Search for Supersymmetry in single-lepton final states with the ATLAS detector

and

Improved background model for the search of new physics

Dissertation zur Erlangung der Doktorwürde

Vorgelegt von

Stephan Horner

Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität Freiburg

Dekan: Prof. Dr. Kay Königsmann
Leiter der Arbeit: Prof. Dr. Gregor Herten
Referent: Prof. Dr. Gregor Herten
Koreferent: Prof. Dr. Horst Fischer
Tag der Verkündigung des Prüfungsergebnisses: 06. Juli 2011

Meinen Eltern, Gudrun, Johannes und Nina gewidmet.

Acknowledgments

This work would not have been possible without the help of several people. I want to thank in particular:

- Prof. Gregor Herten, my doctoral advisor, for continuous support and offering a PhD position, partly funded by the BMBF, in the first place.
- The Graduiertenkolleg Physik an Hadron-Beschleunigern, funded by the DFG and represented by Prof. Karl Jakobs, for the awarded fellowship and associated grant.
- The ATLAS Collaboration, in particular the SUSY Working Group.
- Dr. Sascha Caron and Dr. Jan Erik Sundermann, for guidance and close collaboration.
- My coworkers, especially Florian Ahles, Riccardo-Maria Bianchi, Dr. Renaud Brunelière, Tobias Rave and Dr. Zuzana Rúriková.
- Michael Rammensee, for valuable discussion about statistics issues.
- Valerio Consorti, for providing additional calculations.
- Dr. Till Eifert, for close collaboration during the data analysis.
- Dr. Glen Cowan and Dr. Eilam Gross, for close collaboration resulting in Part III of the thesis.
- Dr. Adil Aktas and Dr. Anton Gamel, for keeping the computer infrastructure up and running in spite of adverse conditions, not only due to the weather.
- All the other people at the Institute of Physics with whom I had a very good time.

Preface

The current understanding of nature at its elementary level is condensed in a physical theory known as the Standard Model. While it has been extraordinarily successful in describing almost all known phenomena in particle physics with astonishing precision since the 1970s, it is also commonly agreed upon that it can only be an approximation of more fundamental laws of physics, which have yet to be discovered. Supersymmetry is an attractive candidate for an extension of the Standard Model since it recovers many of its deficiencies and, maybe, constitutes an important milestone on the road to a theory of everything.

There are good reasons to believe that if supersymmetry exists, it will make itself apparent at the so-called weak energy scale. With the start of the Large Hadron Collider, a new era in particle physics has therefore begun, since it will be possible to thoroughly probe this scale for the first time. The experiments at the Large Hadron collider will also search for the last big missing piece of the Standard Model, the elusive Higgs boson, and potentially unveil the cause for the associated electroweak symmetry breaking, which gives rise to the masses of the particles and may well be connected with supersymmetry. This thesis presents the author's contribution to the search for supersymmetry in final states which contain a single high-energetic electron or muon, using data recorded by the ATLAS experiment until the end of 2010.

A crucial prerequisite for the search for supersymmetry or any other new physics phenomenon is the precise simulation of known Standard Model processes. Due to shortcomings in the modeling or an incomplete understanding of detector effects, the measurements will always systematically deviate from the predictions. A novel method, which incorporates such deviations into an improved model, is presented. It combines information from the simulation and from the data, obtained by control measurements, in a statistically well-defined manner.

The thesis is organized as follows. Part I gives a brief description of the physics at the ATLAS experiment, focusing on the supersymmetric extension of the Standard Model of particle physics and the detector design itself. It serves as an introduction to Part II, which presents the search for supersymmetric particles in single-lepton final states with ATLAS. Part II starts by discussing the prospects of a potential discovery of supersymmetry in proton-proton collisions at the design centre-of-mass energy of 14 TeV, followed by a presentation of the one-lepton search analysis using both early data and all data collected by ATLAS in 2010 at the reduced collision energy of 7 TeV. The results of the analysis have been published in [1, 2, 3, 4]. Part III introduces the proposed method to improve the background model in a search for new physics and has been published in [5, 6]. Part IV closes with a conclusion summarizing the main results of the thesis, followed by an appendix with additional material and the bibliography.

Contents

Part I

Physics at the ATLAS experiment

Chapter 1

The Standard Model and its supersymmetric extension

This chapter gives a brief overview on the Standard Model of particle physics and its supersymmetric extension. It is not a complete description, but introduces the main aspects of the theory necessary to understand the results of the research presented in Part II of the thesis.

Recommendable introductions to the Standard Model and supersymmetry can be found in [7, 8] and [9, 10, 11] respectively, which also constitute the basis of the explanations to follow.

1.1 The Standard Model of particle physics

The Standard Model (SM) describes the elementary particles and their interactions through three out of four known fundamental forces, that is the electromagnetic, the weak and the strong interaction. The current model does not include gravitational interaction, whose effects are negligible at collider energies achievable to date.

1.1.1 Particles and interactions

Particles The observed fundamental particles can be divided into two main categories: so-called *matter particles*, which are fermions with an intrinsic spin of 1/2, and *force particles* with bosonic integer spin. They are assumed to be elementary, i.e. devoid of any substructure, and therefore point-like.

The fermions, listed in table 1.1, are arranged in three groups called *generations*, which only differ by the masses of their particles. In addition, they are divided into quarks, which interact via all four forces, and leptons, which do not interact strongly. Each generation includes two quark and two lepton *flavors*. The first generation comprises the up (u) and down quark (d), the electron (e) and the electron neutrino (ν_e). The particles of the second generation are labeled charm (c) and strange quark (s), muon (μ) and muon neutrino (ν_μ),

Table 1.1: The three generations of quarks and leptons in the Standard Model, their masses and electric charges [12]. The particles are denoted by their symbols as described in the text. The uncertainties on the electron and muon masses are negligible. For details on the neutrino mass limits see [13].

		Quarks			**Leptons**	
Generation	Particle	Mass [MeV]	Charge [e]	Particle	Mass [MeV]	Charge [e]
First	u	1.7 to 3.3	+2/3	e	0.511	-1
	d	4.1 to 5.8	-1/3	ν_e	$< 2 \times 10^{-6}$	0
Second	c	$1.27^{+0.07}_{-0.09} \times 10^3$	+2/3	μ	105.7	-1
	s	101^{+29}_{-21}	-1/3	ν_μ	< 0.19	0
Third	t	$172.0 \pm 0.9 \pm 1.3 \times 10^3$	+2/3	τ	1776.82 ± 0.16	-1
	b	$4.19^{+0.18}_{-0.06} \times 10^3$	-1/3	ν_τ	< 18.2	0

Table 1.2: The force carriers of the three fundamental forces described by the Standard Model, their masses and electric charges [12]. The particles are denoted by their symbols as described in the text. The photon and gluons are assumed massless in the model with the upper limit on the latter taken from [14].

Particle	Interaction	Mass [MeV]	Charge [e]
γ	Electromagnetic	$< 1 \times 10^{-24}$	$< 1 \times 10^{-35}$
Z	Weak	$91.188 \pm 0.002 \times 10^3$	0
$W^\pm$	Weak	$80.399 \pm 0.023 \times 10^3$	± 1
g	Strong	$< \mathcal{O}(1)$	0

while the ones making up the third generation are called top (t) and bottom quark (b), tau (τ) and tau neutrino (ν_τ). The Standard Model does not provide an explanation for the considerable range of the mass spectra, which span more than three orders of magnitude for the quarks and leptons each. Ordinary matter is made up of electrons, up and down quarks since bound states including quarks and charged leptons of the higher generations are not stable. Finally, for each fermion there exists an anti-particle with equal mass but opposite charge.

The interaction between the particles is mediated through the exchange of bosons, which are listed in table 1.2. The photon (γ) and the gluon (g), which have unity spin and are massless, constitute the carriers of the electromagnetic and the strong force respectively. The three spin 1 particles associated with the weak interaction, $W^\pm$ and Z, are very massive with about 80 and 91 GeV respectively, which implies a finite range of the force.

Mathematically, the fermionic and bosonic particles are described as spinor and vector fields respectively, with their dynamics governed by relativistic quantum field theory (QFT). The interaction of the matter particles with the fundamental force fields arises as a direct consequence of the axiomatic requirement that the Lagrangians of the free fields are invariant under local phase transformations of the fields (*locally gauge invariant*). The strengths of the interactions are expressed by *coupling constants*. At ever higher interaction energies, which correspond to smaller space distances, additional *virtual corrections* to the basic interaction have to be considered. By an approach called *renormalization* these corrections can be absorbed into then energy-scale dependent couplings and masses. The energy dependence of these so-called *running couplings* and *running masses* is determined by *Renormalization Group Equations* (RGEs).

Strong force The strong interaction is described by a QFT model called Quantum Chromodynamics (QCD), which derives its name from the fact that the force acts on all particles with color charge. Quarks (anti-quarks) carry one of three possible colors (anti-colors), while gluons always carry color–anti-color combinations. The interaction, mediated by eight gluon fields, is the result of the QCD Lagrangian being invariant to phase transformations by members of the SU(3) symmetry group.

The force between colored particles increases with distance, creating new quark–anti-quark pairs through vacuum oscillations beyond a certain energy threshold. This behavior, termed *color confinement*, renders an observation of isolated colored particles impossible and results in a clustering of quarks in color-neutral combinations of three particles (combining red, green and blue color) called *baryons* or quark–antiquark pairs (color canceling anti-color) called *mesons*. The origin of confinement is still under debate since no analytic proof that it is implied by QCD has been proposed to date (see e.g. [15]).

Since the force decreases at smaller distances, quarks and gluons behave as practically free particles in high-energy interactions (so-called *asymptotic freedom*). The latter fact is crucial for the precise modeling of high-energetic collisions at hadron colliders since it enables calculations using pertubative approaches.

Unified electroweak force and Higgs mechanism Quantum Electrodynamics (QED) describes the interaction between electrically charged particles based on photon exchange. It has been one of the most successful physical theories to date with predictions up to an accuracy of one part in a trillion (cf. *g-factor*, see e.g. [16, 17]). The QED Lagrangian is invariant under a local U(1) transformation which implies non-self-interacting photons and a force decreasing with distance, consistent with predictions by classical electrodynamics in the macroscopic limit.

QED can be combined with the weak interaction to obtain a unified electroweak force, whose Lagrangian is symmetric with respect to the combined SU(2) × U(1) gauge group. However, the four resulting force fields need to be massless to preserve local gauge invariance, which is clearly incompatible with the observed massive $W^{\pm}$ and Z bosons. As a resolution, an additional field, the so-called *Higgs field*, is introduced, which transforms under the same symmetry group and has four degrees of freedom. Its a priori symmetric potential has a degenerate minimum at finite field values. In the ground state of the Lagrangian, the symmetry is broken and the Higgs field mixes with the force fields, thereby making the $W^{\pm}$ and Z bosons massive. The remaining degree of freedom represents a scalar particle, the *Higgs boson*, which is yet to be observed. The search for it is one of the top priorities at the Large Hadron Collider. Since the symmetry of the Lagrangian is broken by selecting one of equally likely degenerate ground states without an apparent external cause, the Higgs mechanism introduces so-called *spontaneous symmetry breaking* in the Standard Model.

Also the fermions acquire mass through the Higgs mechanism, which is realized via so-called *Yukawa couplings*. The values of the couplings, which determine the masses of the fermions, are like the Higgs boson mass not provided by the model and need to be measured.

1.1.2 Limitations of the Standard Model

Although extraordinarily successful in explaining almost all known phenomena in particle physics without major adjustments since the 1970s, the Standard Model has various limitations. Some of the most prominent are briefly discussed in the following:

- **Gravity:** As mentioned at the beginning of the chapter, it has not been possible to include a quantum field theory of gravity consistent with General Relativity in the Standard Model.

- **Higgs boson:** One of the crucial ingredients, the Higgs boson, has not been discovered yet. If it will not be measured at the Large Hadron Collider, other, generally more complicated mechanisms for electroweak symmetry breaking, e.g. due to a new strong interaction [18], need to take its place.

- **Origin of masses and generations:** The Standard Model includes 19 parameters whose values cannot be provided by the theory, including the fermion masses and the gauge couplings. In addition, it does not explain the reason for there being three generations with two different flavors.
- **Grand unification:** Motivated by the electroweak unification, it is widely believed that the Standard Model is only a low-energy approximation, also called *effective theory*, of a more fundamental theory, which is supported by the fact that the running couplings of the forces appear to almost meet at very high energies, as shown by the dashed lines in figure 1.1. Such a *grand unified theory* (GUT) would combine the strong and the electroweak forces into a single interaction (first proposed by Georgi and Glashow in 1974 [19]) at a scale around 10^{16} GeV. By predicting a charge quantization not necessitated in the Standard Model, it would also explain why the charges of electrons and protons seem to cancel each other exactly. At the *Planck scale* of 10^{19} GeV the strength of gravity becomes comparable to that of the other forces and would also need to be included in a *theory of everything*, with *string theory* (see e.g. [20]) being the most prominent candidate to date.
- **Hierarchy problem:** It is unknown why the strengths of gravity and the other forces differ by so many orders of magnitude (roughly 10^{32}). When assuming a fundamental theory of new physics around the Planck scale, these different scales result in huge quantum corrections to the bare masses of the particles, in particular that of the Higgs boson. In order to obtain masses at the weak scale of around 1 TeV, the bare masses need to precisely cancel the corrections over many orders of magnitude. This *fine-tuning* of the bare masses is generally considered unnatural and can be avoided by opposite-signed corrections due to yet unobserved particles, predicted e.g. by weak scale supersymmetry (see next section).
- **Matter–antimatter asymmetry:** Assuming that matter and antimatter have been created in equal amounts after the big bang, there is no mechanism within the Standard Model which can account for the apparent overshoot of matter today. The magnitude of *CP violation* in the Standard Model is not sufficient to provide an explanation.
- **Dark matter:** The Standard Model does not contain a particle which can serve as a candidate for dark matter, which is estimated to make up 80% of all matter in the universe (see e.g. [21, 22]).

1.2 The supersymmetric extension of the Standard Model

The shortcomings of the Standard Model point to a more fundamental and, in terms of group theory, potentially more aesthetic theory at higher energy scales. A promising candidate for an extension of the model, which remedies many of the above-mentioned problems,

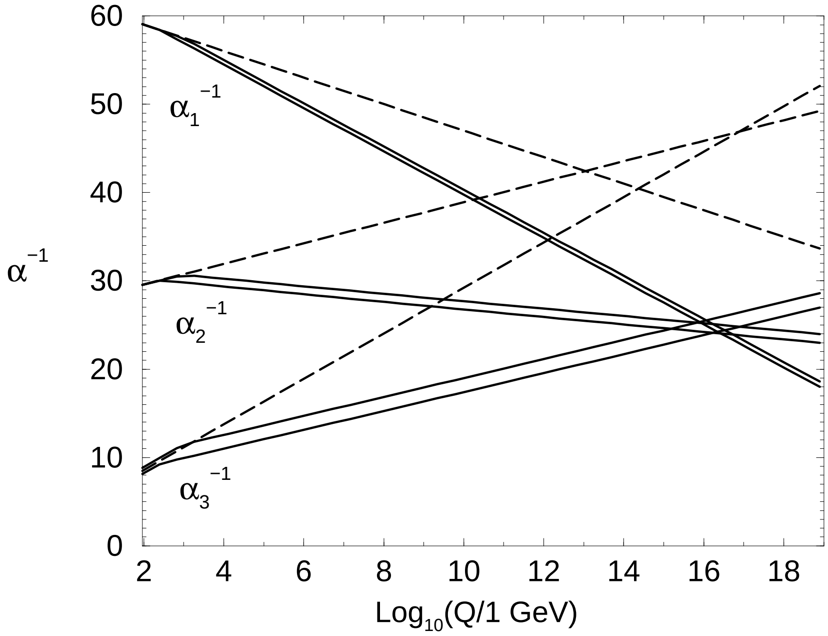


Figure 1.1: Evolution of the inverse gauge couplings $\alpha_{\rm i}$ in the Standard Model (dashed lines) and the MSSM (solid lines) as a function of the energy scale Q [9]. The extended particle content of the MSSM enables an apparant unification of the gauge couplings at a scale of about 2 $\times 10^{16}$ GeV. The indices 1 to 3 correspond to the electromagnetic, the weak and the strong coupling respectively.

is *supersymmetry* (SUSY). It was originally proposed by Miyazawa [23, 24] in 1966 as a symmetry connecting baryons and mesons and was then reformulated for fundamental particles by several groups in the early 1970s (e.g. most prominently by Wess and Zumino [25]). As an extension of the Standard Model, supersymmetry relates bosons with fermions. The symmetry is expressed as a transformation which turns a bosonic state into a fermionic one by means of a spinor operator Q, thereby changing the spin of the particle by half a unit and leaving the other quantum numbers unaltered:

$$Q|\text{Boson}\rangle = |\text{Fermion}\rangle, \qquad Q|\text{Fermion}\rangle = |\text{Boson}\rangle \tag{1.1}$$

Each bosonic degree of freedom is thus related to a fermionic one and vice versa. The corresponding particles are called *superpartners* of each other. The particle content of the Standard Model is not sufficient for such a relation and thus needs to be extended, which implies an obvious problem: None of the superpartners, which have the same masses as their associated Standard Model particles as required by the symmetry, has been observed to date. Therefore, supersymmetry must be broken in the ground state of the theory.

1.2.1 The Minimal Supersymmetric Standard Model

The minimal extension of the Standard Model corresponds to a doubling of the particle content, which constitutes the *Minimal Supersymmetric Standard Model* (MSSM).

Fermions and sfermions The fermions have bosonic superpartners with spin 0, which are called scalar fermions or shortly *sfermions*. Since the numbers of degrees of freedom differ by a factor of two, each fermion is associated with two sfermion superpartners. They are labeled corresponding to the handedness of the fermionic superpartner, i.e. *left-* and *right-handed squarks* and *sleptons*, and usually denoted by a tilde, e.g. $\tilde{e}_R$ represents the selectron whose superpartner is the right-handed electron. As is true for SM fermions, the weak interaction only couples to left-handed sfermions.

Gauge bosons and gauginos The spin 1/2 superpartners of the spin 1 gauge bosons are refered to as *gauginos*. Besides the *gluinos* related to the gluons, there are the superpartners of the electroweak gauge eigenstates (W^+, W^-, W^0 and B^0), which are thus called *winos* and *bino*: $\tilde{W}^+$, $\tilde{W}^-$, $\tilde{W}^0$ and $\tilde{B}^0$. In order to give masses to all SM fermions the SM Higgs sector needs to be extended in the MSSM to contain eight real degrees of freedom, which results in five Higgs bosons and four fermionic superpartner *Higgsinos*, two of which are charged. The neutral Higgsinos mix with the neutral wino and bino to produce four so-called *neutralinos*, often denoted by $\tilde{\chi}^0$, and the charged Higgsinos mix with the charged winos resulting in two pairs of *charginos*: $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_2^{\pm}$.

***R*-parity and dark matter candidate** The Lagrangian of the MSSM allows for the inclusion of baryon- and lepton-number violating terms, which is problematic since the associated processes have not been observed experimentally. One can, however, add a new symmetry, called *R-parity*, which forbids such terms and whose quantum number P_R is conserved:

$$P_R = (-1)^{3(B-L)+2s} \tag{1.2}$$

where B and L are the baryon and lepton number respectively and s is the spin of the considered particle. Thus, Standard Model particles always have an R-parity value of $+1$ and their supersymmetric partners -1. As an immediate consequence, the lightest supersymmetric particle, abbreviated by *LSP*, cannot decay further and is therefore stable. Every other supersymmetric particle must eventually decay into a state with an odd number of LSPs, typically only one. Furthermore, SUSY particles can only be produced in pairs in an interaction. The stability and the substantial mass of the LSP renders it a natural candidate for dark matter [21, 22], which is widely believed to consist of non-baryonic and *cold*, i.e. non-relativistic, massive particles necessary for the observed clustered structure in the universe. If the LSP has a mass around the TeV-scale, it needs to have zero charge to be consistent with the astronomical non-observation. In that case, it fulfills a further requirement for a dark matter candidate.

Supersymmetry benefits As mentioned above, the supersymmetric extension recovers many of the Standard Model's deficiencies, although its historic development was not motivated by this purpose. Besides providing a good candidate for dark matter, it solves the Higgs hierarchy problem. Like the Standard Model particles, their superpartners contribute quantum corrections to the Higgs mass, but with opposite sign. Provided their masses are not too different from the SM partners (order of 1 TeV), the corrections are of roughly the same size and cancel to all orders in pertubation theory, thereby stabilizing the Higgs mass at the weak scale.

Another argument for SUSY masses around the weak scale is their impact on the running of the couplings as displayed in figure 1.1. While the couplings almost meet if the particle content comprises only the known SM particles (dashed lines), they actually can meet if their SUSY partners are also included in the calculation (solid lines), which strongly hints at a unified force at this scale consistent with grand unified theories. The deviation of the couplings from the ordinary SM ones starts at the mass scales of the SUSY particles since processes with on-shell SUSY particles then become possible. In the context of the MSSM, unification of the couplings can be achieved if the SUSY masses are around the TeV-scale as depicted in the figure.

As a bonus, soft supersymmetry breaking, which is discussed below, provides a mechanism to break the electroweak symmetry, which is in the Standard Model postulated to be of spontaneous origin. Finally, current string theories which incorporate both fermions and bosons are necessarily supersymmetric.

1.2.2 Supersymmetry breaking

As stated above, supersymmetry must be broken in the ground state for the masses of the SUSY particles to be different from their SM partner ones. The breaking mechanism must work in a way such that the phenomenology and relationships between the couplings are not affected, which would compromise e.g. the cancelation of the radiative corrections to the Higgs mass. This is called *soft supersymmetry breaking* and is achieved by adding extra SUSY breaking terms to the SUSY Lagrangian.

The inclusion of the SUSY breaking terms introduces 105 additional parameters related to the masses, phases and mixing angles in addition to the 19 parameters of the Standard Model. However, many of these imply flavor mixing or CP violating processes, which are strongly constrained by experiment. Imposing so-called *soft supersymmetry-breaking universality relations* significantly reduces the arbitraryness of the resulting Lagrangian. These relations are presumed to be the consequence of a specific model responsible for supersymmetry breaking, associated with an underlying simplicity or symmetry of the Lagrangian at very high energy scales. It is assumed that supersymmetry breaking occurs in a *hidden sector* of particles, which have (almost) no interaction with the MSSM particles in the *visible sector*. Analogous to the SM spontaneous electroweak symmetry breaking, the Lagrangian density of the underlying model in the hidden sector is invariant under supersymmetry but the vacuum state is not. The SUSY breaking is mediated to the visible sector by particles which interact with both sectors, giving rise to the soft SUSY breaking terms of the Lagrangian. This interaction might be of gravitational original as is the case in the *gravity mediated supersymmetry breaking* scenario. Other possibilities, which are not discussed here further, include the ordinary electroweak and strong interactions as mediators and the hidden sector being located in extra-dimensional spacetime.

Gravity mediated supersymmetry breaking The gravitino, which is the massive superpartner of the yet unobserved graviton, constitutes the mediator in a scenario, where a *super Higgs mechanism* breaks supersymmetry within a supergravity model. Assuming a *minimal* form of the supergravity Lagrangian, one ends up with only 5 additional parameters in the *minimal supergravity* scenario (mSUGRA):

- $\boldsymbol{m_0}$ **:** universal gaugino mass at the GUT scale
- $\boldsymbol{m_{1/2}}$ **:** universal scalar particle mass at the GUT scale
- $\boldsymbol{A_0}$ **:** universal trilinear coupling
- $\boldsymbol{tan\beta}$ **:** ratio of the vacuum expectation values of the two neutral Higgs fields
- $\boldsymbol{sign(\mu)}$ **:** sign of the Higgs(ino) mass parameter

Like couplings, the masses of the MSSM particles at observable scales can be computed using renormalization group equations. Figure 1.2 shows an example of the running of scalar and gaugino mass parameters. The negative value of $(\mu^2 + \mathrm{m}^2_{\mathrm{H_u}})^{1/2}$ at low scales induces electroweak symmetry breaking, which is thus called *radiative electroweak symmetry breaking* since it is driven by quantum corrections decreasing the Higgs mass when evolving to lower scales.

SUSY particle masses While the mass spectra differ substantially depending on the considered SUSY scenario, a few general features within the MSSM are considered likely [9]:

- **Gluinos:** The gluinos are presumed to be heavier than the lighter charginos and neutralinos as is also indicated by the associated mass parameter M_3 in the example shown by figure 1.2.
- **Squarks and sleptons:** In general, the squarks are expected to be substantially heavier than the sleptons. Furthermore, the third generations squarks and sleptons can have considerably different masses compared to the first two generations, which are nearly degenerate. The third generation's left- and right-handed particles mix strongly, whereas the first two generation's gauge and mass eigenstates are practically

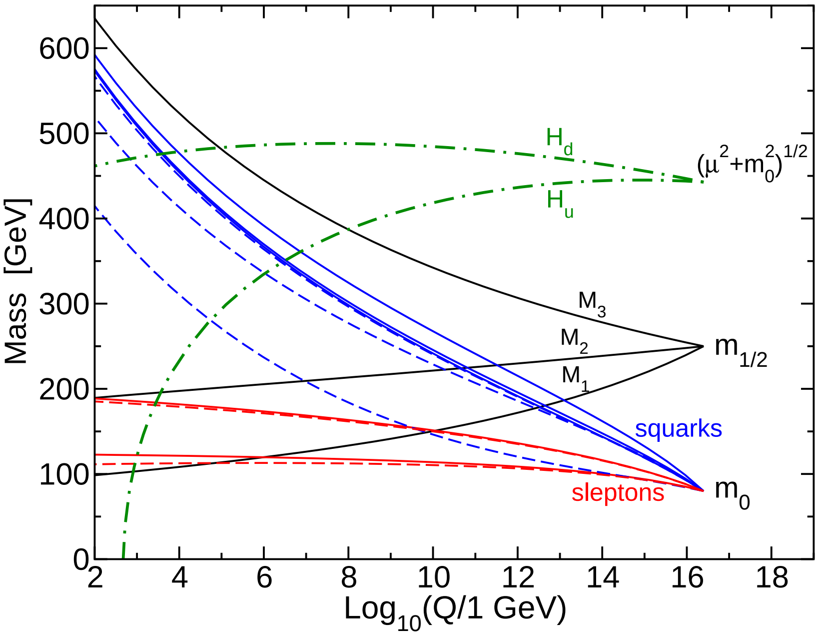


Figure 1.2: Evolution of the scalar ($\mathrm{m_0}$) and gaugino ($\mathrm{m_{1/2}}$) masses assuming universal masses at an energy scale of $\mathrm{Q_0} = 2.5 \times 10^{16}$ GeV for a specific mSUGRA scenario with $m_0 = 80$ GeV, $m_{1/2} = 250$ GeV, $A_0 = -500$ GeV, $\tan\beta = 10$ and positive μ [9]. $\mathrm{M_1}$, $\mathrm{M_2}$ and $\mathrm{M_3}$ are the bino, wino and gluino mass parameters respectively. The $\mathrm{m_0}$ in $(\mu^2 + \mathrm{m_0^2})^{1/2}$ is to be replaced with the Higgs mass parameters $\mathrm{m_{H_u}}$ and $\mathrm{m_{H_d}}$ for the two dashed green lines respectively.

identical. The lightest squarks are probably stops or sbottoms and the lightest charged slepton a stau.

- **LSP:** The lightest supersymmetric particle is presumably a neutralino or the gravitino in order to represent a good dark matter candidate.

- **Higgs particles:** The mass of the lightest Higgs boson h_0 is expected to be smaller than 150 GeV and can be much smaller than the masses of the other four Higgs scalars, which are usually denoted by A^0, $H^\pm$ and H^0.

These features can be observed in the mass spectra of the benchmark signals used in the analyses of Part II of the thesis (cf. table A.1 in the appendix), which have been chosen to cover a wide range of SUSY masses and decay topologies. The potential production of SUSY particles and typical decays are discussed in subsection 1.3.4.

1.2.3 Experimental mass constraints

Due to the huge parameter space of the MSSM it is difficult to place general limits on the SUSY particle masses. Therefore, most limits are determined for parameter value combina-

Table 1.3: Lower limits on MSSM SUSY particle masses at 95% confidence level [12]. For model assumptions see description in the text. Details on the limits can be found in [26].

Particle	Symbol	Mass Limit	Valid Parameter Region
Neutralinos	$\tilde{\chi}_1$	> 46 GeV	all $\tan\beta$, all m_0, all mass differences $m_{\tilde{\chi}_2} - m_{\tilde{\chi}_1}$
	$\tilde{\chi}_2$	> 62.4 GeV	$1 < \tan\beta < 40$, all m_0, all $m_{\tilde{\chi}_2} - m_{\tilde{\chi}_1}$
	$\tilde{\chi}_3$	> 99.9 GeV	$1 < \tan\beta < 40$, all m_0, all $m_{\tilde{\chi}_2} - m_{\tilde{\chi}_1}$
	$\tilde{\chi}_4$	> 116 GeV	$1 < \tan\beta < 40$, all m_0, all $m_{\tilde{\chi}_2} - m_{\tilde{\chi}_1}$
Chargino	$\tilde{\chi}_1^{\pm}$	> 94 GeV	$\tan\beta < 40$, all m_0, $m_{\tilde{\chi}_1^{\pm}} - m_{\tilde{\chi}_1} > 3$ GeV
Selectron	$\tilde{e}$	> 107 GeV	all $m_{\tilde{e}_R} - m_{\tilde{\chi}_1}$
Smuon	$\tilde{\mu}$	> 94 GeV	$1 < \tan\beta < 40$, $m_{\tilde{\mu}_R} - m_{\tilde{\chi}_1} > 10$ GeV
Stau	$\tilde{\tau}$	> 94 GeV	all θ_τ, $m_{\tilde{\tau}_R} - m_{\tilde{\chi}_1} > 15$ GeV
Squark	$\tilde{q}$	> 379 GeV	$A_0 = 0$, all $m_{\tilde{g}}$, $\tan\beta = 3$, $\mu < 0$ (weakly sensitive to $\tan\beta$ and μ over much of parameter space)
Gluino	$\tilde{g}$	> 308 GeV	all $m_{\tilde{q}}$ (sensitivity to $\tan\beta$ and μ as for $\tilde{q}$)
		> 392 GeV	$m_{\tilde{q}} = m_{\tilde{q}}$ (sensitivity to $\tan\beta$ and μ as for $\tilde{q}$)

tions of specific scenarios (mostly mSUGRA) and their validity for other SUSY models is questionable. However, other models often predict comparable phenomenologies, which can be tested using these indicative limits.

Table 1.3 summarizes mass limits for the different SUSY particles, which have been provided by both collider and astrophysical experiments, under the following assumptions [12]:

1. The lighest neutralino is the LSP.
2. R-parity is conserved.
3. With the exception of stop and sbottom, all squarks are degenerate in mass. Left- and right-handed squarks also have the same mass.
4. Gaugino masses are unified at the GUT scale.
5. Limits for sleptons refer to the right-handed particles, which typically have smaller cross-sections and smaller masses provided sfermion masses unify at the GUT scale.

1.3 Modeling of proton-proton collisions

The Large Hadron Collider (LHC), which is introduced in more detail in the subsequent chapter, is a synchrotron built to collide particle beams of protons or lead nuclei. Due to the complex nature of the strong force, the simulation of high-energy proton interactions is a complicated task, which will be explained in the following. The section also introduces the Standard Model processes relevant for the search analysis in Part II and discusses potential supersymmetry production at the LHC.

1.3.1 Theoretical basis

Factorization theorem In proton-proton collision, it is in fact the constituents of the protons which interact with each other, provided the collision energy is high enough to resolve the proton substructure. The constituents, generally called *partons*, are the three *valence quarks*, two of which are up quarks and the third is a down quark, the gluons and the so-called *sea quarks*, which are formed as short-lived quark–antiquark pairs from gluon splits. The QCD factorization theorem (see e.g. [27] for details) states that the hadronic cross-section of two protons interacting to produce the final state X can be calculated as

$$\sigma_{pp\to X} = \sum_{i,j=q,\bar{q},g} \iint dx_i\, dx_j\, f_i(x_i,\mu_F^2) f_j(x_j,\mu_F^2) \cdot \hat{\sigma}_{ij\to X}(x_i x_j s, \mu_R^2, \mu_F^2) \tag{1.3}$$

where the $x_{i/j}$ are the momentum fractions of the partons i/j with respect to the momentum of the proton they are part of, μ_F and μ_R are the factorization and renormalization scale respectively, $f_{i/j}$ are the so-called *parton distribution functions* and $\hat{\sigma}_{ij\to X}$ is the cross-section of the scattering between the partons. $\sqrt{s}$ is the center-of-mass energy of the proton collisions and $\sqrt{x_i x_j s}$ thus the center-of-mass energy in the parton-parton interaction. If the parton scattering is hard enough such that the strong coupling becomes smaller than unity, $\hat{\sigma}_{ij\to X}$ can be calculated using pertubative approaches. The factorization scale μ_F defines the separation between soft and hard QCD processes, while the renormalization scale μ_R is used in regulating divergent terms in the pertubation expansion. The dependence of the cross-section on those scales decreases as more terms are included in the pertubation series. When calculated to all orders in pertubation theory, the scale dependence of $\sigma_{pp\to X}$ disappears entirely. Therefore, the choice of μ_R and μ_F is arbitrary, but different choices produce different results due to different unknown higher-order corrections (see e.g. [28]). In order to avoid large corrections, one usually chooses scale values close to the typical momentum scale of the considered process, e.g. both scales set to the Z mass in electroweak scattering.

Parton distribution functions The parton distribution functions (PDFs) can be interpreted as probability densities of finding a parton with a particular longitudinal momentum fraction x within the proton. However, they do not constitute observables and need not be positive definite. Since the behavior of partons within the proton is governed by soft interactions, the PDFs are not calculable with pertubative QCD. Equation (1.3) can be seen as a splitting, determined by the value of the factorization scale μ_F, of long and short distance physics. Soft interactions involving gluons with a momentum transfer lower than μ_F are included in the PDFs, while the others are treated in the matrix elements of the parton cross-section calculation.

Figure 1.3 shows as an example the leading order PDFs for the various partons as calculated by the MSTW group [29]. At low x the gluons, who carry about half of the total proton momentum, dominate whereas the valence quarks peak at high momentum fractions. Due to their non-pertubative nature and the current limitations of lattice QCD calculations, the PDFs cannot be calculated from first principles and thus need to be determined experimentally by e.g. deep inelastic lepton-proton scattering.

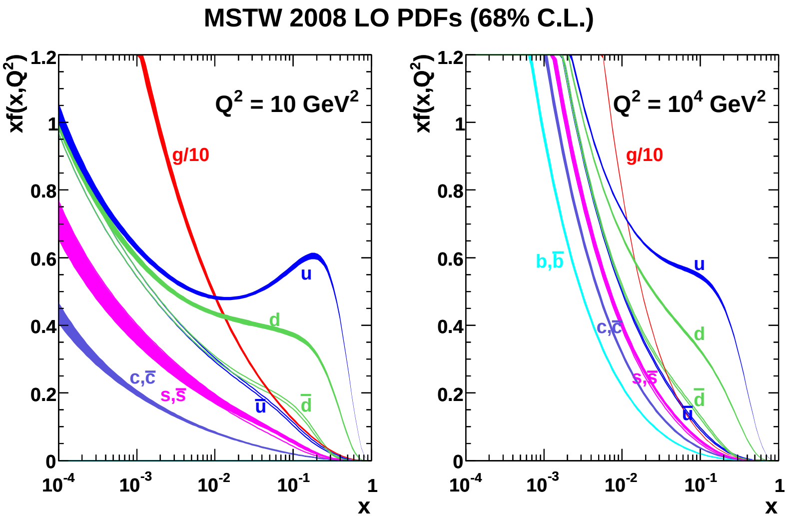


Figure 1.3: MSTW leading order parton distribution functions [30]. x and Q represent the proton momentum fraction of the parton and the momentum transfer of the interaction respectively.

Final state A typical final state of a proton-proton collision, succinctly called *event* in high energy physics jargon, comprises

- **Particles produced in the hard interaction:** In general, these are the interesting features for the analysis. Due to QCD color confinement, produced quarks and gluons fragment into jets of hadrons.

- **Initial (ISR) and final state radiation (FSR):** Colored particles emit bremsstrahlung in the form of gluons prior to and after the collision, which appear as additional jets in the measurement.

- **Proton remnants:** The rest of the protons fragment into color-neutral hadrons resulting in extra calorimeter activitiy. Furthermore, several partons of the incoming protons can undergo scattering. All this activity is embraced by the term *underlying event.*

- **Multiple interaction:** Due to the large number of protons per bunch (up to 10^{11}) several proton pairs can collide in a bunch crossing which results in a *pile-up* of final states. When operating with small distances between the bunches, final states can also get contaminated by collisions in adjacent bunch crossings.

1.3.2 Event simulation

As was outlined above, the final state of an event comprises a multitude of different objects from various origins. The modern approach of simulating such events is based on an interplay of different programs, each providing components of the simulation in a modular fashion. These programs are generally called *Monte Carlo generators*, since stochastic Monte Carlo methods are employed to sample and integrate over phase spaces. After presenting some key concepts of the simulation, the Monte Carlo generators and detector simulations used in the analysis of Part II are introduced. The explanations are based on the introductory texts of [31, 32].

Generalities The starting point of the simulation is the analytical calculation of a *leading order* hard scattering of two incoming partons in two outgoing particles, which can be abbreviated as a "$2 \to 2$ process", combined with a sampling of the phase space of the particles' kinematics yielding candidate events. These events are, however, non-physical since isolated partons do not exist and partons are expected to radiate other particles due to the strong coupling constant associated with the color charge, which make up the initial and final state radiation described above. The first reason necessitates simulating the formation of bound color-neutral states, also called *hadronization*, which is of non-pertubative nature. The second reason calls for the estimation of effects due to higher-order corrections in pertubation theory, which give rise to additional particle radiation. Higher-order corrections can either be included directly in the matrix element calculation or added afterwards using a recursive procedure called *parton showering.* Figure 1.4 illustrates the *evolution* of an event, starting with the hard process of the partons, whose initial momenta are determined by the PDFs. Additional radiation is here provided by parton showers, followed by the hadronization into color-neutral states, which may decay further. In general, there is also radiation from the incoming partons (ISR) present, which is not shown here for simplicity.

Parton shower As indicated in figure 1.4, the parton shower method starts with the hard process and evolves the event by allowing partons to split in pairs of other partons, which split further until the recursive procedure reaches an infrared cutoff value. The splitting

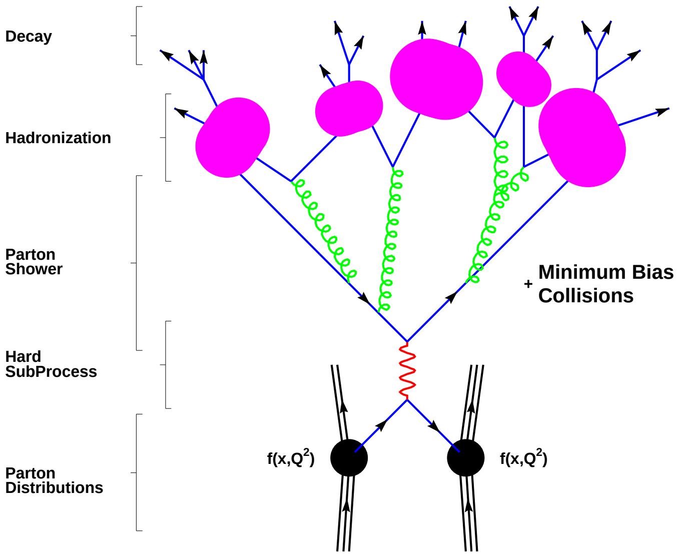


Figure 1.4: Example event evolution with time progressing upwards [33].

probablility is determined by the DGLAP equations (see e.g. [34]). In that way, an initial hard parton is replaced by several parton shower *jets*, expanding a $2 \to 2$ to a $2 \to n$ process. This method yields a good description of soft and collinear emission, but is less suited for hard and wide-angle radiation. It can also be employed to model electromagnetic showers. PYTHIA [35] is a typical shower generator.

Higher-order matrix elements Two approaches exist which describe higher-order effects directly in the matrix element of the hard scattering. The first is realized in so-called *Tree Level Matrix Element generators* such as ALPGEN [36]. It considers only real emissions (Feynman diagrams at *tree* or *Born level*), with the number of emissions coinciding with the pertubative order in the strong coupling constant. The absence of virtual loops and associated difficulties in the regularization of the matrix elements facilitates the calculation of final states possessing up to about ten partons. The second approach includes both real and virtual corrections to provide a complete prediction at a given pertubative order. To date, only generators at the next-to-leading order, i.e. including one extra emission, are available with MC@NLO [37, 38, 39] being a prominent example.

Hadronization The hadronization, often also called *fragmentation*, of colored partons into colorless hadrons takes place in the non-pertubative regime of QCD and is therefore only simulated with phenomenological models, whose parameters are determined by experimental data. A successful model is the *Lund String Model* [40], which is incorporated in PYTHIA. It treats gluons as self-attracting field lines forming tube-like strings between quarks. When the distance is increased, the strings break, thereby creating quark–anti-quark pairs from the vacuum. This implies the production of additional hadrons between the particle jets of the separating quarks, which has been observed experimentally.

Monte Carlo generators The following programs have been employed to generate simulated events for the analysis presented in Part II of the thesis:

- **PYTHIA [35]:** This so-called *general-purpose generator* is equally well capable of simulating hadronic events in proton-proton, electron-proton and electron-positron collision covering Standard Model, as well as supersymmetric and exotic, $2 \rightarrow 2$ (and some $2 \rightarrow 1$ and $2 \rightarrow 3$) processes. Also the modeling of the underlying event, the hadronization via the Lund String Model and subsequent particle decays are included. ISR and FSR are modeled with parton showers. Therefore, PYTHIA is less suited for the prediction of final states with additional high-energetic jets, such as events comprising W or Z boson production together with several jets.

- **HERWIG [41, 42]:** Like PYTHIA, HERWIG is a general-purpose generator. The simulation of the underlying event can be extended by combining with an external package called JIMMY [43], which models multiple scattering in the collision of two protons. HERWIG++[44] is an updated version written completely in C++ featuring a new parton shower and an improved hadronization model.

- **ALPGEN [36]:** This generator has been designed to accurately simulate Standard Model final states with high jet multiplicities. The modeling of additional jets is based on the direct calculation of the corresponding leading order matrix elements. Parton showering, provided by an interface to PYTHIA or HERWIG, produces additional soft and collinear jets for which the tree level matrix elements would get divergent. In order to avoid double-counting in the transitional regime, an algorithm based on the MLM technique [45] is used to merge parton shower with matrix element jets.

- **MC@NLO [37, 38, 39]:** The MC@NLO generator includes the complete next-to-leading order QCD corrections in the calculation of the hard scattering for certain processes, e.g. W/Z boson and top pair production, which improves the prediction of the total rate and the emission of an additional hard parton. It can be interfaced with HERWIG for showering and hadronization. A noteworthy particularity is the occurence of events with negative unit weight (typically 15% of all events), which are per se non-physical but yield, for sufficienctly high statistics, together with the postive weight events the full NLO prediction.

- **JIMMY [43]:** This program is a library of routines, which can be linked with HERWIG to simulate multiple parton scattering events.

- **FEWZ [46]:** FEWZ (*Fully Exclusive W and Z Production*) is used to determine the cross-section of lepton pair production via the Drell-Yan process at next-to-next-to-leading order in the strong coupling constant.

- **ISAJET [47]:** ISAJET is also a general-purpose generator, but is here only used to generate mass spectra of the SUSY particles, which are then fed into HERWIG for event generation.
- **PROSPINO [48, 49, 50]:** This program is used to calculate the next-to-leading order cross-sections for the production of supersymmetric particles.

Detector simulation Two different programs are available to simulate the interaction and measurement of the particles with the ATLAS detector. A time-consuming full simulation and a fast parametrized version, which can be used if the accurate modeling of detector effects is not necessary.

- **Full simulation:** The full detector simulation is carried out with the GEANT4 [51] toolkit. It models the passage of particles through matter under the influence of the applied (non-uniform) magnetic field and accounts for interactions with the various subdetectors and other obstructing material, resulting in e.g. multiple scattering, bremsstrahlung or ionization. In addition, it simulates the digital response of the subdetectors to the measured hits.
- **Fast simulation:** The fast simulation ATLFAST [52] consists of a parametrized smearing of the momenta and energies of the particles based on the expected performance of the ATLAS detector as predicted by the full simulation. Furthermore, a fraction of the particles are removed to reproduce the expected reconstruction efficiency of the subdetectors.

1.3.3 Relevant Standard Model processes

The most relevant Standard Model processes, which constitute an irreducible background to the in Part II presented searches for supersymmetry using one-lepton final states with at least two high-energetic jets, shall be presented in the following. Details on the simulation of each process are given in dedicated subsections of each chapter in Part II (subsections 3.2, 4.3 and 5.2) .

- **Top production:** As will become apparent in the search analysis, the production of top quarks, especially top quark pairs, with at least one quark decaying into an electron or muon makes up most of the selected final states. Due to its very short lifetime of about 0.5×10^{-24} s [12] the top quark decays to a W and a down-type quark (b quark in more than 99% of cases) before forming a hadronic state with other quarks [12]. The decay of the W thus defines the observed final state. Figure 1.5 shows the dominant contributions to top production, which are modeled with the MC@NLO generator including all next-to-leading order corrections. Besides so-called *semi-leptonic top pair decays*, where only one of the W boson decays leptonically and the other to a quark pair, also *di-leptonic decays*, where one of the leptons is out of the measurement acceptance, not reconstructed or mis-identified as a hadronic object, pass the selection cuts of the analysis.

- **Vector boson plus jets production:** The other dominant type of processes resulting in observed isolated electrons or muons are the production of W and to a lesser extent Z bosons, collectively called *vector bosons (V)*, which decay leptonically. The diagrams in figure 1.6 illustrate some example processes of W with associated jet production. The modeling of $V+$ jets for the search analysis of Part II is, wherever possible, done using ALPGEN in order to simulate the jet multiplicities more accurately.
- **QCD multi-jet production:** Requiring the measurement of an isolated lepton rejects most of the QCD multi-jet final states. Still, due its huge overall rate, it remains an important background with isolated leptons being faked by jets or coming from heavy particle decays within jets. Besides the necessary high statistics of to be generated MC events, the simulation of QCD is further complicated by the correct modeling of the missing energy, which arises here mostly from mismeasurements. Therefore, QCD is simulated with PYTHIA, which is less costly compared to ALPGEN in computional terms, but the prediction needs to be verified with data-driven methods.

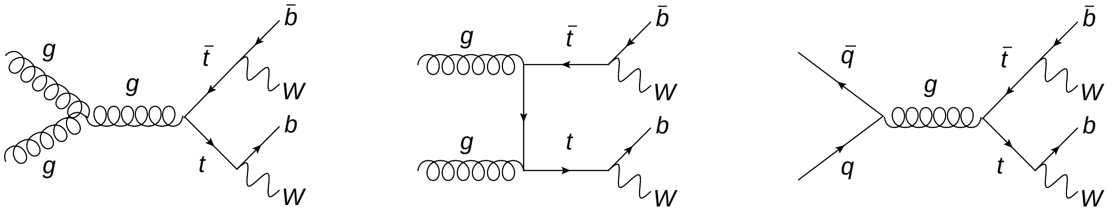


Figure 1.5: Example diagrams for top pair production at tree level via gluon–gluon fusion (left and middle), which is the dominant process at the LHC, and quark–anti-quark annihilation (right). The top quarks decay before they can hadronize. Time progresses to the right.

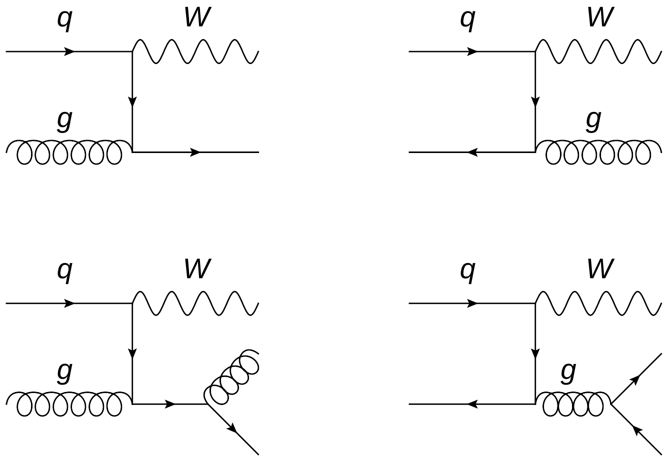


Figure 1.6: Example diagrams for W and associated production of one (top row) and two (bottom row) additional jets at tree level via quark–gluon scattering (left) and quark–anti-quark annihilation (right). Time progresses to the right.

1.3.4 Exemplary supersymmetry signals

SUSY pair production In the research presented in Part II, only R-parity conserving supersymmetry signals are considered. This implies a pair production of SUSY particles, dominantly squarks and gluinos through the strong interaction. Figures 1.7 and 1.8 sketch gluino and squark production at leading order via gluon–gluon and gluon–quark fusion respectively. Figure 1.9 shows the production via quark–anti-quark annihilation and quark–quark scattering. Whichever production is dominant depends on the masses of the squarks and gluinos, which differ strongly in the various SUSY models. Electroweak production of SUSY particles, which can also directly produce gauginos, plays a comparatively minor role at the LHC.

Cascade decays The decays of the produced SUSY particles [9] often involves several steps and are thus called *cascade-decays*. If the squarks are heavier than the gluinos, the strong decay $\tilde{q} \to q\tilde{g}$ will dominate. Otherwise, they will decay weakly to charginos, $\tilde{q} \to q\tilde{\chi}_i^{\pm}$, or neutralinos, $\tilde{q} \to q\tilde{\chi}_i^0$, which decay further to the lightest neutralino (LSP). Gluinos can only decay via the strong interaction, where the process $\tilde{g} \to q\tilde{q}$ is the most dominant, provided the involved squark has a lower mass than the gluino. If all the squarks are heavier than the gluino, it will decay via an off-shell squark producing a chargino, $\tilde{g} \to qq'\tilde{\chi}_i^{\pm}$, or a neutralino, $\tilde{g} \to qq'\tilde{\chi}_i^0$. Some examples for gluino decay are illustrated in figure 1.10.

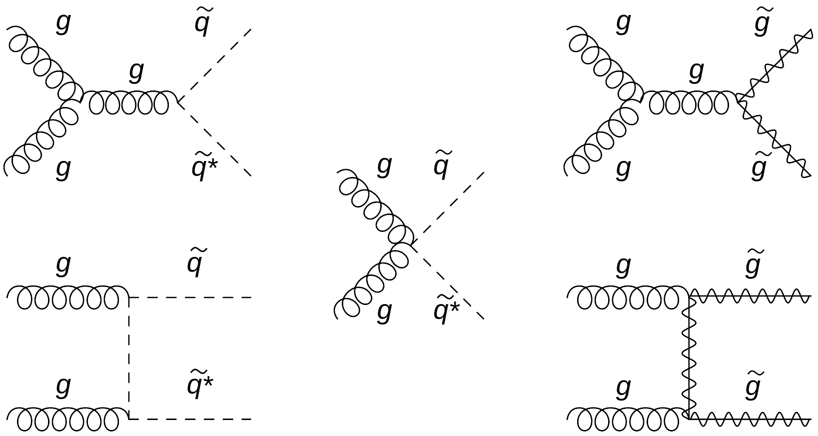


Figure 1.7: Example diagrams for squark and gluino pair production at tree level via gluon fusion. Adapted from [9]. Time progresses to the right.

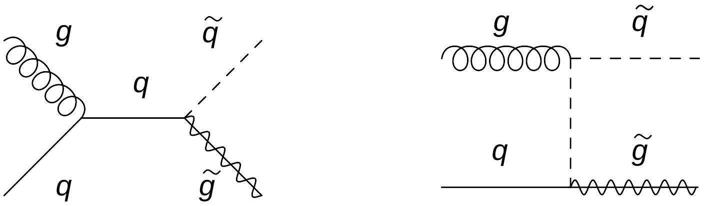


Figure 1.8: Example diagrams for squark and gluino production at tree level via quark–gluon fusion. Adapted from [9]. Time progresses to the right.

SUSY final state Assuming that R-parity is conserved, a typical final state of pair-produced colored SUSY particles therefore includes

- several high-energetic jets,
- two LSPs, which are, in order to meet dark matter constraints, gravitationally and, depending on the SUSY scenario, weakly interacting and thus not directly detectable,
- potentially isolated leptons from weak decays.

Simulation As mentioned in the last subsection, in the analysis of Part II the SUSY particle mass spectra and branching ratios are generated with ISAJET interfaced with HERWIG for event generation. Next-to-leading order production cross-sections are calculated with PROSPINO.

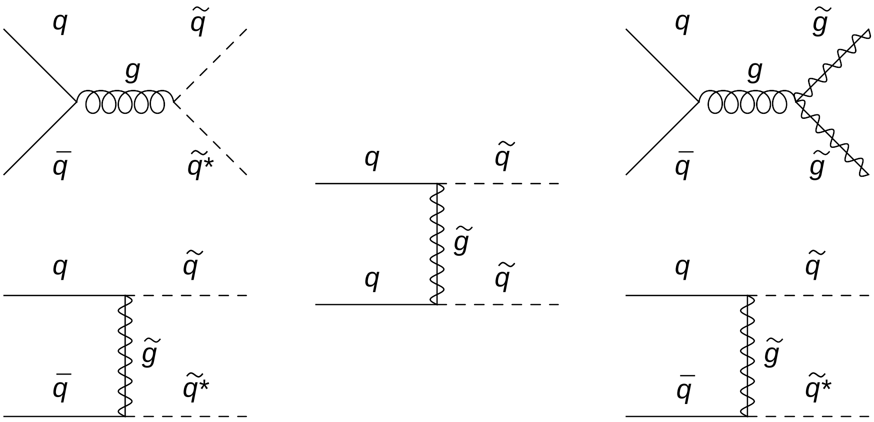


Figure 1.9: Example diagrams for squark and gluino pair production at tree level via quark–anti-quark annihilation and quark–quark scattering (middle). Adapted from [9]. Time progresses to the right.

Figure 1.10: Possible gluino cascade decays resulting in final states with quarks, one neutralino as LSP and additional fermions (f), which can be either quarks or leptons. The involved squarks can be either on- or off-shell depending on the mass spectrum of the considered scenario. Adapted from [9]. Time progresses to the right.

Chapter 2

The ATLAS detector at the Large Hadron Collider

This chapter provides an introduction to the ATLAS (A Toroidal LHC ApparatuS) detector at the Large Hadron Collidor (LHC) focusing on the aspects relevant for the search analysis presented in Part II of the thesis. The explanations follow to a great extent [53, 54, 55].

2.1 The Large Hadron Collider

The Large Hadron Collider is currently the world's largest and highest-energy particle accelerator, built to extend the understanding of the physical world at its most fundamental level and in particular help answer key questions of elementary particle physics. It is designed to collide particle beams of protons and lead nuclei at unprecedented center-of-mass energies of 14 TeV and 1148 TeV, respectively. The experiments conducted at the LHC are expected to shed light on unresolved issues of particle physics, among which the existance of the Higgs boson, new heavy particles due to supersymmetry or elusive extra dimensions are the most prominent. In addition, the LHC will provide precision measurements of known Standard Model phenomena. The LHC is built in the former LEP [56] tunnel at CERN [57] (originally *Conseil Européen pour la Recherche Nucléaire*), Geneva, which lies up to 175 m beneath the surface and has a circumference of approximately 27 km. Operated at the design specifications, it will circulate two particle beams, each containing 2,808 bunches of up to 10^{11} protons per bunch, at a velocity close to the speed of light, resulting in 40 million bunch collisions per second at four interaction points equivalent to a luminosity of $10^{34}\,\mathrm{cm}^{-1}\,\mathrm{s}^{-1}$. On the average, 23 proton-proton scatterings are expected per bunch collision. The beams are kept on track by 1,232 superconducting dipole magnets producing a magnetic field of up to 8.3 T and kept focused by 392 quadrupole magnets.

An accident on Sep. 19th, 2008 involving an electric fault in the connection between two magnets (see e.g. [58]) and its implications delayed the start-up of the LHC by more than a year. As a consequence, it was decided to operate the machine with beam energies at half of their design intensity during the first period of data taking, resulting in proton-proton collisions at a centre-of-mass-energy of 7 TeV. The design power will presumably not be achieved before 2014.

The sketch in figure 2.1 shows the geographical positions of the main experiments at the LHC. Like ATLAS [55], CMS [59] (Compact Muon Solenoid) is a general-purpose detector directly competing with the former in the search for new physics. ALICE [60] (A Large Ion Collider Experiment) is an experiment specialized in studying heavy ion collisions, in particular the properties of *quark-gluon plasma*. b-physics and the related CP violation are the main field of investigation of the LHCb [61] (Large Hadron Collider beauty) experiment. Two much smaller and very specialized detectors are not shown in the sketch. TOTEM [62] (TOTal Elastic and diffractive cross section Measurement), which is technically integrated into CMS, measures the total cross-section as well as elastic and diffractive scattering of proton-proton collisions. LHCf [63] (Large Hadron Collider forward) is an experiment dedicated to the measurement of neutral particles (e.g. pions) emitted in the very forward region of collisions, intended to help explain the origin of very high-energetic cosmic rays. It is located close to ATLAS.

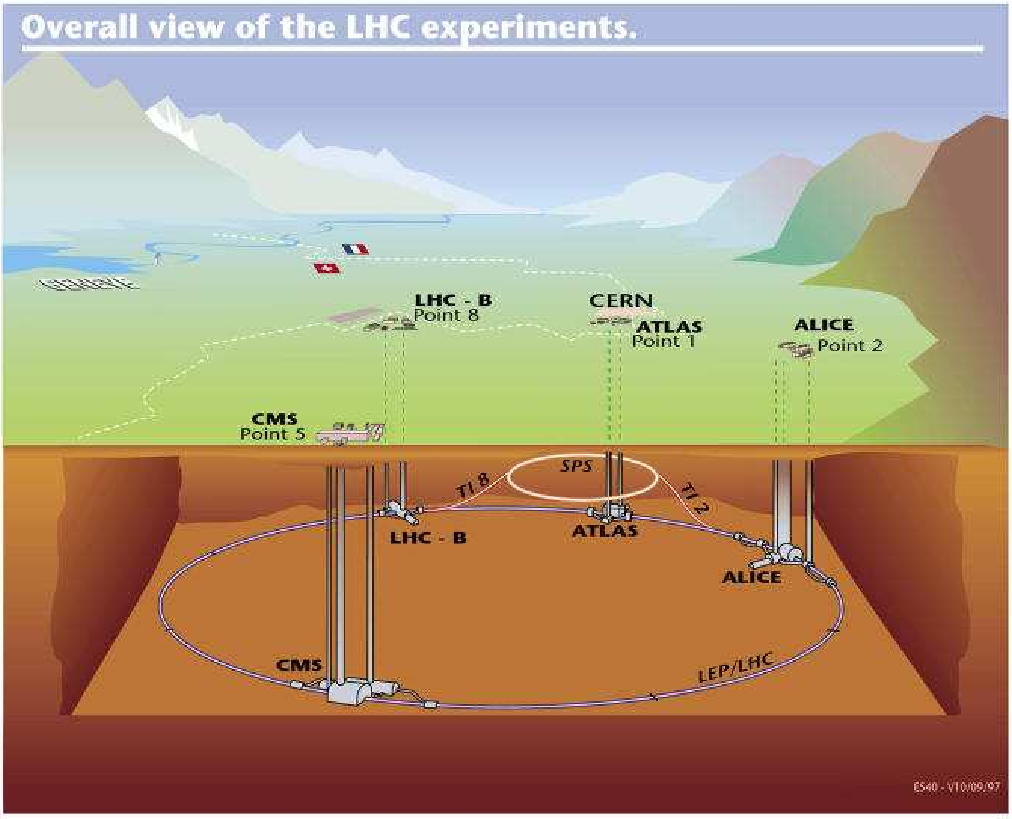


Figure 2.1: The Large Hadron Collider and its main experiments [64].

2.2 The ATLAS detector

ATLAS is a so-called general-purpose detector since it is designed to measure a very broad range of different signals, involving the detection of hadrons, electrons, muons, photons and neutrinos. Each of the stated species is measured by a single or combination of dedicated sub-detectors. Figure 2.2 sketches the measurement of the particles in the typical arrangement of the sub-detectors. Only electrically charged particles leave a track in the inner detector, whose bending caused by a magnet field can be used to determine their momentum. Electrons and photons are measured in the electromagnetic calorimeter producing showers of secondary particles due to bremsstrahlung and pair production, whereas hadrons typically deposit smaller amounts of their energy in this sub-detector due to the lower bremsstrahlung as a consequence of their larger mass. The latter also produce particle showers in the hadronic calorimeter via strong interaction cascades. The minimum ionizing muons lose very little of their energy when passing through all detectors. Magnetic fields are thus used to determine the momentum from the curvature of their trajectories in the inner detector and the muon system. Since neutrinos are only weakly interacting, they cannot be directly detected and their presence can only be inferred from conservation of momentum of all particles involved in a collision.

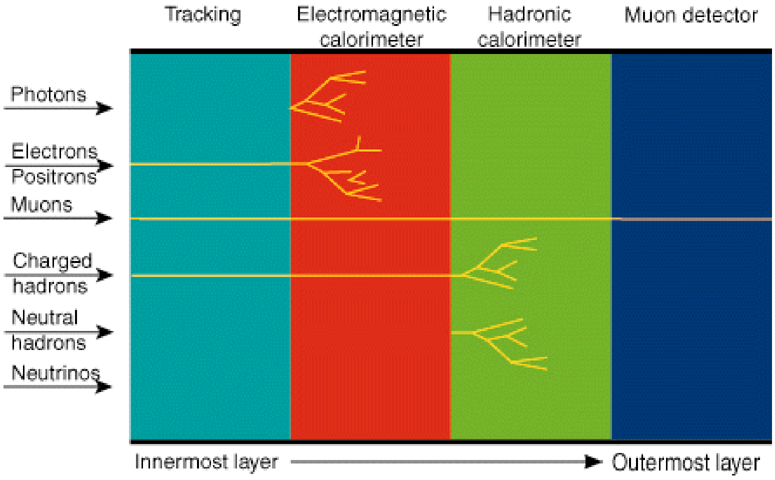


Figure 2.2: The imprints of different particle types in the different layers of a particle detector [65].

The overall layout of the ATLAS detector is shown in figure 2.3. It weighs approximately 7,000 tons, is 44 m long and 25 m tall. The sub-detectors are built as a series of ever-larger concentric cylinders around the nominal interaction point in order to cover almost the whole solid angle (so-called 4π-*detector*). The solenoid magnet surrounds the inner detector, which consists of the semiconductor pixel and strip detectors and the transistion radiation tracker. The toriod magnets are situated outside the electromagnetic and hadronic calorimeters and enclose part of the muon system.

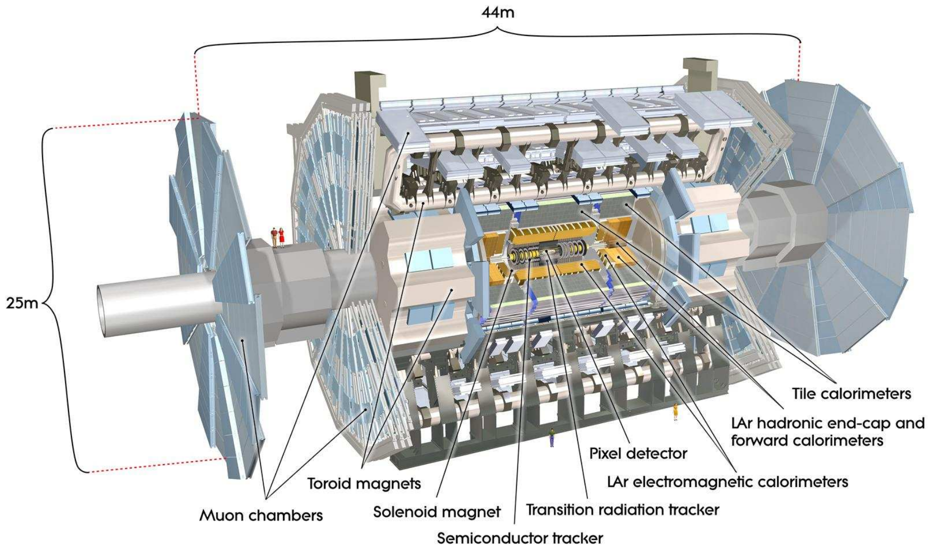


Figure 2.3: The ATLAS detector [66].

Before discussing the various sub-detectors in more detail, the coordinate system used at ATLAS is introduced in the following subsection. Recent performance studies using minimum-bias and cosmic-ray events can be found in [67, 68].

2.2.1 The ATLAS coordinate system

The origin of the coordinate system used at ATLAS is the nominal interaction point of the proton-proton collisions. The z-axis lies along the direction of the beam whereas the y- and x-axis are chosen to point upwards and towards the centre of the ring, respectively. The azimuthal angle ϕ is measured in the x–y plane (or *transverse plane*) and the polar angle θ is calculated with respect to the z-axis. Instead of θ one usually uses the pseudorapidity η, defined as

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

which often yields flatter particle production distributions. The distance ΔR in η–ϕ space is defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}.$$

2.2.2 The inner detector

Since approximately 1,000 particles will emerge from the collision point every 25 ns within $|\eta| < 2.5$ at the design luminosity, the inner detector (ID) needs to be robust against hard radiation and have a very high granularity to resolve the multitude of tracks and vertices. Figure 2.4 illustrates the layout of the inner detector. The whole inner detector is contained in a solenoid magnetic field of 2 T, which bends the particles' trajectories to measure their momenta. Just outside the beryllium beam-pipe, the three layers of the high granularity pixel dectector with cell sizes of $50 \times 400\,\mu$m are situated. The pixel detector has more than 80 million readout channels, which is about 50% of the total readout channels. Its close proximity to the collision point is essential for the measurement of secondary vertices caused by particle (e.g. b quark) decays. The semi-conductor tracker (SCT) is the middle component of the ID with four double layers of silicon strips and approximately 6.3 million readout channels. The reduced particle flux allows for an approximate sensor size of $80\,\mu\mathrm{m} \times 12\,\mathrm{cm}$. Both ID and SCT use silicon p–n junctions operated at reverse bias and cover the region up to $|\eta| < 2.5$. At radii greater than 56 cm, typically 36 tracking points are provided by the 300,000 straw tubes of the transition radiation tracker (TRT), which covers tracks up to $|\eta| < 2.0$. The tubes have a diameter of 4 mm and a length of 144 cm in the barrel and 37 cm in the end-caps, where they are arranged radially in wheels. They provide information on the radius and azimuthal angle of the tracking points with approximately 351,000 readout channels. The TRT assists in discriminating electrons from heavier charged particles (e.g. pions) by efficiently absorbing the transition radiation photons in the xenon-based gas mixture within the straw tubes. For recent results on the alignment of the tracking system using collision and cosmic-ray tracks see [69].

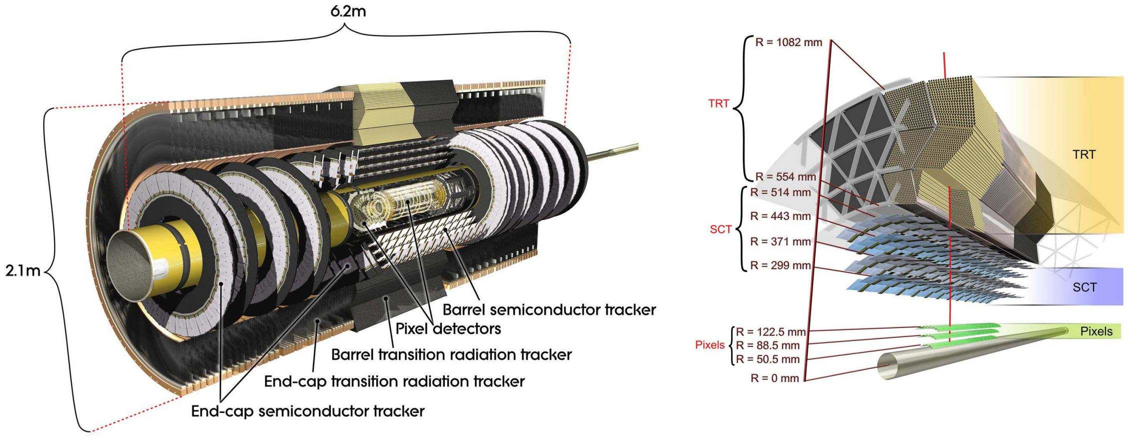


Figure 2.4: The ATLAS inner detector [70].

2.2.3 The calorimeter

The calorimeter system, displayed in figure 2.5, comprises two components: the electromagnetic and the hadronic calorimeter. These are so-called *sampling calorimeters* as they consist of alternating layers of high-density absorption material (metal) and layers which can measure the deposited energy (liquid argon or scintillators), inferring the original particle energy from the sampled particle shower. The calorimeter has a complete azimuthal symmetry and measures particles up to very high pseudorapidity values in order to facilitate an accurate inference of missing transverse momentum associated with non-detectable particles.

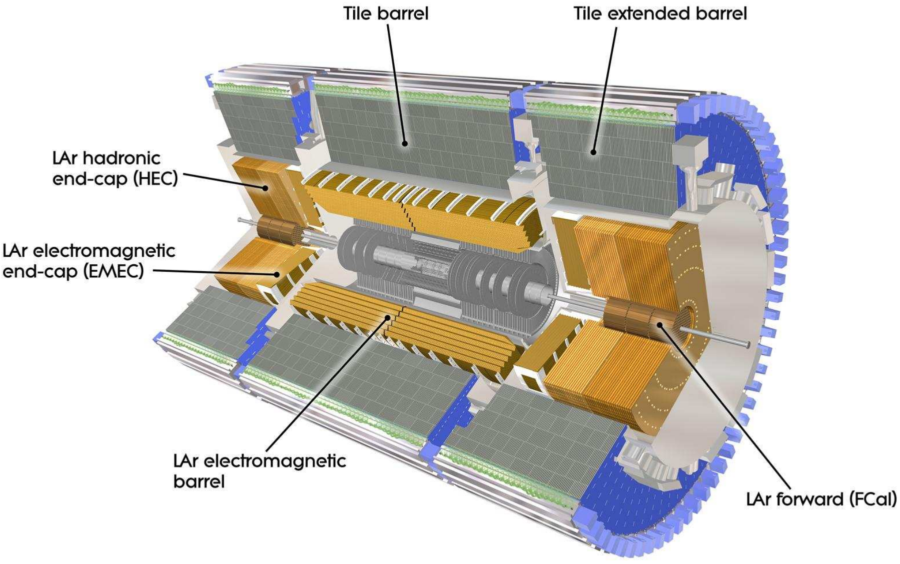


Figure 2.5: The ATLAS calorimeter [71].

The electromagnetic calorimeter covers a pseudorapidity region up to $|\eta| < 3.2$ and consists of lead absorption and liquid argon sampling layers. In the region up to $|\eta| < 1.8$, a liquid argon presampler detector is used to correct for the energy lost by electrons and photons upstream of the calorimeter. The hadronic barrel calorimeter, covering up to $|\eta| < 1.7$, uses plastic scintillator tiles embedded in an iron absorber. The hadronic end-cap calorimeters with a coverage from $|\eta| > 1.5$ to $|\eta| < 3.2$ use again liquid argon technology, due to intrinsic radiation tolerance needed in the forward region, together with copper plates as absorbers. The forward calorimeter in the region $3.1 < |\eta| < 4.9$ comprises three modules where the first, using copper, is optimized for electromagnetic measurements and the other two, using tungsten, measure mainly the energy of hadronic interactions. The readiness of the calorimeters has undergone extensive successful testing [72, 73].

2.2.4 The muon system

The detection of muons is based on the magnetic deflection of their trajectories using three superconducting air-core toroid magnets situated outside the calorimeter system, one of which in the form of eight large loops in the barrel region and the other two inserted in the end-caps consisting of eight coils each. Figure 2.6 shows the positions of the magnets and the four different types of muon chambers, which can be divided in two basics sets with one dedicated to the triggering on muons and the other providing precisions measurements of the muon tracks. The design of the muon system complies with the requirement to measure the transverse momenta of 1 TeV muons with a resolution of less than 10%.

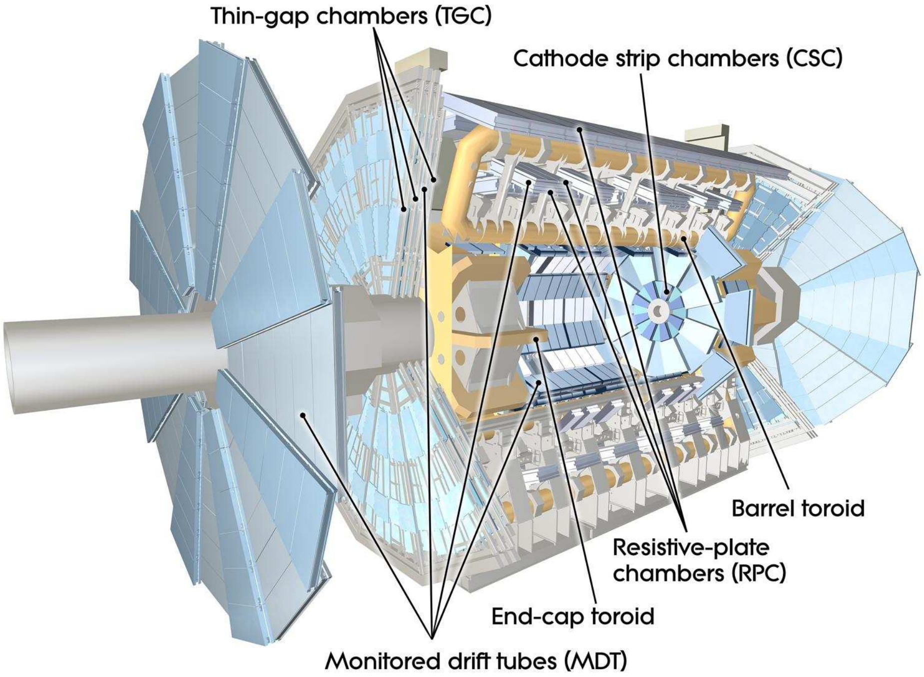


Figure 2.6: The ATLAS muon system [74].

Muon triggering Resistive Plate Chambers (RPC) in the barrel region (up to $|\eta| < 1.05$) and Thin Gap Chambers (TGC) in the end-caps ($1.05 < |\eta| < 2.4$) constitute the fast muon trigger system, which delivers track information on a timescale of a few tens of nanoseconds after the passage of the particle. Complementary to the precision chambers, the trigger chambers measure a second track coordinate orthogonal to the one provided by the former.

Muon track measurement Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSC) provide high precision tracking up to an absolute pseudorapidity of 2.7. The former contain three to eight layers of drift tubes and cover most of the η region, while the latter are multiwire proportional chambers with cathode planes segmented into strips in orthogonal

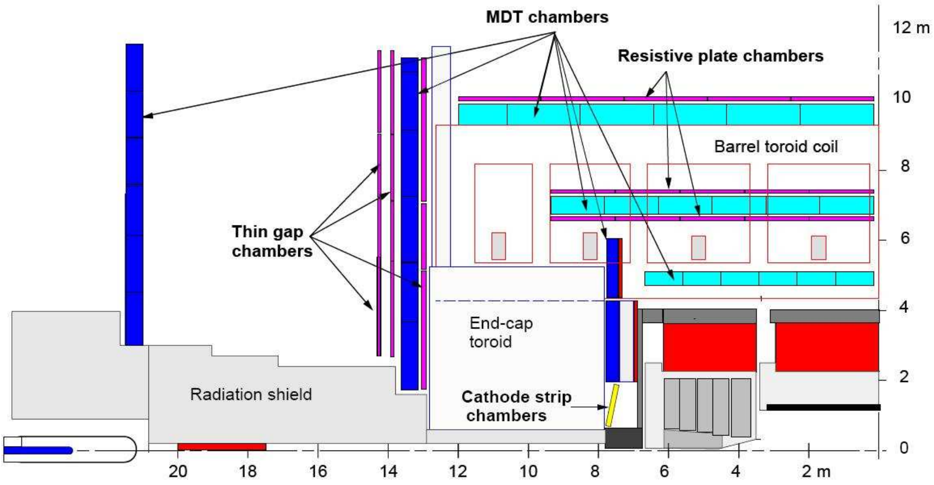


Figure 2.7: The ATLAS muon system [1]. The sub-detectors are arranged in stations at increasing distance from the nominal collision point, which is at the center of the displayed coordinate system.

directions and offer higher granularity needed in the forward region. The subdetectors are located in three separated stations at increasing distance from the nominal collision point as displayed in figure 2.7.

2.2.5 Trigger system

As mentioned above, approximately 40 million proton-proton collision each second are expected when operating the LHC at the design luminosity. Only a tiny fraction of the resulting events, where two partons collide *head-on*, are of interest for the majority of the physics analyses. These events need to be singled out and stored for the so-called *offline analysis*.

This task is handled by the three levels of the ATLAS trigger system. The first level (L1), which is hardware-based, uses only a limited portion of the detector information in order to provide fast decisions within $2.5\,\mu$s, reducing the potentially interesting events to about 75,000 per second. In addition, it flags the regions within the events where its selection process has identified interesting features. The information fed into this trigger includes the transverse momenta of leptons, photons and jets as well as a crude determination of the missing momentum and the total transverse energy. The software-based second (L2) and third (Event Filter) levels exploit the full granularity and precision of the sub-detectors together with gradually more sophisticated algorithms to further reduce the number of interesting events to 3,500 (decision time 40 ms) and 200 (decision time 4 s) per second, respectively.

2.2.6 Object identification and reconstruction

This subsection gives an introduction to the identification and reconstruction of objects – i.e. jets, electrons, muons and missing momentum – used in the supersymmetry search presented in Part II. Photons and taus are not considered. Details on the object definition are given in the dedicated subsections of each chapter (subsections 3.3, 4.5 and 5.4).

Jets

The measurement of quarks and gluinos through the detection of their associated jets of hadrons is driven by several requirements (see e.g. [1]). From a theoretical point of view, the jet reconstruction should reflect the original hard scattering and it should be infrared- and collinear-safe, which means that the reconstruction is to be invariant with respect to additional emission of a soft or collinear partons within the jet. Furthermore, the reconstruction should not be affected by other activity in the event or by signal-dependent detector effects and it should display a high reconstruction efficiency.

Two different jet finding algorithms are used in the analyes of Part II: The *fixed cone jet* finder fed by calorimeter tower signals as described on pages 261ff. of [1] and the anti-k_t *algorithm* [75] using topological cell clusters as input.

In the first case, tower signals, defined by overlaying the calorimeter cells with a fine grid in η and ϕ (e.g. $\Delta\eta = \Delta\phi = 0.1$), are taken as input for the algorithm. The cells with highest transverse momenta surpassing a given threshold act as seeds, around which all objects within a certain cone radius (typically $\Delta R = 0.4$ or 0.7) are assigned to a new jet in a recursive manner. This algorithm is generally not infrared-safe, which can be partly improved upon by using *split and merge* techniques between adjacent jets: After stable cones are established, cones which share energy are either merged into one jet or split in two separate ones depending on the fraction of the energy they have in common.

The anti-k_t algorithm employed uses topological cell clusters as input, which are three dimensional energy deposits constructed by imposing decreasing signal-to-noise thresholds on cells with increasing distance from seed cells. A distance measure between the clusters depending on their transverse momenta and angular separation is used to uniquely assign them to then collinear- and infrared-safe jets.

For the jets considered in the analysis of Part II, the reconstruction efficiency of the employed algorithm is close to 100% (see e.g. [76]).

Electrons

Electrons are reconstructed from measurements taken both in the inner detector and the calorimeters. In principle, candidates can be chosen using a cut-based approach or multivariate analysis techniques. In the analyes of Part II, only the cut-based method is employed, for which three reference sets of cuts have been defined (see pages 75ff. of [1]): *loose*, *medium* and *tight*. Generally, a looser selection exhibits a higher reconstruction efficiency, but also a higher contamination of jets faking electrons.

The electron reconstruction starts by selecting electromagnetic seed clusters using a sliding window algorithm, which are then tried to match to a track in the inner detector subject to the cuts defined by the three different sets:

- **Loose cuts:** These cuts include requirements on the lateral shape and width of the shower in the second layer of the electromagnetic calorimeter, where electrons typically deposit most of their energy. Furthermore, an upper cut on the ratio of the transverse energy in the first layer of the hadronic calorimeter over the transverse energy of the electromagnetic cluster, which quantifies the *hadronic leakage*, is imposed.

- **Medium cuts:** The medium cuts include the loose cuts and require in addition cuts on the energy deposits in the first layer of the electromagnetic calorimeter and cuts on the matched track in the inner detector. The latter includes a minimal number of hits in the pixel and strip subdectectors and a maximum transverse impact parameter.

- **Tight cuts:** The tights cuts make full use of the available identification tools. Among others, the cuts on the track–cluster matching are tightened and TRT measurements are included on top of the medium selection.

The medium selection constitutes the basis for the electron definition in the analysis of Part II. The reconstruction efficiency for isolated, i.e. originating from vector boson decay, medium electrons amounts to about 80% to 90% depending on the details of the cuts applied [1, 77]. The relative energy resolution is estimated to be of the order of a few percent, as shown on pages 56ff. of [1] and in [77], and decreases to about 1% for energies above 200 GeV.

Muons

There exist different approaches to identify muons by using the muon spectrometer, the inner detector or a combination of the two, as decribed on pages 161ff. of [1]. *Standalone muons* are reconstructed by identifying tracks in the muon system and extrapolating them to the beamline. Standalone muons can also be matched to tracks found in the inner detector resulting in so-called *combined muons*. Finally, extrapolated inner detector tracks can be associated with nearby hits in the muon spectrometer yielding *tagged muons*. Two standard algorithms are available for each approach, grouped in two collections, which are labeled by the name of the respective combined alogrithm: *Staco* and *Muid*.

In the analyes of Part II, combined Staco muons are considered (except where stated otherwise), which are reconstructed using a statistical combination of the two independent track measurements in the inner detector and the muon spectrometer by means of their covariance matrices. The difference between the two track vectors weighted by the combined covariance matrix, expressed as a χ^2-value, is used as a quality criterium for the track matching. The reconstruction efficiency for combined muons with $\chi^2 < 100$ is well above 90% (pages 174ff. of [1] and also e.g. [78]). The relative muon transverse momentum resolution is of the order of a few percent, as shown on pages 175ff. of [1] and in [79].

Missing transverse momentum

The missing transverse momentum, which is defined in section 3.4, can be determined by either of two different approaches, both of which combine measurements in the calorimeter and reconstructed muon tracks (see pages 368ff. of [1]). The *cell-based method* starts with calorimeter energy deposits above a certain threshold which are then reweighted according to the objects they are associated with. The *object-based method* starts with the reconstructed objects and considers the remaining deposited energy as originating from charged and neutral pions. More details on the employed definition of the missing transverse momentum are given in the dedicated subsections of each chapter.

Part II

Supersymmetry search in the one-lepton channel with ATLAS

Chapter 3

Prospects for supersymmetry discovery at $\sqrt{s} = 14$ TeV

3.1 Introduction

If supersymmetry is a fundamental symmetry of nature, there is good reason to believe, as argued in chapter 1, that the supersymmetric partners of the Standard Model particles have masses around the TeV-scale. Should they exist, they are expected to be abundantly produced by the proton-proton collisions at the LHC, giving rise to final states with hard jets and possibly high-energetic leptons from weak decays (cf. subsection 1.3.4). If, in addition, R-parity is conserved, SUSY particles will always be produced in pairs and the lightest supersymmetric particle (LSP) will be stable. For reasons outlined in section 1.2, the LSP is expected to be only weakly or gravitationally interacting and therefore escapes direct detection by ATLAS.

Thus, the typical signature of such a realization of supersymmetry is a final state with hard jets, substantial missing energy and potentially one or more isolated leptons.

The strategy of the ATLAS SUSY Working Group [80, 81] for the search for R-parity conserving supersymmetry has been to divide the analyses in exclusive lepton multiplicity final states. This allows for a more specialized estimation of the varying background composition of each subsignature. Some analyses have been further subdivided into different jet multiplicities to cover the different regions of phase space more effectively. Eventually, results from the various selections can be combined to increase the overall sensitivity.

The analysis presented in this thesis, denoted as *one-lepton analysis* in what follows, selects final states with exactly one isolated electron or muon in addition to several jets and missing energy in the plane transverse to the beams. Its aim is to look for significant deviations from the Standard Model expectations. If none are found, possible SUSY scenarios can be excluded and the existing limits for minimal masses of supersymmetric particles can be raised.

More generally, the following analysis can be considered as a search for any hitherto unknown signal which is produced through the strong force and features large missing energy in combination with an isolated lepton.

This chapter deals with the prospects of a supersymmetry search in the one-lepton channel for the LHC design centre-of-mass energy of $\sqrt{s} = 14$ TeV. It introduces the one-lepton search, which is then further refined in the two subsequent chapters where the analysis of the data collected with the ATLAS detector in 2010 is discussed.

The presented results have been published on pages 1589ff. of [1]. The author has significantly contributed to the work on the 4-jet selection and provided the studies on the 2- and 3-jet selections as well as the results on the discovery reach of the one-lepton analysis.

In the context of the analysis, *lepton* refers to either an electron or a muon and not a tau.

3.2 Signal and background simulation

In the following, the modeling of the relevant physics processes, as described in section 1.3, in this analysis is briefly discussed. For further details on the simulation see pages 6f. and 1515ff. of [1].

Signal samples

Since even the minimal supersymmetric extension of the Standard Model spans a 105-dimensional phase space, the mSUGRA realization of SUSY, introduced in section 1.2, is taken as a test scenario due to its low number of free parameters. The following mSUGRA benchmark points are chosen to cover a wide range of SUSY masses and decay topologies while being roughly consistent with the estimated cold dark matter density in the universe:

- SU1: $m_0 = 70$ GeV, $m_{1/2} = 350$ GeV, $A_0 = 0$, $\tan\beta = 10$, $\mu > 0$
- SU2: $m_0 = 3550$ GeV, $m_{1/2} = 300$ GeV, $A_0 = 0$, $\tan\beta = 10$, $\mu > 0$
- SU3: $m_0 = 100$ GeV, $m_{1/2} = 300$ GeV, $A_0 = -300$ GeV, $\tan\beta = 6$, $\mu > 0$
- SU4: $m_0 = 200$ GeV, $m_{1/2} = 160$ GeV, $A_0 = -400$ GeV, $\tan\beta = 10$, $\mu > 0$
- SU6: $m_0 = 320$ GeV, $m_{1/2} = 375$ GeV, $A_0 = 0$, $\tan\beta = 50$, $\mu > 0$
- SU8.1: $m_0 = 210$ GeV, $m_{1/2} = 360$ GeV, $A_0 = 0$, $\tan\beta = 40$, $\mu > 0$

Table 3.1 lists the corresponding cross-sections, calculated with PROSPINO 2.0.6 [48, 49, 50], and the number of Monte Carlo events for each sample. The samples have been generated with HERWIG [41, 42] using the particle spectrum predicted by ISAJET [47] and CTEQ6M [82] for the parton distribution, followed by the full detector simulation provided by GEANT4 [51]. The masses of the SUSY particles predicted for each point are shown in Table A.1 in the appendix.

Table 3.1: SUSY benchmark signals with their leading order (σ^{LO}) and next-to-leading order (σ^{NLO}) cross-sections, number of generated MC events (N) and corresponding integrated luminosity (L).

Signal	σ^{LO} [pb]	σ^{NLO} [pb]	N	L [fb^{-1}]
SU1	7.4	10.9	200,000	18.4
SU2	4.9	7.2	50,000	7.0
SU3	18.6	27.7	500,000	18.1
SU4	262.0	402.2	200,000	0.5
SU6	4.5	6.1	30,000	4.9
SU8.1	6.4	8.7	50,000	12.2

In order to estimate the discovery reach of the one-lepton analysis, two planes in the mSUGRA parameter space, defined by constant values of $\tan\beta$, A_0 and μ, are sampled:

- $\tan\beta = 10$, $A_0 = 0$, $\mu > 0$ grid: 625 points are chosen with 25 steps of 120 GeV from 60 to 2940 GeV in m_0 and 25 steps of 60 GeV from 30 to 1470 GeV in $m_{1/2}$. The computation of 25 points failed due to theoretical reasons.
- $\tan\beta = 50$, $A_0 = 0$, $\mu > 0$ grid: 225 points are chosen with 15 steps of 200 GeV from 200 to 3000 GeV in m_0 and 15 steps of 100 GeV from 100 to 1500 GeV in $m_{1/2}$.

For each point of the grids 20,000 events have been generated, again with HERWIG and ISAJET. Due to limited resources only the simplified detector simulation ATLFAST [52] has been applied. The leading order cross-sections provided by HERWIG have been used, which are generally smaller than the corresponding next-to-leading order ones.

Background samples

The following Standard Model processes contribute to the background in the one-lepton SUSY search with diverse importance: top pair production ($t\bar{t}$), $W+$ jets, $Z+$ jets, QCD multi-jets, and di-boson events.

Top pair production, which turns out to be the dominant background, has been modeled with MC@NLO [37, 38, 39]. HERWIG has been used for the simulation of the parton shower and the fragmentation of the particles in the jets. The underlying event has been modeled with JIMMY [43]. For $W/Z+$ jets ALPGEN [36] has been chosen in order to reliably predict the additional jets in the vector-boson production, again with HERWIG and JIMMY for hadronisation and the underlying event respectively. Next-to-next-to-leading order (NNLO) cross-section have been determined with FEWZ [46]. For the 2- and 3-jet selections, presented in section 3.8, PYTHIA has been used for the $W/Z+$ jets processes as the ALPGEN samples have been produced with a 4-jet generator filter.

Due to the large Monte Carlo statistics required for the QCD multi-jet events, they have been produced with PYTHIA [35]. The di-boson samples have been generated with HERWIG. The full detector simulation using GEANT4 has been applied on all background samples.

CTEQ6L [82] and CTEQ6M parton distributions have been used for cross-sections calculated at leading and next-to-leading order respectively. For all processes except QCD multi-jets, the number of generated events exceeds an equivalent integrated luminosity of $1\,\mathrm{fb}^{-1}$. The cross-sections of the main Standard Model backgrounds are summarized in table 3.2.

Table 3.2: Cross-sections of the most relevant backgrounds. The figures for the next-to-next-to-leading order (NNLO) W/Z cross-sections are stated per lepton flavor. The $t\bar{t}$ next-to-leading order (NLO) cross-section calculation includes a next-to-leading logarithm resummation (NLL) [83].

Process	σ [pb]	Order
$t\bar{t}$	833	NLO & NLL
$W \to \ell\nu$	20,500	NNLO
$Z \to \ell\ell$	2,020	NNLO
QCD	3.67×10^8	LO

3.3 Object definition

The definition of the objects considered in the analysis is based on the identification and reconstruction algorithms outlined in subsection 2.2.6.

Jets

Jets are reconstructed using the fixed cone algorithm, the definition and performance of which are explained in detail on pages 262ff. of [1]. A narrow cone size of $R_{cone} = 0.4$ is chosen to account for the expected high jet multiplicity of supersymmetry events. The jets are required to have a minimum transverse momentum of $p_{\mathrm{T}} > 20$ GeV and to lie within a pseudorapidity of $|\eta| < 2.5$.

Electrons

The medium selection cuts are employed to reconstruct electrons which provide a high efficiency (around 80 %) while retaining sufficient purity (jet background rejection roughly 2000) for isolated electrons in the SUSY analysis presented below (see pages 79 and 86f. of [1]). Electrons are required to have a minimum tranvserse momentum of $p_{\mathrm{T}} > 10$ GeV and a pseudorapidity of $|\eta| < 2.5$. As an additional isolation criterium the transverse energy in a cone of $\Delta R = 0.2$ around the electron, computed using the calorimetric information, is not to exceed 10 GeV. Jets overlapping with such an electron within a radius of $\Delta R < 0.2$ are discarded. If the two objects overlap within $0.2 < \Delta R < 0.4$, the electron is rejected as it mostly originates from a particle decay inside the jet. Finally, events are not considered with an electron reconstructed in the so-called *crack region* $1.37 < |\eta| < 1.52$, where the identification is degraded due to both the large amount of material in front and the crack between the barrel and extended barrel of the calorimeter [55].

Muons

The Staco algorithm is used to reconstruct combined muons (for details see pages 162ff. of [1]). To ensure a reasonable quality of combination, the χ^2 of the track matching is required to be less than 100. If several inner detector tracks match a spectrometer track, the one with the smallest ΔR is taken. Muons are required to have a minimum transverse momentum of $p_{\mathrm{T}} > 10$ GeV and a pseudorapidity of $|\eta| < 2.5$. As for electrons, the total calorimeter energy deposited in a cone of $\Delta R = 0.2$ around the muon must not exceed 10 GeV. Muons identified within $\Delta R < 0.4$ of a jet are rejected.

Missing transverse energy

The missing transverse energy is calculated from calorimeter cells (*cell-based method*), where the cells contribute with different calibration weights depending on the objects (jets, leptons and non-associated clusters) they are associated with (pages 368ff. of [1]). Sources of so-called *fake* missing energy (e.g. noisy or dead calorimeter cells, electronic problems, beam-gas and -halo events, cosmic rays) are not considered here.

3.4 Global variables

The approach of this analysis is to enhance the amount of selected potential signal events compared to Standard Model background events by cutting on variables defined in the plane transverse to the beam. Besides the transverse momenta of the reconstructed objects, the following variables are used throughout this part of the thesis.

Missing transverse momentum

The missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$, synonymously called *missing transverse energy*, is the norm of the vectorial sum of all particle momenta in the transverse plane. Neglecting any initial transverse momentum in the incoming protons, this quantity would be zero owing to conservation of momentum. In the experiment, it always has a finite value arising from undetectable particles and mismeasurements. It is expected to be large in R-parity conserving SUSY events due to the stable lightest SUSY particles, which leave the detector unnoticed. The longitudinal component of the missing momentum cannot be determined since the momenta of the partons prior to the collision are not known.

Effective mass

The effective mass, denoted as m_{eff}, is the scalar sum of transverse momenta of the main objects in the selected events and is defined as

$$m_{\mathrm{eff}} = \sum_{i=1}^{N_{\mathrm{jets}}} p_{\mathrm{T}}^{\mathrm{jet},i} + p_{\mathrm{T}}^{\mathrm{lep}} + E_{\mathrm{T}}^{\mathrm{miss}}\,. \tag{3.1}$$

N_{jets} is the jet multiplicity of the selection, e.g. in the 4-jet one-lepton channel the four hardest jet transverse momenta would be included, neglecting any further jets present in the event. $p_{\mathrm{T}}^{\mathrm{lep}}$ is the transverse momentum of the selected lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ the missing transverse energy. In SUSY events, the effective mass is related to the mass scale of the produced SUSY particles.

Transverse mass

The transverse mass is defined as the invariant mass of the lepton and the vector of the missing energy in the plane transverse to the beam. Assuming that the masses of both objects are negligible, it is calculated by

$$m_{\mathrm{T}} = \sqrt{2\, p_{\mathrm{T}}^{\mathrm{lep}}\, E_{\mathrm{T}}^{\mathrm{miss}}\,(1-\cos(\Delta\phi))}\,, \tag{3.2}$$

where $\Delta\phi$ is the angle between the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$ -vector in the transverse plane. In the case of a W decay, the transverse mass of the lepton and the neutrino cannot exceed the rest mass of the W (within detector resolution).

Transverse sphericity

Due to their high masses, SUSY particles are expected to be produced almost at rest compared to Standard Model particles. The ensuing cascade-decay results in a rather spherical distribution of the decay products, a feature which can be exploited to discriminate against Standard Model events. The transverse sphericity is defined as

$$S_{\mathrm{T}} = \frac{2\lambda_2}{(\lambda_1 + \lambda_2)}\,, \tag{3.3}$$

where λ_1 and λ_2 are the eigenvalues of the 2×2 transverse sphericity tensor

$$S_{ij} = \sum_k p_{ki} p_{kj} / \sum_k p_k^2 \tag{3.4}$$

built from the two-component momenta p_k of all selected jets and leptons in the transverse plane. Spherical distributions are characterized by transverse sphericity values close to unity.

3.5 Trigger

For an analysis investigating events with exactly one isolated lepton, a single lepton trigger is the obvious choice. Since this study has been carried out before the start of data-taking, such a trigger has been only simulated, yielding an efficiency of around 80% (see pages 1589f. of [1]) both for background and signal events. For simplicity, an efficiency of 100% is assumed in this analysis.

3.6 Significance of observation

5σ discovery The one-lepton analysis constitutes a simple counting experiment. The number of observed events is Poisson-distributed with a mean value which is the sum of the expected number of Standard Model background events and the expected number of events of a possible signal.

In order to claim the discovery of new physics, one usually tries to falsify the hypothesis that there is only Standard Model physics present. Thus, one calculates the probability p, also called *p-value*, of observing the N_{data} data events or even more given the expected number of Standard Model events N_b:

$$p = \sum_{i=N_{\mathrm{data}}}^{\infty} \frac{e^{-N_b} {N_b}^i}{i!} \tag{3.5}$$

High energy physics folklore considers a measurement to be a discovery, i.e. one can safely reject the *SM-only-hypothesis*, if the probability p turns out to be smaller than 2.9×10^{-7}, which corresponds to the integrated tail of a Gaussian distribution beyond five standard deviations (so-called 5σ *discovery*).

Incorporating background uncertainties In general, however, the expected number of events N_b is not known with perfect accuracy and the systematic uncertainty δN_b associated with it needs to be accounted for in the calculation. One way to do this is to consider it as a Bayesian prior. Since the uncertainty often arises from various sources, one can assume a Gaussian distribution of the prior as stated by the central limit theorem of probability theory.

The probablity p to observe N_{data} or more events given the expectation $N_b \pm \delta N_b$ is thus calculated by:

$$p = A \int_0^\infty db\ G(b; N_b, \delta N_b) \sum_{i=N_{\text{data}}}^{\infty} \frac{e^{-b} b^i}{i!}\,, \tag{3.6}$$

where $G(b; N_b, \delta N_b)$ is a Gaussian with mean N_b and standard deviation δN_b. The factor

$$A = \left[\int\limits_0^\infty db\ G(b; N_b, \delta N_b) \sum_{i=0}^{\infty} e^{-b} b^i / i! \right]^{-1}$$

ensures normalization to unity.

The probability p can be transformed into Gaussian *standard-deviations*, denoted in the following by the symbol Z_n, using the formula

$$Z_n = \sqrt{2}\ \mathrm{erf}^{-1}(1 - 2\mathrm{p}), \tag{3.7}$$

where erf(x) represents the Gauss error function.

Look-elsewhere-effect The look-elsewhere-effect, also called *multiple-comparison-problem*, denotes that fact that a determined probability does not correspond to the true frequency of occurence any longer if more than one selection is considered. This can easily lead to the misinterpretation of a statistical fluctuation as a real phenomenon. Thus, one has to correct the determined Z_n according to the true frequency. In the absence of an obvious analytic formula one can employ so-called *toy Monte Carlos*: One generates a sufficiently large number of random pseudo-measurements from the SM-only-hypothesis and establishes a relation between the determined probablity and the observed frequency.

3.7 4-jet selection

Given the expected cascade-decay of the produced SUSY particles (cf. subsection 1.3.4), which implies the radiation of additional quarks and gluons, many SUSY final states will exhibit a high jet multiplicity. Thus, a signature believed to be very sensitive for supersymmetry has been proposed in 1999 [84], which contains at least four jets with high transverse momentum.

Based on this proposal, the following event selection criteria constitute the one-lepton analysis in the 4-jet selection:

1. At least four jets with a $p_{\mathrm{T}} > 50$ GeV, with at least one of these having $p_{\mathrm{T}} > 100$ GeV.

2. Exactly one electron or one muon with $p_{\mathrm{T}} > 20$ GeV.

3. No secondary lepton with $p_{\mathrm{T}} > 10$ GeV.
4. Missing transverse energy: $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV as well as $E_{\mathrm{T}}^{\mathrm{miss}} > 0.2\ m_{\mathrm{eff}}$.
5. Transverse sphericity: $S_{\mathrm{T}} > 0.2$.
6. Transverse mass, formed from the lepton and the missing energy: $m_{\mathrm{T}} > 100$ GeV.
7. Effective mass: $m_{\mathrm{eff}} > 800$ GeV.

The required isolated lepton and a minimum of transverse missing energy aim to suppress the background coming from QCD multi-jet events. Cut 3 ensures an exclusive selection with respect to other analyses (see [1]), in particular with the di-lepton selection. The missing transverse energy, which is associated with the non-detectable LSP in the presence of signal, should have at least the same magnitude as the average jet transverse momentum which is realized by the second condition of cut 4. The latter condition is also intended to be effective against fluctuations of the missing energy, which grow with increasing calorimeter activity. The transverse sphericity cut is a requirement on the angular p_{T} distribution of the event in the transverse plane and is introduced to exploit the expected sphericity of SUSY events. Since its impact is rather modest (see below in table 3.3), it is dropped in later stages of the analysis (cf. chapters 4 and 5). Cut 6 is devised as a veto on events where the lepton and the missing energy originate from a W decay and is thus effective against top and $W+$ jets events. Finally, cut 7 selects high-mass final states and defines the signal region.

Table 3.3 shows the expected number of events for a luminosity of $1\,\mathrm{fb}^{-1}$ for the various cuts defined above. The number of QCD multi-jet events is reduced to zero for the available MC statistics after cut 4. $t\bar{t}$ and $W+$ jets constitute the dominant background contributions for all the remaining cuts. In the signal region $t\bar{t}$ represents about 90% of the total Standard Model background. The two plots in figure 3.1 show the effective mass distributions of the Standard Model backgrounds and the mSUGRA benchmark point SU3, which was defined in 3.2, after cuts 4 and 6 respectively. The signal stands out clearly from the dominant $t\bar{t}$ background, especially in the signal region ($m_{\mathrm{eff}} > 800$ GeV) on the right plot.

Table 3.3 also contains the cutflows for the other mSUGRA benchmark points of section 3.2. All of them yield a signal-to-background ratio bigger than unity in the signal region.

In order to properly calculate a significance of observation for these benchmark points, one needs to take into account systematical uncertainties on the expectated events. Data-driven methods, discussed in detail on pages 1525-1588 of [1], are expected to yield estimated errors of ±50% on the rate of QCD and ±20% on $t\bar{t}$, $W+$ jets events and on all the other backgrounds. Table 3.4 lists the significances for each SU point for different cuts on the effective mass calculated with formulas (3.6) and (3.7). Except for SU2 every one of these signals would trigger a discovery claim ($Z_n > 5$) when defining $m_{\mathrm{eff}} > 800$ GeV as the signal region.

Table 3.3: Number of expected events for an integrated luminosity of $1\,\mathrm{fb}^{-1}$ after successively applying the analysis cuts defined in the text.

Sample	Cuts 1–4	Cut 5	Cut 6	Cut 7
SU1	571.7	423.0	259.9	232.3
SU2	86.7	75.6	46.1	39.6
SU3	995.7	767.9	450.5	363.6
SU4	7523.6	6260.4	2974.4	895.8
SU6	342.3	250.9	161.9	147.9
SU8.1	296.4	214.4	151.4	136.3
$t\bar{t}$	2028.5	1546.8	131.7	36.0
W	425.2	314.8	9.9	5.4
Z	39.0	27.3	1.7	0.2
Di-boson	7.3	5.1	0.8	0.0
QCD	0.0	0.0	0.0	0.0
Standard Model BG	2500.1	1894.0	144.1	41.6

3.8 2- and 3-jet selection

If the masses of the produced SUSY particles are comparatively small or if predominantly squark pairs are produced, the associated events can often display a low jet multiplicity. In this section, two analyses selecting events with an isolated lepton and 2 or 3 jets respectively are introduced, which enhance the overall sensitivity for supersymmetry discovery in the one-lepton channel. The cuts concerning the leptons, transverse sphericity and the transverse mass remain unchanged. To compensate the increase of Standard Model background due to the removal of the fourth jet requirement, the minimal transverse momenta are raised to 150 GeV for the hardest and 100 GeV for the other one or two jets. Furthermore, the cut on the missing energy needs to be adjusted to $E_{\mathrm{T}}^{\mathrm{miss}} > \max(100\text{ GeV}, 0.3\,m_{\mathrm{eff},\,2j})$ and $E_{\mathrm{T}}^{\mathrm{miss}} > \max(100\text{ GeV}, 0.25\,m_{\mathrm{eff},\,3j})$ for the 2- and 3-jet selection respectively, since the number of quantities dominantly contributing to the effective mass (jets and missing energy) decreases from 5 to 3 and 4, respectively.

Figure 3.2 shows the effective mass distributions for events with one lepton and three (on the left) or two (on the right) jets after all cuts. The SU3 supersymmetry signal stands out clearly against the background, which arises almost completely from top pair production. It has to be noted though that the total lack of background in some bins is mainly due to insufficient statistics. PYTHIA, which offers a less reliable description of additional hard jets

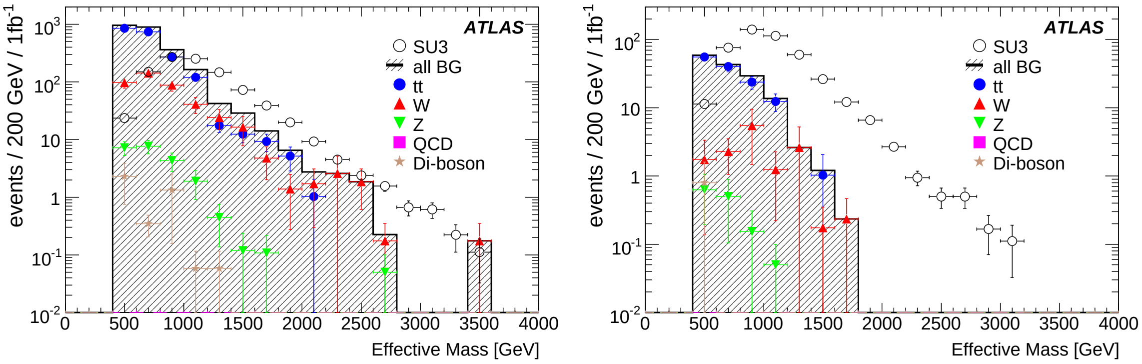


Figure 3.1: Effective mass distributions after cut 4 (left) and cut 6 (right) defined in the text for a luminosity of 1 fb^{-1}. Event numbers for the different processes can be found in table 3.3.

(cf. section 1.3), has been employed to model the $W/Z+$jets background since the ALPGEN events of the 4-jet analysis above pass a 4-jet filter on generator level (see section 3.2).

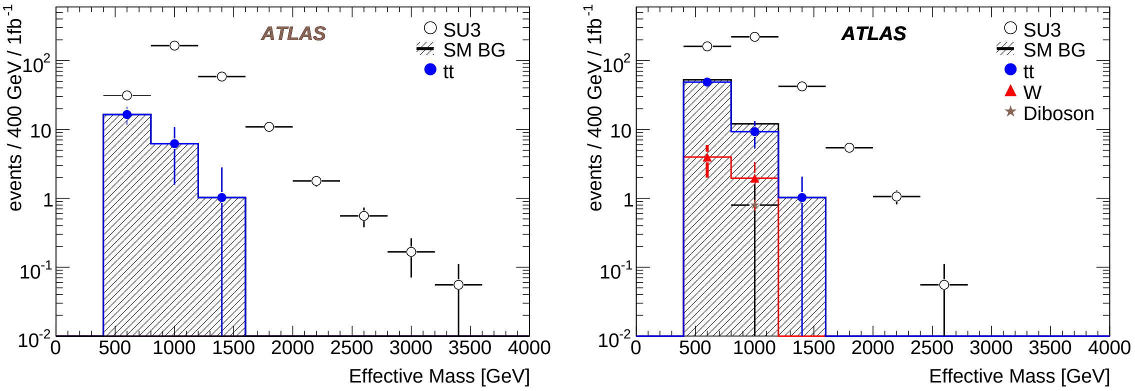


Figure 3.2: Effective mass distributions for events with one lepton and 3 (left) or 2 (right) jets after all cuts as defined in the text.

3.9 Discovery reach

In order to estimate the discovery reach of the one-lepton analysis in terms of squark and gluino masses, two grids in the $m_0 - m_{1/2}$ plane of the mSUGRA scenario are considered, defined by constant values of $A_0 = 0$, $\mu > 0$ and $\tan\beta = 10$ or $\tan\beta = 50$, respectively. Details on the production of the grids are described in section 3.2.

For each point of the grids, the significance of observation is calculated, using formulas (3.6) and (3.7), as a function of the cut on the effective mass, which is increased in steps of 400 GeV. The highest Z_n value obtained is then corrected for this *multiple comparison* procedure (see section 3.6) and the corresponding grid point considered to be within the

Table 3.4: Significance of observation Z_n for the different SUSY benchmark points for an integrated luminosity of $1\,\text{fb}^{-1}$ including the systematic uncertainty on the background prediction.

Sample	$m_{\text{eff}} > 400$ GeV		$m_{\text{eff}} > 800$ GeV		$m_{\text{eff}} > 1200$ GeV	
	Events	Z_n	Events	Z_n	Events	Z_n
Standard Model BG	144		42		2	
SU1	260	7.6	232	12.3	114	18.0
SU2	46	1.5	40	3.4	15	6.0
SU3	450	9.5	364	16.7	110	17.7
SU4	2974	33.7	896	29.4	99	16.6
SU6	162	4.9	148	8.9	76	14.2
SU8.1	151	4.6	136	8.4	66	13.1

5σ discovery reach if $Z_{n,corr} > 5$. The same systematic uncertainties as in section 3.7 are applied: ±50% on QCD multi-jet events and ±20% on all other backgrounds.

Figure 3.3 shows the 5σ reach in the one-lepton channel for different jet multiplicity selections in the $\tan\beta = 10$ scenario, again for an integrated luminosity of $1\,\text{fb}^{-1}$. The three analyses have a comparable performance with slight advantages for the 2- and 3-jet selections in the low m_0 regime, where the pair production of squarks is dominant and therefore in general fewer jets are present in the events. The shaded areas in the upper left and bottom right corner cover excluded regions, where either the electrically charged stau is the LSP or no electroweak symmetry breaking occurs. The dashed green line borders the region where the chargino has a mass smaller than 103 GeV, which has been excluded by direct searches at LEP [85]. The figure shows that squark and gluinos with masses up to around 1 TeV can be excluded with a luminosity of $1\,\text{fb}^{-1}$, as indicated by the dashed grey lines.

In the left plot of figure 3.4 the reach of the 4-jet one-lepton analysis is compared to the one achieved by analyses requiring different lepton multiplicities. The 4-jet zero-, one- and two-lepton analyses apply identical cuts on the jet transverse momenta and on $E_{\text{T}}^{\text{miss}}$ in addition to requiring an isolated lepton veto (for leptons with $p_{\text{T}} > 10$ GeV), exactly one isolated lepton and exactly one oppositely charged isolated lepton pair (lepton $p_{\text{T}} > 10$ GeV each), respectively. In addition, the zero-lepton analysis requires the ϕ-angle between the direction of the transverse missing energy and each of the three hardest selected jets to be larger than 0.2. This cut rejects events where the missing energy arises mainly from an underestimation of the energy of a reconstructed jet. It is not applied on the fourth jet in order to not further reduce the remaining phase space. The 1-jet three-lepton analysis requires

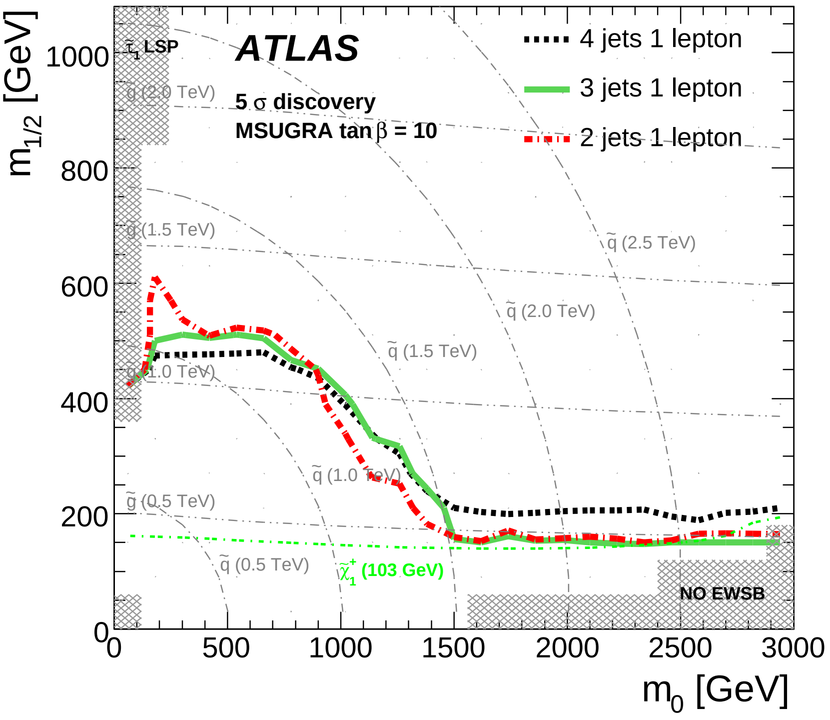


Figure 3.3: The expected 5σ discovery reach contours for an integrated luminosity of $1\,\text{fb}^{-1}$ for the one-lepton analysis with various jet requirements as a function of m_0 and $m_{1/2}$ in the $\tan\beta = 10$ mSUGRA plane. The horizontal and curved grey lines indicate the gluino and squark mass contours, respectively, in steps of 500 GeV. The shaded areas are excluded as described in the text.

at least one jet with $p_T > 200$ GeV and at least three isolated leptons with $p_T > 10$ GeV, but does not impose any requirements on E_T^{miss}. The zero-lepton analysis outperforms the one-lepton selection especially at low m_0. The other two show a substantially worse potential reach.

The right plot of figure 3.4 displays the reach in the $\tan\beta = 50$ plane where enhanced τ production is expected. Also in this scenario the zero-lepton channel seems the most promising. A dedicated analysis requiring at least one τ instead of the isolated lepton does not prove to be more effective, which results from the lower efficiency and purity of the τ identification.

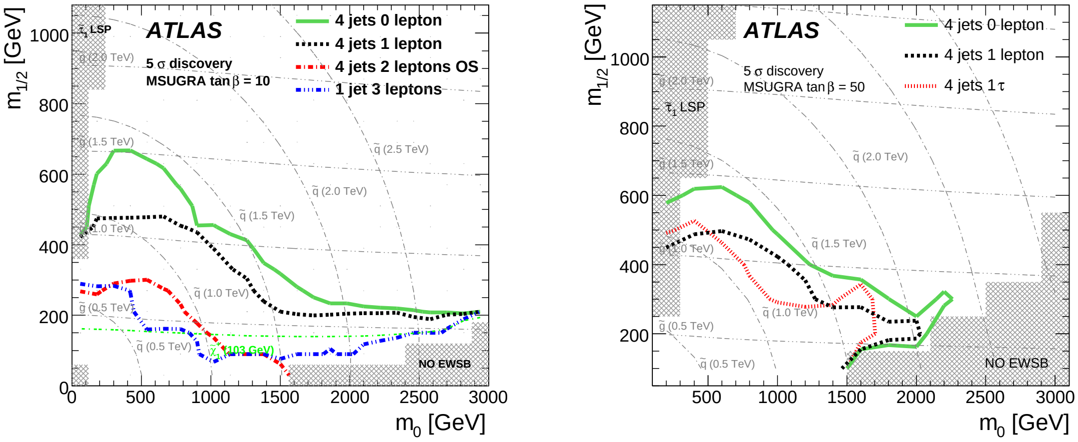


Figure 3.4: The expected 5σ discovery reach contours for an integrated luminosity of $1\,\mathrm{fb}^{-1}$ for the 4-jet selection with different lepton requirements as a function of m_0 and $m_{1/2}$ in the mSUGRA scenarios with $\tan\beta = 10$ (left) and $\tan\beta = 50$ (right). The horizontal and curved grey lines indicate gluino and squark mass contours, respectively, in steps of 500 GeV. The shaded areas are excluded as described in the text.

3.10 Conclusion

The search for supersymmetry in one-lepton final states has tantalizing prospects. With an integrated luminosity of $1\mathrm{fb}^{-1}$, supersymmetric particles with masses up to around 1 TeV can be discovered. While the discovery reach of the one-lepton channel in the considered mSUGRA models is surpassed by that of the zero-lepton selection, the QCD multi-jet background, whose reliable determination is one of biggest challenges at a hadron collider, has a much smaller impact. Even though the LHC is currently only operated at half of its design collision energy, the presented studies keep their relevance for future collisions at ultimately 14 TeV. The discovery reach in other SUSY-breaking models, such as the *non-universal-Higgs model* (NUHM) or the *anomaly-mediated SUSY breaking model* (AMSB), are briefly discussed in [1] with the one-lepton analysis performing not as well as other search analyses.

Chapter 4

Search for supersymmetry at $\sqrt{\mathrm{s}} = 7$ TeV with early data

4.1 Introduction

Based on the analysis introduced in the previous chapter, this chapter discusses the search for supersymmetry in the one-lepton channel using early data, recorded between March and July 2010 with the ATLAS detector. The results of the analysis have been presented at the *35th International Conference on High Energy Physics* in Paris end of July 2010 [2, 3].

Following the incident on Sep. 19th 2008 (see e.g. [58]), which significantly delayed the start-up of the accelerator, the LHC management decided to set the beam energy to only half of its design intensity during the first period of data taking, resulting in proton-proton collisions with a centre-of-mass-energy of 7 TeV. In the following section, the prospects for supersymmetry discovery at 7 TeV are briefly discussed. The rest of the chapter deals with the data analysis, where the author has significantly contributed to the work on the event selection and the background estimation as well as the preparation of the plots.

4.2 Prospective discovery reach at $\sqrt{\mathrm{s}} = 7$ TeV

In light of the reduced beam energy during the first period of data taking, the study presented in the previous chapter has been updated to estimate the supersymmetry discovery reach at $\sqrt{s} = 7$ TeV [86].

In the analyses considered therein, the minimum jet p_{T} requirements have been reduced from 50 to 40 GeV for all but the leading jets. The minimum leading jet p_{T} has been raised to 180 GeV in the 2-jet selection. $E_{\mathrm{T}}^{\mathrm{miss}}$ has been required to exceed 80 instead of 100 GeV in all selections.

In addition, the ϕ-angle cut (> 0.2) between $E_{\mathrm{T}}^{\mathrm{miss}}$ and each of the selected jets (up to the third), which has already been used in the zero-lepton analysis in section 3.9, has been applied in all analyses.

Figure 4.1 shows the expected discovery reach for these analyes for an integrated luminosity of $1\,\text{fb}^{-1}$ in the $\tan\beta = 10$, $A_0 = 0$, $\mu > 0$ mSUGRA scenario (cf. section 1.2) in the $m_0 - m_{1/2}$ plane and as a function of squark and gluino masses. As before, the zero-lepton analysis exhibits the highest sensitivity. The one-lepton selection has a comparable reach over large parts of this phase space. If the gluino and squark masses are of the same magnitude, signals with particle masses up to 700 GeV can be discovered with an integrated luminosity of 1 fb^{-1}.

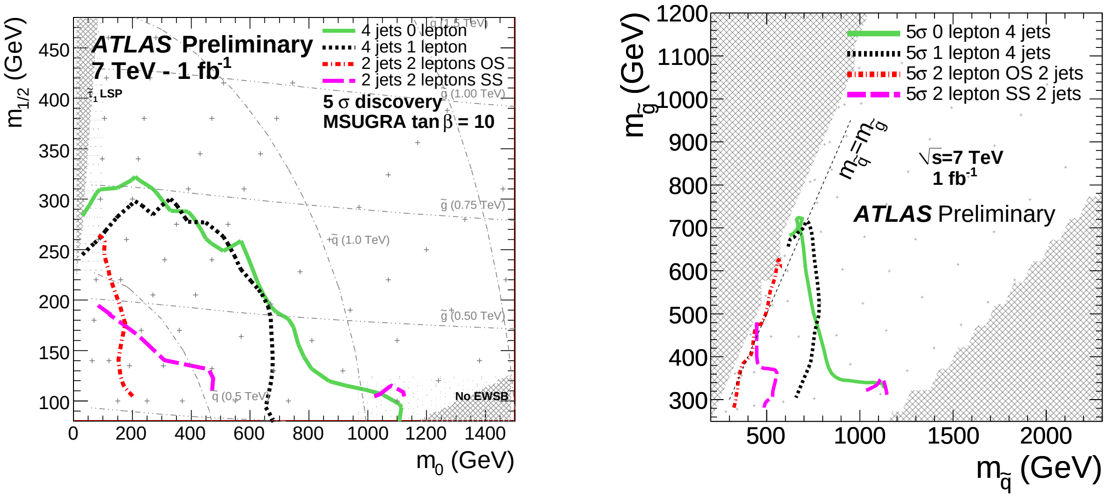


Figure 4.1: The 5σ discovery reach contours for various analyses as a function of m_0 and $m_{1/2}$ (left) as well as squark and gluino masses (right) for the $\tan\beta = 10$, $A_0 = 0$, $\mu > 0$ mSUGRA scenario for an integrated luminosity of $1\,\text{fb}^{-1}$ at the reduced centre-of-mass-energy of 7 TeV [86].

4.3 Signal and background simulation

The same generators as in the analysis in the previous chapter, which is in the following referred to as *prospects-analysis*, have been used for the different signal and background samples (cf. section 3.2). PYTHIA [35], HERWIG [42] and JIMMY [43] have been tuned with CDF and D0 data [87] for an improved underlying event and minimum bias description. The full detector simulation, provided by GEANT4 [51], has been applied on all samples followed by a reconstruction identical to that used for data. The produced samples, their cross-sections and their sizes in terms of integrated luminosity are shown in table 4.1.

Table 4.1: Cross-sections (σ) times branching ratios (BR) and integrated luminosities (L) for the SM backgrounds and SU4. The cross sections are given at NNLO for $W \to l\nu$ and $Z \to \nu\bar{\nu}$, at *approximate* NNLO [88, 89, 90] for $t\bar{t}$ and at LO for the QCD multi-jet events and $Z \to l^+l^-$. $\hat{p}_\mathrm{T}$ is the p_T of the two partons involved in the hard scattering.

Process	σ × BR [nb]	L [nb^{-1}]
QCD $8 \leq \hat{p}_\mathrm{T} < 17$ GeV	9.85×10^6	0.14
QCD $17 \leq \hat{p}_\mathrm{T} < 35$ GeV	6.78×10^5	2.06
QCD $35 \leq \hat{p}_\mathrm{T} < 70$ GeV	4.10×10^4	34.1
QCD $70 \leq \hat{p}_\mathrm{T} < 140$ GeV	2.20×10^3	636
QCD $140 \leq \hat{p}_\mathrm{T} < 280$ GeV	88	1.59×10^4
QCD $280 \leq \hat{p}_\mathrm{T} < 560$ GeV	2.35	5.96×10^5
QCD 560 GeV $\leq \hat{p}_\mathrm{T}$	0.034	4.12×10^7
$W \to e\nu$	10.45	2.0×10^5
$W \to \mu\nu$	10.45	2.0×10^5
$W \to \tau\nu$	10.45	2.0×10^5
$Z \to \nu\bar{\nu}$	5.82	1.0×10^5
$Z \to e^+e^-$	0.79	5.0×10^5
$Z \to \mu^+\mu^-$	0.79	5.0×10^5
$Z \to \tau^+\tau^-$	0.79	5.0×10^5
$t\bar{t}$	0.164	7.0×10^6
SU4	0.060	11.9×10^5

Signal sample

In this analysis, the mSUGRA submodel SU4 ($m_0 = 200$ GeV, $m_{1/2} = 160$ GeV, $A_0 = -400$ GeV, $\tan\beta = 10$, $\mu > 0$, as in section 3.2) is taken as a benchmark signal due to its low predicted squark and gluino masses (cf. table A.1) and thus relatively high cross-section. A sample of 50,000 events have been produced with HERWIG++ [44] version 2.4.2 using the particle spectrum and branching ratios predicted by ISAJET [47] version 7.75 and the MRST2007LO* parton density distributions [91]. The cross-section at $\sqrt{s} = 7$ TeV amounts to 42.3 pb at leading order and 59.95 pb at next-to-leading order as calculated by HERWIG++ and by PROSPINO [48] respectively.

Background samples

MC@NLO [37, 38, 39] version 3.41 has been employed to model top pair production together with HERWIG, JIMMY and the next-to-leading order CTEQ6.6 PDF set [92] for parton showering, fragmentation and the simulation of the underlying event. $W/Z+$ jets events have been produced using ALPGEN [36] version 2.13, HERWIG and JIMMY together with CTEQ6L1 [82]. Their next-to-next-to-leading order (NNLO) cross-sections shown in table 4.1 have been calculated by using the FEWZ program [46]. QCD multi-jet events have been generated with PYTHIA 6.4.21 using the MRST2007LO* set for the parton density functions. For the muon channel the events have been filtered requiring at least one muon with $p_{\mathrm{T}} >$ 10 GeV and $|\eta| < 2.8$ prior to applying the time-consuming detector simulation.

4.4 Data selection

The total integrated luminosity of the data used in the following analysis amounts to $(70 \pm 8)\,\mathrm{nb}^{-1}$. Single lepton triggers are employed to preselect interesting events.

In the electron channel, the hardware-based trigger `L1_EM5` is used together with the high level trigger `EF_g10_loose`, which flags events having deposits in the electromagnetic calorimeter bigger than 10 GeV. This combination displays full efficiency, i.e. $(100^{+0}_{-10})\%$, for data events containing one isolated electron with a $p_{\mathrm{T}} > 20$ GeV. The efficiency modeled in the Monte Carlo simulation is somewhat lower with $(96 \pm 3)\%$.

For the muon selection, the hardware-based trigger `L1_MU6` is chosen, which is tuned to muon chamber hit patterns caused by 6 GeV muons. Depending on the pseudorapidity, its efficiency is measured to be $(73 \pm 5)\%$ for $|\eta| < 1.05$ and $(82 \pm 4)\%$ for $1.05 < |\eta| < 2.4$ for events containing one isolated muon. The corresponding values in Monte Carlo turn out to be higher with $(80 \pm 3)\%$ for $|\eta| < 1.05$ and $(93 \pm 3)\%$ for $1.05 < |\eta| < 2.4$. In both cases, the efficiencies stay constant for muons with $p_{\mathrm{T}} > 8$ GeV. No reweighting of the Monte Carlo is performed since the deviation is small compared to other systematic uncertainties considered (cf. subsection 4.7.3).

Details on the trigger efficiency determination can be found in [93].

4.5 Object definition

The object selection used for this analysis follows closely that presented in section 3.3 with the exception of the jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithms.

Jets

Jets are reconstructed using the anti-k_t algorithm [75] with a radius parameter of $R = 0.4$. The transverse momenta, measured at the electromagnetic scale, are corrected for dead material and the non-compensating nature of the calorimeter using a simulation-based calibration [94]. Only jets with a corrected $p_{\mathrm{T}} > 20$ GeV at the hadronic scale and with $|\eta| < 2.5$ are included. Cleaning cuts are applied to reject events which contain jets probably arising from cosmic rays, calorimeter noise or out-of-time energy deposits [95].

Electrons

Electrons are reconstructed using an improved medium selection, as defined e.g. in [77]. They are required to have a minimum tranvserse momentum of $p_{\mathrm{T}} > 10$ GeV and to lie within $|\eta_{calo}| < 2.47$, where η_{calo} is the pseudorapidity measured in the second sampling layer of the electromagnetic calorimeter. Events are discarded, if a selected electron lies in the crack region defined as $1.37 < |\eta_{calo}| < 1.52$. The expected efficiency is increased to about 90% with a jet rejection of roughly 7,000 for isolated electrons [77]. The same isolation and jet overlap removal conditions as in section 3.3 apply.

Muons

The same definitions as in section 3.3 apply, except that the considered pseudorapidity range is reduced to $|\eta| < 2.4$, beyond which muons cannot be triggered. As above, muons within $\Delta R < 0.4$ of a jet are removed. Events with a muon, whose extrapolated track has a distance of more than 1 cm from the primary vertex in the z-direction, are rejected as those most probably originate from cosmic rays.

Missing transverse energy

The missing transverse energy is calculated from the topological clusters of the calorimeters at the electromagnetic scale, without correcting them for the nature of the objects they are associated with or for losses due to dead material. Cells up to an $|\eta|$-range of 4.5 are included. To account for the contribution coming from well-isolated muons, their vectors are subtracted from the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector.

4.6 Event selection

Due to the smaller centre-of-mass energy of 7 TeV and, more importantly, the low recorded integrated luminosity of $70\,\mathrm{nb}^{-1}$, only the 2-jet one-lepton selection with looser cuts compared to the prospects-analysis is considered. Contrary to the selection of chapter 3, two separate analyses for the electron and the muon channel are carried out. The following cuts are applied:

1. Lepton trigger (see section 4.4).
2. Primary vertex with more than four associated tracks.
3. Exactly one lepton with $p_\mathrm{T} > 20$ GeV.
4. No secondary lepton with $p_\mathrm{T} > 10$ GeV.
5. At least two jets with $p_\mathrm{T} > 30$ GeV.
6. Missing transverse energy: $E_\mathrm{T}^\mathrm{miss} > 30$ GeV.
7. Transverse mass, formed from the lepton and the missing energy: $m_\mathrm{T} > 100$ GeV.

The primary vertex requirement is introduced to reject events potentially unrelated to proton-proton collisions. Again, the second lepton veto is imposed to avoid overlap with the di-lepton analysis. Neither the transverse sphericity nor the effective mass cuts are applied.

4.7 Background estimation

Given the low statistics, a comprehensive background estimation using data-driven techniques is not feasible. However, the available data can be used to estimate the overall rate of the two most significant backgrounds in this analysis, that is QCD multi-jet and $W+$ jets production.

Since the QCD multi-jet background is only modeled at leading order with PYTHIA, its total cross-section is not expected to be well predicted. In addition, a slight mismodeling of the fake lepton rate in the Monte Carlo simulation can result in a significant change of this background due to its large total rate. It is therefore essential to constrain this background.

$W+$ jets events represent the dominant background contribution for the set of cuts defined in section 4.6. ALPGEN takes part of the higher order corrections, which are the source for additional jet radiation, into account. Still, the overall rate for events with high jet multiplicities might be substantially off (see e.g. [96]) and ought to be verified using data.

Suitable control regions for both processes are defined in the $E_\mathrm{T}^\mathrm{miss} - m_\mathrm{T}$ plane after the 2-jet cut. Figure 4.2 shows the distribution of data, QCD multi-jet and $W+$ jets events for both the electron and muon channel. The electron selection is dominated by QCD, which is for both lepton flavours concentrated in the low $E_\mathrm{T}^\mathrm{miss}$ region. The $W+$ jets events are predicted to be most prominent in the medium $E_{\mathrm{T}}^{\mathrm{miss}} - m_{\mathrm{T}}$ region. This behavior is expected, since in the case of $W+$ jets events the transverse mass corresponds to the invariant mass of the W in the transverse plane and the missing energy originates from the non-detectable neutrino, whose momentum distribution peaks around 40 GeV. Contrary to that, the missing energy in QCD events arises mostly from mismeasurements and the transverse mass is thus in general not associated with a particular object.

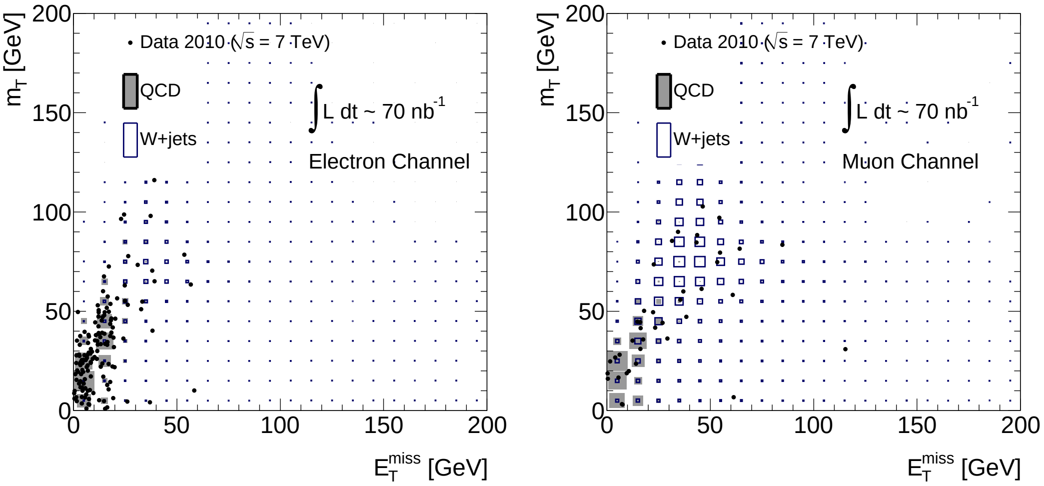


Figure 4.2: $E_{\mathrm{T}}^{\mathrm{miss}}$ versus m_{T} plane for background normalization of QCD multi-jet and $W+$ jets events in the electron (left) and muon (right) channel. The size of the boxes reflects the relative magnitude of each contribution.

4.7.1 QCD normalization

Following the event distribution in figure 4.2, the QCD control region is defined as $E_{\mathrm{T}}^{\mathrm{miss}} <$ 40 GeV and $m_{\mathrm{T}} <$ 40 GeV. The fraction of background events other than QCD is below the 5%-level in this selection. In a conservative approximation, all its data are considered to originate from QCD multi-jet events, which results in a slightly overestimated normalization factor. Table 4.2 lists the measured data events and the associated QCD multi-jet expectation for both lepton channels. Separate normalization factors for the electron and muon channel are determined, in order to account for the different causes, and their modeling in the simulation, of QCD events to contain reconstructed isolated leptons. The last column of table 4.2 shows that PYTHIA overestimates the true rate by more than a factor of two in both cases.

Table 4.2: Measured data and expected QCD multi-jet events in the QCD region defined by $E_{\mathrm{T}}^{\mathrm{miss}} < 40$ GeV and $m_{\mathrm{T}} < 40$ GeV. QCD is overestimated by more than a factor of two. All errors are statistical.

	Data	QCD expectation	Normalization factor
Electron channel	101	245 ± 38	0.41 ± 0.08
Muon channel	15	31.4 ± 0.9	0.48 ± 0.12

4.7.2 $W+$ jets normalization

For the $W+$ jets normalization, a medium region defined by 30 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 50 GeV and 40 GeV $< m_{\mathrm{T}} <$ 80 GeV is chosen. The number of data events as well as the predicted W and other Standard Model background are shown in table 4.3. The purity of this selection is worse than in the QCD case above. The other Standard Model contributions consist of the rescaled QCD (cf. previous subsection) and $t\bar{t}$, which are subtracted from the data before calculating a normalization factor. Combining both channels, a scaling factor is derived and found to be 2.1 ± 1.0. The stated error includes an additional systematic uncertainty of 30% to account for the extrapolation of the factor to the rest of the phase space, but is still dominated by the statistical uncertainty of approximately 0.7.

Table 4.3: Measured data and expected Standard Model events in the W region defined by 30 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 50 GeV and 40 GeV $< m_{\mathrm{T}} <$ 80 GeV. QCD is rescaled with the normalization factors of table 4.2. All errors are statistical. Combining both channels, a normalization factor of 2.1 ± 1.0 is obtained, including a 30% extrapolation uncertainty as described in the text.

	Data	W expectation	Other SM expectation
Electron channel	6	2.20 ± 0.02	0.8 ± 0.3
Muon channel	4	2.13 ± 0.02	0.3 ± 0.1

4.7.3 Systematic uncertainties

The following sources of systematic uncertainties on the background prediction are included in the analysis, reflecting the shortcomings in the modeling of the physics processes and an incomplete detector simulation. The treatment is not exhaustive, but believed to include the major contributions.

One of the most significant contributions concerns the correct determination of the jet energy scale. Its uncertainty is estimated as a function of jet transverse momentum and pseudorapidity, treating all reconstructed jets in a fully correlated way and propagating the effect to the transverse missing momentum accordingly. The relative uncertainty on the jet transverse momentum amounts to about 10% within the range 20 GeV $< p_{\mathrm{T}} <$ 60 GeV and 7% for values above that [94].

In kinematic regions, for which QCD ALPGEN predictions were available, they have been compared to the PYTHIA estimation and good agreement has been found. Nevertheless, a 50% error, in addition to the uncertainty of the normalization factor, is assigned to the QCD background to cover remaining uncertainties on the shapes of the distributions of the investigated variables.

As discussed above, a 30% extrapolation uncertainty is included in the error of the W normalization factor resulting in a total uncertainty of 50% on the $W+$ jets production rate.

An uncertainty of 60% is applied on the ALPGEN-predicted cross-section of $Z+$ jets events motivated by the underestimation of the simulation demonstrated in [96].

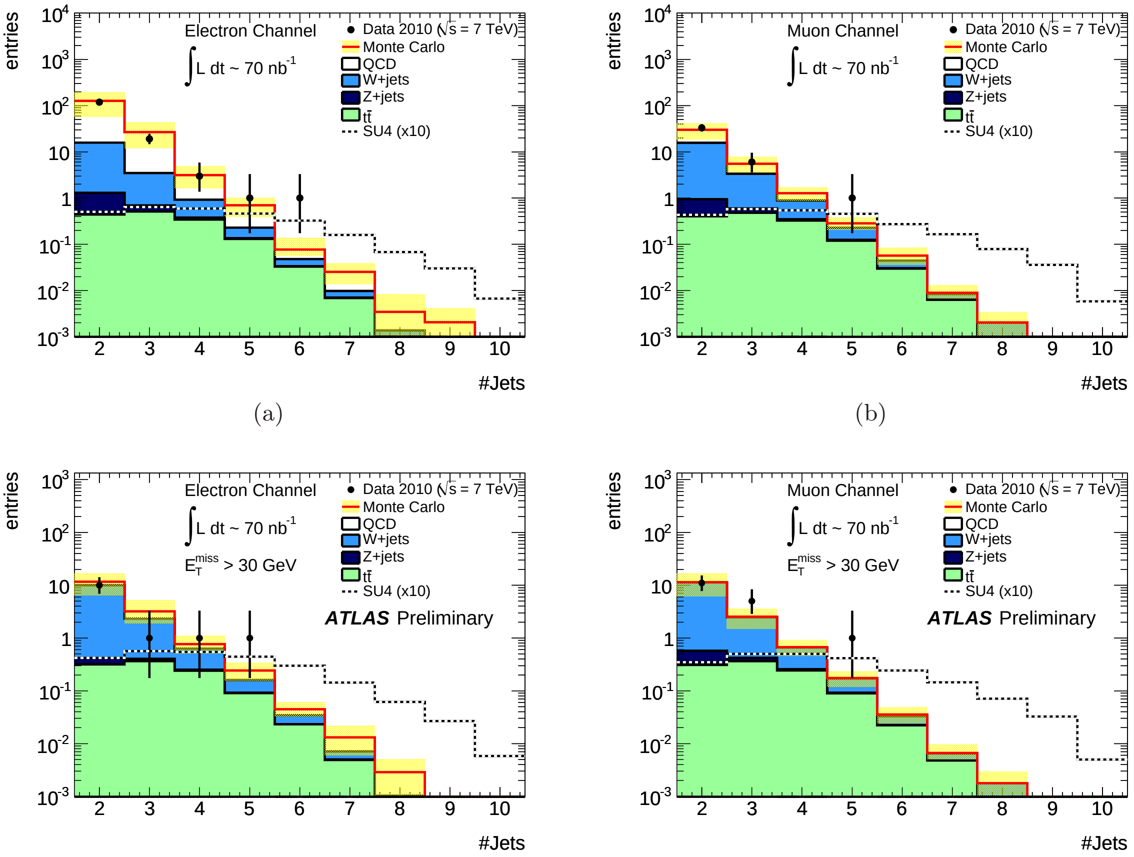
An 11% luminosity uncertainty [97] is applied, in a fully correlated way, on the $Z+$ jets and $t\bar{t}$ backgrounds which are normalized to their respective cross-sections. Remaining uncertainties on the trigger efficiency and the lepton identification efficiency are small in comparison and therefore neglected in this analysis. All systematic errors are added in quadrature including the statistical uncertainty of the Monte Carlo prediction.

4.8 Results

The four plots in figure 4.3 show the jet multiplicity distributions of the events passing cuts 1-5 defined in section 4.6, i.e. they contain at least two jets and an isolated electron or muon. The $W+$ jets and QCD multi-jet backgrounds are scaled to data as described in the previous section. The yellow band represents the total error on the Standard Model background including all the systematic uncertainties discussed in subsection 4.7.3. The errors on the data show the 68% Poisson confidence intervals to indicate the statistical uncertainty in each bin. Good agreement between data and simulation is found both before and after applying the $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV requirement (cut 6), which substantially reduces the QCD background. The benchmark signal SU4 is enhanced by a factor of 10 for illustrative purposes and exhibits a higher multiplicity on the average.

Figures 4.4 and 4.5 show the distributions of the same events, without applying the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut, as a function of the transverse missing energy and the effective mass, respectively. At this selection stage both channels are clearly dominated by QCD, which is concentrated in the low $E_{\mathrm{T}}^{\mathrm{miss}}$-region (cf. figure 4.2). In the range where data are measured, the predicted shapes are reasonably well reproduced.

When applying the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut, $W+$ jets becomes the largest background contribution as displayed in figure 4.6. $t\bar{t}$ and $Z+$ jets remain at a negligible level with their contribution



(a) (b) (c) (d)

Figure 4.3: Jet multiplicities before ((a) and (b)) and after ((c) and (d)) applying a cut on the transverse missing energy $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV, for the electron (left) and muon (right) channel. The SU4 expectation is enhanced by a factor of 10 for illustrative purposes. All distributions show good agreement between data and Monte Carlo prediction.

more than one order of magnitude smaller than $W +$ jets. The transverse mass distribution before and after applying the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut are shown in figure 4.7. The $W +$ jets and $t\bar{t}$ backgrounds peak around the W-mass in both channels. In order to suppress, in particular, the former, the transverse mass is required to be larger than 100 GeV (cut 7). The remaining events are shown in figure 4.8 and constitute the population in the signal region. Two events in the electron and one event in the muon channel pass all selection cuts and are in agreement with the Standard Model predictions of 3.6 ± 1.6 and 2.8 ± 1.2 events respectively, with all systematic uncertainties of subsection 4.7.3 included in the stated errors. No excess hinting at hitherto unknown signals is observed. Figure 4.9 shows the transverse missing energy distributions after the m_{T} cut. The three signal region data events are measured in the region of highest probability as predicted by the Standard Model. Table 4.4 summarizes the number of observed data events and the corresponding background predictions for cuts 5 to 7.

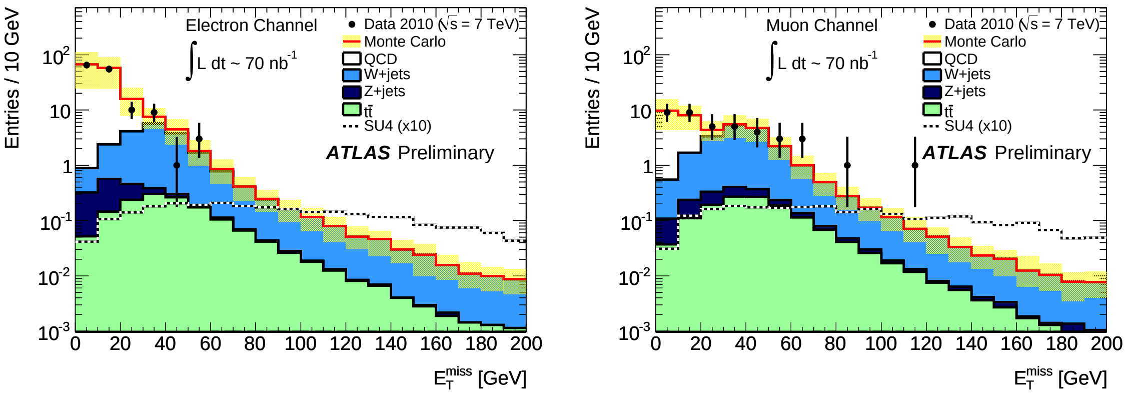


Figure 4.4: Missing transverse energy distributions for the electron (left) and muon (right) channel without a cut on the transverse mass (cf. cut 5 in table 4.4).

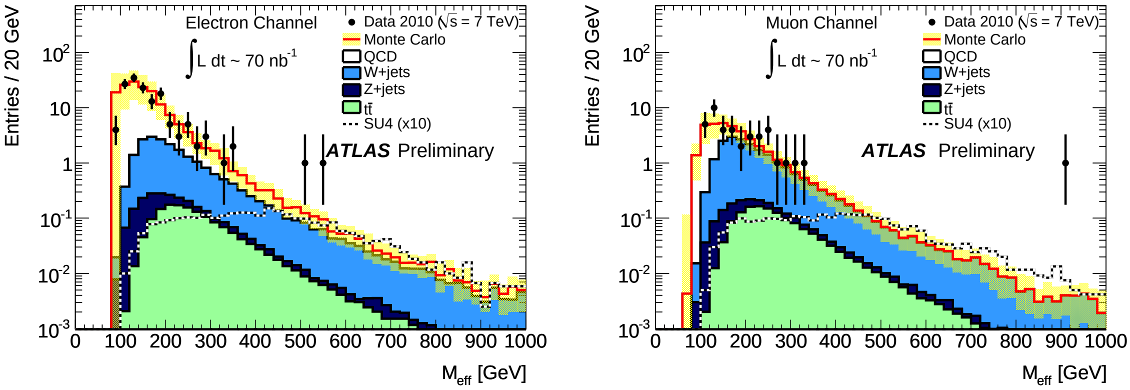


Figure 4.5: Effective mass distributions for the electron (left) and muon (right) channel without cuts on the missing transverse energy or transverse mass (cf. cut 5 in table 4.4). The data are in good agreement with the Standard Model predictions.

4.9 Conclusion

The analysis of the first $(70 \pm 8)\,\mathrm{nb}^{-1}$ of ATLAS data does not show any significant deviation from the Standard Model prediction in final states with one isolated lepton and at least two well-defined jets. The measurements indicate that the backgrounds for consecutive searches in the one-lepton channel are under control.

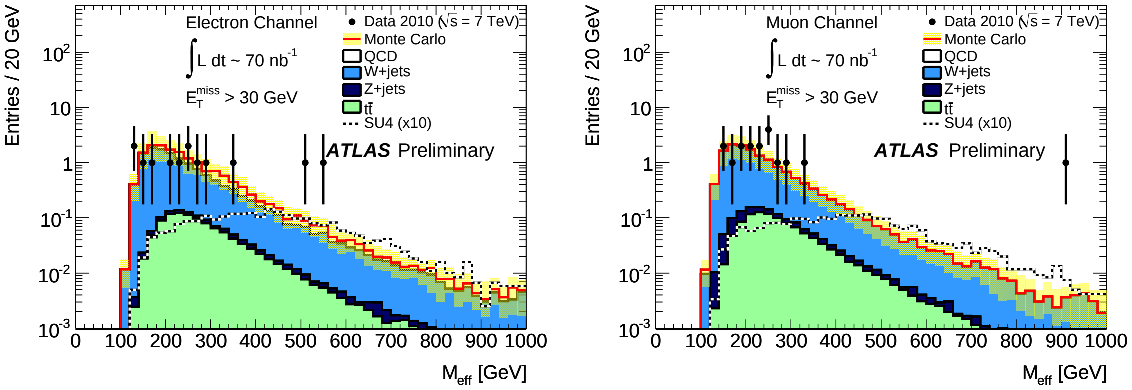
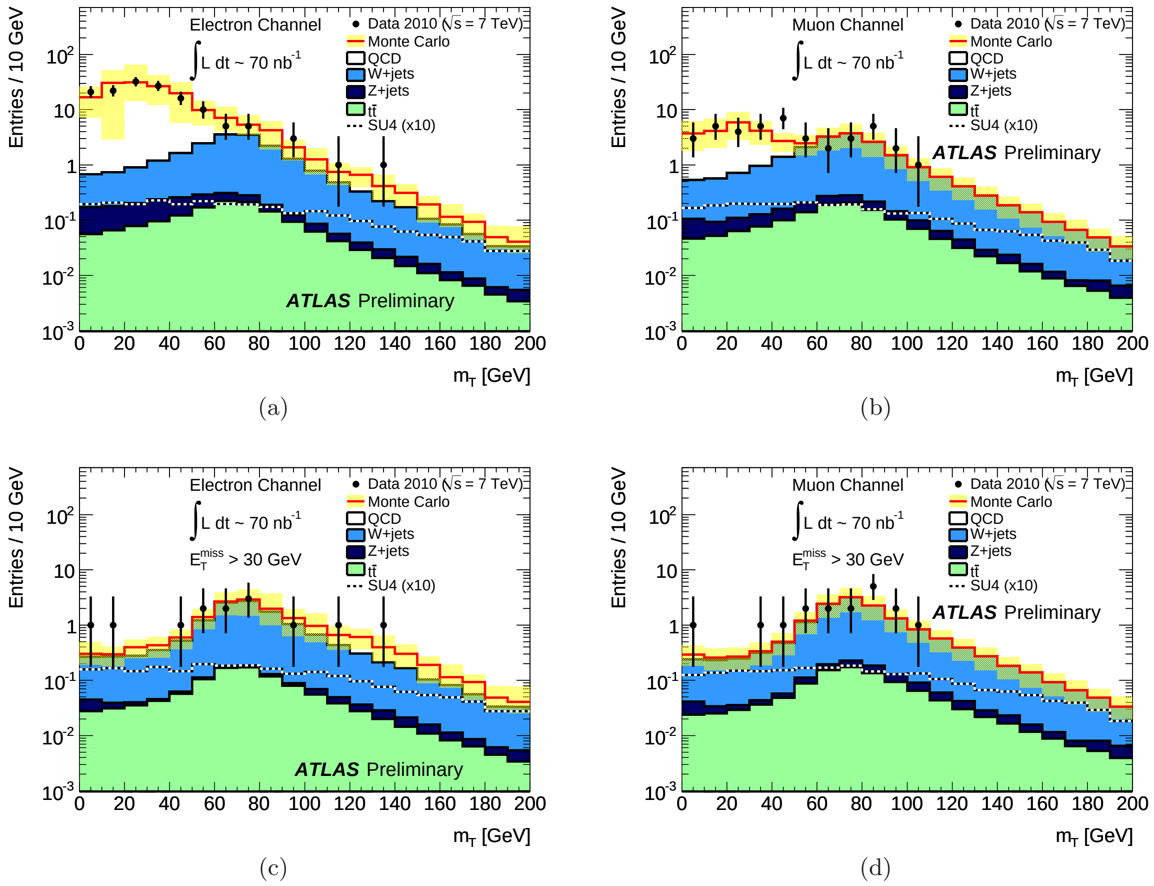


Figure 4.6: Effective mass distributions for the electron (left) and muon (right) channel with a cut on the transverse missing energy $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV (cf. cut 6 in table 4.4). The QCD background has been substantially reduced compared to figure 4.5.

Table 4.4: Observed and predicted number of events after cuts 5, 6 and 7 as defined in section 4.6. Good agreement with the Standard Model prediction and no indication for new physics are found. The stated errors include all systematic uncertainties discussed in subsection 4.7.3.

Cuts applied up to	**Electron channel**		**Muon channel**	
(as defined in 4.6)	Data	Expectation	Data	Expectation
Cut 5: 2 jets, 1 lepton	143	157 ± 85	40	37 ± 14
Cut 6: $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV	13	16 ± 7	17	15 ± 7
Cut 7: $m_{\mathrm{T}} > 100$ GeV	2	3.6 ± 1.6	1	2.8 ± 1.2



(a) (b) (c) (d)

Figure 4.7: Transverse mass distributions before ((a) and (b)) and after ((c) and (d)) applying a cut on the transverse missing energy $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV, for the electron (left) and muon (right) channel. As expected, the $W+$ jets and $t\bar{t}$ backgrounds peak around the W-mass.

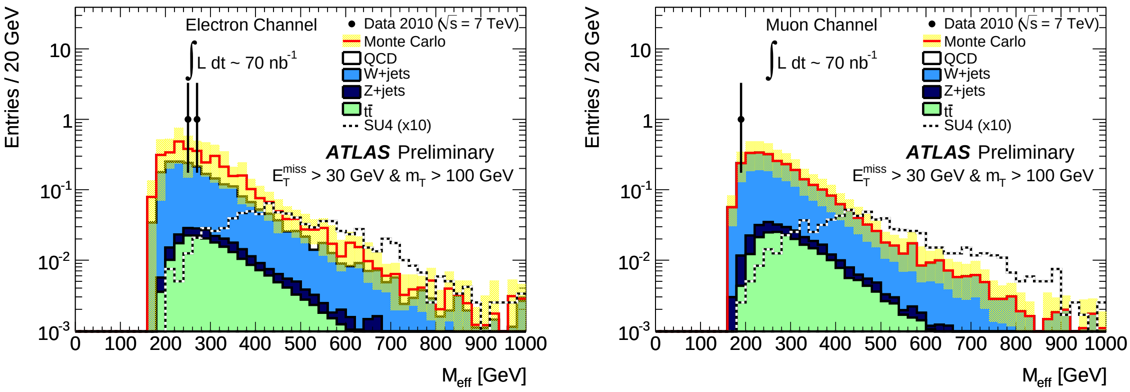


Figure 4.8: Effective mass distributions for the electron (left) and muon (right) channel with a cut on the transverse missing energy $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV and on the transverse mass $m_{\mathrm{T}} > 100$ GeV to better reject leptons from W decays (cf. cut 7 in table 4.4). Only three data events remain in the signal region consistent with the Standard Model prediction.

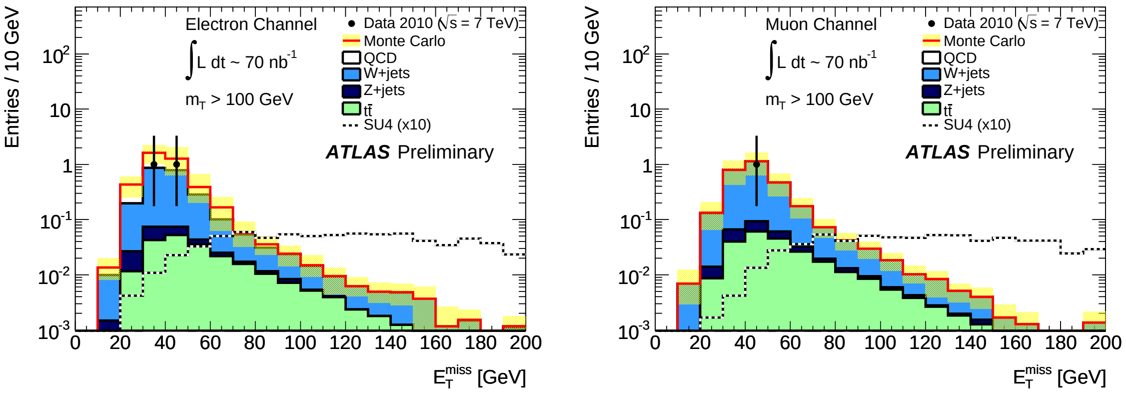


Figure 4.9: Missing transverse energy distributions for the electron (left) and muon (right) channel after a cut on the transverse mass $m_\mathrm{T} > 100$ GeV. The three signal region data events are measured in the region of highest probability as predicted by the Standard Model.

Chapter 5

Search for supersymmetry at $\sqrt{s} = 7$ TeV with all data taken in 2010

5.1 Introduction

In this chapter, which concludes the part on supersymmetry searches, the one-lepton channel analysis including all the data collected in 2010 is discussed, with a focus on the contribution provided by the author.

Thanks to a smoothly running machine, a total of $45\,\mathrm{pb}^{-1}$ of integrated luminosity could be recorded by ATLAS until end of October 2010. This substantial amount of data together with an increase in manpower allowed for a more sophisticated analysis in terms of an improved background determination, culminating in a combined estimation by means of a profile likelihood fit. As in the previous chapter, the subsequent sections first briefly describe the physics simulation, data selection and object definition, followed by the definition of the event selection cuts, the Standard Model background determination focusing on the W estimation and the presentation of the various data distributions. Continuing the practice of the analysis with early data, two parallel selections distinct by the flavor of the event's isolated lepton are investigated for significant deviations from the Standard Model predictions. In the absence of any significant excess, new limits both on a general signal rate and the squark and gluino masses in a mSUGRA scenario can be set and are presented at the end of the chapter.

A large part of the results shown below have been published in [4] and in a supporting ATLAS-internal note [98]. The author has significantly contributed to the work on the event selection and the W background estimation.

Table 5.1: NNLO cross-sections of the most relevant backgrounds and NLO cross-section of the benchmark signal SU4. The cross-section for $t\bar{t}$ is *approximately* NNLO as defined in [88, 89, 90]. The figures for W/Z are stated per lepton flavor. The QCD multi-jets background is scaled to data in an appropriate control region (see subsection 5.6.1). For more details see tables B.1 to B.3 in the appendix.

Process	σ [pb]
$t\bar{t}$	160
$W \to \ell\nu$	10,450
$Z \to \ell\ell$	1,070
SU4	60

5.2 Signal and background simulation

The same samples as in the previous analysis with early data, which is for the rest of the chapter referred to as *early-data-analysis*, are used both for the mSUGRA benchmark signal SU4 and the Standard Model backgrounds (cf. section 4.3). In addition, single top processes simulated with MC@NLO [99, 100], $Wb\bar{b}$ produced with ALPGEN and di-boson samples produced with HERWIG are included. The QCD multi-jet events simulated with PYTHIA are normalized to data in a QCD control region (see subsection 5.6.1). Samples including pile-up effects have been used when available. However, no significant difference on the background estimation compared to non-pile-up samples is observed (for details on the study see appendix A in [98]) and therefore no additional systematic uncertainty is assigned.

For the determination of exclusion limits on the squark and gluino masses, which is presented in subsection 5.7.3, 220 grid points in the $m_0 - m_{1/2}$ plane of the $\tan\beta = 3$, $A_0 = 0$, $\mu > 0$ mSUGRA scenario have been generated with ISAJET [47] version 7.75 together with HERWIG++ [44] and the MRST2007LO* [91] PDF set. The $\tan\beta$ value has been decreased from 10 used in the prospects-analysis (cf. section 3.2) to 3 since most of the limits provided by the Tevatron experiments are available for the latter value.

Table 5.1 summarizes the cross-sections for the most important background processes and the benchmark signal SU4. All $V+$ jets samples are normalized to NNLO cross-sections or, in the case of $t\bar{t}$, to *approximate* NNLO as defined in [88, 89, 90]. A full list of the employed samples can be found in the appendix section B.1. The full detector simulation, provided by GEANT4, has been applied on all samples.

5.3 Data selection

As mentioned above, data with an integrated luminosity of approximately 45 pb^{-1} have been recorded by ATLAS. Requiring good operating conditions of all relevant subsystems, in order to ensure object reconstruction and trigger operation at the expected level of quality, reduces the amount of useable data to about $(35 \pm 4)\,\mathrm{pb}^{-1}$ with a luminosity uncertainty of 11% [97].

Interesting events are flagged with single lepton triggers for the offline analysis. Table 5.2 lists the data taking periods with their corresponding integrated luminosity and the unprescaled triggers used. In the Monte Carlo simulation, the same trigger is used for all data periods.

Table 5.2: Data taking periods with their corresponding integrated luminosity (L) and the unprescaled triggers used to select electron and muon events for the offline analysis.

		Electrons		**Muons**	
Period	L [pb^{-1}]	Data	Simulation	Data	Simulation
A–D	0.3	EF_e10_loose	EF_e10_medium	L1_MU6	EF_mu10_MSonly
E	0.9	EF_g17_etcut	EF_e10_medium	EF_mu10_MSonly	EF_mu10_MSonly
F–H	14.4	EF_e15_medium	EF_e10_medium	EF_mu13	EF_mu10_MSonly
I	19.0	EF_e15_medium	EF_e10_medium	EF_mu13_tight	EF_mu10_MSonly

The electron trigger efficiency is determined via a *tag-and-probe*-method and found to be almost at 100% for electrons with a $p_{\mathrm{T}} > 20$ GeV, being in very good agreement with the efficiency found in Monte Carlo. The muon trigger efficiency, determined using both a muon-independent data stream and tag-and-probe is worse with 0.74 ± 0.01 in the barrel and about 0.91 ± 0.01 in the endcap region for muons in the trigger plateau starting at $p_{\mathrm{T}} \approx 15$ GeV. The efficiency in the simulation is approximately 10% and 1% higher in the barrel and endcap regions respectively, which yields an overall scale factor of 0.95 ± 0.05. Since the factor is compatible with unity, it is only applied in the combined profile likelihood fit discussed in subsections 5.7.2 and 5.7.3 in order to determine the statistical uncertainties more accurately. The underlying studies are presented in detail in [98].

5.4 Object definition

The object definition is based on the one of the early-data-analysis (section 4.5), but has been further refined using the additional collected data.

Jets

As in the analysis of chapter 4, jets are reconstructed using the anti-k_t algorithm with a cone of $R_{cone} = 0.4$ and are calibrated to the hadronic scale. The jets are required to have a minimum transverse momentum of $p_{\mathrm{T}} > 20$ GeV and to lie within a pseudorapidity range of $|\eta| < 2.5$. As above, a set of cleaning cuts is applied to reject events which contain jets probably not originating from proton-proton collision (for details see [98, 101]). The latter concerns less than 0.1% of the events in data and in a typical SUSY signal simulation.

b-Jets

In order to obtain a top quark-enhanced sample a b-jet identification algorithm [102], shortly called *SV0 tagger*, is used on top of the jet reconstruction described above. b-jets originate from secondary vertices, which are significantly displaced from the primary vertex of the collision due to the decay of the b-hadron within the jet. A jet is tagged if its signed decay length significance $L/\sigma(L)$ is greater than 5.72, where L is the decay length, $\sigma(L)$ its resolution and the sign given by the projection of the decay length on the jet axis. This particular value is chosen since it yields a reasonable b-tagging efficiency of about 55% while providing a good rejection against other quark and gluon jets [103, 104].

Based on dedicated SV0 tagger performance studies [105], the b-tag efficiency and mistag rate in the Monte Carlo simulation are corrected by calculating weights for each selected jet, both tagged and non-tagged, and reweighting the whole event by the product of these jet weights, as detailed in [106, 98].

Electrons

The electron definition, detailed in [77, 107], is based on the *medium-level* selection used in the previous chapter. Candidates must fulfill shower shape criteria in the first two layers of the electromagnetic calorimeter and certain quality critera on the loosely matched inner detector track. In addition, no energy deposits in the first layer of the hadronic calorimeter are allowed. Electrons having these properties are labeled *RobustMedium*. If they pass further requirements in the pixel detector, in the TRT and on the ratio of cluster energy to track momentum, they are classified as *RobusterTight* electrons. Since the latter requirements are sufficient for a good rejection against fakes from QCD multi-jet events (roughly 6000 and 80000 for RobustMedium and RobusterTight respectively [107]), no further isolation criteria are imposed.

Electrons are required to fulfill $p_{\mathrm{T}} > 20$ GeV and $|\eta_{cl}| < 2.47$ where only the pseudorapidity of the calorimeter cluster is used. Events are discarded if a selected electron lies in the crack region defined by $1.37 < |\eta_{cl}| < 1.52$. If an electron touches a calorimeter region with a faulty read-out it is dropped, but the event is kept. Jets overlapping with a RobustMedium electron within a radius of $\Delta R < 0.2$ are discarded. If the two objects overlap within $0.2 < \Delta R < 0.4$, the electron is removed. The Monte Carlo expectation is rescaled (depending on the electron's pseudorapidity) to reproduce the electron identification efficiency measured in data [108], which amounts to about 94% (average associated data/MC scale factor ≈ 0.97) for RobustMedium and 76% (average scale factor ≈ 0.98) for RobusterTight electrons.

Muons

As in the previous analyses, muons are identified using the Staco algorithm based on the combination of tracks reconstructed in the inner detector (ID) and the muon system (MS). In addition, to increase the identification efficiency in regions with worse MS performance, also muons, whose ID tracks can be matched to only partially reconstructed MS tracks, are included. The detailed requirements are listed in [98] and follow the recommendations of the Muon Combined Performance Group [109]. In this analysis, muons are considered isolated if the sum of momenta in a cone with $R = 0.2$ around the track is less than 1.8 GeV.

Muons are required to fulfill $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$. Muons overlapping with reconstructed jets within radius of $\Delta R < 0.4$ are removed. Events with a muon whose extrapolated track has a distance of more than 1 cm in the z-direction from the primary vertex are rejected. In order to reproduce the observed momentum resolution in data, which, due to the preliminary alignment and calibration, has not yet fully achieved the expected performance [110], a muon momentum smearing is applied on the Monte Carlo simulation as proposed in [109, 110]. Apart from that, no rescaling of the identification efficiency, which is well above 90%, is necessary.

Missing transverse energy

The missing transverse energy is calculated from the vectorial sum of the reconstructed jets, the isolated signal lepton, additional non-isolated muons or muons overlapping with jets and topological calorimeter clusters not assigned to any of those objects. Only muons which meet all the requirements stated above except for the isolation and jet overlap conditions are considered.

5.5 Event selection

The large amount of available data compared to the situation in the early-data-analysis allows for a reduction of the investigated phase space to higher momentum regions, where the relative population with Standard Model events is even further decreased. Only events with at least three high-energetic jets, exactly one isolated electron or muon and large missing energy are selected for closer inspection. The following cuts define the one-lepton search presented in this chapter:

1. Lepton trigger (see section 5.3).
2. Primary vertex with more than four associated tracks.
3. Exactly one lepton with $p_{\mathrm{T}} > 20$ GeV.
4. At least three jets with $p_{\mathrm{T}} > 30$ GeV with the leading jet $p_{\mathrm{T}} > 60$ GeV.
5. $\Delta\phi$ between jet and $E_{\mathrm{T}}^{\mathrm{miss}} > 0.2$ for the three leading jets.
6. Missing transverse energy: $E_{\mathrm{T}}^{\mathrm{miss}} > 125$ GeV.
7. Transverse mass, formed from the lepton and the missing energy: $m_{\mathrm{T}} > 100$ GeV.
8. $E_{\mathrm{T}}^{\mathrm{miss}}$ of the same magnitude as the average jet p_{T}: $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$.
9. High mass region: $m_{\mathrm{eff}} > 500$ GeV.

The minimal ϕ-angle requirement between jets and the transverse missing energy, already introduced in chapter 3, aims to reject events where the $E_{\mathrm{T}}^{\mathrm{miss}}$ mainly arises from an underestimation of jet momenta. The relatively high $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is a consequence of cuts 8 and 9 and is only kept to validate data–simulation agreement at an intermediate stage of the analysis. Sufficient data statistics also permits the reintroduction of the $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$ cut, which is effective against fake missing energy fluctuations. Given the expected large mass of the colored SUSY particles, a minimum effective mass of 500 GeV is required for events to enter the signal region.

5.6 Background estimation

As already mentioned in the introduction, data-driven methods can be employed to estimate the contribution of the most import Standard Model backgrounds, that is QCD multi-jets, $W+$ jets events and top quark production. Based on the approach in the early-data-analysis two disjunct regions are defined in the $E_{\mathrm{T}}^{\mathrm{miss}} - m_{\mathrm{T}}$ plane, providing control samples of diverse purity, for all events passing cut 5 of section 5.5, i.e. requiring three hard jets and exactly one isolated lepton and applying the $\Delta\phi$ cut between the jets and the missing momentum. Whereas the W control region is now defined to lie within 30 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 80 GeV and 40 GeV $< m_{\mathrm{T}} <$ 80 GeV, the QCD control region, denoted by QR, remains the same with $E_{\mathrm{T}}^{\mathrm{miss}} <$ 40 GeV and $m_{\mathrm{T}} <$ 40 GeV. Two exclusive selections within the W control region, denoted by WR and TR as they are enriched in $W+$ jets and top events respectively, are obtained by requiring that none or at least one of the three selected jets is b-tagged respectively. Figure 5.1 shows the positions of these regions and the signal region in the $E_{\mathrm{T}}^{\mathrm{miss}} - m_{\mathrm{T}}$ plane.

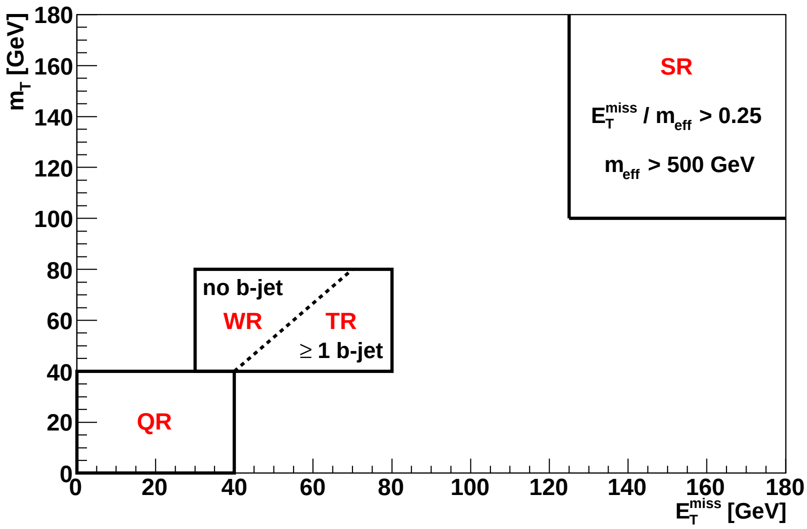


Figure 5.1: Position of the signal (SR) and control regions (QR, WR, TR) in the $E_{\mathrm{T}}^{\mathrm{miss}} - m_{\mathrm{T}}$ plane as described in the text. The W control region can be split in two selections WR and TR, which are enriched in $W+$ jets jets and top events respectively, by the b-tagging requirement. Two additional conditions are imposed on the events in the signal region as depicted in the figure.

5.6.1 QCD background estimation

A convenient method to estimate the QCD multi-jets background is the so-called *Matrix-Method* which has been recently applied in a top pairs observation study [111]. It is based on the selection of two data samples, usually called *loose* and *tight* sample. The loose one differs from the tight by relaxed lepton identification criteria which result in a substantially increased contamination with fake leptons originating from QCD events. The number of fake leptons, and therefore the number of QCD events, passing the tight identification cuts can be determined analytically using the efficiencies of fake and real leptons from the loose sample to also pass the tight criteria. The efficiencies can be taken from the simulation, but are here determined using semi-data-driven methods, as detailed in [98].

In the electron channel, another approach based on a combined fit of templates associated with the different origins of the fake electrons (i.e. bremsstrahlung conversion, heavy flavor decays and misidentified hadrons) is used as an additional cross-check.

Except for the signal region, where the matrix method is used, the QCD background is constrained based on measurements in the QCD control region. Backgrounds other than QCD multi-jet events are subtracted from the data in that region, before dividing by the QCD MC prediction (MC expectations and data stated in table 5.12). The obtained scaling factors are consistent with results from the other methods discussed above. In the following, only the main results of the different methods are presented for the electron and muon selection. The detailed studies can be found in [98].

Electron channel

For the signal region, the matrix method yields an upper limit on the expected number of QCD multi-jet events of 0.26 events. This number is dominated by the statistical uncertainty of the number of events in the loose sample. For all the other selections, a rescaling of QCD to data, determined in the QCD control region, is applied with a scaling factor of 0.76 ± 0.13 [1] where the stated error represents the statistical uncertainty. The rescaled expected QCD events are in agreement with both the values predicted by the matrix method and the combined template fit in other investigated regions.

Muon channel

As in the electron channel, the matrix method is employed to obtain an upper limit of QCD events in the signal region which turns out to be 0.49 and is again dominated by the statistical uncertainty on the number of events in the loose sample. The higher value with respect to the electron channel result arises from the higher probability of loose fake muons also to pass the tight criteria. As above, a QCD rescaling of 0.44 ± 0.08, determined in the QCD control region, is used for all the other selections, again in agreement with matrix method predictions in other investigated regions.

5.6.2 $W +$ jets background estimation

The W control region is used to verify and, if necessary, correct the prediction of the Monte Carlo simulation with data both for $W +$ jets events, which is detailed in the following, and also for top quark processes, which is briefly discussed in the subsequent subsection 5.6.3. The definition of the region's boundaries is motivated by its relatively high portion of W and top events with respect to other processes. An alternative, almost purely data-driven method based on the W charge asymmetry present at proton-proton colliders is used to support the estimation.

Validation of the simulation with data

Table 5.3 summarizes the contributions of the different physics processes in the W control region (30 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 80 GeV, 40 GeV $< m_{\mathrm{T}} <$ 80 GeV) as predicted by the simulation for the electron and muon channel. The QCD prediction has been normalized with the scaling factors stated in section 5.6.1.

In the electron channel, 282 data events are measured whereas 265 ± 6 events are expected by the simulation, with 56% of which originating from $W +$ jets production. This amounts to an underestimation of the data by approximately 6%. In the muon channel, the simulation,

---

[1] At the time of writing, the text of the supporting internal note [98] included obsolete figures for the scale factors (electron channel: 0.70 ± 0.12, muon channel: 0.55 ± 0.09), whereas the results stated in the note were obtained using the correct ones (electron channel: 0.76 ± 0.13, muon channel: 0.44 ± 0.08).

predicting 282 ± 5 events with 59% from $W+$ jets events, apparently overestimates the measured 258 events by about 9%. However, the agreement improves if a scale factor of 0.95 ± 0.05, derived from discrepancies in the muon trigger efficiency between data and simulation (see section 5.3), is applied resulting in a reduced central value of the prediction of 267 events.

Table 5.3: Measured and expected number of events in the W control region for a luminosity of $35\,\text{pb}^{-1}$. The total prediction (SM) comprises the following contributions: $W+$ jets (W), top quark (T), $Z+$ jets together with di-boson processes (Z+DB) and QCD multi-jets (QCD). All stated errors represent the statistical uncertainty of the simulation and the error on the scaling factors in case of QCD.

	Data	**SM**	**W**	**T**	**Z+DB**	**QCD**	**SU4**
Electrons	282	265 ± 6	147.5 ± 3.5	94.0 ± 0.9	9.9 ± 0.8	12.8 ± 4.1	8.4 ± 0.6
Muons	258	282 ± 5	166.4 ± 3.8	100.2 ± 0.9	11.4 ± 0.9	3.5 ± 2.2	7.4 ± 0.6

Both deviations are in principal compatible with the statistical spread of a Poisson process. Furthermore, it has to be noted that the stated errors only represent the statistical uncertainty of the Monte Carlo simulation. Additional systematic effects, associated with detector mismeasurements and an incomplete modeling of the physics processes, are to be included in order to conclude on the goodness of the data–simulation agreement. They are discussed in the following.

Energy scale uncertainty The jet energy scale uncertainty for jets within $|\eta| < 2.8$ and with a $p_{\mathrm{T}} > 20$ GeV is determined to be smaller than 10% per jet as estimated by various studies comparing different physics models, in particular concerning the hadronic showering, and different detector configurations as well as comparisons with data using di-jet events [94, 112]. To account for differences in the quark and gluon content of $W/Z+$ jets with respect to QCD di-jet events, which entails a different calorimeter response, an additional uncertainty of 5% per jet is assigned for these and di-boson processes.

The electron and muon energy scale uncertainties have been determined to be 3% and 1% respectively using reconstructed Z events in data [113, 114]. In all cases, the change on the energy scale is propagated to the missing energy.

Energy resolution uncertainty To correct for deviations in the modeling of the jet resolution, a random jet energy smearing is applied following the prescription detailed in [115].

Each jet p_{T} is smeared independently with a p_{T}-dependent Gaussian resolution of

$$0.55\sqrt{\frac{(4.6\ \mathrm{GeV})^2}{p_{\mathrm{T}}^2}+\frac{0.846^2\ \mathrm{GeV}}{p_{\mathrm{T}}}+0.064^2}, \tag{5.1}$$

which broadens the p_{T} spectra by about 12% and 8% for a p_{T} of 12 and 8 GeV respectively.

Likewise, the electron energy resolution is adjusted to account for inaccuracies in the simulation following the recommandations in [113]. The electron energy is smeared independently with a Gaussian resolution, depending on the cluster energy E_{cl}, of

$$\sqrt{\left(S\,(1+\Delta S)\,\sqrt{E_{\mathrm{cl}}}\right)^2+\left(C\,(1+\Delta C)\,E_{\mathrm{cl}}\right)^2-\left(S\sqrt{E_{\mathrm{cl}}}\right)^2-(CE_{\mathrm{cl}})^2}, \tag{5.2}$$

with $S = 0.1$ the sampling term, $\Delta S = 0.2$ the error on the sampling term, $C = 0.007$ the constant term, and ΔC the error of the constant term. The error on the constant term is $\Delta C = 1$ in the barrel ($|\eta_{\mathrm{CL}}| < 1.37$) and $\Delta C = 4$ in the end cap region ($|\eta_{\mathrm{CL}}| \geq 1.37$). The muon energy is varied randomly by 5% in the barrel and 9% in the end cap to reproduce the resolution observed in data [114]. In all cases the missing energy is corrected accordingly.

Lepton identification efficiency uncertainty For both channels the error on the lepton identification and reconstruction effiency is assumed to be 5% [114, 108].

Luminosity uncertainty As in the early-data-analysis, an uncertainty of 11% is assigned to the luminosity as determined by [97].

Theory uncertainties As detailed in [98], generator studies involving comparisons of different simulations and variations of various theory scales yield theoretical uncertainties on the rates of $W+$ jets and $t\bar{t}$ events of 50% and 8% respectively, which are also applied on $Z+$ jets and single top processes in that order.

Table 5.4 summarizes the impact of these systematic variations on the prediction of top production and $V+$ jets events when applied on the events in the W control region. Aside from the theory error, the jet energy scale (JES) uncertainty constitutes the dominant error contribution for $V+$ jets events followed by the luminosity uncertainty. The reason for the JES uncertainty of the top background to be much smaller is twofold. No additional 5% error is applied and the net migration into or out of the control region when changing the energy scale is much smaller compared to $V+$ jets events. When considering the whole $E_{\mathrm{T}}^{\mathrm{miss}} - m_{\mathrm{T}}$ plane, the expected number of top events change by approximately ±9% in both channels when raising and lowering the JES, respectively.

The opposite and counter-intuitive change in the number of expected events when raising or lowering the lepton energy scale is also a result of the net migration across the control region boundaries. The error on the QCD events is determined in the course of the global profile likelihood fit, which is discussed in subsection 5.7.2.

Table 5.4: Impact of systematic variations, as described in the text, on the predicted number of $V+$ jets and top events in the W control region of the electron and muon channel for a luminosity of $35\,\text{pb}^{-1}$. The following systematic uncertainties are shown: Limited Monte Carlo statistics (MC stat.), jet energy scale (JES), electron/muon energy scale (E/MES), jet energy resolution (JRES), electron/muon energy resolution (E/MRES), electron/muon identification and reconstruction efficiency (E/MEFF), theory uncertainties and the luminosity uncertainty. When relevant, the systematic variations are performed upward and downward, respectively, and the resulting variations are presented in this order. ΔN^{CR} states the absolute total uncertainty on the predicted number of events N^{CR}. The error on the QCD events is determined in the course of the global profile likelihood fit, discussed in subsection 5.7.2.

Electrons	Data	Sum	V+jets	top	QCD
N^{CR}	282	265	158	94	13
ΔN^{CR}		±92	±87	±15	±27
MC stat. [%]			±2	±1	
JES [%]			${}^{+15}_{-20}$	${}^{-1}_{-4}$	
EES [%]			${}^{-4}_{+2}$	${}^{-3}_{+4}$	
JRES [%]			±2	±4	
ERES [%]			±0.9	±0.3	
EEFF [%]			±6	±6	
LUMI [%]			±11	±11	
Theory [%]			±50	±8	

Muons	Data	Sum	V+jets	top	QCD
N^{CR}	258	282	178	100	4
ΔN^{CR}		±103	±101	±17	±13
MC stat. [%]			±3	±1	
JES [%]			${}^{+24}_{-19}$	${}^{+6}_{-6}$	
MES [%]			${}^{-6}_{+5}$	${}^{-3}_{+4}$	
JRES [%]			±3	±6	
MRES [%]			±0.7	< 0.1	
MEFF [%]			±6	±6	
LUMI [%]			±11	±11	
Theory [%]			±50	±8	

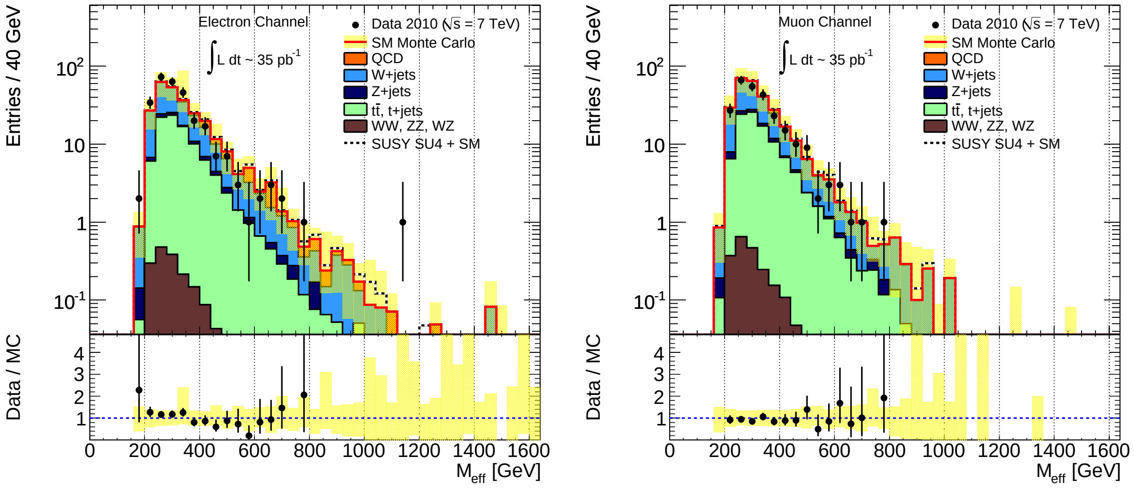


Figure 5.2: Effective mass distributions of events in the W control region for the electron (left) and muon (right) channel. Good agreement between data and simulation is observed, as displayed by the ratio of data over total simulated SM events shown on the bottom.

Combining all the different contributions, the total relative errors on the background prediction amount to 35% and 37% for the electron and muon channel respectively, which cover the observed deviations from the data.

Figure 5.2 shows the effective mass distributions in the W control region for both channels. As in the early-data-analysis, the errors on the data indicate the 68% Poisson confidence intervals of the statistical uncertainty. The QCD multi-jet events, which play only a minor role in this selection, are shown in orange to better distinguish them from the benchmark signal SU4, which is added on top of the Standard Model background. The yellow error band represents the total background uncertainty including the jet energy scale uncertainty, the theory errors, the errors due to limited Monte Carlo statistics, the errors on the QCD scaling factors and the luminosity uncertainty. The distributions show good agreement between data and simulation, displayed explicitly in the ratio plots on the bottom.

Figures 5.3 and 5.4 show the missing transverse energy and transverse mass distributions of the same events. All data points are consistent with the prediction within the indicated uncertainty. The lowest bin in the electron channel displays a large relative error on the Standard Model background, especially in the transverse mass distribution (figure 5.4) where also the total background appears underestimated. Both effects can be traced back to the limited Monte Carlo statistics of the QCD simulation.

A scaling factor for $W+$ jets can be determined by first subtracting all other backgrounds from the data and then dividing by the expected number of $W+$ jets events. Figure 5.5 shows this factor and its statistical uncertainty for the different data-taking periods. Considering also the determined systematic uncertainties, the scaling factors, including all data periods,

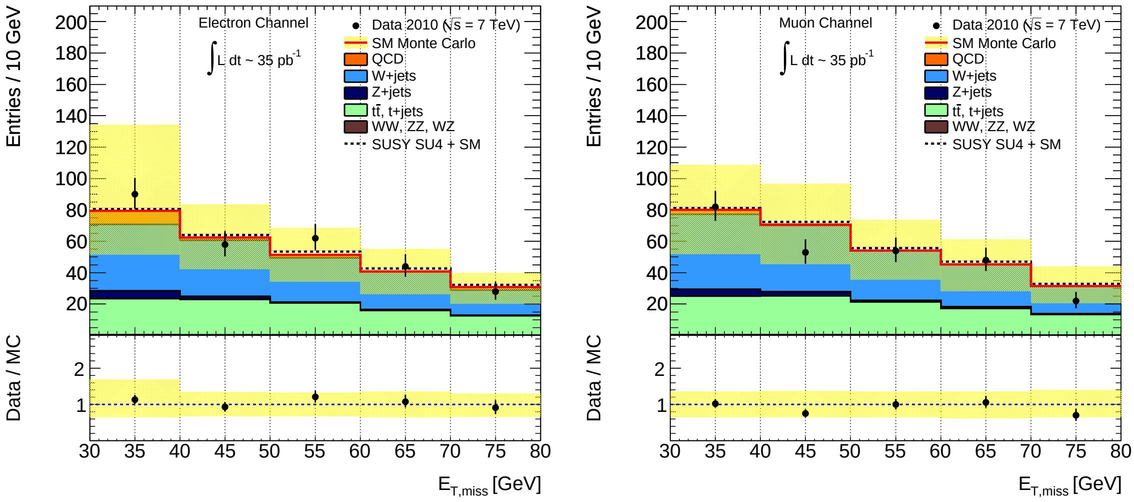


Figure 5.3: Missing transverse energy distributions of events in the W control region for the electron (left) and muon (right) channel. The large uncertainty in the lowest bin of the left plot is due to the limit statistics of the QCD multi-jets simulation.

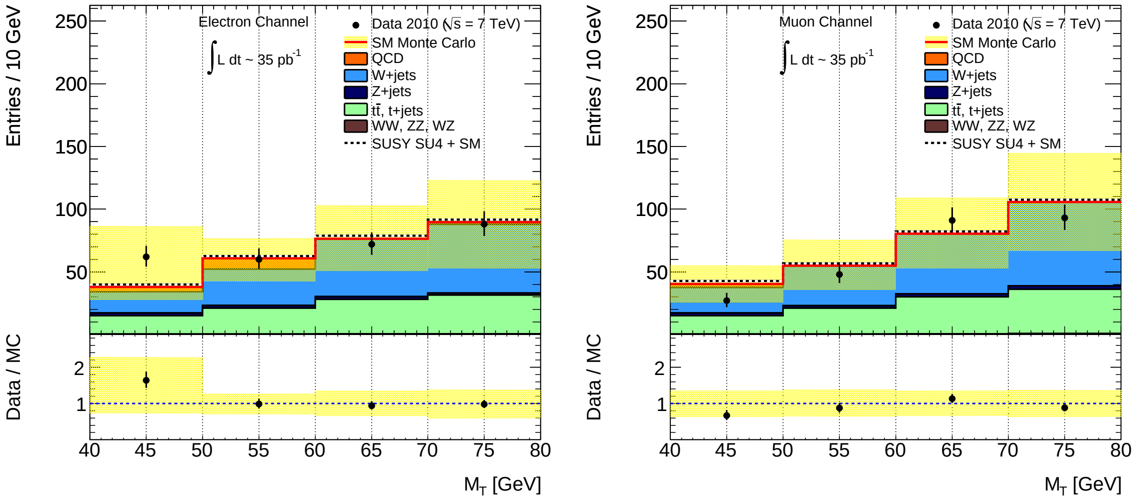


Figure 5.4: Transverse mass distributions of events in the W control region for the electron (left) and muon (right) channel. The deviation and large uncertainty in the lowest bin of the left plot is due to the limit statistics of the QCD multi-jets simulation.

amount to 1.1 ± 0.1 (stat) ± 0.5 (sys) in the electron and 0.9 ± 0.1 (stat) ± 0.5 (sys) in the muon channel. They are compatible with unity and the $W+$ jets background thus does not need to be rescaled.

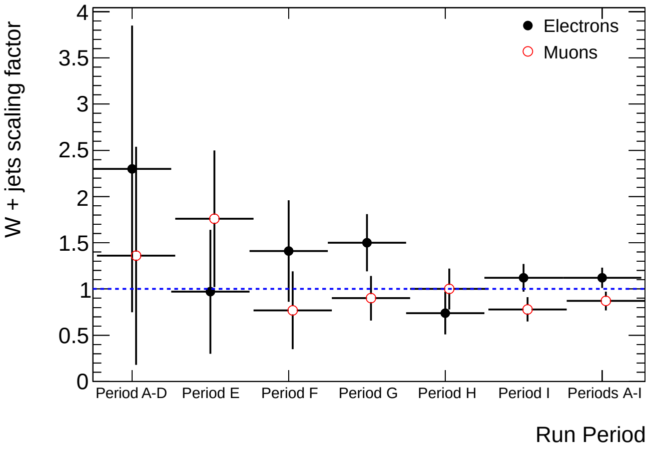


Figure 5.5: $W+$ jets scaling factor for the different data taking periods, determined in the W control region. The factor is obtained by subtracting the other simulated backgrounds from the data and then dividing by the expected number of $W+$ jets events. Only the statistical errors are shown and indicate consistence between the runs. The periods correspond to the following integrated luminosities: 0.3 (A-D), 0.9 (E), 1.7 (F), 5.7 (G), 7.0 (H) and $19.0\,\text{pb}^{-1}$ (I) totaling $34.6\,\text{pb}^{-1}$.

Alternative W and top background estimation using charge asymmetry

Due to the collision of two positively charged protons at the LHC the cross-section for positively charged W bosons, produced mainly in $u\bar{d}$ annihilation, is significantly higher than for negatively charged W bosons, produced mainly by $\bar{u}d$ collisions, which translates into a lepton charge asymmetry. Since all the other Standard Model processes except single top production are charge symmetric, this asymmetry can be exploited to predict the leptonic $W+$ jets background. One measures the difference $\Delta\mathrm{W}$ between the number of single lepton events with positive charges D^{+} and negative charges D^{-} in data:

$$\Delta\mathrm{W} = \mathrm{D}^{+} - \mathrm{D}^{-} \tag{5.3}$$

The estimated number of W events can then be calculated using

$$\mathrm{W}_{\mathrm{est}} = \Delta\mathrm{W}\frac{\mathrm{C}+1}{\mathrm{C}-1}, \tag{5.4}$$

where C is the ratio of the expected number of positive W events over the number of negative ones:

$$C = \frac{W^+_{MC}}{W^-_{MC}} \tag{5.5}$$

This ratio is the only quantity taken from Monte Carlo simulation classifying the method as almost purely data-driven. The inverse of the ratio in equation (5.4) is the quantity usually referred to as asymmetry A:

$$\left(\frac{C+1}{C-1}\right)^{-1} = A = \frac{W^+ - W^-}{W^+ + W^-} \tag{5.6}$$

Tables 5.5 and 5.6 show the estimated numbers of leptonic W events in the W control region and the whole $E_T^{\text{miss}} - m_T$ plane up to 80 GeV using the presented method. Besides the hitherto used 3-jet channel, a softer selection requiring at least two jets with transverse momenta greater 60 and 30 GeV respectively is considered in addition to validate the performance of the method. Obviously, the numbers for the inclusive 2- and 3-jet selections are correlated. Generally, the estimations agree well with the ALPGEN Monte Carlo predictions within the stated errors. In the W control region, the total uncertainties amount to between 25% and 40% for the 2- and 3-jet selections when assuming a 15% systematic uncertainty on C, which is motivated below.

Table 5.5: Number of observed events with positively (D^+) and negatively (D^-) charged lepton as well as $W +$jets predicted by Monte Carlo simulation (W MC) and by the charge asymmetry method (W est.), including the relative statistical (Stat.) and systematical error (Sys.) on the latter, in the electron channel for a luminosity of 35 pb^{-1}. The error on the ratio $C = W^+/W^-$ represents the statistical uncertainty of the simulation. The values of E_T^{miss} and m_T are stated in GeV.

El	E_T^{miss}	m_T	D^+	D^-	Data	W MC	W est.	Stat.	Sys.	C
2 jet	(30, 80)	(40, 80)	512	366	878	616	660 ± 194	21%	21%	1.57 ± 0.04
2 jet	(0, 80)	(0, 80)	5124	4795	9919	1366	2024 ± 902	31%	32%	1.39 ± 0.02
3 jet	(30, 80)	(40, 80)	167	115	282	148	182 ± 65	33%	13%	1.80 ± 0.09
3 jet	(0, 80)	(0, 80)	1260	1164	2424	329	466 ± 265	52%	23%	1.52 ± 0.05

The magnitude of charge asymmetry depends on the contributions from the different inital parton combinations. Two gluons in the initial state give rise to charge-symmetric W production, while gluon–quark and quark–anti-quark interaction result in an overshoot of W^+ compared to W^-. As can be seen e.g. in figure 1.3, for increasing parton momentum, expressed by the proton momentum fraction x, the gluon contribution decreases compared to the quark contributions. In addition, the difference between the up and down quark contributions increases. Both effects result in a higher charge asymmetry. Selecting events

Table 5.6: Number of observed events with positively (D^+) and negatively (D^-) charged lepton as well as $W+$jets predicted by Monte Carlo simulation (W MC) and by the charge asymmetry method (W est.), including the relative statistical (Stat.) and systematical error (Sys.) on the latter, in the muon channel for a luminosity of 35 pb^{-1}. The error on the ratio $C = W^+/W^-$ represents the statistical uncertainty of the simulation. The values of E_T^{miss} and m_T are stated in GeV.

Mu	E_T^{miss}	m_T	D^+	D^-	Data	W MC	W est.	Stat.	Sys.	C
2 jet	(30, 80)	(40, 80)	501	345	846	697	577 ± 139	19%	15%	1.74 ± 0.04
2 jet	(0, 80)	(0, 80)	1702	1397	3099	1484	1543 ± 474	19%	24%	1.49 ± 0.02
3 jet	(30, 80)	(40, 80)	151	107	258	166	146 ± 58	38%	12%	1.86 ± 0.09
3 jet	(0, 80)	(0, 80)	436	377	813	354	278 ± 149	49%	22%	1.54 ± 0.05

with a higher number of additional jets implies higher average momenta of the involved initial partons, which thus increases the charge asymmetry (see also [116, 117]), as reflected by the higher value of C in tables 5.5 and 5.6.

Apart from that, the value of C in the muon channel turns out to be systematically higher than in the electron selection. Additional tracks associated with electrons from bremsstrahlung conversion increase the charge misidentification rate in the electron channel. Since the misidentification rate is charge-independent, it results in an overall decrease of the charge asymmetry with respect to the muon channel. However, the estimated electron charge misidentification probability of about 2% (see below) is not sufficient to explain the observed difference between the two channels.

Inspecting the proton momentum fraction distributions of the selected events, one observes that, due to different energy scales and resolutions associated with the object identification and reconstruction, initial partons with a higher x values are selected in the muon channel compared to the electron channel. This is evident in figure 5.6, which shows the normalized x distributions of the selected muon and electrons events and their ratio in the 2-jet selection. The average increase in x is thus the main cause for a higher charge asymmetry in the muon channel.

Statistical uncertainty Tables 5.5 and 5.6 show that in most cases the statistical error is the dominant uncertainty on the estimated W events for the available luminosity. It is estimated by

$$(\Delta \mathrm{W}_{\mathrm{est}})_{\ \mathrm{stat}} = \sqrt{\mathrm{D}}\,\frac{\mathrm{C}+1}{\mathrm{C}-1} \tag{5.7}$$

and thus scales with the square root of the luminosity. It is noteworthy that the error does

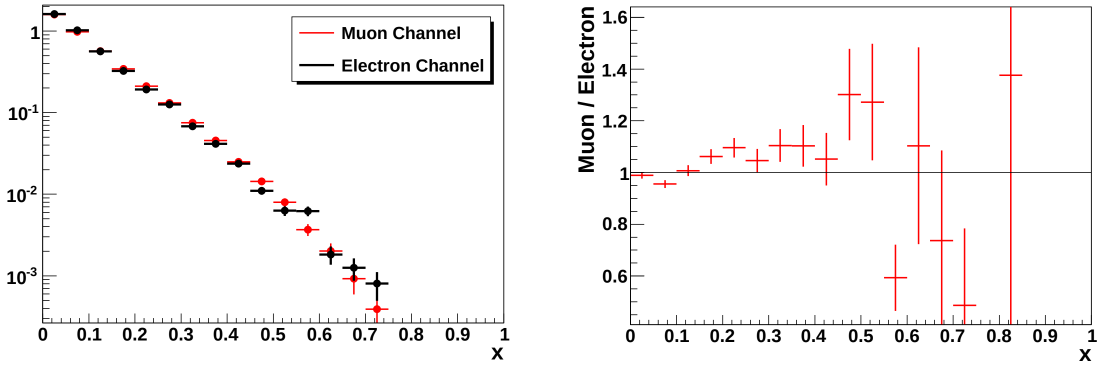


Figure 5.6: Proton momentum fraction (x) distributions, normalized to unity, of the selected events in the electron and muon channel (left) and their ratio (right) for the 2-jet selection. The selected muon channel events have a higher x value on the average, which results in the observed higher C value in tables 5.5 and 5.6.

not depend on the fraction of W events in the selection but only on its total number of data events. The relative error is thus reduced for increased purity of the control sample.

Generator uncertainty on the ratio $\mathbf{W^+/W^-}$ The other major source of uncertainty arises from the determination of C in equation 5.5. In order to estimate its magnitude, simulations with ALPGEN and MCFM [118] have been carried out for the 2-jet channel. No detector simulation has been applied. Table 5.7 shows the values of C for different minimum jet transverse momenta as well as for a combined variation of the renormalization and factorizations scales, whose central value was taken as the scalar sum of the outgoing parton transverse momenta denoted by H_T. Generally, the C values calculated with ALPGEN are higher than those from MCFM, but only by a few percent, and increase for higher minimum jet p_T. The variation of the scale H_T has a negligible effect.

Table 5.7: Values for the ratio W^+/W^- calculated with MCFM and ALPGEN for the 2-jet selection without applying a detector simulation. In the case of MCFM the scale H_T, which is the scalar sum of the outgoing parton p_Ts and has been taken as renormalization and factorizations scale, has been varied to check the stability of the prediction. The ALPGEN values are systematically higher. The calculations have been provided by Valerio Consorti.

		MCFM		**ALPGEN**
minimum jet p_T [GeV]	H_T [GeV]	$H_T \times 0.5$ [GeV]	$H_T \times 2$ [GeV]	
(50, 50)	1.65 ± 0.02	1.63 ± 0.02	1.65 ± 0.01	1.71 ± 0.02
(60, 30)	1.61 ± 0.01	1.61 ± 0.02	1.61 ± 0.01	1.66 ± 0.02
(80, 80)	1.72 ± 0.02	1.73 ± 0.03	1.72 ± 0.01	1.79 ± 0.04

PDF uncertainty on $\mathbf{W^+/W^-}$ The type of PDF used in the generation of the $W +$ jets events may also influence the expected ratio of postively over negatively charged W bosons. Therefore, the ALPGEN samples have been reweighted from CTEQ6ll [82] to MSTW2008 [119]. Tables 5.8 and 5.9 show the obtained values for C.

Table 5.8: Values for the ratio W^+/W^- using three different PDF sets, CTEQ6ll, MSTW2008LO and MSTW2008NNLO, for the electron channel. For MSTW2008LO, the full eigenset was used to include the parameter uncertainty (based on [120]). "upwards" and "downwards" refer to the value of the new event weight in relation to the central weight. These variations are much smaller than the statistical uncertainty. The values of $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T} are stated in GeV.

El			**CTEQ6ll**		**MSTW LO**		**MSTW NNLO**
	$E_{\mathrm{T}}^{\mathrm{miss}}$	m_{T}		central	upwards	downwards	
2 jet	(30, 80)	(40, 80)	1.57 ± 0.04	1.53 ± 0.04	1.52	1.53	1.50 ± 0.04
2 jet	(0, 80)	(0, 80)	1.39 ± 0.02	1.36 ± 0.02	1.36	1.37	1.33 ± 0.02
3 jet	(30, 80)	(40, 80)	1.80 ± 0.09	1.76 ± 0.09	1.75	1.77	1.73 ± 0.09
3 jet	(0, 80)	(0, 80)	1.52 ± 0.05	1.49 ± 0.05	1.49	1.50	1.45 ± 0.05

Table 5.9: Values for the ratio W^+/W^- using three different PDF sets, CTEQ6ll, MSTW2008LO and MSTW2008NNLO, for the muon channel. For MSTW2008LO, the full eigenset was used to include the parameter uncertainty (based on [120]). "upwards" and "downwards" refer to the value of the new event weight in relation to the central weight. These variations are much smaller than the statistical uncertainty.

Mu			**CTEQ6ll**		**MSTW LO**		**MSTW NNLO**
	$E_{\mathrm{T}}^{\mathrm{miss}}$	m_{T}		central	upwards	downwards	
2 jet	(30, 80)	(40, 80)	1.74 ± 0.04	1.68 ± 0.04	1.68	1.69	1.65 ± 0.04
2 jet	(0, 80)	(0, 80)	1.49 ± 0.02	1.46 ± 0.02	1.45	1.46	1.42 ± 0.02
3 jet	(30, 80)	(40, 80)	1.86 ± 0.09	1.82 ± 0.09	1.81	1.83	1.78 ± 0.09
3 jet	(0, 80)	(0, 80)	1.54 ± 0.05	1.50 ± 0.05	1.50	1.51	1.47 ± 0.05

MSTW systematically yields a smaller value than CTEQ. It is noteworthy that the error due to the PDF parameters does not cover this deviation, a fact which has also been observed elsewhere [121] and is presumed to be related with differences in the sea quark flavor structure of the PDF sets.

Table 5.10: Expected number of $V+$ jets events in the W control region as determined by three different methods: a pure Monte Carlo estimate, a semi data-driven estimate using W charge asymmetry and an estimate based on the profile likelihood fit (see table 5.12).

	Electrons	**Muons**
Monte Carlo	158 ± 87	178 ± 101
Charge asymmetry	192 ± 65	157 ± 58
Fit results	172 ± 31	127 ± 29

Charge misidentification Two-lepton-events with opposite and same charge, but having the same flavor, are selected in data if the invariant mass of the two leptons lies within a window of 70 GeV and 110 GeV and the associated jets fulfill the same transverse momentum cuts. Assuming that the leptons in these events orginate from a Z decay, one can estimate the probability for charge mismeasurement. Using the available luminosity, 9 events with both leptons having the same charge are found in the 2-jet electron channel. With 179 opposite charge events in the same selection, a probability for electron charge mismeasurement of $(2.4 \pm 1.1)\%$ follows. No same charge muon pairs and 211 opposite charge pairs are found implying a probability for muon charge mismeasurement below 1%.

Total uncertainty on $\mathbf{W^+/W^-}$ Considering all the uncertainties discussed above, a total systematic error of 15% is assigned to C which leads to the systematic uncertainty on the estimated W background shown in the last but one columns of tables 5.5 and 5.6.

Summary

Table 5.10 summarizes the estimated values of the $V+$ jets background in the W control region. The pure Monte Carlo prediction is in good agreement with the estimation based on the W charge asymmetry and the values obtained through the combined profile likelihood fit (see table 5.12).

5.6.3 Top background estimation

As mentioned at the beginning of the section, the events in the W control region can be split in two samples, denoted by WR and TR, to increase the $V+$ plus jets and top purity by requiring that at least one or none of the selected jets is b-tagged, respectively (for details see [98]). The resulting top control sample TR has a purity of about 82% and 85% for the electron and muon channel respectively and shows good agreement with the selected data. In the profile likelihood fit presented in subsection 5.7.2, a combined fit is carried out to simultaneously determine $V+$ jets and top backgrounds in the two selections.

5.7 Results

This section presents the main results of the analysis starting with a discussion of the event distributions after successively applying the selection cuts. The subsequent subsection briefly describes the combined profile likelihood fit and profile likelihood ratio test which are employed to determine exclusion limits for the mSUGRA scenario and upper limits on the number of expected events from a generic signal, as presented in the last subsection.

5.7.1 Event distributions

Figure 5.7 shows the missing transverse mass distributions after cut 5 of section 5.5, that is requiring three hard jets and exactly one isolated lepton and applying the $\Delta\phi$ cut between the jets and the missing momentum. All backgrounds are normalized to their respective cross-sections except QCD multi-jet events, which are normalized using the scale factors stated in subsection 5.6.1. Especially the electron selection is dominated by QCD multi-jet events at low missing transverse energy, followed by the $W+$ jets and top background. Diboson processes play a negligible role. Besides the jet energy scale error as the dominant detector-related uncertainty, the yellow band representing the total error on the Standard Model background includes the theory errors, the errors due to limited Monte Carlo statistics, the error on the QCD scaling factors and the luminosity uncertainty. The ratio plots on the bottom indicate good agreement between measurement and simulation.

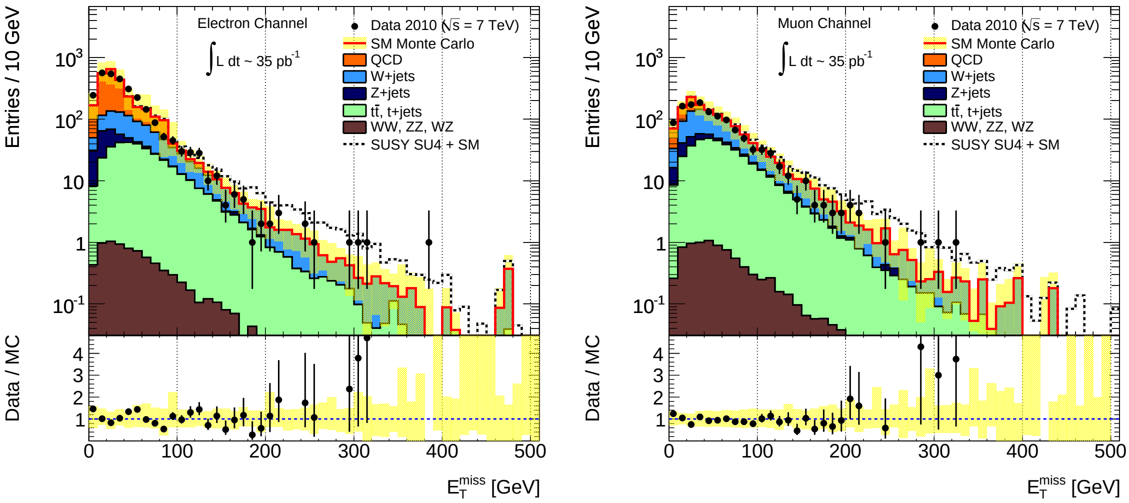


Figure 5.7: Missing transverse energy distributions for the electron (left) and muon (right) channel without a cut on the transverse mass. The predicted shapes are well reproduced as displayed by the ratio of data over total simulated SM events shown on the bottom. The yellow error band includes the uncertainties as described in the text.

The effective mass distributions of the same events are displayed in figure 5.8. The predicted shapes are well reproduced, especially in the muon channel.

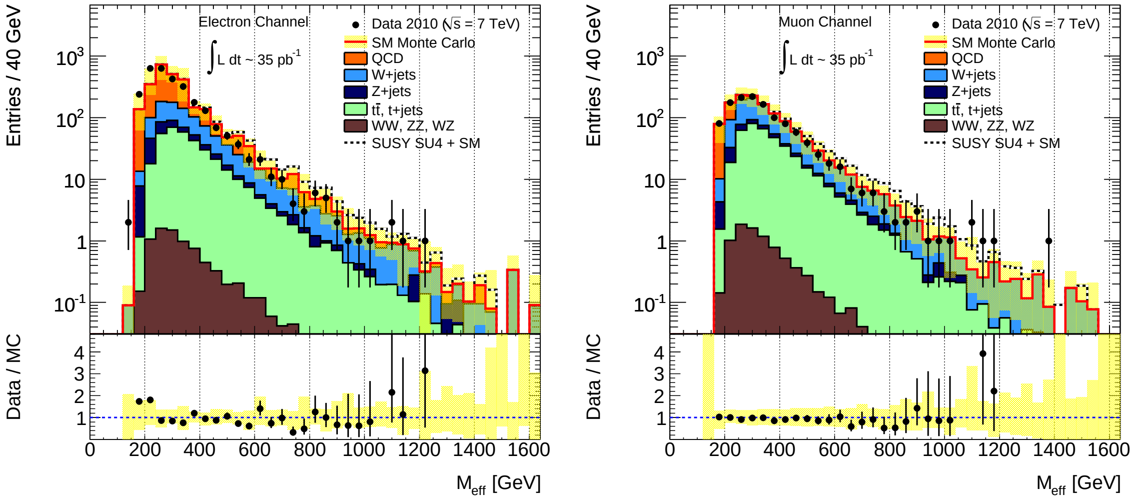


Figure 5.8: Effective mass distributions of the same events as in figure 5.7 for the electron (left) and muon (right) channel without cuts on the missing transverse energy or transverse mass. Good agreement between data and simulation is observed.

Table 5.11 summarizes the observed data and the predicted events separately for each process. The errors represent the statistical uncertainty of the simulation and, in the case of QCD multi-jets events, the error on the scaling factors. The uncertainty on the total SM background comprises all the uncertainties included in the yellow error band, as listed above. The total background is in good agreement with the data and no indication for new physics is observed.

Applying a cut of 125 GeV on the transverse missing energy removes most of the events, in particular QCD multi-jet processes. The remaining background consists of $W+$ jets and top events in almost equal parts as shown in the table and figure 5.9. Especially in the muon channel, the background appears to be slightly systematically overestimated in line with the overestimated trigger efficiency in the simulation (cf. section 5.3), but the deviations are covered by the determined uncertainty.

Figure 5.11 shows the transverse mass distributions before and after applying the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. Again, the shapes are well reproduced and motivate a cut $m_{\mathrm{T}} > 100$ GeV to remove the majority of top and W events. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for events with $m_{\mathrm{T}} > 100$ GeV are shown in figure 5.10, again indicating good description of the data by the simulation. The remaining events after applying both the $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T} cuts are displayed in figure 5.12. The numbers in table 5.11 indicate that top processes constitute almost 80% of the total background, the latter of which is in good agreement with the 5 electron and 6 muon data events left at this stage of the analysis. The requirement on the ratio of $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} (cut 8) does not significantly change the situation as can be seen in figure 5.13 with 5 events

Table 5.11: Observed and predicted number of events in the electron and muon channel after cuts 5 to 9 as defined in section 5.5[1]. The Standard Model background (SM) includes the following contributions: $W+$ jets (W), top processes (T), $Z+$ jets (Z), di-boson process (DB), QCD multi-jet events (QCD). The stated uncertainties represent the statistical error of the simulation and the error on the scaling factors in case of QCD. The uncertainty on the total SM background includes, in addition, the jet energy scale errors, the theory errors and the luminosity error. Good agreement with the Standard Model prediction is observed and no indication for new physics is found.

Cuts applied up to	**Electron channel**							
(as defined in section 5.5)	Data	SM	W	T	Z	DB	QCD	SU4
Cut 5: 3 jets, 1 lepton	2806	$2730\ ^{+980}_{-600}$	510 ± 7	338 ± 2	127 ± 3	7.7 ± 0.1	1750 ± 330	94 ± 2
Cut 6: $E_{\mathrm{T}}^{\mathrm{miss}} > 125$ GeV	70	$73\ ^{+29}_{-29}$	37 ± 2	28.7 ± 0.5	0.7 ± 0.2	0.67 ± 0.04	6 ± 3	50 ± 1
Cut 7: $m_{\mathrm{T}} > 100$ GeV	5	$5\ ^{+3.4}_{-1.9}$	1.1 ± 0.3	3.8 ± 0.2	-	0.07 ± 0.01	0.003 ± 0.001	23 ± 1
Cut 8: $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$	5	$4.2\ ^{+2.6}_{-1.7}$	0.8 ± 0.3	3.3 ± 0.1	-	0.06 ± 0.01	-	19.2 ± 0.9
Cut 9: $m_{\mathrm{eff}} > 500$ GeV	1	$1.8\ ^{+1.0}_{-0.9}$	0.4 ± 0.2	1.3 ± 0.1	-	0.03 ± 0.01	-	15.3 ± 0.8

Cuts applied up to	**Muon channel**							
(as defined in section 5.5)	Data	SM	W	T	Z	DB	QCD	SU4
Cut 5: 3 jets, 1 lepton	1226	$1300\ ^{+420}_{-390}$	612 ± 7	380 ± 1.8	73 ± 2.5	8.3 ± 1.3	230 ± 40	96 ± 2
Cut 6: $E_{\mathrm{T}}^{\mathrm{miss}} > 125$ GeV	62	$75\ ^{+38}_{-32}$	42 ± 2	31.1 ± 0.5	1.2 ± 0.3	0.73 ± 0.04	-	49 ± 1
Cut 7: $m_{\mathrm{T}} > 100$ GeV	6	$6.5\ ^{+3.1}_{-2.6}$	1.5 ± 0.3	4.8 ± 0.2	0.13 ± 0.09	0.07 ± 0.01	-	24 ± 1
Cut 8: $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$	5	$5.6\ ^{+2.5}_{-2.2}$	1.3 ± 0.3	4.2 ± 0.2	0.07 ± 0.07	0.06 ± 0.01	-	19.7 ± 0.9
Cut 9: $m_{\mathrm{eff}} > 500$ GeV	1	$2.1\ ^{+1.1}_{-1.0}$	0.7 ± 0.2	1.4 ± 0.1	-	0.04 ± 0.01	-	15.4 ± 0.8

[1] At the timing of writing, the incorrect errors stated in the corresponding table in [98] had not been corrected.

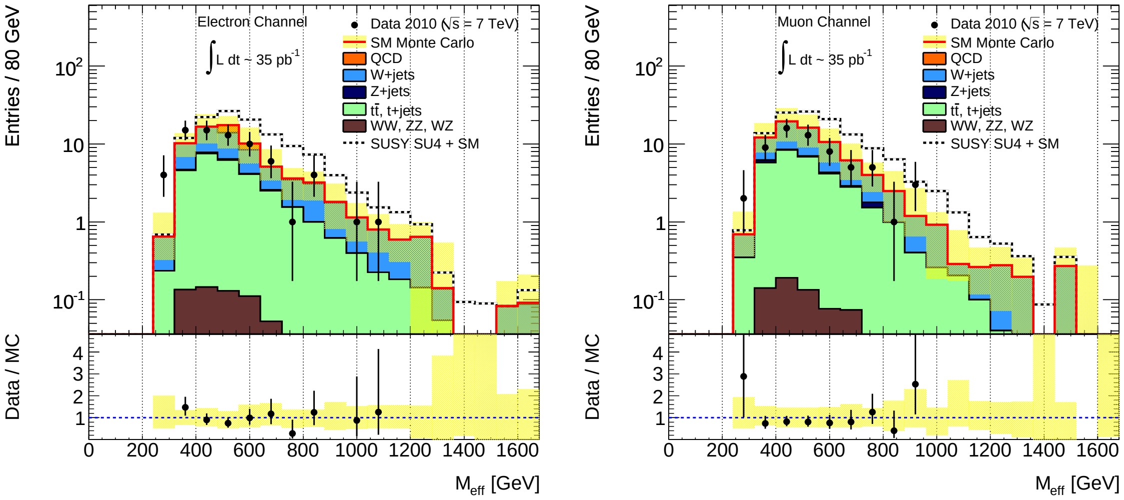


Figure 5.9: Effective mass distributions for the electron (left) and muon (right) channel with a cut on the transverse missing energy $E_{\mathrm{T}}^{\mathrm{miss}} > 125$ GeV. The QCD background has been practically removed compared to figure 5.8.

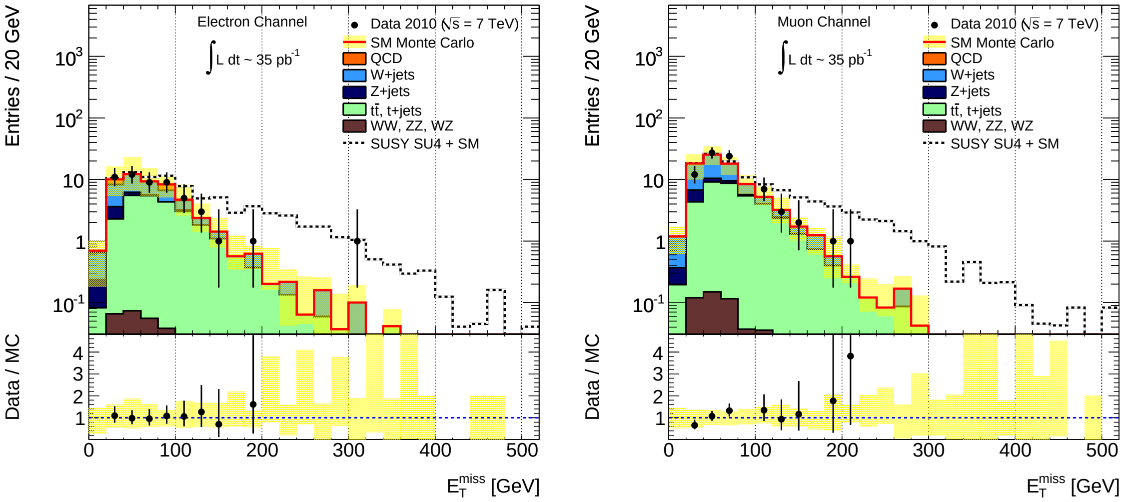
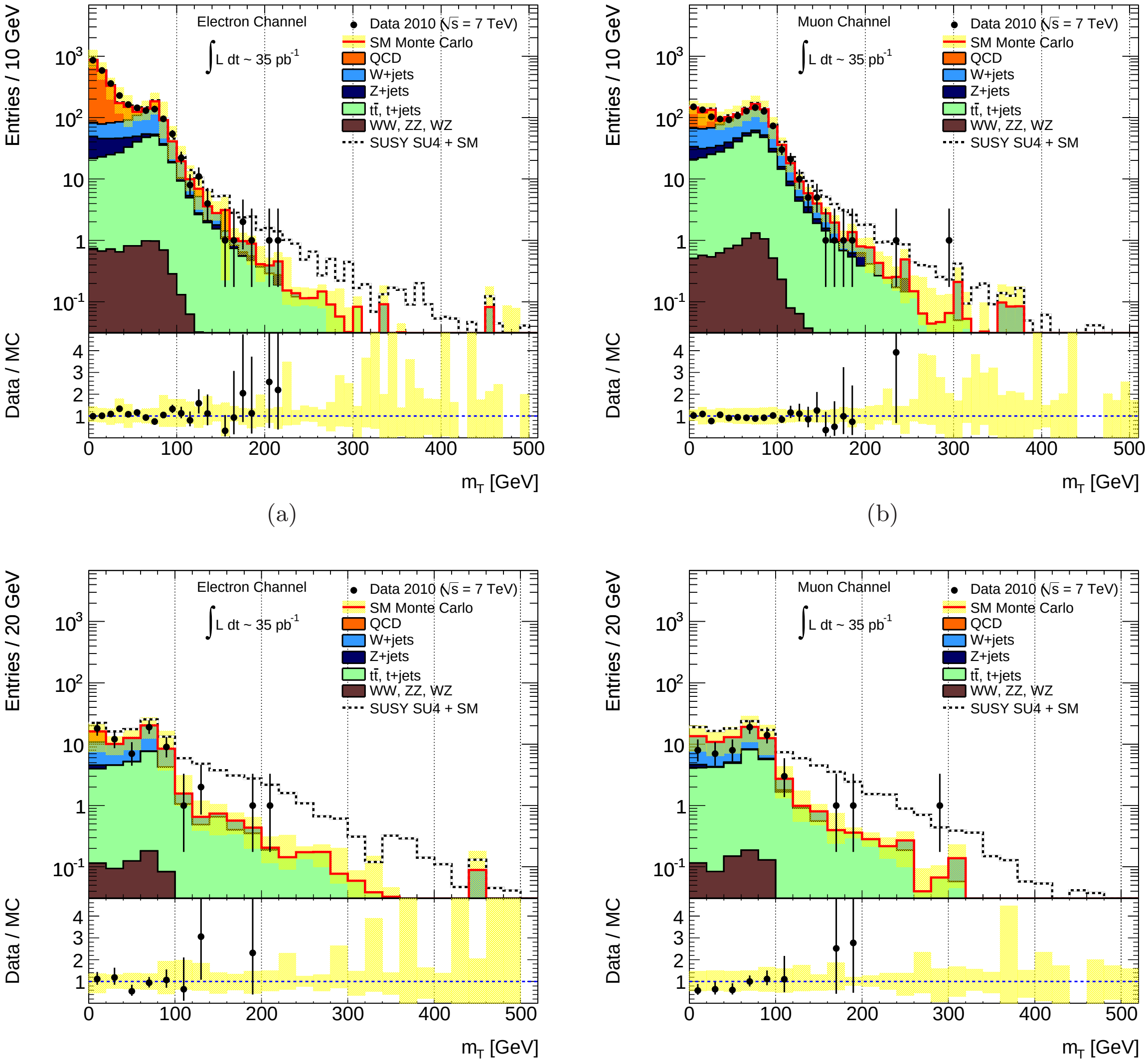


Figure 5.10: Missing transverse energy distributions for the electron (left) and muon (right) channel after a cut on the transverse mass $m_{\mathrm{T}} > 100$ GeV (cf. figure 5.11).

remaining in both selections. As defined by cut 9, the signal region contains all events with an effective mass greater than 500 GeV. In both channels one data event remains which is consistent with $1.8^{+1.0}_{-0.9}$ and $2.1^{+1.1}_{-1.0}$ predicted events respectively. Event displays of the two events can be found in section B.2 of the appendix.



(a) (b)

(c) (d)

Figure 5.11: Transverse mass distributions before ((a) and (b)) and after ((c) and (d)) applying a cut on the transverse missing energy $E_{\mathrm{T}}^{\mathrm{miss}} > 125$ GeV, for the electron (left) and muon (right) channel. As expected, the $W+$ jets and $t\bar{t}$ backgrounds peak around the W mass.

5.7.2 Global profile likelihood fit

This subsection briefly describes the combined profile likelihood fit which is used to optimize the background prediction and to determine limits on a general signal and on squark and gluino masses in the mSUGRA scenario, as presented thereafter. An extensive discussion can be found in [98] and in a dedicated internal note [122].

The combined fit is carried out simultaneously in four disjunct regions, namely the signal region SR and the control regions QR, WR and TR defined in section 5.6 and figure 5.1.

The likelihood function L is therefore the product of four Poisson distributions for the

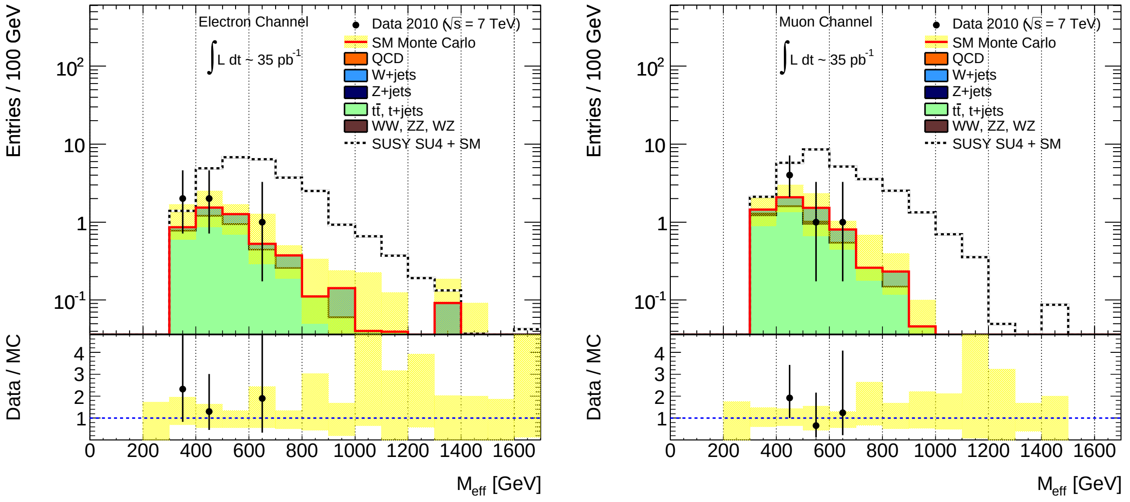


Figure 5.12: Effective mass distributions for the electron (left) and muon (right) channel with a cut on the transverse missing energy $E_{\mathrm{T}}^{\mathrm{miss}} > 125$ GeV and on the transverse mass $m_{\mathrm{T}} > 100$ GeV to better reject leptons from W decays.

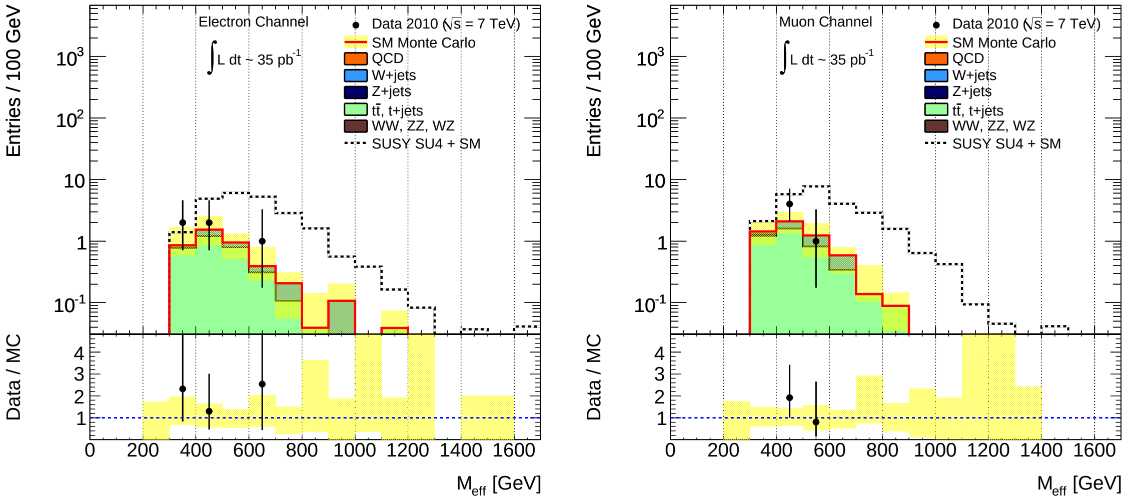


Figure 5.13: Effective mass distributions for the electron (left) and muon (right) channel with the additional cut $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$. The signal region contains the events with $m_{\mathrm{eff}} > 500$ GeV.

four regions and an additional factor C_{Syst}, which includes constraints on the systematic uncertainties:

$$L(\vec{\boldsymbol{n}}|\mu, \vec{\boldsymbol{b}}, \vec{\boldsymbol{\theta}}) \quad = \quad P_{\mathrm{SR}} \times P_{\mathrm{WR}} \times P_{\mathrm{TR}} \times P_{\mathrm{QR}} \times C_{\mathrm{Syst}}\,. \tag{5.8}$$

The vector $\vec{\boldsymbol{n}}$ contains the observed data events in the different regions and the vector $\vec{\boldsymbol{b}}$ the

Standard Model background predictions in the dedicated control regions, e.g. b_{QR} represents the number of QCD multi-jet events in the QCD control region QR. The vector $\vec{\boldsymbol{\theta}}$ comprises the so-called nuisance parameters, which parametrize the systematic uncertainties. The number μ states the signal strength, which can take values between 0, for a turned-off signal, and 1, for the nominal number of events predicted by the signal under consideration. The factor C_{Syst} consists of Gaussian probability densities for the nuisance parameters,

$$C_{Syst}(\vec{\boldsymbol{\theta}}^0, \vec{\boldsymbol{\theta}}) = \prod_{j \in \mathrm{SU}} G(\theta^0_{\mathrm{j}}, \theta_{\mathrm{j}}) \,, \tag{5.9}$$

with j running over the set of systematical uncertainties. By construction, the nuisance parameters are uncorrelated, but their impact in the different regions is not, as implied by formulas (5.10) to (5.12) below.

Extrapolation factors, denoted by $C_{\mathrm{region\ j \to region\ i}}$, relate the same background processes in other regions with respect to their dedicated control regions, e.g. $C_{\mathrm{QR \to WR}} \times b_{QR}$ equals the number of QCD multi-jet events in the WR region. The expectation values of the Poisson distributions are thus defined as

$$\begin{aligned} \lambda_S(\mu, \vec{\mathbf{b}}, \vec{\boldsymbol{\theta}}) &= \mu \cdot s + \sum_j C_{j\mathrm{R \to SR}}(\vec{\boldsymbol{\theta}}) \cdot b_{j\mathrm{R}} \,, & (5.10) \\ \lambda_i(\mu, \vec{\mathbf{b}}, \vec{\boldsymbol{\theta}}) &= \mu \cdot C_{\mathrm{SR \to iR}}(\vec{\boldsymbol{\theta}}) \cdot s + \sum_j C_{j\mathrm{R \to iR}}(\vec{\boldsymbol{\theta}}) \cdot b_{j\mathrm{R}} \,, & (5.11) \end{aligned}$$

where λ_S and λ_i are the number of expected events in the signal region and control region i respectively. The index j runs over the three control regions. The expected number of signal events in the signal region is represented by s. The extrapolation factors depend on the nuisance parameters, which tune the formers' dependency on the systematic uncertainties, i.e. they scale the relative change of the extrapolation factors due to a systematic variation:

$$C_{\mathrm{process\ j,\ region\ j \to i}} = C^{\mathrm{nominal}}_{\mathrm{process\ j,\ region\ j \to i}} \times \left(1 \,+\, \sum_k \Delta_{j,i;k} \, \theta_k \right) , \tag{5.12}$$

where $\Delta_{j,i;k}$ is the relative change in the extrapolation factor for the nuisance parameter θ_k and $C^{\mathrm{nominal}}_{\mathrm{process\ j,\ region\ j \to i}}$ is simply the ratio of expected events of process j in region i over the expected events of the same type in region j as predicted by the simulation. This definition results in the Gaussian θ-distributions to be centered at zero and to have unit width.

The profile log-likelihood ratio (LLR) is defined in the usual way as

$$\Lambda(\mu) = -2 \left(\ln L(\vec{\boldsymbol{n}}|\mu, \hat{\hat{\vec{\boldsymbol{b}}}}, \hat{\hat{\vec{\boldsymbol{\theta}}}}) - \ln L(\vec{\boldsymbol{n}}|\hat{\mu}, \hat{\vec{\boldsymbol{b}}}, \hat{\vec{\boldsymbol{\theta}}}) \right) \,, \tag{5.13}$$

where $\hat{\boldsymbol{\mu}}$, $\hat{\vec{\boldsymbol{b}}}$, $\hat{\vec{\boldsymbol{\theta}}}$ maximize the likelihood function, and $\hat{\hat{\vec{\boldsymbol{b}}}}$, $\hat{\hat{\vec{\boldsymbol{\theta}}}}$ maximize the likelihood for the specific, fixed value of μ. The LLR constitutes the basis for the test statistic, whose distribution is determined using pseudo-experiments. If the p-value of the test statistic for a given signal model (with signal strength $\mu = 1$) is smaller than 0.05, the model is said to be excluded at 95% confidence level.

When determining an upper limit on a generic unknown signal rate, the signal strength μ is set to unity and instead the LLR is minimized for a fixed number of expected signal events s. In addition, the (unknown) signal contamination in the control regions, determined by $C_{\mathrm{SR}\to\mathrm{iR}}$, is turned off. The upper limit is obtained by successively increasing the expected number of signal events until the p-value falls below the 5% threshold.

5.7.3 Exclusion limits

In the following, results using the combined profile likelihood procedure outlined in the previous subsection are presented. For the detailed discussion see [98]. Table 5.12 summarizes the results of the combined fit without assuming a particular signal model and therefore no contamination in the control regions. All quoted uncertainties include both statistical and systematical contributions. In the signal region, no excess is observed. The fitted number of background events is compatible with the observed data events and the fitted number of signal events is compatible with zero. Good agreement between the fitted values and the predictions by the simulation is observed in the electron channel. In the muon channel, the fitted WZ and top events deviate from the predictions by about −30% and +20% respectively, especially in the TR and WR selections where the determined errors are small due to a strong anti-correlation between the b-tagging and the fitted WZ and top events. However, uncertainties on the total rate of the background like global theory errors und the luminosity uncertainty cancel in the fit, but need to be considered when comparing with estimations by the simulation.

In order to further support the assumption that the simulation is able to predict the Standard Model backgrounds to a satisfactory degree, additional control regions at low $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T} are included in the fit and produce very good agreement with the observed data as detailed in [98].

Upper limits on the expected number of events s from a generic signal are derived to be for the

- Electron channel: $s < 2.2$ events at 95% confidence level (CL) and for the
- Muon channel: $s < 2.5$ events at 95% CL.

These values can be translated in model-independent effective cross-section limits

- Electron channel: $\epsilon B\sigma < 6.5 \times 10^{-2}$ pb, at 95% CL, and
- Muon channel: $\epsilon B\sigma < 7.3 \times 10^{-2}$ pb, at 95% CL,

where ϵ is the lepton-channel-dependent selection acceptance and efficiency, B the lepton-channel-dependent leptonic branching fraction, and σ is the total new physics cross-section.

Table 5.12: Fit results for the electron (top part) and muon (bottom part) channels, for an integrated luminosity of 35 pb^{-1} [98]. The results are obtained by assuming a generic signal in the signal region and no signal contamination in the control regions. Nominal MC expectations (normalized to MC cross-sections) are given for comparison.

Electron channel	SR	TR	WR	TR + WR	QR
Observed events	1	80	202	282	1464
Fitted bkg events	1.81 ± 0.75	80 ± 9	202 ± 14	282 ± 17	1464 ± 38
Fitted top events	1.34 ± 0.52	65.0 ± 12.3	31.8 ± 15.8	96.7 ± 25.9	40.1 ± 11.3
Fitted WZ events	0.47 ± 0.40	11.2 ± 4.6	160.9 ± 27.4	172.1 ± 31.2	169.7 ± 34.1
Fitted top+WZ events	1.81 ± 0.69	76.2 ± 11.8	192.6 ± 24.3	268.8 ± 32.1	209.8 ± 33.8
Fitted QCD events	$0.0^{+0.3}_{-0.0}$	3.7 ± 7.6	9.4 ± 19.6	13.0 ± 27.2	1254.2 ± 51.3
MC exp. SM events	1.75	77.76	189.15	266.91	1848.38
MC exp. top events	1.29	62.9	31.0	93.93	38.94
MC exp. WZ events	0.46	10.2	146	156.29	154.10
MC exp. QCD events	0.0	4.67	12.02	16.69	1655.34

Muon channel	SR	TR	WR	TR + WR	QR
Observed events	1	93	165	258	346
Fitted bkg events	2.25 ± 0.94	93 ± 10	165 ± 13	258 ± 16	346 ± 19
Fitted top events	1.76 ± 0.67	85.0 ± 10.5	41.8 ± 18.6	126.8 ± 25.9	49.7 ± 10.2
Fitted WZ events	0.49 ± 0.36	7.7 ± 3.3	119.8 ± 26.0	127.4 ± 29.0	71.4 ± 16.4
Fitted top+WZ events	2.25 ± 0.79	92.6 ± 9.7	161.5 ± 17.8	254.2 ± 21.0	121.0 ± 12.0
Fitted QCD events	$0.0^{+0.5}_{-0.0}$	0.3 ± 1.2	3.4 ± 12.1	3.7 ± 13.3	224.9 ± 22.3
MC exp. SM events	2.10	79.35	206.36	285.71	607.03
MC exp. top events	1.39	67.1	33.0	100.13	39.24
MC exp. WZ events	0.71	11.6	166	177.80	99.51
MC exp. QCD events	0.0	0.70	7.08	7.78	468.28

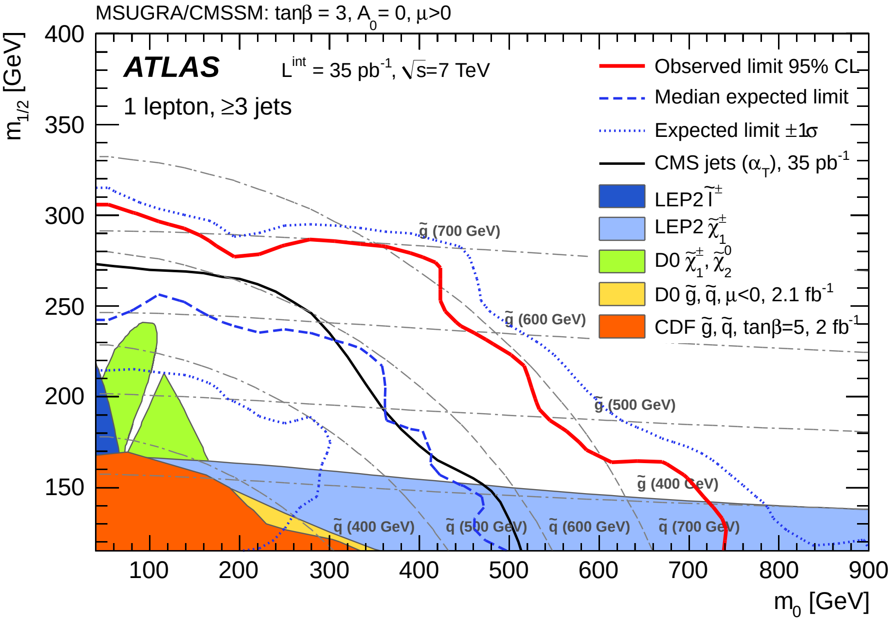


Figure 5.14: Observed and expected 95% CL exclusion limits, as well as the $\pm 1\sigma$ variation on the expected limit, in the combined electron and muon channels for an integrated luminosity of $35\,\mathrm{pb}^{-1}$ [98, 4]. Also shown are the published limits from CMS [123], CDF [124], and D0 [125, 126], and the results from the LEP experiments [127, 85]. The dashed gray lines show the squark (curved lines) and gluino (nearly horizontal lines) mass contour lines for $400, 500, 600, 700\,\mathrm{GeV}$.

Finally, the results of the analysis can be used to set upper limits on squark and gluino masses for specific SUSY scenarios. Figure 5.14 shows the expected and observed exclusion limits at 95% confidence level for the $\tan\beta = 3$ mSUGRA scenario combining both lepton channels. For each point of the mSUGRA grid, the corresponding expected signal events in the signal and all control regions together with all associated systematic uncertainties are included in the profile likelihood fit. The SUSY signal acceptance after all selection cuts ranges between approximately 0.01% and 4%, which is dominated by the leptonic branching fractions for points with $m_{1/2}$ values greater than 150 GeV. Due to the observed deficit of data events compared to the expected Standard Model background in the signal regions of both channels, the observed limit surpasses the expected limit, but stays within the 1σ band. The dashed gray lines indicate the squark and gluino mass contours. Squark masses below approximately 650 GeV can be excluded in this scenario. Gluino masses below 700 GeV can be excluded for points with similar squark and gluino masses.

5.8 Conclusion

The supersymmetry search in the one-lepton channel using the full 2010 data set shows no significant deviation from the Standard Model predictions. Several partially data-driven methods and a global profile likelihood fit are employed to validate the simulation of the various background processes. Good agreement with the observed data is found.

One data event each is observed in the signal regions of the electron and muon channel consistent with an expectation of 1.81 ± 0.75 and 2.25 ± 0.94 Standard Model events respectively. Based on these results, upper limits on the expected number of events from a new physics signal are derived to be 2.2 and 2.5 at 95% confidence level in the electron and muon channel respectively.

The exclusion limit on squark masses can be raised to about 650 GeV for the mSUGRA realization of supersymmetry with $\tan\beta = 3, A_0 = 0$ and $\mu > 0$. For similar squark and gluino masses, gluino masses below 700 GeV can be excluded in this scenario.

Part III

Improved background model for the search of new physics

Chapter 6

Motivation

The way to discover new physics beyond the Standard Model (SM) is to measure a significant deviation from the SM prediction in a signal region, that is a region of phase space where new physics is expected to appear. It is therefore essential that one can have utmost confidence in an estimate on this SM prediction in order to avoid false discoveries and overlooked signals. State-of-the-art Monte Carlo (MC) generators yield such estimates by modeling the relevant physics processes. However, systematic effects due to an imperfect detector and shortcomings in the underlying models of the MC generators lead to an insufficient description of the data. A way to verify and improve the validity of the MC prediction is to compare it with data in a signal-free control region in phase space.

Usually, measurements in the control region have to satisfy certain requirements. An observable of interest x needs to have similar physical meanings and dependencies on systematic effects in both signal and control regions. Remaining differences of these lead to systematic uncertainties which need to be estimated, but are not covered here.

Figure 6.1 shows an example of a data distribution measured in a signal-free control region with its corresponding MC prediction. To keep things simple, it is here assumed that all data originate from a single process and that a sufficiently large MC sample is available resulting in a smooth curve with negligible statistical uncertainty. Within the large systematic uncertainty data and Monte Carlo estimate are in agreement. However, the data are not fully described by the central MC prediction.

In the following chapter, a method, published in [5, 6], is presented which incorporates this systematic deviation between the data and the original model into an improved background model.

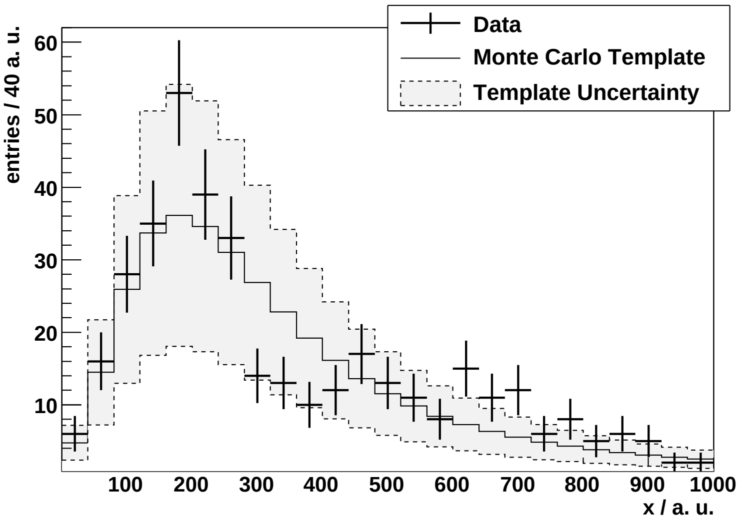


Figure 6.1: Example of a data distribution for a variable x in arbitrary units and corresponding Monte Carlo estimate with its systematic uncertainty in a signal-free control region.

Chapter 7

Absorbing systematic effects to obtain a better background model

7.1 Concept

The idea of the presented method is to reweight the MC estimate by multiplying it with an appropriate correction function. The correction function depends on a set of adjustable parameters, which are determined by fitting the modified estimate to the data in the control region. Then the same function is to be applied on the corresponding template in the signal region. In general, the statistical uncertainty grows with the number of parameters. Since this method starts with the original MC expectation, the number of parameters needed for the correction is generally smaller than in other fitting approaches, where one fits a function or splines to the data.

Another advantage of this method is that the templates in control and signal region need not have identical shapes, only the systematic effects have to influence them in a similar way. This procedure can be easily generalized to the carrying out of a combined fit of several data distributions arising from multiple sources.

In order to avoid mistaking a real signal in the control region as a systematic effect of the detector or the physics modeling, the Monte Carlo templates can be varied according to known sources for systematic deviations (energy scales, efficiencies, PDF uncertainties, etc.) before the fit is carried out. These variations yield constraints on the expected systematic deviation of the original template when there is no signal present. Only if the data is compatible with those, the fit is followed through (cf. the template uncertainty in figure 6.1).

Obviously, the crucial point lies in choosing the right correction functions which, together with the MC template, constitute the *best model* to describe the data.

7.2 Determination of the best model

Were there no systematics present, plain Monte Carlo estimates would suffice in describing the different contribution to the data. Thus, those templates form the natural starting point to determine the best background description. One modifies them with correction functions of increasing complexity, that is, with an increasing number of free parameters, thereby allowing greater flexibility for the adjustment to data, until the goodness-of-fit reaches a certain level. It is useful to take functions forming a complete basis set such as certain kinds of polynomials.

The starting model, the unaltered Monte Carlo template, shall be labeled *zeroth-order model.* The mean number of entries in each bin $\vec{\nu} = (\nu_1, \ldots, \nu_N)$ predicted by this model constitute the template histogram. Assuming that the data is independently Poisson distributed, the probability to observe the data $\vec{n} = (n_1, \ldots, n_N)$ is

$$P(\vec{n}; \vec{\nu}) = \prod_{i=1}^{N} \frac{\nu_i^{n_i}}{n_i!} e^{-\nu_i} \ . \tag{7.1}$$

To quantify the level of compatibility between $\vec{n}$ and $\vec{\nu}$, one could compute Pearson's chi-square statistic,

$$\chi^2_{\mathrm{P}} = \sum_{i=1}^{N} \frac{(n_i - \nu_i)^2}{\nu_i} \ . \tag{7.2}$$

An almost equivalent statistic is based on the likelihood ratio,

$$\lambda(\vec{\nu}) = \frac{L(\vec{\nu})}{L(\hat{\vec{\nu}})}, \tag{7.3}$$

where $L(\vec{\nu}) = P(\vec{n}; \vec{\nu})$ is the likelihood of the hypothesized model $\vec{\nu}$, and $\hat{\vec{\nu}}$ is the maximum likelihood estimator for $\vec{\nu}$, i.e. the values of $\nu_1, \ldots, \nu_N$ which maximize the likelihood. By setting the derivative of $L(\vec{\nu})$ equal to zero and solving, one easily finds

$$\hat{\nu}_i = n_i \tag{7.4}$$

for all i.

If the model $\vec{\nu}$ is correct, then Wilks' theorem [128] states that the distribution of the statistic

$$q_{\vec{\nu}} = -2 \ln \lambda(\vec{\nu}) = 2 \sum_{i=1}^{N} \left(n_i \ln \frac{n_i}{\nu_i} + \nu_i - n_i \right) \tag{7.5}$$

approaches a chi-square distribution for a sufficiently large data sample.[1] The number of degrees of freedom is the difference in the number of free parameters, often called the parameters of interest, of the two likelihood functions in equation (7.3). Here, the likelihood

---

[1] In computing $q_{\vec{\nu}}$, the logarithmic term should be skipped if $n_i = 0$.

function in the numerator has no free parameters whereas in the denominator the number of free parameters is equal to the number of bins N since the mean values ν_i are independently adjusted to the data values n_i for each bin.

In fact in many practical examples the chi-square approximation is extremely good even for moderate samples, e.g. n_i roughly a half dozen or more. Details on the regularity conditions required for Wilks' theorem to be valid are discussed in standard texts such as [129]. Pearson's χ^2_{P} and the statistic $q_{\vec{\nu}}$ are for the present example very similar. $q_{\vec{\nu}}$ will be used here.

For either goodness-of-fit statistic, χ^2_{P} or $q_{\vec{\nu}}$, one would quantify the compatibility between data and model by giving the p-value. This is the probability, under assumption of the model $\vec{\nu}$, to obtain a value of the statistic greater than or equal to that found with the actual data. That is,

$$p = \int_{q_{\vec{\nu},\mathrm{obs}}}^{\infty} f_{\chi^2}(z; N)\, dz \ , \tag{7.6}$$

where

$$f_{\chi^2}(z; N) = \frac{1}{2^{N/2}\Gamma(N/2)} z^{N/2-1} e^{-z/2} \tag{7.7}$$

is the chi-square distribution for N degrees of freedom, and Γ is the Euler gamma function.

If the compatibility between the data and the zeroth-order model turns out to be unsatisfactory, one tries to improve the level of agreement by multiplying the template with a suitable correction function $s(x; \vec{\theta})$ as suggested above. The modified prediction for the mean number of entries in the ith bin is then

$$\nu_i \to \nu_i s(x_i; \vec{\theta}) \ , \tag{7.8}$$

where x_i is the value of the abscissa variable in the centre of the ith bin and $\vec{\theta}$ stands for the set of M adjustable parameters $\vec{\theta} = (\theta_1, \ldots, \theta_M)$ of the function. One can use the same ratio as in equation (7.3) to assess the goodness-of-fit with only its numerator replaced by the likelihood of the modified prediction. Consequently, the number of degrees of freedom is reduced to $N - M$. This test can be applied on correction functions with an increasing number of adjustable parameters until the p-value exceeds a given threshold, say 0.1 or 0.2. Alternatively, one could pick the model with the highest p-value. The corresponding function together with the template would then constitute the model with the smallest complexity to be compatible with the data, using a particular set of correction functions.

One can, however, check whether the next more general model would significantly improve the data description by using another, closely related, test statistic which is based on the ratio of the likelihoods of the two models with m and $m + 1$ parameters respectively:

$$q_{m,m+1} = -2 \ln \frac{L(\hat{\vec{\theta}}^{(m)})}{L(\hat{\vec{\theta}}^{(m+1)})} \ . \tag{7.9}$$

Under the assumption that the more restrictive model in the numerator is correct and providing the data sample is not too small, $q_{m,m+1}$ will follow a chi-square distribution for one degree of freedom. For this to be true, the model in the numerator has to be a sub-model of the one in the demoninator, the same being true for $q_{\vec{\nu}}$ above.

Hence, the proposed strategy to determine an improved model for a certain background is twofold. First, an absolute goodness-of-fit using $q_{\vec{\nu}}$ establishes the minimal number of needed parameters for a sufficient data description. Then a relative test via $q_{m,m+1}$ examines whether or not a substantially better prediction can be obtained by moving on to a more general model.

It has to be noted that the choice of a set of correction functions is obviously not unique. Different situations may call for different sets and a close inspection beforehand might hint at a certain choice. In the following, ordinary polynomials are used as the basis set.

Consider two possible scenarios for measurements in a control region shown in figure 7.1. Both measurements are compatible with the uncertainty of their respective predictions. In the left plot, however, the data strongly deviate from the central prediction which hints at substantial systematic effects being present in that scenario. This assumption is further supported by a $p(q_{\vec{\nu}})$ value of only about 0.3%. On the other hand, the deviations in the right plot seem compatible with statistical fluctations as is reflected by a $p(q_{\vec{\nu}})$ value of about 56%. In both cases, the data shall now be used in an attempt to obtain a better background model following the procedure outlined above.

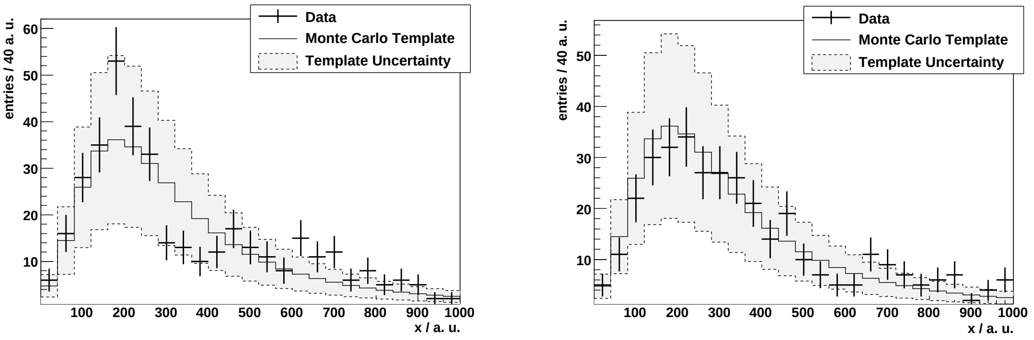


Figure 7.1: Two scenarios for measurements in a control region having the same Monte Carlo prediction.

7.2.1 First scenario: Large systematic effects

Choosing the best polynomial. Starting with the first scenario, the central prediction (zeroth-order model) is modified with polynomials of order 2, 5, and 7, displayed in increasing shades of grey in figure 7.2. The width of the bands corresponds to the respective statistical uncertainty. Table 7.1 shows both types of p-values for functions up to order 10. Using a correction function of degree 5 yields the first model with an acceptable goodness-of-fit of

0.46, which can be improved even further by including more parameters.

The compatibility peaks at a value of 0.69 when using 8 adjustable parameters. The decline for more complex models results from increasing the number of free parameters while not gaining a substantial improvement in terms of data description, as is also reflected by a high $p(q_{8,9})$ value of about 0.8. Following the proposal stated above, the model with the highest p-value is taken as the new improved background model.

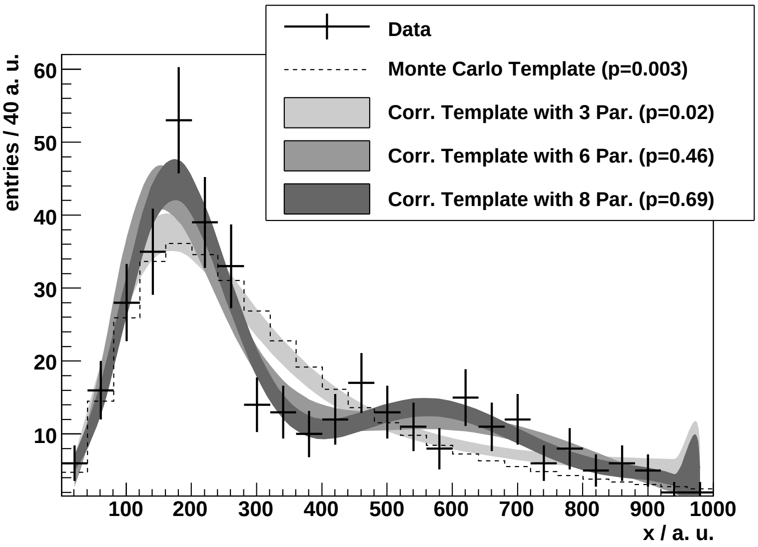


Figure 7.2: The Monte Carlo estimate for the data distribution on the left of figure 6.1 is modified with correction functions of an increasing number of parameters until a satisfactory goodness-of-fit is reached, expressed by the p-value of equation (7.6).

Alternative starting templates. Apart from the statistical error of the fit, an additional uncertainty arises from the choice of the starting template. To investigate the dependency of the corrected model on a particular shape of the original hypothesis, additional starting templates are selected from within the systematic uncertainty of the MC prediction, as shown in the left plot of figure 7.3. As was mentioned above, in the case of real data one would vary the MC prediction according to known systematic effects, thereby obtaining a set of possible starting templates. All those templates are corrected separately with the polynomial yielding the highest absolute goodness-of-fit, shown in the right plot of the figure. In addition, the true model, from which the data have been generated, is displayed as the black solid line. After correction the new models nicely converge to the true model, almost regardless of the shape of the starting template. Figure 7.4 shows the same curves divided by the true data model. The strong deviations of the original prediction (thick red line) from the constant line at unity represent the rather extreme introduced systematic effects, which call for a correction using several parameters.

Table 7.1: p-values for the test statistics defined in equations (7.5) and (7.9) for the data distribution shown in figure 7.2.

Type of correction function	$p(q_{\vec{\nu}})$	$p(q_{m,m+1})$
none (fixed to unity)	0.0027	0.15
Constant	0.0033	0.17
Linear	0.0038	0.0075
Quadratic	0.018	0.019
Cubic	0.052	0.0013
4th degree	0.33	0.072
5th degree	0.46	0.34
6th degree	0.46	0.04
7th degree	0.69	0.80
8th degree	0.63	0.21
9th degree	0.68	0.99
10th degree	0.60	-

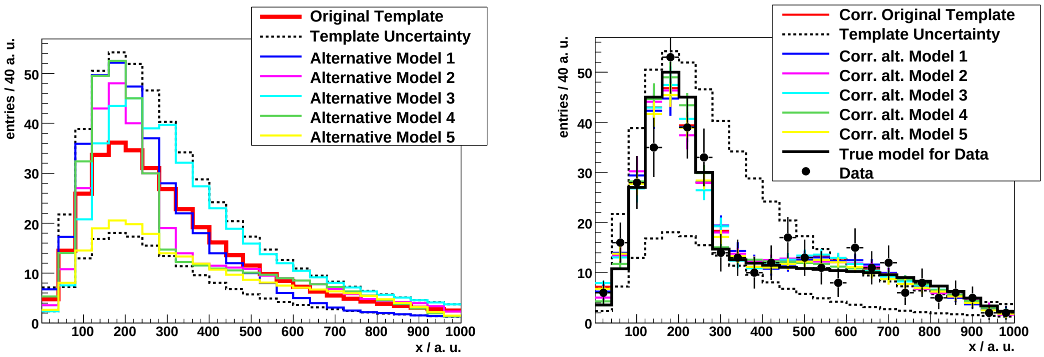


Figure 7.3: Selecting different templates within the systematic uncertainty of the original Monte Carlo prediction to estimate their influence on the corrected model.

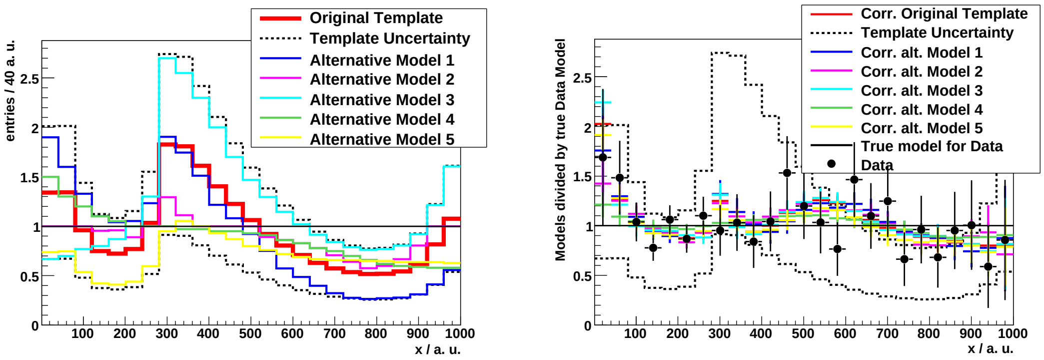


Figure 7.4: Models before (left) and after (right) correction divided by the true model for data.

New background model and its uncertainty. Assuming a flat prior probability for the different starting templates, the best estimated model is finally taken as the mean value of all corrected templates. Its total uncertainty is calculated by generating 2,000 toy data sets from this estimate and applying the proposed method on every one of them.

The bin-wise RMS of the corrected models' distribution (see figure 7.5) together with the inter-bin correlation is then taken as an estimate for the statistical error. Figure 7.6 contrasts the best estimated model with the true model for the data. The true model is nicely reproduced.

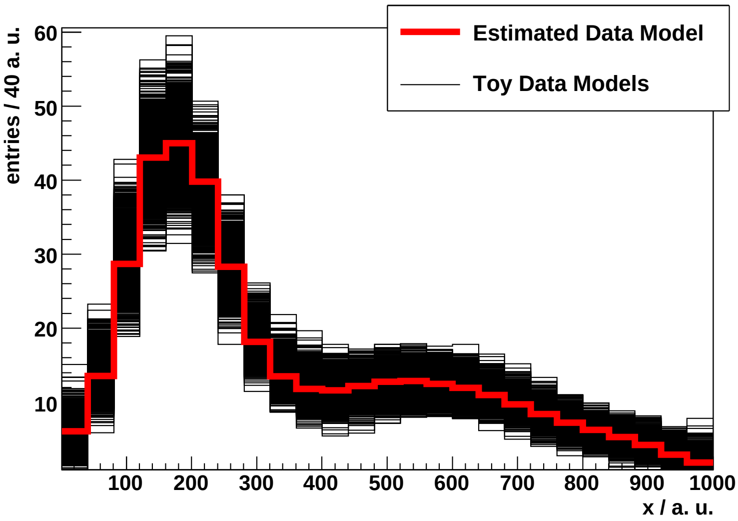


Figure 7.5: Corrected models using 2,000 sets of toy data generated from the "best estimated model" shown in red. The RMS of each bin is taken as an estimate for the statistical uncertainty of the method.

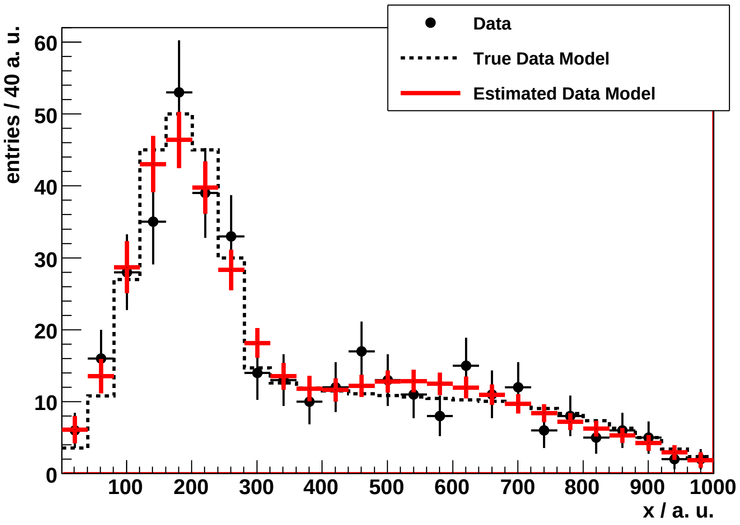


Figure 7.6: The estimated and the true model for the data agree well within the indicated uncertainty.

7.2.2 Second scenario: No systematic effects

The second scenario shown on the right of figure 7.1 shall illustrate the usefulness of the proposed method when there is apparently only little or no systematic deviation present. As before, first the central MC prediction is modified with the correction function giving the highest $p(q_{\vec{\nu}})$ value. This turns out to be a linear function (see table 7.2), which slightly tilts the template as can be seen in figure 7.7. It is noteworthy that a simple rescaling does not improve the data description ($p(q_{0,1}) = 0.96$) in this case, whereas the linear correction returns a significantly better model, as demonstrated by $p(q_{1,2}) = 0.01$.

Figures 7.8 and 7.9 show again the correction of different starting templates with respect to the true model for data. As a limiting case, this scenario has been generated without systematic effects. Hence the central prediction would constitute the best model. Still, the proposed method has the various templates converge to the true model. The offset at higher x-values results from a bias introduced by the data. Again, toy experiments are generated from the mean of the corrected models to obtain the statistical uncertainty (see figure 7.10). Figure 7.11 shows the estimated and the true model for the data, which agree well within the indicated uncertainty.

7.2.3 Extrapolation to signal region

Once the improved background model has been determined following the procedure described above, the same correction is to be applied on the Monte Carlo prediction for the background in the signal region (see also section 7.3). In addition, systematic effects associated with the transfer of the correction from control to signal region need to be considered, such as the

Table 7.2: p-values for the test statistics defined in equations (7.5) and (7.9) for the data distribution shown in figure 7.7.

Type of correction function	$p(q_{\vec{\nu}})$	$p(q_{m,m+1})$
none (fixed to unity)	0.56	0.96
Constant	0.51	0.01
Linear	0.85	0.41
Quadratic	0.84	0.40
Cubic	0.83	0.90
4th degree	0.79	0.38
5th degree	0.78	0.46
6th degree	0.76	0.13
7th degree	0.84	0.37
8th degree	0.80	0.79
9th degree	0.74	0.36
10th degree	0.67	-

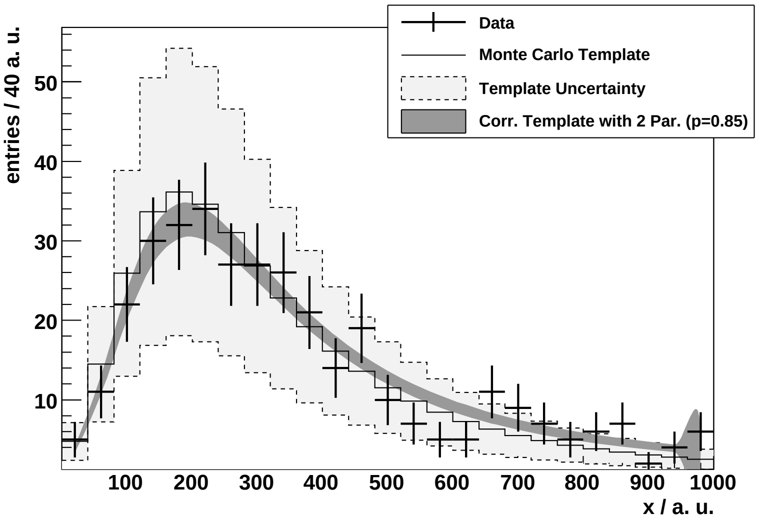


Figure 7.7: The Monte Carlo estimate for the data distribution on the right of figure 6.1 is modified with correction functions of an increasing number of parameters until a satisfactory goodness-of-fit is reached, expressed by the p-value of equation (7.6). In this scenario two parameters are sufficient, yielding a p-value of about 85%.

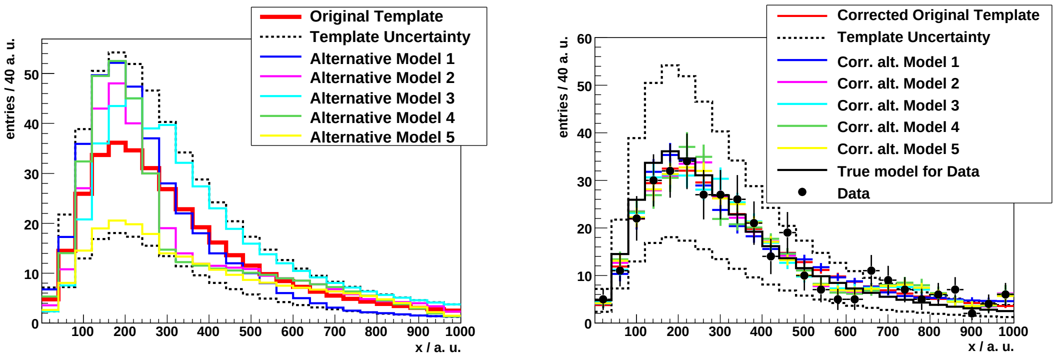


Figure 7.8: Selecting different templates within the systematic uncertainty of the original Monte Carlo prediction to estimate their influence on the corrected model.

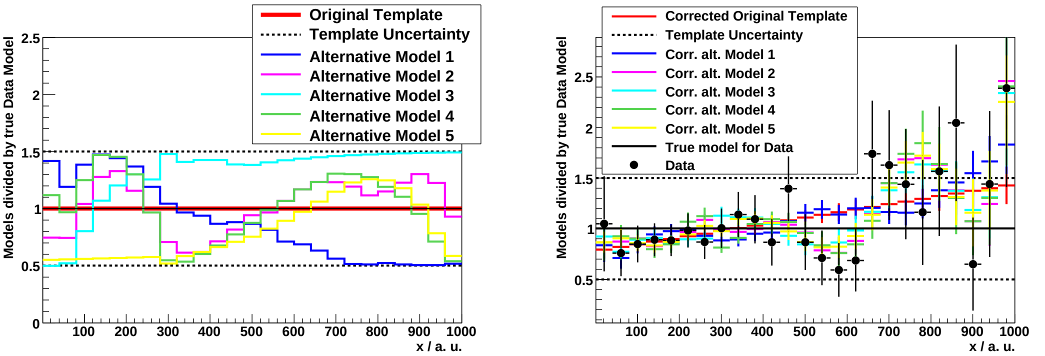


Figure 7.9: Models before (left) and after (right) correction divided by the true model for data.

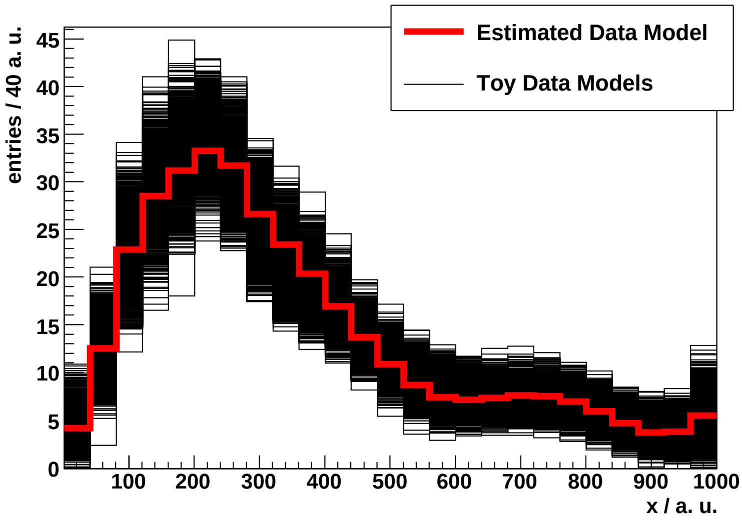


Figure 7.10: Corrected models using 2000 sets of toy data generated from the "best estimated model" shown in red. The RMS of each bin is taken as an estimate for the statistical uncertainty of the method.

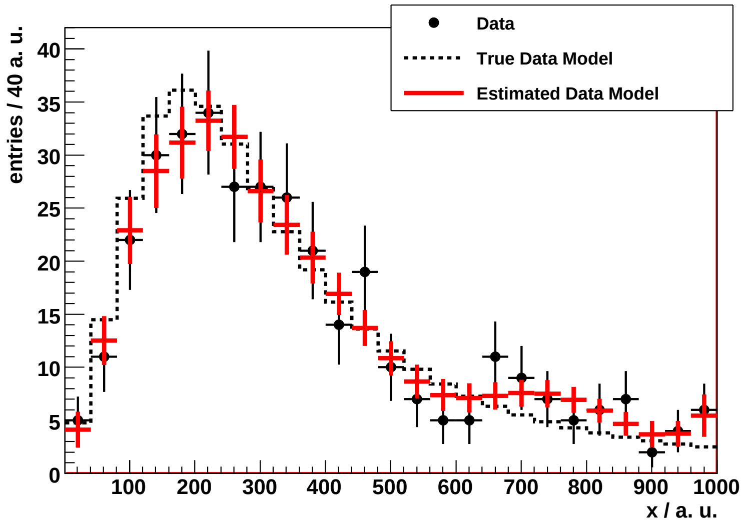


Figure 7.11: The estimated and the true model for the data agree well within the indicated uncertainty.

different influence on the shapes of the distributions by certain systematic sources. These have to be treated on a case-by-case basis and lie beyond the scope of this discussion.

After establishing a signal, one would want to quote its cross-section. It is very likely that the same correction for systematic effects would also make a reliable signal Monte Carlo prediction more realistic. The proposed method also considerably facilitates the search for a suitable control region, since differences between signal and control region due to uncontroversial features of the simulation (phase space, for instance) are accounted for automatically.

7.3 Performance and comparison with other background estimation techniques

This section deals with the performance of the proposed method with respect to other common background estimation techniques in terms of the total errors and the compensation of systematic effects.

7.3.1 Data from control region as a model

If a control region can be defined such that the shapes of the relevant background processes are practically identical to the ones in the signal region, a simple scaling of the data can be used to get a model for the background in the signal region. In a first approximation, the uncertainties of such a model are simply the square root values of the data. For simplicity, assume the efficiency of signal to control region to be unity. In this case the model determined in the control region can be taken as-is for the signal region. Consider again the second scenario discussed in subsection 7.2.2, depicted in figure 7.1. Figure 7.12 compares the relative uncertainties. The corrected model represents a much smoother and thus more realistic model than the data do – see also figure 7.11. In addition, the uncertainties of the corrected model outperform bin-wise the Poisson errors of the data. Still, one has to account for the correlation in the first case when summing events of several bins whereas the data are independently distributed.

Background estimates for a new physics search

In a search for new physics one is often interested in the high mass tails of distributions for being the most sensitive regions to discover new phenomena. Suppose this region to include all x-values greater than 600 a.u. – see figure 7.13. Table 7.3 summarizes the expected number of events and its uncertainty for the original prediction, the corrected model and when taking the data as the model. Both the data model and the corrected model have a comparable and much smaller uncertainty than the original prediction. The abovementioned correlation boosts the error of the corrected model to the level of the data uncertainty in this example.

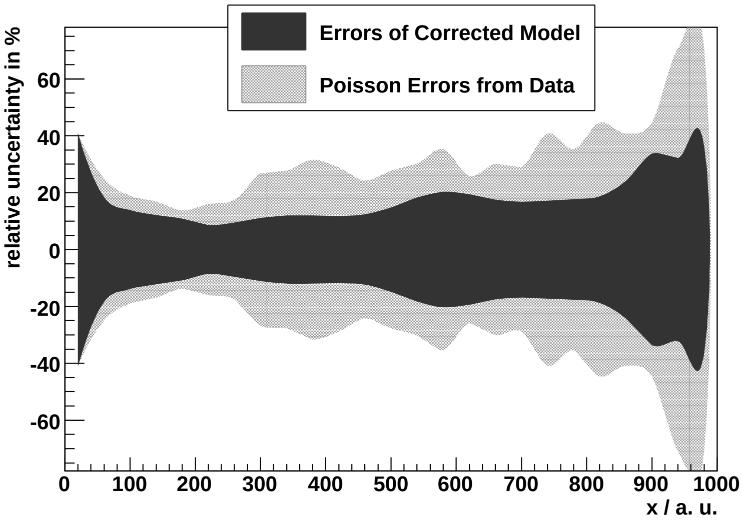


Figure 7.12: Relative total uncertainties of the data and the corrected background model as shown in figure 7.11.

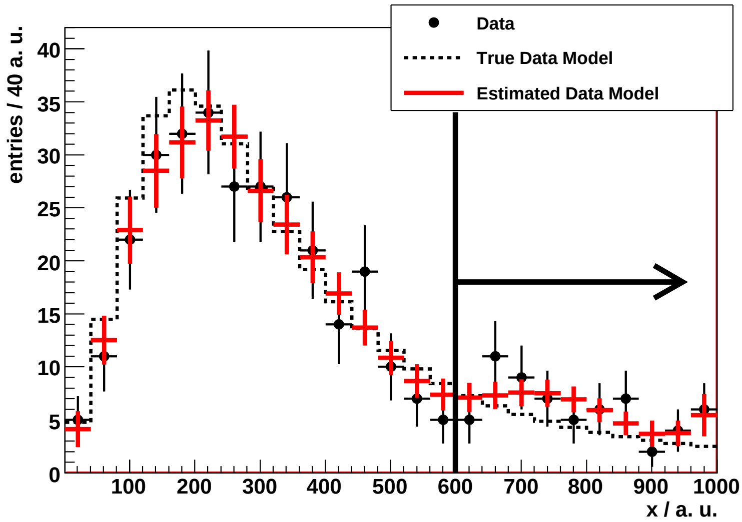


Figure 7.13: Summing events in the region $x > 600$ a.u., which is assumed to be sensitive for new physics.

Table 7.3: Number of expected events for $x > 600$ a.u. predicted by different models (cf. figure 7.13). The error of the corrected model is the same as the one from the data in this case, in general it is smaller (see figure 7.14). The MC template is identical to the true data model since no systematic effects have been introduced in this scenario.

Model	Number of expected events	Relative error
Original prediction (MC template)	43.9 ± 21.9	50%
Corrected model	59.9 ± 7.6	12.7%
Data as model	62.0 ± 7.9	12.7%

In order to obtain a general statement on how the error of the proposed method compares with the one from the data, 10,000 pseudo data sets are created from the true data model. Figure 7.14 shows the distribution of events for x-values greater than 600 a.u. Taking the data as the model, it produces an unbiased prediction of 43.92 events for the mean value with an error of 6.68 events, as expected in agreement with the true values of 43.89 and 6.63 within the statistical limitation of the sample. Applying the proposed method yields on average a value of 44.14 and a reduced error of 6.26. The mean value is slightly positively biased but only by about 4% of the quoted uncertainty. This bias originates from fits with small p-values in the percent regime. It can be reduced by either simply vetoing such fits with highest absolute p-values less than for instance 1% or by allowing fits with more than 10 parameters.

More knowledge about the true shape of the distribution can reduce the uncertainty of the method even further since fewer parameters will be needed for the adjustment of appropriate starting templates. As a limiting case, five templates which only differ in their normalization with respect to the true model are employed. The resulting distribution of expected events is also displayed in figure 7.14 as the dashed red line, demonstrating a further decrease of the error to 5.92.

Significance of a possible signal

When calculating a significance of a number of observed events given the Standard Model expectation, one can include the uncertainty on the latter by using a Bayesian prior. Since the uncertainty often arises from various sources one can assume a Gaussian distribution of the prior as stated by the central limit theorem of probability theory. Therefore, it is desirable for the error of the method to exhibit a Gaussian behaviour.

The six plots of figure 7.15 show the number of expected events for the region $x > 600$ a.u. fitted with a Gaussian function both with linear and logarithmic ordinate for the three different predictions from figure 7.14. As expected the data display an exact Poisson

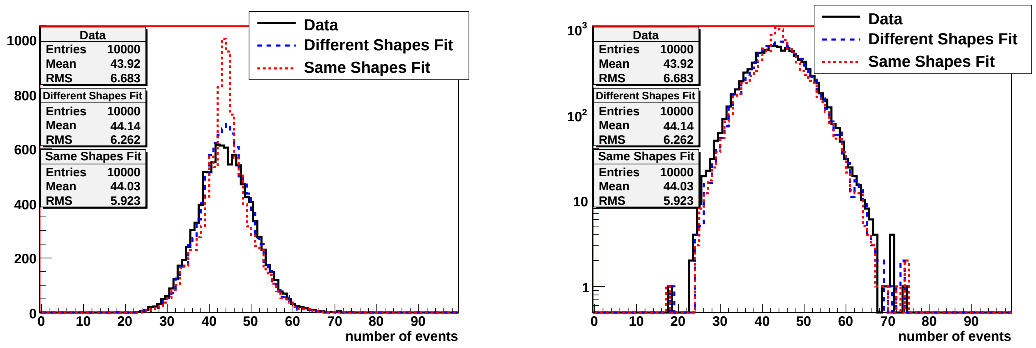


Figure 7.14: Distributions of number of expected events in the region $x > 600$ a.u. of figure 7.13 with linear and logarithmic ordinate. The proposed method using the *different shapes* from figure 7.3 yields a smaller RMS than using the data as a model. The uncertainty can be further reduced by using templates more similar to the true model, in this case only differing in the scale but having the *same shape*.

distribution which can be well approximated with a Gaussian for high enough mean values. Also the uncertainty of the proposed method using the different shapes from figure 7.3 is compatible with a normal approximation except for acceptable deviations in the central and tail regions. Using templates only differing in scale from the true background model entails an uncertainty which varies from the Gaussian fit. However, the RMS value of 5.92 events agrees nicely with the width of the fitted Gaussian. Furthermore, the Gaussian behavior is expected to improve when accounting for additional systematic effects associated with the transfer of the background model from the control to the signal region.

In order to investigate how the different errors affect the discovery potential, two toy measurements for the two regions $x > 600$ a.u. and $x > 800$ a.u. are assumed to be 99 and 52 events respectively as shown in table 7.4. High energy physics folklore considers a measurements to be a discovery if the probability, assuming only known physics, of observing data as or less likely is smaller than 2.9×10^{-7}, which corresponds to the integrated tail of a Gaussian distribution beyond five standard deviations (5σ *discovery*).

Using the data from the control region as the background model, one would claim a discovery since the significance, which is calculated by convoluting the Poisson probability for the data with the Gaussian prior function representing the systematic uncertainty of the background (see e.g. [130] and [1]), surpasses the 5σ threshold. Taking instead the predicted mean value and error of the proposed method using the different starting templates the significance grows to 5.12 and 5.29 for the two regions. It can be even further raised to 5.25 and 5.38 when using the set of *same shape templates*. The jump in significance is equivalent to an increase in luminosity of 4% and 12% for the two regions respectively when using the same shape fit results instead of the data. Thus, by using the proposed method the required integrated luminosity for a discovery is reduced. This effect gets bigger the smaller

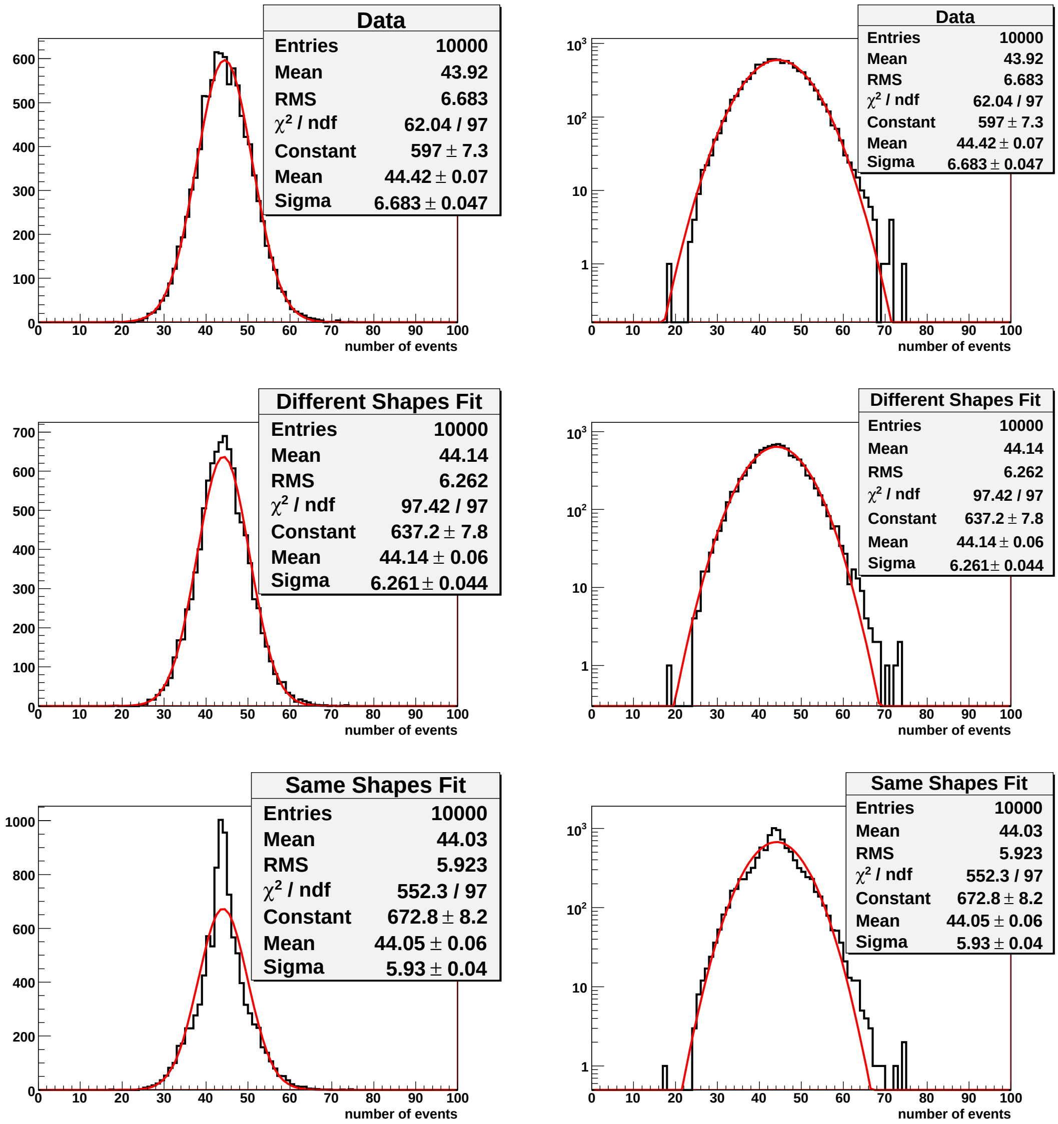


Figure 7.15: Gaussian functions fitted to the distributions of figure 7.14. Both the *data* and the *different shapes fit* display an acceptable Gaussian behaviour whereas the *same shapes fit* does not, but still does produce the same RMS value.

the inspected tail region compared to the region in x which has been used to determine the background model.

Table 7.4: The significance of a discovery can be increased by using the proposed method instead of data from the control region as a background model. The measurements of 99 and 52 events for the two regions have been chosen to allow for a 5σ discovery when using the data. The increase in significance is equivalent to a saving in luminosity as described in the text.

	$x > 600$ a.u.: 99 events		**$x > 800$ a.u.: 52 events**	
	Background:	Significance:	Background:	Significance:
Data	43.92 ± 6.68	5.01	15.62 ± 3.93	5.10
Different Shapes	44.14 ± 6.26	5.12	15.56 ± 3.60	5.30
Same Shapes	44.03 ± 5.92	5.25	15.53 ± 3.45	5.38

7.3.2 Parametrized Monte Carlo shapes

A different approach to estimate the background in the signal region lies in choosing appropriate parametric functions inspired by the shapes of the Monte Carlo estimates and fixing the parameters through a fit to data in the control region. These predictions can then be extrapolated to the signal region by scaling, provided the shapes are very similar in both regions. Unless the functional form of a certain distribution can be inferred from theoretical considerations, the use of such ad hoc parametrizations is in general, however, questionable.

Examining the Monte Carlo template in figure 7.1 one could guess that a Landau function, which depends on three parameters, might describe the background sufficiently. This is the case because the MC template is a discretized Landau function. The outcome of the fit to data of the first scenario is shown in figure 7.16. While having a much smaller uncertainty than the background data scaling method, the predicted shape is incompatible with the data in this example as reflected by a p-value $p(q_{\vec{\nu}}) = 0.002$. Obviously, if the systematic effects cannot be compensated by an adjustment of the function's parameters, the estimate will differ significantly from the true background.

7.3.3 Direct fit to data in control region

Alternatively, one could fit the data in the control region for the first scenario with a suitable polynomial function to compensate for the strong systematic effects, thereby dismissing any prior knowledge about the model. In order to determine the best-suited polynomial, the same statistical test used for the proposed method ($p(q_{\vec{\nu}})$) have been employed, selecting a

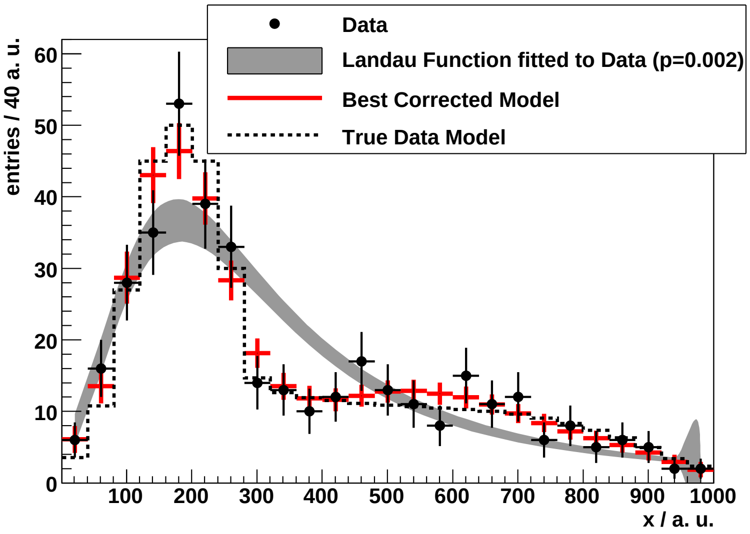


Figure 7.16: Data (black) in control region (cf. figure 7.2) fitted with a Landau function (grey band), the choice of which was inspired by the Monte Carlo estimate (dashed line in figure 7.2). The systematic effects cannot be compensated by an adjustment of the Landau function which is reflected by a poor p-value of about 0.2%.

function with 9 parameters. Table 7.5 shows the expected number of events predicted by the different approaches. The polynomial fit produces a worse result than the proposed method, which yields identical values to the data in this case, both with respect to the predicted value and its uncertainty. So even in this case, where a lot of parameters are required to adjust the templates to the true model, the proposed method outperforms a direct polynomial fit.

Table 7.5: Number of expected events predicted by different models for the scenario with large systematic effects (cf. figure 7.6). The errors of the corrected model and the polynomial fit have been calculated using 10,000 pseudo data sets. The value of the true model amounts to 376.7 events.

Model	Number of expected events	Relative error
Original prediction (MC template)	352.9 ± 176.5	50%
Data as model	380.0 ± 19.5	5.1%
Corrected model	380.0 ± 19.5	5.1%
Polynomial fit of order 8	363.6 ± 21.2	5.8%

7.4 Summary and conclusion

The underlying idea of the method presented is to correct the Monte Carlo background estimates for systematic deviations. To that end, they are multiplied with successively more complex correction functions until a statistical test reports good compatibility with data in a control region. The correction determined that way is then applied on the corresponding templates in the signal region yielding an improved background model to search for new physics.

While systematic effects are absorbed by the correction functions, the total uncertainty of the model can be reduced compared to other common methods. In order to avoid absorbing a possible signal in the fit carried out in the control region, the Monte Carlo estimates can be varied according to known systematic effects, thereby obtaining constraints on the maximal acceptable modification of the templates.

Finally, the usefulness of the proposed method is not restricted to high energy physics. It can be applied in other scientific fields where one uses data from control regions to estimate the background in a signal region and is confronted with large systematic uncertainties.

Part IV

Conclusion

Chapter 8

Summary

The search for supersymmetric particles in final states with one isolated electron or muon, substantial missing momentum and additional high-energetic jets has great potential. By analyzing data equivalent to an integrated luminosity of 1 fb^{-1} at the Large Hadron Collider's design collision energy of 14 TeV, strongly interacting particles with masses up to 1 TeV can be discovered in certain supersymmetry scenarios. At the reduced collision energy of 7 TeV, the discovery reach still extends up to masses of 700 GeV for the same integrated luminosity.

Investigating all the data collected by the ATLAS experiment in 2010, which amounts to approximately 35 pb^{-1}, no significant deviations from the Standard Model predictions have been found. In general, good agreement between the data and the simulation has been observed. Several control measurements have been carried out to verify the most important contributions to the background. In particular, sources for systematic deviations from the prediction of W boson and top processes have been investigated. An additional measurement based on the charge asymmetry of W boson production has been employed to further support the validity of the simulation. In the absence of any signal, supersymmetric partners of quarks and gluons with masses below 700 GeV have been excluded at 95% confidence level for a particular scenario within the mSUGRA realization of supersymmetry and for equal squark and gluino masses.

In a search for new physics one needs to have utmost confidence in the Standard Model background estimations. A novel method has been proposed which, by combining information from the data and the simulation, yields an improved background model. The method is based on a rigorous statistical procedure that accounts for systematic deviations from the simulation and avoids the introduction of ad hoc parametrizations. It is particularly useful when substantial deviations from the predictions are observed, and its field of application is not restricted to high energy physics.

Appendix A

SUSY masses in the $\sqrt{s} = 14$ TeV prospect studies

Table A.1 contains the masses of the SUSY particles for the different SUSY benchmark signals used in chapter 3.

Table A.1: Masses in GeV for the SUSY benchmark signals used in chapter 3 [1]. The left- and right-handed superpartners of the third generation mix strongly to give two mass eigenstates denoted by arabic number subscripts, the masses of which differ substantially. The superpartners of the first and second generation are nearly degenerate.

Particle	SU1	SU2	SU3	SU4	SU6	SU8.1
$\tilde{d}_L$	764.90	3564.13	636.27	419.84	870.79	801.16
$\tilde{u}_L$	760.42	3563.24	631.51	412.25	866.84	797.09
$\tilde{b}_1$	697.90	2924.80	575.23	358.49	716.83	690.31
$\tilde{t}_1$	572.96	2131.11	424.12	206.04	641.61	603.65
$\tilde{d}_R$	733.53	3576.13	610.69	406.22	840.21	771.91
$\tilde{u}_R$	735.41	3574.18	611.81	404.92	842.16	773.69
$\tilde{b}_2$	722.87	3500.55	610.73	399.18	779.42	743.09
$\tilde{t}_2$	749.46	2935.36	650.50	445.00	797.99	766.21
$\tilde{e}_L$	255.13	3547.50	230.45	231.94	411.89	325.44
$\tilde{\nu}_e$	238.31	3546.32	216.96	217.92	401.89	315.29
$\tilde{\tau}_1$	146.50	3519.62	149.99	200.50	181.31	151.90
$\tilde{\nu}_\tau$	237.56	3532.27	216.29	215.53	358.26	296.98
$\tilde{e}_R$	154.06	3547.46	155.45	212.88	351.10	253.35
$\tilde{\tau}_2$	256.98	3533.69	232.17	236.04	392.58	331.34
$\tilde{g}$	832.33	856.59	717.46	413.37	894.70	856.45
$\tilde{\chi}_1^0$	136.98	103.35	117.91	59.84	149.57	142.45
$\tilde{\chi}_2^0$	263.64	160.37	218.60	113.48	287.97	273.95
$\tilde{\chi}_3^0$	466.44	179.76	463.99	308.94	477.23	463.55
$\tilde{\chi}_4^0$	483.30	294.90	480.59	327.76	492.23	479.01
$\tilde{\chi}_1^+$	262.06	149.42	218.33	113.22	288.29	274.30
$\tilde{\chi}_2^+$	483.62	286.81	480.16	326.59	492.42	479.22
h^0	115.81	119.01	114.83	113.98	116.85	116.69
H^0	515.99	3529.74	512.86	370.47	388.92	430.49
A^0	512.39	3506.62	511.53	368.18	386.47	427.74
H^+	521.90	3530.61	518.15	378.90	401.15	440.23
t	175.00	175.00	175.00	175.00	175.00	175.00

Appendix B

Additional material for full 2010 data analysis

B.1 Simulated samples

The following samples, taken from the official mc09 Monte Carlo production, have been used in the full data analysis presented in chapter 5.2.

Tables B.1 and B.2 list all $V+$ jets samples together with the dataset number, the employed generator, the corresponding cross sections as well as the numbers of generated events. Table B.3 contains the same figures for top, QCD multi-jet and and di-boson samples.

B.2 Event displays of signal region events

Figures B.1 and B.2 and tables B.4 and B.5 show the event displays and main event characteristics of the two events in the electron and muon signal region of the full data analysis. respectively.

Table B.1: $W+$ jets Monte Carlo samples with sample ID, generator, cross-section (σ) times branching ratio (BR), k–factor and number of generated events (N_{gen}).

Sample ID	Name	Generator	σ × BR [pb]	k–factor	N_{gen}
107680	WenuNp0_pt20	AlpgenJimmy	$6.9{\cdot}10^3$	1.20	1381931
107681	WenuNp1_pt20	AlpgenJimmy	$1.3{\cdot}10^3$	1.20	258408
107682	WenuNp2_pt20	AlpgenJimmy	$3.8{\cdot}10^2$	1.20	188896
107683	WenuNp3_pt20	AlpgenJimmy	$1.0{\cdot}10^2$	1.20	50477
107684	WenuNp4_pt20	AlpgenJimmy	$2.5{\cdot}10^1$	1.20	12991
107685	WenuNp5_pt20	AlpgenJimmy	6.9	1.20	3449
107690	WmunuNp0_pt20	AlpgenJimmy	$6.9{\cdot}10^3$	1.20	1386038
107691	WmunuNp1_pt20	AlpgenJimmy	$1.3{\cdot}10^3$	1.20	255909
107692	WmunuNp2_pt20	AlpgenJimmy	$3.8{\cdot}10^2$	1.20	187860
107693	WmunuNp3_pt20	AlpgenJimmy	$1.0{\cdot}10^2$	1.20	50877
107694	WmunuNp4_pt20	AlpgenJimmy	$2.5{\cdot}10^1$	1.20	12991
107695	WmunuNp5_pt20	AlpgenJimmy	6.9	1.20	3498
107700	WtaunuNp0_pt20	AlpgenJimmy	$6.9{\cdot}10^3$	1.20	1365491
107701	WtaunuNp1_pt20	AlpgenJimmy	$1.3{\cdot}10^3$	1.20	254753
107702	WtaunuNp2_pt20	AlpgenJimmy	$3.8{\cdot}10^2$	1.20	188446
107703	WtaunuNp3_pt20	AlpgenJimmy	$1.0{\cdot}10^2$	1.20	50472
107704	WtaunuNp4_pt20	AlpgenJimmy	$2.5{\cdot}10^1$	1.20	12996
107705	WtaunuNp5_pt20	AlpgenJimmy	6.9	1.20	3998
106280	WbbNp0_pt20	AlpgenJimmy	3.2	1.20	6499
106281	WbbNp1_pt20	AlpgenJimmy	2.6	1.20	5500
106282	WbbNp2_pt20	AlpgenJimmy	1.4	1.20	2997
106283	WbbNp3_pt20	AlpgenJimmy	0.6	1.20	1500

Table B.2: $Z+$ jets Monte Carlo samples with sample ID, generator, cross-section (σ) times branching ratio (BR), k–factor and number of generated events (N_{gen}).

Sample ID	Name	Generator	σ × BR [pb]	k–factor	N_{gen}
107650	ZeeNp0_pt20	AlpgenJimmy	$6.6{\cdot}10^2$	1.25	304216
107651	ZeeNp1_pt20	AlpgenJimmy	$1.3{\cdot}10^2$	1.25	63440
107652	ZeeNp2_pt20	AlpgenJimmy	$4.0{\cdot}10^1$	1.25	19497
107653	ZeeNp3_pt20	AlpgenJimmy	$1.1{\cdot}10^1$	1.25	5499
107654	ZeeNp4_pt20	AlpgenJimmy	2.9	1.25	1499
107655	ZeeNp5_pt20	AlpgenJimmy	$7.5{\cdot}10^{-1}$	1.25	500
107660	ZmumuNp0_pt20	AlpgenJimmy	$6.6{\cdot}10^2$	1.25	303947
107661	ZmumuNp1_pt20	AlpgenJimmy	$1.3{\cdot}10^2$	1.25	62996
107662	ZmumuNp2_pt20	AlpgenJimmy	$4.0{\cdot}10^1$	1.25	18993
107663	ZmumuNp3_pt20	AlpgenJimmy	$1.1{\cdot}10^1$	1.25	5497
107664	ZmumuNp4_pt20	AlpgenJimmy	2.9	1.25	1499
107665	ZmumuNp5_pt20	AlpgenJimmy	$7.5{\cdot}10^{-1}$	1.25	499
107670	ZtautauNp0_pt20	AlpgenJimmy	$6.6{\cdot}10^2$	1.25	303359
107671	ZtautauNp1_pt20	AlpgenJimmy	$1.3{\cdot}10^2$	1.25	63481
107672	ZtautauNp2_pt20	AlpgenJimmy	$4.0{\cdot}10^1$	1.25	19492
107673	ZtautauNp3_pt20	AlpgenJimmy	$1.1{\cdot}10^1$	1.25	5497
107674	ZtautauNp4_pt20	AlpgenJimmy	2.9	1.25	1499
107675	ZtautauNp5_pt20	AlpgenJimmy	$7.5{\cdot}10^{-1}$	1.25	499

Table B.3: Top, QCD multi-jet and di-boson Monte Carlo samples with sample ID, generator, cross-section (σ) times branching ratio (BR), k–factor and number of generated events (N_{gen}).

Sample ID	Name	Generator	$\sigma \times$ BR [pb]	k–factor	N_{gen}
105200	T1	MC@NLO Jimmy	$8.9{\cdot}10^{1}$		999387
105204	TTbar_FullHad	MC@NLO Jimmy	$7.1{\cdot}10^{1}$		149889
108340	st_tchan_enu	MC@NLO Jimmy	7.0		9993
108341	st_tchan_munu	MC@NLO Jimmy	7.0		9997
108342	st_tchan_taunu	MC@NLO Jimmy	7.0		10000
108343	st_schan_enu	MC@NLO Jimmy	$4.7{\cdot}10^{-1}$		9950
108344	st_schan_munu	MC@NLO Jimmy	$4.7{\cdot}10^{-1}$		9996
108345	st_schan_taunu	MC@NLO Jimmy	$4.7{\cdot}10^{-1}$		9996
108346	st_Wt	MC@NLO Jimmy	$1.3{\cdot}10^{1}$		14995
105009	J0 jetjet	Pythia	$9.7{\cdot}10^{9}$		1399184
105010	J1 jetjet	Pythia	$6.7{\cdot}10^{8}$		1395383
105011	J2 jetjet	Pythia	$4.1{\cdot}10^{7}$		1398078
105012	J3 jetjet	Pythia	$2.2{\cdot}10^{6}$		1397430
105013	J4 jetjet	Pythia	$8.8{\cdot}10^{4}$		1397401
105014	J5 jetjet	Pythia	$2.3{\cdot}10^{3}$		1391612
105015	J6 jetjet	Pythia	$3.4{\cdot}10^{1}$		1347654
109276	J0 jetjet_1muon	Pythia	$8.4{\cdot}10^{5}$		1996528
109277	J1 jetjet_1muon	Pythia	$8.2{\cdot}10^{5}$		1769626
109278	J2 jetjet_1muon	Pythia	$2.2{\cdot}10^{5}$		399911
109279	J3 jetjet_1muon	Pythia	$2.8{\cdot}10^{4}$		199927
109280	J4 jetjet_1muon	Pythia	$2.0{\cdot}10^{3}$		199857
109281	J5 jetjet_1muon	Pythia	$7.1{\cdot}10^{1}$		499174
105985	WW	Herwig	$1.7{\cdot}10^{1}$		249837
105986	ZZ	Herwig	$1.3{\cdot}10^{1}$		249725
105987	WZ	Herwig	$5.5{\cdot}10^{1}$		249830

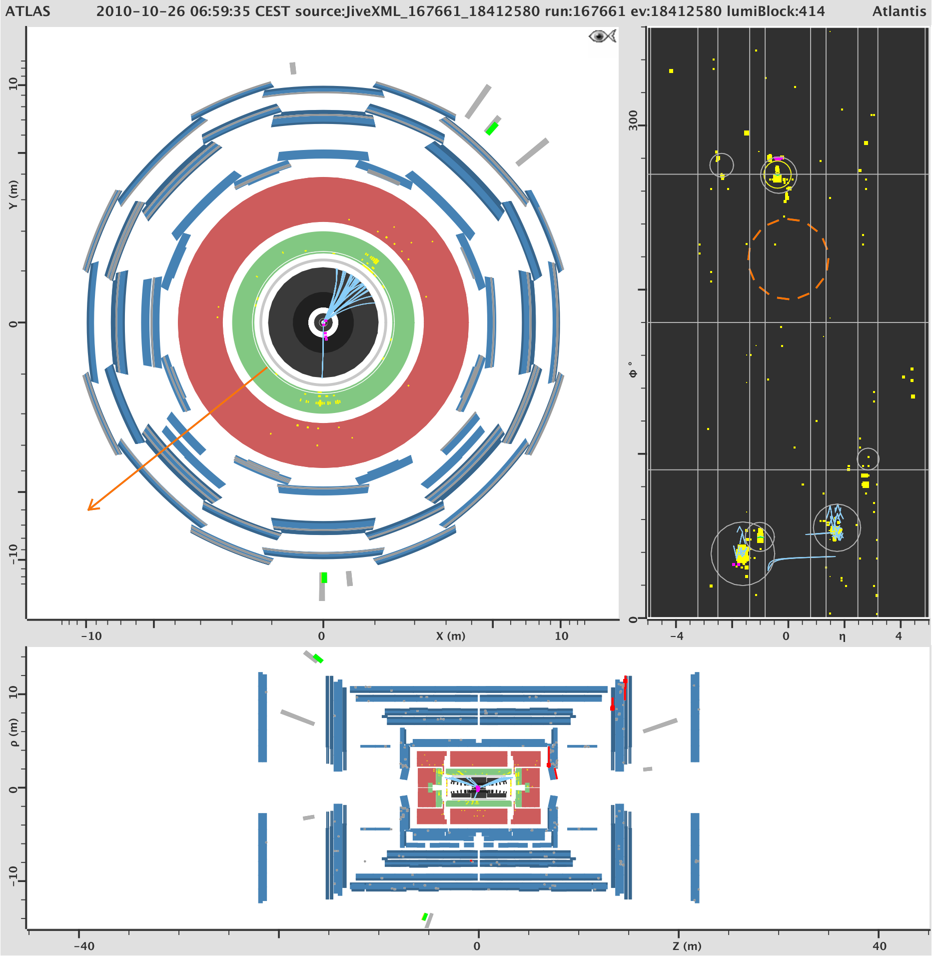


Figure B.1: This electron event is in the signal region (from [98]). The electron emits hard bremsstrahlung in the inner detector.

Table B.4: Event characteristics of the signal region event in the electron channel (from [98]).

$E_{\mathrm{T}}^{\mathrm{miss}}$	Lepton p_{T}	N(jets)	Jet1 p_{T}	Jet2 p_{T}	m_{T}	m_{eff}
303.1	27.8	4	202.0	108.1	182.6	696.3

Table B.5: Event characteristics of the signal region event in the muon channel (from [98]).

$E_{\mathrm{T}}^{\mathrm{miss}}$	Lepton p_{T}	N(jets)	Jet1 p_{T}	Jet2 p_{T}	m_{T}	m_{eff}
140.2	30.2	4	161.5	133.2	108.5	509.6

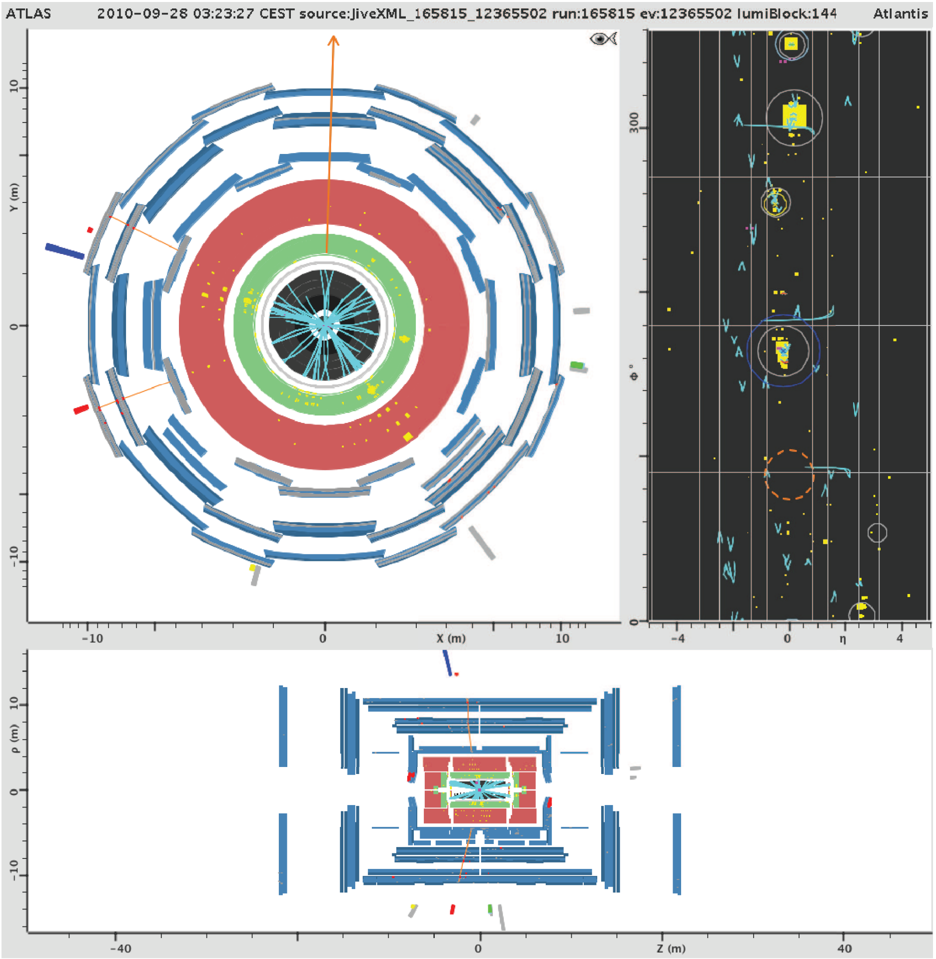


Figure B.2: This muon event is in the signal region (from [98]). It has a second muon track but not muon object with $p_{\mathrm{T}} = 11\,\mathrm{GeV}$; one electron in event passing phase space and electron-identification requirements but touching a calorimeter region with a faulty read-out and is therefore dropped.

Bibliography

[1] The ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, `arXiv:0901.0512 [hep-ex]`.

[2] The ATLAS Collaboration, *Early supersymmetry searches with jets, missing transverse momentum and one or more leptons with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-066, CERN, Geneva, Jul., 2010.

[3] S. Horner, *Inclusive Searches for Supersymmetry with Leptons with the ATLAS detector*, Proceedings of Science **PoS(ICHEP 2010)420** (2010) . `http://pos.sissa.it/archive/conferences/120/420/ICHEP%202010_420.pdf`.

[4] The ATLAS Collaboration, *Search for Supersymmetry Using Final States with One Lepton, Jets, and Missing Transverse Momentum with the ATLAS Detector in* $\sqrt{s} = 7$ *TeV pp Collisions*, Phys. Rev. Lett. **106** (2011) no. 13, 131802.

[5] S. Caron, G. Cowan, E. Gross, S. Horner, and J. E. Sundermann, *Absorbing systematic effects to obtain a better background model in a search for new physics*, Journal of Instrumentation **4** (2009) no. 10, P10009. `http://stacks.iop.org/1748-0221/4/i=10/a=P10009`.

[6] S. Caron, G. Cowan, E. Gross, S. Horner, and J. E. Sundermann, *Absorbing systematic effects to obtain a better background model in a search for new physics*, Proceedings of Science **PoS(ACAT2010)050** (2010) . `http://pos.sissa.it/archive/conferences/093/050/ACAT2010_050.pdf`.

[7] D. Griffiths, *Introduction to Elementary Particles.* John Wiley & Sons, 2^{nd} ed., 2008.

[8] F. Mandl and G. G. Shaw, *Quantum Field Theory.* John Wiley & Sons, revised ed., 1993.

[9] S. P. Martin, *A Supersymmetry Primer*, `arXiv:hep-ph/9709356`.

[10] I. J. R. Aitchison, *Supersymmetry and the MSSM: An Elementary introduction*, `arXiv:hep-ph/0505105`.

[11] J. R. Ellis, *Supersymmetry for Alp Hikers*, `arXiv:hep-ph/0203114`.

[12] K. Nakamura et al. (Particle Data Group), *Review of particle physics*, J. Phys. **G37** (2010) 075021. `http://pdg.lbl.gov/`.

[13] K. Nakamura et al. (Particle Data Group), *Particle Listings - Neutrino Properties*, `http://pdg.lbl.gov/2010/listings/rpp2010-list-neutrino-prop.pdfh`, last accessed Apr. 12th, 2011.

[14] S. Nussinov and R. Shrock, *Upper Limits on a Possible Gluon Mass*, Phys. Rev. **D82** (2010) 034031, `arXiv:1005.0850 [hep-ph]`.

[15] R. Alkofer and J. Greensite, *Quark Confinement: The Hard Problem of Hadron Physics*, J.Phys.G **G34** (2007) S3, `arXiv:hep-ph/0610365 [hep-ph]`.

[16] G. Gabrielse, D. Hanneke, T. Kinoshita, M. Nio, and B. Odom, *New Determination of the Fine Structure Constant from the Electron g Value and QED*, Phys. Rev. Lett. **97** (2006) no. 3, 030802.

[17] G. Gabrielse, D. Hanneke, T. Kinoshita, M. Nio, and B. Odom, *Erratum: New Determination of the Fine Structure Constant from the Electron g Value and QED [Phys. Rev. Lett. 97, 030802 (2006)]*, Phys. Rev. Lett. **99** (2007) no. 3, 039902.

[18] C. T. Hill and E. H. Simmons, *Strong dynamics and electroweak symmetry breaking*, Phys. Rept. **381** (2003) 235–402, `arXiv:hep-ph/0203079`.

[19] H. Georgi and S. L. Glashow, *Unity of All Elementary-Particle Forces*, Phys. Rev. Lett. **32** (1974) no. 8, 438–441.

[20] *The Official String Web Site*, `http://www.superstringtheory.com/`, last accessed Apr. 4th, 2011.

[21] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys. Rept. **405** (2005) 279–390, `arXiv:hep-ph/0404175`.

[22] G. Bertone, *The moment of truth for WIMP Dark Matter*, Nature **468** (2010) 389–393, `arXiv:1011.3532 [astro-ph.CO]`.

[23] H. Miyazawa, *Baryon Number Changing Currents*, Progress of Theoretical Physics **36** (1966) no. 6, 1266–1276.

[24] H. Miyazawa, *Spinor Currents and Symmetries of Baryons and Mesons*, Phys. Rev. **170** (1968) no. 5, 1586–1590.

[25] J. Wess and B. Zumino, *Supergauge transformations in four dimensions*, Nuclear Physics B **70** (1974) 39–50.

[26] K. Nakamura et al. (Particle Data Group), *Particle Listings - Supersymmetric Particles*, `http://pdg.lbl.gov/2010/listings/rpp2010-list-supersymmetric-part-searches.pdf`, last accessed Apr. 12th, 2011.

[27] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1988) 1–91, `arXiv:hep-ph/0409313`.

[28] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept. Prog. Phys. **70** (2007) 89, `arXiv:hep-ph/0611148`.

[29] Martin-Stirling-Thorne-Watt Parton Distribution Functions, `http://projects.hepforge.org/mstwpdf/`, last accessed Mar. 22nd, 2011.

[30] G. Watt, `http://projects.hepforge.org/mstwpdf/plots/plots.html`, last accessed Mar. 22nd, 2011.

[31] M. A. Dobbs et al., *Les Houches guidebook to Monte Carlo generators for hadron collider physics*, `arXiv:hep-ph/0403045`.

[32] T. Sjostrand, *Monte Carlo Generators*, `arXiv:hep-ph/0611247`.

[33] M. Dobbs and J. B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, Computer Physics Communications **134** (2001) no. 1, 41 – 46.

[34] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nuclear Physics B **126** (1977) no. 2, 298 – 318.

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

[36] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, `arXiv:hep-ph/0206293`.

[37] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, `arXiv:hep-ph/0204244`.

[38] S. Frixione, P. Nason, and B. R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, JHEP **08** (2003) 007, `arXiv:hep-ph/0305252`.

[39] S. Frixione and B. R. Webber, *The MC@NLO 3.2 event generator*, `arXiv:hep-ph/0601192`.

[40] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31–145.

[41] G. Corcella et al., *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, `arXiv:hep-ph/0011363`.

[42] G. Corcella et al., *HERWIG 6.5 release note*, `arXiv:hep-ph/0210213`.

[43] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C72** (1996) 637–646, `arXiv:hep-ph/9601371`. `http://projects.hepforge.org/jimmy/`.

[44] M. Bahr et al., *Herwig++ 2.3 Release Note*, `arXiv:0812.0529 [hep-ph]`.

[45] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. **C53** (2008) 473–500, `arXiv:0706.2569 [hep-ph]`.

[46] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. **D69** (2004) 094008, `arXiv:hep-ph/0312266`.

[47] F. E. Paige, S. D. Protopopescu, H. Baer, and X. Tata, *ISAJET 7.69: A Monte Carlo event generator for p p, anti-p p, and e+ e- reactions*, `arXiv:hep-ph/0312045`.

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

[49] W. Beenakker et al., *The Production of charginos / neutralinos and sleptons at hadron colliders*, Phys. Rev. Lett. **83** (1999) 3780–3783, `arXiv:hep-ph/9906298`.

[50] Prospino2, `http://www.ph.ed.ac.uk/~tplehn/prospino/`, last accessed Feb. 21st, 2011.

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

[52] E. Richter-Was, D. Froidevaux, and L. Poggioli, *ATLFAST 2.0 a fast simulation package for ATLAS*, Tech. Rep. ATL-PHYS-98-131, CERN, Geneva, Nov., 1998.

[53] G. Kane and A. Pierce, eds., *Perspectives on LHC Physics.* World Scientific Publishing Company, Oxford, 1st ed., Jun, 2008.

[54] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001.

[55] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08003.

[56] *LEP Working Groups*, `http://delphiwww.cern.ch/offline/lepwgs.html`, last accessed Mar. 17th, 2011.

[57] *CERN - European Organization for Nuclear Research*, `http://www.cern.ch/`, last accessed Mar. 17th, 2011.

[58] L. Rossi, *Superconductivity: its role, its success and its setbacks in the Large Hadron Collider of CERN*, Superconductor Science and Technology **23** (2010) no. 3, 034001. `http://stacks.iop.org/0953-2048/23/i=3/a=034001`.

[59] The CMS Collaboration, *The CMS experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08004.

[60] The ALICE Collaboration, *The ALICE experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08002.

[61] The LHCb Collaboration, *The LHCb detector at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08005.

[62] The TOTEM Collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08007.

[63] The LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08006.

[64] J.-L. Caron, *Overall view of LHC experiments.*, AC Collection. Legacy of AC. Pictures from 1992 to 2002., May, 1998.

[65] Hands on Particle Physics Masterclasses, `http://physicsmasterclasses.org/`, last accessed Mar. 17th, 2011.

[66] J. Pequenao, *Computer generated image of the whole ATLAS detector*, Mar., 2008.

[67] The ATLAS Collaboration, *Performance of the ATLAS Detector using First Collision Data*, JHEP **09** (2010) 056, `arXiv:1005.5254 [hep-ex]`.

[68] The ATLAS Collaboration, *Studies of the performance of the ATLAS detector using cosmic-ray muons*, `arXiv:1011.6665 [physics.ins-det]`.

[69] The ATLAS Collaboration, *Alignment of the ATLAS Inner Detector Tracking System with 2010 LHC proton-proton collisions at sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2011-012, CERN, Geneva, Mar., 2011.

[70] J. Pequenao, *Computer generated image of the ATLAS inner detector*, Mar., 2008.

[71] J. Pequenao, *Computer generated image of the ATLAS calorimeter*, Mar., 2008.

[72] The ATLAS Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, Eur. Phys. J. **C70** (2010) 1193–1236, `arXiv:1007.5423 [physics.ins-det]`.

[73] The ATLAS Collaboration, *Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions*, Eur. Phys. J. **C70** (2010) 723–753, `arXiv:0912.2642 [physics.ins-det]`.

[74] J. Pequenao, *Computer generated image of the ATLAS Muons subsystem*, Mar., 2008.

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

[76] The ATLAS Collaboration, *Performance of Jet Algorithms in the ATLAS Detector*, Tech. Rep. ATL-PHYS-INT-2010-129 (restricted), CERN, Geneva, Dec., 2010.

[77] The ATLAS Collaboration, *Electron and photon reconstruction and identification in ATLAS: expected performance at high energy and results at 900 GeV*, Tech. Rep. ATLAS-CONF-2010-005, CERN, Geneva, Jun., 2010.

[78] The ATLAS Collaboration, *Determination of the muon reconstruction efficiency in ATLAS at the Z resonance in proton-proton collisons at sqrt(s)=7 TeV*, Tech. Rep. ATLAS-CONF-2011-008, CERN, Geneva, Feb., 2011.

[79] The ATLAS Collaboration, *Muon Reconstruction Performance*, Tech. Rep. ATLAS-CONF-2010-064, CERN, Geneva, Jul., 2010.

[80] ATLAS SUSY Working Group, `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SUSYWorkingGroup` (restricted), last accessed Apr. 30th, 2011.

[81] ATLAS Supersymmetry (SUSY) searches, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/SupersymmetryPublicResults`, last accessed Apr. 30th, 2011.

[82] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, `arXiv:hep-ph/0201195`.

[83] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, *NLL resummation of the heavy-quark hadroproduction cross- section*, Nucl. Phys. **B529** (1998) 424–450, `arXiv:hep-ph/9801375`.

[84] The ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 2.* Technical Design Report ATLAS. CERN, Geneva, 1999. Electronic version not available.

[85] LEP SUSY Working Group (ALEPH, DELPHI, L3, OPAL), Note LEPSUSYWG/01-03.1, `http://lepsusy.web.cern.ch/lepsusy/Welcome.html`, last accessed Mar. 12th, 2011.

[86] The ATLAS Collaboration, *Prospects for Supersymmetry discovery based on inclusive searches at a 7 TeV centre-of-mass energy with the ATLAS detector*, Tech. Rep. ATL-PHYS-PUB-2010-010, CERN, Geneva, Jul., 2010.

[87] The ATLAS Collaboration, *ATLAS Monte Carlo tunes for MC09*, Tech. Rep. ATL-PHYS-PUB-2010-002, CERN, Geneva, Mar., 2010.

[88] S. Moch and P. Uwer, *Theoretical status and prospects for top-quark pair production at hadron colliders*, Phys. Rev. **D78** (2008) 034003, `arXiv:0804.1476 [hep-ph]`.

[89] U. Langenfeld, S. Moch, and P. Uwer, *Measuring the running top-quark mass*, Phys. Rev. **D80** (2009) 054009, `arXiv:0906.5273 [hep-ph]`.

[90] U. Langenfeld, S. Moch, and P. Uwer, *New results for t bar t production at hadron colliders*, `arXiv:0907.2527 [hep-ph]`.

[91] A. Sherstnev and R. S. Thorne, *Parton Distributions for LO Generators*, Eur. Phys. J. **C55** (2008) 553–575, `arXiv:0711.2473 [hep-ph]`.

[92] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004, `arXiv:0802.0007 [hep-ph]`.

[93] The ATLAS Collaboration, *Details on: Early supersymmetry searches with jets, missing transverse momentum and one or more leptons with the ATLAS Detector*, Tech. Rep. ATL-PHYS-INT-2010-083 (restricted), CERN, Geneva, Aug., 2010.

[94] The ATLAS collaboration, *Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at* $\sqrt{s} = 7$ *TeV and measured with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2010-056, CERN, Geneva, Jul., 2010.

[95] The ATLAS collaboration, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of* $\sqrt{s} = 7$ *TeV*, Tech. Rep. ATLAS-CONF-2010-038, CERN, Geneva, Jul., 2010.

[96] The D0 Collaboration, *Measurement of differential* Z/γ^* *+ jet + X cross sections in* $p\bar{p}$ *collisions at* $\sqrt{s} = 1.96$*-TeV*, Phys. Lett. **B669** (2008) 278–286, `arXiv:0808.1296 [hep-ex]`.

[97] The ATLAS Collaboration, *Luminosity Determination Using the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-060, CERN, Geneva, Jul., 2010.

[98] The ATLAS Collaboration, *Search for Supersymmetry with jets and missing transverse momentum and one lepton at* $\sqrt{s} = 7$ *TeV (supporting INT note)*, Tech. Rep. ATL-PHYS-INT-2011-008 (restricted), CERN, Geneva, Feb., 2011.

[99] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Single-top production in MC@NLO*, JHEP **03** (2006) 092, `arXiv:hep-ph/0512250`.

[100] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber, and C. D. White, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, `arXiv:0805.3067 [hep-ph]`.

[101] ATLAS Jet and Missing Et Working Group: How to clean jets, `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HowToCleanJets`, last accessed Feb. 21st, 2011.

[102] The ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*, Tech. Rep. ATLAS-CONF-2010-042, CERN, Geneva, Jul., 2010.

[103] The ATLAS Collaboration, *B-tagging for top physics analyses with early ATLAS data at* $\sqrt{s} = 7$ *TeV*, Tech. Rep. ATL-PHYS-INT-2010-133 (restricted), CERN, Geneva, Dec., 2010.

[104] The ATLAS Collaboration, *Background estimation for searches with missing transverse energy and b-jets Supersymmetry Sensitivity/Exclusion Limit Studies in the Missing Transverse Energy and b-Jets channel*, Tech. Rep. ATL-COM-PHYS-2010-1031 (restricted), CERN, Geneva, Dec., 2010.

[105] The ATLAS Collaboration, *Calibrating the b-Tag and Mistag Efficiencies of the SV0 b-Tagging Algorithm in 3* pb^{-1} *of Data with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-099, CERN, Geneva, Dec., 2010.

[106] ATLAS SUSY Etmiss Subgroup: B-tagging Calibration Tool `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BtaggingCalibrationTool`, last accessed Feb. 21st, 2011.

[107] The ATLAS Collaboration, *Measurement of the W→lnu and Z/gamma***→ll production cross sections in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, JHEP **12** (2010) 060, `arXiv:1010.2130 [hep-ex]`.

[108] ATLAS Electron Gamma Working Group: Efficiency Measurements, `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EfficiencyMeasurements`, last accessed Feb. 21st, 2011.

[109] ATLAS Muon Combined Performance Working Group: Guidelines for Analysis in Release 15, `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesRel15`, last accessed Feb. 21st, 2011.

[110] The ATLAS Collaboration, *Muon Momentum Resolution in First Pass Reconstruction of pp Collision Data Recorded by ATLAS in 2010*, Tech. Rep. ATLAS-COM-CONF-2011-003 (restricted), CERN, Geneva, Jan., 2011.

[111] The ATLAS Collaboration, *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV*, `arXiv:1012.1792 [hep-ex]`.

[112] ATLAS Jet and Missing Et Working Group: JES Uncertainty Provider, `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JESUncertaintyProvider`, last accessed Mar. 7th, 2011.

[113] ATLAS Electron Gamma Working Group: EnergyScaleResolutionRecommendations, `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EnergyScaleResolutionRecommendations`, last accessed Mar. 7th, 2011.

[114] The ATLAS Collaboration, *Measurement of the production cross section for W-bosons in association with jets in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector*, `arXiv:1012.5382 [hep-ex]`.

[115] The ATLAS Collaboration, *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s} = 7$ TeV*, Tech. Rep. ATLAS-CONF-2010-054, CERN, Geneva, Jul., 2010.

[116] C.-H. Kom and W. J. Stirling, *Charge asymmetry in W + jets production at the LHC*, Eur. Phys. J. **C69** (2010) 67–73, `arXiv:1004.3404 [hep-ph]`.

[117] The ATLAS Collaboration, *Using the charge asymmetry in W+jet production as a new probe of parton distribution functions in different kinematic regions at the LHC*, Tech. Rep. ATL-COM-PHYS-2010-893 (restricted), CERN, Geneva, Nov., 2010.

[118] J. M. Campbell, R. K. Ellis, and D. L. Rainwater, *Next-to-leading order QCD predictions for W + 2jet and Z + 2jet production at the CERN LHC*, Phys. Rev. **D68** (2003) 094021, `arXiv:hep-ph/0308195`.

[119] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, `arXiv:0901.0002 [hep-ph]`.

[120] The Les Houches Accord PDF Interface: *Example of PDF error calculation*, `http://projects.hepforge.org/~lhapdf/cc/_2examples_2CCTest2_8cc-example.html`, last accessed Mar. 12th, 2011.

[121] G. Watt, `http://projects.hepforge.org/~mstwpdf/pdf4lhc/`, last accessed Mar. 12th, 2011.

[122] The ATLAS Collaboration, *Setting exclusion limits in ATLAS supersymmetry searches with a likelihood ratio based method*, Tech. Rep. ATL-PHYS-INT-2011-032 (restricted), CERN, Geneva, Apr., 2011.

[123] The CMS Collaboration, *Search for Supersymmetry in pp Collisions at 7 TeV in Events with Jets and Missing Transverse Energy*, Phys. Lett. **B698** (2011) 196–218, `arXiv:1101.1628 [hep-ex]`.

[124] The CDF Collaboration, *Inclusive Search for Squark and Gluino Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **102** (2009) 121801, `arXiv:0811.2512 [hep-ex]`.

[125] The D0 Collaboration, *Search for squarks and gluinos in events with jets and missing transverse energy using 2.1 fb^{-1} of $p\bar{p}$ collision data at $\sqrt{s} = 1.96$ TeV*, Phys. Lett. **B660** (2008) 449–457, `arXiv:0712.3805 [hep-ex]`.

[126] The D0 Collaboration, *Search for associated production of charginos and neutralinos in the trilepton final state using 2.3 fb^{-1} of data*, Phys. Lett. **B680** (2009) 34–43, `arXiv:0901.0646 [hep-ex]`.

[127] LEP SUSY Working Group (ALEPH, DELPHI, L3, OPAL), Note LEPSUSYWG/04-01.1, `http://lepsusy.web.cern.ch/lepsusy/Welcome.html`, last accessed Mar. 12th, 2011.

[128] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Statist. **9** (1938) no. 1, 60–62. `http://projecteuclid.org/euclid.aoms/1177732360`.

[129] A. Stuart, K. Ord, and S. Arnold, *Kendall's Advanced Theory of Statistics, Volume 2A: Classical Inference and the Linear Model.* Oxford University Press, Oxford, 1999.

[130] J. T. Linnemann, *Measures of significance in HEP and astrophysics*, `arXiv:physics/0312059`.