Faculteit Wetenschappen
Vakgroep Fysica en Sterrenkunde

Event selection optimization for (b-)jets plus lepton supersymmetry searches with the CMS detector

door

Nadja Strobbe

Promotor: Dr. Michael Tytgat
Begeleider: Lic. Lukas Vanelderen

Masterproef ingediend tot het behalen van de academische graad van
Master in de Fysica & Sterrenkunde

Academiejaar 2010-2011

Errors, like straws, upon the surface flow;
He who would search for pearls must dive below.

John Dryden

Acknowledgements

A year's hard work is presented in this thesis. This work would not have been possible without the help of many people. Therefore, some words of gratitude are in order.

First of all, I would like to thank Lukas for answering all my questions, for helping me with any programming problems, for our brainstorming sessions and for proof-reading this rather long document so thoroughly. I would also like to thank Sezen and Harrison for introducing me to the world of SUSY and all things Bayesian. I really enjoyed our conversations and the nice dinner (eggplants included).

I also wish to thank Michael and professor Ryckbosch for making it possible for me to stay at CERN for several weeks. It was really great to be at the place where it is all happening. You could feel the anticipation in the air. I also want to thank Michael for the feed-back regarding this work.

Another big thank you goes to all my fellow students, especially Karen en Pieter, for the fun we had over the past five years. They would not have been the same without you. The PhD students of the INW also deserve a thank you for the entertaining lunches at the UZ, the nice atmosphere and the occasional birthday cake.

And last, but not least, I would like to thank my mum, dad and brother for their support throughout these years. Although you all had the tendency to fall asleep, or even run away, whenever I wanted to talk about what exactly it was that I was doing, it helped to put things in perspective.

Contents

Preface

One of the *biggest* questions in life is "What are we made of?" Scientists translate this is as "What are the building blocks of matter?" In the nineteenth century, matter was believed to be made of indivisible atoms. With the experiments by E. Rutherford and J.J. Thomson at the end of that century, the realization came that these atoms actually consist of a positively charged nucleus surrounded by negatively charged electrons.

Curious as people are, experiments were set up to study these constituents and the field of particle physics came into existence. During these experiments, physicists accelerated and collided particles and studied what happened. In this manner, a whole bunch of new particles were discovered, the so-called *particle zoo*. At that time, these particles were believed to be elementary, just as people had thought that the atom was elementary.

Of course, the theoretical physicists had not been lying around doing nothing. They had made big progress in quantum mechanics and in the 1960's the first quantum field theories describing the interactions between particles were formulated. Gell-Mann and Zweig noticed that the strange symmetries present in this particle zoo, could be explained if one assumed that many of these particles were not elementary, but consisted of the so-called quarks. Some years later, this was also confirmed experimentally.

From the 1970's onwards, advances in technology allowed ever better accelerators to be built. These experiments, in conjunction with the theoretical advances, paved the way for the construction of the Standard Model of Particle Physics.

The Standard Model provides a description of all known particles and interactions except gravity. It has been very successful in predicting and explaining the results of the experiments done during the past forty years. But Nature would not be Nature if the story ended here. In spite of all the successes of the Standard Model, there are still a few unresolved issues that indicate the need for New Physics, beyond the Standard Model. One issue is related to the question "How do particles acquire mass?" Within the Standard Model, this is explained by the so-called Higgs mechanism. This mechanism predicts the existence of a new massive particle, the Higgs boson. This Higgs boson has, however, not been observed yet and there are also some theoretical issues regarding the mass of the Higgs boson.

Over time, many possible extensions of the Standard Model have been proposed. A very elegant and popular extension is *supersymmetry* or SUSY. This is also the subject of this thesis. Supersymmetry effectively results in a doubling of the particle content of the theory: for each known particle, there is a corresponding superpartner. The answer to the question of whether or not supersymmetry exists in nature, can only be answered by doing experiments. The start-up of the Large Hadron Collider, the latest and most powerful particle accelerator and collider, at CERN in Geneva has provided a boost for the research on supersymmetry. There is, however, one difficulty: supersymmetric models generally have a large number of free parameters. Depending on the values of these parameters, the experimental signatures can be quite different. The question that needs to be answered is thus "How does one search for signals of supersymmetry?" This thesis contributes to finding a solution to this question.

In this work, we present a new method – the sNavigator – to find an optimal event selection for a region in the parameter space of a given New Physics model. This optimal selection should maximize the chance to make a discovery. As it makes no sense to optimize the selection for a paramater space region that has already been excluded, we will take into account the results of previous measurements. These can be other SUSY searches, but also many other measurements contain information about the nature of supersymmetry. We will use a Bayesian approach to include this prior information and navigate our way through the SUSY space.

On a more practical level, the Random Grid Search principle will serve as a basis for our algorithm. The Random Grid Search provides an efficient way to optimize rectangular cuts. This principle will be incorporated in our method to optimize the event selection for a parameter space region. We will also need to generate and simulate events for a large number of points in the chosen parameter space. These points will be sampled via a Markov Chain Monte Carlo method and will be simulated using Delphes, a fast detector simulation tool.

The first part of this thesis consists of two chapters. In the first chapter an overview of the Standard Model will be given. We will then also list some of the unresolved issues. Chapter two provides an introduction to supersymmetry. We will introduce the Minimal SuperSymmetric Model and provide some insight in the experimental signatures of supersymmetry. Several low energy observables and their implications for the SUSY parameter space will also be discussed.

The second part contains everything regarding the analysis that was done. In chapter three we give a short overview of the Large Hadron Collider and the Compact Muon Solenoid (CMS) experiment. We will discuss the various subsystems of the CMS detector and the offline computing system.

In chapter four we come back to the goal of this thesis. The procedure we constructed to maximize the probability of making a discovery will be explained in detail. We will also shortly recall the basics of Bayesian inference.

The Monte Carlo samples and physics object definitions that were used, are listed in chapter five. In chapter six, we apply the Random Grid Search principle in the optimization of the event selection for a single point in a multidimensional parameter space. The performance of the Delphes fast detector simulation is discussed in chapter seven.

Chapter eight provides an extensive test of the sNavigator procedure. What we have learned from this test, is applied in chapter nine, where we perform a case study of the sNavigator procedure for three SUSY analyses involving leptons and b-jets. To conclude this thesis we will also discuss our results and provide an outlook.

Part I

Theory

Chapter 1

The Standard Model of particle physics

1.1 Overview

In this section an overview of the Standard Model of particle physics will be given. We will discuss the particle content of the theory and the interactions. Also, the Brout-Englert-Higgs mechanism will be applied to the Standard Model. This section is largely based on references [1] and [2].

1.1.1 Particles and interactions

The Standard Model of particle physics is a quantum field theory describing the strong and electroweak interaction. It contains three generations (or families) of fermionic matter, each consisting of an up- and down-type quark, a charged lepton and a neutrino (and their corresponding antiparticles). An overview of the Standard Model fermions with some of their properties is given in table 1.1.

The interactions between the particles are described by the gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$. For each generator of the group ($N^2 - 1$ for $SU(N)$, N^2 for $U(N)$), there is an associated gauge field. These are spin-1 particles called gauge bosons. Sometimes they are also called "force carriers", because they are associated with the interactions or forces.

The strong interaction, also known as Quantum Chromodynamics (QCD), is based on the gauge group $SU(3)_c$ and describes the color interaction between the quarks. There are three color charges: red (r), green (g) and blue (b). The eight gauge bosons of QCD are called gluons. Quarks transform as color triplets, while leptons are singlets under $SU(3)_c$. QCD has two important properties: confinement and asymptotic freedom.

Confinement essentially means that quarks cannot exist freely. They are always 'confined' in colorless (white) hadrons, i.e. bound states of quarks and/or antiquarks. A quark and antiquark of opposite color (e.g. r and $\bar{r}$) can form a *meson* and 3 quarks of different color (r, g and b) can form a *baryon*. When one tries to separate two quarks, for example in a high energy collision, the force binding them together becomes stronger and energy is built up between them. At some point it becomes energetically more favourable to use this energy to create extra quarks from the vacuum. This process is called hadronization and it is responsable for the creation of *jets*, the sprays of hadrons that are found at collider experiments. The coupling constant, which describes the strength of the interaction, is large in this regime. Therefore, one cannot use perturbation theory to describe confinement. So far, no analytical explanation of the process exists. Only lattice calculations have provided some insight.

Asymptotic freedom describes the behavior of quarks at small distance scales or equivalently at large momentum scales. When quarks are very close together, the force between them becomes very small. They can then be treated as free quarks and perturbation theory can be used.

The electroweak interaction, as formulated by Glashow, Salam and Weinberg, provides a unified picture of the electromagnetic and weak interactions. It is based on the gauge group $SU(2)_L \times U(1)_Y$.

Fermions	Generation 1	2	3	Spin	Electric charge Q	3rd component of weak isospin T_3	Hyper-charge Y	color
Quarks	$\binom{u}{d}_L$	$\binom{c}{s}_L$	$\binom{t}{b}_L$	$\frac{1}{2}$	$+\frac{2}{3}$ $-\frac{1}{3}$	$\frac{1}{2}$ $-\frac{1}{2}$	$\frac{1}{6}$	r,g,b
	u_R	c_r	t_R	$\frac{1}{2}$	$+\frac{2}{3}$	0	$\frac{2}{3}$	r,g,b
	d_R	s_R	b_R	$\frac{1}{2}$	$-\frac{1}{3}$	0	$-\frac{1}{3}$	r,g,b
Leptons	$\binom{\nu_e}{e}_L$	$\binom{\nu_\mu}{\mu}_L$	$\binom{\nu_\tau}{\tau}_L$	$\frac{1}{2}$	0 -1	$\frac{1}{2}$ $-\frac{1}{2}$	$-\frac{1}{2}$	-
	e_R	μ_R	τ_R	$\frac{1}{2}$	-1	0	-1	-

Table 1.1: Fermions in the Standard Model

In this notation the 'L' stands for left-handed, since the weak interaction only couples to left-handed particles, and 'Y' stands for hypercharge. The gauge bosons for the electroweak theory are the three W-bosons for $SU(2)_L$ and the B-boson for $U(1)_Y$. There exists a relation between the electric charge Q, the hypercharge Y and the third component of the weak isospin T_3. For the convention of the hypercharge as used in table 1.1, this relation takes the form

$$Q = T_3 + Y \ . \tag{1.1}$$

The electroweak theory is a chiral theory, which means that left- and right-handed particles transform differently under the symmetries of the theory. The left-handed quarks and leptons transform as weak isospin doublets, whereas the right-handed quarks and charged leptons transform as singlets under $SU(2)_L$. As a consequence of this chirality, there cannot be any explicit mass terms in the Lagrangian for the fermions. These terms would explicitly break the gauge symmetry. Gauge invariance also prohibits the existence of gauge boson masses in the Lagrangian.

The absence of particle masses is in clear contradiction with experimental observations. In the past decades, we have measured masses to very great precision. It is thus clear that particles must acquire their masses in some way. A very nice way is the Brout-Englert-Higgs mechanism, which explains that particles can get a mass through the process of spontaneous symmetry breaking. During this process a complex, scalar (spin-0) field, the so-called Higgs field, acquires a vacuum expectation value that spontaneously breaks the $SU(2)_L \times U(1)_Y$ symmetry. How it works is explained in more detail in the next section.

1.1.2 The Brout-Englert-Higgs mechanism

The Brout-Englert-Higgs mechanism starts by introducing one or more scalar fields, the Higgs fields. These fields acquire a vacuum expectation value that spontaneously breaks a symmetry in the Lagrangian. From the Goldstone theorem we know that for every spontaneously broken continuous symmetry there will be a massless scalar particle, the Goldstone boson. Hence, the number of Goldstone bosons in the theory is equal to the number of broken generators. In the case of a gauge theory, however, the story does not end here. The so-far massless gauge bosons can become massive by "eating" the Goldstone boson. So in a gauge theory, the number of massive gauge bosons will be equal to the number of broken generators. We will now illustrate this for the case of the Standard Model, more specifically for the electroweak theory.

Before electroweak symmetry breaking, all four gauge bosons (W^1, W^2, W^3, B) are massless. What we observe in experiment is one massless gauge boson, the photon A (sometimes also denoted γ) and three massive gauge bosons (W^+, W^-, Z). We also know that the electric charge is conserved.

The spontaneous symmetry breaking should be of the form

$$SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM} \ .$$

To give mass to three gauge bosons, three Goldstone bosons will have to be "eaten". This means that we will need at least three degrees of freedom for the scalar fields. The simplest way is by introducing a complex, scalar $SU(2)$ doublet with positive hypercharge ($Y = \frac{1}{2}$)

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \ . \tag{1.2}$$

The Standard Model Lagrangian without the strong part is given by

$$\mathcal{L}_{SM} = -\frac{1}{4} W^a_{\mu\nu} W^{\mu\nu}_a - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \overline{L}_i i D_\mu \gamma^\mu L_i + \overline{e}_{Ri} i D_\mu \gamma^\mu e_{Ri} \ , \tag{1.3}$$

where i runs over the three generations, μ, ν are Lorentz indices and a runs over the number of generators in the gauge group. The field strengths are given by:

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g_2 \epsilon^{abc} W^b_\mu W^c_\nu$$
$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

and the covariant derivative for left- and right-handed leptons by:

$$D_\mu L_L = \left(\partial_\mu - i g_2 T_a W^a_\mu - i g_1 Y_l B_\mu\right) L_L$$
$$D_\mu L_R = \left(\partial_\mu - i g_1 Y_l B_\mu\right) L_R \ ,$$

where T_a are the generators of the gauge group and g_1, g_2 are the coupling constants for the electroweak interaction.

Now that we have introduced the scalar doublet Φ, we need to add the scalar part to the Lagrangian

$$\mathcal{L}_S = (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi), \quad \text{with } V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \ . \tag{1.4}$$

The first term is the kinetic term and the second term is the scalar potential, which is often called the "Mexican Hat" potential.

In order for the vacuum to be stable, the parameter λ has to be positive . Depending on the sign of μ^2 one can distinguish two cases. In the case $\mu^2 > 0$, the potential $V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2$ is always positive and the minimum is found at $\langle 0|\Phi|0\rangle \equiv \Phi_0 = \binom{0}{0}$. So in this case no spontaneous symmetry breaking takes place. In case $\mu^2 < 0$, the minimum of the potential is no longer located at the origin. Therefore, the neutral component of the scalar field can acquire a vacuum expectation value (vev) v .

$$\langle 0|\Phi|0\rangle \equiv \Phi_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \ , \quad v = \sqrt{-\frac{\mu^2}{\lambda}} \ . \tag{1.5}$$

By only giving a vev to the neutral component, electromagnetism (with gauge group $U(1)_{EM}$) is conserved, just as we wanted.

We proceed by expanding Φ around its minimum Φ_0

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \ , \tag{1.6}$$

and inserting this in the kinetic part of the scalar Lagrangian (1.4). Redefining the gauge fields as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \tag{1.7}$$

$$Z_\mu = \frac{1}{\sqrt{g_1^2+g_2^2}}(g_2 W_\mu^3 - g_1 B_\mu) \tag{1.8}$$

$$A_\mu = \frac{1}{\sqrt{g_1^2+g_2^2}}(g_2 W_\mu^3 + g_1 B_\mu)\ , \tag{1.9}$$

we find for the kinetic part of the scalar Lagrangian

$$|D_\mu \Phi|^2 = \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{2}g_2^2(v+H)^2 W_\mu^+ W^{\mu-} + \frac{1}{8}(v+H)^2(g_1^2+g_2^2)Z_\mu Z^\mu\ . \tag{1.10}$$

We see that the photon A_μ remains massless. Mass terms for the W and Z bosons have the general form $M_W^2 W_\mu W^\mu$ and $\frac{1}{2}M_Z^2 Z_\mu Z^\mu$. We thus find for the masses

$$M_W = \frac{1}{2}vg_2 \tag{1.11}$$

$$M_Z = \frac{1}{2}v\sqrt{g_1^2+g_2^2} \tag{1.12}$$

$$M_A = 0\ . \tag{1.13}$$

After spontaneous symmetry breaking, three gauge bosons have thus absorbed a degree of freedom from the scalars (corresponding to the would-be Goldstone bosons) and became massive. One massless gauge boson and one scalar remain. The remaining scalar degree of freedom H corresponds to the so-called Higgs boson. Its mass and couplings can be determined from the scalar Lagrangian (1.4) upon substituting equation 1.6. Using $v^2 = -\frac{\mu^2}{\lambda}$ and extracting the parts containing only H, we find for the Lagrangian of the Higgs boson:

$$\mathcal{L}_H = \frac{1}{2}(\partial_\mu H)(\partial^\mu H) - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4}H^4\ . \tag{1.14}$$

Scalar masses have the general form $\frac{1}{2}m\phi^2$; the Higgs boson mass is thus

$$m_H = 2\lambda v^2 = -2\mu^2\ .$$

The only thing we still need to do, is generate masses for the fermions. This can be done by introducing Yukawa coupling terms between the fermions and the Higgs fields. The Yukawa Lagrangian for the first generation is given by

$$\mathcal{L}_F = -\lambda_e \overline{L}\Phi e_R - \lambda_d \overline{Q}\Phi d_R - \lambda_u \overline{Q}\widetilde{\Phi} u_R + h.c.\ , \tag{1.15}$$

where we introduced the conjugate of Φ, $\widetilde{\Phi} = i\tau_2\Phi^*$ which has negative hypercharge. This is needed to be able to couple to the up-type quarks. It is also possible to introduce a completely new Higgs doublet with negative hypercharge. This kind of model is called a two-Higgs doublet model (2HDM) and is needed to introduce supersymmetry (see section 2.3).

Substituting (1.6) in (1.15), we find

$$\mathcal{L}_F = -\frac{1}{\sqrt{2}}\lambda_e(\overline{\nu}_e, \overline{e}_L)\begin{pmatrix}0\\ v+H\end{pmatrix}e_R + \ldots \tag{1.16}$$

$$= -\frac{1}{\sqrt{2}}\lambda_e(v+H)\overline{e}_L e_R + \ldots\ , \tag{1.17}$$

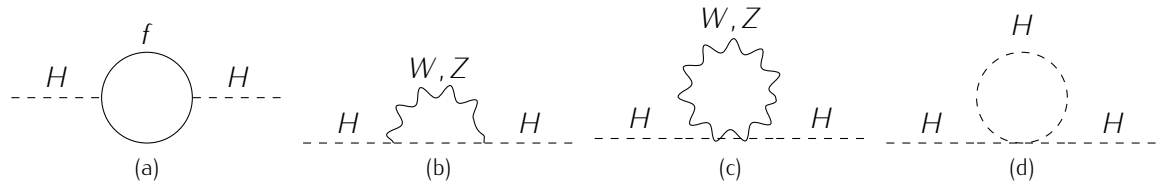


Figure 1.1: One-loop quantum corrections to the Higgs boson mass. (a) Contribution from the Yukawa interaction; (b), (c) Contribution from the gauge interaction; (d) Contribution from the Higgs self-interaction

where we highlighted the electron part. Fermion mass terms have the general form $m_f \bar{f}_L f_R + h.c.$. We find

$$m_e = \frac{\lambda_e v}{\sqrt{2}} \qquad m_u = \frac{\lambda_u v}{\sqrt{2}} \qquad m_d = \frac{\lambda_d v}{\sqrt{2}}$$

Using the same doublet of scalar fields we have thus successfully given mass to the gauge bosons and fermions in our theory. The Brout-Englert-Higgs mechanism also predicts the existence of a massive scalar particle, the Higgs boson.

1.2 The need for physics beyond the Standard Model

The Standard Model as outlined in the previous section has been very successful in describing the measurements done over the past decades. It does not, however, explain everything. There are still several unsolved issues, some of which will be explained in the next few sections. They will illustrate why we need new physics beyond the Standard Model. In this thesis we will focus on supersymmetry as extension of the Standard Model. In chapter 2, a more detailed discussion of supersymmetry is given.

1.2.1 The Higgs boson and the hierarchy problem

The Higgs boson has eluded discovery ever since its existence was predicted in the 1960's. Many attempts have been made to discover the Higgs boson. Direct searches at the Large Electron Positron (LEP) collider at CERN have excluded, at the 95% confidence level, Higgs boson masses up to 114.4 GeV [3]. Searches performed at the Tevatron at FermiLab also did not find any signs of the Higgs boson. They have excluded, at the 95% confidence level, the mass range $158 - 175$ GeV [4]. As long as the Higgs boson is not found, we cannot be certain that the Brout-Englert-Higgs mechanism is the mechanism for mass generation realised in nature.

Another problem related to the Higgs boson is the hierarchy problem. In a quantum field theory, we need to take radiative corrections into account. For fermions and gauge bosons these corrections are at most logarithmically divergent, which can be renormalized without much problems. For the Higgs boson, however, it turns out that when we calculate one-loop quantum corrections, the Higgs mass becomes quadratically dependent on the ultraviolet cut-off scale used to regularize the loop integral. This behavior is unwanted.

Let us illustrate this by considering a fermionic loop correction to the Higgs boson mass, as shown in figure 1.1a. After electroweak symmetry breaking, the Yukawa coupling terms in the Lagrangian are given by (1.17). We can write this for a generic fermion as:

$$\mathcal{L}_{f\bar{f}\phi} = -\frac{\lambda_f}{\sqrt{2}} H \bar{f} f - \frac{\lambda_f v}{\sqrt{2}} \bar{f} f \ , \tag{1.18}$$

where λ_f is the Yukawa coupling between the fermion f and the Higgs field Φ and $m_f = \frac{\lambda_f v}{\sqrt{2}}$ is the mass of the fermion. Neglecting the external momentum of the Higgs boson, one obtains for the one-loop fermion diagram

$$\Delta M_H^2 = N_f \frac{\lambda_f^2}{8\pi^2}\left[-\Lambda^2 + 6m_f^2 \log\left(\frac{\Lambda}{m_f}\right) - 2m_f^2\right] + \mathcal{O}\left(\frac{1}{\Lambda^2}\right) \tag{1.19}$$

There is a clear quadratic dependence on the momentum cut-off Λ. Λ is the scale up to which our theory is valid and is usually taken to be the Planck scale of $\mathcal{O}(10^{19})$ GeV, where a theory of quantum gravity is needed, or the Grand Unification scale of $\mathcal{O}(10^{16})$ GeV. In either case, the one-loop correction to M_H^2 becomes roughly 30 orders of magnitude larger than M_H^2 itself, which is supposed to be of the order of electroweak symmetry breaking scale. For reasons of unitarity and perturbativity, the Higgs boson should stay light, under 1 TeV. To achieve this, we would need to add a counterterm and tune it to an unnaturally high precision. This is called the naturalness or finetuning problem. The question "Why is $\Lambda >> M_Z$?" is the question at the heart of the hierarchy problem.

The hierarchy problem can also be seen as the lack of symmetry protecting the Higgs boson mass against large scales. For gauge bosons, the local gauge symmetry provides this protection, while fermions are protected due to the chiral symmetry. No such symmetry is present for the Higgs boson.

Assume now that there are also N_S scalar particles in our theory, with mass m_S, trilinear coupling $v\lambda_S$ and quadrilinear coupling λ_S. We then get an additional contribution to the one-loop correction of the form

$$\Delta M_H^2 = \frac{N_S\lambda_S}{16\pi^2}\left[-\Lambda^2 + 2m_S^2 \log\left(\frac{\Lambda}{m_S}\right)\right] - \frac{\lambda_S^2 N_S}{16\pi^2}v^2\left[-1 + 2\log\left(\frac{\Lambda}{m_S}\right)\right] + \mathcal{O}\left(\frac{1}{\Lambda^2}\right) \tag{1.20}$$

If we now assume $\lambda_f^2 = -\lambda_S$ and $N_S = 2N_f$, we find upon adding both contributions

$$\Delta M_H^2 = \frac{\lambda_f^2 N_f}{4\pi^2}\left[\left(m_f^2 - m_S^2\right)\log\left(\frac{\Lambda}{m_S}\right) + 3m_f^2 \log\left(\frac{m_S}{m_f}\right)\right] \tag{1.21}$$

We see that all quadratic divergent terms have vanished. Introducing scalar particles with the appropriate couplings has thus technically solved the hierarchy and naturalness problem. The logarithmically divergent term can also vanish in case $m_S = m_f$. In that case the total correction to the Higgs boson mass vanishes completely.

For the other loop diagrams in figure 1.1, an analog reasoning can be made. These diagrams themselves give a divergent contribution, but it can be cancelled by the introduction of new particles. For the W, Z and Higgs bosons we have to introduce fermionic partners with just the right couplings to the Higgs boson. In this way all divergent contributions to the Higgs mass will vanish.

This symmetry between bosons and fermions (and their couplings) is exactly what will happen in supersymmetric theories. In supersymmetry the necessary additional particles are introduced in a natural way.

1.2.2 Dark matter

From cosmology we have a good indication that roughly 25% of the universe is made out of dark matter. Dark matter is called "dark" because it cannot be observed using electromagnetic radiation. It is currently not possible to explain various observations without assuming the existence of dark matter. Among these are the characteristics of the cosmic microwave background radiation, the shape of galactic rotation curves and the structure formation in the universe.

Dark matter is generally divided in two main classes: "hot" and "cold" dark matter. Neutrinos are an example of hot dark matter. Due to their very low masses, they move at relativistic speeds

and cause "free-streaming". As a result, structures at small scales are wiped out. To explain the observed structures, such as galaxies and galaxy clusters, there cannot be too much hot dark matter. Instead, we need large amounts of the so-called cold dark matter. Particles constituting cold dark matter should be stable, have a large mass and should only be subject to the weak interaction. In the Standard Model there is no particle with these properties, whereas in a broad class of supersymmetric theories there is (see section 2.5.2).

1.2.3 Matter-antimatter asymmetry

When we look around in the universe, we see a large asymmetry between matter and antimatter. Everything around us seems to be made from matter. No evidence for large areas of antimatter has ever been found. Assuming that matter and antimatter were produced in equal amounts in the Big Bang and were in thermal equilibrium shortly thereafter, the Standard Model cannot explain how the asymmetry has come to be. Andrei Sakharov postulated three necessary conditions for the existance of a baryon asymmetry [5]:

1. charge (C) and charge-parity (CP) violation
2. no thermal equilibrium
3. at least one baryon number violating process

In the Standard Model, there is a small extent of CP violation present, e.g. in the decay of the neutral kaon. However, even when assuming that the last two conditions could be satisfied, there is not enough CP violation to explain the matter-antimatter discrepancy.

1.2.4 The quest for a Theory Of Everything

For many, the ultimate goal would be to find a theory that can describe all fundamental interactions and particles. This would be a Theory of Everything (TOE). The quest for this TOE started with the unification of electricity and magnetism and can only end when gravity can be included into the theory as well.

A first step would be to find a satisfactory Grand Unified Theory (GUT). At the Grand Unification scale, usually assumed to be $\mathcal{O}(10^{16}$ GeV), nature (gravity not included) would be described by a single gauge group and all Standard Model couplings would be unified. Going to lower energies this gauge symmetry should be broken to $SU(3)_c \times SU(2)_L \times U(1)_Y$, the Standard Model gauge group, and the observed differences between the coupling constants should emerge.

This reasoning can be inverted to test whether the Standard Model predicts unification or not. Using the measured coupling constants and the renormalization group equations to calculate the running of the coupling constants, one can extrapolate the couplings up to the GUT scale. Within the framework of the Standard Model, no point of unification of the coupling constants is found. This result strongly hinders the construction of a GUT.

Chapter 2

Supersymmetry

Supersymmetry, or SUSY, is a symmetry relating bosons and fermions. The operator Q that generates supersymmetry transformations is an anticommuting spinor, with

$$\begin{aligned} Q\,|\text{Boson}\rangle &= |\text{Fermion}\rangle \\ Q\,|\text{Fermion}\rangle &= |\text{Boson}\rangle\,. \end{aligned} \tag{2.1}$$

Q is a fermionic operator and carries a spin angular momentum of 1/2. This immediately gives the link with the spacetime operators of the Lorentz group. The generators Q and $Q^\dagger$ must satisfy the anticommutation and commutation relations

$$\{Q, Q^\dagger\} = P^\mu \tag{2.2}$$

$$\{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0 \tag{2.3}$$

$$[P^\mu, Q] = \left[P^\mu, Q^\dagger\right] = 0, \tag{2.4}$$

with P^μ the four-momentum generator of spacetime translations.

In this chapter, some elements of supersymmetry will be explained. This chapter is based on references [6], [7], [8] and [9].

2.1 Chiral and gauge supermultiplets

Particles in a supersymmetric theory are represented by *superfields*. These are spin multiplets, also called *supermultiplets*, and contain both fermion and boson states. Interactions between particles can be formulated in terms of the superfields. Superfields live in superspace, which is an extension of the usual spacetime with fermionic (Grassmann) variables.

The supersymmetry generators also commute with the generators of gauge transformations. As a consequence, particles in the same supermultiplet also belong to the same representation of the gauge group. They will thus have the same electric charge, weak isospin and color.

The simplest possibility for a supermultiplet consists of a two-component Weyl fermion and a complex scalar field. It is called a *chiral* or *matter supermultiplet*. As a second possibility we can also include a spin-1 vector boson. The resulting supermultiplet is called a *gauge* or *vector supermultiplet* and consists of a massless spin-1 boson and a spin-1/2 Weyl fermion. As gauge bosons transform under the adjoint representation of the gauge group, so must their superpartners, the so-called gauginos. This means that left-handed and right-handed gauginos will transform the same under the gauge group as opposed to the 'ordinary' fermions we know from the Standard Model.

There are other possibilities to construct supermultiplets, but if they have renormalizable interactions they can all be reduced to chiral and gauge supermultiplets. Hence, these will not be considered here.

2.2 Supersymmetry breaking

If supersymmetry would be unbroken, superpartners would have the same mass. This is a consequence of the commutation of the momentum generator P^μ and the supersymmetry generator Q. This means that for every Standard Model particle there would be a supersymmetric partner with the same mass. These particles would have been discovered a long time ago. As none of the new supersymmetric particles have been discovered so far, supersymmetry, assuming it exists, must be broken so that the superpartners have high mass.

This has its influence on the discussion of the hierarchy problem. When assuming perfect supersymmetry, there is no correction to the Higgs boson mass. When supersymmetry is broken, this nice cancellation no longer happens. In order to avoid the reappearance of quadratic divergencies, we need the breaking to be soft. This means that the relations between the dimensionless coupling constants (e.g. $\lambda_S = |\lambda_f|^2$) must still hold. If this is the case we will only have logarithmically divergent terms contributing to the Higgs mass correction. They will be of the form

$$\Delta m_H^2 = m_{soft}^2 \left(\frac{\lambda}{16\pi^2} \ln \frac{\Lambda_{UV}}{m_{soft}} + \dots \right), \tag{2.5}$$

with m_{soft} the mass scale associated with the soft terms, λ a schematic representation of various couplings and where the ellipses include higher order loop corrections and terms independent of Λ_{UV}. The masses of the superpartners cannot be too huge, otherwise the m_{soft}^2 corrections to the Higgs mass would become unnaturally large again. One finds that the masses of at least the lightest superpartners should be at most of the order of 1 TeV.

Up to now, nobody has found a satisfactory way to dynamically break supersymmetry. Therefore, one introduces by hand the possible soft SUSY breaking terms, resulting in a low-energy effective theory. This effective theory can then be used to predict masses, decays et cetera.

2.3 The Minimal SuperSymmetric Model

We are now ready to construct a minimal supersymmetric extension of the Standard Model. The resulting model is called the Minimal SuperSymmetric Model or MSSM. As a first step all the Standard Model particles have to be fitted in supermultiplets. A new conserved quantity will also be imposed. We will then construct the Lagrangian of the MSSM and include the soft SUSY-breaking terms.

Particle content As explained in section 2.1, Standard Model fermions – quarks and leptons – will necessarily be members of chiral supermultiplets. Their spin-0 superpartners are called sfermions – squarks and sleptons – where 's' stands for scalar. The left- and right-handed pieces of a quark or lepton reside in separate supermultiplets and thus have a separate complex scalar superpartner. These superpartners are also assigned a 'handedness', which does not correspond to their helicity (they are spin-0 particles), but to that of their partners.

The Higgs fields will also reside in a chiral supermultiplet since they are spin-0 fields. In supersymmetry it turns out that one Higgs chiral supermultiplet is not enough. Two such multiplets are needed. There are two main reasons for this.

The first is related to anomaly cancellation. An anomaly occurs when a classical symmetry of the Lagrangian is not conserved at the quantum level. If this happens for a gauge symmetry, this causes the theory to be inconsistent. In the Standard Model, the specific particle content results in an anomaly free gauge symmetry. When we now include a fermionic superpartner of the Higgs boson (called a higgsino), this nice cancellation between anomalies will no longer happen. To restore this, we need two Higgs supermultiplets with opposite hypercharge.

Table 2.1: Chiral supermultiplets in the MSSM

Names		Spin 0	Spin 1/2	$SU(3)_c \times SU(2)_L \times U(1)_Y$
squarks, quarks (3 families)	$\widehat{Q}$	$(\widetilde{u}_L \quad \widetilde{d}_L)$	$(u_L \quad d_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	$\widehat{U}^c$	$\widetilde{u}_R^*$	$u_R^\dagger$	$(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	$\widehat{D}^c$	$\widetilde{d}_R^*$	$d_R^\dagger$	$(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
sleptons, leptons (3 families)	$\widehat{L}$	$(\widetilde{\nu} \quad \widetilde{e}_L)$	$(\nu \quad e_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	$\widehat{E}^c$	$\widetilde{e}_R^*$	$e_R^\dagger$	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs, higgsinos	$\widehat{H}_u$	$(H_u^+ \quad H_u^0)$	$(\widetilde{H}_u^+ \quad \widetilde{H}_u^0)$	$(\mathbf{1}, \mathbf{2}, +\frac{1}{2})$
	$\widehat{H}_d$	$(H_d^0 \quad H_d^-)$	$(\widetilde{H}_d^0 \quad \widetilde{H}_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$

Table 2.2: Gauge supermultiplets in the MSSM

Names		Spin 0	Spin 1/2	$SU(3)_c \times SU(2)_L \times U(1)_Y$
gluino, gluon	$\widehat{G}_a$	$\widetilde{g}_a$	g^μ	$(\mathbf{8}, \mathbf{1}, 0)$
winos, W bosons	$\widehat{W}_a$	$\widetilde{W}_a$	W_a^μ	$(\mathbf{1}, \mathbf{3}, 0)$
bino, B boson	$\widehat{B}$	$\widetilde{B}$	B^μ	$(\mathbf{1}, \mathbf{1}, 0)$

The second reason can be found in the structure of supersymmetric theories. Only a $Y = +1/2$ Higgs chiral supermultiplet can have the Yukawa couplings needed to give masses to up-type quarks and similarly only a $Y = -1/2$ Higgs can give mass to down-type quarks and charged leptons. We will denote these Higgs supermultiplets by H_u and H_d, respectively. Due to the chiral nature of the supermultiplets we can no longer use the conjugate of H_u as our second Higgs doublet. This conjugate would have the wrong chirality.

The chiral supermultiplets present in the MSSM are listed in table 2.1. Superpartners of the ordinary Standard Model particles are denoted with a tilde. In the last column, we also listed the representation of the Standard Model gauge group under which the supermultiplet transforms. We defined all chiral supermultiplets in terms of left-handed Weyl spinors. Therefore the conjugates of the right-handed quarks and leptons appear in the table.

The vector bosons of the Standard Model will have to reside in gauge supermultiplets. An overview is given in table 2.2. After electroweak symmetry breaking, which is explained in section 2.5.1, the neutral gauge bosons W^3 and B will mix to form the mass eigenstates Z^0 and A. The corresponding gaugino mixed states are called the zino ($\widetilde{Z}^0$) and photino ($\widetilde{A}$).

R-parity To avoid possible lepton and baryon number violating interactions, a discrete, multiplicative symmetry is imposed. This new symmetry is called R-parity and its quantum number is given by:

$$R = (-1)^{3B+L+2s}, \tag{2.6}$$

with B, L and s the baryon number, lepton number and spin quantum number. Regular Standard Model particles have $R = 1$, their supersymmetric partners have $R = -1$.

This has some important consequences for the phenomenology of the MSSM. At particle colliders which collide Standard Model particles, supersymmetric particles can only be produced in pairs. Also, a SUSY particle will always decay into an odd number of SUSY particles. The lightest supersymmetric particle (LSP) will necessarily be stable.

Lagrangian A Lagrangian consists of kinetic terms and interaction terms. The kinetic terms for the MSSM are the supersymmetric equivalents of those of the SM (written as two-component spinors) and are given by

$$\mathcal{L}_{kin} = -\frac{1}{4} G^a_{\mu\nu} G_a^{\mu\nu} + \widetilde{G}^{\dagger a} i\overline{\sigma}^\mu D_\mu \widetilde{G}_a + f^{\dagger i} i\overline{\sigma}^\mu D_\mu f_i - (D_\mu \phi_i)^\dagger D^\mu \phi^i \ . \tag{2.7}$$

Here G stands for any gauge boson, $\widetilde{G}$ for any gaugino, f for any chiral fermion and ϕ for any scalar.

In the superfield formalism the interaction Lagrangian for the chiral superfields can be obtained from a superpotential $\mathcal{W}$ as

$$\mathcal{L}_{int} = -\sum_i \left| \frac{\partial \mathcal{W}}{\partial z_i} \right|^2 - \frac{1}{2} \sum_{ij} \left(\overline{f}_i \frac{\partial^2 \mathcal{W}}{\partial z_i \partial z_j} f_j + h.c. \right) , \tag{2.8}$$

where z_i are the superfields in the theory. The superpotential giving rise to the supersymmetric interaction Lagrangian of the MSSM is given by

$$\mathcal{W} = \sum_{i,j} -Y^u_{ij} \widehat{U}^c_i \widehat{H}_u \cdot \widehat{Q}_j + Y^d_{ij} \widehat{D}^c_i \widehat{H}_d \cdot \widehat{Q}_j + Y^l_{ij} \widehat{E}^c_i \widehat{H}_d \cdot \widehat{L}_j + \mu \widehat{H}_u \cdot \widehat{H}_d, \tag{2.9}$$

where i, j run over the three generations and Y_{ij} represents the Yukawa coupling among generations. The first three terms are direct generalizations of the Yukawa interactions present in the Standard Model. The last term is a globally supersymmetric mass term for the Higgs fields.

In addition to $\mathcal{L}_{int}$ there are also some interaction terms involving the sfermions and gauginos:

$$\sqrt{2} g \widetilde{G} \phi^* T_a f + h.c. + \frac{g^2}{2} |\phi^* T_a \phi|^2 \tag{2.10}$$

These are the supersymmetric counterparts of the ordinary fermion-gauge boson couplings.

As mentioned before, supersymmetry has to be softly broken. The soft SUSY-breaking terms in the MSSM are listed below.

1. gaugino mass terms:

$$-\mathcal{L}_{gaugino} = \frac{1}{2} \left(M_1 \widetilde{B}\widetilde{B} + M_2 \sum_{a=1}^{3} \widetilde{W}^a \widetilde{W}_a + M_3 \sum_{a=1}^{8} \widetilde{g}^a \widetilde{g}_a + h.c. \right) \tag{2.11}$$

2. scalar fermion mass terms:

$$-\mathcal{L}_{sfermion} = \sum_{i=gen} \left(m^2_{\widetilde{Q}_i} \widetilde{Q}^\dagger_i \widetilde{Q}_i + m^2_{\widetilde{L}_i} \widetilde{L}^\dagger_i \widetilde{L}_i + m^2_{\widetilde{u}_{Ri}} |\widetilde{u}_{Ri}|^2 + m^2_{\widetilde{d}_{Ri}} \left| \widetilde{d}_{Ri} \right|^2 + m^2_{\widetilde{e}_{Ri}} |\widetilde{e}_{Ri}|^2 \right) \tag{2.12}$$

3. mass and bilinear terms for the Higgs bosons:

$$-\mathcal{L}_{Higgs} = m^2_{H_u} H^\dagger_u H_u + m^2_{H_d} H^\dagger_d H_d + B \cdot \mu (H_u \cdot H_d + h.c.) \tag{2.13}$$

4. trilinear couplings between sfermions and Higgs bosons:

$$-\mathcal{L}_{trilinear} = \sum_{i,j=gen} \left(A^u_{ij} Y^u_{ij} \widetilde{u}^*_{Ri} H_u \cdot \widetilde{Q}_j + A^d_{ij} Y^d_{ij} \widetilde{d}^*_{Ri} H_d \cdot \widetilde{Q}_j + A^l_{ij} Y^l_{ij} \widetilde{e}^*_{Ri} H_d \cdot \widetilde{L}_j + h.c. \right) \tag{2.14}$$

2.4 From the Minimal SuperSymmetric Model to Minimal SUperGRAvity

2.4.1 Unconstrained MSSM

As explained in the previous section, the MSSM is based on four main assumptions:

- Minimal particle content
- Minimal gauge group
- Minimal Yukawa interactions and R-parity conservation
- Minimal set of soft SUSY-breaking terms

The Lagrangian that follows from these assumptions contains a large number of free parameters (124 in total) and describes the so-called unconstrained MSSM. Many parameters in this model can take values that would result in phenomenological problems, such as large amounts of CP violation.

To be able to study the MSSM experimentally, this large parameter space needs to be reduced. A way to do this is explained in the next sections.

2.4.2 Phenomenological MSSM

As a first step in reducing the parameter space, the following assumptions are made.

1. There is no CP violation generated by the soft SUSY-breaking terms. Therefore, these terms are all real rather than complex.
2. The matrices for the sfermion masses and trilinear couplings are diagonal. This means that there are no Flavour Changing Neutral Currents (FCNC) at tree level.
3. At low energies, the first and second generation of sfermions have the same soft SUSY-breaking masses and trilinear couplings.

Using these assumptions, there are only 22 free SUSY parameters left.

$\tan\beta$: ratio of the vev's of the two Higgs doublet fields, see section 2.5.1

$m^2_{H_u}, m^2_{H_d}$: Higgs boson masses

M_1, M_2, M_3 : gaugino masses

$m_{\tilde{q}}, m_{\tilde{u}_R}, m_{\tilde{d}_R}, m_{\tilde{l}}, m_{\tilde{l}_R}$: 1st and 2nd generation sfermion masses

A_u, A_d, A_l : 1st and 2nd generation trilinear couplings

$m_{\tilde{Q}}, m_{\tilde{t}_R}, m_{\tilde{b}_R}, m_{\tilde{L}}, m_{\tilde{\tau}_R}$: 3rd generation masses

A_t, A_b, A_τ : 3rd generation trilinear couplings

Sometimes, one completely neglects the trilinear couplings for the first and second generation. Then one remains with a 19-dimensional model. From the experimental side, the 19 or 22 parameters of this model are a lot easier to handle than the 124 parameters of the unconstrained MSSM. This model is called the phenomenological MSSM or pMSSM.

2.4.3 Minimal supergravity

Although the pMSSM is already much more manageable than the unconstrained MSSM, further approximations can still be made. These are usually universal conditions valid at the GUT-scale.

In the minimal supergravity or mSUGRA model, the main assumption is that supersymmetry breaking occurs in a hidden sector which communicates with the visible sector only via gravitational interactions. These interactions are assumed to be flavour blind.

The mSUGRA unification conditions at the GUT-scale are:

1. unification of gaugino masses

$$M_1 = M_2 = M_3 \equiv m_{1/2} \tag{2.15}$$

2. unification of scalar masses

$$m_{\tilde{Q}_i} = m_{\tilde{u}_{Ri}} = m_{\tilde{d}_{Ri}} = m_{\tilde{L}_i} = m_{\tilde{l}_{Ri}} = m_{H_u} = m_{H_d} \equiv m_0 \tag{2.16}$$

3. universal trilinear couplings

$$A^u_{ij} = A^d_{ij} = A^l_{ij} \equiv A_0 \delta_{ij} \tag{2.17}$$

In total the mSUGRA model after electroweak symmetry breaking (see next section) only has four continuous free parameters and 1 unknown sign: $\tan\beta$, m_0, $m_{1/2}$, A_0, sign(μ). The masses and couplings at lower energies can be found by using the renormalization group equations.

The mSUGRA model is the model used most in experiments nowadays. But there are also other possibilities to reduce the parameter space of the MSSM. These are e.g. based on anomaly or gauge mediated SUSY breaking.

2.5 Mass spectra and decay cascades

2.5.1 Electroweak symmetry breaking

In the MSSM there are two complex Higgs doublets, H_u and H_d, which acquire vacuum expectation values and cause spontaneous electroweak symmetry breaking. Three different sources contribute to the scalar Higgs potential V_H:

Firstly there are the terms containing the quartic Higgs interactions. These are the so-called "D-terms" and come from the second term in equation 2.10. The $U(1)_Y$ part is denoted by V^1_D, the $SU(2)_L$ part by V^2_D. We find:

$$V^1_D = \frac{1}{2}\left[\frac{g_1}{2}\left(|H_u|^2 - |H_d|^2\right)\right]^2 , \tag{2.18}$$

$$V^2_D = \frac{1}{2}\left[\frac{g_2}{2}\left(H^{i*}_d \tau^a_{ij} H^j_d + H^{i*}_u \tau^a_{ij} H^j_u\right)\right]^2 , \tag{2.19}$$

with $\tau^a = 2T^a$ and $\tau^a_{ij}\tau^a_{kl} = 2\delta_{il}\delta_{jk} - \delta_{ij}\delta_{kl}$. Summing these two contributions, we get

$$V_D = \frac{g_2^2}{8}\left[4|H_d^\dagger \cdot H_u|^2 - 2|H_d|^2|H_u|^2 + (|H_d|^2)^2 + (|H_u|^2)^2\right] + \frac{g_1^2}{8}\left[|H_u|^2 - |H_d|^2\right]^2 \tag{2.20}$$

The second source are the derivatives of the superpotential, the so-called "F-terms" (see equations (2.8) and (2.9)).

$$V_F = \sum_i \left|\frac{\partial \mathcal{W}(\phi_i)}{\partial \phi_i}\right|^2 = \mu^2\left(|H_d|^2 + |H_u|^2\right) \tag{2.21}$$

The last contribution comes from the soft SUSY-breaking scalar Higgs masses and bilinear terms (see equation (2.13)).

$$V_{\text{soft}} = m^2_{H_u} H^\dagger_u H_u + m^2_{H_d} H^\dagger_d H_d + B \cdot \mu (H_u \cdot H_d + h.c.) \tag{2.22}$$

The total scalar potential after expanding the Higgs fields into their components is thus given by

$$\begin{aligned} V_H = {} & \left(|\mu|^2 + m^2_{H_u}\right)\left(|H^0_u|^2 + |H^+_u|^2\right) + \left(|\mu|^2 + m^2_{H_d}\right)\left(|H^0_d|^2 + |H^-_d|^2\right) \\ & + b\left(H^+_u H^-_d - H^0_u H^0_u + h.c.\right) \\ & + \frac{1}{8}\left(g_1^2 + g_2^2\right)\left(|H^0_u|^2 + |H^+_u|^2 - |H^0_d|^2 - |H^-_d|^2\right) + \frac{1}{2}g_2^2|H^+_u H^{0*}_d + H^0_u H^{-*}_d|^2 , \end{aligned} \tag{2.23}$$

where we defined $b = B\mu$.

The minimum of this potential should break the electroweak symmetry down to electromagnetism $SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM}$. Using the freedom to make gauge transformations, we can choose without loss of generality $\langle 0|H_d^-|0\rangle = 0$ at the minimum. From $\frac{\partial V_H}{\partial H_d^-} = 0$ it follows that then also $\langle 0|H_u^+|0\rangle = 0$. The charged directions are thus indeed unbroken.

Setting $H_u^+ = H_d^- = 0$, the scalar potential becomes

$$\begin{aligned} V_H = &\left(|\mu|^2 + m_{H_u}^2\right)|H_u^0|^2 + \left(|\mu|^2 + m_{H_d}^2\right)|H_d^0|^2 - b\left(H_u^0 H_u^0 + h.c.\right) \\ &+ \frac{1}{8}\left(g_1^2 + g_2^2\right)\left(|H_u^0|^2 - |H_d^0|^2\right) \ . \end{aligned} \tag{2.24}$$

Only the b term depends on the phases of the fields. This means that a possible phase in b can always be absorbed by a redefinition of the phases of H_u or H_d. Hence, we can take b to be real and positive. As a consequence CP cannot be spontaneously broken by the Higgs scalar potential, at least at tree level. Also the Higgs scalar mass eigenstates will have well-defined eigenvalues of CP.

To have a viable theory, we need to make sure that the potential is bounded from below. If this were not the case, the vacuum would not be stable. The quartic interactions in the potential stabilize the potential for large values of H_u^0 and H_d^0. Only in the special case $|H_u^0| = |H_d^0|$, corresponding to a flat direction in field space, the quartic terms vanish. To ensure that the potential is still bounded from below, even in these cases, we find the requirement

$$2b < 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2 \ . \tag{2.25}$$

To have electroweak symmetry breaking, we need a linear combination of H_u^0 and H_d^0 to have a negative squared mass term. This gives the requirement

$$b^2 > \left(|\mu|^2 + m_{H_u}^2\right)\left(|\mu|^2 + m_{H_d}^2\right) \ . \tag{2.26}$$

If this inequality is not satisfied, $H_u^0 = H_d^0 = 0$ will be a stable minimum of the potential and there will be no electroweak symmetry breaking. It is also interesting to note that both inequalities can only be satisfied at the same time when $m_{H_u} \neq m_{H_d}$. In the mSUGRA model at the unification scale, we find at tree level that both masses are equal. Hence no EW symmetry breaking. However due to radiative corrections, which give rise to differences in the renormalization group equation, the masses at lower momentum scales will differ. These quantum corrections trigger electroweak symmetry breaking.

Assuming that the above stated conditions can be satisfied, we can now require the vev's to be compatible with the observed phenomenology of electroweak symmetry breaking. We write $v_u = \langle H_u^0\rangle$ and $v_d = \langle H_d^0\rangle$. They are related to the mass of the Z-boson

$$v_u^2 + v_d^2 = v^2 = \frac{2m_Z^2}{g_1^2 + g_2^2} \approx (174 \text{ GeV})^2 \tag{2.27}$$

The ratio of the vev's is written as $\tan\beta = \frac{v_u}{v_d}$ and is not fixed by current experiments. The minimization equations can be written as

$$m_{H_u}^2 + |\mu|^2 = b\cot\beta + \frac{m_Z^2}{2}\cos(2\beta) \tag{2.28}$$

$$m_{H_d}^2 + |\mu|^2 = b\tan\beta - \frac{m_Z^2}{2}\cos(2\beta) \ . \tag{2.29}$$

Using these equations we can write m_Z and $\tan\beta$ as a function of b and $|\mu|$. Or the other way round, we can always eliminate b and $|\mu|$ and replace them with the variable $\tan\beta$. Only the sign of μ remains undetermined.

The Higgs scalar fields of the MSSM are two complex $SU(2)_L$ doublets, or eight real scalar degrees of freedom. Three of these are used to give mass to the W- and Z-bosons. After electroweak symmetry breaking five Higgs scalar mass eigenstates remain: two CP-even neutral scalars h^0 and H^0, one CP-odd neutral scalar A^0 and two charged scalars H^+ and H^-. By convention h^0 is lighter than H^0. To compute their masses, one can expand the doublet fields around their vacuum expectation value and plug this expansion into the Higgs potential. One finds that the mass of h^0 is bounded from above. Including radiative loop corrections and assuming that the sparticles contributing to the loop diagrams are lighter than about 1 TeV, this bound is $m_{h^0} < 130$ GeV. The masses of the other Higgs bosons can, in principle, become arbitralily large.

2.5.2 Mass eigenstates

In the MSSM, the superpartners listed in tables 2.1 and 2.2 are not necessarily the mass eigenstates of the theory. After electroweak symmetry breaking, gauge eigenstates with the same $SU(3)_c \times U(1)_{EM}$ quantum numbers can mix with each other. To determine the mass eigenstates, one only needs to diagonalize the mass matrices, which can be constructed by gathering all terms quadratic in the fields.

Gauginos and higgsinos The charged electroweak gauginos $\widetilde{W}_1$, $\widetilde{W}_2$ and the charged higgsinos $\widetilde{H}_u^+$,$\widetilde{H}_d^-$ mix to form the so-called charginos, denoted by $\widetilde{\chi}_1^\pm$ and $\widetilde{\chi}_2^\pm$. The same happens for their neutral counterparts. They mix to form neutralinos, denoted by $\widetilde{\chi}_{1,2,3,4}^0$. Neutralinos are very interesting for cosmology. If the lightest supersymmetric particle is a neutralino, this is an ideal candidate for dark matter: it is heavy, only weakly interacting and stable. The gluino, which is a color octet, cannot mix with any other particle. In this respect it is unique among all MSSM sparticles. The mass eigenstate is thus the same as the gauge eigenstate, and has a mass equal to M_3, coming from the soft SUSY breaking part of the Lagrangian.

Squarks and sleptons In principle, any scalars with the same quantum numbers can mix with each other. This would mean that the mass eigenstates of the squarks and sleptons should be determined by diagonalizing three 6×6 mass matrices – one for up-type squarks, one for down-type squarks and one for charged sleptons – and one 3×3 mass matrix – for the sneutrinos.

Due to the hypothesis of flavour-blind soft parameters, mixing angles are generally small. The third generation squarks and sleptons can have very different masses compared to their first and second generation counterparts, because of the effects of the large Yukawa and soft couplings for the third generation. Furthermore, they can also have substantial mixing in pairs $(\widetilde{t}_L, \widetilde{t}_R)$, $(\widetilde{b}_L, \widetilde{b}_R)$ and $(\widetilde{\tau}_L, \widetilde{\tau}_R)$. One can calculate that the off-diagonal terms in the mass matrices are given by

$$m^2_{\widetilde{f}_{L,R}} = m_f \left(A_f + \mu \cot \beta\right) \tag{2.30}$$

for the up-type quarks and by

$$m^2_{\widetilde{f}_{L,R}} = m_f \left(A_f + \mu \tan \beta\right) \tag{2.31}$$

for the down-type quarks and charged fermions. This mixing is thus especially important for the stop-sector because of the large top mass. For the sbottoms and staus it might become significant in the case of large $\tan\beta$. Mixing for the first and second generation sfermions is insignificant due to the small fermion masses.

2.5.3 Sparticle decays

The decay possibilities for SUSY particles depend strongly on the precise SUSY mass spectrum. This is directly related to the question "How is supersymmetry broken?" Depending on the model, this question is answered differently.

In the case of mSUGRA we end up with only five free parameters: m_0, $m_{1/2}$, A_0, μ and b. After electroweak symmetry breaking we can trade $|\mu|$ and b for $\tan\beta$. The entire mass spectrum of mSUGRA is thus determined by four parameters and a sign. To study the phenomenology of mSUGRA, the universal masses m_0 and $m_{1/2}$ need to be evolved to the electroweak scale by using the renormalization group equations.

Once the particle mass spectrum of the SUSY model under study is known, one can investigate which decays are possible. If they are kinematically allowed, two-body decays will always dominate. Otherwise the decays can proceed via virtual sparticles which gives rise to three- and four-body decays.

Let us first consider the neutralinos and charginos. They contain an admixture of the electroweak gauginos and thus inherit their couplings. Possible two-body decays are the decays to lepton + slepton, quark + squark and neutralino/chargino + $W/Z/h$. If these are not allowed, the three-body decay to fermion + fermion + neutralino/chargino will occur. In case the sleptons are relatively light, these fermions will often be leptons. In case the decay proceeds through the lightest Higgs boson, these fermions are often b-quarks. This then gives rise to b-jets in the final state.

For squarks, the decay to quark + gluino will dominate if the gluino is light enough. This is due to the large QCD coupling. In case this decay is not accessible, the decay can proceed via the electroweak interaction to a quark + neutralino/chargino. For right-handed squarks the decay to the lightest neutralino will usually dominate, as in many cases the lightest neutralino is predominantly bino.

Gluinos can only decay through a squark, wich can be on-shell or not. The stop and sbottom can be much lighter than the other squarks, which means that it is quite possible that $\widetilde{g} \to t\widetilde{t}_1$ and $\widetilde{g} \to b\widetilde{b}_1$ are the only available two-body decays. In this case, they will dominate and b-jets will appear in the final states.

2.5.4 SUSY decay cascades and experimental signatures

In this section we will briefly discuss the experimental signatures of supersymmetry at hadron colliders such as the Large Hadron Collider (LHC) (see chapter 3 for more information).

At hadron colliders, sparticles can only be produced in pairs. This is a consequence of the conservation of R-parity. Possible production reactions can proceed via the strong interaction, such as

$$gg \to \widetilde{g}\widetilde{g}, \quad \widetilde{q}_i\widetilde{q}_j^*, \tag{2.32}$$

$$gq \to \widetilde{g}\widetilde{q}_i, \tag{2.33}$$

$$q\overline{q} \to \widetilde{g}\widetilde{g}, \quad \widetilde{q}_i\widetilde{q}_j^*, \tag{2.34}$$

$$qq \to \widetilde{q}_i\widetilde{g}_j, \tag{2.35}$$

or via the electroweak interaction, such as

$$q\overline{q} \to \widetilde{\chi}_i^+\widetilde{\chi}_j^-, \quad \widetilde{\chi}_i^0\widetilde{\chi}_j^0, \quad \widetilde{l}_i^+\widetilde{l}_j^-, \quad \widetilde{\nu}_l\widetilde{\nu}_l^*, \tag{2.36}$$

$$u\overline{d} \to \widetilde{\chi}_i^+\widetilde{\chi}_j^0, \quad \widetilde{l}_L^+\widetilde{\nu}_l, \tag{2.37}$$

$$d\overline{u} \to \widetilde{\chi}_i^-\widetilde{\chi}_j^0, \quad \widetilde{l}_L^-\widetilde{\nu}_l^* \ . \tag{2.38}$$

At the LHC, the production of squarks and gluinos by gluon-gluon and gluon-quark fusion usually dominates. These then decay as explained in the previous section, resulting in a so-called decay cascade in which quarks and leptons can be produced and which ends with the production of at least two LSP's due to the conservation of R-parity. Usually this LSP is a neutralino, which is only weakly interacting and will thus escape from the detector without being detected. This will result in a considerable amount of the so-called missing transverse energy.

Missing transverse energy (MET) is defined as the opposite of the vectorial sum in the transverse plane (with respect to the beam axis) of the momenta of all particles[1]. When all particles in the collision are measured, this sum should be zero. If this sum has a large, non zero value, it usually means that a particle was not detected. This could thus be an indication of supersymmetry.

Promising search topologies at the LHC are thus those topologies that look for large amounts of missing transverse energy, several jets and possibly some leptons.

The classic signal for supersymmetry is missing transverse energy and jets, without isolated leptons. This signal can get contributions from all types of sparticle pair production, except sleptons. Vetoing leptons reduces the background from Standard Model processes involving leptonic W decays. The main background contribution comes from $t\bar{t}$ production and QCD multijet production. This QCD background can be lowered by putting a high enough MET cut to avoid backgrounds from mismeasurements of the jet energy.

The single lepton + jets + MET topology can have a large rate from various sparticle production mechanisms, but also has large Standard Model backgrounds from processes with $W \to l\nu$. The QCD contribution is however substantially lower than in the previous case.

Another search topology is the one with two leptons with the same charge, the so-called same sign dileptons. This signal can occur for example if the gluino decays to hadrons plus a chargino. This chargino can then decay into a charged lepton, a neutrino and a neutralino. In the case where we had gluino pair production, this means that two leptons can be present in the final state. These leptons can have the same charge because the gluino is a Majorana fermion and does not know about electric charge. This signature has only small backgrounds from Standard Model processes. The main SM backgrounds for dilepton signatures come from W^+W^-, $t\bar{t}$ and Z/γ production, and these processes can only give leptons with opposite charge.

In addition to these topologies, it can also be interesting to require b-jets. Due to the potential large mixing effects for third generation squarks, especially for large values of $\tan\beta$, sbottoms and stops can be much lighter than their first and second generation counterparts. This can result in final states which are rich in b-jets. A possibility for this is when the only two-body decay of the gluino is $\widetilde{g} \to b\widetilde{b}_1$ with $\widetilde{b}_1 \to b\widetilde{\chi}^0$. Depending on the decay of the neutralino, additional leptons can be produced as well.

A lot of different signatures are thus present in the SUSY parameter space. Searches for these signatures have so far all come up empty handed. This has allowed to put constraints on the SUSY parameter space. Direct searches are however not the only way to get information about the possible existence of supersymmetry. In the next section we will explain how several low energy observables can also yield valuable information.

2.6 Low energy observables and their implications for supersymmetry

In general, the Standard Model provides a good fit for the existing high precision electroweak measurements. However, there are some observables whose SM predicted values significantly differ from the corresponding experimental measurements. A possible explanation for the discrepancies could be the presence of sparticles in the interactions. These sparticles could however also spoil the very precise agreement between the SM prediction and the experimentally determined values for some other observables. The influence of the sparticles is strongly connected to their masses. For large enough masses, the sparticles essentially decouple from the low-energy observables.

By studying electroweak precision observables, we can thus extract information about the possible scale of supersymmetry breaking. Among these observables are the muon anomalous magnetic

[1] In practice, one often sums the energies deposited in the calorimeter cells. This is why this variable is called missing transverse *energy*.

moment, the mass of top- and bottom-quarks and the mass of the lightest Higgs boson. In addition we can also use B-physics observables, such as the mass mixings of the B-mesons and the branching ratio of the rare decays $b \to s\gamma$, $B_s \to \mu^+\mu^-$ and $B \to \tau\nu$, to constrain SUSY. Below we will discuss some of these low energy observables (LEOs) [10] [11].

Muon anomalous magnetic moment The magnetic moment $\overrightarrow{M}$ of a particle with mass m and charge e is related to the particle spin $\overrightarrow{S}$ by the gyromagnetic ratio g:

$$\overrightarrow{M} = g\frac{e}{2m}\overrightarrow{S} \tag{2.39}$$

At tree level, QED predicts the exact result $g = 2$ for elementary spin-1/2 particles such as muons or electrons. Quantum effects from QED loop diagrams, from strong or weak interactions, or from hypothetical new particles lead to a deviation

$$a \equiv \frac{g-2}{2}\ , \tag{2.40}$$

the so-called anomalous magnetic moment. The theoretical prediction for a is dominated by the one-loop QED contribution, followed by higher-order QED and strong interaction effects. Loop contributions from heavy particles with mass M are generally suppressed by a factor $\frac{m^2}{M^2}$. Therefore, the anomalous magnetic moment of the muon is a factor $\left(\frac{m_\mu}{m_e}\right)^2 \approx 40000$ more sensitive to such contributions than the anomalous magnetic moment of the electron.

The muon anomalous magnetic moment has been measured to a very high accuracy by the E821 experiment at Brookhaven National Laboratory [12] :

$$a_\mu^{exp} = (11\ 659\ 208.9 \pm 5.4 \pm 3.3) \cdot 10^{-10} \tag{2.41}$$

where we quote the statistical and systematic uncertainties. Due to this high accuracy, we are now sensitive to effects of physics at the electroweak scale.

In order to mirror the high experimental accuracy, the Standard Model prediction has been refined a lot over the past years. The main source of uncertainty comes from the hadronic vacuum polarization. Including the most recent results on this, the SM prediction is [13]:

$$a_\mu^{SM} = (11\ 659\ 180.2 \pm 4.9) \cdot 10^{-10}\ . \tag{2.42}$$

This deviates from the experimental value by

$$\Delta a_\mu = (28.7 \pm 8.0) \cdot 10^{-10}\ . \tag{2.43}$$

One can interpret this deviation in the framework of supersymmetry, where in addition to the SM loop diagrams, there are also loops involving sfermions and gauginos. For sparticles of mass M_{SUSY}, we can approximate the additional contribution to a_μ as [14]:

$$a_\mu^{SUSY} = 13 \cdot 10^{-10} \left(\frac{100\ \text{GeV}}{M_{SUSY}}\right) \tan\beta\ \text{sign}(\mu)\ . \tag{2.44}$$

More accurate values for a certain parameter point in a given SUSY model can be computed using various software packages such as `SuperIso` [15] or `micrOMEGAs` [16].

The value of Δa_μ can then also be used to put contraints on the parameter space. From (2.44) we see that positive values for the parameter μ are strongly favoured if SUSY is to explain the discrepancy between a_μ^{SM} and a_μ^{exp}. It is also possible to put both lower and upper bounds on the masses of certain SUSY particles.

Lightest Higgs boson mass For a given SUSY model, we can calculate the mass of the lightest Higgs boson and compare this with the direct search limits obtained at LEP and Tevatron. This tells us whether or not this specific point in the parameter space has been excluded or not. The cross section limits from Higgs searches at LEP and Tevatron are implemented in the software package `HiggsBounds` [17].

Branching ratio ($b \to s\gamma$) In the Standard Model, the radiative flavor changing decay $b \to s\gamma$ is mediated by loops containing charge 2/3 quarks and W-bosons. The theoretically predicted value for the branching ratio agrees very well with the experimentally observed value of

$$(3.55 \pm 0.35) \cdot 10^{-4} \,, \tag{2.45}$$

measured in the decay $B \to X_s\gamma$.

In SUSY theories, there are additional contributions to this decay from loops involving charginos and stops or top quarks and charged Higgs bosons. The measurement of this decay branching ratio is thus a powerful tool to constrain the SUSY parameter space.

Branching ratio ($B_s \to \mu^+\mu^-$) The branching ratio for the flavour changing process $B_s \to \mu^+\mu^-$ is predicted to be $(3.4 \pm 0.5) \cdot 10^{-9}$ within the SM. In the MSSM, interactions involving neutral Higgs bosons can enhance the branching ratio by several orders of magnitude at high $\tan\beta$. The branching ratio is experimentally bounded from above by results of the CDF experiment. The current bound, as stated by the Particle Data Group [18] is given by

$$BR(B_s \to \mu^+\mu^-) < 4.7 \cdot 10^{-8} \,. \tag{2.46}$$

New physics phenomena should thus not exceed the SM prediction by more than an order of magnitude. This again puts constraints on the models.

Branching ratio ($B_u \to \tau\nu$) This purely leptonic decay proceeds through the annihilation of b and $\overline{u}$ quarks into a W-boson. In the MSSM for example, there is an additional contribution from the virtual exchange of a charged Higgs boson. One usually compares the experimental value (as stated by the Heavy Flavor Averaging Group [19]) with the theoretical value via their ratio:

$$R^{exp}_{B\tau\nu} = \frac{BR(B \to \tau\nu)_{exp}}{BR(B \to \tau\nu)_{SM}} = 1.66 \pm 0.54 \tag{2.47}$$

For the MSSM one can calculate (e.g. via `SuperIso`) the ratio

$$R^{MSSM}_{B\tau\nu} = \frac{BR(B \to \tau\nu)_{MSSM}}{BR(B \to \tau\nu)_{SM}} \tag{2.48}$$

and compare with experiment.

Dark matter relic density A final constraint on the SUSY parameter space comes from the Wilkinson Microwave Anisotropy Probe (WMAP) observations. Assuming a cosmological model including dark matter, they have obtained a value for the dark matter relic density in the universe. In SUSY models such as the MSSM, the lightest neutralino is usually a prime candidate for dark matter. One can then also compute the matter density it would give rise to (e.g. with `micrOMEGAs`) and compare this with the WMAP results.

Part II

Analysis

Chapter 3

The Compact Muon Solenoid at the Large Hadron Collider

3.1 The Large Hadron Collider

The Large Hadron Collider, or LHC, is, as its name already tells us, a large particle accelerator and collider [20]. It is located at CERN, the European Organization for Nuclear Research in Geneva. The LHC is situated in an underground tunnel – between 45 and 170m deep – and has a circumference of 26.7 km. It is not a perfect circle, but consists of eight straight and eight curved sections.

Although it can also operate with heavy ion (lead) beams, the LHC mostly collides protons. These beams are kept onto their tracks via superconducting magnets. When operating at the design energy of 7 TeV per beam, the magnets need to provide an extremely strong magnetic field of over 8 T. Due to the fact that the LHC collides particle beams, and not particle-antiparticle beams, both beams need to be steered by opposite magnetic fields. This results in two separate beam pipes which only cross at the points of collision. Due to space limitations, both beam pipes are located within the same magnet. This so-called twin-bore or two-in-one magnet is designed in such a way that the magnetic field in both beam pipes is opposite.

The LHC also achieves intensities that have never been obtained before, with a design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ for proton physics runs and $10^{27}\text{cm}^{-2}\text{s}^{-1}$ when colliding heavy ions. To achieve this luminosity, the LHC will collide 2808 bunches of protons per beam, with a bunch spacing of 25 ns.

LHC experiments There are four main experiments that detect the collisions provided by the LHC: two general purpose detectors ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid), a dedicated detector for b-physics LHCb (Large Hadron Collider beauty) and a dedicated detector for heavy ion physics ALICE (A Large Ion Collider Experiment). There are also two smaller experiments, LHCf (Large Hadron Collider forward) and TOTEM (TOTal Elastic and diffractive cross section Measurement). Figure 3.1 shows a schematic layout of the LHC.

Physics motivation The main physics goal of the LHC and its experiments is to answer some fundamental questions such as "What gives particles mass?", "What is dark matter?", "What happened to all the antimatter?" and "What was the state of matter just after the Big Bang?". The CMS and ATLAS experiments focus on providing insight in the first two questions by investigating the nature of electroweak symmetry breaking (cf. Higgs mechanism) and searching for new physics phenomena such as supersymmetry and extra dimensions. The LHCb experiment mainly focuses on CP violation and antimatter. The ALICE experiment finally, wants to study the properties of the quark-gluon plasma.

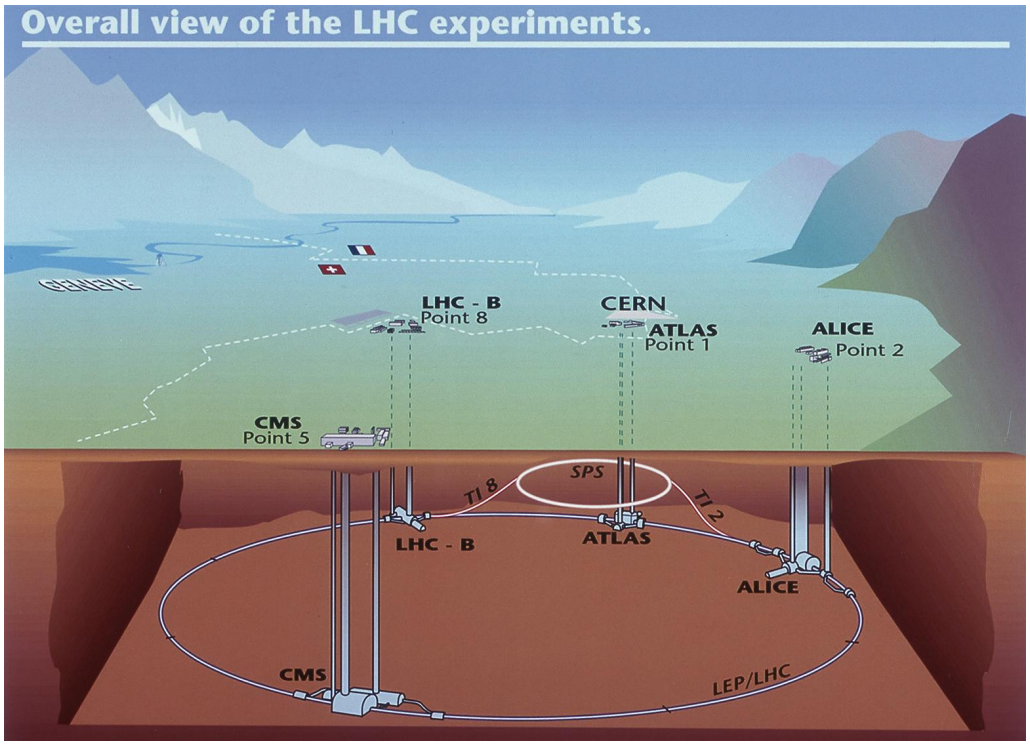


Figure 3.1: Schematic view of the Large Hadron Collider [21]

3.2 The Compact Muon Solenoid

The Compact Muon Solenoid experiment is one of the two general purpose detectors at the LHC. The CMS detector has the typical cylindrical shape with the beam pipe in the center and all subdetectors layered around it. There is a main barrel region and in addition two endcap regions. The CMS detector is 21.6 m long with a diameter of 14.6 m and has a total weight of 12500 tonnes. The main features are a high-field (4 T) superconducting solenoid, an extensive muon detection system, a silicon-based inner tracking system and electromagnetic and hadronic calorimeters [22].

Due to the very high luminosity of the LHC and short time between bunch crossings, detectors need to be carefully designed. At the design luminosity, there will be a lot of pile-up events. On average about 20 inelastic collisions will be superimposed. This means that every 25 ns, of the order of 1000 particles will pass through the detector. In order not to confuse one bunch crossing with another, detector response needs to be as short as possible, and ideally not more than 25 ns. Using detectors with high granularity also helps to mitigate the effects of pile-up. This consecutively results in a large number of detector channels which need to be perfectly synchronized.

Radiation levels in the detector are high because of the large flux of particles coming from the interaction region. This requires radiation-hard detectors and electronics.

In order for CMS to meet the physics goals of the LHC, the following requirements need to be fulfilled:

- Good muon identification and momentum resolution
- Good charged-particle momentum resolution and reconstruction efficiency
- Efficient triggering and tagging of τ's and b-jets
- Good electromagnetic energy resolution and efficient photon and lepton isolation at high luminosities
- Good missing transverse energy and dijet-mass resolution, requiring hadron calorimeters with a large geometric coverage and fine lateral segmentation

The coordinate system of CMS is a right-handed coordinate system with the origin at the nominal

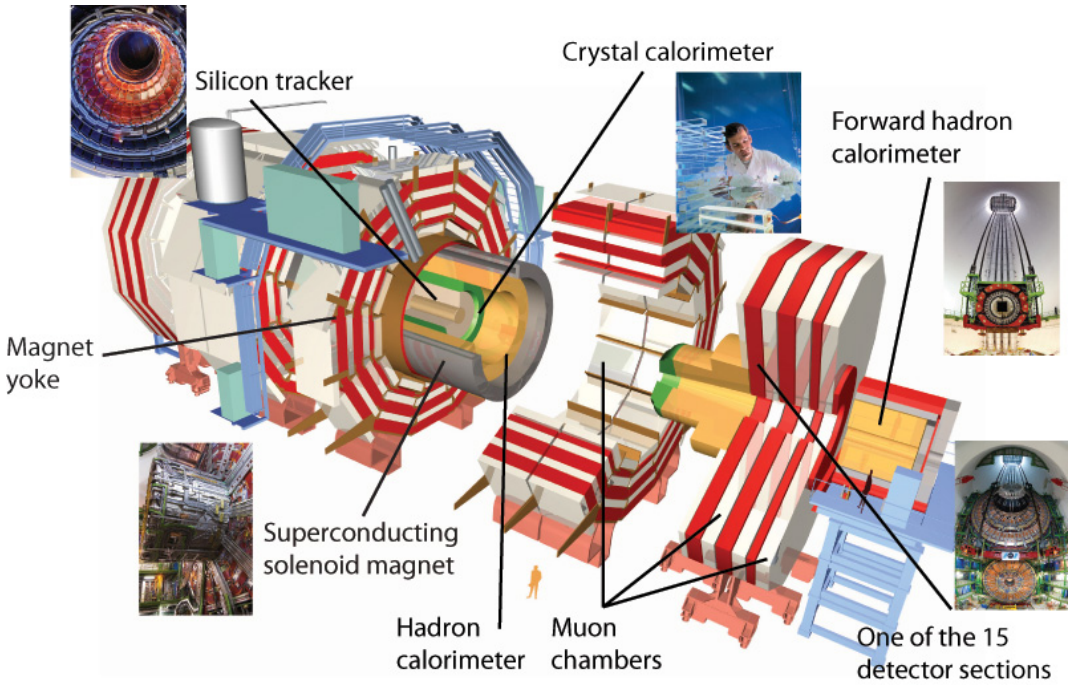


Figure 3.2: Schematic view of the Compact Muon Solenoid [23]

interaction point inside the experiment. The x-axis points radially inwards towards the center of the LHC, the y-axis points vertically upwards and the z-axis points along the beam direction. The azimuthal angle ϕ is measured in the $x-y$ plane, starting from the x-axis. The radial coordinate in this plane is denoted by r. The polar angle θ is measured from the z-axis. Often one uses pseudorapidity $\eta = -\ln\tan(\theta/2)$ instead of θ.

In the next sections some additional information about various subsystems of the CMS detector will be provided. An overview of the CMS detector is shown in figure 3.2.

3.2.1 CMS subsystems

Magnet

The CMS magnet system consists of a superconducting solenoid and an iron return yoke. The solenoid has a mass of 220 tonnes and the huge return yoke has a mass of 10 000 tonnes. The 4 Tesla field provides the large bending power that makes it possible to accurately determine the momentum of highly energetic charged particles.

Inner tracker

The inner tracker is the innermost detector layer. Its purpose is to provide precise measurements of the tracks of charged particles emerging from the LHC collisions and to precisely reconstruct primary and secondary vertices. The inner tracker volume has a length of 5.8 m and a diameter of 2.6 m. The CMS solenoid provides a homogeneous magnetic field of 4 T over the entire volume.

The inner tracker consists of 10 layers of silicon microstrip detectors and an additional 3 layers of silicon pixel detectors close to the interaction region. The pixel detectors improve the measurements of the position of primary and secondary vertices and impact parameter of charged-particle tracks.

Electromagnetic Calorimeter

The electromagnetic calorimeter or ECAL is a lead tungstate crystal calorimeter and is the second main detector layer. The scintillation light that is produced in the crystals is detected by avalanche photodiodes (APD's) in the barrel and by vacuum phototriodes (VPT's) in the endcaps. These photodetectors are fast, are able to operate in the very high magnetic field and can withstand large amounts of radiation.

An electromagnetic calorimeter is designed to absorb the full electromagnetic shower. Very few electrons or photons manage to pass through the ECAL, whereas muons and hadrons can easily get through. One of the most important criteria in the design of the ECAL, is the resolution needed to detect the decay of the Higgs boson to two photons.

In between the inner tracker system and the ECAL, the preshower detector is located. The preshower detector is a sampling calorimeter whose main function is to identify neutral pions and to help determine the position of electrons and photons.

Hadronic Calorimeter

The hadronic calorimeter (HCAL) is located in between the ECAL and the inner part of the magnet coil. Its purpose is to measure the energy of hadrons. The main part of the HCAL is a brass/scintillator sampling calorimeter with coverage up to $\eta < 3.0$. The scintillation light coming from the calorimeter is converted by wavelength-shifting fibres and detected by hybrid photodiodes (HPD's). This restricts the amount of material that can be used to stop particles. Therefore, a second, outer hadronic calorimeter is placed outside the magnet coil. This *tail-catcher* is designed to absorb any residual parts of the hadronic shower.

To further extend the coverage (up to $\eta < 5.0$), a forward iron/quartz-fibre calorimeter was also installed. This addition ensures full geometric coverage, which is needed for the measurement of the (missing) transverse energy in the event.

Muon system

As is evident from its name, the detection of muons is of central importance to CMS. Many signatures of new physics involve muons and in addition, muons are relatively easy to detect. A high performance muon system thus allows one to easily distinguish interesting processes above the enormous background rate.

Muons are the only particles that manage to get through all detector layers described above. Dedicated muon detectors are thus located outside the HCAL and solenoid. Their functions are muon identification, muon momentum measurements and triggering.

Due to the large return field of the CMS magnet, it was possible to install four muon stations in between the layers of the magnet return yoke. Each station consists of several layers of aluminium drift tubes (DT's) in the barrel region and cathode strip chambers (CSC's) in the endcaps. These are complemented with resistive plate chambers (RPC's) that are mainly used for triggering purposes.

3.2.2 Trigger and data acquisition system

The LHC provides collisions at a very high rate. At design luminosity there will be a bunch crossing every 25 ns, with up to 20 proton-proton interactions per crossing. It is impossible to process and store all the data that is produced in these collisions. It is therefore important to have a trigger system to reduce the rate. This trigger has to be designed in such a way that as many interesting events as possible are selected, while at the same time efficiently rejecting the non-interesting events.

The CMS trigger system consists of two stages. The Level 1 Trigger (L1) consists of custom built electronics (placed on trigger boards) and is located partly on the detector and partly in an

underground control room 90m from the experimental cavern. It reduces the rate from 40 MHz to about 100 kHz. The L1 trigger needs to make a decision (keep the event or not) very fast, in 3.2μs. This means that there is no time to do complicated calculations. Also tracker information cannot be used because of the large number of output channels. Information from the calorimeters and the muon system are sent from the detector readout to the trigger boards. Meanwhile all the output of the event is held in buffer pipelines in the front-end electronics. When the L1 trigger has made its decision, this output is either discarded or sent onwards towards the switchboard.

The switchboard assembles the information from the different subdetectors into a complete event and sends this data to the High Level Trigger (HLT). It is also in charge of transporting the information from the experimental cavern to the surface. The HLT is a software system running on a filter farm of a few thousand commercial processors. It has access to all the readout information and can perform more complicated algorithms. The specific algorithms that are used, change over time to take into account the changes in luminosity and analysis strategy. The HLT further reduces the rate to about 100 Hz. The remaining events are stored and can be used for offline analysis.

3.2.3 CMS Offline computing

Once the trigger system has decided which events to keep, they have to be reconstructed and analyzed. The raw detector data needs to be converted to understandable physics objects. Reconstruction is the most CPU-intensive activity in the data processing chain. One needs to reconstruct vertices, run track finding algorithms, identify electrons, muons and photons, reconstruct jets, apply b-tagging algorithms, run detector-specific filtering and cleaning algorithms and so on.

The CMS software framework that does all that is called CMSSW. CMSSW is based on the Event Data Model (EDM), which is centered around the concept of an Event. An Event is essentially a C++ object container for all raw and reconstructed data related to a certain collision. These events are stored in ROOT files [24]. The CMSSW event processing model consists of one executable, called cmsRun, and many plug-in modules. These modules contain all the code needed in the event processing, such as calibration and reconstruction algorithms.

To do an analysis, we also need Monte Carlo (MC) simulations of various Standard Model processes and new physics models. Monte Carlo events are generated at parton level. Then showering and hadronization is applied and finally these events are put through the `Geant4` [25] based CMS detector simulation. The result is data similar to what one obtains from the actual detector. This MC data can then also be reconstructed as if it were actual detector data.

CMSSW is designed in such a way that one can use the same CMSSW executable for both detector and Monte Carlo data. To run an analysis, one has to specify the data and the modules that should be used. This can be done by using the Physics Analysis Toolkit (PAT) [26]. This is a high-level analysis layer providing easy access to many algorithms in the CMSSW framework. Most physics analyses are done using PAT. For this thesis we will also use The NtupleMaker [27] tool. The NTupleMaker creates flat ntuples from data in EDM format. One can specify what parts of the event information should be kept in order to reduce the size of the output files. For the further analysis of the resulting ntuples, the only prerequisite is ROOT, although they can also be used within the CMSSW framework.

In the next paragraphs we will briefly discuss the jet reconstruction and b-tagging algorithms that are used in this document.

Jet reconstruction When high energy partons are produced in a collision, they will undergo fragmentation. This will lead to a collimated bunch of hadrons, the so-called jet. The goal of a jet reconstruction algorithm is to define a consistent way to cluster particles into bunches of collimated particles. Electrons and photons, which leave energy deposits in the ECAL, will generally also be reconstructed as jets.

A jet reconstruction algorithm should be infrared and collinear safe. This means that soft radiation or collinear splitting of the parton should not alter the jet collection. Sequential recombination algorithms fulfill these requirements. This type of algorithm mimics a walk back through the QCD fragmentation by repeatedly recombining two particles or energy deposits into one new particle. The question that needs to be answered is how to decide which particles to combine and when to stop the recombination. We will illustrate this with the so-called anti-kt algorithm [28].

The algorithm starts by calculating the distance d_{ij} between all particles i and j and the distance d_{iB} between all particles i and the beam. For the anti-kt algorithm these distances are given by

$$d_{ij} = \min\left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2}\right)\frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{Ti}^2}\ , \tag{3.1}$$

with $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i + \phi_j)^2$ (y is the rapidity). R is called the jet radius and is a measure of the angular size of the jet.

Next we need to find the smallest value among these distances. If it is a d_{ij}, particles i and j are recombined and the distances are recalculated. If it is a d_{iB}, i is assumed to be a jet and is removed from the list. Then the algorithm is repeated from the beginning. This continues until there are no particles left. At that point the final list of jets is constructed.

Particle Flow Particle Flow (PF) is a technique to reconstruct all stable particles (electrons, muons, photons, charged and neutral hadrons) in the event by combining information from all detector parts [29]. PF combines the charged particle tracks with information from the calorimeters and the muon system. This allows for a more accurate description of the event. The resulting list of particles can then be used to build jets, calculate missing transverse energy et cetera. With PF it is possible to first remove the (isolated) electrons and muons from the list, before proceeding with the jet clustering. In this way overlap between jets and electrons (or muons) is reduced to a minimum.

b-tag algorithms Many signals of new physics can have b-rich final states, while for most Standard Model backgrounds (except top quark production) b-jets are not so common. It is therefore important to be able to distinguish these b-jets from other (lighter flavour) jets. The tagging of b-jets mainly relies on specific properties of B-hadrons such as their large lifetime and mass, their decays to final states with high charged track multiplicities and their relatively large semileptonic branching ratios.

The b-tagger that will be used in this thesis, is within CMS known as the `simpleSecondaryVertexHighEffBJetTags` tagger [30]. It is based on the reconstruction of the secondary vertex corresponding to the decay of the B-hadron and the associated decay length that can be inferred from the location of the secondary vertex relative to the primary vertex.

Chapter 4

sNavigator: a tool to optimize the event selection for searches for New Physics with a multidimensional parameter space

4.1 Introduction

With the help of experiments performed during the past 40 years, we have managed to narrow down the 19-dimensional Standard Model parameter space to essentially one point. This is in sharp contrast with the status of New Physics phenomena such as supersymmetry. At this point, we know very little about SUSY, we do not even know whether or not SUSY exists. As a result, the multidimensional SUSY parameter space is still very large. We only have a very vague idea about how SUSY looks like because this strongly depends on the location in SUSY space.

This lack of knowledge makes searching for supersymmetry rather difficult. In a standard analysis, one devises an event selection in order to remove as much background as possible, while retaining the signal one is looking for. In the case of supersymmetry, this is not directly possible as we do not know how the signal looks like.

Up till now, SUSY searches performed at CMS have not done much optimization but rather rely on accurate knowledge of the detector performance. In these analyses a baseline selection that defines the search topology is applied, plus some additional cuts to remove specific Standard Model backgrounds. Then a search region is defined, based on some general predictions. The exact boundaries of this search region are, however, chosen rather arbitrarily.

Once the selection is decided upon, the expected number of background events in the search region is calculated. At this point the selection is usually also applied to a few points in the SUSY space – the benchmark points – to see the effect. In addition one or more planes in the parameter space of a certain model can be chosen and expected exclusion lines based on Monte Carlo simulation can be plotted. In this context we can also note the recent introduction of so-called simplified models [31]. These models are designed to involve only few particles and interactions and can be viewed as limits of more general New Physics models. Simplified models are characterized mainly by their topology and can be used to express the sensitivity of an analysis in a more model-independent way.

Finally the chosen selection will of course also be applied to the actual data in the hope of seeing an excess of events in the search region.

This type of search is obviously not the most efficient way to search for a particular model of New Physics. If your goal is to discover supersymmetry, you can try to optimize your event selection specifically for supersymmetry. Optimizing for a certain point in the parameter space is relatively straight-forward. This is however far from optimal. There is no guarantee that a selection which is optimal for one point in the parameter space, will also be optimal for other points. This absence

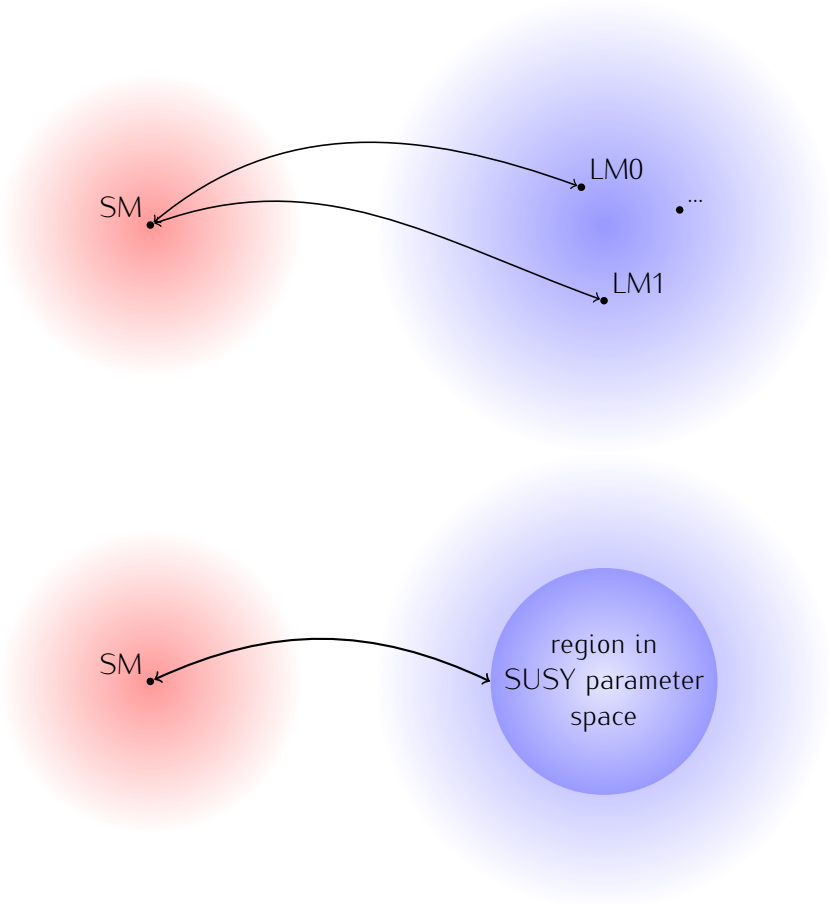


Figure 4.1: Optimizing an event selection for a benchmark point in SUSY space (top figure) or for a SUSY parameter space region (bottom figure) w.r.t. Standard Model (SM) background processes

of a clear-cut way to optimize, is the reason why searches performed so far have not done it. The search has been kept as broad as possible in order not to miss possible signals due to finetuning of the selection for one specific signal. This, however, also means that searches are not as sensitive to certain New Physics models as they could be.

We propose an alternative and complementary approach. We will devise a method to find a selection that is optimized for a specific multidimensional New Physics model, such as supersymmetry. Instead of taking into account only a few points in the parameter space, we want to do the optimization for a region in the parameter space of the chosen model. This concept is illustrated in figure 4.1. In addition, to perform the search as efficiently as possible, we also want to take into account results of previous measurements. It does not make much sense to optimize for a parameter space region which has already been excluded by other analyses. Regions in the parameter space that are more likely should contribute with a larger weight. The used analyses can be other SUSY searches performed at CMS, but also many other measurements contain information about the nature of SUSY. Several of these measurements were described in section 2.6.

This thesis is an effort towards answering the question "How to optimize for a parameter space region?" We can summarize the goal as follows.

> *The goal is to devise a method to find an event selection that* ***maximizes the probability of discovery*** *in a parameter space region of a given New Physics model. This discovery probability should encompass the knowledge obtained by previous measurements.*

A way to include this prior knowledge in a proper way, is to use Bayesian statistics [32] [33]. We

will introduce the concept of the *discovery probability* and we will propose a method to efficiently find an event selection that maximizes this quantity.

In the next section we will first recall a few basic elements of Bayesian statistics. After that we will give a detailed description of the method we constructed.

4.2 Intermezzo: Bayesian statistics

The goal of most scientists is to conduct experiments that yield data from which we can *infer* something interesting about the phenomenon under study. The Bayesian approach to statistical inference is based on the concept of "subjective" probability, a measure of belief that something is true. This is different from the so-called frequentist approach in which probability is just a relative frequency. Observations are viewed as random draws from some ensemble of possible observations such that the probability of a certain observation is the frequency one would observe when repeating the measurement a large (in principle infinite) number of times.

Consider a theory one wishes to test and data obtained in an experiment. For frequentists only a statement of the form

$$P(\text{Data}|\text{Theory}) \tag{4.1}$$

makes sense. They only calculate the probability to observe a certain data set, given a certain theory. By contrast, the Bayesian approach also deals with statements of the form

$$P(\text{Theory}|\text{Data}) \ , \tag{4.2}$$

the probability that a given theory is true, given the observed data. This is exactly the kind of statement we want to make. Given the data provided by experiments, we want to exclude or confirm a theory, in our case supersymmetry.

We can link these two statements using Bayes' theorem

$$P(\text{Theory}|\text{Data}) = \frac{P(\text{Data}|\text{Theory})P(\text{Theory})}{P(\text{Data})} \tag{4.3}$$

$P(\text{Theory})$ is called the prior probability. Everything we believe to know about the theory independently of the data is included in this prior. The probability $P(\text{Data}|\text{Theory})$ is often referred to as the likelihood – for counting experiments in particle physics this is a Poissonian likelihood – and $P(\text{Theory}|\text{Data})$ is called the posterior probability.

The power of the Bayesian approach lies for the most part in the fact that one can meaningfully speak about the probability of a theory or hypothesis, no matter what this might be. Another big advantage is the possibility to easily incorporate more information, once it becomes available. To this end, we can use Bayes' theorem in an iterative way. The posterior of the current step becomes the prior of the next step.

The problem that remains, is how to choose the initial prior. If the only thing one knows about the theory, is that it can be described by a parameter θ in some range, one can for example attribute the same probability $\pi(\theta)$ to all values in this range. We call this a flat prior. But this necessarily means that we have also assigned a prior to all functions of this parameter θ. In most cases this prior will not be flat. The question that arises naturally is to which variable we should assign the flat prior. Depending on the choice, we will obtain different results. In its essence, this is not a bad thing. By using different priors and checking how much the results change, one obtains information about the usefulness of the data for the problem at hand. If the results depend strongly on the prior, one should try to acquire more data.

4.3 Description of the method

4.3.1 Probability of discovery

As stated in the introduction, we define our optimal selection as the selection that maximizes the probability of discovery. First we must state precisely what we mean by this.

We define the probability of discovery P_{disc} of a certain parameter space region R for a given event selection as

$$P_{disc} = \int_R p(\mathcal{S} > 3|\theta)p(\theta)d\theta \ , \tag{4.4}$$

with θ the representation of a parameter set of the given physics model, hereafter referred to as a point in the parameter space, and $\mathcal{S}$ the significance given a certain event selection. We choose as measure of discovery

$$p(\mathcal{S} > 3|\theta) = \begin{cases} 0 & \text{if } \mathcal{S} = \frac{S_\theta}{\sqrt{S_\theta + B}} \leq 3 \\ 1 & \text{if } \mathcal{S} = \frac{S_\theta}{\sqrt{S_\theta + B}} > 3 \end{cases} \ , \tag{4.5}$$

with S_θ and B the number of signal and background events passing a certain event selection. Of course, this measure of discovery can be adapted at will. $p(\theta)$ is the posterior probability of parameter point θ and we use Bayes' theorem to include the information of previous measurements. We can include both CMS analyses and LEO's (see section 2.6).

$$p(\theta) \equiv p(\theta|LEO, \text{CMS measurements}, \ldots) \tag{4.6}$$

$$\propto p(LEO|\theta)p(\text{CMS measurements}|\theta) \ldots p_0(\theta) \ , \tag{4.7}$$

where the ellipsis stand for possible other measurements that can be included and where $p_0(\theta)$ is the prior probability, which we choose to be flat ($p_0(\theta) = 1$).

Substituting (4.7) in our definition of discovery probability, we find

$$P_{disc} = \int_R p(\mathcal{S} > 3|\theta)p(LEO|\theta)p(\text{CMS measurements}|\theta) \ldots d\theta \ . \tag{4.8}$$

This is the quantity we wish to maximize in our procedure. In other words, we want to find the event selection that maximizes P_{disc}.

In practice we will not integrate directly over our chosen parameter space. Instead we will sample points from this parameter space via a Markov Chain Monte Carlo method [34]. We will do this in an efficient way and sample according to the density $p(LEO|\theta)$. Then we can replace the integral by a summation over the sampled points.

$$P_{disc} = \sum_\theta p(\mathcal{S} > 3|\theta)p(\text{CMS measurements}|\theta) \ldots \tag{4.9}$$

We can interpret this sum as the weighted number of parameter points that can be discovered given a certain event selection.

In the next paragraphs we will discuss in more detail how the likelihoods $p(LEO|\theta)$ and $p(\text{CMS measurements}|\theta)$ are calculated.

Likelihood for low energy observables Several low energy observables were discussed in section 2.6. The total likelihood for these LEO's is given by the product of the likelihood for each included observable. Depending on the type of observable and its experimental status, the likelihood takes a different form [35] [36].

Table 4.1: The low energy observables that were included in the calculation of $p(LEO|\theta)$

i	Observable	Limit	Likelihood function
1	Δa_μ	$(28.7 \pm 8.0) \times 10^{-10}$ $[e^+e^-]$ $(19.5 \pm 8.3) \times 10^{-10}$ [taus]	Weighted Gaussian average
2	$BR(b \to s\gamma)$	$(3.55 \pm 0.34) \times 10^{-4}$	Gaussian
3	$BR(B_s \to \mu\mu)$	$\leq 4.7 \times 10^{-8}$	Upper limit
4	$R(B_u \to \tau\nu)$	1.66 ± 0.54	Gaussian
5	m_t	173 ± 1.1	Gaussian
6	$m_b(m_b)$	$4.19^{+0.18}_{-0.06}$	Two-sided Gaussian
7	$\alpha_s(M_Z)$	0.117 ± 0.002	Gaussian
8	m_h	LEP & Tevatron (`HiggsBounds`)	$L_8 = 1$ if allowed. $L_8 = 10^{-9}$ if $m_h^{'}$ sampled from $Gauss(m_h, 1.5)$ is excluded.
9	SUSY mass	LEP & Tevatron (`micrOMEGAs`)	$L_9 = 1$ if allowed $L_9 = 10^{-9}$ if excluded

For observables for which there exists an actual measurement, a Gaussian likelihood is used. If we denote the experimental value of an observable as x_0, the uncertainty on this measurement as σ_x and the prediction for the observable given the parameters of the model as x, the likelihood has the form

$$\mathcal{L}_{Gauss} = \exp\left(-\frac{(x-x_0)^2}{2\sigma_x^2}\right) \tag{4.10}$$

If the uncertainties are quoted with different upper and lower limits, one uses a two-sided Gaussian.

In case there exists only an 95% exclusion upper limit x_0 for an observable, the likelihood takes the form

$$\mathcal{L}_{UL} = \frac{1}{1+\exp\left(\frac{x-x_0}{\Delta x}\right)}\ , \tag{4.11}$$

where Δx is taken to be around 1% of the value of the upper limit to mimic the 95% exclusion limit.

In the calculation of $p(LEO|\theta)$, we included the muon anomalous magnetic moment, the mass of the top quark, the branching ratio $b \to s\gamma$, the branching ratio $B_u \to \tau\nu$ and the strong coupling constant at the Z-mass. These observables have a Gaussian likelihood. We also included the mass of the bottom quark with a two-sided Gaussian likelihood and the branching ratio of $B_s \to \mu^+\mu^-$ with a likelihood of the form $\mathcal{L}_{UL}$.

Finally we also included results from direct searches. For the limits on the higgs boson mass m_h from LEP and Tevatron results, we draw a value m_h' from a Gaussian with mean $m_h(\theta)$ and width 1.5. If m_h' has been excluded, we use as likelihood $\mathcal{L} = 10^{-9}$, otherwise we put $\mathcal{L} = 1$. As we will consider supersymmetry, the limits on the masses of SUSY particles are also included. If the given model gives rise to masses that have been excluded, the likelihood is $\mathcal{L} = 10^{-9}$, otherwise $\mathcal{L} = 1$.

The included low energy observables are summarized in table 4.1. We have opted not to use the WMAP measurements, since the resulting limit is quite stringent and the nature of the cosmological model is not certain.

Likelihood for CMS analyses The likelihood for a (cut-and-count) CMS analysis in its simplest form is just a Poissonian likelihood:

$$p(\text{CMS measurement}|\theta) = Poisson(N|s(\theta), b) = \frac{e^{-(s(\theta)+b)}(s(\theta)+b)^N}{N!} \ , \tag{4.12}$$

with N the number of observed events in the analysis and $s(\theta)$, b the number of expected signal and background events. N and b are taken directly from the results of the CMS analysis. In order to calculate $s(\theta)$, we will need to generate and simulate events for the sampled parameter point θ. Then we will implement exactly the same selection as in the considered analysis and count how many signal events pass this selection. This form of the likelihood will be used in chapter 8 during the validation of the sNavigator method.

We can, however, do better and also include uncertainties on the background prediction [37] [38]. Let us consider the following situation: we have two regions, the region of interest and a control region that provides the background estimation. The contribution of signal events in the control region is taken to be neglibible. Assume next that the number of events in the control region, q, can be related to the number of background events in the signal region, b, by a known scale factor k such that $q = kb$. The probability of observing Q events if one expects kb events is given by

$$P(Q|b) = Poisson(Q|b) = \frac{e^{-kb}(kb)^Q}{Q!} \tag{4.13}$$

This additional information can be included in the likelihood:

$$P(N, Q|s(\theta), b) = Poisson(N|s(\theta), b)Poisson(Q|b) \tag{4.14}$$

$$= \frac{e^{-(s(\theta)+b)}(s(\theta)+b)^N}{N!}\frac{e^{-kb}(kb)^Q}{Q!} \tag{4.15}$$

Observing Q events in the control region also allows to make an estimate of the number of background events in the signal region: $\widehat{b} = B = Q/k$. There is also an associated uncertainty of $\delta B = \sqrt{Q}/k$. We can now reverse this logic and assume that for analysis i, we have obtained a background estimate in some way. B_i and δB_i are thus known, while Q_i and k_i are not. We can still use the above formula as we can compute Q_i and k_i from B_i and δB_i:

$$Q_i = \left(\frac{B_i}{\delta B_i}\right)^2 \ , \qquad k_i = \frac{B_i}{\delta B_i^2} \tag{4.16}$$

As total likelihood for several cut-and-count analyses (i) we thus find:

$$p(\text{CMS measurements}|\theta) = \prod_i \frac{e^{-(s_i(\theta)+b_i)}(s_i(\theta)+b_i)^{N_i}}{N_i!}\frac{e^{-k_i b_i}(k_i b_i)^{Q_i}}{\Gamma(Q_i+1)} \ , \tag{4.17}$$

where i runs over the number of included measurements, N_i is the number of observed events for analysis i and s_i and b_i are the number of expected signal and background events. N_i, B_i and δB_i are taken directly from the CMS analyses. $s_i(\theta)$ is calculated by applying the corresponding selections to the events of parameter point θ and b_i is put equal to its maximum likelihood estimator B_i. As Q_i can be non-integer, we have replaced $Q_i!$ by the Gamma function $\Gamma(Q_i+1)$.

This more complete form of the likelihood will be used for the case study in chapter 9. There, we will also list the analyses for which we calculated this likelihood and which values we used for N_i, B_i and δB_i, see table 9.1.

4.3.2 Maximizing the discovery probability

Now that we have defined what we mean by the discovery probability, we need to construct an algorithm to find the event selection that maximizes this quantity.

In general, an event can be described by a set of variables $\vec{x} = (x, y, \ldots)$. The way we define an event selection is by placing rectangular cuts $x > x_0, y > y_0, \ldots$. We will call $\vec{x_0} = (x_0, y_0, \ldots)$ a cut point. The goal is to find the event selection, or equivalently the cut point, that maximizes the discovery probability. We will do this by using the principle of the Random Grid Search (RGS) [39]. The main idea of RGS is to consider the signal events as possible cut points. A more detailed explanation can be found in chapter 6, where we apply the RGS principle in the optimization for a single point of the chosen parameter space.

In order to find the optimal event selection, we will first need to generate events for the sampled points in the chosen parameter space region. Then these events need to pass through a detector simulation. We will use the Delphes Fast Simulator for this (see chapter 7 for more details). In addition we also need events for various Standard Model background processes. The background samples we used, were produced centrally within the CMS collaboration and are described in section 5.1.

Before proceeding with the actual event selection optimization, we will first calculate the likelihood (4.17) for each sampled point of the parameter space. Then we will apply a preselection to our signal and background samples in order to define a search topology.

To find the optimal event selection for the chosen parameter space region, we have implemented three methods. An explanation of these methods is given below. For the application and comparison of the methods, we refer to chapter 8.

Method A In a first step of this method we use the RGS principle to obtain the best cut point for each sampled point in our parameter space. The definition of 'best' can be adapted at will; we define the best cut point for a certain parameter point as that selection that gives us the largest significance $\mathcal{S}$. The significance is defined as

$$\mathcal{S} = \frac{S}{\sqrt{S+B}}\,, \tag{4.18}$$

with S, B the number of signal and background events passing the cut. So for each parameter point, we pick the cut point that maximizes the significance. As possible cut points we only consider the signal events corresponding to that particular parameter point.

In a second step, we extract the optimal cut point from these 'best' cut points. With optimal cut point we mean that point that maximizes the discovery probability (equation 4.9) for the chosen parameter space region.

Method B In this method we skip the first step of method A and consider all events from all parameter points as possible cut points. For each of these cut points we calculate P_{disc}. The optimal cut point is the one that gives the largest value for P_{disc}.

This method will always give a better or equally good result as method A. The reason for this is the fact that the cut points considered in method A are a subset of those used in this method. As there are more cut points to choose from, there is a chance that we will find a better one. However, this method has the disadvantage of being quite a lot slower than method A. It will therefore be important to investigate whether or not method A yields similar results as method B.

Method C As a third method and to serve as an additional check, we implemented a method based on the Bayesian principle "If you do not know anything about a quantity, average over it." We will combine all events from all sampled parameter points into one large signal sample. Then we use

the RGS principle as if we would optimize for a single parameter point. To pick the best cut point we need to compute the number of signal and background events. For the signal events, we use a weighting according to the cross section and likelihood of the corresponding parameter point.

$$w = \left(\frac{\sigma_\theta L}{n_{\text{gen}}} \right) \frac{p(\text{CMS measurements}|\theta)}{\sum_\theta p(\text{CMS measurements}|\theta)}, \tag{4.19}$$

with σ_θ the cross section of parameter point θ, L the luminosity and n_{gen} the number of generated events per parameter point. We then just pick that cut point that gives us the largest significance.

To compare with methods A and B, we will also compute the discovery probability for the optimal cut point as determined by this method. We can also note that method C is slower than method A, especially when a large amount of parameter points are included.

Chapter 5

Monte Carlo samples and physics object definitions

5.1 Monte Carlo samples

In order to execute the sNavigator procedure, we need Monte Carlo (MC) data for various Standard Model background processes. We list the MC samples that we will use in the next chapters in table 5.1. In this table we have included the process, dataset name as used in CMS and cross section.

We will also need MC data for the signal samples. We will use the mSUGRA model (see section 2.4.3) as signal. Within CMS there is a list of low mass mSUGRA benchmark points (LM points) which exhibit different final states. For these points, MC samples using the full CMS detector simulation are available. As our chosen topology includes b-jets, we will use two benchmark points with b-rich final states, LM9 and LM13. These points correspond to the following mSUGRA parameters:

$$\text{LM9}: \quad m_0 = 1450, \quad m_{1/2} = 175, \quad A_0 = 0, \quad \tan\beta = 50, \quad \mu > 0 \tag{5.1}$$

$$\text{LM13}: \quad m_0 = 270, \quad m_{1/2} = 218, \quad A_0 = -553, \quad \tan\beta = 40, \quad \mu > 0 \tag{5.2}$$

LM9 has heavy squarks with masses ranging from 0.9 to 1.5 TeV, and a much lighter gluino with mass 487.7 GeV. As a result gluino production is dominant. The gluinos decay to 3-body states via a virtual squark. 45% of the decays are in the form of $\widetilde{\chi}_i^0 b\overline{b}$.

LM13 also has a light gluino with mass 543.1 GeV. In this case it is accompanied by a lighter squark spectrum with a mass range of 297 to 561 GeV. Of these squarks, only the sbottoms and stops are lighter than the gluino. This implies that gluinos decay exclusively to two-body states involving sbottoms and stops accompanied with bottom and top quarks. Final states with b's can also arise from direct sbottom or stop production.

The dataset and cross section of these two points is shown in table 5.2.

All MC datasets in tables 5.1 and 5.2 were produced centrally within the CMS collaboration. For the SM backgrounds, the event generation at parton level was done by `MadGraph` [40], the showering and hadronization by `PYTHIA6` [41]. The SUSY signals were completely generated by `PYTHIA6`. The data samples were then further processed with PAT using the special `SUSYPAT V9` recipe, which works with the CMS software release `CMSSW_3_8_7` (see section 3.2.3). Finally, the NtupleMaker tool was run over the data samples to create flat ntuples. These flat ntuples were then used for the analysis.

Of course, to execute the sNavigator procedure, we will need to generate events for a lot more SUSY points than just LM9 and LM13. As explained in section 4.3.1, we will sample points from the four-dimensional (we choose $\mu > 0$) mSUGRA parameter space. We will do this via a Markov Chain Monte Carlo method according to the likelihood of the low energy observables.

Table 5.1: Monte Carlo samples for the Standard Model backgrounds that were used throughout this thesis.

Process	Dataset	Cross section (pb)
$t\bar{t}$ + jets inclusive	/TTJets_TuneD6T_7TeV-madgraph-tauola/Fall10-START38_V12-v2/AODSIM	165
Single top, s-channel	/TToBLNu_TuneZ2_s-channel_7TeV-madgraph/Fall10-START38_V12-v1/AODSIM	1.492332
Single top, t-channel	/TToBLNu_TuneZ2_t-channel_7TeV-madgraph/Fall10-START38_V12-v1/AODSIM	20.957532
Single top, tW-channel	/TToBLNu_TuneZ2_tW-channel_7TeV-madgraph/Fall10-START38_V12-v1/AODSIM	10.6
QCD, $50 < HT < 100$	/QCD_TuneD6T_HT-50To100_7TeV-madgraph/Fall10-START38_V12-v1/AODSIM	$3 \cdot 10^7$
QCD, $100 < HT < 250$	/QCD_TuneD6T_HT-100To250_7TeV-madgraph/Fall10-START38_V12-v1/AODSIM	$7 \cdot 10^6$
QCD, $250 < HT < 500$	/QCD_TuneD6T_HT-250To500_7TeV-madgraph/Fall10-START38_V12-v1/AODSIM	$17.1 \cdot 10^3$
QCD, $500 < HT < 1000$	/QCD_TuneD6T_HT-500To1000_7TeV-madgraph/Fall10-START38_V12-v1/AODSIM	$5.2 \cdot 10^3$
QCD, $1000 < HT$	/QCD_TuneD6T_HT-1000_7TeV-madgraph/Fall10-START38_V12-v1/AODSIM	83
$Z/\gamma(\rightarrow ll)$ + jets, $m(ll) > 50$	/DYJetsToLL_TuneD6T_M-50_7TeV-madgraph-tauola/Fall10-START38_V12-v2/AODSIM	3048
$Z(\rightarrow \nu\nu)$ + jets	/ZinvisibleJets_7TeV-madgraph/Fall10-START38_V12-v1/AODSIM	4500
$W(\rightarrow l\nu)$ + jets	/WJetsToLNu_TuneZ2_7TeV-madgraph-tauola/Fall10-START38_V12-v1/AODSIM	31314

Table 5.2: Monte Carlo samples for the SUSY signals that were used throughout this thesis.

Process	Dataset	Cross section (pb)
SUSY LM9	/LM9_SUSY_sftsht_7TeV-pythia6/Fall10-START38_V12-v1/AODSIM	7.134
SUSY LM13	/LM13_SUSY_sftsht_7TeV-pythia6/Fall10-START38_V12-v1/AODSIM	6.899

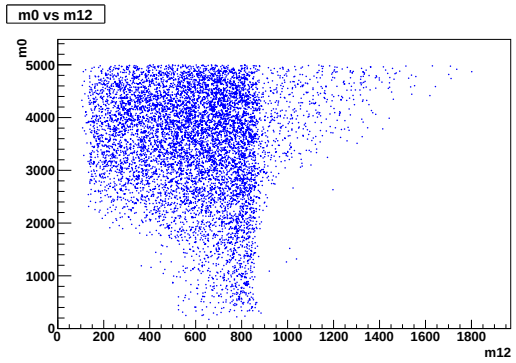


Figure 5.1: m_0 (GeV) versus m_{12} (GeV) for all 8000 SUSY points

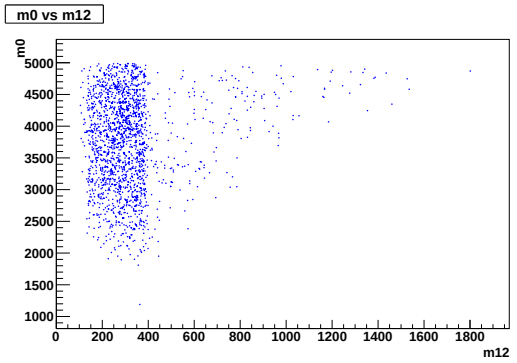


Figure 5.2: m_0 (GeV) versus m_{12} (GeV) for the subset of SUSY points with $\sigma > 0.1$ pb

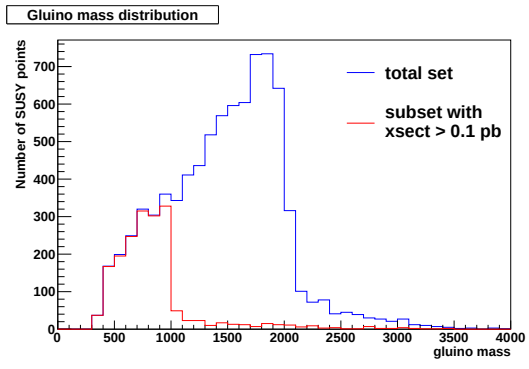


Figure 5.3: Gluino mass (GeV) distribution

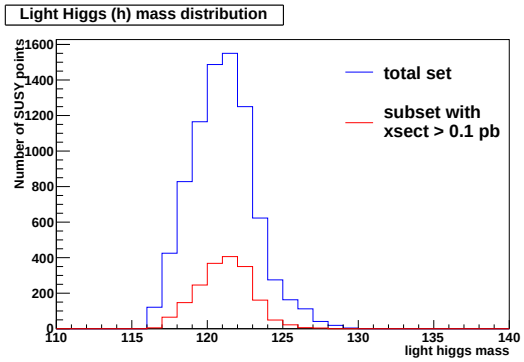


Figure 5.4: Light Higgs mass (GeV) distribution

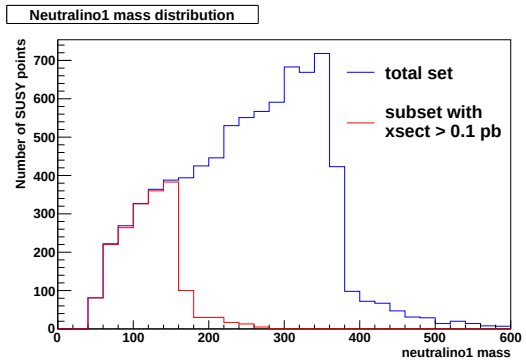


Figure 5.5: Lightst neutralino mass (GeV) distribution

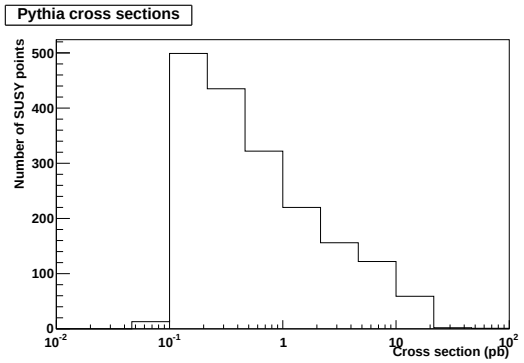


Figure 5.6: Distribution of cross sections for the SUSY points that will be used in the optimization

We sampled 8000 mSUGRA points with cross section (σ) larger than 1 fb. Information on the sampled points is written in `.slha` (SUSY Les Houches Accord) files [42]. These files contain the SUSY model parameters, the Standard Model parameters, the masses of all particles (obtained via `SOFTSUSY` [43]) and the branching ratios of the decays (obtained via `SUSY-HIT` [44]). Event generators are able to use these files to generate events. We have used `PYTHIA6` as event generator.

`PYTHIA6` is also able to calculate cross sections, of which the accuracy depends on the number of generated events. We first generated 100 events for these 8000 points to obtain their cross section. Only the subset of points with cross section larger than 0.1 pb will be used in the application of sNavigator.

In figures 5.1 and 5.2 we show the distribution of the SUSY points in the $m_0 - m_{12}$ plane for the full set and the subset with largest cross sections. To further characterize the SUSY points we show the distribution of the gluino, lightest higgs boson and lightest neutralino mass in figures 5.3, 5.4 and 5.5.

As a next step, we generated 1000 events for all 1829 points in the subset of points with the largest cross section. This also resulted in a more accurate value of the cross section, of which the distribution is shown in figure 5.6. The events for this set of SUSY points were then put through the Delphes detector simulation (see chapter 7). The Delphes output is an ntuple that can be analyzed in the same manner as the CMSSW ntuples which were created with the NtupleMaker.

5.2 Physics object definitions

In this section we will list our physics object definitions. Unless explicitly stated otherwise, these definitions are used throughout this document.

5.2.1 Jets

Jets are defined to be Particle Flow (PF) jets (collection `selectedPatJetsPF`) with jet energy corrections applied up to the `L3Absolute` level. Jet clustering was performed after the removal of all `selectedPatElectronsPF` and `selectedPatMuonsPF` (see below). In addition, jets need to fulfill the following requirements:

- $p_T > 30$ GeV
- $|\eta| < 2.4$
- `Loose` jet identification criteria as defined by the JetMET group [45]:
  - `neutralEmEnergyFraction` < 0.99
  - `numberOfDaughters` > 1
  - `chargedHadronEnergyFraction` > 0
  - `chargedMultiplicity` > 0
  - `chargedEmEnergyFraction` < 0.99
  - (`neutralHadronEnergy`+`HFHadronEnergy`)/`energy` < 0.99

This corresponds to the jet definition of the RA2b[1] analysis [46].

Throughout the following chapters we will regularly use the variable HT. This is defined as

$$HT = \sum_{jets} p_T \ . \tag{5.3}$$

5.2.2 b-jets

b-jets need to fulfill the jet requirements as stated above and need to be b-tagged. In accord with the RA2b analysis, we use the `simpleSecondaryVertexHighEffBJetTags` ($SSHVE$) b-tagger at its medium working point. This working point is designed to have a mis-tag rate of 1% and corresponds to $SSVHE_{disc} \geq 1.74$, with $SSVHE_{disc}$ the value returned by the discriminator algorithm.

5.2.3 Missing transverse energy (MET)

We use uncorrected Particle Flow MET (collection `patMETsPF`). This is defined as the opposite of the vector sum of the p_T's of all particles reconstructed by the Particle Flow algorithm.

[1] In CMS, SUSY analyses are divided into so-called "reference analyses" or RA's. Each reference analysis has a different search topology. The RA2b analysis has as topology jets + b-jet(s) + missing transverse energy.

5.2.4 Electrons

To define an isolated electron, we will use two definitions: a loose one, based on the RA2b analysis and a tighter one, based on the RA4[2] analysis. In both cases we use PF electrons (collection `selectedPatElectronsPF`). To define whether or not an electron is isolated, we will use the *relIso* variable, which is defined as the sum of all transverse energy deposits in the tracker, ECAL and HCAL within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.3$ around the electron, divided by the transverse momentum of the electron:

$$relIso = \frac{\sum_{\Delta R<0.3} E_T(ECAL) + \sum_{\Delta R<0.3} E_T(HCAL) + \sum_{\Delta R<0.3} p_T(tracker)}{p_T(electron)} \tag{5.4}$$

For the RA2b-inspired definition we require:

- $p_T > 8$ GeV
- $|\eta| < 2.5$
- $relIso < 0.2$

For the RA4-inspired definition we require:

- $p_T > 8$ GeV
- $|\eta| < 2.4$
- $|\eta|$ not in the gap between the ECAL barrel and endcap $[1.4442 - 1.566]$
- $relIso < 0.07$ in the barrel and $relIso < 0.06$ in the endcap

5.2.5 Muons

As for electrons, we will use two isolated muon definitions: a looser, RA2b inspired definition and a tighter RA4 inspired definition. We will use PF muons (collection `selectedPatMuonsPF`) in both cases. *relIso* for muons is defined in an analogous way as for electrons. The RA2b-inspired definition is given by:

- $p_T > 8$ GeV
- $|\eta| < 2.5$
- $relIso < 0.2$
- must pass the Muon ID criteria `AllGlobalMuons`

The RA4-inspired definition:

- $p_T > 8$ GeV
- $|\eta| < 2.4$
- $relIso < 0.1$
- impact parameter in the transverse plain w.r.t the beamspot `dB` < 0.02 cm
- number of reconstructed hits in the silicon tracker `innerTrack_numberOfValidHits` ≥ 11

[2]The RA4 analysis has as topology single lepton + jets + missing transverse momentum.

Chapter 6

Application of the Random Grid Search principle to optimize the event selection for the SUSY benchmark point LM9

In this chapter we will first discuss the Random Grid Search principle in a bit more detail. Then we will test the procedure in the optimization with respect to a single SUSY parameter point. This will allow us to acquire a thorough understanding of the RGS principle.

We will do the optimization for the b-jet(s) + lepton topology. Therefore, we will apply a pre-selection to our data samples so that only events of that topology remain. The actual optimization will be done for the SUSY benchmark point LM9 (see section 5.1). We will show the results of this optimization and the conclusions we can draw.

6.1 The Random Grid Search

A basic question in high energy physics is how to classify an event, described by a set of variables $\vec{x} = (x, y, \ldots)$, as a signal or background event. This decision should be made based on the values of the variables in $\vec{x}$. This is not a trivial task. For many variables, signal and background will look quite similar. A first thing to do is thus to choose variables that have some discriminating power between signal and background. One can then try to find a region that best isolates signal from background.

Over time, many approaches have been proposed [39]. The conventional approach, illustrated in figure 6.1, is to apply rectangular cuts, for example $x > x_0, y > y_0, \ldots$. The exact values of the cut point are often determined by eye and common sense. After a procedure of trial and error, one hopes to arrive at a cut point (or equivalently an event selection) that is close to optimal, thus keeping as much signal and rejecting as much background as possible. Since no real optimization is done, this is absolutely not guaranteed.

To improve on the conventional cut selection, one can systematically scan a grid of possible cut points, see figure 6.2, and choose the *best* one among these. One can for instance pick the cut point that maximizes the significance, the signal to background ratio... This technique is, however, not very efficient. A lot time can be wasted in regions where there are few signal or background events. Also the number of grid points to consider rapidly goes up with the number of dimensions of the cut points.

A more efficient version of the regular grid search is to choose the possible cut points according to the distribution of $\vec{x}$ for the signal. So we use the values of $x, y, \ldots$ for each signal event as possible cut points. In this way a grid is constructed with random spacing, hence the name 'Random Grid Search', or RGS. This is illustrated in figure 6.3. The resulting set of cut points will be called a *cut set*.

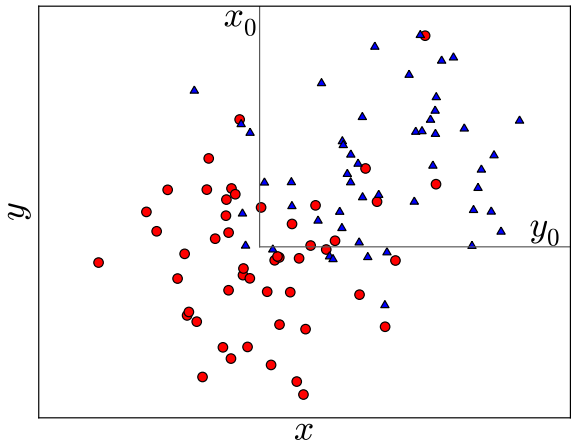


Figure 6.1: Conventional approach to placing rectangular cuts

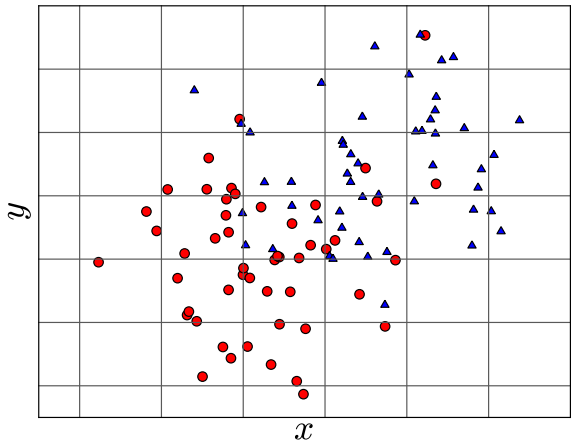


Figure 6.2: Regular Grid Search

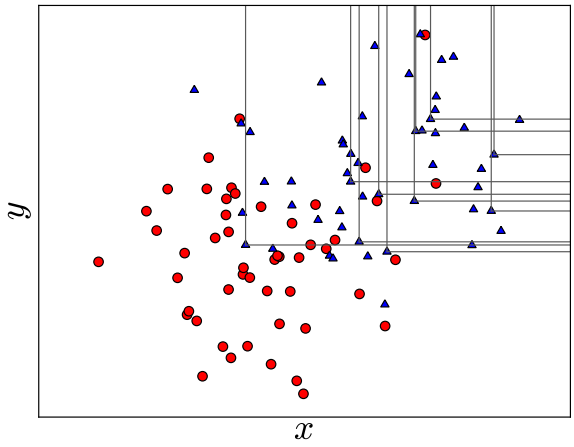


Figure 6.3: Random Grid Search

At the end of the procedure, we will again find an optimal cut point, with the big advantage that the method is relatively fast. It does not scale with the number of variables used for the cut point and only spends time in those places where there actually are signal events.

The RGS principle will be used in the next sections to find an optimal selection for the benchmark point LM9 in the b-jet(s) + lepton topology.

6.2 Preselection

We use the object definitions as explained in section 5.2. For the leptons we use the looser, RA2b-inspired definition. As preselection we require the presence of at least three jets, at least one b-tagged jet and exactly one isolated lepton (electron or muon).

We will use the Monte Carlo samples listed in tables 5.1 and 5.2. The efficiency of the preselection and the resulting effective cross section for these samples are shown in table 6.1. We see that the main background contributions come from $t\bar{t}$ and QCD.

Table 6.1: Efficiency of and effective cross section after the preselection

Sample	σ (pb)	Preselection efficiency	σ_{eff} (pb)
DYJetsToLL_D6T_M-50	3048	1.86e-4	0.57
QCD	3.7e7	2.18e-7	8.06
TTJets_D6T	165	9.65e-2	15.93
Single top	33	3.55e-2	1.17
WJetsToLNu_Z2	31314	7.55e-5	2.36
ZinvisibleJets_Z2	4500	4.59e-7	2.06e-3
LM9_Z2	7.134	2.04e-2	0.15
LM13_Z2	6.899	7.26e-2	0.50

6.3 Optimization results for benchmark point LM9

6.3.1 Cut variables

Before one can start the optimization procedure, one has to choose appropriate cut variables. The way to do that is, however, not straight-forward as we are dealing with multiple variables at the same time. The choice of variables will thus not only depend on their intrinsic discriminating power, but also on the amount of correlation between them. Among the considered cut variables are the missing transverse energy MET, the scalar sum of the p_T's of all jets HT, the number of jets n_{jets}, the number of b-tagged jets $n_{btaggedjets}$, the p_T of the jet with the highest and second highest p_T, the vectorial sum of all jet p_T's MHT, the p_T of the lepton, the transverse mass m_T of the lepton and the MET and the angle between the lepton and the MET.

To get an idea of the discriminating power of the potential cut variables, we plotted their distributions, after applying our preselection cuts, for the signal and background samples. The distributions were normalized to a luminosity of 300 pb^{-1} and are shown in the left-hand plots of figures 6.4 - 6.15.

To investigate the correlations we made two-dimensional density plots of these variables, again after applying the preselection criteria and normalized to a luminosity of 300 pb^{-1}. These plots are shown in figures 6.16 - 6.24 for the LM9 (left-hand side) and $t\bar{t}$ (right-hand side) samples.

From these figures, we can already deduce that the MET, HT and n_{jets} provide a good discriminating power between signal and background. These will thus be prime candidates for cut variables. The number of b-tagged jets and the transverse mass, however, do not have much discriminating

power. We can also note that the lepton p_T spectrum for the signal samples is softer than for $t\bar{t}$. This indicates that lepton p_T will not be a good candidate cut variable. From the density plots we can also conclude that HT is very correlated with the jet p_T's. As HT is the sum of all jet p_T's, this behavior is to be expected.

6.3.2 Optimization

Based on our findings, we proceed to perform the actual optimization. We will optimize for LM9 with respect to $t\bar{t}$, which is the largest background component. As cut variables we used, in various combinations, the HT, MET, n_{jets}, p_T of the jet with the highest p_T and the p_T of the jet with the second highest p_T. We also tried three different definitions of significance to see how this affects the optimization. We used the default definition and two varieties of the so-called Punzi significance [47].

$$\mathcal{S}_{default} = \frac{S}{\sqrt{S+B}} \tag{6.1}$$

$$\mathcal{S}_{Punzi} = \frac{S}{\frac{a}{2} + \sqrt{B}} \quad \text{with } a \text{ a constant} \tag{6.2}$$

In tables 6.3 - 6.5 we show the results of the optimization. For each choice of significance and combination of cut variables we list the optimal values for the cut variables, the significance and the number of signal and background events for a luminosity of 300 pb^{-1}.

From these tables we can conclude that HT is the variable with the best discriminating power. By adding the MET and jet multiplicity, we can improve the results. This makes sense since these variables separately also have a good discriminating power and are not too correlated with HT. One intuitively expects that adding more variables will always improve the results, because you have more options to define the search region. From our results we can see that this is not necessarily the case. Adding the jet p_T to the cut variables did not enlarge the significance one can achieve. This is probably due to the large correlation between jet p_T and HT.

6.3.3 Optimal selection

As a final step in this test of RGS, we can take a closer look at what happens when applying the optimal selection. As optimal set of cut variables, we use the set with cut variables HT, n_{jets} and MET from table 6.4:

- $HT \geq 559$ GeV
- $n_{jets} \geq 5$
- $MET \geq 66.8$ GeV.

This selection gives us a significance of 1.83 for 300 pb^{-1}. For 1 fb^{-1} this would become roughly larger than 3, while for 3 fb^{-1} this would become larger than 5, which means that this SUSY point would be within discovery reach.

In the right-hand plots of figures 6.4 - 6.15, the distributions of various quantities are plotted after applying these cuts. The corresponding effective cross section for all samples is shown in table 6.2. We can see that the contribution to the background coming from QCD and Z is reduced to a negligible level. Top processes remain as the main background source.

Table 6.2: Effective cross sections after applying the optimal selection ($HT \geq 559$, $n_{jets} \geq 5$, $MET \geq 66.8$)

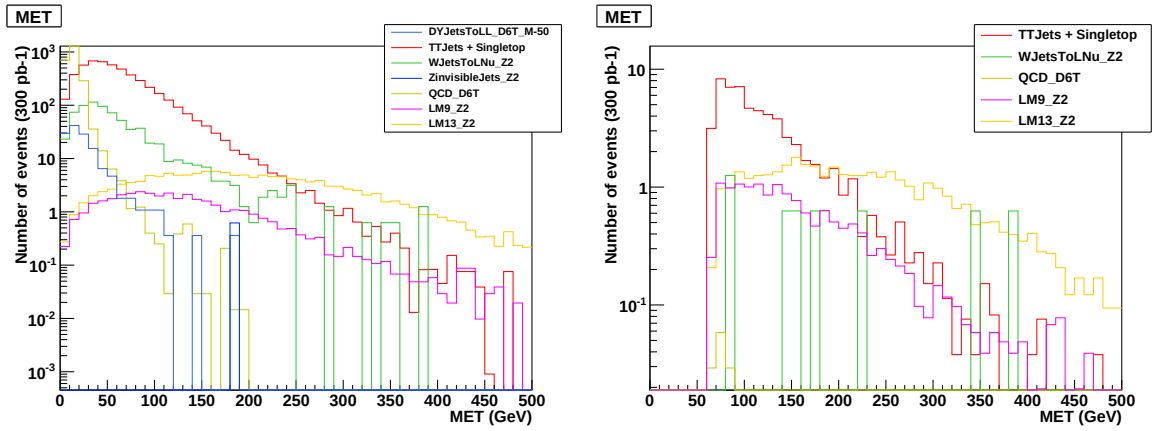
Sample	σ (pb)	σ_{eff} (pb) after applying the optimal selection
DYJetsToLL_D6T_M-50	3048	0
QCD	3.7e7	5.35e-4
TTJets_D6T	165	1.94e-1
Single top	33	3.00e-3
WJetsToLNu_Z2	31314	1.67e-2
ZinvisibleJets_Z2	4500	0
LM9_Z2	7.134	4.99e-2
LM13_Z2	6.899	1.27e-1

6.3.4 Conclusion

We can conclude that RGS works satisfactory. One does have to be careful in choosing the cut variables one wants to use. Correlations between the variables can have a counter-intuitive effect.

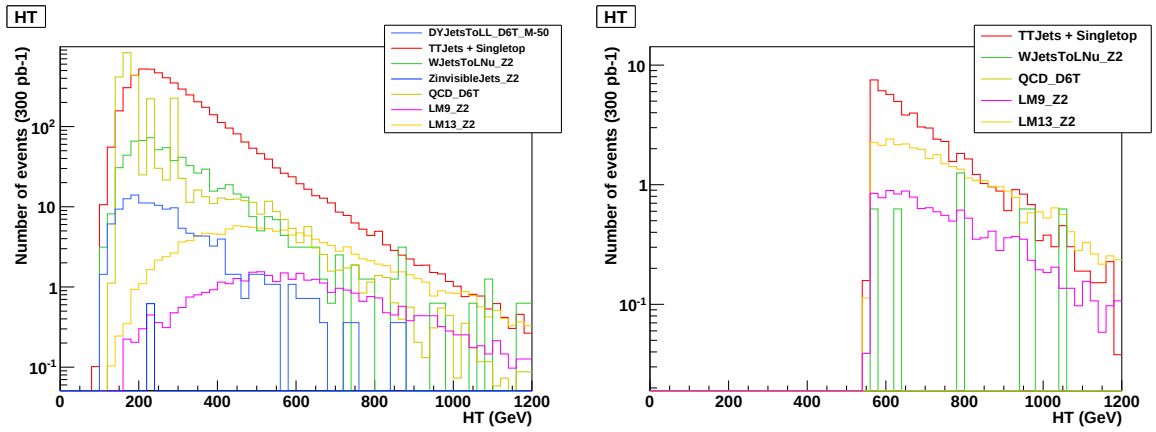
It is interesting to note that the optimal cut does not depend on the luminosity in case the default version of the significance (equation 6.1) was used. When using the Punzi significance (equation 6.2), cut points yielding a higher efficiency are favoured over those yielding a higher purity. In the remainder of this work, we will use the default significance as this is the definition that is most widely used in the particle physics community.

We can also conclude that for the LM9 benchmark point, a good selection requires a hard cut on HT, a more moderate cut on MET and a large number of jets.



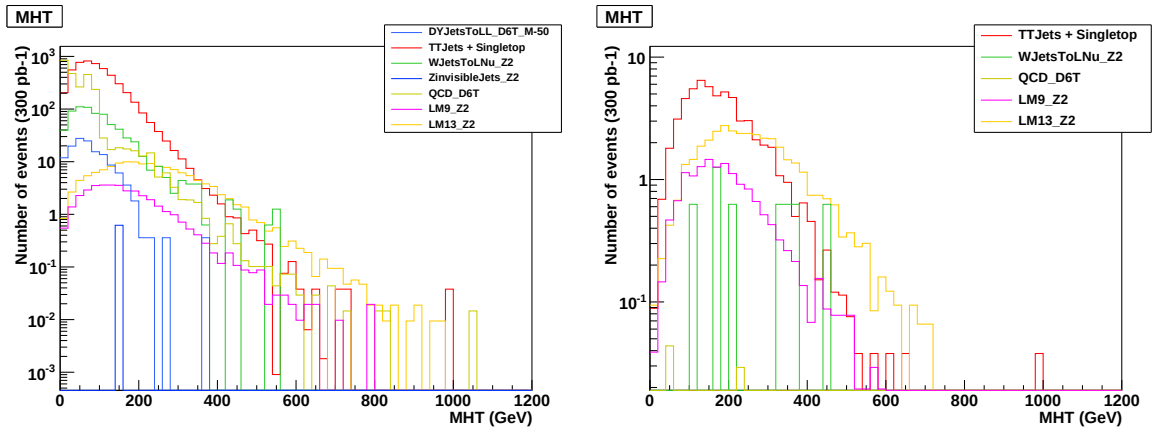
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.4: *MET* distribution for signal and background samples, before and after applying the optimal cuts



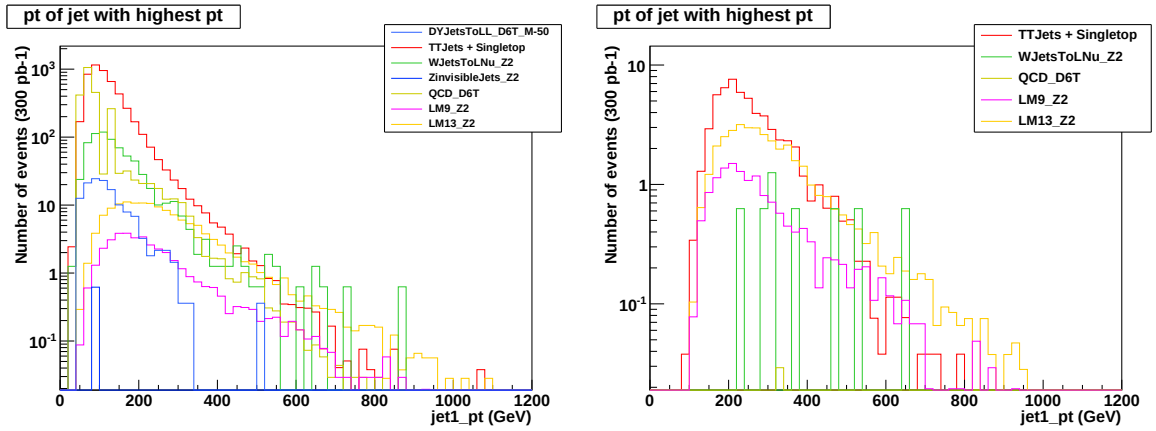
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.5: *HT* distribution for signal and background samples, before and after applying the optimal cuts



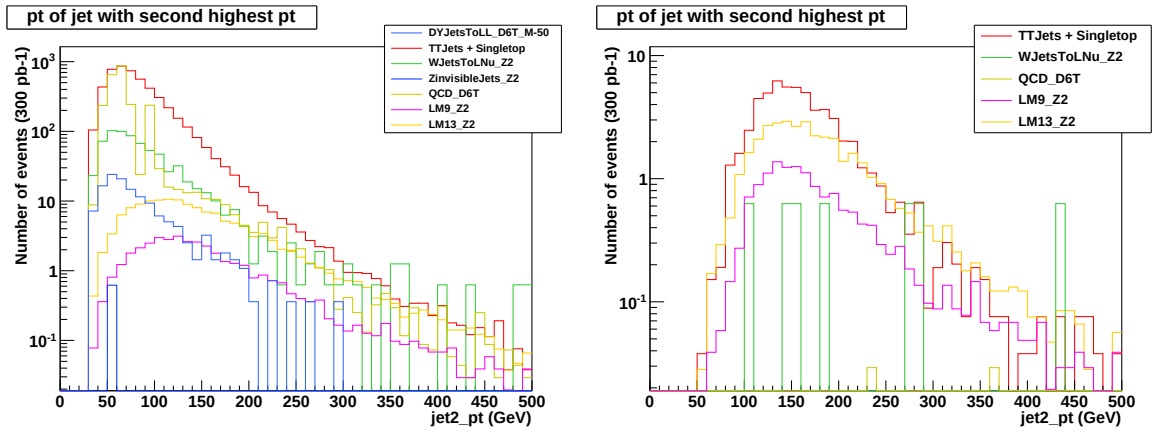
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.6: *MHT* distribution for signal and background samples, before and after applying the optimal cuts



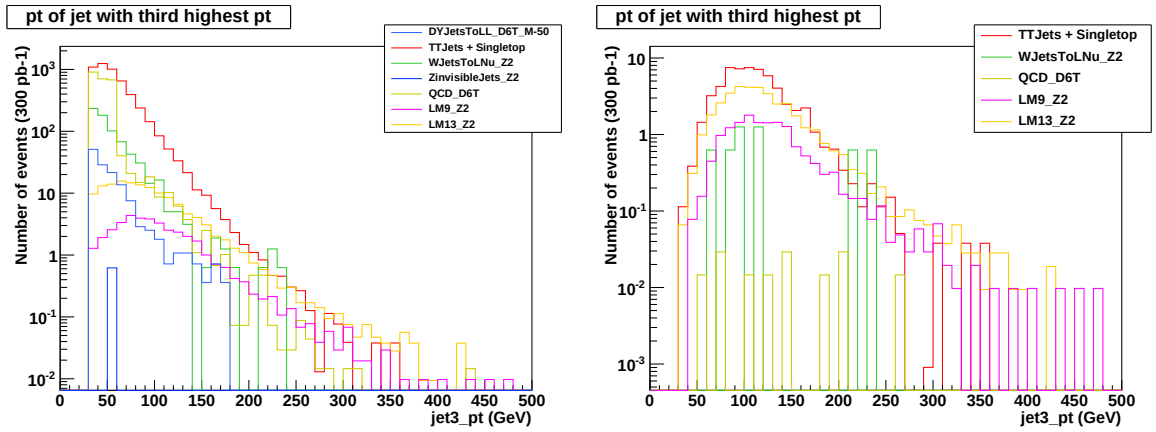
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.7: Distribution of the p_T of the jet with the highest p_T for signal and background samples, before and after applying the optimal cuts



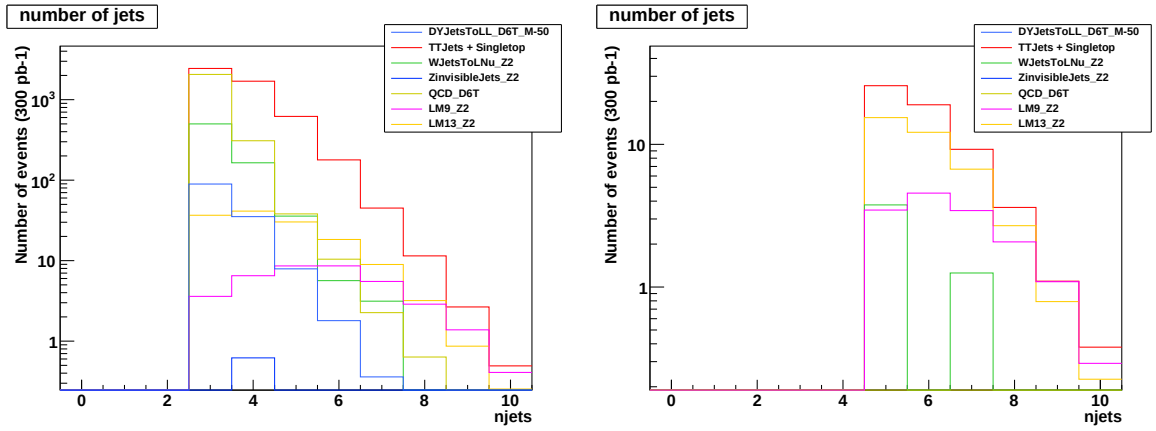
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.8: Distribution of the p_T of the jet with the second highest p_T for signal and background samples, before and after applying the optimal cuts



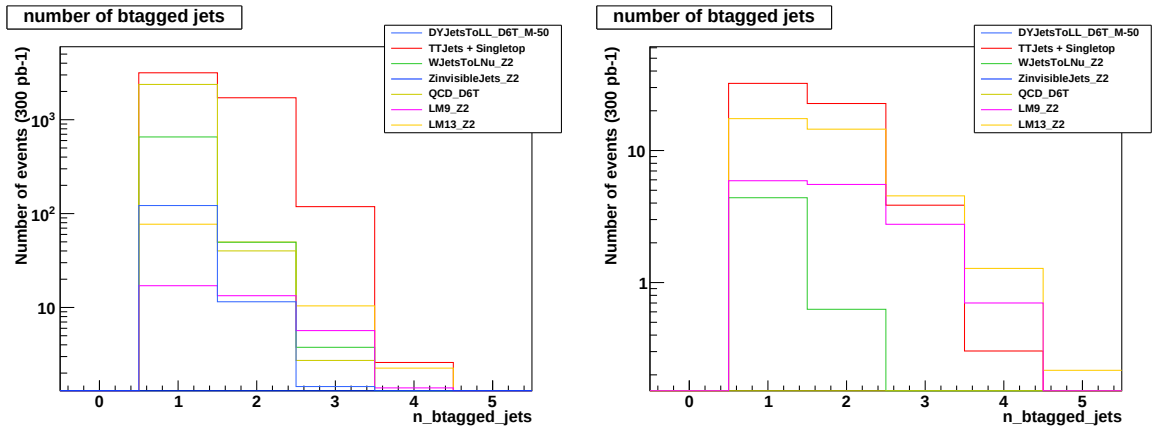
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.9: Distribution of the p_T of the jet with the third highest p_T for signal and background samples, before and after applying the optimal cuts



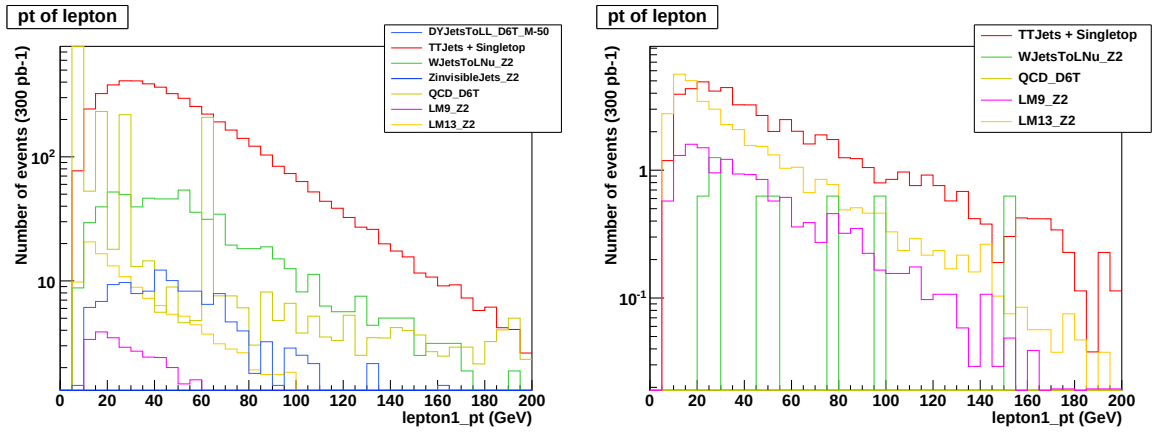
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.10: n_{jets} distribution for signal and background samples, before and after applying the optimal cuts



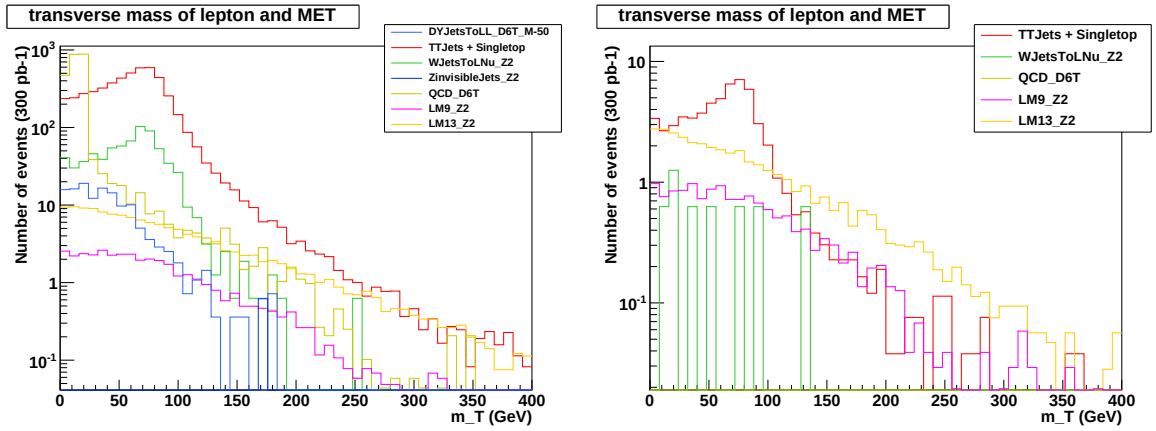
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.11: $n_{btaggedjets}$ distribution for signal and background samples, before and after applying the optimal cuts



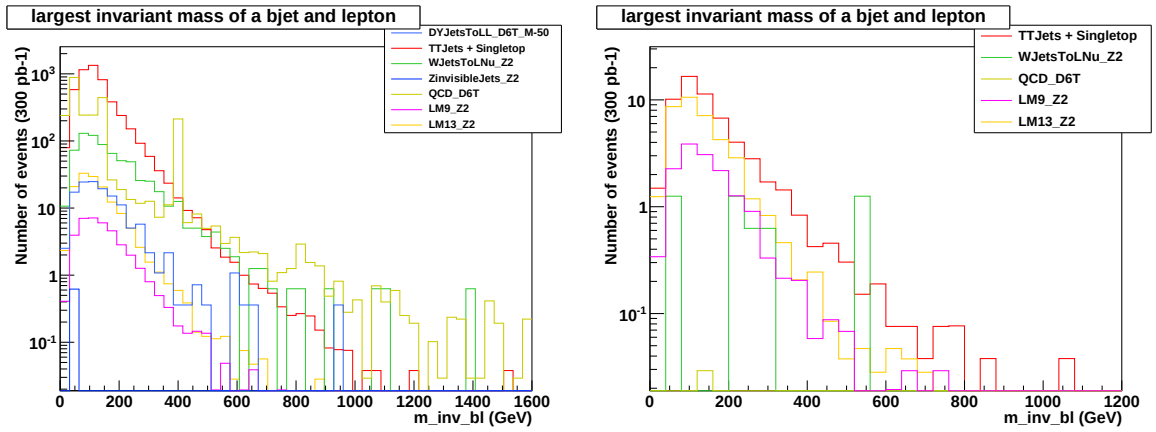
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.12: Lepton p_T distribution for signal and background samples, before and after applying the optimal cuts



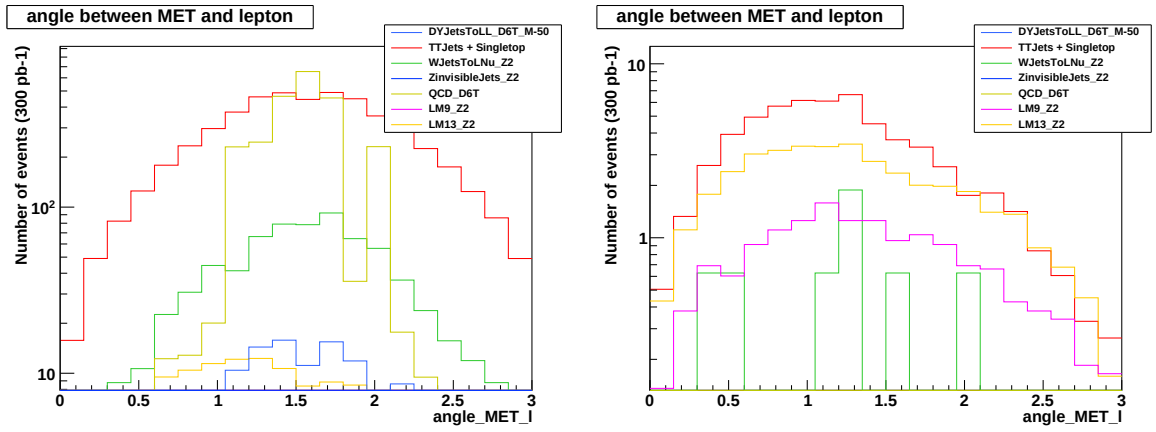
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.13: m_T distribution for signal and background samples, before and after applying the optimal cuts



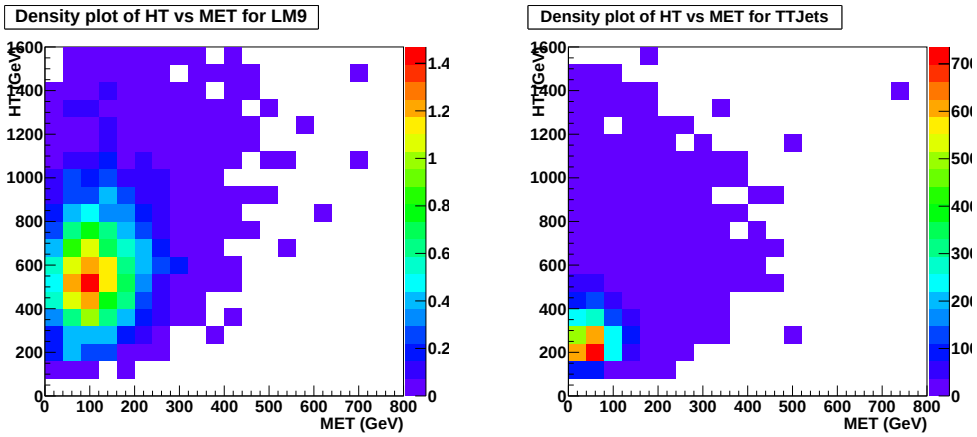
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.14: Distribution of the invariant mass of the lepton and a b-jet, for signal and background samples, before and after applying the optimal cuts



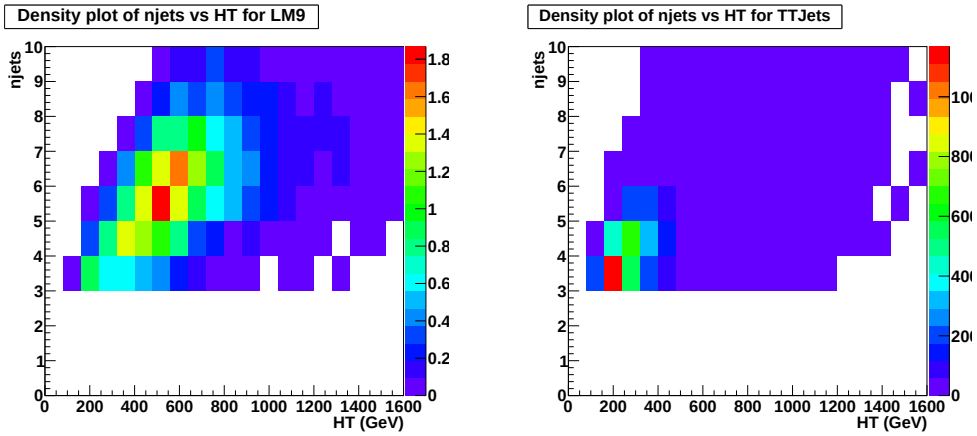
(a) Preselection (b) $MET > 66.8$, $HT > 559$, $n_{jets} > 5$

Figure 6.15: Distribution of the angle between the lepton and the MET for signal and background samples, before and after applying the optimal cuts



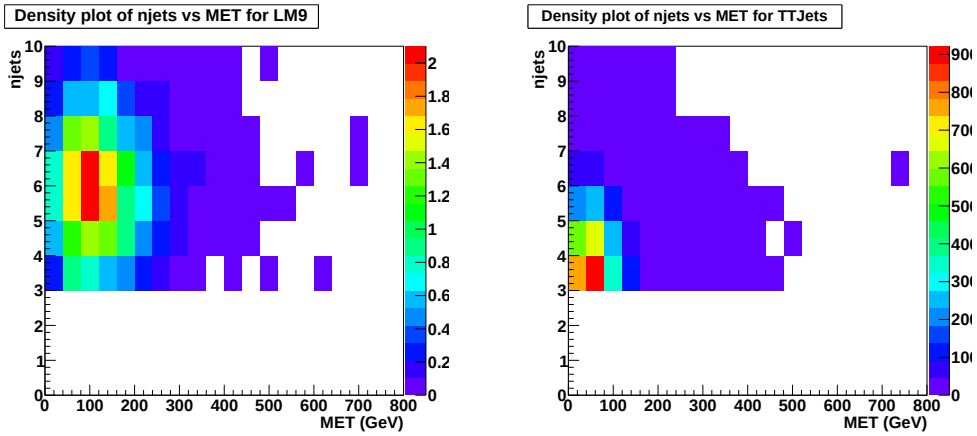
(a) LM9 (b) TTJets

Figure 6.16: Density plot of HT versus MET



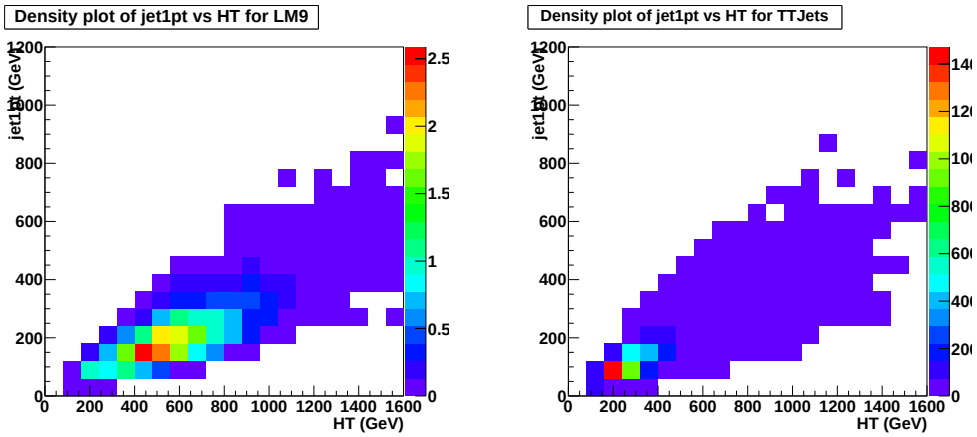
(a) LM9 (b) TTJets

Figure 6.17: Density plot of n_{jets} versus HT



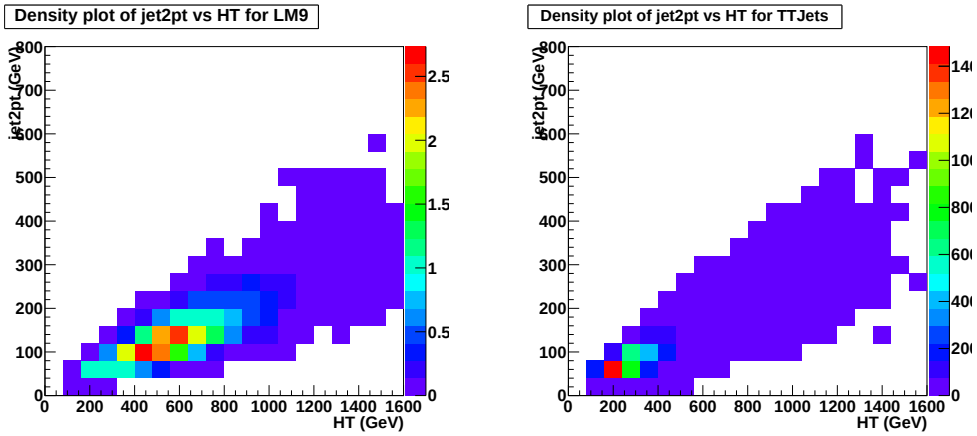
(a) LM9 (b) TTJets

Figure 6.18: Density plot of n_{jets} versus MET



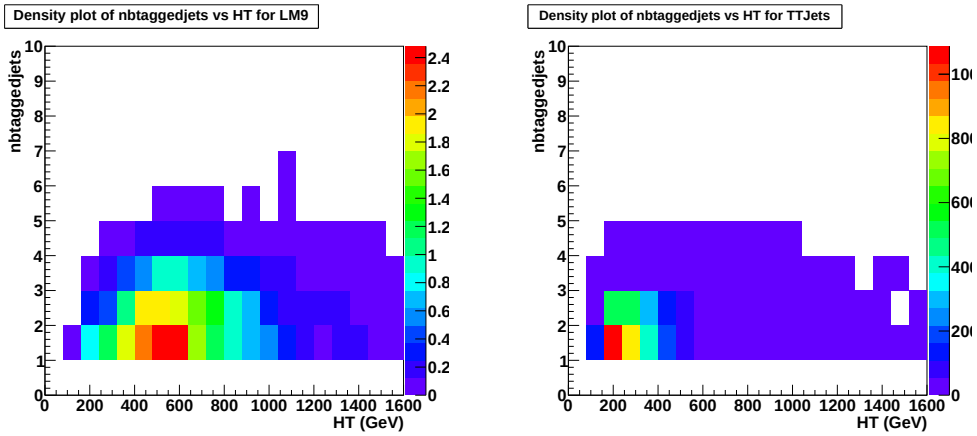
(a) LM9 (b) TTJets

Figure 6.19: Density plot of the p_T of the jet with the highest p_T versus HT



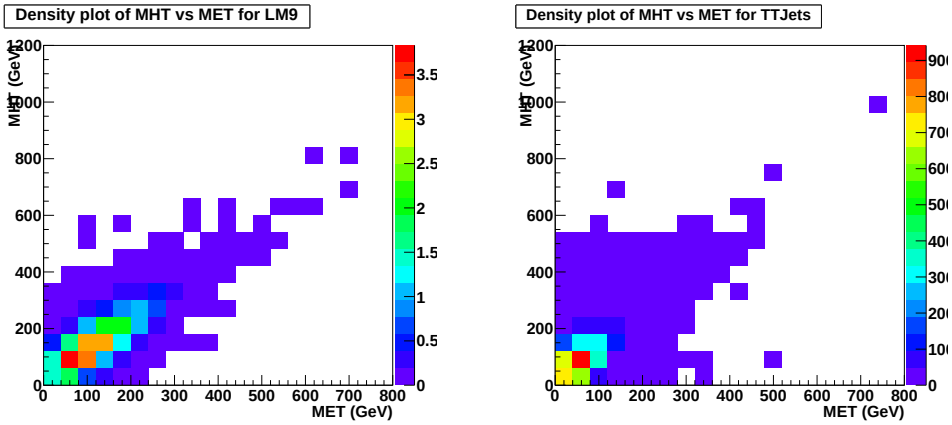
(a) LM9 (b) TTJets

Figure 6.20: Density plot of the p_T of the jet with the second highest p_T versus HT



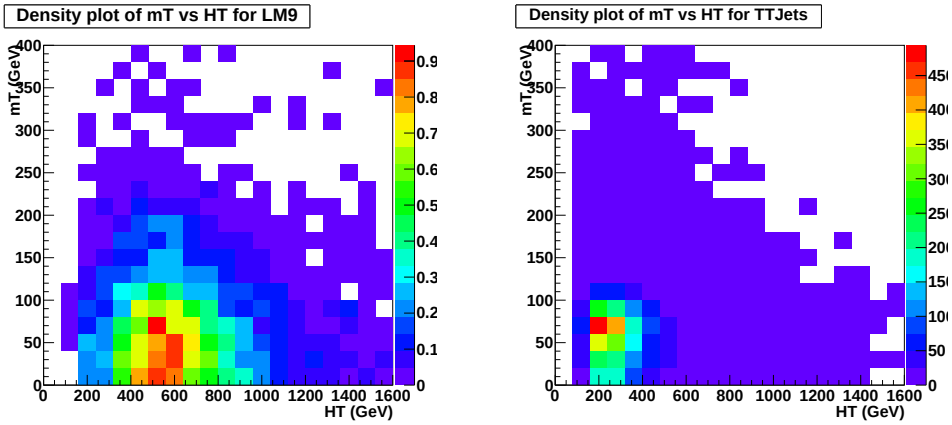
(a) LM9 (b) TTJets

Figure 6.21: Density plot of $n_{btaggedjets}$ versus HT



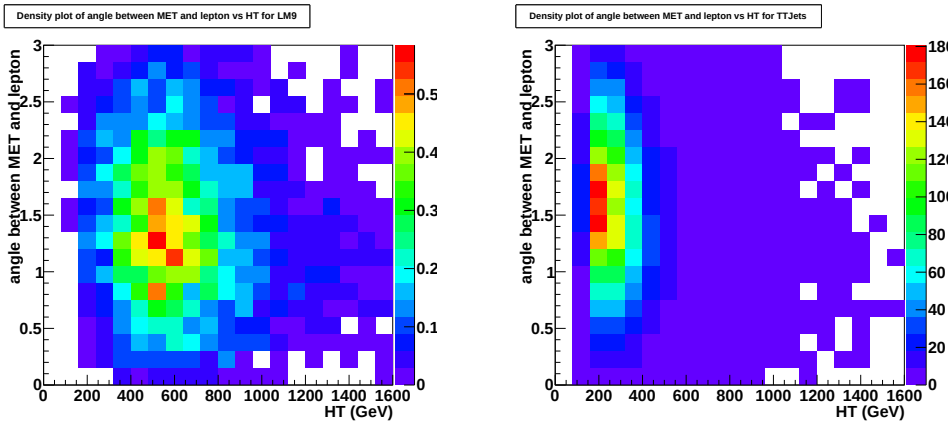
(a) LM9 (b) TTJets

Figure 6.22: Density plot of MHT versus MET



(a) LM9 (b) TTJets

Figure 6.23: Density plot of m_T versus HT



(a) LM9 (b) TTJets

Figure 6.24: Density plot of the angle between the lepton and the MET versus HT

Table 6.3: Optimization results for default significance (300 pb^{-1}).

HT	n_{jets}	MET	jet1_p_T	jet2_p_T	significance	signal	background
592.5	/	/	/	/	1.63	21.1	145.6
/	7	/	/	/	1.34	11.3	59.9
/	/	131	/	/	1.07	18.2	272.7
565.0	/	68.9	/	/	1.80	18.6	88.4
594.0	5	/	/	/	1.73	18.8	100.0
/	6	102.0	/	/	1.66	12.2	41.9
/	/	86.0	169.0	/	1.26	20.7	247.3
593.9	5	66.0	/	/	1.93	15.3	47.8
571.0	/	65.2	117.0	/	1.79	18.4	88.2
577.9	5	76.2	/	96.0	1.93	14.7	43.4
518.4	5	86.0	124.1	/	1.89	15.9	54.9

Table 6.4: Optimization results for Punzi significance with $a = 3$ (300 pb^{-1}). The cut point in bold is the cut point used as optimal cut point, see section 6.3.3.

HT	n_{jets}	MET	jet1_p_T	jet2_p_T	significance	signal	background
557.5	/	/	/	/	1.56	23.8	190.2
/	6	/	/	/	1.25	21.1	237.5
/	/	131	/	/	1.01	18.2	272.7
565.0	/	68.9	/	/	1.71	18.6	88.4
573.6	5	/	/	/	1.64	20.1	115.6
/	6	80.4	/	/	1.54	14.9	66.9
/	/	86.0	169.0	/	1.20	20.7	247.3
558.6	**5**	**66.8**	**/**	**/**	**1.83**	**16.9**	**59.7**
571.0	/	65.2	117.0	/	1.69	18.4	88.2
577.9	5	76.2	/	96.0	1.82	14.7	43.4
518.4	5	86.0	124.1	/	1.79	15.9	54.9

Table 6.5: Optimization results for Punzi significance with $a = 5$ (300 pb^{-1}).

HT	n_{jets}	MET	jet1_p_T	jet2_p_T	significance	signal	background
557.5	/	/	/	/	1.46	23.8	190.2
/	6	/	/	/	1.18	21.1	237.5
/	/	131	/	/	0.96	18.2	272.7
496.0	/	70.0	/	/	1.57	22.8	143.6
557.5	5	/	/	/	1.52	21.1	129.0
/	6	80.4	/	/	1.40	14.9	66.9
/	/	77.6	168.0	/	1.13	22.0	286.2
515.2	5	68.2	/	/	1.66	18.7	76.9
496.0	/	70.0	87.2	/	1.57	22.7	142.6
470.4	5	63.6	/	70.3	1.62	21.2	111.6
518.4	5	86.0	124.1	/	1.61	15.9	54.9

Chapter 7

Performance of the Delphes Fast Simulator

To execute the sNavigator procedure as explained in chapter 4, we need to simulate events for a large number of points in the SUSY parameter space. Therefore, we will not use the CMS Full Detector Simulation. The required CPU time for this would become too large. Instead, we will use Delphes (version 1.9), a framework for fast simulation of a general purpose collider experiment [48]. We opted not to use the CMS Fast Simulator because Delphes is faster and more convenient (we avoid the overhead of working within the full CMSSW framework). We quote the Delphes description from their website [49]:

> *The simulation includes a tracking system, embedded into a magnetic field, calorimetry and a muon system, and possible very forward detectors arranged along the beamline (zero-degree calorimeters and near-beam detectors). The framework is interfaced to standard file format (e.g. Les Houches Event File or HepMC) and outputs observable analysis data objects, like missing transverse energy and collections of electrons or jets. The simulation of detector response takes into account the detector resolution, and usual reconstruction algorithms for complex objects, like FastJet.*

Delphes works with a so-called "detector card". This is a file containing all the information about the specific detector you wish to mimic, in our case the CMS detector. One can specify the geometry of the detector (tracker, calorimeters, muon system) and the magnetic field. Also the energy resolution, the jet reconstruction algorithm to be used and some parameters for the calculation of lepton isolation can be adjusted.

In the remainder of this chapter, we will compare the Delphes simulation with CMSSW FullSim. We will do this for all main physics objects. Within CMSSW these are defined according to the definitions in section 5.2. For the leptons we use the RA2b-based definitions. Within Delphes, we will use the standard Delphes collections and apply the same p_T and η cuts as for CMSSW. For each of the physics objects, we will assess the "out-of-the-box" Delphes performance. If needed, and feasible, we will try to improve this performance.

We will use the benchmark point LM9, for which a CMSSW sample with 220000 events was produced centrally (see table 5.2). For the Delphes sample, we first generated 25000 events via `PYTHIA6` and `HepMC` [50]. The output `.hepmc` file was then used as input for the Delphes simulation. The resulting ntuple allows one to easily access all objects and perform the analysis. In order to be able to do a full comparison, no event selection was applied.

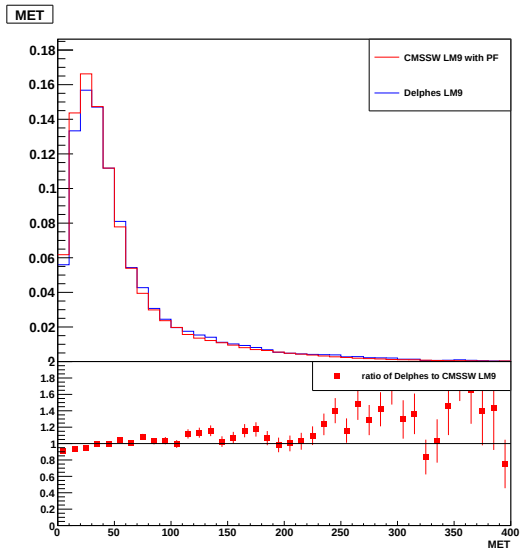


Figure 7.1: (top) Missing transverse energy (in GeV) distribution for CMSSW and Delphes; (bottom) Ratio of Delphes over CMSSW

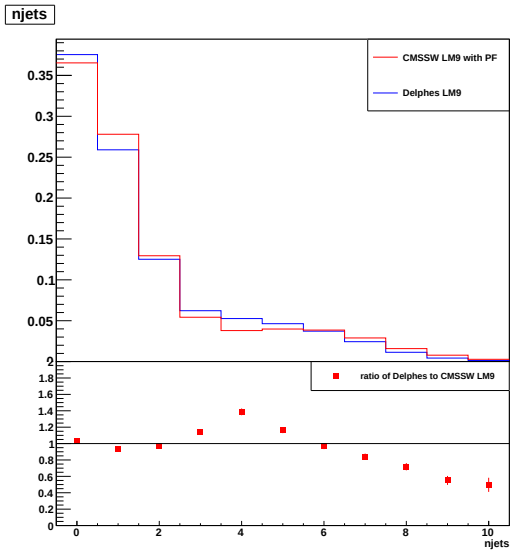


Figure 7.2: (top) Distribution of the number of jets for CMSSW and Delphes; (bottom) Ratio of Delphes over CMSSW

7.1 Missing transverse energy

The missing transverse energy in Delphes is defined as the opposite of the vector sum of all energy deposits in the calorimeter cells and the momentum of the muon.

$$\overrightarrow{E_T}^{\text{miss}} = -\sum_{i}^{\text{cells}} \overrightarrow{E_T}(i) - \overrightarrow{p_T}^{\text{muon}} \tag{7.1}$$

In figure 7.1 we show the missing transverse energy distribution for CMSSW and Delphes and their ratio. As we can see, there is a good agreement. No further improvements are needed at this point.

7.2 Jets

In Delphes, calorimeter towers (as opposed to PF particles in CMSSW) are clustered into jets with the anti-kt algorithm with jet radius 0.5 [28]. As the CMSSW PF jet collection was cleaned to remove all isolated leptons, we will also implement this for Delphes. In Delphes, muons never enter into the jet collection, but electrons do. Therefore, the Delphes jet collection was cleaned to remove all entries that were in fact electrons. This was done by requiring that the distance jet-electron exceeds $\Delta R = 0.3$, where $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The comparison between Delphes and CMSSW is shown in figure 7.2. We see that the agreement is reasonable. As jets are complicated objects, a more thorough study is needed to improve this.

7.3 b-tagging

Another important issue for the b-jet(s) + lepton topology, is to make sure that b-tagging is included in Delphes in a satisfactory way. In Delphes a standard b-tag efficiency of 40% is used. Each jet has its associated `btag`, which is equal to `1` for b-tagged jets and equal to `0` for non b-tagged jets.

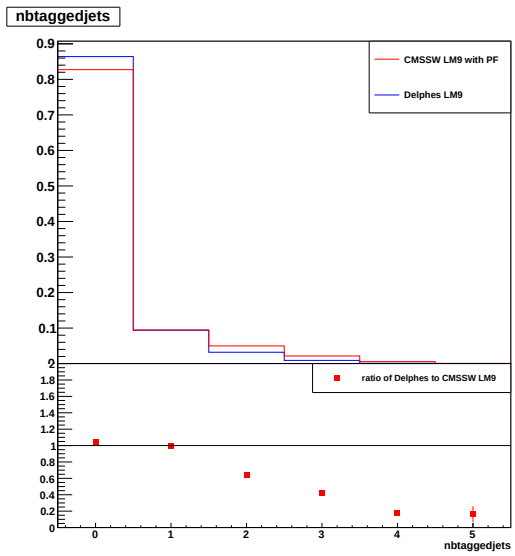


Figure 7.3: (top) Distribution of the number of b-tagged jets for CMSSW and Delphes (with standard b-tag efficiency); (bottom) Ratio of Delphes over CMSSW

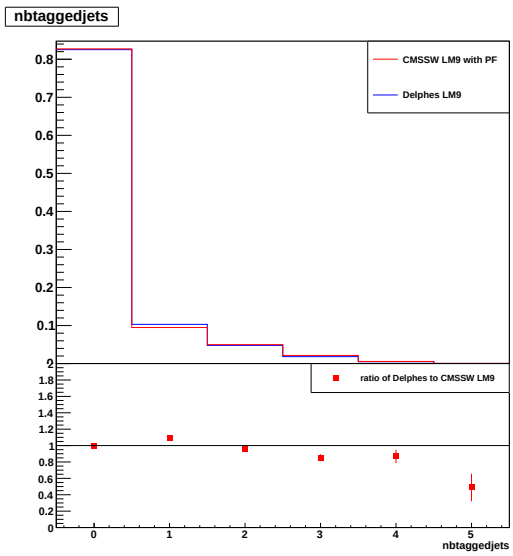


Figure 7.4: (top) Distribution of the number of b-tagged jets for CMSSW and Delphes (with b-tag efficiency map obtained via CMSSW); (bottom) Ratio of Delphes over CMSSW

Comparing Delphes to CMSSW we find that the agreement is not so good (see figure 7.3). We will therefore implement b-tag and b-mistag efficiency maps as functions of p_T and η.

These efficiency maps are created via Bayesian Neural Networks (BNN's) [51], using a CMSSW sample, e.g. TTJets. For each jet in this sample, we determine whether it is a true b-jet or not. In CMSSW, each jet has an associated variable `partonFlavour`, containing the generator level information on the flavour of the jet. This variable allows to distinguish b-jets from non b-jets. For the b-tag efficiency map, we create an input data sample, containing for all true b-jets, their p_T, η and whether or not they are tagged. A BNN is then trained with this sample and returns a functional form for the efficiency map $\epsilon_{\text{b-tag}}(p_T, \eta)$. For the b-mistag efficiency map $\epsilon_{\text{b-mistag}}(p_T, \eta)$, we proceed in the same way and train a BNN with the non b-jet sample. Both functional forms can then be included in Delphes.

To use these maps in practice, we try to match a jet to a generator level B-hadron. To achieve this, we compute the distance in ΔR. If $\Delta R < 0.3$ we take the jet to be a real b-jet and apply the b-tag efficiency map. For this we draw a random number r between 0 and 1. If $r < \epsilon_{\text{b-tag}}(p_T^{\text{jet}}, \eta^{\text{jet}})$, the jet is b-tagged. If $\Delta R > 0.3$, the jet is assumed to be a non b-jet and the b-mistag efficiency map is applied. The performance of this method is shown in figure 7.4. We see that the agreement between Delphes and CMSSW is now excellent.

7.4 Electrons and muons

In our analysis we will only use isolated leptons. We will thus have to check the isolation criteria in Delphes. In Delphes each lepton has an isolation flag, which is equal to one for isolated leptons. We will thus require `IsolFlag = 1`.

Comparing the number of electrons and muons in Delphes and CMSSW, see figures 7.5 and 7.6, we see that there is a good agreement. We must note, however, that if we tighten the isolation criteria in CMSSW (e.g. by using the RA4-based definition instead of the RA2b-based definition), the agreement will worsen. In that case, it shall be worthwile to also implement efficiency maps for lepton reconstruction and isolation.

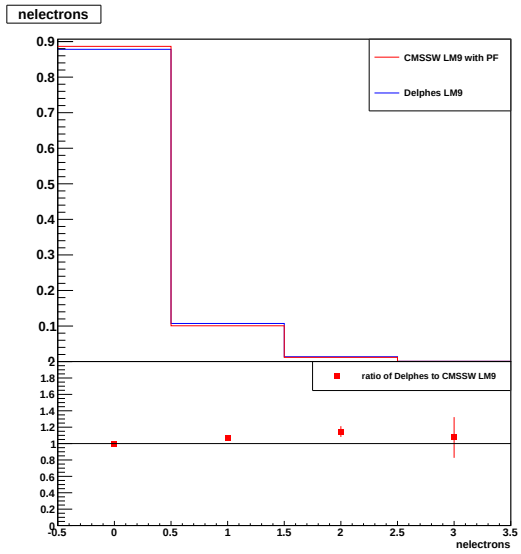


Figure 7.5: (top) Distribution of the number of electrons for CMSSW and Delphes; (bottom) Ratio of Delphes over CMSSW

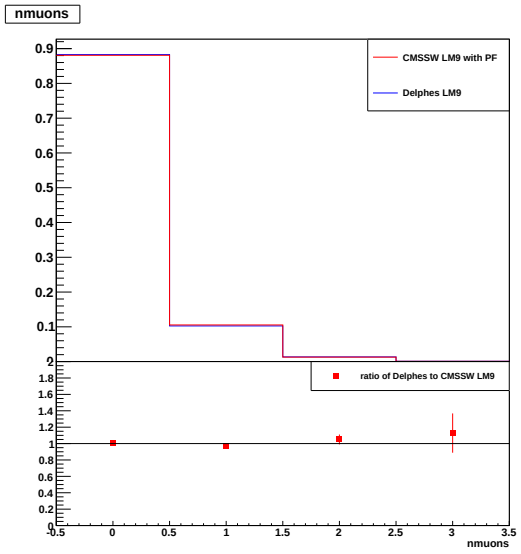


Figure 7.6: (top) Distribution of the number of muons for CMSSW and Delphes; (bottom) Ratio of Delphes over CMSSW

7.5 Conclusions

Overall, the "out-of-the-box" Delphes performance is good. In addition, we were able to significantly improve the b-tagging performance by including efficiency maps obtained via Bayesian Neural Networks.

Chapter 8

Validation of the sNavigator procedure

In this chapter we present results obtained in a first test of the sNavigator procedure as explained in chapter 4. We will discuss the preselection that was applied and the results we obtained for all three methods (see section 4.3.2). We will also give an estimate of the resources (storage, computing time) needed to run the procedure. To conclude, we will compare the three methods and discuss the caveats of the procedure.

8.1 Preselection

We will perform the optimization for the b-jet(s) + lepton search topology. Therefore, we will first apply a preselection. We applied the following preselection in the CMSSW framework (see also section 5.2):

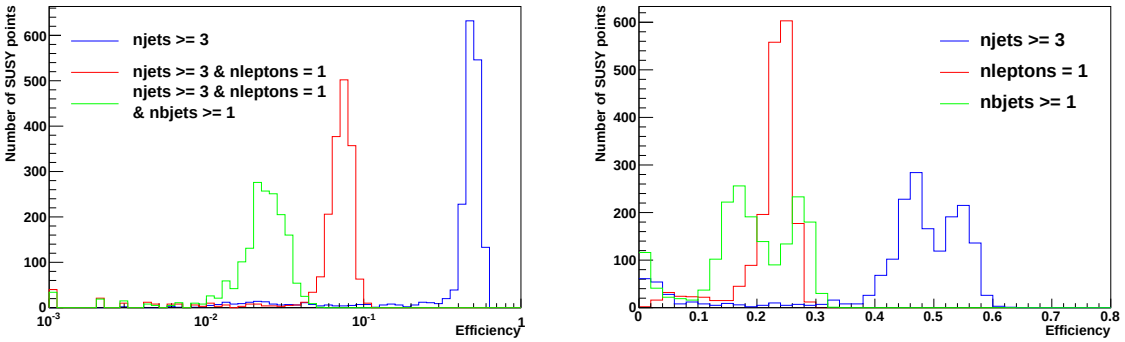
- minimum three jets
- minimum one b-jet
- exactly one isolated lepton (electron or muon) according to the RA2b-based definition

This preselection was applied to all background samples listed in table 5.1. The resulting efficiencies and effective cross section after the preselection are shown in table 8.1.

The same preselection should then also be applied for the signal. As discussed in section 5.1, we use a set of 1829 mSUGRA points with cross section larger than 0.1 pb^{-1} as our standard set of SUSY points. For the characterization of these points, we refer to section 5.1 and figures 5.1 - 5.6. We will use the standard object collections as provided by Delphes (see chapter 7). So we use the jet collection without cleaning and we use the standard Delphes `btag` and `IsolFlag`.

Table 8.1: Cross section and efficiencies for the background samples

Background	σ (pb)	Preselection efficiency			σ_{eff} (pb)
		njets ≥ 3	nleptons $= 1$	nbtaggedjets ≥ 1	
DYJetsToLL_D6T_M-50	3048	5.85e-3	2.04e-3	1.85e-4	0.57
QCD	3.7e7	2.22e-2	1.31e-5	8.89e-7	32.92
TTJets_D6T	165	7.90e-1	2.00e-1	1.54e-1	25.41
Single top	33	3.07e-1	9.34e-2	6.48e-2	2.13
WJetsToLNu_Z2	31314	3.76e-3	1.55e-3	1.50e-4	4.71
ZinvisibleJets_Z2	4500	3.08e-3	1.84e-6	4.59e-7	2.07e-3
Total					65.74



(a) Cumulative cuts (b) Separate cuts

Figure 8.1: Distribution of preselection efficiencies

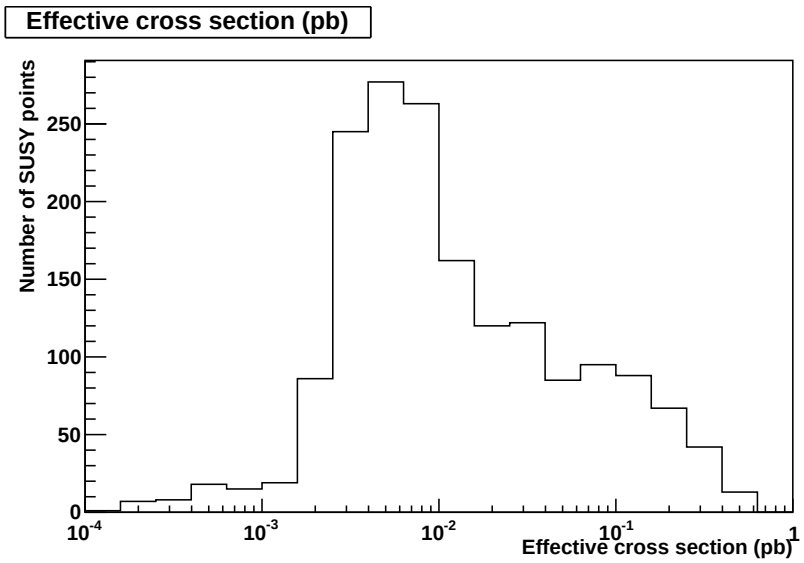


Figure 8.2: Effective cross section after the preselection cuts

Table 8.2: Comparison of preselection efficiencies determined via Delphes and CMSSW.

		njets ≥ 3	nleptons $= 1$	nbjets ≥ 1
LM9	Delphes	0.238	0.033	0.018
	CMSSW	0.227	0.037	0.024
LM13	Delphes	0.671	0.119	0.060
	CMSSW	0.666	0.137	0.097

The distribution of the signal efficiencies for the different steps in the preselection is shown in figure 8.1. The average (total) preselection efficiency is about $2 - 3\%$. The distribution of the effective cross section after the preselection is shown in figure 8.2.

As we have both Delphes and CMSSW samples for the benchmark points LM9 and LM13, we were able to compare the resulting preselection efficiencies in both cases. The results of this can be viewed in table 8.2. For a more detailed study of the performance of Delphes versus CMSSW, we refer to chapter 7. From table 8.2, we can conclude that in general the agreement between Delphes and CMSSW is reasonable. There is still room for improvement, especially for b-tagging. In section 7.3 we showed that we can significantly improve the b-tagging performance by replacing

the standard (uniform) btagging of Delphes by a functional form depending on the η and p_T of the jet.

8.2 Results

We will now proceed with the actual optimization and apply the three methods described in section 4.3.2.

For the calculation of the discovery probability, equation 4.9, we only take into account the RA2b likelihood in its simplest, Poissonian form, see equation 4.12. The values for the number of observed events N^{RA2b} and expected number of background events b^{RA2b} are taken directly from the RA2b analysis [46], which was performed on 36 pb^{-1} of data. We used $N^{RA2b} = 18$ and $b^{RA2b} = 20.4$.

For methods A and B, we use the normalized discovery probability. As stated in section 4.3.2, the optimal cut point is the point that gives the largest value for the discovery probability.

For method C, we also follow the steps as outlined in section 4.3.2. The optimal cut point is the one that gives us the largest significance $\mathcal{S} = \frac{S}{\sqrt{S+B}}$, where we use the weighting as stated in equation 4.19 to calculate S. For this cut point we will then also compute the discovery probability in order to compare method C with methods A and B.

All results shown here were obtained for a luminosity of 300 pb^{-1} and with cut variables MET, HT and number of jets n_{jets}, unless stated otherwise.

The results of the sNavigator procedure are shown in table 8.3. For each of the three used methods, we show the value of the discovery probability, the associated number of SUSY points that have a significance larger than 3 and the values of the optimal cut point. In figure 8.3 we show the distribution of the significance for each SUSY point, weighted with the RA2b likelihood, for the optimal cut point as determined by method A.

Table 8.3: Optimization results for all 1829 SUSY points (lumi = 300 pb^{-1})

		Method A	Method B	Method C
npoints		2	2	1
P_{disc}		0.000719343	0.00075231	0.000661256
Cut set	MET $\geq$	97.2	187.9	216.3
	HT $\geq$	648.8	876.7	566.7
	njets $\geq$	9	6	8

Table 8.4: Number of SUSY points for which the cut set of method B contains a cut that allows the SUSY point to attain a significance larger than some value (for a luminosity of 300 pb^{-1}).

Significance	≥ 1	≥ 2	≥ 3	≥ 4	≥ 5
# SUSY points	460	156	45	7	1

In order to get an idea of how many SUSY points can, in the most optimal case, have a significance $\mathcal{S}$ above a certain value $\mathcal{S}_0$, we checked for each SUSY point what the maximum significance it can achieve is, using the cut set of method B. If this maximum significance is larger than $\mathcal{S}_0$, we count the SUSY point. For several values of $\mathcal{S}_0$, the number of SUSY points that can have a significance larger than this value is listed in table 8.4. We do not expect to achieve these numbers in our optimization as it is unlikely that the cut points allowing the SUSY points to have $\mathcal{S} > \mathcal{S}_0$ are the same for all SUSY points. The distribution of the maximum significance for all sampled SUSY points is displayed in figure 8.4. Comparing tables 8.3 and 8.4, we see that we can only discover 2 out of 45 SUSY points with our optimal selection. The selection is also very tight.

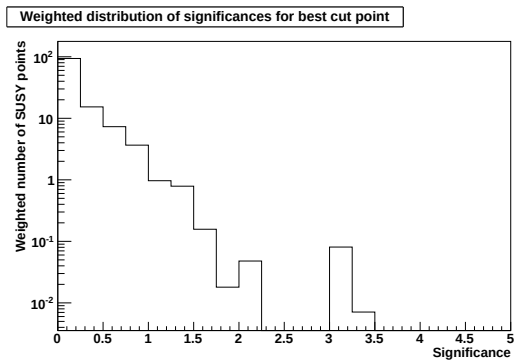


Figure 8.3: Distribution of significances for the optimal cutpoint determined by method A, using the default set of SUSY points and for a luminosity of 300 pb^{-1}

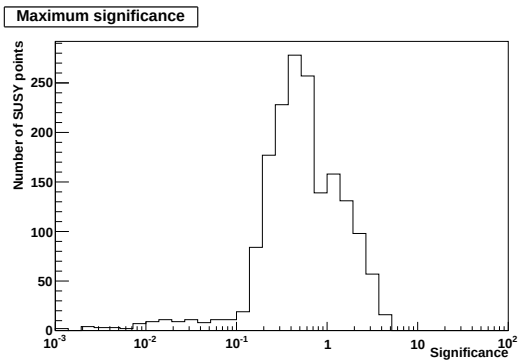


Figure 8.4: Distribution of the maximum significance as calculated with the cut set of method B, using the default set of SUSY points and for a luminosity of 300 pb^{-1}

Table 8.5: Optimization results for luminosity of 1 and 5 fb^{-1}, using the standard set of SUSY points.

		Method A (1 fb^{-1})	Method A (5 fb^{-1})
npoints		86	342
P_{disc}		0.0121447	0.100768
Cut set	MET $\geq$	130.5	119.2
	HT $\geq$	516.7	532.6
	njets $\geq$	6	7

In figures 8.5, 8.6 and 8.7 we show a scatter plot of the RA2b likelihood versus the cross section, the RA2b likelihood versus the maximum significance and the cross section versus the maximum significance, respectively. As expected, the likelihood is lower for points with high cross section and high maximum significance. We also see that there is a strong correlation between the cross section and the maximum significance.

To investigate the effect of the luminosity on the procedure, method A was also executed for a luminosity of 1 fb^{-1} and 5 fb^{-1}. The results can be found in table 8.5 and figure 8.8. To compare, the number of SUSY points with maximum significance larger than 3 is 239 for 1 fb^{-1} and 585 for 5 fb^{-1}. We see that we can now discover a larger portion of the sampled SUSY points. The resulting optimal cut values are also less tight.

Comparing the resulting optimal selections for the different luminosities, we see that there is quite a difference between the selection obtained for 300 pb^{-1} and for the other two luminosities. As there were only two SUSY points that could be discovered for 300 pb^{-1}, we decided to inspect these points in a bit more detail.

We noticed that one of these points has a very large cross section (38 pb) and also a very high likelihood (because no events passed the RA2b selection). Hence, this one SUSY point was able to strongly influence the optimization as there were only few SUSY points that contributed to the discovery probability. For higher luminosities, more points contribute and as a consequence the effect of one point is not so strong. The sum of the likelihoods of many points with moderate likelihood can become larger than the large likelihood of that one point. In theory, this is exactly what we want the procedure to do. In practice, however, we need to make sure that the behavior of such points is not a statistical fluctuation.

In our case, there was only one Monte Carlo event from that point with high cross section and high likelihood that passed our preselection. This event had nine jets, $MET = 225$ GeV and $HT = 883$ GeV. The optimal cut point was drawn towards this event because such a hard cut is

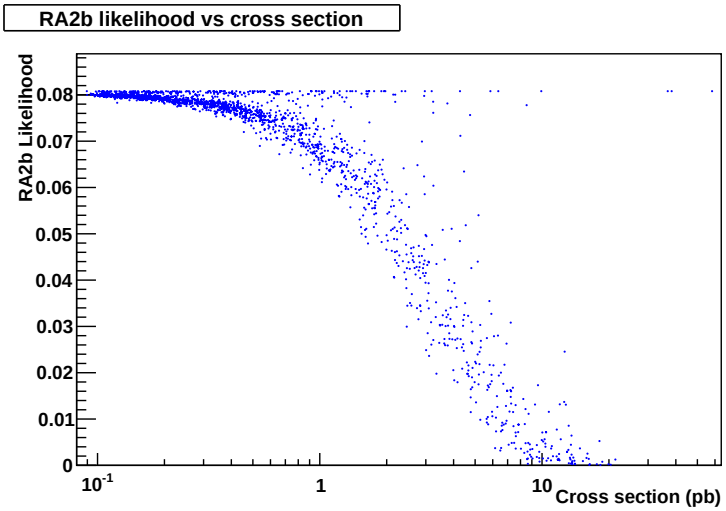


Figure 8.5: RA2b likelihood versus cross section

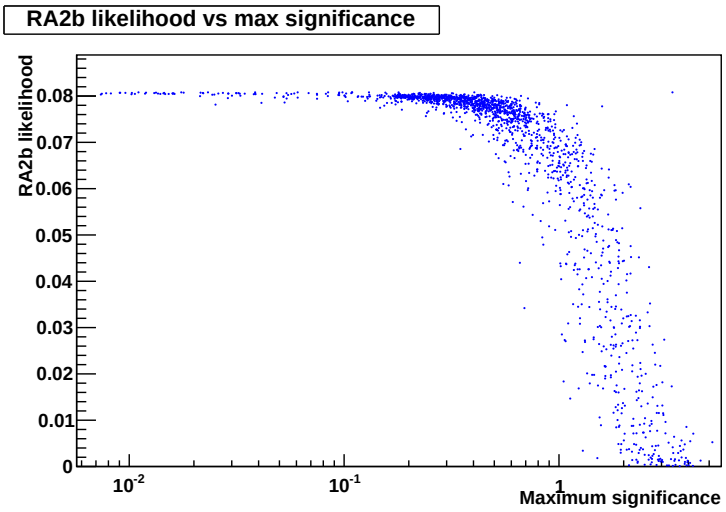


Figure 8.6: RA2b likelihood versus maximum significance

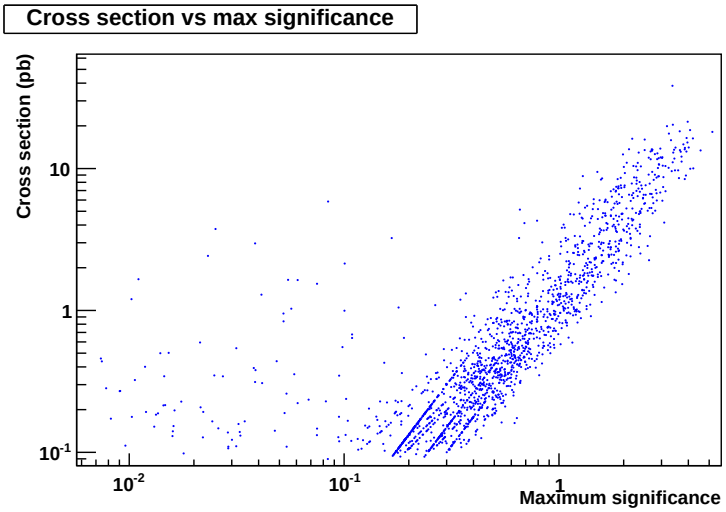


Figure 8.7: Cross section versus maximum significance

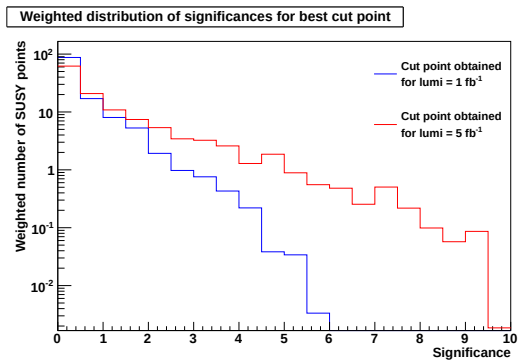


Figure 8.8: Distribution of significances for the two cut points in table 8.5, obtained by performing the optimization via method 1 for luminosities of 1 and 5 fb^{-1}.

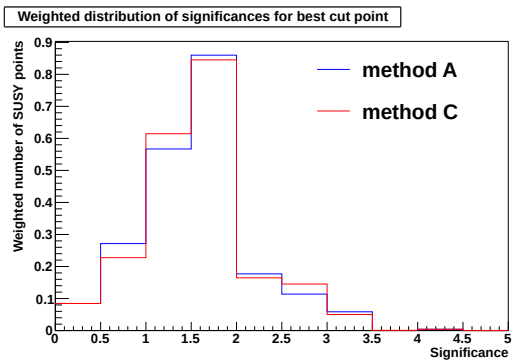


Figure 8.9: Distribution of significances for the two cut points in table 8.6, obtained by performing the optimization on the set of SUSY points with maximum significance larger than 2.

Table 8.6: Optimization results for the subset of SUSY points with maximum significance > 2, after re-generation and re-simulation (lumi = 300 pb^{-1})

		Method A	Method B	Method C
npoints		7	7	5
P_{disc}		0.0293542	0.0293542	0.0256786
Cut set	MET $\geq$	102.2	102.2	98.8
	HT $\geq$	505.8	505.8	500.3
	njets $\geq$	6	6	6

very effective in reducing the background. After investigating, among others, the distribution of the number of jets, we had to conclude that this event was an outlier. This lead us to believe that we were, in fact, dealing with a statistical fluctuation.

To study this further, we repeated the entire procedure (generation, simulation, optimization) for those SUSY points with maximum significance greater than 2 (see figure 8.4). Now, the SUSY point mentioned above did not have any events passing our preselection and hence did not influence the optimization anymore. The new results are shown in table 8.6. Methods A and B give us the same optimal cut point and also method C gives a very similar result. The distribution of the significances for the two optimal cut points is shown in figure 8.9. The optimal selection is quite a lot looser than before and is more in line with the results in table 8.5. Plus, we can now discover up to 7 SUSY points. That one SUSY point was thus indeed responsible for the very tight optimal selection that was obtained before. By performing this cross-check we have learned a valuable lession about the performance of sNavigator.

To see the effect of the optimal selections on the background samples listed in table 8.1, we apply the cut set of method A from table 8.6 and the cut sets from table 8.5 to all of these samples. We also apply these selections to the benchmark points LM9 and LM13, both for CMSSW and for Delphes. The resulting number of events are shown in table 8.7. We see that the main background is $t\bar{t}$+jets. There is also a small contribution of W+jets. For this topology the background from Z+jets and QCD is negligable or even non-existing.

As a final step in this test of the sNavigator procedure, we also did the optimization for the cut variables MET, HT, number of jets and lepton p_T, using the set of 1829 SUSY points. The results are shown in table 8.8. The optimization for 300 pb^{-1} via method A was also done for all SUSY points except that one point with large cross section and likelihood. The result of this is shown in table 8.9. To compare with the case where lepton p_T was not used as cut variable, we plot the

Table 8.7: Number of events for the included background samples and the benchmark SUSY points LM9 and LM13 after applying the optimized selection

Sample	Event yield after optimal selection for 300 pb^{-1}	Event yield after optimal selection for 1 fb^{-1}	Event yield after optimal selection for 5 fb^{-1}
DYJetsToLL_D6T_M-50	0	0	0
QCD	0.01	0.05	0
TTJets_D6T	34.9	64.8	146.5
Single top	0.4	0.7	1.1
WJetsToLNu_Z2	2.5	6.3	20.9
ZinvisibleJets_Z2	0	0	0
Total BG	37.8	71.8	168.5
LM13 CMSSW (Delphes)	31.5 (29.4)	91.2 (70.0)	209.4 (175.0)
LM9 CMSSW (Delphes)	13.3 (12.9)	32.1 (35.7)	104.1 (142.9)

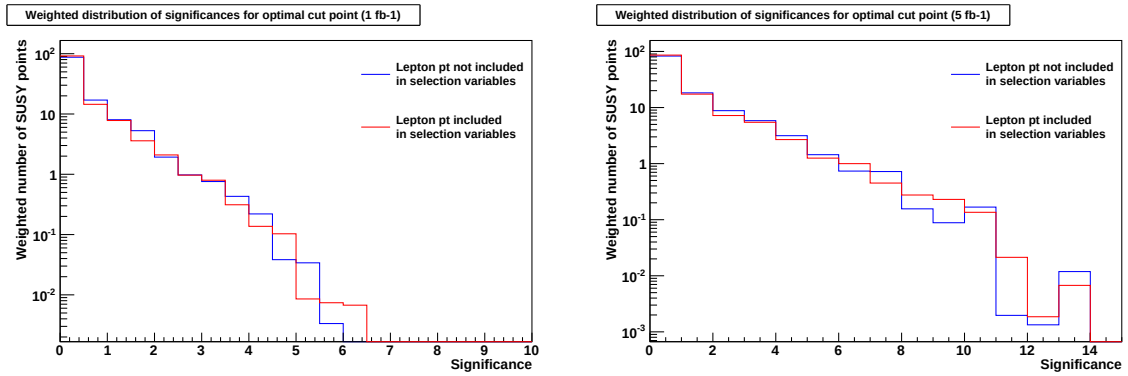
Table 8.8: Optimization results with lepton p_T added to the cut variables

		Method A (300 pb^{-1})	Method B (300 pb^{-1})	Method A (1 fb^{-1})	Method A (5 fb^{-1})
npoints		2	3	92	349
P_{disc}		0.000719343	0.000746678	0.0112294	0.0942412
Cut set	MET $\geq$	139.9	147.8	119.3	114.1
	HT $\geq$	469.7	818.9	421	509
	njets $\geq$	8	7	6	6
	lepton pt $\geq$	16.5	10.0	9.17	8.17

Table 8.9: Optimization results with lepton p_T added to the cut variables, for all SUSY points except the one with large cross section and likelihood

		Method 1 (300 pb^{-1})
npoints		9
P_{disc}		0.000234
Cut set	MET $\geq$	163.1
	HT $\geq$	304
	njets $\geq$	6
	lepton pt $\geq$	10.97

distribution of significances for the optimal cut set in figure 8.10. We find that the lepton p_T is not a good cut variable as it does not improve the optimization. The optimal cut value for the lepton p_T is very low, around the value for our object definition. This illustrates that the p_T spectrum for leptons in mSUGRA is in general rather soft.



(a) 1 fb^{-1} (b) 5 fb^{-1}

Figure 8.10: Distribution of significances for the optimal cut point for the cases where lepton p_T was used or not used in the optimization (see table 8.8). We show the results for a luminosity of 1 fb^{-1} and 5 fb^{-1}.

8.3 Computing resources

Storage Per SUSY point and for 1000 generated events about 30MB of storage is used. The main part of this ($\approx$ 90%) is the Delphes .root file. For the 1829 SUSY points in the processed subset, this amounts to 53GB.

Computing time Most of the procedure can be run in parallel with one job per SUSY point. The average CPU time per job, as a function of the number of sampled SUSY points and the number of generated events, for various parts of the optimization procedure is shown in table 8.10. Each job runs on one node. In real time, it took approximately 2.5 hours on $\mathcal{O}(100)$ nodes for the 1829 jobs to finish.

To determine which cut point is the best one, results for all SUSY points have to be combined. This is done in a ROOT macro. Running this macro for method A takes only about a minute. As the cut set for method B is substantially larger (about 20-30 times as large for 1000 generated events with a preselection efficiency of 2-3%), running the macro for method B can take up to half an hour. Also for method C it takes about half an hour to determine the best cut point.

Table 8.10: Average needed CPU time per job as function of number of included SUSY points N_{SUSY}, number of generated events n_{gen}, number of SUSY MC events after the preselection n_{presel} and total number of background MC events after the preselection n_{BG}

Subprocess	CPU time
Pythia	$\approx 50\text{ms} \cdot n_{gen}$
Delphes	$\approx 15\text{ms} \cdot n_{gen}$
preselection	$\approx 3\text{ms} \cdot n_{gen}$
RA2b selection	$\approx 2.5\text{ms} \cdot n_{gen}$
subtotal	$\approx 70\text{ms} \cdot n_{gen}$
RGS method A step 1	$\approx 0.4\mu\text{s} \cdot n_{presel}(n_{presel} + n_{BG})$
RGS method A step 2	$\approx 0.03\mu\text{s} \cdot N_{SUSY}(n_{presel} + n_{BG})$
RGS method B	$\approx 0.03\mu\text{s} \cdot N_{SUSY} n_{presel}(n_{presel} + n_{BG})$

8.4 Conclusions

From this extensive test of the optimization procedure, we have learned several things:

1. The optimization procedure indeed favours regions of the parameter space with high likelihood.
2. We need to be careful with statistical fluctuations. In the case where there are only few SUSY points with maximum significance larger than 3, a statistical fluctuation can strongly influence the optimization, especially if the corresponding SUSY point also has a high likelihood. Therefore, it is a good idea to generate more events per SUSY point, e.g. 5000 instead of 1000. In order to assess the stability of the optimization procedure in more detail, a bootstrapping method should be executed.
3. Method B will always do as good as method A or better, but it has the down-side of being relatively slow. The higher the number of generated events, the slower method B will be when compared to method A. As method A gives very similar results, we conclude that method A is the preferred method.
4. We can also note that the more Bayesian oriented method C gives similar results as methods A and B. This strengthens our confidence in the optimization procedure.
5. It is important to choose adequate cut variables. With the wrong choice, the efficiency of the optimization will go down.
6. As expected, the optimal cut set depends on the luminosity.
7. In general, for the b-jet(s) + lepton topology, we found that the optimal selection has a high jet multiplicity of $\mathcal{O}(6)$, a high value for HT of $\mathcal{O}(500)$ GeV and a medium value for MET of $\mathcal{O}(100 - 130)$ GeV.

In chapter 9 we will apply the method for the b-jet(s) + single lepton, b-jet(s) + opposite sign dilepton and b-jet(s) + same sign dilepton analyses. We will then also include the likelihood for the SUSY analyses involving leptons.

Chapter 9

A case study with leptons and b-jets

In this chapter the sNavigator method will be applied to three topologies: b-jet(s) + single lepton (section 9.1), b-jet(s) + opposite sign dilepton (section 9.2) and b-jet(s) + same sign dilepton (section 9.3).

Similarly to the validation of the method in chapter 8, we will list the preselection that was applied, the optimal cut point that was obtained and the event count after applying this optimal selection.

We will only use method A as this is the fastest method and gives good results. For the discovery probability, the complete likelihood, including the uncertainties on the background prediction, will be used. Five CMS SUSY analyses will be taken into account. We will again use the RA2b likelihood, now with updated numbers for the background prediction and observed number of events [46]. In addition the likelihoods for the leptonic SUSY analyses will also be included. We will use the single lepton analysis RA4 (both electron and muon channel) [52], the opposite sign dilepton analysis RA5 [53] and the same sign dilepton analysis RA6 [54]. The values we used to calculate the likelihood in equation 4.17 are listed in table 9.1.

As SUSY model, we again consider mSUGRA and use the 1829 signal samples as discussed in section 5.1. Instead of generating 1000 events per sample, we will now generate and simulate 5000 events. This should help reduce the effect of statistical fluctuations and should make the results of the method more robust.

We use the CMSSW physics object definitions as stated in section 5.2. For the leptons, we use the RA4-based definition. In Delphes (see chapter 7) we use the cleaned jet collection, the standard lepton isolation requirement (`IsolFlag = 1`) and the improved version of the b-tagging which was obtained by using b-tag and b-mistag efficiency maps (see section 7.3). All p_T and η requirements are the same as for CMSSW.

Table 9.1: Number of observed events and background prediction for the CMS SUSY analyses that were used in the calculation of the discovery probability. Both statistical and systematic uncertainties are included in the background prediction.

Analysis (i)	Observed count (N_i)	Background estimate ($B_i \pm \delta B_i$)
RA2b	19	10.0 ± 16.1
RA4 electron	17	33.0 ± 13.2
RA4 muon	26	29.2 ± 10.2
RA5	0	1.2 ± 0.8
RA6	1	1.3 ± 0.9

9.1 b-jet(s) + single lepton

9.1.1 Preselection

As preselection for the b-jet(s) + single lepton topology we required at least three jets, at least one b-jet and exactly one isolated lepton (electron or muon).

The efficiency of this preselection for the considered background samples is shown in table 9.2. We can note that the dominant backgrounds are $t\bar{t}$ and QCD and to a lesser extent also single top and W production.

The distribution of the efficiency for the SUSY samples is shown in figure 9.1. From this we can see that the average total preselection efficiency is of the order of a few percent. The distribution of the effective cross section after applying the preselection cuts can be viewed in figure 9.2.

A comparison between the preselection efficiencies obtained via CMSSW and Delphes for the benchmark points LM9 and LM13 is shown in table 9.3. The agreement is reasonable.

Table 9.2: b-jet(s)+ single lepton preselection efficiencies and effective cross section

Background	σ (pb)	Preselection efficiency			σ_{eff} (pb)
		njets ≥ 3	nleptons $= 1$	nbtaggedjets ≥ 1	
QCD	3.7e7	2.2e-2	3.2e-6	5.8e-7	21.5
Singletop	33	3.0e-1	8.8e-2	6.1e-2	2.0
TTJets_D6T	165	7.9e-1	1.9e-1	1.5e-1	24.0
ZinvisibleJets_Z2	4500	3.1e-3	1.4e-6	4.6e-7	0.0021
DYJetsToLL_D6T_M-50	3048	5.9e-3	2.0e-3	1.9e-4	0.57
WJetsToLNu_Z2	31314	3.8e-3	1.5e-3	1.4e-4	4.4
Total					52.4

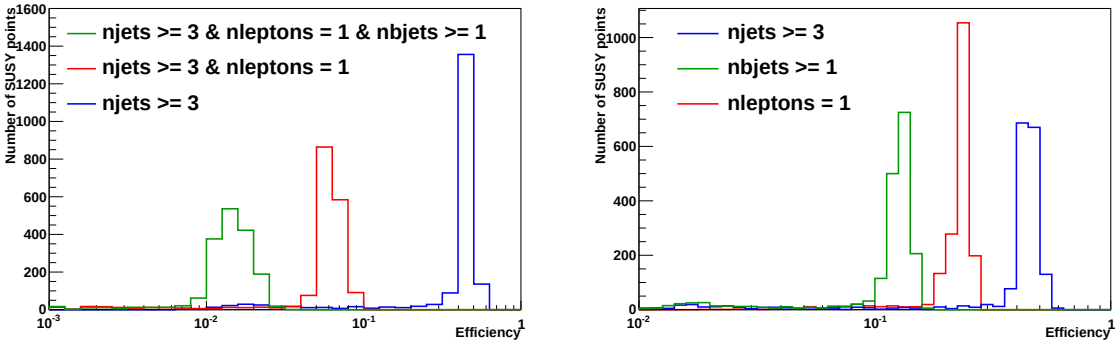
Table 9.3: Comparison of preselection efficiencies in Delphes and CMSSW for the SUSY benchmark point LM9 and LM13.

		njets ≥ 3	nbjets ≥ 1	single lepton	OS dileptons	SS dileptons
LM9	Delphes	0.240	0.140	0.027	0.007	0.0016
	CMSSW	0.227	0.142	0.023	0.004	0.0011
LM13	Delphes	0.648	0.405	0.099	0.001	0.0050
	CMSSW	0.666	0.433	0.090	0.001	0.0031

9.1.2 sNavigator results

We will perform the optimization procedure for three luminosities (300 pb^{-1}, 1 fb^{-1}, 5 fb^{-1}) and for three sets of cut variables:

1. the missing transverse energy MET, the scalar sum of all jet p_T's HT (see equation 5.3) and the number of jets n_{jets}
2. MET, HT, n_{jets} and lepton p_T
3. the so-called missing transverse energy significance $S_{MET} = \frac{MET}{\sqrt{HT_{20}}}$, HT_{20}, n_{jets}. In the calculation of HT_{20} we take all jets with $p_T > 20$ GeV into account.



(a) Cumulative cuts

(b) Separate cuts

Figure 9.1: Distribution of the b-jet(s) + single lepton preselection efficiencies

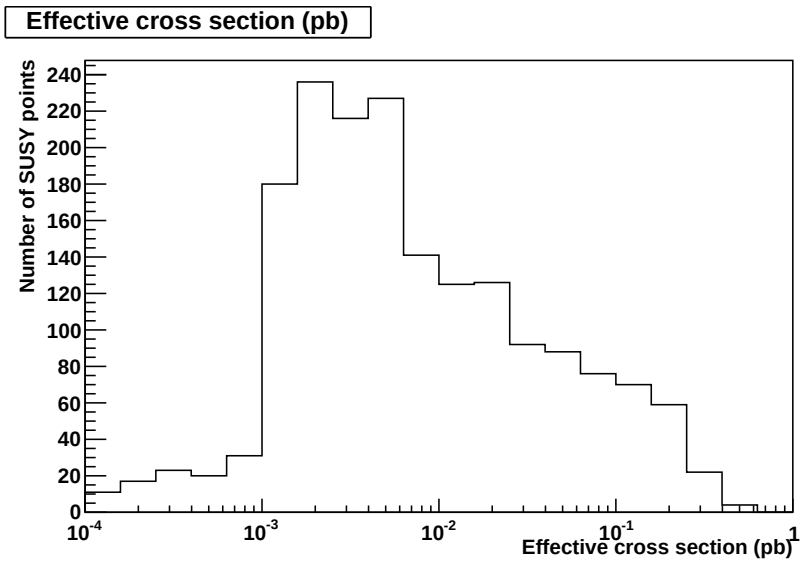


Figure 9.2: Effective cross section after the b-jet(s) + single lepton preselection cuts

This set of variables is based on the selection of the RA4 analysis. The S_{MET} variable is used instead of MET because the correlation between S_{MET} and HT_{20} is smaller than between MET and HT_{20}.

The resulting optimal cut points are shown in tables 9.4, 9.5 and 9.6. In these tables we also list the values of the discovery probability and the corresponding number of SUSY points. In figures 9.3, 9.4 and 9.5 we show the distribution of the significance for the optimal cut points.

Looking at the values of the discovery probability, we can conclude that the cut set that includes the lepton p_T as cut variable, gives the smallest values for P_{disc}. This again illustrates that the lepton p_T spectrum of mSUGRA is in general soft. There is no clear conclusion on which set of cut variables gives the best results. For higher luminosities the first set yields slightly better results, while for a luminosity of 300 pb^{-1} the third set gave the highest value of P_{disc}.

We next proceed by applying the cut point, obtained for 1 fb^{-1} and by using the third set of cut variables, to our background samples and to the LM9 and LM13 benchmark SUSY points. The results of this are shown in table 9.7.

The main Standard Model backgrounds surviving this selection are $t\bar{t}$ and W production. There is also a small contribution of single top. QCD and Z production processes are negligible.

Table 9.4: Optimal cuts for the b-jet(s) + single lepton selection, using MET, HT and n_{jets} as cut variables

Cuts	300 pb^{-1}	1 fb^{-1}	5 fb^{-1}
$MET \geq$	59	91	96
$HT \geq$	340	416	515
$n_{jets} \geq$	6	5	5
P_{disc}	1.21e-11	7.57e-10	7.57e-09
number of SUSY points	1	38	237

Table 9.5: Optimal cuts for the b-jet(s) + single lepton selection, using MET, HT, n_{jets} and lepton p_T as cut variables

Cuts	300 pb^{-1}	1 fb^{-1}	5 fb^{-1}
$MET \geq$	67	92	92
$HT \geq$	373	449	449
$n_{jets} \geq$	6	5	5
lepton $p_T \geq$	8.1	8.0	8.0
P_{disc}	1.21e-11	6.40e-10	7.29e-09
number of SUSY points	1	31	229

Table 9.6: Optimal cuts for the b-jet(s) + single lepton selection, using S_{MET}, HT and n_{jets} as cut variables

Cuts	300 pb^{-1}	1 fb^{-1}	5 fb^{-1}
$S_{MET} \geq$	3.02	**3.99**	3.12
$HT_{20} \geq$	522	**451**	546
$n_{jets} \geq$	5	**4**	5
P_{disc}	2.20e-11	6.20e-10	7.36e-09
number of SUSY points	1	31	230

Table 9.7: Event yield for the b-jet(s) + single lepton selection after applying the optimal cut, determined via cut variables S_{MET}, HT_{20}, n_{jets} and for a luminosity of 1 fb^{-1}.

Sample	Event yield after optimal selection ($S_{MET} \geq 3.99$, $HT_{20} \geq 451$, $n_{jets} \geq 4$) for 1 fb^{-1}
QCD	2.3
Single top	16.9
TTJets_D6T	532.7
ZinvisibleJets_Z2	0
DYJetsToLL_D6T_M-50	1.2
WJetsToLNu_Z2	96.2
Total BG	649.3
LM13 CMSSW	301.4
LM9 CMSSW	69.6

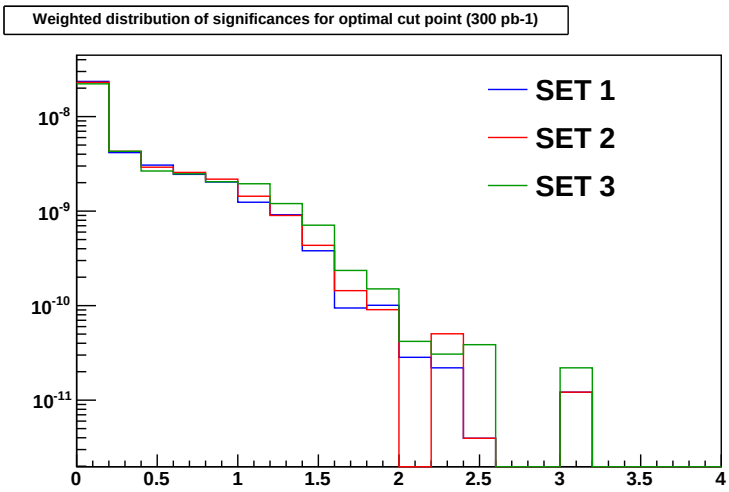


Figure 9.3: Distribution of the significance for the optimal cut points obtained for 300 pb^{-1}

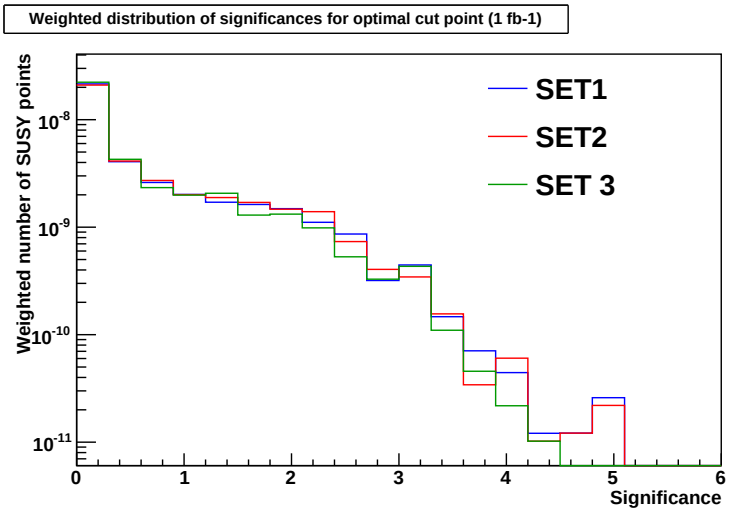


Figure 9.4: Distribution of the significance for the optimal cut points obtained for 1 fb^{-1}

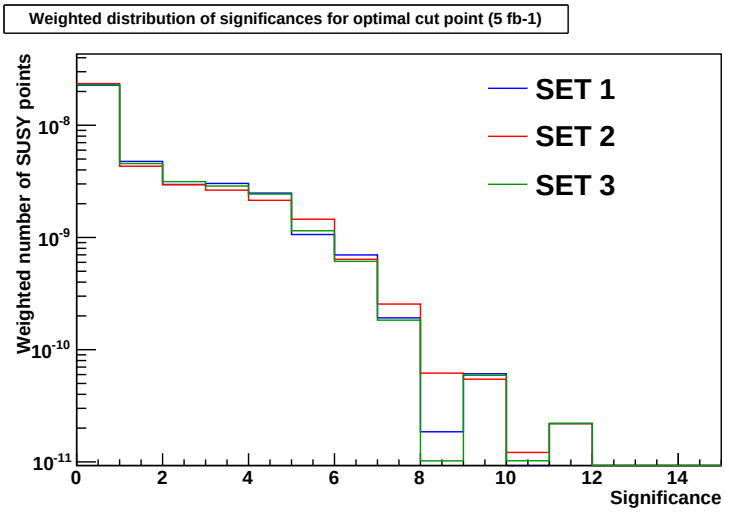


Figure 9.5: Distribution of the significance for the optimal cut points obtained for 5 fb^{-1}

To get an idea of which regions of the parameter space we are sensitive to, we show the significance for the optimal cut point obtained with set 3 for 1 fb^{-1}, as a function of the mSUGRA parameters (m_0, $m_{1/2}$, A_0, $\tan\beta$), see figures 9.6 - 9.9. In addition we also show the significance as a function of the mass of various supersymmetric particles, see figures 9.10 - 9.16. We see that the parameter with the strongest effect on the significance, is $m_{1/2}$, the unified mass of the gauginos at the GUT scale. This is also reflected in the strong dependence on the gluino, neutralino and chargino mass.

In figures 9.21 - 9.28, we show the fraction of SUSY points with significance larger than three, weighted with their likelihood, per bin of the mSUGRA parameters and the masses of the gluino, lightest higgs boson, lightest neutralino and lightest chargino. This fraction was calculated as follows

$$\text{fraction in bin } i = \frac{\sum_{p_i \atop \mathcal{S}>3} \mathcal{L}_{p_i}}{\sum_{p_i} \mathcal{L}_{p_i}} , \tag{9.1}$$

where p_i runs over all SUSY points in bin i and $\mathcal{L}$ represents the product of all included likelihoods (see equation 4.17). From these figures we can conclude that for 1 fb^{-1} and using the optimal cuts determined with cut set 3, we are sensitive to gluino masses up to 450 GeV, lightest neutralino masses up to 65 GeV and lightest chargino masses up to 120 GeV.

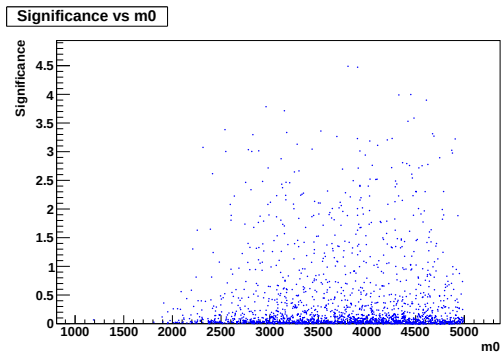


Figure 9.6: Significance versus m_0 (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

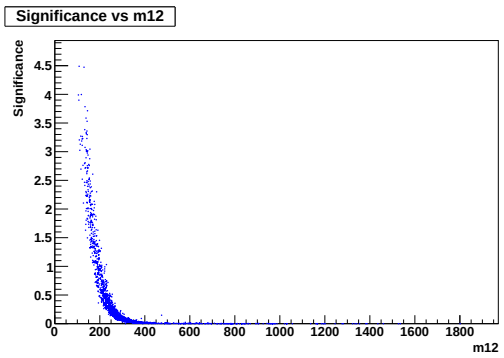


Figure 9.7: Significance versus $m_{1/2}$ (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

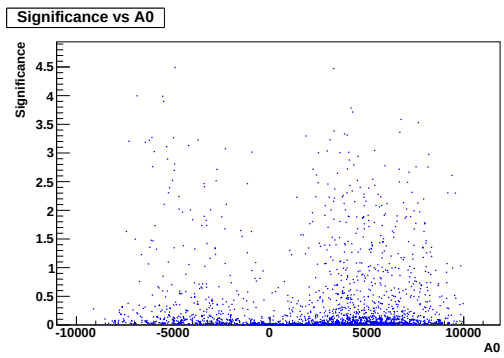


Figure 9.8: Significance versus A_0 for the optimal cut point obtained with set 3 for 1 fb^{-1}

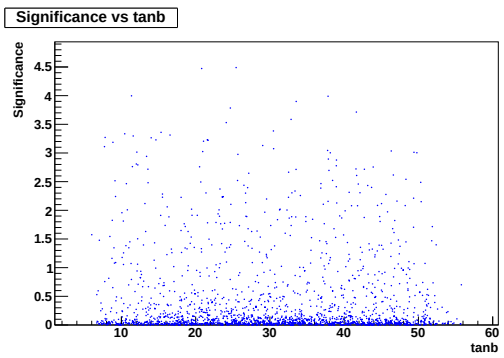


Figure 9.9: Significance versus $\tan\beta$ for the optimal cut point obtained with set 3 for 1 fb^{-1}

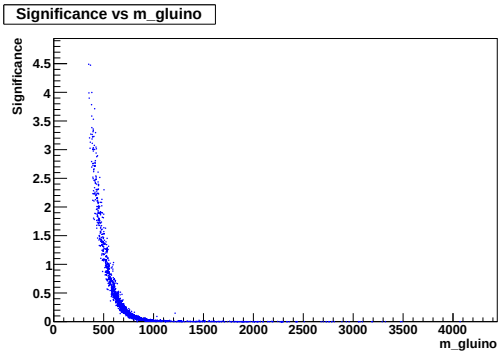


Figure 9.10: Significance versus gluino mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

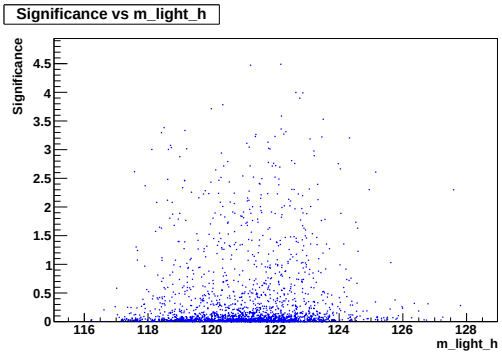


Figure 9.11: Significance versus lightest higgs mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

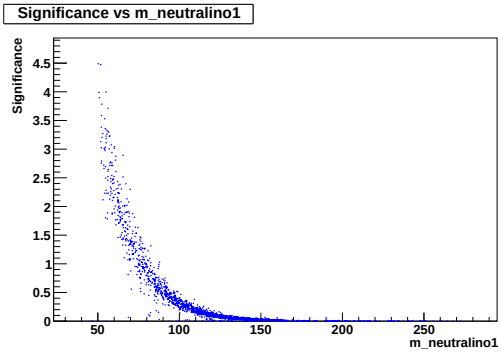


Figure 9.12: Significance versus lightest neutralino mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

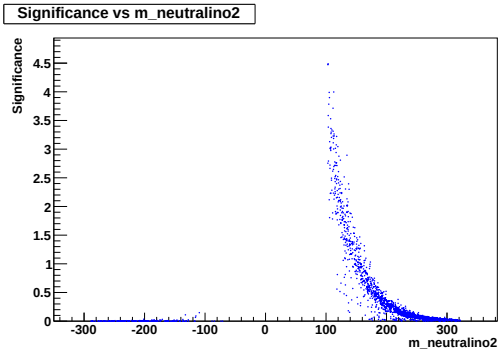


Figure 9.13: Significance versus second lightest neutralino mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

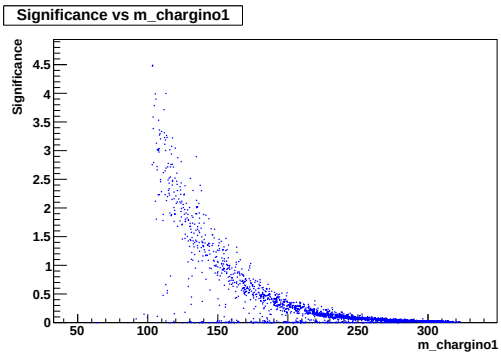


Figure 9.14: Significance versus lightest chargino mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

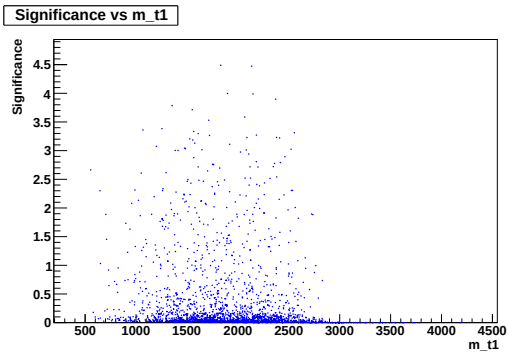


Figure 9.15: Significance versus lightest stop mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

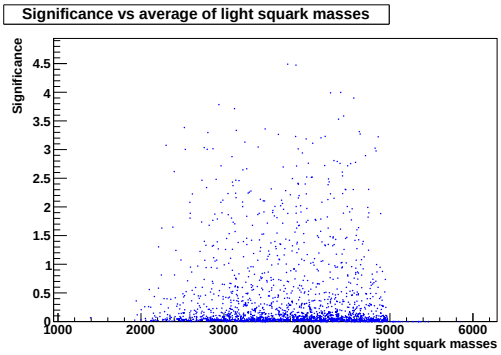


Figure 9.16: Significance versus the average of the light squark ($\widetilde{u}$,$\widetilde{d}$,$\widetilde{c}$,$\widetilde{s}$) masses (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

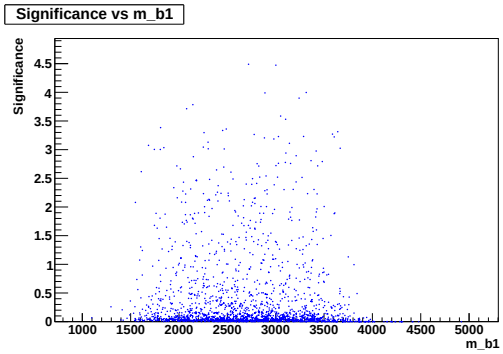


Figure 9.17: Significance versus lightests sbottom mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

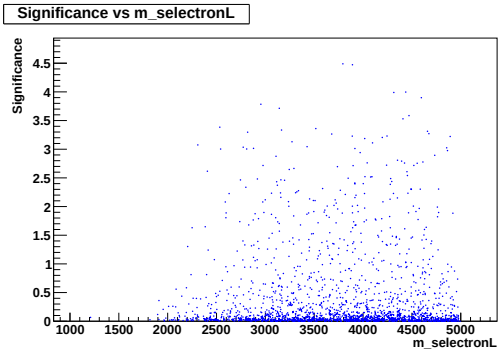


Figure 9.18: Significance versus left-handed selectron mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

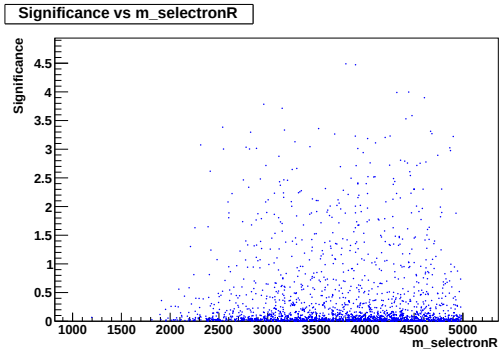


Figure 9.19: Significance versus right-handed selectron mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

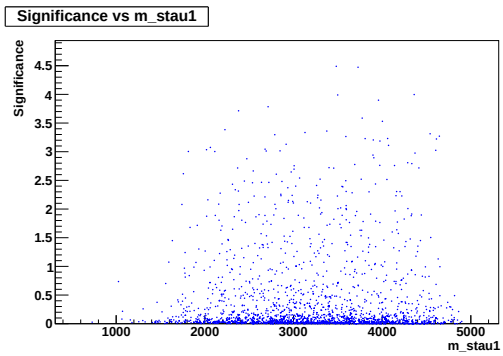


Figure 9.20: Significance versus lightest stau mass (in GeV) for the optimal cut point obtained with set 3 for 1 fb^{-1}

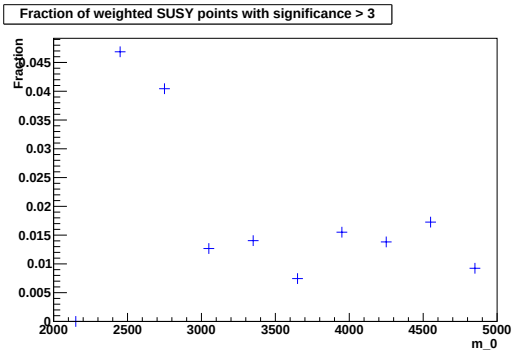


Figure 9.21: Fraction of weighted SUSY points with significance > 3 as a function of m_0 (in GeV) for the optimal cut point determined via set 3 for 1 fb^{-1}

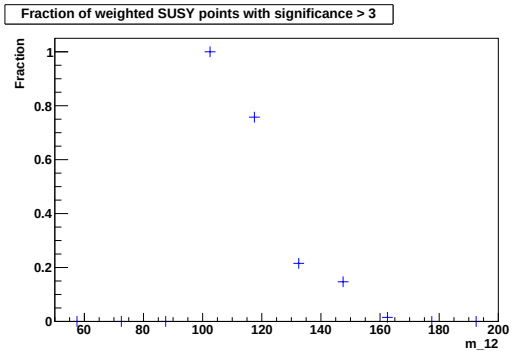


Figure 9.22: Fraction of weighted SUSY points with significance > 3 as a function of $m_{1/2}$ (in GeV) for the optimal cut point determined via set 3 for 1 fb^{-1}

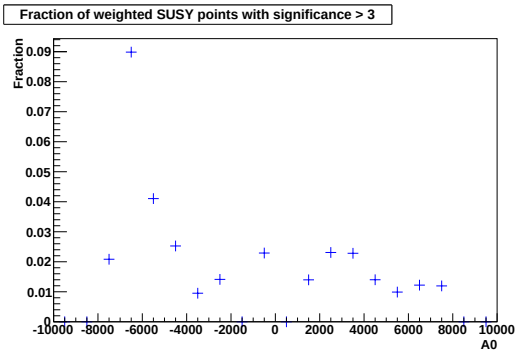


Figure 9.23: Fraction of weighted SUSY points with significance > 3 as a function of A_0 for the optimal cut point determined via set 3 for 1 fb^{-1}

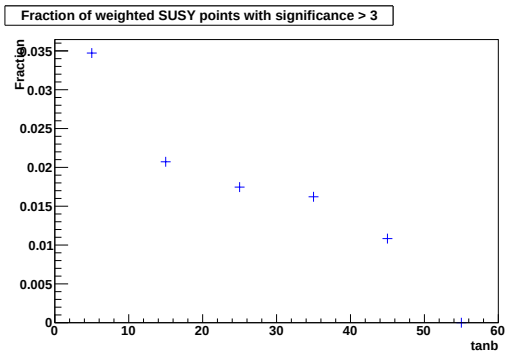


Figure 9.24: Fraction of weighted SUSY points with significance > 3 as a function of $\tan\beta$ for the optimal cut point determined via set 3 for 1 fb^{-1}

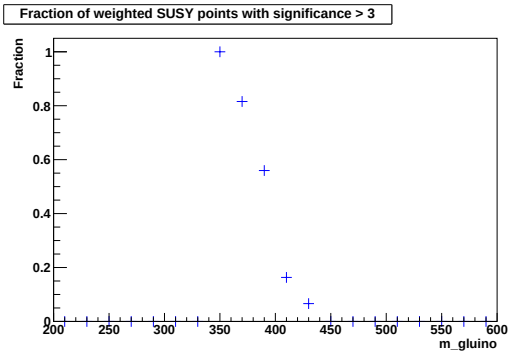


Figure 9.25: Fraction of weighted SUSY points with significance > 3 as a function of gluino mass (in GeV) for the optimal cut point determined via set 3 for 1 fb^{-1}

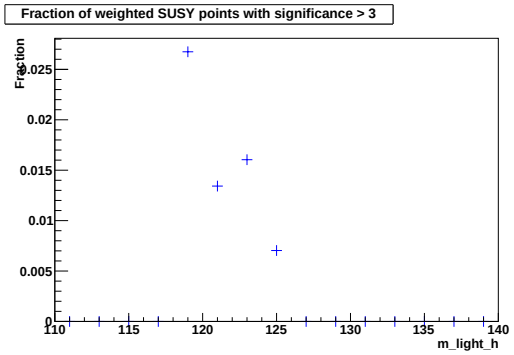


Figure 9.26: Fraction of weighted SUSY points with significance > 3 as a function of lightest higgs mass (in GeV) for the optimal cut point determined via set 3 for 1 fb^{-1}

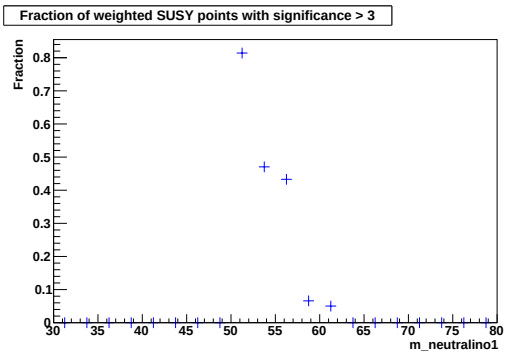


Figure 9.27: Fraction of weighted SUSY points with significance > 3 as a function of lightest neutralino mass (in GeV) for the optimal cut point determined via set 3 for 1 fb^{-1}

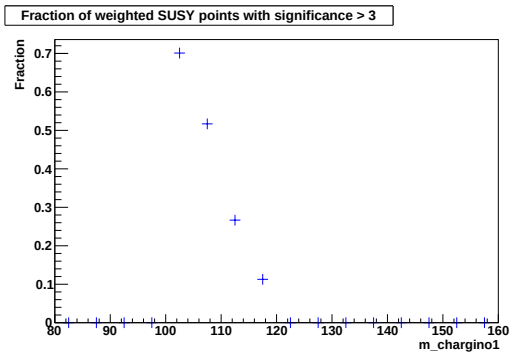


Figure 9.28: Fraction of weighted SUSY points with significance > 3 as a function of lightest chargino mass (in GeV) for the optimal cut point determined via set 3 for 1 fb^{-1}

9.2 b-jet(s) + opposite sign dilepton

9.2.1 Preselection

As preselection for the b-jet(s) + opposite sign dilepton topology we require the presence of at least three jets, at least one b-jet and two leptons (electron or muon) with opposite charge.

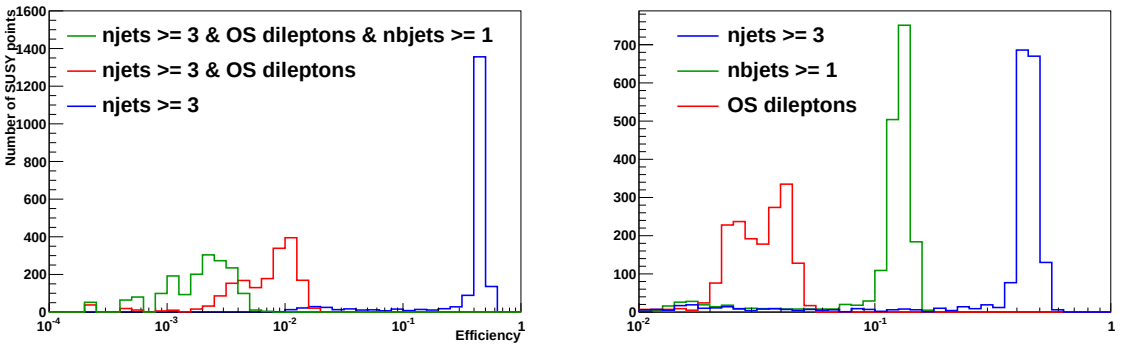
The efficiency of this preselection for the considered background samples is shown in table 9.8. We can note that the dominant background comes from $t\bar{t}$. There is also a smaller contribution from Z/γ + jets.

The distribution of the efficiency for the SUSY samples is shown in figure 9.29. From this we can see that the average total preselection efficiency is of the order of a few per mille. The distribution of the effective cross section after applying the preselection cuts can be viewed in figure 9.30.

For the comparison between the preselection efficiencies obtained via CMSSW and Delphes for the benchmark points LM9 and LM13 we refer to table 9.3. We see that there is a reasonable agreement, but that there is still some room for improvement.

Table 9.8: b-jet(s) + opposite sign dilepton preselection efficiencies and effective cross section

Background	σ (pb)	Preselection efficiency			σ_{eff} (pb)
		njets ≥ 3	OS dileptons	nbtaggedjets ≥ 1	
QCD	3.7e7	2.2e-2	0	0	0
Singletop	33	3.0e-1	1.9e-3	1.4e-3	0.015
TTJets_D6T	165	7.9e-1	6.4e-3	5.1e-3	0.85
ZinvisibleJets_Z2	4500	3.1e-3	0	0	0
DYJetsToLL_D6T_M-50	3048	5.9e-3	9.8e-4	1.1e-4	0.35
WJetsToLNu_Z2	31314	3.8e-3	0	0	0
Total					1.21



(a) Cumulative cuts

(b) Separate cuts

Figure 9.29: Distribution of the b-jet(s) + opposite sign dilepton preselection efficiencies

9.2.2 sNavigator results

We have again performed the optimization procedure for three luminosities (300 pb^{-1}, 1 fb^{-1}, 5 fb^{-1}) and for the same three sets of cut variables:

1. MET, HT, n_{jets}

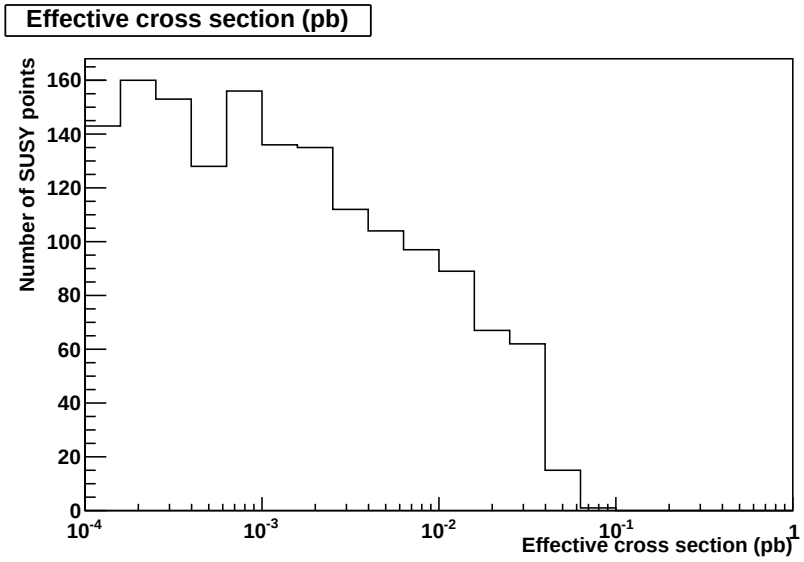


Figure 9.30: Effective cross section after the b-jet(s) + opposite sign dilepton preselection cuts

2. MET, HT, n_{jets}, lepton p_T
3. S_{MET}, HT_{20}, n_{jets}.

The resulting optimal cut points with the corresponding value for the discovery probability and number of SUSY points, are shown in tables 9.9, 9.10 and 9.11.

For a luminosity of 300 pb^{-1}, no cut point for which there were SUSY points with $\mathcal{S} > 3$ was found. This illustrates that the b-jet(s) + opposite sign dilepton topology is a difficult topology and requires a large luminosity to be able to see a possible excess. The first and third set of cut variables can achieve around the same value of the discovery probability, while the second set – with lepton p_T – gives a lower value for P_{disc}.

Compared to the b-jet(s) + single lepton topology, the optimal cut values for MET, HT and S_{MET} are softer. The jet multiplicity is still quite high. The value for the discovery probability is in this case about half of what it is for the b-jet(s) + single lepton topology.

We show the results of the application of the optimal cut point, obtained for 1 fb^{-1} and the third set of cut variables, to our background samples and to the LM9 and LM13 benchmark SUSY points in table 9.12. $t\bar{t}$ is essentially the only background that remains. The Z/γ + jets process that also contributed to the background after the preselection, has now been eliminated completely.

Table 9.9: Optimal cuts for the b-jet(s) + opposite sign dilepton selection, using MET, HT and n_{jets} as cut variables

Cuts	300 pb^{-1}	1 fb^{-1}	5 fb^{-1}
MET $\geq$ HT $\geq$ njets $\geq$	/ / /	33 410 4	50 397 5
P_{disc}	/	1.22e-10	3.72e-09
number of SUSY points	/	7	126

Table 9.10: Optimal cuts for the b-jet(s) + oppposite sign dilepton selection, using MET, HT, n_{jets} and lepton p_T as cut variables

Cuts	300 pb^{-1}	1 fb^{-1}	5 fb^{-1}
MET $\geq$ HT $\geq$ njets $\geq$ lepton pt $\geq$	/ / / /	46 525 4 19.6	61 330 4 16.3
P_{disc}	/	7.01-11	3.14e-09
number of SUSY points	/	4	108

Table 9.11: Optimal cuts for the b-jet(s) + opposite sign dilepton selection, using S_{MET}, HT and n_{jets} as cut variables

Cuts	300 pb^{-1}	1 fb^{-1}	5 fb^{-1}
S_{MET} $\geq$ HT (jet pt $>$ 20) $\geq$ njets $\geq$	/ / /	**1.93** **444** **4**	1.55 414 5
P_{disc}	/	1.25e-10	3.61e-09
number of SUSY points	/	6	123

Table 9.12: Event yield for the b-jet(s) + opposite sign dilepton selection after applying the optimal cut, determined via cut variables S_{MET}, HT_{20}, n_{jets} and for a luminosity of 1 fb^{-1}.

Sample	Event yield after optimal selection ($S_{MET} \geq 1.93$, $HT_{20} \geq 444$, $n_{jets} \geq 4$) for 1 fb^{-1}
DYJetsToLL_D6T_M-50 TTJets_D6T Single top	0 41.7 0.04
Total BG	41.7
LM13 CMSSW LM9 CMSSW	3.9 17.6

9.3 b-jet(s) + same sign dilepton

9.3.1 Preselection

For the b-jet(s) + same sign dilepton topology we require as preselection that there are at least three jets, at least one b-jet and two leptons (electron or muon) with the same charge present in the event.

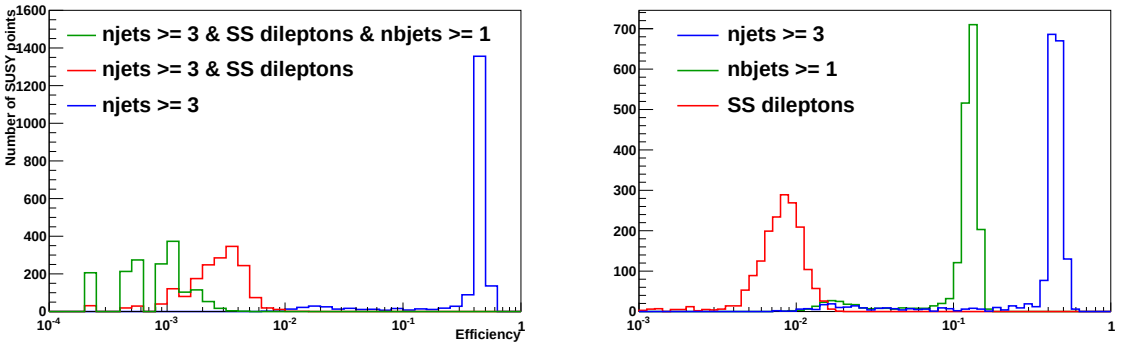
The efficiency of this preselection for our background samples is shown in table 9.13. We observe that the dominant background comes from $t\bar{t}$. There is also a small contribution from W and single top production.

The distribution of the efficiency for the SUSY samples is shown in figure 9.31. From this we can see that the average total preselection efficiency is of the order of one per mille. The distribution of the effective cross section after applying the preselection cuts can be viewed in figure 9.32.

For the comparison between the preselection efficiencies obtained via CMSSW and Delphes for the benchmark points LM9 and LM13 we refer to the last column of table 9.3. We see that there is again some room for improvement for the leptons.

Table 9.13: b-jet(s) + same sign dilepton preselection efficiencies and effective cross section

Background	σ (pb)	Preselection efficiency			σ_{eff} (pb)
		njets ≥ 3	SS dileptons	nbtaggedjets ≥ 1	
QCD	3.7e7	2.2e-2	0	0	0
Singletop	33	3.0e-1	1.2e-4	9.3e-5	0.0031
TTJets_D6T	165	7.9e-1	8.2e-4	6.6e-4	0.11
ZinvisibleJets_Z2	4500	3.1e-3	0	0	0
DYJetsToLL_D6T_M-50	3048	5.9e-3	2.4e-6	0	0
WJetsToLNu_Z2	31314	3.8e-3	5.3e-7	1.3e-7	0.0042
Total					0.12



(a) Cumulative cuts

(b) Separate cuts

Figure 9.31: Distribution of the b-jet(s) + opposite sign dilepton preselection efficiencies

9.3.2 sNavigator results

We have again used the same three luminosities (300 pb^{-1}, 1 fb^{-1}, 5 fb^{-1}) and sets of cut variables in the optimization:

1. MET, HT, n_{jets}

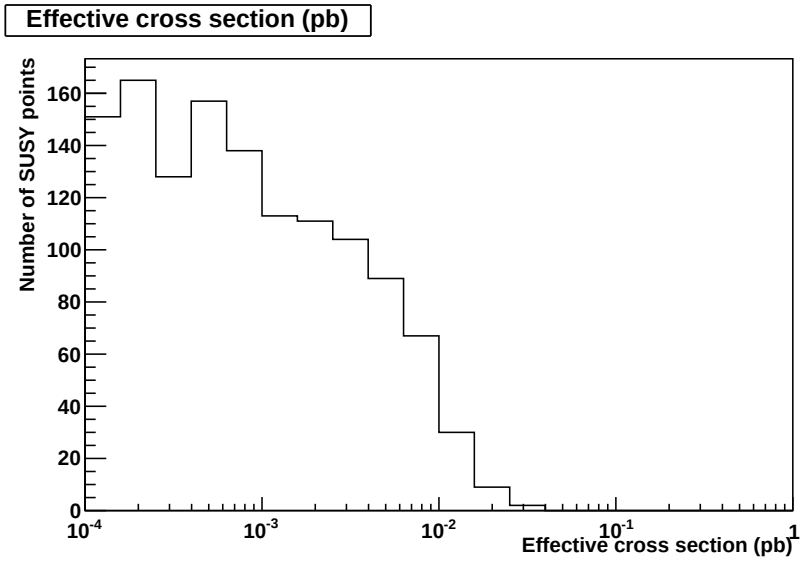


Figure 9.32: Effective cross section after the b-jet(s) + same sign dilepton preselection cuts

2. MET, HT, n_{jets}, lepton p_T
3. S_{MET}, HT_{20}, n_{jets}.

The resulting optimal cut points with the corresponding value for the discovery probability and number of SUSY points, can be viewed in tables 9.14, 9.15 and 9.16. The values for the discovery probability are similar to those that were obtained in the previous section for the b-jet(s) + opposite sign dilepton topology.

The results of the application of the optimal cut point, obtained for 1 fb^{-1} the third set of cut variables, to our background samples and to the LM9 and LM13 benchmark SUSY points are shown in table 9.17. From this table we see that essentially only $t\bar{t}$ remains as background. We can also note that the resulting number of signal events for 1 fb^{-1} is quite low.

Table 9.14: Optimal cuts for the b-jet(s) + same sign dilepton selection, using MET, HT and n_{jets} as cut variables

Cuts	300 pb^{-1}	1 fb^{-1}	5 fb^{-1}
MET $\geq$	/	24	69
HT $\geq$	/	392	370
njets $\geq$	/	4	4
P_{disc}	/	6.01e-11	2.89-09
number of SUSY points	/	3	99

Table 9.15: Optimal cuts for the b-jet(s) + same sign dilepton selection, using MET, HT, n_{jets} and lepton p_T as cut variables

Cuts	300 pb^{-1}	1 fb^{-1}	5 fb^{-1}
MET $\geq$	/	38	60
HT $\geq$	/	357	400
njets $\geq$	/	4	4
lepton pt $\geq$	/	17.9	14.0
P_{disc}	/	4.80e-11	2.80e-09
number of SUSY points	/	3	94

Table 9.16: Optimal cuts for the b-jet(s) + same sign dilepton selection, using S_{MET}, HT and n_{jets} as cut variables

Cuts	300 pb^{-1}	1 fb^{-1}	5 fb^{-1}
S_{MET} $\geq$	/	**0.58**	3.0
HT (jet pt $>$ 20) $\geq$	/	**423**	400
njets $\geq$	/	**4**	4
P_{disc}	/	9.72e-11	2.88e-09
number of SUSY points	/	4	98

Table 9.17: Event yield for the b-jet(s) + same sign dilepton selection after applying the optimal cut, determined via cut variables S_{MET}, HT_{20}, n_{jets} and for a luminosity of 1 fb^{-1}.

Sample	Event yield after optimal selection ($S_{MET} \geq 0.58$, $HT_{20} \geq 423$, $n_{jets} \geq 4$) for 1 fb^{-1}
TTJets_D6T	6.06
Single top	0.04
WJetsToLNu_Z2	0
Total BG	6.1
LM13 CMSSW	10.8
LM9 CMSSW	5.4

Discussion and outlook

The way to search for signals of New Physics models, such as supersymmetry, is not clear-cut. The reason for this is their multidimensional nature. Depending on the specific combination of parameters, the signal can look quite different. Optimizing an analysis for these New Physics phenomena, is thus no trivial task. One could, for example, optimize for a specific point in the parameter space. This is, however, far from optimal. There is no guarantee that a selection which is optimal for one point in the parameter space, will also be optimal for other points or even an entire region. Current searches have therefore opted to base the analysis on accurate knowledge of the detector performance.

In this thesis, we have proposed a method that addresses the question of how to optimize for a parameter space region. Our method, the sNavigator, allows to find an event selection that maximizes the probability to discover a certain multidimensional New Physics model. We used a Bayesian approach to include information obtained from prior measurements. In the calculation of the discovery probability, we gave a larger weight to areas of the parameter space that have not been excluded yet and thus have a higher likelihood.

In order to find the event selection that maximizes this discovery probability, we used the principle of the Random Grid Search (RGS). RGS allows to find an optimal rectangular cut in an efficient way, by considering signal events as possible cut points and choosing the best one among these. We used this principle to first obtain the best cut point for each considered point of the parameter space. For each of these cut points, we then computed the discovery probability for the chosen parameter space region. We define the optimal cut point as the cut point that gives us the largest value for the discovery probability.

We have studied the behavior of RGS in the optimization for a single point in mSUGRA space. We learned that we need to be careful when choosing the variables we wish to use as cut variables. These variables should have a good discriminating power between signal and background. We also need to watch out for correlations between the variables. As we are cutting in multiple dimensions at the same time, correlations can have a counter-intuitive effect. We observed that including more variables does not necessarily mean that we obtain better results.

To simulate events for various points in the chosen parameter space region (sampled via a Markov Chain Monte Carlo method), we used Delphes, a framework for the fast simulation of a generic collider experiment. We have tuned Delphes to mimic the CMS detector as good as possible and assessed its performance. Overall, the out-of-the-box performance was quite good. We were able to improve b-tagging by including b-tag and b-mistag efficiency maps obtained via Bayesian Neural Networks. Isolated leptons are, when defined according to the loose, RA2b-based definition, described well within Delphes. However, the agreement worsens when one uses a tighter isolation cut. In the future this behavior can be improved by also implementing efficiency maps for lepton identification and isolation.

During the validation of sNavigator, we noticed that statistical fluctuations can influence the procedure significantly. This mainly happens for low luminosities when there are few points in the parameter space that can be discovered and in the case one of these points has a high likelihood. In

our test we noticed a point with a large likelihood and cross section, and with only one Monte Carlo event passing our preselection criteria. After some investigating, it was clear that this event was an outlier. When we generated more events for this point, none of these passed the preselection cuts. The fact that this point could be discovered was thus merely a statistical fluctuation. Although the method does what it is supposed to do – favour points with high likelihood –, we need to be careful with such fluctuations. Generating more events per parameter point should help, but might not solve the problem completely. A more thorough test of the stability of the method, e.g. by performing a bootstrapping procedure, is needed. We can also note that for higher luminosities, the effect of such points is mitigated as there are more points that can be discovered. Their combined likelihoods can become larger than the likelihood of that one point.

To conclude this thesis, we have applied the sNavigator method for the mSUGRA model and for three distinct topologies: b-jet(s) + single lepton, b-jet(s) + opposite sign dileptons and b-jet(s) + same sign dileptons. These topologies have hardly been explored within the CMS SUSY group. For each topology we computed the optimal cut point for three luminosities (300 pb^{-1}, 1 fb^{-1}, 5 fb^{-1}) and three sets of cut variables. We found that putting additional, hard, cuts on the lepton p_T is not a good idea. This illustrates that the lepton p_T spectrum for mSUGRA is in general rather soft. For all three topologies, the main background after applying the optimal cuts is $t\bar{t}$. For the b-jet(s) + dilepton searches and for a luminosity of 300 pb^{-1}, no good cut point was found. This illustrates that these channels are very challenging and require large luminosities. In many scenarios these will not be the 'discovery channels', but in case an excess is observed in another channel they can provide valuable insight into the exact nature of the possible signal.

Overall, we can state that the sNavigator method works well. It indeed favours regions of the parameter space with high likelihood and finds sensible cuts. With our optimal event selection, it was possible to discover a large number of the sampled SUSY points, especially taking into account the challenging nature of the chosen topologies.

As the sNavigator method can be applied to any model of new physics and any search topology, it seems natural to also apply it for a different SUSY model, such as pMSSM. This is a much more general model, without the strong assumptions of mSUGRA. It will be very interesting to see what the optimal cut values are for this model and how much they differ from the results obtained for mSUGRA. This will give us insight in the differences between both models. One can of course also apply the method for other search topologies. In case there are existing SUSY analyses for such topologies, it would be instructive to compare their selection to the optimal selection as obtained via sNavigator.

Now that the method is in place, the logical next step is to perform the actual analysis. We should decide on the New Physics model, cut variables and luminosity for which we wish to optimize. Then we should run sNavigator to obtain the optimal cut values. This concludes the analysis design stage. The following step is then to apply the resulting selection and do a thorough background estimation, including systematic uncertainties et cetera. Then we can compare with data. In the case where no excess over the Standard Model prediction is seen, we can proceed to set limits. The result of the analysis can then also be used in a second iteration of the entire procedure. We can simply include it in the calculation of the discovery probability and run sNavigator again.

Nederlandstalige samenvatting

Het Standaard Model van de deeltjesfysica beschrijft alle gekende deeltjes en interacties (behalve de zwaartekracht) en blijkt erg succesvol in het voorspellen en verklaren van zo goed als alle uitgevoerde experimenten. Toch zijn er nog enkele onopgeloste zaken die doen vermoeden dat het Standaard Model onvolledig is. In de loop der tijd zijn veel uitbreidingen voorgesteld. Een elegante en populaire uitbreiding, en het onderwerp van deze thesis, is supersymmetrie (SUSY).

Supersymmetrie is een symmetrie tussen bosonen en fermionen die leidt tot een verdubbeling van het aantal deeltjes. Het bestaan van deze nieuwe deeltjes, de superpartners van de gewone deeltjes, kan enkel bevestigd worden door het uitvoeren van experimenten. Met het opstarten van de Large Hadron Collider in Genève, is de zoektocht naar deze supersymmetrische deeltjes in een stroomversnelling gekomen. Er is helaas een struikelblok: supersymmetrische modellen hebben in het algemeen erg veel vrije parameters. Afhankelijk van de precieze waarden voor deze parameters, kan SUSY er erg verschillend uitzien. Dit maakt het moeilijk om naar signalen van supersymmetrie te zoeken en de hamvraag is dus "Wat is de beste manier om signalen van supersymmetrie op te sporen?"

In deze thesis hebben we een methode voorgesteld die een antwoord biedt op deze vraag. Onze methode, de sNavigator, laat toe een event selectie te vinden die de kans op ontdekking van een bepaald multidimensionsaal model van nieuwe fysica, zoals supersymmetrie, zo groot mogelijk maakt. Om dit te bereiken, gebruiken we een Bayesiaanse aanpak om informatie verkregen uit voorgaande metingen te benutten en om op een efficiënte manier door de parameterruimte te navigeren. Het heeft namelijk geen zin om te optimaliseren voor een gebied dat al is uitgesloten door vorige experimenten. Kort gezegd maximaliseren we het gebied binnen het bereik van ontdekking, waarbij we meer rekening houden met punten die, volgens eerdere metingen, een grotere probabiliteit hebben.

Elk punt in de parameterruimte zal aanleiding geven tot events met welbepaalde kenmerken. Events kunnen beschreven worden door een set van variabelen $x, y, \ldots$, waarvan de distributie er naargelang het precieze proces anders zal uitzien. Door cuts te plaatsen op deze variabelen ($x > x_0, y > y_0, \ldots$), kunnen we het gezochte signaal proberen onderscheiden van de gekende achtergrondprocessen. Deze set van cuts noemt men de event selectie.

Om de selectie die de ontdekkingskans maximaliseert, te vinden, maken we gebruik van het principe van de "Random Grid Search" (of RGS). RGS laat toe om, op een efficiënte manier, een optimale rechthoekige cut te vinden, door events van het signaal dat men zoekt te gebruiken als mogelijke cuts. In de sNavigator methode gebruiken we dit principe om eerst voor elk punt uit de gekozen parameterruimte de beste cut te vinden. Voor elk van deze cuts berekenen we dan de ontdekkingskans voor het volledige parametergebied. De cut die correspondeert met de grootste waarde voor de ontdekkingskans is onze optimale cut.

Vooraleer we de sNavigator toepasten, hebben we eerst het gedrag van RGS bestudeerd in de optimalisatie voor een bepaald punt (in tegenstelling tot een gebied) in de mSUGRA[1] para-

[1] mSUGRA of minimal supergravity is een model voor SUSY met vier vrije parameters en een teken.

meterruimte. We leerden dat we voorzichtig moeten zijn met het kiezen van cut variabelen. Deze variabelen moeten een onderscheid kunnen maken tussen achtergrondprocessen en het gezochte signaal. We moeten ook uitkijken voor correlaties. Aangezien we cuts plaatsen in een multidimensionsale ruimte, kunnen correlaties voor een tegen-intuïtief resultaat zorgen. Zo zagen we dat het toevoegen van meer variabelen niet noodzakelijk betekent dat de resultaten beter worden.

Om sNavigator te kunnen toepassen, moeten we eerst punten in het gekozen gebied van de parameterruimte samplen. Om dit op een efficiënte manier te doen, gebruiken we een Markov Chain Monte Carlo methode waarbij we samplen volgens de waarschijnlijkheidsdistributie voor de zogeheten low energy observables. Enkele voorbeelden zijn het anomaal magnetisch moment van het muon, het verval van bepaalde B-hadronen en de massa van de top-quark. Deze low energy observables bevatten informatie over de aard van supersymmetrie en kunnen op die manier een groter of kleiner gewicht geven aan bepaalde delen van de parameterruimte.

Na het samplen van punten in de parameterruimte, worden voor elk van deze punten events gesimuleerd via Delphes, een programma dat een snelle detector simulatie kan uitvoeren. We hebben Delphes zo geconfigureerd dat de CMS detector gesimuleerd wordt. Daarna bekeken we de verschillen tussen de Delphes simulatie en de volledige CMS detector simulatie (CMSSW FullSim). Over het algemeen kon er een goede overeenkomst vastgesteld worden. We hebben deze bovendien nog kunnen verbeteren door voor b-tagging een functionele vorm voor de b-tag en b-mistag efficiëntie in te voeren. Deze functionele vorm werd verkregen door gebruik te maken van Bayesiaanse neurale netwerken. De overeenkomst op vlak van lepton isolatie zou op deze manier ook verbeterd kunnen worden.

We kunnen dus besluiten dat de componenten van onze ontwikkelde methode naar behoren functioneren. De sNavigator methode zelf werd natuurlijk ook uitvoerig getest. Tijdens deze test stelden we vast dat statistische fluctuaties de procedure sterk kunnen beïnvloeden. Dit kan gebeuren bij relatief lage luminositeit, wanneer er weinig punten uit de parameterruimte ontdekt kunnen worden, en wanneer een van deze punten een hoge likelihood heeft. In ons geval was er een punt met hoge likelihood en hoge werkzame doorsnede waarvoor er slechts één gesimuleerd event door onze preselectie raakte. Dit event bleek een uitschieter te zijn. Wanneer we meer events genereerden voor dit punt, bleek geen enkel event de preselectie te overleven. De vaststelling dat dit specifieke punt ontdekt kon worden, was dus veeleer een statistische fluctuatie. Door de hoge likelihood van dit punt, had deze fluctuatie echter een grote invloed op de optimalisatie.

Hoewel de methode dus doet wat het moet doen – de voorkeur geven aan punten met hoge likelihood –, moeten we op onze hoede zijn voor zulke statistische fluctuaties. Meer events genereren zou moeten helpen, maar zal misschien niet voldoende zijn. Er is dus nog een grondigere test van de stabiliteit nodig, bijvoorbeeld via bootstrapping. We kunnen ook nog opmerken dat voor hogere luminositeiten, zulke punten hun sterke invloed verliezen aangezien er dan meer punten zijn die ontdekt kunnen worden. Hun gecombineerd gewicht kan dan namelijk groter worden dan het gewicht van dat ene punt. Met de bekomen optimale selectie komen dan ook een aanzienlijk groter aantal gesamplede punten binnen het bereik van ontdekking. We kunnen dus concluderen dat de sNavigator globaal gezien goed functioneert.

Als afsluiter van dit werk hebben we de sNavigator procedure toegepast op het mSUGRA model voor drie topologieën: b-jet(s) + één lepton, b-jet(s) + twee leptonen met tegengestelde lading en b-jet(s) + twee leptonen met gelijke lading. Voor elke topologie berekenden we de optimale cut voor drie luminositeiten (300 pb^{-1}, 1 fb^{-1} en 5 fb^{-1}) en drie sets van cut variabelen. We ondervonden dat het geen goed idee is om een extra cut te plaatsen op de p_T van de leptonen. Dit illustreert dat leptonen in mSUGRA over het algemeen een zacht p_T-spectrum hebben. Na toepassen van de optimale cuts, is voor elk van die topologieën $t\bar{t}$ de belangrijkste resterende achtergrond. Voor de topologieën waarin twee leptonen gevraagd worden, kan er voor een luminositeit van 300 pb^{-1} geen goede cut gevonden worden. Dit toont aan dat deze kanalen erg uitdagend zijn en hoge

luminositeiten vragen. In veel gevallen zullen zij niet de "ontdekkingskanalen" zijn, maar ingeval er in een ander kanaal een signaal gevonden wordt, zullen deze kanalen waardevolle informatie verschaffen over de precieze aard van dit signaal.

Aangezien de sNavigator kan toegepast worden op elk model van nieuwe fysica en voor elke topologie, is het een natuurlijke stap om de methode toe te passen voor een ander SUSY model, bijvoorbeeld pMSSM. Dit is een veel algemener model, zonder de sterke aannames van mSUGRA. Het zal erg leerzaam zijn om de resulterende optimale cuts te vergelijken. Dit zal ons inzicht geven in de verschillen tussen beide modellen. Men kan natuurlijk ook de methode toepassen voor een andere topologie. Indien er al bestaande analyses zijn met die topologie, zal het interessant zijn om de bestaande selectie te vergelijken met de optimale selectie zoals verkregen via de sNavigator.

De sNavigator methode laat dus toe om een analyse te ontwerpen. De volgende stap is dan logischerwijze om deze analyse ook effectief uit te voeren. Hiervoor moeten we de verkregen optimale selectie toepassen en het aantal verwachte events voor de Standaard Model achtergrondprocessen bepalen. Op dit punt moeten ook de systematische onzekerheden in rekening gebracht worden. Eenmaal dit gebeurd is, kunnen we ook vergelijken met de data. In het geval er geen teveel aan events geobserveerd wordt, kunnen we de parameterruimte van het gekozen model sterk inperken. Dit resultaat kan dan geïncorporeerd worden in een volgende iteratie van de volledige procedure, waarin we dan een geüpdatete versie van de ontdekkingskans kunnen gebruiken.

Bibliography

[1] A. Djouadi, "The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model," *Phys. Rept.* **457** (2008) 1–216, arXiv:hep-ph/0503172.

[2] M. E. Peskin and D. V. Schroeder, "An Introduction to quantum field theory,". Reading, USA: Addison-Wesley (1995) 842 p.

[3] **LEP Working Group for Higgs boson searches** Collaboration, R. Barate *et al.*, "Search for the standard model Higgs boson at LEP," *Phys. Lett.* **B565** (2003) 61–75, arXiv:hep-ex/0306033.

[4] **CDF and D0** Collaboration, "Combined CDF and D0 Upper Limits on Standard Model Higgs- Boson Production with up to 6.7 fb^{-1} of Data," arXiv:1007.4587 [hep-ex].

[5] A. D. Sakharov, "Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe," *JETP Lett.* **5** no. 1, (1967) 24–27.

[6] S. P. Martin, "A Supersymmetry Primer," arXiv:hep-ph/9709356.

[7] N. Polonsky, "Supersymmetry: Structure and phenomena. Extensions of the standard model," *Lect. Notes Phys.* **M68** (2001) 1–169, arXiv:hep-ph/0108236.

[8] A. Djouadi, "The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model," *Phys. Rept.* **459** (2008) 1–241, arXiv:hep-ph/0503173.

[9] J. D'Hondt and J.-M. Frere, "Extensions of the Standard Model." Course notes, Vrije Universiteit Brussel, 2010.

[10] S. S. AbdusSalam, B. C. Allanach, F. Quevedo, F. Feroz, and M. Hobson, "Fitting the Phenomenological MSSM," *Phys. Rev.* **D81** (2010) 095012, arXiv:0904.2548 [hep-ph].

[11] J. R. Ellis, S. Heinemeyer, K. A. Olive, A. M. Weber, and G. Weiglein, "The Supersymmetric Parameter Space in Light of B-physics Observables and Electroweak Precision Data," *JHEP* **08** (2007) 083, arXiv:0706.0652 [hep-ph].

[12] **Muon G-2** Collaboration, G. W. Bennett *et al.*, "Final report of the muon E821 anomalous magnetic moment measurement at BNL," *Phys. Rev.* **D73** (2006) 072003, arXiv:hep-ex/0602035.

[13] M. Davier, A. Hoecker, B. Malaescu, and Z. Zhang, "Reevaluation of the Hadronic Contributions to the Muon g-2 and to alpha(MZ)," *Eur. Phys. J.* **C71** (2011) 1515, arXiv:1010.4180 [hep-ph].

[14] D. Stockinger, "The muon magnetic moment and supersymmetry," *J. Phys.* **G34** (2007) R45–R92, arXiv:hep-ph/0609168.

[15] 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].

[16] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov, "micrOMEGAs: Version 1.3," *Comput. Phys. Commun.* **174** (2006) 577–604, arXiv:hep-ph/0405253.

[17] P. Bechtle, O. Brein, S. Heinemeyer, G. Weiglein, and K. E. Williams, "HiggsBounds: Confronting Arbitrary Higgs Sectors with Exclusion Bounds from LEP and the Tevatron," *Comput. Phys. Commun.* **181** (2010) 138–167, arXiv:0811.4169 [hep-ph].

[18] **Particle Data Group** Collaboration, K. Nakamura *et al.*, "Review of particle physics," *J. Phys.* **G37** (2010) 075021.

[19] Heavy Flavor Averaging Group http://www.slac.stanford.edu/xorg/hfag/.

[20] L. Evans, (ed.) and P. Bryant, (ed.), "LHC Machine," *JINST* **3** (2008) S08001.

[21] AC Team, "The four main LHC experiments." June 1999, http://cdsweb.cern.ch/record/40525.

[22] **CMS** Collaboration, R. Adolphi *et al.*, "The CMS experiment at the CERN LHC," *JINST* **3** (2008) S08004.

[23] CERN, "CMS design." http://bigscience.web.cern.ch/bigscience/en/cms/cms2.html.

[24] http://root.cern.ch/drupal/.

[25] **GEANT4** Collaboration, S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit," *Nucl.Instrum.Meth.* **A506** (2003) 250–303.

[26] http://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookPAT.

[27] https://twiki.cern.ch/twiki/bin/view/CMS/TheNtupleMaker.

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

[29] "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET,". CMS-PAS-PFT-09-001, http://cdsweb.cern.ch/record/1194487.

[30] http://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookBTagging.

[31] D. Alves *et al.*, "Simplified Models for LHC New Physics Searches," arXiv:1105.2838 [hep-ph].

[32] H. B. Prosper, "Probability and Statistical Inference," arXiv:physics/0606179.

[33] G. D'Agostini, "Probability and measurement uncertainty in physics: A Bayesian primer," arXiv:hep-ph/9512295 [hep-ph]. Notes from lectures in Rome, 5/95 and DESY, 9/95.

[34] R. R. de Austri, R. Trotta, and L. Roszkowski, "A Markov chain Monte Carlo analysis of the CMSSM," *JHEP* **05** (2006) 002, arXiv:hep-ph/0602028.

[35] H. Baer, S. Kraml, A. Lessa, S. Sekmen, and X. Tata, "Effective Supersymmetry at the LHC," *JHEP* **10** (2010) 018, arXiv:1007.3897 [hep-ph].

[36] G. Belanger, F. Boudjema, A. Pukhov, and R. K. Singh, "Constraining the MSSM with universal gaugino masses and implication for searches at the LHC," *JHEP* **11** (2009) 026, arXiv:0906.5048 [hep-ph].

[37] Private communications with Harrison B. Prosper.

[38] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *European Physical Journal C* **71** (Feb., 2011) 1554–+, arXiv:1007.1727 [physics.data-an].

[39] N. A. Amos *et al.*, "The random grid search: A simple way to find optimal cuts,". Prepared for International Conference on Computing in High- energy Physics (CHEP 95), Rio de Janeiro, Brazil, 18-22 Sep 1995.

[40] J. Alwall *et al.*, "MadGraph/MadEvent v4: The New Web Generation," *JHEP* **09** (2007) 028, arXiv:0706.2334 [hep-ph].

[41] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175.

[42] P. Z. Skands *et al.*, "SUSY Les Houches Accord: Interfacing SUSY Spectrum Calculators, Decay Packages, and Event Generators," *JHEP* **07** (2004) 036, arXiv:hep-ph/0311123.

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

[44] A. Djouadi, M. 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.

[45] N. Saoulidou, "Particle Flow Jet Identification Criteria." CMS Analysis Note 2010/003 (2010).

[46] G. Kaufman *et al.*, "Search for New Physics in Events with b-Jets and Missing Transverse Energy." CMS Analysis Note 2011/022 (2011).

[47] G. Punzi, "Sensitivity of searches for new signals and its optimization," arXiv:physics/0308063.

[48] S. Ovyn, X. Rouby, and V. Lemaitre, "Delphes, a framework for fast simulation of a generic collider experiment," arXiv:0903.2225 [hep-ph].

[49] http://www.fynu.ucl.ac.be/users/s.ovyn/Delphes/index.html.

[50] M. Dobbs and J. B. Hansen, "The HepMC C++ Monte Carlo event record for High Energy Physics," *Comput.Phys.Commun.* **134** (2001) 41–46.

[51] P. Bhat and H. Prosper, "Bayesian neural networks,". Prepared for PHYSTAT05: Statistical Problems in Particle Physics, Astrophysics and Cosmology, Oxford, England, United Kingdom, 12-15 Sep 2005.

[52] W. Adam *et al.*, "Search for supersymmetry in proton-proton collisions at $\sqrt{s} = 7$ TeV in events with a single lepton, jets, and missing transverse momentum." CMS Analysis Note 2010/419 (2010).

[53] **CMS** Collaboration, S. Chatrchyan *et al.*, "Search for Physics Beyond the Standard Model in Opposite- Sign Dilepton Events at sqrt(s) = 7 TeV," arXiv:1103.1348 [hep-ex].

[54] **CMS** Collaboration, S. Chatrchyan *et al.*, "Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy at the LHC," arXiv:1104.3168 [hep-ex].