratio  $R_{\text{OptB}}$  is not dependent on the significance of discovery. The  $R_{\text{OptB}}$  is computed for varying  $N_j$  and  $p_T^{1j}$  cuts, which are plotted in fig. 9.2. The last cut, which cut on the required  $\sum E_T^i$ , does not change  $R_{\text{OptB}}$  for varying  $\sum E_T^i$  cut and therefore we use the value cut as before, i.e. we require  $\sum E_T^i \geq 300$  GeV. Thus, the OptB cuts, after the optimization, requires;

1.  $N_j \geq 4$ ,
2.  $p_T^{1j} \geq 300$  GeV,
3. and  $\sum E_T^i \geq 300$  GeV.

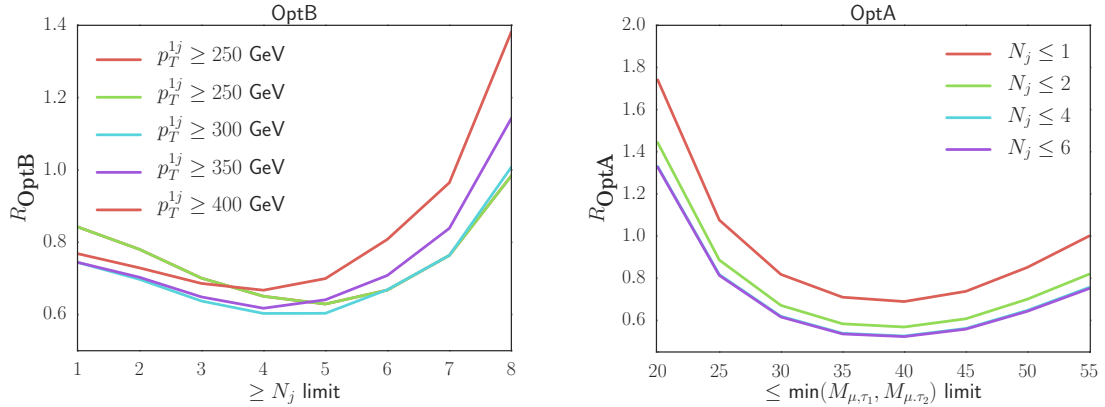


Figure 9.2: The optimization of  $R_{OptB}$  by varying the required number of jets and the required hardest jet transverse momentum  $p_T^{1j}$ s, to the left, and optimization of the  $R_{OptA}$  varying the max number of jets and the invariant mass limit, to the right. This is for the parameter point from Figy et al. [8].

For the OptA cuts the reference value is  $\mathcal{L}_{\text{ref. OptA}} = (231.7 \pm 2.6) \text{ fb}^{-1}$ , using the cut limits presented in section 7.8. We define the ratio, as for OptB, as  $R_{OptA} = \mathcal{L}_{\text{int}}/\mathcal{L}_{\text{ref. OptA}}$ . The two cuts that change  $R_{OptA}$  the most are cut nr. one and three. The optimization of these are shown in the right plot in fig. 9.2 and indicate that the required number of jet cut ( $N_j$ ) lowers  $R_{OptA}$  as it increase. This can be explained by looking at the feature distributions in fig. 9.1, because almost all events from SUSY EW have  $N_b = 0$  and the background has a peak at  $N_b = 1$ . Thus, the  $N_b = 0$  cut alone is effective in removing background. Since the number of jets, distribution for the SUSY EW and the background overlap, the main difference when setting the required  $N_j$ , is how much of the SUSY QCD signal that passes. This lowers  $R_{OptA}$  and the lowest  $R_{OptA}$  is achieved by removing the  $N_j$  cut from OptA. Thus, the OptA cuts, after the optimization, requires;

1.  $N_b = 0$ ,
2.  $\min(M_{\mu,\tau_1}, M_{\mu,\tau_2}) \leq 40 \text{ GeV}$ ,
3.  $p_T^{1j} \leq 200 \text{ GeV}$  and  $|\phi^{\cancel{E}T} - \phi^\mu| \geq 1.5$ .

## 9.4 Neural network optimization

A deep neural network is optimized for the parameter point from Figy et al. [8], as the signal to background ratio is different in this parameter point and some observable

Algorithm	$\mathcal{L}_{\text{int}}/\mathcal{L}_{\text{OptB}}$
MC DNN	$0.49 \pm 0.03$
OptB	$1.0 \pm 0.1$
OptA	$8.32 \pm 0.08$

Table 9.2: The results using the parameter point from [8], where  $\mathcal{L}_{\text{int}}/\mathcal{L}_{\text{OptB}}$  is the ratio of the obtained result to the result from using the OptB cuts.

feature distributions have changed. The network used is the multi class DNN with 4 layers consisting of 250, 200, 100 and 50 nodes. The ‘ReLU’ function, eq. (8.1.5), is used as the activation function for the hidden layer nodes and the output activation function is the ‘soft-max’ function, eq. (8.6.1). The data is pre-processed in the same way as in section 8.3, and training data is distributed according to the relative cross-sections from tables 7.3 and 9.1. The class weights in the deep neural network, which is optimized for this parameter point, are 10 for the  $WWb\bar{b}$ , 3 for the other background processes and 1 for the signals.

## 9.5 Machine learning vs. optimized cuts

The results from the neural network, OptA and OptB, are presented in table 9.2. The values are normalized to the OptB result, i.e.  $\mathcal{L}_{\text{int}}/\mathcal{L}_{\text{OptB}}$ . One can clearly see that the deep neural network performs better than the optimized cuts. The deep neural network reduced the required  $\mathcal{L}_{\text{int}}$  by a factor of 2.0 from the OptB cut, which performs best of the two optimized cuts. The uncertainties are obtained in the same way as in section 7.7, where the machine learning efficiency  $\Delta\epsilon_{\text{ML}}$  is calculated using eq. (7.4.2).

The efficiencies of the three methods are presented in table 9.3. We can see that the neural network cuts away most of the background and is slightly better than the OptB cut. The main difference is that OptB do not cut away as much of the  $WWb\bar{b}$  background as the deep neural network. This is due to the ability to change the class weight of the neural network, which makes it focus on the important processes. However, this class weighting results in a signal efficiency for the neural network of only 33.8% from the SUSY EW signal. This is much lower than the OptA signal efficiency from this signal. The OptA do not perform as well as the other two in removing the background. This is because the SUSY EW signal have more signature similarities with the background.

Signal	$\epsilon_{ML} / 10^{-2}$	$\epsilon_{\text{OptB}} / 10^{-2}$	$\epsilon_{\text{OptA}} / 10^{-2}$
SUSY EW	$33.8 \pm 0.3$	$0.098 \pm 0.008$	$86.9 \pm 0.1$
SUSY QCD	$81.8 \pm 0.2$	$68.5 \pm 0.1$	$0.581 \pm 0.001$
SUSY $\chi\chi$	$32.0 \pm 0.1$	$2.51 \pm 0.04$	$35.56 \pm 0.03$
$WZ$	$0.53 \pm 0.02$	$2.88 \pm 0.03$	$22.08 \pm 0.02$
$ZZ$	$0.390 \pm 0.008$	$0.563 \pm 0.001$	$22.68 \pm 0.01$
$WWb\bar{b}$	$0.29 \pm 0.04$	$0.62 \pm 0.04$	$6.56 \pm 0.01$
$t\bar{t}Z$	$0.7 \pm 0.2$	$2.4 \pm 0.1$	$1.98 \pm 0.01$
$t\bar{t}W$	$0.59 \pm 0.02$	$0.54 \pm 0.02$	$1.09 \pm 0.01$
$WWW$	$5.59 \pm 0.06$	$0.90 \pm 0.02$	$14.95 \pm 0.02$
$ZWW$	$4.05 \pm 0.02$	$2.44 \pm 0.02$	$10.53 \pm 0.02$

Table 9.3: The efficiency for the deep neural network, OptB and OptA (see section 9.3) cuts from the detector cross-section of the parameter point from Figy et al. [8], from the signals see table 9.1, and background see table 7.3.





# Chapter 10

## Summary and conclusion

### 10.1 Overview of results

In this thesis we have studied the detectability of a supersymmetric scenario with tau-sneutrino NLSP and gravitino as the LSP, in particular within the NUHM model. The analysis had a starting point in the published work by Figy et al. [8]. First the tau-sneutrino lifetime for a given gravitino mass was calculated in chapter 4 and showed that the tau-sneutrino is long-lived within the allowed gravitino mass range.

The parameter space was studied to find a parameter point satisfying the mass of the lightest Higgs boson. Due to this constraint, the soft gaugino mass parameter ( $m_{1/2}$ ) had to be increased and because  $m_{1/2}$  contributes to the mass of all sfermions and the Higgs bosons, the total mass spectrum became high. The squark masses, except for the stop, have a mass of  $\sim 2$  TeV, which suppress the squark production channels at the LHC. This is a problem for detection of this scenario because the contribution from squarks are usually dominant and since the electroweak supersymmetry productions are low because of the weaker coupling to  $W$ ,  $Z$  and  $\gamma$ . Note that this is one of the lighter parameter points in the region, which means that points in general have higher mass spectra.

Once the parameter point was in place, we studied the possible collider signatures. The signature ( $\tau_h^\pm \tau_h^\pm \mu^\pm + n \cdot j + \cancel{E}_T$ ) was chosen based on the small background and because it was the same as in Figy et al. [8]. Other possible signatures could be interesting to look into, especially based on the background removal efficiency obtained in chapter 8.

The signature that was used in this thesis did not include electrons, since they require more attention when choosing detection efficiencies and detector cuts. This means that a separate simulation with only electrons should be carried out and would lower the total required  $\mathcal{L}(5\sigma)$ .

Collision events were simulated with a center of mass energy of  $\sqrt{s} = 13$  TeV. A deep neural network was trained on the simulated data to distinguish signal from background events. We explored different methods, such as class weights and classifying each process individually in a multi-class classifier. The network was also compared to a boosted decision tree algorithm. The best performing algorithm was the multi class deep neural network with increased background class weights. The background class weight was set to 3 for all but the dominant background  $WWb\bar{b}$  which had a background class weight of 10. The result from the best performing DNN has a required  $\mathcal{L}_{\text{int}}(5\sigma) = (3.4 \pm 0.7) \times 10^3 \text{ fb}^{-1}$ . The large uncertainty is due to the large background removal efficiency and the small number of simulated events for some of the backgrounds processes. If the tau-sneutrino NLSP scenario is realized in nature, the LHC will have problems detecting it, even at the HL LHC with the planed  $3000 \text{ fb}^{-1}$  [181]. In this case, a electron-positron ( $e^+e^-$ ) collider at  $\sqrt{s} = 500 \text{ GeV}$  would help. In an electron-positron collider the signals are much cleaner, which gives higher detection efficiency. This lowers the required  $\mathcal{L}_{\text{int}}(5\sigma)$ .

In the last chapter, the performance of the machine learning algorithm was compared to optimized cuts. The resulting required  $\mathcal{L}_{\text{int}}(5\sigma)$  for the deep neural network is a factor of 2.0 lower then for the optimized cuts. The cuts used in this thesis are taken from Figy et al. [8] and these cuts were re-optimized for the simulations done in this thesis. The observable feature distributions were different from Figy et al., which means there could be other optimized cuts that do a better job. A full cut analysis is not conducted as this is beyond the scope of this thesis. The reason the deep neural network performed better is mainly due to the complexity of the neural network, i.e. the ability to find complex correlations in the  $N$  dimensional plane, whereas optimized cuts are restricted by the dimensions that can be visualized. A downside with the use of a deep neural network is the lack of understanding related to what it does to separate the signal and background. This means that we do not get any information what the difference in the processes are.

## 10.2 Outlook

Below are some of the interesting further lines of investigation we identified;

- A detector simulation could be carried out to find out whether this affects the required  $\mathcal{L}_{\text{int}}(5\sigma)$ . This would effectively smear out the data, but studies indicate that neural networks are quite robust against such smearing of data [182].
- Simulate more events to train and test the neural network; this would make it possible for the neural network to learn more complex structure correlations because the network can have more nodes in each layer without overfitting [4]. This could help to lower the uncertainty in  $\mathcal{L}_{\text{int}}(5\sigma)$ , because the uncertainty formula, eq. (7.4.2), is inverse proportional to the number of events.
- Because of the high background removal efficiency, other signatures such as the two-lepton signature could be interesting to study. This gets contribution from the dominant supersymmetry production channel,  $\tilde{\tau}^+\tilde{\tau}^-$ . The two-lepton signature would therefore have a larger signal cross-section.
- It would be interesting to investigate possible tau-sneutrino NLSP specific characteristics to show that a supersymmetry signal, if detected, is from a tau-sneutrino NLSP scenario and not another scenario. This can be done by carefully choosing the observable features, which should be specific for this theory, e.g. the invariant mass of the sleptons.
- The neural network algorithm can be further optimized by other hyperparameters, or, especially in the case of multi class classification, tuning the classification threshold.



# Appendices



# Appendix A

## Notation and identities

### A.1 Notation and conventions

- Natural units are used throughout the thesis;  $c = \hbar = 1$ .
- The einstein summation convention is assumed throughout the text if not noted otherwise. These are denoted as greek letters and are summed from 0 to 3.
- The Minkowski metric  $g^{\mu\nu}$  is defined in the  $(1, -1, -1, -1)$  convention.
- The four-vectors are defined, in terms of the space-time and momentum four-vectors, as

$$x_\mu = g_{\mu\nu}x^\nu = (x^0, -x^1, -x^2, -x^3) = (t, -\mathbf{x}), \quad p^\mu = (E, -\mathbf{p}) \quad (\text{A.1.1})$$

- The space-time derivative is denoted as  $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$ .
- The Feynman slash notation is given by

$$\not{A} \equiv A_\mu \gamma^\mu \quad (\text{A.1.2})$$

- The Dirac conjugate of a fermion field is defined by

$$\bar{\psi} \equiv \psi^\dagger \gamma^0 \quad (\text{A.1.3})$$



## A.2 Identities

### A.2.1 Gamma matrices identities

- The Dirac gamma matrices are defined in the *chiral representation*;

$$\gamma^\mu = \begin{pmatrix} 0 & \bar{\sigma}^\mu \\ \sigma^\mu & 0 \end{pmatrix} \quad (\text{A.2.1})$$

where  $\sigma^\mu$  are the pauli matrices;

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (\text{A.2.2})$$

- Anticommutation relation; Clifford algebra

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \quad (\text{A.2.3})$$

- Trace of odd number of  $\gamma$ -matrices gives

$$\text{Tr}[\gamma^\alpha \gamma^\beta \dots \gamma^\gamma \gamma^\delta] = 0 \quad (\text{A.2.4})$$

- Trace identities for the gamma matrices

$$\text{Tr}[\gamma_\mu \gamma_\nu] = 4g_{\mu\nu}, \quad \text{Tr}[\gamma_\mu \gamma_\nu \gamma_\sigma \gamma_\rho] = 4(g_{\mu\nu} g_{\sigma\rho} - g_{\mu\sigma} g_{\nu\rho} + g_{\mu\rho} g_{\nu\sigma}) \quad (\text{A.2.5})$$

- Gamma 5 matrix and its trace identities

$$\gamma_5 = i\gamma^0 \gamma^1 \gamma^2 \gamma^3 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \quad (\text{A.2.6})$$

$$\text{Tr}[\gamma_5] = \text{Tr}[\gamma_5 \gamma^\alpha] = \text{Tr}[\gamma_5 \gamma^\alpha \gamma^\beta] = \text{Tr}[\gamma_5 \gamma^\alpha \gamma^\beta \gamma^\gamma] = 0, \quad (\text{A.2.7})$$

$$\text{Tr}[\gamma_\mu \gamma_\nu \gamma_\rho \gamma_\sigma \gamma_5] = 4i\epsilon_{\mu\nu\rho\sigma}. \quad (\text{A.2.8})$$

# Appendix B

## Lifetime Calculations

### B.1 2-body decay

The decay rate, taken from PDG review [15], in the rest frame of the decaying particle ( $q = (M, 0, 0, 0)$ )

$$d\Gamma = \frac{1}{32\pi^2} |\mathcal{M}|^2 \frac{|\mathbf{p}_1|}{M^2} d\Omega, \quad (\text{B.1.1})$$

where  $M$  is the mass of the decaying particle,  $\mathcal{M}$  is the matrix element and  $\Omega$  is the solid angle. The momentum is given by [15]

$$|\mathbf{p}_1| = |\mathbf{p}_2| = \frac{1}{2M} \left( (M^2 - (m_1 + m_2)^2)(M^2 - (m_1 - m_2)^2) \right)^{1/2}. \quad (\text{B.1.2})$$

The four vector products where  $q = p_1 + p_2$  are the four vector of the decaying particle

$$\begin{aligned} q \cdot p_1 &= q \cdot (q - p_2) = q \cdot q - q \cdot p_2 = q \cdot q - (p_1 + p_2) \cdot (q - p_1) \\ &= q \cdot q - p_1 \cdot q + p_1 \cdot p_2 + p_2 \cdot q - p_1 \cdot p_2 \\ \Rightarrow 2q \cdot p_1 &= q \cdot q + p_1 \cdot p_1 - p_2 \cdot (q - p_1) = q \cdot q + p_1 \cdot p_1 - p_2 \cdot p_2 \\ \Rightarrow q \cdot p_1 &= \frac{1}{2}(M^2 + m_1^2 - m_2^2), \end{aligned} \quad (\text{B.1.3a})$$

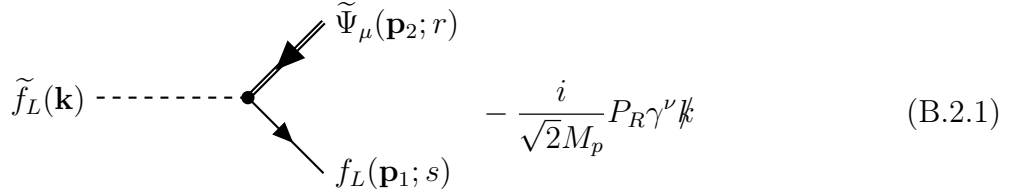
where  $q \cdot q = M^2$  and  $p_i \cdot p_i = m_i^2$  is used. The derivation for the two other products are the same, and the result is

$$q \cdot p_2 = \frac{1}{2}(M^2 - m_1^2 + m_2^2), \quad (\text{B.1.3b})$$

$$p_1 \cdot p_2 = \frac{1}{2}(M^2 - m_1^2 - m_2^2). \quad (\text{B.1.3c})$$

## B.2 Gravitino calculation

The vertex factor is taken from [58]



$$\tilde{f}_L(\mathbf{k}) \text{ --- } \bullet \begin{cases} \nearrow \tilde{\Psi}_\mu(\mathbf{p}_2; r) \\ \searrow f_L(\mathbf{p}_1; s) \end{cases} \quad - \frac{i}{\sqrt{2}M_p} P_R \gamma^\nu \not{k} \quad (\text{B.2.1})$$

The Feynman rules for the external lines are; the initial left-handed sfermion  $\tilde{f}_L(\mathbf{k}) \rightarrow 1$ , the final left-handed fermion  $f_L(\mathbf{p}_1; s) \rightarrow u_s(\mathbf{p}_1)$  and the final gravitino  $\tilde{\Psi}_\mu(\mathbf{p}_2; r) \rightarrow \psi_\mu^{(-)}(\mathbf{p}_2; r)$ . This gives the matrix element

$$\mathcal{M} = -\frac{i}{\sqrt{2}M_p} \bar{u}_s(\mathbf{p}_1) P_R \gamma^\mu \not{k} \psi_\mu^{(-)}(\mathbf{p}_2; r), \quad (\text{B.2.2})$$

and the hermitian conjugate

$$\mathcal{M}^\dagger = \frac{i}{\sqrt{2}M_p} \bar{\psi}_\nu^{(-)}(\mathbf{p}_2; r) \not{k} \gamma^\nu P_L u_s(\mathbf{p}_1). \quad (\text{B.2.3})$$

Squaring the matrix element and summing over the final state polarizations

$$|\mathcal{M}|_{\text{pol}}^2 = \sum_{r,s} \frac{1}{2M_p^2} \bar{\psi}_\nu^{(-)}(\mathbf{p}_2; r) \not{k} \gamma^\nu P_L u_s(\mathbf{p}_1) \bar{u}_s(\mathbf{p}_1) P_R \gamma^\mu \not{k} \psi_\mu^{(-)}(\mathbf{p}_2; r), \quad (\text{B.2.4})$$

using the spin sums

$$\sum_s u_s(p_1) \bar{u}_s(\mathbf{p}_1) = (\not{p}_1 + m_f) \quad (\text{B.2.5})$$

and

$$\begin{aligned}\Pi^{(-)}(p_2)_{\mu\nu} &= \sum_r \psi_\mu^{(-)}(\mathbf{p}_2; r) \bar{\psi}_\nu^{(-)}(\mathbf{p}_2; r) \\ &= -(\not{p}_2 - m_{3/2}) \left\{ g_{\mu\nu} - \frac{1}{3} (\gamma_\mu \gamma_\nu + \gamma_\mu \frac{p_{2\nu}}{m_{3/2}} - \frac{p_{2\mu}}{m_{3/2}} \gamma_\nu + 2 \frac{p_{2\mu} p_{2\nu}}{m_{3/2}^2}) \right\},\end{aligned}\tag{B.2.6}$$

taken from [48]. Taking the trace

$$\begin{aligned}|\mathcal{M}|_{\cancel{P}\cancel{\sigma}}^2 &= \frac{1}{2M_P^2} \text{Tr}[\Pi_{\mu\nu}^{(-)}(p_2) \not{k} \gamma^\nu P_L (\not{p}_1 + m_f) P_R \gamma^\mu \not{k}] \\ &= \frac{1}{2M_P^2} \text{Tr}[\Pi_{\mu\nu}^{(-)}(p_2) \not{k} \gamma^\nu \not{p}_1 \gamma^\mu \not{k} P_R]\end{aligned}\tag{B.2.7}$$

where  $P_L \gamma^\mu = \gamma^\mu P_R$ ,  $P_R P_R = P_R$  and  $P_L P_R = 0$  have been used. Using the anticommutation relation for gamma matrices eq. (A.2.3), and the Rarita-Swinger equation [42]  $\gamma^\mu \Psi_\mu^{(\pm)r} = 0$  and  $\bar{\Psi}_\mu^{(\pm)r} \gamma^\mu = 0$  making the identities for the spinsum  $\gamma^\mu \Pi_{\mu\nu}^{(\pm)} = 0$  and  $\Pi_{\mu\nu}^{(\pm)} \gamma^\nu = 0$ . The expression can be reduced to

$$\begin{aligned}|\mathcal{M}|_{\cancel{P}\cancel{\sigma}}^2 &= \frac{2}{M_P^2} \text{Tr}[\Pi_{\mu\nu}^{(-)}(p_2) k^\mu k^\nu \not{p}_1 P_R] \\ &= \frac{1}{M_P^2} \text{Tr}[-(\not{p}_2 - m_{3/2}) \left\{ g_{\mu\nu} - \frac{1}{3} (\gamma_\mu \gamma_\nu + \gamma_\mu \frac{p_{2\nu}}{m_{3/2}} - \frac{p_{2\mu}}{m_{3/2}} \gamma_\nu + 2 \frac{p_{2\mu} p_{2\nu}}{m_{3/2}^2}) \right\} k^\mu k^\nu \not{p}_1 (1 + \gamma_5)] \\ &= \frac{1}{M_P^2} \text{Tr}[-(\not{p}_2 - m_{3/2}) \left\{ k^2 - \frac{1}{3} (\not{k} \not{k} + \not{k} \frac{(p_2 k)}{m_{3/2}} - \frac{(p_2 k)}{m_{3/2}} \not{k} + 2 \frac{p_{2\mu} p_{2\nu}}{m_{3/2}^2}) \right\} \not{p}_1 (1 + \gamma_5)] \\ &= \frac{1}{M_P^2} \text{Tr}[-(\not{p}_2 - m_{3/2}) \left\{ k^2 - \frac{1}{3} (\not{k} \not{k} + 2 \frac{p_{2\mu} p_{2\nu}}{m_{3/2}^2}) \right\} \not{p}_1 (1 + \gamma_5)]\end{aligned}\tag{B.2.8}$$

Using that the trace identities eqs. (A.2.4) and (A.2.7), multiplying out the parenthesis yields

$$|\mathcal{M}|_{\cancel{P}\cancel{\sigma}}^2 = \frac{1}{M_P^2} \left\{ -\text{Tr}[-\not{p}_2 (k^2 - \frac{2}{3} \frac{(p_2 k)^2}{m_{3/2}^2}) \not{p}_1] + \frac{1}{3} \text{Tr}[\not{p}_2 \not{k} \not{k} \not{p}_1 (1 + \gamma_5)] \right\}\tag{B.2.9}$$

The  $\gamma_5$  term is zero since using the  $\gamma_5$  trace identity eq. (A.2.8), and that the four-momentum to the sneutrino is  $k = (m_{\tilde{\nu}}, 0, 0, 0)$  the levi-cevita will always become

$\epsilon_{\mu 00\sigma} = 0$  which is zero. We get the 4 momentum products

$$\begin{aligned}
|\mathcal{M}|_{\mathcal{P}\sigma}^2 &= \frac{4}{M_P^2} \left\{ -k^2(p_1 p_2) + \frac{2}{3m_{3/2}^2} (p_2 k)^2 (p_1 p_2) + \frac{1}{3} ((p_2 k)(p_1 k) - (p_2 k)(p_1 k) + (p_1 p_2)k^2) \right\} \\
&= \frac{4}{M_P^2} \left\{ \left(\frac{1}{3} - 1\right)k^2(p_1 p_2) + \frac{2}{3m_{3/2}^2} (p_2 k)^2 (p_1 p_2) \right\} \\
&= \frac{4(p_1 p_2)}{M_P^2 m_{3/2}^2} \left\{ -\frac{2}{3}k^2 m_{3/2}^2 + \frac{2}{3}(p_2 k)^2 \right\},
\end{aligned} \tag{B.2.10}$$

using the four vector products (eqs. (B.1.3b) and (B.1.3c)) we arrive at the expression

$$\begin{aligned}
|\mathcal{M}|_{\mathcal{P}\sigma}^2 &= \frac{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)}{3M_P^2 m_{3/2}^2} \left( -4m_{\tilde{f}}^2 m_{3/2}^2 + (m_{\tilde{f}}^2 + m_{3/2}^2 - m_f^2)^2 \right) \\
&= \frac{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)}{3M_P^2 m_{3/2}^2} \left( -4m_{\tilde{f}}^2 m_{3/2}^2 + m_{\tilde{f}}^4 + m_{3/2}^4 + m_f^4 + 2m_{\tilde{f}}^2 m_{3/2}^2 - 2m_{\tilde{f}}^2 m_f^2 - 2m_{3/2}^2 m_f^2 \right) \\
&= \frac{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)}{3M_P^2 m_{3/2}^2} \left( (m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2) - 4m_{3/2}^2 m_f^2 \right) \\
&= \frac{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)^3}{3M_P^2 m_{3/2}^2} \left( 1 - \frac{4m_{3/2}^2 m_f^2}{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)^2} \right).
\end{aligned} \tag{B.2.11}$$

The differential decay rate (eq. (B.1.1)) becomes after integrating over the solid angle  $d\Omega \sim 4\pi$

$$\Gamma = \frac{1}{8\pi} |\mathcal{M}|^2 \frac{|\mathbf{p}_1|}{M^2}. \tag{B.2.12}$$

Using the momentum (eq. (B.1.2))

$$\begin{aligned}
|\mathbf{p}_1| = |\mathbf{p}_2| &= \frac{1}{2m_{\tilde{f}}} \left( (m_{\tilde{f}}^2 - (m_f + m_{3/2})^2)(m_{\tilde{f}}^2 - (m_f - m_{3/2})^2) \right)^{1/2} \\
&= \frac{1}{2m_{\tilde{f}}} \left( (m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2 - 2m_f m_{3/2})(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2 + 2m_f m_{3/2}) \right)^{1/2} \\
&= \frac{1}{2m_{\tilde{f}}} \left( (m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)^2 - 4m_f^2 m_{3/2}^2 \right)^{1/2} \\
&= \frac{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)}{2m_{\tilde{f}}} \left( 1 - \frac{4m_{3/2}^2 m_f^2}{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)^2} \right)^{1/2},
\end{aligned} \tag{B.2.13}$$

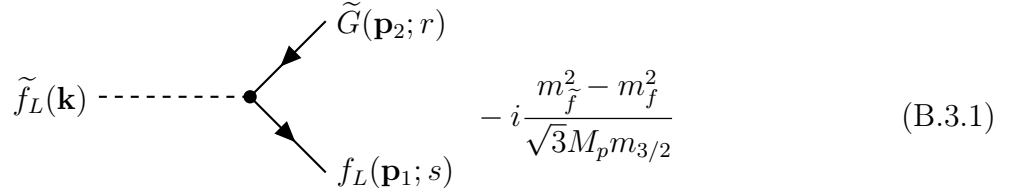
putting this and the matrix element into the decay rate formula (eq. (B.2.12)) and the

decay rate for a sfermion decaying to a gravitino and fermion

$$\Gamma = \frac{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)^4}{48\pi M_P^2 m_{3/2}^2 m_{\tilde{f}}^3} \left( 1 - \frac{4m_{3/2}^2 m_f^2}{(m_{\tilde{f}}^2 - m_{3/2}^2 - m_f^2)^2} \right)^{3/2}. \quad (\text{B.2.14})$$

### B.3 Goldstino calculation

By using the gravitino-goldstino theorem, the sfermion decaying to a gravitino and a fermion can be calculated only using the spin 1/2 part of the gravitino i.e the using the goldstino. The Feynman diagram and vertex factor is given below [\[59\]](#)



$$\tilde{f}_L(\mathbf{k}) \text{ --- } \bullet \begin{cases} \nearrow \tilde{G}(\mathbf{p}_2; r) \\ \searrow f_L(\mathbf{p}_1; s) \end{cases} \quad -i \frac{m_{\tilde{f}}^2 - m_f^2}{\sqrt{3} M_p m_{3/2}} \quad (\text{B.3.1})$$

The Feynman rules for the initial sfermion  $\tilde{f}_L(\mathbf{k}) \rightarrow 1$ , the final fermion  $f_L(\mathbf{p}_1; s) \rightarrow \bar{u}_s(\mathbf{p}_1)$  and the final goldstino  $\tilde{G}(\mathbf{p}_2; r) \rightarrow v_r(\mathbf{p}_2)$ . The matrix element is with its complex conjugate

$$\mathcal{M} = -i \frac{m_{\tilde{f}}^2 - m_f^2}{\sqrt{3} M_p m_{3/2}} \bar{u}_s(\mathbf{p}_1) P_L v_r(\mathbf{p}_2) \quad (\text{B.3.2})$$

$$\mathcal{M}^\dagger = i \frac{m_{\tilde{f}}^2 - m_f^2}{\sqrt{3} M_p m_{3/2}} v_r^\dagger(\mathbf{p}_2) P_L \gamma^0 u_s(\mathbf{p}_1) = i \frac{m_{\tilde{f}}^2 - m_f^2}{\sqrt{3} M_p m_{3/2}} \bar{v}_r(\mathbf{p}_2) P_R u_s(\mathbf{p}_1), \quad (\text{B.3.3})$$

where  $P_L \gamma^0 = \gamma^0 P_R$  is used to get the hermitian conjugate. Multiplying with the complex conjugate and summing over the polarizations

$$\sum_{s,r} |\mathcal{M}|_{\mathcal{P}\mathcal{O}}^2 = \frac{(m_{\tilde{f}}^2 - m_f^2)^2}{3M_p^2 m_{3/2}^2} \sum_{s,r} \bar{v}_r(\mathbf{p}_2) P_R u_s(\mathbf{p}_1) \bar{u}_s(\mathbf{p}_1) P_L v_r(\mathbf{p}_2). \quad (\text{B.3.4})$$

Taking the trace, and using the projection operator identity  $P_L P_L = P_L$  (using  $P_L \gamma^0 = \gamma^0 P_R$  again)

$$\sum_{s,r} |\mathcal{M}|_{\mathcal{P}\mathcal{O}}^2 = \frac{(m_{\tilde{f}}^2 - m_f^2)^2}{3M_p^2 m_{3/2}^2} \text{Tr} \left[ \sum_r v_r(\mathbf{p}_2) \bar{v}_r(\mathbf{p}_2) \sum_s u_s(\mathbf{p}_1) \bar{u}_s(\mathbf{p}_1) P_L \right]. \quad (\text{B.3.5})$$

Using the spin sums for fermions [48]  $\sum_r v_r(p)\bar{v}_r(p) = (\not{p} - m)$  and  $\sum_s u_s(p)\bar{u}_s(p) = (\not{p} + m)$

$$|\mathcal{M}|_{\mathcal{P}\mathcal{A}}^2 = \frac{(m_{\tilde{f}}^2 - m_f^2)^2}{3M_p^2 m_{3/2}^2} \text{Tr}[(\not{p}_2 - m_{3/2})(\not{p}_1 + m_f) \frac{(1 - \gamma_5)}{2}]. \quad (\text{B.3.6})$$

Evaluating the traces by using the trace identities (eqs. (A.2.4), (A.2.5) and (A.2.7)) and multiplying out the parenthesis yields

$$|\mathcal{M}|_{\mathcal{P}\mathcal{A}}^2 = \frac{1}{2} \frac{(m_{\tilde{f}}^2 - m_f^2)^2}{3M_p^2 m_{3/2}^2} \left( \text{Tr}[\not{p}_2 \not{p}_2] - 4m_{3/2} m_f \right) = \frac{1}{2} \frac{(m_{\tilde{f}}^2 - m_f^2)^2}{3M_p^2 m_{3/2}^2} (4(p_1 \cdot p_2) - 4m_{3/2} m_f). \quad (\text{B.3.7})$$

Since  $k = p_1 + p_2$  and  $k^2 = m_{\tilde{f}}^2$ ,  $p_1^2 = m_f^2$ ,  $p_2^2 = m_{3/2}^2$ , the four vector product (eq. (B.1.3c)) gives

$$|\mathcal{M}|_{\mathcal{P}\mathcal{A}}^2 = \frac{m_{\tilde{f}}^2 (m_{\tilde{f}}^2 - m_f^2)^2}{3M_p^2 m_{3/2}^2} \left( 1 - \frac{(m_{3/2} + m_f)^2}{m_{\tilde{f}}^2} \right). \quad (\text{B.3.8})$$

The decay rate and the momentum are the same as in previous chapter, putting this and the matrix element into the decay rate

$$\begin{aligned} \Gamma &= \frac{m_{\tilde{f}}^2 (m_{\tilde{f}}^2 - m_f^2)^2}{24\pi M_p^2 m_{3/2}^2} \left( 1 - \frac{(m_f + m_{3/2})^2}{m_{\tilde{f}}^2} \right) \frac{m_{\tilde{f}}^2}{2m_{\tilde{f}}^3} \left( \left( 1 - \frac{(m_f + m_{3/2})^2}{m_{\tilde{f}}^2} \right) \left( 1 - \frac{(m_f - m_{3/2})^2}{m_{\tilde{f}}^2} \right) \right)^{1/2} \\ &= \frac{m_{\tilde{f}} (m_{\tilde{f}}^2 - m_f^2)^2}{48\pi M_p^2 m_{3/2}^2} \left( 1 - \frac{(m_f + m_{3/2})^2}{m_{\tilde{f}}^2} \right)^{3/2} \left( 1 - \frac{(m_f - m_{3/2})^2}{m_{\tilde{f}}^2} \right)^{1/2}. \end{aligned} \quad (\text{B.3.9})$$

This is the decay rate for a sfermion decaying to a goldstino/gravitino and a fermion.

# Bibliography

- [1] R. Santos, M. Nguyen, J. Webster, S. Ryu, J. Adelman, S. Chekanov, and J. Zhou, “Machine learning techniques in searches for  $t\bar{t}h$  in the  $h \rightarrow b\bar{b}$  decay channel,” JINST **12** (2017), no. 04, P04014, [1610.03088](#).
- [2] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, “Parameterized neural networks for high-energy physics,” Eur. Phys. J. **C76** (2016), no. 5, 235, [1601.07913](#).
- [3] P. Baldi, P. Sadowski, and D. Whiteson, “Enhanced Higgs Boson to  $\tau^+\tau^-$  Search with Deep Learning,” Phys. Rev. Lett. **114** (2015), no. 11, 111801, [1410.3469](#).
- [4] P. Baldi, P. Sadowski, and D. Whiteson, “Searching for Exotic Particles in High-Energy Physics with Deep Learning,” Nature Commun. **5** (2014) 4308, [1402.4735](#).
- [5] T. Falk, K. A. Olive, and M. Srednicki, “Heavy sneutrinos as dark matter,” Phys. Lett. **B339** (1994) 248–251, [hep-ph/9409270](#).
- [6] C. Arina and N. Fornengo, “Sneutrino cold dark matter, a new analysis: Relic abundance and detection rates,” JHEP **11** (2007) 029, [0709.4477](#).
- [7] J. R. Ellis, K. A. Olive, and Y. Santoso, “Sneutrino NLSP Scenarios in the NUHM with Gravitino Dark Matter,” JHEP **10** (2008) 005, [0807.3736](#).
- [8] T. Figy, K. Rolbiecki, and Y. Santoso, “Tau-Sneutrino NLSP and Multilepton Signatures at the LHC,” Phys. Rev. **D82** (2010) 075016, [1005.5136](#).
- [9] L. Covi and S. Kraml, “Collider signatures of gravitino dark matter with a sneutrino NLSP,” JHEP **08** (2007) 015, [hep-ph/0703130](#).
- [10] A. Katz and B. Tweedie, “Leptophilic Signals of a Sneutrino (N)LSP and Flavor Biases from Flavor-Blind SUSY,” Phys. Rev. **D81** (2010) 115003, [1003.5664](#).



- [11] M. E. Peskin and D. V. Schroeder, An Introduction to quantum field theory. Addison-Wesley, Reading, USA, 1995.
- [12] F. Mandl and G. Shaw, Quantum Field Theory. A Wiley-Interscience publication. John Wiley & Sons, 2010.
- [13] S. P. Martin, “A Supersymmetry primer,” [hep-ph/9709356](#). [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [14] **Super-Kamiokande** Collaboration, Y. Fukuda et al., “Evidence for oscillation of atmospheric neutrinos,” Phys. Rev. Lett. **81** (1998) 1562–1567, [hep-ex/9807003](#).
- [15] **Particle Data Group** Collaboration, C. Patrignani et al., “Review of Particle Physics,” Chin. Phys. **C40** (2016), no. 10, 100001.
- [16] S. Weinberg, “A Model of Leptons,” Phys. Rev. Lett. **19** (1967) 1264–1266.
- [17] A. Salam and J. C. Ward, “Electromagnetic and weak interactions,” Phys. Lett. **13** (1964) 168–171.
- [18] S. L. Glashow, “Partial Symmetries of Weak Interactions,” Nucl. Phys. **22** (1961) 579–588.
- [19] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, “Advantages of the Color Octet Gluon Picture,” Phys. Lett. **47B** (1973) 365–368.
- [20] P. W. Higgs, “Broken symmetries, massless particles and gauge fields,” Phys. Lett. **12** (1964) 132–133.
- [21] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons,” Phys. Rev. Lett. **13** (1964) 321–323.
- [22] E. Noether, “Invariante variationsprobleme,” Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1918** (1918) 235–257.
- [23] S. Weinberg, “Effective Gauge Theories,” Phys. Lett. **91B** (1980) 51–55.
- [24] S. R. Coleman, J. Wess, and B. Zumino, “Structure of phenomenological Lagrangians. 1.,” Phys. Rev. **177** (1969) 2239–2247.
- [25] H. Georgi and S. L. Glashow, “Unity of All Elementary Particle Forces,” Phys. Rev. Lett. **32** (1974) 438–441.

- [26] N. Jeevanjee, Groups, Lie Groups, and Lie Algebras, pp. 87–143. Birkhäuser Boston, Boston, 2011.
- [27] S. R. Coleman and J. Mandula, “All Possible Symmetries of the S Matrix,” Phys. Rev. **159** (1967) 1251–1256.
- [28] R. Haag, J. T. Lopuszanski, and M. Sohnius, “All Possible Generators of Supersymmetries of the s Matrix,” Nucl. Phys. **B88** (1975) 257. [,257(1974)].
- [29] H. Baer and X. Tata, Weak Scale Supersymmetry: From Superfields to Scattering Events. Cambridge University Press, 2006.
- [30] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries,” Phys. Rev. **127** (1962) 965–970.
- [31] **ATLAS, CMS** Collaboration, G. Aad *et al.*, “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. **114** (2015) 191803, [1503.07589](#).
- [32] J. Ellis, J. Hagelin, and D. Nanopoulos, “Spin-zero leptons and the anomalous magnetic moment of the muon,” Physics Letters B **116** (1982), no. 4, 283 – 286.
- [33] R. Barbieri, S. Ferrara, D. Nanopoulos, and K. Stelle, “Supergravity,  $r$  invariance and spontaneous supersymmetry breaking,” Physics Letters B **113** (1982), no. 3, 219 – 222.
- [34] A. H. Chamseddine, R. L. Arnowitt, and P. Nath, “Locally Supersymmetric Grand Unification,” Phys. Rev. Lett. **49** (1982) 970.
- [35] S. Dimopoulos and S. Raby, “Supercolor,” Nucl. Phys. **B192** (1981) 353–368.
- [36] Z. Chacko, M. A. Luty, A. E. Nelson, and E. Ponton, “Gaugino mediated supersymmetry breaking,” JHEP **01** (2000) 003, [hep-ph/9911323](#).
- [37] P. Fayet and J. Iliopoulos, “Spontaneously Broken Supergauge Symmetries and Goldstone Spinors,” Phys. Lett. **51B** (1974) 461–464.
- [38] T. Lee and G.-H. Wu, “Interactions of a single goldstino,” Phys. Lett. **B447** (1999) 83–88, [hep-ph/9805512](#).
- [39] S. Weinberg, “A model of leptons,” Phys. Rev. Lett. **19** (Nov, 1967) 1264–1266.

- [40] P. Nath and R. Arnowitt, “Generalized super-gauge symmetry as a new framework for unified gauge theories,” Physics Letters B **56** (1975), no. 2, 177 – 180.
- [41] D. Z. Freedman, P. van Nieuwenhuizen, and S. Ferrara, “Progress Toward a Theory of Supergravity,” Phys. Rev. **D13** (1976) 3214–3218.
- [42] W. Rarita and J. Schwinger, “On a theory of particles with half-integral spin,” Phys. Rev. **60** (Jul, 1941) 61–61.
- [43] S. Ferrara, J. Scherk, and P. van Nieuwenhuizen, “Locally Supersymmetric Maxwell-Einstein Theory,” Physical Review Letters **37** (oct, 1976) 1035–1037.
- [44] V. S. Rychkov and A. Strumia, “Thermal production of gravitinos,” Phys. Rev. **D75** (2007) 075011, [hep-ph/0701104](#).
- [45] H. Nilles, “Supersymmetry, supergravity and particle physics,” Physics Reports **110** (aug, 1984) 1–162.
- [46] S. D. Deser and B. Zumino, “Broken supersymmetry and supergravity,” Phys. Rev. Lett. **38** (1977), no. PRINT-77-0426-BRANDEIS, 1433.
- [47] E. Cremmer, B. Julia, J. Scherk, S. Ferrara, L. Girardello, and P. van Nieuwenhuizen, “Spontaneous symmetry breaking and higgs effect in supergravity without cosmological constant,” Nuclear Physics B **147** (1979), no. 1, 105 – 131.
- [48] N. D. Christensen, P. de Aquino, N. Deutschmann, C. Duhr, B. Fuks, C. Garcia-Cely, O. Mattelaer, K. Mawatari, B. Oexl, and Y. Takaesu, “Simulating spin- $\frac{3}{2}$  particles at colliders,” Eur. Phys. J. **C73** (2013), no. 10, 2580, [1308.1668](#).
- [49] P. Fayet, “Mixing Between Gravitational and Weak Interactions Through the Massive Gravitino,” Phys. Lett. **70B** (1977) 461.
- [50] R. Casalbuoni, S. De Curtis, D. Dominici, F. Feruglio, and R. Gatto, “A gravitino - goldstino high-energy equivalence theorem,” Phys. Lett. **B215** (1988) 313–316.
- [51] M. Kawasaki and T. Moroi, “Gravitino decay into a neutrino and a sneutrino in the inflationary universe,” Phys. Lett. **B346** (1995) 27–34, [hep-ph/9408321](#).

- [52] R. A. Alpher, H. Bethe, and G. Gamow, “The Origin of Chemical Elements,” Physical Review **73** (apr, 1948) 803–804.
- [53] R. V. Wagoner, W. A. Fowler, and F. Hoyle, “On the Synthesis of Elements at Very High Temperatures,” The Astrophysical Journal **148** (apr, 1967) 3.
- [54] T. Moroi, H. Murayama, and M. Yamaguchi, “Cosmological constraints on the light stable gravitino,” Physics Letters B **303** (apr, 1993) 289–294.
- [55] S. Weinberg, “Cosmological Constraints on the Scale of Supersymmetry Breaking,” Phys. Rev. Lett. **48** (1982) 1303.
- [56] L. Covi, J. Hasenkamp, S. Pokorski, and J. Roberts, “Gravitino Dark Matter and general neutralino NLSP,” JHEP **11** (2009) 003, [0908.3399](#).
- [57] M. Pospelov, “Particle physics catalysis of thermal Big Bang Nucleosynthesis,” Phys. Rev. Lett. **98** (2007) 231301, [hep-ph/0605215](#).
- [58] M. Grefe, Neutrino signals from gravitino dark matter with broken R-parity. PhD thesis, Hamburg U., 2008. [1111.6041](#).
- [59] M. Klasen and G. Pignol, “New Results for Light Gravitinos at Hadron Colliders: Tevatron Limits and LHC Perspectives,” Phys. Rev. **D75** (2007) 115003, [hep-ph/0610160](#).
- [60] D. Matalliotakis and H. P. Nilles, “Implications of nonuniversality of soft terms in supersymmetric grand unified theories,” Nucl. Phys. **B435** (1995) 115–128, [hep-ph/9407251](#).
- [61] S. P. Martin and M. T. Vaughn, “Two loop renormalization group equations for soft supersymmetry breaking couplings,” Phys. Rev. **D50** (1994) 2282, [hep-ph/9311340](#). [Erratum: Phys. Rev.D78,039903(2008)].
- [62] J. R. Ellis, T. Falk, K. A. Olive, and Y. Santoso, “Exploration of the MSSM with nonuniversal Higgs masses,” Nucl. Phys. **B652** (2003) 259–347, [hep-ph/0210205](#).
- [63] W. Porod, “SPheno, a program for calculating supersymmetric spectra, SUSY particle decays and SUSY particle production at e+ e- colliders,” Comput. Phys. Commun. **153** (2003) 275–315, [hep-ph/0301101](#).

- [64] W. Porod and F. Staub, “SPheno 3.1: Extensions including flavour, CP-phases and models beyond the MSSM,” *Comput. Phys. Commun.* **183** (2012) 2458–2469, [1104.1573](#).
- [65] T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, and G. Weiglein, “High-precision predictions for the light CP-even Higgs Boson Mass of the MSSM,” [arXiv:1312.4937v1](#).
- [66] S. Borowka, T. Hahn, S. Heinemeyer, G. Heinrich, and W. Hollik, “Momentum-dependent two-loop QCD corrections to the neutral Higgs-boson masses in the MSSM,” [arXiv:1404.7074v2](#).
- [67] J. Heisig, J. Kersten, N. Murphy, and I. Strümke, “Trilinear-augmented gaugino mediation,” *JHEP* **17** (may, 2017) 3.
- [68] **ATLAS** Collaboration, G. Aad *et al.*, “Searches for heavy long-lived charged particles with the ATLAS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV,” *JHEP* **01** (2015) 068, [1411.6795](#).
- [69] **CMS** Collaboration, S. Chatrchyan *et al.*, “Searches for long-lived charged particles in pp collisions at  $\sqrt{s}=7$  and 8 TeV,” *JHEP* **07** (2013) 122, [1305.0491](#).
- [70] L. Selbuz and Z. Z. Aydin, “Tau sleptons and Tau sneutrino Decays in MSSM under the Cosmological Bounds,” [0808.2540](#).
- [71] D. Dercks, N. Desai, J. S. Kim, K. Rolbiecki, J. Tattersall, and T. Weber, “CheckMATE 2: From the model to the limit,” *Comput. Phys. Commun.* **221** (2017) 383–418, [1611.09856](#).
- [72] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual,” *Eur. Phys. J.* **C72** (2012) 1896, [1111.6097](#).
- [73] M. Cacciari, G. P. Salam, and G. Soyez, “The Anti-k(t) jet clustering algorithm,” *JHEP* **04** (2008) 063, [0802.1189](#).
- [74] M. Cacciari and G. P. Salam, “Dispelling the  $N^3$  myth for the  $k_t$  jet-finder,” *Phys. Lett.* **B641** (2006) 57–61, [hep-ph/0512210](#).
- [75] Y. Bai, H.-C. Cheng, J. Gallicchio, and J. Gu, “Stop the Top Background of the Stop Search,” *JHEP* **07** (2012) 110, [1203.4813](#).
- [76] M. L. Graesser and J. Shelton, “Hunting Mixed Top Squark Decays,” *Phys. Rev. Lett.* **111** (2013), no. 12, 121802, [1212.4495](#).

- [77] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP* **05** (2006) 026, [hep-ph/0603175](#).
- [78] **DELPHES 3** Collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi, “DELPHES 3, A modular framework for fast simulation of a generic collider experiment,” *JHEP* **02** (2014) 057, [1307.6346](#).
- [79] A. L. Read, “Presentation of search results: The CL(s) technique,” *J. Phys.* **G28** (2002) 2693–2704. [,11(2002)].
- [80] **ATLAS** Collaboration, G. Aad *et al.*, “Search for supersymmetry at  $\sqrt{s} = 13$  TeV in final states with jets and two same-sign leptons or three leptons with the ATLAS detector,” *Eur. Phys. J.* **C76** (2016), no. 5, 259, [1602.09058](#).
- [81] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for new phenomena in events with a photon and missing transverse momentum in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector,” *JHEP* **06** (2016) 059, [1604.01306](#).
- [82] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for squarks and gluinos in final states with jets and missing transverse momentum at  $\sqrt{s} = 13$  TeV with the ATLAS detector,” *Eur. Phys. J.* **C76** (2016), no. 7, 392, [1605.03814](#).
- [83] **ATLAS** Collaboration, G. Aad *et al.*, “Search for gluinos in events with an isolated lepton, jets and missing transverse momentum at  $\sqrt{s} = 13$  TeV with the ATLAS detector,” *Eur. Phys. J.* **C76** (2016), no. 10, 565, [1605.04285](#).
- [84] **ATLAS** Collaboration, G. Aad *et al.*, “Search for pair production of gluinos decaying via stop and sbottom in events with  $b$ -jets and large missing transverse momentum in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector,” *Phys. Rev.* **D94** (2016), no. 3, 032003, [1605.09318](#).
- [85] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for top squarks in final states with one isolated lepton, jets, and missing transverse momentum in  $\sqrt{s} = 13$  TeV  $pp$  collisions with the ATLAS detector,” *Phys. Rev.* **D94** (2016), no. 5, 052009, [1606.03903](#).
- [86] **ATLAS** Collaboration, T. A. collaboration, “Search for  $W' \rightarrow tb$  in the hadronic final state with the ATLAS Detector in  $\sqrt{s} = 13$  TeV  $pp$  collisions,”

- [87] **ATLAS** Collaboration, T. A. collaboration, “Search for new phenomena in a lepton plus high jet multiplicity final state with the ATLAS experiment using  $\sqrt{s} = 13$  TeV proton–proton collision data,”.
- [88] **ATLAS** Collaboration, T. A. collaboration, “Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in  $36.1 \text{ fb}^{-1}$  of  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector,”.
- [89] **ATLAS** Collaboration, T. A. collaboration, “Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a  $b\bar{b}$  pair in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector,”.
- [90] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for dark matter at  $\sqrt{s} = 13$  TeV in final states containing an energetic photon and large missing transverse momentum with the ATLAS detector,” *Eur. Phys. J.* **C77** (2017), no. 6, 393, [1704.03848](#).
- [91] D. Borah, A. Dasgupta, U. K. Dey, S. Patra, and G. Tomar, “Multi-component Fermionic Dark Matter and IceCube PeV scale Neutrinos in Left-Right Model with Gauge Unification,” *JHEP* **09** (2017) 005, [1704.04138](#).
- [92] **ATLAS** Collaboration, T. A. collaboration, “Search for Supersymmetry in events with photons, jets and missing transverse energy with the ATLAS detector in 13 TeV  $pp$  collisions,”.
- [93] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for the direct production of charginos and neutralinos in final states with tau leptons in  $\sqrt{s} = 13$  TeV  $pp$  collisions with the ATLAS detector,” *Eur. Phys. J.* **C78** (2018), no. 2, 154, [1708.07875](#).
- [94] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for photonic signatures of gauge-mediated supersymmetry in 13 TeV  $pp$  collisions with the ATLAS detector,” [1802.03158](#).
- [95] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for squarks and gluinos in final states with jets and missing transverse momentum using  $36 \text{ fb}^{-1}$  of  $\sqrt{s} = 13$  TeV  $pp$  collision data with the ATLAS detector,” [1712.02332](#).
- [96] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for electroweak production of supersymmetric states in scenarios with compressed mass spectra at  $\sqrt{s} = 13$  TeV with the ATLAS detector,” *Phys. Rev.* **D97** (2018), no. 5, 052010, [1712.08119](#).

- [97] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for supersymmetry in a final state containing two photons and missing transverse momentum in  $\sqrt{s} = 13$  TeV  $pp$  collisions at the LHC using the ATLAS detector,” *Eur. Phys. J.* **C76** (2016), no. 9, 517, [1606.09150](#).
- [98] **ATLAS** Collaboration, T. A. collaboration, “Search for squarks and gluinos in events with an isolated lepton, jets and missing transverse momentum at  $\sqrt{s} = 13$  TeV with the ATLAS detector,”.
- [99] **ATLAS** Collaboration, T. A. collaboration, “Further searches for squarks and gluinos in final states with jets and missing transverse momentum at  $\sqrt{s} = 13$  TeV with the ATLAS detector,”.
- [100] **ATLAS** Collaboration, T. A. collaboration, “Search for electroweak production of supersymmetric particles in the two and three lepton final state at  $\sqrt{s} = 13$  TeV with the ATLAS detector,”.
- [101] **ATLAS** Collaboration, T. A. collaboration, “Search for direct top squark pair production in events with a Higgs or  $Z$  boson, and missing transverse momentum in  $\sqrt{s} = 13$  TeV  $pp$  collisions with the ATLAS detector,”.
- [102] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for dark matter produced in association with bottom or top quarks in  $\sqrt{s} = 13$  TeV  $pp$  collisions with the ATLAS detector,” *Eur. Phys. J.* **C78** (2018), no. 1, 18, [1710.11412](#).
- [103] **ATLAS** Collaboration, G. Aad *et al.*, “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* **10** (2013) 130, [1308.1841](#). [Erratum: *JHEP*01,109(2014)].
- [104] **ATLAS** Collaboration, G. Aad *et al.*, “Search for direct third-generation squark pair production in final states with missing transverse momentum and two  $b$ -jets in  $\sqrt{s} = 8$  TeV  $pp$  collisions with the ATLAS detector,” *JHEP* **10** (2013) 189, [1308.2631](#).
- [105] **ATLAS** Collaboration, G. Aad *et al.*, “Search for direct production of charginos and neutralinos in events with three leptons and missing transverse momentum in  $\sqrt{s} = 8$  TeV  $pp$  collisions with the ATLAS detector,” *JHEP* **04** (2014) 169, [1402.7029](#).



- [106] **ATLAS** Collaboration, G. Aad *et al.*, “Search for direct top-squark pair production in final states with two leptons in pp collisions at  $\sqrt{s} = 8\text{TeV}$  with the ATLAS detector,” *JHEP* **06** (2014) 124, [1403.4853](#).
- [107] **ATLAS** Collaboration, G. Aad *et al.*, “Search for direct top squark pair production in events with a Z boson, b-jets and missing transverse momentum in  $\sqrt{s}=8$  TeV pp collisions with the ATLAS detector,” *Eur. Phys. J.* **C74** (2014), no. 6, 2883, [1403.5222](#).
- [108] **ATLAS** Collaboration, G. Aad *et al.*, “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* **06** (2014) 035, [1404.2500](#).
- [109] **ATLAS** Collaboration, G. Aad *et al.*, “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* **09** (2014) 176, [1405.7875](#).
- [110] **ATLAS** Collaboration, G. Aad *et al.*, “Search for top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in  $\sqrt{s} = 8$  TeV *pp* collisions with the ATLAS detector,” *JHEP* **11** (2014) 118, [1407.0583](#).
- [111] **ATLAS** Collaboration, G. Aad *et al.*, “Search for pair-produced third-generation squarks decaying via charm quarks or in compressed supersymmetric scenarios in *pp* collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector,” *Phys. Rev.* **D90** (2014), no. 5, 052008, [1407.0608](#).
- [112] **ATLAS** Collaboration, G. Aad *et al.*, “Search for new phenomena in events with a photon and missing transverse momentum in *pp* collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector,” *Phys. Rev.* **D91** (2015), no. 1, 012008, [1411.1559](#). [Erratum: *Phys. Rev.*D92,no.5,059903(2015)].
- [113] **ATLAS** Collaboration, G. Aad *et al.*, “Search for direct pair production of a chargino and a neutralino decaying to the 125 GeV Higgs boson in  $\sqrt{s} = 8$  TeV *pp* collisions with the ATLAS detector,” *Eur. Phys. J.* **C75** (2015), no. 5, 208, [1501.07110](#).
- [114] **ATLAS** Collaboration, G. Aad *et al.*, “Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp

- collisions at  $\sqrt{s}=8$  TeV with the ATLAS detector,” Eur. Phys. J. **C75** (2015), no. 7, 299, [1502.01518](#). [Erratum: Eur. Phys. J.C75,no.9,408(2015)].
- [115] **ATLAS** Collaboration, G. Aad *et al.*, “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,” Eur. Phys. J. **C75** (2015), no. 7, 318, [1503.03290](#). [Erratum: Eur. Phys. J.C75,no.10,463(2015)].
- [116] **ATLAS** Collaboration, G. Aad *et al.*, “ATLAS Run 1 searches for direct pair production of third-generation squarks at the Large Hadron Collider,” Eur. Phys. J. **C75** (2015), no. 10, 510, [1506.08616](#). [Erratum: Eur. Phys. J.C76,no.3,153(2016)].
- [117] **ATLAS** Collaboration, “Search for supersymmetry at  $\sqrt{s} = 8$  TeV in final states with jets, missing transverse momentum and one isolated lepton,”.
- [118] **ATLAS** Collaboration, “Search for direct production of the top squark in the all-hadronic  $t\bar{t}b + e\text{miss}$  final state in 21 fb-1 of p-p collisions at  $\sqrt{s}=8$  TeV with the ATLAS detector,”.
- [119] **ATLAS** Collaboration, “Search for direct production of charginos and neutralinos in events with three leptons and missing transverse momentum in 21 fb<sup>-1</sup> of pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector,”.
- [120] **ATLAS** Collaboration, T. A. collaboration, “Search for direct-slepton and direct-chargino production in final states with two opposite-sign leptons, missing transverse momentum and no jets in 20/fb of pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector,”.
- [121] **ATLAS** Collaboration, T. A. collaboration, “Search for strongly produced supersymmetric particles in decays with two leptons at  $\sqrt{s} = 8$  TeV,”.
- [122] **ATLAS** Collaboration, T. A. collaboration, “Measurement of jet fragmentation in 5.02 TeV proton-lead and proton-proton collisions with the ATLAS detector,”.
- [123] **ATLAS** Collaboration, G. Aad *et al.*, “Search for photonic signatures of gauge-mediated supersymmetry in 8 TeV pp collisions with the ATLAS detector,” Phys. Rev. **D92** (2015), no. 7, 072001, [1507.05493](#).

- [124] **ATLAS** Collaboration, G. Aad *et al.*, “Search for the direct production of charginos, neutralinos and staus in final states with at least two hadronically decaying taus and missing transverse momentum in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector,” *JHEP* **10** (2014) 096, [1407.0350](#).
- [125] J. Bellm *et al.*, “Herwig 7.0/Herwig++ 3.0 release note,” *Eur. Phys. J.* **C76** (2016), no. 4, 196, [1512.01178](#).
- [126] **ATLAS, CMS** Collaboration, P. de Jong, “Supersymmetry searches at the LHC,” in *Proceedings, 32nd International Symposium on Physics in Collision (PIC 2012): Strbske Pleso, SI* pp. 241–254. 2012. [1211.3887](#).
- [127] **CMS** Collaboration, A. M. Sirunyan *et al.*, “Search for third-generation scalar leptoquarks and heavy right-handed neutrinos in final states with two tau leptons and two jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV,” *JHEP* **07** (2017) 121, [1703.03995](#).
- [128] **ATLAS** Collaboration, G. Aad *et al.*, “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13$  TeV,” *Eur. Phys. J.* **C76** (2016), no. 5, 292, [1603.05598](#).
- [129] **ATLAS** Collaboration, G. Aad *et al.*, “Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data,” *Eur. Phys. J.* **C74** (2014), no. 7, 2941, [1404.2240](#).
- [130] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, “Parton distributions and the LHC:  $W$  and  $Z$  production,” *Eur. Phys. J.* **C14** (2000) 133–145, [hep-ph/9907231](#).
- [131] F. Cascioli, T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhöfer, A. von Manteuffel, S. Pozzorini, D. Rathlev, L. Tancredi, and E. Weihs, “ZZ production at hadron colliders in NNLO QCD,” *Phys. Lett.* **B735** (2014) 311–313, [1405.2219](#).
- [132] M. Grazzini, S. Kallweit, S. Pozzorini, D. Rathlev, and M. Wiesemann, “production at the LHC: fiducial cross sections and distributions in NNLO QCD,” *JHEP08* (2016).
- [133] A. Buckley *et al.*, “General-purpose event generators for LHC physics,” *Phys. Rept.* **504** (2011) 145–233, [1101.2599](#).

- [134] R. K. Ellis, W. J. Stirling, and B. R. Webber, “QCD and collider physics,” Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8** (1996) 1–435.
- [135] W. A. Bardeen, A. J. Buras, D. W. Duke, and T. Muta, “Deep Inelastic Scattering Beyond the Leading Order in Asymptotically Free Gauge Theories,” Phys. Rev. **D18** (1978) 3998.
- [136] R. Placakyte, “Parton Distribution Functions,” 2011. [1111.5452](#).
- [137] J. D. Bjorken and E. A. Paschos, “Inelastic Electron Proton and gamma Proton Scattering, and the Structure of the Nucleon,” Phys. Rev. **185** (1969) 1975–1982.
- [138] S. Höche, “Introduction to parton-shower event generators,” pp. 235–295. 2015. [1411.4085](#).
- [139] M. Bengtsson and T. Sjostrand, “Coherent Parton Showers Versus Matrix Elements: Implications of PETRA - PEP Data,” Phys. Lett. **B185** (1987) 435.
- [140] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, “QCD matrix elements + parton showers,” JHEP **11** (2001) 063, [hep-ph/0109231](#).
- [141] L. Lonnblad, “Correcting the color dipole cascade model with fixed order matrix elements,” JHEP **05** (2002) 046, [hep-ph/0112284](#).
- [142] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, “Parton Fragmentation and String Dynamics,” Phys. Rept. **97** (1983) 31–145.
- [143] G. C. Fox and S. Wolfram, “A Model for Parton Showers in QCD,” Nucl. Phys. **B168** (1980) 285–295.
- [144] G. P. Salam and G. Soyez, “A Practical Seedless Infrared-Safe Cone jet algorithm,” JHEP **05** (2007) 086, [0704.0292](#).
- [145] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, “Longitudinally invariant  $K_t$  clustering algorithms for hadron hadron collisions,” Nucl. Phys. **B406** (1993) 187–224.
- [146] **CMS Collaboration** Collaboration, “A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging,” Tech. Rep. CMS-PAS-JME-09-001, CERN, Geneva, Jul, 2009.

- [147] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, “Event generation with SHERPA 1.1,” *JHEP* **02** (2009) 007, [0811.4622](#).
- [148] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, “LHAPDF6: parton density access in the LHC precision era,” *Eur. Phys. J.* **C75** (2015) 132, [1412.7420](#).
- [149] F. Krauss, R. Kuhn, and G. Soff, “AMEGIC++ 1.0: A Matrix element generator in C++,” *JHEP* **02** (2002) 044, [hep-ph/0109036](#).
- [150] T. Gleisberg and S. Hoeche, “Comix, a new matrix element generator,” *JHEP* **12** (2008) 039, [0808.3674](#).
- [151] S. Hoeche, F. Krauss, S. Schumann, and F. Siegert, “QCD matrix elements and truncated showers,” *JHEP* **05** (2009) 053, [0903.1219](#).
- [152] M. Bahr et al., “Herwig++ Physics and Manual,” *Eur. Phys. J.* **C58** (2008) 639–707, [0803.0883](#).
- [153] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, “Parton distributions in the LHC era: MMHT 2014 PDFs,” *Eur. Phys. J.* **C75** (2015), no. 5, 204, [1412.3989](#).
- [154] **NNPDF** Collaboration, R. D. Ball et al., “Parton distributions for the LHC Run II,” *JHEP* **04** (2015) 040, [1410.8849](#).
- [155] F. Cascioli, P. Maierhofer, and S. Pozzorini, “Scattering Amplitudes with Open Loops,” *Phys. Rev. Lett.* **108** (2012) 111601, [1111.5206](#).
- [156] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, “A critical appraisal of NLO+PS matching methods,” *JHEP* **09** (2012) 049, [1111.1220](#).
- [157] M. Grazzini, S. Kallweit, D. Rathlev, and M. Wiesemann, “ $W^\pm Z$  production at hadron colliders in NNLO QCD,” *Phys. Lett.* **B761** (2016) 179–183, [1604.08576](#).
- [158] A. Broggio, A. Ferroglia, G. Ossola, B. D. Pecjak, and R. D. Sameshima, “Associated production of a top pair and a Z boson at the LHC to NNLL accuracy,” *JHEP* **04** (2017) 105, [1702.00800](#).

- [159] A. Broggio, A. Ferroglia, G. Ossola, and B. D. Pecjak, “Associated production of a top pair and a W boson at next-to-next-to-leading logarithmic accuracy,” *JHEP* **09** (2016) 089, [1607.05303](#).
- [160] S. Dittmaier, A. Huss, and G. Knippen, “Next-to-leading-order QCD and electroweak corrections to WWW production at proton-proton colliders,” *JHEP* **09** (2017) 034, [1705.03722](#).
- [161] D. T. Nhung, L. D. Ninh, and M. M. Weber, “NLO corrections to WWZ production at the LHC,” *JHEP* **12** (2013) 096, [1307.7403](#).
- [162] A. Denner, S. Dittmaier, S. Kallweit, and S. Pozzorini, “NLO QCD corrections to WWbb production at hadron colliders,” *Phys. Rev. Lett.* **106** (2011) 052001, [1012.3975](#).
- [163] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz, and F. Siegert, “Rivet user manual,” *Comput. Phys. Commun.* **184** (2013) 2803–2819, [1003.0694](#).
- [164] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for direct top squark pair production in final states with two leptons in  $\sqrt{s} = 13$  TeV  $pp$  collisions with the ATLAS detector,” *Eur. Phys. J.* **C77** (2017), no. 12, 898, [1708.03247](#).
- [165] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for B-L R-parity-violating top squarks in  $\sqrt{s} = 13$  TeV  $pp$  collisions with the ATLAS experiment,” *Phys. Rev.* **D97** (2018), no. 3, 032003, [1710.05544](#).
- [166] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in  $36 \text{ fb}^{-1}$  of  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector,” *JHEP* **01** (2018) 055, [1709.07242](#).
- [167] G. Cowan, *Statistical Data Analysis*. Oxford Science Series. OUP, 1998.
- [168] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The bulletin of mathematical biophysics* **5** (Dec, 1943) 115–133.
- [169] J. Han and C. Moraga, “The influence of the sigmoid function parameters on the speed of backpropagation learning,” in *From Natural to Artificial Neural Computation*, J. Mira and F. Sandoval, eds., pp. 195–201. Springer Berlin Heidelberg, Berlin, Heidelberg, 1995.

- [170] F. Rosenblatt, The Perceptron, a Perceiving and Recognizing Automaton Project Para. Report: Cornell Aeronautical Laboratory. Cornell Aeronautical Laboratory, 1957.
- [171] S. Marsland, Machine Learning: An Algorithmic Perspective, Second Edition. Chapman & Hall/CRC, 2nd ed., 2014.
- [172] F. Chollet et al., “Keras.” <https://github.com/keras-team/keras>, 2015.
- [173] M. Abadi and others., “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015. Software available from tensorflow.org.
- [174] R. H. R. Hahnloser, R. Sarpeshkar, M. A. Mahowald, R. J. Douglas, and H. S. Seung, “Erratum: Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit,” Nature **405** (jun, 2000) 947–951.
- [175] X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier neural networks,” in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, G. Gordon, D. Dunson, and M. Dudík, eds., vol. 15 of Proceedings of Machine Learning Research, pp. 315–323. PMLR, Fort Lauderdale, FL, USA, 11–13 Apr, 2011.
- [176] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” CoRR **abs/1412.6980** (2014) [1412.6980](#).
- [177] J. J. Shi, “Reducing prediction error by transforming input data for neural networks,” Journal of Computing in Civil Engineering **14** (2000), no. 2, 109–116.
- [178] A. Kvellestad, S. Maeland, and I. Strümke, “Signal mixture estimation for degenerate heavy Higgses using a deep neural network,” [1804.07737](#).
- [179] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” Journal of Computer and System Sciences **55** (1997), no. 1, 119 – 139.
- [180] Pedregosa et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research **12** (2011) 2825–2830.
- [181] B. Schmidt, “The High-Luminosity upgrade of the LHC: Physics and Technology Challenges for the Accelerator and the Experiments,” Burkhard Schmidt J. Phys.: Conf. Ser **706** (2016).

- [182] E. Barberio, B. Le, E. Richter-Was, Z. Was, D. Zanzi, and J. Zaremba, “Deep learning approach to the Higgs boson CP measurement in  $H \rightarrow \tau\tau$  decay and associated systematics,” Phys. Rev. **D96** (2017), no. 7, 073002, [1706.07983](#).