On the road to supersymmetry with ATLAS

CERN-THESIS-2011-120
05/07/2011

Front cover: photograph by *Mateus Mondini* taken at Schiphol Airport, a place Nikheffers in transit to/from CERN know well.

On the road to supersymmetry with ATLAS

Een wetenschappelijke proeve op het gebied van
Natuurwetenschappen, Wiskunde en Informatica

Proefschrift

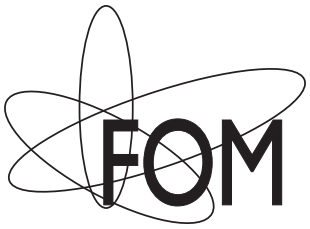
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen op dinsdag 5 juli 2011
om 10.30 uur precies

door

Aleksej Jakovlevich Koutsman

geboren op 02 december 1980
te Moskou

Promotor:	Prof. dr. N. de Groot	
Copromotor:	Dr. W. Verkerke	FOM Instituut Nikhef
Manuscriptcommissie:	Prof. dr. S.J. de Jong	
	Prof. dr. T. Peitzmann	Universiteit Utrecht
	Prof. dr. ir. P.J. de Jong	Universiteit van Amsterdam
	Dr. L. Pontecorvo	INFN - Roma I, Università La Sapienza
	Dr. A.C. König	



ISBN: 978-90-6464-476-4

This work is part of the research programme of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). It was carried out at the National Institute for Subatomic Physics (Nikhef) in Amsterdam, The Netherlands.

Посвящается моим родителям.
Dedicated to my parents.

Contents

Introduction

The twentieth century saw an incredible progress of science and technology. In the first half of the century, the theories of *quantum mechanics* and special and general *relativity* were developed. These theories solved some longstanding questions in physics, such as the ordering of elements in the periodic system with the discovery of protons and neutrons and the constancy of the speed of light. The successive discovery of nuclear fusion finally also solved the problem of the source of solar energy.

The scientific progress did not only benefit humanity, but also brought it close to the brink of complete annihilation in a possible nuclear war. Established in 1954 near Geneva, Switzerland, the new European Organization for Nuclear Research (CERN) focused on "nuclear research of a pure scientific and fundamental character" and stated in the original convention that "the Organization shall have no concern with work for military requirements". The work at the laboratory soon went beyond the study of the atomic nucleus into high-energy particle physics, which was accomplished by construction and operation of particle colliders.

In the 1960s and 1970s, the theory of fundamental particles and their interactions, known as the Standard Model was established. It describes three of the four fundamental forces of nature, and puts quarks and leptons as the fundamental building blocks of nature. A great success of the Standard Model was the prediction and subsequent discovery of the massive weak interaction $W^{\pm}$ and Z bosons at CERN in 1983 and of the heaviest quark, the top quark, in 1995 at the Tevatron accelerator in USA. Despite great triumphs, the Standard Model is not considered to be the final fundamental theory of particle physics. Several theoretical problems and cosmological measurements require an explanation, which the Standard Model cannot give.

One prediction of the Standard Model, the existence of a Higgs boson, is the last missing piece of the puzzle. The Higgs mechanism, postulated already in 1964, is needed to describe the different masses of the fundamental particles, as well as the spontaneous electroweak symmetry breaking, which leads to the existence of the observed massive bosons. The Higgs mechanism introduces a new particle, the Higgs boson, the mass of which must lie below 1 TeV, limited by calculations on weak boson scattering. However when calculating the Higgs boson mass, taking into account its interactions with all other particles, the mass seems to diverge at very high energies. This theoretical problem is called the *naturalness problem*.

During the course of the twentieth century, technological progress allowed astronomers to look at the universe in greater detail as well as observe larger and more distant structures, such as clusters of galaxies. One of the most surprising discoveries is that matter as we know it makes up only a small portion ($\sim 4.6\%$) of mass-energy density of the observable universe. A much larger portion of the universe ($\sim 23\%$) is accounted for by *dark matter*, while the rest is described as *dark energy*.

None of the particles described by the Standard Model fit the description of dark matter. Different extensions of the Standard Model have already been conceived, *supersymmetry* being one favored by many theoreticians. It implies a new symmetry, which predicts that for every particle in the Standard Model, there exists a supersymmetric partner particle, that differs in spin by half a unit and in mass. The difference in spin naturally solves the naturalness problem of the Higgs boson. The difference in mass and the postulation of the existence of a stable

lightest supersymmetric particle, provides a credible dark matter candidate.

The LHC accelerator at CERN takes particle physics into a new domain, as it is designed to supersede its predecessors in both energy and intensity. With enough accumulated data the Standard Model Higgs boson must be found, and if it is not found some new physics phenomenon should be detected. To date no supersymmetric particles have been observed, but with the high collision energy of the LHC expectations are high to find (a hint of) new physics beyond the Standard Model.

Thesis outline

This thesis is concerned with physics performance of the ATLAS experiment at the LHC, and can roughly be divided into two parts. The first part has to do with the discovery of supersymmetry in events with one lepton, specifically one muon. Chapter 4 describes a data-driven method to estimate the Standard Model backgrounds to supersymmetry searches, describing the techniques and possible results using simulated data. Due to the delay in the LHC start-up, it was not possible to reproduce this study on collision data, on the time scale of this thesis.

However it was possible and very exciting to take part in studies with the very first data coming out of ATLAS, since the LHC started colliding protons at 7 TeV collision energy from march 2010. In Chapter 5 the inclusive muon spectrum is studied and compared to simulation. The focus lies on the composition of muons, distinguishing muons coming from pion and kaon decays inside the detector, from the muons coming from the interaction point. First the results based on the very first single muon data are described, while in the second part of the chapter we extend the composition measurement to di-muon events, showing results that include among others the known J/ψ and Z boson resonances. The last chapter, Chapter 3, is concerned with the period of time just before the start of collisions, when the ATLAS detector was commissioned using cosmic rays. In there we focus on the precision chambers of the Muon Spectrometer, which were retrofitted with specialized *twin tube* boards, to measure not only the precision coordinate but the orthogonal coordinate (in the plane of the chamber) as well.

As the LHC and ATLAS embark on a campaign to acquire a large dataset of a few fb^{-1} in the year 2011, all eyes are focused on the results the scientists at CERN will be publishing, as it is often said nowadays that the time of theoretical predictions should be over and the time for observations has come again.

1 Theory

1.1 The Standard Model at proton-proton collisions

1.1.1 Ingredients of the SM

The Standard Model of particle physics is a theory developed in the sixties and seventies of last century describing the fundamental building blocks of nature and their interactions. Quantum field theory provides the mathematical framework for the Standard Model, in which a Lagrangian controls the dynamics and kinematics of the theory. Each particle is described in terms of a dynamical field that pervades space-time. The local $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetries are the symmetries that define the Standard Model. The three symmetry groups give rise to three fundamental interactions. According to Noether's theorem each of these symmetries have an associated conserved charge. The SM describes three of the four fundamental forces known in nature: electromagnetic, weak and strong. The electromagnetic and the weak forces are described together by the electroweak theory.

Forces of the Standard Model Each force is described by a mediator or force carrier particle(s). The photon is the force carrier for the electromagnetic force, which couples to electric charge. The weak force is mediated by the $W^{\pm}$ and the Z gauge bosons that couple to weak charge. A unique feature of the charged weak interaction under the exchange of $W^{\pm}$ bosons, is that it affects only left-handed[1] fields (and right-handed antifields) [1]. The gluons are the force carriers of the strong force that couple to the quantum number colour, which comes in three values labeled as green, red and blue. Unlike the electrically neutral photons for the electromagnetic force , gluons themselves have colour charge. Because of this, the strong interaction behavior is different from the other two forces of the SM, as will be discussed in Section 1.1.2.

Matter in the Standard Model All of matter around us is built up of fundamental particles called quarks and leptons. These quarks and leptons appear in three families or generations. The first family contains the up (u) quark and the down (d) quark, together with the electron (e^-) and the electron neutrino (ν_e). The quarks are the only fermions that carry colour charge and thus couple to the strong force carrier, the gluon. The first family is responsible for all the visible matter on earth, as for example a proton consists of two u and one d quarks. Three families of particles have been discovered, with the last missing lepton, the tau neutrino (ν_τ), escaping direct observation until the year 2000 [2]. The second family is heavier than the first

[1] The helicity of a particle is **L**eft-handed if the direction of its spin is opposite to the direction of its motion.

Family	Fields/Particles *I*	*II*	*III*	Spin	Electric charge
Quarks	$(u,d)_L$	$(c,s)_L$	$(t,b)_L$	$(1/2,1/2)$	$(+2/3,-1/3)$
	u_R	c_R	t_R	1/2	+2/3
	d_R	s_R	b_R	1/2	-1/3
Leptons	$(\nu_e,e^-)_L$	$(\nu_\mu,\mu^-)_L$	$(\nu_\tau,\tau^-)_L$	$(1/2,1/2)$	$(0,-1)$
	e^-_R	μ^-_R	τ^-_R	1/2	-1

Table 1.1: The fermions of the Standard Model, listed with their spin, electrical charges and helicity.

	Particles	Force	Spin	Electric charge
Gauge bosons	g	strong	1	0
	$W^\pm$ and Z	weak	1	±1 and 0
	γ	electromagnetic	1	0
Scalar boson	H		0	0

Table 1.2: The bosons of the Standard Model, listed with their spin and electrical charges.

and the third is heavier than the second. It has been experimentally demonstrated that no more than three species of light neutrinos exist by studying the width of the Z resonance.

Fermions and bosons Tables 1.1 and 1.2 list all the fundamental matter (fermion) and force carrier (boson) particles of the SM. Fermions are by definition particles with half-integer spin and bosons have integer spin, obeying respectively Fermi-Dirac and Bose-Einstein statistics. Fermions are subject to the Pauli exclusion principle, which states that no two identical fermions can simultaneously occupy the same quantum state, in contrast to bosons of which any number can be in the same state. To have a complete picture of the SM particle zoo we would have to mirror Table 1.1 to the *antimatter* world. A particle and its antiparticle have opposite values for all non-zero quantum number labels.

In the SM fermion fields described by flavour eigenstates generally mix between families to form mass eigenstates. For quarks the mixing is described by the CKM matrix [3], while for neutrinos it is described by the PMNS matrix [3]. The quarks are usually named by their flavour eigenstates, while for neutrinos the custom is to use the weak eigenstates.

As mentioned before, the W boson only couples to left-handed fermion fields, which is the meaning of the L-subscript for the doublets in Table 1.1. The right-handed (R-subscript) quarks and charged leptons transform as singlets under the weak interaction.

The Higgs mechanism Mass terms for fermions in the SM can only be realized in a renormalization safe way by the Higgs mechanism [4–6]. The Higgs mechanism spontaneously breaks the electroweak symmetry, transforming the originally massless weak interaction bosons into massive bosons and introducing an extra single physical massive neutral scalar particle, the Higgs boson, while leaving the photon massless. The demonstration of the existence of the Higgs boson is a crucial test for the Standard Model, which is pending verification with the detectors at the LHC.

The Higgs mechanism generates masses by introducing a new field in the theory, the Higgs field, with the unusual property of a non-zero vacuum expectation value. Fermions acquire mass through a Yukawa coupling to the Higgs field, where the coupling constant and thus the mass can be different for each particle. The mass of the Higgs boson is a free parameter in the Standard Model and has been the subject of study for many experiments. The large electron positron collider (LEP) at CERN set a lower limit on the mass of the Higgs of 114.4 GeV at 95% confidence level by direct searches. In 2009 the Tevatron experiments CDF and D0 have come out with their first narrow exclusion range around 170 GeV at 95% confidence level.

An indirect upper limit on the mass of the Higgs boson is set by calculation of longitudinal boson scattering processes in the SM. These grow with energy and break down due to unitarity violation[2] if the Higgs boson mass is larger than $\sim$ 1 TeV [7]. With a center-of-mass collision energy of 14 TeV, the LHC should hence be able to either discover the Higgs boson with $m_H \lesssim 1$ TeV or disprove its existence, with sufficient data. In the case that the Higgs boson is not found, investigation of gauge bosons scattering at high energy should yield insights into new physics.

Gravity and the Standard Model Gravity is the fourth fundamental force of nature, but it is not described by the Standard Model. It is many orders of magnitude weaker than the other three forces at the distances used in elementary particle physics, but becomes dominant at astronomical scales. Many scientists have tried to unite all the four fundamental forces together in a Grand Unified Theory (GUT), string theory currently being a favorite among many theoreticians. Problematic however with these GUT theories is the fact that they do not allow (yet) to make calculations of phenomena we can observe, either to disprove or justify their assumptions.

1.1.2 Quantum Chromodynamics and the parton model

Quantum Chromodynamics (QCD) is the theory describing the strong interaction between particles carrying colour charge, quarks and gluons. Unlike photons, gluons carry the colour charge that they couple to, hence gluons interact with each other, giving rise to unusual behavior of the strong interaction. Effective energy stored in a gluon flux tube between two colour charges is constant per unit of distance. Thus the strong forces increases linearly with the distance between the two charges. The effective coupling constant α_S thus depends on the distance between the charges or the energy scale of the interaction. The constant is said to run, being large at low energy (large distance) and becoming smaller at high energy (small distance).

The net effect of this feature of the strong coupling is *confinement* or in other words coloured particles cannot exist freely. Only colour-neutral bound states of multiple coloured particles can be observed in nature to travel over macroscopic distances. A colour and its anti-colour together are colour-neutral, as well as a combination of three different (anti-)colours. Bound states consisting of a quark-antiquark pair are collectively named mesons. For example a neutral pion π^0 is a bound state of an up and an anti-up quarks with opposite colour charges (e.g. red and anti-red). Bound states consisting of three quarks are collectively named baryons. The proton is a baryon consisting of two up and one down-quarks, where each quark takes one of the three possible colour charges (red, green and blue) to deliver a colour-neutral state. These three are called the valence quarks of the proton, but they are surrounded by a sea of gluons and quark-antiquark pairs that arise from quantum fluctuations.

Another consequence of the structure of strong interaction is that perturbative calculations

[2] Violating unitarity would implicate that the probability of a process occurring rises above 1.

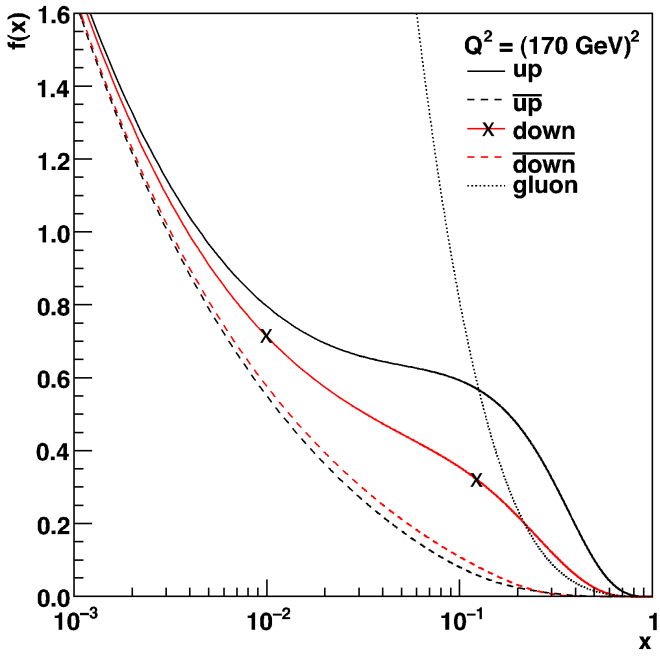


Figure 1.1: Parton distribution functions for the proton according to the CTEQ6.5 parametrization. We see that the distributions for the up and down quarks are higher than those of the anti-quarks due to the contribution of valence quarks. The gluon distribution can be seen to rapidly increase for lower momentum fraction x.

are not possible in the regime of large α_S . Thus at low energy ($\lesssim 1$ GeV) QCD calculations are very difficult, e.g. we cannot calculate the mass of π^0 (~ 135 MeV) from first principles.

The parton model

As the LHC collides protons, an accurate description of the proton is essential, as a high energy pp collision is really a collision of two proton constituents. Thus a description of the constituents is needed at a given energy scale Q of the hard scattering.

The valence quarks and the sea (anti)quarks and gluons are collectively called the *partons* of a proton. Each of these partons carries only a fraction of the momentum and energy of the proton. For the measurement of a hard scattering cross section involving quarks and gluons in the initial state, we need to know the momenta of the incoming particles. As the partons carry only a fraction of the momentum and are in constant interaction with each other, these momenta are unknown, hence the Q of collisions at the LHC varies. Furthermore the outgoing quarks and gluons cannot be observed directly because of confinement but enter the detector as jets, as discussed in Section 1.1.3. Thus we cannot measure a partonic cross section such as $\sigma(qg \to qg)$, but we can make an inclusive measurement, such as the hadronic cross section $\sigma(pp \to jj)$ with two outgoing jets.

To be able to go back in perturbation theory from the hadronic cross section to the partonic cross section, we need to know the probability that a parton of type n is encountered with momentum fraction x. Many experiments have been performed to measure the parton distribution functions $f_n(x_n, Q^2)$ [3] describing these probabilities for a proton, as they cannot be determined

[3] More customary is to write $f_n(x_n, \mu_F^2)$, where μ_F is the factorization scale separating the hard scattering gluon from soft collinear gluon emission for which the probability diverges.

from theory [8]. Most importantly these were measured at the HERA $e^{\pm}p$-accelerator in Hamburg. These measurements will have to be extrapolated to the higher collision energy at LHC, which introduces uncertainties that will have to be studied in due course.

In Figure 1.1 $f_n(x_n, Q^2)$'s are shown where x_n is the fraction of the total momentum carried by parton n, under the assumption that the energy we are measuring at is $Q^2 = (170\ \text{GeV})^2 \simeq m_{top}^2$. For all partons the probability decreases with increasing x. The contribution of the valence quarks inside the proton is apparent as for all x the probability to find an up(down) is higher than an anti-up(down). The gluon distribution dominates the quarks for lower x values, leading to the conclusion that for low momentum transfer collisions gluons are mostly responsible.

The $f_n(x_n, Q^2)$'s are also dependent on interaction scale Q^2. The partons taking part in the hard interaction are not only the valence quarks, but also the sea gluons and (anti-)quarks. The larger the scale Q^2 (in other words the smaller the wavelength of the probe or the larger the probing resolution), the more quantum fluctuations can be observed and hence the amount of $q\bar{q}$ pairs and gluons in the partonic sea increases [9]. Although these partons carry only a small fraction of the proton momentum, their increasing number leads to a relative softening of the valence quark contribution as Q^2 increases.

1.1.3 Description of pp collisions at the LHC

Colliding two protons at high energies at the LHC (see Chapter 2) means studying the interactions of the constituents of the protons, i.e. quarks and gluons. The total interaction of two protons is complicated, as shown schematically in Figure 1.2. In all diagram components except the hard scattering (which will be described in detail in Section 1.1.4) QCD processes dominate.

Fragmentation, hadronisation and decay

Besides the hard scattering final state partons, the total parton multiplicity depends on the amount of (hard) radiation in an event. Extra partons can be produced by radiation of a gluon (quark) either before the hard interaction called *Initial State Radiation* (ISR) or after the hard interaction in the fragmentation phase called *Final State Radiation* (FSR).

After the hard scattering process, each of the coloured final state partons start to loose energy through radiation of gluons. These gluons fragment into additional gluons and quark-antiquark pairs. This continues up to the point where the energy is low enough to recombine all the colour charged particles into mesons and baryons, called *hadronisation*. Hadrons subsequently decay into other hadrons, leptons and neutrinos. Each hard parton from the hard interaction (and ISR/FSR) results in a shower of particles, which are collectively called a *jet*. The jets can be observed by empirically clustering calorimeter cell measurements in a cone algorithm, as explained in Section 2.4.4. Care must be taken when comparing measured jets and theoretically predicted partons, as the produced shower neutrinos escape detection and the ambiguity in defining a jet ('out of cone' particles), cause the measured jet energy to be in general smaller than the initial parton.

Underlying event

The remains of the two protons, supplying the hard scattering partons, are also colour charged and are thus unstable. These unstable states interact in mostly soft scatterings, called the *underlying event*. The interactions in the underlying event are connected to final state particles in the hard scattering, so that the total collision remains colour neutral. The remnants of the

Figure 1.2: Illustration of associated production of a top quark, an anti-top quark and a Higgs boson in a *pp* collision. The ellipses indicate the various stages of the interaction and decay: the hard scattering (HS) involving partons from the incoming protons (PDF), initial state radiation (ISR), final state radiation (FSR), hadronisation and decay of particles and the underlying event (UE). Figure (modified) taken from [10, 11].

protons produce two more jets in the directions of the original proton momenta along the beam line.

Multi-parton interactions

Multiple parton-parton interactions can occur in a single *pp* collision. Most of these extra interactions (other than the hard scatter that triggered this event) will be soft, but occasionally multiple hard scatters in a single *pp* collision will take place.

As the partonic cross section scales with transverse energy E_{T} like $1/E_{\mathrm{T}}^2$ [12, 13], it diverges at low transverse energy transfer. Hence a lower cut-off scale (around $E_{\mathrm{T}} = 2$ GeV in simulation

software) must be introduced in the theory, below which the perturbative cross section is taken to be zero or strongly dampened [14].

Describing multiple hard interactions is complicated as we have to take into account many effects, such as colour reconnection of underlying event partons with multiple hard scattering partons. The number of interactions per collision is not just a Poisson distribution, but is usually described by a double Gaussian, which is dependent on impact parameter between the incoming hadrons and correlations. Also rescattering, or in other words a parton interacting multiple times in a single hard scattering [11], has to be taken into account. Experimental evidence exists (e.g. measurement of charged particle multiplicities [15]) that the simple model without multi-parton interactions is inadequate, when studying the charged particle multiplicities in hadronic collisions, as well as hard interactions such as $t\bar{t}$ pair production [16]. In the latter case the predicted transverse momentum of the $t\bar{t}$ system varies (up to 50% in the lower p_{T} -regions) when incorporating multiple interactions in the underlying event.

1.1.4 Characterization of main hard scattering processes

In Figure 1.3 the predicted cross sections of the most important hard scattering processes at hadron colliders are shown as a function of the collision energy. In the next sections we will discuss a few processes in detail that are most relevant in the context of this thesis.

The $\sqrt{s}$ dependent behavior of all the cross sections in Figure 1.3 can be understood by considering the momentum fractions needed for production and the available diagrams of each specific process. Production of massive particles, such as top quark pair above the threshold of $2 \times m_{top} \sim 350$ GeV, requires smaller momentum fractions (x) at the LHC than at the Tevatron. Since Figure 1.1 shows the parton density functions to be higher at lower x, the cross section at LHC is higher than at the Tevatron. Below $\sqrt{s} = 4$ GeV the cross sections are calculated for $p\bar{p}$ collisions, while above that energy the calculations are for a pp collider. For processes dependent on the $\bar{q}$ contribution inside the proton, this division shows up as a discontinuity.

The cross section of multijet production $p_{\mathrm{T}} > E_{\mathrm{T}}/\sqrt{20}$ is the only one that falls with increasing collision energy. This can be understood by considering that as the momentum fractions x_1, x_2 are equal at all values of $\sqrt{s}$, the cross section mainly depends on the partonic cross section that falls like $1/E_{\mathrm{T}}^2$ [12], hence the total cross section falls with increasing $\sqrt{s}$.

1.1.5 QCD multijets

The overwhelming majority of events at hadron-colliders are QCD multijet events. Though production of *dijet* events is dominant at the LHC, due to the large QCD cross section, events with three or more jets are also copiously produced. The total QCD multijet cross section is estimated at approximately 6×10^9 pb for jets with $p_{\mathrm{T}} > 10$ GeV [20]. At the same time these are also the processes that have the highest uncertainties in cross section calculations. The determination of these cross sections, as well as the evaluation of the probabilities of QCD multijets mimicking other processes, is important at the start of the LHC, as they constitute an important background for most other hard interactions.

The simplest two parton scattering resulting into two jets has ten leading order diagrams that need to be summed to calculate the total cross Section [7]. If only two partons are produced in the interaction, the two corresponding jets will be back-to-back in azimuth, and momentum conservation makes sure that they are balanced in the transverse plane of the laboratory frame.

A challenge lies in the calculation of events with multiple jets in the final state. For every extra hard parton in the final state, the predictions of leading order QCD deviate from a simple $2 \to 2$ scattering, by a factor α_{S}. The uncertainty of scale Q at every interaction point (for

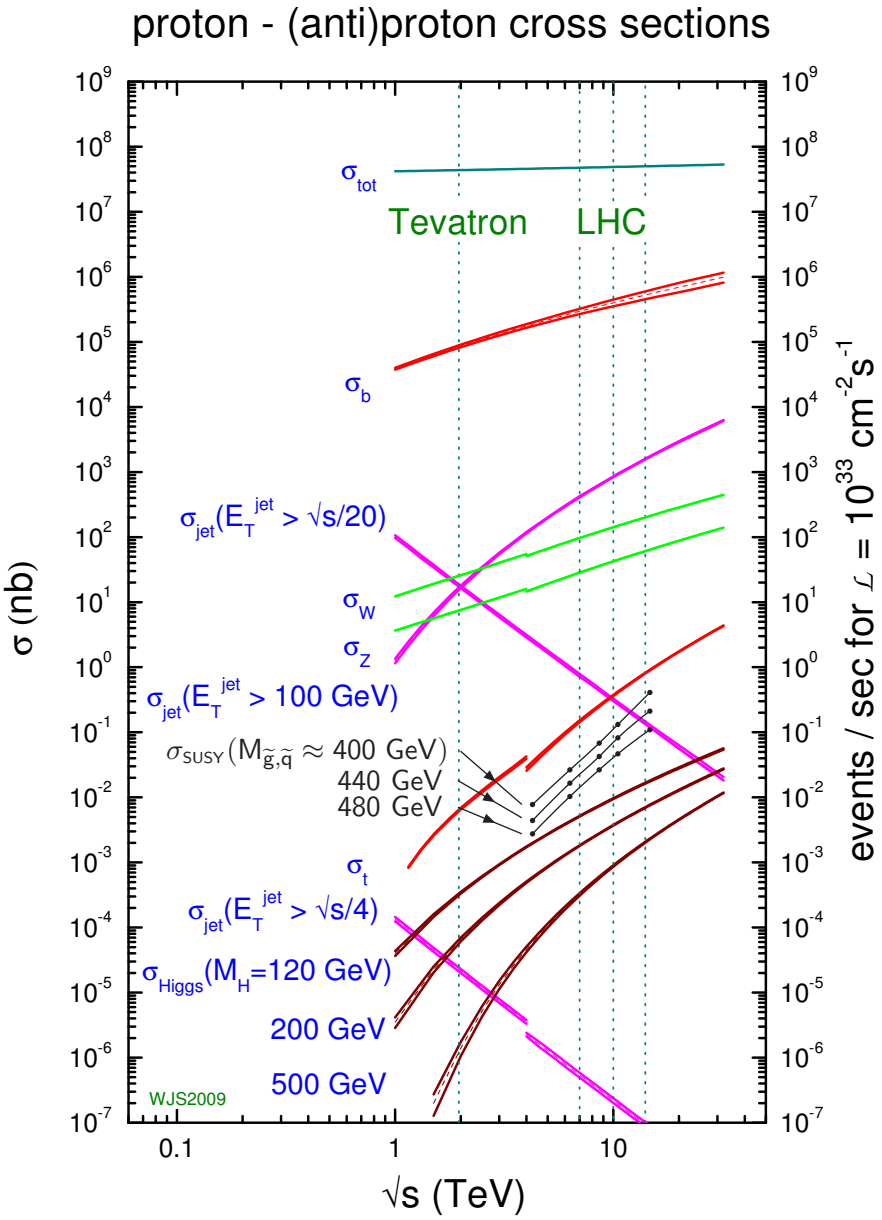


Figure 1.3: Standard Model cross sections at the Tevatron $p\overline{p}$-collider and three possible energies of the LHC pp-collider, taken from [17]. Superimposed are cross sections for three supersymmetric mSUGRA models, studied by ATLAS [18], with masses of lightest coloured $s$particles at respectively $M_{\tilde{g},\tilde{q}} \approx 400,\ 440,\ 480$ GeV. The dots represent NLO cross sections calculated with the PROSPINO [19] program at center-of-mass energies of 4, 6, 8, 10, 14 TeV while the lines are a simple interpolation.

every extra α_S) leads to a greater uncertainty in the total cross section. Scale uncertainties are reduced by taking into account loop corrections for next-to-leading (higher) order. However the higher order calculation becomes increasingly difficult with more outgoing legs in the final state, as the number of possible diagrams expands dramatically. These difficulties in calculating cross sections with associated jets do not only pertain to QCD multijet events, but for all processes with extra radiated partons.

QCD multijet events constitute a background not only for events with jets in the final state, but also for events with leptons and missing energy. The enormous amount of jets leads some jets to fluctuate and mimic the more dense electromagnetic shower in the detector. This can lead the software to wrongly reconstruct an electron in this event instead of a jet. Besides light flavoured jets, QCD multijets produce $b\bar{b}$ pairs at the LHC. The resulting B-hadron can decay semileptonically, producing both a real lepton and real missing energy in the detector. Such an event can mistakenly be perceived as a W/Z boson decay with associated jets.

1.1.6 W/Z production with associated jets

The $W^{\pm}$ and Z gauge bosons were discovered in 1983 by the UA1 and UA2 experiments at CERN [21–23]. The masses of the bosons are given by the Particle Data Group [3] to stand at a current world average of:

$$m_W = 80.398 \text{ GeV}, \tag{1.1}$$
$$m_Z = 91.188 \text{ GeV}. \tag{1.2}$$

The gauge bosons can be produced directly in a hard interaction at the LHC, as diagrams of Figure 1.4 show. The branching ratios to the shown decay modes are:

$$\begin{array}{lcll} W^+ & \to & l^+\nu & : \quad 10.80 \pm 0.09\% \quad (3\times) \\ & \to & hadrons & : \quad 67.60 \pm 0.27\% \\ Z & \to & l^+l^- & : \quad 3.366 \pm 0.002\% \quad (3\times) \\ & \to & \nu\overline{\nu} & : \quad 20.00 \pm 0.06\% \\ & \to & hadrons & : \quad 69.91 \pm 0.06\% \end{array}$$

where l stands for lepton: e, μ or τ. The branching ratios for the W^- are charge conjugates of the W^+-modes above. Taus predominantly decay into hadrons, and with a smaller branching ratio of 17.61 ± 0.05 % into an electron or muon and two neutrinos, thus giving rise to a higher rate of events with an e/μ in the final state than predicted by direct decay.

The expected cross sections [20] times the branching ratio into a final state containing an electron at a center-of-mass energy of 10 TeV:

$$\sigma(W^- \to e^-\overline{\nu_e}) = 12.4 \cdot 10^3 \text{ pb} \tag{1.3}$$
$$\sigma(Z \to e^+e^-) = 1.10 \cdot 10^3 \text{ pb}, \tag{1.4}$$

where the respective cross sections for μ, τ final states are the same. The uncertainties on these cross sections are known to the 1% level or better. The given cross sections are calculated at leading order by the ALPGEN program [24], multiplied by the corresponding k-factor [20], to scale the cross section to NLO. The inclusive boson production cross sections can be calculated up to NNLO with a precision of a few percent by the state of the art FEWZ program [25]. The calculated FEWZ cross sections agree with the preliminary 2010 results from ATLAS [26].

Boson production with associated jets

Partons jets produced in ISR/FSR can add jets to W/Z production, making it a potential background to physics processes such as production of Higgs bosons, supersymmetric particles and models with extra dimensions. Production of vector bosons in association with an arbitrary number of jets are collectively named $W + jets$ and $Z + jets$.

Just as for QCD multijets, cross section calculations for the $W/Z+jets$ processes have proved

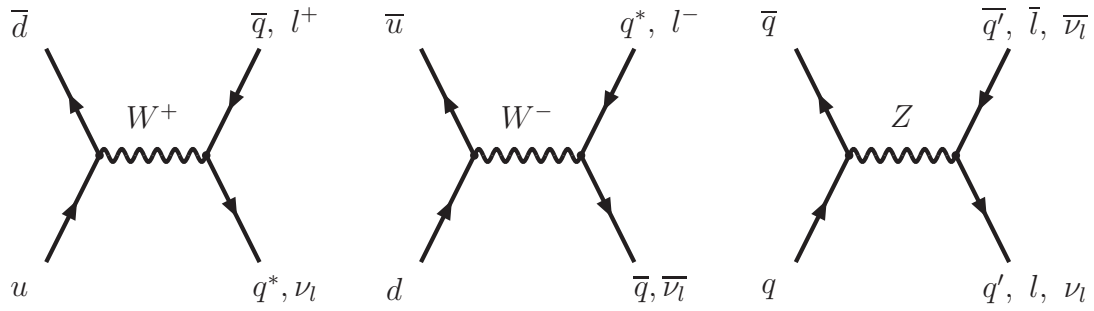


Figure 1.4: Diagrams for $W^{\pm}$ and Z production and main decay channels at the LHC, where l stands for charged lepton and q for quark.

to be increasingly difficult with increasing number of associated jets. NLO cross sections have been calculated up to $W/Z + 3\ jets$, but going to four or more associated jets has proven quite a challenge. At leading order it is possible to calculate the cross sections with up to six associated jets. These cross sections will have to be measured by data-driven methods at the LHC.

W-production charge asymmetry

The production of W bosons at the LHC exhibits a charge asymmetry, unlike other SM processes, because it requires a $q\bar{q}$ pair at medium to high momentum fraction x. This high fraction is needed due to the relation $sx_1^2x_2^2 > M_W^2$ between W-mass M_W, center-of-mass energy $\sqrt{s}$ and momentum fractions x_1, x_2.

In this regime the q contribution in the proton is dominated by the valence quarks, while the $\bar{q}$ is only available as a sea-quark. The production of W^+ boson requires a u quark, while the W^- boson requires a d quark, as seen in Figure 1.4. Since there are two valence u quarks, as opposed to only one valence d quark, the W^+ boson enjoys a higher production rate than the W^- boson at the LHC.

If we take a simplified parton model where the momentum distribution for the sea quarks of all flavours is equal, only leading order diagrams are taken into account and only the valence quarks are responsible for the u and d contribution with equivalent momentum distributions the asymmetry is:

$$A_W = \frac{N_{W^+}}{N_{W^+} + N_{W^-}} - \frac{N_{W^-}}{N_{W^+} + N_{W^-}} = \frac{2}{3} - \frac{1}{3} = \frac{1}{3} \tag{1.5}$$

where N_{W^+} is the number of W^+ bosons produced. The actual asymmetry depends on x_1, x_2, which in turn increase as the multiplicity of hard associated jets rises.

This W-charge asymmetry translated into a lepton charge asymmetry can be measured as described in [27]. The W-production charge asymmetry can be used to isolate this process in an inclusive single lepton data sample.

1.1.7 Top quark production

Predicted by the Standard Model as the partner of the bottom (b) quark in a weak doublet, the top quark (t) was the last quark to be discovered in 1995 by the CDF [28] and D0 [29] collaborations at Tevatron. The top quark is the heaviest fundamental particle measured with a mass of $m_{top} = 173.1 \pm 0.6\ (\text{stat.}) \pm 1.1\ (\text{syst.})$ GeV [30], five orders of magnitude larger than the first family quarks. At this mass the theoretically calculated decay width is $\Gamma_{top} = 1.35$ GeV [8].

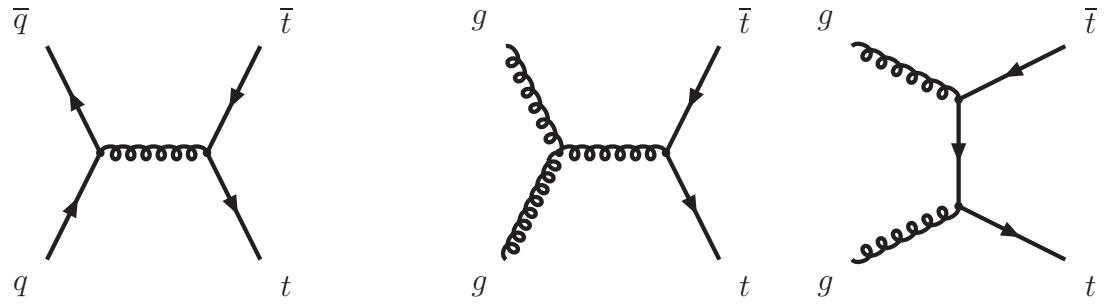


Figure 1.5: Leading order top quark pair production. Quark-antiquark annihilation (*left*) and gluon-gluon fusion (*middle and right*), the dominant production channel at the LHC.

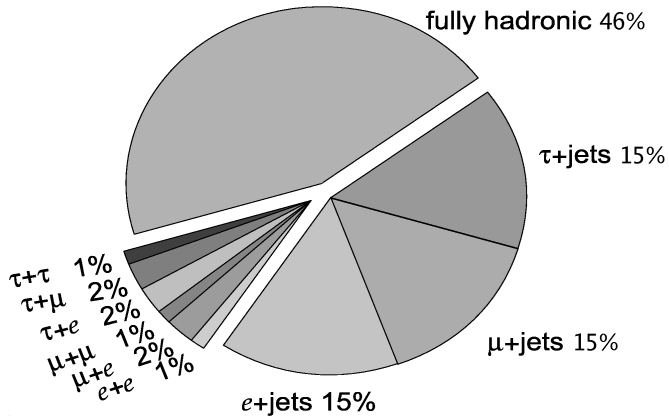


Figure 1.6: Branching fractions of a top quark pair illustrated in a pie-chart. The pair decay is categorized by the decay of the W boson, either hadronically or leptonically.

Top quark pair production

The top quarks will be predominantly produced at the LHC in pairs through the strong interaction. The leading order production processes, $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$, are shown in Figure 1.5. The approximate NNLO calculated cross section for $\sqrt{s} = 10$ TeV proton-proton collisions is:

$$\sigma(pp \to t\bar{t}) = 401.6 \ {}^{+3.7\%}_{-4.3\%} \text{ (scales) } {}^{+4.6\%}_{-4.5\%} \text{ (PDF) pb.}$$

The next-to-next-to-leading order (NNLO) calculations are approximated by resumming large logarithmic terms in the NLO calculation. The resulting cross section is less sensitive to the factorization and renormalization *scales* [11]. As the anti-quark is only available as a sea quark in a pp-collider, gluon-gluon fusion is the dominant production channel accounting for up to 90% of the total. As the gluon parton distribution functions (PDF) are uncertain specifically at very high x, the resulting cross section uncertainties are quoted separately above.

The top quark decays before it hadronizes, thus no bound states (mesons or baryons) with top quarks exist. The top quark decays nearly always ($> 99.8\%$) into a W boson and a b-quark. So the signature of a $t\bar{t}$ decay is determined by the decays of the two W bosons. Three cases can be distinguished: *fully hadronic*, *semileptonic* and *dileptonic* decays. In the fully hadronic case both W's decay hadronically, which results in six quarks in the final state. The experimental signature is hence at least six jets and no leptons. In the semileptonic case one W decays leptonically,

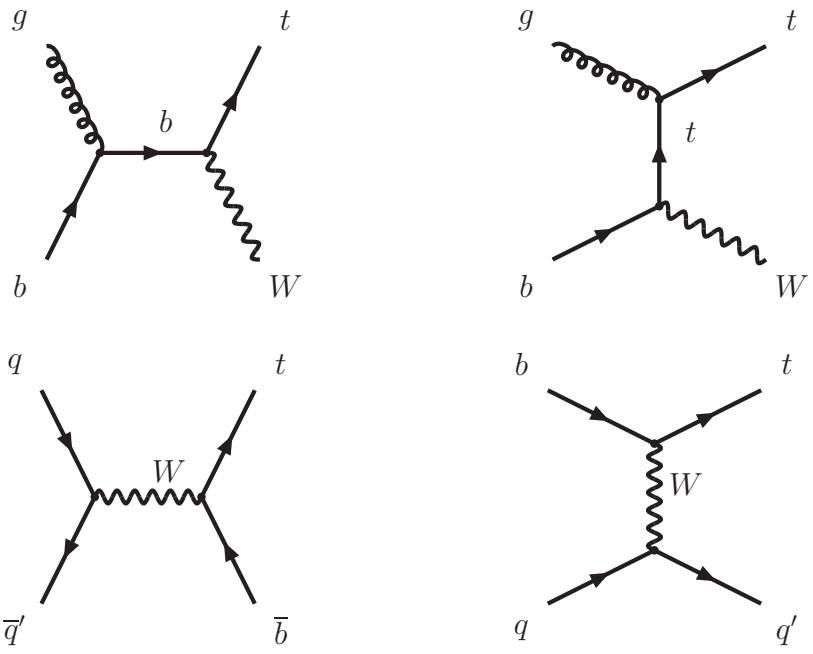


Figure 1.7: Leading order single top quark production. The upper two diagrams show the *Wt-channel*, while the *s-channel* is shown in the lower left and the *t-channel* in the lower right diagram.

thus giving four quarks, one lepton and one neutrino in the final state. The corresponding experimental signature is at least four jets, one lepton and large missing energy carried away by the neutrino. Finally the dileptonic case has both W's decaying leptonically adding up to two quarks, two leptons and two neutrinos. Dileptonic events have thus an experimental signature of at least two jets, two leptons and large missing energy. ISR and FSR can create additional jets in all $t\bar{t}$ event topologies.

The pie-chart of Figure 1.6 displays the branching fractions of a top pair. In a search for supersymmetric final states with multiple jets, one lepton and large missing transverse energy $t\bar{t}$ will be a major background as approximately 30% of all top quark pair events match this signature.

Jets coming from b-quarks can be experimentally distinguished from jets originating from light quarks or gluons by exploiting the relatively long decay time of B-hadrons, of the order of 1-2 ps. The long decay time may result in a reconstructible secondary vertex for the jet with a comparatively large mass of ~ 5 GeV. Alternatively, one can exploit the high semileptonic branching fraction of b-quark decays, which make that roughly 10% of b-quark decays result in a secondary muon (or electron) from a $b \to c$ decay, with a few GeV of transverse momentum relative to the jet axis.

Single top production The D0 collaboration reported first evidence for electroweak single-top production in 2007 [3, 31, 32]. The diagrams for single-top production are shown in Figure 1.7. A single top quark is either accompanied by a W in the *Wt-channel* or by another quark in the *s-* or *t-channels*. At $\sqrt{s} = 10$ TeV the NLO cross sections for the three channels at the LHC are respectively calculated to be 32.7, 6.6 and 122.1 pb [20]. The identification of single top quarks is much more difficult than in the top pair channel, due to a less distinctive signature and significantly larger backgrounds. In searches for supersymmetric final states it plays a much

smaller role as compared to top pair production.

1.1.8 Simulation strategy

To compare theoretical predictions to experimental data, simulation is an invaluable tool. Monte Carlo (MC) generators are the starting point for the simulation. They can produce distributions, such as cross sections, or even complete events as dictated by the physics model using random numbers.

The standard event generation technique in particle physics follows a chain approach, where in chronological order the following steps are taken: first the desired hard interaction process is simulated. Then the 'soft' physics of fragmentation, hadronisation and decay, discussed in Section 1.1.3, are simulated.

The particles remaining after the previous step must then be propagated through the simulated detector, taking into account the interactions with detector material and magnetic fields. The geometry of the detector and propagation through the detector is taken care of by the GEANT4 program [33]. The simulation of particle interactions with the detector material is most extensive and CPU intensive for the calorimeters, where very many particles are produced in the hadronic and electromagnetic showers. The ATLFAST2 [34] program has been designed to be able to produce large numbers of models or events, with less computing power than would be needed for the full GEANT4 simulation of the whole detector. It includes GEANT4 simulation of the inner detector and muon system supplemented by a fast calorimeter simulation.

The next step in the simulation process needs to take the detector response into account, which is very detector specific, where in ATLAS specific *digitization* software has been written. The last step in the simulation chain is to use standard data reconstruction software, as would be normally done for data. Only when all these steps are correctly modeled and validated, can we reliably compare our predictions to data measurements.

Physics simulation

Two main generators PYTHIA [35] and HERWIG [36,37] are available to perform the complete simulation chain until detector propagation in ATLAS. Specialized generators can be interfaced with these general purpose generators to replace or improve a part of the chain.

Hard scattering simulation The *matrix elements* (ME) can be calculated perturbatively using Feynman calculus, where Feynman diagrams are used for the graphical representation of a specific hard interaction. MC@NLO [38, 39] is an event generator framework for a variety of physics processes that can include diagrams up to one-loop level. For the analysis in this thesis, MC@NLO is used to generate $t\bar{t}$ events, for which it includes a full NLO treatment for events with up to one extra parton. MC@NLO uses negative event weights to compensate for double counting due to overlap in the description of parton radiation by the NLO matrix element and parton showering [11]. The sum of the positive and negative event weights must be used for inclusive distributions.

The ALPGEN generator [24] is a collection of LO generators for a variety of processes with up to six associated jets. In this thesis it is used as the default generator for $W + jets$ events. For these events separate samples of $W + 0, 1, 2, 3, 4, 5$ partons are generated, that have to be summed together to get a complete $W + jets$ set of events. Each of these subprocesses is post-fixed with additional jets in the parton showering phase in the subsequent simulation step.

Parton showering The second step is *parton showering* (PS), since incoming and outgoing coloured objects will radiate. In parton showering the radiation is approximated in PYTHIA and HERWIG by the use of DGLAP splitting functions [40–43] together with Sudakov form factors [44]. DGLAP splitting functions give the chance for one quark or gluon to split into two coloured partons, depending on the number of flavours and the momentum fractions carried away by the two outgoing functions. Only three such splitting functions exist for QCD at LO accounting respectively for $q \to qg$, $g \to gg$ and $g \to q\bar{q}$ branching. The Sudakov form factors provide the probability of evolving from an initial time t_0 to a later time t without branching [7]. Both ISR and FSR are part of the parton showering step.

Hadronisation is the next step, that describes the formation of colourless hadrons. PYTHIA uses the string fragmentation model while HERWIG uses the cluster fragmentation model. In string fragmentation the colour field is represented as a string between a quark-antiquark pair, while gluons cause kinks in the strings. The potential energy grows as partons move further apart, until the string breaks creating new $q\bar{q}$ pairs. Hadrons form from all the available quarks when the potential energy is low enough. In cluster fragmentation gluons remaining after parton showering are split into quark-antiquark pairs. All available (anti)quarks combine into massive colour singlet clusters, which decay into lighter clusters if phase space allows, or into a pair of hadrons. If a cluster is too light to decay into two hadrons, it is taken to represent the lightest single hadron of its flavour [36].

After the hadrons have formed the unstable ones are *decay*ed into stable particles, according to branching ratios given by experimental data as much as possible. Some experimental decay information is still not available, in which case decay properties like branching ratios are assumed. Particles with a lifetime of more than 30 picoseconds in the lab frame are deferred to the GEANT4 simulation step.

The underlying event (UE) model is part of PYTHIA, while HERWIG is usually interfaced to JIMMY [45] to model the UE. Both models use the idea that the $2 \to 2$ QCD process dominates, because it is the lowest order in α_S. However much effort is currently being done to correctly include all the different parts of multi-parton interactions into the description of UE [11, 13].

Matrix elements to parton showers matching strategy

Matrix elements correctly describe the wide angle, high-p_{T} emission of partons. Emissions at the soft and collinear limit are well approximated by parton showers. Techniques have been developed to use the best of both worlds by merging matrix elements with parton showering.

However a *double counting* complication arises, as emissions are accounted for in some regions of phase space by both ME and PS. Some techniques have been developed to sidestep the double counting. The basic idea of all these techniques is dividing the phase space into a region handled by ME and a region handled by PS, as shown for a cone of size ΔR_{MS} around the emitting parton in Figure 1.8, taken from [11]. The default technique used on ALPGEN generated processes in ATLAS is called MLM matching [46]. It uses a veto on events which contain emissions from the parton shower in the region already covered by the matrix elements. In ATLAS the phase space for MLM matching is divided in jet transverse energy and the radial distance between two jets R_{jj}, as discussed in more detail in Section 4.1.1. It has been shown that the resulting kinematic distributions do not depend strongly on the choice of the specific merging scale [20].

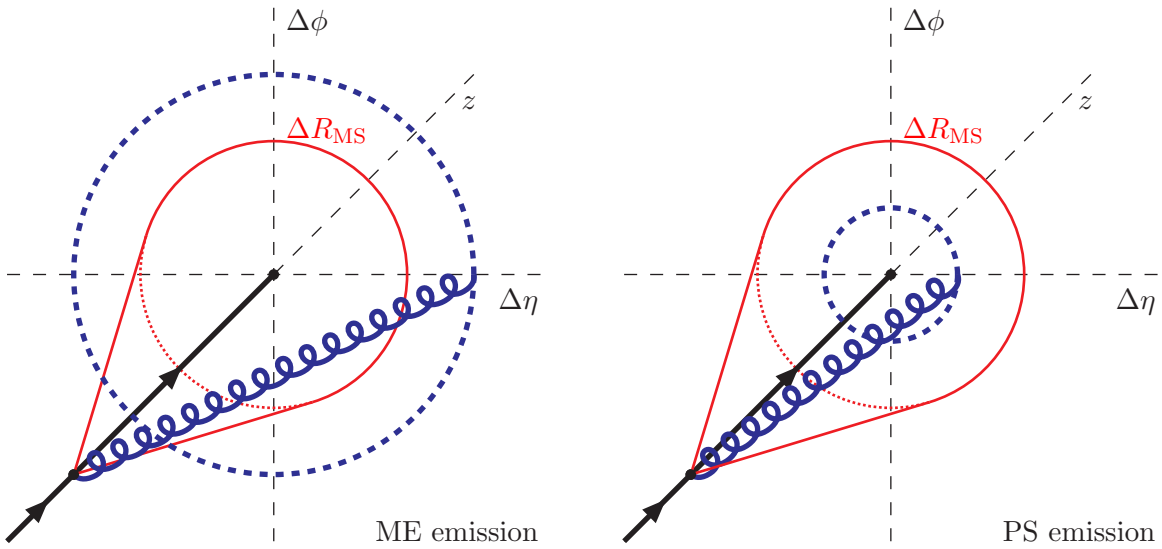


Figure 1.8: Illustration of dividing the phase space into a region for emissions described by the matrix elements (left) and a region for emissions described by the parton shower (right). In this example, the phase space is divided by a cone with size $\Delta R_{MS} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, taken from [11].

1.2 Supersymmetry

Although the Standard Model is in agreement with a great number of experimental measurements, still some theoretical problems exist and some cosmological measurements require an explanation where the SM fails. Supersymmetry (SUSY) is a theoretically attractive extension of the SM that solves some of its outstanding problems, which will be described in the next section. On top of that it supplies us with predictions for new particles in the mass range of up to a few TeV, and as such SUSY was one of the main reasons to build the LHC.

1.2.1 Shortcomings of the SM

Naturalness, fine-tuning and hierarchy problems

Despite the successes of the SM, it is incomplete and there must be a larger, more complete theory which introduces unknown 'new physics' at high(er) energy. The energy scale at which this new physics appears and the SM must be modified is usually denoted as Λ. As the SM does not incorporate gravity, there must be some kind of new physics at the scale when quantum gravity becomes important. This scale is indicated by the Planck mass M_P, which is deduced from Newtonian gravitational constant G_N as:

$$M_P = (G_N)^{-1/2} \simeq 10^{19} \text{ GeV}. \tag{1.6}$$

The very large difference between the electroweak scale ($M_Z \simeq 100$ GeV) and the Planck scale, as $M_Z \ll M_P$, is known as the *hierarchy problem* as this contrast is not understood [47,48].

If no new physics exist between the electroweak and the Planck scales, a problem arises with the Higgs mechanism. When we compute the Higgs mass at one loop level up to the cutoff scale Λ, the quartic selfinteraction[4] of the Higgs boson generates a quadratically divergent contribution

[4]The scalar potential for the Higgs boson, h, is given schematically by $V \sim M_{h0}^2 h^2 + \lambda h^4$.

in Λ as:

$$M_h^2 \sim M_{h0}^2 + \frac{\lambda}{4\pi^2}\Lambda^2. \tag{1.7}$$

Assuming the new physics scale Λ to be the Planck scale, the one-loop correction to the Higgs mass is of the order of $(10^{19})^2$ GeV^2. It is formally possible to fine tune parameters to cancel this large contribution, but it must happen up to the 16th decimal, because the SM Higgs must obey $M_h \lesssim 1$ TeV to prevent unitarity violation, as discussed in Section 1.1.1. As this cancellation must occur at every order in perturbation theory, the parameters must be fine tuned again and again. This is called the *fine-tuning problem* or the *problem of (un)naturalness*, as the Higgs boson is the only particle that suffers from these quadratic divergences. All fermion masses are 'natural' as they do not exhibit quadratic divergences [47].

1.2.2 SUSY as possible solution of SM problems

If we reconsider nature to contain, besides the Higgs field, also massive fermions (F) and massive scalars (S) the one-loop contribution to M_h^2 becomes [47]:

$$\begin{aligned} M_h^2 \;\sim\; & M_{h0}^2 - \frac{g_F^2}{4\pi^2}(\Lambda^2 + m_F^2) + \frac{g_S^2}{4\pi^2}(\Lambda^2 + m_S^2) \\ & + \text{logarithmic divergences} \;+\; \text{uninteresting terms}, \end{aligned} \tag{1.8}$$

where g_F, g_S are the respective couplings to the Higgs. The relative minus sign between the two contributions is the result of spin-statistics, mentioned in Section 1.1.1. In the case that the two coupling strengths are equal, $g_F = g_S$, the quadratic divergences in Λ cancel out exactly. The terms that are left are logarithmic divergences, that do not need extravagant fine tuning, and a term $\frac{g_F^2}{4\pi^2}(m_S^2 - m_F^2)$. As long as this mass difference is not larger than approximately 1 TeV, we are left with a well behaved Higgs boson mass. Hence if supersymmetry exists and solves the Higgs mass fine-tuning problem, which is the underlying assumption of $g_F = g_S$, we should find some supersymmetric particles at the LHC.

Couplings unification

An appealing class of theories of nature are the so-called *Grand Unified Theories* (GUT), that postulate that all forces in nature at low energy are different manifestations of a single force at a very high energy. The relation between the coupling constants at LHC energy scales and those at very high energies are given by the renormalization group equations, which evolve these constants as a function of the energy scale. A characteristic feature of GUT theories is that all coupling constants unite in a single point at a given critical energy.

The evolution of the inverse coupling constants is shown in Figure 1.9 as a function of energy scale. If only the SM particles exist in nature we see that the couplings do not unify, where if SUSY is added to the equations the three couplings cross each other nicely. The two scenarios for SUSY shown in the figure vary by the supersymmetry breaking scale from 250 to 1000 GeV. On top, supersymmetry might help in unification of gravity with the other three forces, but for a detailed discussion the reader is referred to [8, 49].

Relation to cosmology

Over the last fifty years cosmologists have accumulated many measurements that show that visible matter describes only a small part of the energy budget of the universe. Most clear indication comes from rotation curves of spiral galaxies, where a galactic halo of *dark matter* that interacts

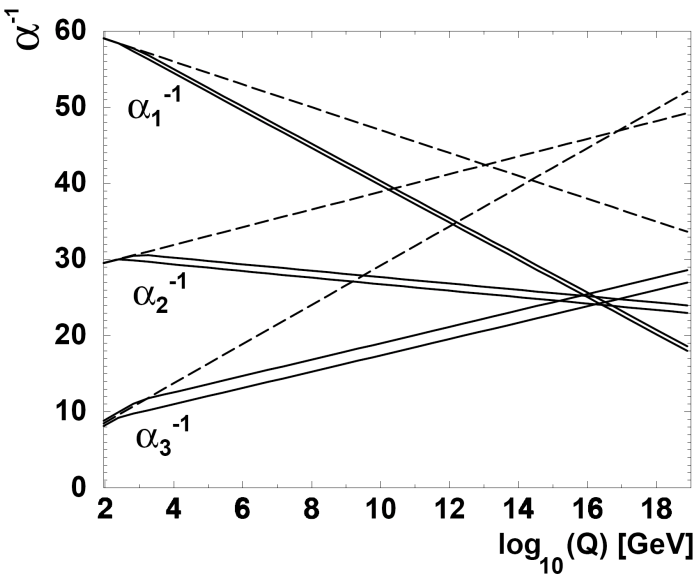


Figure 1.9: Evolution of the three inverse coupling constants in the Standard Model (dashed lines) and in supersymmetric theories (solid lines). Two SUSY scenarios with a varying supersymmetry breaking scale are depicted at either 250 GeV (lower) or at 1 TeV (upper).

gravitationally must be assumed to explain the curve shapes. According to recent data [50, 51], the matter content of the universe can be subdivided into three categories. Baryonic matter that earth and earthlings are made of constitutes 4% of the total, dark matter holds 23% and dark energy accounts for the missing 73%.

Dark matter must be electrically neutral and carry no colour charge. It must be massive, absolutely stable and weakly interacting. The lightest supersymmetric particle in R-parity conserving SUSY models, discussed in the next section, has exactly these characteristics and is considered a good dark matter candidate. If SUSY meets the cosmological constraints, finding SUSY at the LHC will thus not only take care of problems in the Standard Model but will also shed light on the dark matter content of the universe.

1.2.3 Supersymmetry at the fundamental level

Before we go into details of the specific supersymmetric models that were used in this thesis and possible manifestations of SUSY at the LHC, we briefly outline some fundamental concepts of a supersymmetric model.

Boson fermion symmetry

Supersymmetry implies a symmetry between fermions and bosons. It states that for every particle in the SM there is a corresponding superpartner which differs only in spin by half a unit, but with all other characteristics the same. So a fermion in the SM becomes a boson under the supersymmetric transformation and vice versa. Examining the particle content of the Standard Model in Tables 1.1 and 1.2, we see that none of them can be each others superpartners. This means that we must introduce new fields into our model, hence ending up with twice as many fundamental particles as we have now.

Mixing of superfields

It is customary to apply the supersymmetric transformation to the SM fields before unification of the electromagnetic and the weak interactions, and before electroweak symmetry breaking. Hence the U(1) B-field and the SU(2) W_i gauge bosons get their supersymmetric partners, and not the photon or the W/Z bosons.

Similarly to the situation for SM quarks and leptons, where the weak eigenstates are not identical to the mass eigenstates and are related through a mixing matrix, supersymmetric particles (for short *s*particles) with the same quantum numbers will in general mix to form mass eigenstates. As opposed to the SM quark sector, the mixing of supersymmetric weak eigenstates produces significantly different mass eigenstates that are named differently, as will be discussed in Section 1.2.4.

R-parity

When writing down the interactions in the Lagrangian of a supersymmetric model, nothing prevents us from writing down the most general form. However this general form has a problem as lepton and baryon number are no longer conserved. This means that a proton could decay through the exchange of the scalar partner of the down quark. This is contradicted by the measured upper limit on the mean life time of a proton, that is $> 10^{30}$ years at 90% confidence level [3]. There are different strategies to deal with lepton and baryon number violation, but the most common is to forbid these by introducing a new symmetry.

The symmetry which is employed is named *R-parity conservation* and the according new quantum number R-parity can be written down as:

$$R \equiv (-1)^{3(B-L)+2S}, \tag{1.9}$$

where B and L are respectively baryon and lepton number and S is the spin of a particle. R-parity is a discrete multiplicative symmetry, that must be conserved in all interactions. For all the particles of the Standard Model the eigenvalue of $R = 1$ and for all their superpartners $R = -1$. This has profound experimental consequences:

- Being a multiplicative symmetry the number of SUSY particles in any given interaction is always conserved modulo two. Hence supersymmetric particles can only be produced in pairs from Standard Model particles.
- At the LHC SUSY particles with strong interaction couplings will have highest production rates. A coloured supersymmetric particle will decay in a chain until the lightest SUSY particle is produced. These *cascade* decays have as signature multiple high momentum jets and possibly leptons.
- The Lightest Supersymmetric Particle (LSP) is absolutely stable as it has no particle with $R = -1$ to decay to. A clear signature from SUSY particle decay is the large missing energy that is carried away by the LSP.

The LSP is considered a good candidate for the dark matter content of the universe.

SUSY breaking

As no *s*particles have been observed in any experiment to date, SUSY only remains a viable extension of the SM if there exists a mechanism of 'SUSY breaking', that causes the *s*particles

to have much higher masses than their SM counterparts.

There is a variety of suggested mechanisms to generate the breaking of SUSY, e.g. gravity-mediated [52], anomaly-mediated [53] or gauge-mediated [54] breaking, which is assumed to occur in a hidden sector at much higher energies[5]. But in practice one can also model SUSY breaking through the introduction of generic 'soft' mass terms into the theory, termed soft as they do not re-introduce the quadratic divergences which motivated the introduction of supersymmetry in the first place [47]. In the context of this thesis, we use gravity-mediated breaking SUSY models, discussed in Section 1.2.5, that assume that the hidden sector couples to the visible sector through gravitational interactions.

1.2.4 Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model (MSSM) is the simplest supersymmetric extension of the SM with minimal particle field content and R-parity conservation.

SUSY particles

For every particle in the SM there is a supersymmetric partner, as can be seen by comparing the SM particles in Tables 1.1 and 1.2 to Table 1.3(a) showing the superpartner content of MSSM. The *gauginos* carrying spin-1/2 are superpartners of the gauge bosons, and the superpartners of quarks and leptons carrying spin-0 are called *squarks* and *sleptons*.

Supersymmetry dictates that there be at least two Higgs doublets (H_u, H_d) in the MSSM, instead of a single one in the SM. This is needed to generate masses for both up- and down-type quarks. In the same way as in the SM, the weak gauge bosons acquire mass through electroweak symmetry breaking, except that we are now left with five scalar Higgs particles: h^0, H^0, A^0, H^+ and H^-. Hence the non-supersymmetric Higgs sector is larger in MSSM models than in Table 1.2. The superpartners of the scalar Higgses are spin-1/2 *Higgsinos*.

Once SUSY and electroweak symmetry are broken, *s*particles in flavour eigenstates will in general mix to form mass eigenstates. All the resulting mass eigenstates are shown in Table 1.3(b). The only one that stays untouched is the gluino, being the only gaugino with colour charge. The colour neutral gauginos and Higgsinos mix to form two charged states called charginos $\widetilde{\chi}^{\pm}_{1,2}$, and four neutral states called neutralinos $\widetilde{\chi}^{0}_{1,2,3,4}$, where number labeling is based on increasing mass.

As the *s*quarks and *s*leptons are scalar bosons, the *L*-/*R*-handedness does not refer to their helicity state but to that of their SM partner field, though the labels are kept to display the relation. The mixing of *s*bosons happens mainly in the third family, because the mixing matrix off-diagonal contributions are proportional to the SM partner mass [55], hence most relevant for third family partners. The third family mass eigenstates are labeled with *1,2* again according to their increasing mass. This does not happen in the Standard Model, because the t_L and t_R must have the same mass under Lorentz invariance. There can also be flavour mixing of families of *s*fermions in a super-CKM matrix, but for simplicity it is ignored.

1.2.5 Minimal supergravity

We know that supersymmetry must be broken as no supersymmetric partners of the SM particles have been discovered at low energies. Parameterizing the breaking and taking all possible SUSY interactions into account, the MSSM has 105 parameters extra compared to the SM. Having no

[5] You could argue that we replace the hierarchy problem of the SM with the SUSY breaking problem.

	Fields		
Squarks	$(\widetilde{u},\widetilde{d})_L$	$(\widetilde{c},\widetilde{s})_L$	$(\widetilde{t},\widetilde{b})_L$
	$\widetilde{u}_R$	$\widetilde{c}_R$	$\widetilde{t}_R$
	$\widetilde{d}_R$	$\widetilde{s}_R$	$\widetilde{b}_R$
Sleptons	$(\widetilde{\nu}_e,\widetilde{e})_L$	$(\widetilde{\nu}_\mu,\widetilde{\mu})_L$	$(\widetilde{\nu}_\tau,\widetilde{\tau})_L$
	$\widetilde{e}_R$	$\widetilde{\mu}_R$	$\widetilde{\tau}_R$
Gauginos	$\widetilde{g}$		
	$\widetilde{B},\widetilde{W}_{1,2,3}$		
Higgsinos	$\widetilde{H_u},\widetilde{H_d}$		

(a)

	Mass eigenstates		
	$\widetilde{u}_L,\widetilde{u}_R$	$\widetilde{c}_L,\widetilde{c}_R$	$\widetilde{t}_1,\widetilde{t}_2$
	$\widetilde{d}_L,\widetilde{d}_R$	$\widetilde{s}_L,\widetilde{s}_R$	$\widetilde{b}_1,\widetilde{b}_2$
	$\widetilde{\nu}_e$	$\widetilde{\nu}_\mu$	$\widetilde{\nu}_\tau$
	$\widetilde{e}_L,\widetilde{e}_R$	$\widetilde{\mu}_L,\widetilde{\mu}_R$	$\widetilde{\tau}_1,\widetilde{\tau}_2$
	$\widetilde{g}$		
	$\widetilde{\chi}^{\pm}_{1,2}$	(Charginos)	
	$\widetilde{\chi}^{0}_{1,2,3,4}$	(Neutralinos)	

(b)

Table 1.3: The supersymmetric partners of the Standard Model particle fields (a). Mass eigenstates (b) after mixing of the supersymmetric fields, where number labeling is based on increasing mass. The mixing is negligible in the first two generations of $s$quarks and $s$leptons.

strong limits on these parameters creates an unworkable situation from the experimentalist point of view, as one would have to examine the phenomenologies of SUSY models in 105 dimensions.

In gravity mediated MSSM (mSUGRA) models, gravity is the sole messenger between the hidden sector and the MSSM sector. The following assumptions are added to reduce the number of parameters: just like for the coupling constants, it is assumed that at the GUT scale all scalars ($s$quarks, $s$leptons, Higgs bosons) have a common mass m_0, all gauginos and Higgsinos have a common mass $m_{1/2}$ and all the Higgs-$s$fermion-$s$fermion couplings have a common value A_0. There are only five parameters left in total that define a mSUGRA model: $\mathbf{m_0}$, $\mathbf{m_{1/2}}$, $\mathbf{A_0}$, $\boldsymbol{\tan\beta}$ and $\boldsymbol{sign(\mu)}$. The requirement on the Z boson mass to come out at the measured value leads to the definition of $\tan\beta = v_u/v_d$ as the ratio of vacuum expectation values of the two Higgs doublets [56, 57]. And μ is the bilinear coupling between the two Higgs fields, but only the sign is a parameter in mSUGRA models. Using the renormalization group equations one can exactly compute all the particle masses that one expects to measure back at the LHC scale.

The masses of $s$particles in mSUGRA have the following general dependencies on the five parameters. With an increasing m_0 all $s$particle masses rise, except the gauginos and the lightest Higgs. Masses of all $s$particles are raised by an increasing $m_{1/2}$. The mass of the lightest Higgs decreases for small $\tan\beta$. For increasing $\tan\beta$ the heavy Higgses and primarily the third family $s$fermions become lighter. The main effect of lowering A_0 is a small drop in the mass of the lightest Higgs. The last parameter $sign(\mu)$ is the least significant, as it has little effect compared to the other parameters.

1.2.6 Phenomenology of SUSY

The phenomenology of supersymmetry at the LHC is determined by three ingredients [58]. The first is the type of particles produced in the initial interaction, largely determined by the couplings. The most abundantly produced particles are the ones that couple the strongest to the proton beams, hence the supersymmetric sector that carries a colour charge will be the most

important.

The supersymmetric cross section is the second ingredient, which determines the amount of SUSY particles produced with a specific integrated luminosity. This cross section mostly depends on the masses of the *s*particles produced the most in the initial hard interaction. The lighter these particles are, the higher the cross section of the process, the easier it is to find a significant deviation from the SM cross section with the same final state.

The last ingredient describing the phenomenology are the decay chains of supersymmetric particles. These are mainly determined by the available kinematics, the possible couplings and the mass spectrum for each SUSY model. The lighter particles are produced more copiously in the chain than the heavier ones, if the couplings are allowed.

SUSY production mechanisms

Since the typical mass of a SUSY particle lies above 100 GeV, the partons needed to produce these heavy particles at the LHC must carry a high proton momentum fraction $x > 1\%$ [58]. Going back to Figure 1.1 we can see that such high momentum fractions are mostly carried by valence quarks or gluons.

Due to the strong interaction, colour charged particles are produced at the LHC with the highest rates, thus the masses of supersymmetric particles carrying colour charge define the type of particles produced most in the collisions. If gluino masses are small, while the *s*quark masses are large, the gluinos are produced the most in the initial interaction. This is the case for most points in mSUGRA parameter space. As the valence quarks have the highest probability to carry high momentum fraction, first family u/d *s*quarks are produced more often than other flavours, because no strong flavour-changing couplings are predicted for the mSUGRA models.

The exception to this rule is the production of *s*tops, which can have quite a small mass for large values of $|A_0|$, especially if $\tan\beta$ is large. For certain choices of mSUGRA parameter values *s*top production can even become the dominant process [58].

The second phenomenological ingredient, the supersymmetric cross section, for most (some) points in mSUGRA parameter space thus heavily depends on the gluino (*s*top) mass. As the gluino (*s*top) mass decreases, the cross section generally increases, and vice versa. All the other *s*particle masses have a much smaller effect on the total SUSY cross section.

From the Tevatron we have a lower limit of $M_{\tilde{g},\tilde{q}} \gtrsim 400$ GeV at 95% confidence level. The three SUSY models shown in Figure 1.3 are just above this limit with $M_{\tilde{g},\tilde{q}} \approx 400,\ 440,\ 480$ GeV regulated by changes of the m_0 and $m_{1/2}$ parameters. To be able to find or exclude SUSY at the LHC, the SM processes must be understood to a very high precision, as most SUSY model cross sections are orders of magnitude lower than corresponding SM backgrounds.

We repeat that due to R-parity conservation the supersymmetric particles produced from SM particles, can only be produced in pairs. The discussion of the production mechanisms and available kinematics above must be understood in terms of *s*particle production in pairs.

SUSY decay signatures

As mentioned in Section 1.2.3, as a consequence of R-parity conservation, all supersymmetric decay chains must end with the LSP. If the decay chain starts from a gluino, the gluino will typically decay to a *s*quark and a quark, where decays to the lightest *s*tops and *s*bottoms are preferred because of their low mass. However if the *s*quarks are heavier than gluinos, the gluino will generally have a three-body decay into a $q\bar{q}$ pair and a gaugino. The decay chain continues, as *s*quarks decay into quarks of the same family, as mixing between families is negligible, and a chargino or a neutralino. The charginos and neutralinos decay into the LSP, usually $\tilde{\chi}_1^0$, together

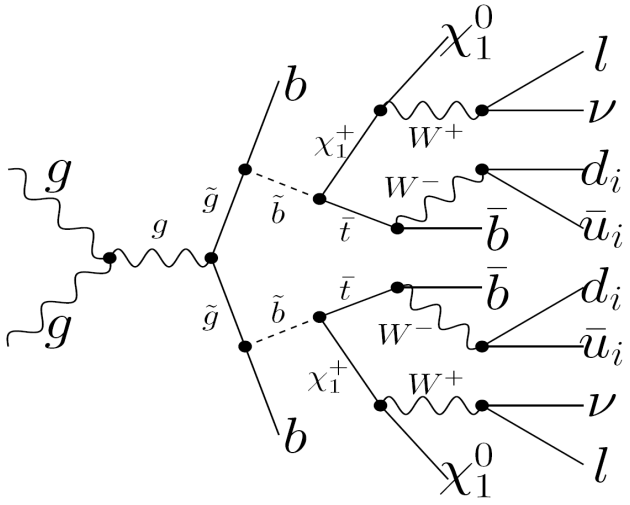


Figure 1.10: An example of a possible production process of two gluinos and subsequent decay cascades resulting into two leptons, eight jets, two neutrinos and two neutralinos in the final state, taken from [49].

with a W or Z boson. If the mass difference between the chargino (neutralino) and the LSP is too small to produce a gauge boson, this decay will be a tri-prong via an offshell boson.

The most important factor in determining the allowed decay chains is the spectrum of low-mass particles, which can be very different for different parts of the mSUGRA parameter space. If the mass of the mother particle barely exceeds the mass of the daughters, due to phase space effects this decay channel will be suppressed. As more phase space is available, the decay becomes more important, that is if the coupling to the mother particle is strong enough. An important consequence of this is that SUSY phenomenology depends on the value of SUSY parameters in a highly non-linear way.

Figure 1.10 shows an example of the production and subsequent long cascade decays of two gluinos at the LHC, taken from [49]. Many energetic final state particles are created in such cascade decays, as in the example shown in the figure a total of eight jets (out of which four are b-jets), two leptons and large missing E_{T} would be measured in a detector. The example shown in Figure 1.10 does not take into account initial/final state radiation nor the underlying event, which lead to even more hard jets in the detector. Supersymmetric events at the LHC are characterized by multiple energetic jets, together with large missing E_{T} not pointing in the direction of the jets, and possibly one or more leptons.

SUSY simulation strategy

The simulation strategy for SUSY models in ATLAS is exactly the same as the strategy followed for the SM processes. The only exception are the programs used to calculate the supersymmetric mass spectrum and the corresponding branching ratios.

The SUSY mass spectrum and branching ratios are generated by the specialized program ISASUSY, part of the ISAJET [59] package, for each mSUGRA model. These are then used as input for HERWIG, that already has a general description of allowed interactions between SUSY and SM particles in the case of R-parity conserving mSUGRA models. Hence the only input from ISASUSY are the exact masses and the specific coupling strengths. HERWIG

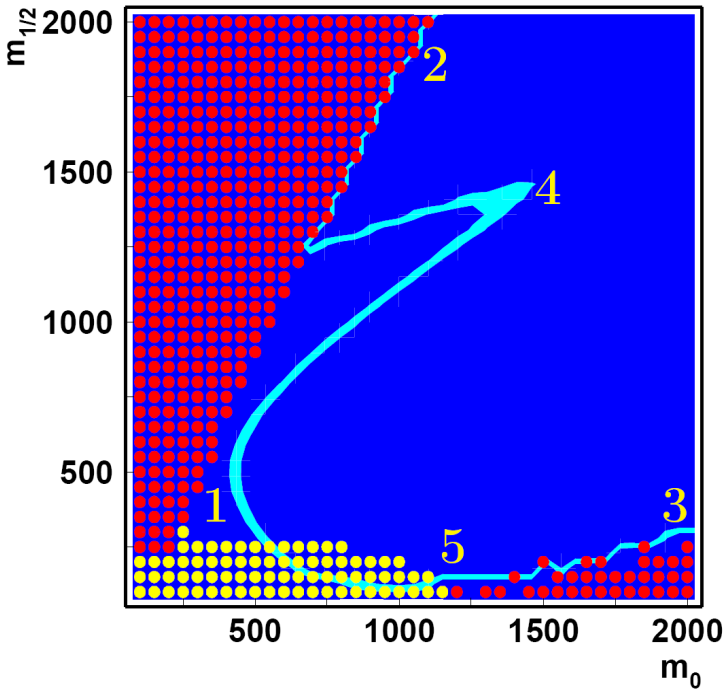


Figure 1.11: The light band portrays the WMAP allowed region. The excluded regions in dots where the $s$tau is the LSP (upper left), due to radiative corrections to the Z boson (lower right) and the LEP limit on the Higgs mass (lower left). The numbers denote the different regions: the bulk region(1), the coannihilation region(2), the focus point region(3), the funnel region(4) and the low mass region(5). Figure is taken from [49] for $\tan\beta = 51$.

completes all the other aspects of simulation, as was explained in Section 1.1.8.

The cross section for each mSUGRA model is calculated using the PROSPINO [19] program. Besides the leading order cross section it also delivers the k-factor, that gives the ratio between the LO and NLO cross sections to compensate for missing higher order terms. Once SUSY models go through the complete simulation procedure, their cross sections are multiplied by the respective k-factor.

1.2.7 Constraints on SUSY parameter space

Several experimental constraints exist on the available mSUGRA parameter space, of which we will now discuss the most important ones.

The strongest bound on allowed mSUGRA space is the dark matter relic density, as measured by the WMAP collaboration [51]. If R-parity is conserved and therefore the LSP is stable, it must contribute to the total amount of dark matter in the universe. The supersymmetric contribution to the total amount of dark matter is equal to the number of LSPs times its mass. Hence to obtain the low dark matter relic density measured by WMAP, one either needs a light LSP or one needs to decrease the number of LSPs in the early universe, through annihilation into SM particles. The annihilation process does not break R-parity conservation, since two SUSY particles are in the initial state and only SM particles are in the final state, but it does reduce the number of LSPs.

Another strong bound is given by the limit set by LEP on the mass of the Higgs boson to be $m_H > 114.4$ GeV. Finally the last bound we will discuss is on the nature of the LSP. In some mSUGRA models the *s*tau turns out to be the LSP, however there are very strong constraints on the dark matter candidate being electrically neutral. Hence this region is also excluded.

Figure 1.11 shows the allowed and the excluded regions in the $m_0 - m_{1/2}$ plane for $\tan\beta = 51$. The light band is the only allowed region, mostly due to the WMAP constraint. The upper left dotted region is excluded since the *s*tau is the LSP for these values of m_0 and $m_{1/2}$, while the lower left dotted region is excluded due to LEP measurements on the Higgs mass. The lower right dotted region is excluded because the radiative electroweak symmetry breaking mechanism would not work there, or in other words because the Z boson mass has been measured and mSUGRA models must predict it correctly. The allowed and excluded regions are strongly dependent on the choice of the mSUGRA parameters, specifically on $\tan\beta$, hence Figure 1.11 should be viewed as an example for large $\tan\beta$ values.

Allowed regions

Besides the allowed and excluded regions, Figure 1.11 shows also the benchmark regions denoted by numbers. The benchmark regions find their naming mostly from the way the dark matter constraints are fulfilled, these regions are defined as:

- *Coannihilation region*: Coannihilation region is defined by a loophole in the calculation of the total dark matter relic density. In this region the masses of the *s*tau and the LSP are very close, so close in fact that the *s*tau decay to a tau and an LSP is suppressed. This means that in the early universe the LSP and the light *s*tau coexist. As a result of the *s*tau-neutralino cross section being much higher than the neutralino-neutralino cross section, the *s*tau and neutralino coannihilate. This leads to a lower dark matter relic density.

- *Focus Point region*: Focus Point region is characterized by the high Higgsino component of the LSP. This enhances the $\tilde{\chi}_1^0\tilde{\chi}_1^0 \to WW$ annihilation in the early universe.

- *Bulk region*: In the Bulk region the LSP annihilation in the early universe happens through the exchange of light *s*leptons.

- *Low Mass region*: Low mass region is close to the Tevatron excluded parameter space. It has a much higher cross section compared to the other regions.

- *Funnel region*: Funnel regions are another loophole to lower the dark matter relic density. If the mass of the LSP is of the order of half the mass of the Higgs boson, then annihilation is enhanced because it takes place near a resonance in the cross section.

Particle spectra

The allowed regions are used by ATLAS to define *benchmark points* listed in Table 1.4, with the aim of exploring sensitivity to a wide class of final state signatures [60]. Besides the m_0 and $m_{1/2}$ values, also the other mSUGRA model parameters vary between the different benchmark points, hence the connection between the allowed regions in Figure 1.11 and specific points is only approximate.

We will now briefly discuss the phenomenologies of the ATLAS benchmark points. The different points with corresponding characteristics are:

Name (Label)	m_0	$m_{1/2}$	A_0	$\tan\beta$	sgn(μ)	σ^{NLO} (pb)
Coannihilation (SU1)	70	350	0	10	+	10.86
Focus Point (SU2)	3550	300	0	10	+	7.18
Bulk (SU3)	100	300	-300	6	+	27.68
Low Mass (SU4)	200	160	-400	10	+	402.19
Funnel (SU6)	320	375	0	50	+	6.07

Table 1.4: Five of the ATLAS benchmark points in the mSUGRA parameter space, where m_0, $m_{1/2}$ and A_0 are given in GeV. The cross sections listed are at 14 TeV center-of-mass energy.

- SU1 (*Coannihilation region*): Although this point is characterized by a relatively low *s*top mass, the main production goes via gluinos and u/d *s*quarks, as *s*top production knows, at leading order, only a few diagrams compared to many possible diagrams to produce gluinos and u/d *s*quarks [61, 62]. Besides that, u/d *s*quarks have both $\widetilde{u}_L/\widetilde{d}_L$ and $\widetilde{u}_R/\widetilde{d}_R$ mass eigenstates that are nearly degenerate in mass, while the light *s*top has only the $\widetilde{t}_1$ light mass eigenstate. The produced gluinos/*s*quarks have many different decay channels involving neutralinos/charginos and SM quarks/gluons. Specific to this benchmark point is the small mass difference between the LSP and the lightest *s*lepton (*s*tau). Thus soft leptons (taus) are characteristic decay products.

- SU2 (*Focus Point region*): This point has very high scalar masses, so all the production goes via gluinos and gauginos. The gluinos decay through three-body decays to a lighter gaugino and a W/Z or Higgs boson. To recognize this scenario at the LHC we would have to look for $E_{\mathrm{T}}^{\mathrm{miss}}$ signals combined with signals from the SM bosons.

- SU3 (*Bulk region*): This point has a *s*top/light-*s*quark/gluino mass ratio comparable to SU1, so the production characteristics described for SU1 apply here as well. However the mass difference between the LSP and the lightest *s*lepton is not as small, so no soft leptons are expected. Thus these events are marked by a plethora of decay possibilities into leptons, multiple jets and missing transverse energy. As the total cross section for SU3 is approximately a factor two higher than for SU1, the sensitivity of ATLAS to find it is raised, so that less integrated luminosity is needed to either discover or exclude it.

- SU4 (*Low Mass region*): Although the *s*top mass is as low as 206 GeV for this point near the Tevatron exclusion bound, the dominant initial process is still gluino production. However *s*tops and tops are copiously produced in gluino decays, leading to a typically high number of produced *b*-jets. This point is characterized by high multiplicity of jets due to long decay chains and many leptons, as the light gauginos can only decay to the LSP through a three-body decay.

- SU6 (*Funnel region*): For the SU6 point the light gauginos are heavier than *s*taus, but lighter than all other *s*fermions, thus *s*taus are produced abundantly. These decay predominantly to a tau and the LSP, so the characteristic of this point is an abundance of taus in the final state.

Although many different regions are addressed by the benchmark points, to study the evolution of phenomenologies and to cover a larger parameter space a grid of mSUGRA points was produced by ATLAS, as described in more detail in Chapter 4.

2

The LHC and the ATLAS detector

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [63] is a 26.7 km long circular machine built at CERN to accelerate and collide protons. It is situated 45-170 m underground in the circular former LEP [64] accelerator tunnel on the border of Switzerland and France, close to the city of Geneva. The LHC is designed to study the Standard Model (SM) at a new energy frontier and reveal physics beyond the SM. As discussed in Chapter 1 this new physics has small cross sections which are proportional to the center of mass energy of the collisions. The exploration of these rare events requires the machine to run with both high beam energy and high beam intensity.

Being a proton-proton collider, the LHC must have two rings to accommodate counter rotating beams, as opposed to a particle-antiparticle collider such as the Tevatron [65] that can share the phase-space for both beams in one ring. However the cost-saving choice of using the 3.7 m diameter LEP tunnel put strong restrictions on the design of a proton-proton collider. This hard limit on tunnel space led to the adoption of the twin-bore magnet design. The two proton beams are bent in the dipole magnets, which are embedded in one single iron yoke. The disadvantage of this scheme is that beam flexibility is diminished.

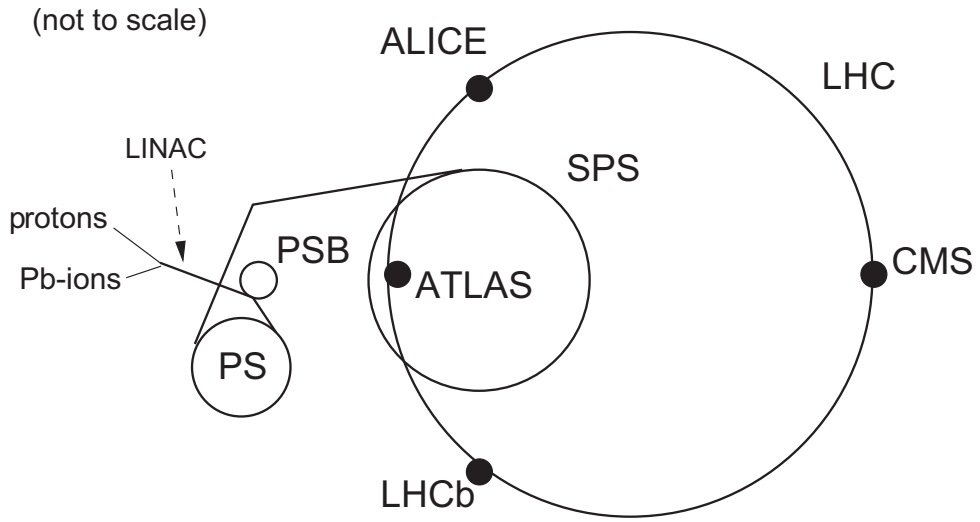


Figure 2.1: Overview of the CERN accelerators and four experiments. ALICE: A Large Ion Collider Experiment. LHCb: Large Hadron Collider Beauty experiment. CMS: Compact Muon Solenoid. ATLAS:A Toroidal LHC ApparatuS.

The design specifications of the LHC are to collide protons on protons at a maximum center of mass energy of 14 TeV. Hydrogen gas atoms are ionized to deliver the protons. The bunches of protons are first accelerated in a dedicated linear accelerator (LINAC) to an energy of 50 MeV. Injected into the Proton Synchrotron Booster (PSB) they are accelerated to 1.4 GeV. The PSB passes the protons on into the Proton Synchrotron (PS), that passes them into Super Proton Synchrotron (SPS) while increasing the beam energy to 26 and 450 GeV respectively. By this point in time the protons are moving at a speed very close to the speed of light. The protons from SPS are then injected into the LHC, where the last acceleration step before collisions takes place in the RF-cavities that accelerate them to a maximum of 7 TeV per beam. A schematic view of the whole accelerator complex is given in Figure 2.1. To keep the protons in a circular trajectory more than 1200 dipole magnets need to provide a magnetic field of 8.36 Tesla. The only technologically feasible option is to use superconducting magnets. In the LHC the super-conducting coils are made of copper-clad niobium-titanium, that is cooled to 1.9 K by super-fluid Helium.

2.1.1 The physics at the LHC

The number of events produced at the LHC is given by:

$$N = \mathcal{L} \cdot \sigma, \tag{2.1}$$

where σ is the cross section of the physics process and $\mathcal{L}$ is the machine instantaneous luminosity. The instantaneous luminosity is defined as the number of protons that pass by per unit area per unit time. The design of LHC foresees to increase the instantaneous luminosity to 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which corresponds to an integrated luminosity of 100 fb^{-1} per year, almost two orders of magnitude higher than Tevatron. For a typical mSUGRA cross section of 10 *pb*, this would amount to 10^6 supersymmetric events at each interaction point every year. To reach such high instantaneous luminosity each bunch in the LHC will consist of 10^{11} protons and close to 3000 bunches will be circulated, while the bunch spacing will be 25 ns. The downside is that with an inelastic proton-proton cross-section of 80 *mb*, on average 23 inelastic collisions will occur per bunch-bunch crossing, thus crowding up the detectors and setting very stringent criteria on design and performance for the experiments.

Unlike a lepton collider like the LEP, the effective center-of-mass energy of the interactions at the LHC is different for each event. The maximum design center-of-mass s is 14 TeV, but all the hard interactions will be of lower energy as we collide protons that are composite particles. The two partons, either quarks and gluons, carry a fraction (x_1 and x_2) of the proton momentum thus providing the effective collision energy of $\sqrt{x_1 x_2 s}$. As the fraction carried by one of the two partons is higher, the particles produced in the hard interaction will have a boost along the beam axis. Hence looking from the detector-frame the total momentum of the final state particles will only be conserved in the plane transverse to the beam direction.

On September the 10th of 2008 the first proton beams were successfully circulated inside the LHC in both directions. Nine days later a major accident occurred caused by bad soldering between two of the superconducting magnets, causing the LHC to shut down, investigate, repair and restart a year later. The careful investigation showed that more safety features need to be installed on the LHC magnets before it can operate at the design energy of 14 TeV. In March 2010 for the first time LHC accelerated protons to 3.5 TeV per beam, the highest energy ever reached by an accelerator. Shortly after the proton beams were collided at a center-of-mass energy of $\sqrt{s} = 7$ TeV. At this energy the LHC has been colliding protons in 2010 and will continue to do so in 2011 and 2012 until the experiments have collected an integrated luminosity

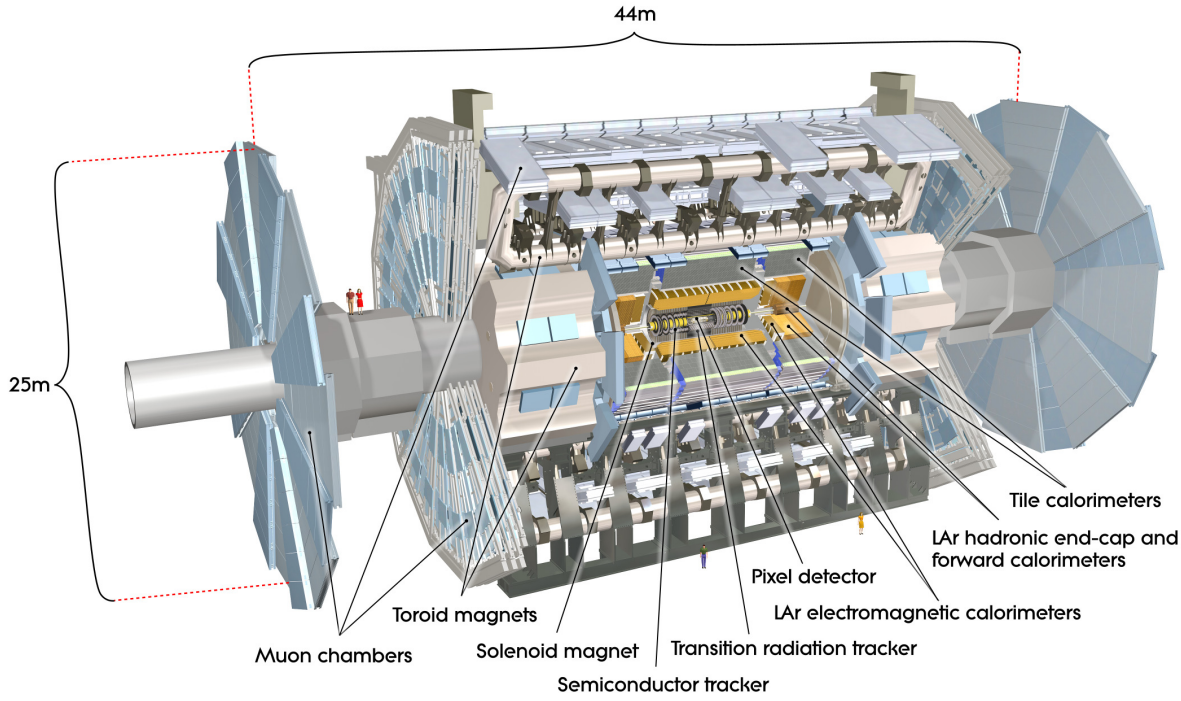


Figure 2.2: Cut-away view of the ATLAS detector. The dimensions of the detector are 25 m in height and 44 m in length.

of $\int \mathcal{L}dt \sim 1$ fb^{-1} , while stepwise increasing the instantaneous luminosity to 10^{33} cm^{-2}s^{-1}. The dataset collected so far by the experiments has shown them to be in very good shape for discoveries as performance and re-discovery of the Standard Model proceed at a swift pace. After the experiments have collected enough integrated luminosity, the LHC will go through an additional maintenance shutdown, to be able to operate at the design center-of-mass-energy and instantaneous luminosity.

Six experiments have been constructed to take data at the LHC. Two multipurpose experiments, ATLAS [66] and CMS [67], are to explore the full spectrum of proton-proton physics. In addition to protons LHC can also accelerate and collide lead ions. The ALICE [68] experiment is set up to study the physics of quark-gluon plasma formation occurring with ion collisions. LHCb [69] is a B-physics experiment that will study CP-violation and rare decays in B-hadron decay. Two smaller experiments TOTEM [70] and LHCf [71] will respectively study the total pp cross section on one side, ultimately important for all LHC experiments, and the energy distributions of very forward produced particles on the other, which will help scientists to interpret and calibrate large-scale experiments for ultra-high energy cosmic rays by analyzing cascades produced at the LHC.

2.2 The ATLAS detector

The ATLAS experiment (A Toroidal LHC ApparatuS) is designed to study a wide range of physics processes in the TeV region. As a typical colliding beam detector it has an approximate cylindrical symmetry. Untypical however is its enormous size that can be appreciated by the reader in Figure 2.2, where human figures are displayed for comparison. These proportions

are mostly due to the air-core toroid magnet in the muon spectrometer, further discussed in Section 2.5. The detector is organized in a central *barrel* where the detector elements form cylindrical layers around the beam pipe, and the two *endcaps* where the detector elements are organized in wheels. In the next sections we will discuss each sub-detector system, starting from the inner detector which is closest to the beam-pipe.

The origin of the ATLAS coordinate system is defined as the center of the detector, coinciding by design with the nominal LHC interaction point. The z-axis points along the beam-pipe in the anti-clockwise direction. The x-axis is oriented towards the center of the LHC ring and the y-axis is perpendicular to it and points upward.

The symmetric cylindrical design of the detector suggests the use of a polar coordinate system. The z-axis remains as before, the azimuthal angle ϕ is the angle in the xy-plane originating from the x-axis and R is the radial component. The polar angle θ is defined as the angle with the positive z-axis, however the variable commonly used is the pseudorapidity $\eta = -\log(\tan(\theta/2))$. In the limit of a massless particle, it is equal to the rapidity $y = \frac{1}{2}\log\frac{E+p_z}{E-p_z}$. The reason for this transformation of the polar coordinate is the fact that particle multiplicity is approximately constant as a function of η (further referred to as rapidity), and the difference in the rapidity of two particles is invariant under Lorentz boosts along the beam axis.

The data taking period of 2010 was very successful for the ATLAS experiment, and the relevant performance results for each sub-detector will be discussed in the respective sections. For now we mention that the fraction of operational channels was 99% on average for the whole of ATLAS, all of the sub-detectors reaching an fraction of operational channels above 97%. Furthermore, the percentage of recorded luminosity as compared to the total luminosity delivered by the LHC was above 93%.

2.3 Inner detector

At design luminosity approximately 1000 particles will emerge from the collision point within $|\eta| < 2.5$ every 25 ns, creating a very large track density in the detector. To achieve the momentum and vertex resolution requirements imposed by the benchmark physics processes, high-precision measurements must be made with fine detector granularity. Determining particle momenta with a high precision and reconstructing the primary and secondary vertices is the purpose of the inner detector (ID). The inner detector is immersed in a 2 T solenoidal field. Pattern recognition, momentum and vertex measurements, and electron identification are achieved with a combination of discrete, high-resolution semiconductor pixel and strip detectors in the inner part of the tracking volume, and straw-tube tracking detectors in its outer part. The outer tracking volume also has the capability to generate and detect transition radiation for charged particles passing through, helping in particle identification. The lay-out of the ATLAS inner detector is illustrated in Figure 2.3.

2.3.1 Beam pipe

The 38 m long beam pipe section in the ATLAS experimental area consists of seven parts, bolted together to form a ultra-high vacuum system. The central chamber is centered around the interaction point and is integrated and installed with the pixel detector. It has an inner radius of 29 mm and a nominal outer radius of 34.3 mm. To reduce the amount of material to an absolute minimum, the central beam-pipe has been manufactured from beryllium with a thickness of 0.8 mm. The remaining six chambers, made of stainless steel, are installed symmetrically on both sides of the interaction point.

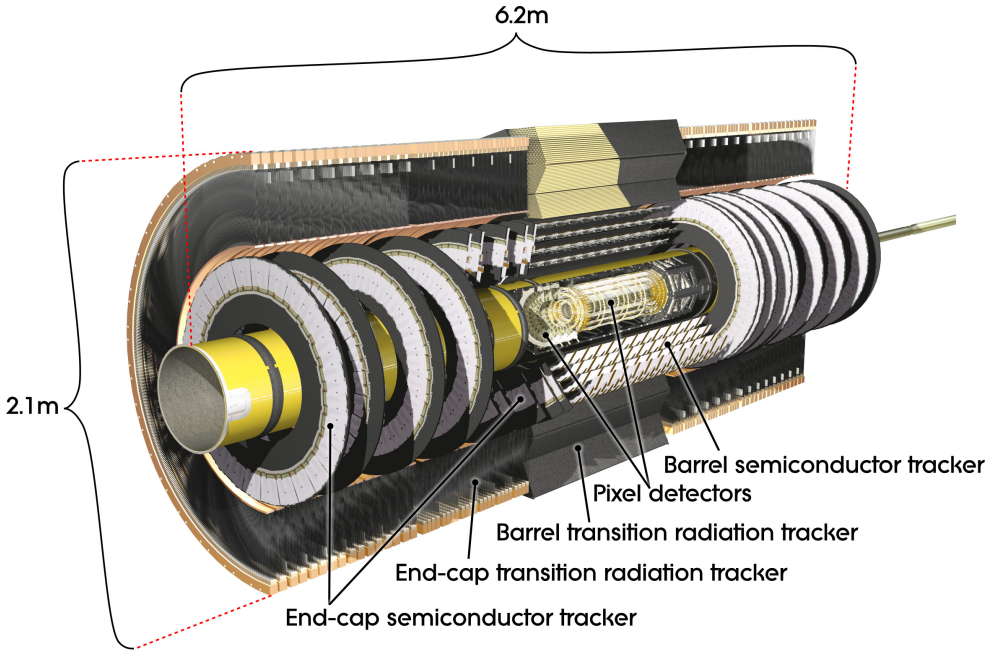


Figure 2.3: Cut-away view of the ATLAS inner detector.

2.3.2 The pixel detector

Purpose The pixel detector was built to measure the position of charged particle tracks at the highest possible precision and is of great importance for a good vertex resolution and the performance of b-jet identification algorithms, which rely on the relatively large life time of approximately 1.5 ps ($c\tau \sim 450$ μm) of B-hadrons. If the B-hadrons are produced from b quarks with a significant boost γ, as is the case in decay of top quarks, the resulting flight path leads to the characteristic signature of a displaced secondary vertex for b-jets.

Detection Principle The detection principle for charged particles is the measurement of charge deposition induced by ionization in a charge depleted layer of silicon. The amount of charge deposited in a single pixel is recorded by measuring the *time-over-threshold* of a signal with a nominal threshold of 0.5 fC (approximately 3000 e^-). As each ionizing particle will deposit some charge on adjacent pixels, the hit position is determined by locating the center of the struck pixels cluster.

Design Considerations Being the closest detector to the interaction point, the pixel detector has very high granularity. It is made up of three concentric cylindrical layers in the barrel and three disks perpendicular to the beam axis in each forward region. While the beam-pipe extends to 34.3 mm in radial direction, the three barrel layers get as close as possible to the beam pipe with 50.5, 88.5 and 122.5 mm respectively, while keeping some distance to allow for services. The layers and disks are equipped with silicon sensors that are segmented into *pixels*. Most pixels have a size of $50 \times 400 \mu m^2$ and in total 80.4 million pixels are read out.

The high-radiation environment imposes stringent conditions on the inner-detector sensors, on-detector electronics, mechanical structure and services. Over the ten-year design lifetime of the experiment, the pixel inner vertexing layer must be replaced after approximately three years of operation at design luminosity.

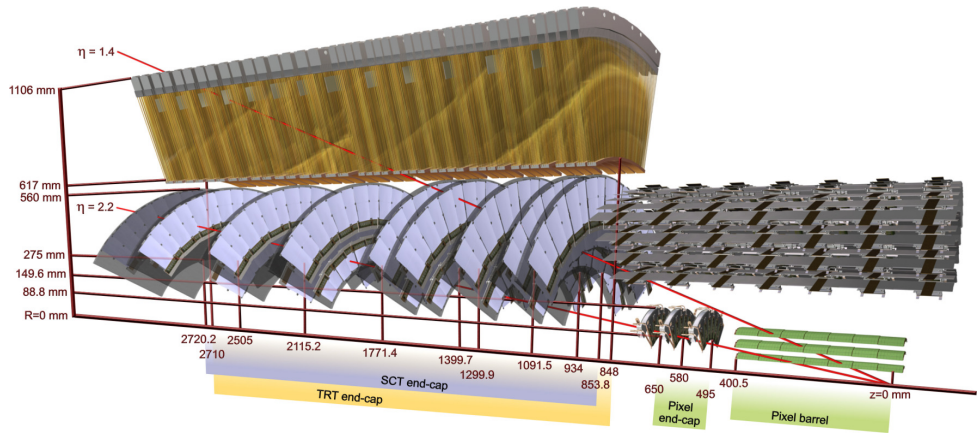


Figure 2.4: Drawing showing the sensors and structural elements traversed by two charged tracks with a p_T of 10 GeV in the inner detector (η=1.4 and 2.2). The track at η=1.4 traverses successively the beryllium beam pipe, the three cylindrical silicon-pixel layers, four of the SCT disks with double layers and approximately 40 straws contained in the end-cap transition radiation tracker wheels. In contrast, the endcap track at η=2.2 traverses successively the beryllium beam pipe, only the first of the cylindrical silicon-pixel layers, two endcap pixel disks and the last four disks of the end-cap SCT. The coverage of the endcap TRT does not extend beyond $|\eta|$=2.

The pixel modules have an intrinsic resolution of 12 μm in the $R-\phi$ coordinate and 110 μm in the z coordinate according to testbeam results [66]. For very low momentum tracks below 1 GeV , such as $K^{\pm}$ coming from decay of the ϕ meson, the time-over-threshold can be used as a specific energy loss dE/dx measurement [72].

2.3.3 The Semi-Conductor Tracker

Purpose The Semi-Conductor Tracker (SCT) surrounds the pixel detector. The SCT contributes to the tracking of charged particles, especially at higher momenta where the curvature of the track in the magnetic field is small, the bigger lever arm of the SCT compared to the pixel detector is important.

Detection Principle Just as the pixel detector it uses silicon sensors, but for the SCT these are segmented in *strips*. This design was chosen reflecting the lower particle density, allowing the reduction of the number of readout-channels and the budget.

An SCT module consists of two sensors glued back-to-back. Each sensor has 768 strips with a strip pitch of 80 μm. The strips of the two sensors on each module have a relative stereo angle of 40 mrad, used for a position measurement along the strip length by finding the intersection of the two strips hit by the traversing particle.

Design Considerations In the barrel the SCT modules are arranged in four cylindrical layers recording hits up to $|\eta| < 1.4$ and each endcap is equipped with modules on nine disks covering the $1.4 < |\eta| < 2.5$ region as shown for two simulated tracks in Figure 2.4. The SCT consists of 2112 modules in the barrel and 988 modules in each endcap, amounting to a total number of more than 6 million strips.

The readout of the SCT modules is binary, thus only recording whether the strip was hit or not, limiting the single strip resolution to about 20 μm. The nominal resolution on the intersection coordinate of the two sensors on a single module, or the coordinate parallel to the strip orientation, is $\sim$ 800 μm due to the small stereo angle.

2.3.4 The Transition Radiation Tracker

Purpose The Transition Radiation Tracker (TRT) is the outermost component of the ID. It is designed to provide to the pattern recognition and tracking algorithms many space points on top of the silicon hits, while at the same time it is useful in identifying electrons through transition radiation.

Detection Principle On average a track traversing the TRT will give 36 hits as it passes the straws. The straws are gas-filled tubes with radius of four millimeters and length up to 144 cm. The gas mixture used is $70 : 27 : 3$ of $Xe : CO_2 : O_2$ that is ionized by the passage of a particle. In the middle of each straw is a 31 μm diameter gold-plated tungsten anode wire kept at ground potential. The straw tube is held at approximately -1530 V. The ionized electron-clusters generated by a passing particle drift towards the wire under the applied potential.

In addition the straws are embedded in polypropylene fibers with different indices of refraction. A traversing charged particle will emit transition radiation in the X-ray regime, where the radiation intensity is proportional to the relativistic factor $\gamma = E/m$. This relation is used to differentiate between electrons and pions, as the electron has a mass that is 273 times smaller than a charged pion. The X-rays are absorbed by the sensitive Xenon in the straw gas and their relatively high energy deposits are distinguished from ionization by applying two thresholds in the readout electronics. So when the track momentum is measured by the TRT straws and the silicon detectors, the transition radiation signal can be used as a discriminant between electrons and pions, cross-checking and complementing the calorimeter.

Design Considerations In the barrel the straws are arranged axially along z in 73 cylindrical layers adding up to $52,544$ straws. In the endcaps the straws are pointing towards the beamline, each endcap counting 18 wheels with $319,488$ straws in total. Though the TRT straws are further away from the interaction point, the size of each straw is so much bigger than a pixel or a strip in the SCT, that some straws are expected to have an occupancy of 50% for the LHC running at the nominal design luminosity. To reduce the occupancy rate some straw wires in the barrel are electrically separated in the middle by a glass wire-joint. These straws are read-out on both ends resulting in half the occupancy. The downside of this modification is the inefficiency around $|\eta| = 0$.

The time that it takes for the clusters to reach the wire is converted to a measurement of the distance from the track to the wire. The attained resolution for this drift radius is around 130 micron.

2.3.5 The solenoid magnet

Purpose The inner detector is contained in a superconducting solenoid magnet. The magnet generates an axial field of around 2 Tesla, is 5.3 meters long and is 2.5 meters in diameter. The purpose of the magnetic field is to curve the tracks of charged particles in the inner detector, making the measurement of particle momentum possible by measuring the amount of track curvature.

Design Considerations To achieve the desired calorimeter performance, the layout was carefully optimized to keep the material thickness in front of the calorimeter as low as possible, hence the solenoid windings and LAr calorimeter share a common vacuum vessel. The solenoid magnet is shorter than the ID itself, which makes the field inhomogeneous and weaker in the forward and rear regions. The resolution on the particle momentum measurement is affected by this. The effect can be expressed by looking at the bending power of the magnet. The bending power

is given by the field integral $\int Bdl$ and drops from about 2 Tm at $|\eta| = 0$ to about 0.5 Tm at $|\eta| = 2.5$. For the high rapidity ($|\eta| > 1.85$) tracks a second effect comes into play. The lever-arm or the length of the measured trajectory in the $R - \phi$ plane is reduced, also worsening the resolution.

The solenoid field outside the ID is guided through a return yoke, which is described in more detail in Section 2.4. The return yoke serves the purpose of keeping the stray field outside the solenoid to a minimum, so as not to alter the trajectories of charged particles outside the ID and keeping the magnetic field in other detectors low.

2.3.6 Material budget

Since the inner detector is required to reconstruct the produced particles very precisely, the amount and the sort of non-active material traversed by the particles had to be carefully chosen to have the right balance for minimal interaction, radiation hardness and structural stiffness. For example the silicon sensors are prone to radiation damage, while this effect can be minimized by keeping the silicon cool [8,73]. Keeping the ID permanently cool slows down reverse annealing of the silicon, that increases the charge collection time and worsens the signal efficiency and timing. Besides that, the inner detector has to be cooled because the front-end electronics produce a heat load of ~ 20 kW. The pipes of the cooling services however could also affect the precision of the measurements if the particles interact with them. Another example of material budget considerations is the choice for the SCT end-cap support structure of carbon-fiber reinforced epoxy, which maintains adequate stiffness while minimizing the material in the detector.

The interaction of the traversing particles with the material of the detector has three main effects. It can lead to the deviation of the particle trajectory through multiple Coulomb scattering on one side. The second effect, specifically in the case of electrons, is that energy can be lost through Bremsstrahlung, which is radiation as a consequence of acceleration in the electromagnetic field of the nuclei or the electrons of the traversed material. The last effect is conversion of a photon into an electron-positron pair, which happens to approximately 40% of photons before they reach the electromagnetic calorimeter [66].

The adverse effects of material interaction can be quantified by two important properties:

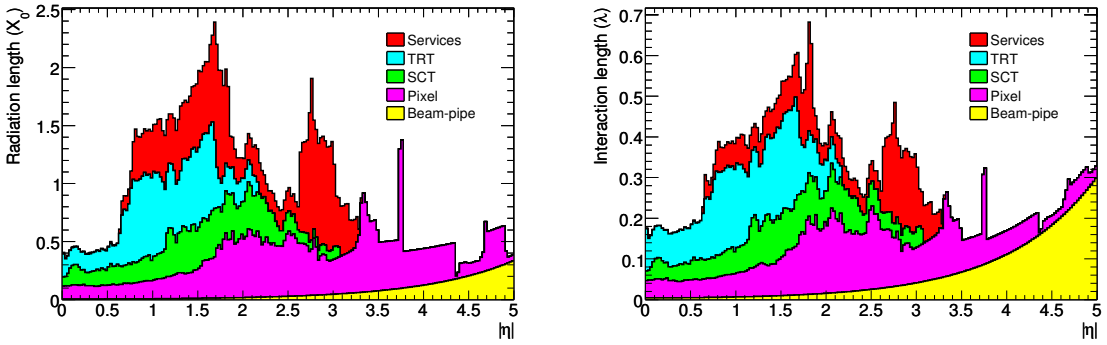
- The radiation length, $\mathbf{X_0}$, is the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung. It is also $\frac{7}{9}$ of the mean free path for pair production by a high-energy photon.

- The nuclear interaction length, $\boldsymbol{\lambda}$, is the mean free path between inelastic collisions. For a homogeneous material $\lambda = A/N\rho\sigma$, where A is the atomic weight, N is the Avogadro number, ρ is the density of the material and σ is the cross section of the incoming particle on the nucleus with weight A.

Figure 2.5 shows the distributions of X_0 and λ as a function of $|\eta|$ for each subdetector as well as the beam-pipe and services. The most striking features are the service and structural material at the interface of the barrel and end-cap regions at $|\eta| \approx 0.7$, and the contribution from the pixel services at $|\eta| > 2.7$.

2.3.7 Track reconstruction in the inner detector

A charged particle traversing a tracking detector will generate a space point in that detector. The task of pattern recognition is to determine which points belong to which tracks and to give an estimate of the track parameters for each track. This information is then given to the



(a) (b)

Figure 2.5: Material distribution (X_0, λ) at the exit of the ID envelope, including the services and thermal enclosures. The distribution is shown as a function of $|\eta|$ and averaged over ϕ. In (a) and (b) the breakdown indicates the contributions of external services and of individual sub-detectors, including services in their active volume.

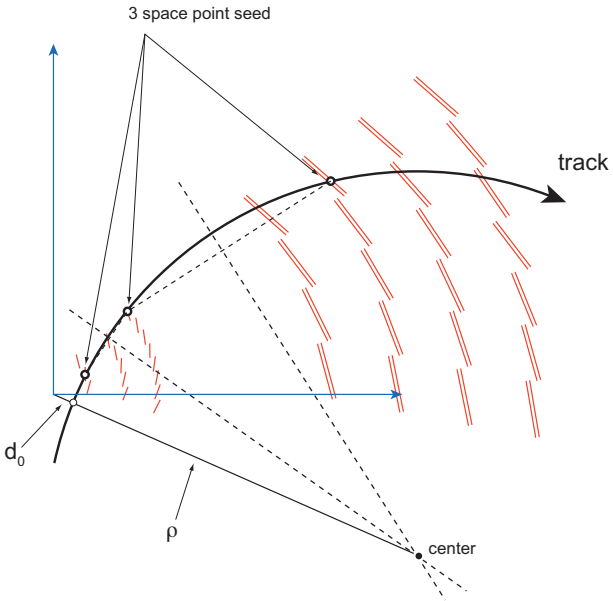


Figure 2.6: A sketch of the technique used to estimate the track parameters of the seeds.

track fitting algorithm, that produces a track trajectory that is as close to the true trajectory as possible.

The primary inside-out track reconstruction sequence used to reconstruct tracks in the ATLAS ID begins with seed finding in the silicon layers of the pixels detector and the SCT. The seeds are then used to build roads, within which hits may be found while moving towards the outer edge of the silicon detector. Finally, an extension to the TRT is probed and the collection of hits is fit to obtain the final track parameters.

Seeds are formed from sets of three space points with each space point originating from a unique layer of the silicon detector. The default of three maximizes the possible number of combinations, while still allowing a first crude momentum estimate to be made. An estimate of the perigee parameters can be made by assuming a perfect helical track model in a constant

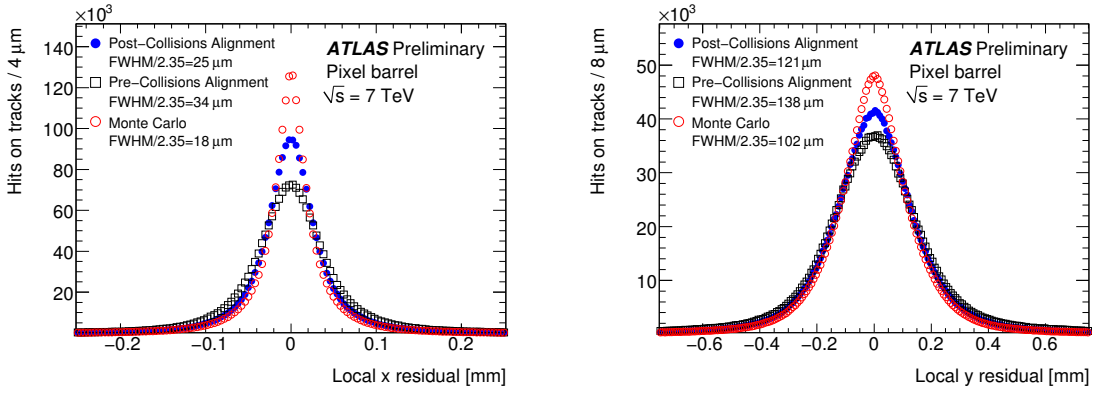
magnetic field. Figure 2.6 illustrates the circle that can be obtained from three space points [74]. The track projected into the transverse plane follows a circular trajectory, which is uniquely described by three parameters: the transverse momentum, p_{T} , the transverse impact parameter, d_0, and the azimuthal angle, ϕ_0. A perfect helical track model ignores effects from multiple scattering and energy loss, which depend on the amount of material the particle has traversed, but is a good first approximation.

The longitudinal parameters are determined by assuming that the track propagates without bending in the rz-plane. The pseudorapidity, η, of the seed is estimated from the average η position of the three space points, for which the angle θ is taken as the angle of the track with the z-axis in the rz-plane. The longitudinal impact parameter, z_0, is estimated from the intersection of a straight line, with the same average η value, with the nominal interaction point.

Once a first estimate of the track parameters is made and ambiguities due to track candidates sharing hits, holes or fake tracks are solved, the track fitting algorithms take over. The two track fitting techniques which are widely used in high energy physics, the *global least-squares fit* and the *Kalman filter* are both implemented in ATLAS as described in [73]. The two different track fitting methods are expected to give almost identical results as they are attempting to find the optimum track trajectory by minimizing the hit residuals of the tracks. Kalman filter is faster and less CPU intensive, but the global least-squares fit has the advantage of explicit calculation of the scattering angles, which can be used to study the material distributions along the tracks. At event reconstruction stage a choice can be made to use one of the two track fitters.

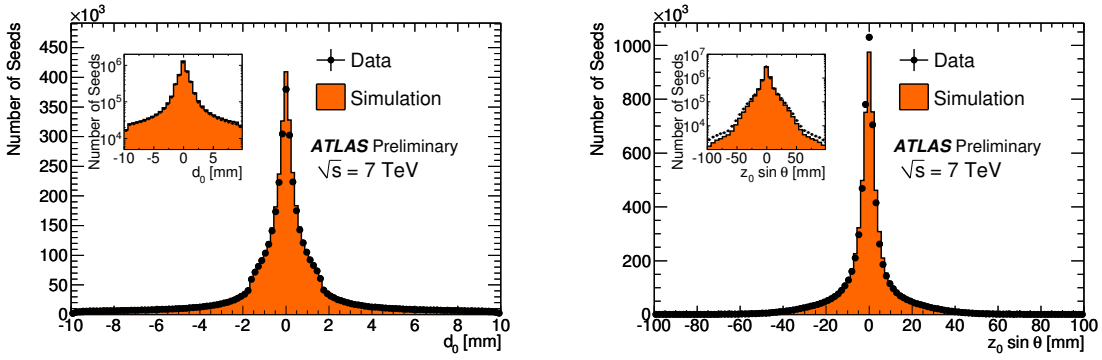
2.3.8 Alignment of the inner detector

After the assembly of the detector, the position of the individual modules is known with much worse accuracy than their intrinsic resolution. Therefore, a track-based alignment procedure has to be applied to determine the absolute position of the sensitive devices. The baseline goal of the alignment of the ATLAS ID is to determine the position and orientation of the modules with such precision that the determination of track parameters is not worsened by more than 20% with respect to those derived with perfect knowledge of the detector geometry. This translates into a requirement on position precision for physics measurements of 10 μm. To reach



(a) (b)

Figure 2.7: Pixel residual distributions integrated over all hits-on-tracks in barrel modules for the local x coordinate (a) and for the local y coordinate (b). In both plots the *Post-Collisions Alignment* shows an improved resolution as compared to the *Pre-Collisions Alignment*, while the perfectly aligned Monte Carlo simulation sample shows the best attainable resolution.



(a) (b)

Figure 2.8: The transverse (a) and longitudinal (b) impact parameter distributions of the seeds in data and simulation. The p_{T} spectrum of the simulation has been reweighted to agree with that for data. The distributions are normalized to the same number of seeds.

a precision of 10 μm on the silicon-module positions, approximately one million good tracks with various topologies are needed. All the track-based alignment approaches are based on the minimisation of hit residuals from high-momentum tracks, which are preferred because of their lower multiple-scattering distortions. The residual is calculated by re-fitting the track with the hit-on-track under study removed.

During the commissioning phases in 2008 and 2009, several million cosmic ray tracks were recorded. These data were used to perform the first alignment of the detector and prepare for the first LHC collisions in 2009. In [75] the alignment using the tracks from first collision events is described in detail. The data sample analyzed comprises 1 million events collected using the ATLAS minimum bias trigger from a single $\sqrt{s} = 7$ TeV proton-proton collision run taken on 23rd April 2010. The events are reconstructed using the *Post-Collisions Alignment*, but also using the *Pre-Collisions Alignment*, in order to illustrate the improvement in the understanding of the ID alignment that was brought about by first collisions. In all other respects the reconstruction of the collisions events is identical between the two. The tracks used in this study are required to pass the following selection criteria: track $p_{\mathrm{T}} > 2$ GeV and number of (SCT + Pixel) hits is above 5. The collision data results are compared to a minimum bias Monte Carlo simulation sample generated using `PYTHIA` with a perfectly aligned inner detector geometry.

Figure 2.7 shows the local residual distributions (the projection of the residual onto the module local direction) for all hits-on-tracks in pixel barrel modules for the x coordinate 2.7(a) and the y coordinate 2.7(b). Quoted is the full width half maximum (FWHM) of each distribution divided by a factor 2.35, as for a Gaussian distribution the FWHM and standard deviation, σ, are related by $\sigma = \mathrm{FWHM}/2.35$. The intrinsic resolution of the detector elements and the track extrapolation uncertainty combine to give the observed width of the Monte Carlo residual distributions. One can see that in general the width of the data residual distributions is reduced using *Post-Collisions Alignment* compared with *Pre-Collisions Alignment*, indicating a significant improvement in the ID alignment after collision tracks have been used. The same behavior is also shown for the end-cap modules of the pixel detector and the SCT and TRT barrel and end-cap modules in [75].

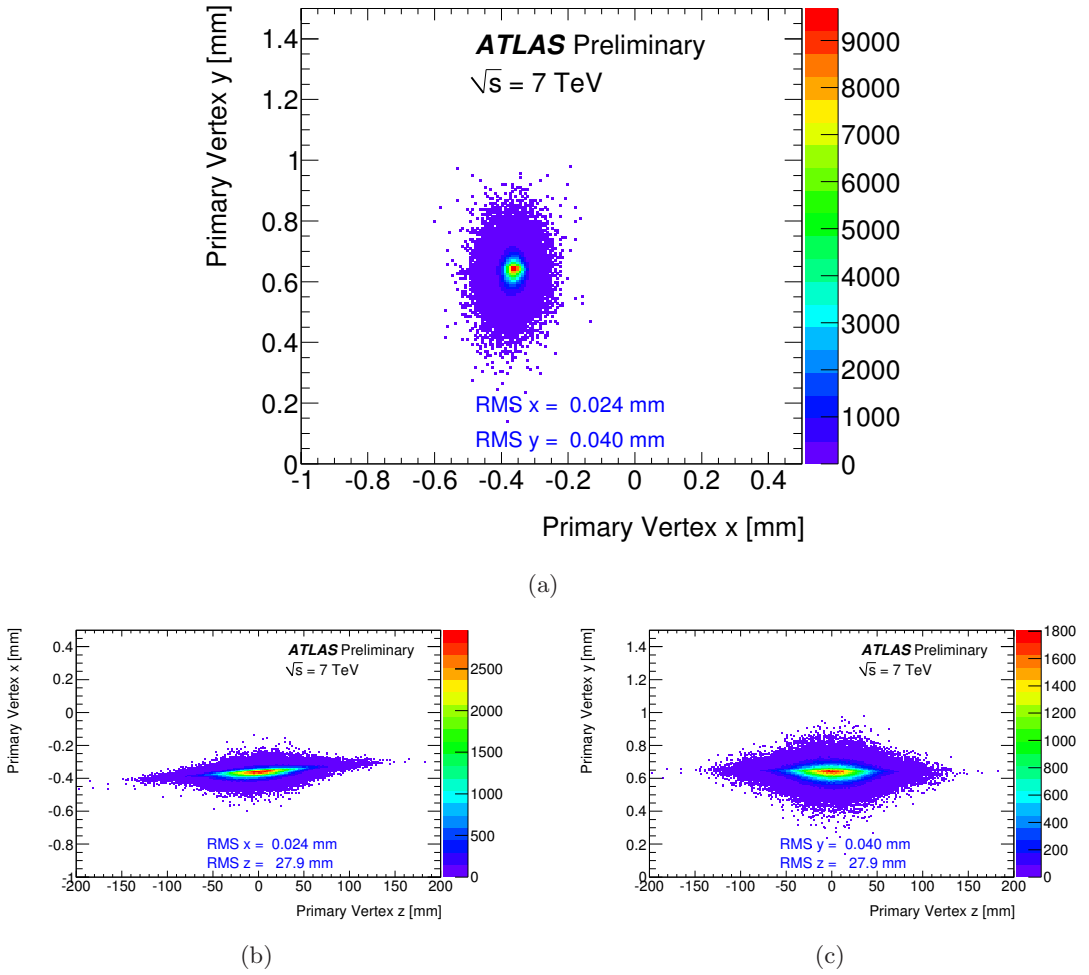


Figure 2.9: Two-dimensional distributions of reconstructed primary vertices in 7 TeV data, in the $x-y$ (a) plane, the $x-z$ (b) plane and the $y-z$ (c) plane.

2.3.9 Performance of the inner detector

Two important track parameters are the point of closest approach in the transverse plane, the transverse impact parameter d_0, and the point of closest approach in the longitudinal plane, which is given by the multiplication of the longitudinal impact parameter z_0 with $\sin\theta$. In [74] a first comparison of the performance of the track seeding algorithms in simulation and data using data taken at $\sqrt{s} = 7$ TeV in April 2010 is performed. Figure 2.8 shows good agreement between simulation and data for both of the parameters, where for simulation non-diffractive minimum bias Monte Carlo simulation was used produced with the `PYTHIA` generator. The impact parameter distributions are calculated with respect to the beam spot, as the primary vertex is not reconstructed at this stage in the pattern recognition. As expected, the width of the transverse impact parameter distribution is smaller than that of the longitudinal impact parameter, which is due to the narrower beam spot width in the transverse plane. The discontinuity in the d_0 distribution, well described by the simulation, is caused by momentum dependent cuts that are applied in the seed-finding algorithm. A discrepancy is observed in the tails of the $z_0 \cdot \sin\theta$ distribution due to the crude θ estimate.

In [76] first estimates are presented of the coordinate resolutions of primary vertices recon-

structed in ATLAS for minimum bias events in 7 TeV proton-proton collisions. About $3.4 \cdot 10^6$ events taken in spring 2010 were used for this analysis, corresponding to an integrated luminosity of approximately 6 nb^{-1} . The reconstruction of primary vertices is organized in two steps: a) the primary vertex finding algorithm, dedicated to associate reconstructed tracks to the vertex candidates, and b) the vertex fitting algorithm, dedicated to reconstruct the vertex position and its corresponding error matrix. It also refits the associated tracks constraining them to originate from the reconstructed interaction point.

Figure 2.9 shows the two-dimensional distributions of the reconstructed primary vertices in the $x-y$, $x-z$ and $y-z$ planes for tracks surviving track quality cuts on the impact parameters, number of silicon hits and $p_{\text{T}} > 150$ MeV. These distributions mostly reflect the size of the beam spot during the studied collisions. It can be noted that the beam widths in x and in y are slightly different, but in z the beam width is orders of magnitude larger. A significant tilt of the luminous region in the $x-z$ plane is also observed. The vertex resolution in the transverse and longitudinal directions has been estimated as a function of the number of tracks at the vertex and of the accumulated transverse momentum. For events with 70 tracks or $\sqrt{\Sigma_{track} p_{\text{T}}^2}$ over 8 GeV the resolution has been measured to be about 30 μm in the transverse plane and about 50 μm in the longitudinal direction. The vertex resolution is expected to improve even more when moving out from the minimum bias regime to higher track multiplicities and values of $\sqrt{\Sigma_{track} p_{\text{T}}^2}$, characteristic for hard collisions.

Although some improvements in performance are to be expected with higher collected integrated luminosity, already at the current stage presented among others in [74–76] the inner detector of ATLAS has performed extremely well and is understood to a very large extent.

2.4 Calorimeters

Purpose The calorimeters of ATLAS identify and measure the energy of both charged and neutral particles. The only particles not directly measured by the calorimeters are the weakly interacting particles, such as neutrinos and the supersymmetric LSP discussed in Chapter 1. However the net energy carried away by these can be indirectly measured by studying the balance sum of transverse energy in an event, for which a hermetic calorimeter is a prerequisite. If the calorimeter system would have an uncovered region, a particle carrying high energy could escape detection and lead to a mismeasurement of the missing transverse energy. The missing transverse energy E_T^{miss} can be seen as the sum of the transverse momenta of all non-interacting particles. Hence one of the main design goals for ATLAS calorimeters is hermetic coverage over $|\eta| < 4.9$.

Design Considerations The calorimeter system is divided into two parts: the electromagnetic (EM) for detecting electrons and photons and the hadronic part for all strongly interacting particles. Using dedicated reconstruction software ATLAS calorimeters can also identify energy depositions of muons traversing its volume as described in [77]. In Figure 2.10 we see a schematic view of the ATLAS calorimeter system surrounding the inner detector.

Detection Principle The calorimeters designed for ATLAS are *sampling* calorimeters. The incoming particles interact with the dense absorber material creating a shower of charged and neutral particles. Interleaved with the absorber is the active material, which detects the energy depositions of the shower particles. The total signal in the active material (*sampling fraction*) is proportional to the real energy of the incoming particle. Ideally all the electron and photon

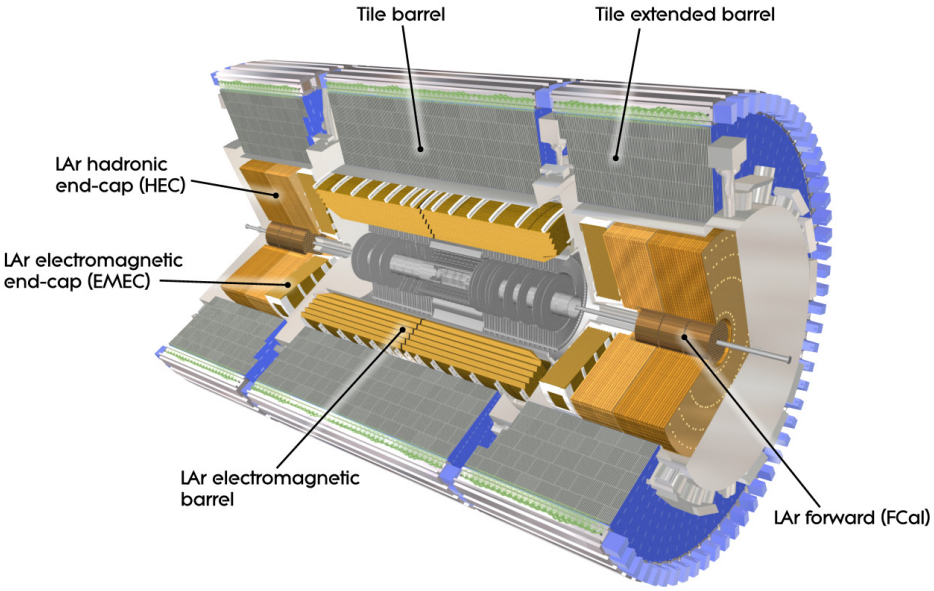


Figure 2.10: Cut-away view of the ATLAS calorimeter system.

showers are contained by the EM calorimeter and all the hadronic showers are stopped by the hadronic calorimeter as to prevent punch-throughs into the muon system.

2.4.1 Electromagnetic Calorimetry

Purpose The mission of high granularity liquid-argon (LAr) electromagnetic calorimeters in ATLAS is high precision measurement and identification of electrons and photons.

Detection Principle The barrel part ($0 < |\eta| < 1.475$) and the two endcaps ($1.375 < |\eta| < 3.2$) shown in Figure 2.10 use the same sampling calorimeter technology. The active medium is liquid argon (LAr) and lead plates covered with thin stainless steel sheets are the absorber. Using LAr as active material gives the benefit of being able to replace it without taking the complete detector apart, whenever deemed necessary. As shown in Figure 2.11(a) the lead plates are accordion-shaped to provide full ϕ coverage without cracks. In between the lead plates electrodes made out of copper and kapton are installed for power supply and read-out. The particles in the shower ionize the liquid argon and the freed charges are collected on the high voltage electrodes. The main difference between the barrel and the endcaps is the orientation of the modules. In the barrel the accordion-wave runs radially, while in the endcap it runs parallel to beam.

Design Considerations The EM calorimeter is designed to stop all electrons and photons coming from the interaction point, while at the same time keeping a low interaction length λ depth. The electromagnetic calorimeter material amounts to $\sim 1.5\lambda$ as compared to approximately 10λ for the hadronic calorimeter. While in terms of radiation length X_0 the total thickness of the EM calorimeter is > 22 radiation lengths in the barrel and $> 24X_0$ in the endcaps (EMEC).

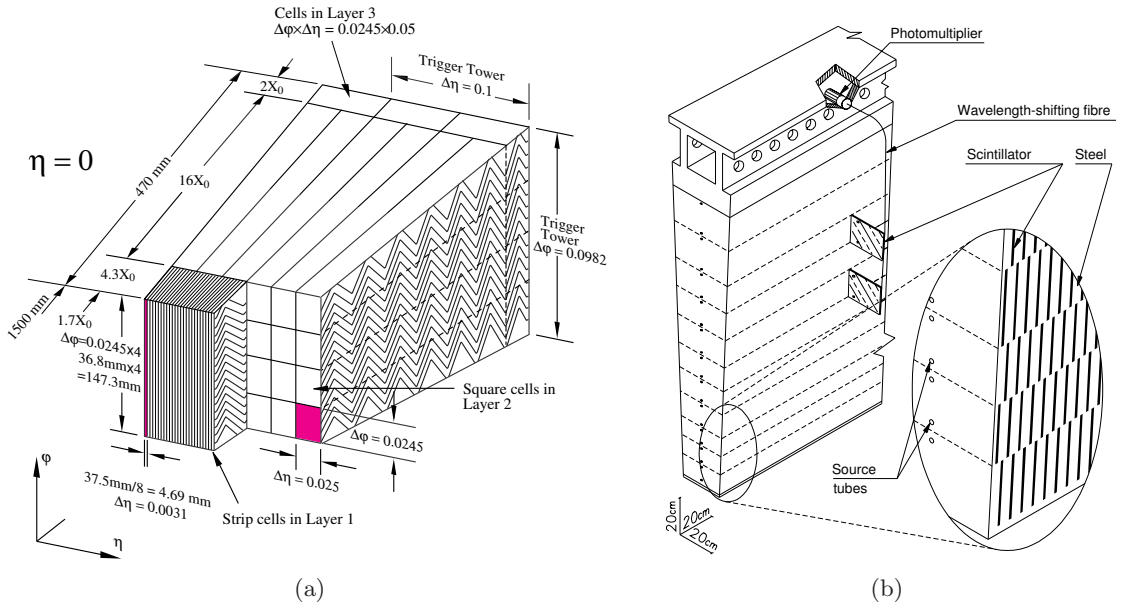
The barrel modules have three layers of sampling with decreasing granularity at larger radius. The inner-most layer is finely segmented in η with granularity of $\delta\eta \times \delta\phi = 0.003 \times 0.1$ and has a thickness of $4.3X_0$. The fine granularity makes for a good separation of γ/e and e/π_0, as photons

and pions decay into a pair of particles while the electron just starts showering. The second layer has an increased $\delta\eta \times \delta\phi = 0.025 \times 0.025$ granularity, but receives the bulk of the energy deposit with a thickness of $16X_0$. The last layer with a granularity of $\delta\eta \times \delta\phi = 0.05 \times 0.025$ and a thickness of $2X_0$ has the purpose of measuring the high energy shower tail and distinguishing electromagnetic showers from the hadronic ones, that deposit most of their energy further away from the beamline.

2.4.2 Hadronic Calorimetry

Purpose The *jets* of particles coming from hadronisation of quarks and gluons (and also hadronic τ-lepton decay) are the subject of measurement for the hadronic calorimeter. The hadronic showers are longer, wider and vary in their development more than the electromagnetic ones and are measured by a combination of the electromagnetic and hadronic calorimeters. The electromagnetic calorimeter is more precise, but expensive to produce and operate, while the hadronic is more coarse, but for a much lesser cost.

Detection Principle The absorber material for the tile calorimeter is steel and scintillator is the active medium as shown in Figure 2.11(b) for a tile module. The steel of the barrel also serves as the return yoke for the solenoid of the inner detector as discussed earlier in Section 2.3. Traversing shower particles create light in the scintillator, which is then collected using wavelength-shifting fiber on each side of the scintillating tile. The fibers are grouped together and are readout by photomultiplier tubes (PMTs), which are housed inside a steel girder to keep the interference from the magnetic field low. A three-dimensional cell structure is defined by the grouping of fibers, with three distinct sampling depths (1.5λ,4.1λ and 1.8λ for the barrel). For the first two layers the rectangular cells are $\delta\eta \times \delta\phi = 0.1 \times 0.1$ big and for the last



(a) (b)

Figure 2.11: Sketch of a LAr barrel module (**a**) where the different layers are clearly visible with the ganging of electrodes in phi. The granularity in η and ϕ of the cells of each of the three layers and of the trigger towers is also shown. The tile calorimeter schematic (**b**) showing how the mechanical assembly and the optical readout are integrated together. The various components of the optical readout, namely the tiles, the fibers and the photomultipliers, are shown.

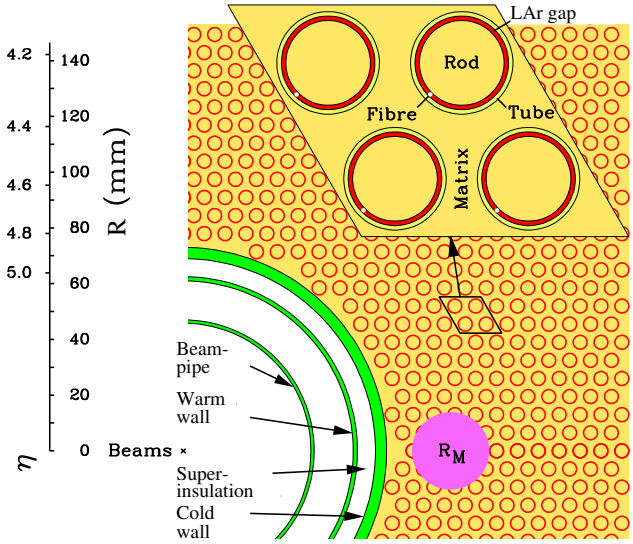


Figure 2.12: Electrode structure of FCal1 with the matrix of copper plates and the copper tubes and rods with the LAr gap for the electrodes.

$\delta\eta \times \delta\phi = 0.2 \times 0.1$. The readout through both sides of the tile give the possibility of averaging the signal and provides a redundant readout link. The orientation of the scintillator tiles radially and normal to the beam line allows for almost seamless azimuthal calorimeter coverage.

The hadronic calorimeter is divided into a barrel part, the *tile* calorimeter, and the endcap calorimeter (HEC). The hadronic endcap calorimeter uses liquid argon as the active material like the EM, but copper as absorber arranged in parallel-plate geometry. Liquid argon is a suitable material for the higher radiation levels, as it can be easily replaced. In rapidity HEC covers the range of $1.5 < |\eta| < 3.2$ and it is subdivided into two wheels. The cell readout granularity is $\delta\eta \times \delta\phi = 0.1 \times 0.1$ for $1.5 < |\eta| < 2.5$ and increases to 0.2×0.2 for $|\eta| > 2.5$.

Design Considerations The material of the barrel tile calorimeter amounts to $\sim 7.5\lambda$ and that of the HEC to $\sim 10\lambda$ which is enough to contain the most energetic showers. The tile calorimeter itself is sectioned into three parts to leave space for cables and services to the ID, the central barrel ($0 < |\eta| < 1.0$) and two extended barrels ($0.8 < |\eta| < 1.7$) as can be seen in Figure 2.10.

2.4.3 Forward Calorimeter

Purpose and Design Considerations The forward calorimeter (FCal) covers the $3.1 < |\eta| < 4.9$ range and serves for both electromagnetic and hadronic calorimetry. During the design phase the extremely high particle fluxes in this region had to be considered. Covering these difficult regions is necessary for hermiticity, that must be observed for precise measurements of missing transverse energy.

Detection Principle The FCal is approximately 10 interaction lengths (λ) deep and consists of three modules in each end-cap: the first (FCal1), made of copper, is optimized for electromagnetic measurements, while the other two (FCal2/FCal3), made of tungsten, measure predominantly the energy of hadronic jets. Each module consists of a metal matrix, with regularly spaced longitudinal channels as can be seen in Figure 2.12. Each channel is filled with concentric rods and tubes parallel to the beam axis with LAr in the gap between the rod and the

tube. The copper rod and tube are separated by a radiation-hard plastic fiber wound around the rod. The three modules (FCal1/FCal2/Fcal3) have a depth of $27.6X_0$, $91.3X_0$ and $89.2X_0$ respectively, which is needed for the extremely high particle fluxes in the forward regions.

2.4.4 Jet reconstruction

The default algorithm employed for jet finding in ATLAS is anti-k_T [78] with distance parameter R = 0.6 [11] or R = 0.4. It is a sequential recombination jet finder that is both collinear and infrared safe, that is important when comparing theoretical predictions with partons to measurements of reconstructed objects. The main idea of the anti-k_T algorithm is that softer input objects are merged with harder objects in order of their closeness in $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

A different algorithm, based on iterative seeded fixed-cone procedure, was the default in the ATLAS collaboration before 2009 and is used in Chapter 4. For this algorithm first all input is ordered in decreasing order in transverse momentum, p_T. If the object with the highest p_T is above the seed threshold and has at least $p_T > 1$ GeV, all objects within a cone in pseudorapidity η and azimuth ϕ with $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < R_{cone}$, where R_{cone} is the fixed cone radius, are combined with the seed. A new direction is calculated from the four-momenta inside the initial cone and a new cone is centered around it. Objects are then (re-)collected in this new cone, and again the direction is updated. This process is re-iterated until the direction of the cone does not change anymore, at which point the cone is considered stable and is called a *jet*. At this point the next seed is taken from the input list and a new cone jet is formed with the same iterative procedure. The jets found this way can share constituents, and a split-merge procedure is implemented in ATLAS. Jets which share constituents with more than half of the fraction of the p_{T} of the less energetic jet are merged, while they are split if the amount of shared p_{T} is below half of the fraction.

Topological cell clusters (*TopoClusters*) are the default input for the jet finding algorithms. Seed cells with a signal-to-noise ratio above a certain threshold, $|E_{cell}/\sigma_{cell,noise}| > 4$, are used to start the clustering. After all directly neighboring cells of the seed cells are collected into the cluster, it is extended with all neighboring cells that cross a lower threshold $|E_{cell}/\sigma_{cell,noise}| > 2$. A ring of guard cells without a specific threshold is added to the cluster to finalize the procedure. When all cell clusters have been identified in an event, the collection is given to the jet finding algorithm.

The jets found by the algorithm are constructed from the raw signals of the calorimeter cells. Since the ATLAS calorimeter is non-compensating, this raw signal has to be calibrated, to account for the difference in electromagnetic and hadronic response. In ATLAS a calibration scheme for calorimeter jets is applied based on cell signal weighting, called *H1* calibration after the experiment that helped develop and refine this approach [79]. To each cell a weight w is applied, which is a function of its location and its signal density, that is defined as electromagnetic energy signal divided by the cell volume. The weighting factor is ~ 1 for high density signals assumed to come from electromagnetic showers, and rising up to 1.5, the typical e/π signal ratio for the ATLAS calorimeters, with decreasing cell signal densities associated to hadronic showers. The weighting factor is applied to both the energy and the momentum terms of the cell four-momentum. All the calibrated cells of the jet are then summed up into a calibrated jet.

The reconstruction of jets in the wide variety of physics processes of interest at the LHC demands among others a very precise understanding of the performance of the jet algorithms. The software not only has to deal with enormous amounts of readout channels, but complications also arise from physics. First of all there is the *underlying event* which produces two jets in the direction of the original protons. Secondly we need to keep *pile-up* events in mind, which

are caused by the 40 MHz design frequency of collisions at the LHC. Particles created at the previous bunch-crossing (25 ns earlier) might interfere with the energy deposition and readout of the interesting hard collision. Lastly most of the interactions between two protons at the LHC are soft or also called *minimum bias* events. At nominal design luminosity on average 23 inelastic collisions will occur per bunch-bunch crossing, which can negatively affect the readout either as pile-up or disbalance in the transverse plane.

2.4.5 Electron reconstruction

Electrons are reconstructed using the dedicated `egamma` [80] algorithm. Electron reconstruction begins with the creation of a preliminary set of clusters in the EM calorimeter. The size of these seed clusters corresponds to 3×5 cells in $\eta \times \phi$ in the middle layer of the EM calorimeter. Electron reconstruction is seeded from such clusters with $E_T > 2.5$ GeV, using a sliding window algorithm over the full acceptance of the EM calorimeter. Electrons are reconstructed from the sliding window clusters if there is a suitable match with a track of $p_{\mathrm{T}} > 0.5$ GeV. The chosen track is the one lying with an extrapolation closest in $\eta - \phi$ space to the cluster centre.

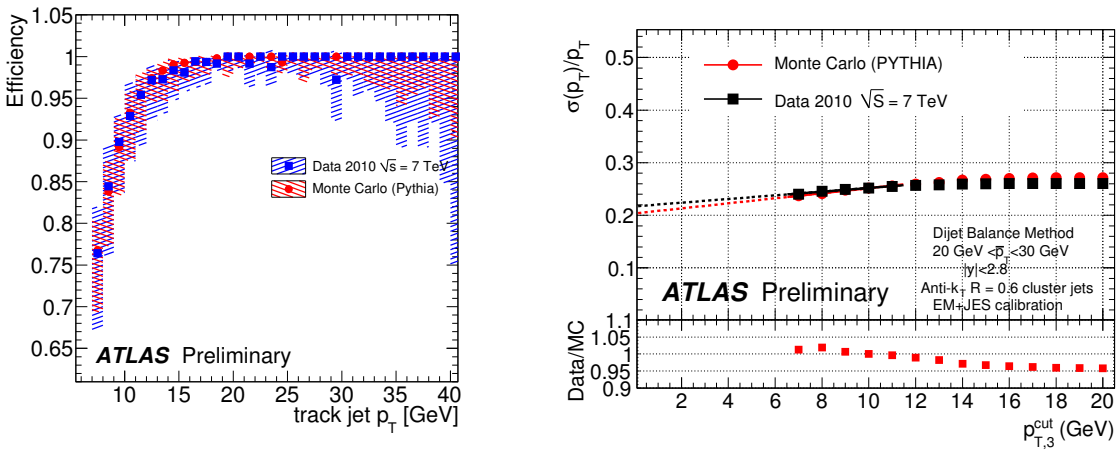
The baseline electron identification algorithm in ATLAS relies on variables which deliver good separation between isolated electrons and fake signatures from hadronic jets. These variables include information from the calorimeter, the tracker and the matching between tracker and calorimeter. Three reference sets of cuts with increasing quality checks have been defined for electrons: *preselection*, *medium* and *tight* [81]. Preselection mostly takes care of cutting out problematic regions in the EM calorimeter and requiring $E_T > 7$ GeV. Medium electron candidates are chosen by looking at the quality of the associated track and the calorimeter deposit. Finally tight cuts make sure that it is an electromagnetic shower by looking at the ratio of energy deposits in EM and hadronic calorimeters, high threshold transition radiation hits and tighter track to calorimeter deposit matching quality cuts.

2.4.6 Performance of the ATLAS calorimeters

Jet energy resolution and jet efficiency A first in-situ measurement of the jet energy resolution and selection efficiency relative to track jets using a sample of proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV is presented in [82] for a total integrated luminosity of 6 nb^{-1} . Track-based jets (track jets) can be built using tracks reconstructed in the inner detector as inputs to the anti-k_T jet finding algorithm. Calorimeter jets are reconstructed with a p_{T} threshold of 7 GeV, and track-jets are reconstructed with a threshold of 4 GeV. A tag-and-probe method is implemented to measure, in-situ, the jet reconstruction and selection efficiency relative to track jets. This technique allows to determine the efficiency to match calorimeter to track jets in a di-jet back-to-back event topology.

Figure 2.13(a) shows the efficiency to match calorimeter jets to probe track jets in data and simulation. The total error is the quadratic sum of the statistical and systematic errors. The calorimeter jet reconstruction is found not to be fully efficient for track jet p_{T} smaller than 20 GeV. The drop in efficiency is mainly due to calorimeter jet energy and angular resolution, and to the effect of the minimum jet p_{T} threshold for reconstructed calorimeter jets of 7 GeV. However above track jet p_{T} of 20 GeV the jet reconstruction efficiency is very close to 100 %.

The di-jet balance method for the determination of the jet p_{T} resolution is based on momentum conservation in the transverse plane. The asymmetry resolution (σ_A) between the transverse momenta of the two leading jets is used to calculate the relative jet resolution as $\sqrt{2}\sigma_A = \sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}}$. Events with two back-to-back leading jets, satisfying at least $\delta\phi \geq 2.8$ between them are required, together with a veto on any extra jets with $p_{\mathrm{T},3}$ above 10 GeV. Although these requirements



(a) (b)

Figure 2.13: Jet selection efficiency (a) relative to track jets as a function of probe track jet p_{T}, in data and Monte Carlo. Resolution (b) versus the upper threshold applied on the third jet for p_{T} bin $20 < \bar{p}_T < 30$ GeV. The solid line corresponds to the linear fit applied while the dashed line shows the extrapolation to $p_{T,3} = 0$. Errors are statistical only and usually smaller than 0.6 %.

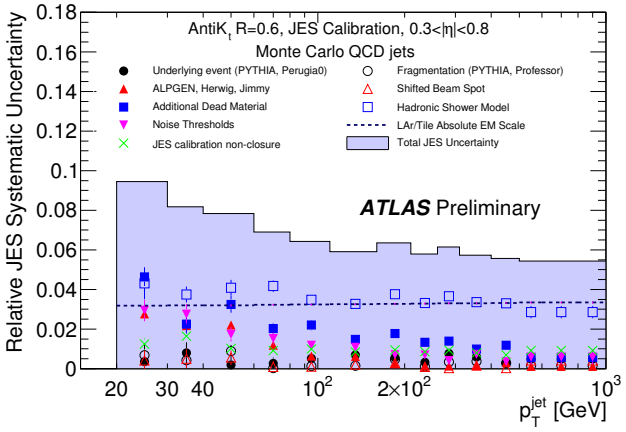


Figure 2.14: Relative jet energy scale systematic uncertainty as a function of p_T^{jet} for jets in the pseudorapidity region $0.3 < |\eta| < 0.8$ in the calorimeter barrel. The total uncertainty is shown as the filled area. The individual sources are also shown, with statistical errors if applicable.

are designed to enrich the purity of the back-to-back jet sample, it is important to account for the effects due to the presence of additional soft particle jets. The dependence of the di-jet balance asymmetry on the presence of a third jet is illustrated in Figure 2.13(b) both for data and Monte Carlo simulation. For the average p_{T} bin $20 < \bar{p}_T < 30$ GeV the resolution is shown as a function of $p_{\mathrm{T},3}$. The jet energy resolutions obtained with the different $p_{\mathrm{T},3}$ cuts are fitted with a straight line and extrapolated to $p_{\mathrm{T},3} \to 0$, to find the resolution for pure di-jet events.

Jet energy scale A correct estimate of the energy of jets (jet energy scale, or JES) is input to many physics analyses and its uncertainty is the dominant experimental uncertainty for measurements such as the di-jet cross section, the top quark mass measurements and new physics searches with jets in the final state.

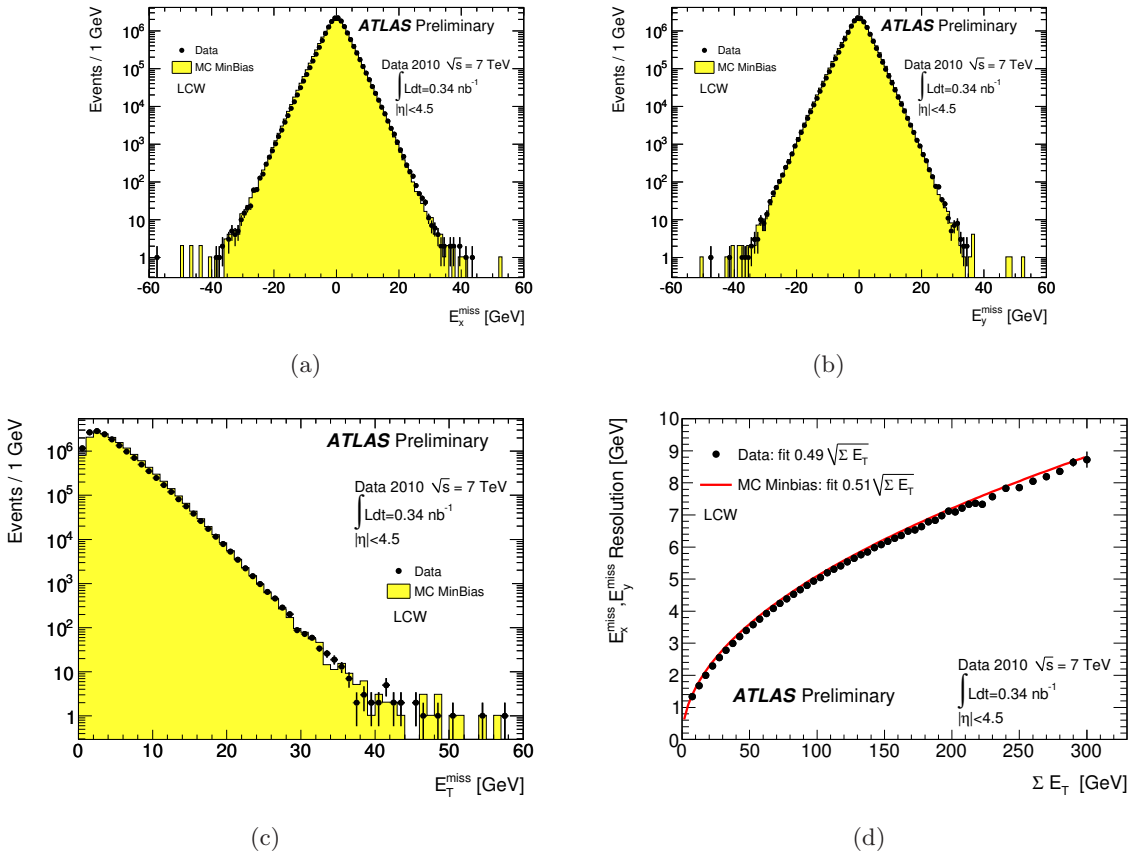
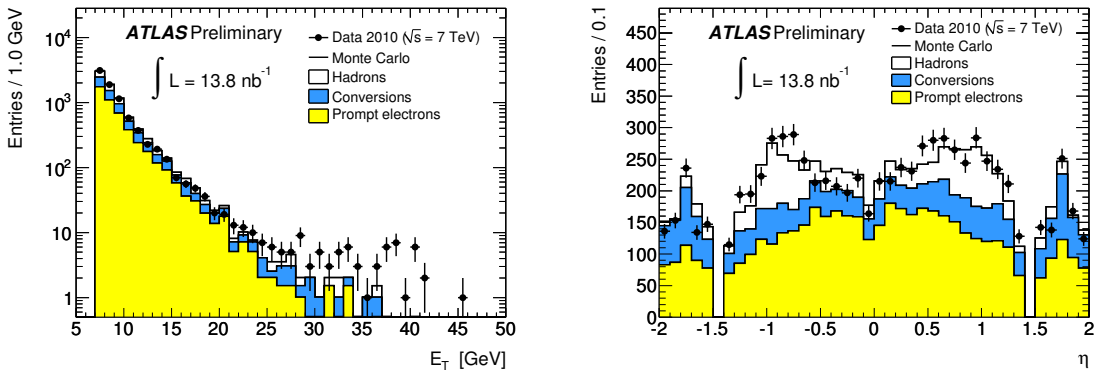


Figure 2.15: Distributions of E_x^{miss}(a), E_y^{miss}(b) and $E_{\text{T}}^{\text{miss}}$(c) as measured in a data sample of 15.2 million selected minimum bias events (dots) at 7 TeV center-of-mass energy, recorded in April 2010. In the calculation only TopoCluster cells are used, with energies calibrated with the LCW. The expectation from Monte Carlo simulation is superimposed (histogram) and normalized to the number of events in data. E_x^{miss} and E_y^{miss} resolution (d) as a function of the total transverse energy ($\Sigma \text{E}_{\text{T}}$) for minimum bias events. The line represents a fit to the resolution obtained in the Monte Carlo simulation and the full dots represent the results from data taken at $\sqrt{s}$= 7 TeV.

In [83] a first determination of the jet energy scale and the evaluation of its systematic uncertainty for inclusive jets measured in ATLAS from proton-proton collisions at $\sqrt{s} = 7$ TeV are described. Reconstructed jets are calibrated as a baseline to the energy scale measured by the calorimeters, called the electromagnetic (EM) scale. The jet energy as measured from the ATLAS calorimeters is corrected for calorimeter non-compensation, energy loss in the material upstream of the calorimeters, shower leakage and out-of-cone effects. The choice of jet energy scale calibration for the first ATLAS data is a jet by jet correction applied as a function of the jet transverse momentum and pseudorapidity.

For inclusive jets with $p_{\text{T}} > 20$ GeV and within $|\eta| < 2.8$, the jet energy scale is determined with an uncertainty smaller than 10% [83]. The JES uncertainty is derived combining information from single pion test-beam measurements, uncertainties on the material budget of the calorimeter, the description of the electronic noise, the theoretical model used in the Monte Carlo generation, the comparison of test beam data for the hadronic shower model used in the simulation and other effects such as a shifted beam spot and the electromagnetic scale uncertainty for the calorimeters. The total JES uncertainty (filled area) as a function of jet p_{T} as well



(a) (b)

Figure 2.16: Distributions of cluster transverse energy E_{T} (a) and pseudorapidity η (b), for electron candidates passing the tight identification cuts. The Monte Carlo sample used here does not contain any electrons from W/Z-boson decay.

as the relative contributions of individual effects are shown in Figure 2.14. The most prominent contributions come from the hadronic shower model and the absolute EM energy scale.

Performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction presently used in ATLAS for physics analysis includes contributions from transverse energy deposits in the calorimeters, corrections for energy loss in the cryostat and measured muons as discussed in detail in Section 4.1.3. Performance of ATLAS $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction in 7 TeV proton-proton collisions is described in [84].

In minimum bias proton-proton collisions the average $E_x^{\mathrm{miss}}(E_y^{\mathrm{miss}})$ is expected to be compatible with zero. This is confirmed by the data and the Monte Carlo simulation of minimum bias events as shown in Figures 2.15(a) and 2.15(b). $E_{\mathrm{T}}^{\mathrm{miss}}$ is the length of the missing transverse energy vector and is non-zero by design. Figure 2.15(c) shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in collision data and `PYTHIA` simulated Monte Carlo sample after applying calibrations based on local calorimeter cell weighting (LCW) [84]. The comparison between data and simulation shows very good agreement.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution approximately follows stochastic behavior in its dependence on total transverse energy ($\Sigma \mathrm{E_T}$), reconstructed from calorimeter cells of TopoClusters as $\Sigma \mathrm{E_T} = \sum_{i=1}^{N_{cell}} E_i \sin\theta_i$. In Figure 2.15(d) the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is shown as a function of the total transverse energy in the event. Superimposed is a fit to the resolution obtained from the simulated sample, which shows that we understand our detector response to a great extent already at this early stage.

Electrons in collisions data In [81] the first observation of inclusive electrons in the collision data collected by the ATLAS experiment at $\sqrt{s} = 7$ TeV is presented, corresponding to a total integrated luminosity of 13.8 $\pm$1.5 nb^{-1} . The Monte Carlo sample used throughout the note is non-diffractive minimum-bias events generated with `PYTHIA`, filtered for total transverse energy from generated particles (except muons and neutrinos) greater than 6 GeV in an area of 0.4×0.4 in $\eta \times \phi$ space, characteristic for high density electromagnetic showers. Figure 2.16 shows the kinematic properties compared to the Monte Carlo predictions for electron candidates passing the tight selection criteria. The Monte Carlo sample is sub-divided into its three dominant components: hadrons, secondary electrons dominated by photon conversions in the detector material, and prompt electrons from semi-leptonic decays of charm and beauty hadrons. One

can note in particular the enhancement of the predicted hadron contribution in the $|\eta| \sim 1$ range corresponding to the transition region between the barrel and end-cap TRT, where electron identification is less powerful. Also noticeable is the dominance of the predicted conversion background at high $|\eta|$, where the amount of traversed material in the inner detector is largest. At the high end of the E_{T} -spectrum, the signal expected from electrons from W/Z-boson decay is clearly visible as an excess above the contributions expected from the three simulation components. Note that the Monte Carlo sample used for this comparison does not contain any electrons from W/Z-boson decay.

2.5 Muon spectrometer

Purpose The muon spectrometer (MS) is the outermost subdetector of ATLAS and is for a great part responsible for the overall dimensions of the experiment. Its purpose is precise measurement of muon momenta, while delivering muon triggers with dedicated trigger chambers. Figure 2.18 gives an overview of the muon spectrometer layout.

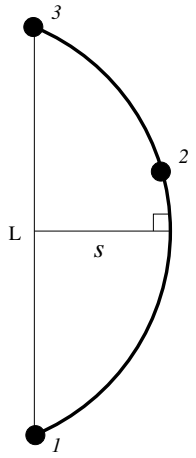


Figure 2.17: Sagitta (s) in three-point measurement. L is the distance between the outer measurements 1 and 3.

Detection Principle The muon momentum can be determined by measuring the position of the muon at three points in space. The trajectory of the muon is curved due to the magnetic field provided by the toroid magnet. From the curvature its momentum is derived. The curvature is measured in the track fit where the magnetic field is known in detail. However for a good approximation and practical application the *sagitta* (s) is used. The sagitta is defined as the maximum deviation of a circle from a straight line, see Figure 2.17, and it is linked to transverse momentum by the equation $p_{\mathrm{T}} = \frac{L^2 B}{8s}$ where B is the magnetic field strength. Note that the sagitta is larger and can be measured with higher relative accuracy when the distance L is larger, explaining the choice of ATLAS design. Also the relative error on the momentum is proportional to the relative error on the sagitta.

The precise measurement of muon momenta is provided by Monitored Drift Tube (MDT) chambers for both barrel and endcaps, the only exception being in the very forward regions of the endcaps where Cathode Strip Chambers (CSC) take care of the task. Triggering in the barrel region is done by Resistive Plate Chambers (RPC) and in the endcaps by Thin Gap Chambers (TGC). The total number of readout channels in the muon spectrometer is above one million as can be seen from Table 2.1, which details the exact number of chambers and channels for the four technologies, as well as the η coverage.

Figure 2.19 shows the schematic cross-section of a quadrant of the spectrometer in the bending plane. A three letter naming scheme is used for the MDT chambers according to their position. The first letter indicates if the chamber is in the *barrel* (B) or in the *endcap* (E). The second letter refers to the layer of the chamber which can be *inner* (I), *middle* (M) or *outer* (O). The third letter is defined by the size of the station, differentiating between *large* (L) and *small* (S) chambers. For example a BOL is a large chamber in the outer layer of the barrel. Some stations that do not follow this naming scheme are placed in low coverage and transition regions, such as EEL in Figure 2.19.

Technology	Function	Coverage	# Chambers	# Channels
MDT	tracking	$\|\eta\| < 2.7$	1150	354k
CSC	tracking	$2.0 < \|\eta\| < 2.7$	32	30.7k
RPC	trigger	$\|\eta\| < 1.05$	544	373k
TGC	trigger	$1.05 < \|\eta\| < 2.7$	3588	318k

Table 2.1: Detector technologies of the muon spectrometer.

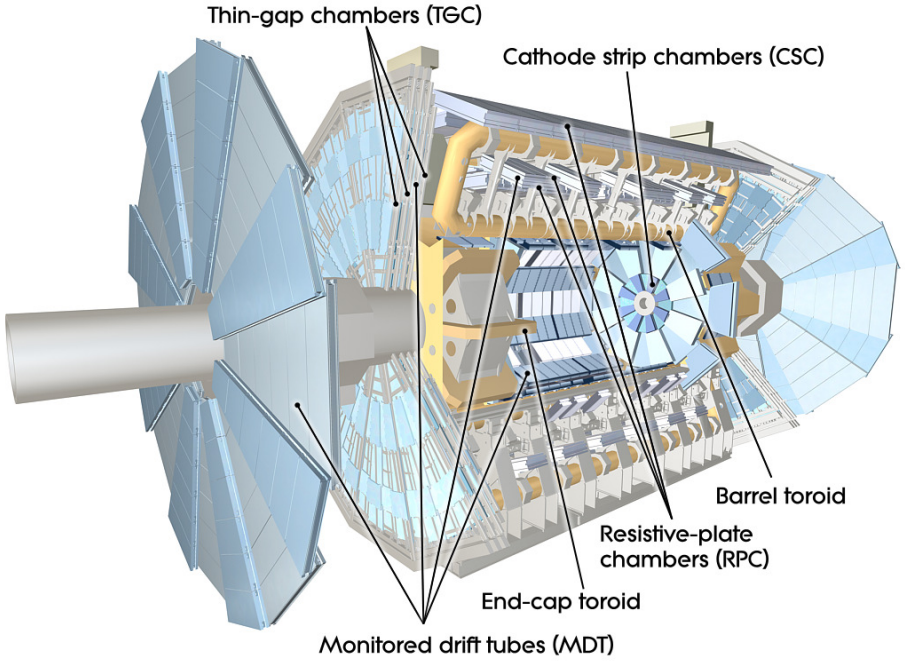


Figure 2.18: Cut-away view of the ATLAS muon system.

Design Considerations The design resolution for measuring high-p_{T} muons is of the order of $1 - 10\%$ for muon transverse momenta of respectively $10 - 1000$ GeV. The air-core toroid system, with a long barrel and two inserted endcap magnets, generates strong bending power in a large volume within a light and open structure. Multiple-scattering effects are thereby minimised and stringent muon momentum resolution is achieved. The muon instrumentation includes, as a key component, trigger chambers with timing resolution of the order of $1.5 - 4$ ns. The maximum signal drift time in the precision MDT chambers is two orders of magnitude larger, hence they cannot be used to trigger upon in the LHC environment.

2.5.1 Toroid magnet

The toroid magnet system consists of three superconducting systems, one for the barrel and two for the endcaps. Each of them consists of eight coils, which are positioned symmetrically around the z-axis. The system has an average field strength of 0.5 T. The bending power ($\int Bdl$) is shown in Figure 2.20(a) and the field strength in Figure 2.20(b) for two instructive ϕ-angles.

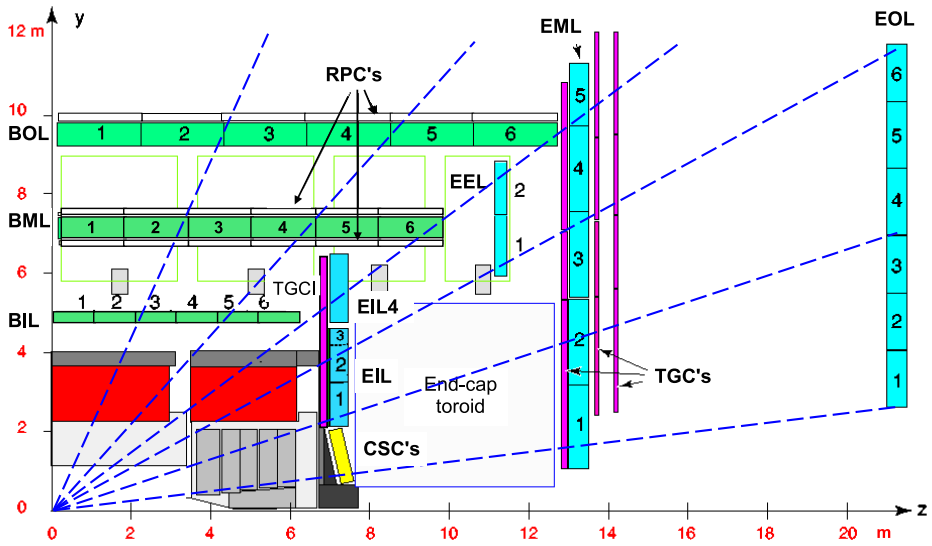


Figure 2.19: Cross-section of the muon system in a plane containing the beam axis (bending plane). Infinite-momentum muons would propagate along straight trajectories which are illustrated by the dashed lines and typically traverse three muon stations.

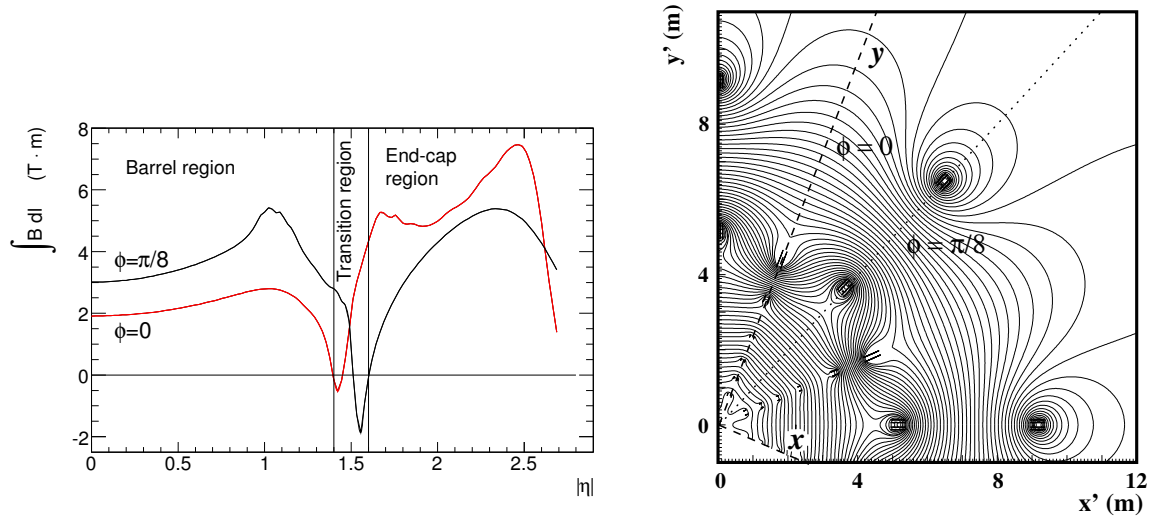
For the barrel region ($|\eta| < 1.4$) the bending power varies between 1.5 and 5.5 Tm, while in the endcap toroid ($1.6 < |\eta| < 2.7$) the variation is between 1 and 7.5 Tm. In the transition region the bending power is smaller. The strength and the uniformity of the magnetic field could be enhanced by an iron core as was chosen by the CMS collaboration, but the resolution of muon momenta measurement would be strongly degraded by multiple scattering.

2.5.2 Monitored drift tube chambers

Purpose MDT chambers are responsible for most of the precision measurements in the Muon Spectrometer. They are positioned in such a way as to maximize the precision of the measurements in the bending plane.

Detection Principle An MDT is an aluminium tube with a diameter of 30 mm filled with a drift gas mixture of $Ar : CO_2 = 93 : 7$ at a pressure of 3 bar. Through the middle of the tube runs a gold plated tungsten anode wire with a diameter of 50 μm. Between the wire and the tube wall a voltage of 3080 V is applied. When a particle traverses the tube the gas will be ionized and the electron clusters will drift towards the wire creating an avalanche as is schematically shown in Figure 2.22(a). The distance between the particle and the anode wire is determined by measuring the arrival time of the first cluster that breaches a predefined threshold as seen in Figure 2.22(b). This drift time is transformed into a drift radius via an *rt-relation* that we show in Figure 2.22(d), that is obtained by combining the times of a large number of crossings shown in Figure 2.22(c) into a drift time spectrum. The *rt-relation* is not linear and is sensitive to external factors such as temperature, magnetic field, gas mixture and high voltage. These factors have to be carefully monitored by the installed magnetic and temperature sensors.

A schematic view of a typical barrel chamber is shown in Figure 2.21. Each station consists of two multilayers which in turn consist of three or four layers of MDTs. Only inner stations have four tube layers per multilayer to improve the local pattern recognition.



(a) (b)

Figure 2.20: Predicted field integral (**a**) as a function of $|\eta|$ from the innermost to the outermost MDT layer in one toroid octant, for infinite-momentum muons. The curves correspond to the azimuthal angles $\phi = 0$ (red) and $\phi = \pi/8$ (black). Calculated magnetic field map (**b**) in the transition region between barrel and endcap. The coordinate system of the magnetic field is rotated by $\frac{\pi}{8}$ with respect to the ATLAS system.

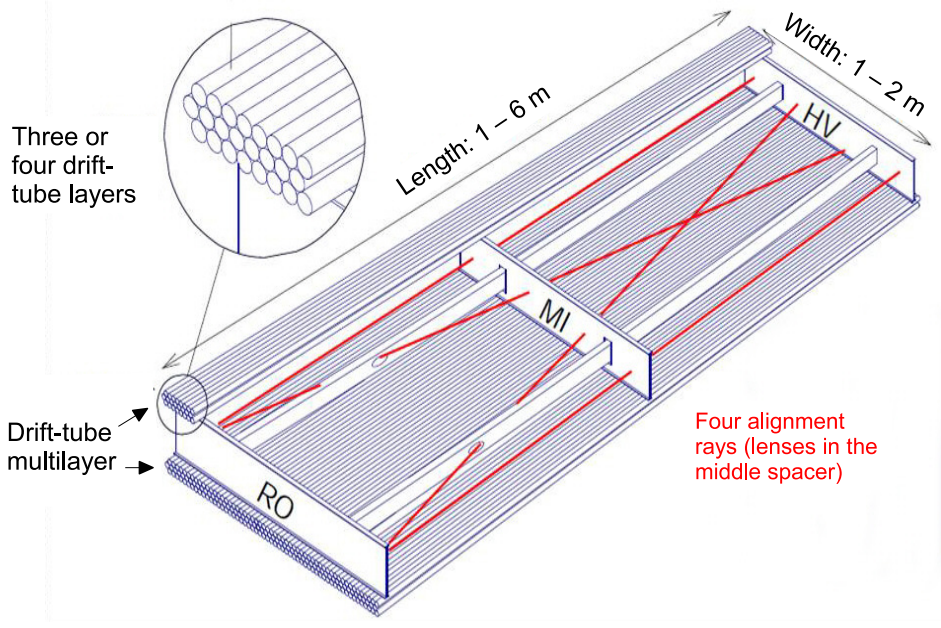


Figure 2.21: Mechanical structure of an MDT chamber. Three spacer bars connected by longitudinal beams form an aluminium space frame, carrying two multi-layers of three or four drift tube layers. Four optical alignment rays, two parallel and two diagonal, allow for monitoring of the internal geometry of the chamber. RO and HV designate the location of the readout electronics and high voltage supplies, respectively.

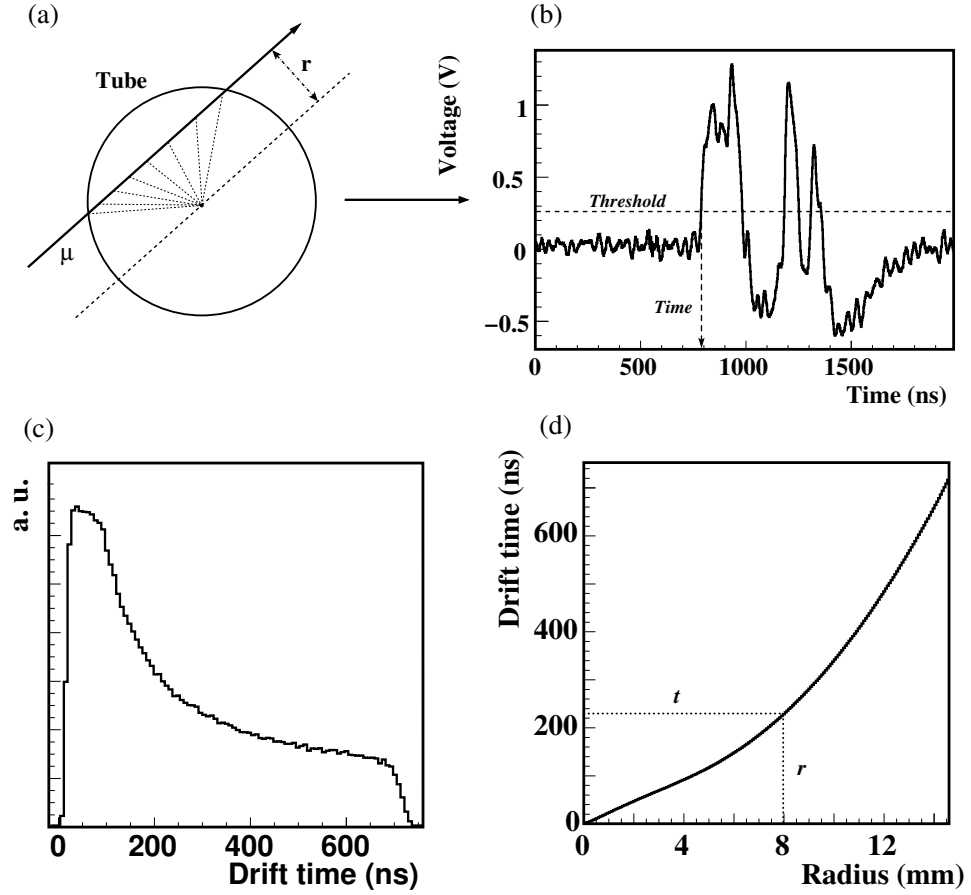


Figure 2.22: Schematic overview of the operational principle of an MDT tube. (a) Schematic overview of the creation of charged clusters by a muon. (b) Measured signal pulse. (c) Typical drift time spectrum. (d) Typical rt-relation. Taken from [85].

Design Considerations The support structure of the muon spectrometer as well as the chambers were made from aluminum, to prevent interaction with the magnetic field. The MDTs are designed to provide a precision measurement with a typical resolution of 80 μm per tube or 35 μm per chamber in the bending plane or η-plane shown in 2.22(a), but deliver no information on the position of the traversing particle along the tube. For the biggest BOL chambers the tubes are almost 5 meters long, so the hit position along the tube can have an effect on the muon momentum measurement. Measuring the coordinate along the tube is mostly done by RPCs discussed in Section 2.5.4, but the MDTs themselves can also be used for this purpose in a Twin Tube setup as will be discussed in Chapter 3.

2.5.3 Cathode strip chambers

Purpose and Design Considerations As can be seen from Figure 2.19 the CSCs are installed in the inner wheel of the endcaps, as the expected particle rate exceeds 150 kHz/cm^2 giving too high occupancy for the safe and correct operation of MDTs. The CSCs operate correctly up to a rate of 1000 kHz/cm^2 while still achieving a resolution of 60 μm in the precision plane and a resolution of 5 mm in the non-bending plane.

Detection Principle The CSCs are multiwire proportional chambers filled with $Ar : CO_2 = 80 : 20$ gas mixture. The anode wires are oriented radially and have cathode strips oriented either perpendicular to them in η or parallel to them in ϕ. Interpolation of the charges induced on the neighboring cathode strips provides a measurement of the position. Each CSC chamber contains four CSC planes resulting in four independent measurements in η and ϕ for each traversing muon. Due to pairing of measurements in both coordinates the CSCs are good in resolving ambiguities if more than one track is present, which is important in this high particle density region.

2.5.4 Resistive plate chambers

Purpose The trigger for muons in the barrel region is provided by the RPCs. Besides the trigger RPCs provide measurements of muon tracks in the non-bending plane.

Detection Principle The RPC is a gaseous detector with 2 mm gas gaps between two parallel resistive plates. The gas mixture is $C_2H_2F_4 : Iso - C_4H_{10} : SF_6 = 94.7 : 5 : 0.3$ with a voltage of 9.8 kV between the plates. When a particle passes through the gas gap, it will create an electron avalanche to the anode plate. Metallic readout strips are mounted onto the plates either in η or ϕ direction with a pitch of respectively 23 and 35 mm in-between.

Design Considerations As shown in Figure 2.19 the middle MDT chambers have two RPCs, one on each side, and the outer MDT chambers have only one RPC. An MDT chamber with corresponding RPC chamber(s) is called a station. Each RPC chamber provides two measurements per traversing particle, one in η and one in ϕ. Hence in total a muon going through the barrel is provided with six RPC measurements, making a momentum dependent trigger possible. The RPC measurement in the non-bending plane (ϕ) provides the missing coordinate for the muon, as we mentioned at the end of Section 2.5.2. In both the bending and non-bending plane the RPCs deliver a typical resolution of 10 mm.

2.5.5 Thin gap chambers

Purpose and Design Considerations In the endcaps TGCs deliver the muon trigger. TGCs are positioned in four planes around the beam axis as shown in Figure 2.19, without being physically attached to a precision chamber like the RPCs.

Detection Principle The TGCs are multiwire proportional chambers operating with $CO_2 : n\text{-}C_5H_{12} = 55 : 45$ gas mixture. The wires oriented in η are operated at 2.9 kV. Both the wires and the pick-up strips positioned perpendicular to the wires are readout, unlike the CSCs, giving both an η and a ϕ measurement. The azimuthal coordinate from TGCs complements the muon measurement by the MDTs in the bending direction. The typical resolution of the TGC chambers is 2-6 mm in the bending plane and 3-7 mm in the non-bending plane.

2.5.6 Muon reconstruction

ATLAS will detect and measure muons in the muon spectrometer, but will also exploit the measurements in the inner detector and the calorimeters to improve the muon identification efficiency and momentum resolution. In ATLAS four kinds of muon candidates are distinguished depending on the way they are reconstructed: *stand-alone muons, combined muons, segment tagged muons* and *calorimeter tagged muons.*

Stand-alone muons are reconstructed using only the hits in the muon spectrometer to find tracks and extrapolate these to the beam line. The standalone algorithm first builds track segments in each of the three muon stations by performing a straight line fit. Then all possible combinations of at least two segments are linked to form tracks, that are refitted from the drift time measurements of the hits belonging to the segments. This global fit returns a χ^2 value that is used for quality selection. The resulting standalone tracks are extrapolated to the beam line, correcting for the energy loss in the material in front of the muon spectrometer.

Combined muons are stand-alone muons that are combined with matching tracks from the inner detector. In reverse, a trajectory in the inner detector is identified as a segment tagged muon if the trajectory extrapolated to the muon spectrometer can be associated with straight track segments in the MDT chambers. Finally a trajectory in the inner detector is identified as a calorimeter muon if the associated energy depositions in the calorimeters are compatible with the hypothesis of a minimum ionizing particle. In the early phase of the LHC operation, ATLAS uses two reconstruction algorithms for each muon category following different pattern recognition strategies, further referred to as Chain 1 and Chain 2 as described in [86]. The complementarity of the different muon types and the corresponding alternative reconstruction algorithms makes it possible to evaluate the muon performance in detail.

2.5.7 Alignment of the Muon Spectrometer

The design transverse momentum resolution at 1 TeV of the MS is about 10%, this translates into a sagitta resolution of 50 μm. The intrinsic resolution of MDT chambers contributes a 40 μm uncertainty to the track sagitta, hence other systematic uncertainties (alignment and calibration) should be kept at the level of 30 μm or smaller. Since longterm mechanical stability in a large structure such as the MS cannot be guaranteed at this level, a continuously running alignment monitoring system has been installed, which relates the position of each chamber to that of its neighbours, both within an MDT layer and along $R - z$ trajectories within MDT towers. This system is based on optical and temperature sensors and detects slow chamber displacements, occurring at a timescale of hours or more. The information from the alignment system is used in the offline track reconstruction to correct for the chamber misalignment.

The RASNIK optical sensor [87] has a simple design principle as schematically shown for an MDT chamber with in-plane alignment sensors in Figure 2.21: a source of light is imaged through a lens onto an electronic image sensor acting as a screen. The system shown in Figure 2.21 can record internal chamber deformations of a few μm, while the MS optical sensor network is able to reliably detect relative changes in chamber position at the 20 μm level. In addition to optical position measurements, it is also necessary to determine the thermal expansion of the chambers. In total, there are about 12000 optical sensors and a similar number of temperature sensors in the system.

The optical alignment system is insufficient to reconstruct, on its own, the absolute positions of the MDT chambers: only variations in relative position can be determined with the required precision. Track-based alignment algorithms must therefore be used in combination with the optical system to achieve the desired sagitta accuracy, and also to determine the global positions of the barrel and end-cap muon-chamber systems with respect to each other and to the inner detector. Several hundred million cosmic ray events collected during 2008 and 2009 were used to commission the Muon Spectrometer and to study the performance of among others the alignment as described in [88].

Data with the toroidal field off were used to measure the alignment precision in the barrel and to validate the alignment corrections in relative mode. The method is to use straight

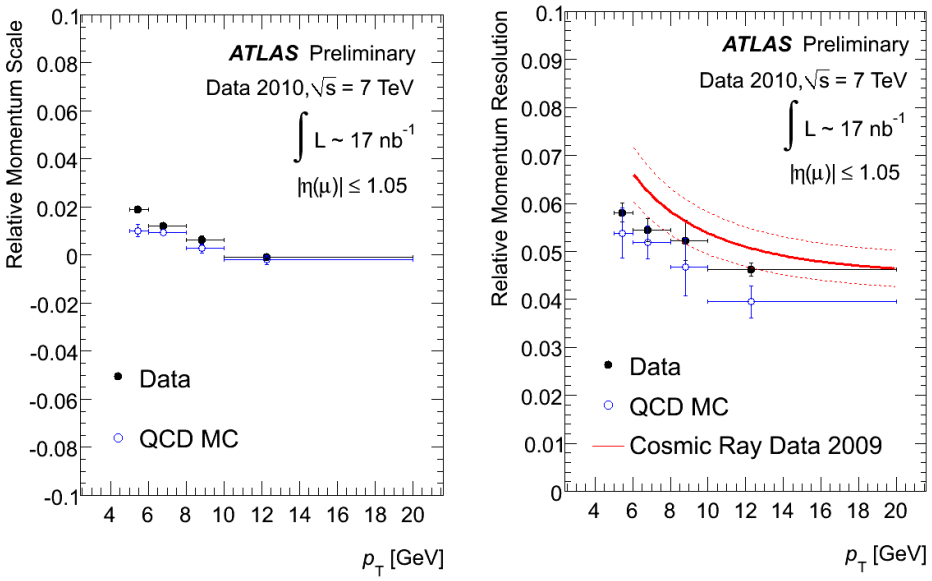


Figure 2.23: The muon momentum scale (left) and resolution (right) for the barrel ($|\eta| < 1.05$) as a function of the combined muon p_{T}, obtained from collision data from 2010 (full dots). The expectations from QCD Monte-Carlo (empty dots) are overlaid. The curve gives the fitted relative resolution of the MS obtained using cosmic ray data, with an uncertainty band corresponding to $\pm 1\sigma$.

muon tracks to determine in absolute mode the initial spectrometer geometry and, once this geometry is determined, to use the optical alignment system to trace all chamber displacements in a relative mode, while the toroid magnet field is ramped up. For a perfect alignment, the reconstructed sagitta of straight tracks should be zero. The alignment procedure with straight tracks is based on the so-called MILLEPEDE fitting method [89] that uses both alignment and track parameters inside a global fit. The preliminary studies with cosmic rays [88] indicate that the method of track-based alignment in combination with the optical system is robust and with sufficient muon data from collisions the design alignment precision will be achieved.

2.5.8 Performance of the muon spectrometer

The performance of the ATLAS muon reconstruction and identification is studied with 17 nb^{-1} of LHC proton-proton collision data at $\sqrt{s} = 7$ TeV collected with muon triggers in [90].

The combined muon efficiency is a product of three efficiencies, namely the efficiency of reconstructing a muon track in the inner detector, the efficiency of reconstructing a muon track in the muon spectrometer, and the efficiency of matching the reconstructed inner detector and muon spectrometer tracks. One can try to obtain efficiency estimates by exploiting the complementarity of the muon reconstruction algorithms. Ideally the efficiency of finding a calorimeter tagged muon is independent of the muon reconstruction efficiency in the muon spectrometer and the efficiency of matching the inner detector with the muon spectrometer track. However due to a higher misidentification probability of the calorimeter taggers, it is necessary to complement the calorimeter tagged muons with requirements on the activity in the muon spectrometer in the region of the tagged muon. In the *segment-enhanced* approach it is required that the calorimeter tagged muon is also segment tagged in the muon spectrometer. The relative efficiency of a combined reconstruction algorithm is defined as the fraction of calorimeter tagged

muons, which are found by the combined reconstruction algorithm. The selection of calorimeter tagged muons is further enhanced by selecting muons satisfying $\frac{p_{\mathrm{ID}}-p_{\mathrm{MS}}-p_{\mathrm{param}}}{p_{\mathrm{ID}}} < 0.3$, which makes sure that muons coming from late decays in flight of pions and kaons are disregarded. The relative efficiency is measured to be $97.4 \pm 0.1\%$ ($95.8 \pm 0.1\%$) for Chain 2 (Chain 1) as compared to respectively $97.7 \pm 0.3\%$ ($97.2 \pm 0.3\%$) from minimium bias Monte Carlo simulation after requiring at least one jet with $p_{\mathrm{T}} > 6$ GeV in the event.

To provide an estimate of the combined muon resolution and scale, the $\frac{p_{\mathrm{ID}}-p_{\mathrm{MS}}-p_{\mathrm{param}}}{p_{\mathrm{ID}}} = \frac{\Delta p_{\mathrm{loss}}}{p_{\mathrm{ID}}}$ variable is used in a template-based likelihood fit, detailed in Chapter 5. The shape of the $\frac{\Delta p_{\mathrm{loss}}}{p_{\mathrm{ID}}}$ distribution is the result of the Gaussian form of the average muon energy loss due to the instrumental resolution, convolved with a Landau distribution accounting for large energy loss fluctuations (with respect to such average) due to traversing of the detector material. Hence a convolution of a Gaussian with a Landau function is used as the parametrized description of the $\frac{\Delta p_{\mathrm{loss}}}{p_{\mathrm{ID}}}$ distribution for the muons. The mean of the Gaussian is extracted as the muon momentum scale, as one expects $\frac{\Delta p_{\mathrm{loss}}}{p_{\mathrm{ID}}}$ to be centered around zero for combined muons. The quoted muon momentum resolution is obtained by adding linearly the Gaussian and Landau widths. In Figure 2.23 the muon momentum scale and resolution are shown as a function of combined muon p_{T} in the barrel region for collision data, minimum bias Monte Carlo simulation and the data collected in 2009 with cosmic ray muons. Generally a good agreement between data and simulation is found.

2.6 Trigger system

The Level-1 (L1) trigger system uses a subset of the total detector information to make a decision on whether or not to continue processing an event. L1 reduces the data rate to approximately the design value of 75 kHz. The subsequent two levels, collectively known as the high-level trigger, are the Level-2 (L2) trigger and the event filter which use all the detector information. With an average event size of 1.5 MB and a final design data-taking rate of approximately 200 Hz, this still corresponds to a data storage rate of 300 MB/s which must be available each second of every hour of every day of LHC running, least we miss the rare new physics events.

2.6.1 Level-1 trigger

The L1 trigger searches for signatures from high-p_T muons, electrons/photons, jets, and τ-leptons decaying into hadrons. It also selects events with large missing transverse energy and large total transverse energy. The L1 trigger uses reduced-granularity information from the muon trigger chambers (RPC and TGC) and the calorimeter sub-systems. The detector readout systems can handle a maximum L1 accept rate of 75 kHz (upgradeable to 100 kHz). The Level-1 trigger is a hardware implemented trigger using custom-built electronics.

An essential function of the L1 trigger is unambiguous identification of the bunch-crossing of interest. The very short (25 ns) bunch-crossing interval makes this a challenging task. The physical size of the muon spectrometer implies muon times-of-flight exceeding the bunch-crossing interval. For the calorimeter trigger, a serious complication is that the width of the calorimeter module signals extends over many (typically four) bunch-crossings. While the trigger decision is being formed, the information for all detector channels has to be retained in custom pipeline memories. The L1 latency, which is the time from the proton-proton collision until the L1 trigger decision, has a target latency of 2.0 μs with a 0.5 μs contingency. About 1 μs of this is already occupied by cable-propagation delays alone.

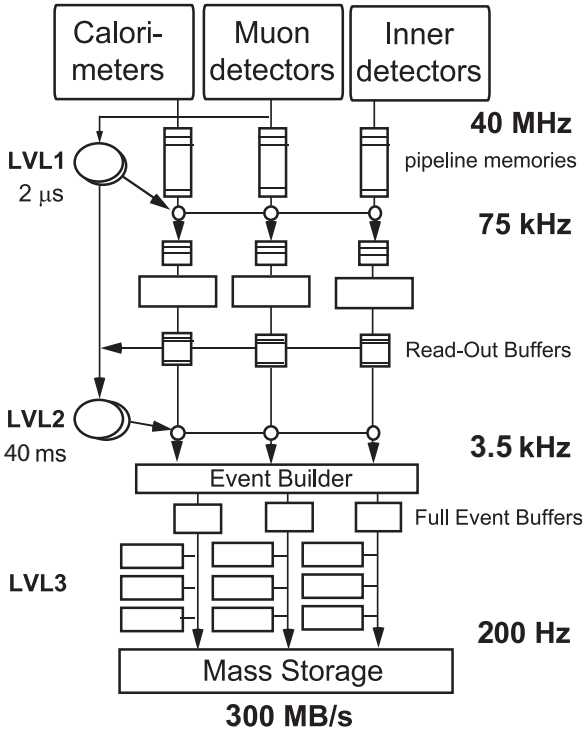


Figure 2.24: Schematic view of the ATLAS trigger system.

2.6.2 Level-2 trigger

The L2 trigger is seeded by *Regions-of-Interest* (RoI's). These are regions of the detector where the L1 trigger has identified possible trigger objects which carry information on coordinates in η-ϕ, energy and type of signatures. The L2 trigger reduces the event rate to below 3.5 kHz, with an average event processing time of approximately 40 ms. It is a software based trigger and it has access to the full granularity information of all sub-detectors for the RoI's.

2.6.3 Event Filter

Past the L2 decision, the third and last trigger level called the Event Filter (EF) is activated. The Event Builder rebuilds the full event using standard ATLAS event reconstruction and analysis software. The required event processing time is four seconds on average, which sets stringent limits on the performance of these applications. A farm with approximately 1500 computers reduces the final output rate to the design value of 200 Hz.

2.6.4 Data streams and trigger menu

The decision for accepting or rejecting an event is based on a *trigger menu*. A flexible trigger menu scheme is used in ATLAS, where the acceptance thresholds for each specific chain can be adjusted at will. With the LHC instantaneous luminosity increasing over time, the trigger menu will be accordingly adapted to keep the output rate reasonable.

At the end of the trigger cycle each event is classified into one or more physics *streams* based on the EF results. This grouping of events simplifies further analysis or a possible reprocessing step. ATLAS uses inclusive streaming, which means that the same event can end up in

several streams and care must be taken when analyzing multiple streams as not to use duplicates. A possible scenario for early running is using four physics streams: electrons and photons (*egamma*), muons, jet/τ/E_T^{miss} and minimum bias. Although minimum bias events are classified as background, their expected rate at the LHC has a large uncertainty, hence this special stream to study these events. Next to the physics streams there are also the calibration, express and debug streams. These are used for problem solving, calibration and almost real-time monitoring of the data quality as the names suggest.

3

Twin Tubes in the ATLAS Muon Spectrometer

3.1 Introduction

The ATLAS muon spectrometer measures the deflection of the muon trajectories in the magnetic field with high precision in the bending direction by Monitored Drift Tube (MDT) chambers. As explained in Section 2.5.2, a single MDT consists of an aluminum tube filled with drift gas which holds a wire at its centre. The tube serves as the cathode, while the central wire is the anode of the drift tube. High-Voltage (HV) is applied to the wires on one end of the tubes and the signals are collected at the other ends, known as the Read-Out (RO) side. A MDT chamber consists of two multilayers (with each 3 or 4 layers of tubes) separated by a spacer frame.

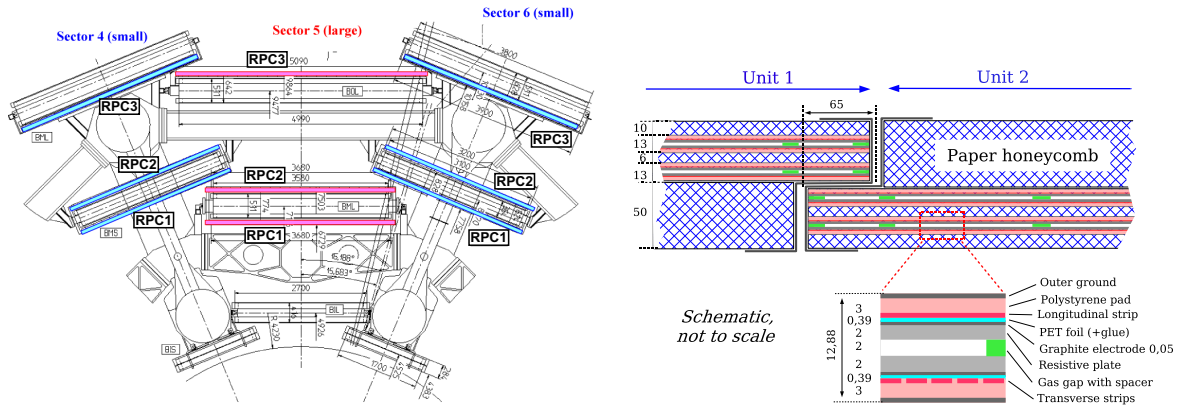
A muon traversing the MDT at a certain distance of the wire ionizes the gas. The electrons drift to the anode wire and after gas amplification a signal is generated that propagates to both ends of the wire. At the RO side, the on-chamber electronics shape the signal and when it passes the (adjustable) threshold, measure the arrival time by a Time-to-Digital Converter (TDC), as well as charge information using a Wilkinson Analog-to-Digital Converter (ADC). The distance of the muon trajectory with respect to the wire is accurately determined from the measured time using a space-time relation, discussed in Section 2.5.2.

The time (t_{TDC}) is measured in bins of 0.78 ns with respect to the time of the proton-proton collision, and has several contributions: the drift time (t_{drift}), the propagation delay (t_{prop}) of the signal along the anode wire, the muon time-of-flight (t_{ToF}) from the interaction point to the impact point, and t_0 that depends on many fixed delays like cable lengths, front-end electronics response and Level-1 trigger latency. The t_0 offset has to be determined for each drift tube. So as to not degrade the precision coordinate resolution of the MDTs, the precision of the t_0 offset expected in LHC collision data is better than 1 ns with a dataset of about 10K muons crossing the drift tube. However due to some additional time jitter present in cosmic ray data discussed later, the calibrated resolution in our dataset is 2-4 ns [88].

The propagation delay is proportional to the distance x from the impact point to the Read-Out end of the tube: $t_{\mathrm{prop}} = x/v$, where v is the effective signal propagation speed (see Figure 3.2(a)). The drift time is then given by:

$$t_{\mathrm{drift}} = t_{\mathrm{TDC}} - t_0 - t_{\mathrm{ToF}} - \frac{x}{v}\,. \tag{3.1}$$

The correction for the delay due to the signal propagation along the wire turns out to be one of the larger corrections, notably for the 5 m long MDTs in the Barrel Outer Large (BOL) chambers. This correction can only be made once the location along the wire at which the muon



(a) (b)

Figure 3.1: Cross-section through the upper part of the barrel (a) with the RPCs marked as filled bands. In the middle chamber layer, RPC1 and RPC2 are below and above their respective MDT partner. In the outer layer, the RPC3 is above the MDT in the large and below the MDT in the small sectors. Cross-section through an RPC (b), where two units are joined to form a chamber. Each unit has two gas volumes supported by spacers (the distance between successive spacers is 100mm), four resistive electrodes and four readout planes, reading the transverse and longitudinal direction. The sandwich structure (hashed) is made of paper honeycomb. The ϕ-strips are in the plane of the figure and the η-strips are perpendicular to it. All dimensions are in mm.

passes through the MDT, is known. In the baseline design of the ATLAS Muon Spectrometer this so-called *second coordinate* is extracted from the Resistive Plate Chamber (RPC) data.

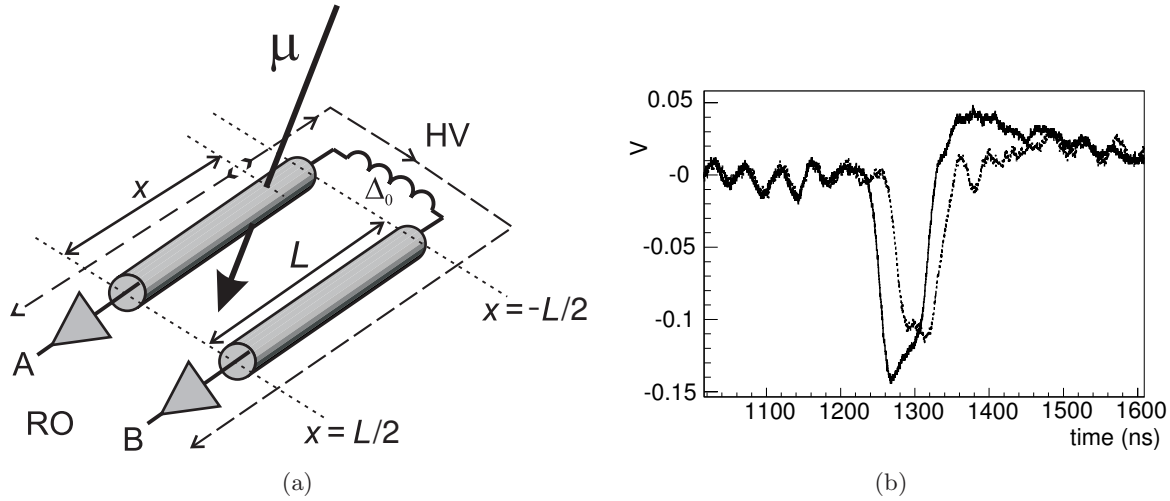
To achieve this, the RPC chambers also include strip planes segmented in the non-bending plane with 30 mm wide strips. Figure 3.1(b) shows a schematic view of an RPC chamber, that is built up of two units, each unit in turn containing two gas volumes. Each gas volume has strips in both η and ϕ-planes. The η-plane strips are necessary for a momentum dependent trigger measurement (which is measured with greater precision by the MDTs, which are however too slow to deliver a workable trigger in the LHC environment), and strips in the ϕ-plane deliver the aforementioned second coordinate or the coordinate along the MDT tubes. A RPC produces signals of 5 ns full width at half-maximum with a time jitter of 1.5 ns [66].

Figure 3.1(a) shows a standard barrel sector and the location of the RPCs (filled bands) relative to the MDTs. The outer MDT chambers have only one RPC each attached, while the middle MDT chambers are enclosed between two RPC units. Just as with the precision chambers, three measurements along the track are enough to calculate the track sagitta that is related to the momentum of the traversing particle.

The second coordinate can also be measured by pairing two MDTs at the HV side forming a so-called **twin tube**. This modification endows the MDTs with full 3D track reconstruction using specially designed electronics boards. The principle, calibration and performance of twin tubes in ATLAS is the subject of the rest of this chapter.

3.2 Twin tube principle, motivation and hardware

Using a twin-tube, i.e. a pair of MDTs, the second coordinate can be determined from the two registered times. This principle is shown schematically in Figure 3.2(a). A pair of tubes



(a) (b)

Figure 3.2: Schematic view of a twin-tube (a), consisting of two MDTs. A muon and the coordinate system are indicated. The dashed arrows indicate the propagation of the signals and the chamber is read out at both A and B. Example of measured twin-tube pre-amplifier output signals (b). The continuous line represents the prime signal and the dashed line the delayed twin-partner signal.

is interconnected at the HV end via an impedance-matched delay line. The prime muon signal generated in tube A, propagates to the Read-Out end of tubes A and B. In Figure 3.2(b) the measured raw muon signals on a twin-tube pair are shown, taken from [91]. The muon traverses tube A and it records a drift time $t^{\mathrm{A}}_{\mathrm{TDC}}$ as usual. Its twin-partner (tube B) records a drift time $t^{\mathrm{B}}_{\mathrm{TDC}}$ with an extra delay of $\Delta_0 \simeq 6$ ns (in our implementation). The built-in delay is needed to distinguish prime and twin-partner signals for muons passing near the HV end of a MDT. The measured times in the two MDTs are related to the drift time t_{drift}, the t_0's and the time-of-flight delay via:

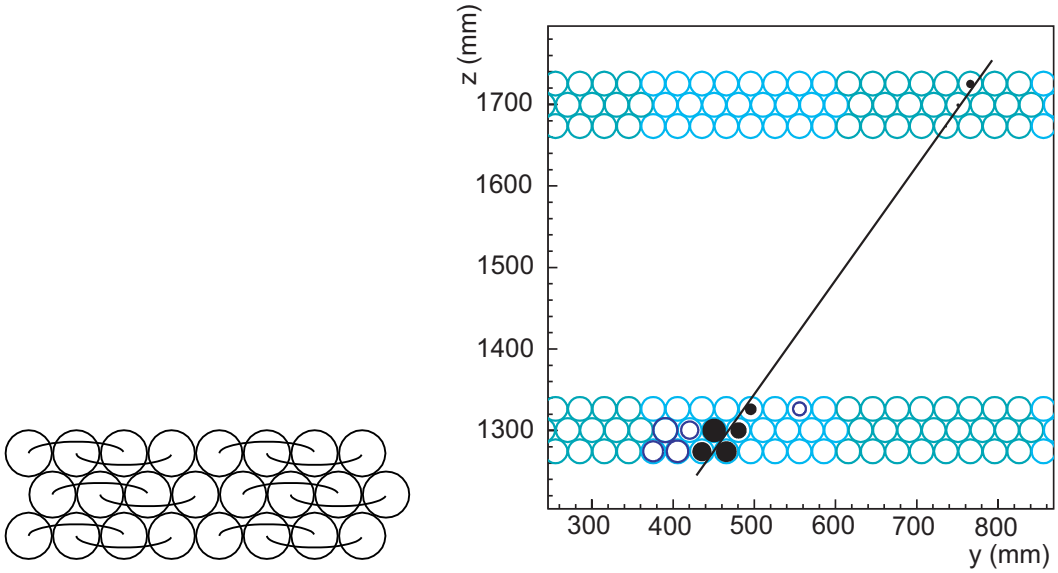
$$t^{\mathrm{A}}_{\mathrm{TDC}} = t_{\mathrm{drift}} + t^{\mathrm{A}}_0 + \frac{\frac{L}{2} - x}{v_A} + t_{\mathrm{ToF}} \,, \tag{3.2}$$

$$t^{\mathrm{B}}_{\mathrm{TDC}} = t_{\mathrm{drift}} + t^{\mathrm{B}}_0 + \frac{\frac{L}{2} + x}{v_A} + \frac{L}{v_B} + t_{\mathrm{ToF}} + \Delta_0 + \Delta t^B \,, \tag{3.3}$$

where L is the chamber length and x the coordinate along the wire, which we also refer to as *twin coordinate*. As shown in Figure 3.2(a), $x = 0$ is in the middle of the chamber and x is equal to $L/2$ ($-L/2$) at the RO (HV) sides. The variable Δt^B is a time slewing correction, caused by damping of the signal as it travels through tube B. In case the amplitude of the signal is infinite and there is no damping, the velocities v_A and v_B are equal to the speed of light and Δt^B is zero.

A model for the twin coordinate was developed assuming signal propagation with the speed of light. The time slewing caused by the exponential damping of the signal and the different gains of the amplifiers were added as time corrections, assumed to be zero before calibration. The time shift induced by reflections at the HV side was measured to be about 0.7 ns [91]. The electronics gain of the amplifier was found to vary by 12 %, consistent with the electronics specifications [91].

If this model is linearized it can be written as Equations 3.2 and 3.3. The speed of light has to be replaced by two effective propagation speeds v_A and v_B and one needs to add a constant



(a) (b)

Figure 3.3: Twin-tube connections on one High-Voltage board (a) for a chamber with three layers per multilayer. Typical twin-tube event (b) in a chamber with the bottom multilayer equipped with twin-tube High-Voltage boards showing the original (prime) hits by solid circles and the twin-partner hits by open circles.

Δt^B to the time delay. The data will be calibrated using these linear expressions. Defining the corrected time difference Δt as:

$$\Delta t \;=\; (t^{\mathrm{B}}_{\mathrm{TDC}} - t^{\mathrm{B}}_0) - (t^{\mathrm{A}}_{\mathrm{TDC}} - t^{\mathrm{A}}_0)\;, \tag{3.4}$$

and solving for the second (twin) coordinate x yields:

$$x \;=\; \frac{v_A}{2}\left(\Delta t - \Delta_0 - \Delta t^B - \frac{L}{v_B}\right)\;. \tag{3.5}$$

The x-coordinate is linear in the corrected time difference. For deciding which tube was traversed by the muon it is sufficient to require that Δt is positive.

3.2.1 Twin-tube implementation in ATLAS

In the implementation for an MDT chamber consisting of two times three layers, the MDTs are paired as shown in Figure 3.3(a). This layout excludes that both twin-partners are directly hit by the same traversing muon coming from the interaction point. Additionally, it is also compatible with the segmentation (three layers of eight tubes) of the baseline on-chamber HV distribution boards, also known as mezzanines. The delay lines are integrated in the six layer printed circuit board and the delay is 6 ns.

At the Cetraro ATLAS Muon Workshop in 2005, it was decided to equip the inner multilayer of two BOL chambers in ATLAS with twin-tube HV boards, as a test case. Both chambers are in the lower most sector of the muon spectrometer, known as sector 13 in the ATLAS ϕ-plane. Both chambers are placed as fourth in η-plane (counting from $\eta = 0$), one at positive η (known as side A) and another at negative η (known as side C). In the rest of this chapter, we will

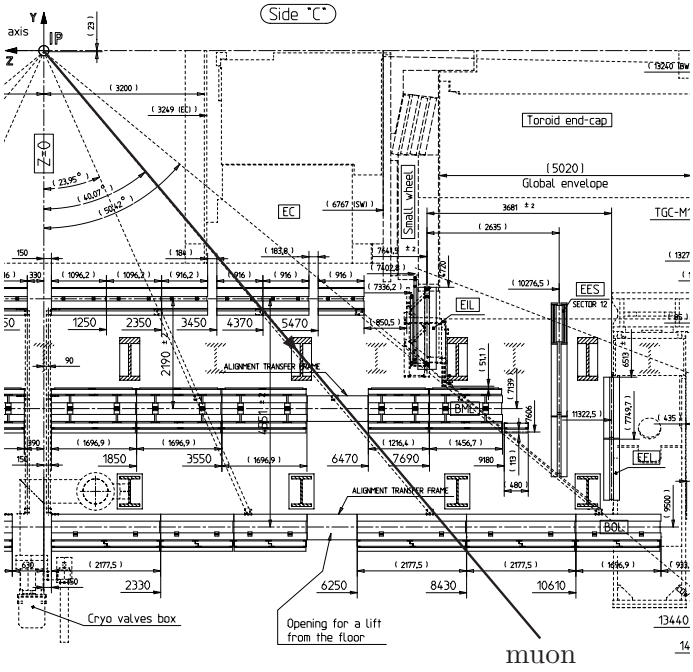


Figure 3.4: Part of a technical drawing of sector 13, where twin tube HV boards are installed on **BOL4A13** and **BOL4C13**. The thick line shows a muon coming from the interaction point (IP) and hitting a twin tube chamber. The limited trigger chamber coverage in this region is seen, as the middle chambers with two RPCs could not be placed, because of the access shaft for the elevator.

hence refer to these two chambers as **BOL4A13** and **BOL4C13**.

The motivation for installing these two chambers with twin tubes is the limited trigger chamber coverage in this region, as shown in Figure 3.4. A muon coming from the interaction point towards the twin chamber, does not hit an RPC on its path until it reaches the twin chamber, as shown by the thick line in the figure. The middle chamber could not be installed because a shaft is needed for the elevator. If the RPC installed on the twin chamber would stop functioning for some reason, then ATLAS would not have a second coordinate measurement for muons passing this chamber.

On top of that the twin tube setup can assist in resolving ambiguities when an RPC chamber is hit by multiple muons in the same event. Since for each hit the RPC delivers an η-strip and a ϕ-strip signal, for multiple muons pattern recognition might associate the η-strip of one muon to the ϕ-strip of the other muon. In that case the *ghost*-hits, given by the wrong association, are indistinguishable from the correct strips crossing. Twin tubes can solve this ambiguity by delivering an independent 3D measurement in both η and ϕ-planes.

Figure 3.3(b) shows a typical event in a BOL chamber, where one of the multilayers is equipped with twin-tube HV boards. The original hits, which we refer to as the *prime* hits, are shown by solid circles with respective drift-radii, while the corresponding twin hits are shown by the open circles. The y, z-axes displayed in Figure 3.3(b) are the local chamber coordinates. The distance between two multilayers is measured in z, the drift-radius is given by y and z, while the hit position along the tube is on the x-axis.

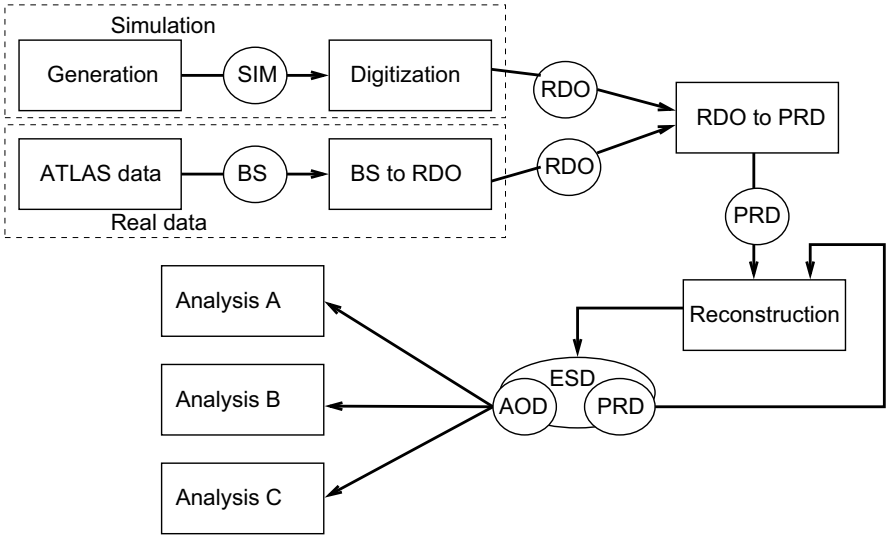


Figure 3.5: The modularity of an ATHENA job is ensured by using data classes. Any input data can be handled by the reconstruction as long as its format is PRD (Prepared Raw Data). The output of the reconstruction is of the format ESD (Event Summary Data), from which an AOD (Analysis Object Data) is extracted.

3.3 Software implementation

The complexity of the ATLAS experiment calls for a flexible and modular software framework. A framework called ATHENA [92] has been developed, which is able to handle different tasks such as event generation, simulation, reconstruction and analysis. The software is extendable and flexible such that it can adapt to the need of the users and allow for software development. The software inside the framework is robust and maintainable by a large community.

The ATHENA framework is realized as a component model-based framework. A distinction is made between algorithmic classes performing dedicated tasks and data classes for communication between the different algorithmic modules. By defining common interfaces of the algorithms and using a well defined Event Data Model (EDM) for the data classes, ATHENA ensures commonality between lower level sub-detector specific algorithms and for higher level combinations between various systems used by the physics groups.

By defining the format of input/output data objects to be processed by the algorithm, interchangeability of modules is ensured inside ATHENA. Figure 3.5 shows a flow diagram of a full job in ATLAS (taken from [93]), starting from either simulation or real ATLAS data to physics analysis.

Data from ATLAS is realized as a *byte-stream* (BS) which is translated to *Raw Data Objects* (RDOs), C++ classes representing raw hits. For example an MDT RDO contains, on top of all the raw datawords connected to the channel, the channel id and the recorded TDC and ADC values. The RDO illustrates the concept of modular design and the use of an EDM explicitly, as simulated data are translated to RDO objects as well. The module responsible for translating the RDO to *Prepared Raw Data* (PRD) objects takes any RDO as input. It does not need to know the source of the RDO enabling the same reconstruction flow for both real and simulated data in ATHENA. In the step from RDO to PRD software related to MDT hits had to be adjusted, to be able to process the twin tube data correctly, as will be discussed in the next section. As well the digitization step software of the simulation chain was extended to simulate

twin tube hits for the two chambers in ATLAS equipped with twin tube HV boards.

As required by EDM, the reconstruction step takes any PRD as input. It produces *Event Summary Data (ESD)* objects and *Analysis Object Data (AOD)* objects [92]. These objects can be analyzed by several user analysis modules. The ESD is the output of the reconstruction job and contains the full event information. Its content is suitable for re-reconstruction and calibration as well. Though for analysis purposes the ESD is too large and a slimmed object, the AOD, is extracted from the ESD. The AOD contains among others containers of physics objects such as four-momenta of the particles (muons, electrons, jets) in the event.

3.3.1 Software for twin tubes

One of the last steps in the simulation process is to take the detector response into account, which happens for MDT hits in *MdtDigitizationTool*. Here the software was extended for the two multilayers with twin tubes. For a simulated MDT hit in these multilayers, a twin hit is produced using information of the prime hit and equations from Section 3.2 to calculate the arrival time of the twin signal. The calculated twin hit time is then smeared with a Gaussian resolution, as measured for the x-coordinate to be 17 cm in a test setup described in [91]. For each twin hit a random number is drawn from the Gaussian resolution around the calculated twin hit time. The propagation speed here is assumed to be the speed of light and no delay due to slewing is taken into account, hence presenting the simulated twin tubes as ideal. Although addition of a twin hit is by default on in simulation software, it can be simply switched off by setting the *useTwin* flag in the job options to false.

The muon PRD are the transient representation of muon RDO, as they are not written out to ESD. For each technology in the muon spectrometer a dedicated PRD class exists, that inherits from the EDM base class *PrepRawData*. The standard *MdtPrepRawData* objects are constructed from the muon RDOs by performing a crude calibration of the drift circle, and saving the hit TDC and ADC values for a possible later recalibration. The calibration of the drift circle happens inside the *MdtCalibrationSvc* class.

The *MdtRdoToPrepDataTool* is extended with functionality to deal with twin tube data. If an MDT RDO hit is found from one of the two multilayers equipped with twin tube HV boards, the collection of MDT RDOs is scanned to find the corresponding twin MDT RDO. If none is found the MDT RDO is treated as a regular MDT RDO and is made into a PRD as discussed above. On the other hand if the twin hit has also been recorded, then a new function is called upon in *MdtCalibrationSvc* to calculate, besides the crude drift radius of the prime hit, the twin coordinate of the RDO twin pair. If the reconstructed twin coordinate is unphysical, in the sense that it lies outside the physical dimensions of the tube taking into account five standard deviations of x-resolution, it is put at the tube end. From the sign of time difference between the two hits, a decision is made which of the two is the prime hit and which is the twin hit.

When all the information has been collected, a new *MdtTwinPrepRawData* object is created. This object saves the same information for the prime hit as is done for non twin tube MDT hits, except that instead of the one dimensional drift radius, it saves a two dimensional position given by the drift radius as the first coordinate and the twin coordinate as the second one with respective errors. In addition it also saves the twin hit TDC and ADC values, to be able to recalibrate the twin coordinate later as well. The *MdtTwinPrepRawData* class inherits from the *MdtPrepRawData* class, hence giving the pattern recognition and other upstream algorithms the choice of using or discarding the saved twin coordinate and the twin hit TDC/ADC values.

In the time window allowed by the muon software, multiple hits can be recorded on the same tube within one RDO collection, possibly coming from other muons in cosmic showers

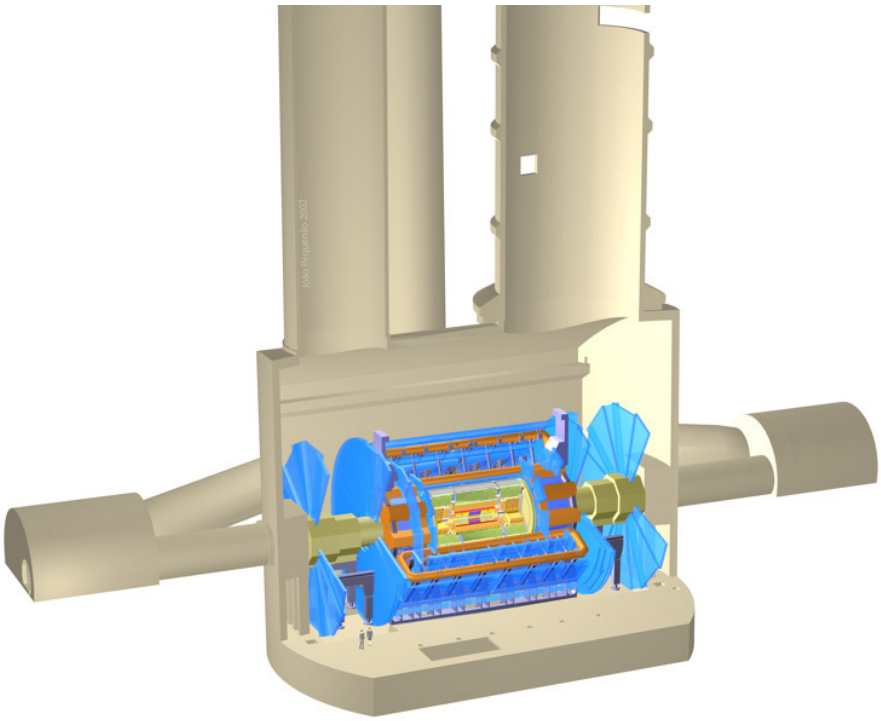


Figure 3.6: Schematic view of the ATLAS detector in its cavern and the main service shafts as used in the cosmic simulation.

or δ-electrons being knocked out by the original muon. These secondary hits are collected and stored in the same way as described above. For a very small percentage ($< 0.01\%$) even tertiary hits were found within one RDO collection. A study performed on these tertiary hits showed them to come from noisy tubes, as they were out-of-time and had very low ADC counts. Hence it was chosen to throw these tertiary hits away in the twin tube analysis.

Just as for the digitization, also in *MdtRdoToPrepDataTool* the twin tube extension can be simply switched off by usage of the *useTwin* flag in the job options. However if real data coming from ATLAS is being handled while this flag is set to false, the twin hits will be stored as normal MDT PRDs, which might mislead the pattern recognition and track reconstruction. For the purpose of disregarding the twin hits altogether, an extra flag *use1DPrepDataTwin* can be set, that will make sure that the prime hits are saved as standard *MdtPrepRawData* objects, while the twin hits are not saved at all.

3.4 Calibration and performance

In preparation for LHC collisions, the ATLAS detector has acquired a large number of 'cosmic ray muon' events, which were used to align and calibrate the detector. Cosmic ray muons result from the interaction of energetic cosmic rays (protons and nuclei) in the upper atmosphere and subsequent particle showers and decays. In the rest of this chapter a subset of data corresponding to about half a million cosmic ray muon events is presented, that are used to calibrate and study the performance of twin tubes.

Most of the cosmic ray muons reach the ATLAS detector underground via the two big shafts, shown schematically in Figure 3.6. The cosmic ray muons have incident angles close to

the vertical axis and are triggered by the RPCs. For the autumn 2009 run used in this analysis the toroid magnet was on.

Besides being recorded as data, cosmic ray muons were also simulated in ATLAS using Monte Carlo (MC) techniques, muons being the dominant particle source. From previous measurements we know the energy spectrum and the total rate of cosmic ray muons [94]. In the simulation, muons are generated at the surface of the Earth on a 600 by 600 m^2 area, centered above the nominal ATLAS interaction point. The GEANT4 program [33] is used to implement the ATLAS detector geometry as well as the shape of the cavern, the main service shafts and the rock.

Muon reconstruction was handled by the *Moore* algorithm [95], also known as Chain2 in [86]. The general strategy is that trajectories in individual chambers can be approximated by straight lines over a short distance, where bending in the magnetic field has little effect, and are fit to track *segments*. Combining the segments from multiple chambers in a global fit, tracks are formed. The algorithm had to be adapted for the cosmic ray muon conditions, because cosmic ray muons do not point to the interaction point as collision muons do, and they are asynchronous with the detector clock. In addition during commissioning the trigger detectors were not timed with sufficient precision and originally alignment of the detectors was not very precise [88]. Hence the standard tracking requirements were relaxed, such as hit to track association, and the procedure to accommodate the timing conditions.

For the rest of the chapter, whenever we speak of RPC and twin tube hits, we refer to *hits-on-track* [85]. These hits-on-track are used first in the reconstruction of the segments and in the global fit they must be associated to the track. The hits that are not on track might be caused by noisy tubes or strips and will thus degrade the resolution and efficiency.

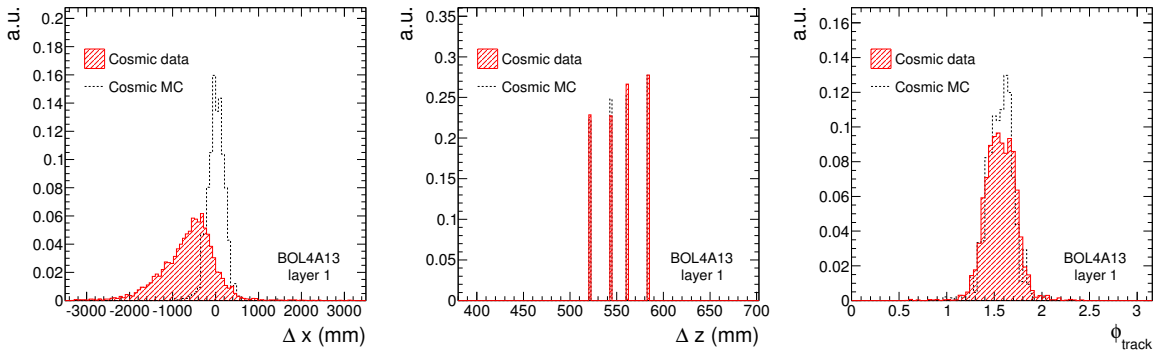
Although each chamber has one multilayer equipped with twin tubes, hence six MDT twin layers in total in ATLAS, in the rest of the chapter we use the first layer of BOL4A13 to showcase and discuss performance. It is representable for all the other twin tube layers, unless we specifically mention otherwise.

All the following performance studies are performed after selecting segments with at least one RPC and one twin tube MDT hit-on-track. On top the segments must satisfy the requirement that the tubes on it have had their t_0 fitted at least once, which means that they were at least coarsely calibrated in time. Furthermore we apply some cleaning cuts on the recorded data, such as disregarding events where a cosmic shower of multiple particles may disturb our measurements by wrong combinations of RPC and MDT hits. This is achieved by requiring that we have only events with a maximum of 3 tracks going through the muon spectrometer. Low momentum tracks suffer a lot more from multiple scattering in the detector material, so these are also cut out by asking for $p_{track} > 3$ GeV.

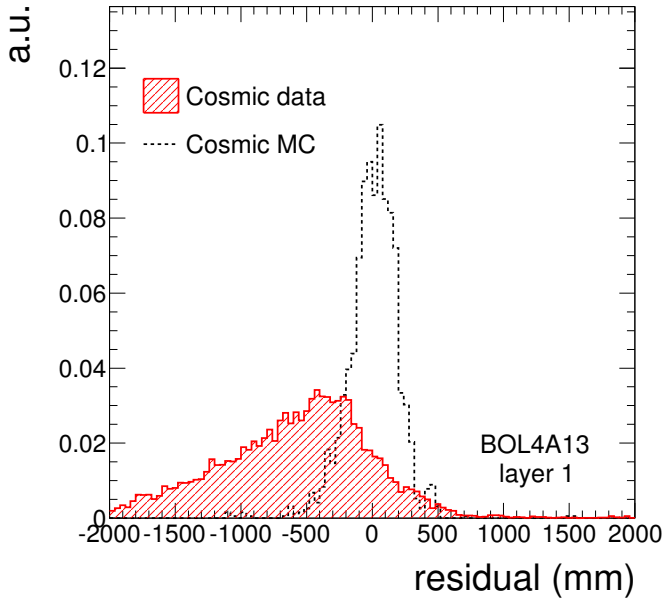
3.4.1 Twin coordinate residual

One of the main quality assessments for the twin tube setup is the precision with which we can measure the twin coordinate, i.e. the twin coordinate resolution. As we know the RPCs to have a spatial resolution of 3 cm for the second coordinate, while the best twin tube resolution was measured to be 17 cm in [91]. We can use the RPCs as our independent baseline measurement to calibrate and assess the twin tube performance.

Hence we start with two independent two-dimensional measurements set in the chamber local coordinate system: one from RPC $(x_{\mathrm{RPC}}, z_{\mathrm{RPC}})$ and the other from MDT with twin setup $(x_{\mathrm{twin}}, z_{\mathrm{twin}})$. Because the RPC and the twin tube measurements are done at different z-planes,



(a) (b) (c)



(d)

Figure 3.7: Uncalibrated residual subparts Δx (a), Δz (b) and ϕ_{track} (c) for data and simulated samples, as well as the calculated residual (d). The figures show the normalized distributions for layer-1 of chamber BOL4A13, but are representative for all twin tube layers of both chambers.

we extrapolate the RPC hits to the twin hit z-plane, using the difference ($\Delta z = z_{\text{RPC}} - z_{\text{twin}}$) as:

$$x_{\text{RPC}\uparrow\text{twin}} = x_{\text{RPC}} - \Delta z \cdot \frac{\cos\phi_{\text{track}}}{\sin\phi_{\text{track}}} \ , \tag{3.6}$$

and in the same manner x_{twin} can be extrapolated to the RPC hit plane as:

$$x_{\text{twin}\downarrow\text{RPC}} = x_{\text{twin}} + \Delta z \cdot \frac{\cos\phi_{\text{track}}}{\sin\phi_{\text{track}}} \ , \tag{3.7}$$

where ϕ_{track} is the global angle ϕ of the track transformed to the local coordinates of the chamber. The local chamber coordinates of each muon spectrometer sector are dependent on the specific orientation of the sector, hence one must take care when comparing these. The correct mathematical expressions for the transformation are given in [93].

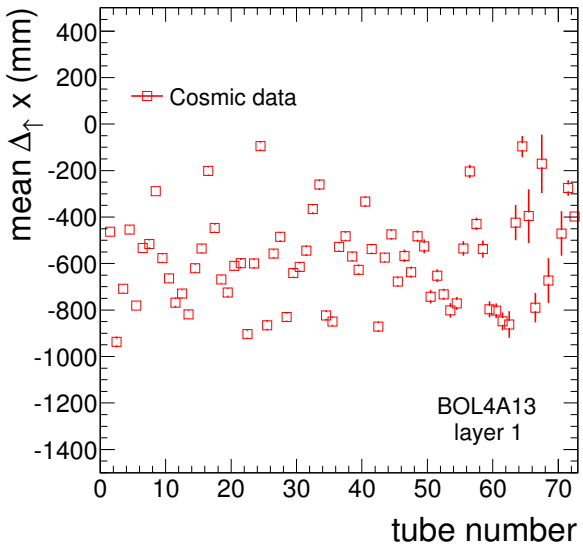


Figure 3.8: The mean of $\Delta_{\uparrow}x$ distribution as a function of tube number for layer-1 of chamber BOL4A13.

We define the *residual* of a twin tube hit as compared to an RPC hit by using the formula used for 'point of closest approach' measurement:

$$\text{residual} = \Delta x \cdot \sin\phi_{\text{track}} - \Delta z \cdot \cos\phi_{\text{track}}, \tag{3.8}$$

where Δx is the difference between the RPC and twin hits, $x_{\text{RPC}} - x_{\text{twin}}$. This formula has the advantage over a purely one-dimensional residual calculation, as it is independent of the specific orientation of the local coordinate system with respect to the global coordinate system.

Figure 3.7(d) shows the twin tube residual calculated for simulated and recorded cosmic ray muons. The simulated residual is equal to what we put in, as described in Section 3.3.1, while the data residual is clearly much wider and its mean is not centered around zero.

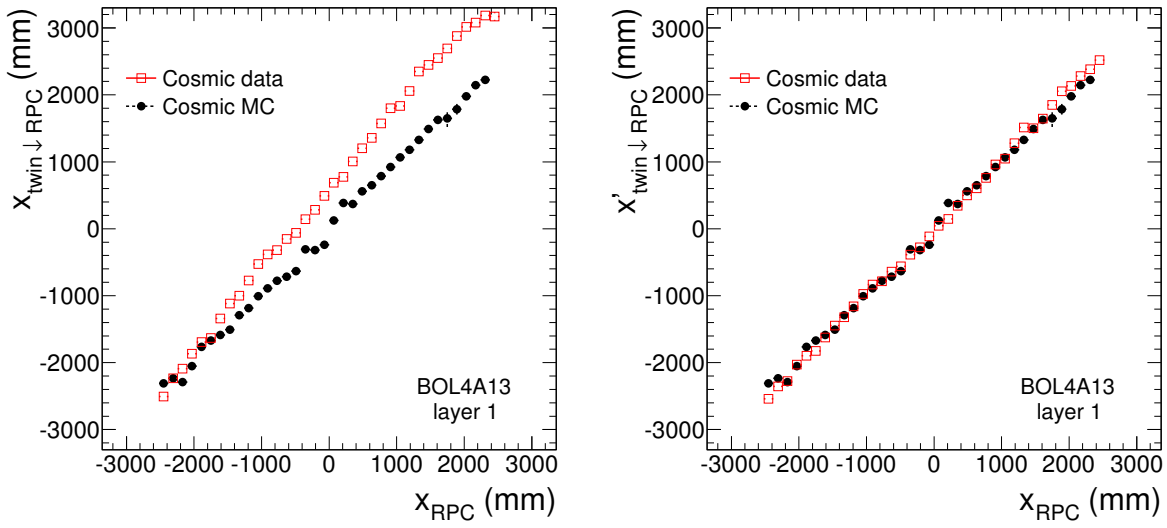
When we compare the three variables that make up the residual, in Figures 3.7(a), 3.7(b) and 3.7(c), it is clear that the Δx distribution is responsible for the residual difference between data and simulation. Both Δz and ϕ_{track} compare well between data and MC samples. The Δz distribution shows the four RPC gas volumes, discussed in Section 3.1, each with their own η and ϕ measurements. Because twin tubes are installed in the bottom most sector of ATLAS, in local coordinates ϕ_{track} is equal to $\pi/2$ for vertically descending muons, as expected from cosmic ray muons, as is shown for data and simulation in Figure 3.7(c).

3.4.2 Calibration of twin tubes

As discussed earlier the t_0 resolution after global calibration of the cosmic ray muon dataset has a precision of 2-4 ns. However for the twin tube setup this is unacceptable, as 4 ns translates for us into 1200 mm, when we take propagation speed to be the speed of light ($c = 300$ mm/ns). We must calibrate the t_0's of our tubes with a greater precision, for which we use the Δx distribution.

The difference between x_{twin} and $x_{\text{RPC}\uparrow\text{twin}}$ we refer to as $\Delta_{\uparrow}x$ to clarify that we have extrapolated the RPC measurement. Now we can study the $\Delta_{\uparrow}x$ distribution for every tube, as is shown for layer-1 of BOL4A13 by the profile distribution in Figure 3.8.

For all the tubes the mean of $\Delta_{\uparrow}x$ is negative, which shows how much t_0 is off. But since we are interested in a correct x_{twin}, we calibrate x_{twin} by adding the mean of $\Delta_{\uparrow}x$ for every tube,



(a) (b)

Figure 3.9: The profile distributions of x_{RPC} versus $x_{\mathrm{twin}\downarrow\mathrm{RPC}}$ before (a) and after (b) the calibration described in the text, for both cosmic data and simulated samples. For clarity the calibrated twin coordinate is denoted with an apostrophe $x'_{\mathrm{twin}\downarrow\mathrm{RPC}}$.

which we call δx, hence calibrating $\Delta_{\uparrow}x$ to be centered around zero for each tube. What is also interesting to see from Figure 3.8 is the structure as a function of tube number. As the tubes are read out by mezzanines in groups of 8, we see a repetitive structure every 8 tubes. This points to the fact that possibly cable lengths and electronics response of the mezzanines is dependent on the tube number in the corresponding group of 8.

Besides calibrating the offset of the $\Delta_{\uparrow}x$ for every tube, it is important to study how the extrapolated twin coordinate is correctly correlated to the coordinate measured by the RPC. Figure 3.9(a) shows the profile distribution of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$ as a function of x_{RPC} for layer-1 of BOL4A13. Because these are two independent measurements of the same coordinate, we expect to see on average a 100% correlation, if they are calibrated well. Since our simulation is for an ideal case where the speed propagation was put to be the speed of light, the profile distribution shows a perfect correlation between $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$ and x_{RPC}. For data we see that this is not the case, as the actual signal speed in the tubes is slower than the speed of light.

For every tube we calibrate the signal propagation speed, by fitting a linear function to the profile distribution of x_{RPC} versus $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$, not taking side acceptance effects into account by fitting in the $|x_{\mathrm{RPC}}| < 2100$ mm range. The slope of the fit function α is used to calibrate the effective signal speed by dividing the speed of light by α, and putting the new value into Equation 3.5. The fit values of α are in between 1 and 1.5 for all tubes, which is good agreement with two propagation speeds of 266 and 274 mm/ns found in a test setup [91]. A few tubes returned unphysical results for the value of α due to the low statistics, that will be discussed later, and these tubes were not used in the plots after calibration.

After applying the calibration of δx and α on every tube, we again plot x_{RPC} versus $x'_{\mathrm{twin}\downarrow\mathrm{RPC}}$ in Figure 3.9(b). For clarity we denote the calibrated twin coordinate with an extra apostrophe x'_{twin}. Although we have calibrated tube by tube, the distribution in Figure 3.9(b) is for a whole layer containing 72 tubes. The almost perfect agreement between data and simulation gives us confidence that we have calibrated the twin tubes correctly.

3.4.3 Twin position dependence of ADC value

As explained before, the damping of the signal is the main ingredient to an effective speed of propagation which is smaller than the speed of light. It is caused by the impedance of the wire and it can differ between tubes, as small wire thickness differences will lead to different impedances. The question is however whether we measure the damping of the signal with the twin tube setup. It is answered by Figure 3.10(a), which shows the profile distribution of ADC counts as a function of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$.

As a reminder we would like to mention that the RO side of the tubes is situated at $x \sim 2500$ mm, while the HV side is at $x \sim -2500$ mm. Figure 3.10(a) shows the ADC counts as a function of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$ for both the prime and twin hits, as RPC hits are an independent measurement of the muon hit position. If the muon passes near the HV side, the signal travels through the prime tube almost the same distance as it does through the twin tube. Hence we expect the ADC counts of the twin and prime hits for the HV side muons to be almost equal. If on the other side the muon passes near the RO side, the prime signal is almost immediately recorded and does not get dampened at all, while the twin signal has to travel the length of almost two 5 meter long tubes and gets dampened a lot. The difference between ADC counts of prime and twin hits should hence be the largest for the RO side muons. In between these two extremes the dampening of both signals happens, but the twin signal always get dampened more as it must travel at least one whole tube length.

This behavior is exhibited by the cosmic ray muon data as shown by Figure 3.10(a), while for the simulation this was not implemented and is hence not seen. However what is important is that our linearized model seems to be confirmed by the data, hence with the exception of the small slewing correction we have calibrated our twin tubes.

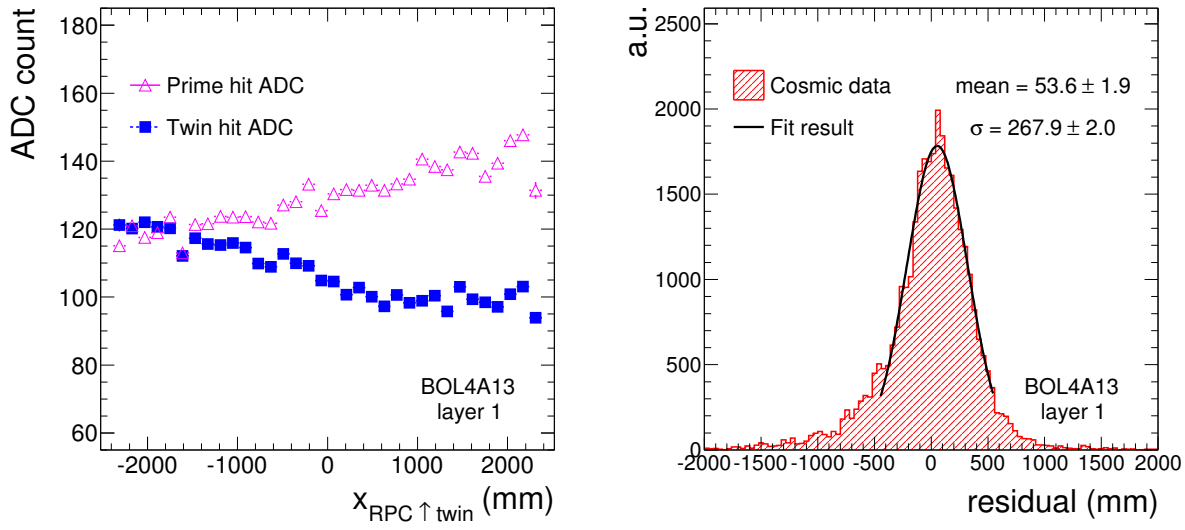
The slewing effect is caused by the finite height of the signal in conjunction with the applied noise threshold of approximately 50 ADC counts. For larger signals the TDC recorded time-over-threshold is somewhat shorter than for smaller signals. However as the primary pulse slope is quite steep, as can be seen from Figure 3.2(b), we expect the effect of slewing to be small. Figure 3.10(a) suggests that a slewing calibration could be performed, but such precision falls outside the scope of this thesis.

3.4.4 Calibrated twin coordinate resolution

If we now plot the residual for data again after calibration, we get the distribution shown in Figure 3.10(b). Comparing this distribution to the residual distribution of Figure 3.7(d), we can see the effects of calibration.

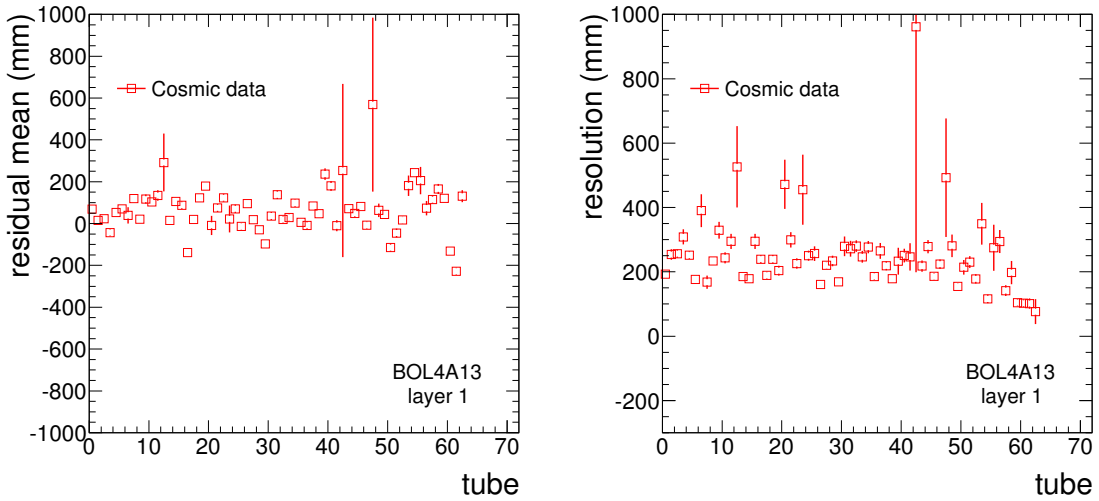
We define the twin coordinate resolution as the standard deviation of a Gaussian function fit to the residual distribution. The Gaussian fit is performed only to the central region ($|\mathrm{residual}| < 500$ mm), as the tails are mostly due to propagation dampening fluctuations, noisy tubes and multiple scattering. The twin coordinate resolution is measured to be 26.7 cm. As expected, it is somewhat higher than the 17 cm resolution that was measured in the test setup. However in the test setup the signals were enhanced by working at a higher HV of 3300 V as compared to 3080 V used for data taking, as well some strong cleaning cuts were applied to the test data before calculating the resolution.

The resolution is composed of a few (systematic) effects, of which the respective contributions are estimated in Table 3.1 [85, 96]. Firstly the discrete nature of our time measurements with TDC counts of finite precision of 0.78 ns, may lead to a worsening of the twin coordinate resolution. If a signal falls just inside or outside the 0.78 ns window, it can have a significant change of twin coordinate, as it corresponds to a shift of 117 mm, for signal propagation with



(a) (b)

Figure 3.10: Mean ADC counts of prime and twin hits (a) as a function of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$. For muons passing close to the HV side the ADC counts of prime and twin hits are very similar as expected, while the difference increases as muons hit closer to the RO side. The residual after calibration (b) for layer-1 of chamber BOL4A13 with a Gaussian fit to the central region.



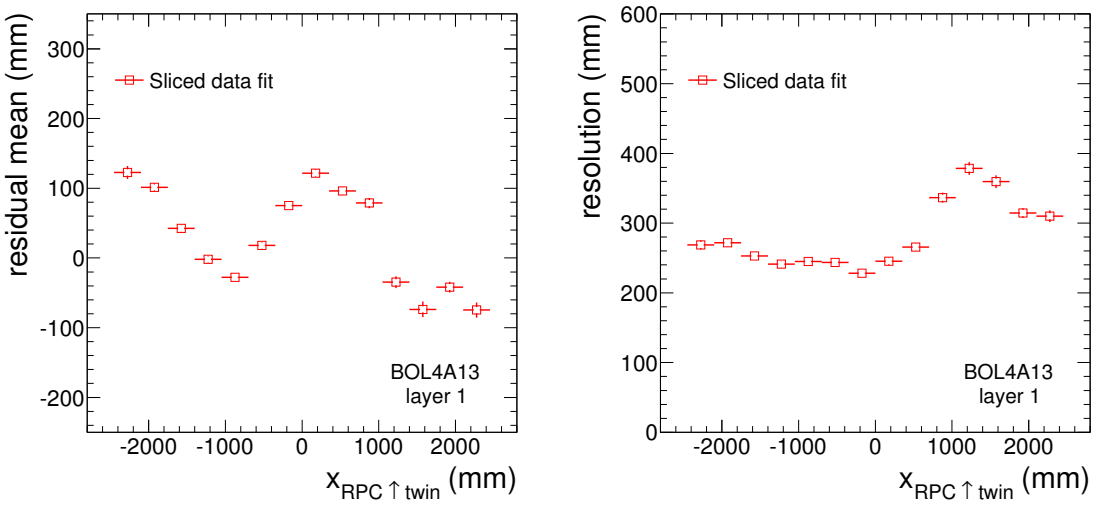
(a) (b)

Figure 3.11: Fitted residual mean and resolution as a function of tube number for layer-1 of chamber BOL4A13. The spread of resolution over all the tubes is quite large, however statistics does not allow to study resolution in each tube as a function of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$.

the speed of light. Also due to the lack of precise alignment and global t_0 calibration our tracks might have a systematic shift in their angle, which also degrades the resolution, as well as the relaxed hit to track association. The aforementioned slewing effect can worsen the resolution, as well as the fluctuation of the signal due to propagation. On top of that we have assumed our RPC hit measurement to be the baseline, but the RPC (x, z)-measurement of course also has a

Effect	TDC bin size	Alignment & t_0 calibration	Slewing	RPC resolution	Delay & reflections
Contribution (ns)	< 1.6	< 1	< 3	1	< 1

Table 3.1: The estimated contribution of different components to the resolution.



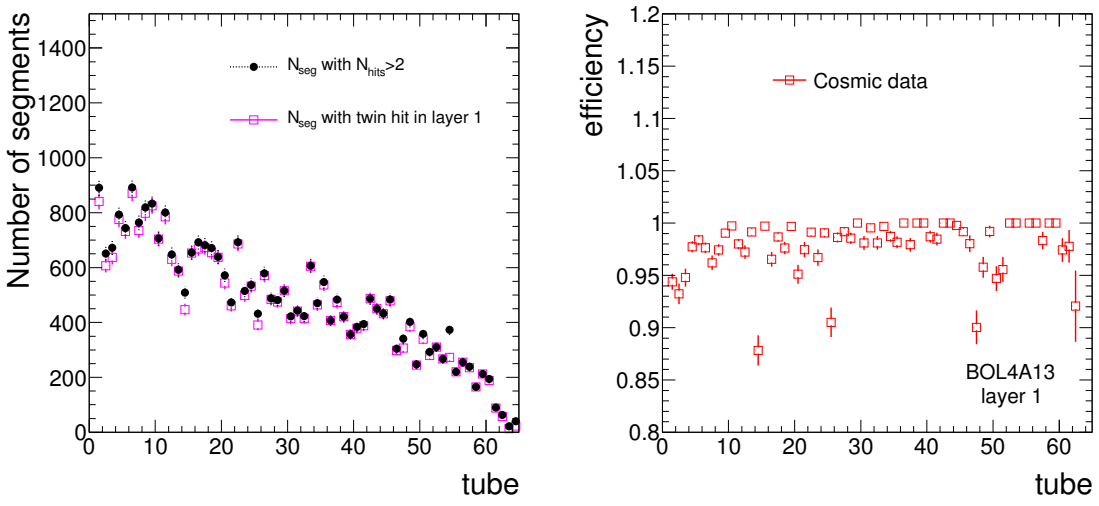
(a) (b)

Figure 3.12: Fitted residual mean and resolution as a function of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$ for layer-1 of chamber BOL4A13. As the hits move closer to the RO side the slewing effect increases and the resolution becomes worse.

non-zero resolution that needs to be taken into account. Lastly the built-in delays may have a jitter around the average of 6 ns, as well as reflections will affect the resolution. The residual fitted mean of ~ 5.4 cm away from zero can also be caused by a number of the effects mentioned above.

As we are summing over all the tubes of layer-1 in the residual plot, we could benefit by studying the residual for each tube separately. Figures 3.11(a) and 3.11(b) show respectively the fitted residual mean and twin coordinate resolution as a function of tube number. The residual mean has values between -10 and 20 cm, while the resolution is between 15 and 30 cm. This shows that the resolution of 26.7 cm quoted above is probably on the high side, as many Gaussians with different means and resolutions are summed together, leading to an average resolution that is overestimated. If we perform a horizontal line fit to the resolution per tube distribution, we get an average resolution of 19.9 cm.

We can also study the evolution of the twin coordinate residual in the global ϕ-dimension (along the tube length), as opposed to the global η-dimension (increasing with increasing tube number). Statistics are too low to study the resolution dependence on $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$ for every tube, hence we must again go back to a whole layer. Figures 3.12(a) and 3.12(b) show respectively the residual mean and resolution as a function of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$ for layer-1 of BOL4A13. We expect the resolution to become worse as the signal gets more dampened, so for muons passing at the RO side the resolution should be worse than for muons passing close to the HV side. This behavior is shown by Figure 3.12(b), however it is not a linear rise as the ADC value difference between



(a) (b)

Figure 3.13: Efficiency (b) of layer-1 for BOL4A13 chamber as a function of tube number. The two parts of the efficiency calculation (a), number of segments with at least three hits and number of segments with a twin hit in layer-1, as a function of tube number. Some noisy tubes can be seen, as well as decreased efficiency for low tube numbers due to acceptance effects.

prime and twin hits. This can be understood as again we are summing different tubes with different resolutions and residual offsets together. This can also be seen by comparing the average resolution as a function of tube number (Figure 3.11(b)) and as a function of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$ (Figure 3.12(b)). On top of that the residual mean varies itself for different $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$ values, as shown by Figures 3.12(a).

When we study each of the $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$-slices that were used for the residual fits, the following can be noted. For hits close to the HV side, a small fluctuation that can be caused by reflections, can have a relatively big influence on Δt and disturb the measurement. As we move away from the HV side, reflections will have less effect, thus though the tails (due to fluctuations) get wider the resolution as measured in the central region becomes somewhat better. As we move even further towards the RO side, the expected damping effect becomes more prominent, degrading both the tails and the central residual region, thus producing the worst resolution. A few cross checks were performed, to make sure that for example we are not mislead by a noisy RPC strip or a miscalibration of the residual subparts. None of the cross checks showed anything unexpected, and the resolution evolution is as shown.

3.4.5 Twin hit efficiency

The efficiency of the twin tube setup is the second important measure of performance, besides the resolution. We define the efficiency per layer to be:

$$\varepsilon = \frac{N_{segments} \text{ with twin hit in layer}}{N_{segments} \text{ with } > 2 \text{ hits}}, \tag{3.9}$$

where a *twin hit* is defined as an MDT prime hit on track, for which the second coordinate has been calculated using the time difference with the twin partner. In the numerator we count the number of segments with a twin hit in that specific layer. The pairing probability of a hit

with the corresponding twin hit was measured to be 99.8% [91]. In the denominator we count the number of segments that have at least 3 MDT hits, which can be in either the twin tube multilayer or the other multilayer.

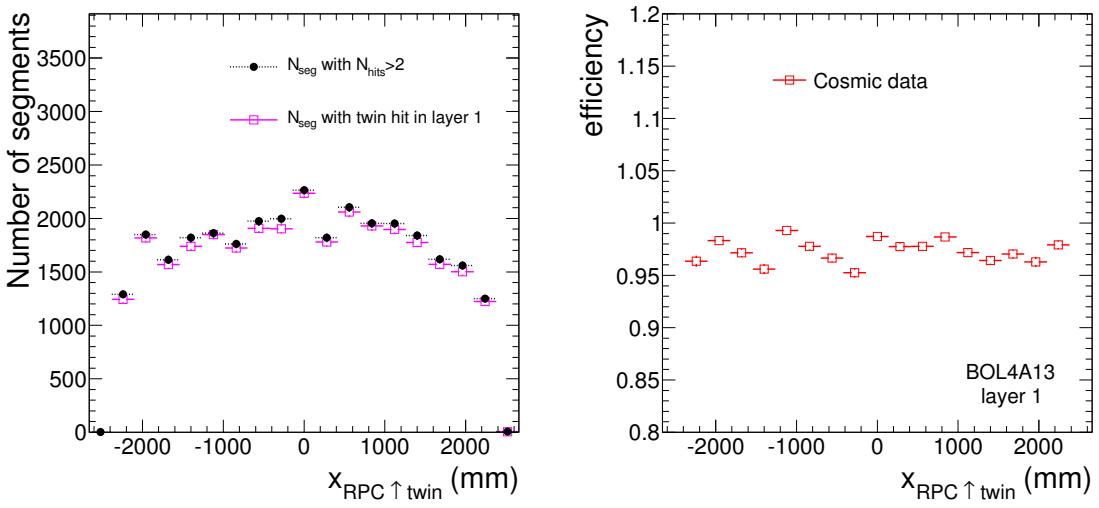
Since the second coordinate of a segment must come from the RPC hits, we select only those segments that have an RPC hit on track from a ϕ-strip. Otherwise the ϕ coordinate of the track is only coming from the middle chambers, and since there were known misalignments, the hit to track association for the extrapolated track might be incorrect. As a consequence of this selection and the way the RPCs are attached to the MDT chambers, as shown in Figure 3.1(a), we at the same time apply a geometrical acceptance cut. As we are looking at the lower most sector of ATLAS, the RPC is under the MDT chamber (if you rotate Figure 3.1(a) by 180 degrees), hence vertically descending muons should in principle hit all the MDT layers if it also hits the RPC.

Figure 3.13(b) shows the efficiency as a function of tube number for layer-1 of BOL4A13. The efficiency errors are calculated using binomial statistics as implemented in ROOT [97]. On average an efficiency of 98.8% is measured, even though a few noisy tubes are seen such as numbers 14, 25 and 47. Although our selection criteria make sure that geometrical effects are minimized, we still see an efficiency reduction towards the very low tube numbers. As these tubes are on the outside of the chamber, and muons might still have a little angle to the vertical axis, it is possible to have the required 3 MDT hits in the second (non twin tube) multilayer and an RPC hit on a segment, while non of the twin tubes are hit. You expect this geometrical effect to be only of importance for the first few tubes as can be seen in Figure 3.13(b), since ϕ_{track} was shown to differ from vertical only slightly.

The denominator and the numerator of the efficiency calculation are shown in Figure 3.13(a) as a function of tube number. The apparent drop in number of segments with increasing tube number was not expected a priori. However since we are in the lower most sector of ATLAS, we have the whole detector above the twin tube chambers. Furthermore, the support structures such as the feet of the detector are all around the lower most sector. However we expect all this to be correctly implemented in the geometry description. When we compare the distributions of Figure 3.13(a) to the simulated distributions, almost exactly the same evolution is seen. Hence we conclude that the twin chambers fall under the cosmic ray muon shadow of the support structure and the rest of the detector, and thus the higher tube numbers are less illuminated. A similar shadowing effect is also seen on the other side in the BOL4C13 chamber.

To explain why the efficiencies are not exactly 100% we need to look at possible sources of inefficiencies. Two different types of inefficiencies can be defined: hardware inefficiency due to absence of a hit in the tube, and tracking inefficiency when a hit is not associated to the segment because its residual is larger than the association cut [88]. Hardware inefficiency is very small, mostly occurring at large drift distances near the tube wall, where the short track length results in fewer primary electrons or due to the track passing through the dead material between adjacent tubes. The tracking inefficiency is dominated by δ-electrons, produced by the muon itself, which can mask the muon hit if the δ-electron has a smaller drift time than the muon. Tube noise can be an additional source of this type of inefficiency.

As we did for the resolution, we can also study the evolution of efficiency in the other dimension than the tube number dimension. Figure 3.14(b) shows efficiency and Figure 3.14(a) shows the denominator and numerator of the efficiency calculation as a function of $x_{\text{RPC}\uparrow\text{twin}}$. For all values of $x_{\text{RPC}\uparrow\text{twin}}$ efficiency is between 95% and 100%. Although you expect efficiency to be worse for muons passing close to the RO side, as the twin signal gets dampened, this is not seen in data. This can understood by looking at Figure 3.10(a), where it can be seen that even for muons passing at the RO side, the twin signal ADC count on average is well above



(a) (b)

Figure 3.14: Efficiency (b) of layer-1 for BOL4A13 chamber as a function of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$. The two parts of the efficiency calculation (a), number of segments with at least three hits and number of segments with a hit in layer-1, as a function of $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$.

the ADC noise pedestal of approximately 50 ADC counts. The shadow effect seen in the tube number dimension can also be seen in $x_{\mathrm{RPC}\uparrow\mathrm{twin}}$, though it is not as striking, but tubes seem to be less illuminated away from the center of the tube.

3.5 Conclusions

The twin tube MDT setup has been installed in two chambers of the ATLAS muon spectrometer on the inner multilayer. Using cosmic ray muons it has been calibrated and its performance has been assessed. The coordinate along the wire direction can be measured with an average resolution of 19.9 cm by a single twin tube. When we would combine the measurements of each layer together a resolution of 11.5 cm can be obtained. The efficiency of the twin tube setup has been measured to be 98.8%, for which some small geometrical effects and noisy tubes are not taken into account and hence degrade the total efficiency.

The twin tube setup could be installed in the future in more ATLAS MDT chambers, by which full 3D standalone tracking could be done offline with the MDTs only. Since twin tubes increase MDT occupancy by a factor 2, they should only be installed in areas where the required occupancy allows. The RPC chambers would still be unmissable to deliver a fast momentum dependent trigger. However twin tubes could help pattern recognition in solving ambiguities for multiple hits on one chamber.

4

Combined Fit Method for Background Estimation to Supersymmetry

4.1 Introduction

Any discovery of new physics, supersymmetry or other, can only be claimed when the Standard Model backgrounds are understood and are under control. It is expected that at the LHC Monte Carlo predictions are not sufficient to achieve this, as these have never been tested at such high energies: the backgrounds will have to be derived from the data itself, possibly helped by Monte Carlo. The development and description of such a *data-driven* background estimation technique is the topic of this chapter, discussed as well in [98, 99].

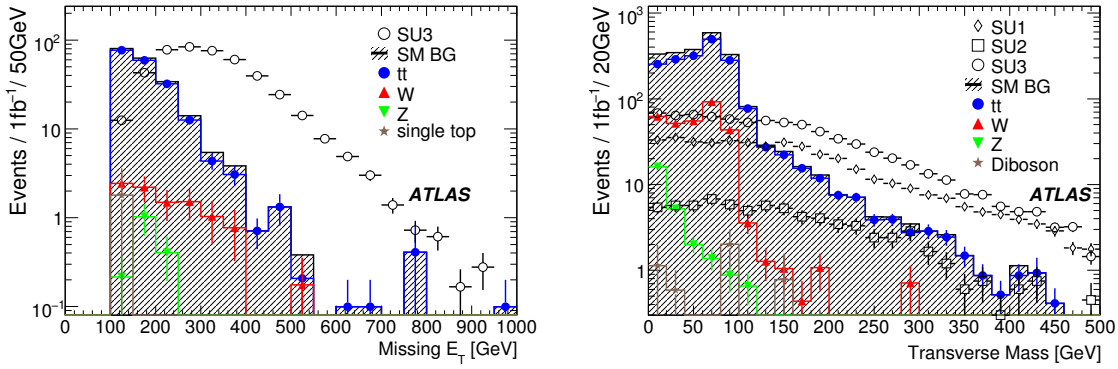
The general SUSY signature under study is multiple energetic jets and large missing transverse energy, possibly accompanied by one or more energetic and isolated leptons. Because of the high uncertainties on the multijet QCD cross sections, we focus here on the study of events with exactly *one isolated lepton*. The background from QCD multijet events should be strongly suppressed by the isolated lepton requirement, while keeping most signal events. Requiring two or more isolated leptons will further suppress this background, but it will also eat into the signal. The trade off between background suppression and signal significance is thus best for one lepton analysis.

General strategy

The starting point for our strategy is a counting experiment: we want to determine the number of SM events we expect in a certain phase space region, and we compare this expected number of events to the number of events found in the data. To fix the region where we search for the excess of events, called the *signal region*, it is natural to look at two observables for which the SUSY spectrum typically extends to much higher values than the spectrum of SM backgrounds. These observables are the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse mass M_{T}, see Figure 4.1. By cutting at a certain large value of $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T}, we can get a high signal over background ratio in our signal region.

To determine the SM background contribution in the high $E_{\mathrm{T}}^{\mathrm{miss}}$-$M_{\mathrm{T}}$ region, we use the fact that this region is 2-dimensional. We define an *L-shaped control region* in the low $E_{\mathrm{T}}^{\mathrm{miss}}$-$M_{\mathrm{T}}$ plane, shown in Figure 4.2, to estimate the shapes and relative fractions of the SM backgrounds. Thus we can use the full reach of the M_{T} distribution of the data at low $E_{\mathrm{T}}^{\mathrm{miss}}$, and the full $E_{\mathrm{T}}^{\mathrm{miss}}$ reach at low M_{T}. Extrapolating into high $E_{\mathrm{T}}^{\mathrm{miss}}$-$M_{\mathrm{T}}$ region gives us an estimate of the number of SM background events in our SUSY signal region.

The dominant backgrounds to SUSY in the one lepton channel are semileptonic and dilep-



(a) (b)

Figure 4.1: Missing transverse energy (a) and transverse mass (b) for the main backgrounds in one lepton channel and some ATLAS SUSY benchmark points (SU1, SU2, SU3). Though these distributions were made for $\sqrt{s} = 14$ TeV [60], the 10 TeV distributions are quite similar.

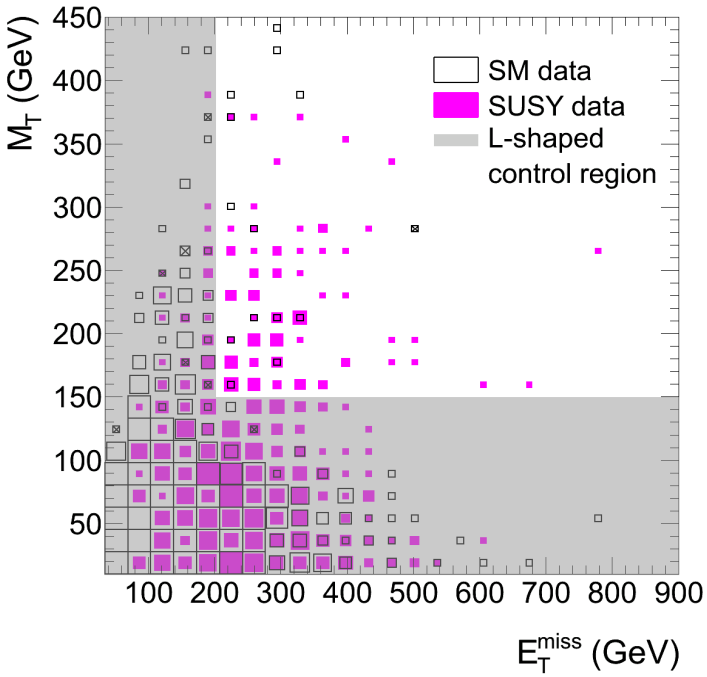


Figure 4.2: The L-shaped control region visualized as a semi-transparent band on top of the simulated SM backgrounds data in empty squares and a SUSY model (reference point 2 as discussed in Section 4.4) in filled squares.

tonic $t\bar{t}$, and $W + jets$ events. It is not possible to distinguish between these processes and SUSY production on an event by event basis. Therefore we construct a model that combines specific physics features of these backgrounds, such as mass peaks or phase space cutoffs, to estimate the probability that an event is either one or the other.

We model the backgrounds by their behavior in three observables: $E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T} and M_{jjj}, defined in Section 4.1.3. The data distributions in these observables combine features specific

to SUSY (hard $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} spectra) with features that are unique to the SM backgrounds. The top mass peak in the M_{jjj} distribution is a distinguishing trait of semileptonic $t\bar{t}$ events. A W-mass Jacobian peak in the M_{T} distribution is exhibited by both the $W + jets$ and the semileptonic $t\bar{t}$ events. We have not been able to define an observable that shows a unique signature for dileptonic $t\bar{t}$ events. Our model depends on a combination of the behavior of dileptonic $t\bar{t}$ in all three observables.

To keep dependence on MC as small as possible, we parametrize the shapes of the backgrounds using *probability density functions* with parameters that can be floated, and perform an unbinned maximum likelihood fit in the control region to constrain the model parameters from data. The mathematics involved and partly developed specifically by our method are described in Section 4.2.

Problems of a simple counting experiment

There are two issues with the naive approach of a counting experiment. First, if SUSY events are produced at the LHC, there will be a non-negligible amount of SUSY events in our control region. To account for this *SUSY contamination* we need to model our backgrounds *and* our signal in the control region to be able to separate these two contributions, and be capable to extrapolate the backgrounds model only to the signal region.

As it turns out, it is possible to model a wide range of different mSUGRA models with a simple two-parameter model in the control region, as described in Section 4.3.6. By combining our SUSY model with our background model, we can fit the shape parameters and the ratio of signal to background events in our control region in one iteration.

The second issue with a naive counting experiment is the underlying assumption that the shape of the M_{T} ($E_{\mathrm{T}}^{\mathrm{miss}}$) distribution does not change with increasing $E_{\mathrm{T}}^{\mathrm{miss}}$ (M_{T}). This assumption is incorrect for some backgrounds and we have to take these correlations into account, as described in Section 4.3.3.

4.1.1 Signal and Background generation

In the following section we briefly describe the signal and background simulated samples used in this study, produced following the strategy described in Sections 1.1.8 and 1.2.6. All datasets were simulated with a center of mass energy of $\sqrt{s} = 10$ TeV, which in 2009 was assumed to be the collision energy for the LHC once it would restart. For this analysis we do not expect a large change of strategy to be necessary to be able to apply it to the current LHC collision energy of 7 TeV. Some changes to event selection would most likely be enough to maximize the significance reach of our analysis.

Signal generation and mSUGRA grid

We focus on R-parity conserving supersymmetric models where SUSY breaking is mediated by gravitational interaction (mSUGRA), as described in Section 1.2.5. Since there is no unique model of SUSY breaking, these models should be viewed only as possible patterns of LHC signatures, not as complete theories.

mSUGRA grid In order to cover a large parameter space and different phenomenologies, but to reduce the number of SUSY points an mSUGRA grid was produced [100] using the ATLFAST2 simulation package, discussed in Section 1.1.8. The grid was made in "radial coordinates", i.e. points on outgoing radial lines in the $(m_0, m_{1/2})$ plane for $\tan\beta = 10$ and 50.

N partons	Cross Section (pb)	
	$t\bar{t}(l\nu l\nu) + jets$	$t\bar{t}(l\nu qq) + jets$
0	12.7	51.8
1	13.7	57.1
2	9.4	38.3
3	7.1	27.6

Table 4.1: Cross sections for ALPGEN simulated $t\bar{t}$ processes with different generators. The cross sections are given after MLM matching and include the calculated NLO k-factor.

N partons	Cross Section (pb)	
	$W(l\nu) + jets$	$W(l\nu) + b\bar{b} + jets$
0	12196.5	6.2
1	2549.9	6.1
2	812.4	3.5
3	243.2	2.0
4	66.8	
5	19.9	

Table 4.2: Cross sections for simulated ALPGEN $W + jets$ and $W + b\bar{b} + jets$ processes, where l stands for lepton (e, μ or τ), multiplied by the respective k-factors to normalize to NNLO. The cross sections are given after MLM matching was applied.

For each $\tan\beta$ value, lines with different slopes in the $(m_0, m_{1/2})$ plane were produced, while the other parameters were kept at $A_0 = 0$ and $\mu > 0$.

In our analysis we used approximately 60 models produced in the mSUGRA grid with m_0 values ranging from 50 to 2000 GeV and $m_{1/2}$ values in-between 100 and 450 GeV as can be seen from Figure 4.25. The leading order HERWIG cross sections in this SUSY parameter grid extend from ~ 0.3 pb for the highest mass SUSY point to approximately 500 pb for the lowest mass SUSY point, with a notable exception of one point with a cross section of 2818 pb with $m_0 = 150$ GeV, $m_{1/2} = 150$ GeV. As discussed in Section 1.2.6, after producing the simulated samples these cross sections are multiplied by the respective k-factor to compensate for missing higher order terms.

Backgrounds

Different Monte Carlo generators were used for different background samples. This was done in an attempt to optimize the reliability of the estimate and cross check the results for the Standard Model samples.

Top quark pair production The $t\bar{t}$ (and single top) simulation sample was produced using the MC@NLO generator. The cross section for the joint semileptonic and dileptonic $t\bar{t}$ sample as calculated by MC@NLO is 217 pb [20].

Besides a MC@NLO simulated sample, alternate $t\bar{t}$ samples were produced using the ALPGEN generator. We use the ALPGEN $t\bar{t}$ events for a study of systematic uncertainties. The cross sections for all the separate ALPGEN samples given in Table 4.1 are calculated after MLM matching, which removes duplicate final states as discussed in Section 1.1.8.

QCD light jets				
N partons	min−max p_{T} (GeV)			
	35−70	70−140	140−280	280−∞
2	30114237	1116549	31872	750
3	9835390	1486726	65509	1945
4	1494832	552311	49028	2150
5	249185	189793	24249	1393
6			11572	973

QCD($b\bar{b}$) + jets				
N partons	min−max p_{T} (GeV)			
	35−70	70−140	140−280	280−∞
0	137665	5398.1	147.9	3.2
1	193821	27239.6	1078.6	25.2
2	53806.5	18592	1430.1	50.1
3	13470.9	9460.5	1021.2	52.9
4			706.5	55.5

Table 4.3: Cross sections in pb for simulated ALPGEN QCD multijet processes, that were split according to quark flavour and p_{T} of the leading jet. The last two columns give the minimal and maximal transverse momentum of the leading jet as was used for the separation of the samples. The cross sections are given after MLM matching was applied.

W + jets The cross sections of ALPGEN simulated $W+jets$ samples after applying the MLM matching technique are given in Table 4.2, where l stands for leptons of all possible flavours (e, μ or τ). All the different flavours must be summed to get a complete sample. The k-factors are calculated by comparing the NNLO cross sections, computed by the FEWZ program, to the LO simulation cross sections. Because the higher order terms are missing in the simulation, the leading order produced sample cross sections are multiplied by the k-factor.

W + $b\bar{b}$ + jets The $W + jets$ processes simulated by ALPGEN only take into account light flavour (*u, d, s* and *c* quarks) jets. A separate ALPGEN process takes care of the $W + b\bar{b} + jets$ production. Although these processes have relatively low cross sections, they must be considered to correctly estimate the total cross section and if b-tagging is to be used. A small overlap is expected between the $W + b\bar{b} + jets$ and $W + light\ jets$ samples, as the latter may contain $b\bar{b}$ pairs generated by parton showering. This small amount of double counting is minimized by the choice of the generator level cuts [20].

QCD multijet Although the requirement of a single isolated lepton and missing transverse energy strongly suppresses QCD multijet events, the cross sections of these processes are orders of magnitude higher than the processes involving top quarks and W bosons, and as such still have to be taken into account. ALPGEN was again the generator of choice, for it can calculate matrix elements of events with up to 6 partons in the final state.

The generation of QCD events was split according to the transverse momentum of the lead-

ing jet, to be able to produce useful amounts of integrated luminosity. The lowest produced p_{T} of the leading jet was set at 35 GeV, due to practical limitations related to cross sections of lower p_{T} samples. For the same reason the $35-70$, $70-140$ and $140-280$ GeV p_{T} samples could only be produced with an integrated luminosity of 10 pb^{-1}, while the highest ($280-\infty$ GeV) p_{T} sample was produced with an integrated luminosity of 300 pb^{-1}.

The produced QCD samples with corresponding cross sections and allowed leading jet transverse momenta are given in Table 4.3. Just as for the $W+jets$ simulation, the QCD ALPGEN process with $b\bar{b}$ pairs has to be produced separately from the light jets process. All the samples in Table 4.3 must be summed to arrive at a complete simulated QCD multijet sample.

4.1.2 Object identification for SUSY analysis

Supersymmetric events are characterized by several high momentum jets, missing transverse energy and in our case an isolated lepton. In this section we will describe the particle identification criteria.

Jets The principal detector for jet reconstruction in ATLAS is the calorimeter system, described in detail in Section 2.4. The jet definition used in this analysis is a fixed-cone jet with $R_{cone} = 0.4$, seeded by topological cell clusters and calibrated using the H1 scheme, called `Cone4H1TopoJets` in ATLAS nomenclature. Specifically the narrow cone size was chosen because of the large multiplicity of jets in SUSY events.

Electrons In ATLAS electrons are reconstructed by the dedicated `egamma` algorithm [80], as described in Section 2.4.5. In this analysis we use *medium* electrons with an isolation requirement. The transverse isolation energy in a cone of $\Delta R < 0.2$ around the electron is required to be smaller than 10 GeV. In the available simulation datasets, a bias is incorrectly introduced in the crack region $1.37 < |\eta| < 1.52$. Events with an electron reconstructed in this region are therefore rejected.

As jets and electrons are both objects reconstructed from the calorimeters an overlap removal procedure is defined. Jets reconstructed within a cone $\Delta R = 0.2$ of an identified electron are discarded from the jet list, to prevent double counting the same object as both a jet and an electron.

Muons Combining the measurements in the muon spectrometer with both the inner detector and the calorimeter information delivers the best possible quality of muon reconstruction, as described in detail in Section 2.5.6. In this analysis, combining of the inner detector and the stand-alone muon spectrometer tracks is done by the `Staco` algorithm [101], also known as Chain 1 in [86].

The match χ^2, defined as the difference between outer and inner track vectors weighted by their combined covariance matrix, provides a measure of the quality of this match and is used to decide which pairs are retained as *combined muons*. We require that the tracks should match with $\chi^2 < 100$. Just as for electrons, the muons are required to be isolated, by demanding that the total calorimeter energy deposited in a cone of $\Delta R < 0.3$ around the muon be less than 10 GeV. Muons found to overlap with a jet within $\Delta R < 0.4$ are removed from the collection.

4.1.3 Global Variables

Most SUSY analyses use some global event variables that have good signal discriminating or descriptive power. In this section we define those used in our analysis.

Missing Transverse Energy

Measurement of missing transverse energy requires very detailed calibration and measurement of the calorimeters and the muon spectrometer. The particles not interacting with the detector, such as neutrinos in the SM and the LSPs for SUSY, carry away energy. By summing all energy deposits in the detector, we get the opposite vector of this missing energy, when sub-detectors are calibrated correctly. Since at the LHC we do not know the boost of the produced particles along the beam-axis, we can only speak of missing energy measurement in the transverse plane, or $E_{\mathrm{T}}^{\mathrm{miss}}$.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction used in ATLAS starts by summing all calorimeter deposits, then correcting for the measured energy of the muons and the energy lost in inactive material of the cryostat. The $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithm first calculates the sum in the two transverse directions, x and y as follows:

$$E_{x,y}^{\mathrm{Final}} = E_{x,y}^{\mathrm{Calo}} + E_{x,y}^{\mathrm{Muon}} + E_{x,y}^{\mathrm{Cryo}}, \tag{4.1}$$

where

$$E_{x,y}^{\mathrm{Calo}} = -\sum_{\mathrm{TopoCells}} E_{x,y}, \tag{4.2}$$

$$E_{x,y}^{\mathrm{Muon}} = -\sum_{\mathrm{muons}} E_{x,y} \tag{4.3}$$

$$E_{x,y}^{\mathrm{Cryo}} = -\sum_{\mathrm{jets}} w^{\mathrm{Cryo}} \sqrt{E_{x,y}^{\mathrm{EM3}} \times E_{x,y}^{\mathrm{HAD}}}. \tag{4.4}$$

$E_{x,y}^{\mathrm{Calo}}$ term Only calorimeter cells that are associated to a TopoCluster contribute to $E_{x,y}^{\mathrm{Calo}}$, to suppress noise. The ATLAS calorimeters are non-compensating, which means that the response to hadrons is lower than to electrons, hence a calibration must be applied. ATLAS has two different calibration schemes. One is based on global cell energy-density weighting (GCW), where by applying cell-level weights low density signals are enhanced, which are more likely coming from hadronic activity. The other approach is local cluster weighting (LCW) scheme, that uses properties of TopoClusters to calibrate them individually, also called H1 scheme in Section 2.4.4. The LCW scheme is used to calibrate the calorimeter contribution to $E_{\mathrm{T}}^{\mathrm{miss}}$ in our method.

The $E_{x,y}^{\mathrm{Calo}}$ calculation is further refined by calibrating each contribution according to the reconstructed object it is assigned to. The assignment adheres to the following order: electrons, photons, muons, hadronically decaying taus, b-jets and finally light jets. Thus $E_{x,y}^{\mathrm{Calo}}$ ($E_{x,y}^{\mathrm{Final}}$) is replaced by the refined $E_{x,y}^{\mathrm{RefCalo}}$ ($E_{x,y}^{\mathrm{RefFinal}}$) defined as:

$$E_{x,y}^{\mathrm{RefCalo}} = E_{x,y}^{\mathrm{RefEle}} + E_{x,y}^{\mathrm{RefGamma}} + E_{x,y}^{\mathrm{RefTau}} + E_{x,y}^{\mathrm{RefJet}} + E_{x,y}^{\mathrm{RefMuon}} + E_{x,y}^{\mathrm{CellOut}}. \tag{4.5}$$

As before each term in 4.5 is calculated as the negative sum of calibrated cells inside the specific object. All TopoClusters calorimeter cells without an object assignment are collected in the $E_{x,y}^{\mathrm{CellOut}}$ term.

$E_{x,y}^{\mathbf{Muon}}$ **term** For non-isolated muons, the calorimeter term already accounts for the muon energy deposits in the calorimeter, as the energy lost in the calorimeter cannot be separated from the nearby jet energy. So for the muon term $E_{x,y}^{\mathrm{Muon}}$, the momenta as measured in the muon spectrometer by the combined algorithm are used to prevent double counting. All the combined muons in the $|\eta| < 2.5$ are summed, while for the region uncovered by the inner detector $2.5 < |\eta| < 2.7$ the standalone muon spectrometer momenta are used. For isolated muons the combined muon momentum is used for the muon term $E_{x,y}^{\mathrm{Muon}}$ in Equation 4.2. In the refined calibration, for the $E_{x,y}^{\mathrm{RefMuon}}$ term only non-isolated muons are taken into account, as the energy of the isolated ones are already in the $E_{x,y}^{\mathrm{Muon}}$ term.

$E_{x,y}^{\mathbf{Cryo}}$ **term** The last term of Equation 4.1 accounts for the energy lost in the dead material of the cryostat between the electromagnetic and hadronic calorimeters. The last layer of the electromagnetic LAr calorimeter, $E_{x,y}^{\mathrm{EM3}}$, is compared to the first layer of the hadronic calorimeter, $E_{x,y}^{\mathrm{HAD}}$, and for each jet a calibration weight, w^{Cryo}, is applied.

$\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ Finally the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is calculated by taking the length of the $E_{x,y}^{\mathrm{RefFinal}}$ vector:

$$E_{\mathrm{T}}^{\mathrm{miss}} = E_{x,y}^{\mathrm{RefFinal}} = \sqrt{(E_x^{\mathrm{RefFinal}})^2 + (E_y^{\mathrm{RefFinal}})^2} \tag{4.6}$$

$E_{\mathrm{T}}^{\mathrm{miss}}$ is the first observable used in our analysis, as SUSY events are characterized by high missing transverse energy.

Transverse mass

The second observable used in this analysis is *transverse mass*, M_{T}, that is defined as:

$$M_{\mathrm{T}} = \sqrt{2(p_T^{lep} E_{\mathrm{T}}^{\mathrm{miss}} - \overrightarrow{p}_T^{lep} \cdot \overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}})}, \tag{4.7}$$

where p_T^{lep} is the transverse momentum of the lepton.

For events that contain a single leptonic W boson decay, such a semileptonic $t\bar{t}$ and $W + jets$ events, the M_{T} distribution displays a characteristic Jacobian peak around the W mass. A good approximation is that the neutrino from the W is the sole responsible for the missing transverse energy. When combined with the lepton from the W it results into the transverse W mass, which is at maximum equal to m_W. We assume here that the masses of neutrino and lepton are negligible in comparison to their momenta, which is correct for W boson decays. For one lepton SUSY events where $E_{\mathrm{T}}^{\mathrm{miss}}$ is a superposition of the two LSPs and possibly neutrinos from decay of SM particles in the chain, the M_{T} distribution shows no characteristic cut-offs, but is purely defined by the kinematics of the specific supersymmetric model and event selection criteria.

Three-jet mass

The hadronic decay of a top quark results in three hard jets. If we would reconstruct and identify these three jets correctly, we expect to find the top mass peak by calculating the invariant mass of the three-jet system. The invariant *three-jet mass*, M_{jjj}, is the final observable of our analysis and is defined as:

$$M_{\mathrm{jjj}}^2 = (\sum_{i=1}^{3} p_{max\sum p_T}^{jet,i})^2. \tag{4.8}$$

The $p^{jet,i}_{max\sum p_T}$ is the four-momentum of the jets. A simple empirical algorithm is used to identify the three jets coming from the hadronic top decay. Denoted by the $max\sum p_T$ subscript in the equation, the algorithm selects the three jets, which give the three-jet sum combination with the highest transverse momentum to collectively represent our top quark candidate.

There is an ambiguity in choosing the correct three-jet combination among the reconstructed jets of a $t\bar{t}$ pair, even when b-tagging is used (which is not the case in this analysis). The ambiguity arises from the fact that semileptonic $t\bar{t}$ production has 4 partons carrying colour charge in the final state and the two incoming protons, that can all radiate extra partons. The combination with the highest transverse momentum chooses the correct pairing in approximately 25% of all events [102], considerably better than random, given the fact that most events have more than 4 reconstructed jets.

On top of that some partons are not reconstructed correctly into a jet or not with the right momentum or not reconstructed at all if they fall outside the detector acceptance. Another possibility for the algorithm not to reconstruct the top mass, is when all the partons are reconstructed correctly, but if for example the hadronic side bottom quark radiates a hard gluon, the b-jet four-momentum will not represent the original top decay daughter. We would need to take into account the fourth jet of the radiated gluon to correctly reconstruct the top quark mass. The last possibility is that the two partons from a W boson decay will be merged into one jet if the boost of the W is very large, pushing the two partons (quarks) close to each other in $\eta - \phi$ space.

Effective Mass

The effective mass, M_{eff}, is not used as an observable in this analysis but plays a role in our event selection. The definition of effective mass is:

$$M_{\text{eff}} = \sum_{i=1}^{4} p_T^{jet,i} + \sum_{i=1} p_T^{lep,i} + E_{\text{T}}^{\text{miss}}, \tag{4.9}$$

where the sums run respectively over the four highest p_{T} jets within $|\eta| < 2.5$ and over all the identified leptons. This variable has the interesting property that for SUSY events the M_{eff} distribution peaks at a value which is strongly correlated with the mass of the pair of SUSY particles produced in the initial proton-proton interaction, which can be used to quantify the SUSY mass-scale.

4.1.4 Event Selection

Our event selection rests on two principles: we want to keep the highest possible SUSY signal significance in the signal region, while keeping enough of the backgrounds in the control region to correctly estimate their fractions and shapes. The exception to the second principle is the uncertain contribution from QCD multijet events that we try to suppress as much as possible. The event selection criteria can be summarized by:

1. Exactly one isolated muon[1] with $p_{\text{T}} > 20$ GeV and $|\eta| < 2.5$, satisfying identification criteria described earlier.

2. No additional leptons, either electrons or muons, with $p_{\text{T}} > 10$ GeV, known as *2nd lepton veto*.

[1] Due to practical constraints we focus on events with one isolated muon only, though the method described here has been shown to work on events with one isolated electron [60].

3. $E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 0.2 \times M_{\mathrm{eff}}$.

4. At least four jets with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 20$ GeV, out of which the leading jet has to have $p_{\mathrm{T}} > 80$ GeV and the two sub-leading jets must have $p_{\mathrm{T}} > 40$ GeV.

The first cut defines the one-lepton analysis, while the second ensures that the overlap with the zero- and two-lepton analyses is absent, as well as cutting most of the dileptonic $t\bar{t}$ background. Cuts 3 and 4 both strongly reduce the Standard Model backgrounds, while keeping much of the SUSY signal intact. Our choice for the $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold of 40 GeV is somewhat lower than the baseline ATLAS SUSY cut, as we are focused not only on maximizing the signal significance but also the event count of the backgrounds in the control region. For the same reason we do not use a $M_{\mathrm{T}} > 100$ GeV cut usually used by other analyses studying SUSY in the one-lepton case.

The $E_{\mathrm{T}}^{\mathrm{miss}} > 0.2 \times M_{\mathrm{eff}}$ cut is effective in getting rid of most SM backgrounds, specifically the multijet QCD events [60]. For SUSY events the values of $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{eff} are quite correlated, while much less so for backgrounds. Two other variables can be used to suppress the QCD background, transverse sphericity (S_T), and azimuthal angle difference between jets and missing transverse energy ($\Delta\phi(jet, E_{\mathrm{T}}^{\mathrm{miss}})$). The first assumes that the heavy SUSY particles produced in the initial interaction are almost at rest, hence their decay products would be distributed isotropically, while for the QCD events the direction of the two partons from the hard scattering provides a privileged direction. The second is based on the fact that for QCD events $E_{\mathrm{T}}^{\mathrm{miss}}$ is closely associated to one of the leading jets, either fake from mis-measurement or real from decay of shower particles, thus requiring $\Delta\phi(jet, E_{\mathrm{T}}^{\mathrm{miss}}) > 0.2$ for the leading three jets would suppress the QCD background. However these two variables are not used in our analysis, as we already get enough QCD suppression from the event selection criteria described above.

Table 4.4 summarizes the event counts before and after event selection for an integrated luminosity of 200 pb^{-1} in the full observables range (not in control/signal regions). The $W+jets$ background is the sum of all the ALPGEN produced samples described earlier. The QCD sample is only the sum of all the samples with highest allowed leading jet p_{T}[2]. As can be seen in the table, after our event selection the QCD background can be considered negligible in comparison with the other background samples. Though care must be taken, as most QCD samples are not produced with enough luminosity to affirm this claim. However our requirement of 4 hard jets suppresses the lower p_{T} QCD samples stronger than the harder p_{T} samples, as asymmetric p_{T} distributions are more common than events where two jets have the same p_{T}. Besides that this requirement is strongly biased towards events with high jet multiplicity, thus mainly samples with many associated partons survive the event selection.

The MC@NLO $t\bar{t}$ sample here is split into its dileptonic and semileptonic constituents, on the basis of information of the simulated matrix elements that can be accessed. The dileptonic $t\bar{t}$ remains a sizable component of the SM background, even after the 2nd lepton veto. In these events one of the leptons from the W decay must be not identified as an electron or a muon. These dileptonic $t\bar{t}$ events can be classified into three categories [60]: 50% of these events contain a $W \to \tau\nu$ decay where the tau decays hadronically, in 20% of these events the lepton is misidentified as a jet, and for the rest of the events the lepton is either lost inside a jet or falls outside the p_{T} or η acceptance. The single top samples are not mentioned in the table as

[2]These are event counts only for the samples with the highest allowed leading jet $280 < p_{\mathrm{T}} < \infty$ GeV as they are the only ones that could be produced with enough integrated luminosity for this study. The lower p_{T} samples are produced with an integrated luminosity of 10 pb^{-1}. They contain in total 4641500 (heavy flavour) and 9340000 (light flavour) events, but do not contribute a single event after our selection, hence our choice not to include them in this table.

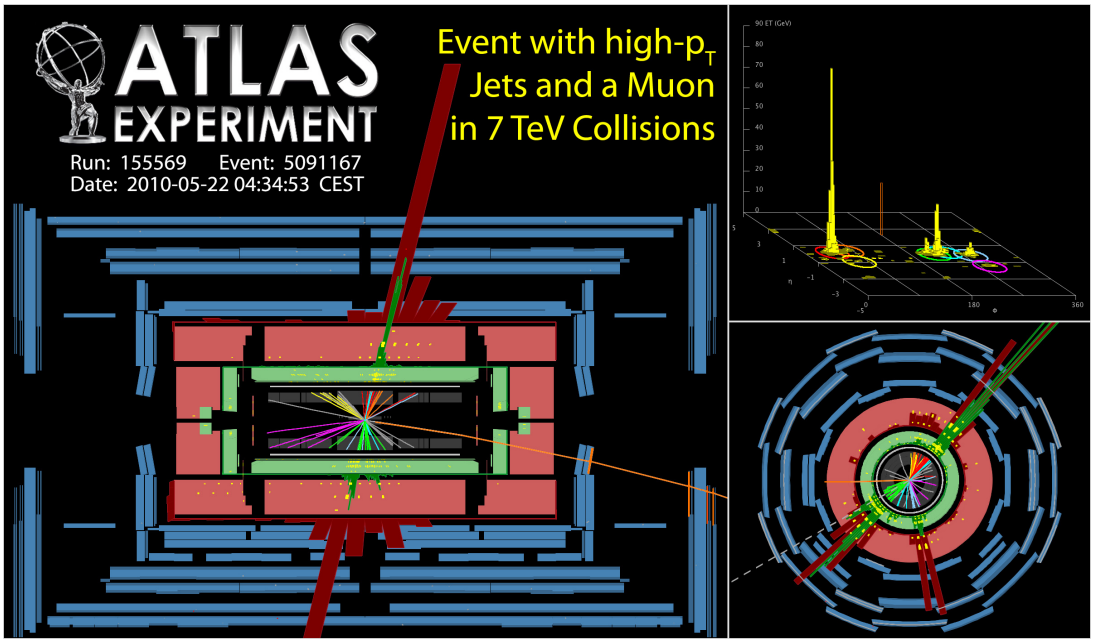


Figure 4.3: Event display of a collision (run number 155569, event number 5091167) with M_{eff} of 915 GeV when only the leading two jets are included in the scalar sum, increasing to 1156 GeV if all jets are included. There are a total of 145 tracks associated with the primary vertex; no second vertex is reconstructed. The missing transverse momentum is 118 GeV. There is one well isolated positively charged muon with p_{T} of 25 GeV and $\eta = 2.33$. That muon is cleanly selected with 11 hits on the monitored drift tubes, 6 on the cathode strip chambers, 5 pixel hits and 8 silicon strip hits.

Sample	N_{events} Before	N_{events} After
$t\bar{t}$ $(l\nu qq)$	32328	471
$t\bar{t}$ $(l\nu l\nu)$	8768	102
$W + jets$	7946800	599
QCD	1481000^2	7
SU4	21480	975
SU3	1092	27

Table 4.4: Event counts in the full observables region before and after selection for an integrated luminosity of 200 pb^{-1} of the backgrounds and two SUSY models, one high mass (SU3) and one low mass (SU4).

no events survive our selection.

For comparison we include two ATLAS SUSY benchmark points in Table 4.4, a high mass SUSY point SU3 and a low mass SUSY point SU4, that are discussed in Section 1.2.7. For an integrated luminosity of 200 pb^{-1}, it would be very demanding to find high mass SUSY such as SU3 in LHC data. On the other hand the SU4 model still retains many events after our selection, the caveat being that the shapes of distributions in our observables for such low mass SUSY are very close to the SM. Figure 4.3 shows a display of an ATLAS recorded 7 TeV collision event with high-p_{T} jets, an isolated muon and missing transverse energy.

Trigger efficiency

The trigger efficiencies for the initial LHC running scenario are assessed in [60, 100]. A trigger on isolated muons with p_{T} of 10 GeV, called `mu10`, was found to be efficient well above 95% once you are past the turn-on curve. In our event selection we require muons with $p_{\mathrm{T}} > 20$ GeV, so we are safely in the `mu10` trigger plateau. For simplicity in the rest of the study we assume a trigger efficiency of 100%.

4.2 Mathematics

4.2.1 Introduction

In particle physics we deal with measurements of events. Each event is a discrete occurrence in time and has one x or more $\vec{x}$ measured observables associated with it. This analysis looks at the distribution of an observable (e.g. $E_{\mathrm{T}}^{\mathrm{miss}}$) for many events, combining the information of several observables into one model to maximize the effectiveness of the background estimate. To model such distributions it is natural to talk about *probability density functions* or *PDFs* in short. We will use PDFs to define a model of signal and backgrounds in Section 4.3, and use the maximum likelihood estimation technique to determine the parameters of this model. This section deals with the mathematics of working with PDFs.

A PDF $F(\vec{x};\vec{p})$ gives the probability density of a distribution of observables $\vec{x}$ where $\vec{p}$ are parameters of the model describing this distribution. A probability density function obeys two rules: it must be unit normalized and it must be positive definite for all possible values of $\vec{x}$ and $\vec{p}$. Normalization implies a domain in the observables phase space, where the PDF is defined.

As an example we take a Gaussian distribution in observable x. The parameters of the model are the mean (m) and standard deviation (σ). The PDF $G(x;m,\sigma)$ following the Gaussian function can be written as:

$$G(x;m,\sigma) = \frac{e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}}{\int_{x_{min}}^{x_{max}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx} \ , \tag{4.10}$$

where $x_{min}(x_{max})$ are the lower (upper) limit of our observable x.

Before we continue, some clarification of used notation: we use small cap letters $f(x)$ to denote an unnormalized PDF, technically a function. Large caps $F(x)$ are reserved for a properly normalized function, also known as a PDF. Where needed we denote the range or domain R over which the PDF is normalized as $F_R(x)$.

There are many advantages to using PDFs. First of all they are a prerequisite for the use of the (unbinned) maximum likelihood estimation technique. For a given set of measurements $\vec{x}$ the likelihood is defined as a product:

$$L(\vec{p}) = \prod_i F(\vec{x}_i;\vec{p}). \tag{4.11}$$

The likelihood function thus gives the probability of finding the set of data points x_i for a given set of parameters. Finding the maximum of the likelihood means finding the set of parameters for which this measurement is most likely. For convenience the negative log of the likelihood is often used:

$$-\ln L(\vec{p}) = \sum_i \ln F(\vec{x}_i;\vec{p}) \ , \tag{4.12}$$

as it is computationally easier to minimize this sum which is equivalent to maximizing the likelihood $L(\vec{p})$.

Another advantage to the use of PDFs is that we can add them together with an intuitive interpretation of fraction coefficients. If - for example - we want to describe a signal distribution on top of a background with a fraction α, we write a sum of PDFs:

$$F(x) = \alpha S(x) + (1-\alpha)B(x) \ , \tag{4.13}$$

where we define coefficient $\alpha[0,1]$, which can be another parameter in our model. Because both the signal and the background PDFs are normalized to one, by construction the sum PDF is also unity normalized. For a shorter notation we will from now on use $F(x)$ to describe a PDF leaving out the parameters wherever this is convenient.

Just as we can write a sum PDF, we can also extend our model to multiple observables or dimensions by simply writing a product of lower or one-dimensional PDFs as:

$$H(x,y) = F(x) \cdot G(y) \tag{4.14}$$

where $F(x)$ and $G(y)$ are factorizing PDFs, or in other words the PDF $F(x)$ ($G(y)$) is independent of y (x). Factorizing PDFs are used in a product PDF, when no correlations are present. A product PDF of normalized PDFs is also a PDF itself by construction.

Conditional PDF If needed, correlations can be introduced with *conditional* PDFs, which are denoted with a vertical line as $F(x|y)$. This means that $F(x|y)$ describes the distribution in x for a given value of y, but the PDF itself has no information on the y distribution.

A conditional PDF $F(x|y)$ is different from a standard 2-dimensional PDF $K(x,y)$ by the normalization:

$$\int F(x|y)dx = 1 \ , \tag{4.15}$$

$$\int\int K(x,y)dxdy = 1 \ . \tag{4.16}$$

The conditional PDF integrates to unity over x for each value of y. This is achieved through the normalization factor, which we denote as $N(y)$, specific to conditional PDFs:

$$F_N(x|y) = \frac{f(x,y)}{\int_N f(x,y)dx} = f(x,y) \cdot N(y) \ , \tag{4.17}$$

where $f(x,y)$ is the unnormalized conditional function. This normalization factor of course depends generally on y and on the range over which the PDF is to be normalized. The latter is a crucial part of the problems we solve in this section, as discussed in Section 4.2.2.

As a conditional PDF $F(x|y)$ describes the shape of the model in only one (x) dimension, external information on the distribution in observable y must be added or assumptions about the shape must be made, if one needs to derive a distribution from the PDF $F(x|y)$, for example for a toy MC study. Computationally it is easier to add a second observable, for which the conditional PDF $F(x|y)$ can be multiplied by a PDF $G(y)$ to form, a *conditional product PDF*:

$$H(x,y) = F(x|y) \cdot G(y) = \frac{f(x,y)g(y)}{\int f(x,y)dx \int g(y)dy} \ . \tag{4.18}$$

This is a proper PDF, as it is unit normalized if we integrate it over x and y. In the case of a conditional product PDF each component is normalized separately, thus the normalization of

$H(x, y)$ is defined as a product of two 1D integrals. This is nice as numerical integration of 2(or higher)-dimensional PDFs is computationally extensive, while for 1D integrals it is relatively fast. Besides it can introduce correlations in your model that can be described in a straight forward fashion.

All the descriptions of the models, and all fitting in this chapter was done using the `RooFit` framework [103], which is part of the `ROOT` framework [97] commonly used in high energy physics.

Problems of extrapolating from the control region to the signal region

In the combined fit analysis we use all the aspects of working with PDFs that were mentioned above, as will be detailed in Section 4.3. We set up the analysis such that we fit our model in a control region, extrapolating the result to the signal region. A consequence of our model containing conditional product PDFs, is that the shape of the PDF is dictated by the range over which it is normalized, as we will show in the next section. As we want to have an unambiguously defined shape of the extrapolated PDF independent of the specific normalization region, this is undesired behavior.

Another problem arises from the entanglement of PDF shape and normalization range when we want to use a composite normalization range, such as our L-shaped control region. From the computational point of view, only square regions are easy to integrate over, hence we must build our L-shaped control region from two squares, shown as A and B in Figure 4.7. As the shape is different when the PDF is normalized in A from the shape of the PDF normalized in B, there is a discontinuity at the border of the two control ranges, if not explicitly handled. As we want to effectively describe our data in the complete composite control region by a single continuous PDF, the discontinuity poses another problem.

4.2.2 The conditional product PDF in a single subrange

We illustrate the problem of shape and normalization range entanglement for conditional product PDFs using an example shown in Figure 4.4(a). The model shown in the figure is a Gaussian PDF in x multiplied by a uniform PDF in y, where the corresponding function product can be written as:

$$f(x, y) \cdot g(y) = e^{-\frac{1}{2}\left(\frac{x-\mu \cdot y}{\sigma}\right)^2} \cdot C_y \ .$$

Here the mean (parameter m) of the Gaussian in Equation 4.10 is replaced by $\mu \cdot y$, while the value of μ is set at 1. Thus the mean of the Gaussian moves linearly as a function of y, which gives the sliding peak visible in Figure 4.4(a). The uniform PDF in y is denoted as the constant C_y.

The normalization factor, as defined in Equation 4.17, of the conditional component $F(x|y)$ of this product PDF is shown as a function of y in Figure 4.5 by the solid curve. As $F(x|y)$ has to be unit normalized for every value of y, the normalization factor increases slightly at low and high y values, because there the Gaussian is only partly contained inside the x range, the normalization factor deviates from the value of the full integral over a Gaussian.

If we take the same conditional product PDF, but normalize it now on the subrange SR = x $\in [-10, 0]$, the shape is strikingly different. This can be seen in Figure 4.4(b), where the position of the saddle point has moved. For illustrative comparison we still plot the PDF in the full range FR = x $\in [-10, 10]$, although technically it is only valid as a PDF in the normalization range SR. This is visualized by obscuring the right half of Figure 4.4(b). We conclude that the formulation of Equation 4.18 is not a robust way to define a 2D shape, as the shape is dependent

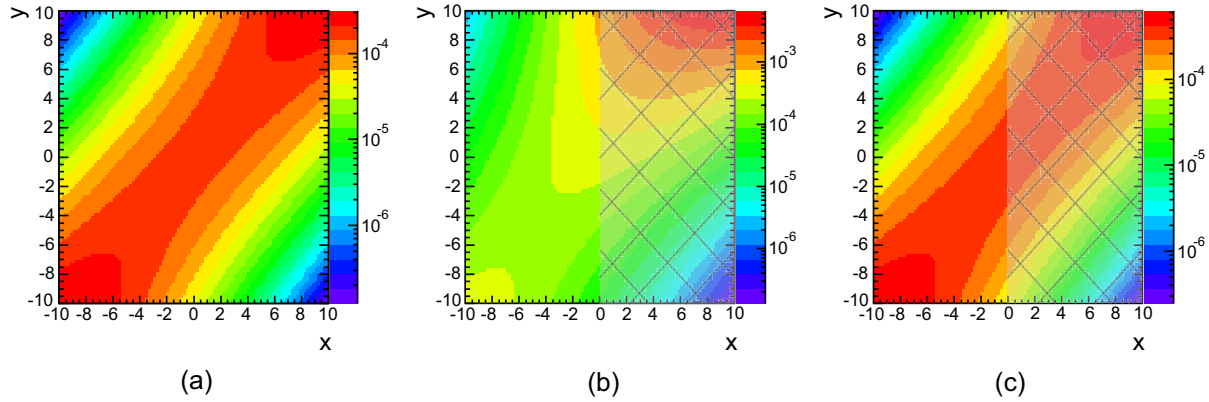


Figure 4.4: In (a) is the original model of Equation 4.2.2 defined in the full range FR = x $\in [-10, 10]$. In (b) is the same model, but now normalized to a subrange SR = x $\in [-10, 0]$. In (c) is again the same model normalized in the subrange SR, but with the shape definition range set to the full range FR. In (b) and (c), the right half is obscured as the PDF is technically not valid outside its normalization range.

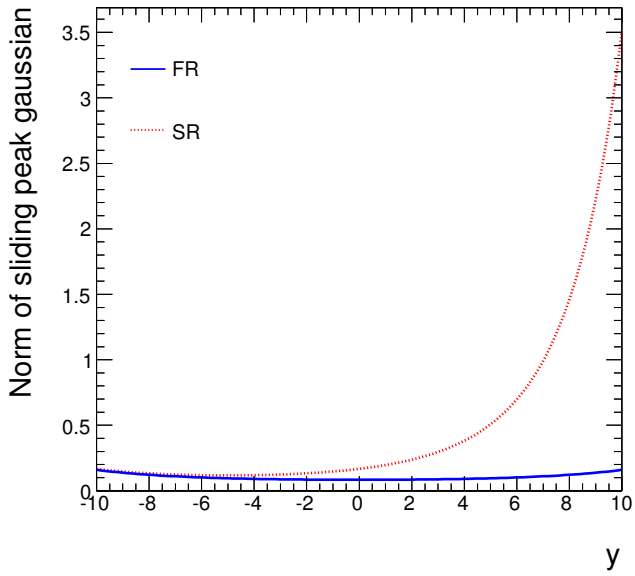


Figure 4.5: The normalization factor for the model normalized on SR (dotted curve) compared to the normalization factor of the model normalized on FR (solid curve), as a function of y.

on the choice of normalization range.

The difference between Figure 4.4(a) and Figure 4.4(b) is caused by the y-dependence of the normalization factor of $F(x|y)$, which is shown for both cases in Figure 4.5, solid curve for FR normalization and dotted curve for SR normalization. The difference is clear: for higher values of y, the normalization factor for the subrange SR becomes very large. The reason is that at higher y the peak of the Gaussian model in x lies outside the normalization range. Thus to keep the normalization to unity for all values of y as required for a conditional PDF, the normalization factor must strongly increase as a function of y.

As we will need to extrapolate our model from one range to another, this causes problems because the extrapolated shape is dependent on the choice of the control range. We will fix them

through the introduction of the concept of a *shape definition range* of the PDF, which can be chosen different from the normalization range. In the next section we work out the mathematics, but in Figure 4.4(c) we show the result of shape definition range implementation. There the normalization range is set to SR, but the shape definition range of the model is set to FR.

Comparison of Figures 4.4(a) and 4.4(c) shows that the overall normalization is different, but the shape of the two distributions is clearly the same. The overall normalization is different by a factor 2 as expected, as the PDF is symmetric and thus the integral over SR (Figure 4.4(c)) is half of the integral over FR (Figure 4.4(a)).

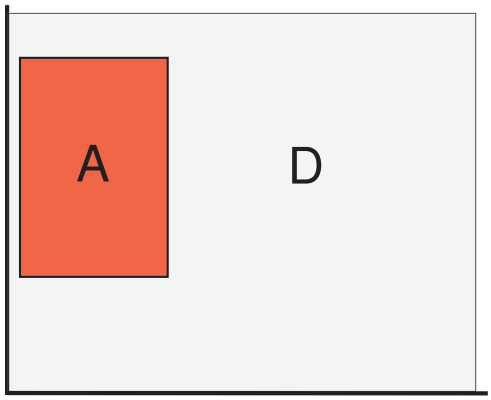


Figure 4.6: A subrange A inside the full range D.

4.2.3 Conditional product PDF with shape definition range

In this section we show how the concept of a shape definition range ds and a separate normalization range N can be implemented. For a conditional product PDF $H(x,y) = F(x|y) \cdot G(y)$ to retain its shape under the change of the normalization range, it must be rewritten as:

$$H_N^{\mathrm{ds}}(x,y) = \frac{F_{\mathrm{ds}}(x|y)g(y)}{\iint_N F_{\mathrm{ds}}(x|y)g(y)dxdy} \ , \tag{4.19}$$

where $\mathrm{F_{ds}(x|y)}$ indicates PDF $F(x|y)$ normalized in shape definition range ds.

A useful variable for our explanation is the ratio R, which is defined as the fraction of a PDF $\mathrm{H(x,y)_X}$ ($H(x,y)$ normalized in certain range X) that is contained in the subrange A compared to the full range D. A schematic example of the subrange A and the full range D is shown in Figure 4.6.

Let us define two different normalization ranges[3] for $H(x,y)$, D and A, that give respectively ratios R_1 and R_2 defined as:

$$R_1 = \frac{\iint_A H_D(x,y)dxdy}{\iint_D H_D(x,y)dxdy} = \frac{\int_A \left(\int_A F_D(x|y)dx\right) G_D(y)dy}{\int_D \left(\int_D F_D(x|y)dx\right) G_D(y)dy} \tag{4.20}$$

$$R_2 = \frac{\iint_A H_A(x,y)dxdy}{\iint_D H_A(x,y)dxdy} = \frac{\int_A \left(\int_A F_A(x|y)dx\right) G_A(y)dy}{\int_D \left(\int_D F_A(x|y)dx\right) G_A(y)dy} \tag{4.21}$$

[3]The most general case would be if for the two components of the product PDF, the two integration ranges and the two normalization ranges are all different. This would however mean longer, more dense formulas without changing the result.

If the shape of the PDF is invariant under a change of normalization range, then the ratios R_1 and R_2 are identical. We have seen in the previous section that this is not the case, so let us focus first on R_1. Since the integral over D of the PDF normalized to D equals 1 by construction (in the denominator of Equation 4.20), expanding PDFs F and G with functions f and g, we get:

$$R_1 = \iint\limits_A \left(\frac{f(x,y)g(y)}{\int_D f(x,y)dx \int_D g(y)dy} \right) dxdy \tag{4.22}$$

Since we define our model on the full range in our analysis, we define this as the 'reference' ratio, and D as our shape definition range. Now we calculate the ratio R_2, with the PDFs normalized[4] to A instead of D:

$$R_2 = \frac{\int \left(\int_A F_A(x|y)dx \right) G_A(y)dy}{\int \left(\int_D F_A(x|y)dx \right) G_A(y)dy} = \frac{1}{\iint_D \left(\frac{f(x,y)g(y)}{\int_A f(x,y)dx \int_A g(y)dy} \right) dxdy} \ . \tag{4.23}$$

Here we see the issue: if the function f had no y-dependence, then $\int f dx \int g dy = \iint (fg) dxdy$ would hold, and the ratios R_1 and R_2 would be the same, as the last equation could be rewritten as:

$$R_2^\dagger = \frac{1}{\left(\frac{\iint_D f^\dagger(x)g(y)dxdy}{\int_A f^\dagger(x)dx \int_A g(y)dy} \right)} = \frac{\iint_A f^\dagger(x)g(y)dxdy}{\iint_D f^\dagger(x)g(y)dxdy} = R_1 \ , \tag{4.24}$$

where the dagger in $R_2^\dagger$ indicates that for this R a function $f^\dagger(x)$ is assumed that does not depend on y. However as our original function f is dependent on y, as shown in Equation 4.2.2, the ratios R_1 and R_2 are different.

We solve this by replacing $g(y)$ in $G_A(y) = g(y)/\int_A g(y)dy$ with $g'(y)$, that is dependent on shape definition range D and normalization range A as:

$$g(y) \to g'(y) = g(y) \cdot \frac{\int_A f(x,y)dx}{\int_D f(x,y)dx} = g(y) \int_A F_D(x|y)dx \tag{4.25}$$

Applying this substitution to the PDF $G_A(y)$ gives:

$$\begin{aligned} G_A(y) &= \frac{g(y)}{\int_A g(y)dy} \to \\ G'_A(y) &= \frac{g(y) \int_A F_D(x|y)dx}{\iint_A g(y)\ F_D(x|y)dxdy} = G^*_A(y) \int_A F_D(x|y)dx \ , \end{aligned} \tag{4.26}$$

where we define a new function $G^*_A(y)$ as:

$$G^*_A(y) = \frac{g(y)}{\iint_A g(y) F_D(x|y)dxdy} \ . \tag{4.27}$$

When we apply the substitution of Equation 4.25 to R_1, where the normalization range is D and the shape definition range is D as well, this is multiplication by 1. Thus the transformed R'_1 stays the same as R_1 in Equation 4.22.

[4]Note that D now extends outside the normalization range and $F(x,y)$ is no longer a PDF, since it is no longer unit normalized.Technically, $F(x,y)$ is a function here. The goal of this exercise is to be able to carefully construct a PDF such that ratio R can be calculated.

If we now calculate R_2' where $g'(y)$ is substituting $g(y)$ in R_2, the numerator of ratio R_2' becomes:

$$\int_A \left(\int_A F_A(x|y)dx \right) G_A(y)dy \rightarrow \iint_A F_D(x|y)G_A^*(y)dxdy \ , \tag{4.28}$$

as $\int_A F_A(x|y)dx = 1$ by construction. The denominator of R_2' gets an extra term, so we are left with:

$$\begin{aligned} &\int_D \left(\int_D F_A(x|y)dx \right) G_A(y)dy \rightarrow \\ &\int_D \left(\int_D F_A(x|y)dx \int_A F_D(x|y)dx \right) G_A^*(y)dy = \int_D G_A^*(y)dy \ . \end{aligned} \tag{4.29}$$

Dividing the R_2' numerator (Equation 4.28) by the R_2' denominator (Equation 4.29) gives after some rearranging of terms:

$$R_2' = \iint_A \frac{f(x,y)g(y)}{\int_D f(x,y)dx \int_D g(y)dy} dxdy \tag{4.30}$$

which is the same as the ratio R_1', given by Equation 4.22.

This means that we can ensure that the shape of the total PDF is invariant under the change of the normalization range, if we implement a shape definition range. This can easily be extended to N dimensions.

Implications for conditional product PDFs

Now that we have constructed a formalism to define the shape of a conditional product PDF independent of its normalization range, we redefine the conditional product PDF $H(x,y)$ by inserting Equation 4.25 in Equation 4.18, and renaming full range D to shape definition range ds, which gives:

$$\begin{aligned} H_A^{\mathrm{ds}}(x,y) &= \frac{f(x,y)g(y)\frac{\int_A f(x,y)dx}{\int_{\mathrm{ds}} f(x,y)dx}}{\int_A f(x,y)dx \int_A \frac{\int_A f(x,y)dx}{\int_{\mathrm{ds}} f(x,y)dx} g(y)dy} \\ &= \frac{f(x,y)g(y)}{\int_{\mathrm{ds}} f(x,y)dx \int_A (\int_A F_{\mathrm{ds}}(x|y)dx)g(y)dy} \ . \end{aligned} \tag{4.31}$$

If we rewrite this in terms of the conditional PDF $F(x|y)$ normalized on the shape definition range ds:

$$H_A^{\mathrm{ds}}(x,y) = \frac{F_{\mathrm{ds}}(x|y)g(y)}{\iint_A F_{\mathrm{ds}}(x|y)g(y)dxdy} \ . \tag{4.32}$$

This equation resembles a PDF of a 2-dimensional object $F_{\mathrm{ds}}(x|y)g(y)$, where the normalization of this object is defined by an integral in two dimensions. Hence our conclusion is that to keep the shape of a conditional product PDF invariant under the change of normalization range, the product PDF must be redefined as a single 2-dimensional object with a shape definition range ds for the conditional PDF.

Computationally this means that on a subrange of observable x, the denominator of Equation 4.32 can only be calculated numerically, as the conditional PDF $F(x|y)$ is not analytically calculable over a subrange.

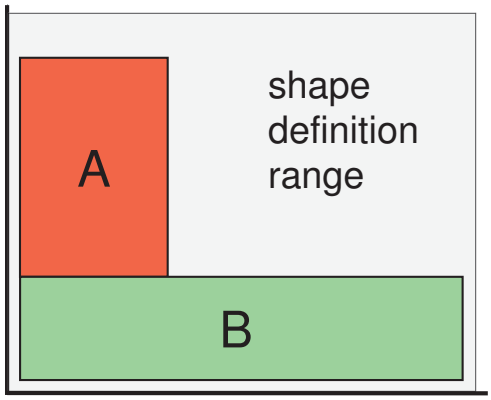


Figure 4.7: Composite range $A+B$ within the shape definition range.

4.2.4 A composite normalization range

An L-shaped control region, as is used in the combined fit method, is computationally easiest to describe as a composite range $A+B$ shown schematically in Figure 4.7. For a simple factorizable product PDF, the normalization condition for a composite range $A + B$ is:

$$\begin{aligned}\iint_{A+B} F_{A+B}(x)G_{A+B}(y)dxdy &= \frac{\int\int_{A+B} f(x)g(y)dxdy}{\int_{A+B} f(x)dx \int_{A+B} g(y)dy} \\ &= \frac{\int_{A+B} f(x)dx \int_{A+B} g(y)dy}{\int_{A+B} f(x)dx \int_{A+B} g(y)dy} = 1 \end{aligned} \tag{4.33}$$

which holds because the integrals over x and y can be separated.

This integral factorization however does *not* work for conditional product PDFs, as normalization of $F(x|y)$ and $G(y)$ are tied together, as shown in the previous section by comparison of Equations 4.23 and 4.24. Because of the entanglement of PDF shape and normalization, the shape of $H(x, y)$ in normalization range A is different from the shape in normalization range B, which is problematic if we describe $H(x, y)$ by a single continuous PDF.

The solution to this is to normalize the conditional product PDF $H(x, y)$ as a single 2-dimensional object with the prescription of the previous section, instead of normalizing each component separately. So instead of writing:

$$H_{A+B}(x, y) = F_{A+B}(x|y)G_{A+B}(y) \tag{4.34}$$

the PDF must be written as in Equation 4.32 with a shape definition range ds. Normalized over a composite range it becomes:

$$H^{\mathrm{ds}}_{A+B}(x, y) = \frac{F_{\mathrm{ds}}(x|y)g(y)}{\iint_A F_{\mathrm{ds}}(x|y)g(y)dxdy + \iint_B F_{\mathrm{ds}}(x|y)g(y)dxdy} \tag{4.35}$$

This normalization equation can be extended to an arbitrary sum of ranges (Σ_R) that make up the composite normalization range:

$$H^{\mathrm{ds}}_{\Sigma R}(x, y) = \frac{F_{\mathrm{ds}}(x|y)g(y)}{\Sigma_R \iint_R F_{\mathrm{ds}}(x|y)g(y)dy} \tag{4.36}$$

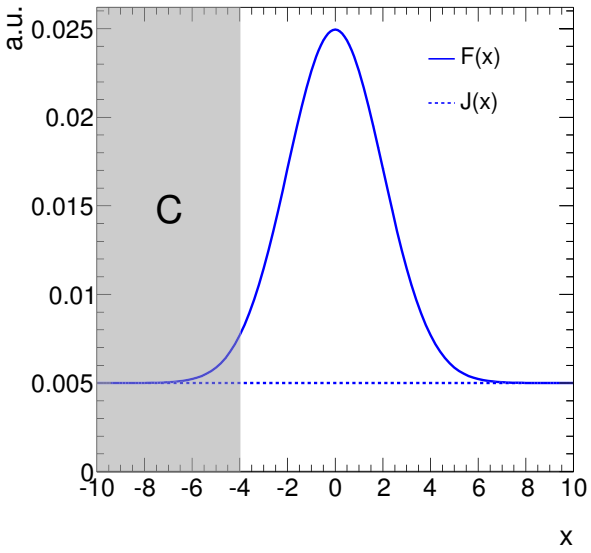


Figure 4.8: Example of a sum PDF $F(x)$ and the uniform component PDF $J(x)$ as described in the text, while the coefficient range C is shown by the gray rectangle.

4.2.5 Normalization and the coefficient range

Having shown in the last section that we can normalize conditional product PDFs on a composite range, the question we want to address in this section is how to *add* PDFs in a composite range. The focal point of this discussion is how do we interpret the coefficient α of a sum PDF on a subrange, which shows another case of PDF shape entanglement with normalization range.

We define a PDF F in x, not conditional for now, which is a sum of two PDFs:

$$F_N(x) = \alpha_N J_N(x) + (1 - \alpha_N) I_N(x) \ , \tag{4.37}$$

noting that such a sum of PDFs is only a PDF itself, if the components J and I are normalized on the same region N in phase space. Let us take an example, where $J(x)$ is a uniform distribution and $I(x)$ is a Gaussian, and we set the value of α_N to be 0.5. The sum PDF $F(x)$ and the uniform component $J(x)$ are shown in Figure 4.8 for the full normalization range $N \in x[-10, 10]$.

Now we want to evaluate the PDF on a subrange of N, that we call range $C \in x[-10, -4]$, which is shown in Figure 4.8 as the gray rectangle. We can see from Figure 4.8 that the fraction of the uniform PDF $J(x)$ in range C (α_C) is much greater than the fraction in range N (α_N). If the value of coefficient α_C is kept the same as α_N, we can intuitively predict that the shape of the PDF would be completely different, thus the shape of the sum PDF $F(x)$ is entangled with the normalization range.

We solve this by introducing the concept of a new range, called the *coefficient range*, which defines the range on which the coefficient α of a sum PDF is evaluated. If we take subrange C to be the coefficient range, the coefficient α_C is different from α_N by a factor that depends on the shape of the PDF as:

$$\alpha_C = \frac{\int_C \alpha_N J_N(x) dx}{\int_C F_N(x) dx} \ . \tag{4.38}$$

The value of α_C in this example is $\sim$0.93, clearly greater than α_N=0.5. The ranges N and C can be made composite in a trivial manner, as well as the component PDFs can be made conditional.

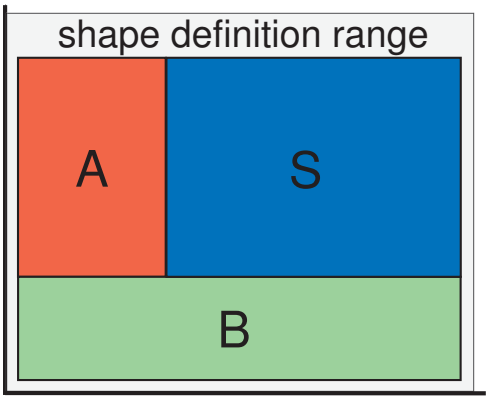


Figure 4.9: Composite control range $A+B$ and the signal range S within the shape definition range.

In summary when we are normalizing a sum PDF on a (possibly composite) range N, but we want to evaluate it on a different range, known as the coefficient range C, the coefficient α must be transformed as in Equation 4.38 for (each subrange of) N.

4.2.6 Product of sum PDFs with conditional terms

In this section we put all the prescriptions of previous sections together, to show what a 2-dimensional model with correlations describing two separate processes should look like. Starting from the sum PDF given by Equation 4.37, we can add dependence on a second observable as follows:

$$K_N(x,y) = F_N(x) \cdot G_N(y) = (\alpha_N J_N(x) + (1-\alpha_N) I_N(x)) G_N(y) \ . \tag{4.39}$$

Writing out the normalization explicitly gives:

$$K_N(x,y) = \left(\alpha_N \frac{j(x)}{\int_N j(x)dx} + (1-\alpha_N) \frac{i(x)}{\int_N i(x)dx} \right) \frac{g(y)}{\int_N g(y)dx} \ . \tag{4.40}$$

If any of the three PDFs in $K(x,y)$ is a conditional PDF, we must use the prescription of Section 4.2.3 to keep shape invariance. If in the example above, we take $J(x|y)$ as a conditional PDF, then the entire expression F becomes a conditional PDF as well. Hence we must change the normalization of $K(x,y)$ according to the recipe of Equation 4.36:

$$K_N^{\mathrm{ds}}(x,y) \quad = \left(\alpha_{\mathrm{ds}} \frac{j(x,y)}{\int_{\mathrm{ds}} j(x,y)dx} + (1-\alpha_{\mathrm{ds}}) \frac{i(x)}{\int_{\mathrm{ds}} i(x)dx} \right) \frac{g(y)}{\iint_N F_{\mathrm{ds}}(x|y) g(y) dxdy} \ . \tag{4.41}$$

Note that everywhere the sum PDF remains normalized consistently, but the normalization range and the coefficient range of $F(x|y)$ have been changed to the shape definition range ds. It is straightforward to extend the normalization of $K(x,y)$ to a composite range, by replacing N in the numerator with a sum over ranges ΣR.

If $K(x,y)$ is a 2D model of one process, and we introduce a second process in our model that is also described by a conditional product PDF $M(x,y)$, then the total model $P(x,y)$ is given by:

$$P_N(x,y) = \beta_N K_N(x,y) + (1-\beta_N) M_N(x,y) \ . \tag{4.42}$$

We are interested in the fraction of events described by $K(x,y)$ in the signal region S, but we fit the total $P(x,y)$ model to data in a composite (L-shaped) control region $A+B$, schematically

shown in Figure 4.9. This means we have to set three different ranges for our model. The shape definition range *ds* is set at the full phase space of x and y, so the PDFs are continuous and unambiguously defined throughout all the ranges. The normalization range is set at the composite control range $A + B$. Lastly the coefficient range of β is set at the signal region S, which means we can perform the fit in $A + B$ and the extrapolation to S in one iteration.

The *coefficient range*, the *normalization range* and the *shape definition range* have all been implemented in the `RooFit` framework, the latter specifically for this analysis. Now we are ready to define our background and signal models, which are 3-dimensional products of sums of conditional and non-conditional PDFs.

4.3 Fit method models

4.3.1 Introduction

This section is concerned with modeling physics processes in our three observables: $E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T} and M_{jjj}, and we start with the models for the SM backgrounds. The semileptonic $t\bar{t}$, dileptonic $t\bar{t}$ and $W + jets$ processes are the dominant physical backgrounds to SUSY searches with one lepton. We build a model for each background process separately, that we describe by a 3D PDF. But we reorganize the models to improve fit stability, because we are not primarily interested in determining the relative yields of $t\bar{t}$ and W, but in determining the complete background contribution in the signal region.

The reorganization concerns the semileptonic $t\bar{t}$ and the $W + jets$ samples, as these events have very similar (component) shapes in our observables, that are difficult to separate in a combined fit. The M_{jjj} distribution of the semileptonic $t\bar{t}$ sample (Figure 4.21(a)) shows a pronounced Gaussian peak, where the three jets with the highest Σp_{T} form the mass of the top. These events we call *top peak* (TP) events. Besides the peak there is an exponentially decaying component where the three jets that were selected by the M_{jjj} algorithm do not correctly identify the original top quark, which we call the *top combinatorics* (TC) events. We split the semileptonic top sample into a TP and a TC sample, as detailed in Section 4.3.5, and add the latter to the $W + jets$ sample. Therefore we (re)label our backgrounds:

- **T2**: Dileptonic $t\bar{t}$
- **TP**: Correctly reconstructed semileptonic $t\bar{t}$, top peak
- **CW**: $W + jets$ merged with top combinatorics

We construct the complete model for all the SM backgrounds by adding the 3D PDFs for all processes together. The total combined model is thus expressed as:

$$\mathrm{PDF}_{SM}^{\mathrm{3D}} = N_{obs} \left\{ (1 - f_{TT}) \mathrm{PDF}_{CW}^{\mathrm{3D}} + f_{TT} \left[(1 - f_{T2}) \mathrm{PDF}_{TP}^{\mathrm{3D}} + f_{T2} \mathrm{PDF}_{T2}^{\mathrm{3D}} \right] \right\} \tag{4.43}$$

Here f_{T2} denotes the fraction of T2 events relative to the number of TP+T2 events, and f_{TT} denotes the fraction of TP+T2 events relative to the total number of background events. These fractions are parameters of the model. By defining them recursively, as done here, we make sure that the combined PDF is well defined for all fraction values between 0 and 1. Defining the fractions non-recursively can lead to the remainder of the fractions becoming negative, which poses a problem when interpreting it in terms of physical processes.

The complete 3D model is quite elaborate, so we describe it in steps. First, in the next section, we discuss the analytical 1D models that are used in each observable for all the backgrounds.

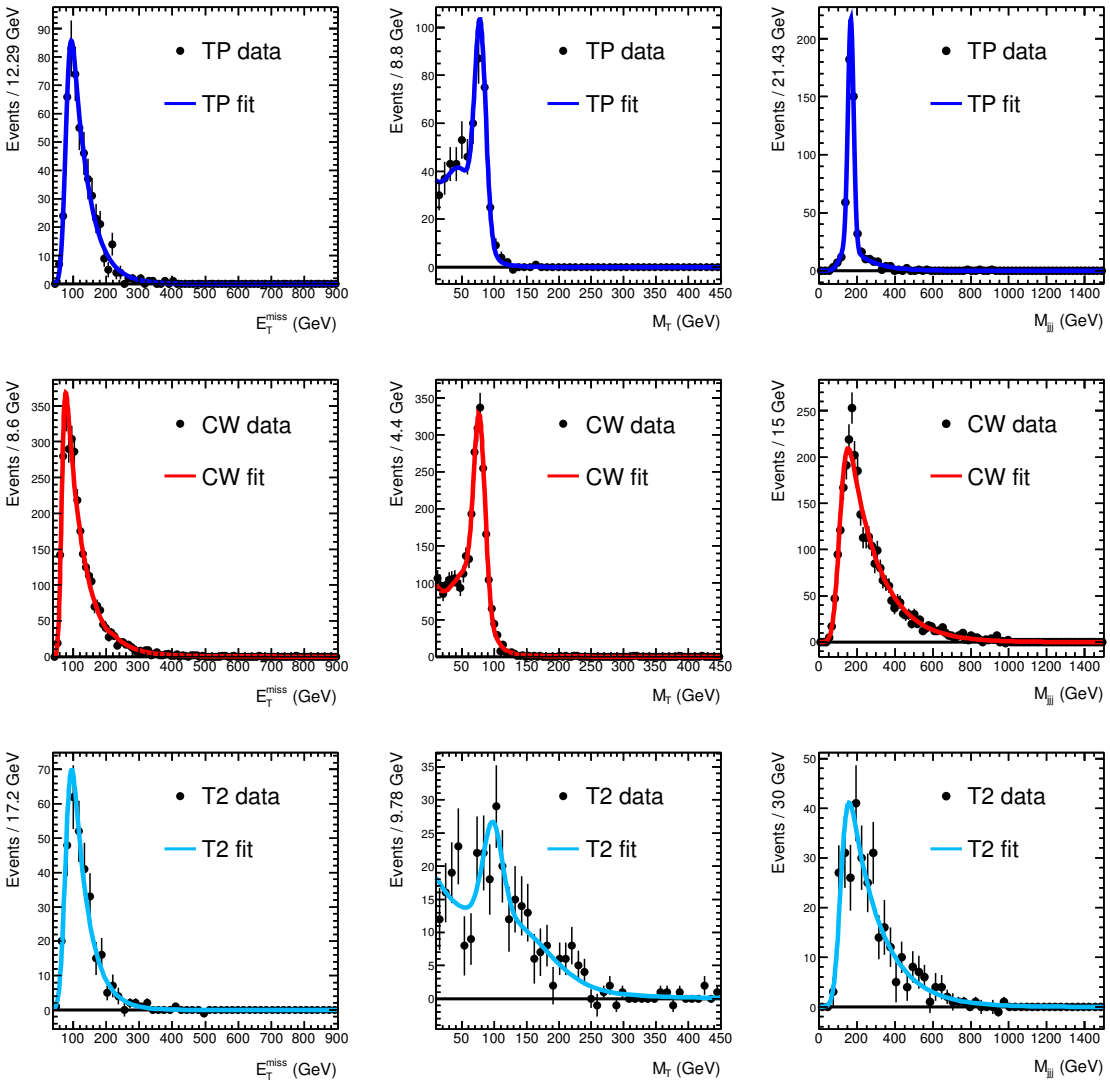


Figure 4.10: 1D projections of all the background models in the three observables. The left column showing the $E_{\mathrm{T}}^{\mathrm{miss}}$, the middle the M_{T} and the right column the M_{jjj} distributions.

Then, in Section 4.3.3, we examine the existing correlations in the simulated samples. We show how by adjusting the 3D models with the use of conditional PDFs, these correlations can be described. Afterwards, in Section 4.3.4, we discuss the physics behind the strong M_{T} evolution of the T2 sample. Next, in Section 4.3.5, we show the technique used for the separation of TP/TC events, and demonstrate that it does not introduce a method dependency in our model. The last section before going to the results, Section 4.3.6, describes the generic model of SUSY, that we use to describe the contamination in the control region.

4.3.2 Analytical description of background models

The 3D models of each background are a product of 1D models in each observable. Figure 4.10 shows the 1D projections of simulated data samples for all three backgrounds, that are described

Background process	Observable		
	$E_{\mathrm{T}}^{\mathrm{miss}}$	M_{T}	M_{jjj}
TP	TTComb	MTFunc	Gaussian (+TTComb)
CW	$\mathrm{TTComb}_1 + \mathrm{TTComb}_2$	MTFunc	TTComb
T2	TTComb	MTFunc	TTComb

Table 4.5: The analytical description used in each observable of the three backgrounds.

Description	Parameter 1	Parameter 2
Mean of the $E_{\mathrm{T}}^{\mathrm{miss}}$ TTComb	$E_{\mathrm{T}}^{\mathrm{miss}}$::mean(TP)	$E_{\mathrm{T}}^{\mathrm{miss}}$::mean (T2)
Width of both CW $E_{\mathrm{T}}^{\mathrm{miss}}$ TTCombs	$E_{\mathrm{T}}^{\mathrm{miss}}$::$\sigma_1$(CW)	$E_{\mathrm{T}}^{\mathrm{miss}}$::$\sigma_2$(CW)
Fraction of core to wide Gaussians in M_{T}	M_{T}::$\mathrm{f}_{\mathrm{core}}$(TP)	M_{T}::$\mathrm{f}_{\mathrm{core}}$(CW)
Width of the core M_{T} Gaussian	M_{T}::σ(TP)	M_{T}::σ(CW)
Ratio of Gaussian widths in M_{T}	M_{T}::r_σ(TP)	M_{T}::r_σ(CW)
Width of the M_{jjj} TTComb	M_{jjj}::σ(CW)	M_{jjj}::σ(T2)

Table 4.6: Sets of parameters that are combined into one global parameter.

by the 1D models shown by the solid curves. The analytical descriptions of each 1D model are detailed in Table 4.5, subdivided by the observable and the sample in the same pattern as Figure 4.10 for easier comparison.

A *TTComb* shape, detailed in the next section, is used for the TP and T2 $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions. For the CW $E_{\mathrm{T}}^{\mathrm{miss}}$ model a sum of two TTCombs is used to account for a small kink in the tail of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, caused by the difference between TC and $W + jets$ events.

All the backgrounds have an *MTFunc* shape, detailed in next section, describing the M_{T} distributions. The M_{jjj} distributions for T2 and CW are described by a TTComb. Finally, the M_{jjj} distribution of the TP sample shows the characteristic Gaussian top mass peak. The definition of the TP model also contains a small TTComb component, to account for the fact that the splitting of TP and TC in simulation samples is not perfectly achievable. However we assume that in the fit of the combined background model to data, the small amount of fake-TP gets absorbed in the CW fraction. For the combined fit the TTComb fraction in the TP M_{jjj} model is fixed to zero.

In total this gives 13 parameters for the $E_{\mathrm{T}}^{\mathrm{miss}}$ models, 21 parameters for the PDFs modeling the M_{T} distributions, and 8 parameters for the M_{jjj} models (not taking the TTComb for fake-TP into account). Thus the background model has 42 shape parameters, plus 2 recursive fractions when the three models are combined.

A way to reduce the total number of parameters is by defining a *global* parameter that describes the same behavior in multiple background models, or even within one background model. Two separate parameters, that have their value within the error margins from each other and have the same physical interpretation, can be replaced by a single global parameter. Parameters that have been replaced by a global parameter are shown in Table 4.6. This reduces the number of shape parameters to 36.

There is a subtlety in handling the CW (double TTComb) model in a fit that has to be addressed: having two TTCombs with the same functional form in one model introduces an am-

biguity for the fitting algorithm, as it is not clear which TTComb describes the distribution after the kink. We avoid this ambiguity by replacing e.g. the parameter mean_2 with $\text{r}_{\text{mean}} \times \text{mean}_1$, and forcing the ratio r_{mean} to be larger than one.

The *TTComb* and *MTFunc* functions

Before we continue, we define two functions which are used widely in our models. The first is called *TTComb*, since it was first used to describe the combinatorics background of the $t\bar{t}$ sample. It is defined as:

$$TTComb\ (x; mean, \sigma, \alpha) = (1 + erf\ (x; mean, \sigma)) \times e^{-\alpha x}. \tag{4.44}$$

The shape of this PDF can be seen in Figure 4.11, where also the effect of changing the parameters is shown. The TTComb is an exponential at high values of x, while at low values the error function $erf(x)$ gives a smooth cutoff. Width and position of the cutoff are controlled by the σ and *mean* parameters, while the exponential decay is determined by the α parameter.

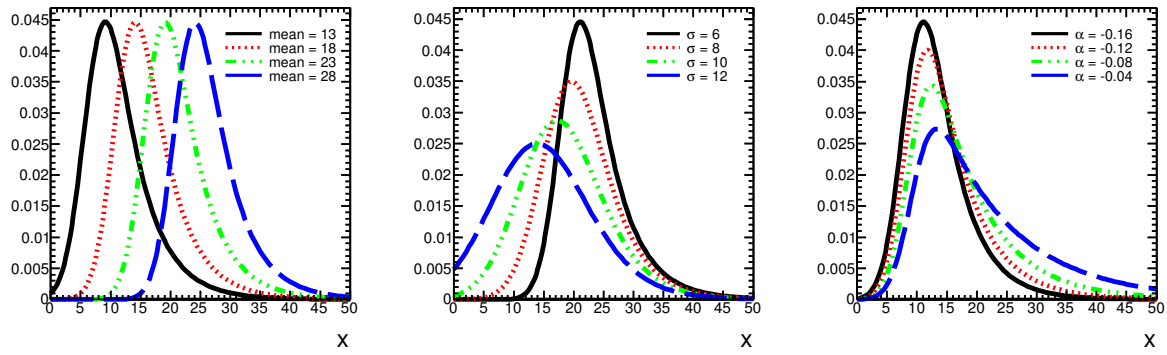


Figure 4.11: Figures showing the shape of the TTComb function for different parameter values. The values of *mean*, σ and α increase from solid to dashed curves.

The second new function that we define is called *MTFunc*. It is used to describe the M_{T} models of all SM backgrounds, and it is created to fit the three main features of these distributions: a Gaussian core, a large exponential tail at high M_{T}, and a plateau at low M_{T}. This function is a sum of two Gaussians and an exponential written as:

$$MTFunc\ (x; mean_{core}, \sigma_{core}, mean_{wide}, \sigma_{wide}, \alpha, f_{peak}, f_{core}) = \tag{4.45}$$
$$(1 - f_{peak})e^{-\alpha x} + f_{peak}\left[f_{core}G(x; mean_{core}, \sigma_{core}) + (1 - f_{core})G(x; mean_{wide}, \sigma_{wide})\right]\ ,$$

with the fractions such that the normalization of the function can be maintained to unity and the interpretation of the components is physical. The shape of this distribution and its components can be seen in Figure 4.12.

The same technique, as was described for the double TTComb model of the CW sample in $E_{\text{T}}^{\text{miss}}$, is used for the two Gaussians in MTFunc to avoid ambiguities. Here both the *mean* and the σ of the wide Gaussian are replaced by ratio terms (r_{mean} and r_{σ}) multiplied with respective core Gaussian parameters.

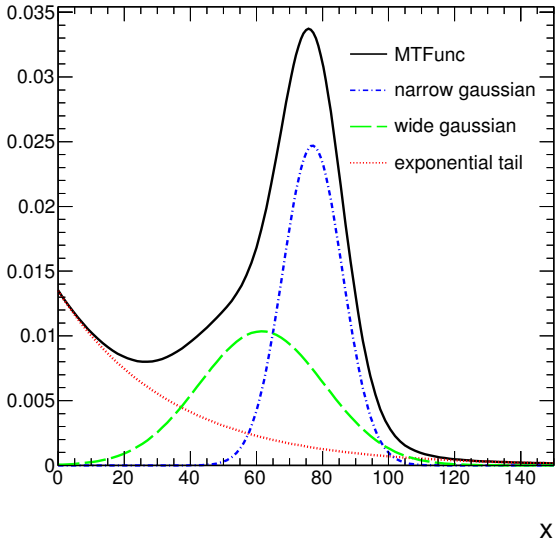


Figure 4.12: The MTFunc distribution and its components.

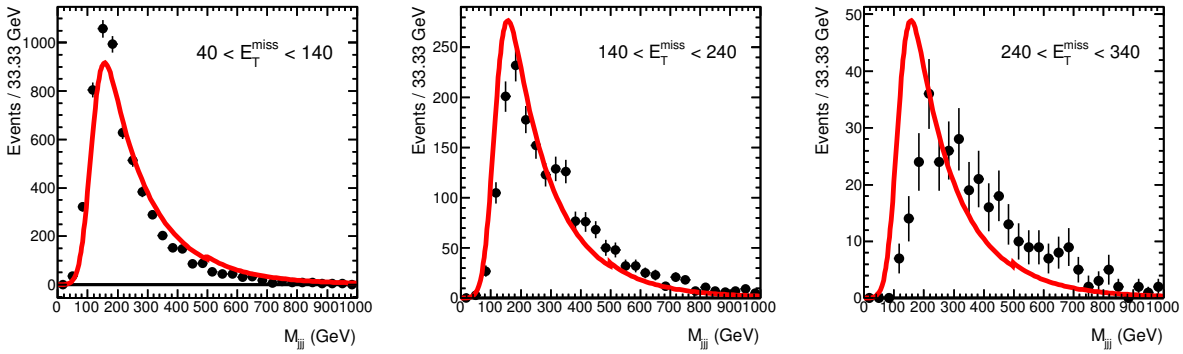


Figure 4.13: The M_{jjj} distribution of the CW sample in three slices of $E_{\mathrm{T}}^{\mathrm{miss}}$, while the uncorrelated model is projected onto every slice as the curve.

4.3.3 Examining correlations in simulation samples

We have shown that we can model the 1D projections in the three observables, but they may hide correlations, which we examine in this section. If the PDF F(x) is uncorrelated with observable y, then the shape of F(x) should be the same for every slice in y.

We focus first on the shape of CW in M_{jjj}. Figure 4.13 shows the M_{jjj} distribution of the CW sample in slices of $E_{\mathrm{T}}^{\mathrm{miss}}$. To guide the eye, the 1D projection of the CW model without correlations fitted to the unbinned data is plotted in each slice. Figure 4.13 shows clearly that the shape of the M_{jjj} distribution depends on $E_{\mathrm{T}}^{\mathrm{miss}}$, and that correlations cannot be ignored. The changes to the models that are required to handle this are described next.

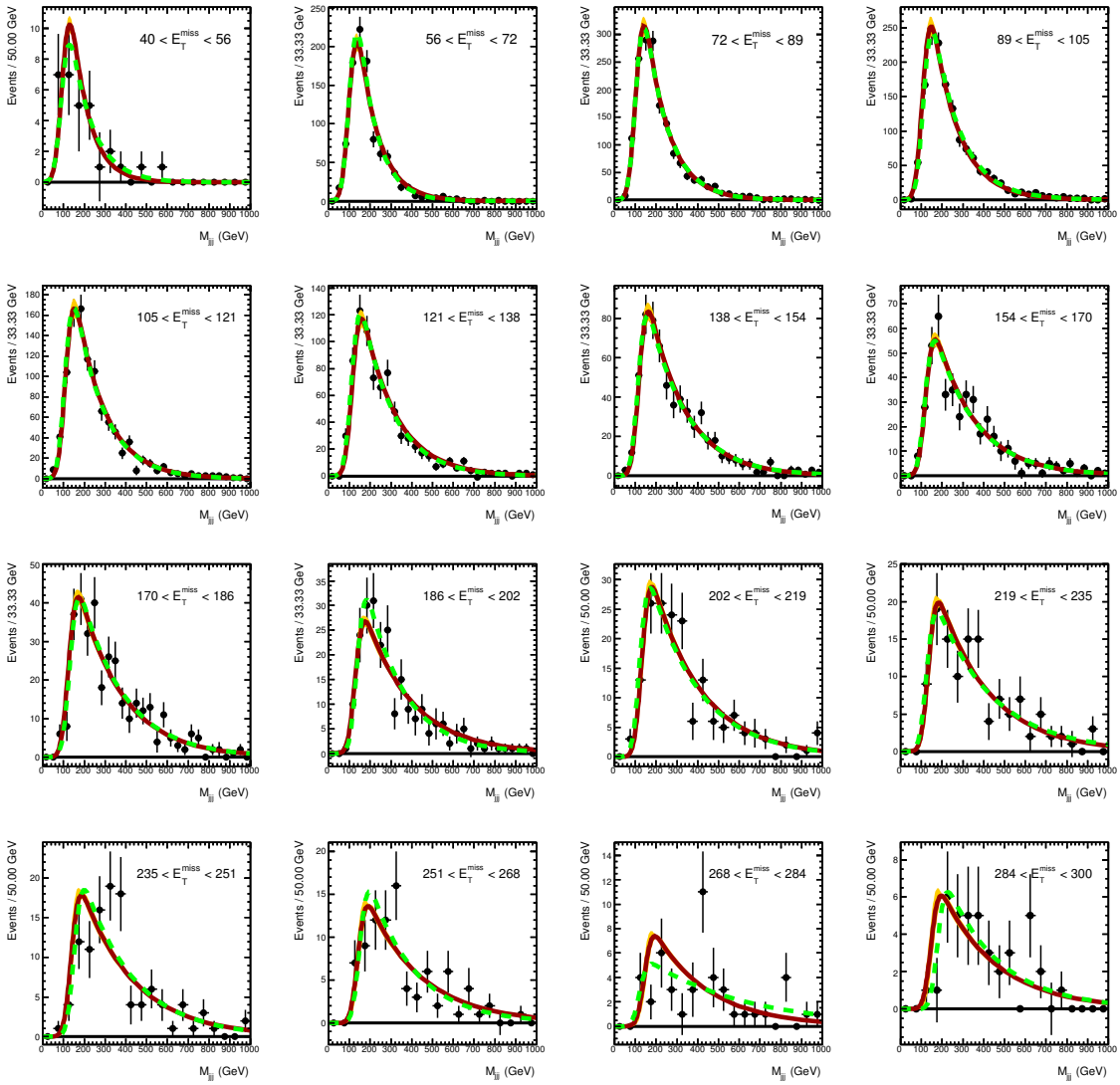
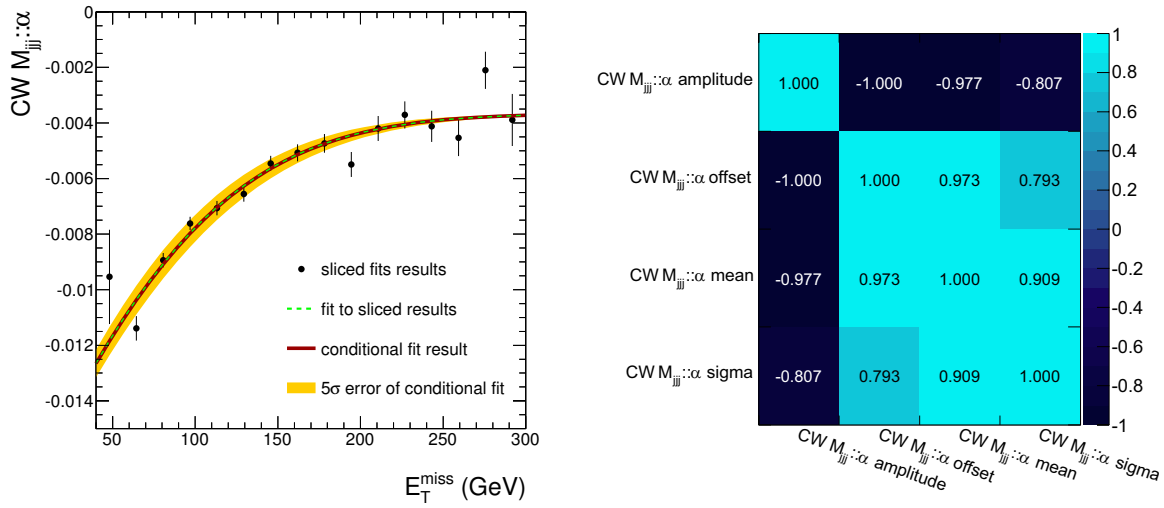


Figure 4.14: The M_{jjj} distribution of the CW sample sliced in $E_{\mathrm{T}}^{\mathrm{miss}}$ (markers). A fit of the M_{jjj} model to the data in each separate bin is given by the dashed curve, while the conditional PDF fit to the full data is given by the solid curve. The propagated errors of the conditional model fit are given as a light band with 5 standard deviations, however these are very small.

Using conditional PDFs to describe correlations

To describe correlations the PDF F(x;p) can be adjusted to match the data in each y-slice by changing the value of parameter p. Thus we replace parameter p by a function $p(y)$, and PDF F(x) becomes a conditional PDF F(x|y). This change we will now try to describe quantitatively for $\mathrm{PDF}_{CW}(M_{\mathrm{jjj}})$.

The shape of CW M_{jjj} distribution is described by the TTComb function. The TTComb function has three parameters: α, *mean* and σ, out of which we pick α as an example. If we slice the CW M_{jjj}-distribution in bins of missing transverse energy, as shown in Figure 4.14, we can fit our TTComb model to each $E_{\mathrm{T}}^{\mathrm{miss}}$-bin separately, displayed by the dashed curve in each subfigure. Taking the value and error of parameter α for each separate $E_{\mathrm{T}}^{\mathrm{miss}}$-bin, we can see if



(a) (b)

Figure 4.15: Evolution of the CW $M_{\mathrm{jjj}} :: \alpha$ parameter as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ (a). The dots are results of separate bin fits, which were fitted by the dashed curve. The solid curve is the result of the correlated model fit to the full CW sample, while 3 standard deviations of the full sample fit with propagated errors are given by the light band. Correlation matrix (b) of the fit to separate bin results shown by the dashed curve in the left plot.

and how it evolves with increasing $E_{\mathrm{T}}^{\mathrm{miss}}$. This evolution of parameter α as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ is shown in Figure 4.15(a) by the markers.

We parametrize the evolution of $\alpha(E_{\mathrm{T}}^{\mathrm{miss}})$ with an analytical function. The evolution can be described using the error function (*erf*) as follows:

$$\alpha(E_{\mathrm{T}}^{\mathrm{miss}}) = \mathit{amplitude} \times \mathrm{erf}\ ((\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} - \mathrm{mean})/\mathrm{sigma}) + \mathit{offset}, \tag{4.46}$$

where one PDF parameter α is replaced by four function parameters *amplitude, offset, mean* and *sigma*, that may not all need to be floating parameters in the fit. We rewrite $\mathrm{PDF}_{CW}(M_{\mathrm{jjj}})$ as $\mathrm{PDF}_{CW}(M_{\mathrm{jjj}}|E_{\mathrm{T}}^{\mathrm{miss}})$ by substituting parameter α with the function $\alpha(E_{\mathrm{T}}^{\mathrm{miss}})$.

The fit of the $\alpha(E_{\mathrm{T}}^{\mathrm{miss}})$ function to the evolution of α in Figure 4.15(a) is shown by the dashed curve. The correlation matrix of this $\alpha(E_{\mathrm{T}}^{\mathrm{miss}})$ fit is shown in Figure 4.15(b), from which we learn that many of these parameters are highly correlated with each other. These strong correlations suggest redundancy in the parametrization and can be eliminated by setting some of these parameters constant. We have chosen to set the *amplitude*, *offset* and *mean* constant, which leaves a function $\alpha(E_{\mathrm{T}}^{\mathrm{miss}})$ that has only one floating parameter: *sigma*. Thus we can replace the original M_{jjj} parameter α in our model by the $E_{\mathrm{T}}^{\mathrm{miss}}$-dependent function $\alpha(E_{\mathrm{T}}^{\mathrm{miss}})$ without increasing the total number of parameters.

To validate the conditional model $\mathrm{PDF}_{CW}(M_{\mathrm{jjj}}|E_{\mathrm{T}}^{\mathrm{miss}})$, we compare the shape of the conditional PDF fit to the full data, in slices of $E_{\mathrm{T}}^{\mathrm{miss}}$, to the non-conditional PDF ($\mathrm{PDF}_{CW}(M_{\mathrm{jjj}})$) fit only to data in that slice. In Figure 4.14 the conditional PDF fit is shown by the solid curve, while the non-conditional PDF fit to an individual slice is shown by the dashed curve, as mentioned before. In the same figure also the propagated error of the conditional PDF fit is shown with 5 standard deviations by the band. At the parameter α level, the shape of $\alpha(E_{\mathrm{T}}^{\mathrm{miss}})$ out of the conditional PDF fit to the full data is summarized by the solid curve and the error band

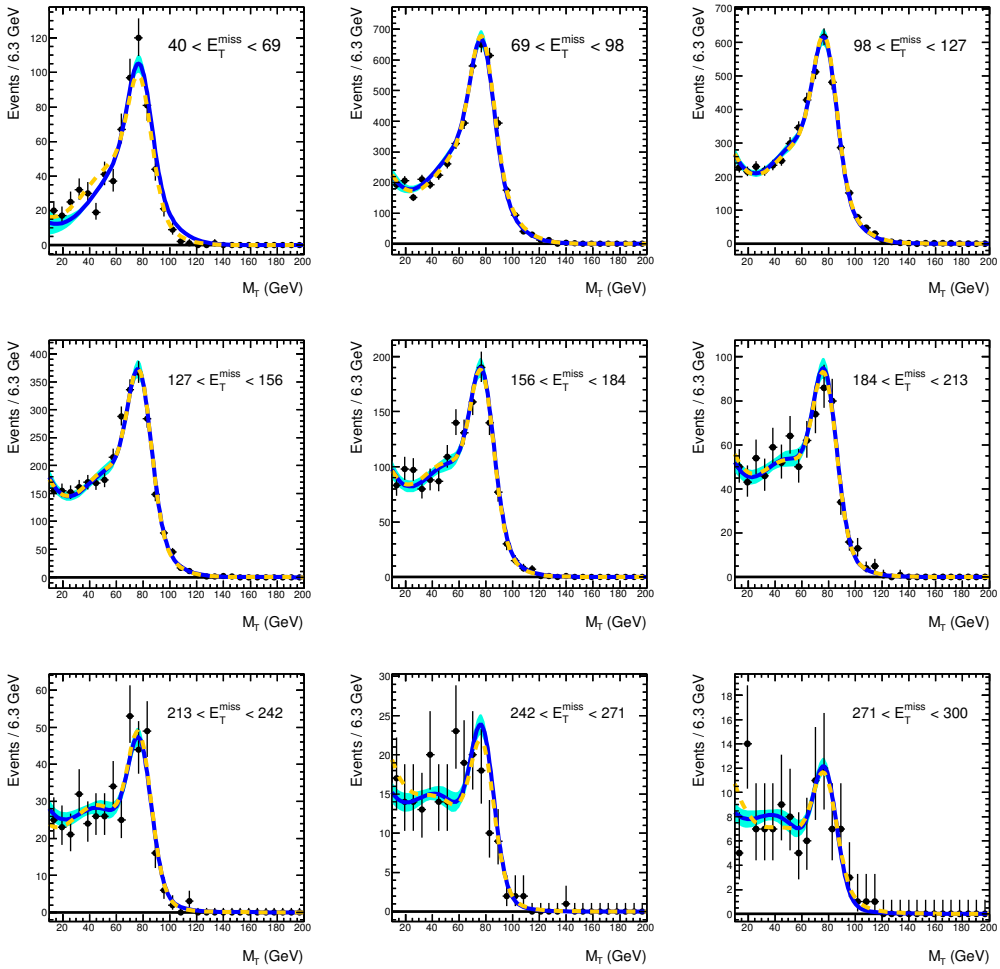


Figure 4.16: M_{T} distribution of the TP sample sliced in $E_{\mathrm{T}}^{\mathrm{miss}}$ in markers. A fit in every separate $E_{\mathrm{T}}^{\mathrm{miss}}$ slice is given by the dashed curve, while the conditional model fit to the full (non-sliced) data is given by the solid curve. The propagated errors of the conditional model fit are given as a light band with 3 standard deviations. The M_{T}-range in these plots is reduced from $[10, 450]$ to $[10, 200]$ for easier viewing.

in Figure 4.15(a). We conclude that the evolution of α is correctly described by the conditional model, while the parameter that is left floating, *sigma*, will account for a possible deviation between data and simulation.

In Figure 4.16 we show another example of a model with parameters replaced by conditional functions. This figure shows the M_{T} distribution for the TP sample in slices of $E_{\mathrm{T}}^{\mathrm{miss}}$. The parameters of MTFunc, described by Equation 4.45, that have been replaced by functions are f_{peak} and r_{mean}. The sliced M_{T} distributions are shown by the markers, while the separate bin fits are given again by the dashed curves. The conditional PDF model fit to the complete TP sample is shown by the solid curve, with three standard deviation errors given by the light band, that is in good agreement with the separate bins fit.

After having examined all possible correlations between observables and all background components, we conclude that for the TP model only the M_{T} PDF depends on $E_{\mathrm{T}}^{\mathrm{miss}}$, while for the CW and T2 models both the M_{T} and the M_{jjj} PDFs have a dependency on $E_{\mathrm{T}}^{\mathrm{miss}}$. Thus the

Parameter	Replacing Function
	CW
M_{T}::f_{peak}	$\min(1, \mathrm{erf}((\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} - \mu^{\mathrm{c}})/\sigma^{\mathrm{c}}) + \mathrm{b})$
M_{T}::μ	$b + E_{\mathrm{T}}^{\mathrm{miss}} \times a^{c}$
M_{T}::r_{μ}	$b^{c} + E_{\mathrm{T}}^{\mathrm{miss}} \times a$
M_{jjj}::α	$A^{c} \times \mathrm{erf}((E_{\mathrm{T}}^{\mathrm{miss}} - \mu^{c})/\sigma) + b^{c}$
M_{jjj}::μ	$b + E_{\mathrm{T}}^{\mathrm{miss}} \times a$
	TP
M_{T}::f_{peak}	$\min(1, \mathrm{erf}((\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} - \mu^{\mathrm{c}})/\sigma) + \mathrm{b}^{\mathrm{c}})$
M_{T}::r_{μ}	$\max(0.5, \mathrm{b}^{\mathrm{c}} + \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} \times \mathrm{a})$
	T2
M_{T}::μ	$\min(250, \mathrm{b} + \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} \times \mathrm{a}^{\mathrm{c}})$
M_{T}::σ	$\min(35, \mathrm{b} + \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} \times \mathrm{a}^{\mathrm{c}})$
M_{T}::f_{peak}	$\min(1, \mathrm{A}^{\mathrm{c}} \times \mathrm{erfc}((\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} - \mu)/\sigma^{\mathrm{c}})$
M_{T}::α	$\min(-1 \cdot 10^{-4}, \mathrm{b} + \mathrm{erf}((\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} - \mu^{\mathrm{c}})/\sigma^{\mathrm{c}}) + \mathrm{a}^{\mathrm{c}} \times \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}})$
M_{jjj}::μ	$b^{c} + E_{\mathrm{T}}^{\mathrm{miss}} \times a$
M_{jjj}::α	$A^{c} \times \mathrm{erf}((E_{\mathrm{T}}^{\mathrm{miss}} - \mu^{c})/\sigma) + b^{c}$

Table 4.7: Model parameters with corresponding functions that replaced them. Conditional parameters with a *c*-superscript are kept constant in the fit, as they are found to be redundant. The error function is denoted as *erf*, while *erfc* is the complementary error function. In the formulas a denotes a slope, b an offset and A an amplitude, while μ and σ are the mean and standard deviation of the Gaussians, from which the error functions are derived. Although for the sake of simplicity these variables have the same name in each formula above, they are different parameters of the model.

shape of the 3D PDF for the TP background is described by:

$$\mathrm{PDF}_{TP}^{\mathrm{3D}}(E_{\mathrm{T}}^{\mathrm{miss}}, M_{\mathrm{T}}, M_{\mathrm{jjj}}) = \mathrm{PDF}_{TP}^{\mathrm{1D}}(E_{\mathrm{T}}^{\mathrm{miss}}) \times \mathrm{PDF}_{TP}^{\mathrm{1D}}(M_{\mathrm{T}}|E_{\mathrm{T}}^{\mathrm{miss}}) \times \mathrm{PDF}_{TP}^{\mathrm{1D}}(M_{\mathrm{jjj}}) \tag{4.47}$$

and for CW and T2 by:

$$\begin{gathered}\mathrm{PDF}_{BG}^{\mathrm{3D}}(E_{\mathrm{T}}^{\mathrm{miss}}, M_{\mathrm{T}}, M_{\mathrm{jjj}}) = \\ \mathrm{PDF}_{BG}^{\mathrm{1D}}(E_{\mathrm{T}}^{\mathrm{miss}}) \times \mathrm{PDF}_{BG}^{\mathrm{1D}}(M_{\mathrm{T}}|E_{\mathrm{T}}^{\mathrm{miss}}) \times \mathrm{PDF}_{BG}^{\mathrm{1D}}(M_{\mathrm{jjj}}|E_{\mathrm{T}}^{\mathrm{miss}}) \ ,\end{gathered} \tag{4.48}$$

In Table 4.7 we list all the parameters that are substituted with the corresponding conditional functions, itemized by the background sample. As explained earlier, most of the conditional function parameters are redundant due to high correlations with each other, and can be kept constant. The *c*-superscript in Table 4.7 denotes the parameters that were made constant. After all substitutions the total number of floating parameters in the combined fit grows only by one to a total of 37, due to elimination of redundant parameters.

As mentioned in Section 4.3.2, the total number of parameters can also be reduced by identifying *global* parameters that describe the same behavior in multiple background models and have the same value within error margins. After performing the substitutions of model parameters by conditional functions, three more global parameters are identified and shown in Table 4.8. Thus we reduce the total shape parameter count for the combined background fit to 34.

Description	Parameter 1	Parameter 2
Offset of peak fraction in M_{T}	M_{T}::$\mathrm{f}_{\mathrm{peak}}$::b(TP)	M_{T}::$\mathrm{f}_{\mathrm{peak}}$::b(CW)
Slope of ratio of Gaussian means in M_{T}	M_{T}::r_{μ}::a(TP)	M_{T}::r_{μ}::a(CW)
Offset of ratio of Gaussian means in M_{T}	M_{T}::r_{μ}::b(TP)	M_{T}::r_{μ}::b(CW)

Table 4.8: Sets of conditional function parameters listed in Table 4.7, that are combined into a single global parameter.

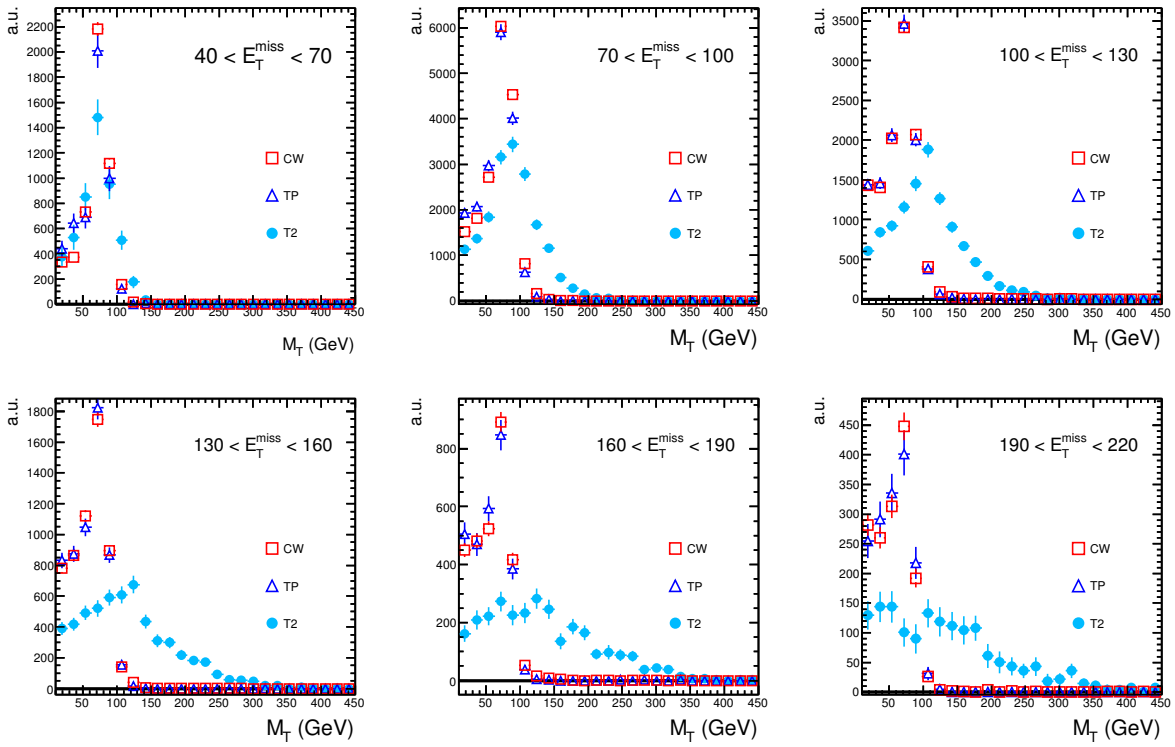


Figure 4.17: Normalized M_{T} distributions of the three main background samples sliced in $E_{\mathrm{T}}^{\mathrm{miss}}$. Dileptonic $t\bar{t}$ distribution has a remarkably strong evolution comparing to the other two samples, that serves to distinguish it in a combined fit.

4.3.4 Evolution of dileptonic $t\bar{t}$

Figure 4.17 shows the M_{T} distribution for the three background samples in slices of $E_{\mathrm{T}}^{\mathrm{miss}}$. In all slices of $E_{\mathrm{T}}^{\mathrm{miss}}$ the CW and TP events show the characteristic W-mass Jacobian peak. Furthermore the dileptonic $t\bar{t}$ (T2) background shows a strong evolution as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ compared to the other two samples.

The T2 background is difficult to describe, and has been a source of trouble for many data-driven methods developed in the ATLAS SUSY working group, but the behavior shown in Figure 4.17 gives the combined fit method a handle on its shape and yield. The advantage of the combined fit method is that it is a multidimensional fit method, and by using the conditional $\mathrm{PDF}_{T2}(M_{\mathrm{T}}|E_{\mathrm{T}}^{\mathrm{miss}})$ we can describe the correlations.

The question we answer in this section is whether the T2 peak in the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ slice is

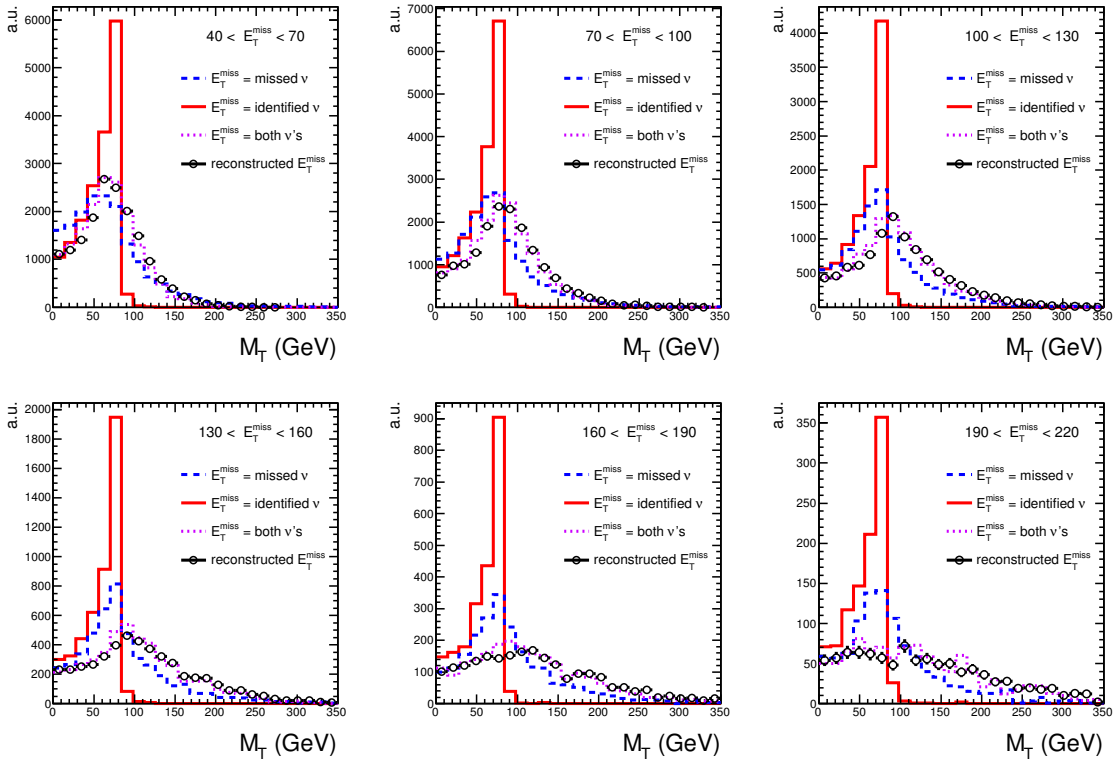


Figure 4.18: Normalized M_{T} distributions in slices of $E_{\mathrm{T}}^{\mathrm{miss}}$, where the missing transverse energy is given by the measured side neutrino (solid curve), missed side neutrino (dashed curve), both neutrinos (dotted curve) and finally the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ (round markers).

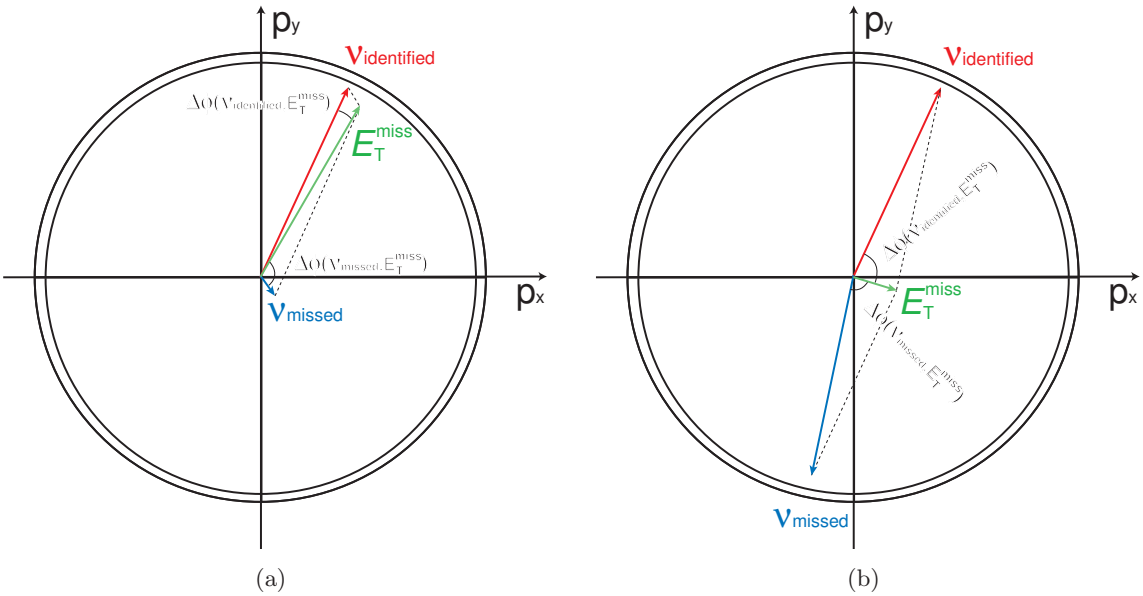


Figure 4.19: Schematic visualization of two $E_{\mathrm{T}}^{\mathrm{miss}}$ configurations, where the missed neutrino has very small p_{T} (a), and where the two neutrinos are produced almost back-to-back (b).

physics inspired, or in other words whether it is a W-mass Jacobian peak. The answer is that it is not a Jacobian, but is merely a consequence of our event selection.

Missing transverse energy in dileptonic $t\bar{t}$ events As was pointed out in Section 4.1.4, dileptonic $t\bar{t}$ events have two leptons but only one is *identified* and the other is *missed*. The two neutrinos produced in the decay of the two respective W bosons, we from now on refer to as the *'identified'* neutrino and the *'missed'* neutrino. If reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$, as measured in the detector, is the result of only the identified neutrino, we expect to see the Jacobian peak in M_{T}. If on the other hand $E_{\mathrm{T}}^{\mathrm{miss}}$ is given only by the missed neutrino, then in M_{T} we expect to see a smeared shape without any resemblance to the Jacobian: the (identified) lepton from one W boson combined with the (missed) neutrino from the other W boson should not provide any kind of peak in the M_{T} distribution, unless event selection makes it so.

The above statements are proven by Figure 4.18, that shows recalculated M_{T} distributions if $E_{\mathrm{T}}^{\mathrm{miss}}$ is given by the identified neutrino (solid curve) and the missed neutrino (dashed curve). The M_{T} distributions, with the identified neutrino providing $E_{\mathrm{T}}^{\mathrm{miss}}$, show nice Jacobian peaks in all slices of missing transverse energy. The M_{T} distributions recalculated with solely the missed neutrino providing $E_{\mathrm{T}}^{\mathrm{miss}}$ are as expected a smeared shape. Although it peaks at the same value as the real Jacobian, this is only an effect of our event selection and not of underlying physics.

To understand the composition of $E_{\mathrm{T}}^{\mathrm{miss}}$ in dileptonic $t\bar{t}$ events, we assume that reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ for dileptonic $t\bar{t}$ events is purely the superposition of the two neutrinos. This assumption is proven also in Figure 4.18, where recalculated M_{T} distributions are shown if transverse missing energy is given by both neutrinos (dotted curve) and the original reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ (round markers). The M_{T} distributions where both of the neutrinos are vectorially summed to form the recalculated transverse missing energy, show an almost identical behavior to the reconstructed M_{T} distributions, affirming our assumption on the composition of $E_{\mathrm{T}}^{\mathrm{miss}}$. From this we conclude that the missed lepton and possible neutrinos coming from decays in parton showers or τ lepton decays, do not play an important role in reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$.

Transverse mass in dileptonic $t\bar{t}$ events If we follow the proven assumption, that $E_{\mathrm{T}}^{\mathrm{miss}}$ in dileptonic $t\bar{t}$ events is given by the superposition of the two neutrinos, then we can postulate the following two hypothesis. Firstly, if the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ slice of Figure 4.17 is a real Jacobian peak, then $E_{\mathrm{T}}^{\mathrm{miss}}$ is mainly given by the identified neutrino. If this hypothesis is true, then the distance in azimuthal angle ϕ between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the identified neutrino must be very small for the low $E_{\mathrm{T}}^{\mathrm{miss}}$ events. Secondly, if the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ slice is a real Jacobian peak, then the missed neutrino is produced with very low p_{T} in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ events, or otherwise it would significantly alter the total $E_{\mathrm{T}}^{\mathrm{miss}}$. Thus the distance in ϕ between the missed neutrino and the identified neutrino (or $E_{\mathrm{T}}^{\mathrm{miss}}$) should be randomly distributed, as the missed neutrino does not contribute almost anything to the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$. A visualization of the two hypotheses and the following statements on $\Delta\phi$ is shown schematically in Figure 4.19(a).

Both of these hypotheses are proven false by Figure 4.20, where the $\Delta\phi$ distributions are shown again in slices of $E_{\mathrm{T}}^{\mathrm{miss}}$. In the upper left plot of Figure 4.20, the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ slice is shown. The solid curves give $\Delta\phi$ between the identified neutrino and $E_{\mathrm{T}}^{\mathrm{miss}}$, while the dashed curves give $\Delta\phi$ between the missed neutrino and $E_{\mathrm{T}}^{\mathrm{miss}}$. The distance between $E_{\mathrm{T}}^{\mathrm{miss}}$ and the identified neutrino is not small in the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ slice, as expected from the hypotheses, in contrary, the $\Delta\phi(\nu_{\mathrm{identified}}, E_{\mathrm{T}}^{\mathrm{miss}})$ distribution is as broad as the $\Delta\phi(\nu_{\mathrm{missed}}, E_{\mathrm{T}}^{\mathrm{miss}})$ distribution.

The dotted curves of Figure 4.20 show the $\Delta\phi$ distribution between the two neutrinos, which shows that they are predominantly produced back-to-back in the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ slice. As the neutrinos are produced (almost) back-to-back, the sum of the two delivers an event with low

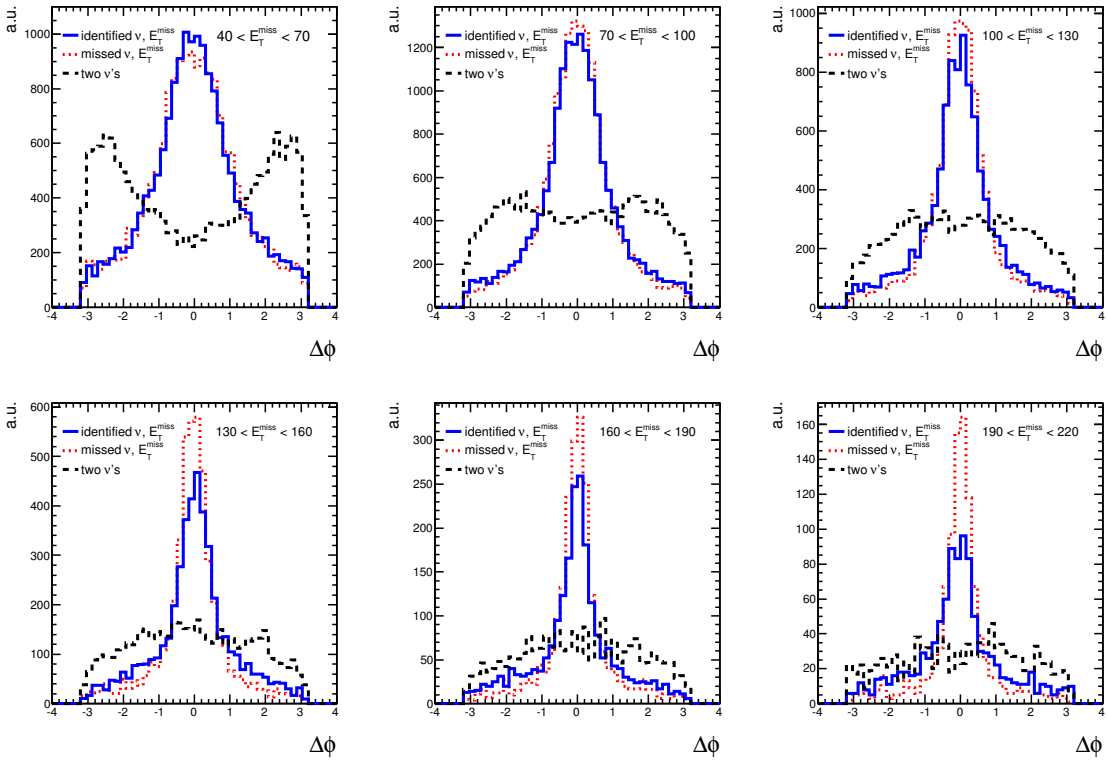


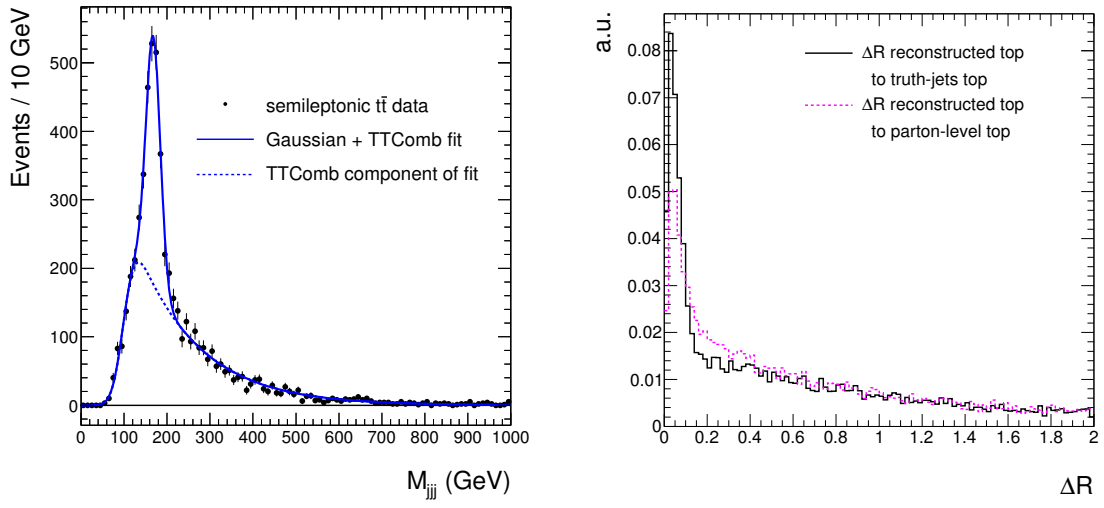
Figure 4.20: Normalized $\Delta\phi$ distributions in slices of $E_{\mathrm{T}}^{\mathrm{miss}}$, between the identified-side neutrino and $E_{\mathrm{T}}^{\mathrm{miss}}$ (solid curve), missed-side neutrino and $E_{\mathrm{T}}^{\mathrm{miss}}$ (dotted curve) and the two neutrinos (dashed curve).

$E_{\mathrm{T}}^{\mathrm{miss}}$, that points away from both neutrinos. This configuration is visualized schematically in Figure 4.19(b). Both hypotheses are thus incorrect, as the identified neutrino is not the dominant component of $E_{\mathrm{T}}^{\mathrm{miss}}$, and neither is the missed neutrino produced with very low p_{T}.

We conclude that the M_{T} distribution in the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ slice of Figure 4.17 is not a Jacobian peak, but an effect of our event selection. If for example our event selection would require $E_{\mathrm{T}}^{\mathrm{miss}}$ above 100 GeV, the upper right plot of Figure 4.17 would be the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ slice, which is clearly not the W-mass Jacobian peak. Despite the fact that the peak is not a W-mass Jacobian, the M_{T} behavior as shown in Figure 4.17 is representative of dileptonic $t\bar{t}$ events with one identified lepton, and still gives a handle on estimating their contribution in a combined fit.

4.3.5 Top peak, combinatorics separation

The M_{jjj} distribution of the semileptonic $t\bar{t}$ sample, shown in Figure 4.21(a), has two distinct parts: a *top peak* (TP) and a *top combinatorics* (TC) component. As we try to separate these two components in our model two questions arise. The first question is whether the three jets found by the M_{jjj} algorithm can be matched unambiguously to the three matrix element (ME) partons from the hadronic top decay. As matching reconstruction-side objects such as jets, to ME partons is ambivalent, we cannot solve this issue completely, but in the discussion we introduce a parameter that can be used to split TP from TC events in practice. The second question arises from the fact that our TP/TC separation is based on a truth-level information, which can possibly introduce a bias in our reconstructed distributions. We will show however



(a) (b)

Figure 4.21: The semileptonic $t\bar{t}$ M_{jjj} distribution (a) before the TP/TC split, with the corresponding Gaussian+TTComb model fit (solid curve) and the TTComb component (dashed curve) separately. Normalized distributions of ΔR (b) between the reconstructed top quark and the truth jets top quark (solid histogram) and between the reconstructed top quark and the partons top quark (dashed histogram).

that this is not the case.

Ambiguity in matching ME partons to reconstructed jets The production and hadronic decay of a top quark is accompanied by radiation of associated gluons (and quarks). Hard gluon radiation results in an extra observable jet, but ambiguities are introduced due to collinear and soft gluon radiation, as jet reconstruction algorithms and detectors are not perfect. Thus a simulated parton or the sum of three partons are not observables as such. An illustrative example is shown in Figure 4.21(b), where the $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ distribution between the *reconstructed top* and the *parton-level top* is shown by the dashed curve. The reconstructed top quark is the sum of three jets with the highest summed p_{T} found by the M_{jjj} algorithm, and the parton-level top quark is defined as the sum of the three truth-level decay products of the top quark.

If the ambiguities concerning soft gluon radiation were not present, we would expect to see a clear distinction in ΔR between correctly matched top events (low ΔR) and combinatorial top events (high ΔR). In contrast the ΔR distribution has a smooth transition from the peak at low ΔR to the top combinatorics events at higher ΔR values.

We improve the quality of parton-to-jet matching by replacing simulated partons by *truth-jets*. A truth-jet is made by running the jet algorithm on all interacting simulation particles, skipping only neutrinos, thereby also taking soft gluon radiation into account. Truth-jets are what we expect to measure in the limit that we have an ideal detector and an ideal jet algorithm[5], and are therefore likely to result in a better defined match than simulated gluons and quarks. We define the *truth-jets top* as the sum of three truth-jets, that are individually best matched in ΔR to the three partons from hadronic top.

Figure 4.21(b) shows the ΔR distribution between the reconstructed top and the truth-jets

[5]To make sure that we do not have double counting, we do overlap removal between truth-jets and truth-electrons just the way we do it for reconstruction objects.

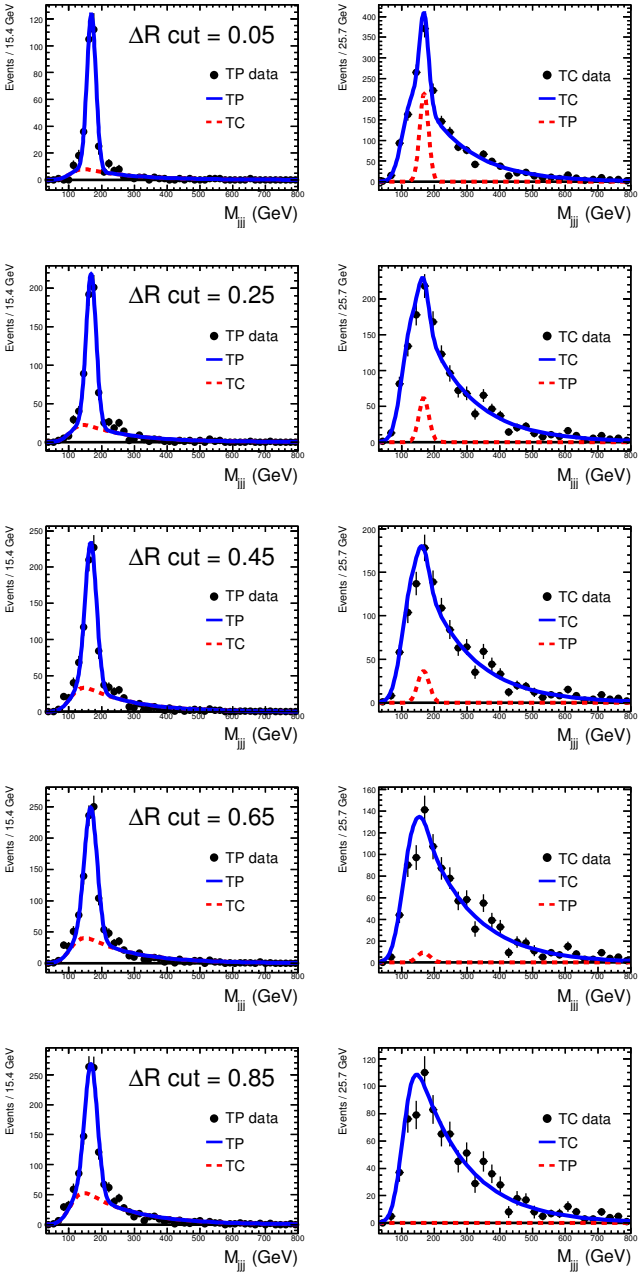


Figure 4.22: The results of simultaneous fits to the TP/TC divided semileptonic $t\bar{t}$ sample, separated by the ΔR cut denoted in the plot. The column on the left shows the data that has been tagged as TP and on the right as TC. Data is represented by the dots and the fit result by the solid curve, while the dashed curve shows the contribution of TC(TP) in the opposite TP(TC) data sample.

top by the solid curve. In this distribution we do see a clearer distinction between correctly matched and combinatorial reconstruction top events at $\Delta R \sim 0.1$, thus we can use this variable to separate the TP events from TC events.

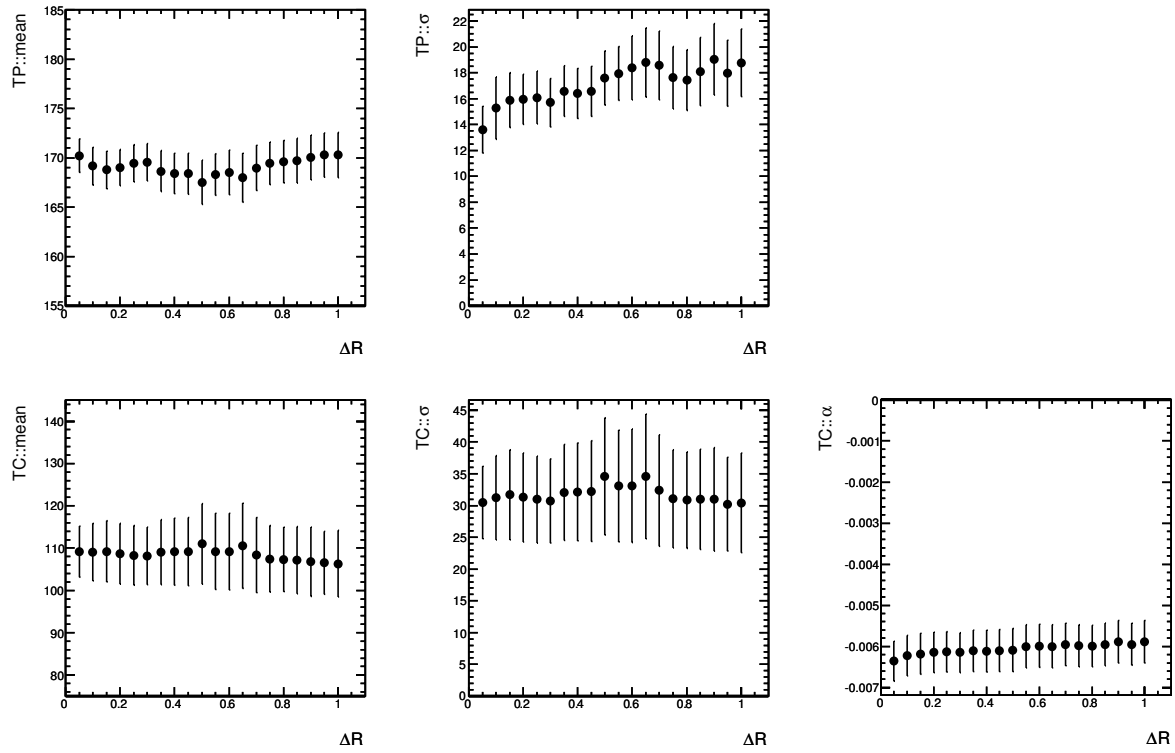


Figure 4.23: The evolution of simultaneous fit parameters as function of ΔR. On top are two parameters describing our top peak model *TP::mean* and *TP::*σ, and on the bottom are the three parameters of the combinatorics model *TC::mean*, *TC::*σ and *TC::*α. Within the error margins the values of all parameters stay constant, hence the shape is invariant under the change of ΔR.

Splitting $t\bar{t}$ events into TP/TC components without introducing a bias Splitting the semileptonic $t\bar{t}$ sample into TP/TC components for the combined fit method, is based on ΔR between reconstructed top and truth-jets top, which we from here on refer to just as ΔR. This split is however not perfect as some mixing occurs, because even a combinatorial top occasionally has $\Delta R < 0.1$ with respect to the truth-jets top. But this does not pose a problem for our analysis, if we can prove that the TP/TC shape in the split samples is invariant under the change of ΔR. If that is true, we can deduce the total fraction of TP (TC) events in the full semileptonic $t\bar{t}$ sample.

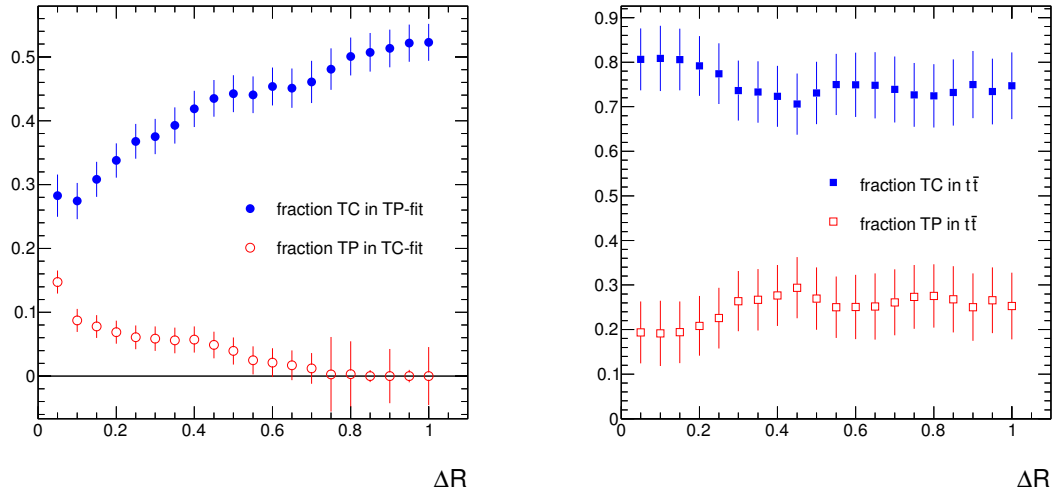
We study the shape dependence on the ΔR cut by setting up a simultaneous fitting procedure. For each ΔR value the TP and TC (pseudo-) data samples are then fitted simultaneously to the same model, while splitting only the parameter that describes the relative fraction of TP in each data sample. We thus assume that the shapes of TP and TC are the same in both data samples.

Our total model has five shape parameters:

- top peak described by a Gaussian: *TP::mean* and *TP::*σ
- top combinatorics described by a TTComb: *TC::mean*, *TC::*σ and *TC::*α,

and two yield parameters:

- fraction TP in TP-fit: f_{TP}^{TP}



(a) (b)

Figure 4.24: Fraction of TP(TC) events in the opposite TC(TP) data sample (a) as fitted simultaneously for different ΔR cuts. Total TP(TC) fraction in the complete semileptonic $t\bar{t}$ sample (b) as a function of ΔR. Within the error margins, the fractions stay constant, giving us confidence that our model is correct and choosing a specific cut in ΔR will not bias our results.

- fraction TP in TC-fit: f_{TP}^{TC}.

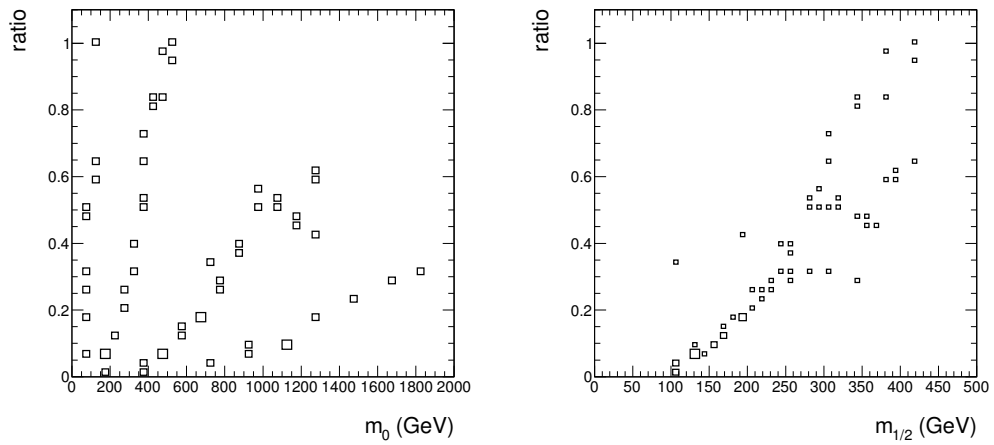
Figure 4.22 shows the results of several simultaneous fits for increasing values of ΔR. Each simultaneous fit is presented by two adjacent plots, the left one showing the TP-data (dots) and fit result (solid curve), while the right one shows the TC-data and fit result. Each plot contains also the fitted contribution of the opposite data sample, shown by a dashed curve, so in the left plot we see the TC component in TP-data and in the right one the TP component of the TC-fit is shown. The data is correctly described by our model for all ΔR values.

The fitted values and errors of the shape parameters are shown in Figure 4.23 as a function of ΔR. The five shape parameters stay constant within the error margins, hence we conclude that the TP/TC shapes are invariant under the change of ΔR.

The small increase in the value of $TP::\sigma$ is understood, as for higher values of ΔR events are allowed into the TP-data, where the jet algorithm fails to collect all the energy deposits for one of the jets. This can occur in events where a gluon is radiated off one of the final state quarks, but this gluon is too soft to change the reconstructed jet momentum and direction significantly, though its energy deposit is missed by the jet algorithm. Hence if this misreconstructed jet is selected by the M_{jjj} algorithm, the reconstructed M_{jjj} value is slightly lower than the top mass. Such events give a small tail on the low side of the Gaussian TP peak, that slightly increases the fitted $TP::\sigma$ value.

The evolution of the two yield parameters in the simultaneous fit is shown in Figure 4.24(a) as a function of ΔR. In empty dots is f_{TP}^{TC}, while in full dots is $(1-f_{TP}^{TP})$ which is the combinatorics fraction in TP-fit or f_{TC}^{TP}. As expected we see the fraction of TC events in TP-fit increase with ΔR, as more events with a worse match contribute to our TP sample. We also see that with increasing ΔR our TC sample becomes gradually free of correctly reconstructed TP events, reaching a completely pure combinatorics sample ($f_{TP}^{TC} = 0$) around $\Delta R \simeq 0.8$.

In Figure 4.24(b) we show the TP (TC) fitted fraction in the complete semileptonic $t\bar{t}$ sample. The total fractions are the sum of the yields from the TP and the TC fitted samples, calculated



(a) (b)

Figure 4.25: *Ratio* versus m_0 (a) and $m_{1/2}$ (b) for the mSUGRA grid points.

as:

$$f_{TP}^{\text{total}} = \alpha \cdot f_{TP}^{TP} + (1-\alpha) \cdot f_{TP}^{TC} \quad (4.49)$$

$$f_{TC}^{\text{total}} = \alpha \cdot (1 - f_{TP}^{TP}) + (1-\alpha) \cdot (1 - f_{TP}^{TC}) \ , \quad (4.50)$$

where α is the fraction of events in TP-data compared to the total semileptonic $t\bar{t}$ sample for a specific ΔR cut. Although the fit yields f_{TP}^{TP} and f_{TC}^{TP} evolve with ΔR in Figure 4.24(a), the total contribution of TP(TC) in Figure 4.24(b) is independent of the specific ΔR value within error margins.

In summary, we conclude that the shape of TP/TC models in invariant under the ΔR cut, and that the total fraction of TP/TC events in semileptonic $t\bar{t}$ events is independent of the ΔR value. Hence we have a technique, based on ΔR, to split the semileptonic $t\bar{t}$ sample into a combinatorics and a peak sample, that is free of shape bias. For the combined fit analysis we choose $\Delta R = 0.1$, based on the shape of the distribution shown in Figure 4.21(b).

4.3.6 Generic model of SUSY in the control region

If SUSY exists and supersymmetric particles are produced at the LHC, then they are also likely to contribute events to the control region. This SUSY contamination has been the subject of much work in the ATLAS collaboration, as it leads to an overestimation of the background in the signal region, if unaccounted for. The combined fit method pioneered the contamination assessment by constructing an empirical *Ansatz model* that describes SUSY in the control region. Defining an analytical model for SUSY in the full phase space is impossible, as different models have very different signatures, but in the control region all the SUSY models are at the low edge of the available kinematics. This gives us the possibility to model the variation of SUSY shape in the control region.

We define the L-shaped control region by $E_{\text{T}}^{\text{miss}} < 200$ GeV $\bigvee M_{\text{T}} < 150$ GeV, as shown schematically in Figure 4.2, where we perform the fit to data. The fit results are extrapolated to the SUSY-rich signal region, which is defined by reversing the control region definition to $E_{\text{T}}^{\text{miss}} > 200$ GeV $\bigwedge M_{\text{T}} > 150$ GeV. The signal region boundaries choice is based on Figure 4.2 and a previous study [60]. After performing the significance reach estimation of the combined fit

method, as discussed in Section 4.4.7, these boundaries could be refined in an iterative procedure to maximize the reach on simulated samples, but this falls outside the scope of this thesis.

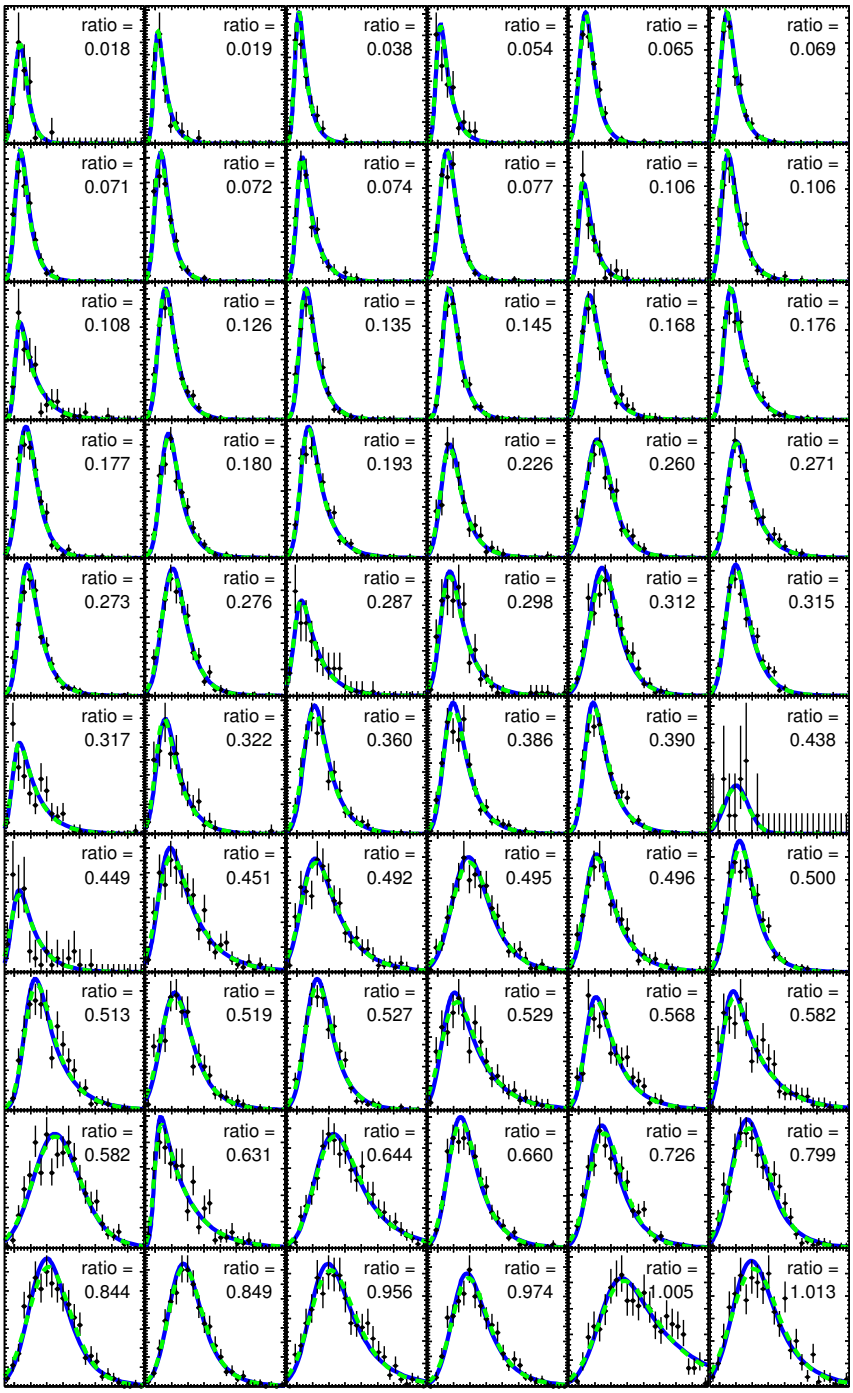


Figure 4.26: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in $M_{\mathrm{T}} < 150$ GeV region for all mSUGRA grid points sorted by increasing *ratio*. The lowest *ratio* is in top left, while the highest is in the bottom right. For each mSUGRA point on top of the data (markers) is the 1D $E_{\mathrm{T}}^{\mathrm{miss}}$-fit (dashed curve) and the 3D fit (solid curve) in the control region.

Classification of SUSY mass scale

To study different SUSY scenarios and the dependence on SUSY model parameters, a grid of mSUGRA points has been produced by the SUSY working group [100], described in Section 4.1.1. The grid of models has both high and low mass SUSY. To be able to grossly classify all these mSUGRA points in terms of 'SUSY mass', we introduce the *ratio* defined as:

$$\textbf{ratio} = \frac{N_{\mathrm{SIG}}}{N_{\mathrm{CTL}}}, \tag{4.51}$$

where N_{SIG} (N_{CTL}) is the number of events in the signal (control) region. Low value of *ratio* means that most SUSY events are in the control region, while a high value means that most events are in the signal box. Figure 4.25 shows the distributions of *ratio* versus the m_0 and $m_{1/2}$ parameters for the grid of mSUGRA points. The radial m_0, $m_{1/2}$ lines, used to produce the grid, are clearly visible hence our argument that *ratio* is a rough measure of SUSY mass.

Shape evolution with SUSY mass scale

Figure 4.26 shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the low M_{T} region ($M_{\mathrm{T}} < 150$ GeV) for all mSUGRA grid points sorted by increasing *ratio*. For each mSUGRA grid point are also shown: a 1D $E_{\mathrm{T}}^{\mathrm{miss}}$-fit in the low M_{T} region (dashed curve) described by the TTComb function, and a 3D fit in the full L-shaped control region (solid curve). Figure 4.26 shows that the same shape can be used to describe all mSUGRA points, with a finite number of parameters. Besides that, the evolution of the shape is gradual with increasing *ratio*, as the $E_{\mathrm{T}}^{\mathrm{miss}}$distribution becomes wider and the peak moves to higher values with successive rising *ratio* values.

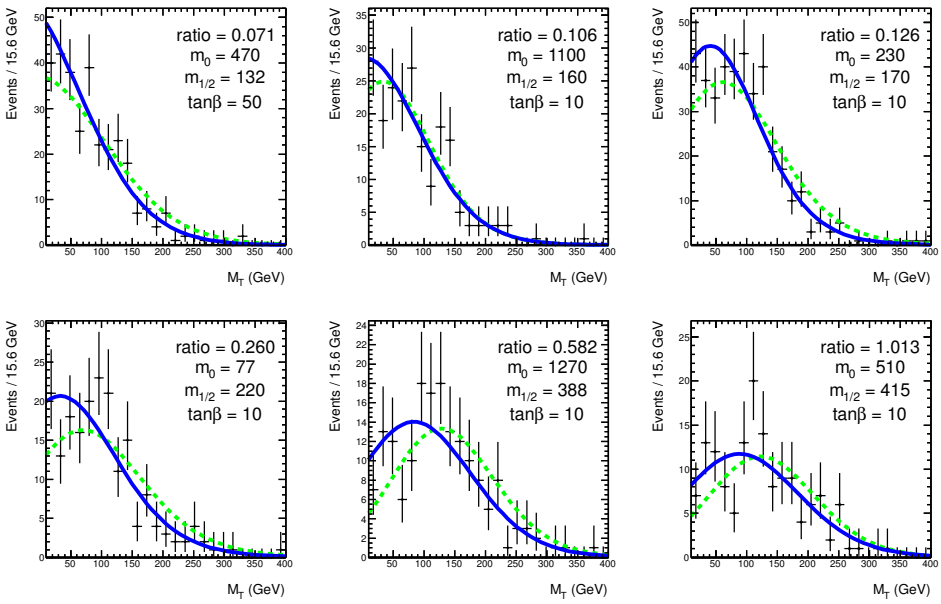
We see a gradual evolution in all three observables, as shown for M_{T} and M_{jjj} by Figures 4.27(a) and 4.27(b) for six representative mSUGRA grid points. The M_{T} and M_{jjj} distributions are shown respectively in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ region ($E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV) and the full L-shaped control region. Again superimposed are the 1D fits (dashed curve) and the 3D fits (solid curve), where the PDF for M_{T} is the TTComb function, while for M_{jjj} it is a convolution of an exponential with a Gaussian. The fits show that the same 3D PDF can be used for all grid points. The small deviations in M_{T} of the 3D fit are mainly due to the technical difficulty of plotting a projection of a L-shaped region in a single observable. The simulated data distribution is shown in the $E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV region, while the fit is performed on the L-shaped control region but projected onto the $E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV region.

Correlations in mSUGRA grid

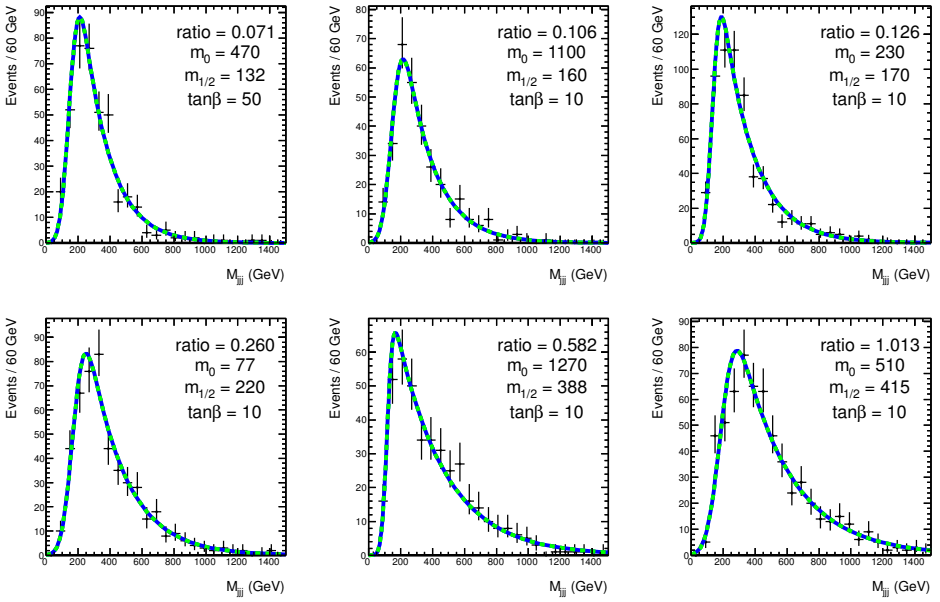
The great advantage of having a grid of mSUGRA models is that we can study how the parameters of our model are correlated with each other and with the *ratio*. If there is a strong correlation between two parameters in a fit, then one of them is redundant. In the case of the background models, we had no other choice but to set the redundant parameter constant. But with the grid of mSUGRA points we can remove this redundancy in the parametrization by describing one parameter as a function of another, thus also reducing the total number of parameters and simplifying our fitting procedure.

Our initial 3D Ansatz model of SUSY contamination in the control region has 9 parameters:

- For missing transverse energy (TTComb): $E_{\mathrm{T}}^{\mathrm{miss}}$::$\alpha$, $E_{\mathrm{T}}^{\mathrm{miss}}$::*mean*, $E_{\mathrm{T}}^{\mathrm{miss}}$::$\sigma$
- For transverse mass (TTComb): M_{T}::α, M_{T}::*mean*, M_{T}::σ



(a)

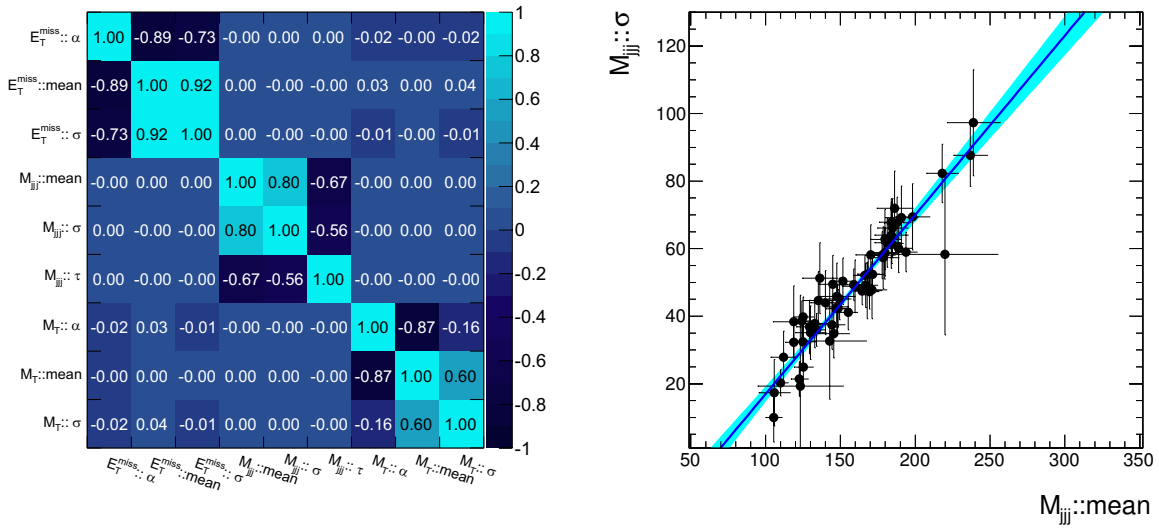


(b)

Figure 4.27: M_{T} distributions in $E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV region (a) and M_{jjj} distributions in the control region (b) for 6 grid points sorted by increasing *ratio*. The lowest *ratio* is in top left, while the highest in the bottom right. For each mSUGRA point superimposed on top of the data is the 1D fit as the dashed curve and the 3D fit as a solid curve.

- For three-jet mass (Exponential $\otimes$ Gaussian): M_{jjj}::τ, M_{jjj}::*mean*, M_{jjj}::σ.

Figure 4.28(a) shows the correlation matrix of the 3D Ansatz model fit for a single mSUGRA point with $\tan\beta = 10$, $m_0 = 470$ GeV and $m_{1/2} = 380$ GeV, which shows that many parameters are strongly correlated with each other. If we compare the values of M_{jjj}::*mean* and M_{jjj}::σ for



(a) (b)

Figure 4.28: Correlation matrix of the fit for a single mSUGRA model (a) with $\tan\beta = 10$, $m_0 = 470$ GeV and $m_{1/2} = 380$ GeV. Distribution of fitted values of $M_{\rm jjj} :: mean$ and $M_{\rm jjj} :: \sigma$ for all the grid points (b). Overlaid in a solid curve is a straight line fit, while the light band gives the one standard deviation error of the fit.

all grid points we get the distribution seen in Figure 4.28(b). We see a clear linear dependence that we can describe by a straight line, shown by the solid curve fit in Figure 4.28(b). We can replace the $M_{\rm jjj}$::σ parameter in our model by a function, that describes the linear dependence on $M_{\rm jjj}$::*mean*. Hence we are left with a 3D model with one parameter less, so 8 parameters in total.

We consecutively repeat this procedure for the other parameters of our Ansatz model to come to the surprising conclusion that we only need 2 parameters, $E_{\rm T}^{\rm miss}$ ::α and $M_{\rm T}$::α, to describe all mSUGRA grid models in the control region in three dimensions. The other 7 initial parameters can be described by simple functional relations to the remaining two and each other as:

$$\begin{aligned}
{\rm E_T^{miss}} ::\sigma &= 7659.78 \times {\rm E_T^{miss}} ::\alpha + 156.85, \\
{\rm E_T^{miss}} ::{\rm mean} &= 2.4402 \times {\rm E_T^{miss}} ::\sigma - 16.459, \\
{\rm M_T} ::{\rm mean} &= -7815.7 \times {\rm M_T} ::\alpha - 52.5861, \\
{\rm M_T} ::\sigma &= 101.61 \quad (constant), \\
{\rm M_{jjj}} ::\tau &= 17614.7 \times {\rm E_T^{miss}} ::\alpha + 436.45, \\
{\rm M_{jjj}} ::{\rm mean} &= 0.4339 \times {\rm M_{jjj}} ::\tau + 39.567, \\
{\rm M_{jjj}} ::\sigma &= 0.5398 \times {\rm M_{jjj}} ::{\rm mean} - 37.67. \qquad (4.52)
\end{aligned}$$

The 2-parameter Ansatz model in three dimensions describes the data correctly, as shown for $E_{\rm T}^{\rm miss}$($M_{\rm T}$ and $M_{\rm jjj}$) by the solid curve in respectively Figure 4.29 (Figures 4.30(a) and 4.30(b)) for the six representative mSUGRA points. For comparison the original 9-parameter model fit is shown in the same figures by the dotted curve, while the light band shows the propagated three standard deviations error of this original fit. We claim that our reduced Ansatz model fits reasonably well to all the grid mSUGRA points and compares well to the original 9-parameter

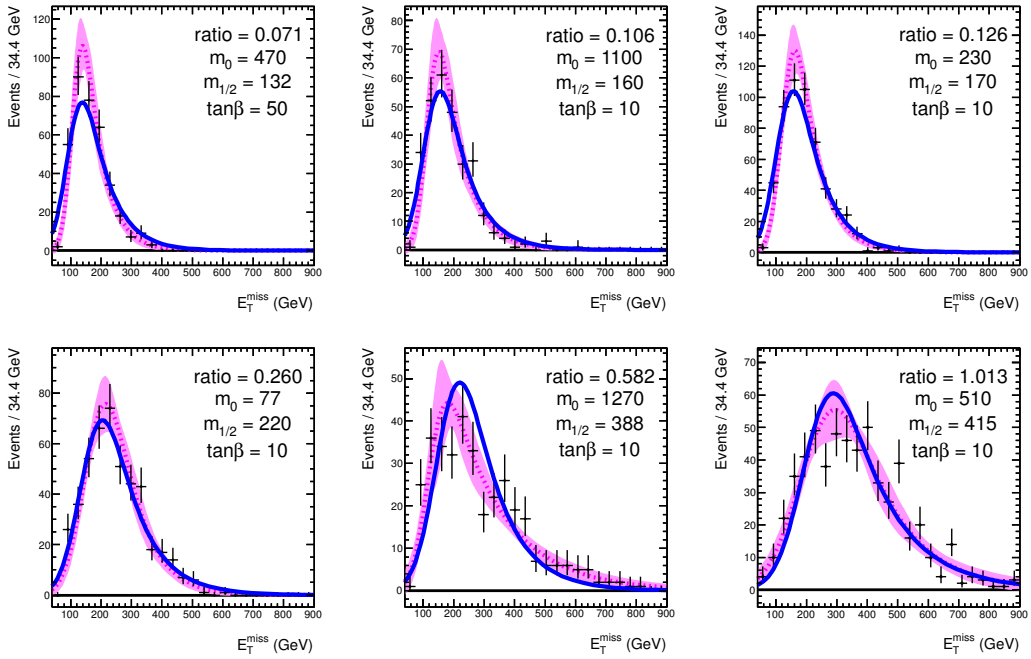


Figure 4.29: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in $M_{\mathrm{T}} < 150$ GeV region for six representative mSUGRA grid points sorted by increasing *ratio*. The lowest *ratio* is in top left, while the highest is in the bottom right. For each mSUGRA point on top of the data (markers) is the original 9-parameter fit (dotted curve) with corresponding propagated three standard deviations error (light band), and the 2-parameter fit (solid curve).

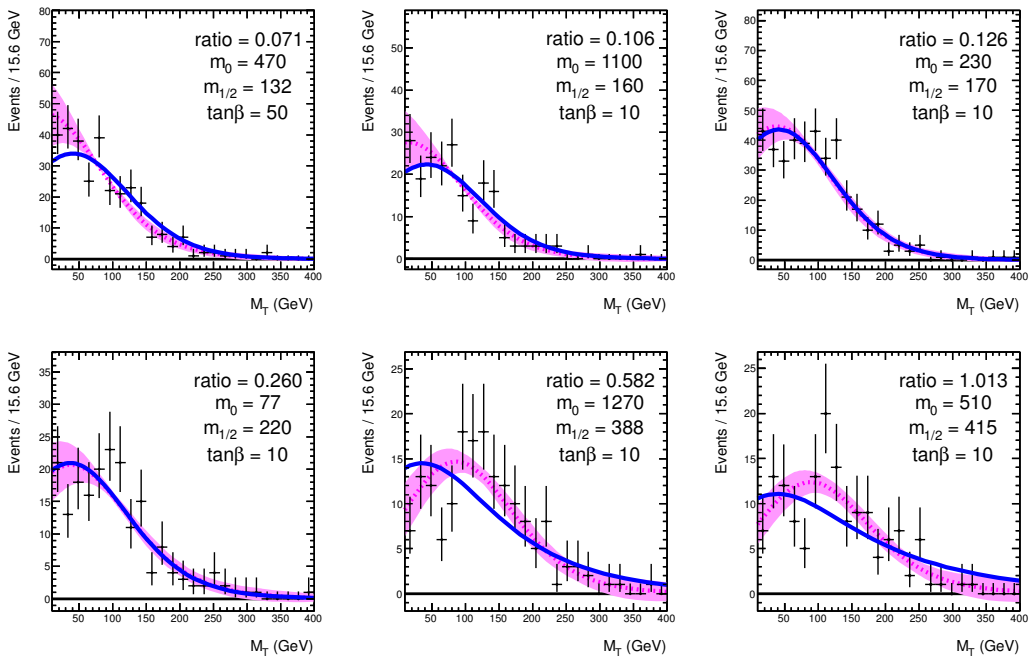
model. As discussed earlier, the technical difficulty of plotting a projection of a L-shaped region in a single observable gives small deviations between the shown $E_{\mathrm{T}}^{\mathrm{miss}}/M_{\mathrm{T}}$ distributions and the fitted 3D model.

In the SUSY analyses, there are two possible uses for the Ansatz model: the first one, as used in the combined fit method, is when the parameters of the Ansatz model are floated in a fit to the L-shaped control region together with the SM background parameters. Here the Ansatz SUSY model serves to keep the SM background measurement free of bias due to SUSY contamination. Only the SM backgrounds are then extrapolated into the signal region. The Ansatz model makes no assumptions about the shape or abundance of SUSY events in the signal region.

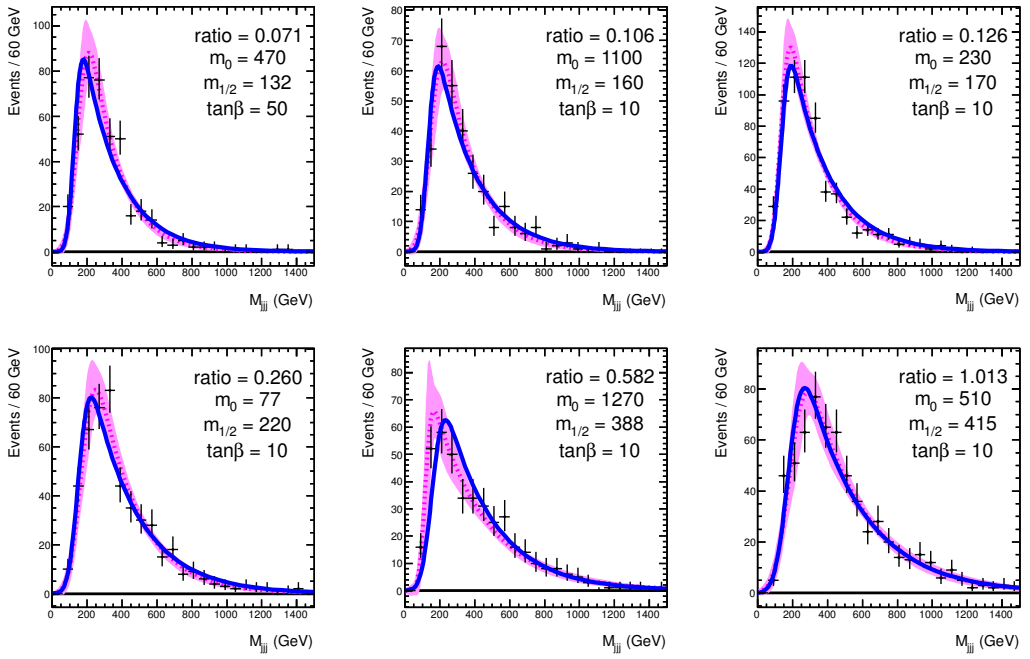
A second possible use of the Ansatz model, is where the shape parameters of the Ansatz model are constrained to values consistent with the $\frac{N_{\mathrm{SIG}}}{N_{\mathrm{CTL}}}$ *ratio*, estimated by the fit after extrapolation. This can be achieved by describing the correlations of the Ansatz model parameters with the *ratio*, shown in Figure 4.31. The estimated *ratio* after extrapolation compared to the expected *ratio* from the Ansatz shape can be added to the likelihood as a Poisson term, to be able to fit the SM backgrounds and the Ansatz model together in one iteration. However this second use of the Ansatz model is only meaningful, when SUSY is actually observed in the data.

4.4 Results

In this section we show that the combined fit method accurately estimates the parameters of the combined model in the control region and gives the correct yields when extrapolated to the



(a)



(b)

Figure 4.30: M_{T} distributions in $E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV region (a) and M_{jjj} distributions in the L-shaped control region (b) for six representative mSUGRA grid points sorted by increasing *ratio*. The lowest *ratio* is in top left, while the highest is in the bottom right. For each mSUGRA point on top of the data (markers) is the original 9-parameter fit (dotted curve) with corresponding propagated three standard deviations error (light band), and the 2-parameter fit (solid curve).

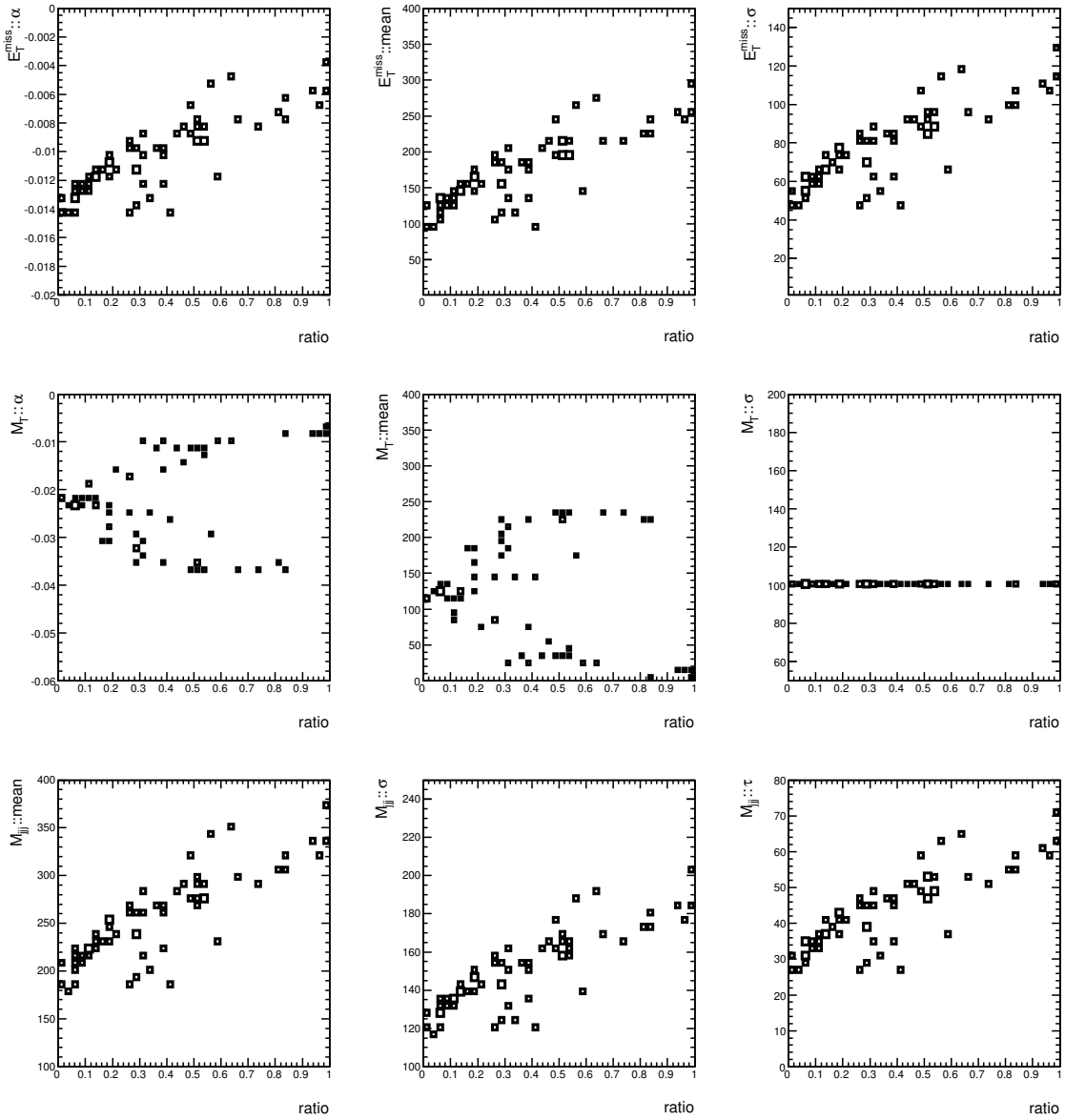


Figure 4.31: Distributions of the 9 Ansatz model parameters versus *ratio* for all mSUGRA grid points. The correlations suggest a Poisson term in the likelihood comparing the extrapolated *ratio* to the expected *ratio* could help assess the Ansatz parameters.

signal region, on a combined simulated (pseudo) data sample with an integrated luminosity of 1 fb^{-1}. The goal of the combined fit method is to have a fully data-driven estimation technique of the SM backgrounds, where none of the model parameters are taken from Monte Carlo (MC) simulation, but are constrained by the fit.

Because the full model is quite involved, we show the results of the fits in steps of increasing complexity as follows:

1. First we estimate the shape parameters from MC samples.

2. Then we fit the background model with fixed shapes to a combined MC sample containing

only the SM backgrounds.

3. Next we float a few shape parameters, that completes the TP/TC separation.
4. Then we show some cross checks performed on the combined data sample with and without SUSY.
5. As the next step we show results on extrapolation to the signal region for the model with and without SUSY.
6. Finally we describe the model with as many shape parameters as possible floating freely in the fit procedure.
7. This is followed by results on the extrapolation of the complete model (with SUSY and maximal floating shapes) to the signal region and the possible significance reach in SUSY phase space.
8. Then we show results of 'toy' MC studies performed to assess a possible bias.
9. We end by estimating the generator dependence of the combined fit method and the systematics due to detector effects.

SUSY reference points

Although we show results on the significance reach of the combined fit method for all the simulated SUSY samples, mentioned in Section 4.1.1, for simplicity the rest of the results are shown only for three mSUGRA *reference points*, representative of the mSUGRA grid. These points are chosen for their different values of *ratio* (defined as $N_{\mathrm{SIG}}/N_{\mathrm{CTL}}$ in Section 4.3.6) and their different SUSY yields in the signal region. The three reference points have the following parameters:

- Point 1: $m_0 = 1100$, $m_{1/2} = 160$, $N_{\mathrm{SIG}}/N_{\mathrm{CTL}} = 24/208$
- Point 2: $m_0 = 77$, $m_{1/2} = 220$, $N_{\mathrm{SIG}}/N_{\mathrm{CTL}} = 94/397$
- Point 3: $m_0 = 91$, $m_{1/2} = 300$, $N_{\mathrm{SIG}}/N_{\mathrm{CTL}} = 45/77$

4.4.1 Initial estimate of the shape parameters

The full combined model of SM backgrounds and SUSY contamination that was used for the fits is described in Appendix A in terminology of the `RooFit` framework, that was used to perform the study.

Given the complexity of the full model, initial values of the model parameters are determined from a fit to the simulated samples, that we call the *prefit* stage. The three SM backgrounds TP, T2 and CW are fitted simultaneously, each to their corresponding simulated sample in the full phase space region $40 < E_{\mathrm{T}}^{\mathrm{miss}} < 900 \bigwedge 10 < M_{\mathrm{T}} < 450$. The simultaneous procedure is applied, because of joint or global parameters that need to be estimated from multiple background samples. The SUSY model is fitted separately, since it is only defined in the L-shaped control region $E_{\mathrm{T}}^{\mathrm{miss}} < 200 \bigvee M_{\mathrm{T}} < 150$.

The result of the SM backgrounds prefit is shown in Figure 4.32 as 1D projections of the 3D models. The SUSY prefit on reference point 2 is shown in Figure 4.33, separately for each leg of the L-shaped control region. Both the SM backgrounds and the SUSY reference point are correctly described by the models.

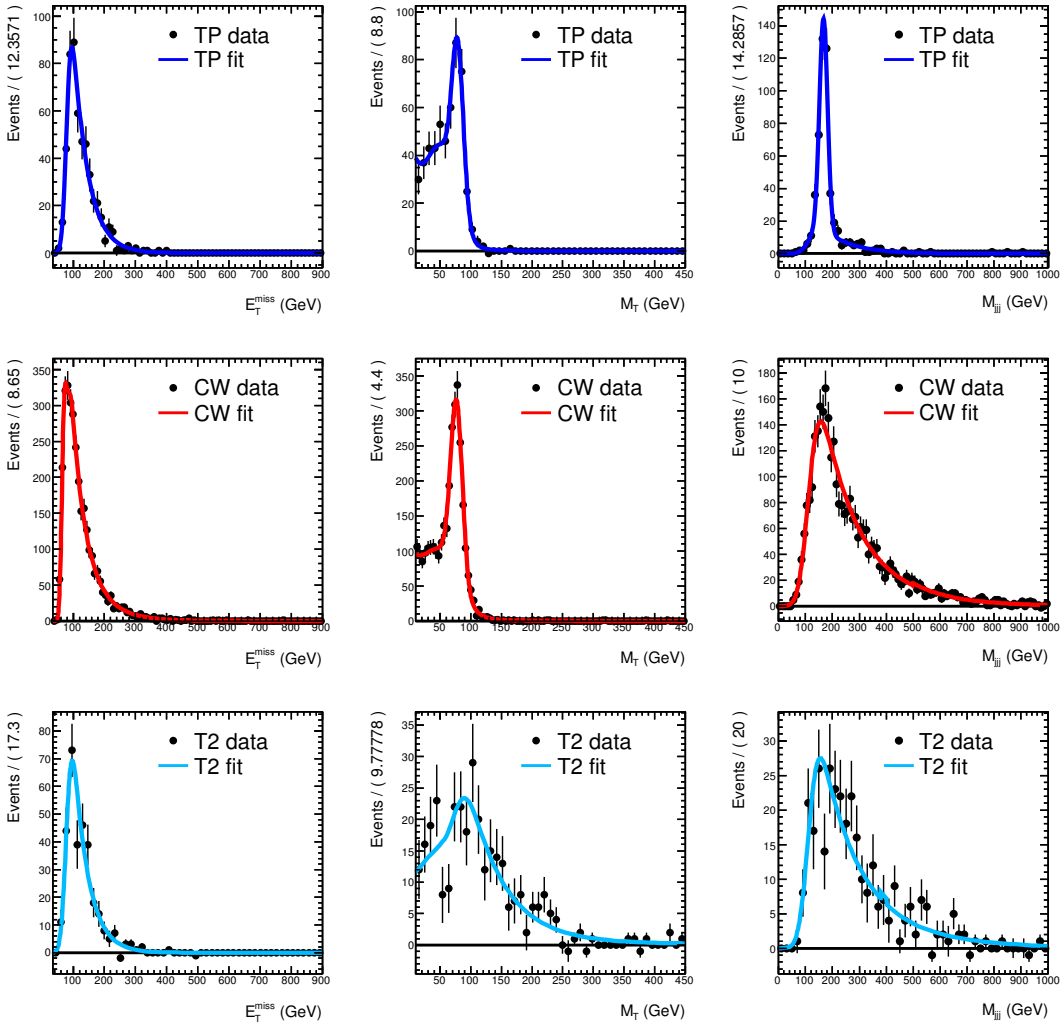


Figure 4.32: The prefit result for the backgrounds. For each background this figure shows three 1D projections of the 3D PDF. The three 3D models are fitted simultaneously to their respective background sample.

4.4.2 Combined background fit with fixed shapes

After getting an estimate of all the shape parameters, we add the three background models to form one combined model, as discussed in Section 4.3.1. The first test of this model is to leave all shape parameters constant and float only the yields of the three background contributions in a fit to the combined MC sample without SUSY. If our model is correct, the fitted yields should be comparable to the true yields, as the same MC samples are used in this fit as in the prefit stage.

The results of the simplest combined fit are shown in Table 4.9, where the yields as obtained from the fit are compared to the true number of events in each background sample. The fit is done twice: once in the full phase space region and once in the L-shaped control region. As can be seen from Table 4.9, the fitted yields are correct within approximately one standard deviation.

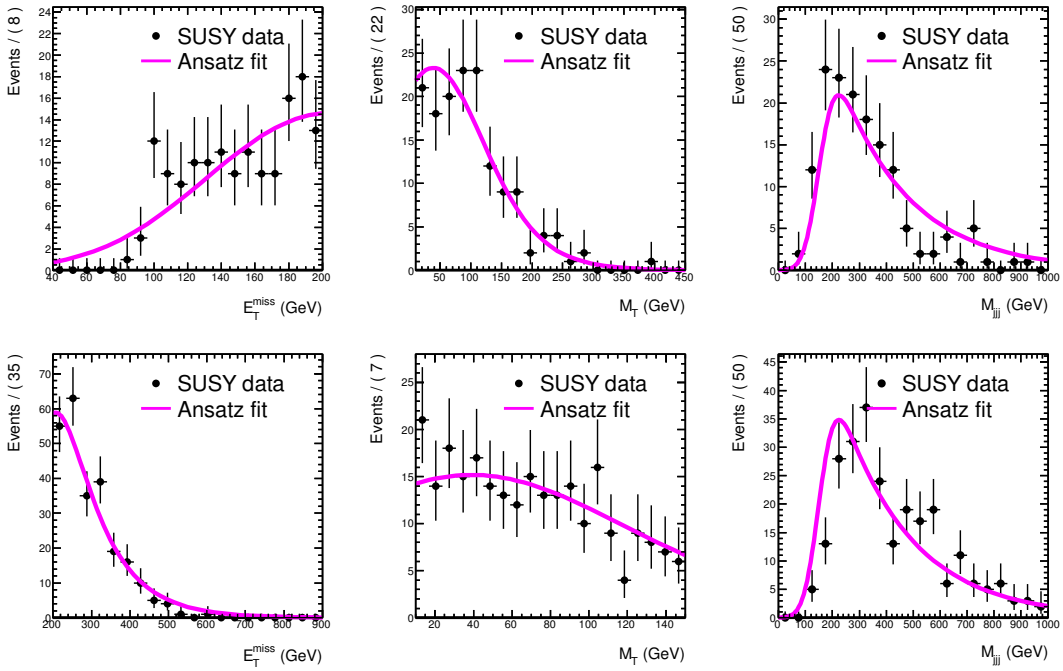


Figure 4.33: The prefit result for the reference SUSY grid point 2. The three top plots are for the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ leg of the L-shaped region ($E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV), while the bottom plots are for the low-M_{T} leg ($M_{\mathrm{T}} < 150$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV).

Sample name	Fitted Yield	True Yield
Full region $40 < E_{\mathrm{T}}^{\mathrm{miss}} < 900 \bigwedge 10 < M_{\mathrm{T}} < 450$		
TP	566 ± 51	514
T2	320 ± 32	311
CW	3046 ± 60	3107
Control region $E_{\mathrm{T}}^{\mathrm{miss}} < 200 \bigvee M_{\mathrm{T}} < 150$		
TP	565 ± 51	514
T2	288 ± 32	299
CW	3065 ± 59	3105

Table 4.9: Fitted yields using the combined model with fixed shapes in the full phase space region and in the L-shaped control region.

4.4.3 Floating shape fit to complete the TP/TC separation

The splitting of the semileptonic $t\bar{t}$ sample in TP/TC components is not perfect, as was discussed in Section 4.3.5, because some TC events are still present in the TP sample and vice versa. To complete the TP/TC separation in the combined fit, we float the M_{jjj} parameters of the CW and TP models, since these are expected to change with respect to the prefit. Besides floating the M_{jjj} shapes, we manually set the TTComb (TC contamination in TP sample) fraction to zero in the TP model. This setup of the model parameters in the fit we call the *minimal floating shapes* fit.

Sample name	Fitted Yield	True Yield
Control region $E_{\mathrm{T}}^{\mathrm{miss}} < 200 \bigvee M_{\mathrm{T}} < 150$		
TP	412 ± 54	400
T2	293 ± 31	299
CW	3213 ± 61	3219

Table 4.10: The results of the fit in the L-shaped control region with TP and CW M_{jjj} shapes floating, also called the minimal floating shapes fit in the text.

Table 4.10 shows the combined fit results with minimal floating shapes performed in the L-shaped control region on the combined MC sample without SUSY. The minimal floating shape fit correctly estimates the SM backgrounds, as the fitted yields are within the error margins of the true yields.

When comparing the results in Table 4.10 to those in Table 4.9, note the fact that for the minimal floating shapes fit we set the true yields at different values. From the study in Section 4.3.5 we have an estimate of the true number of correctly reconstructed top (f_{TP}^{total}) events and combinatorics background (f_{TC}^{total}) events in the semileptonic $t\bar{t}$ sample. Our choice of $\Delta R = 0.1$ gives the following fractions: $f_{TP}^{\mathrm{total}} = 0.1915$ and $f_{TC}^{\mathrm{total}} = 0.8085$. These fractions are used to calculate the true yields following $N_{true}^{TP/TC} = N_{true}^{T1} * f_{TP/TC}^{\mathrm{total}}$, where N_{true}^{T1} is the total number of events in the semileptonic $t\bar{t}$ sample.

The variation of $f_{TP/TC}^{\mathrm{total}}$ with ΔR is approximately $\sigma_{f_{TP/TC}^{\mathrm{total}}} \pm 0.0158$, which we estimate by fitting a horizontal line to the distributions of Figure 4.24(b). This translates in a variation of ± 33 events for the true yields $N_{true}^{TP/TC}$, which is much smaller than the variation of true yields between Table 4.10 and Table 4.9. Besides that, a variation of 33 events is small compared to the statistics of the samples, hence we ignore this uncertainty in the next sections.

4.4.4 Estimating the SM backgrounds in the presence of SUSY

The first aim of the combined fit method is to correctly estimate the SM backgrounds in the L-shaped control region, before extrapolating them to the signal region. However if SUSY events contaminate the control region, the combined model describing only the SM backgrounds is inaccurate, thus we add the SUSY Ansatz PDF to the combined model. On the other hand adding the SUSY PDF must not interfere with SM background estimation, or in other words we do not want to find a false positive if there is no SUSY present in data.

The results of the four cross checks are listed in Table 4.11, where all the fits are done in the L-shaped control region. The minimal floating shapes setup is used, and wherever applicable we also float the SUSY Ansatz PDF parameters.

First off we repeat the fit of the previous section, where the model of the backgrounds only is fitted to a MC sample of backgrounds only. The yields are estimated correctly. The next step is to add the SUSY PDF to the model, but leave the backgrounds only MC data sample untouched, thus studying whether we see a false positive. The fitted yield of SUSY in this cross check is 0±8, while the background yields are almost identical to the first fit. We conclude that in the absence of SUSY in data, the combined fit with the SUSY Ansatz PDF included will find the correct background yields and will not falsely discover SUSY.

The next cross check is to see what the combined fit with a backgrounds only model does, when it is fit to a data sample with SUSY. As expected this fit returns incorrect yields for the

Sample name	Fitted Yield	True Yield
Fit with no SUSY in data, no SUSY in model		
TP	412 ± 54	400
T2	293 ± 31	299
CW	3213 ± 61	3219
Fit with no SUSY in data, SUSY in model		
TP	411 ± 54	400
T2	293 ± 31	299
CW	3213 ± 61	3219
SU	0 ± 8	0
Fit with SUSY in data, no SUSY in model		
TP	326 ± 52	400
T2	455 ± 35	299
CW	3137 ± 61	3219
Fit with SUSY in data (reference point 1), SUSY in model		
TP	383 ± 57	400
T2	309 ± 48	299
CW	3278 ± 67	3219
SU	157 ± 48	208
Fit with SUSY in data (reference point 2), SUSY in model		
TP	379 ± 53	400
T2	321 ± 41	299
CW	3255 ± 67	3219
SU	360 ± 39	397
Fit with SUSY in data (reference point 3), SUSY in model		
TP	405 ± 53	400
T2	318 ± 34	299
CW	3212 ± 63	3219
SU	61 ± 17	77

Table 4.11: The cross checks of fitting the combined model with and without a SUSY PDF on a MC data sample with and without SUSY in the L-shaped control region. As expected only the fit of the model without a SUSY PDF on a MC data sample with SUSY returns an incorrect result, hence we conclude that you must add the SUSY PDF to the combined model to accurately assess the SM backgrounds in the control region. All the other permutations of model versus data sample show correctly estimated yields.

SM backgrounds. Specifically the T2 component gets overestimated, as it is the only background with high $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} tails somewhat like SUSY models.

The final cross check is whether the fit of the model with SUSY Ansatz included, performed on a MC data sample also with SUSY will find the correct yields. Table 4.11 shows results of this cross check for the three mSUGRA grid reference points. For all the reference points the

SM background yields as well as the SUSY yield are correctly estimated, within error margins. We conclude that the combined fit method correctly estimates the SM background and SUSY yields in the L-shaped control region, if the model used in the fit includes the SUSY Ansatz PDF.

4.4.5 Combined fit with extrapolation

Having shown that we estimate the yields correctly in the control region, the next step for the combined fit validation is to show that we extrapolate the SM backgrounds to the signal region accurately. In this section we show the extrapolation performance of the combined fit with the backgrounds only model on a MC data sample of SM backgrounds, and also with the model containing the SUSY PDF fit to a MC data sample with mSUGRA reference point 2. Both fits are performed with the minimal floating shapes setup, and the two SUSY parameters floating if SUSY PDF is included.

Table 4.12 shows the results of the fits *and* extrapolation. The extrapolation yields very accurate results for the estimated background contributions in the signal (SIG) region, both with and without SUSY in the model (data sample). This shows that our model estimates both the yields *and* the shapes of the backgrounds correctly in the control region, in the absence and in the presence of SUSY.

The quoted errors on the yields in the signal region, in Table 4.12 and in the rest of this chapter, are statistical errors on the extrapolation of the model from the fit result in the control region only. They do not include the Poisson errors on the event counts in the signal region. The quoted error on the SUSY yield in the signal region is associated to the statistical errors of the extrapolated background yields only. The SUSY model is only used in the control region to correctly estimate the background yields, but it is not extrapolated, as we do not want to make assumptions about the shape of the signal in the signal region.

Sample name	Fitted $\text{Yield}_{\text{SIG}}$	True $\text{Yield}_{\text{SIG}}$
Fit with no SUSY in data, no SUSY in model		
TP	0.12 ± 0.003	0
T2	12 ± 0.2	12
CW	1.9 ± 0.2	2
Fit with SUSY in data, SUSY in model		
TP	0.06 ± 0.011	0
T2	12 ± 0.18	12
CW	2 ± 0.17	2
SU	94 ± 0.11	94

Table 4.12: The results of the fit in the control region extrapolated to the signal region (SIG), with TP and CW M_{jjj} shapes floating, as well as the SUSY PDF parameters when applicable.

4.4.6 Combined fit with maximal floating shapes

The goal of the combined fit method is to fit all the shapes and yields from data, so that we rely on MC simulation as little as possible. The full model however has too many parameters to

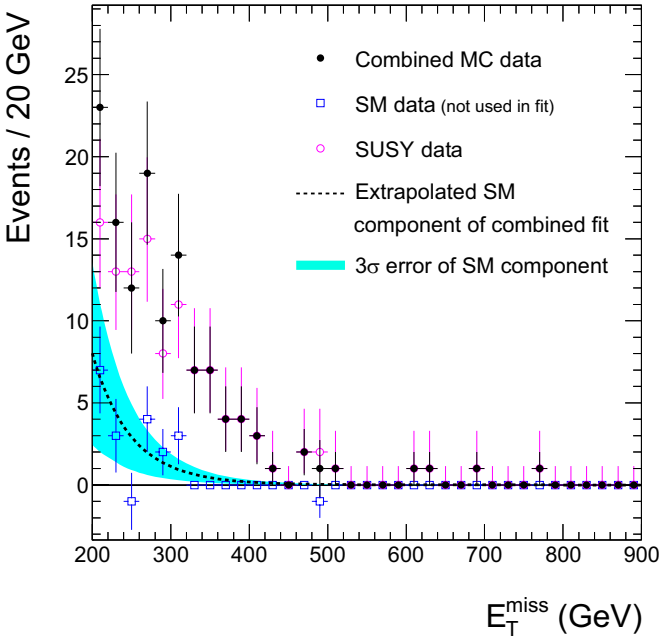
Sample name	Fitted Yield$_{\mathrm{CTL}}$	True Yield$_{\mathrm{CTL}}$	Fitted Yield$_{\mathrm{SIG}}$	True Yield$_{\mathrm{SIG}}$
		Reference point 1		
TP	428 ± 61	400	0.03 ± 0.092	0
T2	406 ± 68	299	13 ± 0.56	12
CW	3113 ± 87	3219	0.9 ± 0.58	2
SU	179 ± 52	208	23 ± 0.14	24
				Significance = 4.3
		Reference point 2		
TP	395 ± 56	400	0.02 ± 0.072	0
T2	365 ± 64	299	12 ± 0.66	12
CW	3174 ± 88	3219	1 ± 0.68	2
SU	380 ± 42	397	94 ± 0.14	94
				Significance = 13
		Reference point 3		
TP	448 ± 61	400	0.03 ± 0.079	0
T2	392 ± 55	299	13 ± 0.59	12
CW	3088 ± 86	3219	1 ± 0.6	2
SU	67 ± 19	77	44 ± 0.11	45
				Significance = 7.7

Table 4.13: The results of the combined fit in the control region for the three mSUGRA reference points, with the largest possible subset of parameters floating. The estimated and true yields are shown for the control region (CTL) as well as extrapolated to the signal region (SIG). The errors quoted are the statistical errors on the extrapolation from the fit. Note that the Poisson error on the event count in the signal region is not included. Significance is calculated in the signal region taking Poisson errors into account, as explained in the text.

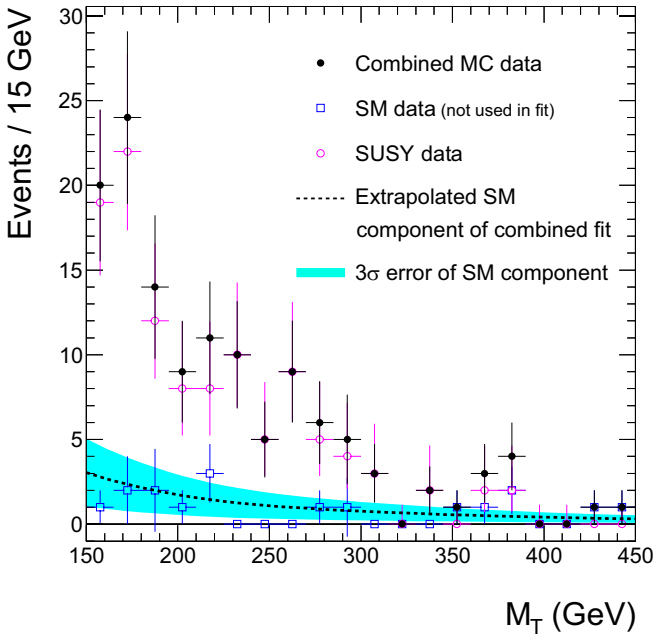
float them all without destabilizing the fit, but we float as many as possible while keeping the fit stable, which we call the *maximal floating shapes* setup.

Determining which parameters can be released in the fit requires careful handling. To avoid biased results we do not want to introduce flexibility in one component, while keeping the shape of another fixed. We determine what set of floating parameters is acceptable by looking at the relative error of the estimated SUSY yield after the extrapolation. As the same MC samples are used for the initial parameter estimates in the prefit as in the combined fit, the estimated SUSY yield has the smallest error after the fit with the minimal floating shapes and floating the two SUSY PDF parameters. Parameters are kept fixed, if floating them increases the error by a factor $\gtrsim 5$ compared to the minimal floating shapes. Besides that, we require that the fit converges with the extra parameter left floating for the standard MC data samples, as well as the sample where MC@NLO generated $t\bar{t}$ is replaced by the ALPGEN generated $t\bar{t}$ samples. The final set of parameters floated in the maximal floating shapes fit is shown in Appendix B. Out of the 36 shape parameters in the combined fit model, 23 are left floating in the maximal floating shapes fit, while many of the fixed parameters are concentrated in the T2 model, more specifically in the M_{T} observable.

In Table 4.13 we quote the fit results for the three mSUGRA reference points, while we float



(a)



(b)

Figure 4.34: The extrapolation to the signal region of the background component of the model after fitting the complete model in the L-shaped control region, with 3σ error band in both $E_{\mathrm{T}}^{\mathrm{miss}}$ (a) and M_{T} (b). As reference the respective SM and SUSY simulated data components are shown, but they are not used in the fit. The overlap between the SM data and the extrapolated model validates the complete procedure.

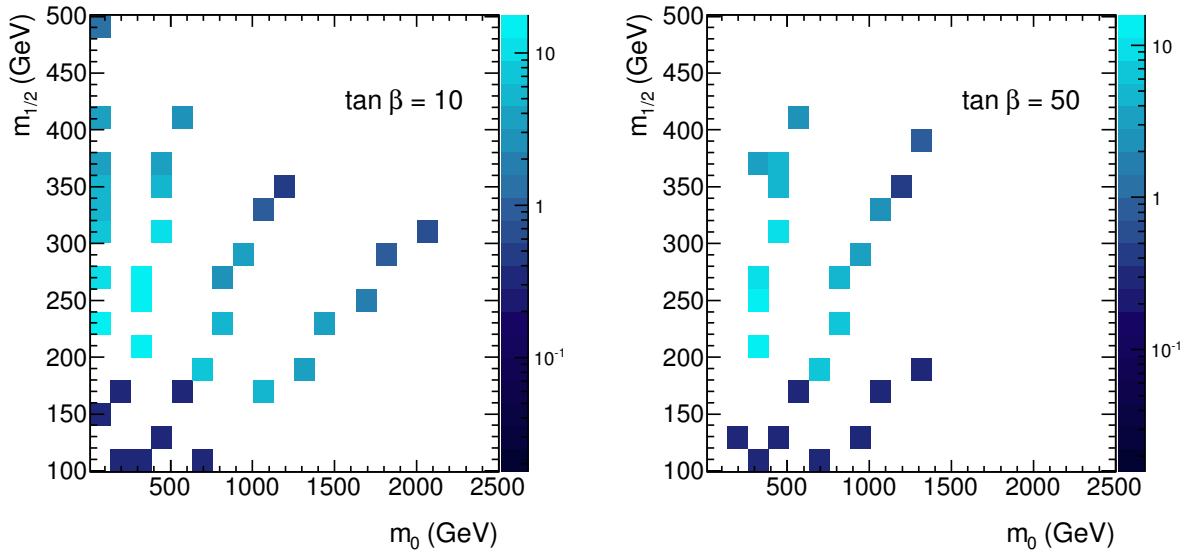


Figure 4.35: Significance that can be obtained with the combined fit method as a function of m_0 and $m_{1/2}$, for $\tan\beta = 10$ (left) and $\tan\beta = 50$ (right).

the largest set of shape parameters possible. The yields shown in Table 4.13 are quoted in both, the control region (CTL) and after extrapolation in the signal region (SIG). The fitted yields for the SM backgrounds and for SUSY are in good agreement with the true yields for all the three mSUGRA reference points.

Also quoted in Table 4.13 is the *significance* of detecting the SUSY signal over the estimated background for the observed number of events in the signal region. Significance is calculated using the tools provided by the `RooStats` framework [104, 105], where we approach the calculation as if it is purely a counting experiment in the signal region. The problem is treated in a fully frequentist fashion by interpreting the relative background uncertainty as being due to the auxiliary control observation, or fit result in the control region, while the number of observed events and the number of expected background events are distributed as Poissons. It can be expressed as $\mathrm{Poisson}(N_{\mathrm{SIG}}|\mathrm{s{+}b})\cdot\mathrm{Poisson}(N_{\mathrm{CTL}}|\tau\mathrm{b})$, where N_{SIG} is the measured number of events in the signal region distributed as a Poisson around $s{+}b$ (signal+background), and N_{CTL} is the estimated number of background events in the control region distributed as a Poisson around τb, where τ is the scale factor used in the extrapolation of the background b to the signal region. Thus we estimate significance by looking at the probability that an observed event count in the signal region, is produced solely by fluctuation of the background, where significance is given in terms of equivalent standard deviations of the normal distributions [106].

The last (and possibly most convincing) result of the combined fit method, the extrapolated SM background model into the signal region is shown in Figure 4.34. It shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} distributions of the MC data sample (in black filled circles) and the SM and SUSY data separately (in magenta empty circles and blue empty squares respectively), together with the background model extrapolated to the signal region. Besides the extrapolated model, also the uncertainties on the extrapolated model with three standard deviations are shown by the light blue band. We conclude that the extrapolated model, fitted in the control region, describes both the integrated event count as well as the shape of SM backgrounds in the signal region correctly. As a future improvement, this can be used to enhance the significance reach of the combined fit

method, as we could calculate the significance of the SUSY excess for every bin in Figure 4.34 and add these together, rather than treating the entire signal region as a single bin.

4.4.7 Significance reach in mSUGRA phase space

Figure 4.35 shows the estimated significance of the excess, found by the combined fit method with maximal floating shapes after extrapolation, for all points in the mSUGRA phase space for which we have a simulated sample. The mSUGRA points expected to yield high significance are the points that have a low enough mass scale to be copiously produced at the LHC, but not so low that all SUSY events end up in the control region. This can be clearly seen in Figure 4.35 as for low m_0 and $m_{1/2}$ the estimated significance is below 1, rising with increasing scalar (m_0) and fermion ($m_{1/2}$) masses well above the discovery potential of significance equal to 5. Until for higher values of m_0 and $m_{1/2}$, the production cross section and consecutively the number of SUSY events in the signal region drops down to levels that cannot be measured with enough significance at the chosen integrated luminosity.

4.4.8 Closure test

We perform a closure test of the combined fit method by comparing the estimated number of SUSY events ($\mathrm{N}^{\mathrm{fit}}_{\mathrm{SUSY}}$) to the true number of SUSY events ($\mathrm{N}^{\mathrm{true}}_{\mathrm{SUSY}}$) for each simulated mSUGRA point. The difference between the two is divided by the error on the estimate ($\sigma_{\mathrm{N}^{\mathrm{fit}}_{\mathrm{SUSY}}}$). As we are sampling from different mSUGRA points, this is not a true estimation of bias as performed the next section, but it shows us how well we estimate the number of SUSY events, and thus if the calculated significance of the method is correct.

The distribution for the estimated number of SUSY events in the control region ($\mathrm{N}^{\mathrm{fit}}_{\mathrm{SUSY,CTL}}$) is shown in Figure 4.36(a), and in the signal region ($\mathrm{N}^{\mathrm{fit}}_{\mathrm{SUSY,SIG}}$) in Figure 4.36(b). To guide the eye a unit Gaussian centered at zero is drawn in both figures.

Also quoted are the *mean* and the root mean square (*rms*) of the distributions in Figures 4.36(a) and 4.36(b). The *mean* of the signal region distribution shows that $\mathrm{N}^{\mathrm{fit}}_{\mathrm{SUSY,SIG}}$ is slightly overestimated, however the *rms* tells us that the overestimation is on average only half of the estimate error. As we have only ∼60 mSUGRA points simulated a correct estimation of a bias is statistically limited, thus we validate the fit procedure by running toy Monte Carlo studies.

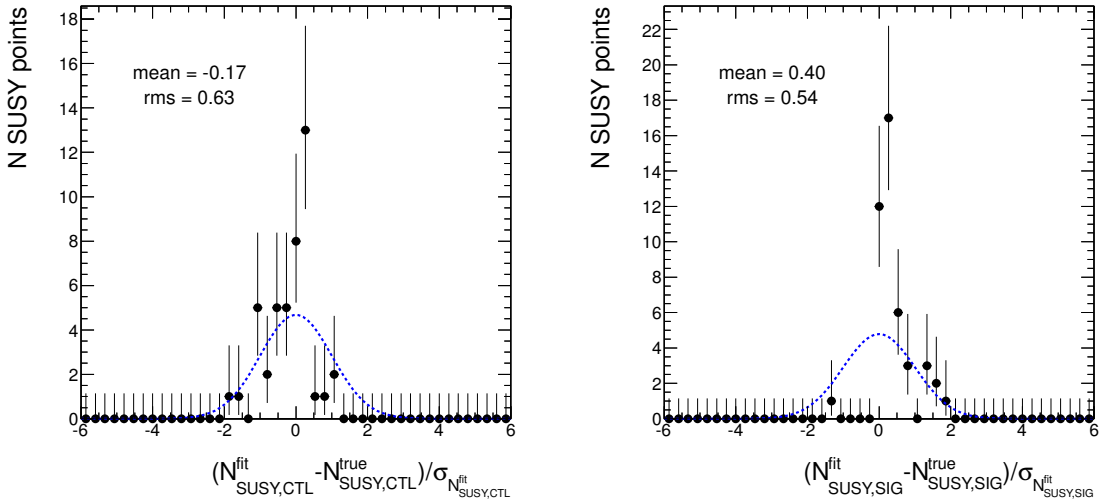
4.4.9 Validation of the fit procedure

To verify the absence of bias in the combined fit method and to study the correctness of the quoted errors, we run a series of 'toy' Monte Carlo studies on the reference mSUGRA grid points. For each reference mSUGRA point we run 1000 experiments. We sample the shape of the combined PDF, after performing the minimal floating shapes fit, to generate simulated events. The generated data is fit to the combined fit model with minimal floating shapes[6].

The fraction of SUSY events (f^{fit}_{SU}) and the corresponding error ($\sigma_{f^{fit}_{SU}}$) that are estimated by the combined fit are used to calculate the *pull*, defined as:

$$\text{pull}\ (f_{SU}) = \frac{f^{fit}_{SU} - f^{true}_{SU}}{\sigma_{f^{fit}_{SU}}}. \tag{4.53}$$

[6]For the lack of available computing power we cannot use the maximal floating shapes model.



(a) (b)

Figure 4.36: Pull distributions for the estimated number of SUSY events in the control region (a) and in the signal region (b) for all simulated mSUGRA samples. To guide the eye a unit Gaussian centered at zero is drawn.

If the fit is unbiased and the parameter error estimate is correct, the pull distribution has an rms of 1 and is centered at 0, or in other words it is a unit Gaussian. Maximum likelihood estimators using simple PDFs and medium or large event counts are generally unbiased. If low event yields are involved, this is not generally the case as the bias is $\propto \frac{1}{N}$, while the error on the measurement is $\propto \frac{1}{\sqrt{N}}$. Hence one typically must worry about biases when N is small [107, 108], as is the case for the combined fit method, where the yield of SUSY is quite small.

Besides small N effects, boundaries on the model parameters can be a separate source of bias, that can even happen at high N. Setting a boundary on a parameter means that MINUIT [109], the program used by `RooFit` for minimization of the log likelihood, transforms the parameter with an arcsin to map the parameter internally to an open domain. Thus internally the boundaries are approached by an asymptotic shape.

For example in the combined fit model we define the fractions of 3D PDFs recursively, because we try to keep the yields physical, hence fractions are bounded between 0 and 1. As the fraction approaches 0 in the fit procedure, the low-side error becomes smaller than the high-side error, leading to the probability of an upward fluctuation becoming greater than the probability of a downward one, when asymmetric errors are calculated. This can give rise to a small overestimation of the fraction and an asymmetric pull distribution.

Figure 4.37(a) shows the f_{SU}^{fit} distribution of 1000 toy experiments fit to the combined model with minimal floating shapes in the control region for mSUGRA reference point 2, while Figure 4.37(b) shows the corresponding $\sigma_{f_{SU}^{fit}}$, and Figure 4.37(c) shows the resulting pull distribution. Also shown in Figure 4.37(c) is a Gaussian function fit to the pull distribution. The corresponding Gaussian *mean* and σ are quoted, as well as the *mean* and the *rms* of the pull distribution. The values of the Gaussian function parameters are in good agreement with the pull distribution values.

Figure 4.37(c) shows that there is a small positive bias in the estimate of f_{SU}^{fit}, as the pull *mean* is significantly away from 0. The pull distribution is slightly asymmetric, caused by the asymmetric distributions in Figures 4.37(a) and 4.37(b), but the Gaussian fit displayed in Fig-

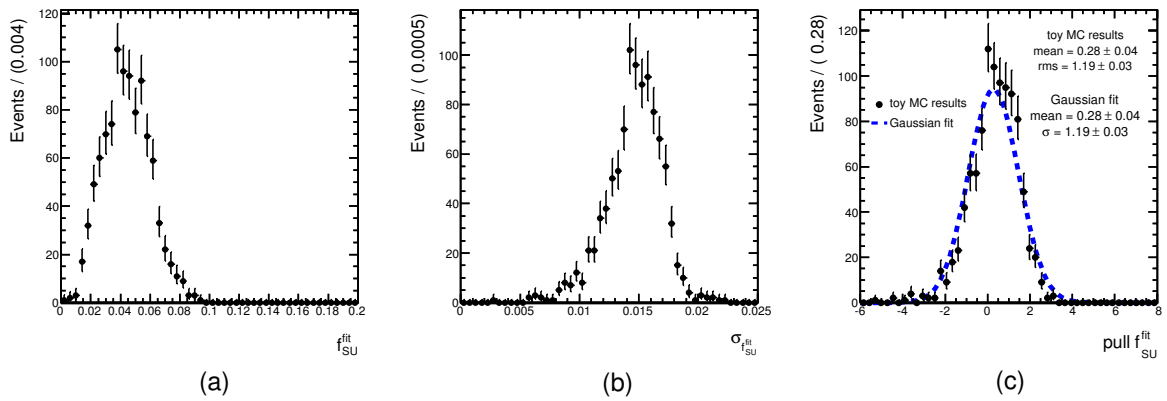


Figure 4.37: The fitted fraction of SUSY events (a), the error on the fitted SUSY fraction (b) and the pull distribution of the fitted SUSY fraction (c) for the mSUGRA reference point 2. A Gaussian function fit to the pull distribution is shown as the dashed curve. The *mean* and the *rms* of the pull distribution are quoted, as well as the Gaussian function parameters *mean* and σ, with respective errors.

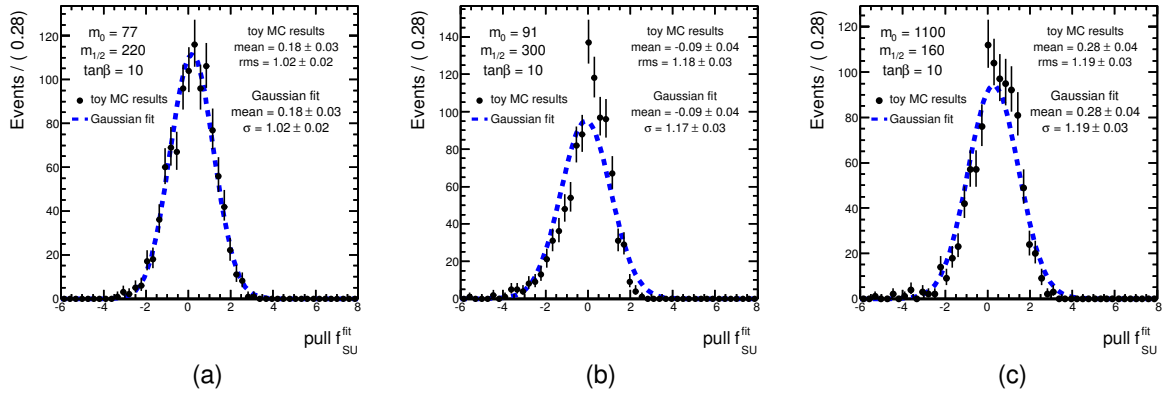


Figure 4.38: The pull distributions of f_{SU}^{fit} for the three mSUGRA reference points. The values of Gaussian function fit (dashed curve) parameters are quoted, as well as the pull *mean* and *rms*. A small bias is observed in the estimation of f_{SU}^{fit} for the combined fit method.

ure 4.37(c) describes of the pull distribution reasonably well. The pull *rms* shows us that the bias is only $\sim 20\%$ of the estimated error $\sigma_{f_{SU}^{fit}}$.

To make sure that it is not only the mSUGRA reference point 2 that is somewhat overestimated, in Figure 4.38 we show the pull distributions of 1000 toy experiments for all three mSUGRA reference points. There is a small bias in the estimation of the SUSY fraction for the reference points, however this bias is on average only $\sim 20\%$ of the estimated error. We conclude that although the estimation of the SUSY yield by the combined fit method is slightly positively biased, the bias is only a fraction of the estimated error.

Generator independence

An important test of the combined fit method is whether it is dependent on the generator used in the prefit stage. To test this, we do the prefit on the data sample with $t\bar{t}$ events generated by MC@NLO, and fit this model on a data sample with $t\bar{t}$ events generated by ALPGEN. We use

Sample name	Fitted Yield$_{\mathrm{CTL}}$	True Yield$_{\mathrm{CTL}}$	Fitted Yield$_{\mathrm{SIG}}$	True Yield$_{\mathrm{SIG}}$
MC@NLO prefit model, Alpgen $t\bar{t}$ in data sample				
TP	261 ± 46	255	0.002 ± 0.028	0
T2	361 ± 59	328	13 ± 0.46	10
CW	2992 ± 85	3046	0.8 ± 0.47	2
SU	413 ± 42	397	92 ± 0.14	94
			Significance = 12	
Alpgen prefit model, MC@NLO $t\bar{t}$ in data sample				
TP	343 ± 51	400	0.01 ± 0.041	0
T2	391 ± 57	299	11 ± 0.53	12
CW	3254 ± 80	3219	1 ± 0.54	2
SU	327 ± 42	397	95 ± 0.12	94
			Significance = 14	

Table 4.14: The results of the combined fit tested on data with a different $t\bar{t}$ generator, on a data sample created using reference point 2. The test with the ALPGEN prefit model comes out slightly worse in the control region, but still performs well after extrapolation. The errors quoted for the signal region are the propagated errors of the extrapolation only, and do not include the Poisson errors of the event count in the signal region.

the maximal floating shapes setup, so that most of the parameters of the model are constrained from the fit. If our model is indeed generator independent, it should be able to correctly estimate the yield of ALPGEN $t\bar{t}$ events. Unfortunately we do not have an alternative $W + jets$ generator available.

The results of the combined fit on the MC data sample with ALPGEN $t\bar{t}$ events are shown in Table 4.14. The yields in the control region and the signal region are estimated correctly. To get the proper true yields, we redo the TP/TC splitting technique in semileptonic $t\bar{t}$ for the ALPGEN samples, as described in Section 4.3.5.

Table 4.14 also shows the reverse test, where we start with a prefit on the ALPGEN $t\bar{t}$ sample, but the combined fit is performed on a combined data sample with MC@NLO $t\bar{t}$, while floating the largest possible set of shape parameters. The result of this combined fit in the control region shows a slight underestimation of the SUSY yield, due to an overestimation of the dileptonic $t\bar{t}$ yield, but in the signal region the yields are correct within error margins, as the shapes (and yields) extrapolated from the control region give the correct number of events.

4.4.10 Systematic uncertainties

Finally we estimate the impact of detector effects like misreconstruction of jets, leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ have on the power of the combined fit method. We vary the reconstruction parameters listed in Table 4.15 and rerun the analysis, where the parameter variation is taken as in [60]. The systematic uncertainties quoted in Table 4.15 are calculated as the variation of the measured SUSY cross section, defined as:

$$\sigma_{SUSY} = \frac{N_{SUSY,SIG}^{fit}}{\epsilon_{sel}} \ . \tag{4.54}$$

Source	Systematic ± statistical uncertainty [%]
Jet energy scale up 5%	11 ± 2.4
Jet energy scale down 5%	7.2 ± 2.6
Jet energy resolution 10% (relative)	12 ± 2.9
Muon energy scale up 0.2%	1.2 ± 0.8
Muon energy scale down 0.2%	0.72 ± 0.46
Muon energy resolution 1% (relative)	1.4 ± 0.94
Soft $E_{\mathrm{T}}^{\mathrm{miss}}$ scale up 10%	2.6 ± 1.1
Soft $E_{\mathrm{T}}^{\mathrm{miss}}$ scale down 10%	1.9 ± 1.5
MC@NLO vs. ALPGEN	2.6 ± 1.6

Table 4.15: Different sources of systematic uncertainty and their effect on the SUSY combined fit performance.

Again $N_{SUSY,SIG}^{fit}$ is the number of SUSY events in the signal region, after extrapolating the SM backgrounds. The event selection efficiency ϵ_{sel} is determined as the number of events after selection divided by the number of events before selection for reference point 2.

The number of events for each SM background in the signal region is dominated by small event counts. Hence the number of events that migrate to the signal region, due to the systematic check, may limit the precision of the estimate of the systematic uncertainty. Therefore the statistical uncertainty on the estimate of the systematic uncertainty is explicitly quoted in Table 4.15.

Variations in the jet energy scale have the largest effect on the outcome of the method, giving a systematic uncertainty around 10%, while the other systematic uncertainties are at the level of a few percent. This effect can be explained by the following two observations. Firstly SUSY event selection is strongly dependent on hard jet cuts, thus a variation of the jet energy effects the selection efficiency. Secondly also $E_{\mathrm{T}}^{\mathrm{miss}}$ depends strongly on the jet energy scale, thus fluctuating the energy of the jets has an effect on the number of events that migrate in and out of the signal region.

We include the generator dependence, discussed in the previous section, as a systematic uncertainty. However as the yield of SUSY events in the signal region is estimated correctly in Table 4.14, the resulting uncertainty is not very large.

4.5 Conclusion

In this chapter we have presented the *combined fit method* approach to a data-driven SUSY search with one lepton. Within the `RooFit` framework we have developed tools to deal with the necessary mathematics. Using these tools we have established a model which combines models for each relevant SM background, i.e. $t\bar{t}$ and $W + jets$ production, with a SUSY Ansatz model valid for most of the mSUGRA phase space. Through parameterization we have reduced the dependence on specific MC generators. Our method was first to address the SUSY contamination with an SUSY Ansatz model inside the ATLAS SUSY working group.

We have shown that the combined fit method is capable of finding the correct relative yields of the different background processes, and at the same time estimating the proper excess of observed events over the expected background yield in the signal region, also when SUSY contamination is present in the control region. In addition we checked that our method does not find SUSY

when no SUSY is present in the data. For an integrated luminosity of 1 fb^{-1} we have shown on simulated data, that the combined fit method can discover mSUGRA with a statistically significant excess for $m_{1/2}$ above approximately 175 GeV at all values of m_0.

4.6 Ideas for future improvements

The combined fit approach still has a partial dependence on simulation input, primarily because a limited number of shape parameters are fixed from simulated samples. It may be possible to reduce or perhaps altogether eliminate this dependence by fitting to additional sideband and control samples.

4.6.1 Using events with 3 jets as an additional control sample

SUSY events are characterized by multiple highly energetic jets. So if we want a control sample with reduced SUSY contamination, it makes sense to look at events with lower jet multiplicity. We separate our simulated samples in two jet multiplicity bins, exactly 3 jets versus 4 or more (4+) jets.

The added value of events with exactly 3 jets is shown by Figure 4.39, where shown are the normalized distributions in the three fit observables for the CW sample separated by jet multiplicity. As event selection requires three hard jets, the M_{jjj} distribution in the 3 jets bin is harder than in the 4+ jets bin. But the shapes of $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} distributions are very similar for 3 and 4+ jet multiplicities, as expected, because for the SM backgrounds in the one lepton search, $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} are at first order the result of leptonic W decay. The small discrepancies in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution are due to the different W/TC content of the CW sample. Although we only show the distributions for the CW sample, for the other background samples the shapes of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} distributions are also very similar for the two jet multiplicities.

The information on the shape of models contained in the 3 jet control sample can be used in an extended version of the combined fit method. A simultaneous fit in the two jet multiplicities can be performed, where many of the parameters describing the $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} shapes are shared. In the limit that all the shape parameters are shared between the two jet multiplicities, then we only have to add the three fraction parameters in the 3 jet bin to our combined fit. In practice some shape parameters in $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} must be split to account for differences between the

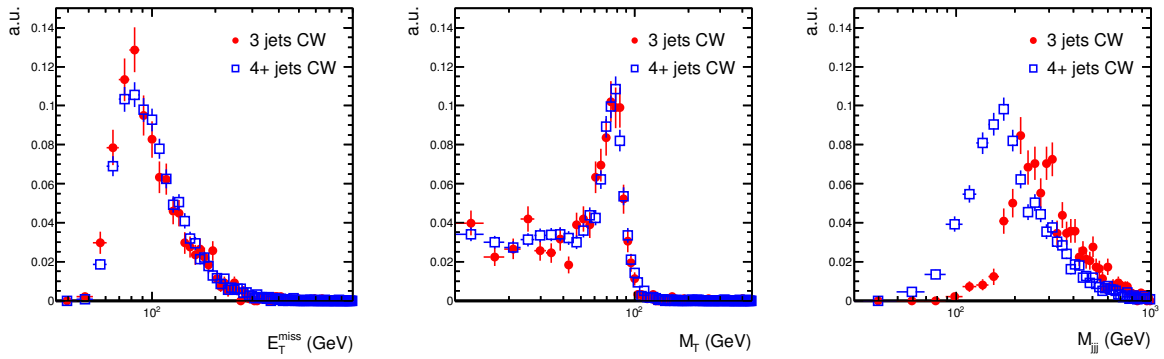


Figure 4.39: Normalized distributions for the CW sample subdivided by jet multiplicity, exactly 3 jets in full round markers and 4 or more jets in empty square markers.

Sample name	3 jets	4+ jets
CW	980	3107
TP	103	514
T2	111	311
SU	64	491
W	879	1532
TC	101	1575

(a)

μ^-	μ^+
1204	1903
260	254
163	148
246	245
417	1115
787	788

(b)

Table 4.16: The event counts of simulated samples divided by the jet multiplicity (a) and by the charge of the muon (b). SUSY is taken as the second mSUGRA grid reference point. The two parts of the CW sample (TC and W) are also shown separately.

two jet multiplicities. In return for a small increase in the number of parameters, much greater stability of the fit is ensured, as the shapes are determined from two independent samples simultaneously, while the fractions in the 4+ jets bin have as much information as before.

In Table 4.16(a) the event counts are shown when we separate our simulated samples in two jet multiplicity bins. What is clear is that SUSY events are represented less in the 3 jets sample than in the 4+ jets events, for the mSUGRA reference point 2 that is shown.

At leading order semileptonic $t\bar{t}$ decay produces four partons, hence the 3 jets bin event counts for TP and TC samples are much smaller than in the 4+ jets bin. For semileptonic $t\bar{t}$ events with 3 reconstructed jets, one of the leading order final state partons is not reconstructed. If we sort the 4 decay partons by p_{T}, we find that for $\sim 61\%$ of semileptonic $t\bar{t}$ events the fourth parton falls outside the jet selection acceptance of $|\eta| < 2.5$ and $p_{\mathrm{T}} > 20$ GeV. In addition, for $\sim 5\%$ of the 3 jet semileptonic $t\bar{t}$ events the first to third partons fall outside the jet acceptance. As reconstruction algorithms are not perfect and the ATLAS detector does not have perfect coverage, we expect these 66% to increase if we would look at $\eta - p_{\mathrm{T}}$ acceptance of jets instead of partons..

Another possible cause for semileptonic $t\bar{t}$ events to be reconstructed with only 3 jets, is jet-merging. The two jets coming from a hadronic W decay with a large boost, are so close in $\eta - \phi$ space that they are reconstructed as one single jet. These merged W jets can be found by looking at the invariant mass of each separate jet. In only $\sim 2\%$ of the 3 jets events, we find a jet with a mass $70 < m_W < 90$ GeV around the W mass.

We can correctly reconstruct the top mass in semileptonic $t\bar{t}$ events with 3 jets, only if the b ($\bar{b}$) quark of the leptonic t ($\bar{t}$) decay side is outside the acceptance, which happens in $\sim 12\%$ of the 3 jet events. In contrast, the event counts of TP/TC samples in Table 4.16(a) in the 3 jets bin are almost equal, which is an effect of the way we split TP events from TC events (ΔR between reconstructed top and truth-jets top). The reconstructed top is just the sum of the three reconstructed jets, while the truth-jets top is the sum of three truth-jets, individually best matched to the hadronic top decay partons. But since the truth-jets also have to be inside the $\eta - p_{\mathrm{T}}$ acceptance, the truth-jets top does not match the hadronic parton top. Thus the event counts of TP/TC samples in Table 4.16(a) in the 3 jets bin are misleading as the TP events do not represent only the events with a reconstructed top peak, but also have a big combinatorics component.

The decomposition of the CW sample in Table 4.16(a) shows the event count of W events in the 3 jets bin to be lower than in the 4+ jet bin. The cross section for production of W bosons

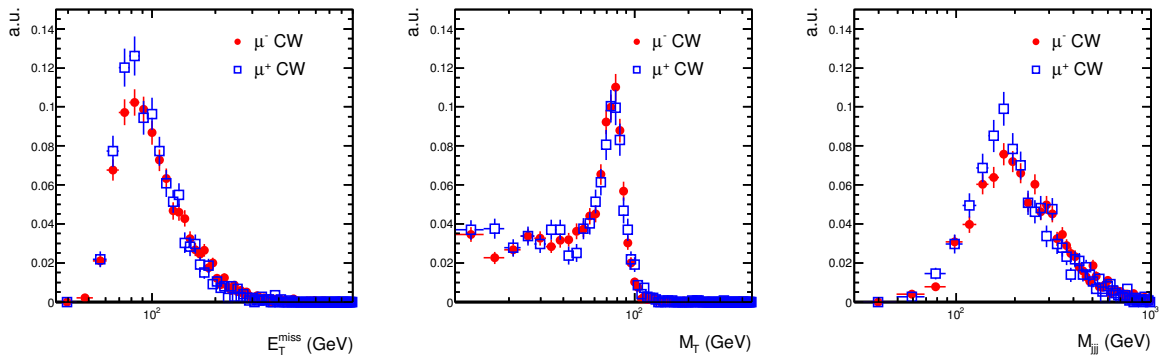


Figure 4.40: Normalized distributions for the CW sample subdivided by muon charge, positively charged in empty square markers and negatively charged in full round markers.

with 3 associated jets is greater than for W bosons with 4+ associated jets. But the probability to produce only 3 very hard jets ($p_{\mathrm{T}}^{1,2,3} > 80, 40, 40$ GeV) and no additional jets is very small. Still however the fraction of W events in comparison to the total number of SM events is much larger in the 3 jets bin (~ 0.75) than in the 4+ jets bin (~ 0.4).

The ratio of dileptonic $t\bar{t}$ events in comparison to the semileptonic $t\bar{t}$ is much greater in 3 jets bin (~ 0.55) than in the 4+ jets bin (~ 0.15). However the absolute number of dileptonic $t\bar{t}$ events is still greater in the 4+ jets bin, because the event selection requires three very hard jets, which biases it towards events with higher jet multiplicity.

4.6.2 Lepton charge asymmetry

The production of W bosons at the LHC has a charge asymmetry, as was discussed in Section 1.1.6. The production of $t\bar{t}$ has no charge asymmetry, hence observed lepton charge asymmetry gives an extra handle on the background composition.

Figure 4.40 shows the normalized distributions of the three fit observables for the CW sample now compared by the charge of the muon. The shapes of the distributions are rather similar for the two muon charges in all three observables. The decay products l/ν (up/down-type) of the longitudinally polarized W bosons, that are produced in approximately 70% of top quark decays, have a p_{T} asymmetry, that leads to the small deviations in the $E_{\mathrm{T}}^{\mathrm{miss}}$-spectra for the two muon charge split CW samples.

It can thus be beneficial to fit the distributions of μ^- and μ^+ events separately in a simultaneous fit procedure. Again a simultaneous combined fit should have more stability, as the shape parameters are determined from two independent data samples with different background component yields. Further stability can be assured by constraining the ratio of yield parameters of the positively and negatively charged muons for each background sample to their predicted values, which have a relatively small theoretical uncertainty.

5

Muons in LHC collision data

Since march 2010 the Large Hadron Collider has been colliding protons at a center-of-mass energy of 7 TeV. This chapter presents analysis of the very first chunk of data delivered by the LHC. In here we analyze the inclusive muon content of this sample in terms of prompt muons versus muons from in-flight decays of kaons and pions.

The analysis shown is based on a total integrated luminosity of approximately 17 nb^{-1} [110]. However the last section is dedicated to showing an example of how this analysis could be used for a physics measurement such as the di-muon composition with an integrated luminosity of 1.5 pb^{-1}, as is described in an ATLAS published note [111].

5.1 Introduction

Pions and kaons decaying into muons constitute a source of background to measurements and searches involving muons in the final state, such as the observation of $W \to \mu\nu$. In this chapter, we present a method to estimate the pion and kaon contamination in events with at least one combined muon, *i.e.*, associated to both one track in the inner detector and one in the muon spectrometer, and we apply it to measure the prompt component of the inclusive muon spectrum at $\sqrt{s} = 7$ TeV. We also discuss the possibility of making this analysis more data-driven and less dependent on simulation input with a larger data sample.

5.2 Data sets and event selection

This analysis is based on an integrated luminosity of approximately 17 nb^{-1}, obtained with stable LHC beams and on-line muon triggers, together with high-quality data from the ATLAS detector. Events are recorded and luminosity is measured in blocks of time typically lasting about two minutes and the detector status and data quality are evaluated for each such block. Data from a block are not considered for this analysis if problems are found in the inner detector, muon spectrometer, trigger or if the solenoid and toroid magnetic fields are not on. The quality is assessed by both the detector subsystems and the offline combined performance groups. The same quality cuts also apply to the di-muon events discussed in Section 5.8.

Collision events are selected by requiring the timing information of the event to be in coincidence with a paired LHC proton bunch and the unprescaled first-level trigger from the muon system without momentum threshold. Furthermore event selection requires at least three inner detector tracks associated with a reconstructed primary vertex. The subset of data used in this chapter is then obtained by requiring one muon with $p_{\text{T}} > 4$ GeV and $|\eta| < 2.5$. The muons used throughout this whole chapter are identified by the match of an inner detector track with a

track reconstructed in the muon spectrometer (Chain 2 [86]). The muon parameters are derived from a common track fit to the hits in the two sub-detectors. The associated inner detector track must for all muons satisfy the conditions of having at least one hit in the pixel detector and at least six hits in the semi-conductor tracker. In the first 17 nb^{-1} of LHC data, a sample of 157466 muons is left after this selection.

Two sets of Monte Carlo (MC) simulated data, based on the Pythia 6.4 generator [35], are produced to perform the analysis and comparison with data:

- A sample of 20 million non-diffractive minimum bias events produced with the MC09 tune [112];
- A sample of 10 million QCD di-jet events, filtered requiring $p_{\text{T}}^{\text{cell}} > 17$ GeV, where $p_{\text{T}}^{\text{cell}}$ is the transverse momentum in a region of approximate size $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$.

Additional three simulated sets are produced to cross-check and validate the procedure:

- A sample of 40 million non-diffractive minimum bias events (MC09 tune), filtered at truth level by the requirement of having at least one region of size $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ containing a total transverse momentum larger than 6 GeV;
- Two samples of 5 million events each produced with the Perugia0 [113] and DW tunes [114].

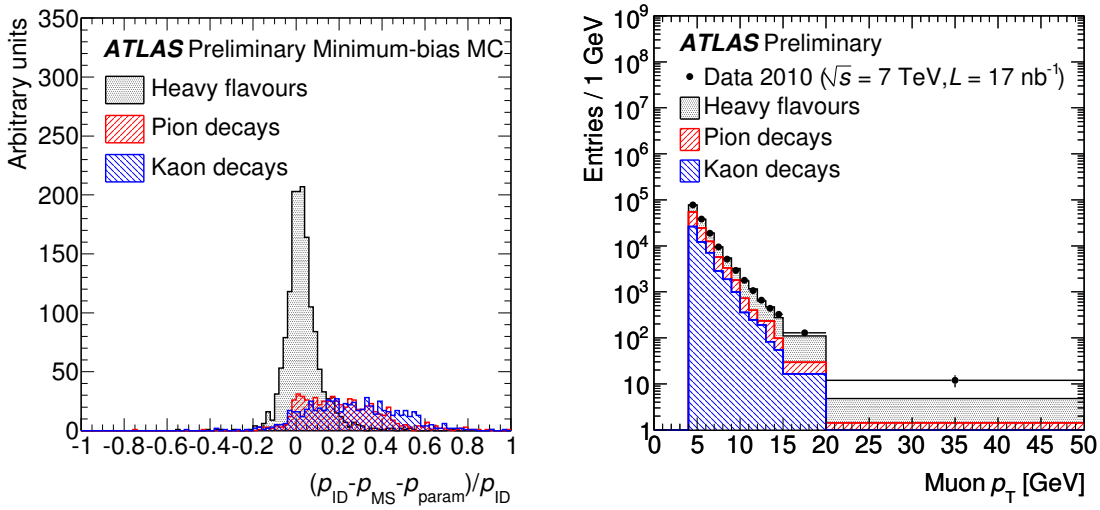
The Perugia0 tune was developed to present the LHC community with an optimized set of parameters that can be used as default settings in Pythia 6.4, specifically tuning the parton shower and underlying event model. The data sets used to constrain the models include hadronic Z decays at LEP, Tevatron minimum bias and Drell-Yan data, and SPS minimum bias data. The DW tune was put forward by the Tevatron-for-LHC workshop, that was conceived to pass on the expertise of the Tevatron and to test new analysis ideas coming from the LHC community. Just as for the Perugia0 tune, Pythia was tuned for correct description of parton showering and underlying event.

Throughout the whole chapter, if not stated otherwise, the term "minimum bias simulation" refers to the MC09 tune sample of 20 million non-diffractive minimum bias events without any filters.

5.3 Analysis method

5.3.1 Description

At a centre-of-mass of 7 TeV, the main sources of muon production are decays of W and Z bosons, charm and bottom hadrons, and pions and kaons. Because of their longer lifetime, pions and kaons may cross a large part of the detector before decaying. Although the muon is emitted isotropically in the rest frame of the pion (kaon), the angles between the decaying particle and the muon in the lab system are usually small due to the Lorentz boost and the small mass difference. Because of this, the tracker hits from the two particles are often associated to the same track. In general, the momentum measurement in the muon spectrometer will correspond to the muon trajectory. Instead, the measurement in the inner detector is dominated either by the pion (kaon) momentum or by the muon momentum, depending on the decay distance. Hence for a combined muon coming from a late pion (kaon) decay, the inner detector and the muon spectrometer tracks have different momenta. For muons produced close to the interaction point, called *prompt* muons, such as those coming from the decay of heavy-flavoured hadrons or



(a) (b)

Figure 5.1: Figure 5.1(a): the distribution of $\Delta p_{\text{loss}}/p_{\text{ID}}$ for different components, as predicted by the minimum bias simulation and for reconstructed muons with $p_{\text{T}} > 6$ GeV. Figure 5.1(b): expected transverse momentum spectra of the different components, overlaid with the observed data. The simulation is rescaled to have the same number of entries as in data.

W and Z bosons, this discrepancy between inner detector and muon spectrometer measurement does not occur.

These considerations have led us to define

$$\frac{\Delta p_{\text{loss}}}{p_{\text{ID}}} = \frac{p_{\text{ID}} - p_{\text{MS}} - p_{\text{param}}}{p_{\text{ID}}}, \tag{5.1}$$

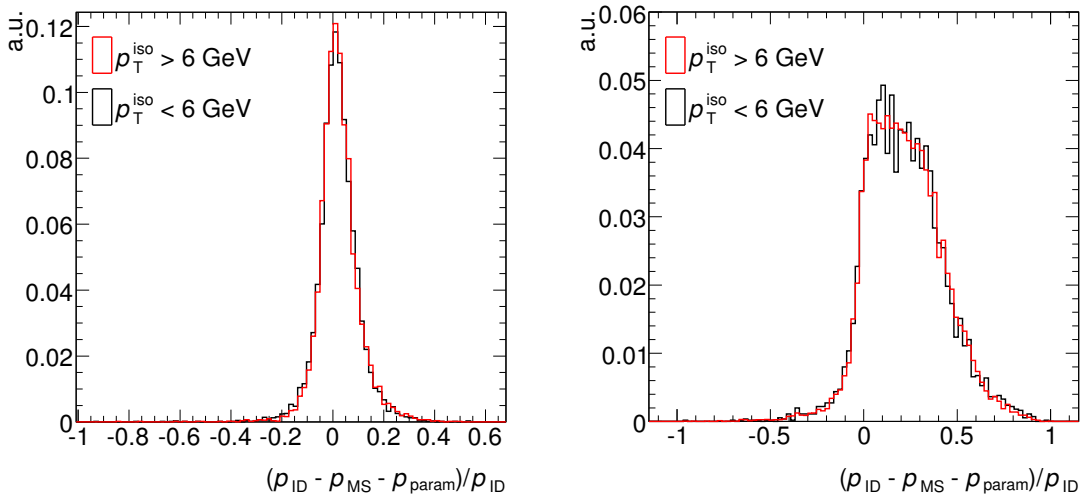
where p_{ID} and p_{MS} are the momenta as measured in the inner detector and in the muon spectrometer respectively, while p_{param} is the parameterized estimation of the energy lost by a muon crossing the material between the two devices. The parameterized estimation is preferred to the measured energy in the calorimeter since the muons considered are usually not isolated. The distribution of this variable for the different components, as predicted by the minimum bias simulation, is shown in Figure 5.1(a). The expected muon transverse momentum spectra are plotted in Figure 5.1(b) and overlaid with the observed data.

In Table 5.1 the various sources of muon production in the minimum bias simulation are broken up into different categories. Early decays, upstream of the first inner detector measurement $r < 400$mm, are essentially indistinguishable from prompt muons, though they only contribute a few percent to the total rate due to the smaller available path length. Specifically for higher momentum tracks, pions and kaons will on average travel further than the first inner detector layer before decaying. A small underestimation of the material between the inner detector and the muon spectrometer was discovered during studies, which leads to the $\Delta p_{\text{loss}}/p_{\text{ID}}$ distribution for prompt muons not being centered around zero, as one would expect for a precisely estimated p_{param}. This is corrected for in later releases of ATLAS software, but in this study the discrepancy remains.

The method that we present is based on a likelihood fit of the $\Delta p_{\text{loss}}/p_{\text{ID}}$ variable, providing us with the yields of the prompt and the pion/kaon components. The main ingredients of the fit procedure are the models used to describe their shape distributions, usually called templates. In the present work, they are derived from simulated events. Alternatively, the models can be

Muon origin	Percentage
Heavy-flavour decay	43 %
Pion decay within ID volume $r < 400$mm	6 %
Pion decay within ID volume $r > 400$mm	13 %
Pion decay in calorimeter	12 %
Kaon decay within ID volume $r < 400$mm	4 %
Kaon decay within ID volume $r > 400$mm	12 %
Kaon decay in calorimeter	10 %
Fake	$< 1\%$
Other prompt muons	$< 1\%$

Table 5.1: Various sources of muon production in the minimum bias simulation.



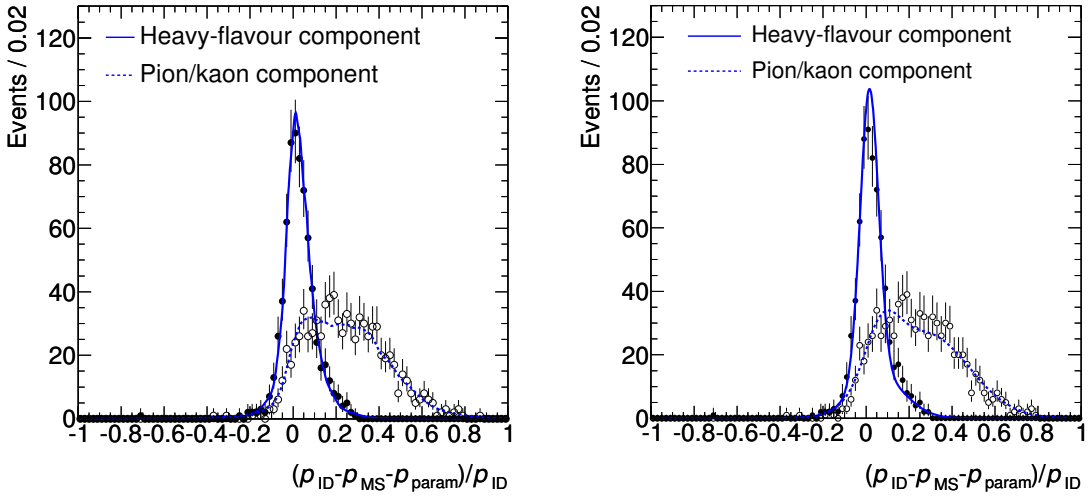
(a) (b)

Figure 5.2: The dependence of the $\Delta p_{\text{loss}}/p_{\text{ID}}$ distribution on muon transverse energy isolation collected in a cone of radius $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$. It is shown that a cut on transverse energy isolation has a negligible effect on the signal component (a) or pion and kaon component (b).

constructed from data by selecting muons from J/ψ and Υ decays and pions from K^0_S and Λ resonances. At the time of publication, this is not possible due to the limited available integrated luminosity, though in Section 5.3.2 we use such resonances to validate the simulation-based templates with data.

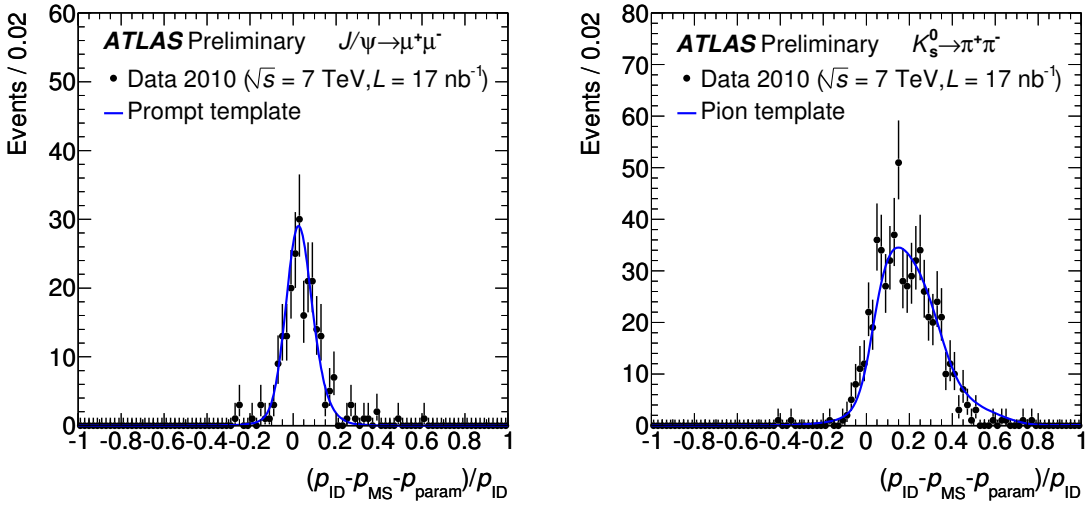
The templates are built from a sample of muons extracted from simulated QCD events. The topology of events in this MC sample, in general, differs from data in particular because the muons from QCD events are more likely to be produced inside high-p_T jets. As a consequence, the muon energy isolation distribution from the QCD sample is different from the minimum bias sample. However, from Figure 5.2 we can safely assume that this will have a negligible impact on the template shapes.

We build both a non-parametric and a parametric description of the templates. In the non-parametric approach the shape of each component is represented by a probability density



(a) (b)

Figure 5.3: Non-parametric (Figure 5.3(a)) and parametric (Figure 5.3(b)) templates for heavy-flavour and pion/kaon components for muons with transverse momentum between 6 and 8 GeV, as obtained from simulated QCD events, as described in the text. For the sake of comparison, also the minimum bias simulated data are shown, represented by full and open points.



(a) (b)

Figure 5.4: Template for prompt-muon component (Figure 5.4(a)) superimposed on top of muons from J/ψ decays. Template for pion component (Figure 5.4(b)) superimposed on top of muons from $K^0_S \to \pi^+\pi^-$ decays.

function derived using the kernel estimation technique [115]. The resulting distribution is the best possible continuous function represented as a superposition of Gaussians with equal surface but varying width, depending on the local event density. The parametric model of the prompt-muon component is described by the convolution of a Gaussian and a Landau function, that describe respectively the average muon energy loss and accounting for large energy loss fluctuations, as described in Section 2.5.8. The pion/kaon component is given by the sum of

three Gaussians.

Figure 5.3(a) shows the non-parametric templates for heavy-flavour and pion/kaon components respectively for muons with transverse momentum between 6 and 8 GeV, as obtained from simulated QCD events. For the sake of comparison, also the minimum bias simulated data are shown, represented by full and open points. In Figure 5.3(b) the parametric templates are shown for the same p_{T} range.

In the rest of the chapter, the non-parametric templates are used to perform the fits. The parametric models are used as an alternative to estimate a systematic uncertainty due to the choice of the template-building technique, as explained in the next sections.

5.3.2 Template validation with data

In this section we discuss the validation of the simulation-based templates with data.

To validate the prompt-muon component we select muons coming from decays of J/ψ. We select opposite-sign di-muon events, where one muon has to pass the selection criteria listed in Section 5.2, while for the second muon we relax the kinematic cuts to $p_{\mathrm{T}} > 2$ GeV. Finally, we select an almost pure J/ψ sample by requiring that the invariant mass of the di-muon pair lies between 2.5 GeV $< M_{\mu\mu} <$ 3.5 GeV. Figure 5.4(a) shows the $\Delta p_{\mathrm{loss}}/p_{\mathrm{ID}}$ distribution for the surviving events as well as the prompt-muon template from simulated QCD events. The template describes the prompt muons from J/ψ accurately, as $\chi^2 = 0.21$ is found when comparing the J/ψ data to the MC template.

Similarly we validate the pion template by selecting $K^0_{\mathrm{S}} \to \pi^+\pi^-$ decays. A muon, again identified as in Section 5.2, is paired with an inner detector track. The two are required to have opposite-sign charges and to originate from the same secondary vertex. An almost pure sample of K^0_{S} decays is obtained by requiring the invariant mass of the di-track candidate to be inside the window 475 MeV $< M_{\pi\pi} <$ 520 MeV. Figure 5.4(b) shows the $\Delta p_{\mathrm{loss}}/p_{\mathrm{ID}}$ distribution for the selected K^0_{S} events. The pion template built from simulated QCD events is superimposed on top and shows good agreement with data with a χ^2 equal to 0.29.

5.4 Systematic uncertainties

In order to account for differences in shape between data and simulation, the templates are extended with three parameters that model several detector effects. A translation in $\Delta p_{\mathrm{loss}}/p_{\mathrm{ID}}$ is accounting for uncertainty in the momentum scale (shift parameter), while a Gaussian smear describes a worsening of the momentum resolution. In addition, the distribution can be dilated, with respect to its mean, along the abscissa (stretch parameter). To estimate the pion and kaon contamination in data, we construct an unbinned profile likelihood as function of the prompt-muon fraction where the three distortion parameters are left free to float and are treated as nuisance parameters.

Figure 5.5 shows the template distributions for different values of the three nuisance parameters, as well as for different pion and kaon contents. In simulation the decay in flight muons originate for $\sim 60\%$ from pions and the rest from kaons [111].

Additional systematic uncertainties are estimated as the variation of the best fit value after each of the following changes.

- *Template shape uncertainty.* Instead of the QCD di-jet sample, the templates are derived from minimum bias simulated events.

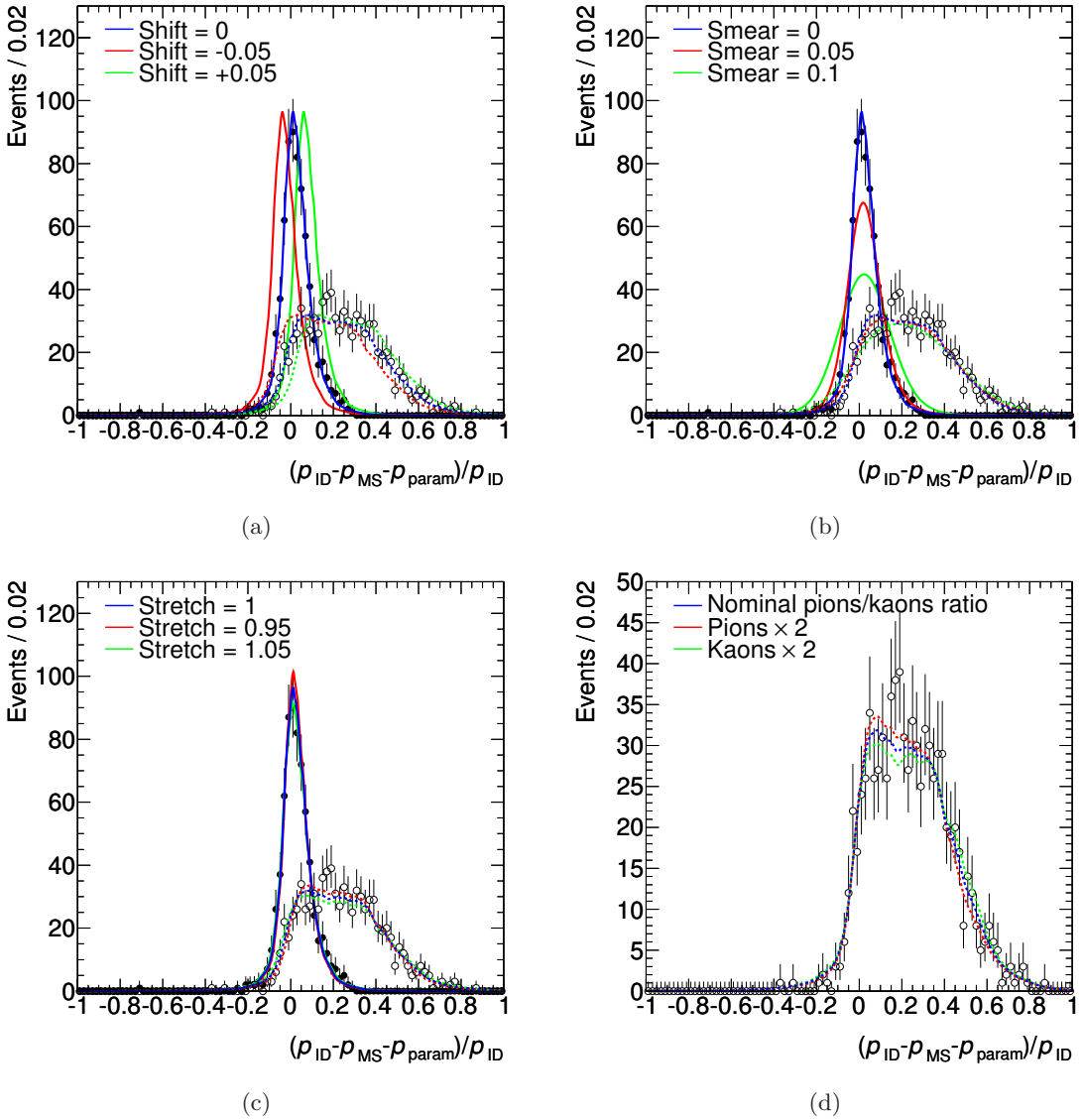
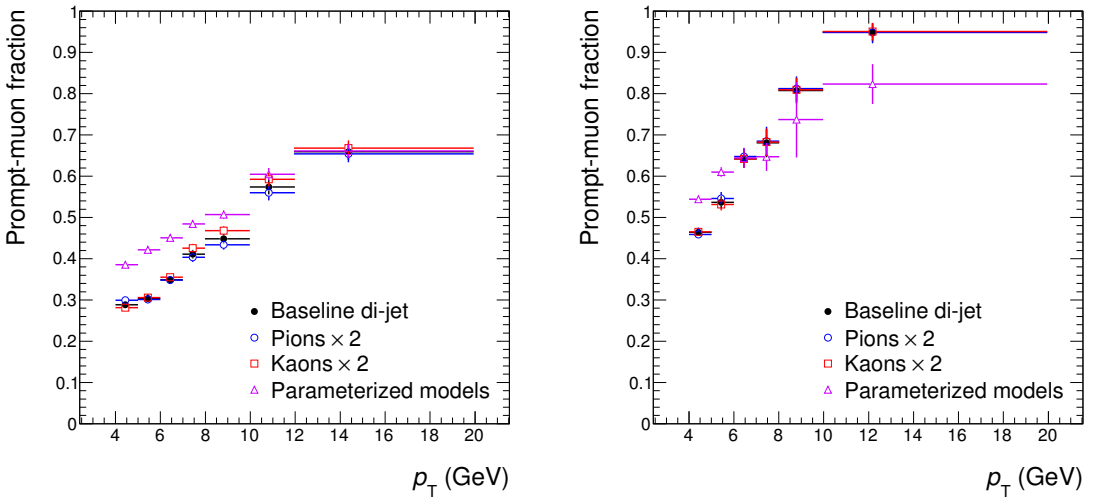


Figure 5.5: Template distributions for different values of the nuisance parameters described in the text (Figures 5.5(a), 5.5(b), 5.5(c)), and for different values of the pion and kaon content (Figure 5.5(d)). All the distributions are shown for reconstructed muons with 6 GeV $< p_{\mathrm{T}} <$ 8 GeV.

- *Template uncertainty on relative pion/kaon content.* The pion (kaon) content in the non-prompt muon template is increased by a factor two.

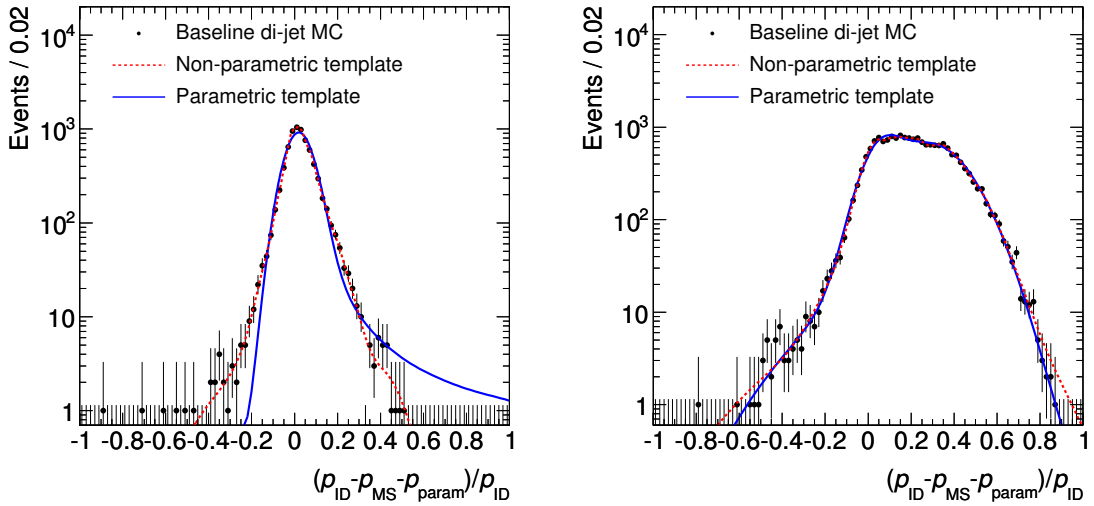
- *Template method uncertainty.* The kernel-estimated templates are replaced by the parametric models as described in Section 5.3.

Figure 5.6 shows the variations of the best fit value of the prompt-muon fraction for each systematic mentioned above. The largest systematic comes from the method of template description with the parametrized models. As an example we take a closer look at the difference between parametric and the baseline non-parametric templates in Figure 5.7 in the $6-8$ GeV p_{T} bin for barrel muons. Although in the central regions of the distributions the two methods agree quite well, in the tails the template descriptions are different, especially for the signal



(a) (b)

Figure 5.6: Variations of the best fit value of the prompt-muon fraction after multiplying the pion/kaon content by a factor of two and replacing the non-parameterized templates by parameterized templates for the barrel (Figure 5.6(a)) and for the end-cap (Figure 5.6(b)).

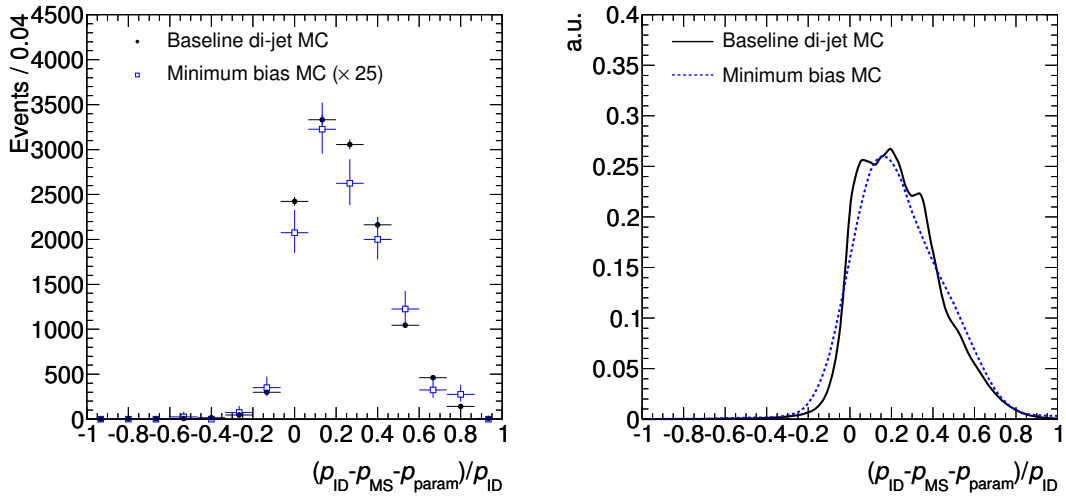


(a) (b)

Figure 5.7: Non-parametric (dashed curve) and parametric (solid curve) templates compared for the prompt component (a) and pion/kaon components (b) for muons with transverse momenta between 6 and 8 GeV in the barrel, as obtained from simulated QCD di-jet events. Note the log scale in comparison to Figure 5.3.

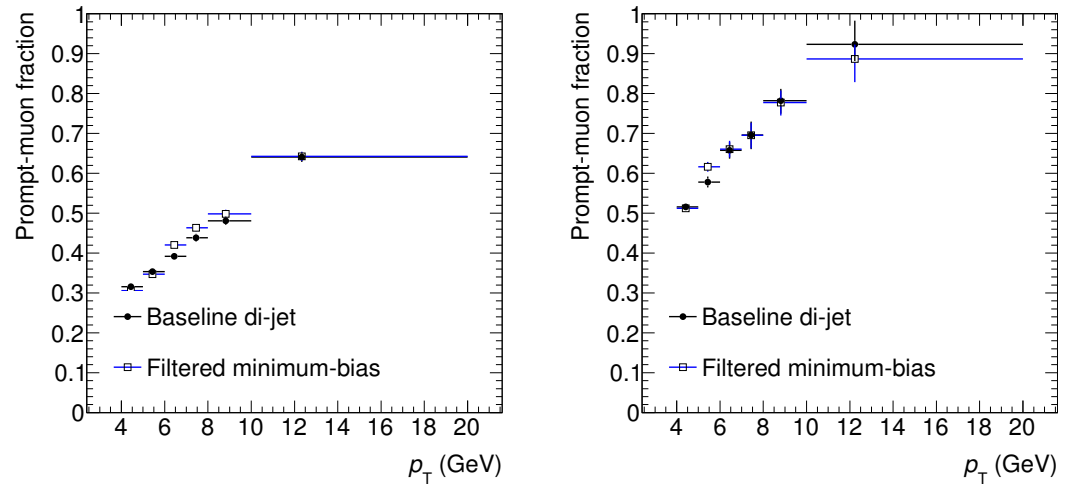
template in Figure 5.7(a). Although a few extensions of the prompt-muon parametrized model have been tried, none showed strongly enhanced shape description or fit stability. The signal parametrized template is well motivated by underlying physics and because we add three distortion parameters that account for detector effects, we use this systematic for a somewhat conservative estimate of the fit uncertainties.

In Figure 5.8(a) the difference in statistics between minimum bias and QCD di-jet samples is



(a) (b)

Figure 5.8: The minimum bias data (a) and derived template (b) in squared markers/dashed curve compared to the baseline QCD di-jet data and template in round markers/solid curve. The results are shown in the $7-20$ GeV p_{T} bin for barrel muons that come from in-flight decays of kaons and pions. For comparison the minimum bias sample was rescaled by a factor 25 in (a), while in (b) a smooth template was created by using interpolation.



(a) (b)

Figure 5.9: Comparison of the best fit values of the prompt-muon fraction based on the di-jet templates (baseline, full circles) and on the p_{T}-filtered minimum bias templates (empty squares). Barrel and end-cap are respectively shown in Figure 5.9(a) and 5.9(b).

shown in the $7-20$ GeV p_{T} bin for barrel muons coming from pions/kaons, where the minimum bias sample is rescaled by a factor 25. The templates derived from these two simulation samples are shown in Figure 5.8(b), where a clear difference can be noticed. Due to the low statistics in the minimum bias sample, the derived templates are used to perform fits on data only in two p_{T} bins, $4-7$ GeV and $7-20$ GeV, both for endcap and barrel muons. The relative error of these two p_{T} bins is propagated to the corresponding bins in the final systematic uncertainty

estimation.

Since the p_{T}-threshold in the baseline di-jet sample, which is set to 17 GeV, is higher than the typical muon transverse momenta considered in this analysis, a test is needed in order to ensure that no biases are introduced in the template building. The low statistics of the default minimum bias sample prevent a careful check, thus the template shapes obtained from the di-jet sample, together with the fit results derived from them, have been carefully cross-checked with the 6 GeV p_{T}-filtered minimum bias events described in Section 5.2. Figure 5.9 shows the fit results based on this p_{T}-filtered minimum bias templates compared with the di-jet templates. The two sets of results are compatible. The di-jet sample is then kept as baseline for template building as it provides the best muon statistics especially at high transverse momenta.

The systematics are summarized in Tables 5.2 and 5.3 for muons in the barrel and the end-cap respectively. For the barrel muons the systematic is dominated by the parametrized template results in the low p_{T} bins, while in the high p_{T} bins the systematic coming from minimum bias simulation takes over. In the endcap the parametrized model systematic dominates in all p_{T} bins except the $6-7$ GeV bin.

All the sources of systematic uncertainty are added in quadrature to give the total systematic uncertainty. Although this is a conservative estimate, for the first collision data sample with an integrated luminosity of 17 nb^{-1} it is considered the correct technique.

Source of Systematic	p_{T} bin [GeV]						
Error [%]	4−5	5−6	6−7	7−8	8−10	10−12	12−20
Pions × 2	3.8	0.7	0.1	1.9	3.2	2.4	0.8
Kaons × 2	2.5	0.7	1.9	3.6	4.5	3.3	1.3
Parametric template	33.6	38.8	29.2	17.8	13.2	5.4	0.2
Minimum bias template	7.7	7.7	7.7	17.4	17.4	17.4	17.4
Total	34.8	39.6	30.3	25.2	22.5	18.7	17.5

Table 5.2: Different sources of systematic error and their effect on the prompt-muon fraction in every p_{T} bin for the barrel muons.

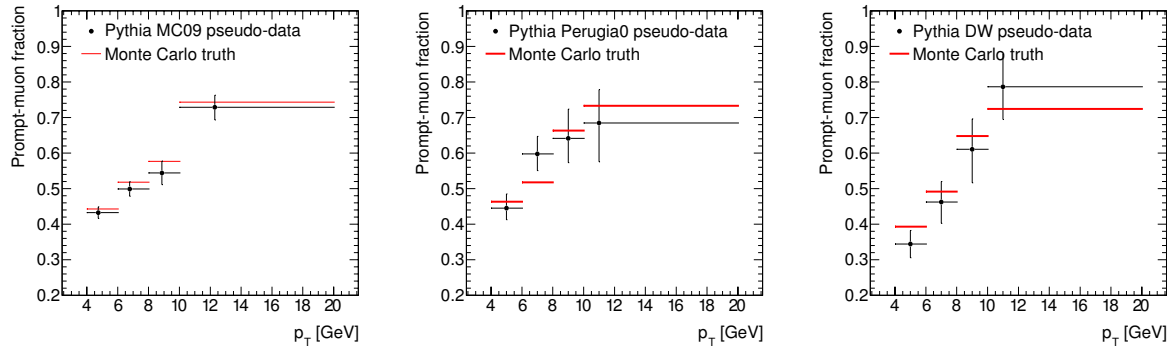
5.5 Closure test

As a closure test, we apply the developed analysis method to different samples of simulated minimum bias events. After fitting the simulated samples we compare the MC truth information expected and estimated fractions of prompt muons.

The fit is repeated on the three minimum bias models with different tunes listed in Section 5.2 in different p_{T} bins. The templates used for all the fits are obtained from the di-jet QCD sample (Section 5.3) using the kernel-estimation technique. The results are summarized in Figure 5.10. As can be concluded, within the error margins the estimations are in good agreement with the expected yields. All the best-fit values of the nuisance parameters are checked to be compatible with the ideal case values.

Source of Systematic Error [%]	p_T bin [GeV]					
	4−5	5−6	6−7	7−8	8−10	10−20
Pions × 2	1.1	1.7	0.9	0.7	0.6	0.1
Kaons × 2	0.3	0.9	0.1	0.3	0.2	0.2
Parametric template	17.4	13.7	0.3	4.9	8.7	13.3
Minimum bias template	12.5	12.5	12.5	2.8	2.8	2.8
Total	21.5	18.6	12.5	5.7	9.2	13.6

Table 5.3: Different sources of systematic error and their effect on the prompt-muon fraction in every p_T bin for the endcap muons.



(a) (b) (c)

Figure 5.10: Estimated prompt-muon component (round markers) as function of p_T, for three different simulated samples as described in the text. The error bars are derived from the 68% confidence level of the profile likelihood. The lines without markers represent the predictions obtained from the minimum bias simulated model. Note that for all fits the templates are obtained from the QCD sample.

5.6 Results

The presented method is applied to muon triggered data subdivided in p_T bins, ranging from 4 to 20 GeV, and then fitted with the template distributions (built from QCD simulated events). Four rapidity regions are fitted separately because of the different detector instrumentation: a central region bounded by $\eta = 1.8$, where the inner-detector transition between barrel and end-cap takes place, and a forward region covering up to $\eta = 2.5$. Both regions are divided in positive and negative rapidity values.

Four examples of best-fit distributions obtained are reported in Figure 5.11. As detector instrumentation is not perfectly symmetric for positive and negative rapidity, the fitted $\Delta p_{\text{loss}}/p_{\text{ID}}$ distributions are somewhat different. However the resulting prompt-muon fraction is statistically compatible, which serves as another cross check of our procedure. The results shown in the rest of this section are hence only split between barrel and end-cap muons, as the negative and positive rapidity regions are summed together.

The fit results as function of the muon p_T are summarized in Figure 5.12(a) for muons within the pseudo-rapidity region $|\eta| < 1.8$. Similarly, Figure 5.12(b) reports the fit results for $|\eta| > 1.8$.

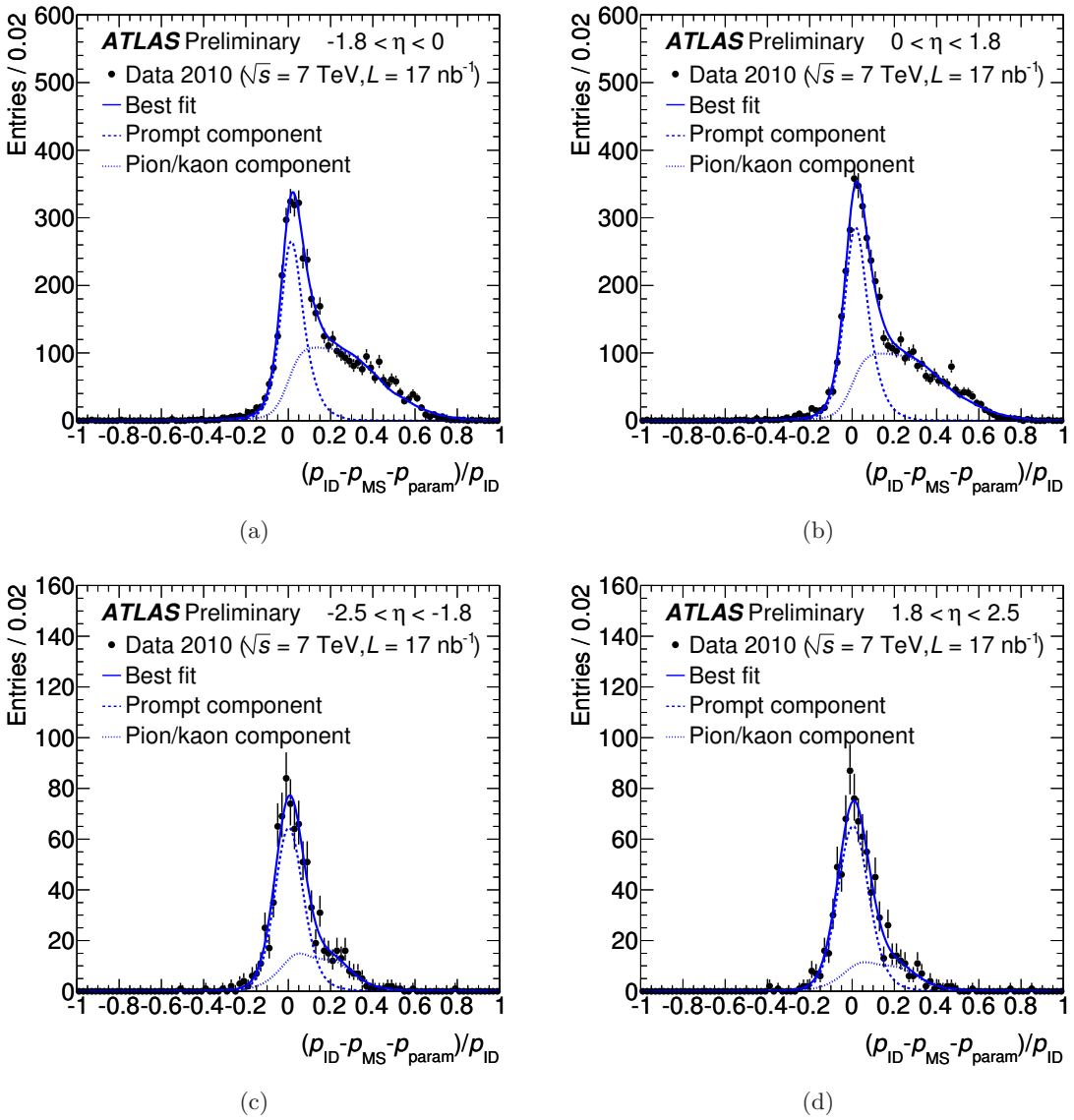


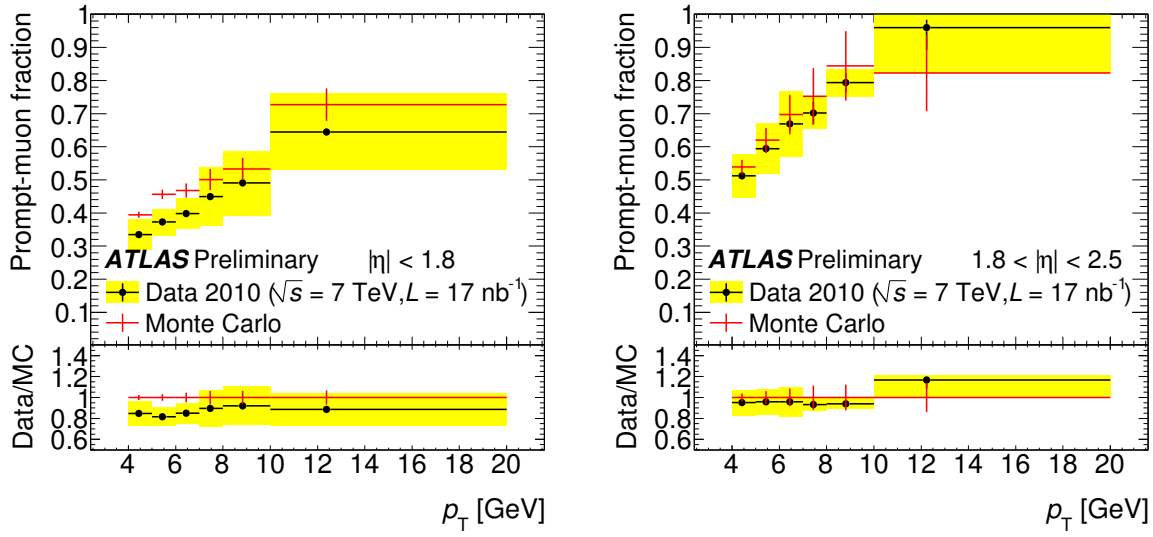
Figure 5.11: Best-fit distributions obtained for 7 GeV $< p_{\mathrm{T}} <$ 8 GeV in four different pseudo-rapidity regions.

The error bars are derived from the 68% confidence level of the profile likelihood. The bands are instead calculated by summing in quadrature the fit and the systematics uncertainties on the templates, as described in Section 5.4.

The lines without markers in Figure 5.12 represent the predictions obtained from the minimum bias simulation truth information with their statistical uncertainties. The fraction of prompt muons is calculated by requesting prompt muons to come from a heavy flavour hadron (either charmed or bottom) decay and to pass our kinematic cuts, divided by the total number of true simulated muons passing our cuts.

The muon sample considered here contains a large contamination from pion and kaon decays, although the fraction of prompt muons becomes larger than 50% above 10 GeV in the central region and above 5 GeV in the forward region.

The prompt muon fraction in the data agrees within systematic uncertainties with the pre-



(a) (b)

Figure 5.12: Measured prompt component (round markers) as function of p_{T} for muons with $|\eta| < 1.8$ in (a) and $|\eta| > 1.8$ in (b). The error bars are derived from the 68% confidence level of the profile likelihood. The bands are calculated by summing in quadrature the fit and the systematics uncertainties on the templates. The lines without markers represent the predictions obtained from the minimum bias simulated model, with their statistical uncertainties.

diction of the minimum bias simulation, although the measured central value is mostly below the expectation. This is not unexpected as the simulation is not tuned specifically to reproduce the heavy-flavour content. Moreover the simulation is based on a leading-order generator which might imply a variation with respect to bottom and charm production as compared to the data. Another effect could be a lower probability in simulation to detect a muon coming from the decay of pions and kaons, which though has been studied in ATLAS [90] and shows a good agreement between data and simulation. The overall evolution as function of p_{T} of the prompt muon fraction in data is comparable to the expectation.

5.7 Conclusion

We have described a data-driven method to estimate the pion and kaon contamination in a data sample with a muon in the final state. The method has been applied to a set of muon triggered events, finding a fraction of prompt muons from heavy-flavour decays compatible with the value predicted by the minimum bias simulation within systematic uncertainties. Although the fraction of prompt muons increases at higher p_{T}, in the range of values considered in this study, we can conclude that the muon production is dominated by pion and kaon decays.

This method can be applied to many different studies involving muons, such as the muon momentum scale and resolution measurement described in Section 2.5.8 and [90], pion and kaon contamination in measurements of leptonically decaying W bosons, or to understand the composition of di-muon events as we will briefly show in the next section.

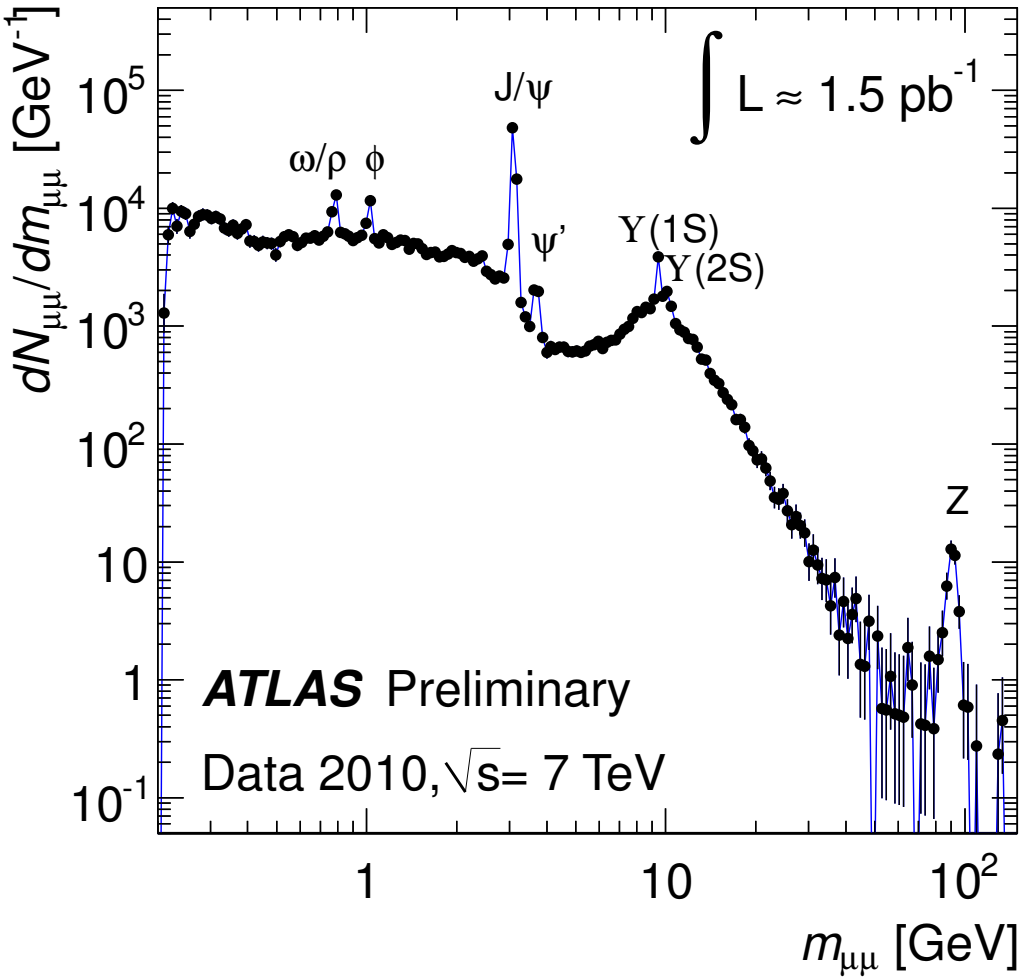


Figure 5.13: Di-muon invariant mass spectrum for data, from fully combined opposite sign muons applying the event and muon selection as described in the text. The labels in the figure indicate $\mu^+\mu^-$ resonances which are visible with an integrated luminosity of 1.5 pb^{-1}.

5.8 Application to Di-muon Composition

The invariant mass spectrum of $\mu^+\mu^-$ pairs is composed of decays from neutral particles such as J/ψ, Υ and Z-bosons, but also from dileptonic decays of $b\bar{b}$ and $c\bar{c}$ pairs. Pions and kaons decaying into muons also contribute to this spectrum. In the note [111], we present a data driven method to distinguish the contribution of events with two prompt muons, from events where at least one muon is coming from π or K decay in flight, and where both muons are the product of pions and kaons.

Collision events were selected using an event filter trigger requesting a muon with transverse momentum above 4 GeV. For the latest runs this trigger was prescaled, giving 1.5 pb^{-1} collected out of 3.4 pb^{-1} delivered integrated luminosity. The offline event selection required at least three inner detector tracks associated with a reconstructed primary vertex. The subset of data used in this section was then obtained by requiring one combined muon (Chain 2 [90]) with offline p_{T} larger than 4 GeV and a second muon at $p_{\mathrm{T}} > 3$ GeV. We restricted this study in

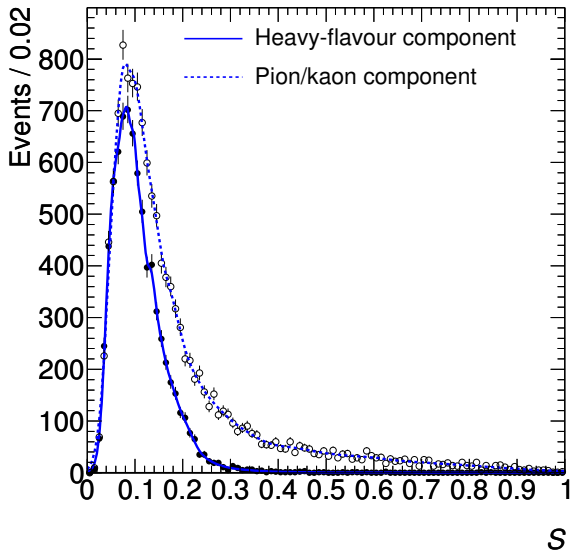


Figure 5.14: Scattering angle significance, S, templates from simulated QCD di-jet events for heavy-flavour and pion/kaon components for muons with transverse momentum between 3 and 4 GeV.

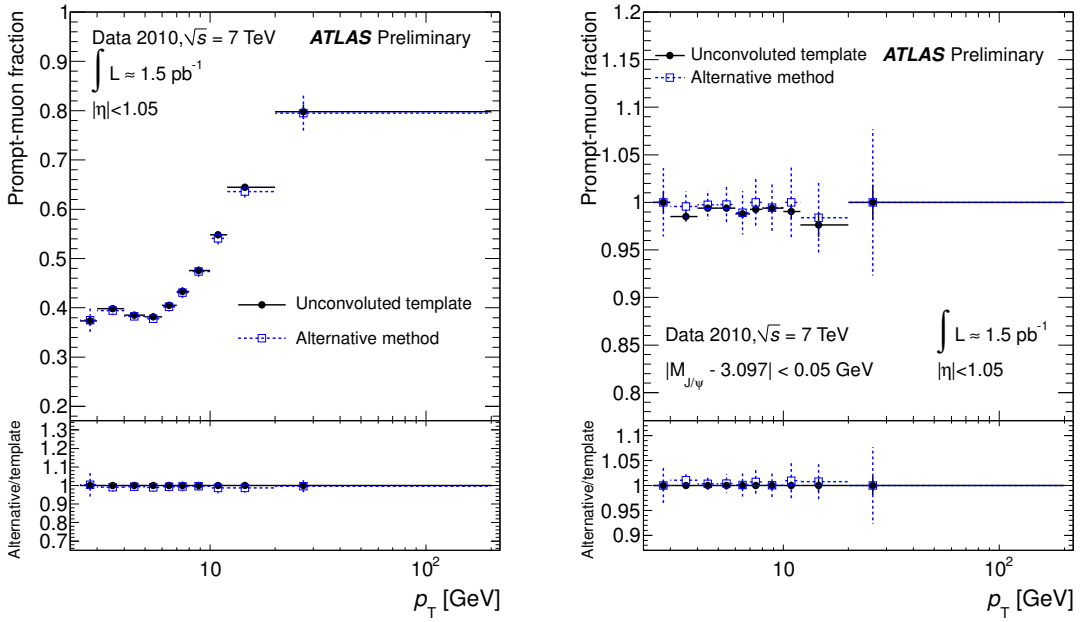
pseudorapidity to the barrel region $|\eta| < 1.05$ of the muon spectrometer. A successful extension of the track in the transition radiation tracker was required and the vertex fit of the two candidate tracks must have a $\chi^2/ndf < 6$. A sample of 33575 $\mu^+\mu^-$ and 8772 $\mu^\pm\mu^\pm$ pairs were left after this selection.

Figure 5.13 shows the di-muon mass spectrum of the data using the described selection. The plot indicates the known resonances for which the π/K contribution was studied as well as the continuous shape of the invariant mass of the muon pairs. The binning of this plot was enhanced specifically to show all the resonances in a clear way on the double logarithmic scale. On top of that our muon selection criteria, more specifically the requirements of combined muons with specific p_{T}, cause the lower mass pairs to be underrepresented in the figure. For example the muons coming from the ∼1 GeV ϕ resonance are only accepted if the ϕ had a large boost, as both energy and momentum must be conserved in the decay. A ϕ standing still in the laboratory frame could never decay to two muons with p_{T}'s above respectively 3 and 4 GeV.

As the main discriminant the $\Delta p_{\mathrm{loss}}/p_{\mathrm{ID}}$ variable was used in a likelihood fit described in Section 5.3. However since a muon requires a momentum of approximately 3 GeV to penetrate the calorimeter and make a track in the muon spectrometer the potential difference in momentum balance between $\pi \to \mu$ and prompt μ is too small to be used as discriminant for very low energy tracks. On the other hand, the smaller boost implies that the decay often produces a noticeable change in direction of the track. A measure to quantify the change in direction, maximum scattering angle significance (S) [111], could thus be used to discriminate even in the low momentum region. Figure 5.14 shows the template difference between the heavy flavour and the pion/kaon components for muons in the $3-4$ p_{T} bin in barrel region as derived from QCD di-jet simulation.

The two discriminants, S and $\Delta p_{\mathrm{loss}}/p_{\mathrm{ID}}$, are summed in a linear combination to form a composite discriminant variable adequate for muons of all momenta as follows:

$$c(r) = \left|\frac{\Delta p_{\mathrm{loss}}}{p_{\mathrm{ID}}}\right| + r\left|S\right| \tag{5.2}$$



(a) single muons

(b) J/ψ mass region

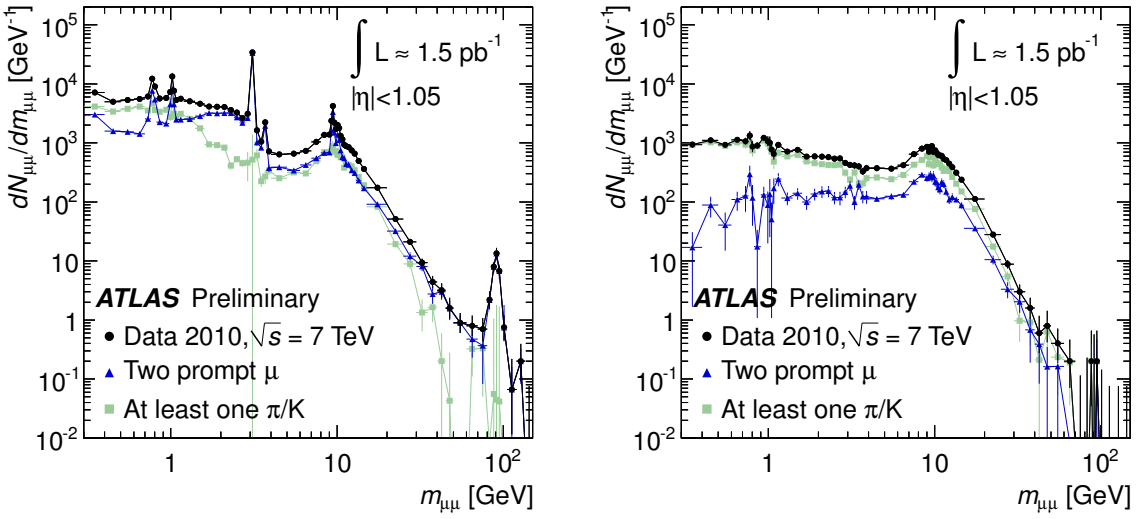
Figure 5.15: Prompt muon fraction as function of the muon p_{T} evaluated using different fitting methods, for (a) single muons and (b) muons in the J/ψ resonance region. The circles are the results using unconvoluted templates, which should in principle give the same prompt muon fraction as the alternative fitting method described in Section 5.8 (squares).

which extends the useful p_{T} range to the full phase space of interest for this analysis. The choice of the value of r was given by optimizing the separation of $c(r)$ for prompt and non-prompt muons in the QCD di-jet sample. Though $\Delta p_{\mathrm{loss}}/p_{\mathrm{ID}}$ is a more powerful discriminant than the scattering angle significance, the optimal separation is obtained by the combination with $r = 0.07$ [111].

The templates, which were built from MC using single muon events, were used to fit the fraction of prompt muons on events in the recorded data with two muons of opposite charge and same charge independently. The data was first split in subsets that corresponded to bins in invariant mass m of the di-muons and transverse momentum of each individual muon. These subsets were then fitted using the single muon templates, thereby creating a probability $P(p_{\mathrm{T}}, m)$ for every muon in a pair that it originates from a π/K given its p_{T} and m. For a di-muon pair the probability that both muons are from a pion/kaon decay in flight is the product of the two muon probabilities. By summing over all the muon pairs in a specific mass bin, and correctly propagating uncertainties, including the systematic uncertainties, you arrive at the measured number of events of a given category at the specified mass [111].

Validation by changing fitting method

The systematic uncertainties, as discussed in Section 5.4, were added in quadrature with the statistical uncertainty and the total uncertainty was later used in the di-muon composition determination. For fits which given these uncertainties would produce confidence intervals outside the physically allowed region the confidence intervals were truncated to the physical boundaries.



(a) opposite sign

(b) same sign

Figure 5.16: The invariant mass distribution of (a) $\mu^+\mu^-$ and (b) $\mu^\pm\mu^\pm$ between 0 and 150 GeV. The filled circles are raw data, while the squares are the contribution of di-muons that contain at least one π or K determined from data.The triangles are the contribution of two prompt muons. The scale on the y-axis takes the bin width into account.

Though the systematic uncertainties at this integrated luminosity are very small compared to the statistical uncertainties.

On top of the systematics already described, the extracted fraction of prompt muons was validated by changing the fitting method. Instead of unbinned likelihood template fits, a series of fits using the fractions of two histograms were performed, following the technique described in [116]. The two input histograms were derived from the same simulation sample that was used to derive the baseline unbinned templates. Hence the main difference compared to the baseline method is that our baseline templates are unbinned probability density functions, while the alternative method used binned histograms and minimization per bin. An advantage of the alternative method is that it also uses the statistical error of the simulation in the minimization, but since it minimizes per bin its disadvantage is that it disregards the full shape information of the probability density function of the discriminant being used for the fit.

Since we convoluted the original template during the fit to account for detector effects we could only compare the results of this alternative fraction fitting method with the unconvoluted template fitting method. The results therefore deviate slightly from the values used elsewhere in this study, but the two methods gave consistent results as Figure 5.15 shows. From the discrepancies of the two methods for all subsets of muon p_{T} we determined that the choice of fitting method introduced an overall systematic uncertainty which is approximately 1%. This was accounted for by adding a 1% uncertainty in quadrature to all results of the template fits.

Results on di-muon composition

The overall π/K contamination for the full mass range was measured at 71% for events where the two muons have the same sign, whereas the corresponding value for the opposite sign di-muon events is 32%. The fraction of di-muon events where both muons originate from π or K is 23% for the same sign and 5% for the opposite sign spectrum. Figures 5.16(a) and 5.16(a) show

respectively the opposite sign and same sign di-muon composition as a function of di-muon pair invariant mass. The specific content of these distributions, such as the exact production and decay processes are well understood and described in detail in [111], but fall outside the scope of this section.

We have described a method to estimate the pion and kaon contamination in a data sample with two muons in the final state. No evidence for previously unknown resonances in either same sign or opposite sign muon pair mass spectrum was found, and apart from a small excess in the π/K contribution at the Υ mass region in the opposite sign spectrum, the findings of this study do not contradict the current understanding of the production mechanisms of muon pairs in a collider experiment.

A

The combined fit model

This appendix contains a print of the entire 3D model for background and signal as used in the analysis in Section 4.4. The format is that of `RooAbsPdf::printCompactTree()`, and should be read as follows: each indentation means the object on that line is a dependent of the line above. For instance a Gaussian PDF over observable `x` with parameters `mean` and `sigma` is denoted as:

```
RooGaussian::ModelName
  RooRealVar::x
  RooRealVar::mean
  RooRealVar::sigma
```

A sum of two Gaussians in `x` would be written as:

```
RooAddPdf::ModelName
  RooGaussian::Model1
    RooRealVar::x
    RooRealVar::mean1
    RooRealVar::sigma1
  RooRealVar::Model1_fraction
  RooGaussian::Model2
    RooRealVar::x
    RooRealVar::mean2
    RooRealVar::sigma2
  RooRealVar::x
```

Since it depends directly on `x` and the two component PDFs. More details on the components of this PDF can be found in the `RooFit` manual [103]. For shorter notation we use the following acronyms:

- `RooRealVar`→`RRV`, `RooFormulaVar`→`RFV`
- `RooProdPdf`→`RPP`, `RooAddPdf`→`RAP`
- `RooTTComb`→`RTT`, `RooGaussian`→`RGAU` (or `RooGaussModel`→`RGM`) and `RooExponential`→`REXP`.

In this format the combined fit model looks like:

RAP::ALL

RPP::SU RRV::ALL_fracSU	RAP::SM		
	RPP::CW RRV::SM_fracCW	RAP::TT	
		RPP::T2 RRV::TT_fracT2	RPP::TP
RTT::SU_etmiss RRV::etmiss RRV::SU_etmiss_base RFV::SU_etmiss_mean RFV::SU_etmiss_sigma RRV::SU_etmiss_base RFV::SU_etmiss_sigma RRV::SU_etmiss_base RAP::SU_mtrans RRV::SU_mtrans_frac RTT::SU_mtrans_real RRV::mtrans RRV::SU_mtrans_base RFV::SU_mtrans_mean RRV::SU_mtrans_base RRV::SU_mtrans_sigma RRV::mtrans RooDecay::SU_mjjj RGM::SU_mjjj_3_conv _exp(-@0/@1)_mjjj _SU_mjjj_tau_[SU_mjjj] RooConstVar::1 RRV::mjjj RFV::SU_mjjj_mean RFV::SU_mjjj_tau RRV::SU_mjjj_tau_slope RRV::SU_etmiss_base RRV::SU_mjjj_tau_offset RFV::SU_mjjj_sigma RFV::SU_mjjj_mean RFV::SU_mjjj_tau RRV::SU_mjjj_tau_slope RRV::SU_etmiss_base RRV::SU_mjjj_tau_offset RFV::SU_mjjj_tau RRV::SU_mjjj_tau_slope RRV::SU_etmiss_base RRV::SU_mjjj_tau_offset RRV::mjjj RFV::SU_mjjj_tau RRV::SU_mjjj_tau_slope RRV::SU_etmiss_base RRV::SU_mjjj_tau_offset	RAP::CW_etmiss RTT::CW_etmiss1 RRV::etmiss RRV::CW_etmiss_base RRV::CW_etmiss_mean RRV::CW_etmiss_sigma RRV::CW_etmiss_frac RTT::CW_etmiss2 RRV::etmiss RooProduct::CW_etmiss2_2 RRV::CW_etmiss_base_ratio RRV::CW_etmiss_base RooProduct::CW_etmiss2_3 RRV::CW_etmiss_mean_ratio RRV::CW_etmiss_mean RRV::CW_etmiss_sigma RRV::etmiss RAP::CW_mtrans_j4 RAP::CW_mtrans_peak_j4 RRV::G_mtrans_corefrac_j4 RGAU::CW_mtrans_core RRV::mtrans RFV::CW_mtrans_mean RRV:: CW_mtrans_mean_slope RRV::etmiss RRV:: CW_mtrans_mean_offset RRV::G_mtrans_sigma RGAU::CW_mtrans_tail RRV::mtrans RooProduct::CW_mtrans_tail_2 RFV:: CW_mtrans_mean_ratio RRV::etmiss RRV:: G_mtrans_mean_ratio_slope RRV:: G_mtrans_mean_ratio_offset RFV::CW_mtrans_mean RRV::CW_mtrans_mean_slope RRV::etmiss RRV::CW_mtrans_mean_offset RooProduct::CW_mtrans_tail_3 RRV::G_mtrans_sigma_ratio RRV::G_mtrans_sigma RRV::mtrans RFV::CW_mtrans_fpeak RRV::G_mtrans_fpeak_offset RRV::G_mtrans_fpeak_amp RFV::CW_mtrans_fpeak_erf RRV::etmiss RRV::CW_mtrans_fpeak_mean RRV::CW_mtrans_fpeak_sig REXP::CW_mtrans_comb RRV::mtrans RRV::CW_mtrans_base RRV::mtrans RTT::CW_mjjj RRV::mjjj RFV::CW_mjjj_base RRV::CW_mjjj_base_offset RRV::CW_mjjj_base_amp RFV::CW_mjjj_base_erf RRV::etmiss RRV::CW_mjjj_base_mean RRV::CW_mjjj_base_sigma RFV::CW_mjjj_mean RRV::CW_mjjj_mean_slope RRV::etmiss RRV::CW_mjjj_mean_offset RRV::G_mjjj_sigma	RTT::T2_etmiss RRV::etmiss RRV::T2_etmiss_base RRV::G_etmiss_mean RRV::T2_etmiss_sigma RAP::T2_mtrans_j4 RAP::T2_mtrans_peak_j4 RRV::T2_mtrans_corefrac_j4 RGAU::T2_mtrans_core RRV::mtrans RFV::T2_mtrans_mean RRV::etmiss RRV::T2_mtrans_mean_slope RRV::T2_mtrans_mean_offset RFV::T2_mtrans_sigma RRV::etmiss RRV::T2_mtrans_sigma_slope RRV::T2_mtrans_sigma_offset RGAU::T2_mtrans_tail RRV::mtrans RooProduct::T2_mtrans_tail_2 RRV::T2_mtrans_mean_ratio RFV::T2_mtrans_mean RRV::etmiss RRV::T2_mtrans_mean_slope RRV::T2_mtrans_mean_offset RooProduct::T2_mtrans_tail_3 RRV::T2_mtrans_sigma_ratio RFV::T2_mtrans_sigma RRV::etmiss RRV:: T2_mtrans_sigma_slope RRV:: T2_mtrans_sigma_offset RRV::mtrans RFV::T2_mtrans_fpeak RRV::T2_mtrans_fpeak_offset RRV::T2_mtrans_fpeak_amp RFV::T2_mtrans_fpeak_erfc RRV::etmiss RRV::T2_mtrans_fpeak_mean RRV::T2_mtrans_fpeak_sig REXP::T2_mtrans_comb RRV::mtrans RFV::T2_mtrans_base RRV::etmiss RRV::T2_mtrans_base_slope RRV::T2_mtrans_base_offset RRV::T2_mtrans_base_amp RFV::T2_mtrans_base_erf RRV::etmiss RRV::T2_mtrans_base_mean RRV::T2_mtrans_base_sig RRV::mtrans RTT::T2_mjjj RRV::mjjj RFV::T2_mjjj_base RRV::T2_mjjj_base_offset RRV::T2_mjjj_base_amp RFV::T2_mjjj_base_erf RRV::etmiss RRV::T2_mjjj_base_mean RRV::T2_mjjj_base_sig RFV::T2_mjjj_mean RRV::etmiss RRV::T2_mjjj_mean_slope RRV::T2_mjjj_mean_offset RRV::G_mjjj_sigma	RTT::TP_etmiss RRV::etmiss RRV::TP_etmiss_base RRV::G_etmiss_mean RRV::TP_etmiss_sigma RAP::TP_mjjj RGAU::TP_mjjj_peak RRV::mjjj RRV::TP_mjjj_peak_mean RRV::TP_mjjj_peak_sigma RRV::TP_mjjj_fpeak RTT::TP_mjjj_comb RRV::mjjj RRV::TP_mjjj_base RRV::TP_mjjj_comb_mean RRV::TP_mjjj_comb_sigma RRV::mjjj RAP::TP_mtrans_j4 RAP::TP_mtrans_peak_j4 RRV::TP_mtrans_corefrac_j4 RGAU::TP_mtrans_core RRV::mtrans RRV::TP_mtrans_mean RRV::G_mtrans_sigma RGAU::TP_mtrans_tail RRV::mtrans RooProduct::TP_mtrans_tail_2 RFV::TP_mtrans_mean_ratio RRV:: G_mtrans_mean_ratio_offset RRV:: G_mtrans_mean_ratio_slope RRV::etmiss RRV::TP_mtrans_mean RooProduct::TP_mtrans_tail_3 RRV::TP_mtrans_sigma_ratio RRV::G_mtrans_sigma RRV::mtrans RFV::TP_mtrans_fpeak RRV::TP_mtrans_fpeak_offset RRV::G_mtrans_fpeak_amp RFV::TP_mtrans_fpeak_erf RRV::etmiss RRV::TP_mtrans_fpeak_mean RRV::TP_mtrans_fpeak_sig REXP::TP_mtrans_comb RRV::mtrans RRV::TP_mtrans_base RRV::mtrans

B

Parameters in maximum floating shapes model

The complete model as described in Appendix A contains 74 distinct parameters. This is more then the minimum amount of parameters necessary to describe the signal and backgrounds of the combined fit analysis. This is clearly visible through large correlations between parameters of the model, and this redundancy in the parametrization has been addressed in Sections 4.3.6 and 4.3.6.

Only 36 shape parameters plus 3 yield parameters are *real* parameters of the fit. An initial estimate for the 36 shape parameters is made in the prefit. In the combined fit stage, we would ideally release all these parameters in order to minimize the dependence on Monte Carlo generators. In practice, this is not possible. Every fit parameter increases the chance of instability in the minimization procedure, thus there is a practical maximum to this number, that can only be found by trial and error. This is the list of 23 shape parameters that combined the largest freedom in shape with a stable fit result.

`TP_mjjj_base`	`TP_mtrans_base`
`TP_mjjj_peak_mean`	`CW_mtrans_base`
`TP_mjjj_peak_sigma`	`T2_mtrans_corefrac`
`CW_mjjj_base_amp`	`T2_mtrans_sigma_ratio`
`CW_mjjj_base_mean`	`G_mtrans_corefrac`
`CW_mjjj_base_offset`	`G_mtrans_sigma`
`CW_mjjj_base_sigma`	`G_mtrans_sigma_ratio`
`CW_mjjj_mean_offset`	`SU_mtrans_base`
`CW_mjjj_mean_slope`	
`T2_mjjj_base_sig`	`CW_etmiss_mean`
`T2_mjjj_mean_slope`	`G_etmiss_mean`
`G_mjjj_sigma`	`SU_etmiss_base`

This list includes the global M_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ parameters, most CW and TP parameters, most T2 M_{jjj} and $E_{\mathrm{T}}^{\mathrm{miss}}$ parameters, and some T2 M_{T} parameters.

References

[1] D. Griffiths, *Introduction to Elementary Particles*. Wiley, February, 2008.

[2] DONUT Collaboration, K. Kodama et al., *Observation of tau-neutrino interactions*, Phys. Lett. **B504** (2001) 218–224, arXiv:hep-ex/0012035.

[3] Particle Data Group Collaboration, C. Amsler et al., *Review of particle physics*, Phys. Lett. **B667** (2008) 1.

[4] P. W. Higgs, *Broken Symmetries and the masses of gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[5] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132–133.

[6] P. W. Higgs, *Spontaneous Symmetry Breakdown Without Massless Bosons*, Phys. Rev. **145** (1966) 1156–1163.

[7] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*. Cambridge Univ. Pr., UK, 1996.

[8] E. van der Kraaij, *First Top Quark Physics with ATLAS - A Prospect*. PhD thesis, Nikhef, Universiteit van Amsterdam, 2009.

[9] S. Grijpink, *Charged Current Cross Section Measurement at HERA*. PhD thesis, Nikhef, Universiteit van Amsterdam, 2004.

[10] T. Gleisberg, S. Hoeche, F. Krauss, M. Schoenherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007.

[11] M. Gosselink, *Radiating Top Quarks*. PhD thesis, Nikhef, 2010.

[12] G. Altarelli and M. Mangano, *Proceedings of the workshop on standard model physics (and more) at the LHC*. CERN, 2000.

[13] T. Sjostrand and P. Z. Skands, *Multiple interactions and the structure of beam remnants*, JHEP **03** (2004) 053, arXiv:hep-ph/0402078.

[14] P. van Mulders, *Calibration of the jet energy scale using top quark events at the LHC*. PhD thesis, Vrij Universteit Brussel, Universiteit Antwerpen, 2010.

[15] T. Sjöstrand and M. van Zijl, *A multiple-interaction model for the event structure in hadron collisions*, Phys. Rev. D **36** (1987) no. 7, 2019–2041.

[16] H. Jung, *Multiparton interactions and underlying events at HERA and the LHC*, http://www.desy.de/~jung/talks/multiple-interactions.pdf.

[17] http://projects.hepforge.org/mstwpdf/plots/plots.html.

[18] R. Bruneliere, S. Caron, J. Dietrich, P. de Jong, G. Polesello, and Z. Rurikova, *ATLAS prospects for the discovery of low mass SUSY for different LHC centre-of-mass energies*, ATL-PHYS-INT-2009-020, CERN, Geneva, Feb, 2009.

[19] http://www.thphys.uni-heidelberg.de/∼plehn/prospino/.

[20] A. Shibata et al., *Understanding Monte Carlo Generators for Top Physics*, ATL-COM-PHYS-2009-334, CERN, Geneva, Jun, 2009.

[21] UA1 Collaboration, U. Collaboration, *Experimental observation of isolated large transverse energy electrons with associated missing energy at* $\sqrt{s} = 540\ GeV$, Phys. Lett. **B122** (1983) 103–116.

[22] UA1 Collaboration, U. Collaboration, *Experimental observation of lepton pairs of invariant mass around 95* GeV/c^2 *at the CERN SPS collider*, Phys. Lett. **B126** (1983) 398–410.

[23] UA2 Collaboration, U. Collaboration, *Evidence for* $Z^0 \to e^+e^-$ *at the CERN* $\overline{p}p$ *collider*, Phys. Lett. **B129** (1983) 130–140.

[24] 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.

[25] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, *FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order*, arXiv:1011.3540 [hep-ph].

[26] 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*, arXiv:1010.2130 [hep-ex].

[27] ATLAS Collaboration, *Measurement of W→munu charge asymmetry in proton-proton collisions at sqrt(s)=7 TeV with the ATLAS detector*, ATL-COM-PHYS-2011-099, CERN, Geneva, Feb, 2011.

[28] CDF Collaboration, F. Abe et al., *Observation of Top Quark Production in* $p\bar{p}$ *Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626, arXiv::hep-ex/9503002v2.

[29] D0 Collaboration, S. Abachi et al., *Observation of the Top Quark*, Phys. Rev. Lett. **74** (1995) 2632, arXiv:hep-ex/9503003v1.

[30] Tevatron Electroweak Working Group, *Combination of CDF and D0 Results on the Mass of the Top Quark*, arXiv:0903.2503 [hep-ex].

[31] P. Houben, *A Measurement of the Mass of the Top Quark using the Ideogram Method*. PhD thesis, Nikhef, Universiteit van Amsterdam, 2009.

[32] J. Hegeman, *Measurement of the top quark pair production cross section in proton-antiproton collisions at* $\sqrt{s} = 1.96\ TeV$, *hadronic top decays with the D0 detector*. PhD thesis, Nikhef, Universiteit Twente, 2009. FERMILAB-THESIS-2009-07.

[33] S. A. et al., *GEANT4 - A Simulation Toolkit*, Nuclear Instruments and Methods **A506** (2003) 250–303.

[34] K. Assamagan et al., *The ATLAS Monte Carlo Project*, ATL-SOFT-INT-2010-002, CERN, Geneva, Feb, 2010.

[35] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026.

[36] 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.

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

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

[39] 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.

[40] V. Gribov and L. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys **15** (1972) 428 and 678.

[41] L. Lipatov, *The parton model and perturbation theory*, Sov. J. Nucl. Phys **20** (1975) 94.

[42] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126** (1977) 298.

[43] Y. Dokshitser, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Pert urbation Theory in Quantum Chromodynamics*, Sov. Phys. JETP **46** (1977) 641.

[44] V. Sudakov, *Vertex parts at very high-energies in quantum electrodynamics*, Sov. Phys. JETP **30** (1956) 65. see also: arXiv:hep-ph/0312336v1.

[45] 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.

[46] S. Hoeche et al., *Matching parton showers and matrix elements*, arXiv:hep-ph/0602031.

[47] S. Dawson, *SUSY and such*, NATO Adv. Study Inst. Ser. B Phys. **365** (1997) 33–80, arXiv:hep-ph/9612229.

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

[49] A. V. Gladyshev and D. I. Kazakov, *Supersymmetry and LHC*, Phys. Atom. Nucl. **70** (2007) 1553–1567, arXiv:hep-ph/0606288.

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

[51] WMAP Collaboration, J. Dunkley et al., *Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP data*, Astrophys. J. Suppl. **180** (2009) 306–329, arXiv:0803.0586 [astro-ph].

[52] E. Cremmer, P. Fayet, and L. Girardello, *Gravity-induced supersymmetry breaking and low energy mass spectrum*, Physics Letters B **122** (1983) no. 1, 41 – 48.

[53] J. L. Feng and T. Moroi, *Supernatural supersymmetry: Phenomenological implications of anomaly-mediated supersymmetry breaking*, Phys. Rev. **D61** (2000) 095004, arXiv:hep-ph/9907319.

[54] G. F. Giudice and R. Rattazzi, *Theories with gauge-mediated supersymmetry breaking*, Phys. Rept. **322** (1999) 419–499, arXiv:hep-ph/9801271.

[55] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, Volume-2*. CERN, Geneva, 1999.

[56] S. P. Martin, *A Supersymmetry Primer*, arXiv:hep-ph/9709356v5.

[57] F. Gianotti and G. Ridolfi, *Searching for supersymmetry at the LHC*, http://cdsweb.cern.ch/record/601655. CERN Academic Training, 2003.

[58] I. Niessen, *Supersymmetric Phenomenology in the mSUGRA Parameter Space*, arXiv:0809.1748 [hep-ph].

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

[60] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics, ch. Supersymmetry*, arXiv:hep-ex/0901.0512. Geneva, 2008, CERN-OPEN-2008-020.

[61] W. Beenakker et al., *Soft-gluon resummation for squark and gluino hadroproduction*, JHEP **12** (2009) 041, arXiv:hep-ph/0909.4418.

[62] W. Beenakker et al., *Supersymmetric top and bottom squark production at hadron colliders*, JHEP **08** (2010) 098, arXiv:hep-ph/1006.4771.

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

[64] Vol. I LEP Design Report, *The LEP injector chain*, , CERN, 1983. CERN-LEP/TH/83-29.

[65] TeVI Group, *Design Report Tevatron 1 project*, FERMILAB-DESIGN-1984-01.

[66] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **S08003** (2008) .

[67] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **S08004** (2008) .

[68] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **S08002** (2008) .

[69] LHCb Collaboration, *The LHCb detector at the LHC*, JINST **S08005** (2008) .

[70] TOTEM Collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **S08007** (2008) .

[71] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, JINST **S08006** (2008) .

[72] ATLAS Collaboration, *Performance of the ATLAS Detector using First Collision Data*, arXiv:hep-ex/1005.5254.

[73] M. Limper, *Track and vertex reconstruction in the ATLAS inner detector.* PhD thesis, Nikhef, Universiteit van Amsterdam, 2009.

[74] ATLAS Collaboration, *Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2010-072, CERN, Geneva, Jul, 2010.

[75] ATLAS Collaboration, *Alignment Performance of the ATLAS Inner Detector Tracking System in 7 TeV proton-proton collisions at the LHC*, ATLAS-CONF-2010-067, CERN, Geneva, Jul, 2010.

[76] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at* $\sqrt{s}$ *=7 TeV in the ATLAS experiment*, ATLAS-CONF-2010-069, CERN, Geneva, Jul, 2010.

[77] G. Ordonez Sanz, *Muon Identification in the ATLAS Calorimeters.* PhD thesis, Nikhef, Radboud Universiteit Nijmegen, 2009.

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

[79] H1 Collaboration, C. Schwanenberger, *The jet calibration in the H1 liquid argon calorimeter*, arXiv:physics/0209026.

[80] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics, ch. Electrons and Photons*, arXiv:hep-ex/0901.0512. Geneva, 2008, CERN-OPEN-2008-020.

[81] ATLAS Collaboration, *Observation of inclusive electrons in the ATLAS experiment at* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2010-073, CERN, Geneva, Jul, 2010.

[82] 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 sqrts = 7 TeV*, ATLAS-CONF-2010-054, CERN, Geneva, Jul, 2010.

[83] 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*, ATLAS-CONF-2010-056, CERN, Geneva, Jul, 2010.

[84] ATLAS Collaboration, *Performance of the Missing Transverse Energy Reconstruction and Calibration in Proton-Proton Collisions at a Center-of-Mass Energy of 7 TeV with the ATLAS Detector*, ATLAS-CONF-2010-057, CERN, Geneva, Jul, 2010.

[85] N. van Eldik, *The ATLAS muon spectrometer: calibration and pattern recognition.* PhD thesis, Nikhef, Universiteit van Amsterdam, 2007.

[86] ATLAS Collaboration, *Muon Performance in Minimum Bias pp Collision Data at* $\sqrt{s} = 7$ TeV *with ATLAS*, ATLAS-CONF-2010-036, CERN, Geneva, Jul, 2010.

[87] H. van der Graaf et al., *RASNIK technical system description for ATLAS*, NIKHEF Note ETR-2000-04. http://cdsweb.cern.ch/record/1073160.

[88] ATLAS Collaboration, *Commissioning of the ATLAS Muon Spectrometer with Cosmic Rays*, arXiv:hep-ex/1006.4384.

[89] V. Blobel, *Millepede: Linear Least Squares Fits with a Large Number of Parameters*, http://cdsweb.cern.ch/record/1073160.

[90] ATLAS Collaboration, *Muon Reconstruction Performance*, ATLAS-CONF-2010-064, CERN, Geneva, Jul, 2010.

[91] M. Woudstra et al., *Twin-tubes: 3D tracking based on the ATLAS muon drift tubes*, Nuclear Instruments and Methods in Physics Research A **560** (2006) .

[92] ATLAS Computing Group, *ATLAS Computing Technical Design Report*, CERN-LHCC-2005-022, CERN, 2005.

[93] Z. van Kesteren, *Identification of muons in ATLAS*. PhD thesis, Nikhef, Universiteit van Amsterdam, 2010.

[94] J. Snuverink, *The ATLAS muon spectrometer: commissioning and tracking*. PhD thesis, Nikhef, Universiteit van Amsterdam, 2009.

[95] D. Adams et al., *Track reconstruction in the ATLAS Muon Spectrometer with MOORE*, ATL-SOFT-2003-007, CERN, Geneva, May, 2003.

[96] Private communication with P. Kluit.

[97] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81–86.

[98] F. Koetsveld, A. Koutsman, W. Verkerke, N. de Groot, and P. de Jong, *The combined fit method for 1-lepton SUSY searches*, ATL-COM-PHYS-2010-899, CERN, Geneva, Nov, 2010.

[99] F. Koetsveld, *Searching Supersymmetry with ATLAS*. PhD thesis, Radboud Universiteit Nijmegen, Nikhef, forthcoming.

[100] ATLAS Collaboration, *Prospects for Supersymmetry and Univeral Extra Dimensions discovery based on inclusive searches at a 10 TeV centre-of-mass energy with the ATLAS detector*, ATL-PHYS-PUB-2009-084, CERN, Geneva, Jul, 2009.

[101] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics, ch. Muons*, arXiv:hep-ex/0901.0512. Geneva, 2008, CERN-OPEN-2008-020.

[102] W. Verkerke and I. Van Vulpen, *Commissioning ATLAS using top-quark pair production*, ATL-COM-PHYS-2007-023, CERN, Geneva, Apr, 2007.

[103] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, arXiv:physics/0306116.

[104] L. Moneta, K. Belasco, K. Cranmer, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, *The RooStats Project*, arXiv:1009.1003v1 [physics.data-an].

[105] K. Cranmer, *Statistical Challenges for Searches for New Physics at the LHC*, in *Statistical Problems in Particle Physics, Astrophysics and Cosmology*, L. Lyons & M. Karagöz Ünel, ed. 2006. arXiv:physics/0511028.

[106] J. Linnemann, *Measures of Significance in HEP and Astrophysics*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*, L. Lyons, R. Mount, & R. Reitmeyer, ed. 2003. arXiv:physics/0312059.

[107] R. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*. Wiley, July, 1989.

[108] W. Verkerke, *Data Analysis lectures*, Amsterdam, 2004. http://www.nikhef.nl/∼verkerke/talks/data_analysis/data_analysis_2004_v17.pdf.

[109] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343–367.

[110] ATLAS Collaboration, *Extraction of the prompt muon component in inclusive muons produced at* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2010-075, CERN, Geneva, Jul, 2010.

[111] M. Consonni, A. Koutsman, E. van der Poel, M. Raas, N. Ruckstuhl, and R. Sandstrom, *Dimuon composition in ATLAS at 7 TeV*, ATLAS-CONF-2011-003, CERN, Geneva, Aug, 2010.

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

[113] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. **D82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[114] TeV4LHC QCD Working Group Collaboration, M. G. Albrow et al., *Tevatron-for-LHC Report of the QCD Working Group*, arXiv:hep-ph/0610012.

[115] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198–207, arXiv:0011057 [hep-ex].

[116] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219–228.

[117] A. Bosma, *The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types*. PhD thesis, Rijksuniversiteit Groningen, 1978.

[118] ATLAS Collaboration, *Search for supersymmetry using final states with one lepton, jets, and missing transverse momentum with the ATLAS detector in sqrts = 7 TeV pp*, arXiv:hep-ex/1102.2357.

[119] ATLAS Collaboration, *Search for squarks and gluinos using final states with jets and missing transverse momentum with the ATLAS detector in sqrt(s) = 7 TeV proton-proton collisions*, arXiv:hep-ex/1102.5290.

Summary

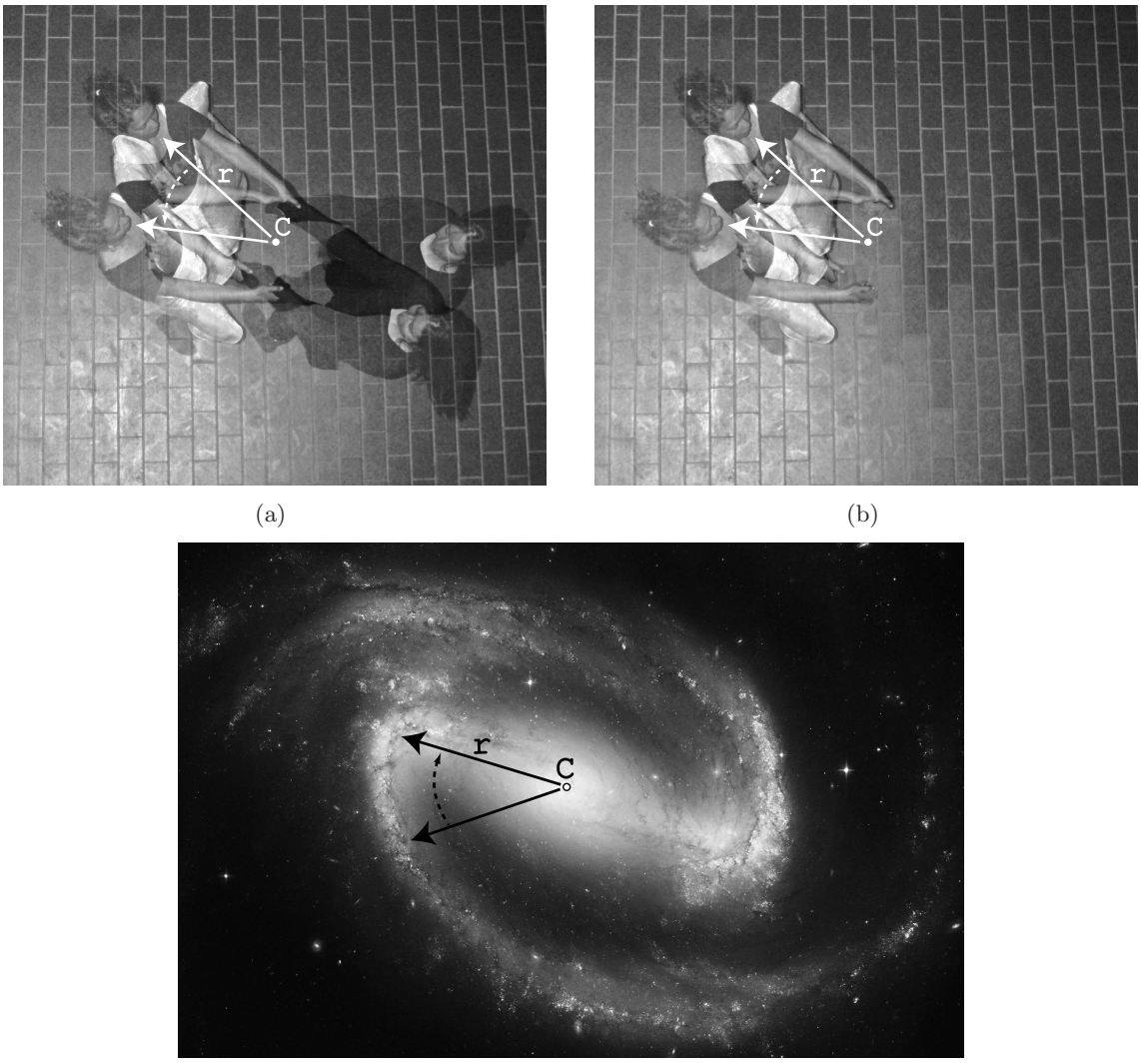
One of the great mysteries of physics today is the nature of *dark matter* particles. Unlike visible matter that stars, earth and earthlings are made of, dark matter does not emit or scatter light, nor does it interact in any other way except through gravity. Astronomers have shown that visible matter accounts only for 4.6% of the mass-energy density of the observable universe, while dark matter accounts for 23%. Thus physicists can not explain what a large part of the universe is made of, quite an embarassing situation to find oneself in.

How do we know dark matter exists?

One of the clearest indications of the existence of dark matter comes from measurements to the rotational curves of galaxies. Let us first consider a familiar example of two friends twirling around each other, such as shown in Figure 1(a). The point around which they rotate, called the *center of mass*, and schematically depicted as C in Figure 1(a), is right in the middle if the two friends have the same weight. Just like in a centrifuge, the speed of rotation depends on the length of their arms, or the distance r from the center of mass. If we know the weight of the girls and the amount of force they are using to twirl, the speed of rotation can be calculated using the laws of motion published for the first time by Isaac Newton in 1687. If the girls use more force to twirl, the speed will increase. If the girls put on backpacks with heavy books in them and use the same force, the speed will decrease.

Now let us consider a different situation, where one of the girls is invisible, as shown in Figure 1(b). The question is, what can we find out about the invisible girl, just by studying the photograph. From the Figure 1(b), we can measure the speed with which the visible girl is rotating and where the center of mass C is located. Because we know the weight of the visible girl and the rotation speed, we know the amount of force she is using. The laws of motion prescribe that circular motion shown by the girl, can only be accomplished if a person/object is pulling at her from the other side with equal but opposite force. If we assume that the other person/object is at an equal distance r from the center, then it must be rotating at the same speed. If we know the force, the distance from the center and the speed we can calculate the mass of the invisible person/object. Thus from the photograph we can deduce the weight of the invisible girl as a function of the assumed length of her arms.

Astronomers performed the same experiment, except they were measuring the rotation speed of stars in spiral galaxies [117], an example of which is shown in Figure 1(c). By counting the number of stars (measuring the amount of light) in a galaxy as a function of distance r from the galactic center, they can calculate the mass distribution. Again on account of Newton and his second great contribution to physics in the form of the law of universal gravitation, astronomers can calculate the gravitational force each star is feeling, and thus predict the speed with which the stars should be rotating around the galaxy center. To the great surprise of astronomers, the measured rotation speeds show a large discrepancy with the predicted rotation speeds. The explanation that requires the least adjustment to the laws of physics is that there is a substantial amount of matter away from the center of the galaxy, that is unaccounted for by measuring the amount of light. Hence the astronomers face the problem we saw in Figure 1(b), where an



(a) (b)

(c)

Figure 1: A familiar example of two girls twirling around each other (a), while in (b) one of the girls has been made invisible. An example of a spiral galaxy (c) called NGC 1300.

invisible object must be inferred to explain the rotation of the visible objects. This new form of matter was christened 'dark matter', though probably 'transparent' or 'invisible' matter would have been a better name for it.

Supersymmetry to the rescue

The theory of fundamental particles and their interactions, known as the Standard Model of particle physics was established in the 1960s and 1970s. All the particles observed so far in nature and in accelerator experiments can be explained by the Standard Model, but none of them fit the description of dark matter. That is why dark matter is considered to be a 'new' form of matter, or a form of matter that goes beyond the Standard Model.

One of the attractive theoretical extensions of the Standard Model is the theory of *Supersymmetry*. Not only does it give a credible dark matter particle candidate, but it also solves other theoretical issues connected to the Higgs particle. Before we can discuss Supersymmetry,

we must introduce the quantum mechanical concept of *spin*.

Spin Just like electric charge, *spin* is a fundamental characteristic of each elementary particle. This means that it is a specific and immutable property of the particle that can not be altered. All electrons in nature have an electric charge of -1, but besides that they also have spin[1] equal to 1/2.

In contrast to classical mechanics, where observables such as speed can take on any values in a continuous range, in quantum mechanics observables such as spin can take on only very specific values. This contra-intuitive behavior is called the quantization of observable quantities. Spin for example can take on values only as a non-negative integer multiple of 1/2, hence allowed values are: 0, 1/2, 1, 3/2, 2, etc.

Fermions and bosons Besides the quantization of spin, quantum mechanics also classifies all particles in two distinct clans. The clan of particles with half-integer (1/2, 3/2, etc.) spin are called *fermions*, while the other clan of integer (0, 1, etc.) spin particles are called the *bosons*. The spin-statistics theorem tells us, that fermions and bosons behave differently when grouped together. If two fermions occupy the same physical space, at least one property such as spin orientation must be different. Without this law of nature, the world around us would look very different, as more than two electrons would be able to occupy the lowest energy state of an atom, hence completely transforming the periodic table of elements and all of chemistry. Bosons on the other hand are 'social' particles, as the same physical space can be occupied by multiple bosons.

All the elementary *matter* particles, such as electrons (leptons and quarks), in the Standard Model are fermions, while all the *force carrier* particles (photon, gluon, $W^{\pm}/Z$) are bosons. It is not understood why there should be only fermionic matter particles in nature. The theory of Supersymmetry answers this question by introducing a new symmetry, that says that for every Standard Model particle there exists a supersymmetric partner that differs only by half an integer in spin. Thus according to Supersymmetry there are also bosonic matter particles and fermionic force carrier particles in nature. The rules of Supersymmetry prescribe that each supersymmetric particle must decay into one other supersymmetric particle and one (or more) non-supersymmetric particle(s). As the *lightest supersymmetric particle* (LSP) does not have a supersymmetric particle to decay to, if it obeys the above rule, it is stable by construction. This LSP is considered to be a good candidate for the dark matter particle.

Because we have never measured the supersymmetric partner of an electron, we say that the symmetry is not perfect but slightly broken, like a bent mirror in a fun house. Thus the supersymmetric partner of an electron is not only different in spin, but it is also much heavier. From cosmological dark matter measurements and theoretical calculations we have clues that suggest that we should be able to produce supersymmetric particles at the Large Hadron Collider, if Supersymmetry indeed is responsible for the dark matter.

The ATLAS experiment and Supersymmetry

The Large Hadron Collider (LHC) is a particle accelerator at the European Center for Nuclear Research (CERN), close to the city of Geneva. It is a machine situated 100 meters under the ground and has a circumference of 27 km. The LHC is designed to collide protons at an energy and intensity that surpasses all previous accelerators by orders of magnitude. At one of the LHC collision points, the ATLAS detector has been built to look for among others supersymmetric

[1]Here we use the so called 'natural units', where the Planck constant $\hbar$ and the elementary charge e are set to 1. This simplifies our notation, as spin 1/2 would otherwise have to be written as $1/2\hbar$ and $-1e$ would be the charge of the electron.

particles.

Because the supersymmetric particles are so heavy, we need more energy to create them, as the famous equation by Einstein shows $E = mc^2$. But because they are so heavy, they are also unstable and they decay. Each decay chain of a supersymmetric particle ends with the absolutely stable LSP. This LSP is weakly interacting, which means that it will fly through the ATLAS detector without leaving a trace. The only way we can find out that it has escaped detection, is by looking at the energy and momentum of all the other particles produced in the event. We know that energy and momentum must be conserved, which means that if we correctly add up all the detected particles and energy depositions, we should find a disbalance, as a lot of energy has been taken away by the LSP. The amount and the direction of this disbalance is called the *missing energy* of a collision, which is one of the most clear signatures of supersymmetric particles in ATLAS. Besides high missing energy, supersymmetric events in ATLAS are also characterized by high *transverse mass*.

Detection of Supersymmetry in ATLAS by the method described in this thesis, rests upon finding a statistically significant excess of events with high missing energy and high transverse mass. Also in events where heavy Standard Model particles are created, such as the top quark, missing energy is a characteristic signature. However this is well described by theory, thus we can model these background events and the missing energy/transverse mass distributions we expect to see, if no supersymmetric particles are produced in the collisions.

We estimate the Standard Model backgrounds in the low missing energy and low transverse mass regions by fitting the model taken from theory. Then we extrapolate the background model to the *signal* region, with high missing energy and high transverse mass, to predict an expected number of Standard Model events. If we measure a large excess of events above this background model, we have discovered some 'new' (beyond the Standard Model) physics. To find out whether we see supersymmetric particles or something completely different will take years, but the thrill of discovery will be overwhelming, as we might finally be able to explain what 23% of the universe is made of.

At the time of writing of this thesis, first results on searches for Supersymmetry with ATLAS have been published [118, 119], using the dataset acquired during 2010. No evidence for supersymmetric particles has been found, but the expectations are high for the dataset of 2011/2012, that is anticipated to be orders of magnitude larger. The larger dataset will give us excess to even more rare events, where the energy of the collision is even higher, possibly high enough to produce the heavy supersymmetric particles.

Samenvatting

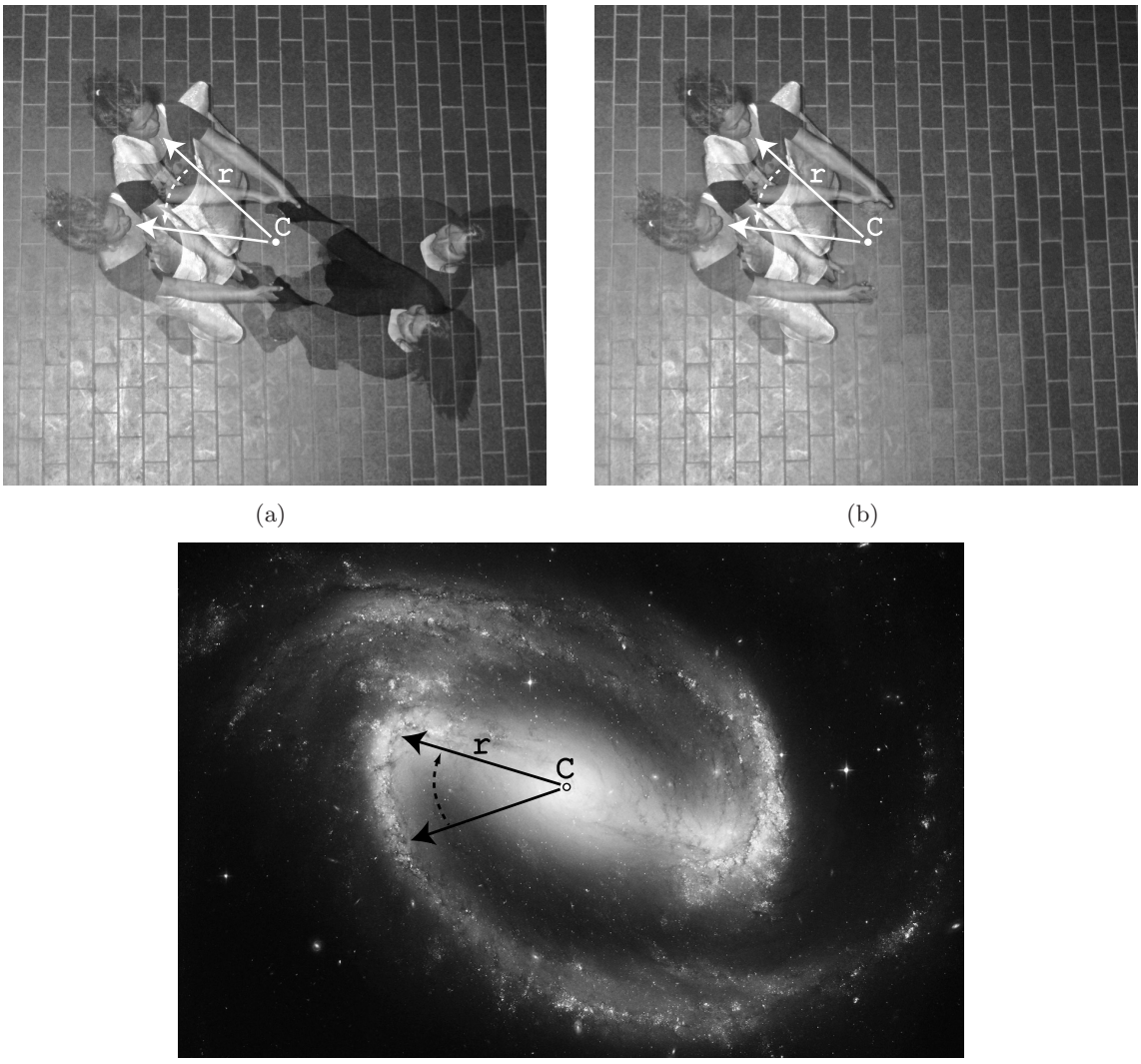
De aard van *donkere materie* deeltjes is een van de grote vraagstukken van hedendaagse natuurkunde. In tegenstelling tot zichtbare materie waar de sterren, de aarde en de aardbewoners uit bestaan, straalt donkere materie geen licht uit en verstrooit ook geen licht. Donkere materie gaat op geen enkele manier interactie aan, behalve door zwaartekracht. Uit metingen aan het heelal hebben de sterrenkundigen vastgesteld dat zichtbare materie alleen maar 4% van de energiedichtheid van het universum bevat, terwijl donkere materie ongeveer 23% voor zijn rekening neemt. Dus de fysici weten niet waar een groot gedeelte van het universum uit bestaat, een behoorlijk gênante positie om zich in te vinden anno 21ste eeuw.

Hoe weten we dat donkere materie bestaat?

Metingen aan rotatie curves van sterrenstelsels geven een van de duidelijkste indicaties van het bestaan van donkere materie. Maar laten we eerst beginnen met een bekend voorbeeld, twee vriendinnen die om elkaar heen draaien zoals afgebeeld in Figuur 1(a). Het *massamiddelpunt* waar ze omheen roteren, schematisch weergegeven als C in Figuur 1(a), ligt precies in het midden als de twee vriendinnen hetzelfde gewicht hebben. Net als in een centrifuge, hangt de snelheid waarmee ze draaien af van de lengte van hun armen, die in Figuur 1(a) wordt weergegeven als afstand r ten opzichte van het massamiddelpunt. Als we het gewicht, de lengte van hun armen en de kracht die ze gebruiken om te draaien kennen, dan kunnen we de rotatiesnelheid berekenen met behulp van de bewegingswetten van Newton, die hij in 1687 voor het eerst heeft gepubliceerd. Zouden de vriendinnen meer kracht zetten, dan zou de rotatiesnelheid toenemen. Zouden ze zware rugzakken omdoen maar evenveel kracht zetten, dan zou de rotatiesnelheid afnemen.

Laten we nu een andere situatie beschouwen, waarbij een van de vriendinnen onzichtbaar is geworden, zoals in Figuur 1(b) te zien valt. De vraag die we willen beantwoorden is, wat kunnen we over de onzichtbare vriendin te weten komen als we alleen van de foto kunnen uitgaan. Uit Figuur 1(b) kunnen we de snelheid waarmee de zichtbare vriendin draait afleiden, net als ook waar het massamiddelpunt C ligt. Omdat we van de zichtbare vriendin het gewicht en de rotatiesnelheid kennen, weten we hoeveel kracht zij zet. De bewegingswetten schrijven voor dat cirkelbeweging, zoals die vertoond wordt door de zichtbare vriendin, kan alleen tot stand worden gebracht wanneer een ander persoon of object aan haar trekt met een evenredige maar tegenovergestelde kracht. Als we aannemen dat de andere persoon zich op een even grote afstand r bevindt ten opzichte van C, dan moet de andere persoon even snel draaien. Maar als we de kracht, de afstand tot het massamiddelpunt en de snelheid weten, dan kunnen we de massa van de onzichtbare persoon berekenen. We kunnen dus het gewicht van de onzichtbare vriendin afleiden uit de foto, onder een aanname over de lengte van haar armen.

Sterrenkundigen hebben een vergelijkbaar experiment uitgevoerd, waarbij ze de rotatiesnelheden van sterren in spiraalvormige sterrenstelsels hebben gemeten [117]. Een voorbeeld van een spiraalvormige sterrenstelsel is weergegeven in Figuur 1(c). Door het aantal sterren te tellen (hoeveelheid licht te meten), kunnen de sterrenkundigen de massa verdeling in een sterrenstelsel berekenen. Alweer dankzij Newton en zijn tweede grote bijdrage aan de natuurkunde



(a) (b)

(c)

Figuur 1: Een bekend voorbeeld van twee om elkaar draaiende vriendinnen (a), waar in (b) een van de meisjes onzichtbaar is gemaakt. Een voorbeeld van een spiraalstelsel (c), genaamd NGC 1300.

in de vorm van de algemene gravitatiewet, kunnen ze ook berekenen hoeveel zwaartekracht elke ster zal voelen, en dus ook een voorspelling doen over de rotatiesnelheid van de ster om het centrum van de sterrenstelsel. Tot hun grote verbazing ontdekten de sterrenkudigen dat de gemeten snelheden een groot verschil laten zien ten opzichte van de voorspelde snelheden. De verklaring die de minste aanpassing van de wetten van fysica nodig heeft, is dat een aanzienlijke hoeveelheid materie zich buiten het centrum van de sterrenstelsel bevindt, die niet wordt meegenomen als we de hoeveelheid licht meten. De sterrenkundigen worden dus geconfronteerd met hetzelfde probleem dat wij zagen in Figuur 1(b), waar een onzichtbaar object verondersteld moet worden om de rotatie van de zichtbare objecten te verklaren. Deze nieuwe vorm van materie werd 'donkere materie' genoemd, hoewel 'transparante' of 'onzichtbare' materie de lading waarschijnlijk beter zou dekken.

Supersymmetrie als oplossing

De wisselwerkingen tussen fundamentele deeltjes worden beschreven door het Standaard Model, een theorie ontwikkeld in de jaren 60 en 70 van vorige eeuw. Alle deeltjes die in de natuur en bij versnellers zijn gemeten kunnen verklaard worden vanuit het Standaard Model, maar geen enkel deeltje voldoet aan de eisen voor de beschrijving van donkere materie. Dit is ook de reden waarom donkere materie als een 'nieuwe' vorm van materie wordt beschouwd, of een materie vorm van buiten het Standaard Model.

Verschillende uitbreidingen van het Standaard Model zijn op de markt, waarvan *Supersymmetrie* het meest theoretisch aantrekkelijk lijkt. Het geeft een geloofwaardige kandidaat deeltje voor donkere materie, maar lost tegelijkertijd een aantal andere theoretische problemen op, die te maken hebben met het Higgs mechanisme. Voordat we SUSY kunnen bespreken, moet we eerst het kwantummechanische concept van spin introduceren.

Spin Net als elektrische lading is *spin* een fundamentele kenmerk van elk elementair deeltje. Dit betekent dat het een specifieke en onveranderlijke eigenschap van een deeltje is. Alle elektronen in de natuur hebben een elektrische lading van -1, maar ze hebben ook allemaal een spin[2] van 1/2.
In tegenstelling tot de klassieke mechanica, waarin fysische grootheden als snelheid continue variabelen zijn, kunnen in de kwantummechanica grootheden als spin alleen hele specifieke waarden aannemen. Dit contra-intuïtief gedrag heet de quantisatie van observabele grootheden. Spin bijvoorbeeld kan alleen waarden aannemen die een niet-negatief veelvoud zijn van 1/2, dus toegestane waarden zijn: 0, 1/2, 1, 3/2, 2, etc.

Fermionen en bosonen Behalve quantisatie van spin, verdeelt kwantummechanica alle deeltjes ook in twee stammen. De stam van deeltjes met halftallige (1/2, 3/2, etc.) spin heet *fermionen*, terwijl de heeltallige (0, 1, etc.) spin deeltjes behoren tot de stam van *bosonen*. De stelling van spin statistiek vertelt ons dat fermionen zich anders gedragen dan bosonen als ze gegroepeerd worden. Als twee fermionen dezelfde fysische ruimte innemen, dan moet minstens een van hun eigenschappen, zoals spin richting, anders zijn. Als deze natuurwet niet zou gelden, dan zou de wereld om ons heen er heel anders uit zien, aangezien meer dan twee elektronen dan in de laagste energie toestand om een atoom zouden kunnen draaien. Dit zou het periodiek systeem en scheikunde compleet transformeren. Bosonen zijn aan de andere kant 'sociale' deeltjes, omdat het toegestaan is aan meerdere bosonen om dezelfde fysische ruimte in te nemen.

Alle elementaire *materie*-deeltjes zoals elektronen (leptonen en quarks) in het Standaard Model zijn fermionen, terwijl alle *kracht*-deeltjes (foton, gluon, $W^{\pm}/Z$) bosonen zijn. Het is niet begrepen waarom er alleen maar fermionische materie deeltjes bestaan. De theorie van Supersymmetrie geeft een antwoord op deze vraag door een nieuwe symmetrie te introduceren, die zegt dat voor elk Standaard Model deeltje er een supersymmetrische partner bestaat die alleen een half in spin verschilt. Dus volgens Supersymmetrie bestaan er wel degelijk bosonische materie- en fermionische kracht-deeltjes in de natuur. De regels van Supersymmetrie shrijven verder voor dat alle supersymmetrische deeltjes horen te vervallen in een ander supersymmetrisch deeltje en èèn (of meerdere) Standaard Model deeltje(s). Als alle supersymmetrische deeltjes aan deze regel voldoen, dan bestaat er een absoluut stabiel supersymmetrisch deeltje. Namelijk het *lichste supersymmetrische deeltje* (LSP uit het Engels), omdat er geen lichtere supersymmetrische deeltje nog is om naartoe te vervallen. De LSP wordt algemeen beschouwd als een goede kandidaat voor donkere materie.

[2]Ter vergemakkelijking gebruiken we hier de zogenaamde 'natuurlijke eenheden', waar we de Planck constante $\hbar$ en de elektrische constante e gelijk aan 1 stellen. Anders zouden we spin als $1/2\hbar$ en elektrisch lading als $-1e$ moeten opschrijven.

Omdat we nog nooit een supersymmetrische partner van een elektron hebben waargenomen, zeggen we dat de symmetrie lichtelijk gebroken is, zoals een lachspiegel op een kermis. De supersymmetrische partner van een elektron heeft dus niet aleen een andere spin, maar is ook veel zwaarder. Vanuit de kosmologie en de theoretische berekeningen hebben we aanwijzingen die suggereren dat we bij de Large Hadron Collider supersymmetrische deeltjes zouden moeten kunnen produceren, mits Supersymmetrie voor donkere materie verantwoordelijk is.

De ATLAS detector en Supersymmetrie

Vanaf maart 2010 zijn de ogen van de wereld gericht op het Europees Centrum voor Nucleair Onderzoek (CERN), waar de Large Hadron Collider (LHC) protonen op elkaar botst met de hoogste energie ooit dat door een mens gebouwd apparaat is bereikt. De botsingen vinden plaats bij 7 biljoen elektronvolt (eV) en een intensiteit die alle eerdere versnellers ver overtreft. Op een van de punten waar de deeltjes op elkaar botsen is de ATLAS detector gebouwd, om onder andere naar supersymmetrische deeltjes te zoeken.

Omdat de supersymmetrische deeltjes zo zwaar zijn, hebben we meer energie nodig om ze te creëren, zoals de befaamde vergelijking van Einstein $E = mc^2$ laat zien. Maar tegelijkertijd omdat ze zwaar zijn, zijn supersymmetrische deeltjes instabiel en vervallen ze, waarbij elk vervalsketen eindigt met het absoluut stabiele LSP. De LSP gaat zo weinig interactie met materie aan dat het dwars door de ATLAS detector vliegt. De enige manier om uit te vinden dat het niet is gedetecteerd, is door te kijken naar de energie- en impuls-balans van alle andere deeltjes in het event. Aangezien we weten dat energie en impuls behouden grootheden zijn, zouden we door het optellen van alle gemeten deeltjes en energie afzettingen een energie-disbalans moeten meten. De hoeveelheid en de richting van deze disbalans noemen wij de *missende energie* ($E_{\mathrm{T}}^{\mathrm{miss}}$) van een botsing, en hoge missende energie is een van de duidelijkste signaturen van supersymmetrische deeltjes in ATLAS. Behalve door hoge missende energie, worden supersymmetrische events in ATLAS ook gekenmerkt door hoge *transversale massa* (M_{T}).

De methode beschreven in dit proefschrift om bij ATLAS Supersymmetrie te ontdekken berust op het meten van een statistisch significant overschot aan events bij hoge missende energie en hoge transversale massa. Ook bij events waar zware Standaard Model deeltjes, zoals de top quark, worden geproduceerd is missende energie een karakteristieke signatuur. Echter dit wordt goed beschreven door theorie, waardoor we een model kunnen opstellen van de $E_{\mathrm{T}}^{\mathrm{miss}} - M_{\mathrm{T}}$ verdelingen voor deze achtergrond events, die we verwachten te zien als er geen supersymmetrische deeltjes worden geproduceerd.

De gecombineerde fit methode schat alle Standaard Model achtergronden af door het theoretisch geinspireerd model te fitten aan data in de lage $E_{\mathrm{T}}^{\mathrm{miss}} - M_{\mathrm{T}}$ regio. Daarna wordt het achtergrond model geextrapoleerd naar het *signaal* regio in hoge $E_{\mathrm{T}}^{\mathrm{miss}} - M_{\mathrm{T}}$, om het aantal verwachte Standaard Model events te berekenen. Als we een groot overschot meten aan events bovenop het achtergrond model, dan kunnen we spreken van ontdekking van 'nieuwe' (buiten het Standaard Model) fysica. Het zal daarna jaren duren voordat we met zekerheid kunnen zeggen of we supersymmetrische deeltjes of compleet iets anders hebben gemeten, maar de sensatie van een ontdekking zal overweldigend zijn, aangezien we dan eindelijk kunnen uitleggen waar 23% van het heelal uit bestaat.

Tijdens het schrijven van dit proefschrift, zijn de eerste resultaten van de zoektocht naar Supersymmetrie bij ATLAS gepubliceerd [118, 119] met de data verkregen in 2010. Er is nog geen bewijs van supersymmetrische deeltjes gevonden, maar de verwachtingen zijn hoog gespannen voor de data van 2011/2012, die orders groter verondersteld worden. Met een grotere hoeveelheid data kunnen we nog zeldzamere events bestuderen, waar de botsingsenergie nog hoger zal zijn, misschien wel hoog genoeg om de zware supersymmetrische deeltjes te produceren.

Acknowledgements

Many years have passed since I first came to Nikhef for my bachelor project in 2003, but time flies when you are having fun. Many people have contributed to this sense of time flying by, that I will try to thank in here shortly.

First and foremost, I must dedicate this book to my parents. Not only were they always there with help and advice, they sacrificed their lives and their country so that me and my brother had a better life. If there is one decision that changed my life and made this thesis possible, then it is my parents moving to Holland, for which I deeply bow to them. Besides my parents I must thank my grandmother, that always kept a watchful eye on me and reminded me to stay myself. Also I would like to thank my brother and the rest of the family for encouraging me.

My academic life would not have been the same without the Nikhef ATLAS group. Especially I must thank my supervisor, Wouter Verkerke, who is the most complete physicist I have ever met, with an amazing understanding of the theoretical, statistical and computational challenges of our work. Also my gratitude goes out to my advisor, Nicolo de Groot, who always stayed enthusiastic, as well as professors Stan Bentvelsen and Paul de Jong who guided me through my PhD at Nikhef. There is one more person that I need to thank here, that is Peter Kluit, for all those hours spent debugging, explaining about and showing me around in the world of the muon at ATLAS.

Now come the really important people, the PhDs and postdocs that make life at Nikhef come alive. The list is long and I apologize to the people I surely forgot: Aras, Patrick, Marcello, Duncan, Joana, Zdenko, Gustavo, Caroline, Gordon, Gossie, Manouk, Manuel, Erik, Eric, Menelaos, Jörg, Faab, Daan, Robin, Serena, Chiara, Priscilla, Hegoi, Tristan, Egge, Giuseppe, Barbara, Gijs, John, Ido, Nicole and Corey. Thank you all for the good times at breaks, outings, meetings and beyond! I owe Jochem a great debt for: appropriately introducing me to life as a PhD by falling asleep in his office chair on my first working day, familiarizing me with bb5 and the ATLAS pit while installing muon chambers (also thanks to Gerrit and Rene), and spending hours answering questions about the complex world of muon software. My partner in (s)crime, pardon the lack of humor in physics, Folkert Koetsveld gets my hat off for our year-long struggles thRoough the math and the mess of a SUSY analysis. Additionally, I want to thank Michele Consonni and Marcel Raas for the exciting work we did together on the very first real LHC data. The twelve months I spent at CERN were made even more exciting by the BPs, for which I must thank Ohm, Jack, Eddie, Gabe, Cookey and the extended family of awesomeness. The unavoidable attraction of the snow slopes was enhanced by the great company of Carolina, Nicola and the CERN ski club.

I must start a whole new paragraph to emphasize the special mention for my two wingmen, my colleagues and my friends, Sipho van der Putten and Alexander Doxiadis. The last year would have been much harder if you guys would not have been around to talk to, help out and stay at work later than should be allowed. Who's afraid of a bench full of professors, when two such paranymphs are at your side. Also, I must thank a few people that assisted in the very last never ending list of bits that needed doing, so thanks go to Ahkin, Mateus, Wolfi, Przemek, Irene, Kees Huyser, Auke Korporaal.

Beyond any doubt I must mention here my friends from the 'outside' world. The ones that

know me the longest, the ones that keep on amazing me, the ones that always kept it real: Mali (v.v.kaboutah), Mark (sinush/tata), Nikita (zloi), Johannes (TAFKAJLA), Potlood (P.P), Matteo (op ouwen!), Marcel (mod), Esmir, Noël. I drink to that, but most of all...I drink to my friends! (if that will ever happen). In addition, I want to thank the comrades of the International Socialists, cause they know what's up. Lastly, two people that have travelled with me all over the globe, spent countless hours having to listen to awful (opening) bands and shared the excitement of hearing great new music, partied and farted with me, met an incredible amount of awesome people with me, the two still original members of Vitamin X, Marko and Marc, we have had a great run, let us keep on rolling.

As one should, I leave the best for last. Solkin has seen the worst of the last year of writing this thesis, the cranky moods, the long hours, the short temper, but she still stayed by my side. I do sincerely hope that now comes the good part, but looking back, for me all our times together, the good, the bad and the ugly have been pinche great, si no? *Te queiro mucho!*