Chasing SUSY Through Parameter Space

Anders Kvellestad

12th August 2015

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1674

ISSN 1501-7710

Cover: Hanne Baadsgaard Utigard.
Print production: John Grieg AS, Bergen.

Produced in co-operation with Akademika Publishing.
The thesis is produced by Akademika Publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

"Drøyme du om nytt univers, meir orden, mindre trash?"

— Stein Torleif Bjella

Abstract

In this thesis we explore several physics scenarios based on the Minimal Supersymmetric Standard Model (MSSM) and its extension in terms of R-parity violation. We derive predictions for, and analyse results from, the ongoing experiments at the Large Hadron Collider (LHC). The important role played by statistics and parameter scanning techniques in analysis of large parameter spaces is emphasized.

We investigate a class of models where the superpartners of the Higgs and electroweak bosons, the charginos and neutralinos, are nearly degenerate in mass. When R-parity is conserved, this degeneracy can potentially make the lightest chargino long-lived on the timescales of collider physics. However, through a Bayesian analysis of the relevant parameter space, we find that this possibility is disfavoured by current data. In models where R-parity violating interactions are included, supersymmetric particles can decay to final states involving only Standard Model particles. We demonstrate that a small chargino–neutralino mass difference in this case may lead to interesting signals in LHC searches, including resonances of three charged particles.

Although the first run of the LHC produced no clear sign of physics beyond the Standard Model, some small, yet intriguing, excesses have been observed in the data. We investigate whether two excesses seen by the CMS and ATLAS experiments in searches for dileptons, jets and missing energy can be interpreted as early hints of supersymmetry. After taking into account the null-results of other LHC searches, we find that the supersymmetric scenarios considered are not viable explanations of the observed excesses.

Acknowledgements

First of all I would like to express my sincere gratitude to my PhD supervisor, Professor Are Raklev. This thesis would not have been possible without your expertise and guidance. Thank you for always keeping the office door open, for patiently answering questions on everything from quantum field theory to tailor-made shirts, and for helping me finish those cookies in my office.

I also would like to thank my co-supervisor, Professor Alex Read, whose expertise in both statistics and physics have helped spark my own interest in the interplay between these two fields.

To Ben Allanach, Per Osland, Smaragda Lola and Nils-Erik Bomark, I am very grateful for having had the opportunity to collaborate with you on the papers included in this thesis.

I have thoroughly enjoyed my time at the Theory Group in Oslo, and that has little to do with the dimly lit hallways and everything to do with all the great people I have met there. May your lunch breaks never end. In particular I wish to thank my office mate Lars A. Dal for much help, many interesting discussions and general good company.

I would further like to thank everyone in the `GAMBIT` collaboration. To learn from, and collaborate with, so many talented people has been a great experience, and I am looking forward to the continuation. Especially I wish to thank Pat Scott for invaluable help and advise.

During my time at the Department of Physics in Oslo I have had the pleasure to work with Arnt Inge Vistnes on teaching and Hilde Lynnebakken on outreach activities. Thank you both — it has been highly rewarding.

To my family, thank you for your constant support, and for always showing an interest in my work. A special thank you also to the Bergen/Gothen-

burg physics crew: Siri, Ørjan, Anders and Daniel.

Finally, I would like to thank Kaja for all her support, and for sticking up with me despite my absentmindedness and unorthodox working hours during the completion of this work.

List of papers

This thesis is based on the following papers:

Paper 1 N.-E. Bomark, A. Kvellestad, S. Lola, P. Osland and A. R. Raklev, "Long lived charginos in Natural SUSY?," *JHEP* **1405** (2014) 007, `arXiv:1310.2788 [hep-ph]`.

Paper 2 N. E. Bomark, A. Kvellestad, S. Lola, P. Osland and A. R. Raklev, "R-parity violating chargino decays at the LHC," *JHEP* **1412** (2014) 121, `arXiv:1410.0921 [hep-ph]`.

Paper 3 B. Allanach, A. R. Raklev and A. Kvellestad, "Interpreting a CMS excess in $lljj + missing$-transverse-momentum with the golden cascade of the minimal supersymmetric standard model," *Phys. Rev.* **D91** (2015) 115022, `arXiv:1409.3532 [hep-ph]`.

Paper 4 B. Allanach, A. Raklev and A. Kvellestad, "Consistency of the recent ATLAS $Z + E_T^{\text{miss}}$ excess in a simplified GGM model," *Phys. Rev.* **D91** (2015) 095016, `arXiv:1504.02752 [hep-ph]`.

Contents

Chapter 1

Introduction

As of writing this thesis, the second run of the Large Hadron Collider (LHC) has just started. For the first time in history we are colliding particles at a center-of-mass energy of 13 TeV — and we are doing so at an impressive rate. The first LHC run, with collisions taking place at 7 and 8 TeV, brought some truly remarkable results, with the discovery of the Higgs boson in 2012 as the crowning achievement.

Supersymmetry has long been one of the most promising theories for physics beyond the Standard Model. Yet, despite having performed a large number of complementary searches, the LHC experiments have not found any clear evidence for the existence of supersymmetric particles in the 7 and 8 TeV data. However, with higher energies the experiments will now be able to probe even further into the parameter space of supersymmetric theories. Time, and a lot of hard work, will eventually tell us what — if anything — is hiding there.

This thesis explores a few of the many possible ways in which supersymmetry may show up in LHC searches. In particular, the first two papers focus on scenarios where a small mass difference between the lightest chargino and neutralino may determine the expected collider phenomenology. In the latter two papers we investigate whether two small, yet interesting, excesses observed in LHC searches for dileptons, jets and missing energy can be interpreted as early hints of a supersymmetry signal.

We consider in some detail the methods used to derive predictions from

the typically large parameter spaces of supersymmetric theories. On the one hand, theoretical assumptions can be used to narrow down the parameter space, with the idea of *naturalness* being one important example. On the other hand, we can investigate large parameter spaces probabilistically in parameter scans, typically requiring powerful sampling algorithms and efficient computer codes for calculating model predictions. We will discuss both foundational and practical aspects of this approach.

The remainder of the thesis is organized as follows. In Chapter 2 we introduce supersymmetry and its realization in the form of supersymmetric field theories. This is followed by an introduction to the Minimal Supersymmetric Standard Model in Chapter 3. In Chapter 4 we turn to the topic of naturalness and its role as a guiding principle for exploring supersymmetric scenarios. As much of the work presented in this thesis is linked to statistics, we devote Chapter 5 to a discussion of statistical inference and parameter scanning, giving special emphasis to Bayesian methods and their interpretation. Part of the work behind this thesis has been to contribute to the development of `GAMBIT`, a new tool for performing statistical fits to particle physics theories. In Chapter 6 we discuss some of the practical challenges related to such analyses, and describe the approach we have taken with `GAMBIT` to overcome these challenges. Finally, in Chapter 7 we summarize the papers this thesis is based on. The published papers follow at the end.

Chapter 2

Supersymmetry

In this chapter we review the concept of supersymmetry. We start in Section 2.1 by presenting the algebra that defines supersymmetry, and illustrate a few central physical consequences. In Section 2.2 we connect these ideas to current problems in particle physics, motivating why supersymmetry might be relevant for physics at the energies of present day experiments. Section 2.3 concerns how to construct supersymmetric field theories. Finally, in Section 2.4, the phenomenologically important topic of supersymmetry breaking is discussed. For a more thorough treatment of the topics covered in this chapter, we refer the reader to Refs. [1–3].

2.1 The superalgebra

Relativistic field theories are invariant under the spacetime transformations of the *Poincaré group*, which is defined as the group of all transformations of the form

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \tag{2.1}$$

that leaves invariant the spacetime interval $(x-y)^2$. Here $\Lambda^\mu{}_\nu$ represents a Lorentz transformation and a^μ a constant translation. The generators of this group, $M_{\mu\nu}$ for Lorentz transformations and P_μ for translations, satisfy the

Poincaré algebra:

$$[P^\mu, P^\nu] = 0, \tag{2.2}$$
$$[M^{\mu\nu}, P^\rho] = i(g^{\nu\rho}P^\mu - g^{\mu\rho}P^\nu), \tag{2.3}$$
$$[M^{\mu\nu}, M^{\rho\sigma}] = i(g^{\nu\rho}M^{\mu\sigma} + g^{\mu\sigma}M^{\nu\rho} - g^{\nu\sigma}M^{\mu\rho} - g^{\mu\rho}M^{\nu\sigma}). \tag{2.4}$$

The idea of supersymmetry originated in the late sixties and early seventies from attempts to find non-trivial extensions to the Poincaré group. Through the work of, amongst others, Coleman, Mandula [4]; Golfand, Likhtman [5]; and Haag, Lopuszanski, Sohnius [6], it became clear that this could only be accomplished by allowing anticommutators in Lie algebras, so-called *superalgebras*, and further, that supersymmetry represents the most general such extension possible. Here we consider $N = 1$ supersymmetry, where a single set of four supersymmetry generators are introduced through a two-component Weyl spinor Q_a and its Hermitian conjugate $(Q_a)^\dagger \equiv Q^\dagger_{\dot{a}}$. The supersymmetric extension to the Poincaré algebra is then given by the anti-commutation relations

$$\{Q_a, Q_b\} = \{Q^\dagger_{\dot{a}}, Q^\dagger_{\dot{b}}\} = 0, \tag{2.5}$$
$$\{Q_a, Q^\dagger_{\dot{a}}\} = 2(\sigma^\mu)_{a\dot{a}}P_\mu, \tag{2.6}$$

and the commutation relations

$$[Q_a, P^\mu] = [Q^\dagger_{\dot{a}}, P^\mu] = 0, \tag{2.7}$$
$$[Q_a, M^{\mu\nu}] = \tfrac{1}{2}(\sigma^{\mu\nu})_a{}^b Q_b, \tag{2.8}$$
$$[Q^\dagger_{\dot{a}}, M^{\mu\nu}] = \tfrac{1}{2}(\overline{\sigma}^{\mu\nu})_{\dot{a}}{}^{\dot{b}} Q^\dagger_{\dot{b}}. \tag{2.9}$$

Here σ^μ, $\sigma^{\mu\nu}$, and the related $\overline{\sigma}^\mu$, $\overline{\sigma}^{\mu\nu}$, are constructed from the regular Pauli matrices σ^i as follows:

$$\sigma^\mu = (1, \sigma^i), \qquad \overline{\sigma}^\mu = (1, -\sigma^i), \tag{2.10}$$
$$\sigma^{\mu\nu} = \tfrac{i}{2}(\sigma^\mu\overline{\sigma}^\nu - \sigma^\nu\overline{\sigma}^\mu), \quad \overline{\sigma}^{\mu\nu} = \tfrac{i}{2}(\overline{\sigma}^\mu\sigma^\nu - \overline{\sigma}^\nu\sigma^\mu). \tag{2.11}$$

To illustrate the nature of the supersymmetry generators, we consider the case of a state $|m, j_3\rangle$ of mass m and spin j_3 along the z axis. For the rotation generator $J^3 = M^{12}$, Eq. (2.8) reads

$$[Q_a, J^3] = \tfrac{1}{2}(\sigma^3)_a{}^b Q_b. \tag{2.12}$$

Specializing to the Q_1 generator, we have

$$[Q_1, J^3] = \tfrac{1}{2} Q_1, \tag{2.13}$$

meaning that

$$J^3 Q_1 |m, j_3\rangle = (j_3 - \tfrac{1}{2}) Q_1 |m, j_3\rangle . \tag{2.14}$$

Thus, Q_1 has the effect of lowering j_3 by $\frac{1}{2}$. This illustrates the more general result that the supersymmetry generators alter the spin of a state by $\pm\frac{1}{2}$, transforming fermions into bosons and vice versa. Also, the fact that the supersymmetry generators commute with P^μ, Eq. (2.7), implies that states related through a supersymmetry transformation have identical mass,

$$P_\mu P^\mu Q_a |m, j_3\rangle = m^2 Q_a |m, j_3\rangle . \tag{2.15}$$

The single-particle states that transform into each other via some combination of the Q_a and $Q_a^\dagger$ operators are called *superpartners*. These states combine in *supermultiplets* that form irreducible representations of the supersymmetry algebra, with each supermultiplet containing an equal number of fermionic and bosonic degrees of freedom. As the supersymmetry generators also commute with gauge transformations, all particles in a supermultiplet will have identical gauge quantum numbers.

Of course, among the particles of the Standard Model, there are no fermion–boson pairs with identical mass and gauge quantum numbers. Thus, supersymmetry must be a broken symmetry in Nature's current vacuum state, a topic we return to in Section 2.4.

2.2 Motivations

There are several good reasons to expect supersymmetry to be relevant for physics at the TeV energy scale. The most important motivation comes from the *hierarchy problem* of the Standard Model, which we will focus on here.

In the Standard Model, the only fundamental scalar is the Higgs doublet field Φ, and the corresponding scalar potential is given by

$$V = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \tag{2.16}$$

For $\mu^2 < 0$ this potential obtains a degenerate minimum away from the origin, with the consequence that the gauge symmetry of the Standard Model is broken in the vacuum state. The energy scale of electroweak physics is determined by the vacuum expectation value v of the neutral Higgs component. If we minimize V and require agreement with the experimentally determined value $v \approx 174$ GeV, we obtain the relation

$$v = \sqrt{\frac{-\mu^2}{2\lambda}} \approx 174 \text{ GeV}. \tag{2.17}$$

However, the bare mass parameter μ^2 receives loop corrections $\delta\mu^2$ that are quadratically divergent in the loop momentum cut-off Λ. For instance, a fermion f with a coupling λ_f to the Higgs field generates a one-loop contribution

$$(\delta\mu^2)_{\text{one-loop},f} = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda^2 + \mathcal{O}\left(m_f^2 \ln \frac{\Lambda}{m_f}\right). \tag{2.18}$$

Similarly, a scalar s with a Higgs coupling λ_s gives rise to a contribution

$$(\delta\mu^2)_{\text{one-loop},s} = \frac{\lambda_s}{16\pi^2}\Lambda^2 - \mathcal{O}\left(m_s^2 \ln \frac{\Lambda}{m_s}\right). \tag{2.19}$$

In the Standard Model, such scalar loops are due to the quartic self-interaction in Eq. (2.16). The cut-off Λ is interpreted as the energy scale where new physics must be taken into account.

Now, using the loop-corrected Higgs potential, the minimization relation

in Eq. (2.17) becomes

$$v = \sqrt{\frac{-\mu_{\text{eff}}^2}{2\lambda_{\text{eff}}}} = \sqrt{\frac{-(\mu^2 + \delta\mu^2)}{2\lambda_{\text{eff}}}} \approx 174 \text{ GeV}. \tag{2.20}$$

If the theory is to remain perturbative, λ_{eff} cannot be much larger than unity. Consequently, if Λ is much larger than the electroweak scale of $\sim 10^2$ GeV, the above relation can only be satisfied by choosing the bare Lagrangian parameter μ^2 to be of order Λ^2, and tune its value to almost exactly cancel $\delta\mu^2$. This amount of parameter fine-tuning is regarded as highly unnatural, and taken as a sign that the Standard Model cannot remain valid up to very high scales. We will discuss the naturalness concept in much greater detail in Chapter 4.

The fermion and scalar loop corrections in Eqs. (2.18) and (2.19) have opposite signs. This provides a clue that a symmetry relating fermions and bosons may solve the hierarchy problem by facilitating a natural cancellation of the divergent diagrams. Supersymmetry provides exactly this symmetry. Due to the balance of fermionic and bosonic degrees of freedom in a supersymmetric theory, every fermion f will have two scalar superpartners $\tilde{f}$, with masses $m_{\tilde{f}} = m_f$. Further, because they are part of the same supermultiplet, their couplings to the Higgs turn out to be related through $|\lambda_f|^2 = \lambda_{\tilde{f}}$. This ensures the exact cancellation of loop contributions from f and $\tilde{f}$ to all orders in perturbation theory,

$$(\delta\mu^2)_{f+\tilde{f}} = 0. \tag{2.21}$$

Similarly, the divergent loop corrections due to Standard Model bosons are exactly cancelled by their fermionic superpartners.

When supersymmetry is broken, as it must be, we have $m_{\tilde{f}} \neq m_f$. However, the cancellation of quadratic divergences still holds,[1] as illustrated by the one-loop results in Eqs. (2.18) and (2.19) where the Λ^2 terms are independent of the masses. We are left with a contribution where the leading

[1]Here we are assuming that supersymmetry is only *softly* broken, see Section 2.4.1.

term only depends logarithmically on Λ,

$$(\delta\mu^2)_{f+\tilde{f}} \propto (m_f^2 - m_{\tilde{f}}^2) \ln \frac{\Lambda}{m_{\tilde{f}}}. \tag{2.22}$$

The dependence on $m_{\tilde{f}}^2$ implies that the superpartner mass scale cannot be too large if this solution to the hierarchy problem is to avoid fine-tuning. This constitutes the main motivation for expecting superpartners to appear around TeV energy scales.

Another popular argument for TeV-scale supersymmetry comes from the running of the three gauge couplings with the renormalisation scale. If superpartners are introduced around the TeV scale, the running is modified in such a way that it may allow for a unification of all three couplings around the scale $\Lambda_{\text{GUT}} \sim 10^{16}$ GeV.

Finally, the strongest observational evidence we currently have for physics beyond the Standard Model is the existence of dark matter. Supersymmetric theories can provide several viable dark matter candidates, including the *neutralino*, a mass eigenstate of the superpartners of the B, W^3 and Higgs bosons; and the *gravitino*, the superpartner of the graviton in supergravity theories.

2.3 Supersymmetric field theories

The construction of supersymmetric field theories is usually accomplished using the language of *superfields*, with one superfield per supermultiplet. Superfields are defined as functions on a *superspace*, constructed by extending spacetime with a set of four anticommuting, or "fermionic", coordinates. In this formalism, the regular spacetime dependent Lagrangian is generated by integrating the superspace Lagrangian over the fermionic coordinates, a process which ensures that only supersymmetry-invariant terms survive.

However, this formalism necessitates the introduction of a certain amount of mathematical machinery. Since the focus of this thesis is on the phenomenological aspects of supersymmetry, we here choose the less elegant, but more explicit, formulation in terms of regular spacetime dependent fields. In do-

ing this we follow the presentations given in Chapter 3 of Martin [1] and Chapters 3–5 of Aitchison [2].

2.3.1 Chiral supermultiplets

We start by considering the supersymmetric theory of *chiral* supermultiplets.[2] Each supermultiplet, labeled by an index i, contains a complex scalar field ϕ_i and a left-handed Weyl spinor ψ_i. Being a complex two-component object, ψ_i represents four real degrees of freedom. However, the equations of motion will eliminate two of these, meaning that on-shell the degrees of freedom of ψ_i are balanced by the two bosonic degrees of freedom in ϕ_i, as required by supersymmetry. To balance the degrees of freedom also off-shell, we introduce an unphysical complex scalar field F_i, known as an *auxiliary* field, with the simple Lagrangian density $\mathcal{L}_{\text{free},F} = F_i^* F_i$. This implies that F_i has mass dimension 2 and vanishes on-shell, with the equations of motion simply being $F_i = F_i^* = 0$. The non-interacting part of the Lagrangian is then given by

$$\mathcal{L}_{\text{free,chiral}} = -\partial^\mu \phi_i^* \partial_\mu \phi_i + i\psi_i^\dagger \overline{\sigma}^\mu \partial_\mu \psi_i + F_i^* F_i, \tag{2.23}$$

where all spinor indices have been suppressed and repeated supermultiplet indices i should be summed.

We parametrize the supersymmetry transformation $X \to X + \delta_\epsilon X$ of a field X by a constant, infinitesimal Weyl spinor parameter ϵ^α of mass dimension $-\frac{1}{2}$. This will allow scalar and spinor fields to transform into each other, as terms $\sim \epsilon^\alpha \psi_\alpha$ will be scalar objects of mass dimension 1, while terms $\sim \epsilon^\alpha \partial_\mu \phi$ will be spinors of mass dimension $\frac{3}{2}$.

The free-field Lagrangian in Eq. (2.23) is invariant under the supersymmetry transformations

$$\delta\phi_i = \epsilon\psi_i, \qquad \delta\phi_i^* = \epsilon^\dagger \psi_i^\dagger, \tag{2.24}$$

$$\delta(\psi_i)_\alpha = -i(\sigma^\mu \epsilon^\dagger)_\alpha \partial_\mu \phi_i + \epsilon_\alpha F_i, \qquad \delta(\psi_i^\dagger)_{\dot\alpha} = i(\epsilon\sigma^\mu)_{\dot\alpha} \partial_\mu \phi_i^* + \epsilon_{\dot\alpha}^\dagger F_i^*, \tag{2.25}$$

$$\delta F_i = -i\epsilon^\dagger \overline{\sigma}^\mu \partial_\mu \psi_i, \qquad \delta F_i^* = i\partial_\mu \psi_i^\dagger \overline{\sigma}^\mu \epsilon, \tag{2.26}$$

[2] Chiral supermultiplets are also commonly referred to as *scalar* or *matter* supermultiplets.

where repeated spinor indices have been suppressed. From the above transformations it follows that the commutator of two supersymmetry transformations, parametrized by ϵ and ϵ', is another symmetry transformation, namely a translation:

$$(\delta_{\epsilon'}\delta_{\epsilon} - \delta_{\epsilon}\delta_{\epsilon'})X = i(-\epsilon\sigma^{\mu}\epsilon'^{\dagger} + \epsilon'\sigma^{\mu}\epsilon^{\dagger})\partial_{\mu}X, \tag{2.27}$$

for any field X in the supermultiplet.

Next, we extend the Lagrangian to also include renormalizable interaction terms for the supermultiplet fields. It turns out that this part of the Lagrangian can be fully determined by a holomorphic function W of the scalar fields ϕ_i,

$$W = L_i\phi_i + \frac{1}{2}M_{ij}\phi_i\phi_j + \frac{1}{6}y_{ijk}\phi_i\phi_j\phi_k, \tag{2.28}$$

where both M_{ij} and y_{ijk} are symmetric under interchange of the supermultiplet indices. This function is known as the *superpotential*. The most general form of the interaction Lagrangian satisfying supersymmetry and renormalizability can now be expressed as

$$\mathcal{L}_{\text{int}} = \left(-\frac{1}{2}W_{ij}\psi_i\psi_j + W_iF_i\right) + \text{c.c.}, \tag{2.29}$$

where W_i and W_{ij} are determined through W as

$$\begin{aligned} W_i &= \frac{\delta W}{\delta\phi_i} = M_{ij}\phi_j + \frac{1}{2}y_{ijk}\phi_j\phi_k, \\ W_{ij} &= \frac{\delta^2 W}{\delta\phi_i\delta\phi_j} = M_{ij} + y_{ijk}\phi_k. \end{aligned} \tag{2.30}$$

Due to the terms W_iF_i and $W_i^*F_i^*$ in Eq. (2.29), the equations of motion for the auxiliary fields now become $F_i = -W_i^*$ and $F_i^* = -W_i$, meaning that F_i and F_i^* can be expressed in terms of the scalar fields. The complete

Lagrangian then becomes

$$\begin{aligned}\mathcal{L}_{\text{chiral}} = &\quad \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{int}} \\ = &\quad -\partial^\mu \phi_i^* \partial_\mu \phi_i + i\psi_i^\dagger \overline{\sigma}^\mu \partial_\mu \psi_i \\ &\quad -\frac{1}{2}\left(W_{ij}\psi_i\psi_j + W_{ij}^*\psi_i^\dagger\psi_j^\dagger\right) - V(\phi,\phi^*),\end{aligned} \tag{2.31}$$

where $V(\phi,\phi^*) = W_i W_i^*$ is the scalar potential.

2.3.2 Gauge supermultiplets

We now turn to *gauge*, or *vector*, supermultiplets. These contain a real vector field A_μ^a for a massless gauge boson, and a Weyl spinor field λ^a for the fermionic superpartner, referred to as the *gaugino*. The a index labels the gauge fields according to the adjoint representation of the respective gauge group. Due to the gauge transformation, an off-shell gauge field A_μ^a represents only three degrees of freedom. Thus, in order to balance the four degrees of freedom of the off-shell gaugino field, we must introduce an auxiliary real scalar field D^a.

Given a gauge group with structure constants f^{abc}, the gauge transformations of the supermultiplet fields are given by

$$A_\mu^a \to A_\mu^a + \partial_\mu \Lambda^a + g f^{abc} A_\mu^b \Lambda^c, \tag{2.32}$$

$$\lambda^a \to \lambda^a + g f^{abc} \lambda^b \Lambda^c, \tag{2.33}$$

$$D^a \to D^a + g f^{abc} D^b \Lambda^c, \tag{2.34}$$

where Λ^a is an infinitesimal transformation parameter and g is the gauge coupling. The gauge-invariant Lagrangian for these fields is then

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} F_{\mu\nu}^a F^{\mu\nu a} + i\lambda^{\dagger a}\overline{\sigma}^\mu \nabla_\mu \lambda^a + \frac{1}{2} D^a D^a, \tag{2.35}$$

with $F_{\mu\nu}^a$ being the usual field strength tensor, and $\nabla_\mu \lambda^a$ the covariant de-

rivative of λ^a introducing interactions with the A^a_μ fields,

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu, \tag{2.36}$$

$$\nabla_\mu \lambda^a = \partial_\mu \lambda^a + g f^{abc} A^b_\mu \lambda^c. \tag{2.37}$$

This Lagrangian is left invariant under the supersymmetry transformations,

$$\delta A^a_\mu = -\frac{1}{\sqrt{2}} \left(\epsilon^\dagger \overline{\sigma}_\mu \lambda^a + \lambda^{\dagger a} \overline{\sigma}_\mu \epsilon \right), \tag{2.38}$$

$$\delta \lambda^a_\alpha = \frac{i}{2\sqrt{2}} (\sigma^\mu \overline{\sigma}^\nu \epsilon)_\alpha F^a_{\mu\nu} + \frac{1}{\sqrt{2}} \epsilon_\alpha D^a, \tag{2.39}$$

$$\delta D^a = \frac{i}{\sqrt{2}} \left(-\epsilon^\dagger \overline{\sigma}^\mu \nabla_\mu \lambda^a + \nabla_\mu \lambda^{\dagger a} \overline{\sigma}^\mu \epsilon \right). \tag{2.40}$$

From these transformation laws it can be shown that a commutator of two supersymmetry transformations satisfy a relation analogous to Eq. (2.27).

2.3.3 Supersymmetric gauge theory

We can now construct a complete supersymmetric gauge theory with both chiral and gauge supermultiplets. Since supersymmetry and gauge transformations commute, the members $X_i = \phi_i, \psi_i, F_i$ of a chiral supermultiplet must all transform in the same way under a gauge transformation,

$$X_i \to X_i + i g \Lambda^a (T^a X)_i. \tag{2.41}$$

Here T^a are the generators of the group in the representation under which X transform, for instance the fundamental representation. As usual, to restore gauge invariance, derivatives of the matter fields must be replaced by covariant derivatives involving the vector fields A^a_μ,

$$\nabla_\mu \phi_i = \partial_\mu \phi_i - i g A^a_\mu (T^a \phi)_i, \tag{2.42}$$

$$\nabla_\mu \phi^*_i = \partial_\mu \phi^*_i + i g A^a_\mu (\phi^* T^a)^i, \tag{2.43}$$

$$\nabla_\mu \psi_i = \partial_\mu \psi_i - i g A^a_\mu (T^a \psi)_i. \tag{2.44}$$

We must further consider all other possible renormalizable interactions

between the members of the chiral and the gauge supermultiplets. This introduces two new terms $(\phi_i^* T^a \psi_i)\lambda^a$ and $\lambda^{\dagger a}(\psi_i^\dagger T^a \phi_i)$ representing "supersymmetrized gauge interactions", *i.e.*, interactions between the gauginos and the chiral fields. A third term, $(\phi_i^* T^a \phi_i) D^a$, combines with the $\frac{1}{2} D^a D^a$ term in $\mathcal{L}_{\text{gauge}}$ to generate a four-scalar term in the scalar potential once the auxiliary D^a field is eliminated through its equation of motion, $D^a = -g(\phi_i^* T^a \phi_i)$. From the requirement that the total Lagrangian must be invariant under supersymmetry, the couplings for all these terms turn out to be given by the gauge coupling g. Further, the supersymmetry transformations for the chiral supermultiplet in Eqs. (2.24)–(2.26) must be modified to take into account gauge interactions,

$$\delta\phi_i = \epsilon\psi_i, \tag{2.45}$$

$$\delta(\psi_i)_\alpha = -i(\sigma^\mu \epsilon^\dagger)_\alpha \nabla_\mu \phi_i + \epsilon_\alpha F_i, \tag{2.46}$$

$$\delta F_i = -i\epsilon^\dagger \overline{\sigma}^\mu \nabla_\mu \psi_i + \sqrt{2} g (T^a \phi)_i \epsilon^\dagger \lambda^{\dagger a}. \tag{2.47}$$

Combining the above results with $\mathcal{L}_{\text{chiral}}$ and $\mathcal{L}_{\text{gauge}}$, we arrive at the complete Lagrangian for a renormalizable supersymmetric gauge theory:

$$\begin{aligned}
\mathcal{L} = & -\nabla^\mu \phi_i^* \nabla_\mu \phi_i + i\psi_i^\dagger \overline{\sigma}^\mu \nabla_\mu \psi_i && \phi,\ \psi \text{ kinetic terms,} \\
& && \phi\phi A(A),\ \psi\psi A \text{ interactions} \\
& -\tfrac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} + i\lambda^{\dagger a} \overline{\sigma}^\mu \nabla_\mu \lambda^a && A,\ \lambda \text{ kinetic terms,} \\
& && A^3,\ A^4,\ \lambda\lambda A \text{ interactions} \\
& -\tfrac{1}{2}\left(W_{ij} \psi_i \psi_j + W^*_{ij} \psi_i^\dagger \psi_j^\dagger \right) && \psi \text{ mass terms,} \\
& && \phi\psi\psi \text{ interactions} \\
& -\sqrt{2} g (\phi_i^* T^a \psi_i)\lambda^a - \sqrt{2} g \lambda^{\dagger a} (\psi_i^\dagger T^a \phi_i) && \phi\psi\lambda \text{ interactions} \\
& - W_i W_i^* - \tfrac{1}{2} g (\phi_i^* T^a \phi_i)^2 && \phi \text{ mass terms,} \\
& && \phi^3,\ \phi^4 \text{ interactions}
\end{aligned} \tag{2.48}$$

Here the last line corresponds to the scalar potential, $-V(\phi, \phi^*)$. For theories based on a direct product of several gauge groups, for instance the $SU(3)_c \times SU(2)_L \times U(1)_Y$ symmetry of the Standard Model, all terms involving a summation over the representation index a should be repeated for each group.

A supersymmetric gauge theory is fully determined by the field content, the superpotential W and the fields' gauge transformations. The requirement

of gauge invariance will constrain what parameters in W are allowed to be non-zero. First, tadpole terms, $L_i\phi_i$, are only allowed if ϕ_i is a gauge singlet. Further, superpotential mass terms of the form $M_{ij}\phi_i\phi_j$ can only appear if the representations that ϕ_i and ϕ_j transform under are conjugates of each other. Finally, the Yukawa terms, $y_{ijk}\phi_i\phi_j\phi_k$, require that the fields ϕ_i, ϕ_j and ϕ_k transform under representations that can combine to a gauge singlet.

The fact that the superpotential simultaneously determines the $\phi\psi\psi$ Yukawa terms through W_{ij} and "half" the scalar potential through W_i, ensures some of the important properties we have already associated with supersymmetric theories. First, that the same M_{ij} or Yukawa parameters enter in the mass terms of both the scalars and the fermions, so that the members of each supermultiplet are mass degenerate. Second, that the same set of couplings y_{ijk} enter linearly in $\phi\psi\psi$ interactions and quadratically in ϕ^4 interactions, explaining the coupling relation $|\lambda_f|^2 = \lambda_{\tilde{f}}$ that ensured the cancellation of quadratic divergences in Section 2.2.

2.4 Supersymmetry breaking

As already noted, supersymmetry must be a broken symmetry at present energies; otherwise, superpartners mass degenerate with the Standard Model particles would have been discovered long ago. If we believe supersymmetry to be an exact symmetry of the fundamental theory, the symmetry must be broken spontaneously, that is, that the vacuum state is not invariant under the symmetry of the theory. This implies that there must be degenerate vacuum states, transforming among themselves under supersymmetry transformations.

To determine the conditions for spontaneous supersymmetry breaking we therefore start by examining the vacuum state $|0\rangle$. From the connection between the supersymmetry generators and P_μ in Eq. (2.6) we find that the Hamiltonian operator can be expressed as

$$H = P^0 = \frac{1}{4}\left(Q_1 Q_1^\dagger + Q_1^\dagger Q_1 + Q_2 Q_2^\dagger + Q_2^\dagger Q_2\right). \tag{2.49}$$

If the vacuum is invariant under supersymmetry, we must have

$$Q_a |0\rangle = Q_{\dot{a}}^\dagger |0\rangle = 0, \tag{2.50}$$

implying that $H|0\rangle = 0$. Thus, as long as supersymmetry is preserved there will be zero vacuum energy. Keeping in mind that the Hamiltonian in general must satisfy $H \geq 0$, the converse argument then becomes our criterion for spontaneous supersymmetry breaking: if there is a positive vacuum energy, supersymmetry is necessarily broken.

For a positive vacuum energy to be guaranteed independent of kinetic contributions to H, the scalar potential $V(\phi, \phi^*)$ must have a positive vacuum expectation value, $\langle 0|V|0\rangle > 0$. From the equations of motion for the auxiliary fields F_i and D^a, the potential can be written as

$$V = -W_i W_i^* - \frac{1}{2} g(\phi_i^* T^a \phi_i)^2 = F^{i*} F_i + \frac{1}{2} \sum_a D^a D^a, \tag{2.51}$$

where we in the last term have made explicit the sum over all relevant gauge groups. Models for spontaneous supersymmetry breaking can therefore be classified as either "F-term" or "D-term", depending on in which part of the potential the non-zero expectation value lives.

However, even if supersymmetry is spontaneously broken, the tree-level particle masses are connected. In particular, a weighted sum over tree-level squared-mass eigenvalues known as the *supertrace*, $\mathrm{STr}(m^2)$, can be shown to vanish in theories of non-anomalous gauge symmetries:

$$\mathrm{STr}(m^2) \equiv \sum_j (-1)^{2j} (2j+1) \mathrm{Tr}(m_j^2) = -2g \mathrm{Tr}(T^a) D^a = 0. \tag{2.52}$$

Here m^2 is the total squared-mass matrix of the Lagrangian and m_j^2 is the squared-mass matrix for spin-j particles. This has the consequence that, given only the known particles of the Standard Model and spontaneous symmetry breaking, the tree-level masses of some superpartners should be smaller than the corresponding mass terms in the Standard Model, which is very difficult to reconcile with observations. Another difficulty is the fact that the renormalizable supersymmetric Lagrangian in Eq. (2.48) does not contain

any $\phi\lambda\lambda$ term that can turn into a tree-level mass term for gauginos if the scalar field acquires a vacuum expectation value.

In general it seems difficult to spontaneously break supersymmetry using fields that are coupled at tree-level to the supermultiplets of the known particles. The usual assumption is therefore that the vacuum expectation value needed to break supersymmetry is generated in a *hidden sector* of fields with minimal direct interactions with the *visible sector* of known fields. If the effects of symmetry breaking are mediated from the hidden to the visible sector via non-renormalizable interactions or loop processes, phenomenologically viable superpartner masses can be generated despite the tree-level restrictions mentioned above.

Among the most studied frameworks for supersymmetry breaking are *Planck-scale-mediated supersymmetry breaking* (PMSB) and *gauge-mediated supersymmetry breaking* (GMSB). In the PMSB scenario it is assumed that the interaction connecting the hidden and visible sectors is due to new physics related to gravity at the Planck scale, $\Lambda_{\mathrm{Pl}} \sim 10^{18}$ GeV. Given an F-term vacuum expectation value $\langle F \rangle$ in the hidden sector, the expected scale of the mediated symmetry-breaking effects is then $\sim \frac{\langle F \rangle}{\Lambda_{\mathrm{Pl}}}$. In the GMSB approach, one assumes the existence of a set of *messenger fields* in the form of new chiral supermultiplets, which are charged under the gauge group of the visible sector. If the messenger fields also couple to the vacuum expectation value $\langle F \rangle$ of the hidden sector, mass terms for the visible sector can be generated at one-loop for gauginos and two-loop for the chiral scalars. The characteristic scale for these terms is then $\sim \frac{\alpha}{4\pi}\frac{\langle F \rangle}{\Lambda_{\mathrm{mess}}}$, where $\frac{\alpha}{4\pi}$ is a loop factor for the relevant gauge group and Λ_{mess} is the mass scale of the messenger fields.

2.4.1 Soft supersymmetry breaking

In lieu of a preferred model for exactly how supersymmetry ends up being broken, we can take the bottom-up approach of "parametrising our ignorance": We add terms to the Lagrangian that explicitly violate supersymmetry and treat the coefficients of these terms as free model parameters. Any specific model for supersymmetry breaking should imply some pattern on the space of these parameters, something we can hope to uncover through

experiments.

Of course, if we naively add *any* supersymmetry-breaking term to our Lagrangian, we end up spoiling all the properties that made low-scale supersymmetry attractive in the first place. In particular, we want to preserve the cancellation of quadratic divergences that allowed us to solve the hierarchy problem in Section 2.2. We therefore only allow so-called *soft-breaking terms*, which do not generate quadratically divergent contributions, with a total mass dimension for the interacting fields of 3 or less. The complete set of such possible soft terms is

$$\begin{aligned}\mathcal{L}_{\text{soft}} = &-\frac{1}{2}M\lambda^a\lambda^a \\ &-\frac{1}{6}a_{ijk}\phi_i\phi_j\phi_k + \frac{1}{2}b_{ij}\phi_i\phi_j + t_i\phi_i + \text{c.c.} \\ &-(m^2)_{ij}\phi_j^*\phi_i \\ &-\frac{1}{2}c_{ijk}\phi_i^*\phi_j\phi_k + \text{c.c.}, \qquad\qquad \text{(maybe soft)}\end{aligned} \tag{2.53}$$

where c.c. denotes the complex conjugation of all terms on the corresponding line. The M terms, repeated for each gauge group, will give masses to the gauginos, while the $(m^2)_{ij}$ and b_{ij} terms provide additional mass terms for the scalars. The a_{ijk} and c_{ijk} terms are couplings of three scalar fields. In the case where one field acquires a vacuum expectation value, they will turn into mass terms for the remaining fields. However, the c_{ijk} terms are denoted "maybe soft" as they may lead to quadratic divergences if the theory contains a chiral supermultiplet which is a singlet under all gauge symmetries. These terms are often ignored as most models of supersymmetry breaking predict them to be negligible. Finally, the tadpole term $t_i\phi_i$ is only allowed if ϕ_i is a gauge singlet. Taken together, the set of additional mass terms in $\mathcal{L}_{\text{soft}}$ will allow for realistic mass splittings between known particles and their superpartners.

Several of the parameters introduced in $\mathcal{L}_{\text{soft}}$ are heavily constrained by experiments. In particular, complex phases and parameters that are non-diagonal in the supermultiplet indices will often lead to predictions of CP-violation and flavour-changing neutral currents in conflict with observations. A realistic model of supersymmetry breaking must therefore be able to ex-

plain why these terms should be small.

Chapter 3

The Minimal Supersymmetric Standard Model

The *Minimal Supersymmetric Standard Model* (MSSM) is a supersymmetric extension of the Standard Model containing the minimal number of new fields required to make such a theory viable. It forms the theoretical framework for much current research on supersymmetry phenomenology, including the work presented in this thesis, where we include some phenomenologically interesting extensions. In this chapter we provide a brief introduction to the MSSM, giving special emphasis to topics relevant for subsequent chapters. A more detailed introduction to the MSSM can be found in Refs. [1, 2, 7].

3.1 Field content

To construct a supersymmetric theory based on the $SU(3)_c \times SU(2)_L \times U(1)_Y$ gauge symmetry of the Standard Model, all the known particles must be placed in appropriate supermultiplets. The supermultiplets are then completed by postulating the existence of new supersymmetric particles, collectively called *sparticles*.

A list of all the supermultiplets of the MSSM is given in Table 3.1. The Standard Model leptons and quarks are, together with their scalar superpartners, members of chiral supermultiplets. For the left-handed fermions, the supermultiplets are organized in the $SU(2)_L$ doublets Q_i and L_i, while the

right-handed fermions live in the $SU(2)_L$ singlet supermultiplets $\bar{u}_i$, $\bar{d}_i$ and $\bar{e}_i$. Here i is an index running over the three generations of the Standard Model. It is conventional to only use left-handed chiral supermultiplets when formulating the theory. We therefore use the conjugates of right-handed fields in the $\bar{u}_i$, $\bar{d}_i$ and $\bar{e}_i$ supermultiplets. The scalar components of the quark and lepton supermultiplets are referred to as *squarks* and *sleptons*, or collectively as *sfermions*. The gauge supermultiplets g, W and B respectively contain the Standard Model gluons, W bosons and B boson, along with their fermionic superpartners, the *gluinos*, *winos* and the *bino*.

For the Higgs sector, matters are a little more complicated. In the Standard Model, with the Higgs doublet $\Phi = (\phi_a, \phi_b)$, mass terms for the upper components of fermion $SU(2)_L$ doublets are generated from Yukawa interactions involving the conjugated Higgs field through $\tilde{\Phi} = (\phi_b^*, -\phi_a^*)$. As we have seen in Chapter 2, in a supersymmetric theory, Yukawa terms for the chiral fermions originate from the superpotential. However, the superpotential is a holomorphic function of the scalar fields, meaning that it cannot depend on the conjugated fields. We therefore need *two* Higgs doublets, H_u and H_d, for generating mass terms for the upper and lower components of $SU(2)_L$ doublets, respectively. Also, having two Higgs doublets is required to ensure cancellation of gauge anomalies in the electroweak sector. The scalar Higgs fields live in chiral supermultiplets along with their fermionic superpartners, the *higgsinos*.

3.2 R-parity

In the Standard Model, the requirements of gauge invariance and renormalizability rule out any Lagrangian terms that can violate baryon number (B) or lepton number (L). As we will see in the next section, this "accidental" symmetry is no longer present in supersymmetric theories. This may have severe consequences for the viability of supersymmetric models, as there are strong experimental limits on the rates of B- and L-violating processes. In particular, searches for proton decay through the process $p \to l^+\pi^0$ ($l = e, \mu$), in which both baryon and lepton number is violated by one unit, has set a lower limit on the lifetime of the proton of $\sim 10^{34}$ years [8], which translates

Supermultiplet	scalars	fermions	vectors	$SU(3)_c$	$SU(2)_L$	$U(1)_Y$
Q_i	$(\tilde{u}_{iL}, \tilde{d}_{iL})$	(u_{iL}, d_{iL})		$\mathbf{3}$	$\mathbf{2}$	$\frac{1}{6}$
$\bar{u}_i$	$\tilde{u}^*_{iR}$	$u^\dagger_{iR}$		$\bar{\mathbf{3}}$	$\mathbf{1}$	$-\frac{2}{3}$
$\bar{d}_i$	$\tilde{d}^*_{iR}$	$d^\dagger_{iR}$		$\bar{\mathbf{3}}$	$\mathbf{1}$	$\frac{1}{3}$
L_i	$(\tilde{\nu}_{iL}, \tilde{e}_{iL})$	(ν_{iL}, e_{iL})		$\mathbf{1}$	$\mathbf{2}$	$-\frac{1}{2}$
$\bar{e}_i$	$\tilde{e}^*_{iR}$	$e^\dagger_{iR}$		$\mathbf{1}$	$\mathbf{1}$	1
H_u	(H^+_u, H^0_u)	$(\tilde{H}^+_u, \tilde{H}^0_u)$		$\mathbf{1}$	$\mathbf{2}$	$\frac{1}{2}$
H_d	(H^0_d, H^-_d)	$(\tilde{H}^0_d, \tilde{H}^-_d)$		$\mathbf{1}$	$\mathbf{2}$	$-\frac{1}{2}$
g		$\tilde{g}$	g	$\mathbf{8}$	$\mathbf{1}$	0
W		$\tilde{W}^{1,2,3}$	$W^{1,2,3}$	$\mathbf{1}$	$\mathbf{3}$	0
B		$\tilde{B}$	B	$\mathbf{1}$	$\mathbf{1}$	0

Table 3.1: The chiral and gauge supermultiplets in the Minimal Supersymmetric Standard Model. The index $i = 1, 2, 3$ runs over the three generations of quarks and leptons.

to stringent bounds on any interactions that allow this process.

In the MSSM, B- and L-violating Lagrangian terms are avoided by postulating an additional discrete symmetry: All interactions are required to conserve a multiplicative quantum number called *R-parity*, defined as

$$P_R = (-1)^{3(B-L)+2s}, \tag{3.1}$$

where s is the particle spin. From this definition it follows that the known Standard Model particles and the additional MSSM Higgs bosons have $P_R = +1$, while the sparticles all have $P_R = -1$.

Besides protecting against problematic interaction terms, requiring R-parity conservation (RPC) has some important phenomenological implications. First, that sparticles must be created and annihilated in pairs. Second, that sparticles decay to states with an odd number of lighter sparticles. And third, that the lightest supersymmetric particle (LSP) must be absolutely

stable, with the consequence that all sparticle decay processes eventually end with the LSP.

The fact that the stable LSP so far has avoided detection implies that it should have zero electric and colour charge. This ensures the basic properties required for it to be a viable dark matter candidate. For instance, this is the case for the neutralino, which we discuss in Section 3.6. Since a stable, at most weakly-interacting particle will pass unseen through particle detectors, most collider searches for supersymmetry are based on event signatures with a large imbalance in the conserved momentum, so called *missing energy*.

Going beyond the MSSM, an alternative to postulating R-parity conservation is to allow also the B- and L-violating terms in the Lagrangian. This approach is referred to as *R-parity violation* (RPV). The assumption is then that there should be some structure on the space of the new couplings, explaining why large violations of baryon and lepton number have not been detected. In particular, allowing only either B- or L-violating terms will render the proton stable, since the decay process violates both quantum numbers.

When R-parity violation is allowed the lightest sparticle is no longer absolutely stable. In order to explain dark matter, we then need an LSP whose interactions are highly suppressed, leading to a lifetime longer than the age of the universe. This can arise naturally in supergravity theories.[1] Here the gravitino, which becomes massive as a result of supersymmetry breaking, has interactions that are suppressed by the Planck scale. Another possible supersymmetric dark matter candidate with naturally suppressed interactions is the *axino*. Being the superpartner of the axion, a hypothetical scalar arising within the Peccei–Quinn solution to the "strong CP problem" [9], it has interactions suppressed by the Peccei–Quinn scale $f_a \sim 10^{-11}$ GeV.

[1]Supergravity theories are based on *local* supersymmetry transformations, meaning that the transformation parameter ϵ^α introduced in Section 2.3.1 is taken to be a function of spacetime.

3.3 Superpotential

The superpotential for the MSSM is given by

$$W_{\text{MSSM}} = \mu H_u H_d + y^u_{ij} \tilde{u}^*_{iR} \tilde{Q}_j H_u - y^d_{ij} \tilde{d}^*_{iR} \tilde{Q}_j H_d - y^e_{ij} \tilde{e}^*_{iR} \tilde{L}_j H_d, \tag{3.2}$$

where H_u, H_d, $\tilde{Q}$ and $\tilde{L}$ are the scalar $SU(2)_L$ doublets for the corresponding supermultiplets, all listed in the second column of Table 3.1. All colour and weak isospin indices have been suppressed.[2] By comparing the expression in Eq. (3.2) to the general superpotential in Eq. (2.28), we first note that W_{MSSM} does not contain any tadpole term, since there are no total gauge singlets in the theory. Next, we see that we only have a single term of the general form $M_{ij}\phi_i\phi_j$, namely the "μ term" $\mu H_u H_d$. From the third line in Eq. (2.48) we see that this term will appear in the Lagrangian as a mass term for the higgsinos. The remaining terms are all of the general form $y_{ijk}\phi_i\phi_j\phi_k$. When the neutral components of H_u and H_d obtain vacuum expectation values through electroweak symmetry breaking, these terms will, among other things, generate the familiar mass terms and CKM mixings for the Standard Model fermions.

It is worth pointing out that, given the above superpotential, the only new Lagrangian parameter introduced by supersymmetrizing the Standard Model is the higgsino mass parameter μ. On the other hand, the terms contained in W_{MSSM} are not sufficient to explain electroweak symmetry breaking. From the last line in Eq. (2.48) it follows that W_{MSSM} implies the following Higgs squared-mass terms in the scalar potential:

$$V(\phi, \phi^*) \ni |\mu|^2 \left(|H^0_u|^2 + |H^+_u|^2 + |H^0_d|^2 + |H^-_d|^2 \right). \tag{3.3}$$

As this contribution is non-negative and has a minimum when the neutral fields vanish, we will not obtain a symmetry-breaking potential. Thus, electroweak symmetry breaking in the MSSM requires scalar potential contributions from the soft supersymmetry breaking terms.

[2] If we write out all indices in, for instance, the second term, we get $(y^u_{ij})(\tilde{u}^{*a}_{iR})(\tilde{Q}_{ja})_\alpha (H_u)_\beta \epsilon^{\alpha\beta}$. Here $i = 1, 2, 3$ is the generation index, $a = 1, 2, 3$ is the $SU(3)_c$ colour index, and $\alpha, \beta = 1, 2$ are $SU(2)_L$ weak isospin indicies contracted using the $\epsilon^{\alpha\beta}$ antisymmetric tensor with $\epsilon^{12} = 1$.

If we allow R-parity violation the superpotential is extended by

$$W_{\rm RPV} = \mu'_i \tilde{L}_i H_u + \frac{1}{2}\lambda_{ijk}\tilde{L}_i\tilde{L}_j\tilde{e}^*_{kR} + \lambda'_{ijk}\tilde{L}_i\tilde{Q}_j\tilde{d}^*_{kR} + \lambda''_{ijk}\tilde{u}^*_{iR}\tilde{d}^*_{jR}\tilde{d}^*_{kR}, \qquad (3.4)$$

where the first three terms violate lepton number and the last term violates baryon number. Many of the new couplings introduced here are strongly constrained by experiments. In particular, the limit on the proton lifetime implies that products of the λ' and λ'' couplings involving the first generation quarks must be highly suppressed.

3.4 Soft-breaking terms

Following the general structure of Eq. (2.53), we now break supersymmetry by adding to the MSSM Lagrangian the set of soft-breaking terms allowed by gauge invariance and renormalizability:

$$\begin{aligned}
\mathcal{L}_{\rm MSSM,soft} = &-\frac{1}{2}\left(M_1\tilde{B}\tilde{B} + M_2\tilde{W}^a\tilde{W}^a + M_3\tilde{g}^a\tilde{g}^a + \text{c.c}\right) \\
&- \left(a^u_{ij}\tilde{Q}_i H_u\tilde{u}^*_{jR} - a^d_{ij}\tilde{Q}_j H_d\tilde{d}^*_{jR} - a^e_{ij}\tilde{L}_i H_d\tilde{e}^*_{jR} + \text{c.c}\right) \\
&- (m^2_u)_{ij}\tilde{u}^*_{iR}\tilde{u}_{jR} - (m^2_d)_{ij}\tilde{d}^*_{iR}\tilde{d}_{jR} - (m^2_e)_{ij}\tilde{e}^*_{iR}\tilde{e}_{jR} \\
&- (m^2_Q)_{ij}\tilde{Q}^\dagger_i\tilde{Q}_j - (m^2_L)_{ij}\tilde{L}^\dagger_i\tilde{L}_j \\
&- m^2_{H_u}H^*_u H_u - m^2_{H_d}H^*_d H_d - (bH_u H_d + \text{c.c}).
\end{aligned} \qquad (3.5)$$

For the chiral fields, colour and weak isospin indices have been suppressed. In $\mathcal{L}_{\rm MSSM,soft}$ we note in particular that the first line provide mass terms for the bino, winos and gluinos, while the last line contain additional Higgs mass terms that can help generate the scalar potential needed for electroweak symmetry breaking.

The gaugino mass parameters $M_{1,2,3}$, three-scalar couplings $a^{u,d,e}_{ij}$, and sfermion mass matrices $m^2_{u,d,e,Q,L}$ are all generally complex valued, although the sfermion mass matrices are required to be hermitian. The Higgs mass parameters $m^2_{H_u}$ and $m^2_{H_d}$ are real, while b_{ij}, and the μ parameter in the superpotential, can be taken to be real by appropriate complex rotations of

the Higgs supermultiplet fields. In total, after having chosen a basis that removes as much freedom as possible, the MSSM Lagrangian contains 105 new real-valued parameters.

So while the supersymmetry-preserving part of the MSSM only introduced a single new parameter, our parametrization of supersymmetry breaking has left us with what at first may look like a model without much predictive power. However, as noted in Section 2.4.1, there are strong experimental limits on most of the new parameters due to their potential role in flavour-changing and CP-violating processes. As an example, consider a scenario with a sizeable off-diagonal element $(m_e^2)_{21}$ in the slepton mass matrix. The term $(m_e^2)_{21}\tilde{e}_{2R}^*\tilde{e}_{1R} = (m_e^2)_{21}\tilde{\mu}_R^*\tilde{e}_R$ then allows for mixing between the selectron and the smuon. Through a slepton–gaugino loop, this may induce the flavour-changing process $\mu \to e\gamma$, which so far has not been observed experimentally.[3] Similar contributions can come from off-diagonal elements in $(m_L^2)_{ij}$ and a_{ij}^e.

In light of the stringent experimental constraints, and in the interest of reducing the number of free parameters, a set of assumptions is usually employed on the space of the soft-breaking parameters: First, the slepton and squark mass matrices are taken to be diagonal,

$$(m_x^2)_{ij} = \mathrm{diag}(m_{\tilde{x}_1}^2, m_{\tilde{x}_2}^2, m_{\tilde{x}_3}^2), \quad x = u, d, e, Q, L. \tag{3.6}$$

This is the expected result if the mechanism responsible for breaking supersymmetry is flavour-blind. Further, as there is a one-to-one correspondence between the three-scalar terms in $\mathcal{L}_{\mathrm{MSSM,soft}}$ and the three-scalar Yukawa terms in the superpotential, the corresponding couplings are assumed to be related through three proportionality parameters A_0^u, A_0^d and A_0^e:

$$a_{ij}^u = A_0^u y_{ij}^u, \qquad a_{ij}^d = A_0^d y_{ij}^d, \qquad a_{ij}^e = A_0^e y_{ij}^e. \tag{3.7}$$

Finally, the gaugino mass parameters and three-scalar couplings are all taken

[3]The current best limit on the branching ratio is $BR(\mu \to e\gamma) < 5.7 \times 10^{-13}$ at the 90% confidence level [10].

to be real,

$$\mathrm{Im}(M_1) = \mathrm{Im}(M_2) = \mathrm{Im}(M_3) = \mathrm{Im}(A_0^u) = \mathrm{Im}(A_0^d) = \mathrm{Im}(A_0^e) = 0, \quad (3.8)$$

to avoid introducing large CP-violating phases. The sfermion mass parameters in Eq. (3.6) are automatically real due to the hermiticity of the underlying mass matrix.

It is important to note that the above assumptions are enforced at the, usually high, energy scale where the free parameters are defined. When the parameters are evolved down to the electroweak scale using the Renormalization Group Equations (RGE), the relations in Eqs. (3.6) and (3.7) are generally broken due to corrections from Yukawa interactions. However, as the Yukawa couplings are large only for the third generation, the three-scalar couplings and off-diagonal mass matrix elements for the first two generations remain small also at the weak scale. Further, the RGE running does not introduce any new CP-violating phases. The above set of assumptions can therefore protect against dangerous flavour-changing and CP-violation also after RGE running has been taken into account.

3.5 Summary of parameters

Given the assumptions in Eqs. (3.6)–(3.8), the number of free soft-breaking parameters is reduced to 24. Thus, including the μ parameter from the superpotential, we end up with a 25-dimensional MSSM parameter space. The Higgs sector then has four real Lagrangian parameters: μ, $m_{H_u}^2$, $m_{H_d}^2$ and b. By requiring that the electroweak symmetry breaking in the MSSM reproduces the observed electroweak scale, either $(m_{H_u}^2, m_{H_d}^2)$ or $(|\mu|, b)$ can be traded against the precisely measured M_Z and the unknown ratio of the vacuum expectation values for H_u^0 and H_d^0, denoted $\tan\beta \equiv v_u/v_d$. This effectively reduces the number of free MSSM parameters by one. When $(m_{H_u}^2, m_{H_d}^2)$ are eliminated, the b parameter is often replaced by a parameter $m_A^2 = 2b/\sin 2\beta$.

We can now summarize the 24 most common input parameters for phenomenology studies in the MSSM:

Gaugino mass parameters:

$$M_1,\ M_2,\ M_3.$$

Trilinear couplings:

$$A_0^u,\ A_0^d,\ A_0^e.$$

Higgs parameters:

$$\begin{aligned} & \tan\beta,\ m_{H_u}^2,\ m_{H_d}^2,\ \mathrm{sgn}(\mu) \\ \textit{or}\quad & \tan\beta,\ \mu,\ b \\ \textit{or}\quad & \tan\beta,\ \mu,\ m_A^2. \end{aligned}$$

Squark mass parameters:

$$\begin{aligned} & m_{\tilde{Q}_1}^2,\ m_{\tilde{Q}_2}^2,\ m_{\tilde{Q}_3}^2, \\ & m_{\tilde{u}_1}^2,\ m_{\tilde{u}_2}^2,\ m_{\tilde{u}_3}^2, \\ & m_{\tilde{d}_1}^2,\ m_{\tilde{d}_2}^2,\ m_{\tilde{d}_3}^2. \end{aligned}$$

Slepton mass parameters:

$$\begin{aligned} & m_{\tilde{L}_1}^2,\ m_{\tilde{L}_2}^2,\ m_{\tilde{L}_3}^2, \\ & m_{\tilde{e}_1}^2,\ m_{\tilde{e}_2}^2,\ m_{\tilde{e}_3}^2. \end{aligned}$$

3.6 The neutralino and chargino sector

The higgsinos and electroweak gauginos play a central role in all the work presented in this thesis. We therefore end our brief tour of the MSSM with a closer look at this sector.

In the general supersymmetric Lagrangian in Eq. (2.48) we find interactions between gauginos and chiral supermultiplet members of the form

$$\mathcal{L} \ni -\sqrt{2}g(\phi_i^* T^a \psi_i)\lambda^a - \sqrt{2}g\lambda^{\dagger a}(\psi_i^\dagger T^a \phi_i). \tag{3.9}$$

Let the fields ϕ_i, ψ_i belong to one of the Higgs supermultiplets, and take ϕ_i

to be the neutral scalar component, $\phi_i = H^0_{u,d}$. Once ϕ_i acquires a vacuum expectation value, the terms in Eq. (3.9) will turn into mass-mixing terms for the λ^a and ψ_i fields. Therefore, due to electroweak symmetry breaking, we can expect mixing between the electroweak gauginos and the higgsinos. The only constraint comes from the unbroken $U(1)_{\text{em}}$ symmetry, which ensures that only fields of equal electric charge can mix.

The neutral electroweak gauginos $\tilde{B}$ and $\tilde{W}^3$, here denoted as $\tilde{B}^0$ and $\tilde{W}^0$, will mix with the neutral higgsinos $\tilde{H}^0_d$ and $\tilde{H}^0_u$ to form four mass eigenstates $\tilde{\chi}^0_i$ called neutralinos. If we define a vector $\tilde{\chi}^0$ in the gauge eigenstate basis,

$$\tilde{\chi}^0 = \begin{pmatrix} \tilde{B}^0 \\ \tilde{W}^0 \\ \tilde{H}^0_d \\ \tilde{H}^0_u \end{pmatrix}, \tag{3.10}$$

the relevant mass terms in the Lagrangian can be combined as

$$\mathcal{L}_{\tilde{X}^0,\text{mass}} = -\frac{1}{2}(\tilde{\chi}^0)^T \mathbf{M}_{\tilde{\chi}^0} \tilde{\chi}^0 + \text{c.c.}, \tag{3.11}$$

where the mass matrix $\mathbf{M}_{\tilde{\chi}^0}$ is

$$\mathbf{M}_{\tilde{\chi}^0} = \begin{pmatrix} M_1 & 0 & -\frac{1}{\sqrt{2}}g'v_d & \frac{1}{\sqrt{2}}g'v_u \\ 0 & M_2 & \frac{1}{\sqrt{2}}gv_d & -\frac{1}{\sqrt{2}}gv_u \\ -\frac{1}{\sqrt{2}}g'v_d & \frac{1}{\sqrt{2}}gv_d & 0 & -\mu \\ \frac{1}{\sqrt{2}}g'v_u & -\frac{1}{\sqrt{2}}gv_u & -\mu & 0 \end{pmatrix}. \tag{3.12}$$

In addition to the mixing terms discussed above, we recognize the M_1 and M_2 gaugino mass parameters from the soft-breaking terms in Eq. (3.5), and the higgsino mass parameter μ originating from the $\mu H_u H_d$ term in the superpotential, Eq. (3.2).

To find the neutralino mass eigenstates $\tilde{\chi}^0_i$ and the corresponding masses $m_{\tilde{\chi}^0_i}$, we determine the unitary matrix $\mathbf{N}$ that diagonalizes $\mathbf{M}_{\tilde{\chi}^0}$,

$$\mathbf{N}^* \mathbf{M}_{\tilde{\chi}^0} \mathbf{N}^{-1} = \text{diag}(m_{\tilde{\chi}^0_1}, m_{\tilde{\chi}^0_2}, m_{\tilde{\chi}^0_3}, m_{\tilde{\chi}^0_4}). \tag{3.13}$$

The mass eigenstates are then given by

$$\tilde{\chi}_i^0 = \mathbf{N}_{i1}\tilde{B}^0 + \mathbf{N}_{i2}\tilde{W}^0 + \mathbf{N}_{i3}\tilde{H}_d^0 + \mathbf{N}_{i4}\tilde{H}_u^0, \tag{3.14}$$

conventionally labeled $i = 1, 2, 3, 4$ according to increasing mass.

From electroweak symmetry breaking we have that the vacuum expectation values v_u and v_d must be related to the Z boson mass as

$$(v_u^2 + v_d^2) = v^2 = \frac{2M_Z^2}{g^2 + g'^2}, \tag{3.15}$$

which implies that all the mixing elements in $\mathbf{M}_{\tilde{\chi}^0}$ are proportional to M_Z. Thus, when the magnitudes of M_1, M_2 and μ are large compared to M_Z, the mixing terms typically play a minor role. In this limit the composition of the neutralino states is determined by the hierarchy of the mass parameters. For instance, if $|M_1| < |M_2| \ll |\mu|$, we will have $\tilde{\chi}_1^0 \approx \tilde{B}^0$, $\tilde{\chi}_2^0 \approx \tilde{W}^0$ and $\tilde{\chi}_{3,4}^0 \approx \frac{1}{\sqrt{2}}(\tilde{H}_d^0 \pm \tilde{H}_u^0)$, with masses similar to the respective parameters.

The charged winos $\tilde{W}^{\mp} = \frac{1}{\sqrt{2}}(\tilde{W}^1 \mp i\tilde{W}^2)$ and the charged higgsinos $\tilde{H}_u^+$ and $\tilde{H}_d^-$ will similarly mix to form charged mass eigenstates known as *charginos*. With the gauge eigenstate basis

$$\tilde{\chi}^{\pm} = \begin{pmatrix} \tilde{W}^+ \\ \tilde{H_u}^+ \\ \tilde{W}^- \\ \tilde{H_d}^- \end{pmatrix}, \tag{3.16}$$

the chargino mass terms become

$$\mathcal{L}_{\tilde{X}^{\pm},\text{mass}} = -\frac{1}{2}(\tilde{\chi}^{\pm})^T \mathbf{M}_{\tilde{\chi}^{\pm}} \tilde{\chi}^{\pm} + \text{c.c.}, \tag{3.17}$$

with a mass matrix

$$\mathbf{M}_{\tilde{\chi}^{\pm}} = \begin{pmatrix} 0 & 0 & M_2 & gv_d \\ 0 & 0 & gv_u & \mu \\ M_2 & gv_u & 0 & 0 \\ gv_d & \mu & 0 & 0 \end{pmatrix}. \tag{3.18}$$

By diagonalizing this matrix, and using the relations $gv_d = \sqrt{2}\cos\beta M_W$ and $gv_u = \sqrt{2}\sin\beta M_W$, we find that the chargino mass eigenvalues are given by

$$\begin{aligned} m_{\tilde{\chi}^{\pm}_{1,2}} = \quad & \frac{1}{2}\left(|M_2|^2 + |\mu|^2 + 2M_W^2\right) \\ & \mp \frac{1}{2}\sqrt{\left(|M_2|^2 + |\mu|^2 + 2M_W^2\right)^2 - 4\left|\mu M_2 - M_W^2 \sin^2\beta\right|^2}. \end{aligned} \tag{3.19}$$

If $|\mu|$ and $|M_2|$ are much larger than M_W, and taking for example $|M_2| < |\mu|$, the above expression reduces to

$$m_{\tilde{\chi}^{\pm}_1} \simeq |M_2| - \frac{M_W^2}{\mu}\sin 2\beta, \tag{3.20}$$

$$m_{\tilde{\chi}^{\pm}_2} \simeq |\mu| + \frac{M_W^2}{\mu}\sin 2\beta, \tag{3.21}$$

and the states are given by $\tilde{\chi}^{\pm}_1 \approx \tilde{W}^{\pm}$ and $\tilde{\chi}^{\pm}_2 \approx \tilde{H}^{+}_u / \tilde{H}^{-}_d$.

We note that mass degeneracies are expected between the lightest neutralinos and charginos, when their masses are dominantly set by either M_2 or μ, so-called wino and higgsino scenarios. Such scenarios are the focus of Papers 1 and 2 of this thesis.

Chapter 4

Naturalness considerations

Over the past few decades, the concept of *naturalness* has emerged as an important guiding principle in the search for theories that go beyond the Standard Model. In the context of this thesis, naturalness is of interest for several reasons: First, it forms the basis for the hierarchy problem of the Standard Model, perhaps the most widely recognized motivation for TeV-scale supersymmetry. Second, within models of *Natural SUSY*, which we study in Paper 1, naturalness is further used as a principle for formulating more predictive models from the full parameter space of the MSSM. Finally, as we employ Bayesian methods in Papers 1 and 2, it is worth highlighting the fundamental connection between naturalness arguments and Bayesian reasoning.

4.1 Naturalness in particle physics

Historically, the criterion of naturalness in relation to particle physics theories has been formulated in several different, but related, ways.[1] In 1979, Susskind stated that naturalness requires

> *the observable properties of a theory to be stable against minute variations of the fundamental parameters.* [12]

[1] We refer the reader to Ref. [11] for an interesting historical and sociological analysis of the role played by the "naturalness narrative" in particle physics.

Around the same time, 't Hooft formulated a naturalness criterion concerning small parameters and symmetries, stating that a dimensionless theory parameter

> *is allowed to be much smaller than unity only if setting it to zero increases the symmetry of the theory. If this does not happen, the theory is unnatural.* [13]

For the purpose of the present discussion we will adopt an understanding of naturalness rather similar to Susskind's formulation, namely that a theory is unnatural if it requires fine-tuned cancellations of large numbers in order to obtain predictions in agreement with observation. In this sense, if the Standard Model remains valid up to very high energies, *e.g.* the Planck scale $\Lambda_{\text{Pl}} \sim 10^{18}$ GeV, it is a highly unnatural theory since, in order to predict a physical Higgs mass at the electroweak scale, the bare Higgs mass parameter must be tuned to cancel radiative corrections of the size of Λ_{Pl}.

To avoid such — from a naturalness perspective, catastrophic — fine tuning, some new physics is required in order to modify the theory and stabilize the theoretical predictions at a lower energy scale. Underlying this argument is the concept of effective field theory, where physics at high energies only appear as effective couplings when the theory is studied at lower energies.[2] A classic example is Fermi theory, in which the charged current of weak interactions is treated as an effective four-fermion coupling. Just as this approximation breaks down at energies comparable to the W mass, every effective theory comes with a cut-off scale Λ, representing the highest momentum scale where the theory remains valid.

The history of particle physics has seen several examples where a physical prediction, which in an effective theory approach had a divergent dependence on the cut-off Λ, turned out to be tamed by the discovery of new particles. One such example is the mass difference between the K_L^0 and K_S^0 mesons. When computed in Fermi theory valid at the energy scale of the kaon mass,

[2] See Ref. [14] for a discussion of the connection between naturalness arguments and effective field theory, and the limitations that follow from it.

sans any quarks besides u, d and s, the prediction is

$$\frac{m_{K_L^0} - m_{K_S^0}}{m_{K_L^0}} = \frac{G_F^2 f_K^2 \sin^2\theta_c}{6\pi^2}\Lambda^2. \tag{4.1}$$

Here G_F is the Fermi coupling constant, f_K is the kaon decay constant and $\sin\theta_c$ is the Cabibbo angle. To avoid the need for fine-tuned cancellations, the expression in Eq. (4.1) should not be (much) larger than the observed value of 7×10^{-15}. This implies an upper bound on Λ of $\Lambda < 2$ GeV. Thus, a postulated absence of fine-tuning predicts that some new physics affecting the kaons should appear below this energy scale. In this case, the answer was provided by the discovery of the charm quark at $m_c \approx 1.3$ GeV.

On the other hand, there are cases where Nature seems to care little about our naturalness criterion. One important example is the cosmological constant, whose scale is constrained by observations to be around 10^{-3} eV. Our current understanding of particle physics is clearly valid to much higher energies than this, and yet the predictions we obtain for the vacuum energy diverges with the cut-off.[3]

4.2 Natural supersymmetry

As seen in Section 2.2, the introduction of superpartners provides a solution to the big hierarchy problem in the Standard Model. However, we are left with a "little hierarchy problem", coming from the tension between the electroweak scale and the sparticle mass scale, M_{SUSY}.

One manifestation of this tension appears when we require that electroweak symmetry breaking in the MSSM agrees with observations: The vacuum expectation values v_u and v_d of the neutral Higgs fields H_u^0 and H_d^0

[3] As we saw in Section 2.4, a vanishing vacuum energy is a prediction of unbroken supersymmetry, and we could therefore hope to solve this problem in supersymmetric models. However, when supersymmetry is broken the predicted scale for the vacuum energy is the mass scale of the superpartners, meaning that even with TeV-scale supersymmetry we are still left with a huge discrepancy relative to the observed value.

should, at tree-level, be related to the observed Z boson mass as

$$M_Z^2 = \frac{1}{2}(g^2 + g'^2)(v_u^2 + v_d^2), \tag{4.2}$$

where $v_u^2 + v_d^2 = v^2 \approx (174 \text{ GeV})^2$. Requiring that this corresponds to a minimum of the scalar potential generates a relation between M_Z and the relevant MSSM parameters. Recalling that $\tan\beta = v_u/v_d$ and expanding in large $\tan\beta$, the tree-level relation is

$$\begin{aligned} M_Z^2 =& -2(m_{H_u}^2 + |\mu|^2) + \frac{2}{\tan^2\beta}(m_{H_d}^2 - m_{H_u}^2) + \mathcal{O}(1/\tan^4\beta). \\ =& -2(m_{H_u}^2 + |\mu|^2) + \mathcal{O}(1/\tan^2\beta). \end{aligned} \tag{4.3}$$

Thus, the terms on the right-hand side must combine to give the correct value for M_Z. Since $m_{H_u}^2$ is a parameter in the soft supersymmetry-breaking part of the MSSM Lagrangian, its scale is expected to be M_{SUSY}. Consequently, μ must also be of the order of M_{SUSY}. However, μ is a parameter in the supersymmetry-respecting part of the Lagrangian, and there is no *a priori* explanation in the MSSM for why its scale should be correlated with that of the soft-breaking terms. This has become known as the "μ problem", and has inspired model building beyond the MSSM.

Assuming that an explanation can be found for why μ is at M_{SUSY}, Eq. (4.3) still leaves us with the fine-tuning problem of having parameters at M_{SUSY} cancel to produce M_Z at the weak scale. This has lead to the development of Natural SUSY models [15], based on the criteria that the amount of fine-tuning in Eq. (4.3) be kept as low as possible. In general, this is accomplished by ensuring that $m_{H_u}^2$ and μ do not become too large.

Although Eq. (4.3) should be modified by higher-order corrections, which in fact help to alleviate the fine-tuning somewhat [16], we use it here to summarize what has become the standard set of phenomenology predictions for Natural SUSY models.

First, as μ is the higgsino mass parameter, Natural SUSY models predict light higgsinos. In terms of the mass eigenstates of the electroweak gauginos, this implies that, unless either M_1 or M_2 is also small, the two lightest neutralinos and the lightest chargino will all be dominantly higgsino and have

masses of similar size.

Second, $m_{H_u}^2$ receives radiative corrections at the one-loop level from the soft-breaking parameters $m_{Q_3}^2$, $m_{U_3}^2$ and A_t governing the masses of the stops and the left-handed sbottom. It is therefore reasonable to expect these squarks to be rather light in scenarios with low fine-tuning.

Third, the gluino mass parameter M_3 enters in two-loop corrections to $m_{H_u}^2$. Thus, gluinos are not expected to be too heavy in a Natural SUSY scenario either. This is especially relevant for LHC phenomenology, as gluinos are among the sparticles receiving the most stringent constraints from LHC searches.

Based on a quantified measure for what amount of fine-tuning is regarded as acceptable, a topic which we will discuss in the next section, approximate upper bounds on the sparticle masses can be derived. A recent study along these lines is Ref. [17], which for a MSSM scenario finds that higgsinos should be lighter than $\sim$ 600 GeV, stops and the left-handed sbottom should be lighter than $\sim$ 1 TeV, and gluinos lighter than $\sim$ 1.4 TeV. Also, the upper bound on winos is found to be similar to that of gluinos, but this is somewhat less relevant for LHC physics due to the much lower production cross section. The other sparticles are less constrained by naturalness arguments.

The above set of general predictions typically form the starting point for phenomenological studies of Natural SUSY models, including our analysis of the Natural SUSY parameter space in Paper 1. However, it has been shown that such "naturalness bounds" on sparticle masses may heavily depend on the details of the underlying model, especially the question of what is considered to be the fundamental parameters, and at what scale these are defined, see *e.g.* Refs. [17,18] and references within. In particular, in [18] it is shown that having light stops is not necessarily a robust prediction of scenarios with low fine-tuning. For instance, in models with a unified scalar mass parameter m_0 at high scales, the low-scale value of $m_{H_u}^2$ relevant for Eq. (4.3) may depend very weakly on m_0, meaning that a large value for m_0 can give heavy stops without leading to much fine-tuning. Also, the prediction of a light average stop mass may be weakened if other sources of fine-tuning besides Eq. (4.3) are taken into account. For example, in a MSSM scenario with relatively light stops, a certain amount of tuning of the stop mixing

parameters is necessary to obtain a predicted Higgs mass of ~ 125 GeV. We will come back to some of the problems related to naturalness bounds on sparticles in the next section.

4.3 Measuring fine-tuning

Exactly how unnatural or fine-tuned is a given model? For addressing this question, several fine-tuning measures have been constructed and applied in the literature, *e.g.* see Refs. [19–22]. Here we present the widely used Barbieri–Giudice measure: Let $\{\theta_i\}$ be the set of model parameters. For each parameter θ_i, the sensitivity of the predicted M_Z^2 to infinitesimal variations in θ_i is quantified as

$$\Delta_{\theta_i} \equiv \frac{\partial \ln M_Z^2}{\partial \ln \theta_i} = \frac{\theta_i}{M_Z^2} \frac{\partial M_Z^2}{\partial \theta_i}. \tag{4.4}$$

The overall fine-tuning Δ for a point in the model's parameter space is then determined by the most sensitive parameter,

$$\Delta = \max_i \{|\Delta_{\theta_i}|\}. \tag{4.5}$$

To illustrate in what sense this is a measure of fine-tuning, we consider the case of the μ parameter. From Eq. (4.3), taking μ positive and $m_{H_u}^2$ negative, we find that Δ_μ is given by

$$\Delta_\mu = \frac{-4\mu^2}{M_Z^2} \approx \frac{2\mu^2}{\mu^2 - |m_{H_u}^2|}. \tag{4.6}$$

Rearranging this expression, we get

$$\frac{\sqrt{|m_{H_u}^2|}}{\mu} = \sqrt{1 - \frac{2}{\Delta_\mu}} \simeq 1 - \frac{1}{\Delta_\mu}, \tag{4.7}$$

where the last step assumes $\Delta_\mu > 2$. Thus, a sensitivity measure of, for instance, $\Delta_\mu = 100$ would correspond to μ and $\sqrt{|m_{H_u}^2|}$ being equal to an accuracy of $\sim 1\%$. Note that, for simplicity, we have here worked directly

with the parameters at the low energy scale. If the model is defined in terms of parameters at some high scale, the low-scale parameters in Eq. (4.3) should be regarded as functions of these, and the sensitivity should be calculated in terms of the high-scale parameters.

Many analyses of Natural SUSY scenarios are based on first defining what upper bound $\Delta_{\max}$ on the fine-tuning measure is regarded as acceptable, and from this derive bounds on the theory parameters and sparticle masses. The sparticle mass limits quoted in the previous section were all based on $\Delta_{\max} = 100$. But, as emphasized in [17] from which these limits were taken, such results should be interpreted only as rough indications.

One reason for this is the wide range of choices and caveats connected with the use of naturalness measures: First, any study of naturalness is inherently model-dependent. Since correlations among parameters, or the lack thereof, is precisely what determines the value of a fine-tuning measure, working with the full MSSM parameter space is no less model-dependent than choosing a more constrained model. A related concern is the question of exactly what set of parameters to include in the fine-tuning analysis. In particular, should Standard Model parameters such as the top Yukawa coupling be included? Also, should other potential fine-tunings, such as the Higgs mass prediction, be taken into account? Finally, there is the question of what fine-tuning measure to choose; the Barbieri–Giudice measure presented here is only one of several suggested in the literature.

Another, and more fundamental, reason for why parameter and sparticle limits based on naturalness bounds should be interpreted with care, is the fact that the definition of naturalness itself has no well-defined statistical interpretation. Even though naturalness can be quantified using sensitivity measures such as Δ_{θ_i}, these measures cannot — by themselves — tell us anything about the viability of the model being studied. However, as the question of how natural a model is ultimately concerns how *convincing* we find the model to be, a possible solution to the interpretation problem can be found in Bayesian statistics, which provides a well-defined statistical framework for posing exactly these kinds of questions. Therefore, we will revisit the topic of naturalness in Section 5.2, after having introduced the basic concepts of Bayesian statistics.

Chapter 5

Statistics

In physics we construct mathematical models to explain, or at least describe, how nature works. The success of a model depends on how well it agrees with past measurements and its ability to predict the outcomes of future experiments. When confronting models with observations, we seek to answer questions of two basic categories: *hypothesis testing* and *parameter estimation*.

In hypothesis testing we ask whether the predictions of our model are consistent with observations. The model can be a theory with a set of free parameters, or a fully specified scenario, such as a single point in the parameter space of that theory. In questions of parameter estimation we make the assumption that the overall theory is correct, and then try to determine what values for the free parameters are preferred by the data.

Hypothesis testing and parameter estimation are central topics of statistical inference. Broadly speaking, statistical inference is inverted probability theory: Whereas probability theory is mainly concerned with predicting outcomes for random variables given a known model, statistical inference is the attempt to pin down the correct model based on a set of known outcomes. This close connection to probability theory has given rise to two different approaches, based on different interpretations of probability: *frequentist* and *Bayesian* statistics.

In Section 5.1 we review some of the basic differences between frequentist and Bayesian statistics. We devote the most attention to the Bayesian

approach, as it forms the basis for the methods used in Papers 1 and 2. In Section 5.2 we take another look at the concept of naturalness, this time from a Bayesian point of view. Finally, in Section 5.3 we discuss parameter scanning, and in particular the *nested sampling* algorithm used in Papers 1 and 2. For a general introduction to Bayesian methods, see Refs. [23, 24].

5.1 Frequentist and Bayesian statistics

In the frequentist interpretation, probability is defined in terms of the relative frequencies of outcomes in the limit of infinitely many repeated trails. That is, if n_A is the number of trails with the outcome A and N is the total number of repeated trails, the probability $P(A)$ for outcome A is

$$P(A) = \lim_{N\to\infty} \frac{n_A}{N}. \tag{5.1}$$

Thus, probabilities can only be meaningfully assigned to outcomes of a *repeatable* and *random* process. For parameter estimation and hypothesis testing, this means that the questions we are allowed to ask are such that probabilistic statements always refer to quantities that are expected to vary randomly upon (hypothetically) repeated measurements.

As an example, consider a model with a single free parameter θ, predicting a probability density function (p.d.f.) $f(x;\theta)$ for some observable x.[1] Assuming that our model is correct, measurements of x can be used to estimate the true value of the parameter θ. The result of such an analysis is often presented as an interval $(\theta_{\min}, \theta_{\max})$ said to contain the true parameter with a given confidence, *e.g.* 95%. This does *not* mean that there is a 95% chance that the true parameter value is contained in this exact interval — it either is or it is not. Rather, it means that if we could repeat the entire experiment many times, each time constructing a confidence interval in a similar way, the true value of θ should be contained in 95% of these intervals. That is, the probabilistic statement refers to the interval limits since they are functions

[1] For simplicity, we will use the abbreviation p.d.f. also when the observable in question is discrete, although *probability mass function* is the more correct terminology for such cases.

of the data and therefore proper random variables in the frequentist sense.

We find the same principle at play in the case of hypothesis testing, where we ask whether a given model is consistent with the observed data. In order to quantify the deviation between the data set $\{x_1, x_2, \ldots, x_n\}$ and the model expectations, we define a measure $q(x_1, x_2, \ldots, x_n)$ called the *test statistic*. Being a function of the data, the value of q is itself regarded as a random variable with an associated p.d.f., $g(q)$, whose form depends on the data p.d.f. and the definition of $q(x_1, x_2, \ldots, x_n)$. In most practical applications $g(q)$ must be determined from Monte Carlo simulation or by assuming the validity of some limit where the distribution of q is known analytically. A common example of the latter is when the data is assumed to be Gaussian distributed, $x_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, and q is defined as $q = \sum_i (x_i - \mu_i)^2/\sigma_i^2$, in which case $g(q)$ will be a χ^2 distribution. Once the distribution of q is determined, we calculate q_{obs} from the observed data and interpret this as a sample drawn from $g(q)$. If q_{obs} is sufficiently far out in the tail of $g(q)$, meaning that the observed data show a surprisingly large deviation from the model expectations, the model is rejected.[2] Thus, as for parameter estimation, our conclusion is based on interpreting a function of the observed data probabilistically based on how we expect this quantity to vary for repeated experiments.

Before moving on to Bayesian statistics we take the time to introduce the *likelihood* as it is one of the most important quantities in both frequentist and Bayesian analysis, and central to all the work presented in this thesis. The likelihood describes how the predicted probability for the observed data varies as a function of the model parameters. Consider a model with two free parameters θ_1 and θ_2, predicting a joint p.d.f. $f(x, y, z; \theta_1, \theta_2)$ for three observables x, y and z. To obtain the likelihood we simply insert the observed data set $\{x_{\text{obs}}, y_{\text{obs}}, z_{\text{obs}}\}$ into $f(x, y, z; \theta_1, \theta_2)$ and interpret the resulting expression as a function of θ_1 and θ_2:

$$\mathcal{L}(\theta_1, \theta_2) = f(x_{\text{obs}}, y_{\text{obs}}, z_{\text{obs}}; \theta_1, \theta_2). \tag{5.2}$$

Often the observables can be regarded as independent, in which case the

[2] A common criterion for rejecting the model is that the probability contained in the tail of $g(q)$ outside q_{obs}, *i.e.* the p-value, is less than 5%.

joint p.d.f. can be factorized as $f_x(x;\theta_1,\theta_2)f_y(y;\theta_1,\theta_2)f_z(z;\theta_1,\theta_2)$, leading to a similarly factorized likelihood function. For parameter estimation, the likelihood can be used to construct confidence intervals and obtain point estimates, such as the *maximum likelihood estimator* $(\hat{\theta}_1,\hat{\theta}_2)$ found by maximising $\mathcal{L}$ over the (θ_1,θ_2) space. In hypothesis testing the likelihood often serves as the basis for constructing the test statistic.

Although $\mathcal{L}(\theta_1,\theta_2)$ is constructed from a p.d.f., it is important to note that it cannot be interpreted as some sort of p.d.f. for the parameters. First, this would not be defined in the frequentist approach. Second, the likelihood function generally does not have the properties required by a p.d.f., as the functional dependence of $\mathcal{L}$ on the parameters (θ_1,θ_2) will not be the same as the dependence of f on (x,y,z). The distinction between a likelihood and a p.d.f. is important to make as likelihood functions are often described by referring to the form of the underlying p.d.f. For example, the Poisson distribution

$$f(n;\nu) = \frac{\nu^n e^{-\nu}}{n!} \tag{5.3}$$

is a discrete function of the observable n, while the corresponding Poisson likelihood $\mathcal{L}(\nu)$ is a continuous function of ν.

If we are only interested in making inferences about one of the parameters, θ_1, the other parameter θ_2 is regarded as a *nuisance parameter* and we seek to eliminate it from our analysis. Several approaches exist, but for frequentist analyses in high-energy physics the most common solution is to maximize $\mathcal{L}(\theta_1,\theta_2)$ over θ_2 for each point in θ_1, producing what is known as the *profile likelihood*

$$\mathcal{L}_p(\theta_1) = \mathcal{L}(\theta_1,\hat{\hat{\theta}}_2). \tag{5.4}$$

The double hat notation indicates that the maximisation over θ_2 is conditional on the value of θ_1.

We now turn to Bayesian statistics, which employs a very different interpretation of probability. Here the probability $P(A)$ is defined as a *degree of belief* in hypothesis A or a *state of knowledge* about hypothesis A. Thus, probability is no longer based on the concept of repeatable trails, and this greatly extends the range of problems that can be treated in a probabilistic manner. In fact, starting from an axiomatic approach to *plausible reasoning*,

the rules of probability theory can be derived [25], implying that the mathematics of probability theory is well suited for the sort of reasoning under uncertainty that is so central to scientific research. This wider interpretation of probability allows us to ask probabilistic questions concerning models and the values of physical parameters, such as "What is the probability for sub-TeV sparticles in the MSSM, given the available information?". In a frequentist framework this question is nonsensical as it does not involve any repeatable experiment.

The central tool of Bayesian statistics is the well-known Bayes' theorem,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{5.5}$$

which expresses the probability of A given B, $P(A|B)$, in terms of the probability for B given A and the unconditional probabilities for A and B. Being a direct consequence of the sum and product rules of probability theory, this theorem is equally valid in both the frequentist and Bayesian interpretation. But if we replace A with some hypothesis H and B with some data D, and condition all probabilities on any background information I, we arrive at a formulation of the theorem that is only valid in — and completely central to — the Bayesian approach:

$$P(H|D, I) = \frac{P(D|H, I)P(H|I)}{P(D|I)}. \tag{5.6}$$

The left-hand side of this equation, $P(H|D, I)$, is called the *Bayesian posterior probability* or simply the *posterior*, and it is the central quantity we wish to obtain in a Bayesian analysis. It represents our degree of belief in the hypothesis after confronting it with some new data D.

On the right-hand side, the first factor in the nominator, $P(D|H, I)$, represents the probability for obtaining the observed data under the assumption that H is true. When interpreted as a function of H, this is simply the likelihood introduced above, and it is through this quantity that the physical predictions of H enter the equation.

The second factor in the nominator, $P(H|I)$, is known as the *prior probability* (or simply the *prior*) and is the source of much heated debate between

frequentists and Bayesians. It expresses our degree of belief in the hypothesis based on background information alone, *i.e.* prior to taking the data D into account. Thus, Eq. (5.6) can be understood as a prescription for how we should update our degree of belief in H in light of new data, *i.e.*, how to go from $P(H|I)$ to $P(H|D, I)$. While the prior provides a way of including additional information in our inference, it at the same time raises difficult questions regarding how exactly this information should be encoded, and whether it is acceptable that priors can differ from one person to the next. In some sense, having to deal with these difficulties is the price to pay for the wider probability interpretation of Bayesian statistics. We will return to these questions in subsequent sections.

The denominator of Eq. (5.6), $P(D|I)$, is referred to as the *Bayesian evidence*. It corresponds to the expression in the nominator *marginalized* over all possible realisations of H. If we are considering a discrete set of mutually exclusive hypothesis, H_i, for instance different values of a discrete parameter, marginalisation amounts to calculating the sum

$$P(D|I) = \sum_i P(D|H_i, I)P(H_i|I), \tag{5.7}$$

where the set of possible hypothesis must be exhaustive, $\sum_i P(H_i|I) = 1$. More often, H will represent a given point in a space of continuous parameters, such as the parameter space of the MSSM. In this case the prior and posterior probabilities in Bayes' theorem become p.d.f.'s, and the marginalisation sum in Eq. (5.7) turns into an integral over the entire parameter space of the model. Since the dependence on the model parameters is integrated out, the evidence does not affect the shape of the posterior p.d.f. across the parameter space. This means that for parameter estimation purposes it is nothing more than a constant ensuring the proper normalisation of the posterior p.d.f. On the other hand, the evidence takes center stage in applications of Bayesian hypothesis testing.

For an illustration of Bayesian parameter estimation, consider a model with a single free parameter θ, being confronted with a set of measurements D. If we rewrite Eq. (5.6) using a more compact notation, the posterior p.d.f

for θ can be expressed as

$$P(\theta|D) = \frac{\mathcal{L}(\theta)\pi(\theta)}{\mathcal{Z}}. \tag{5.8}$$

Here $P(\theta|D)$ is the posterior, $\mathcal{L}(\theta)$ is the likelihood, $\pi(\theta)$ denotes the prior and $\mathcal{Z}$ is the evidence, given by

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)\, d\theta. \tag{5.9}$$

Despite the simplified notation it should be kept in mind that all the above quantities are conditioned on the overall model being true, as well as any background information. (In Bayesian statistics, truly unconditional probabilities are hard to come by.)

The most complete way of presenting results in Bayesian inference is to present the posterior distribution directly. However, it is often necessary to condense the result down to a point estimate or interval for the parameter of interest. The mean, median or maximum of $P(\theta|D)$ are commonly used point estimates, while for intervals we construct what is called a *credible interval*. This is simply any interval $(\theta_{\min}, \theta_{\max})$ that contains a given fraction p of the posterior probability:

$$p = \int_{\theta_{\min}}^{\theta_{\max}} P(\theta|D)\, d\theta. \tag{5.10}$$

As this requirement is not enough to uniquely determine the interval, a common additional requirement is to choose the interval such that it maximizes the posterior p.d.f. inside the interval.

When the posterior is a joint p.d.f. for several parameters, *e.g.* $P(\theta_1, \theta_2|D)$, the concept of a credible interval generalizes to a *credible region* in the higher-dimensional parameter space. If θ_2 is an uninteresting nuisance parameter that we want to eliminate, the fact that the posterior is a p.d.f. implies that integration over θ_2 is the natural procedure:

$$P_{\theta_1}(\theta_1|D) = \int P(\theta_1, \theta_2|D)\, d\theta_2. \tag{5.11}$$

In many applications we are more interested in the posterior of some func-

tion of model parameters rather than the posterior of the model parameters themselves. This is the case in Papers 1 and 2, where we mainly focus on posterior distributions for observables such as particle masses and lifetimes. For illustration, consider a function $f(\theta)$ of some model parameter θ. Our task is now to perform a change of variable in the posterior from θ to f. Since the value of f is uniquely determined by the value of θ, a p.d.f. for f conditional on θ takes the simple form of a delta function. This is of course independent of any data D, allowing us to express the joint posterior in the space of f and θ as

$$\begin{aligned} P(f, \theta|D) &= P(f|\theta, D)P(\theta|D) \\ &= \delta(f(\theta) - f)P(\theta|D). \end{aligned} \tag{5.12}$$

To obtain the posterior for f we now simply marginalize over the parameter θ,

$$P(f|D) = \int \delta(f(\theta) - f)P(\theta|D)\, d\theta. \tag{5.13}$$

Thus, the posterior for $f(\theta)$ is obtained by weighting all values of f by the posterior probability for the corresponding value(s) of θ.

In numerical studies we often work with a set of samples drawn from $P(\theta|D)$ rather than with the distribution itself. Equations (5.12) and (5.13) then tell us that the posterior for f can be approximated in the following way: First we determine where our samples fall in the (f, θ) space by calculating $f_i = f(\theta_i)$ for each sample θ_i. Then we marginalize over θ by histogramming the samples in terms of f only. This simple procedure readily generalizes to situations with multiple parameters and observables.

Finally, we consider Bayesian hypothesis testing. So far the hypothesis H in Eq. (5.6) has referred to one out of many possible realisations, *e.g.*, a point in the parameter space of some overall theory T assumed to be true. In this case, the Bayesian evidence $P(D|I)$, which we more properly should have denoted $P(D|T, I)$, could be calculated by marginalising over all possible H, as shown in Eqs. (5.7) and (5.9). If we now turn to the case where the theory

T itself is the hypothesis under study, Bayes' theorem becomes

$$P(T|D,I) = \frac{P(D|T,I)P(T|I)}{P(D|I)}, \tag{5.14}$$

where we recognize that the evidence $P(D|T,I)$ now appears in the nominator. However, the denominator $P(D|I)$ is now problematic: As long as we have not specified any alternative theories to T, the only well-defined value for $P(D|I)$ is the trivial case $P(D|I) = P(D|T,I)P(T|I)$ with the "certainty prior" $P(T|I) = 1$, which, of course, also leads to a posterior $P(T|D,I) = 1$. Thus, Bayes' theorem tells us that we cannot asses how well a hypothesis agrees with data without having at least one well-defined alternative hypothesis.

Hypothesis testing within the Bayesian framework is therefore a question of *model comparison*: Given two models T_1 and T_2, the central quantity is the *posterior odds*,

$$\frac{P(T_1|D,I)}{P(T_2|D,I)} = \frac{P(D|T_1,I)}{P(D|T_2,I)}\frac{P(T_1|I)}{P(T_2|I)}, \tag{5.15}$$

representing our relative degree of belief between the two models after the data D is taken into account. On the right-hand side, $P(T_1|I)/P(T_2|I)$ is the prior odds, while $P(D|T_1,I)/P(D|T_2,I)$ is the evidence ratio, usually called the *Bayes factor*, which encodes how the new information contained in the data should update our relative degree of belief. Analogous to how a set of conventional p-values are used to draw conclusions in frequentist hypothesis testing, the results of Bayesian model comparison are often interpreted using *Jeffreys' scale*, see Ref. [26].

5.1.1 Subjective beliefs and objective frequencies?

The debate between the frequentist and Bayesian view on probability and statistical inference is a long and at times heated one, and we cannot do it full justice here. Still, as the work presented in this thesis makes use of both approaches, a few comments are in order.

As a degree of belief can differ from person to person, critics of the

Bayesian approach argue that this interpretation of probability leads to an unacceptable level of subjectivity in Bayesian inference. In particular, having to include a prior belief for the hypothesis under study is seen as problematic. The Bayesian response to this criticism comes in two main flavours: The *subjective Bayesian* approach prefers to fully embrace the subjective nature of Bayesian priors, arguing that the application of a subjective prior simply reflects what we all do when interpreting data anyway. The Bayesian formalism simply forces us to be honest about any assumptions.

The other response has taken the approach of developing objective rules and principles for how priors should be assigned. Under this *objective Bayesian* approach, a difference in assigned prior should only arise from a difference in background information; the objectivity of the method is ensured as long as two people with identical information assign identical priors [24]. We come back to some of the results of this approach when discussing prior assignment in the next section.

An interesting question is what happens when the tools of Bayesian inference are employed on a problem where all information is of a purely frequentist type. For instance, consider an experiment that measures a forward-backward symmetry in particle collisions. We are interested in determining the long-run relative frequency f of forward events,

$$f = \lim_{N\to\infty} \frac{n_F}{N}, \tag{5.16}$$

where n_F is the number of forward events and N denotes the total number of events. In frequentist terms this is of course nothing but the probability for a forward event, $P(F)$. Our only piece of prior information is the simple observation that there are only two possible outcomes per event: forward or backward. Based on this information alone a flat prior is assigned to f. In Jaynes [24] a similar example is worked out in detail. The result is that, after observing n_F forward events in N events total, the posterior distribution for f is given by

$$P(f|n_F, N) = \frac{(N+1)!}{N!(N-n_F)!} f^{n_F}(1-f)^{N-n_F}. \tag{5.17}$$

The f value of maximum posterior probability, *i.e.* the long-run relative frequency with the highest degree-of-belief, is $f = n_F/N$, in perfect agreement with frequentist reasoning. The mean value for f, which is often a better estimator for small samples, is given by

$$\bar{f} = \frac{n_F + 1}{N + 2}. \tag{5.18}$$

This turns out to be identical to the central value of a frequentist confidence interval for the same problem. The +1 and +2 terms can be understood from the fact that our prior knowledge told us that *both* outcomes actually are possible, which amounts to the information contained by two "prior events", one of which was forward. If we did not even know this much, our prior should not have been flat.

We can also consider a different question: In the limit of $N \to \infty, n_F/N \to f$, that is when the frequentist probability for forward events is known from an infinite number of observations, what is the probability for observing m_F forward events in the next M trails? The posterior distribution for $P(m_F|M, f)$ turns out to be the binomial distribution. Thus, in the case of "pure" frequentist knowledge, the degree of belief one should assign to the outcome m_F is precisely the same as the probability a frequentist would calculate.

The above examples suggest that if relative frequencies are all we care about, a Bayesian and a frequentist analysis can reach very similar, and at times identical, conclusions. Arguing from the Bayesian point of view, Jaynes states that

> *this equivalence shows why it is so easy to confuse the notion of probability and frequency, and why in many problems this confusion does no harm.* [24](p. 578)

Regardless of whether there is any underlying confusion, it seems plausible that at least part of the scepticism towards Bayesian methods may come down to habit: When most of the problems we work on are those in which relative frequencies and degrees of belief behave similarly, it is perhaps easy to conclude that relative frequencies is the preferred basis for all inference.

Returning to the case where the Bayesian analysis includes background information of any type, we can ask to what extent the assigned prior will affect the resulting posterior p.d.f. This degree of *prior dependence* will depend on the strength of the data and we can investigate it by repeating the analysis with a different assigned prior. A large variation between the resulting posteriors tells us that the current data are not strong enough to dominate our prior information, and that any conclusions based on these data should be interpreted with care, regardless of analysis method. The other extreme case is when the information in the data completely dominates the prior. For instance, this would be the case if we were determining the mass of a new particle in the range 100 – 1000 GeV, and the likelihood from the experiment was a peak with a width of a few GeV. Unless we had very strong prior information, any reasonable prior should be approximately flat over such a small interval, meaning that the shape of the posterior would be completely determined by the shape of the likelihood. In this limit of overwhelming data, the different questions asked by the frequentist and the Bayesian approach will end up giving very similar answers, even if they technically should still be interpreted differently. This is not an argument for or against either statistical approach, but for physicists it may be reassuring to know that if we are successful enough in gathering data, our knowledge about nature should ultimately not depend on our choice of statistical philosophy.

5.1.2 Objective priors

For the work presented in this thesis we are following the objective Bayesian approach to assigning prior probability distributions. In short, this means that priors are assigned according to a set of principles for how information, or the lack thereof, should be encoded in a probability distribution.

The simplest principle is what is known as *transformation group invariance*, and is an extension of the *principle of insufficient reason* dating back to Bernoulli in 1713 [27]. It is applicable when we have no information about a parameter apart from what role it plays in our theory. The principle holds that the prior should be invariant under any transformation that is considered irrelevant to the problem. For a *location parameter*, *e.g.* the mean

of a Gaussian distribution, performing a coordinate translation $x' = x + a$ should not affect our prior degree of belief. We therefore require that the prior p.d.f. $\pi(x|I)$ satisfies

$$\pi(x|I)\,dx = \pi(x+a|I)\,d(x+a). \tag{5.19}$$

Since $d(x+a) = dx$, this reduces to the requirement

$$\pi(x|I) = \pi(x+a|I), \tag{5.20}$$

which means that $\pi(x|I)$ must be a uniform p.d.f. We often refer to this as a *flat prior*.

Another important class of parameters is *scale parameters*, for example the standard deviation of a Gaussian distribution. Scale parameters are dimensionful parameters that introduce a definite scale in the problem. For the work presented in this thesis, a scale parameter will typically be some Lagrangian mass parameter m. Ignorance as to what is the correct scale for m implies that our prior should not change under a scaling $m' = cm$. The requirement on our prior p.d.f. then becomes

$$\begin{aligned} \pi(m|I)\,dm &= \pi(cm|I)\,d(cm) \\ &= \pi(cm|I)\,c\,dm, \end{aligned} \tag{5.21}$$

which is satisfied if $\pi(m|I)$ has the form $\pi(m|I) \propto 1/m$. A simple change of variable shows that this corresponds to $\pi(\log(m)|I)$ being uniform and we therefore call this a *log prior*.

Technically, both flat and log priors are what is called *improper priors* in that they do not integrate to unity. Such priors are still used in Bayesian analysis as they can produce proper normalized posterior p.d.f.'s given that the likelihood falls off quickly enough. However, all models studied in this thesis are motivated by their possible relevance for physics at the LHC, meaning that the mass scale must be in the GeV–TeV range. The problem of improper priors can therefore be overcome simply by setting hard cuts on the allowed parameter ranges. For instance, for a mass parameter m we will typically construct a prior that has the shape of a log prior on some range $(m_{\min}, m_{\max})$

and is 0 everywhere else. In a Bayesian analysis such limiting of the allowed parameter ranges are formulated in terms of the parameter priors, but as illustrated by our analyses in Papers 3 and 4, similar choices must be made in a frequentist analysis.

In the above cases our lack of information combined with the invariance requirement was enough to uniquely determine the shape of the prior p.d.f. But how should we assign priors to parameters that we know should satisfy some set of constraints? This will be the case for standard model parameters such as the Z mass, for which we know the mean μ_{M_Z} and variance $\sigma^2_{M_Z}$. Any valid prior $\pi(M_Z|I)$ should therefore satisfy

$$\mathrm{E}[(M_Z - \mu_{M_Z})^2] = \int (M_Z - \mu_{M_Z})^2 \pi(M_Z|I)\, dM_Z = \sigma^2_{M_Z}. \tag{5.22}$$

One possible approach for uniquely determining the prior is that of *maximum entropy* due to Jayens [28]. The basis of this approach is the use of Shannon entropy as a measure for information, with higher entropy corresponding to less information. As the name suggests, the maximum entropy approach tells us to choose the prior that maximize entropy while still satisfying the given constraints. In other words, we seek the prior that adds the minimum amount of extra a priori information. For the common case where all we know is a mean and variance, such as for M_Z, it turns out that the maximum entropy prior is nothing but the Gaussian distribution,

$$\pi(M_Z|I) = \frac{1}{\sigma_{M_Z}\sqrt{2\pi}} \exp\left[-\frac{(M_Z - \mu_{M_Z})^2}{2\sigma^2_{M_Z}}\right]. \tag{5.23}$$

From a Bayesian perspective, the common application of Gaussian p.d.f.'s in inference problems can therefore be justified on the grounds that it is the most conservative description of our state of knowledge, given that all we know is a mean and variance. For a review of the topic of objective priors, see Ref. [29].

5.2 Naturalness from a Bayesian perspective

We can now return to the concept of naturalness introduced in Chapter 4. As the Bayesian framework allows us to ask probabilistic questions at the level of the entire parameter space of a model, it has been suggested that the starting point for an analysis of naturalness should simply be *the probability for the observed data in the model* [30]. This is nothing but $P(D|T, I)$. Interpreted as a function of D, this is the so-called *prior predictive probability*, that is, the *a priori* probability for the data D, given that the model is true. If we take D to represent the actual observed data and interpret $P(D|T, I)$ as a function of the model T, it becomes the Bayesian evidence for the model T provided by the data D.[3]

The above approach to naturalness seems to agree well with our intuition: The Standard Model (SM) is a highly unnatural theory because the extreme fine-tuning required in Eq. (2.20) implies that only an exceedingly narrow range of parameter values will predict a value for v close to the observed value, $v \approx 174$ GeV. Given that we had no *a priori* reason to prefer exactly this range of parameter values, $P(v = 174\,\text{GeV}|\text{SM}, I)$ will turn out vanishingly small.

The dynamics captured by the Bayesian evidence can be further illustrated if we express it as a marginalization over the model parameter space,

$$P(D|T, I) = \int \mathcal{L}(\theta)\pi(\theta)\, d\theta, \tag{5.24}$$

where we have used the simplified notation for the likelihood and prior introduced in Eq. (5.9). We see that the evidence becomes large only when there is reasonable agreement between predictions and observations, that is, a sizeable likelihood $\mathcal{L}(\theta)$, over a region of parameter space containing a significant fraction of prior probability $\pi(\theta)\, d\theta$. This also means that the evidence represents a formalized version of *Occam's razor*: Given comparable agreement with data, a model in which the prior probability is spread across a huge parameter space will obtain smaller evidence compared to a model

[3]Note that when interpreted as a function of T, $P(D|T, I)$ is no longer a probability. The situation is exactly analogous to how a parametrized p.d.f. becomes a likelihood function when viewed as a function of the parameters.

with fewer parameters or a more focused prior distribution.

Interestingly, the Bayesian evidence also contains within it the Barbieri–Giudice sensitivity measure defined in Eq. (4.4) [31]. To illustrate this, we consider the case where our model T is the MSSM. Denoting as θ_i all MSSM parameters except μ, the evidence is given by a marginalization over the MSSM parameter space:

$$P(D|T,I) = \iint \mathcal{L}(\theta_i,\mu)\pi(\theta_i,\mu)\, d\theta_i\, d\mu. \tag{5.25}$$

What is usually done in MSSM studies is to use the precise measurement of M_Z to solve for one theory parameter, often μ. A more formal Bayesian approach is to rather include the M_Z measurement in the likelihood on equal footing with all other measurements. Approximating the likelihood contribution from M_Z by a delta function, we can write

$$\mathcal{L}(\theta_i,\mu) = \delta(M_Z(\theta_i,\mu) - M_Z^{\rm obs})\mathcal{L}_{\rm rest}(\theta_i,\mu), \tag{5.26}$$

where $M_Z^{\rm obs}$ is the measured Z boson mass and $\mathcal{L}_{\rm rest}(\theta_i,\mu)$ is the remaining part of the likelihood function. Further, we take the parameter priors to be independent and choose a log prior for μ,

$$\pi(\theta_i,\mu) = \pi(\mu)\pi(\theta_i) \propto \frac{1}{\mu}\pi(\theta_i). \tag{5.27}$$

If we now perform a change of integration variable from μ to M_Z in Eq. (5.25) and carry out the integration over M_Z, we find

$$P(D|T,I) \propto \int \mathcal{L}_{\rm rest}(\theta_i,\mu_Z)\pi(\theta_i)\frac{1}{|\Delta_\mu|_{\mu_Z}}\, d\theta_i, \tag{5.28}$$

where μ_Z is the μ value that predicts the observed $M_Z^{\rm obs}$ for a given θ_i. We see that the Barbieri–Giudice sensitivity measure for μ appears, through the Jacobian, as a suppression factor in the evidence. In [30] it is shown that the Barbieri–Giudice measure is actually a special limit of a more general sensitivity measure automatically appearing in the Bayesian approach.

If we accept that naturalness can be described by the Bayesian evidence,

what evidence values should we regard as acceptable? As it turns out, this is not a well-defined question. Since the evidence is a p.d.f. when viewed as a function of the data D, it is a dimensionful quantity whose value will depend on our parametrisation of the data. As we saw when discussing model comparison in Section 5.1, the Bayesian evidence for a model T_1 can only be meaningfully interpreted *relative* to the evidence for an alternative model T_2, that is, through the Bayes factor,

$$B_{12} = \frac{P(D|T_1, I)}{P(D|T_2, I)}. \tag{5.29}$$

Thus, from the Bayesian point of view, we should not attempt to asses the naturalness of individual models, but rather compare the evidence for competing models through Bayesian model comparison, for which Jeffrey's scale provides a conventional guide for drawing conclusions.

5.3 Exploring parameter spaces

One of the most common tasks of high-energy phenomenology is to identify the most interesting parameter regions within some theory. Usually, what is meant by "interesting parameter regions" is regions of parameter space that are preferred by comparison to current data; this is the interpretation we will assume below. But parameter regions can also be regarded as interesting for other reasons, such as the prediction of special experimental signatures or a high expected sensitivity for some future experiment. Papers 1 and 2 touch on aspects of the former, while an example of the latter can be found in [32]. In any case, the underlying task is that of parameter estimation.

The basic challenge in parameter estimation problems is to explore the likelihood function across the model's parameter space. Once we know this function we can start making inferences about the preferred parameter values. For a frequentist analysis inferences will be based on the likelihood alone, while a Bayesian analysis will combine the likelihood with the prior p.d.f. to make inferences based on the posterior p.d.f.

In all but the simplest cases parameter estimation problems have to be treated numerically, meaning that we also need to worry about whether we

can explore the parameter space with sufficient accuracy. If the likelihood is computationally cheap and the model only has a few free parameters, we can easily evaluate the likelihood over a dense grid of points covering the entire parameter space. All else being equal, an increase in the computational cost of the likelihood must be compensated by a decreased density of grid points, resulting in a less accurate description of how the likelihood function behaves. How large the grid spacing can be without severely affecting our results depends on how quickly the likelihood varies across parameter space. Our analysis in Paper 4 illustrates the simple grid scan approach for a model with only two free parameters.

When the number of parameters increase, the grid scan method quickly falls victim to the "curse of dimensionality": If n is the desired number of evaluation points per parameter and d is the dimensionality of the parameter space, the required number of likelihood evaluations is n^d. For models with more than a few parameters this quickly becomes computationally intractable. The same applies if we instead of a grid scan choose to sample the parameter space randomly.

The solution to this problem is to realize that we do not need to know the behaviour of the likelihood with high accuracy across the entire parameter space. In order to determine the preferred parameter values we only need to accurately explore regions where the likelihood is reasonably high. The challenge then becomes one of optimization: how to identify and explore these parameter regions with the required accuracy, using as few likelihood evaluations as possible.

In determining the algorithm best suited for solving this problem, two factors turn out to be of special importance: First, the likelihood will often be multimodal, as separate regions in parameter space can give predictions that are in comparable agreement with the data. Also, when Monte Carlo methods are used for simulation or integration, false local maxima can occur due to statistical fluctuations. Second, the gradient of the likelihood is usually not known and can only be approximated through additional likelihood evaluations. Such complications imply that Monte Carlo-based sampling algorithms are better suited than deterministic ones. A probabilistic sampling method will be less prone to getting stuck in the region of a local likelihood

maxima, and the algorithms generally do not rely on knowing the gradient. For the analyses in Papers 1 and 2 we have used the Monte Carlo-based technique of nested sampling as implemented in the `MultiNest` package [33]. We present this algorithm in the next section.

5.3.1 Nested sampling and the MultiNest algorithm

The nested sampling algorithm due to Skilling [34] was developed to solve the computationally difficult problem of evaluating the Bayesian evidence $\mathcal{Z}$ in models with multidimensional parameter spaces,

$$\mathcal{Z} = \int \mathcal{L}(\boldsymbol{\Theta})\pi(\boldsymbol{\Theta})\, d\boldsymbol{\Theta}, \tag{5.30}$$

where $\boldsymbol{\Theta}$ is vector notation for a collection of parameters. As noted previously, the evidence is mainly of importance for model comparison studies. But as a by-product of nested sampling we also obtain a set of parameter space samples $\boldsymbol{\Theta}_i$ drawn according to the posterior p.d.f. $P(\boldsymbol{\Theta}|D)$, and this makes the technique suitable also for parameter estimation studies. It is for this purpose that nested sampling has been used in Papers 1 and 2.

The starting point of nested sampling is the introduction of a new variable called the *prior mass*, ξ, defined by

$$d\xi = \pi(\boldsymbol{\Theta})\, d\boldsymbol{\Theta}, \tag{5.31}$$

such that

$$\xi(\lambda) = \int\limits_{\mathcal{L}(\boldsymbol{\Theta})>\lambda} \pi(\boldsymbol{\Theta})\, d\boldsymbol{\Theta}. \tag{5.32}$$

That is, the prior mass is the amount of prior probability contained within the region(s) of parameter space where the likelihood is greater than some value λ, see Fig. 5.1. This means that $\xi(\lambda)$ is a decreasing function of λ, from $\xi(0) = 1$ to $\xi(\mathcal{L}_{\max}) = 0$. The inverse function, which conventionally is denoted $\mathcal{L}(\xi)$,[4]

$$\mathcal{L}(\xi(\lambda)) \equiv \lambda, \tag{5.33}$$

[4]Note that $\mathcal{L}(\xi)$, taking the scalar argument ξ, should not be confused with $\mathcal{L}(\boldsymbol{\Theta})$, which depends on the vector argument $\boldsymbol{\Theta}$.

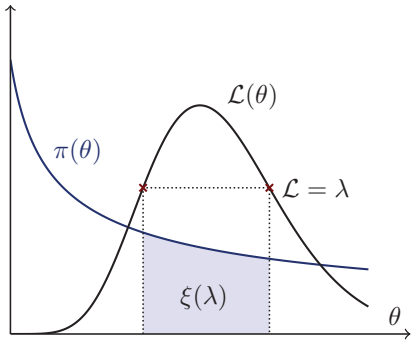

Figure 5.1: The prior mass $\xi(\lambda)$ is the integrated prior p.d.f. $\pi(\theta)$ over the region of parameter space where the likelihood $\mathcal{L}(\theta)$ is greater than λ.

tells us the value for the likelihood contour that contains a given prior mass ξ.

The reason for introducing the prior mass is that the multidimensional evidence integral in Equation (5.30) now can be expressed as a one-dimensional integral over ξ:

$$\mathcal{Z} = \int_0^1 \mathcal{L}(\xi)\, d\xi. \tag{5.34}$$

If we can obtain a set of ordered samples ξ_i along with the corresponding likelihood values $\mathcal{L}(\xi_i)$, the above integral can be solved by standard methods of numerical integration. The name "nested sampling" derives from the way in which this set of samples is obtained:

1. Draw a set of N "live points" $\boldsymbol{\Theta}_j$ according to the prior distribution $\pi(\boldsymbol{\Theta})$.

2. Evaluate the corresponding likelihoods $\mathcal{L}_j = \mathcal{L}(\boldsymbol{\Theta}_j)$.

3. Discard the point with the lowest likelihood, $\boldsymbol{\Theta}_{\text{discard}}$.

4. Pick a new point $\boldsymbol{\Theta}_{\text{new}}$ from $\pi(\boldsymbol{\Theta})$ under the constraint $\mathcal{L}(\boldsymbol{\Theta}_{\text{new}}) > \mathcal{L}(\boldsymbol{\Theta}_{\text{discard}})$.

5. Repeat from step 3.

By iteratively replacing the point of lowest likelihood by a new point with higher likelihood, the entire set of live points will gradually move towards regions of higher likelihood. Thus, the *discarded* points will form a set of samples of increasing likelihood,

$$0 < \mathcal{L}_1 < \mathcal{L}_2 < \dots . \tag{5.35}$$

If ξ_i is the amount of prior mass contained within the likelihood contour of $\mathcal{L}_i$, the likelihood samples will correspond to samples decreasing in prior mass,

$$1 > \xi_1 > \xi_2 > \dots . \tag{5.36}$$

This is exactly the sort of ordered sampling in ξ and $\mathcal{L}(\xi)$ that we wanted for evaluating the integral in Equation (5.34) numerically. However, the exact values ξ_i are not known and must be approximated.

The relation $d\xi = \pi(\boldsymbol{\Theta})\, d\boldsymbol{\Theta}$ implies that sampling parameter points according to $\pi(\boldsymbol{\Theta})$ will correspond to a uniform sampling in ξ. The additional sampling constraint $\mathcal{L}(\boldsymbol{\Theta}_{\text{new}}) > \mathcal{L}(\boldsymbol{\Theta}_{\text{discard}})$ ensures that the prior mass associated with the new sample is less than the prior mass for the discarded sample, $\xi_{\text{new}} < \xi_{\text{discard}}$. Thus, at the beginning of iteration i we have N samples uniformly distributed on the interval $(0, \xi_{i-1})$. The prior mass ξ_i associated with the next point to be discarded is therefore a random variable,

$$\xi_i = t_i\, \xi_{i-1}, \tag{5.37}$$

where t_i follows the p.d.f. $f(t)$ for the largest of N numbers drawn uniformly from $[0, 1]$:

$$f(t) = N t^{N-1}. \tag{5.38}$$

Since we start by sampling from the entire prior mass, $\xi_0 = 1$, we can express ξ_t as

$$\xi_i = t_i\, t_{i-1} \dots t_1, \tag{5.39}$$

or in terms of $\ln \xi_i$,

$$\ln \xi_i = \ln t_i + \ln t_{i-1} + \cdots + \ln t_1. \tag{5.40}$$

From the p.d.f. in Equation (5.38) we find the expectation value and variance for $\ln t$ to be

$$\begin{aligned} \mu_{\ln t} &= -\frac{1}{N}, \\ \sigma^2_{\ln t} &= \frac{1}{N^2}. \end{aligned} \tag{5.41}$$

We can now approximate $\ln \xi_i$ by[5]

$$\ln \xi_i \approx -\frac{i}{N} \pm \frac{\sqrt{i}}{N}. \tag{5.42}$$

Based on this approximation the prior mass ξ_i contained within the contour of $\mathcal{L}_i$ is assumed to be

$$\xi_i = \exp\left[-\frac{i}{N}\right]. \tag{5.43}$$

With known values for both the likelihood and prior mass at each iteration we can approximate the evidence integral in Equation (5.34). After M iterations we have

$$\mathcal{Z} \approx \sum_{i=1}^{M} \mathcal{L}_i w_i = \sum_{i=1}^{M} \mathcal{L}_i \frac{1}{2}(\xi_{i-1} - \xi_{i+1}), \tag{5.44}$$

where the weight $w_i = \frac{1}{2}(\xi_{i-1} - \xi_{i+1})$ represents the unique slice of prior mass associated with the likelihood value $\mathcal{L}_i$, here chosen according to the trapezoidal rule. The sampling algorithm stops when the largest possible contribution $\Delta\mathcal{Z}$ from the current set of live points is negligible compared to the estimate in Eq. (5.44). The uncertainty in the final estimate for $\mathcal{Z}$ is dominated by the approximation made in Eq. (5.42). For a discussion of this, see Ref. [23].

Once the evidence $\mathcal{Z}$ has been determined, the complete set of parameter space samples $\boldsymbol{\Theta}_i$, discarded points and live points, can be turned into a

[5]Where nothing sinister is meant by $\sqrt{i}$.

set of posterior samples by assigning each sample a posterior weight p_i in accordance with Bayes' theorem in Eq. (5.8):

$$p_i = \frac{\mathcal{L}_i w_i}{\mathcal{Z}} \tag{5.45}$$

These samples can then be used for further analysis, as outlined in Section 5.1.

The most challenging part of the original nested sampling algorithm is how to efficiently sample the prior distribution under the additional hard likelihood constraint $\mathcal{L}(\boldsymbol{\Theta}_{\text{new}}) > \mathcal{L}(\boldsymbol{\Theta}_{\text{discard}})$. If new samples are drawn from the entire prior distribution the acceptance rate will decrease steadily as the likelihood constraint grows stronger. Mukherjee et al. improved on this by constraining new samples to be drawn from an ellipsoid containing the set of live points [35]. To tackle multimodal distributions, Shaw et al. introduced clustering algorithms to assign one ellipsoid to each cluster of live points [36]. Feroz and Hobson further improved on this approach and released an implementation in the form of the `Fortran` package `MultiNest` [33, 37].

Chapter 6

Parameter scans in high-energy physics

In this chapter we will focus on how parameter estimation analyses are performed in the context of high-energy physics phenomenology. In addition to parameter estimation we will frequently make use of the terms *parameter scan*, highlighting the exploration of the parameter space, and *global fit*, emphasising the comparison of model predictions to a wide range of experimental observables.

We start by outlining a typical scan setup in Section 6.1, before considering some of the limitations of current global fits in Section 6.2. In Section 6.3 we present the Global And Modular Beyond the Standard Model Inference Tool (`GAMBIT`), an ongoing project that aims to overcome these limitations. We focus in particular on one aspect of the `GAMBIT` project, namely how to make the best possible use of existing physics tools. As part of this effort, we have developed a new tool called `BOSS`, which we present in Section 6.4.

6.1 A vanilla scan setup

The first component of any parameter scan is the algorithm responsible for exploring the parameter space. As discussed in Section 5.3, for simple models with only a few free parameters this can be accomplished by choosing points to lie on a grid or simply pick points at random, *i.e.* according to

a flat probability distribution. More sophisticated algorithms are needed to efficiently explore higher-dimensional parameter spaces, with the nested sampling algorithm presented in Section 5.3.1 being one alternative.

Once a specific parameter point has been chosen our model is fully specified. Our next task is to determine the value of the likelihood function at the given point. As long as the observables considered, $\mathcal{O}_i$, are statistically independent, the total likelihood can be expressed as a factorized composite likelihood,

$$\mathcal{L}(\boldsymbol{\Theta}) = \prod_i \mathcal{L}_{\mathcal{O}_i}(\boldsymbol{\Theta}). \tag{6.1}$$

The p.d.f., and consequently also the likelihood, for an observable depends not only on the underlying theory, but also on the assumptions and uncertainties that go into the experimental measurement. In some cases information about the likelihood function is published when a new measurement is presented, see *e.g.* Ref. [38], but often we are forced to assume a likelihood function based on the nature of the observable and measurement. When all we know about a measured quantity is a central value with error estimates, we usually assume a gaussian likelihood, with a width determined by combining experimental and theoretical uncertainties. For event count observables a Poisson likelihood is typically assumed, while upper or lower limits are often treated using, possibly smeared, step functions.

The factorisation of the total likelihood in Eq. (6.1) makes it easy to arrange the calculations of the different $\mathcal{L}_{\mathcal{O}_i}$ factors in order of increasing computational expense. This way, if one of the easily computable predictions is in huge disagreement with observations, we can simply assign a very small likelihood value to the parameter point — effectively discarding it — and not waste time on the more expensive calculations. A related issue is that of unphysical points in the parameter space, for instance supersymmetric models with no electroweak symmetry breaking, for which further predictions are unreliable or simply not calculable. Such points are usually discarded right away, thus representing an additional factor

$$\mathcal{L}_{\text{physical}}(\boldsymbol{\Theta}) = \begin{cases} 1 & \text{if } \boldsymbol{\Theta} \text{ is physical} \\ 0 & \text{if } \boldsymbol{\Theta} \text{ is unphysical,} \end{cases} \tag{6.2}$$

in the total likelihood, or a similar requirement on the Bayesian prior.

To illustrate a typical, but far from exhaustive, chain of likelihood calculations we consider the study of a supersymmetric model. First the mass spectrum must be calculated, as it forms the basis for most other model predictions. Common tools for this purpose are `SOFTSUSY` [39], `IsaJet` [40], `SuSpect` [41] and `SPheno` [42, 43], often supplemented by a tool like `FeynHiggs` [44–48] for more detailed calculations of the Higgs sector. Next follows observables that are relatively cheap to compute, such as decay rates (`SUSY-HIT` [49]), flavour physics observables (`SuperIso` [50, 51]) and various electroweak precision observables (`FeynHiggs`, `MicrOMEGAS` [52]). The computational expense related to the relic density of dark matter (`DarkSUSY` [53], `MicrOMEGAS`) can vary greatly depending on the number of possible coannihilation processes considered. If included, computationally heavy observables such as NLO cross sections (`Prospino`[1] [58]) and simulated collider searches (`Pythia` [59, 60], `HERWIG++` [61]) are done last.

6.2 Limitations of current global fits

In order to carry out a complete global fit, the scanning algorithm and all the observable calculations must be combined in a joint framework. There exists a few public tools for accomplishing this, most notably `Fittino` [62] and `SuperBayes` [63,64]. In addition, many groups develop their own fitting tools which are not released publicly, with `MasterCode` [65] being one important example. However, all the above mentioned tools are subject to one or more of the following limitations: First, they are typically designed to study only a small subset of New Physics models, with constrained SUSY scenarios being the most common. Second, relating to the previous section, the fitting codes are often strongly linked to a chosen set of theory tools for calculating the model predictions. This makes it difficult for a user to replace or add tools in the likelihood calculation chain. Finally, there is usually a limited selection of available scanning algorithms, with variations of nested sampling or Markov

[1] When only coloured sparticle production is considered, the tool `NLL-fast` [54–57] provides a fast alternative combining pretabulated results from `Prospino` with NLL resummation of soft gluon emission in interpolation tables.

Chain Monte Carlo being the most common. So while the above mentioned scanning packages are efficient in performing the analyses they were designed for, their applicability beyond this is limited by a lack of modularity and flexibility.

For the work presented in this thesis we developed a private `Python`-based scanning code where every external tool is associated with a separate `Python` module. This module is responsible for running the external code as an independent executable and manage all input and output. With this setup it is straightforward to add, remove or rearrange tools in the likelihood calculation chain. However, this modularity comes at the cost of reduced code efficiency and flexibility: Most communication of data must proceed via the reading and writing of files, which is much slower than handling data in memory. And since the possibility of interaction with the external tools is limited, one often finds that new executables must be written when the requirements of the analysis change.

A situation where most analyses rely on privately developed scanning tools may in the long term have some negative consequences for the field as a whole. Much valuable research time can be wasted on "reinventing the wheel", when possibly better solutions already exist. This may lead to the adoption of unnecessarily coarse approximations in the likelihood calculation chain, either to compensate for a scanning tool of suboptimal efficiency, or simply because little time is left to improve the physics calculations. Another important cause for concern is that reproducibility is likely to suffer when the tools used can differ substantially between analyses.

6.3 GAMBIT

In 2012 the Global and Modular Beyond the Standard Model Inference Tool (`GAMBIT`)[2] collaboration was founded. With `GAMBIT` we aim to create an open-source fitting tool that improves significantly on the above mentioned limitations. The collaboration currently consists of 26 members from both the experimental and theoretical side of particle and astroparticle physics, and the first results are scheduled for release in the second half of 2015.

[2] `http://gambit.hepforge.org/`

One of the main features of GAMBIT is a very general treatment of theoretical models. This allows for performing global fits in any New Physics model, not just variations of SUSY. Models can either be defined from scratch, or as a subset of an already implemented model, similar to how the CMSSM is related to the MSSM.

The design of GAMBIT is based on an overarching idea of modularity at all levels. Scanner algorithms, likelihood calculations and external tools should all be replaceable without having to make fundamental changes to GAMBIT itself. At the top level, GAMBIT consists of a set of independent modules called Bits, that all connect and communicate via a central Core module. There is a ScannerBit responsible for running the sampling algorithm, while physics modules such as ColliderBit, DarkBit, FlavBit and HiggsBit take care of calculating physics observables. All modules are written in C++ and the user is free to add, modify or remove modules as desired.

To ensure the desired code modularity, the modules are never directly coupled together, but communicate by presenting the rest of GAMBIT with a set of required inputs, called *requirements*, and a list of possible outputs, called *capabilities*. For instance, DarkBit may need to know the mass of the Higgs, m_H, in order to perform its calculations. It therefore lists m_H as a requirement. By checking the list of capabilities, the Core ensures that m_H is provided by one of the other modules, and sets up the necessary communication. Thus, each individual module remains indifferent to exactly how its requirements are fulfilled, as this is the responsibility of the Core.

6.3.1 Interfacing multiple physics tools

There exists a multitude of specialized tools for computing physics observables in particular models of New Physics, and new tools are continuously being developed. Any global fit package must therefore find a way to interface to these tools, in order to make use of them in the likelihood calculations. In Section 6.2 we described two different solutions to this problem: The first was to choose a limited set of external tools and link them directly with the main scanning code. This approach allowed for efficient communication of data but lacked modularity. In the second approach, all tools were called as

```
Backends              Version    Path to lib                                  Status          #funcs  #types  #ctors
-------------------------------------------------------------------------------------------------------------------
DDCalc0               0.0        Backends/lib/libDDCalc0.so                   OK                 44       0       0
DarkSUSY              5.1.1      Backends/lib/libdarksusy.so                  OK                 44       0       0
FastSim               1.0        Backends/lib/libfastsim.so                   OK                  1       0       0
FeynHiggs             2.10       no path in config/backend_locations.yaml     absent/broken      14       0       0
HiggsBounds           4.1        Backends/lib/libhiggsbounds.so               OK                 10       0       0
HiggsSignals          1.2        Backends/lib/libhiggssignals.so              OK                 11       0       0
LibFarrayTest         1.0        Backends/lib/libFarrayTest.so                OK                  9       0       0
LibFirst              1.0        Backends/lib/libfirst.so                     OK                  8       0       0
                      1.1        Backends/lib/libfirst.so                     OK                 12       0       0
LibFortran            1.0        Backends/lib/libfortran.so                   OK                  6       0       0
MicrOmegas            3.5.5      no path in config/backend_locations.yaml     absent/broken      14       0       0
Pythia                8.186      Backends/lib/libpythia8.so                   absent/broken       0      27     105
                      8.209      Backends/lib/libpythia8.so                   OK                  0      27     103
SUSYPOPE              0.2        no path in config/backend_locations.yaml     absent/broken       3       0       0
SUSY_HIT              1.5        Backends/lib/libsusyhit.so                   OK                 55       0       0
SuperIso              3.4        no path in config/backend_locations.yaml     absent/broken      31       0       0
gamLike               1.0.0      no path in config/backend_locations.yaml     absent/broken       5       0       0
nulike                1.0.0      ../extras/nulike/lib/libnulike.so            OK                  4       0       0
-------------------------------------------------------------------------------------------------------------------
```

Figure 6.1: Screen capture of the backend system in `GAMBIT`.

independent programs. While providing much more modularity, this solution suffered in terms of efficiency.

The approach taken in `GAMBIT` share many similarities with the use of *plug-ins* in modern software. External physics tools, in `GAMBIT` referred to as *backends*, are provided as *shared libraries*[3] that `GAMBIT` load, use and unload at run-time. Known as *dynamic loading*, this allows for efficient communication of data via memory, while at the same time ensuring that `GAMBIT` itself can function without any particular backend.

Similar to how the physics modules list what input they require from other modules, they can also request information from external tools by specifying a set of *backend requirements*. It is then up to the `Core` to check that a backend able to provide this information is connected to `GAMBIT`. As an example, `DarkBit` may list the relic density of dark matter as a backend requirement, which can then be fulfilled by connecting external tools like `DarkSUSY` or `MicrOMEGAS` to `GAMBIT`'s backend system. Figure 6.1 shows a screen capture from `GAMBIT` listing information on what backends the user currently has connected.

[3]On Linux systems the file extension *.so* is often used for shared libraries, while on OS X the *.dylib* extension is common.

6.3.2 Dynamic loading of backends

A shared library appears to other programs as a collection of connection points called *symbols*. These typically refer to functions or variables in the library that are accessible from the outside. When GAMBIT loads the library and specifies a symbol, it obtains a pointer to the underlying function or variable, which can then be used on more or less equal footing with any internal part of GAMBIT.

On UNIX-based systems, dynamic loading is managed via the dl library. This library provides a basic set of functions for working with shared libraries: dlopen, which loads the shared library into memory; dlsym, which obtains the pointer for a given symbol; and dlclose, which unloads the shared library from memory. The pointer provided to GAMBIT by the dlsym function is of unspecified type, or, in C terminology, a pointer to type void, meaning that GAMBIT must cast the pointer to the correct type before using it.

The main piece of information a user must provide to GAMBIT in order to successfully connect a backend is therefore a map between symbols and the corresponding types. This is accomplished through the use of a pair of macros, BE_FUNCTION and BE_VARIABLE, defined in GAMBIT. Here we illustrate the use of BE_FUNCTION with an example: Consider a shared library written in C, containing a function matrixElement with the following signature:

```
double matrixElement(int, int);
```

If the library is built using a C compiler, the library symbol referring to this function will be identical to the function name. If a C++ compiler is used, the symbol will typically be a mangled version of the function name, where the exact structure of the symbol depends on the compiler. Using the GNU C++ compiler we get the symbol "_Z13matrixElementii". The necessary information is then provided to GAMBIT via the BE_FUNCTION macro as follows:

```
BE_FUNCTION(matrixElement, double, (int, int),
            "_Z13matrixElementii", "matrixElement")
```

The last argument, "matrixElement", is what a GAMBIT module must list as its backend requirement in order to be connected to this function. Similarly, the BE_VARIABLE macro can be used to connect to a variable in a shared library.

6.3.3 Reverse engineered plug-ins: dynamic loading of classes

Typically, physics libraries written in C or some version of the FORTRAN language present the user with an interface that is fundamentally just a collection of functions and variables of standard types. The functionality provided by the dl library, itself written in C, is therefore sufficient to make full use of these libraries. However, for libraries written in C++ the situation can become much more complicated.

One of the defining differences between C and C++ is the introduction of classes in C++. Most C++ libraries define a number of classes unique to that particular library. The intended workflow is usually that a user first creates variables, or *instances*, of these classes, and then performs operations on, and extracts information from, these instances. [4]

This poses a problem when a C++ code is to be used via dynamic loading, as class definitions are not accessible through the shared library system of symbols. Thus, even if GAMBIT loads the library into memory using dlopen, it will in general remain useless as all the library classes are unknown to GAMBIT. The naive solution to this would be to include all the relevant classes as part of the GAMBIT source code. However, this would simply reintroduce the problem we are trying to solve, as GAMBIT would start depending on the library, and possibly its dependencies again, just in order to build properly. Further, this approach could easily lead to name clash problems when multiple external libraries are connected, since GAMBIT would have no way of controlling the use of class names and namespaces. So, if we want the user to be able to connect a large number of external physics tools without increasing the complexity of building GAMBIT, this naive solution is clearly not viable.

A lot of modern C++ software allow for the use of plug-ins to extend the functionality of the main application. Although the plug-ins make use of C++ classes, they are dynamically loaded by the main application at run-time. How is this achieved? The secret lies in the C++ concept of *polymorphism*. A base class is used to define a class interface by declaring a

[4] Concepts similar to C++ classes have existed in the FORTRAN language since FORTRAN90, but the user interface of FORTRAN-based high-energy physics tools are rarely based on these language features. We therefore focus on the C++ case.

set of *virtual* member functions. These are functions for which the signature is defined, but where the actual implementation is expected to be overridden in classes that inherit from the bass class. Based on this "interface class", a whole range of new *derived* classes can be constructed. Each of these classes provide a unique implementation for the set of virtual member functions. The result is thus a whole family of classes that all adhere to a common interface defined by the base class.

To illustrate the idea, consider a base class called `Polygon` containing a virtual member function `calculateArea`. From this base class we may define two derived classes, `Triangle` and `Square`. Both classes should contain a `calculateArea` member function, but their implementations of this function would differ.

In plug-in systems the main application provides the base class, while plug-ins are supposed to provide the specialized derived classes. Since the class interface is settled, the main application can be built on the assumption that any future class passed in from a plug-in will have the predefined set of member functions.[5] The actual implementations of these member functions will be contained in the plug-in shared library, which can be dynamically loaded at run-time. For the main application to be able to request class instances from the plug-in, so-called *factory functions* are defined in the plug-in. These are simply functions that return a pointer to a newly created class instance.

The fundamental difference between a plug-in system as described above and the problem we face in `GAMBIT` is this: For a normal plug-in system, the main application is developed first and all plug-ins are developed later, adhering to the predefined interface. With `GAMBIT` it is the other way around: There already exists a number of `C++` physics tools, and we would like turn them into plug-ins, or backends, for `GAMBIT`. Thus, given a `C++` library with a set of classes defined, we must reverse-engineer the surrounding parts of a plug-in system, such as base classes and factory functions. For this purpose we have developed a tool called `BOSS`, which we introduce next.

[5]Technically, in the main application any instance of an unknown derived class is interpreted as an instance of the known base class.

6.4 BOSS: a Backend-On-a-Stick Script.

The Backend-On-a-Stick Script (BOSS) is a Python-based tool for accomplishing the following task: Given the C++ source code of some library, make the required modifications and additions to the library in order to allow its classes to be dynamically loaded from GAMBIT. Starting from a list of classes that the user would like to load, the main steps of BOSS can be summarized as follows:

1. Parse and analyse the original source code, and ensure that the requested classes can be loaded successfully.

2. Make the necessary modifications to the original source code.

3. Generate additional source code, including factory functions for all loaded classes.

4. Determine what the library symbols for the newly created factory functions will be, and summarize this in a file for GAMBIT.

5. Move all generated files to their correct locations inside the original source tree and within GAMBIT.

After BOSS has finished, the user builds a shared library from the modified source code and connects it to GAMBIT in the usual way.

Due to the complexity of the C++ language, parsing of C++ source code is highly non-trivial. For this task, BOSS employs the open-source tool GCC-XML [66], which produces a representation of the source tree in the form of an XML file. GCC-XML itself is based on the open-source GCC compiler. The XML file produced by GCC-XML forms the basis on which BOSS can further analyse the library source code.

To explain the process in a bit more detail, we consider the simplest possible example, namely that of a backend library containing a single class X, shown in Fig. 6.2. As described in the preceding section, a requirement for dynamically loading a library class is that it derives from a base class defining the class interface in terms of virtual member functions. This base class must be common to both the library and the main application. One of the first

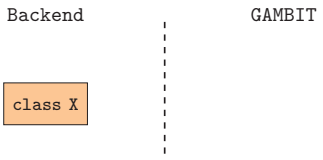

Figure 6.2: The starting point is a backend library containing a class not known to GAMBIT.

tasks of BOSS is therefore to determine what the member functions of class X are, and construct a base class with a corresponding set of virtual member functions. We name this base class Abstract_X, since C++ classes containing virtual functions are generally known as *abstract classes*. BOSS makes several additions to the source code for the original class X. However, the single most important change is to insert the newly generated Abstract_X class in the inheritance list for X, such that Abstract_X takes the role of a base class for X.

The next step for BOSS is to generate source code for factory functions to be included in the backend library. One such function is generated for each constructor in class X. These factory functions construct a new X instance and return its pointer. On the GAMBIT side, this pointer can be interpreted as a pointer to an Abstract_X instance. When this step is done, we have established the basic requirements for imitating a plug-in system, illustrated in Fig. 6.3.

With the basic plug-in structure in place, GAMBIT is able to retrieve a pointer to an X instance and call its member functions. However, there are still a number of problems and inconveniences with this system, most of which are related to one of two limitations: First, we are forced to work with a pointer to a class instance instead of directly with the instance itself, and new instances must be created via the factory functions. This marks a departure from how a user would typically work with the original library, and may cause problems when passing in function arguments or returning function results. Second, there is no way of accessing class member variables directly. This would be no problem for library codes that make use of "getter" and

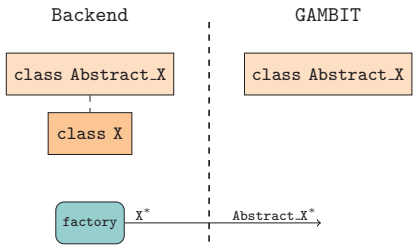

Figure 6.3: The basic requirements needed to set up a plug-in structure is an abstract base class common to both the backend and `GAMBIT`, and a factory function in the backend library for returning pointers to class instances.

"setter" member functions for accessing variables, but far from all libraries follow this convention. When multiple library classes are to be loaded, further complications typically arise, especially relating to type conversions of function arguments and return types.

Ideally, a user that is familiar with the original library should be able to access its functionality in a familiar way from within `GAMBIT` , without having to understand the intricacies of plug-in systems or worry about such problems as outlined above. With `BOSS` we therefore try to construct a solution that will reproduce the user interface of the original library, making the transition from using the original library directly to using it as a `GAMBIT` backend as smooth as possible. To accomplish this, we introduce what we call *wrapper classes*.

A wrapper class is basically a container for the pointer to the original class instance, constructed to provide a more familiar interface to the user. In our simple example, the wrapper class corresponding to class `X` is called `Wrapper_X`, and holds a pointer of type `Abstract_X`. Although we use the name `Wrapper_X` here to distinguish it from the original class `X`, in `GAMBIT` the wrapper class is given the name of the original class. Each constructor in `Wrapper_X` is connected to a factory function that provides the `Abstract_X` pointer, see Fig. 6.4. Thus, instead of calling the factory function directly and keep working with a pointer, the user can create an instance of `Wrapper_X` by calling constructors, similar to how the original library would be used.

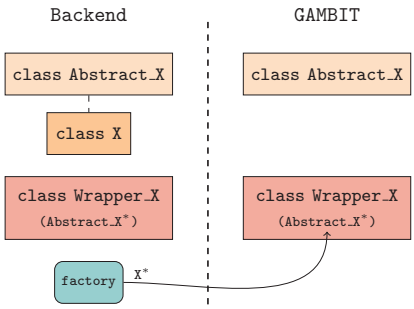

Figure 6.4: To present the user with a more familiar interface, `BOSS` introduces a wrapper class for each original class. A wrapper class holds a pointer to an instance of the original class, used to call member functions from and perform operations on.

Further, `Wrapper_X` contains member functions that imitate the function signatures of the original member functions, which are ultimately called using the contained `Abstract_X` pointer.

Through a system of helper functions inserted by `BOSS` into the original class, we can populate the wrapper class with references to the class member variables. When the type of the variable is a standard `C++` type, known in both the library and `GAMBIT`, this is all we need in order to reproduce the original interface of how the variables are accessed. However, if class `X` contains a member variable that is an instance of another library class, say class `Y`, we can reproduce the original class interface by letting `Wrapper_X` have an instance of `Wrapper_Y` as a member variable. The `Wrapper_Y` instance would then hold a pointer pointing to the original `Y` member variable.

With the above infrastructure in place, class instances can be constructed and used in the usual way from within `GAMBIT`. Existing code that makes use of the original library will therefore require little modification before it can be used in `GAMBIT`. However, it is important to remember that the wrapper class a `GAMBIT` user is provided with is not the original class itself, but rather an "outer shell" used to reconstruct the original user interface. If a user wants to extend or subclass a library class, this should be done in the original library

itself, before running it through BOSS and connecting it to GAMBIT.

In addition to requesting a set of classes to load, the user can also provide BOSS with a list of global library functions that GAMBIT should connect to. If such a function makes use of a library class as return type or argument type, BOSS will insert into the library an accompanying *wrapper function*, which rather uses the corresponding wrapper class. This ensures that the original library interface can be preserved also when global library functions are used from GAMBIT. The BE_FUNCTION macro calls required by GAMBIT to connect to these functions are automatically generated by BOSS.

BOSS is currently capable of providing dynamic loading of most "vanilla" C++ classes. Still, there are several scenarios that the current version cannot deal with properly. In such cases BOSS will generate a wrapper class where the problematic elements of the original class are left out, providing the user with a restricted class interface. Some of the current main limitations of BOSS are related to the use of template classes and functions, arrays containing instances of a library class, and function pointers. However, these problems can sometimes be worked around by small additions to the original library code, such as introducing "getter" and "setter" functions instead of working directly with an array. Some of the current limitations of BOSS are due to limitations in GCC-XML. As GCC-XML was recently succeeded by a new tool CastXML [67], future versions of BOSS will likely be based on CastXML, with the expectation that this will help extend the functionality of BOSS.

BOSS is currently developed in connection to the GAMBIT backend system, and the first public version of BOSS will be released as part of the first GAMBIT release. But the ability to turn existing C++ libraries into plug-ins may prove useful also for other projects where modularity and efficiency are important factors. In the future BOSS may therefore be considered for release as a public tool independent of GAMBIT.

Chapter 7

Summary of thesis results

Supersymmetric models with a small mass difference between the lightest neutralino and chargino can give rise to some interesting signatures at the LHC. In Paper 1 this scenario is explored within a Natural SUSY model with R-parity conservation, while Paper 2 studies the possible consequences of having a small chargino–neutralino mass splitting in models where R-parity is violated. Both analyses employ a Bayesian parameter scan, using the `MultiNest` algorithm introduced in Section 5.3.1. In Paper 1 the resulting posterior distribution constitutes the main result, while the posterior distribution obtained from the scan in Paper 2 more serves as a basis for further phenomenological study.

In Papers 3 and 4 we shift our attention to a couple of interesting, albeit fairly small, excesses reported in CMS and ATLAS searches for dileptons, jets and missing energy. The two papers follow similar outlines: First, we define a simple SUSY model capable of explaining the excesses in question. Next, a parameter scan is performed in order to map out the model's predictions across parameter space, both for the search where the excess is observed and for other relevant SUSY searches. As the number of free parameters is kept at a minimum, the models studied in these two papers can be explored using grid scans. Finally, the scan result is used in a frequentist assessment of the viability of the proposed model.

7.1 Choice of statistical approach

It is worth commenting on the differences between the two first and the two latter papers in terms of the statistical methods used. In the first two papers, the fundamental question asked in the statistical analysis is that of parameter estimation, or rather, phenomenology prediction: Assuming that the overall theory is true, what are the probable phenomenological scenarios? This question fits naturally within a Bayesian approach.

Furthermore, these two papers deal with high-dimensional parameter spaces, meaning that efficiency in the exploration of the parameter space is key. Fundamentally, Bayesian parameter estimation depends on the integral of the posterior p.d.f. over some region of parameter space. A reasonable answer can therefore be obtained even though the exact best-fit parameter point may be missed by the scan. In contrast, the frequentist profile likelihood approach to parameter estimation is based on comparing the likelihood value at each point to the likelihood value of the global best-fit point. This dependency on accurately determining the best-fit point renders the approach more computationally challenging when dealing with high-dimensional parameter spaces. A comparison of the Bayesian and frequentist approach in the context of parameter scans with `MultiNest` can be found in Ref. [68].

The parameter estimation analyses of the first two papers both start from the assumption that the underlying theory is correct. Thus, such analyses can only tell us what parameter regions the current data prefer *relative* to the rest of the model's parameter space. But independent of other parameter regions, should a given parameter point be excluded in light of the data? This is the question asked by the statistical analyses in Papers 3 and 4, where we reinterpret a set of collider searches within two supersymmetry scenarios. Here we perform a frequentist hypothesis test for each point in the parameter space of the model. One practical reason for preferring the frequentist approach in this case is that it simplifies the comparison of our results to those presented in the original experimental analyses.

In Paper 3 the hypothesis test simply consists of comparing the predicted signal yields for a set of collider searches to the corresponding observed upper limits. A parameter point is regarded excluded if it predicts a signal yield

greater than the observed limit for at least one of the searches. This is a common approach in phenomenology analyses investigating the level of tension between a proposed model and a set of experimental searches. It is also an intuitive approach, as it nicely illustrates how different experimental analyses exclude different regions of the model parameter space. However, at the level of individual parameter points, the exact statistical interpretation of this procedure is not entirely clear. For instance, if a parameter point fails one out of several independent tests, all performed at the 95% C.L., we cannot conclude that a single test combining all the data would exclude the point at the 95% C.L. In Paper 4 we use the same approach to establish a tension between the proposed model and the observations, but then extend the analysis by combining all the data and derive well-defined exclusion limits in the parameter space of the model.

7.2 LHC signals from light charginos

In the MSSM, the most important dimensionful parameters of the electroweak gaugino sector are the bino mass, M_1; the wino mass, M_2; and the higgsino mass parameter, μ. As seen in Section 3.6, if M_2 or μ is the smaller of the three parameters, a small mass difference Δm between the lightest neutralino and the lightest chargino is expected. The exact size of this mass splitting is an important factor for determining what type of collider signals to expect. If R-parity is conserved and the neutralino is the LSP, a small chargino–neutralino mass splitting may lead to a relatively long lifetime for the chargino. Thus, the chargino can potentially show up as a metastable charged particle, producing displaced vertices or kinked tracks in the detector. In Paper 1 we ask the question of whether this is an expected signal in models of Natural SUSY.

If, on the other hand, R-parity is not conserved, the chargino can potentially decay to final states with three Standard Model particles. When the chargino is either lighter than, or only slightly heavier than, the lightest neutralino, these decays can have substantial branching fractions, making them interesting as potential collider signatures. This is the possibility we explore in Paper 2.

The aim of Paper 1 is to perform Bayesian parameter estimation on a Natural SUSY subset of the MSSM parameter space, focusing on the posterior distribution for the chargino–neutralino mass difference, and, following from that, the expected chargino lifetime. The analysis adopts a conventional interpretation of naturalness in terms of the Barbieri–Giudice sensitivity measure discussed in Section 4.3.

The complete list of free parameters, and their prior distributions, is given in Table 1 of Paper 1. Most notably, $|\mu|$ and the soft breaking mass parameters for the third generation squarks are restricted to lie below 1 TeV. Parameters with less impact on the naturalness of the model are allowed wider prior ranges. Log priors are used for all dimensionful SUSY parameters. In Paper 1 this is argued based on the connection to the Barbieri–Giudice sensitivity measure demonstrated in Section 5.2. However, as seen in Section 5.1.2, a more general argument for using log priors for dimensionful parameters is that it represents ignorance as to what is the preferred scale. The Standard Model parameters are assigned gaussian priors, since here we have prior information in the form of measurements. Sleptons and first and second generation squarks are decoupled by fixing the corresponding mass parameters at 3 TeV.

The likelihood function for the scan is constructed based on a combination of experimental measurements and limits on sparticle masses. The measured observables include the W boson mass, M_W; the Higgs boson mass, m_h; the two branching ratios $BR(B_s \to \mu\mu)$ and $BR(b_s \to s\gamma)$; and the branching ratio for $B \to \tau\nu$ normalized to the Standard Model prediction, $R(B \to \tau\nu)$. The MSSM parameter space contains the possibility that the chargino is lighter than the neutralino, as demonstrated by Kribs et al. [69]. To allow the scan to explore this scenario, the dark matter relic density is not included in the likelihood function. The existence of dark matter can be explained by assuming that the gravitino is the actual LSP, with the lightest of the chargino and neutralino being the effective LSP on detector timescales.

A set of 95% C.L. sparticle mass limits is applied as hard cuts. Care is taken not to include experimental limits that are based on assumptions incompatible with our choice of free parameters, such as GUT-inspired relations for the gaugino mass parameters. Due to the computational expense

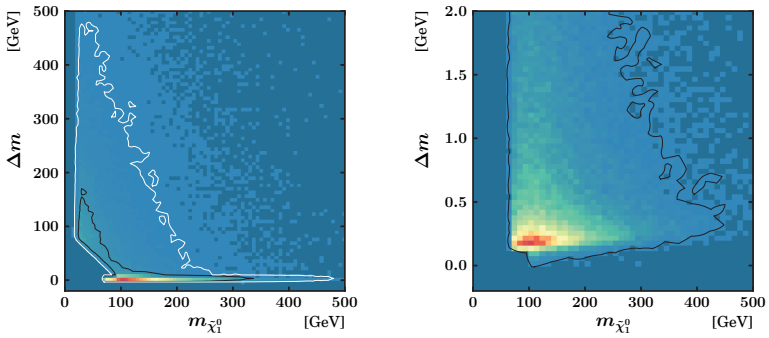

Figure 7.1: The posterior distribution in the plane of the chargino–neutralino mass difference, Δm, and the mass of the lightest neutralino, $m_{\tilde{\chi}_1^0}$. The black and white contours show 68% and 95% Bayesian credible regions, respectively. The right panel depicts the same posterior distribution, but focused on the region of small Δm values.

of LHC simulations and the size of the parameter scan, we do not include LHC sparticle searches in the scan likelihood. As argued in the paper, this is not expected to significantly impact our conclusions concerning the region of small chargino–neutralino mass differences. Nevertheless, having to leave out LHC simulations from the scan is suboptimal, and it is exactly this sort of limitation that `GAMBIT`, and in particular the `ColliderBit` module, aims to improve on. The complete list of constraints included in the scan likelihood is given in Tables 2 and 3 of Paper 1. In general, the chosen set of constraints should be regarded as conservative.

The possibility of having a chargino lighter than the neutralino is shown to be disfavoured by the 95% credible region, as seen in Fig. 7.1 depicting the joint posterior distribution for the chargino–neutralino mass difference and the mass of the lightest neutralino. This result can be traced back to an interplay between the Higgs mass constraint, preferring larger values for $\tan\beta$, and the LEP and Tevatron sparticle mass limits.

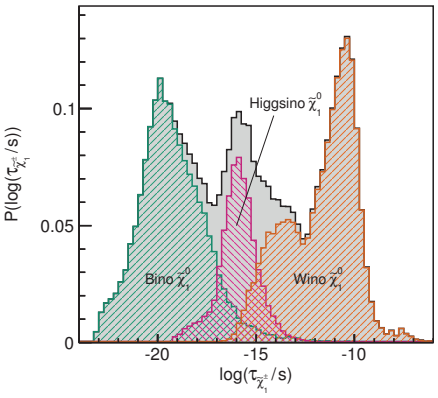

Figure 7.2: The posterior distribution of the chargino lifetime. Also shown are the subsamples with a lightest neutralino that is dominantly bino (green), higgsino (magenta) and wino (brown).

The strongest chargino–neutralino mass degeneracy occurs for scenarios where the lightest chargino and neutralino are predominantly wino, that is when $M_2 < |\mu|, |M_1|$. In this case, the dominant contribution to Δm can come from loop processes involving gauge bosons [70, 71]. The typical size of this contribution is $\alpha_2 M_W/4\pi \sim 200$ MeV, explaining the peak in the posterior distribution around these Δm values. The importance of this effect is enhanced by the Higgs mass constraint, as the tree-level contribution to Δm decreases with increasing $\tan\beta$ in the wino limit.

The fact that the wino-limit loop contribution to Δm typically is slightly larger than the pion mass has an important impact on the lifetime of the chargino, as it keeps the decay channel $\tilde{\chi}_1^\pm \to \tilde{\chi}_1^0 \pi^\pm$ open. As a consequence, there is only a small posterior probability for chargino lifetimes longer than $\sim 10^{-10}$ s. This can be seen in Fig. 7.2, where we also indicate the subset of posterior samples where the lightest neutralino is dominantly bino, higgsino or wino. These three neutralino scenarios give rise to three modes in the

posterior distribution for the chargino lifetime. In the higgsino limit, one expects mass differences larger than a few GeV, leading to expected lifetimes shorter than $\sim 10^{-15}$ s. Finally, when the lightest neutralino is dominantly bino, a chargino–neutralino degeneracy is not generally expected, leading to even shorter chargino lifetimes.

The preference for lifetimes shorter than $\sim 10^{-10}$ s means that charginos that are long-lived on detector timescales is not to be expected. Still, lifetimes of this magnitude, which corresponds to typical decay lengths of the order of 1 cm, may give rise to a significant number of events with kinked tracks. In Fig. 7.3 (left) we show the posterior distribution in the plane of the chargino lifetime and mass, overlaid with the excluded region from an ATLAS search for kinked tracks [72]. The search excludes part of the parameter space where the mass degeneracy is most severe, which, as shown in Fig. 7.2, corresponds to a scenario with a wino-dominated neutralino. In this limit, the chargino–neutralino mass splitting decreases with increasing $|\mu|$. Thus, the most mass degenerate scenarios are expected to be associated with higher levels of fine-tuning. This is illustrated in the right panel of Fig. 7.3, where we show the joint posterior distribution of the chargino lifetime and the Barbieri–Giudice fine-tuning measure introduced in Section 4.3.

The results of Paper 1 show that, within a Natural SUSY subspace of the MSSM, the possibility of having a chargino lighter than the neutralino is disfavoured, and detector-stable charginos is not an expected signature. Also, given the model studied here, a signal in future searches for long-lived charged particles would point towards a wino-dominated chargino and neutralino.

Paper 2 explores the possible consequences of having a small chargino–neutralino mass splitting in a scenario where the following trilinear R-parity violating operators are included in the superpotential:[1]

$$W \ni \lambda_{ijk} L_i L_j \bar{E}_k + \lambda'_{ijk} L_i Q_j \bar{D}_k + \lambda''_{ijk} \bar{U}_i \bar{D}_j \bar{D}_k. \tag{7.1}$$

Considering gauge symmetries this gives a total number of 45 new couplings. In particular, we investigate possible LHC signals from charginos decaying

[1] We here adopt the superfield notation used in the paper. In Eq. (3.4) the same superpotential terms are given in the notation of regular scalar component fields.

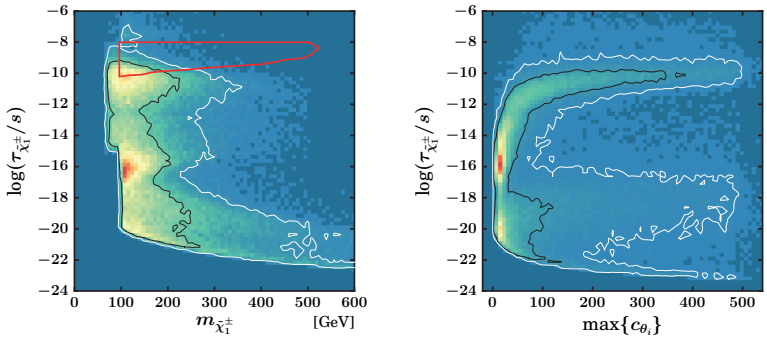

Figure 7.3: Posterior distribution of the chargino lifetime versus mass (left) and fine-tuning measure (right). The black and white contours show 68% and 95% Bayesian credible regions, respectively. The red contour is an overlaid limit from an ATLAS search for kinked tracks [72].

to three SM fermions in processes involving the above operators.

A Bayesian parameter scan is used to identify interesting parameter regions. The list of free parameters and their prior distributions is found in Table 1 of Paper 2. For the main scan, log priors are chosen for all dimensionful SUSY parameters. A smaller scan based on flat priors is used to investigate the prior dependence of the scan results.

Due to the large number of possible RPV couplings, they are not included as free parameters in the scan. Rather, we take the posterior samples resulting from the scan as a starting point, introduce a single dominant RPV operator, and then recalculate the chargino decay rate and branching ratios. For the resulting posterior distributions to be valid, we must assume that the predictions for the observables included in the scan likelihood are not strongly dependent on the given RPV operator.

The constraints going into the scan likelihood function are listed in Table 2 of Paper 2. As for Paper 1, no LHC simulations are included. Further, direct sparticle mass limits from collider searches that rely on large amounts

of missing energy are not generally applicable in a RPV scenario, as the amount of missing energy expected will depend on the size and nature of the dominant RPV coupling.

In order to concentrate the scan on parameter regions where the chargino–neutralino mass difference Δm is small, we include in the scan likelihood the hard requirement that $\Delta m < 1$ GeV. Since this constraint does not originate from an experimental measurement, it can equally well be interpreted as an extra condition on the prior parameter distributions, in particular for μ, M_1 and M_2. As seen in Paper 1, requiring the chargino–neutralino mass splitting to be less than 1 GeV amounts to a preference for the lightest neutralino and chargino to be dominantly wino or higgsino, with the wino scenario being the most probable. The arbitrariness in the choice of this Δm requirement may seem to complicate the interpretation of the posterior distributions resulting from the scan. However, the bound we impose on Δm is in principle no different from other assumptions imposed to reduce the full MSSM parameter space down to a more manageable size.

We study the chargino branching ratios after including a single dominant RPV operator. For each point in the posterior sample we take the value of the RPV coupling to be the upper bound as given in [73]. For instance, for the $L_1 L_2 \bar{E}_1$ operator, the upper bound on the coupling λ_{121} is $\lambda_{121} < 0.049 \times \frac{m_{\tilde{e}_R}}{100\ \mathrm{GeV}}$, coming from charged-current universality. The resulting joint posterior distributions for Δm and the most important chargino branching ratios are shown in Fig. 7.4. The top panels show the R-parity conserving decays to the lightest neutralino plus leptons (left) or hadrons (right). For the hadronic decay, the dominant decay mode is $\tilde{\chi}_1^{\pm} \to \tilde{\chi}_1^0 \pi^{\pm}$. The bottom panels depict the two most important RPV decay channels. The RPC modes dominate for Δm values down to the pion threshold, $\Delta m \sim m_\pi$. For even smaller values of Δm, direct RPV chargino decays become important, with similar branching ratios for the $e^+\mu^+e^-$ and $\nu_e\nu_\mu e^+$ final states. Similar results are found for the other $L_i L_j \bar{E}_k$ operators.

Lowering the chosen λ value will reduce the decay widths for the RPV processes, affecting the branching ratios in the region $\Delta m \in (0, m_\pi)$, where RPV and RPC processes compete. Also, as the RPV decay processes proceed via a sfermion propagator, the decay widths fall as $m_{\tilde{f}}^{-4}$ with increasing

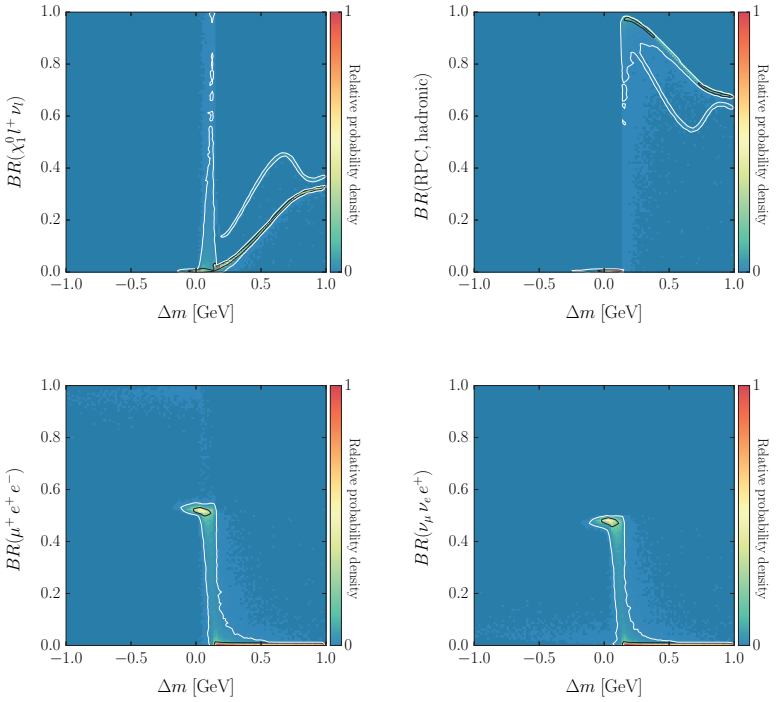

Figure 7.4: Posterior distribution of Δm versus branching ratio for the relevant chargino decay modes in a scenario with a dominant $L_1 L_2 \bar{E}_1$ RPV operator. The 68% and 95% C.R. contours are shown in black and white, respectively.

sfermion mass. On the other hand, the RPV decay widths depend on the chargino mass as $m^5_{\tilde{\chi}^\pm_1}$ while the RPC widths only depend on Δm^3, so the relative importance of the RPV processes will increase for higher chargino masses.

The results for the $L_iQ_j\bar{D}_k$ operators are similar: Below the pion threshold, the RPV decays to the $l_i^+ d_j \bar{d}_k$ and $\nu_i^+ u_j \bar{d}_k$ final states dominate, with approximately equal branching ratios for the two processes. For the $\bar{U}_i\bar{D}_j\bar{D}_k$ operators, however, we find that the RPV decay modes are typically greatly suppressed even for Δm values below the pion mass. This is due to the preference for heavy squarks following from the Higgs mass constraint, leading to a large propagator suppression in the RPV processes.

Considering the relatively weak set of constraints going into the scan likelihood, the resulting posterior distributions are expected to exhibit some prior dependence. Repeating the scan using flat priors for the dimensionful SUSY parameters, we find that the posterior probability for scenarios with large RPV branching ratios increases slightly. With a wider range of probable chargino and sfermion masses, the scaling of RPV widths with these masses become more important, with the net result that the RPV decays can dominate also into the region where the chargino–neutralino mass difference is larger than the pion mass.

Thus, assuming a scenario with a small chargino–neutralino mass splitting, there are regions of the preferred parameter space where the chargino decays dominantly through RPV processes. These decays can produce interesting collider signatures, in particular *lll* and *lqq* final states.

We also study the prospects for observing RPV chargino decays at the 13 TeV LHC. We devote the most attention to the $L_iL_j\bar{E}_k$ operators, as this is where one can expect the strongest deviations from signals previously studied in the context of neutralino RPV decays [74]. In particular, the $L_iL_j\bar{E}_k$ operators allow the chargino to decay via a sneutrino propagator to a final state of three charged leptons. Still, significant amounts of missing energy can be expected from neutrinos originating from the decay of the other chargino or neutralino in the event, and possibly from tau decays in scenarios where the dominant RPV coupling involves the third generation.

Chargino and neutralino pair production and subsequent decay is simu-

lated for selected benchmark parameter points, details are found in Table 3 of Paper 2. In order to properly illustrate the distribution shapes, 10^6 events are generated for each point. The total yield is normalized to an integrated luminosity of 1 fb^{-1}. Assuming a dominant RPV operator of the $L_iL_j\bar{E}_k$ type, events are required to contain at least three isolated leptons with $p_T > 70, 20, 20$ GeV and a missing transverse energy $E_T^{\text{miss}} > 100$ GeV.

Figure 7.5 shows the resulting trilepton invariant mass distributions for the operators $L_1L_2\bar{E}_1$, $L_1L_2\bar{E}_3$, $L_2L_3\bar{E}_2$ and $L_1L_3\bar{E}_3$ (top left to bottom right), using a benchmark parameter point with $m_{\tilde{\chi}_1^\pm} = 526$ GeV and $\Delta m = 0.18$ GeV. Table 4 of Paper 2 contains the values assumed for the various RPV couplings. The most important Standard Model backgrounds, shown in Fig. 7.5, are expected to come from diboson and $t\bar{t}$ production [75]. The NLO cross section for chargino and neutralino pair-production is 49.9 fb for this parameter point [76].

The $ee\mu$ triplet resulting from chargino decays through the $L_1L_2\bar{E}_1$ operator produces a clear resonance peak at the chargino mass.[2] For operators involving the third generation, the trilepton distributions is more smeared out due to the presence of either hadronically or leptonically decaying taus, but still these distributions show identifiable features such as end points and kinks. Given the low expected background and sizeable production cross section, the benchmark point studied here clearly represents a possibility for an early discovery at the 13 TeV run of the LHC. But also in a less optimistic scenario, with a smaller RPV coupling or a reduced cross section due to heavier charginos and neutralinos, the peak in the $ee\mu$ distribution and the features in the other distributions are promising tools for discovering these scenarios at the LHC. With enough statistics, a detailed comparison of kinematical distributions can help identify which RPV operator is dominating, and also whether the signal is mainly due to RPV decays of neutralinos, or a combination of chargino and neutralino RPV decays.

Since the charge of light-quark jets cannot easily be determined experimentally, the expected LHC signals from charginos decaying through $L_iQ_j\bar{D}_k$ operators are mostly similar to that of neutralinos decaying via the same op-

[2]For a dominant $L_1L_2\bar{E}_2$ operator, a similar peak would of course be found in the $e\mu\mu$ distribution.

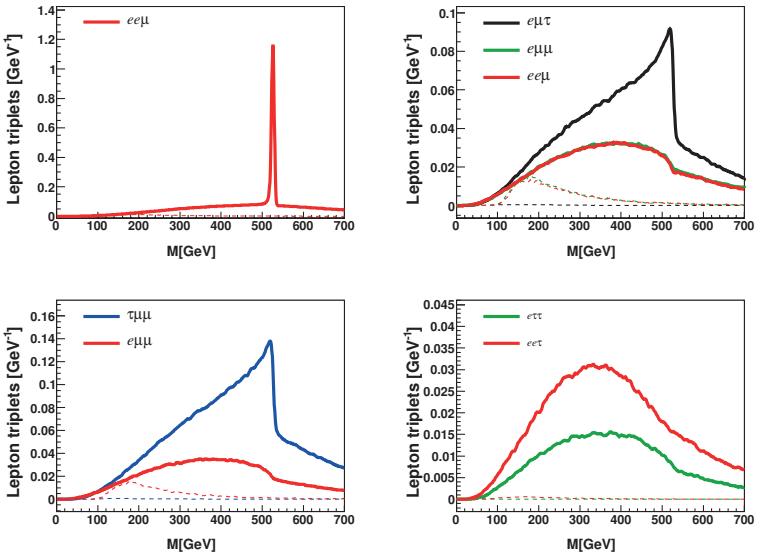

Figure 7.5: Trilepton invariant mass distributions obtained from processes involving the $L_1L_2\bar{E}_1$ (top left), $L_1L_2\bar{E}_3$ (top right), $L_2L_3\bar{E}_2$ (bottom left) and $L_1L_3\bar{E}_3$ (bottom right) operators. For each solid line, the thin dashed line of the same colour shows the dominant Standard Model background. A τ in the lepton triplet refers to a jet from a hadronically decay tau lepton. All distributions are normalized to an integrated luminosity of 1 fb^{-1}.

erator. One exception worth pointing out is that for a dominant $L_iQ_3\bar{D}_k$ operators, a neutralino lighter than the top quark will always decay to $\nu_i b\bar{d}_k$, while a chargino can decay to a $l_i b\bar{d}_k$ final state, where the combination of a charged lepton and a b-tag can improve detectability. As for $L_iQ_j\bar{D}_k$ operators, chargino decays via dominant $\bar{U}_i\bar{D}_j\bar{D}_k$ operators will produce the same signals as for neutralino decays, except possibly when third generation quarks are involved.

In sum, Paper 2 shows that direct decays of charginos through RPV operators can be important in scenarios with chargino–neutralino mass degeneracy, and that for a dominant operator of the $L_iL_j\bar{E}_k$ type the expected signals can differ greatly from that of neutralino decays. Also, a detailed study of the kinematics and flavour content of a signal can provide important information on the flavour structure of the underlying realisation of supersymmetry.

7.3 Intriguing excesses in dilepton searches

Based on the data collected during the first run of the LHC, the CMS and ATLAS experiments have both reported interesting excesses in searches for dileptons, jets and missing energy [77, 78]. While the CMS excess is consistent with a kinematic edge in the dilepton invariant mass spectrum around $m_{ll} \sim 79$ GeV, the excess seen by ATLAS is found on the Z-peak of the dilepton spectrum. An interpretation of the excess in one search as a possible New Physics signal may therefore be in tension with the lack of a corresponding excess in the other search, but as the details of the two analyses differ, the severity of this tension depends on the specific New Physics scenario. However, to asses the validity of a New Physics explanation of either dilepton excess, it is important to also take into account the results of other collider searches for which deviations from the Standard Model can be expected. In Papers 3 and 4 we carry out such analyses for two SUSY models that may explain the excesses in respectively the CMS and ATLAS dilepton searches.

The CMS dilepton search requires events to contain a pair of opposite-sign same flavour (OSSF) leptons (e or μ) with $p_T > 20$ GeV and pseudorapidity $|\eta| < 1.4$ ("central" region) or $1.6 < |\eta| < 2.4$ ("forward" region). At least two jets are required, where the jets are reconstructed using the anti-k_T

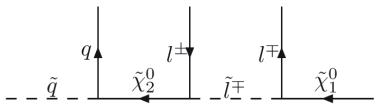

Figure 7.6: The golden cascade decay.

algorithm [79], with the jet radius parameter set to $R = 0.5$. For events with exactly two jets, a missing transverse energy $E_T^{\text{miss}} > 150$ GeV is required, while for events with three or more jets this requirement is lowered to $E_T^{\text{miss}} > 100$ GeV. The dominant Standard Model backgrounds are $t\bar{t}$ production and Drell-Yan production of γ^*/Z.

The resulting data is analysed in two different ways: First, a search for a kinematic edge in the dilepton invariant mass spectrum is performed by fitting the data to a model combining a background shape with a right-triangular signal shape. Such a triangular signal in the dilepton mass spectrum is a classic prediction of SUSY models where the two leptons originate from subsequent two-body steps in a sparticle decay chain, with the intermediate sparticle being the corresponding slepton. CMS finds that for events where both leptons fall within the central pseudorapidity region, the model that best fits the dilepton spectrum contains a contribution with a significance of 2.4σ, and with the kinematic edge located at $m_{ll} = 78.7 \pm 1.4$ GeV. Second, a simpler counting experiment is performed for the three dilepton mass regions $20 < m_{ll} < 70$ GeV, $81 < m_{ll} < 101$ GeV, and $m_{ll} > 120$ GeV. For the lower mass window, and still with both leptons in the central pseudorapidity region, an excess of 130^{+48}_{-49} events over the background expectation is observed, corresponding to a local significance of 2.6σ.

Paper 3 investigates whether this excess can be explained in a simple SUSY model with the production of a pair of first or second generation squarks, one of which decays via the much studied "golden cascade" depicted in Fig. 7.6. Additional jets can come from the decay of the second squark, for instance through $\tilde{q} \to \tilde{\chi}_1^0 q$, and from initial and final state radiation.

We investigate a model where the only free parameters are the soft-mass

parameters for the sparticles participating in the golden cascade scenario. The squark masses are governed by a common first- and second-generation mass parameter $m_{\tilde{q}}$, while for the two lightest neutralinos we take the bino mass M_1 and wino mass M_2 as free parameters. A common mass parameter $m_{\tilde{l}_R}$ is used for the right-handed sleptons of the first two generations, with the corresponding left-handed mass parameter $m_{\tilde{l}_L}$ set to $m_{\tilde{l}_L} = 2m_{\tilde{l}_R}$. The remaining sparticles are decoupled by setting their soft-mass parameters to 3500 GeV, except the gluino mass which is set to 1600 GeV. Finally, the trilinear couplings are set to zero and $\tan\beta$ is fixed at $\tan\beta = 10$. All dimensionful parameters are specified at the scale $\sqrt{m_{\tilde{t}_1} m_{\tilde{t}_2}}$.

This leaves a model with four free soft-mass parameters: $m_{\tilde{q}}$, M_1, M_2 and $m_{\tilde{l}_R}$. We constrain this parameter space using the two pieces of information provided by the CMS analysis, namely the position of the kinematic edge in the m_{ll} spectrum and the overall signal yield in the counting experiment. From energy-momentum conservation, the edge position $m_{ll}^{\max}$ giving the maximum invariant mass predicted by the golden cascade, can be expressed in terms of the slepton and neutralino masses,

$$m_{ll}^{\max} = \sqrt{\frac{(m^2_{\tilde{\chi}^0_2} - m^2_{\tilde{l}})(m^2_{\tilde{l}} - m^2_{\tilde{\chi}^0_1})}{m^2_{\tilde{l}}}}. \tag{7.2}$$

Thus, by choosing mass values for the slepton and one of the neutralinos, the other neutralino mass can be determined by requiring that the predicted $m_{ll}^{\max}$ equals the edge position inferred in the CMS fit. In terms of the input mass parameters, we accomplish this by scanning the plane of $m_{\tilde{l}_R}$ and $\Delta m = M_2 - m_{\tilde{l}_R}$, and for each point determine M_1 from the $m_{ll}^{\max}$ constraint using a modified version of `SOFTSUSY`.

With the decay kinematics in the last steps of the cascade fixed by the neutralino and slepton masses, the squark mass parameter $m_{\tilde{q}}$, which controls the production cross section, can be determined by requiring a given signal yield. This is done efficiently using an iterative procedure, relying on the fact that the event selection efficiency, determined by Monte Carlo simulation of the CMS search, varies much more slowly as a function of $m_{\tilde{q}}$ compared to the production cross section.

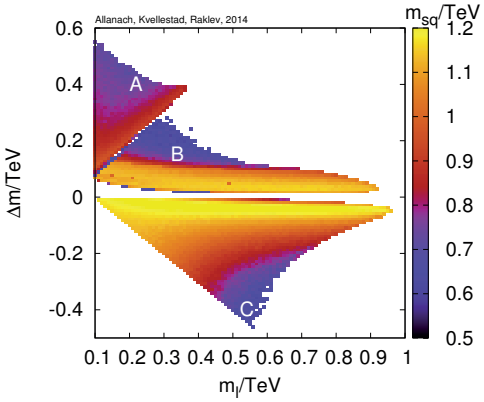

Figure 7.7: The regions of parameter space in the plane of $m_{\tilde{l}_R}$ and $\Delta m = M_2 - m_{\tilde{l}_R}$ that predict a signal yield equal to the 95% CL lower bound on the CMS dilepton signal. The corresponding physical squark mass $m_{\tilde{q}}$ is given by the colour map.

In sum, by scanning only the two parameters M_2 and Δm and employing the above mentioned techniques for determining M_1 and $m_{\tilde{q}}$, we can identify the regions in the full four dimensional input parameter space that provide a given number of expected signal events, and produce a m_{ll} edge at the observed value.

However, we find that the entire parameter space consistent at the 95% CL with the golden cascade interpretation is in conflict with other searches. To illustrate this, for each point in the M_2, Δm plane, we choose $m_{\tilde{q}}$ to be so large that the expected signal yield corresponds to the 95% CL lower bound on the observed signal. Lower values of $m_{\tilde{q}}$ will only lead to a higher production cross section, and thus be even more in tension with the null results of other searches. The region of parameter space identified by this scan is shown in Fig. 7.7. For $\Delta m > 0$, that is when $M_2 > m_{\tilde{l}_R}$, the next-to-lightest neutralino is dominantly wino. Thus, the decay to left-handed sleptons is preferred as long as it is kinematically allowed. This is the case in region A of Fig. 7.7. In region B this decay channel is no longer available,

so the $\tilde{\chi}_2^0$ decays through a right-handed slepton. The slope of the line separating regions A and B is due to our choice of taking $m_{\tilde{l}_L} = 2m_{\tilde{l}_R}$. In region C, where $\Delta m < 0$, it is the lightest neutralino which is wino-dominated. Squark masses as high as 1200 GeV are allowed by the scan. For higher squark masses the predicted signal yield is less than the observed 95% CL lower bound across the entire plane of M_2 and Δm.

The set of parameter points plotted in Fig. 7.7 are passed through our simulations of other relevant collider searches: an ATLAS search for jets and E_T^{miss} [80], the ATLAS dilepton search mentioned at the beginning of this section, a CMS search for multilepton events [81], and an ATLAS search for stop pair production with two leptons in the final state [82]. We find that all parameter points consistent with the golden cascade interpretation predict a signal yield above the observed upper bound for at least one of the other searches, as illustrated in Fig. 7.8. The strong sensitivity of the ATLAS stop search to the golden cascade scenario was first pointed out in [83]. In our scan, all but seven points are excluded by this constraint alone.

In summary, while the golden cascade scenario is in itself a straightforward explanation of the CMS dilepton excess, in its simplest realization it is in tension with several other supersymmetry searches at the LHC. By repeating the above analysis under varying assumptions, *e.g.* increasing the gluino mass or changing the ratio between $m_{\tilde{l}_L}$ and $m_{\tilde{l}_R}$, we find that this conclusion is insensitive to the specific parameter relations chosen.

The CMS dilepton excess has also been studied in a MSSM scenario with cascade decays initiated by bottom squark production [84]. However, also this scenario is in tension with other LHC sparticle searches, as shown in [83]. Some possible non-MSSM interpretations of the CMS excess can be found in Refs. [85–87].

The other dilepton excess is seen in an ATLAS search requiring events to contain two leading OSSF leptons with $p_T > 25, 10$ GeV, at least two jets with $p_T > 35$ GeV and $|\eta| < 2.5$, a missing energy of $E_T^{\text{miss}} > 225$ GeV, and a total transverse energy $H_T > 600$ GeV. The further requirement $81 < m_{ll} < 101$ GeV ensures that the invariant mass of the lepton pair is constrained to the region around the Z-peak in the dilepton spectrum. Jets are reconstructed using the anti-k_T algorithm with $R = 0.4$. Using this

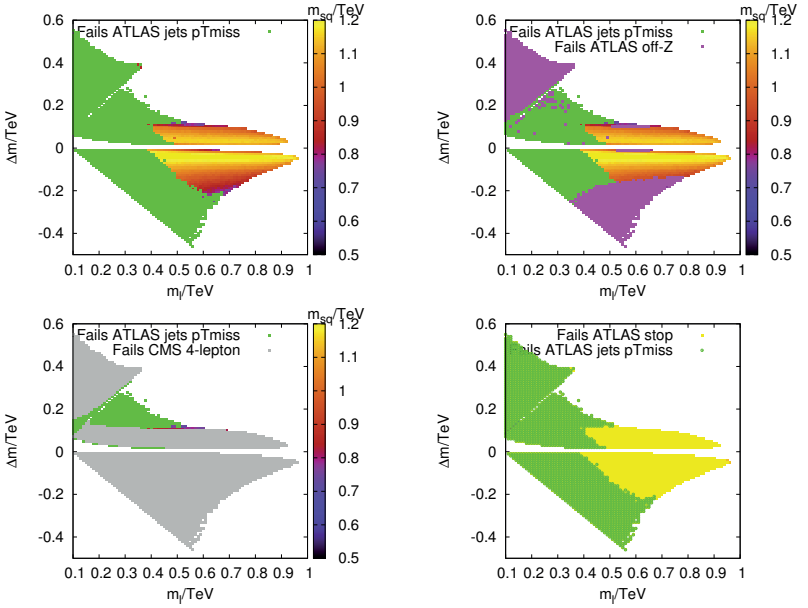

Figure 7.8: The regions of parameter space excluded by the searches "ATLAS jets pTmiss" (green) [80], "ATLAS off-Z" (magenta) [78], "CMS 4-lepton" (gray) [81], and "ATLAS stop" (yellow) [82].

set of selection criteria, a total of 29 events are observed. With an expected background from Standard Model processes of 10.6 ± 3.2 events, the result amounts to an excess of 3.0σ local significance.

ATLAS interprets this result in a general gauge mediation (GGM) model where a light gravitino is the LSP providing the missing energy [88]. In this scenario, the observed excess can potentially be explained by a decay chain starting from a gluino decay producing jets and a higgsino $\tilde{\chi}_1^0$ NLSP. With the neutralino subsequently decaying to Z plus gravitino, the required signature can be obtained for events where the Z decays leptonically (e^+e^- or $\mu^+\mu^-$). In Paper 4 we investigate whether the ATLAS interpretation is still viable when other SUSY searches are taken into account.

The only free parameters of the model are the gluino mass parameter M_3, the higgsino mass parameter μ, and $\tan\beta$. The gravitino is effectively taken to be massless to ensure prompt neutralino decays. The two remaining gaugino mass parameters are set to $M_1 = M_2 = 1.5$ TeV and all sfermions are decoupled by setting $m_{\tilde{f}} = 4.5$ TeV. All dimensionfull parameters are defined at the scale $\sqrt{m_{\tilde{t}_1} m_{\tilde{t}_2}} \sim 4.5$ TeV.

The branching ratio for $\tilde{\chi}_1^0 \to Z\tilde{G}$ is close to 100% for small values of $\tan\beta$, and then decreases with increasing $\tan\beta$ due to stronger competition from the processes $\tilde{\chi}_1^0 \to h\tilde{G}$ and $\tilde{\chi}_1^0 \to \gamma\tilde{G}$. Thus, if the predicted signal rate is to remain unchanged, an increase in $\tan\beta$ must be compensated by an increased production cross section through a lower gluino mass. We investigate $\tan\beta = 1.5$ and $\tan\beta = 30$.

As long as the necessary jets are produced, the exact decay chain from the gluino down to the higgsino $\tilde{\chi}_1^0$ is not important for the ATLAS dilepton search.[3] However, this part of the decay chain plays a crucial role in determining what predictions the model makes for other SUSY searches. In particular, a large production of third-generation quarks can be expected due to the higgsino nature of the $\tilde{\chi}_1^0$ and the slightly heavier $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$. The most important gluino branching ratios are shown in Fig. 7.9 as functions of the gluino–neutralino mass difference Δm for $\tan\beta = 1.5$ (solid lines) and $\tan\beta = 30$ (dashed lines), with the gluino mass fixed at $m_{\tilde{g}} = 900$ GeV. A significant production of top quarks is expected for mass differences down to $\Delta m \sim 350$ GeV. As this will lead to additional leptons from leptonic top decays, searches for multilepton final states are relevant for constraining the parameter space of the model. However, given the small branching ratios for W and Z decaying to leptons, for a large fraction of events there will be no isolated leptons with significant p_T. Thus, searches for zero-lepton final states can also constrain the model. In particular, this is important in the region $\Delta m < 350$ GeV, where the main source of leptons with significant p_T is the Z-boson from the decay $\tilde{\chi}_1^0 \to Z\tilde{G}$.[4]

We perform two grid scans of the (μ, M_3)-plane, one for $\tan\beta = 1.5$ and

[3] Electroweak production alone is not feasible.

[4] Additional leptons can come from off-shell W's and Z's produced in the decays of the $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^\pm$ down to $\tilde{\chi}_1^0$, but these leptons will typically be soft due to the $\mathcal{O}(1)$ GeV mass differences among the higgsino states.

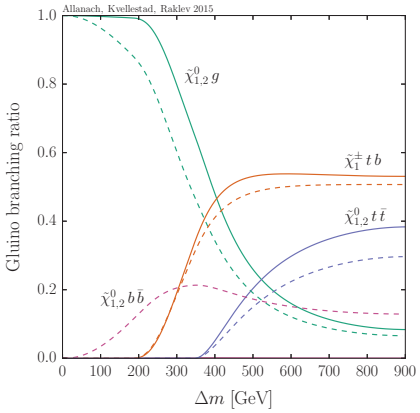

Figure 7.9: Branching ratios for the most important gluino decays as functions of $\Delta m = m_{\tilde{g}} - m_{\tilde{\chi}^0_1}$ for $\tan\beta = 1.5$ (solid lines) and $\tan\beta = 30$ (dashed lines). The gluino mass is set to $m_{\tilde{g}} = 900$ GeV.

one for $\tan\beta = 30$. At each parameter point we simulate the same set of SUSY searches as was used in Paper 3, except that for the ATLAS and CMS dilepton searches and the CMS multilepton search, the relevant signal regions are now those sensitive to the leading dilepton pair coming from a decaying Z. The region in the $(m_{\tilde{\chi}^0_1}, m_{\tilde{g}})$-plane preferred by the ATLAS dilepton excess at the 95% CL is determined using a Poisson likelihood profiled over a Gaussian background uncertainty. Similarly, 95% CL exclusion limits are determined for the other searches. The results are shown in Fig. 7.10 for $\tan\beta = 1.5$ (left) and $\tan\beta = 30$ (right). The region preferred by the ATLAS dilepton search is the bright band marked `ATLAS_onZ`. The CMS search for multilepton final states, in the figure referred to as `CMS_multilepton`, is mainly sensitive to the regions of large gluino–neutralino mass differences, while the ATLAS search for zero-lepton events, `ATLAS_jMET`, excludes regions where this mass difference is smaller. The overall interpretation of the result is similar to

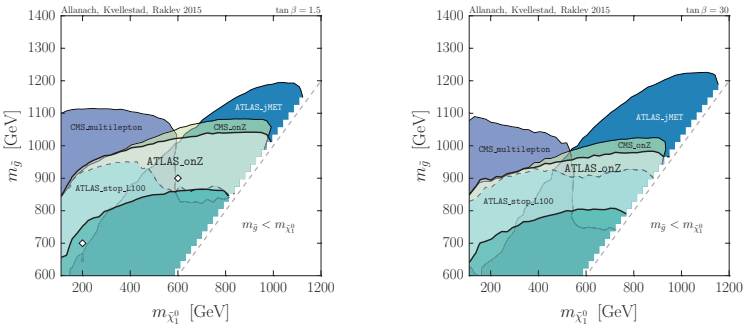

Figure 7.10: The region in the $(m_{\tilde{\chi}_1^0}, m_{\tilde{g}})$-plane preferred by the ATLAS dilepton search at the 95% CL (lighter band, marked `ATLAS_onZ`), compared to the 95% CL exclusion regions from the other searches. Results are given for $\tan\beta = 1.5$ (left) and $\tan\beta = 30$ (right). The white diamond markers in the left-hand plot correspond to two benchmark points used in the ATLAS dilepton analysis [78].

that of the result in Paper 3: The entire parameter space consistent at the 95% CL with a signal interpretation of the excess is in tension with at least one of the other searches.

A combined 95% CL exclusion limit in the $(m_{\tilde{\chi}_1^0}, m_{\tilde{g}})$-plane is derived from a joint likelihood function including all searches. The result is given in Fig. 7.11, where the white and black contours are the limits for $\tan\beta = 1.5$ and $\tan\beta = 30$, respectively. The colour map depicts the predicted signal yield for the ATLAS dilepton search for $\tan\beta = 1.5$. All regions predicting more than 6 expected signal events are excluded.

In conclusion, Paper 4 shows that a GGM model with light gluinos, higgsinos and gravitinos cannot provide a satisfactory explanation of the ATLAS dilepton excess. In particular, when the probability for final-state leptons is increased due to production of top quarks, the constraints from multilepton searches are strong. On the other hand, without such an enhanced probability for lepton production, the model is in conflict with searches for zero-lepton

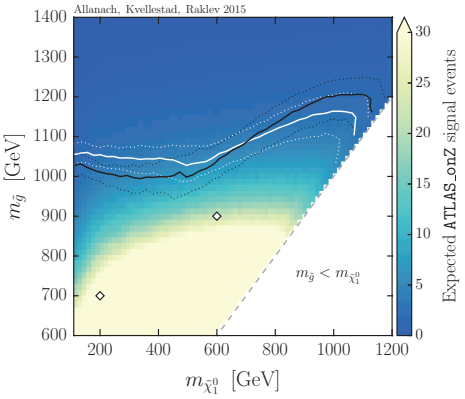

Figure 7.11: The 95% CL exclusion curves in the plane of $m_{\tilde{\chi}_1^0}$ and $m_{\tilde{g}}$ for $\tan\beta = 1.5$ (white) and $\tan\beta = 30$ (black). The region below each contour is excluded. For $\tan\beta = 1.5$, the predicted signal yield for the ATLAS dilepton analysis is given by reference to the colour bar on the right. The dotted contours show the exclusion limits obtained upon varying the NLO production cross section up or down by 20%, as an estimate for the systematic uncertainty.

final states.

The ATLAS dilepton excess has also been interpreted in several other models, *e.g.* see Refs. [89–92]. In particular, the authors of [91] demonstrate that the excess can be explained by 500–700 GeV squarks decaying through a ~ 350 GeV bino $\tilde{\chi}_3^0$ down to higgsino neutralinos around 150–200 GeV.

7.4 Summary

Motivated by naturalness arguments, attractive dark matter candidates and the prospect for high-scale unification of the strong and electroweak forces, supersymmetry has long been one of the most promising theories for phys-

ics beyond the Standard Model. The breaking of supersymmetry, as parametrized in the MSSM, introduces a vast range of possible realisations of supersymmetry at the TeV scale. In order to derive experimentally testable predictions for the LHC and other experiments, well-motivated theoretical assumptions must be combined with smart and efficient numerical exploration of the parameter space. The latter aspect has become even more important in light of the non-discovery of supersymmetry during the first run of the LHC, as theorists have been forced to consider increasingly general models.

In this thesis we emphasize the role statistical methods and their interpretation play in this research program. The high dimensionality of the MSSM parameter space necessitates the use of probabilistic scanning techniques. Thus, any conclusions we draw concerning what the likely and less likely physical scenarios are, will be conditional on the underlying statistical interpretation of the method. We may also need to pay attention to statistical foundations if we use the concept of naturalness as a theoretical guiding principle.

Paper 1 touches on several of the above points. Here we perform a Bayesian analysis of the parameter space in a Natural SUSY scenario. In this subset of the MSSM parameter space the mass difference between the lightest chargino and neutralino can be small, possibly leading to a detectable LHC signal in the form of long-lived charginos. However, our analysis finds that this scenario does not seem probable in light of existing experimental constraints.

In Paper 2 we move beyond the MSSM by allowing also R-parity violating interactions. This opens up a large number of possible processes by which sparticles can decay to final states with only Standard Model particles. Most studies of the collider phenomenology of R-parity violation have focused on the decay of the lightest neutralino. We show that in regions of parameter space where the chargino–neutralino mass difference is small, direct R-parity violating decays of the chargino may be dominant. This scenario may give rise to unusual detector signatures, such as a resonance of three charged leptons.

Although no evidence for supersymmetry, or indeed any physics beyond the Standard Model, have been found so far at the LHC, several smaller

excesses have been observed in the data. It is interesting to investigate whether such excesses point towards some specific New Physics scenario. In Papers 3 and 4 we study two recent excesses seen in CMS and ATLAS searches for dileptons, jets and missing energy. The CMS excess, which hints at the presence of a kinematical edge in the dilepton spectrum, is analysed in a scenario where the two signal leptons originate from subsequent steps in a supersymmetric cascade decay. While this can provide a fairly minimal explanation of the excess, the model is found to be ruled out by other LHC searches for supersymmetry.

The ATLAS search observes an excess of events on the Z-peak of the dilepton spectrum, and interprets this result in a supersymmetry scenario where additional Z bosons arise from the decay of a neutralino down to a light gravitino. In Paper 4 we show that this interpretation of the excess is excluded when the null-results of other searches are taken into account. However, other possible explanations for these two excesses have been suggested in the literature. With the second run of the LHC under way, we should soon know whether these excesses are due to interesting physics or simply random fluctuations in the data.

Finally, much of the work behind this thesis has been devoted to the collaborative development of `GAMBIT` and the related tool `BOSS`, both soon to be publicly released. With `GAMBIT` we hope to provide the high-energy physics community with a powerful open-source tool for extracting as much information as possible from the results of ongoing and future experiments, regardless of whether the data point towards supersymmetry or something completely different.

Bibliography

[1] S. P. Martin, "A Supersymmetry primer," `arXiv:hep-ph/9709356 [hep-ph]`. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[2] I. Aitchison, *Supersymmetry in Particle Physics: An Elementary Introduction*. Cambridge University Press, 2007.

[3] H. Müller-Kirsten and A. Wiedemann, *Introduction to Supersymmetry*. World Scientific lecture notes in physics. World Scientific, 2010.

[4] S. R. Coleman and J. Mandula, "All Possible Symmetries of the S Matrix," *Phys. Rev.* **159** (1967) 1251–1256.

[5] Yu. A. Golfand and E. P. Likhtman, "Extension of the Algebra of Poincare Group Generators and Violation of p Invariance," *JETP Lett.* **13** (1971) 323–326. [Pisma Zh. Eksp. Teor. Fiz.13,452(1971)].

[6] R. Haag, J. T. Lopuszanski, and M. Sohnius, "All Possible Generators of Supersymmetries of the s Matrix," *Nucl. Phys.* **B88** (1975) 257.

[7] H. Baer and X. Tata, *Weak Scale Supersymmetry: From Superfields to Scattering Events*. Cambridge University Press, 2006.

[8] **Super-Kamiokande** Collaboration, K. Abe *et al.*, "Search for proton decay via $p \to \nu K^+$ using 260 kiloton $\cdot$ year data of super-kamiokande," *Phys. Rev. D* **90** (Oct, 2014) 072005.

[9] R. D. Peccei and H. R. Quinn, "CP conservation in the presence of pseudoparticles," *Phys. Rev. Lett.* **38** (Jun, 1977) 1440–1443.

[10] **MEG** Collaboration, J. Adam *et al.*, "New constraint on the existence of the $\mu^+ \to e^+\gamma$ decay," *Phys. Rev. Lett.* **110** (2013) 201801, arXiv:1303.0754 [hep-ex].

[11] A. Borrelli, "Narratives of 'naturalness' in today's particle physics community," in *Narrated Communities – Narrated Realities: Narration as Cognitive Processing and Cultural Practice*, H. Blume, C. Leitgeb, and M. Rössner, eds. Brill, 2015.

[12] L. Susskind, "Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory," *Phys. Rev.* **D20** (1979) 2619–2625.

[13] G. 't Hooft, C. Itzykson, A. Jaffe, H. Lehmann, P. K. Mitter, I. M. Singer, and R. Stora, "Recent Developments in Gauge Theories. Proceedings, Nato Advanced Study Institute, Cargese, France, August 26 - September 8, 1979," *NATO Sci. Ser. B* **59** (1980) pp.1–438.

[14] G. F. Giudice, "Naturally Speaking: The Naturalness Criterion and Physics at the LHC," arXiv:0801.2562 [hep-ph].

[15] M. Papucci, J. T. Ruderman, and A. Weiler, "Natural SUSY Endures," *JHEP* **09** (2012) 035, arXiv:1110.6926 [hep-ph].

[16] B. de Carlos and J. A. Casas, "One loop analysis of the electroweak breaking in supersymmetric models and the fine tuning problem," *Phys. Lett.* **B309** (1993) 320–328, arXiv:hep-ph/9303291 [hep-ph].

[17] J. L. Feng, "Naturalness and the Status of Supersymmetry," *Ann. Rev. Nucl. Part. Sci.* **63** (2013) 351–382, arXiv:1302.6587 [hep-ph].

[18] J. A. Casas, J. M. Moreno, S. Robles, K. Rolbiecki, and B. Zaldívar, "What is a Natural SUSY scenario?," *JHEP* **06** (2015) 070, arXiv:1407.6966 [hep-ph].

[19] J. R. Ellis, K. Enqvist, D. V. Nanopoulos, and F. Zwirner, "Observables in Low-Energy Superstring Models," *Mod. Phys. Lett.* **A1** (1986) 57.

[20] R. Barbieri and G. F. Giudice, "Upper Bounds on Supersymmetric Particle Masses," *Nucl. Phys.* **B306** (1988) 63.

[21] G. W. Anderson and D. J. Castano, "Measures of fine tuning," *Phys. Lett.* **B347** (1995) 300–308, arXiv:hep-ph/9409419 [hep-ph].

[22] P. Athron and D. J. Miller, "A New Measure of Fine Tuning," *Phys. Rev.* **D76** (2007) 075010, arXiv:0705.2241 [hep-ph].

[23] D. Sivia and J. Skilling, *Data analysis: a Bayesian tutorial*. Oxford science publications. Oxford University Press, 2006.

[24] E. Jaynes and G. Bretthorst, *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.

[25] R. T. Cox, "Probability, Frequency and Reasonable Expectation," *American Journal of Physics* **14** (Jan., 1946) 1–13.

[26] R. E. Kass and A. E. Raftery, "Bayes factors," *Journal of the American Statistical Association* **90** no. 430, (1995) 773–795.

[27] J. Bernoulli, *Ars conjectandi.* Thurnisiorum, 1713.

[28] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.* **106** (May, 1957) 620–630.

[29] R. E. Kass and L. Wasserman, "The selection of prior distributions by formal rules," *Journal of the American Statistical Association* **91** no. 435, (1996) 1343–1370.

[30] S. Fichet, "Quantified naturalness from Bayesian statistics," *Phys. Rev.* **D86** (2012) 125029, arXiv:1204.4940 [hep-ph].

[31] M. E. Cabrera, J. A. Casas, and R. Ruiz de Austri, "Bayesian approach and Naturalness in MSSM analyses for the LHC," *JHEP* **03** (2009) 075, arXiv:0812.0536 [hep-ph].

[32] T. Burgess, J. Ø. Lindroos, A. Lipniacka, and H. Sandaker, "Finding viable Models in SUSY Parameter Spaces with Signal Specific Discovery Potential," *JHEP* **08** (2013) 098, arXiv:1210.7020 [hep-ph].

[33] F. Feroz, M. P. Hobson, and M. Bridges, "MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics," *Mon. Not. Roy. Astron. Soc.* **398** (2009) 1601–1614, arXiv:0809.3437 [astro-ph].

[34] J. Skilling, "Nested sampling," *Bayesian inference and maximum entropy methods in science and engineering* **735** (2004) 395–405.

[35] P. Mukherjee, D. Parkinson, and A. R. Liddle, "A nested sampling algorithm for cosmological model selection," *Astrophys. J.* **638** (2006) L51–L54, arXiv:astro-ph/0508461 [astro-ph].

[36] J. Shaw, M. Bridges, and M. Hobson, "Efficient bayesian inference for multimodal problems in cosmology," *Monthly Notices of the Royal Astronomical Society* **378** no. 4, (2007) 1365–1370.

[37] F. Feroz and M. Hobson, "Multimodal nested sampling: an efficient and robust alternative to markov chain monte carlo methods for astronomical data analyses," *Monthly Notices of the Royal Astronomical Society* **384** no. 2, (2008) 449–463.

[38] **CMS, LHCb** Collaboration, V. Khachatryan *et al.*, "Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data," *Nature* (2015) , arXiv:1411.4413 [hep-ex].

[39] B. Allanach, "SOFTSUSY: a program for calculating supersymmetric spectra," *Comput. Phys. Commun.* **143** (2002) 305–331, arXiv:hep-ph/0104145 [hep-ph].

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

[41] A. Djouadi, J.-L. Kneur, and G. Moultaka, "SuSpect: A Fortran code for the supersymmetric and Higgs particle spectrum in the MSSM," *Comput. Phys. Commun.* **176** (2007) 426–455, arXiv:hep-ph/0211331 [hep-ph].

[42] 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, arXiv:hep-ph/0301101 [hep-ph].

[43] 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, arXiv:1104.1573 [hep-ph].

[44] T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, and G. Weiglein, "High-Precision Predictions for the Light CP -Even Higgs Boson Mass of the Minimal Supersymmetric Standard Model," *Phys. Rev. Lett.* **112** no. 14, (2014) 141801, arXiv:1312.4937 [hep-ph].

[45] M. Frank, T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, *et al.*, "The Higgs Boson Masses and Mixings of the Complex MSSM in the Feynman-Diagrammatic Approach," *JHEP* **0702** (2007) 047, arXiv:hep-ph/0611326 [hep-ph].

[46] G. Degrassi, S. Heinemeyer, W. Hollik, P. Slavich, and G. Weiglein, "Towards high precision predictions for the MSSM Higgs sector," *Eur. Phys. J.* **C28** (2003) 133–143, arXiv:hep-ph/0212020 [hep-ph].

[47] S. Heinemeyer, W. Hollik, and G. Weiglein, "The Masses of the neutral CP - even Higgs bosons in the MSSM: Accurate analysis at the two loop level," *Eur. Phys. J.* **C9** (1999) 343–366, arXiv:hep-ph/9812472 [hep-ph].

[48] S. Heinemeyer, W. Hollik, and G. Weiglein, "FeynHiggs: A Program for the calculation of the masses of the neutral CP even Higgs bosons in the MSSM," *Comput. Phys. Commun.* **124** (2000) 76–89, arXiv:hep-ph/9812320 [hep-ph].

[49] A. Djouadi, M. Muhlleitner, and M. Spira, "Decays of supersymmetric particles: The Program SUSY-HIT (SUspect-SdecaY-Hdecay-InTerface)," *Acta Phys. Polon.* **B38** (2007) 635–644, arXiv:hep-ph/0609292 [hep-ph].

[50] F. Mahmoudi, "SuperIso: A Program for calculating the isospin asymmetry of B $\to$ K* gamma in the MSSM," *Comput. Phys. Commun.* **178** (2008) 745–754, arXiv:0710.2067 [hep-ph].

[51] F. Mahmoudi, "SuperIso v2.3: A Program for calculating flavor physics observables in Supersymmetry," *Comput. Phys. Commun.* **180** (2009) 1579–1613, arXiv:0808.3144 [hep-ph].

[52] G. Bélanger, F. Boudjema, A. Pukhov, and A. Semenov, "micrOMEGAs4.1: two dark matter candidates," *Comput. Phys. Commun.* **192** (2015) 322–329, arXiv:1407.6129 [hep-ph].

[53] P. Gondolo, J. Edsjo, P. Ullio, L. Bergstrom, M. Schelke, *et al.*, "DarkSUSY: Computing supersymmetric dark matter properties numerically," *JCAP* **0407** (2004) 008, arXiv:astro-ph/0406204 [astro-ph].

[54] A. Kulesza and L. Motyka, "Threshold resummation for squark-antisquark and gluino-pair production at the LHC," *Phys. Rev. Lett.* **102** (2009) 111802, arXiv:0807.2405 [hep-ph].

[55] A. Kulesza and L. Motyka, "Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC," *Phys. Rev.* **D80** (2009) 095004, arXiv:0905.4749 [hep-ph].

[56] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, *et al.*, "Soft-gluon resummation for squark and gluino hadroproduction," *JHEP* **0912** (2009) 041, arXiv:0909.4418 [hep-ph].

[57] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, *et al.*, "Squark and Gluino Hadroproduction," *Int. J. Mod. Phys.* **A26** (2011) 2637–2664, arXiv:1105.1110 [hep-ph].

[58] W. Beenakker, R. Hopker, M. Spira, and P. Zerwas, "Squark and gluino production at hadron colliders," *Nucl. Phys.* **B492** (1997) 51–103, arXiv:hep-ph/9610490 [hep-ph].

[59] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP* **0605** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[60] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.* **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[61] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J.* **C58** (2008) 639–707, arXiv:0803.0883 [hep-ph].

[62] P. Bechtle, K. Desch, and P. Wienemann, "Fittino, a program for determining MSSM parameters from collider observables using an iterative method," *Comput. Phys. Commun.* **174** (2006) 47–70, arXiv:hep-ph/0412012 [hep-ph].

[63] R. R. de Austri, R. Trotta, and L. Roszkowski, "A Markov chain Monte Carlo analysis of the CMSSM," *JHEP* **0605** (2006) 002, arXiv:hep-ph/0602028 [hep-ph].

[64] R. Trotta, F. Feroz, M. P. Hobson, L. Roszkowski, and R. Ruiz de Austri, "The Impact of priors and observables on parameter inferences in the Constrained MSSM," *JHEP* **0812** (2008) 024, arXiv:0809.3792 [hep-ph].

[65] K. de Vries, E. Bagnaschi, O. Buchmueller, R. Cavanaugh, M. Citron, *et al.*, "The pMSSM10 after LHC Run 1," arXiv:1504.03260 [hep-ph].

[66] http://gccxml.github.io. [Online; accessed 23-June-2015].

[67] https://github.com/CastXML/CastXML. [Online; accessed 23-June-2015].

[68] F. Feroz, K. Cranmer, M. Hobson, R. Ruiz de Austri, and R. Trotta, "Challenges of Profile Likelihood Evaluation in Multi-Dimensional SUSY Scans," *JHEP* **06** (2011) 042, arXiv:1101.3296 [hep-ph].

[69] G. D. Kribs, A. Martin, and T. S. Roy, "Supersymmetry with a Chargino NLSP and Gravitino LSP," *JHEP* **01** (2009) 023, arXiv:0807.4936 [hep-ph].

[70] J. L. Feng, T. Moroi, L. Randall, M. Strassler, and S.-f. Su, "Discovering supersymmetry at the Tevatron in wino LSP scenarios," *Phys. Rev. Lett.* **83** (1999) 1731–1734, arXiv:hep-ph/9904250 [hep-ph].

[71] H.-C. Cheng, B. A. Dobrescu, and K. T. Matchev, "Generic and chiral extensions of the supersymmetric standard model," *Nucl. Phys.* **B543** (1999) 47–72, arXiv:hep-ph/9811316 [hep-ph].

[72] **ATLAS** Collaboration, G. Aad *et al.*, "Search for charginos nearly mass degenerate with the lightest neutralino based on a disappearing-track signature in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Phys. Rev.* **D88** no. 11, (2013) 112006, arXiv:1310.3675 [hep-ex].

[73] B. C. Allanach, A. Dedes, and H. K. Dreiner, "Bounds on R-parity violating couplings at the weak scale and at the GUT scale," *Phys. Rev.* **D60** (1999) 075014, arXiv:hep-ph/9906209 [hep-ph].

[74] N.-E. Bomark, D. Choudhury, S. Lola, and P. Osland, "Flavour Structure of R-violating Neutralino Decays at the LHC," *JHEP* **07** (2011) 070, arXiv:1105.4022 [hep-ph].

[75] "Search for Supersymmetry at the high luminosity LHC with the ATLAS experiment," Tech. Rep. ATL-PHYS-PUB-2014-010, CERN, Geneva, Jul, 2014. http://cds.cern.ch/record/1735031.

[76] W. Beenakker, M. Klasen, M. Kramer, T. Plehn, M. Spira, and P. M. Zerwas, "The Production of charginos / neutralinos and sleptons at

hadron colliders," *Phys. Rev. Lett.* **83** (1999) 3780–3783, arXiv:hep-ph/9906298 [hep-ph]. [Erratum: Phys. Rev. Lett.100,029901(2008)].

[77] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for physics beyond the standard model in events with two leptons, jets, and missing transverse momentum in pp collisions at sqrt(s) = 8 TeV," *JHEP* **04** (2015) 124, arXiv:1502.06031 [hep-ex].

[78] **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** no. 7, (2015) 318, arXiv:1503.03290 [hep-ex].

[79] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP* **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[80] **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, arXiv:1405.7875 [hep-ex].

[81] **CMS** Collaboration, S. Chatrchyan *et al.*, "Search for anomalous production of events with three or more leptons in *pp* collisions at $\sqrt{s} =$ 8 TeV," *Phys. Rev.* **D90** (2014) 032006, arXiv:1404.5801 [hep-ex].

[82] **ATLAS** Collaboration, G. Aad *et al.*, "Search for direct top-squark pair production in final states with two leptons in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *JHEP* **06** (2014) 124, arXiv:1403.4853 [hep-ex].

[83] P. Grothaus, S. P. Liew, and K. Sakurai, "A closer look at a hint of SUSY at the 8 TeV LHC," *JHEP* **05** (2015) 133, arXiv:1502.05712 [hep-ph].

[84] P. Huang and C. E. M. Wagner, "CMS kinematic edge from sbottoms," *Phys. Rev.* **D91** no. 1, (2015) 015014, arXiv:1410.4998 [hep-ph].

[85] B. Dutta, Y. Gao, T. Ghosh, T. Kamon, and N. Kolev, "Dilepton Mass Endpoint in the NMSSM," arXiv:1506.04336 [hep-ph].

[86] B. Allanach, A. Alves, F. S. Queiroz, K. Sinha, and A. Strumia, "Interpreting the CMS $\ell^+\ell^- jj\not{E}$_T Excess with a Leptoquark Model," arXiv:1501.03494 [hep-ph].

[87] M. Dhuria, C. Hati, R. Rangarajan, and U. Sarkar, "Explaining the CMS $eejj$ and e missing p_T jj excess and leptogenesis in superstring inspired E_6 models," *Phys. Rev.* **D91** no. 5, (2015) 055010, arXiv:1501.04815 [hep-ph].

[88] P. Meade, N. Seiberg, and D. Shih, "General Gauge Mediation," *Prog. Theor. Phys. Suppl.* **177** (2009) 143–158, arXiv:0801.3278 [hep-ph].

[89] A. Kobakhidze, A. Saavedra, L. Wu, and J. M. Yang, "ATLAS Z-peaked excess in MSSM with a light sbottom or stop," arXiv:1504.04390 [hep-ph].

[90] S. P. Liew, A. Mariotti, K. Mawatari, K. Sakurai, and M. Vereecken, "Z-peaked excess in goldstini scenarios," arXiv:1506.08803 [hep-ph].

[91] M. Cahill-Rowley, J. L. Hewett, A. Ismail, and T. G. Rizzo, "The ATLAS Z + MET Excess in the MSSM," arXiv:1506.05799 [hep-ph].

[92] N. Vignaroli, "Z-peaked excess from heavy gluon decays to vectorlike quarks," *Phys. Rev.* **D91** no. 11, (2015) 115009, arXiv:1504.01768 [hep-ph].