Universität Bonn

Physikalisches Institut

Search for Supersymmetry in final states with jets, missing transverse momentum and at least one τ lepton with the ATLAS Experiment

Michel Janus

The first search for Supersymmetry (SUSY) in final states with at least one τ lepton, two or more jets and large missing transverse energy based on proton-proton collisions recorded with the ATLAS detector at the LHC is presented. The τ leptons are reconstructed in the hadronic decay mode. To search for new physics in final states with hadronic τ-lepton decays a reliable and efficient reconstruction algorithm for hadronic τ decays is needed. As part of this thesis two existing τ-reconstruction algorithms were further developed and integrated into a single algorithm that has by now become the standard algorithm for τ reconstruction in ATLAS. The suppression of jet background, one of the crucial aspects of τ reconstruction and the SUSY analysis in this thesis, was studied and the probabilities of misidentifying quark- or gluon-initiated jets as hadronic τ-lepton decays were measured in both the 2010 and 2011 ATLAS data at a centre-of-mass energy of $\sqrt{s} = 7$ TeV, using samples of di-jet events. The ATLAS data used for the SUSY search in this thesis was recorded between March and August 2011 and corresponds to an integrated luminosity of 2.05 fb^{-1}. Eleven events are observed in data, consistent with the total Standard Model background expectation of 13.2 ± 4.2 events. As no excess of data over the expected backgrounds is observed, 95% confidence level limits are set within the framework of gauge mediated SUSY breaking (GMSB) models as a function of the GMSB parameters Λ and $\tan\beta$, for fixed values of the other GMSB parameters: $M_{\text{mess}} = 250$ TeV, $N_5 = 3$, sign(μ) = + and $C_{\text{grav}} = 1$. In addition to the GMSB interpretation, a model-independent upper limit of 8.5 on the number of events from potential non-Standard Model sources is derived at the 95% confidence level.

Physikalisches Institut der
Universität Bonn
Nußallee 12
D-53115 Bonn

BONN-IR-2013-01
April 2013
ISSN-0172-8741

Universität Bonn

Physikalisches Institut

Search for Supersymmetry in final states with jets, missing transverse momentum and at least one τ lepton with the ATLAS Experiment

Michel Janus

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen und ist auf dem Hochschulschriftenserver der ULB Bonn `http://hss.ulb.uni-bonn.de/diss_online` elektronisch publiziert.

1. Gutachter:	Prof. Dr. Jochen Dingfelder
2. Gutachterin:	Prof. Dr. Klaus Desch
Tag der Promotion:	13.12.2012

Abstract

A search for Supersymmetry (SUSY) in final states with at least one τ lepton and large missing transverse energy based on proton-proton collisions recorded with the ATLAS detector at the LHC is presented. The τ leptons are reconstructed in the hadronic decay mode. Final states with τ leptons offer a good sensitivity for SUSY models where the coupling to the third generation of fermions is enhanced, e.g. gauge mediated SUSY breaking (GMSB) models, for which the supersymmetric partner of the τ lepton is the next-to-lightest SUSY particle (NLSP) and its decay to electrons or muons is strongly suppressed.

To search for new physics in final states with hadronic τ-lepton decays a reliable and efficient reconstruction algorithm for hadronic τ decays is needed to separate real τ decays and backgrounds from quark- or gluon-initiated jets and electrons. As part of this thesis two existing τ-reconstruction algorithms were further developed and integrated into a single algorithm that has by now become the standard algorithm for τ reconstruction in ATLAS. The suppression of jet background, one of the crucial aspects of τ reconstruction and the SUSY analysis in this thesis, was studied and the probabilities of misidentifying quark- or gluon-initiated jets as hadronic τ-lepton decays were measured in both the 2010 and 2011 ATLAS data at a centre-of-mass energy of $\sqrt{s} = 7$ TeV, using samples of di-jet events.

The ATLAS data used for the SUSY search in this thesis was recorded between March and August 2011 and corresponds to an integrated luminosity of 2.05 fb^{-1}. Events are selected that contain at least two jets with high transverse momentum, one identified hadronic τ-lepton decay and large missing transverse energy. The Standard Model background processes that contribute most to the selected event sample are W- and Z-boson production in association with jets as well as top quark production. Eleven events are observed in data, consistent with the total Standard Model background expectation of 13.2 ± 4.2 events. As no excess of data over the expected backgrounds is observed, 95% confidence level limits are set within the framework of GMSB models as a function of the GMSB parameters Λ and $\tan\beta$, for fixed values of the other GMSB parameters: $M_{\text{mess}} = 250\,\text{TeV}$, $N_5 = 3$, $\text{sign}(\mu) = +$ and $C_{\text{grav}} = 1$. In addition to the GMSB interpretation, a model-independent upper limit of 8.5 on the number of events from potential non-Standard Model sources is derived at the 95% confidence level. This limit corresponds to an upper limit on the visible cross section of 4.0 fb, where the visible cross section is defined as the product of production cross section of the non-Standard Model process, its branching fraction to at least one τ lepton, acceptance, and efficiency of the event selection applied. These limits are the first results of SUSY searches at the LHC that include events with one hadronic τ-lepton decay and significantly extend those placed by previous experiments at LEP and the Tevatron.

Contents

Chapter 1

Introduction

The story of particle physics over the last decades has been one of astounding success. Ever since the formulation of the Higgs mechanism [1] and its inclusion in the theoretical framework of electroweak interactions by Glashow, Weinberg and Salam [2–4] in 1967 there have been no fundamental changes to the "Standard Model" (SM) of particle physics. Since then most observations from experiment, including new particles like the W boson, the Z boson or the top quark, were predicted well in advance, contrary to the time of the infancy of particle physics, in which the discovery of new particles was greeted with surprise[1]. So it is not surprising that many physicists believe that the new boson recently discovered at CERN [5, 6] is the last, missing particle predicted by the Standard Model, the Higgs boson. Finding the Higgs particle was one of the main motivations for the physics program of the Large Hardon Collider (LHC) at CERN in Geneva, Switzerland.

While the Higgs boson offers a mechanism of electro-weak symmetry breaking and makes the Standard Model internally consistent, with respect to particle masses, this model can only be understood as an intermediate step in our understanding of the laws of particle physics. The Standard Model, being an "effective field theory" is heavily dependent on the energy scale up to which the calculations are carried out and it is clear that new physical phenomena will enter at some point. The ultimate scale at which this will eventually happen is the so-called Planck scale, at which gravity is expected to become comparable in strength to the forces described by the Standard Model. Since gravity is not yet included in the Standard Model, the theory will have to be generalized to include it. Unfortunately the Planck scale is - and will probably forever be - outside the reach of "physics in the laboratory" and the only hints at what physics at this scale may look like come from astronomical observations, which are inherently dominated by gravitational effects.

There are however some compelling reasons to believe that new physics might already appear at much lower scales that have become experimentally accessible in recent years with the start of proton-proton collisions in the LHC experiments at a centre-of-mass energy of $\sqrt{s} = 7$ TeV in 2010-11 and the increase in energy to $\sqrt{s} = 8$ TeV in 2012. Especially the CMS and ATLAS experiments, whose designs were partly motivated by the search for the Higgs Boson, are also capable of detecting many of the signatures associated with physics beyond the Standard Model, assuming these new physics processes appear at the TeV scale. One well-motivated extension of the Standard Model is the idea of so-called Supersymmetry (SUSY), which introduces a symmetry between fermionic and bosonic particles. As a result the particle content of the theory would at least double. Depending on the parameters of the theory, SUSY would be able to account for some of the largest shortcomings of the Standard Model, e.g. the fine-tuning problem of the Higgs boson mass, the non-unification of the gauge couplings and the so far unexplained existence of dark matter in the universe. While SUSY is by no means the only idea on the market, the fact that Supersymmetry is a necessary ingredient to many so-called "grand unified theories" makes it an interesting candidate for an extension to the Standard Model. Another important reason why TeV-scale

[1] a famous quote to this effect, which is attributed to Isidor Rabi concerning the discovery of the muon: "Who ordered that?"

SUSY is an interesting field of study, is that in order for SUSY to solve the fine-tuning problem of the Higgs boson mass, the masses of the SUSY particles may not exceed a few TeV.

While several studies [7–10] have already been published by the ATLAS and CMS collaborations, documenting "SUSY searches" at the LHC, most of these have focused on topologies without any leptons or with only light leptons[2]. Final states with hadronic τ-lepton decays pose a greater challenge, because of the difficult task of separating the hadronic τ-lepton decays from possible background signatures from hadronic jets or light leptons. These final states are, however, a very interesting probe for SUSY, because in many SUSY scenarios the decay of SUSY particles to third-generation fermions and therefore to τ leptons is significantly enhanced compared to the fermions of the first and second generation. Therefore several studies of τ-lepton identification at the ATLAS experiment have been performed as part of this thesis, before embarking on the search for SUSY in τ-lepton final states. More recently, I contributed to the first published measurement at the LHC of a SUSY search in final states with one or more τ leptons [11]. At the same time, the ATLAS collaboration published results from a search for SUSY in final states with more than one hadronic τ-lepton decay per event [12] and the CMS collaboration published a search in events where a single hadronic τ-lepton decay occurs in combination with light leptons [13].

This thesis is structured as follows: Chapter 2 gives a short overview of the theoretical concepts needed to appreciate the context of SUSY searches in τ-lepton final states. This includes a short summary of the Standard Model, electroweak symmetry breaking and physics at hadron colliders. The fundamental concepts of SUSY are also discussed in this chapter. In Chapter 3, the ATLAS Experiment is introduced, followed by a description of the physics objects that are reconstructed from the signals recorded with the ATLAS detector in Chapter 4. The reconstruction of the hadronic decay modes of τ leptons at the ATLAS Experiment is described in detail in Chapter 5. One crucial aspect of the τ-identification performance, namely the measurement of the background efficieny of the τ-identification algorithms for hadronic jets, is further described in Chapter 6. Finally Chapters 7, 8 and 9 outline the application of the knowledge collected in the previous two chapters to a search for new physics beyond the Standard Model in τ-lepton final states. The results of this search in 2.05 fb^{-1} of ATLAS data taken between March and August of 2011 are presented in Chapter 10 and Chapter 11 gives a summary of all results gathered in this thesis. Chapter 12 concludes the body of this text with an outlook to future studies and applications of τ identification and searches for physics beyond the Standard Model in τ-lepton final states at the ATLAS Experiment.

[2] light leptons in this case meaning either electrons or muons

Chapter 2

Theoretical Overview

The central topic of this thesis is the search for phenomena beyond the Standard Model of particle physics with the ATLAS detector at the LHC. Since excellent reviews [14] of the Standard Model exist, the introduction given in Sec. 2.1 is not meant to be a comprehensive overview of the theory. Instead only the fundamental concepts of the theory are introduced, insofar as they are necessary for understanding the context of the measurements presented in Chapters 6-10 and their interpretation. Sec. 2.2 will then introduce SUSY as an attractive theoretical concept, which can remedy many of the shortcomings of the Standard Model. While an understanding of the Standard Model and SUSY are important for the interpretation of the results presented in this thesis, there are still several intermediate steps between these concepts and the actual theoretical input to the measurements performed. As the data, from which the results in this thesis are derived, were taken from proton-proton collisions, the concept of so-called parton distribution functions (PDF's), needed to predict the initial states of the partons involved in the collisions, is touched upon in Sec. 2.3. This section also explains the basics of the phenomenological framework of non-perturbative Quantum-Chromodynamics (QCD), which is necessary to describe the interactions of the scattering partons when their energy becomes too small to be properly described using perturbation theory. While both these topics are very important, the meat of the theoretical predictions about the different physics processes under study comes from so-called event generators, programs which predict the kinematics of the hard scattering process using probabilistic methods. The most important of these processes (both signal and background) in the context of this thesis are introduced in Sec. 2.4 together with the programs used to generate the predictions of their kinematics.

2.1 The Standard Model of Particle Physics

This introduction of the Standard Model starts with an overview of its particle content in Sec. 2.1.1 and continues to describe the interactions between these particles in Sec. 2.1.2.

2.1.1 Matter Particles of the Standard Model

One of the fundamental principles of the Standard Model is that all matter is made up of elementary particles, so-called fermions and that the interaction between these matter particles is mediated by another set of particles, called bosons. The full fermionic particle content of the Standard Model is listed in table 2.1. The fermions can be further subdivided by the interactions in which they participate. Quarks take part in the strong, electromagnetic and weak interactions, while leptons do not take part in strong interactions. Both quarks and leptons appear in pairs (or generations), of which there are three. In each generation there is one electrically charged lepton, one neutral lepton and two quarks. As the Standard Model is a quantum theory, the properties of the particles by which they can be distinguished are called quantum numbers. The electrical charge, which is an example of such a quantum number, is obviously

	Generation I	Generation II	Generation III
quarks	u	c	t
	d	s	b
leptons	e	μ	τ
	ν_e	ν_μ	ν_τ

Table 2.1: The fermionic particle content of the Standard Model. The full names of the charged leptons for the first, second and third generation are electron, muon (μ) and tau (τ) lepton, respectively, while the neutral leptons are all named neutrino with a prefix denoting the charged lepton in their generation, i.e. electron neutrino for the first generation. The names of the quarks are up, down, charm, strange, top and bottom and their symbols are simply the first letter of their name.

zero for neutral leptons, while the charged leptons carry one unit of negative charge. The quarks carry non-integer charge and are separated into two types that carry either a charge of $+2/3$ or $-1/3$, respectively. In the first generation, which comprises the building blocks of almost all matter in the universe, the charged lepton is called the electron, the neutral lepton is called the electron neutrino, and the quarks are called the up and the down quark. The names of the particles of the second and third generation of fermions are given in table 2.1.

Each first generation particle has one counterpart in the second and third generation, which has the same properties except for the mass. Every particle in the Standard Model also has another counterpart which is denoted its anti-particle and has the same mass as the particle but all additive quantum numbers inverted. While the leptons can be observed as free particles, the quarks always appear as bound states of one quark and one anti-quark (mesons) or of three quarks (baryons). The mesons and baryons together are called hadrons.

2.1.2 Particle Interactions in the Standard Model

In the Standard Model the interactions between two fermions are described by the exchange of intermediary particles with integer spin, so-called "bosons". The three interactions (or forces) described by the Standard Model are the electromagnetic, the strong and the weak interaction and they are all described in more detail in the following sections, together with their intermediary (or "gauge") bosons.

The Electromagnetic Interaction

The theory of electromagnetism, called Quantum Electrodynamics (QED), was the first properly formulated quantum field theory. In quantum field theory the equations of motion follow from minimisation of the action, which is the integral over time and space of the so-called Lagrangian density (sometimes simply called the Lagrangian). Therefore one can say that if the Lagrangian of a theory is known, all properties of the theory, including the particle content and their interactions, is determined. The simple Langrangian for a single free massive fermion, described by a 4-component spinor ψ (called a "Dirac field"), is of the form:

$$\mathcal{L}_{EM} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \tag{2.1}$$

where $\partial_\mu(\mu = 0, 1, 2, 3)$ is the partial derivate along the space-time coordinate x_μ and γ^μ are the 4×4 Dirac matrices. This Lagrangian density is invariant under global gauge transformations of the form $\psi \rightarrow e^{i\theta}\psi$, where "global" means that the transformation parameter θ does not depend on the

position in space-time. It is however not invariant under transformations which depend on the space-time position x like $\psi \rightarrow e^{i\theta(x)}\psi$, so-called "local" gauge transformations. If one wants to introduce this property to the Lagrangian density, this requires doing a gauge transformation of the partial derivative like $\partial_\mu \rightarrow \partial_\mu - iqA_\mu$. This new, so-called "covariant" derivative introduces a new "gauge field" A_μ. This gauge field is connected with the Dirac field ψ with a strength q. A term of the form $F^{\mu\nu}F_{\mu\nu}$, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, describes the dynamics of the gauge field A_μ. Note that adding a mass term for the gauge field would destroy the invariance of the Lagrangian, for which the gauge field itself was introduced in the first place. The full Lagrangian density now looks like:

$$\mathcal{L}_{EM} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi + q\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{2.2}$$

One can interpret q as the electrical charge of the fermion and the gauge field as the field of a massless bosonic particle, the photon, in which case this Lagrangian indeed describes the dynamics of an electrically charged particle in the presence of an electromagnetic potential. So the requirement of invariance of the Lagrangian of a free Dirac field under local gauge transformations has lead to the full field theory of electrodynamics. Since the local gauge transformations of the form $e^{i\theta}$ generate the group $U(1)$ of unitary one-dimensional transformations, it is also said that the underlying symmetry of electromagnetic interactions is the $U(1)_q$ symmetry, where q stands for charge. Two particles, which interact via exchange of a photon, can either attract or repel each other, depending on the relative sign of their charges.

The Strong Interaction

Because of the commensurate success of QED, the subsequent theories of the other interactions, e.g. the theory of the so-called "strong" interaction between quarks, are formulated in the framework of quantum field theories. As detailed above the Lagrangian of QED is determined by the underlying symmetry $U(1)_q$. When formulating the quantum theory of the strong interaction, also called Quantum Chromodynamics (QCD), it is natural to ask what underlying symmetry governs the theory. In QCD the defining symmetry is $SU(3)_c$, the group of special, unitary transformations in three dimensions and the c stands for color, which is just a name for the charge related to the strong interaction, as "electrical" was the prefix for the charge in QED. Just as in the case of QED, one can require the Lagrangian of a free particle of the theory, in this case denoted as $q(x)$, where q stands for quark, to be invariant under local $SU(3)_c$ gauge transformations of the form:

$$q(x) \rightarrow q'(x) = U(x)q(x) = e^{i\alpha_a(x)\frac{\lambda_a}{2}}q(x) \tag{2.3}$$

where $U(x)$ is an arbitrary 3×3 matrix which is parametrised using the eight Gell-Mann matrices λ_a [15], which are the generators of $SU(3)_c$. The Lagrangian density of QCD, which is invariant under these transformations is of the form:

$$\mathcal{L}_{QCD} = i\bar{q}_f\gamma^\mu D_\mu q_f - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{2.4}$$

The subscript f appears here as a reminder that all six quarks in the theory obey this Lagrangian. The covariant derivative D_μ is again constructed to preserve the underlying symmetry by adding a coupling term for the gauge field, in this case the gluon field:

$$D_\mu = \partial_\mu + ig_3\frac{\lambda_a}{2}G^a_\mu, \tag{2.5}$$

where g_3 is the coupling constant of the strong interaction. To complete the picture one also needs to give the transformation rule of the gauge field G^a_μ:

$$G^a_\mu \to G^a_\mu + \frac{1}{g_3}\partial_\mu \alpha_a - f_{abc}\alpha_b G^c_\mu. \tag{2.6}$$

The field tensor $G^a_{\mu\nu}$ is defined as:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_3 f_{abc} G^b_\mu G^c_\nu, \tag{2.7}$$

where f_{abc} are structure constants, which need to fulfill $[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}] = i f_{abc}\frac{\lambda_c}{2}$. This symmetry introduces three types of charge, which - in keeping with the color analogy - are called red, blue and green. As was the case in QED, particles can carry color and anticolor. By convention quarks are said to carry color, while antiquarks carry anticolor. Because (anti)quarks may change color when interacting strongly, the force carrier, which is called the gluon, must also carry color charge, namely one unit of color and one of anticolor. Having a charge carrying gauge boson gives rise to some rich phenomenology, which is not encountered in QED, as the gluon can interact with itself. This is due to the non-Abelian nature of $SU(3)$, meaning that the eight 3×3 Gell-Mann matrices, that represent the eight different gluons in the algebraic formulation of QCD, do not commute.

Another large difference, which was long thought to be an insurmountable problem of QCD, was that the coupling strength in QCD for low-energy interactions is numerically larger than one. As almost all observables of quantum field theory are calculated using perturbation theory in orders of the coupling strength, this proves a serious threat to the feasibility of QCD computations, because the perturbation series would not converge.

The notion of QCD was however saved by the discovery that the coupling strength of all elementary interactions is not fixed, but has a dependence on the amount of energy transferred between the two interacting particles. This phenomenon is called the "running" of the coupling constants. In the case of QCD the coupling grows weaker with increasing energy, making QCD "asymptotically free" [16, 17] and therefore allows for perturbative calculations of hard scattering processes. For low-energy transfer (also called "soft") interactions, the coupling constant becomes very large, which leads to the concept of "confinement".

The Weak Interaction and Electroweak Unification

The last interaction in the Standard Model is the so-called weak interaction. As the name suggests the coupling strength of the weak interaction is small at low momentum transfer compared to the strong or electromagnetic interactions. It follows therefore that weak processes are often overshadowed by the stronger effects of QED or QCD. However there are significant differences as to which processes are allowed in the three different interactions. While all interactions conserve the number of leptons from each generation, it was found that the weak force is the only interaction in which a charged lepton can be converted directly into its neutrino. It follows therefore that some of the bosons that mediate this weak interaction have to be charged. Indeed the weak interaction has three kinds of intermediary particles: the W^+, the W^- and the Z^0, of which only the first two are electrically charged. All three gauge bosons of the weak interaction are massive, with the mass of the two W bosons being $M_W = (80.399 \pm 0.023)$ GeV and the mass of the Z boson being $M_Z = (91.1876 \pm 0.0021)$ GeV [18]. Depending on which boson is exchanged in a weak interaction one speaks of charged or neutral current interactions. in 1967 Glashow, Weinberg and Salam [2–4] not only succeeded in describing the weak interaction as a field theory, but also in unifying it with the electromagnetic interaction using the algebraic form of an $SU(2)_I \times U(1)_Y$

gauge symmetry. This is achieved by relating the electrical charge of QED with the unified electroweak quantum numbers of weak hypercharge Y_W and weak isospin I_3 as follows:

$$e = I_3 + \frac{Y_W}{2}, \tag{2.8}$$

where Y_W is the weak hypercharge operator, the generator of $U(1)$, which was introduced in analogy to strong hypercharge. We know from experiment that the charged weak current couples differently to left- and right handed fermions, such that in the algebraic formulation left-handed fermions are assigned to $SU(2)_I$ doublets with weak isospin $I_3 = \pm\frac{1}{2}$, while the right-handed fermions are singlets with $I = 0$, which carry only weak hypercharge. From this follows that the $SU(2) \times U(1)$ gauge transformations for the right-handed singlet states (ψ_R) and left-handed doublet states (χ_L) are different:

$$\psi_R \rightarrow \psi'_R = e^{i\beta(x)Y}\psi_R, \tag{2.9}$$

$$\chi_L \rightarrow \chi'_L = e^{i\alpha_a(x)\tau_a}e^{i\beta(x)Y}\chi_L, \tag{2.10}$$

where $\alpha(x)$ and $\beta(x)$ are phases that depend on x and $\tau_a/2(a = 1, 2, 3)$ are the generators of $SU(2)$. The covariant derivative again needs to be constructed from the partial derivative ∂_μ:

$$D_\mu = \partial_\mu + igW^a_\mu\frac{\tau_a}{2} + ig'B_\mu\frac{Y}{2}, \tag{2.11}$$

where g is the coupling constant of the gauge field W^a_μ $(a = 1, 2, 3)$ that couples to $SU(2)_I$ weak isospin and g' is the coupling constant of the gauge field B_μ, which couples to $U(1)_Y$ hypercharge. With these one can again construct the Lagrangian which is invariant under local $SU(2) \times U(1)$ gauge transformations:

$$\mathcal{L}_{EW} = i\bar{\chi}^i_L\gamma^\mu D_\mu\chi^i_L + i\bar{\psi}^i_R\gamma^\mu D_\mu\psi^i_R - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{2.12}$$

and the dynamics of the gauge fields W^a_μ and B_μ are described by the field tensors

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\epsilon_{abc}W^b_\mu W^c_\nu, \tag{2.13}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.14}$$

where the structure constant of $SU(2)$ is the completely antisymmetric tensor ϵ_{abc} and the structure constant of $U(1)$ is zero as seen previously in the pure QED Lagrangian. This means that $SU(2)$ is non-Abelian just like $SU(3)$ and therefore the gauge field W^a_μ does have self-interaction terms in the Lagrangian, while the field B_μ does not.

The question arises how the gauge fields above relate to the observable bosons of the weak interaction, the W^+, W^- and the Z^0. Also the photon γ, or as it was introducted in the QED Lagrangian, the gauge field A_μ, needs to be accounted for. It turns out that the physically observable particles are indeed not the gauge bosons of the GWS theory, but they are linear combinations of them:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \tag{2.15}$$

$$Z_\mu = \cos\theta_W W^3_\mu - \sin\theta_W B_\mu \tag{2.16}$$

$$A_\mu = \sin\theta_W W_\mu^3 - \cos\theta_W B_\mu, \tag{2.17}$$

where a new observable θ_W (called the Weinberg or weak-mixing angle) was introduced, which relates the coupling strengths g and g':

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}. \tag{2.18}$$

Having these two coupling strengths, which determine the mixing of the theoretical gauge fields W_μ^3 and B_μ into the physically observable Z and γ, allows for unification of weak and electromagnetic interactions in one theoretical framework. Just as the charged weak interaction changes a charged lepton to its neutral counterpart, it can also change quarks from up- to down-type. But unlike the lepton case, where lepton number is conserved in each generation, the weak interaction allows for violations of quark generation (also called flavour). The framework of flavour violation was formulated by Cabbibo, Kobayashi and Maskawa [19], by introducing the notion that the mass eigenstates of the quarks are not eigenstates of the weak interaction, but are "skewed" in flavour space. In this picture the eigenstates of the weak interaction mix to make up the quarks which are eigenstates of the strong interaction. The 3×3 quark mixing matrix (also called CKM-matrix) [18] contains the mixing parameters which need to be determined from experiment.

While the $SU(2) \times U(1)$ model allows for a unified description of the electromagnetic and weak interactions, it is important to note, that again mass terms for the gauge bosons are not allowed, as they would break local gauge invariance. Including simple fermion mass terms of the form $m(\bar{\chi}_L\psi_R + \bar{\psi}_R\chi_L)$ is also not possible, due to the singlet (doublet) nature of right-handed (left-handed) fermion states. Another mechanism needs to be invented to give masses to the fermions as well as to the massive gauge bosons of the weak interaction.

The Higgs Mechanism of Electroweak Symmetry Breaking

In order to introduce particle masses into the Lagrangian of the Standard Model, a completely different mechanism from the one used so far to introduce the different gauge fields has to be applied. For this a fundamental scalar field, the so-called Higgs field, is postulated. First consider the simplest Lagrangian of a massive, scalar field in a potential of the form:

$$\mathcal{L}_{Higgs} = \frac{1}{2}(\partial_\mu\phi)^2 - (\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4) \tag{2.19}$$

where the first is the kinetic term $(\partial_\mu\phi)^2$, the second term describes the interaction with a potential and the last, quartic term describes the self interaction of the scalar field. If one constrains the parameters to $\lambda > 0$ and $\mu^2 < 0$ one obtains a potential that obeys reflection symmetry. The miminum of the potential (also called the vacuum expectation value of ϕ) lies at

$$\phi_0 = \pm\sqrt{-\frac{\mu^2}{\lambda}}. \tag{2.20}$$

In this case one can expand the Lagrangian around one of the minima by defining a new field $\phi = \phi_0 + \sigma$, where σ is the perturbation around the minimum. In terms of this quantity the expanded Lagrangian becomes:

$$\mathcal{L}_{Higgs} = \frac{1}{2}(\partial_\mu\sigma)^2 - (-\mu^2)\sigma^2 - \sqrt{-\mu^2\lambda}\sigma^3 - \frac{1}{4}\lambda\sigma^4 + const.. \tag{2.21}$$

The Lagrangian for σ now contains a mass term $m^2 = -2\mu^2$ and both cubic and quartic self-interaction terms. Due to the cubic terms the reflection symmetry is no longer apparent and one says that the symmetry was spontaneously broken.

To apply this mechanism of spontaneous symmetry breaking to the Standard Model, one must construct a field Φ, which is a complex $SU(2)$ doublet of scalar fields ϕ with weak isospin $I = 1$ and one charged and one neutral component:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.22}$$

To preserve $U(1)$ electromagnetic symmetry, one chooses that only the neutral component acquires a vacuum expectation value:

$$\Phi_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \text{ with } v = \sqrt{-\frac{\mu^2}{\lambda}}. \tag{2.23}$$

The field $\Phi(x)$ can then be parametrised using four real fields $\theta_a(x)$ ($a = 1, 2, 3$) and $H(x)$:

$$\Phi(x) = \frac{e^{i\tau_a\theta_a(x)/v}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{2.24}$$

The first three real fields θ_a (Goldstone bosons) do not appear as physical particles but are absorbed by the longitudinal components of the massive vector bosons. This can be achieved by an $SU(2)$ gauge transformation, what is called going to unitary gauge:

$$\Phi'(x) \to e^{-i\tau_a\theta_a(x)/v}\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.25}$$

The fourth field $H(x)$ represents a massive scalar particle, the so-called Higgs boson. When applying the covariant derivative of the electroweak theory to this parametrisation of the Higgs field one obtains a contribution to the Lagrangian of the form:

$$\mathcal{L}_{Higgs} = (\frac{1}{2}vg)^2 W_\mu^+ W^{-\mu} + \frac{1}{2\cos\theta_W^2}(\frac{1}{2}vg)^2 Z_\mu Z^\mu \tag{2.26}$$

which contains mass terms for the physical fields of the electroweak interaction of the form $M_W^2 W_\mu^+ W^{-\mu}$ and $\frac{1}{2}M_Z^2 Z_\mu Z^\mu$. No term of the form $\frac{1}{2}M_\gamma^2 A_\mu A^\mu$ appears in the Lagrangian, so the photon remains massless. The masses of the electroweak gauge bosons equate to the other parameters of the theory as follows:

$$M_W = \frac{1}{2}vg \tag{2.27}$$

$$M_Z = \frac{M_W}{\cos\theta_W} \tag{2.28}$$

In this way the Higgs mechanism spontaneously breaks $SU(2) \times U(1)$ symmetry down to $U(1)$, leaving the symmetry of QED intact and the photon massless. In the process three degrees of freedom are created (the Goldstone bosons) which are absorbed in the theory to give masses to the W^+, the W^- and the Z^0 bosons and one scalar particle is predicted. The masses of the fermions can also be generated via the Higgs mechanism by including Yukawa couplings between each fermion and the scalar Higgs

field. The strengths of these couplings, which equate to the particle masses, are free parameters of the theory and need to be chosen to reflect the particle masses observed in experiment.

2.2 Supersymmetry

This section introduces the basic concepts of Supersymmetry, which is a symmetry that relates fermions to bosons. This means that SUSY introduces an operator that replaces all bosonic particles with some, as of yet not known, fermionic particles and vice-versa. To be a real symmetry, this operation must leave the equations of motion derived from the theory unchanged. The generator Q of this transformation is supposed to be an anticommuting spinor which acts on physical states like this:

$$Q\,|Boson\rangle = |Fermion\rangle \;\;;\;\; Q\,|Fermion\rangle = |Boson\rangle\,. \tag{2.29}$$

As spinors are by nature complex objects, $Q^\dagger$ (the Hermitian conjugate of Q) also needs to be a SUSY generator. From further considerations [20, 21] on the desired particle content of the theory and the possibility for violation of charge-parity (CP) conservation, as it is observed in nature, it follows that the generators of SUSY must satisfy the following algebra of commutation and anticommutation relations:

$$\{Q, Q^\dagger\} = P^\mu \tag{2.30}$$

$$\{Q^\dagger, Q^\dagger\} = \{Q, Q\} = 0 \tag{2.31}$$

$$[Q^\dagger, P^\mu] = [Q, P^\mu] = 0, \tag{2.32}$$

where P^μ is the four-momentum generator of space-time translations. The single particle states in a supersymmetric theory must, to be eigenstates of Q and $Q^\dagger$, consist of both bosonic and fermionic states and it can indeed be shown that the number of fermionic and bosonic states need to be equal within each SUSY state. The states of different spin belonging to one SUSY state are called superpartners and are said to belong to one supermultiplet. To illustrate what is meant by this, the two simplest supermultiplets are described here:

- a chiral supermultiplet, consisting of one “Weyl” fermion, which is a spinor with two components representing a massless particle with spin $\frac{1}{2}$, and a complex scalar field with spin 0.
- a vector supermultiplet, consisting of one massles vector field (spin 1) and one massless Weyl fermion (spin $\frac{1}{2}$).

In the simplest supersymmetric extension of the Standard Model, the quarks, leptons and the Higgs boson would all fall into their own chiral supermultiplet and the gauge bosons would each have their own vector supermultiplets, which are all completed by new particles differing in spin by one half from their Standard Model counterparts. The scalar SUSY particles are assigned the same name as their Standard Model partners with and s prepended to their names, making the supermultiplet partner of the electron the scalar selectron. To derive the names of the SUSY partners of the gauge bosons the syllable -ino is appended to the name of the gauge boson, leading to a list of "gauginos” which contains such particles as the *W*ino and the gluino. The Higgs sector becomes also more complex in SUSY. Two supermultiplets are needed to give mass to left-handed and right-handed fermions separately. Good introductions to the theory of the Higgs boson in the context of Supersymmetry can be found elsewhere [22, 23].

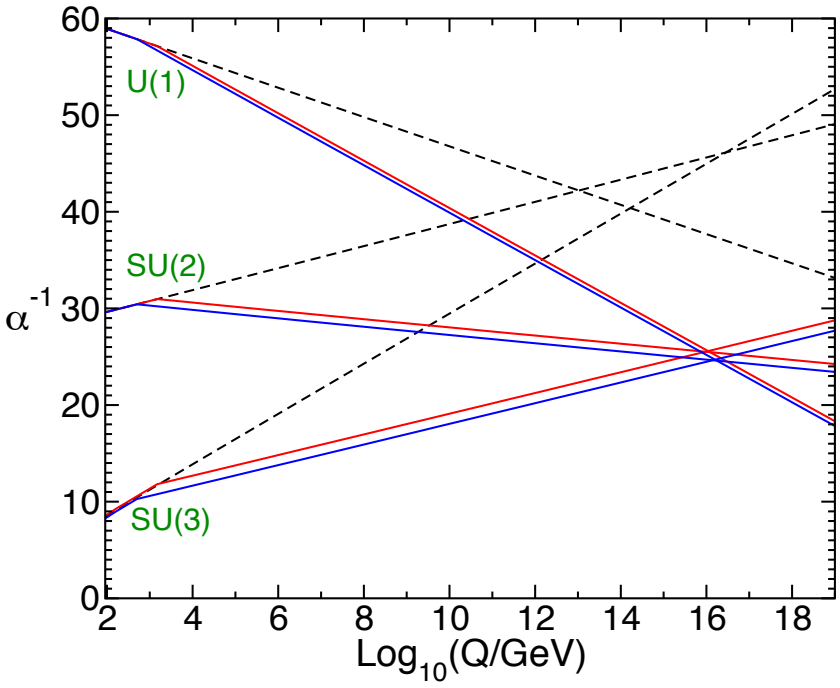


Figure 2.1: Predictions for the running of the gauge couplings from the renormalisation group equations. The evolution of coupling constants is shown by the dashed lines for the Standard Model particle content and by the solid lines for the particle content of the MSSM [24]. The difference between the blue and red solid lines comes from varying the SUSY particle masses between 500 GeV and 1.5 TeV and from varying the coupling strength of the strong interaction within its uncertainties at the Z boson mass.

Considering the possibility of extending the Standard Model in this way, one may ask for motivations for doing so. In fact there are several limitations to the Standard Model to which SUSY provides a possible solution. The most well-known of these are:

2.2.1 Grand Unification

One of the main goals in the development of the Standard Model of particle physics is to unify the mathematical description of all forces in nature. A large step has already been made in unifying electromagnetism and the weak force, by realising that at high energies their coupling constants converge and that they can then be described by one common theory. This running of the coupling constants is described by the renormalisation group equations (RGE) [24] and depends on the particle content of the theory. With the particle content of the Standard Model the coupling constants of the electroweak and strong forces do not converge before the Planck scale. Also no quantum theory of gravity has yet been successfully formulated, although there are attempts to achieve this in the context of string theory. While a theory that combines all four forces in one mathematical framework may still be a long way off, there are several extensions of the Standard Model available which at least allow for unification of the gauge couplings of the three forces already contained in the Standard Model. As can be seen in Fig. 2.1, the additional particle content of Supersymmetry allows for gauge coupling unification and in fact most theories aiming for unification of all four forces, so-called grand-unified theories or GUT's, are supersymmetric.

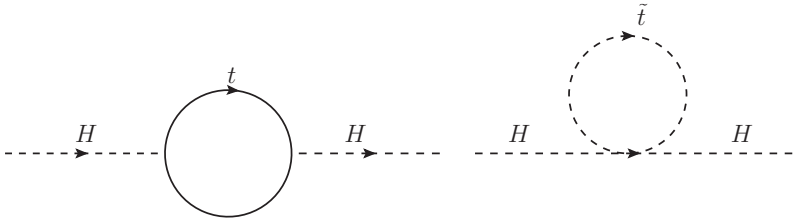


Figure 2.2: Feynman diagrams of loop corrections to the bare Higgs boson mass. Both the fermionic loop correction for the top quark (left) as well as the bosonic loop correction for the SUSY partner of the top (right) are shown.

2.2.2 Dark Matter

While the particle content of the Standard Model is quite varied, the bulk of observable matter is made up of baryons (protons and neutrons) and electrons, with most of the mass accounted for by the baryons. Predictions about the total mass of baryonic matter in the universe from measurements of the cosmic background radiation are, however, in crass contrast to measurements of gravitational effects, such as the rotation curves of galaxies or gravitational lensing. Therefore it is accepted fact that only about 4% of all mass in the universe is actually accounted for by baryons. The rest is attributed to so-called dark matter and dark energy. Dark in this case means that it does not radiate electromagnetic waves and therefore only reveals itself through its gravitational interaction. The Standard Model gives no explanation for the amount of dark matter present in the universe, but certain supersymmetric models introduce a new multiplicative quantum number named R-parity which is defined as:

$$R_P = (-1)^{3(B-L)+2S}, \tag{2.33}$$

where B is the baryon number, L is the lepton number and S is the spin of a given state in the theory. R-parity conservation, which was originally introduced as a way to prohibit unrealistically fast decay rates of the proton, does also imply the existence of new, stable, massive particles which only interact weakly with the particles of the Standard Model. The reason for this is that R-parity is always negative for SUSY particles and always positive for Standard Model particles and therefore additional particles introduced by SUSY can only be produced in pairs. To conserve R-parity, the decay of each SUSY particle must again produce at least one SUSY particle and possibly one or more Standard Model particles. The lightest SUSY particle (LSP) would therefore need to be stable and would be an ideal candidate for dark matter, assuming it does not carry electrical or color charge.

2.2.3 The Hierarchy Problem

Another widely discussed, problematic aspect of the Standard Model is the numerical value of the Higgs boson mass. The Higgs boson, being a scalar particle, aquires radiative corrections to its mass, coming from so-called loop diagrams (see Fig. 2.2) in which the Higgs boson can in principle emit and reabsorb any massive particle contained in the theory.

Since the coupling of particles to the Higgs field is proportional to their mass squared, the most important contribution comes from the top quark. The full mass squared of the Higgs boson m_h^2, with

corrections for radiating a fermion can be written as:

$$m_h^2 = m_{h,bare}^2 + \Delta m_{h,F}^2, \tag{2.34}$$

where $m_{h,bare}$ is the "bare" Higgs mass and $\Delta m_{h,F}$ is the shift due to loop corrections where a fermion is produced and directly recaptured. This turns out to be of the form:

$$\Delta m_{h,F}^2 = \frac{\lambda_F^2}{8\pi^2}[-\Lambda^2 + 3m_F^2 \ln \Lambda/m_F + ...], \tag{2.35}$$

where λ_F is the Yukawa coupling of the fermion to the Higgs field. Since the radiative corrections are dependent on the square of the energy of the virtual particle, a cut-off to this energy (Λ) is introduced. Above this cut-off one expects new physical phenomena to enter which regularise the divergence of the correction terms. The highest possible value for Λ is assumed to be the Planck scale, at which gravitational effects must be taken into account, which are not yet contained in the theory. Bosonic loops also produce a correction $\Delta m_{h,B}^2$ to the Higgs mass but with a different relative sign:

$$\Delta m_{h,B}^2 = \frac{\lambda_B^2}{16\pi^2}[+\Lambda^2 - 2m_B^2 \ln \Lambda/m_B + ...]. \tag{2.36}$$

Given that the heaviest particle in the Standard Model is the top quark and that the correction is proportional to the mass squared of the virtual particle, one can focus on the behaviour of the top loop correction to understand the main features of the problem. If no new physics enters into the theory below the Planck scale, the radiative corrections would contain terms of size $m_{Planck}^2 \approx 10^{38}\ \mathrm{GeV}^2/c^4$, which would need to be cancelled by a bare Higgs mass of the same size, but opposite sign, tuned down to the digits for the observed Higgs mass. If the Higgs boson had a mass of 125 GeV this would mean that the corrections and the bare Higgs mass would need to match in the first 34 decimal places. This is called the "fine-tuning" problem. While such tuning is of course not impossible it offends the desire for "naturalness". In perfect Supersymmetry, where each fermion of the Standard Model has a partner boson with exactly the same mass and vice-versa, this problem naturally disappears as the radiative correction of each Standard Model particle is cancelled by its SUSY partner(s). Some theorists [24] do put this forth as the strongest motivation and also point in favor of the concept of Supersymmetry.

2.2.4 SUSY Breaking

While the reasons stated above make SUSY a very attractive theoretical concept, it is immediately apparent that SUSY cannot be an exact symmetry. If it was, the superpartners of the Standard Model particles would have exactly the same mass and would have been observed already. An easy way to explain why we have not yet observed for example a scalar electron is that the masses of the SUSY particles are higher than a few hundred GeV. In order for SUSY to still provide a solution to the hierarchy problem the masses of the SUSY particles may, however, not exceed a few TeV. For this reason it is also necessary that SUSY-breaking terms in the Lagrangian diverge at most logarithmically with the cut-off scale Λ. If a supersymmetric theory contains only such terms in the Lagrangian one speaks of "soft SUSY breaking". Since no supersymmetric particles have yet been discovered the mechanism of SUSY breaking is also not known. However, several different possible mechanisms have been theorised about. The most common of these theories contains a hidden sector in which SUSY is broken, which couples to the rest of the theory via renormalisable interactions like gravity (for example minimal SuperGravity, mSUGRA) or the gauge interactions of the Standard Model (Gauge Mediated SUSY Breaking - GMSB). The signal scenarios investigated in this thesis all belong to the family of GMSB, which is why only this

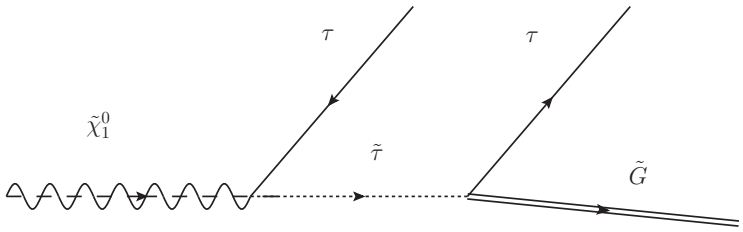


Figure 2.3: Feynman diagram of the decay $\tilde{\chi}_1^0 \to \tau\tilde{\tau} \to \tau\tau\tilde{G}$.

mechanism of SUSY breaking is described in more detail in the following:

In GMSB models [25–31] the breaking of Supersymmetry is communicated through a flavor-blind SM gauge interaction via messenger fields at a scale M_{mess}, which is small compared to the Planck mass. In the minimal GMSB model, the messenger fields form complete representations of $SU(5)$ and therefore preserve the unification of the coupling constants. Squarks, sleptons, and gauginos obtain their masses radiatively from the gauge interactions with the massive messenger fields in such a way that the superpartner masses are proportional to the breaking scale. The LSP is an almost massless particle called the gravitino $\tilde{G}$, while the next-to-lightest SUSY particle (NLSP) dictates the phenomenology of the specific SUSY model. In the simplest GMSB models there are only five free parameters, whose values dictate both the mass spectra of the SUSY particles as well as their interactions with the Standard Model particles:

- Λ: the scale of the SUSY breaking; typically it has values of 10-100 TeV and sets the overall mass scale for all MSSM superpartners, which depend linearly on Λ.
- M_{mess}: the messenger mass scale; it has to be larger than Λ in order to prevent color and charge breaking in the messenger sector.
- N_5: the number of equivalent messenger fields; the gaugino masses depend linearly on N_5 while the sfermion masses are proportional to $\sqrt{N_5}$.
- $\tan\beta$: the ratio of the vacuum expectation values of the two Higgs doublets at the electroweak scale.
- $\text{sign}(\mu) = \pm$: the sign of the Higgsino mass term appearing in the neutralino and chargino mass matrices or in the superpotential.
- $C_{\text{grav}} \geq 1$: the ratio of the gravitino mass to its value for a breaking scale Λ; it determines the lifetime of the NLSP.

Over large ranges of these parameters the NLSP is the stau ($\tilde{\tau}$), which means that one can expect to see two τ leptons being produced at the end of each decay chain, see Fig. 2.3.

In the context of the SUSY studies performed in this thesis, the GMSB parameters except for Λ and $\tan\beta$ are constrained to values for which this is the case, namely: $M_{\text{mess}} = 250\,\text{TeV}$, $N_5 = 3$, $\text{sign}(\mu) = +$ and $C_{\text{grav}} = 1$.

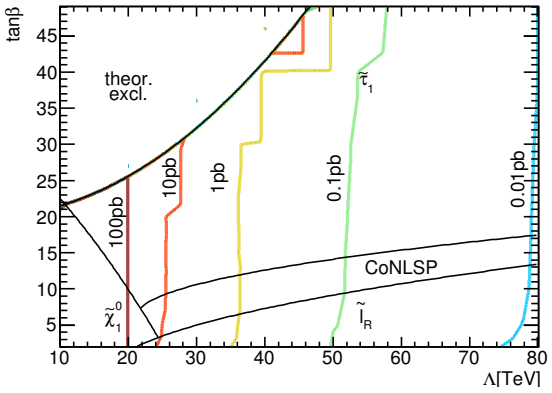


Figure 2.4: The GMSB LO cross section in the Λ-tanβ-plane for $M_{\text{mess}} = 250$ TeV, $N_5 = 3, C_{\text{grav}} = 1$. The black lines indicate the boundaries between regions with different NLSPs.

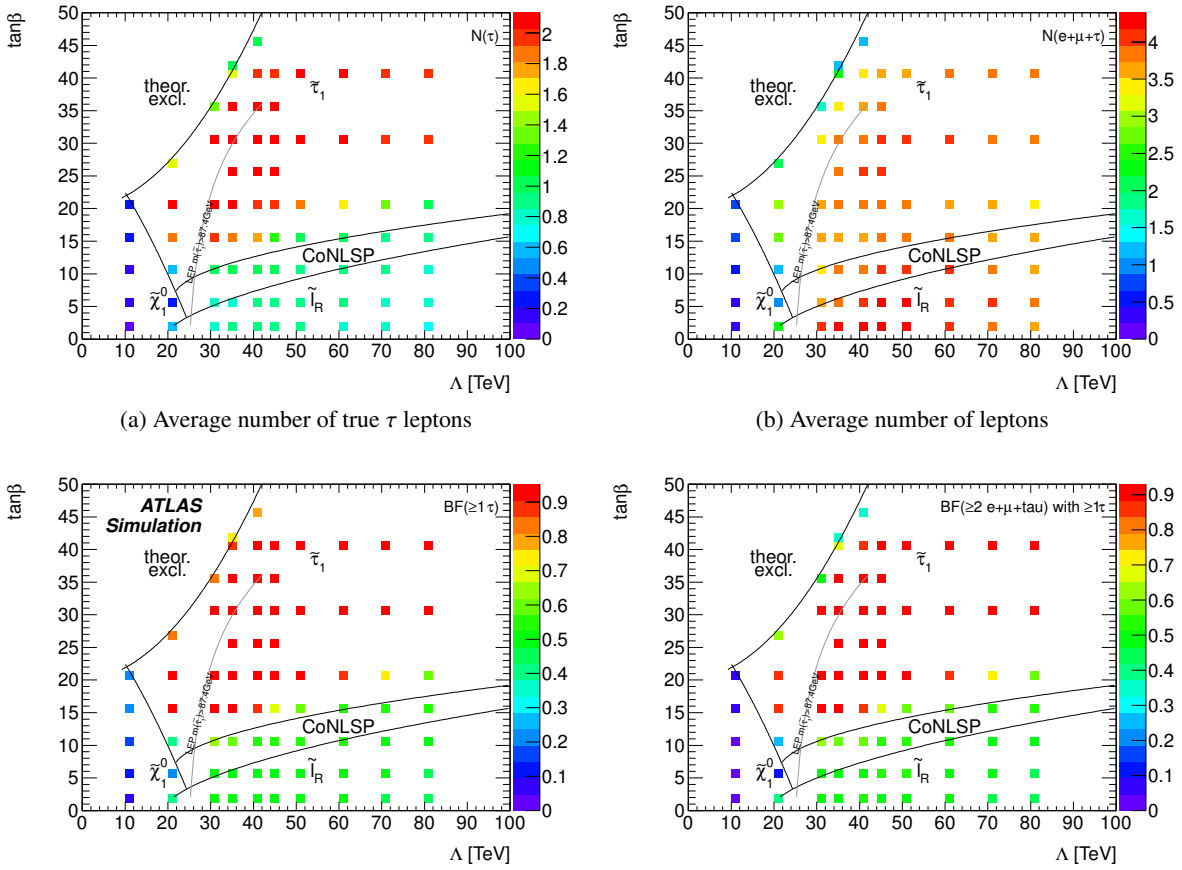
The chosen parameter values restrict the analysis to specific final states and promptly decaying NLSPs. Allowing C_{grav} to be greater than one would prolong the lifetime of the NLSP. For $N_5 = 1$ and small values of tanβ the NLSP is the lightest neutralino $\widetilde{\chi}_1^0$ which decays into a photon and a gravitino. For higher values of tanβ the $\widetilde{\tau}_1$ is the NLSP. For $N_5 \geq 2$ the NLSP is a slepton in a wide range of the parameter space as indicated in Fig. 2.4 for the example of $N_5 = 3$ by the black lines. The NLSP is the $\widetilde{\tau}_1$ ($\widetilde{\ell}_R$) for large (small) values of tanβ while for medium tanβ values the $\widetilde{\tau}_1$ and the right-handed sleptons ($\widetilde{e}_R, \widetilde{\mu}_R$) are almost degenerate in mass (*CoNLSPs*)[1]. The region of small Λ and large tanβ is theoretically excluded since it leads to tachyonic states. GMSB models with $\widetilde{\tau}_1$ NLSP have been searched for at LEP using the pair production of $\widetilde{\tau}_1$ and the subsequent decay $\widetilde{\tau}_1 \rightarrow \tau\widetilde{G}$. For prompt decays, $\widetilde{\tau}_1$ NLSPs with masses below 87 GeV have been excluded [32]. Other searches have been carried out more recently at the ATLAS experiment in final states with either multiple light leptons or multiple hadronic τ lepton decays[12, 33].

To study the dependence of the SUSY models on the two free parameters Λ and tanβ, a set of signal scenarios were generated that differ by varying step-sizes in the Λ – tanβ plane. Fig. 2.4 shows the leading order GMSB production cross section in the Λ – tanβ plane. While only a small dependence of the cross section on tanβ can be observed, it strongly depends on Λ due to the increase of the masses of the SUSY particles with increasing Λ. For example, in the range $\Lambda = 10$ TeV to $\Lambda = 50$ TeV the cross section drops by four orders of magnitude. The lepton production rates across this grid of GMSB signal scenarios are shown in Fig. 2.5. The effects of the changing NLSP are clearly visible on the τ production. Compared to mSUGRA models with high values of tanβ, which also lead to a $\tilde{\tau}_1$NLSP, GMSB models often yield a higher number of τ leptons in the final state since only the $\tilde{\tau}_1$ can couple to the LSP in this case. However, in most cases their cross section is significantly lower. Fig. 2.6 shows the cross section of events with at least one produced τ across the GMSB grid, i.e., the product of Fig. 2.4 and 2.5(c).

2.3 Phenomenology of Proton-Proton Collisions

Within the Standard Model (with SUSY extensions or not) one should be able to derive observables such as event rates and kinematics of all known scattering processes from the Lagrangian density. In order

[1] The *CoNLSP*-region is defined as the region where the mass difference of the $\widetilde{\tau}_1$ and the right-handed sleptons ($\widetilde{e}_R, \widetilde{\mu}_R$) is smaller than the mass of the τ lepton



(a) Average number of true τ leptons

(b) Average number of leptons

(c) Branching fraction to $\geq 1\tau$ lepton

(d) Branching fraction to ≥ 2 leptons, of which ≥ 1 is a τ

Figure 2.5: Lepton production in the GMSB Λ-tanβ-grid. (a) the average number of τ leptons produced in an event, (b) average number of all leptons produced, including e and μ, (c) branching fraction for events that have at least one τ in the final state, (d) branching fraction for multilepton final states with at least one τ.

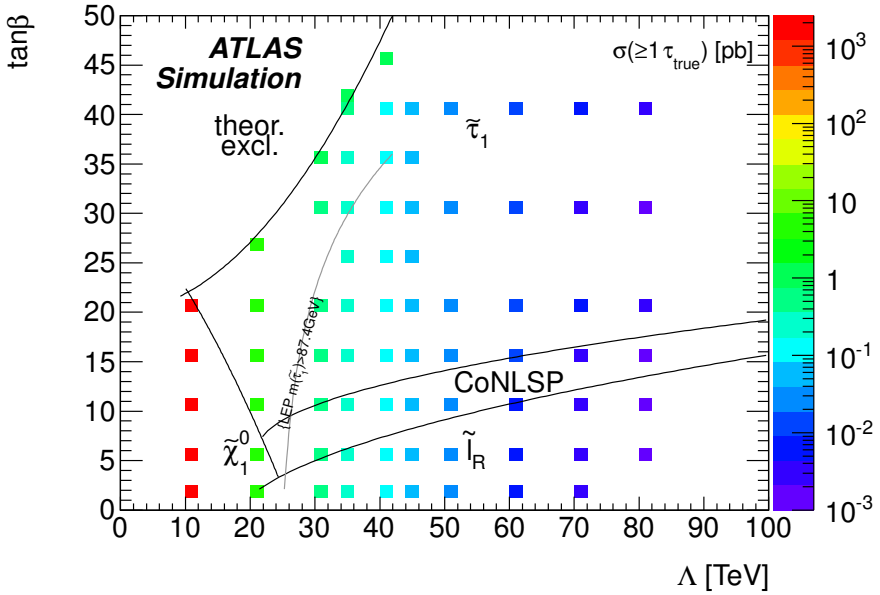


Figure 2.6: Cross section of events with at least one produced τ lepton across the GMSB Λ-tanβ-grid.

to find the total number of events N for a given process that occur over a period of time at a collider experiment, it is easiest to first consider the instantaneous event rate dN at which a certain process will occur. To calculate this, two quantities need to be introduced:

$$dN = Ld\sigma \tag{2.37}$$

These quantities are the so-called instantaneous luminosity L, which describes the rate at which possible scattering partners encounter each other, and the differential cross section $d\sigma$, which gives the probability for a given process to occur each time two scattering partners cross paths. While the luminosity is a quantity of the colliding beams and not of a specific scattering process, the differential cross section depends on the physics process in question and can be calculated from:

$$d\sigma = |\mathcal{M}|^2 dQ \tag{2.38}$$

where $\mathcal{M}$ is the so-called transition matrix element. It corresponds to the square root of the probability of a transition from a given initial state in the theory to the desired final state. The phase-space factor dQ encodes how many kinematically different final states contribute to the desired final state.

Within the Standard Model one can calculate the matrix element from the Lagrangian density, using pertubation theory. At a hadron collider there are additional aspects, which are briefly summarised in the following list and described in more detail thereafter.

- **Parton Distribution Functions:** The Lagrangian of QCD is formulated in terms of elementary quarks and gluons, while at the LHC these partons are contained within the colliding protons. In order to describe the initial state of a proton-proton collision, all possible combinations of initial state partons and their momenta, which contribute to the desired final state, must be summed over using so-called Parton Distribution Functions (PDFs), which need to be determined by experiment.

- **Higher-order Corrections:** As the coupling constant of QCD is large, the higher order terms

in perturbative calculations contribute significantly to the total matrix element. However most strong processes have only been calculated to leading (LO) or next-to-leading (NLO) order in perturbation theory. This means that higher order corrections, that include contributions with additional final state partons, need to be accounted for by phenomenological models. These corrections are especially important at hadron colliders as the initial state partons are strongly interacting.

- **Hadronisation** Due to QCD confinement, the final state quarks and gluons are not observed as free particles but hadronise into colour-neutral hadrons. The energy scale at which this occurs is far below the energy scales at which perturbative QCD calculations are valid and therefore phenomenological models need to be applied.
- **Underlying Event and Pileup:** In addition to the hard interaction of interest, interactions between the other partons in the initial state protons (Underlying Event) and additional proton-proton interactions that happen simultaneously with the hard interaction (Pileup) need to be considered.

To make predictions of observables for any physics process, one uses programs, so-called event generators, which provide a list of final state particles with their four-momenta according to a standardised interface, the so-called Les Houches accord [34]. While each of these programs uses slightly different implementations to address all the requirements listed above, first a generic description of each of these problems is given in Sections 2.3.1, 2.3.2, 2.3.3 and 2.3.4, while the list of important physics processes in the context of this thesis is then given in Sec. 2.4 together with a brief description of the event generators used to generate the simulated samples for these processes.

2.3.1 Parton Distribution Functions

In proton-proton collisions what it measured is the differential cross section $d\sigma(PP \to X)$ for the two colliding protons to produce the final state X. Since the matrix element is only known in terms of the differential partonic cross section $d\sigma(pp \to X)$, one needs to give a prescription how to derive the one from the other:

$$d\sigma(P_A P_B \to X) = \sum_{a,b} \int dx_a dx_b f_{a/P_A}(x_a, \mu_F^2) f_{b/P_B}(x_b, \mu_F^2) d\sigma(pp \to X) \tag{2.39}$$

Here the functions $f_{a/P_A}(x_a, \mu_F^2)$ and $f_{b/P_B}(x_b, \mu_F^2)$ are called parton distribution functions (PDFs) and give the probability density of finding a parton $a(b)$ with momentum fraction $x_a(x_b)$ in proton $A(B)$ at a scale μ_F. All parton combinations that can produce the final state X are summed over. This means that the hadronic cross section can be derived by convolving the partonic cross section with the parton distribution functions. This is known as the QCD-factorisation theorem. Since these PDFs depend on the interactions of the partons within the proton, a regime where perturbative QCD breaks down, they cannot be determined from first principles, but need to be deterimed from experiment.

Many different experiments offer measurements of the PDFs, among them fixed-target experiments like neutrino scattering, electron-proton scattering experiments such as HERA and hadron collider experiments such as the Tevatron and the LHC. The need for determining the PDFs with different experiments becomes apparent when one considers that they do not only depend on the fraction of the proton momentum $x_{a,b}$, carried by the parton, but also on the momentum transfer of the measured collision. It is possible to derive the PDFs for momentum transfers relevant for the LHC from previous experiments, by parametrising PDFs measured at lower momentum transfer and then evolving the parametrisation using the DGLAP [35–37] evolution equations. Many different functional forms can be

used to parametrise the proton PDFs, and indeed there are many different collaborations each providing their own parametrisation. Some of the most well-known of these collaborations are the ABKM [38], CTEQ [39, 40], HERAPDF [41, 42], MSTW [43] and NNPDF [44, 45]. Since both the measurement as well as the evolution of the measured PDFs to LHC scales yield uncertainties on the form of the PDFs, all collaborations provide so-called error eigenvector PDF sets. These PDF sets are obtained by the Hessian method, in which an orthonormal basis of eigentvectors of the fit parameters, together with their uncertainties is determined. Uncertainties due to the PDF fit can then be determined by comparing the measured observables in samples simulated once with the default PDF set and once with the error eigenvector set.

2.3.2 Higher-Order QCD Corrections/Parton Showers

Since the transition matrix element is not completely known analytically, it needs to be calculated using perturbation theory in orders of the strong coupling constant α_s. This calculation can only be performed up to a fixed order, which turns out to be leading or next-to-leading order for most processes (depending on the event generator). This means that additional partons, radiated off the initial state or final state particles, are not accounted for. To provide a description of these addition partons so-called parton shower algorithms are employed that, in the case of QCD, introduce additional splittings of the final and initial state partons of the form $q \to qg$, $g \to gg$ or $g \to q\bar{q}$. The function of a parton-shower algorithm is explained here, using the parton shower algorithm implemented in the Pythia [46] event generator as an example: The probability for not encountering a split, when evolving a parton energy from a higher scale Q^2 to a lower scale q^2, is given by the so-called Sudakov form factor $\Delta(Q^2, q^2)$, which depends on the difference between the two scales, and the process of parton showering is described probabilistically by generating random numbers between zero and one and comparing them to these probabilities. Going down from the higher scale in small intervals, the algorithm finds the first Sudakov form factor which is larger than the generated random number. The splitting corresponding to this form factor occurs and this process is repeated all the way down to a minimum scale $Q_0^2 \approx 1$ GeV at which the showering algorithm is stopped. Another parton shower algorithm employed in simulating the samples used in this thesis is called `Herwig` [47], which uses a similar showering scheme, but instead of ordering the splittings in the momentum transfer, `Herwig` orders the splittings by the emission angle.

2.3.3 Hadronisation

Quarks and gluons cannot be observed as free particles. Since the list of final state particles provided by the event generator together with parton showers still contains quarks and gluons, these partons need to be combined into hadrons. While this process of hadronisation cannot be described by perturbative QCD, several phenomenological models exist which can be used to describe it. The Lund model [48] uses colour strings, which connect neighbouring partons and can be broken up by producing a new $q\bar{q}$ pair. Using the colour information from these strings the outgoing partons are then combined into colour neutral mesons and baryons. `Herwig` again uses a different approach, called the cluster model, where each gluon is split into a $q\bar{q}$ pair, after which the list of quarks is clustered into colour singlet states, which are subsequently decayed into hadrons [49]. Since both of these approaches cannot derive their predictions from an underlying theory, their parameters need to be tuned to measurements.

2.3.4 Underlying Event and Pileup

While the hard scattering events of interest happen only between one parton in each of the colliding protons, there are additional interactions within and between the two proton remnants. This is known as the "underlying event". One way to account for this is the model of multiple partonic interactions in which additional interactions between the remaining partons are simulated. This model is implemented in both the PYTHIA and `JIMMY` [50] programs, and probabilistically generates additional interactions, that normally occur at scales much smaller than the hard process. The `Herwig` generator is usually interfaced with `JIMMY` to simulate the underlying event, although it does have its own model to describe it.

In addition to the interactions of the proton remnants, there can be several additional proton-proton collisions within the same bunch crossing. This phenomenon is called in-time pileup. At design parameters the bunch spacing in the LHC will be 25 ns, which is shorter than the readout time for the full ATLAS detector. Therefore some of the signals recorded in one bunch crossing will be due to collisions occuring in the previous bunch crossing. This is know as out-of-time pileup. Just like the underlying event, most of the additional collisions from pileup will be soft compared to the hard process. These events are often referred to as "minimum-bias" events, because one needs to be careful to avoid a process-specific selection to gather a sample of these events in data. To account for in-time and out-of-time pileup in simulation, such minimum-bias events are simulated and a Poisson-distributed number of them is overlayed on top of the hard process in question. Because the full range of processes and of initial state and final state momenta are available in simulating these events, it is most probable that the bulk of them will be made up of QCD $2 \rightarrow 2$ processes at low momentum transfer.

2.4 Simulated Samples

This final section of this chapter gives an overview of the physics processes, both Standard Model and SUSY, studied as part of this thesis. While the event generators used to simulate these processes only provide a list of final state particles together with their four-momenta, these lists were passed through a simulation of the ATLAS detector based on GEANT4 [51], using the Athena software framework [52, 53], to generate a data format, which can then be further processed with the same reconstruction and analysis algorithms used for real collision data. In order to match the pile-up conditions obversed in data, the simulated samples of both background and signal processes are overlayed with a Poissonian-distributed number of pile-up events [54, 55] and the resulting events are reweighted such that the distribution of the number of interactions per bunch crossing agrees with the data.

2.4.1 Monte Carlo Background Samples

The important Standard Model processes in the context of the studies presented in this thesis are described here in the order in which they appear.

Multi-jet event production is simulated using the PYTHIA [46] event generator with the AMBT1 tune [56]. In the AMBT1 tune, measurements of the charged-particle multiplicity in early ATLAS data are used to improve the description of low-energy collisions and the underlying event by the generator. Since the cross section for multi-jet production at the LHC is only known from theory with large uncertainties the normalisation for this process was treated differently in the different studies presented in this thesis. In the performance studies, where the focus lies on comparing the shapes between data and simulation, the normalisation of the simulated sample was fixed to the data integral. In the SUSY search described later, a cross section of ≈ 11 mb was assumed for the optimisation of the event selection. For the final background estimate in this search, the normalisation was estimated directly from data.

Production of vector bosons in association with jets (W + jets and Z + jets) is simulated with the ALPGEN [57] generator, using the CTEQ6L1 [58] PDF set and the two processes are normalised to a cross section of 31.4 nb and 9.02 nb [59–61], respectively. Up to five additional hard partons with more than 20 GeV of transverse momentum are simulated as part of the matrix element, while softer additional final state partons are generated by the parton shower in HERWIG. A dedicated matching algorithm prevents double-counting and HERWIG is interfaced to JIMMY [50] to simulate the underlying event.

Top quark pair, single top quark and diboson production are generated with MC@NLO [62] and the CTEQ6.6 [39] PDF set and are normalised using cross sections of 0.165 nb, 0.085 nb [63–65] and 0.071 nb [66, 67], respectively. MC@NLO provides the matrix element at next-to-leading-order in QCD perturbation theory. The parton showering is also handled by HERWIG for these processes and the underlying event is also simulated using JIMMY. Due to the matching that needs to be performed between the NLO calculation and the parton shower to avoid double-counting, a significant fraction of events are simulated with negative event weights.

The programs TAUOLA [68, 69] and PHOTOS [70] are used to model the decays of τ leptons and the radiation of photons, respectively.

2.4.2 Monte Carlo Signal Samples

Samples of simulated GMSB events are generated with the Herwig++ [71] generator for values of Λ and $\tan\beta$ in the ranges $10 < \Lambda < 85$ TeV and $2 < \tan\beta < 45$, with the SUSY mass spectra generated using ISAJET 7.80 [72]. The MRST2007 LO* [73] PDF set is used and the production cross sections are calculated with PROSPINO [74–77] to next-to-leading order in the QCD coupling constant using the next-to-leading-order CTEQ6.6 [39] PDF set. Two benchmark points with $\Lambda = 30$ (40) TeV and $\tan\beta = 20$ (30) are chosen as representative, signal scenarios. They are used to optimise the event selection and in plotting as a stand-in for the SUSY signal. These benchmark points have moderate cross sections of 1.95 (0.41)pb, while not yet being excluded by former experiments. They also respect current constraints from SM measurements as $b \to s\gamma$ and $(g_\mu - 2)/2$.

Chapter 3

The ATLAS Experiment

This chapter gives an overview of the experiment at which the work for this thesis was performed. A brief introduction to the LHC machine, which provides the proton beams used to produce the collision data analysed in this thesis, is given in Sec. 3.1. These data were recorded with the ATLAS detector, one of the two large, multi-purpose detectors installed at the LHC. There is another large, multi-purpose detector at the LHC (CMS) and two smaller detectors, one dedicated to b-quark physics (LHCb) and one dedicated to heavy-ion collision measurements (ALICE). The ATLAS detector and its subsystems are described in Sec. 3.2, to give an understanding of the large challenges posed to the detector design by the high luminosity and beam energy provided by the LHC machine and the need to measure not only supersymmetric final states but also all important decay modes of the Higgs boson and most known Standard Model processes with a precision not previously achieved.

3.1 The LHC

The Large Hadron Collider (LHC) at CERN near Geneva, Switzerland is the most powerful particle accelerator built to date. According to its design specifications it is built to accelerate protons up to an energy of 7 TeV, resulting in a centre-of-mass energy of $\sqrt{s} = 14$ TeV in up to 2808 bunches with a spacing of about 7.5 m, which corresponds to a time of 25 ns between two bunch-crossing or a frequency of 40 MHz. The number of protons per bunch was foreseen to be 1.1×10^{11} with a transverse bunch radius of 16.6 µm (RMS) and a longitudinal spread of 7.55 cm (RMS). At the interaction points the beams are crossed at an angle of about 300 µrad to avoid unwanted collisions outside the nominal interaction points. The resulting nominal instantaneous luminosity is $10^{34}\text{cm}^{-2}\text{s}^{-1}$. At this high luminosity it is expected that not only one proton-proton collision per bunch crossing takes place. Rather, the average number of collisions within each bunch crossing μ is expected to be on the order of $\mu \approx 25$.

While most of these design goals have not yet been achieved at the time of writing this thesis, the LHC has broken both the world record for highest centre-of-mass energy as well as the one for highest instantaneous luminosity since its inauguration at a centre-of-mass energy of $\sqrt{s} = 900$ GeV in September 2008. Even though the machine suffered a major incident shortly after inauguration, in which a faulty connection between the superconducting magnets used to bend the beam caused a year-long shutdown for repairs, one can say that the LHC has exceeded expectations as to the speed and reliability with which the accelerator was commissioned. Since March of 2010 the LHC is delivering stable beams of $\sqrt{s} = 7$ TeV with an instantaneous luminosity of up to 3.6×10^{33} $\text{cm}^{-2}\text{s}^{-1}$. The time profile of the integrated luminosity delivered by the LHC and recorded by ATLAS in 2010 and 2011 can be seen in Fig. 3.1. The number of colliding bunches was increased from 1 to 348 in 2010 and was up to 1332 in 2011. At the same time the number of protons per bunch was increased from 0.1×10^{11} to 1.2×10^{11} in 2010. With 1.3×10^{11} protons per bunch in 2011 the LHC already exceeds the expected nominal proton-bunch density. From these numbers of protons per bunch and the resulting densities at the interaction spot, one can derive the resulting average number of interactions per bunch crossing which rose

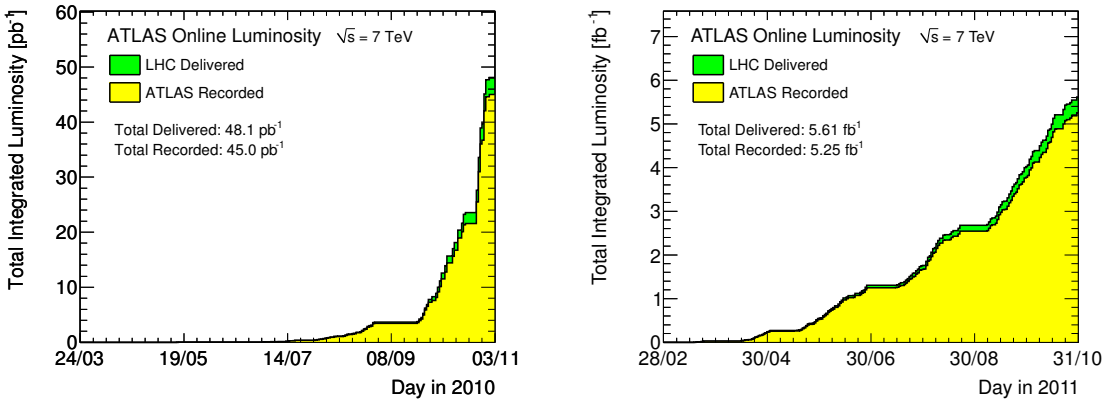


Figure 3.1: Total delivered and recorded integrated luminosity in 2010 (left) and 2011 (right). The cumulative luminosity is shown versus time delivered by the LHC (green), and recorded by ATLAS (yellow) during stable beams and for pp collisions at 7 TeV centre-of-mass energy. The delivered luminosity accounts for the luminosity delivered from the start of stable beams until the LHC requests ATLAS to turn the sensitive detector off to allow a beam dump or beam studies. Shown is the integrated luminosity as determined from counting rates measured by the luminosity detectors. These figures were taken from the public ATLAS luminosity results website [78].

from below one in 2010 to $\mu \approx 3$, while the average value in 2011 rose from $\mu \approx 8$ in March to $\mu \approx 16$ after August.

3.2 The ATLAS Detector

With the experimental conditions provided by the LHC in mind, the ATLAS detector was designed as a multi-purpose detector to search for the Higgs boson, rare SM processes and also for signatures of physics beyond the Standard Model. Considering the cross section of inelastic proton-proton collsions of 80 mb at the design centre-of-mass energy of $\sqrt{s} = 14$ TeV, the detector needs to be able to withstand the fluence produced by 10^9 events per second over a period of several years and measure hadronic particles with high precision. Further considering that many of the rare Standard Model (and beyond the Standard Model) processes involve charged leptons and/or invisible particles escaping the detector, such as neutrinos or long-lived neutral particles, the design goals of the ATLAS detector can be summarised as follows [79]:

- Fast, radiation-hard electronics and sensor elements are required, as well as high detector granularity to handle the particle fluxes and to reduce the influence of additional proton-proton collisions occuring in the same bunch crossing as the collisions of interest.
- The largest possible acceptance in pseudorapidity and full coverage in azimuthal angle is desired.
- A high efficiency in the reconstruction of charged particles with a high resolution of the measurement of their momenta is required. In order to successfully detect τ leptons and b-jets, high-precision vertexing detectors close to the interaction region are needed to reconstruct secondary vertices.
- High-resolution electromagnetic (EM) calorimetry with a high granularity is required to identify electrons and photons. Accurate measurements of jet energies and missing transverse energy

are also required, which in turn requires full-coverage hadronic calorimetry with a good energy resolution for hadrons.

- It is necessary to identify muons with a good momentum resolution over the range of 1 GeV up to 1 TeV and to reconstruct their charge correctly.
- To achieve a manageable trigger rate, a trigger system is needed that selects also low-p_{T} objects with a high efficiency, while at the same time providing sufficient background rejection.

The ATLAS experiment is described in detail in Ref. [79]. The descriptions in this section are adopted from or based on this reference, if not stated otherwise.

The ATLAS coordinate system is a right-handed system with the x-axis pointing to the centre of the LHC ring, the y-axis pointing upwards, and the z-axis following the beam line. The spherical coordinates ϕ and θ are defined in the usual way, with the azimuthal angle, ϕ, measuring the angle in the xy-plane from the positive x-axis, increasing towards positive y. The polar angle θ is measured with respect to the positive z-axis. At hadron colliders the pseudorapidity η, defined as:

$$\eta = -\ln(\tan\frac{\theta}{2}), \tag{3.1}$$

is often used instead of the polar angle. For massless particles, differences in pseudorapidity are invariant under Lorentz boosts along the z-axis, as η is equal to the rapidity y of such particles. As most particles whose momenta are measured with ATLAS like electrons, muons or light hadrons have masses much less than a GeV, η can be said to be a good approximation of y in most cases. Differences in solid angle are usually given as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{3.2}$$

Another important quantity that will be used throughout this thesis is the momentum of a particle transverse to the beam axis p_{T}. This is defined as:

$$p_T = \sqrt{p_x^2 + p_y^2}, \tag{3.3}$$

where $p_{x,y}$ are the x, y-components of the particles four-momentum.

A cutaway view of the entire detector is shown in Fig. 3.2, where all major components of the detector are pointed out separately. It can be seen that the ATLAS detector is built in layers around the interaction point, where several sets of layers can be identified as belonging to each of the sub-detectors. In the following a short overview of the different sub-detectors is given, in the order as seen from the perspective of a particle traversing the detector from the interaction point outwards.

The Inner Detector (ID), which comprises the first set of layers in the detector, is used to detect the passage of charged particles and reconstruct their trajectories. In order to be able to measure also the momenta of the charged particles the ID is immersed in the 2 T magnetic field of a solenoid, which surrounds the central tracker volume. Precision tracking and vertex resolution is offered by the innermost tracking system denoted the semi-conductor tracker (SCT). It consists of three inner layers of semiconductor pixel detectors and several layers of silicon strip detectors, both with a coverage of $|\eta| < 2.5$. The outermost part of the inner detector consists of the Transition Radiation Tracker (TRT) made of straw tubes, which covers up to $|\eta| < 2.0$.

Energies of both charged and neutral particles are measured by two calorimeter systems. The inner one, the electromagnetic calorimeter (ECAL) uses liquid-argon (LAr) as the active material and lead as absorber material and covers the range $|\eta| < 3.2$. Hadronic calorimetry in the range $|\eta| < 1.7$ is provided

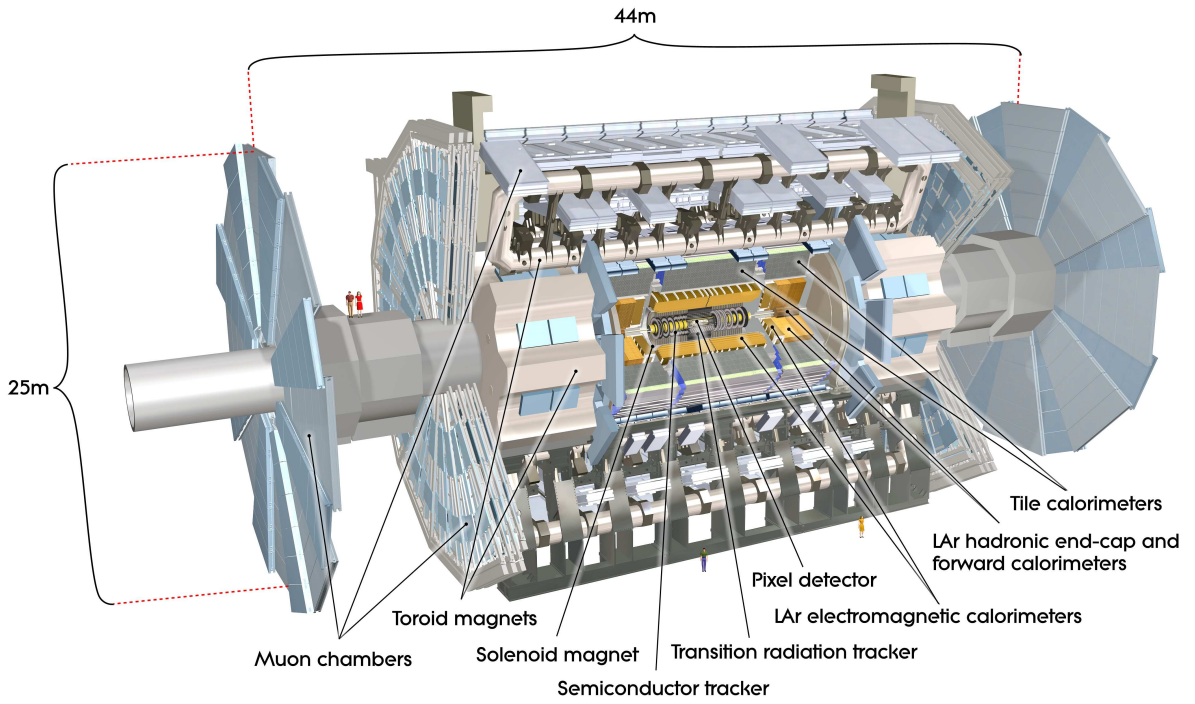


Figure 3.2: Cut-away view of the ATLAS detector. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tons [79].

by a scintillator-tile calorimeter. A LAr-based hadronic calorimeter covers the range $1.5 < |\eta| < 3.2$. The forward calorimeters, which again employ LAr as active material, provide both electromagnetic and hadronic energy measurements, and extend the coverage up to $|\eta| < 4.9$.

The muon spectrometer comprises the outermost layers of the detector, providing muon identification and momentum measurements with a coverage up to $|\eta| < 2.7$. The muon trajectories are bent using an air-core toroid system.

Selection of interesting events over the dominant QCD multi-jet processes, occurs via a three-level trigger system, which consists of a hardware-based trigger (Level 1) and a higher level software-based trigger system (denoted HLT) which consists of two subsystems called Level 2 and Event Filter trigger respectively. Additional detectors are installed in the forward region along the beam pipe to measure the luminosity delivered by the LHC. The following sections describe the different sub-detector systems in more detail.

3.3 Inner Detector

At the very high luminosity foreseen for the nominal beam conditions at the LHC, roughly one thousand particles will emerge from the interaction point during every bunch crossing. This leads to a very large particle flux density close to the interaction point, in which the tracking system needs to distinguish single charged particles and survive the high radiation levels for several years. Silicon pixel and strip detectors in the SCT and, at larger distances from the interaction point, also the straw tubes of the TRT, offer both the precision and radiation hardness necessary for these tasks. The different components of the inner detector are shown in Fig. 3.3 und described in detail hereafter.

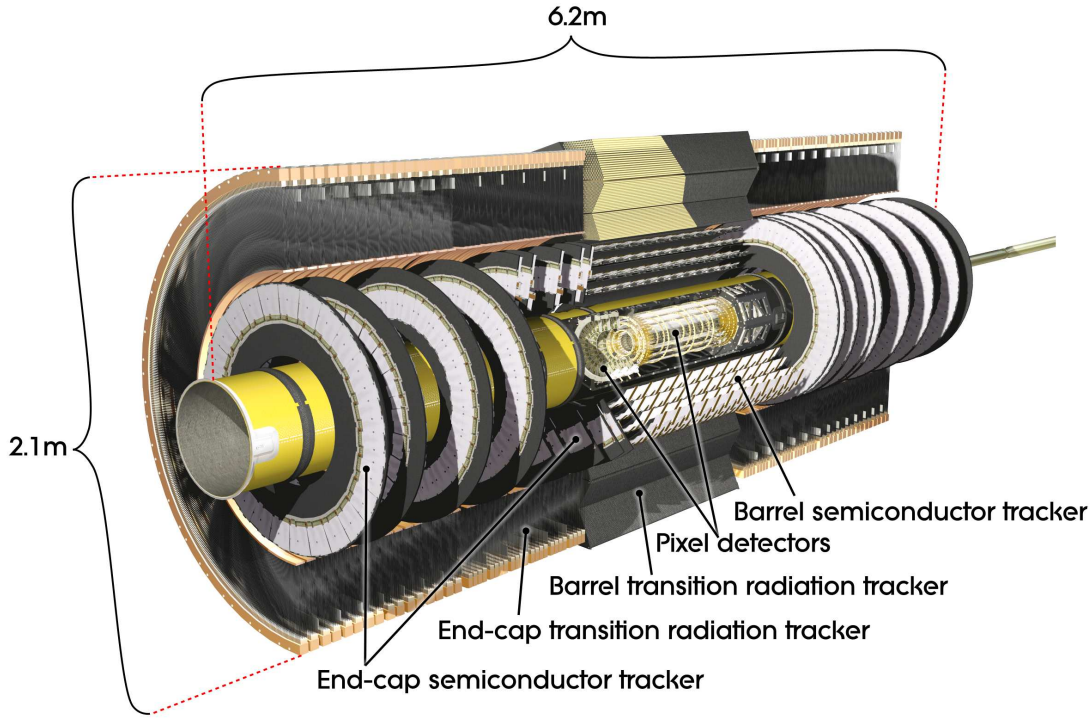


Figure 3.3: Cut-away view of the ATLAS Inner Detector [79].

3.3.1 The Semiconductor Tracker (SCT)

Precision tracking and vertexing is provided by the silicon pixel and strip detectors which are arranged in concentric, cylindrical layers around the interaction point in the barrel region of the detector and in disks perpendicular to the beam axis in the endcaps.

The three innermost layers, which provide the precision vertexing capabilities, are instrumented with silicon pixel detectors, segmented in $R-\phi$ and z with a minimum pixel size of $50 \times 400\,\mu\text{m}^2$ which leads to an intrinsic accuracy in $R-\phi$ of 10 µm and 115 µm in z. To achieve the best possible vertexing resolution the innermost layer, often called the b-layer as vertexing is crucial in the identification of b-quark jets, is positioned as close to the beam pipe as possible at a distance $R = 50.5$ mm.

The eight remaining layers of the SCT consist of silicon strip detectors, where two layers are always joined, such that each strip detector is overlayed by another at an angle of 40 mrad to provide measurements of hit positions in two dimensions, with the measurement of position in the third dimension being offered by the position of the two modules along the beam pipe. This combination of layers yields four space point measurements per track in the strip detectors, with an intrinsic resolution of 17 mm in $R-\phi$ and 580 mm in z in both the barrel and the endcaps.

3.3.2 The Transition Radiation Tracker (TRT)

The TRT is made up of straw tubes filled with xenon gas. The xenon gas is ionised by traversing charged particles. The straws in the barrel region lie along the beam axis and are 144 cm long. They are segmented at $\eta = 0$ with one sensing wire extending into each hemisphere of the detector. In the endcaps the straws extend radially away from the beam pipe up to a length of 37 cm. This means that each straw only measures the position of charged particles in the $R-\phi$ direction, for which it has an

intrinsic resolution of 130 µm per straw. The large number of straws (≈ 351 thousand readout channels) means that an average charged particle is recorded in the TRT with 36 separate hits. Together with the larger distance to the interaction point compared to the SCT, this means that the TRT contributes significantly to the precision of the momentum measurement of charged particles. Additional material interleaved with the straw tubes causes high-p_T electrons to emit transition radiation photons which are absorbed in the straw tubes to produce higher signal amplitudes for electrons than for other particles traversing the TRT. This additional information is used in electron identification.

3.4 Calorimeters

ATLAS employs sampling calorimeter technology, which is separated in two parts: an inner part used to contain and measure electromagnetic showers and an outer part to contain and measure hadronic particle showers. The calorimeter with its separate components is displayed in Fig. 3.4. One of the main design goals of the calorimeter is hermeticity to improve the resolution of missing transverse momentum. Therefore the calorimeter coverage extends up to $|\eta| < 4.9$. Different technologies are used in the pseudorapidity region covered by the inner detector. The very high granularity of the electromagnetic calorimeter (ECAL) allows for high precision electron and photon reconstruction and identification. The rest of the calorimeter still has a good enough resolution, even with a coarser granularity, to measure both the energy and position of hadronic jets with sufficient precision. Apart from high-precision energy measurements and particle identification, the calorimeters must also provide good containment for electromagnetic and hadronic showers and must also limit punch-through into the muon system. Therefore the amount of active and absorber material budgeted in the ECAL amounts to more than 22 radiation lengths X_0 in the barrel and more than $24X_0$ in the end-caps. To contain hadronic showers the combination of ECAL and hadronic calorimeter (HCAL) amounts to at least 9.7 hadronic interaction lengths λ in the barrel (10λ in the endcaps) and an additional thickness of 1.3λ is provided by the outer support structure of the calorimeters. This amount of material in the calorimeter has been shown by measurements as well as simulation to reduce punch-through into the muon spectrometer to below the irreducible levels due to prompt muons or muons from in-flight decays. The two components of the calorimeter system are described in more detail hereafter.

3.4.1 Electromagnetic Calorimeter (ECAL)

The electromagnetic calorimeter is a sampling calorimeter in which liquid argon is used as the active material interlaced with lead absorber plates. The barrel part ($|\eta| < 1.475$) is segmented into three sections in depth and two half barrels, one for each η-hemisphere, while each endcap consists of two coaxial wheels: one outer wheel, segmented into three layers in depth, that overlaps with the inner detector over the range $1.375 < |\eta| < 2.5$, and one inner wheel, segmented into two layers in depth, with coverage over $2.5 < |\eta| < 3.2$. An "accordeon-like" geometry was chosen for the layers of LAr and lead, interlaced with kapton electrodes to collect the signals from the active material, see Fig. 3.5. This geometry ensures absolute symmetry in ϕ without any cracks over the full coverage of the ECAL.

In the region $|\eta| < 1.8$, used for precision measurements of electrons and photons, the ECAL is equipped with a presampler calorimeter, which can be used to correct for the energy lost by electrons and photons upstream of the calorimeter. The presampler consists of an active LAr layer of thickness 1.1 cm (0.5 cm) in the barrel (end-cap) region. The first actual layer ($\approx 4X_0$) of the ECAL (called the η-strip layer) has the highest granularity of up to $\Delta\eta = 0.0031$ in the range of $|\eta| < 1.8$ and somewhat coarser beyond that. This allows for a separation of prompt single photons from the two overlapping

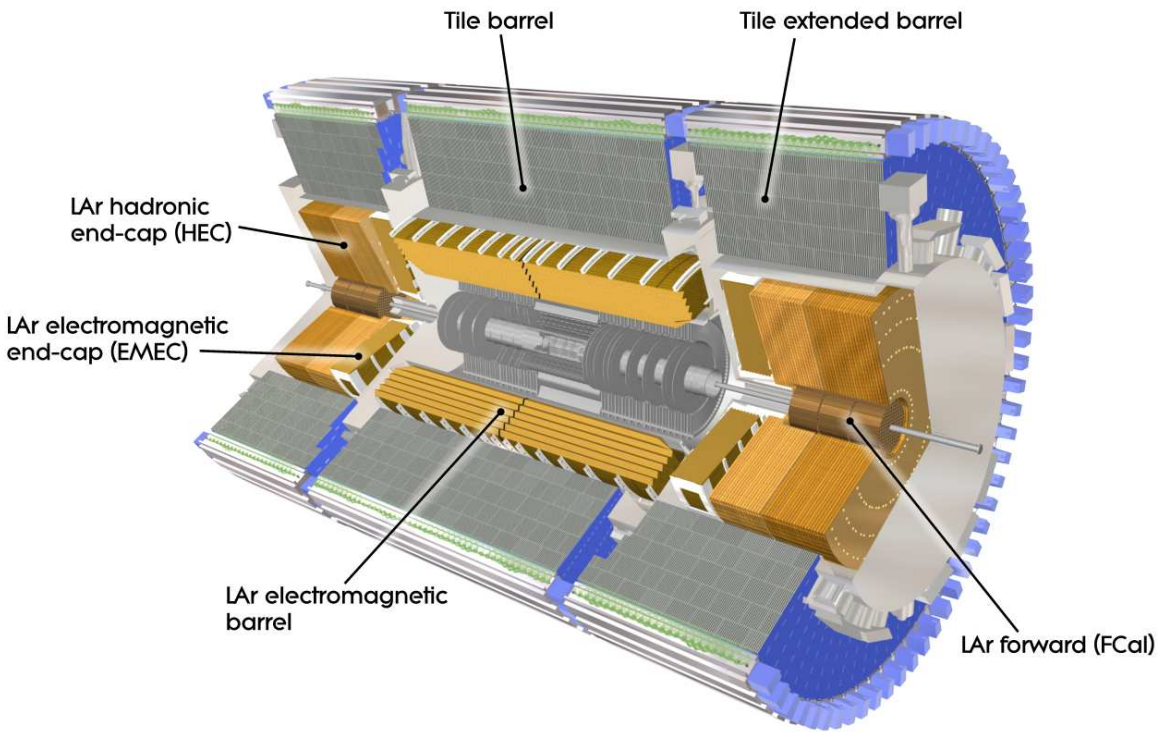


Figure 3.4: Cut-away view of the ATLAS calorimeter system [79].

showers of photons from $\pi^0 \rightarrow \gamma\gamma$ decays or electrons from the energy deposits of hadrons in the ECAL. The granularity in ϕ is somewhat coarser with a cell size of $\Delta\phi$ =0.1 rad. While the η-strip layer is important for position measurements and particle identification, most of the energy of high energy electrons and photons is deposited in the second layer which has a thickness of approximately $16X_0$ and a cell size of 0.025×0.025 in η and ϕ. The third layer (the so-called Tail Catcher), with two interaction lengths in thickness, is used to contain the tails of the showers from very highly energetic electrons and photons and has a coarser granularity of 0.050×0.025 in the region $|\eta| < 2.5$. The design energy resolution of the ECAL is given as:

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E(\text{ GeV})}} \oplus 0.7\%. \tag{3.4}$$

3.4.2 Hadronic Calorimeter (HCAL)

The hadronic calorimeter, placed directly outside the ECAL envelope, is also a sampling calorimeter employing steel as absorber and scintillating tiles as the active material in the barrel region. Two sides of the scintillating tiles are read out by wavelength shifting fibres into two separate photomultiplier tubes. The barrel part, covering the region $|\eta| < 1.7$, is separated into three layers, the first two with a granularity of $\Delta\eta \times \Delta\phi$ of 0.1×0.1 rad and the third one with a granularity of $\Delta\eta \times \Delta\phi$ of 0.1×0.2 rad.

The Hadronic End-Cap (HEC) calorimeter overlaps with the barrel tile calorimeter and covers the region $1.5 < |\eta| < 3.2$. In contrast to the barrel tile calorimeter, the HEC shares a cryostat with the end-cap ECAL and also uses liquid argon as the active material. The absorber is, however, made of

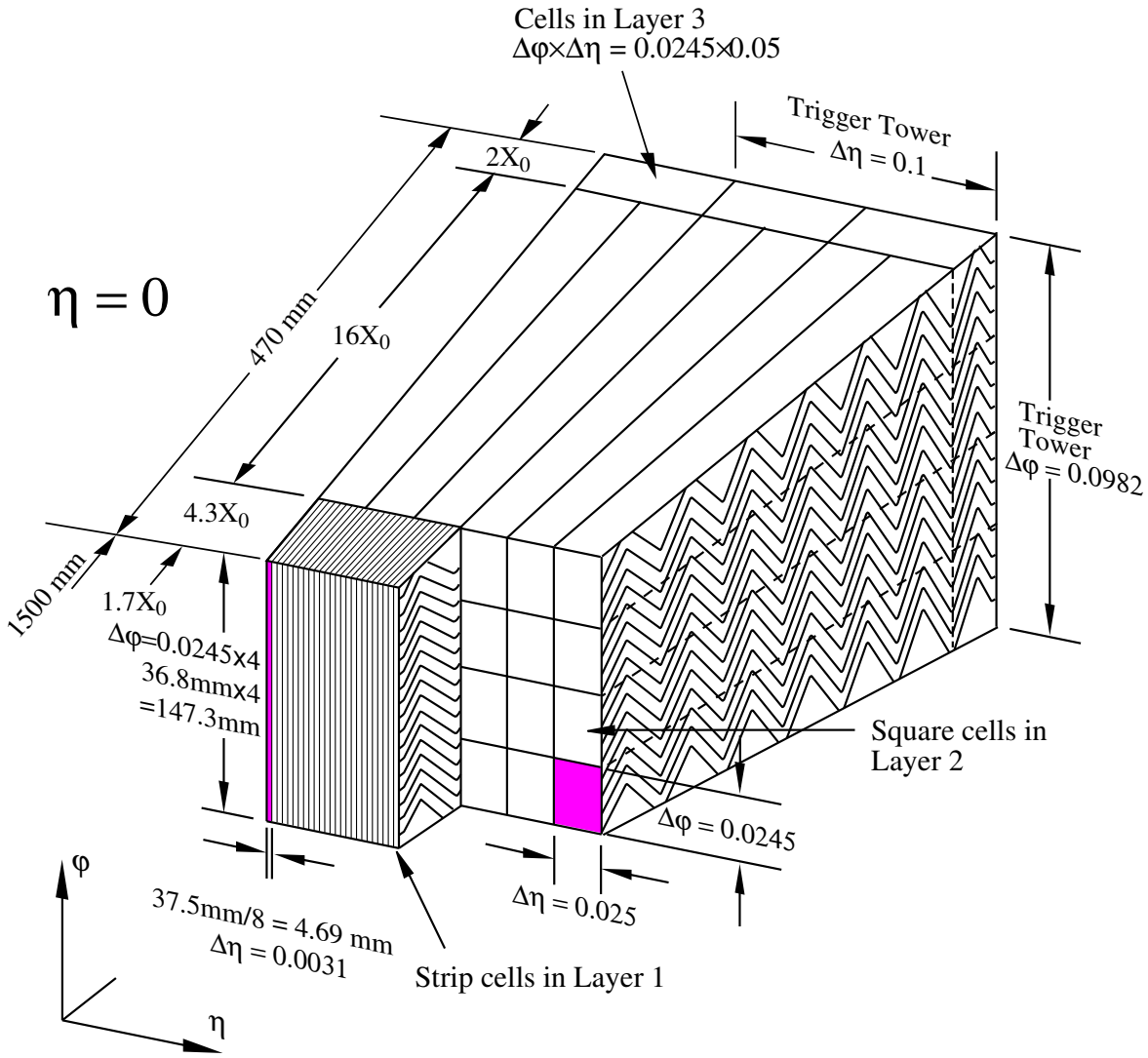


Figure 3.5: Sketch of a barrel module where the different layers are clearly visible with the ganging of electrodes in ϕ. The granularity in η and ϕ of the cells of each of the three layers and of the trigger towers is also shown [79].

copper and the design energy resolution of the hadronic barrel and end-cap for jets is:

$$\frac{\sigma(E)}{E} = \frac{50\%}{\sqrt{E(\text{ GeV})}} \oplus 3\%. \tag{3.5}$$

The Forward Calorimeter (FCal), covering $3.1 < |\eta| < 4.9$, is segmented into three layers in depth, the first of which is optimised for electron and photon shower measurements, while the second and third layer are meant to measure hadronic jets. Therefore copper plates are used in the first layer, while the second and third layer consist of tungsten. Each layer contains longitudinal channels, regularly spaced, and filled with the electrode structure consisting of concentric rods and tubes parallel to the beam axis. LAr in the gap between the rod and the tube is used as the sensitive medium. The design resolution of the FCal for jets is:

$$\frac{\sigma(E)}{E} = \frac{100\%}{\sqrt{E(\text{ GeV})}} \oplus 10\%. \tag{3.6}$$

3.5 Muon System

Muons are the only charged particles that are able to pass through the ID and calorimeters without being stopped. Therefore another set of three layers of detector material surrounds the calorimeter in order to detect muons and measure their momenta. The layout of the muon system (also sometimes called the muon spectrometer, or MS) is shown in Fig. 3.6.

To be able to measure the muon momenta in this sub-detector, three large, superconducting air-toroid magnets are installed around the muon detector layers. Each of the three arrays of magnets consists of eight coils arranged radially and symmetrically around the beam axis. In the range $|\eta| < 1.4$ the field of the large barrel toroid alone achieves the bending of muon tracks, while in the range $1.4 < |\eta| < 1.6$ the two endcap toroids, which are installed at an offset to the central toroid, contribute to the magnetic field encountered by the muons. In the region $1.6 < |\eta| < 2.7$ the muons only encounter the magnetic field of the endcap toroids alone. Since the toroid coils are arranged parallel to the beam axis the magnetic field of the muon system is orthogonal to the field of the central solenoid and bends the muon trajectories in the $R-z$ plane. The performance in terms of bending power is given by the integral of the magnetic field along a hypothetical trajectory of a muon with infinite momentum $\int B_{\perp} dl$. The barrel toroid provides a bending power between 1.5 and 5.5 T m up to $|\eta| < 1.4$, while the end-cap toroids offer between 1 and 7.5 T m in the range $1.6 < |\eta| < 2.7$. In the transition region where the two magnet systems overlap the bending power is somewhat lower.

Different types of detector technologies are used in the muon system, due to the dual requirements of measuring the muon positions with high precision, as well as having fast detector readout to associate a muon signature with a specific bunch crossing and use this information for online selection of events. Monitored Drift Tube chambers (MDTs) are used for the second and third muon spectrometer layers over the entire range of $\eta < 2.7$, while they are only used up to $\eta < 2.0$ in the first layer. In the first layer they are complemented by multiwire proportional chambers, with cathode strips in orthogonal direction, so-called Cathode Strip Chambers (CSC), because these have a better time resolution. The CSCs measure the muon position in the $R-\phi$ plane with an intrinsic precision of $40\,\mu\text{m} \times 10\,\text{mm}$, while the MDTs only measure the track positon in the z direction with a resolution of $35\,\mu\text{m}$.

Since both these types of detectors are mostly designed for precision tracking and the MDTs do not measure position in ϕ, they need to be complemented by faster detectors for the online selection of muon events. This is done in the central region $|\eta| < 1.05$ with so-called Resistive Plate Chambers (RPC) with

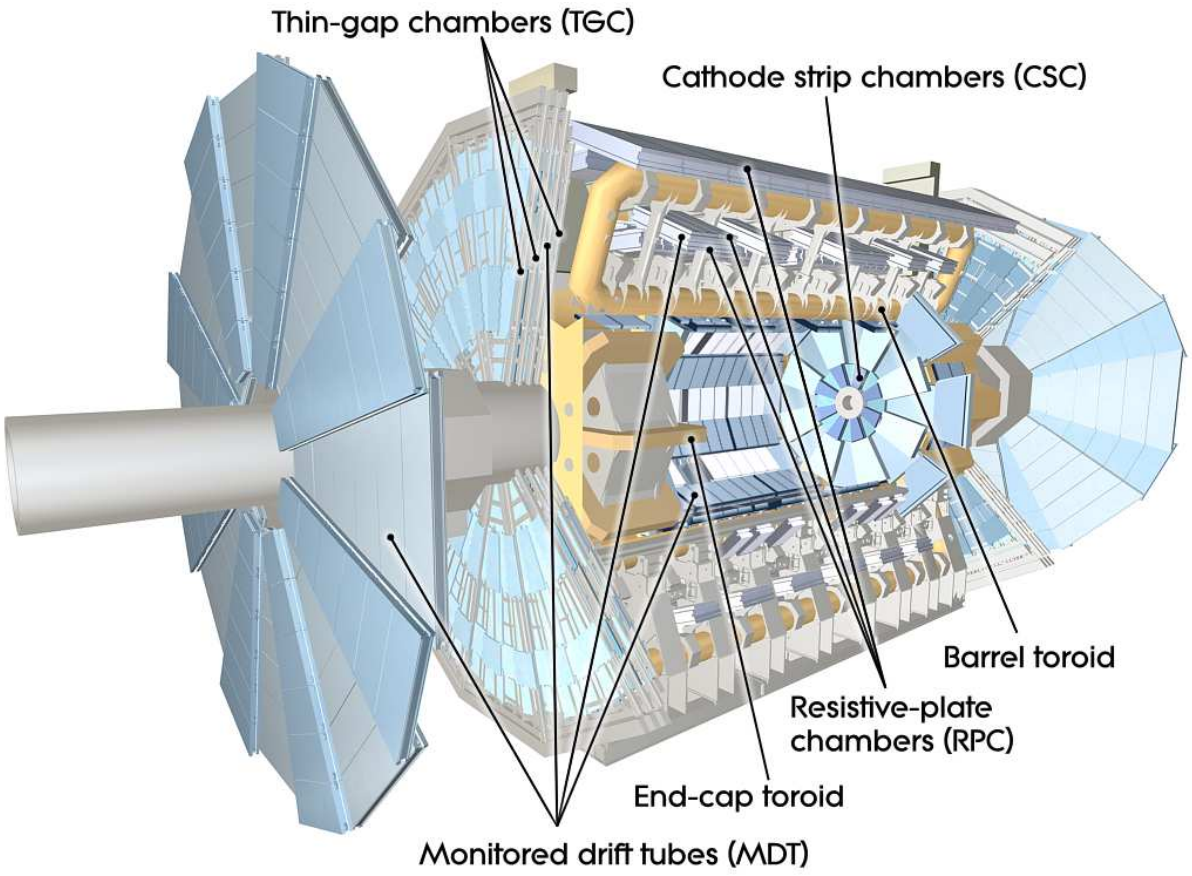


Figure 3.6: Cut-away view of the ATLAS muon system [79].

a time resolution of 1.5 ns and in the region $1.05 < |\eta| < 2.4$ with Thin Gap Chambers (TGCs) with a time resolution of 4 ns.

With this arrangement of muon track detectors good momentum resolution (better than 10%) can be reached for muon tracks up to 1 TeV. For low-momentum muon tracks, the momentum measurement usually needs to be complemented by the p_{T}-measurement in the inner detector due to the strong bending of the muons by the magnetic field of the toroids.

3.6 Trigger

Given the high event rate of about 40 MHz and an average event size of the order of a few MB on disk, a trigger system is needed to reduce the number of recorded events significantly. A design output rate of about 200 Hz is achieved by employing a combination of hardware- and software-based triggers in three levels. The hardware-based level-one trigger (L1) selects single objects with high p_{T}, such as muons, electrons, photons, jets or hadronically decaying τ leptons as well as the sum of the transverse momenta of these objects to identify significant amounts of missing transverse momentum. The information used to identify these interesting signatures comes only from a subset of the detectors, with a coarser granularity to achieve the shortest possible calculation time of less than 2.5 µs. This reduces the event rate down to 75 – 100 kHz. In addition to the pass-or-fail decisions, the L1 trigger system also provides a list of so-called "Regions of Interest" (RoIs), in which the high-p_{T} objects, that passed any of the trigger thresholds, lie. For each RoI the following information is provided: the type of object that fired the L1 trigger, the thresholds that were passed and the location in the detector.

These RoIs are passed as input to the software-based level 2 trigger (L2). Using the full set of sub-detectors at full granularity within these RoIs, the L2 trigger has about 40 ms to make further decisions. Due to the time constraint, the algorithms employed to identify and select objects have to be simplified versions of the offline selection algorithms. This reduces the event rate to about 3.5 kHz. The final layer of trigger selection is called the Event Filter (EF), which also has the full granularity and the full information from all sub-detectors available. Since the EF trigger employs algorithms that are very similar to the offline selection of particles, the computing time per event amounts to about 4 s. The processing of events with the Event Filter is done on separate computing clusters inside the ATLAS cavern and reduces the event output to the desired 200 Hz.

3.7 Forward Detectors

For the measurement of the luminosity delivered by the LHC an online LUminosity measurement using the Cherenkov Integrating Detector (LUCID) is performed with detector elements situated at $z = \pm 17$ m and very close to the beam pipe ($|\eta| \approx 5.8$). This is the main luminosity monitor for the ATLAS experiment, which measures primarily the relative luminosity, by detecting inelastic proton-proton scattering events in the forward region. The Zero Degree Calorimeter (ZDC) at $z = \pm 140$ m ($|\eta| > 8.3$), the Minimum Bias Trigger Scintillators (MBTS) at $z = 3.65$ m ($2.09 < |\eta| < 3.84$), the Beam Conditions Monitor (BCM) [80] at $z = \pm 1.84$ m ($|\eta| = 4.2$), the inner detector ($|\eta| < 2.5$) and the electromagnetic calorimeter ($|\eta| < 4.9$) also all offer capabilities for luminosity measurements to cross-check the measurement by LUCID. The ALFA detector (Absolute Luminosity For ATLAS), a scintillating fibre tracker located inside roman pots at $z = \pm 240$ m, also measures the luminosity by detecting elastic proton-proton scattering at very small angles [81]. The relative uncertainty on the integrated luminosity recorded with the ATLAS experiment in 2010 and 2011 is estimated as 3.4% [82].

Chapter 4

Analysis Objects

When performing a particle physics measurement it is clear that one does not observe the particles described in Chapter 2 directly but only through their interaction with the detector, or the interaction of their decay products. In order to correctly separate the detector signatures of different elementary particles, many different algorithms have been developed. These algorithms are not only designed to identify different elementary particles, but also to measure their kinematical quantities. At the ATLAS Experiment specific algorithms exist to identify muons, electrons, photons, hadronic jets, b-quark jets and hadronic τ-lepton decays and brief descriptions of the methods used to identify each but the last type of object are given in this chapter. A more detailed description of the τ-lepton reconstruction and identification algorithms, together with some original studies performed as part of the work for this thesis, is given in Chapters 5 and 6. In Sec. 4.1 an introduction to the basic detector-level objects, which are used in forming the higher-level physics objects, is given. The higher-level physics objects are described in Sec. 4.2. Apart from hadronic τ-lepton decays, the objects of particular interest in the measurementes presented in this thesis are hadronic jets, electrons and muons. Photons are not used explicitly. Sec. 4.2.4 introduces another important quantity in SUSY searches, the transverse momentum imbalance in the detector, called $E_{\mathrm{T}}^{\mathrm{miss}}$. This chapter closes with an explanation of how ambiguities caused by multiple physics objects reconstructed from the same detector signature of a real particle are resolved in Sec. 4.3.

4.1 Basic Objects in the Detector

All particle reconstruction and identification algorithms at the ATLAS Experiment start from two types of basic reconstructed objects: charged particle tracks and calorimeter cell clusters. The reconstruction of these objects is explained in the following section, as far as it is necessary to understand the more complicated particle reconstruction algorithms.

4.1.1 Track Reconstruction

The reconstruction of charged particle tracks proceeds in several steps. In the first step, three-dimensional space points are built from the signals in the SCT pixel and strip layers. Also the signals in the individual straw tubes in the TRT are combined into so-called calibrated drift circles [83]. In the next step seeds for track finding are built from all clusters of three or more space points which are close to each other in the first four layers of the SCT. These seeds are extrapolated to the outer layers of the SCT and matched with reconstructed space points in these layers. A track fit is performed, after which outlier clusters are removed, cluster-to-track association ambiguities are resolved, and tracks are rejected according to quality criteria like the number of hits in the different layers and the goodness of the track fit. In the next step, these track candidates are associated with the reconstructed drift circles. These extended tracks are refitted using the information from SCT and TRT and compared to the SCT-only tracks. TRT hits

associated with the tracks which significantly reduce the quality of fit compared to the SCT-only fit are labeled as outliers, meaning that they stay associated with the track, but are not used in the fit.

After this procedure is completed another algorithm, which is called "backtracking", is applied, in which track segments in the TRT that have not been associated with any SCT hits by the default tracking algorithm are extended back into the SCT and combined with SCT track fragments which were not selected as track seeds in the SCT. These kind of track fragments are often created by a particle decaying into charged daughters outside the first few layers of the SCT and therefore the backtracking algorithm improves the reconstruction efficiency for electron tracks from photon conversions and tracks from long-lived particle decays.

After both tracking algorithms are completed, a dedicated vertex finding algorithm is executed to reconstruct the primary collision vertex. This is followed by separate algorithms intended to identify photon conversions and secondary decay vertices.

4.1.2 Calorimeter Cluster Reconstruction

As the ATLAS calorimeter is a very finely segmented sampling calorimeter, there are several different possibilities to reconstruct physics objects from the three-dimensional calorimeter shower information. The standard method of reconstructing particle showers in the ATLAS calorimeter is to collect the energy in individual cells into so-called topological clusters (topoclusters) [84]. This clustering algorithm selects seed cells that exceed a given energy threshold. By default this threshold is set to four times the average noise signal σ_{noise} of the cell. Cells which are adjacent and in which more energy than twice their σ_{noise} is measured are added to the cluster. Cells are added to the cluster until no more adjacent cells are found which exceed their σ_{noise} in this way and in a final step all adjacent cells are added which have recorded positive energy. In a final step the energy depositions of neighbouring cells are compared to find local maxima of the energy depositions within a cluster. If more than one maximum is found, the cluster is split into two smaller clusters until only one local maximum of energy deposition is left within each cluster. Because the ATLAS calorimeter is non-compensating, all cell energies are measured at a scale appropriate for measuring electromagnetic showers (EM scale). Therefore each cluster is classified as being of either hadronic or electromagnetic origin using cluster shape variables and the clusters are calibrated accordingly. This procedure is know as local hadron calibration (LC) [85].

4.2 Physics Objects

In the context of the physics measurements detailed in the following chapters the tracks and cell clusters described above are interpreted as individual particles. It is necessary therefore to give an overview of how these basic reconstruction objects are combined and what selection criteria are applied to derive the actual physics objects used in further analysis.

4.2.1 Jet Selection

Different jet-finding algorithms [86–88] can be used to cluster the energy depositions from hadronic jets, using either tracks, calorimeter cells, towers or clusters as input. The standard method at the ATLAS Experiment is to apply an anti-k_T jet-finding algorithm [87] with a distance parameter $R = 0.4$ on the topoclusters described above. The four-momenta of the clusters associated to each jet are added assuming zero mass for each of the four-vectors. Unless stated otherwise, the hadronic jets used in the following chapters are selected from the list of calorimeter jets produced by the jet-finding algorithm by requiring $|\eta| < 2.8$ and $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, where p_{T} is the component of the jet momentum transverse

to the beam axis. For the requirement on p_{T} an additional calibration is applied to the jets, which uses numerical inversion of the true response of hadronic jets determined from simulation [85]. All of the jets within an event that pass the η and p_{T} requirements after calibration, along with further so-called cleaning criteria designed to reject jets reconstructed from detector noise and other instrumental effects, are considered as hadronic jet candidates.

Due to the long lifetime of b-hadrons, it is possible to separate jets that contain b quarks from jets that only contain light quarks and gluons. This can be exploited to separate SM processes like $t\bar{t}$ from other processes without b-quark jets (see Chapter 7). In this thesis, a neural network called `JetFitterCOMBNN` [89] is used for this purpose. The efficiency and purity of this "b-tagging" algorithm at the chosen working point were measured in ATLAS data to be 60% and 94.6%, respectively [90], and a slight mismodelling of these values in the simulation is corrected for.

4.2.2 Electron Selection

Isolated electrons are reconstructed using a sliding-window algorithm to find seed cell clusters in the electromagnetic calorimeter. One isolated high-p_{T} track is matched to the seed cluster by comparing the cluster position with the track position extrapolated to the calorimeter surface. The track momentum p_{track} is also required to be consistent with the cluster energy E_{clus} within $E_{clus}/p_{track} < 10$.

Electron candidates found in this way are identified using information on the track-to-cluster matching, calorimeter shower shapes and the number of high-threshold hits in the TRT. For the direction of electron candidates the direction of the track is used, unless the sum of hits and dead sensors traversed by the track in the silicon tracking detector is below four. In this case, the direction of the cluster associated with the electron candidate is used. This is then used as the direction of the electron momentum and the magnitude of the momentum is set equal to the total energy in the electron cluster, neglecting the electron mass. The transverse momentum p_{T} is then obtained in the usual way from the x- and y-components of the momentum four-vector. The following acceptance requirements are applied: $|\eta| < 1.37$ || $1.52 < |\eta| < 2.47$ and $p_{\mathrm{T}} > 20$ GeV.

4.2.3 Muon Selection

Muons are reconstructed from a combination of inner detector (ID) tracks and tracks in the muon spectrometer. These so-called combined muons use the MS tracks by extrapolating them back to the interaction point, applying corrections for energy loss in the detector material, and combining them with the tracks reconstructed in the inner detector. A χ^2 fit of the combined track is evaluated to classify the quality of the match. The momentum of the muons is taken from either the MS (for high p_{T}) or ID (for low p_{T}) measurement.

To improve the muon selection efficiency, also such muons which are found by extrapolating inner detector tracks outwards to the MS and matching them to incomplete segments of MS tracks, so-called segment-tagged muons, are used. For these muon candidates, the momentum measurement of the MS track segment is not reliable, which is why only the ID measurement is used to reconstruct their momenta.

Muons must satisfy $|\eta| < 2.4$ and $p_{\mathrm{T}} > 20$ GeV and fulfill further identification criteria. The most important of these are: at least one hit in the pixel detector and at least six hits in the SCT, of which one has to be in the b-layer, are required for the ID track. For muons in the central region, $|\eta| < 1.9$, at least six hits in the TRT are required as well. For muons with more than five hits in the TRT, the number of TRT outlier hits needs to be less than 10% of the total number of TRT hits (both associated

and outliers). For all the criteria on the number of hits in the tracking detector, dead modules traversed by the extrapolated muon track are counted as hits.

4.2.4 $E_{\mathrm{T}}^{\mathrm{miss}}$ Reconstruction

The missing transverse energy, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, is of great importance in searches for R-parity conserving SUSY scenarios, because it represents the vectorial sum of the transverse momenta of all particles in a collision that escape the detector unobserved. R-parity conserving SUSY predicts two LSPs per event that escape the detector unseen if they are only weakly interacting.

Missing transverse energy is reconstructed as the negative vector sum of the total transverse energy in an event and is therefore a two-component vector:

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = \begin{pmatrix} E_{\mathrm{x}}^{\mathrm{miss}} \\ E_{\mathrm{y}}^{\mathrm{miss}} \end{pmatrix}. \tag{4.1}$$

In principle this quantity can be computed by projecting the momenta of all calorimeter objects onto the transverse plane. However due to the different calibrations that need to be applied to different physics objects, better results can be obtained by associating cells with identified physics objects and applying the proper calibration for these physics objects. In addition, the transverse momenta of all identified muons are taken into account by adding in the p_{T}-measurements of muon tracks. The full formula for the computation of the two $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ components then reads:

$$E_{\mathrm{x(y)}}^{\mathrm{miss}} = -E_{\mathrm{x(y)}}^{\mathrm{e}} - E_{\mathrm{x(y)}}^{\gamma} - E_{\mathrm{x(y)}}^{\mathrm{jets}} - E_{\mathrm{x(y)}}^{\mathrm{softjets}} - E_{\mathrm{x(y)}}^{\mu} - E_{\mathrm{x(y)}}^{\mathrm{cellout}}. \tag{4.2}$$

This formula includes one term for jets with $p_{\mathrm{T}} > 20$ GeV and one for soft jets with 7 GeV $< p_{\mathrm{T}} <$ 20 GeV, because a special calibration needs to be applied to the soft jets. The $E_{\mathrm{x(y)}}^{\mathrm{cellout}}$ term includes all cells and clusters that cannot be associated with a physics object. Assuming perfect particle identification and energy/momentum resolution, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ should be equal to the vectorial sum of transverse momenta of all non-detectable particles. In the following both the vector $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and its magnitude are denoted by the symbol $E_{\mathrm{T}}^{\mathrm{miss}}$.

It is important to note that in this definition hadronic τ-lepton decays are treated as hadronic jets, because their calibration methods are similar enough that the performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement in events with real hadronic τ-lepton decays is not significantly affected.

4.3 Overlap Removal

Since all the particle selection algorithms detailed above operate independently on the same basic reconstruction output of the ATLAS detector, i.e. lists of tracks, clusters and jets, it is common for these objects to be associated several times with different physics objects. To avoid double-counting of the detector signature of one single real particle (or jet), it is necessary to apply a so-called "overlap removal" algorithm. The aim of this algorithm is to remove any duplicate reconstructed physics objects, which stem from the same true physics object. This is achieved by assigning priorities to the different reconstructed objects, by which any ambiguity between two overlapping objects can be resolved. The following algorithm is used in this thesis:

1. if both an electron and a τ candidate are selected within $\Delta R < 0.2$, the electron is kept and the τ candidate is rejected;

2. if both a muon and a τ candidate are selected within $\Delta R < 0.2$, the muon is kept and the τ candidate is rejected;
3. if both an electron and a jet are selected within $\Delta R < 0.2$, the electron is kept and the jet is rejected;
4. if both a τ candidate and a jet are selected within $\Delta R < 0.2$, the τ candidate is kept and the jet is rejected;
5. if both a jet and an electron are selected within $0.2 < \Delta R < 0.4$, the electron is assumed to be from a secondary decay within the jet and is rejected;
6. if both a jet and a muon are selected within $\Delta R < 0.4$, the muon is assumed to be from a secondary decay within the jet and is rejected.

This algorithm is identical to that used in other SUSY searches at the ATLAS Experiment, except for the addition of the steps specific to τ candidates; these are given lower priority than electrons and muons since the misidentification probabilities for electrons and muons are much lower than the ones for τ candidates. The efficiency for real, hadronic τ-lepton decays, that were reconstructed as τ candidates, to pass these overlap removal steps is very high (above 99.9%), while up to 10% of true electrons and muons are removed. Muons very rarely fake jets, so no dedicated muon-jet overlap removal step is included. The final two steps listed are not, strictly speaking, overlap removal, since the reconstructed objects may be sufficiently distinct. Instead, these steps reject leptons from secondary decays, in particular of heavy quarks, within the jet; the jets are retained in the object selection while the leptons, assumed to be unassociated with the physics process of interest, are discarded.

Chapter 5

Reconstruction and Identification of Hadronic τ-Lepton Decays

This chapter gives an overview of the methods used at the ATLAS Experiment to reconstruct hadronic τ-lepton decays and to separate them from signatures of other physics objects like hadronic jets from quarks or gluons and lighter leptons. The leptonic decay modes of the τ lepton are not considered here, since these would be classified as electrons or muons during reconstruction and the identification algorithms used for light leptons (see Sec. 4.2.2 and Sec. 4.2.3) can be applied to them.

The algorithms used to find hadronic τ-lepton decays in the detector can be split into two categories. The first category, which is described in Sec. 5.1, consists of so-called reconstruction algorithms. These are designed to collect all detector responses, like tracks in the inner detector and energy deposits in the calorimeter, to the visible part of the τ-lepton decay and construct from these a so-called "τ candidate". Since the reconstruction algorithms are not designed to separate real τ-lepton decays from other physics objects, the term τ candidate only denotes the reconstructed object and does not imply a decision on which particle(s) actually were reconstructed as a τ candidate. From the tracks and calorimeter deposits the four-momentum of the τ candidate is reconstructed and identification variables are computed that encode tracking and shower-shape information characteristic for hadronic τ-lepton decays.

This information is used by the second class of algorithms to identify hadronic τ-lepton decays and separate them from other particles. The LHC being a hadron collider, the most abundant process in the ATLAS detector is the production of hadronic jets, with a production cross sections several orders of magnitude larger than that of the most copious source of real τ leptons, which is the decay of W or Z bosons. Several algorithms have been developed to provide the necessary suppression against misidentified τ candidates coming from hadronic jets, using rectangular cuts, a projective likelihood or a boosted decision tree [91] (henceforth denoted as BDT_j). Since only the BDT_j method is used in the studies presented in this thesis, Sec. 5.2 will focus on describing this way of discriminating against hadronic jets, with only a brief comparison of its performance with the performance of the cut-based and projective likelihood identification. The next most probable source of misidentified τ candidates are light leptons, especially electrons. Two methods for suppressing electrons are commonly used at the ATLAS Experiment; a simple variant, applying rectangular selection criteria on shower shape and tracking variables that discriminate between electrons and charged pions (which are mainly produced in hadronic τ-lepton decays), and a boosted decision tree classifier. Again, bearing in mind the application to the physics analysis, only the simpler, cut-based variant for electron rejection (henceforth denoted as $e^{\pm}$-veto) is explained in detail in Sec. 5.2. Details on the other identification methods can be found in [91]. Sec. 5.3 gives a short overview of the methods used at the ATLAS Experiment to determine the selection efficiency of the τ reconstruction and identification algorithms for real hadronic τ-lepton decays.

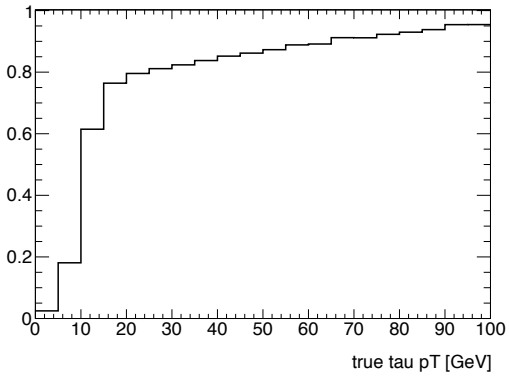


Figure 5.1: Efficiency of the calorimeter-seeded τ-reconstruction algorithm as a function of the transverse momentum of the visible τ-lepton decay products. The efficiency calculation includes a requirement for at least one reconstructed track with $p_{\mathrm{T}} \geq 1$ GeV being associated to the candidate.

5.1 Tau Reconstruction

The reconstruction of hadronic τ-lepton decays starts with the selection of so-called "seeds", which can be either energy deposits in the calorimeter or tracks of charged particles in the inner detector. For both seed types there exists a dedicated algorithm, which applies selection criteria, measures the kinematic variables of the seed and calculates additional variables that can be used as input for the identification algorithms described later. A large part of the technical work performed for this thesis was to maintain and improve the calorimeter-seeded τ-candidate reconstruction, as well as to integrate it together with the track-seeded reconstruction into one single joint algorithm, which has become the standard in the ATLAS Experiment. Therefore only this algorithm will be discussed here in detail. Further discussion of the two separate τ-lepton reconstruction algorithms can be found in [83]. The validation of τ-lepton reconstruction since the beginning of data-taking at the ATLAS Experiment at a centre-of-mass energy of $\sqrt{s} = 900$ GeV, which was also part of the work for this thesis, is documented in [92–94].

In the calorimeter-seeded algorithm, the anti-k_T topocluster jets introduced in Sec. 4.2.1 are used as seeds for reconstructing τ candidates, albeit with different acceptance criteria. Any reconstructed jet with $p_{\mathrm{T}} \geq 10$ GeV and $|\eta| \leq 2.5$, which corresponds to the acceptance of the tracking system, is considered a viable τ candidate. The efficiency of this selection, including a requirement for at least one reconstructed track with $p_{\mathrm{T}} \geq 1$ GeV to be associated with the τ candidate, as determined for simulated $W \to \tau\nu$ decays, can be seen in Fig. 5.1.

5.1.1 Track Reconstruction

Tracks of charged particles found in the inner detector are then associated with the τ candidates by requiring a maximum distance of $\Delta R = 0.4$ in the η and ϕ plane to the energy-weighted barycentre of the calorimeter jet that seeded the τ candidate. These tracks are then required to satisfy certain quality criteria [91] to ensure that they come from particles produced in a highly energetic proton-proton collision. These quality criteria are:

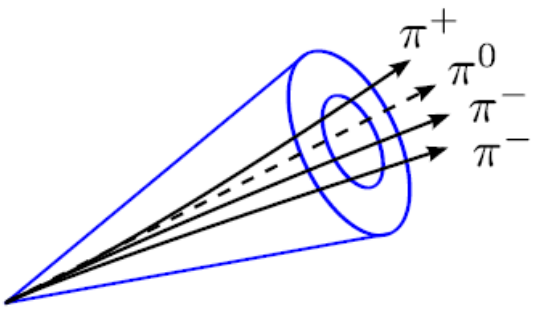


Figure 5.2: Illustration of track selection around the calorimeter seed of a τ candidate. Both the inner and outer cone are centered around the energy-weighted barycentre of the calorimeter jet seed. The neutral pion as well as the tracks from the charged pions fall into the core cone, which is suggested by the inner circle. The space between the inner and the outer cone is called the isolation annulus, which is used to calculate identification variables for the suppression of hadronic jet background.

$$p_{\mathrm{T}} > 1\ \mathrm{GeV}, \tag{5.1}$$
$$\text{number of pixel hits} \geq 2, \tag{5.2}$$
$$\text{number of pixel hits + number of SCT hits} \geq 7, \tag{5.3}$$
$$|d_0| < 1.0\,\mathrm{mm}, \tag{5.4}$$
$$|z_0 \sin\theta| < 1.5\,\mathrm{mm}, \tag{5.5}$$

where d_0 is the distance of closest approach in the transverse plane between the reconstructed vertex and the track and z_0 denotes the distance of closest approach in longitudinal direction. All tracks fulfilling these quality criteria are classified as belonging to the core cone ($\Delta R < 0.2$) or the isolation annulus ($0.2 \leq \Delta R \leq 0.4$) of the candidate, as illustrated in Fig. 5.2. Tau candidates are classified as single- or multi-prong, depending on the number of tracks counted in the core cone, since the products of a real hadronic τ-lepton decay are expected to be contained within the core cone, while the isolation annulus is expected not to contain any decay products from real hadronic τ-lepton decays. Information on the amount of energy deposited, or charged particles found in the isolation annulus with respect to the core cone is therefore used in several of the discriminating variables for separating real hadronic τ-lepton decays from candidates reconstructed from quark- or gluon-initiated jets. Fig. 5.3 shows the distributions of two of these identification variables. The definition of the two variables is given in Sec. 5.1.3.

5.1.2 Calibration

Although the transverse momentum of the jet seed is calculated from the sum of the four-vectors of all clusters associated with the jet, the computation of the transverse momentum of τ candidates only uses the sum of the four-vectors of all clusters within $\Delta R < 0.2$ of the jet seed direction. This has been found to make the p_{T} measurement more resistent to pile-up effects (see Sec. 2.3) and also to yield a good resolution for real hadronic τ-lepton decays, as the visible decay products of τ-lepton decays are known to be well contained within the smaller core cone for all the p_{T} ranges considered here. A comparison of the reconstructed transverse momentum with the true transverse momentum of the hadronic τ-lepton decay products was performed for a mixed sample of simulated $Z \to \tau\tau$ and $Z' \to \tau\tau$ decays. In this

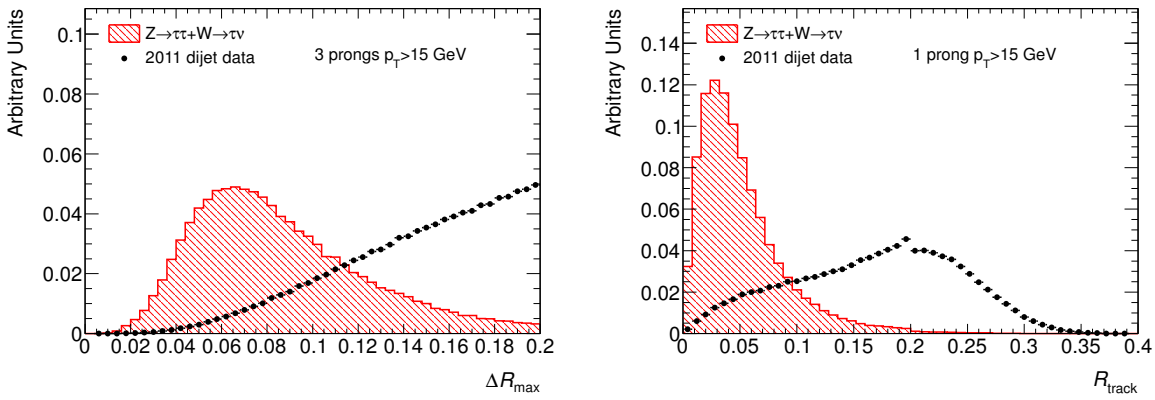


Figure 5.3: Distributions of identification variables that characterize the distance of reconstructed tracks associated with τ candidates from the jet seed axis. The maximum distance of tracks found in the core cone is shown on the left and the average distance of all tracks, including the isolation annulus, is shown on the right. The definition of the two variables is given in Sec. 5.1.3. The distributions are shown both for τ candidates reconstructed from simulated hadronic τ-lepton decays (red dashed histogram) and a data sample enriched in quark- and gluon-initiated jets (black markers).

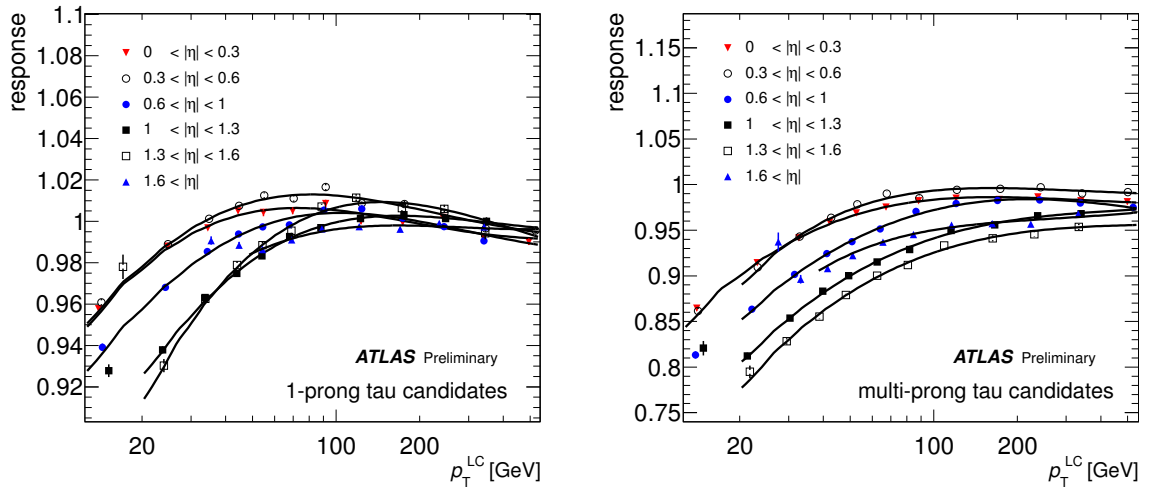


Figure 5.4: Response curves as a function of reconstructed τ-candidate p_T before applying the truth-based correction factors for one-prong (left) and multi-prong (right) τ candidates in various η bins [91].

sample, the distribution of the ratio of generated over reconstructed p_{T} is fitted with a Gaussian function in bins of true p_{T}, reconstructed η and reconstructed number of prongs. Using the mean of the Gaussian fit in each true p_{T} bin, response curves as a function of reconstructed p_{T} for all η and number of prong bins are constructed and shown in Fig. 5.4. These response curves are used to correct the reconstructed τ candidate p_{T} to ensure that the response for real hadronic τ-lepton decays is as close to unity as possible. Further studies of the performance and uncertainties of this procedure are presented in [91].

5.1.3 Identification Variables

While the actual identification of hadronic τ-lepton decays is performed with separate algorithms, which are based on multi-variate classifiers of different types, the input distributions for these classifiers are already prepared within the reconstruction algorithms. Many different variables are used in separating hadronic τ-lepton decays from hadronic jets and electrons. The full list is given in [91]. Here only the subset of variables used by the BDT$_j$ classifier for the suppression of quark- and gluon-initiated jets and the $e^{\pm}$-veto for the suppresion of τ candidates from electrons or positrons is shown.

Identification Variables: Jet Rejection

The BDT$_j$ classifier uses eleven different variables to separate real hadronic τ-lepton decays from hadronic jets in the ATLAS detector. Most of these variables describe the multiplicity and relative proximity of tracks and energy depositions within the τ candidate. For real hadronic τ-lepton decays, one expects a small number of charged and neutral pions with a smaller spread of tracks and showers compared to the high-multiplicity of spread-out particles in hadronic jets from quarks or gluons. The invariant mass of the products of the τ-lepton decay can be reconstructed from calorimeter clusters as well as charged tracks (neglecting the neutral component). Finally, the secondary vertex created by the decay of a τ lepton into multiple charged pions can be reconstructed from the tracks associated with the τ candidate. Example distributions of these variables are shown in Fig. 5.5 for simulated $Z \to \tau\tau$ and $W \to \tau\nu$ samples and a sample of di-jet events collected from 2011 ATLAS data. The full list of the identification variables with their definitions is given below:

Number of isolation tracks ($N^{\mathrm{iso}}_{\mathrm{track}}$): the number of tracks reconstructed in the isolation annulus ($0.2 \leq \Delta R \leq 0.4$) around the jet-seed axis.

Track radius (R_{track}): the p_{T} weighted average distance of tracks to the jet-seed direction:

$$R_{\mathrm{track}} = \frac{\sum_i^{\Delta R_i < 0.4} p_{\mathrm{T},i}\, \Delta R_i}{\sum_i^{\Delta R_i < 0.4} p_{\mathrm{T},i}}, \tag{5.6}$$

where i runs over all core and isolation tracks of the τ candidate, $p_{\mathrm{T},i}$ is the track transverse momentum and ΔR_i is the distance in the $\eta - \phi$-plane between the track i and the jet-seed axis.

Maximum ΔR (ΔR_{max}): the maximal ΔR between a core track and the jet-seed axis.

Track mass (m_{tracks}): the invariant mass of the track system, where the tracks used for the invariant mass calculation are both core and isolation tracks:

$$m_{\mathrm{tracks}} = \sqrt{\left(\sum_{\mathrm{tracks}} E\right)^2 - \left(\sum_{\mathrm{tracks}} \mathbf{p}\right)^2}. \tag{5.7}$$

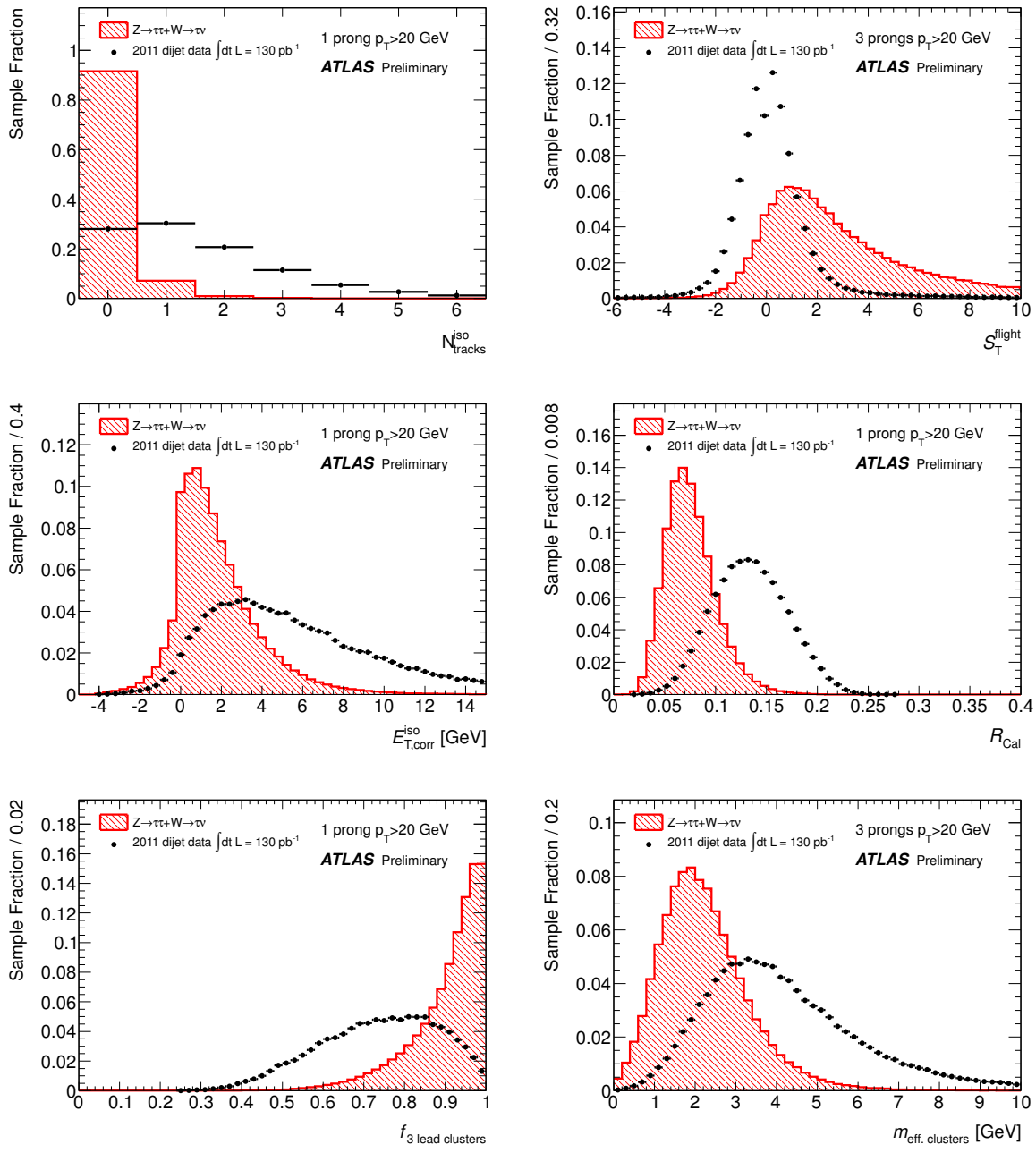


Figure 5.5: Example distributions of identification variables used in the BDT$_j$ classifier. The red dashed histograms show the expected distribution from a simulated sample mix of $Z \to \tau\tau$, $Z' \to \tau\tau$ and $W \to \tau\nu$ decays, while the black markers are for a sample of QCD di-jet events collected from 2011 data [91].

Leading track momentum fraction (f_{track}):

$$f_{\text{track}} = \frac{p_{\text{T},1}^{\text{track}}}{p_{\text{T}}^{\tau}}, \tag{5.8}$$

where $p_{\text{T},1}^{\text{track}}$ is the transverse momentum of the leading or highest-p_{T} track in the core region and p_{T}^{τ} is the transverse momentum of the τ candidate, calibrated at the EM energy scale. Even though real hadronic τ-lepton decays can have a significant contribution from neutral particles, this variable is expected to exhibit higher values for real hadronic τ-lepton decays than for hadronic jets, because of the smaller charged and neutral particle multiplicity.

Leading clusters energy ratio ($f_{\text{3 lead clusters}}$): the ratio of the energy of the three most energetic clusters over the total energy of all clusters associated with the τ candidate.

$$f_{\text{3 lead clusters}} = \frac{\sum_{i\in\{1,2,3\}}^{\Delta R_i<0.2} E_{\text{T},i}}{\sum_{j\in\{\text{all}\}}^{\Delta R_j<0.2} E_{\text{T},j}}, \tag{5.9}$$

where i and j run over all clusters associated with the τ candidate within $\Delta R < 0.2$. ΔR_i is defined as the distance between a barycentre of the cluster and the jet-seed axis. $E_{\text{T},i}$ is the cluster transverse energy, calibrated at the LC scale.

Core energy fraction (f_{core}): the fraction of transverse energy within ($\Delta R < 0.1$) of the jet-seed axis:

$$f_{\text{core}} = \frac{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.1} E_{\text{T},i}}{\sum_{j\in\{\text{all}\}}^{\Delta R_j<0.4} E_{\text{T},j}}, \tag{5.10}$$

where i runs over all cells associated with the τ candidate within $\Delta R < 0.1$ and j runs over all cells in the wider cone of $\Delta R < 0.4$. The calorimeter cells associated with a τ candidate are those which are contained in the topological clusters that are constituents of the jet seeding the τ reconstruction. ΔR_i is defined as the distance between a calorimeter cell and the jet-seed axis. $E_{\text{T},i}$ is the cell transverse energy, calibrated at the EM scale.

Calorimetric radius (R_{Cal}): the energy-weighted, average distance to the jet-seed axis of cells in both the electromagnetic and the hadronic calorimeter within $\Delta R < 0.4$ of the jet seed:

$$R_{\text{Cal}} = \frac{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.4} E_{\text{T},i}\,\Delta R_i}{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.4} E_{\text{T},i}}, \tag{5.11}$$

Cluster mass ($m_{\text{eff. clusters}}$): the invariant mass computed from the constituent clusters of the seed jet, calibrated at the LC energy scale:

$$m_{\text{eff. clusters}} = \sqrt{\left(\sum_{\text{clusters}} E\right)^2 - \left(\sum_{\text{clusters}} \mathbf{p}\right)^2}. \tag{5.12}$$

To minimise the effect of pileup, only the first N leading E_{T} clusters (so-called "effective clusters")

are used in the calculation, defined as

$$N = \frac{(\sum_i E_{\mathrm{T}i})^2}{\sum_i E_{\mathrm{T}i}^2}, \tag{5.13}$$

where i runs over all clusters associated with the τ candidate, and N is rounded up to the nearest integer.

Transverse flight path significance ($S_{\mathrm{T}}^{\mathrm{flight}}$): the decay length significance of the secondary vertex for multi-prong τ candidates in the transverse plane:

$$S_{\mathrm{T}}^{\mathrm{flight}} = \frac{L_{\mathrm{T}}^{\mathrm{flight}}}{\delta L_{\mathrm{T}}^{\mathrm{flight}}}, \tag{5.14}$$

where $L_{\mathrm{T}}^{\mathrm{flight}}$ is the reconstructed signed decay length, and $\delta L_{\mathrm{T}}^{\mathrm{flight}}$ is its estimated uncertainty. Only core tracks are used for the secondary vertex fit.

Leading track IP significance ($S_{\text{lead track}}$): the impact parameter significance of the leading track of the τ candidate:

$$S_{\text{lead track}} = \frac{d_0}{\delta d_0}, \tag{5.15}$$

where d_0 is the distance of closest approach of the track to the reconstructed primary vertex in the transverse plane, and δd_0 is its estimated uncertainty.

Identification Variables: Electron Rejection

As opposed to the BDT$_j$ for the rejection of quark- and gluon-initiated jets, the $e^{\pm}$-veto classifier uses only four different variables. The situation is further simplified by the fact that electrons, producing only a single track in the inner detector, are very unlikely to be misreconstructed as three-prong τ candidates. Therefore they need only be compared to τ-lepton decay modes with a single charged hadron. The variables used to separate electrons from single-prong hadronic τ-lepton decays are similar to those used in regular electron identification. Distributions of these variables for a simulated signal sample of $Z \to \tau\tau$ decays and a simulated background sample of $Z \to ee$ decays are shown in Fig. 5.6. The full list of these variables with their definitions is given below:

TRT HT fraction (f_{HT}): the ratio of high-threshold to low-threshold hits (including outlier hits), in the Transition Radiation Tracker, for the leading p_{T} core track.

$$f_{\mathrm{HT}} = \frac{\text{\# High-threshold TRT hits}}{\text{\# Low-threshold TRT hits}}. \tag{5.16}$$

Since electrons are lighter than pions, and therefore have higher Lorentz γ factors, they are more likely to produce transition radiation that causes high-threshold hits in the TRT.

Electromagnetic track fraction ($f_{\mathrm{EM}}^{\mathrm{track}}$): the ratio of the transverse energy deposited in the electromagnetic calorimeter (at EM scale) to the transverse momentum of the leading track:

$$f_{\mathrm{EM}}^{\mathrm{track}} = \frac{\sum_{i\in\{\mathrm{EM}\}}^{\Delta R_i<0.4} E_{\mathrm{T},i}}{p_{\mathrm{T},1}^{\mathrm{track}}}, \tag{5.17}$$

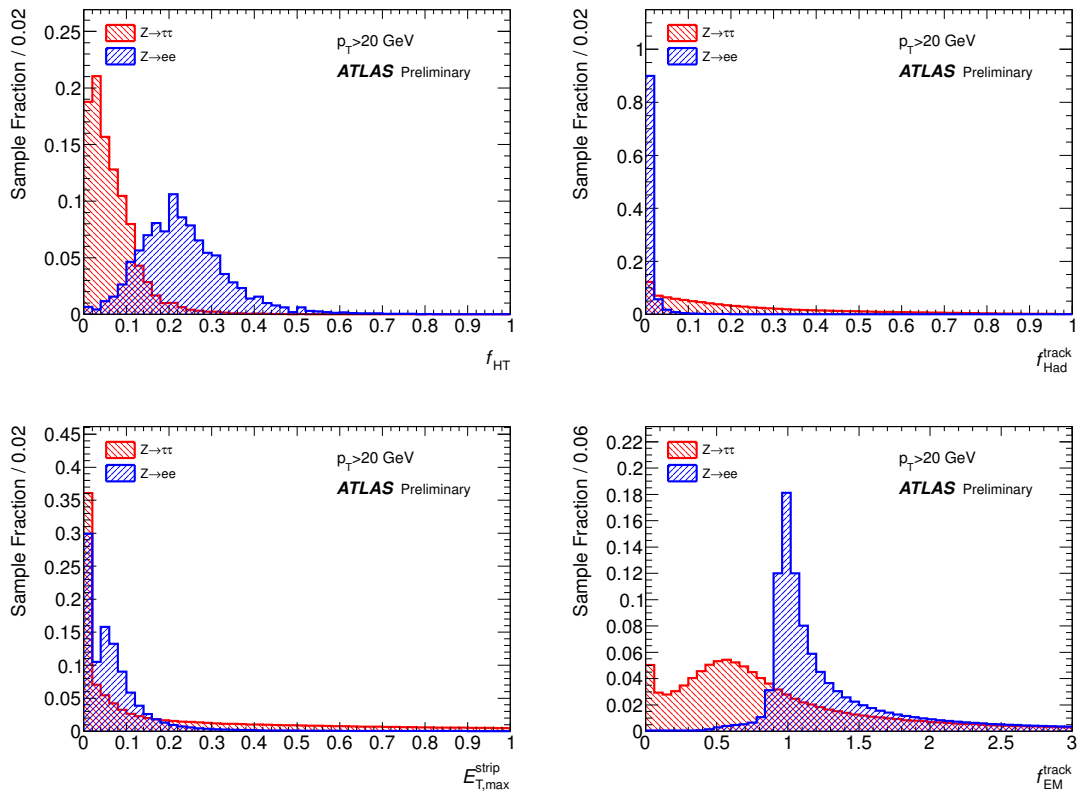


Figure 5.6: Distributions of the input variables to the $e^{\pm}$-veto classifier. Shown are the fraction of high-threshold TRT hits (top left), the hadronic track fraction (top right), the maximum strip E_{T} (bottom left) and the electromagnetic track fraction (bottom right) for a simulated signal sample of $Z \rightarrow \tau\tau$ decays (red dashed histogram) and a simulated background sample of $Z \rightarrow ee$ decays (blue dashed histogram) [91].

where i runs over all cells in the EM calorimeter within $\Delta R < 0.4$ of the jet-seed axis.

Hadronic track fraction ($f_{\text{Had}}^{\text{track}}$): the ratio of the hadronic transverse energy (at EM scale) over the transverse momentum of the leading track:

$$f_{\text{Had}}^{\text{track}} = \frac{\sum_{i\in\{\text{Had}\}}^{\Delta R_i<0.4} E_{\text{T},i}}{p_{\text{T},1}^{\text{track}}}, \tag{5.18}$$

where i runs over all cells in the hadronic calorimeter within $\Delta R < 0.4$ of the jet-seed axis.

Maximum strip E_{T} ($E_{\text{T,max}}^{\text{strip}}$): the maximum transverse energy deposited in a cell in the pre-sampler layer of the electromagnetic calorimeter, which is not associated with that of the leading track.

5.2 Tau Identification

After the identification variables as well as the calibrated four-momentum and the associated tracks of the τ candidate have been provided by the core reconstruction algorithms, a second class of algorithms, so-called identification algorithms, come into play. Since the analysis presented in this thesis uses the BDT$_j$ and cut-based $e^{\pm}$-veto, only the working principles of these methods are illustrated here. A useful measure of the performance of these identification algorithms is the efficiency for a reconstructed τ candidate to pass the additional selection requirements of the identification algorithm. Depending on whether the reconstructed τ candidate is known to come from a real hadronic τ-lepton decay or from another source, a hadronic jet or an electron, one speaks about signal or background efficiency, respectively. The information on the real source of the reconstructed τ candidate can either be obtained from truth-level information in the simulation or from a selected high-purity data sample of signal or background candidates. Once this information is available the signal and background efficiencies can be computed as follows:

Signal efficiency:

$$\varepsilon_{\text{sig}}^{\text{n-prong}} = \frac{\left(\begin{array}{l}\text{\# of } \tau \text{ candidates with } n \text{ reconstructed tracks, passing ID,}\\ \text{and truth-matched to a simulated n-prong decay}\end{array}\right)}{(\text{\# of simulated hadronic } \tau\text{-lepton decays with n prongs})} \tag{5.19}$$

Background efficiency:

$$\varepsilon_{\text{bkg}}^{\text{n-prong}} = \frac{(\text{\# of } \tau \text{ candidates with } n \text{ reconstructed tracks, passing ID})}{(\text{\# of } \tau \text{ candidates with } n \text{ reconstructed tracks})} \tag{5.20}$$

For all identification algorithms, there exist predefined selections with different "levels of tightness" that are called "loose", "medium" and "tight". For the algorithms that are used to suppress quark- or gluon-initiated jets, these selections were trained to provide a p_{T}-independent signal efficiency of approximately 70%, 50% and 30% percent, respectively. The corresponding background efficiencies are shown in Fig. 5.8. The signal and background efficiencies for the different levels and methods of electron suppression are shown in Fig. 5.10.

5.2.1 Hadronic Jet Rejection using Boosted Decision Trees

A boosted decision tree is a multivariate classifier that uses several simple yes-and-no decisions on single variables at a time to classify data as either signal or background. The toolkit for training the BDT_j classifier used for separating hadronic τ-lepton decays from hadronic jets at the ATLAS Experiment is documented in detail in [95]. This toolkit offers many different options for training decision tree classifiers to improve performance and avoid overtraining. The actual configuration used for training the BDT_j classifier is documented in [91]. For the training (growing) of the BDT_j classifier two samples of events, one with real hadronic τ-lepton decays (signal) and one with τ candidates reconstructed from hadronic jets (background) were provided. The signal sample was obtained from a combination of $W \to \tau\nu$, $Z \to \tau\tau$ and $Z' \to \tau\tau$ Monte Carlo samples generated with PYTHIA and the background samples were obtained from a selection of di-jet events in ATLAS data, collected in the spring of 2011 that amount to about $0.8\,\mathrm{fb}^{-1}$. Together with the list of eleven discriminating variables detailed above, these samples constitute the input to the training of the BDT_j classifier. A more detailed description of the sample selection can also be found in [91]. With this input a non-boosted decision tree is grown by the following algorithm:

1. The list of input variables is iterated over and the single cut value for each variable is determined that gives the best separation between signal- and background-type events.

2. The single variable that gives best separation is selected and the sample of signal and background events is separated into two daughter nodes (leaves) according to the cut on this variable.

3. This procedure is repeated on each of the daughter leaves until a stopping criterion, usually a minimum number of events of each type in one leaf, is reached.

4. If the stopping criterion has been reached for a given leaf, it is classified as either a signal or background leaf, depending on which category the majority of events in this leaf belong to.

Note that this algorithm can possibly use a single variable more than once and others not at all, depending on the relative separation power and correlations between the input variables. The sequence of cuts defined by the path from the root node to each leaf node within the tree defines a hypercube in the eleven-dimensional space spanned by the τ-identification variables used as input. Therein also lies the large advantage of decision tree classifiers to simple rectangular cuts. A set of rectangular cuts defines only one hypercube in this space for the signal hypothesis and one for the background hypothesis. The decision tree on the other hand defines as many hypercubes as it has leafs and can therefore in theory (and with unlimited number of training events) approximate any volume in the variable space, even several, disjoint volumes. This also means that it can take into account any type of correlation, even non-linear ones. This big advantage of the decision tree algorithm is offset by its instability to statistical fluctuations in the training samples and its tendency for overtraining. To circumvent these problems a procedure called "boosting" [96] is applied to the single simple decision tree obtained above by applying the so-called "AdaBoost" algorithm [97]. The principle of this boosting algorithm is to train a succession of decision trees - a decision forest - using the same input variables and samples as were used for growing the first tree. But for each successive tree signal and background events that were misclassified by the previously grown tree are weighted by a so-called "boost weight" or "boost factor". The weights of all events are then renormalised to ensure that the sum of weights remains constant over all iterations. Any single event is then classified by the majority vote of all trained decision trees, rather than by the decision of a single tree. The strength of this majority vote can be expressed in a number between 0 and 1 and the distribution of these values for the specific decision tree trained for τ

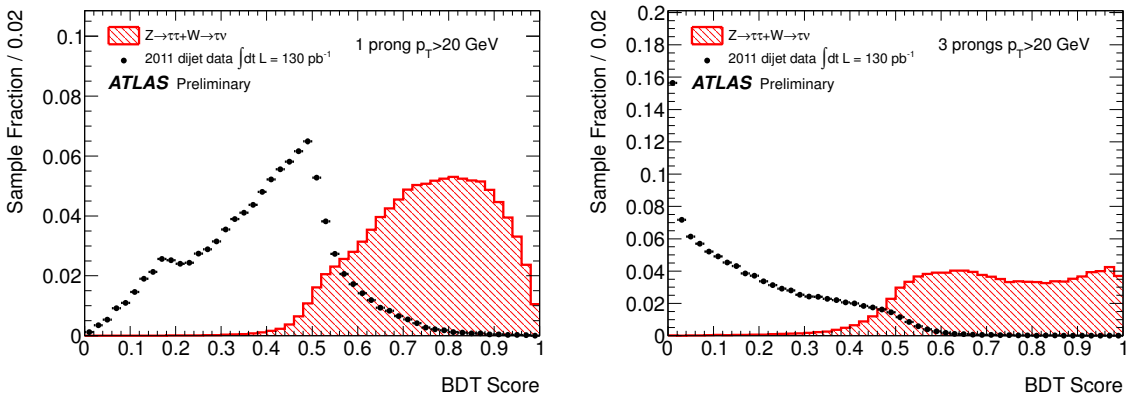


Figure 5.7: Distribution of BDT$_j$ scores for one-prong (left) and three-prong (right) candidates. The distribution for simulated real hadronic τ-lepton decays from $Z \to \tau\tau$ and $W \to \tau\nu$ (red dashed histogram) is compared to the distribution for τ candidates reconstructed from hadronic jets (black markers) selected in 2011 ATLAS data using a tag-and-probe selection [91].

identification described here is shown in Fig. 5.7 for signal and background τ candidates with one or three reconstructed tracks separately. The identification of τ candidates with the BDT$_j$ classifier can then be performed by requiring τ candidates to have a BDT$_j$ score greater than a certain value.

The expected performance of the BDT$_j$ classifier is shown in Fig. 5.8 compared with the simple cut based and projective likelihood approach. It can clearly be seen that the BDT$_j$ classifier gives the best rejection for a given signal efficiency out of these three methods, especially for low-p_{T} τ candidates, which is very important as the cross section for the production of hadronic jets falls strongly with the p_{T} of the jet. Therefore improved suppression of hadronic jet background at low p_{T} is a significant factor in favour of this classifier.

5.2.2 Rejection of Electrons using a Simple Decision Tree

In this algorithm the rejection of τ candidates reconstructed from electrons is performed by applying consecutive selection criteria on the four identification variables: $f_{\mathrm{Had}}^{\mathrm{track}}$, $f_{\mathrm{EM}}^{\mathrm{track}}$, $E_{\mathrm{T,max}}^{\mathrm{strip}}$ and f_{HT}. While a boosted decision tree classifier trained to reject τ candidates from electrons that uses nine different variables also exists, it was found that a simple decision tree using fewer variables is sufficient in the context of the analysis presented in this thesis. Although this simpler algorithm acts differently on τ candidates in the central pseudorapidity region $|\eta| < 1.7$ and the forward region $|\eta| > 1.7$, it starts in both cases with a selection on $f_{\mathrm{Had}}^{\mathrm{track}}$, followed by successive critera on $f_{\mathrm{EM}}^{\mathrm{track}}$, $E_{\mathrm{T,max}}^{\mathrm{strip}}$ and f_{HT}, which depend on whether the previous criteria were fulfilled or not. This selection process also has a tree-like structure (see Fig. 5.9). However, in contrast to the BDT$_j$ classifier no boosting is applied and the decision of the $e^{\pm}$-veto classifier is based on a single decision tree.

The expected performance in terms of real τ lepton efficiency versus inverse electron efficiency of the $e^{\pm}$-veto classifier is shown in Fig. 5.10 in four different regions of pseudorapidity. For comparison, also the expected performance of the boosted decision tree classifier for electrons is shown. While one can clearly see that the boosted variant outperforms the simpler decision tree, it will be shown that the simpler version is more than sufficient to reduce the τ candidate background from electrons in the context of the SUSY analysis.

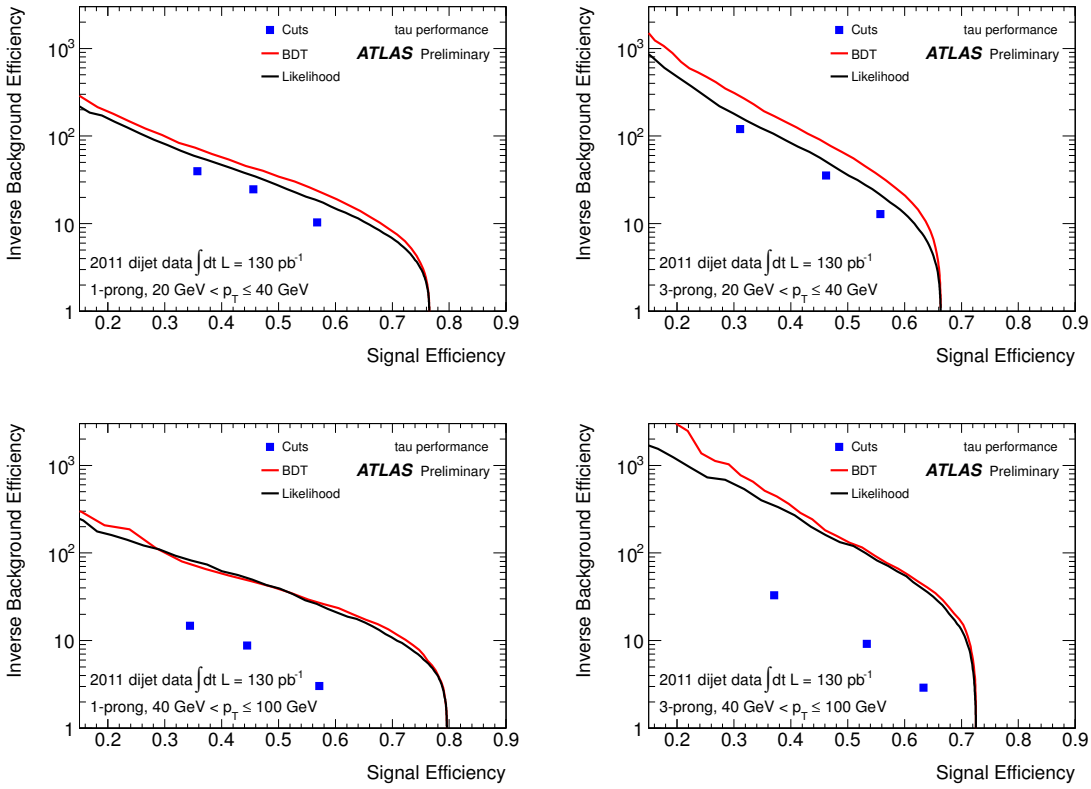


Figure 5.8: Expected performance of all three commonly used τ-identification algorithms (cuts,BDT$_j$ and projective likelihood) used to reject τ candidates reconstructed from hadronic jets. Since the performance of all three methods depends strongly on the p_{T} of the candidate, the expected performance is shown in two ranges of p_{T}: 20 – 40 GeV (top) and 40 – 100 GeV (bottom). Because all three methods are trained separately for candidates with one reconstructed track or more than one reconstructed track, the performance is also shown separately for single-prong (left) and multi-prong (right) candidates [91].

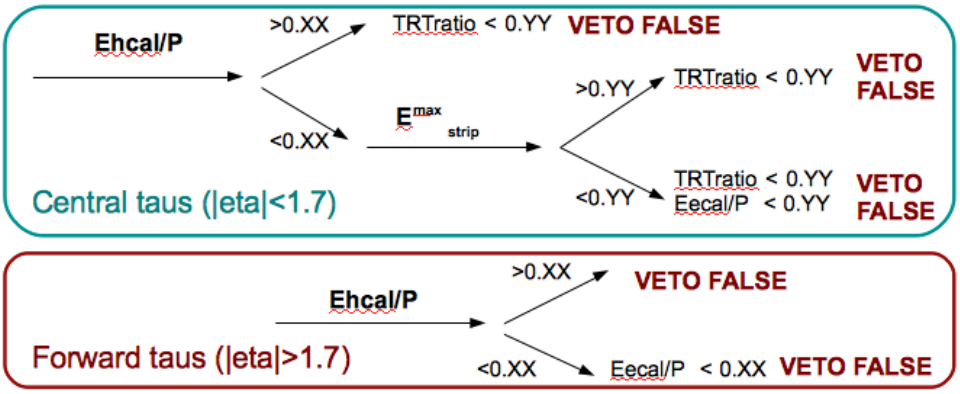


Figure 5.9: Structure of the electron veto algorithm using simple cuts.

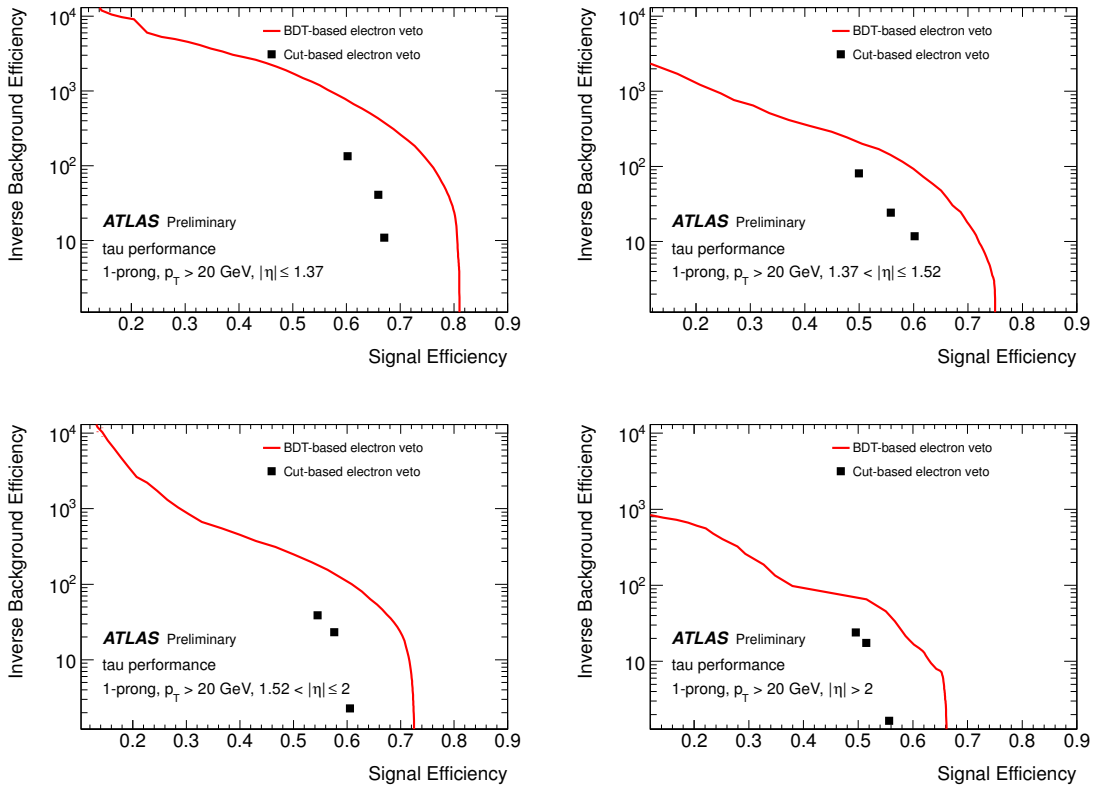


Figure 5.10: Expected performance of both commonly used τ identification algorithms (BDT,$e^{\pm}$-veto) used to reject τ candidates reconstructed from electrons. Since both methods are trained separately in different pseudorapidity regions, the expected performance is shown in four ranges of η [91].

5.3 Tau Identification Efficiency Measurement in Data

The expected efficiencies of the τ-identification algorithms for signal and background have been shown in the previous section. All these efficiencies, except for the background efficiency for hadronic jets, were calculated using truth-level information from simulated samples to classify the reconstructed τ candidates as either real hadronic τ lepton decays or electrons. However the τ-identification algorithms depend on many different variables which are sensitive to almost all aspects of the detector performance, including precision tracking and detailed description of shower shapes in the calorimeter. Mismodelling in the detector simulation will therefore affect the performance of the τ identification in simulation, and it is desirable to verify this performance with data-driven methods. Studies to verify the signal efficiency of the τ identification are described briefly in this section. Because a large part of the work involved in this thesis went into determining the background efficiencies, especially for hadronic jets, a more detailed discussion of these studies is given in Chapter 6.

There are two processes which were looked for at the ATLAS Experiment as "standard candles" for τ-lepton production, namely: $Z \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$. Both decays have been observed at the ATLAS Experiment [98, 99] and cross sections consistent with the Standard Model have been measured. Both processes were also used to derive an estimate of the signal efficiency of the τ-identification algorithms in a "tag-and-probe" (see also Chapter 6) approach. This approach consists of selecting events containing real τ leptons from vector boson decays using a "tag" requirement that is uncorrelated to the identification of the τ-lepton decays in the event. The real hadronic τ-lepton decays in such events can then be used to "probe" the identification efficiency. In a study which utilises $W \rightarrow \tau\nu$ decays, the tag is a large amount of missing transverse momentum due to the neutrino. A detailed description of this study can be found in [100] and an update to current τ-identification algorithms is presented in [91]. While this study is at present still competitive with the one utilising $Z \rightarrow \tau\tau$ decays, it suffers from two significant drawbacks. On the one hand the level of background contamination necessitates a more complicated fitting procedure to extract the signal efficiency for τ identification. On the other hand it is necessary to use a trigger that includes identification requirements on the probe τ candidate used to extract the signal efficiency. Especially the second problem is the reason why this study can be considered at a disadvantage to the efficiency measurement utilising $Z \rightarrow \tau\tau$ decays, detailed in the following section, as it is always preferable not to apply any selection to the probe object in a tag-and-probe study. It should however be noted that both studies give consistent results, suggesting that the trigger bias on the probe candidate does not noticeably bias the results at the current level of precision with which these studies can be performed.

5.3.1 Tau Identification Efficiency Measurement in $Z \rightarrow \tau\tau$ Decays

In this tag-and-probe study the signal process offers two τ-lepton decays, one of which is used as the tag to select the events and the other is used as a probe to determine the efficiency to identify hadronic τ-lepton decays with any given identification algorithm. The tag that provides the highest purity of $Z \rightarrow \tau\tau$ events has been found to be a leptonic τ decay, namely the decay into a highly energetic muon, which can be used to trigger and select the event. The full selection has been adapted from the $Z \rightarrow \tau\tau$ cross section measurement [98] and only the identification requirements on the hadronic τ lepton decay on the probe side have been dropped to obtain a sample of probe τ candidates before identification. The full selection together with the estimation of backgrounds and associated uncertainties is detailed in [91]. A dataset of $800\,\text{pb}^{-1}$ has been analysed with this selection and the invariant mass of the visible decay products of the two τ leptons (i.e. the muon and the hadronic τ decay products) is displayed in Fig. 5.11, before and after identification with the BDT_j classifier.

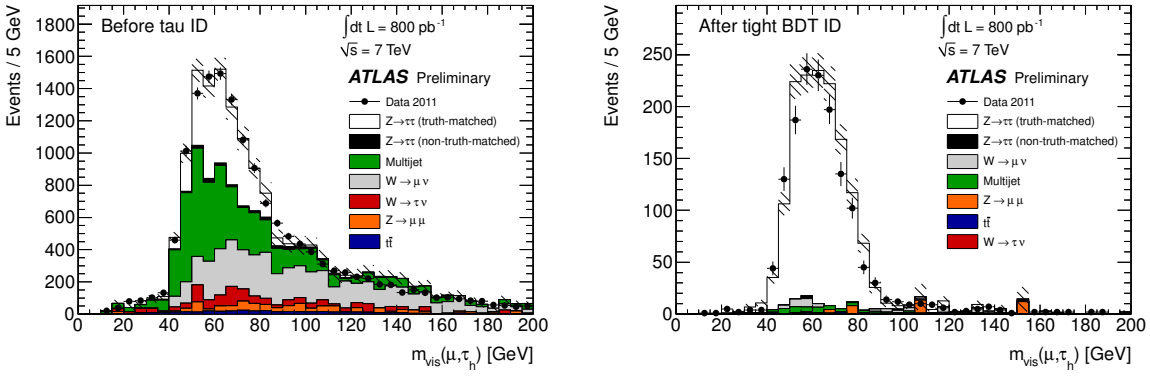


Figure 5.11: Visible mass distribution for events selected with a $Z \to \tau\tau$ tag-and-probe selection before (left) and after (right) applying the BDT$_j$ classifier identification on the probe τ candidate [91].

This yields the signal efficiency of the τ identification after one subtracts off the expected background contributions. The signal efficiency for the τ-identification algorithms can then be calculated by dividing the number of signal candidates in these two distributions (as in equation 5.19). It can clearly be seen that the candidates after identification are dominated by real $Z \to \tau\tau$ decays, but the candidates before identification have a significant contribution from QCD-multi-jet production and W + jets production, where the W decays via $W \to \mu\nu$ and the probe τ candidate is reconstructed from a jet. The uncertainty on the amount of contamination with QCD-multi-jet events is at present the dominant uncertainty in this study, limiting the precision of the measurement to 8 – 12% depending on the τ-identification method applied. As can be expected from the good description of the data distributions in Fig. 5.11 by the simulation, the measured efficiencies in data and simulation agree well within this uncertainty. This is true for all τ-identification algorithms at all levels of "tightness".

Chapter 6

Study of τ-Lepton Misidentification

In any physics measurement involving particle identification there are two aspects of the performance of the identification to consider: the efficiency to identify the particles correctly and the probability to misidentify other particles as the particle in question. The first aspect has been thoroughly studied by others and is described in Chapter 5. The study of the probability to mistakenly identify hadronic jets as hadronic τ-lepton decays represents a large part of the work related to τ identification performed in this thesis and it is also an important aspect of the SUSY analysis presented in Chapter 7. In the case of τ-lepton identification this probability (hereafter denoted as f_{ID}) can be as large as 10%, in contrast to the identification of e.g. electrons or muons where the probability for mistakenly accepting a hadronic jet as a lepton is well below the percent level [101, 102]. Considering the abundance of hadronic jet background at the ATLAS Experiment, the determination of f_{ID} is a crucial ingredient in any analysis involving τ-lepton final states.

In this chapter measurements of f_{ID} with a tag-and-probe selection of di-jet events in data are presented. The idea behind this tag-and-probe method and the event selection are described in Sec. 6.1, together with the simulated samples used for comparisons with the data. The most current, published result [103] using $35\,\mathrm{pb}^{-1}$ of data taken at the ATLAS Experiment in 2010 is summarised in Sec. 6.2. This publication also contains measurements of f_{ID} performed by others in different event topologies and a comparison of the results. The differences in the event selection and results of these studies are discussed in Sec. 6.3, because they illuminate the greatest difficulty in determining a topology-independent parametrisation of f_{ID}, namely the treatment of the origin of the hadronic jets. It will be shown that f_{ID} depends strongly on whether the hadronic jet in question originates from a quark or a gluon. Additionally, because the τ-identification algorithms have changed since this measurement was performed, the same measurement had to be redone with the τ-identification methods used in the SUSY analysis. The measurement of f_{ID} in di-jet topologies is presented in Sec. 6.5. A similar measurement was also performed with a selection of $Z(\rightarrow \ell\ell)$ + jets events. This study is presented in Sec. 6.6. The dataset taken with the ATLAS Experiment in 2011 and analysed in the context of these two studies amounts to $1.04\,\mathrm{fb}^{-1}$. Both the smaller $35\,\mathrm{pb}^{-1}$ dataset and the most current one were taken at a centre-of-mass energy of $\sqrt{s} = 7$ TeV.

6.1 Di-jet Tag-and-Probe Method

6.1.1 Selection and Method

The goal of this method is to measure the misidentification probability f_{ID}, using an unbiased sample of τ candidates in data. This is achieved using a tag-and-probe selection. In this tag-and-probe selection, events are required to exhibit a common signature by which they can be "tagged" and they must contain a "probe" jet that can be used to measure f_{ID}. The requirements on the selected sample can be summarized as follows:

- The sample must be large enough to ensure sufficient statistical precision. This is difficult, since f_{ID} depends on several variables of the τ candidate like p_{T}, η, the number of reconstructed tracks and the amount of pile-up activity (usually represented by the number of reconstructed primary vertices in the event).

- No significant contribution from real τ leptons or other processes that bias the measurement of f_{ID} may be present.

- The events in the sample must share a clear "tag" signature by which they can be selected from collision data.

- The tag signature may not bias the values of f_{ID} measured for the probe jet.

One process that fulfills all these criteria is QCD-di/multi-jet production. These processes have a very high cross section that allows binning f_{ID} simultaneously in several variables. They have the further advantage of being almost completely free of real τ-lepton decays or isolated light leptons, such that the measurement of the probability to misidentify a hadronic jet as a τ lepton is not biased by contributions from signal or other background signatures. In addition the kinematic correlations between the two leading jets in such events gives us a handle on which to select a specific sample of so-called "probe jets", using the jet on the other side of the di-jet pair as a tag. This correlation is clearest in di-jet events that contain exactly two quark- or gluon-initiated jets that are exactly balanced in p_{T} and ϕ. To select such a sample of events, the following selection requirements are applied:

1. Jets are required to have $|\eta| < 2.5$ and $p_{\mathrm{T}} > 15$ GeV.

2. Any pair of jets which is balanced in ϕ ($\Delta\phi \geq \pi - 0.3$ radians) and p_{T} ($|\Delta p_{\mathrm{T}}| \leq p_{\mathrm{T}}^{\mathrm{max}}/2$), where $p_{\mathrm{T}}^{\mathrm{max}}$ is the p_{T} of the leading jet, is considered a valid "tag-and-probe pair".

3. Out of these two jets, one is chosen randomly as the tag jet and required to be matched to the trigger that selected the event.

4. In order to remove the very small contribution from events with real τ-lepton pairs like $Z \rightarrow \tau\tau$, which would pass this selection as well, it is required in addition that the tag jet has at least four tracks associated with it.

5. The other jet is then picked as probe jet and subjected to the τ-reconstruction and -identification algorithms. No further requirements are imposed on the probe jet.

6. After a tag-and-probe jet pair that fulfills all requirements above is found, a reconstructed τ candidate with at least one track associated with it and with $p_{\mathrm{T}} > 15$ GeV is searched for within $\Delta R = 0.2$ of the probe jet as the last stage of the selection.

For the sample of probe-τ candidates obtained with this selection the misidentification probability f_{ID} can be calculated as

$$f_{\mathrm{ID}} = \frac{\text{Number of probe jets identified as } \tau \text{ leptons}}{\text{Number of probe jets reconstructed as } \tau \text{ candidates}}. \tag{6.1}$$

6.1.2 Simulated Samples

Where appropriate the measurements performed with data have been compared with simulated di-jet events from QCD processes generated with PYTHIA [46] with the ATLAS AMBT1 underlying event tune [56]. The events were processed with the ATLAS simulation framework [52] based on GEANT4 [51]. For the $Z(\rightarrow \ell\ell)$ + jets measurement, discussed in Sec. 6.6, the data are compared with simulated events, generated with `ALPGEN` [57]. Since the amount of pile-up interactions in data was not known before the production of the simulated samples, events were generated with a distribution in number of additional interactions from pile-up which only roughly matched the expected pile-up conditions in the data. To further improve the Monte Carlo description of the average number of interactions per bunch crossing the simulated samples were separately reweighted to match the distribution of the number of reconstructed primary vertices in the data.

6.2 Analysis on 35 pb^{-1}

6.2.1 Dataset and Trigger Selection

For all data used in this analysis certain data quality requirements [104] were applied, requiring that all detector components needed for reconstructing τ candidates were operational and events were selected using a calorimeter jet trigger. Because the data sample recorded in 2010, corresponding to an integrated luminosity of 35 pb^{-1}, was rather small, a combination of several jet triggers with different thresholds in transverse momentum of the calorimeter jet were required. In the beginning of data-taking, only the L1 trigger system was operational and the calorimeter jets were selected with thresholds of 5, 10, 15, 30, 55 or 75 GeV of transverse momentum, measured at the electromagnetic scale. In the later stages of data-taking in 2010, the software-based trigger system had been commissioned and the L2 and Event Filter trigger systems were used to select events with p_{T}-thresholds of 20, 30, 35, 50, 75 or 95 GeV at the calibrated energy scale appropriate for hadronic jets. Due to the non-compensating nature of the hadronic calorimeter, the thresholds on the calibrated scale are higher, but indeed each of the six high-level triggers corresponds directly to one L1 trigger, which provides the RoI seed for the software-based triggers. The gain in the number of selected data events from the use of a combination of triggers stems from the fact that for each jet trigger only a randomly chosen fraction of events selected by the trigger are recorded at the ATLAS Experiment. The inverse of this fraction is called the "prescale factor". Therefore it is possible that an event that passed the jet trigger threshold of 50 GeV is not contained in the sample of events selected by the jet trigger with a threshold of 20 GeV, as one would naively expect.

6.2.2 Trigger Dependence

While the tag-and-probe selection described above is straightforward to apply on the level of reconstructed jets and τ candidates, the combination of jet triggers described in 6.2.1 has an effect on the selected probe jets. Over the course of 2010 data-taking the prescales for the different triggers were gradually increased, slowly phasing out the low threshold trigger for the benefit of recording more integrated luminosity with higher threshold triggers. This process continued until at the end of data-taking in 2010 only less than one event in ten thousand selected by the 20 GeV EF trigger was recorded. This trigger is from here on denoted as the 10 GeV L1 trigger, because the largest fraction of the data selected with this trigger was actually taken before the EF trigger system was operational. In contrast to this the average fraction of events passing the 95 GeV EF trigger threshold which were recorded was one in five. To

illustrate the effect of the trigger on the measurements, a comparison was made between the data taken with these two triggers.

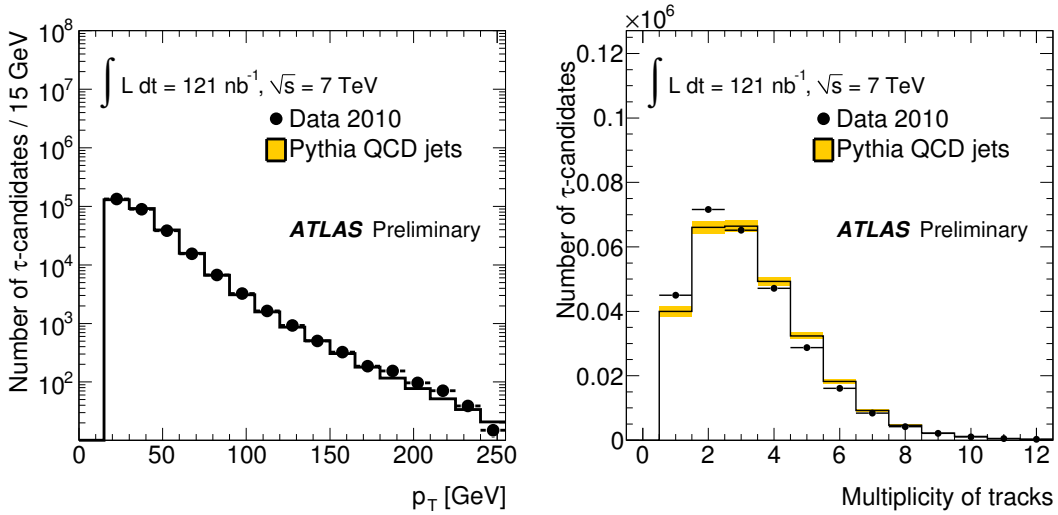


Figure 6.1: The p_{T} distribution (left-hand side) and the number of tracks (right-hand side) of the τ candidates associated with the probe jets for the L1 10 GeV trigger threshold [103]. The integrated luminosity quoted is the integrated luminosity recorded in the early data-taking period multiplied by the prescale factor for this trigger. The solid markers are the data, the open histogram shows the prediction from Monte Carlo and the statistical uncertainty on this prediction is represented by the shaded area.

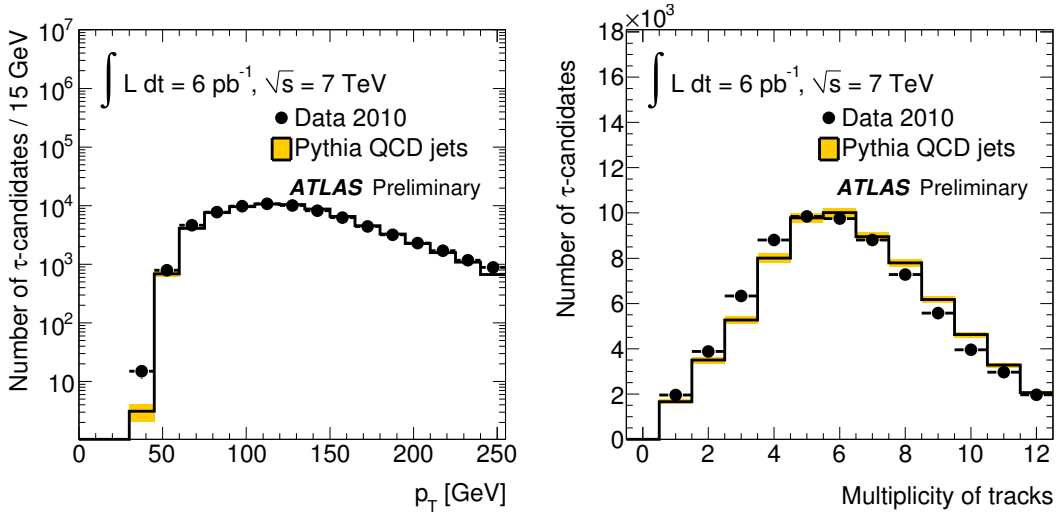


Figure 6.2: The p_{T} distribution (left-hand side) and the number of tracks (right-hand side) of the τ candidates associated with the probe jets for the EF trigger with a 95 GeV threshold [103]. The integrated luminosity quoted is the luminosity recorded in the late 2010 data-taking period multiplied by the prescale factor for this trigger. The solid markers are the data, the open histogram shows the prediction from Monte Carlo and the statistical uncertainty on this prediction is represented by the shaded area.

It was found that, even though the trigger that selected the event is required to be matched to the tag jet, the clear correlation between the two jets in di-jet-like events causes the p_{T} of the probe jet to be close to the p_{T} of the tag jet. The distribution of p_{T} of the probe-τ candidates and the distribution

of their number of reconstructed tracks are shown in Fig. 6.1 for events selected with the 10 GeV L1 trigger threshold and in Fig. 6.2 for events selected with the 95 GeV EF trigger threshold. The effective integrated luminosities quoted include the prescale factors for the different triggers. One can clearly see the difference in the p_{T} of the probe-τ candidate. Higher-p_{T} probe jets result in a higher charged particle multiplicity, which is evident when comparing the number of reconstructed tracks of the τ candidate between the two trigger conditions. Good agreement between data and simulation is seen for the p_{T} distributions in both cases, while the number of reconstructed tracks distributions shows a shift to higher values in simulation.

Since the τ-identification algorithms depend not only on the number of reconstructed tracks and the p_{T} of the candidate, some input variables and the output of the BDT$_j$ classifier are also shown in Fig. 6.3 for candidates selected with the 10 GeV trigger threshold and can be compared to the ones selected with the 95 GeV trigger threshold in Fig. 6.4. While the distributions in data and simulation agree reasonably well, clear differences can again be seen for almost all variables when comparing between the two trigger selections. The reason for this is apparent, as most of the variables e.g. R_{EM} or R_{track} describe the width of the hadronic jet being reconstructed as a τ candidate. The width of a hadronic jet scales roughly with $\frac{1}{\sqrt{E_{jet}}}$ and it is therefore understandable that most identification variables of the τ candidate follow this trend.

The measured misidentification probability f_{ID} is displayed in Fig. 6.5 as a function of p_{T} for both data and Monte Carlo simulation, for the cut-based τ-identification algorithm for the tighter working point as defined in [105]. On the left hand side of Fig. 6.5 one can see that f_{ID} differs significantly between the lowest- and highest-threshold trigger, even though the agreement between data and simulation is good for each trigger threshold separately. This disagreement extends over the full range of p_{T}, meaning that it cannot be caused by threshold effects of the 10 GeV trigger, which can be said to be unbiased starting at an offline p_{T} of $20-30$ GeV. The observed behaviour is readily understood by two effects introduced by the higher threshold trigger. One is the already described p_{T} balancing in di-jet events, which causes the p_{T} distribution of the probe-τ candidates selected with the higher threshold trigger to be different within each p_{T}-bin compared to the unbiased p_{T}-spectrum of the 10 GeV trigger. Since the most important identification variables scale with $E_{jet}^{-1/2}$ there is still a dependence of f_{ID} on the shape of the p_{T}-spectrum within each bin. The other, more significant effect is the trigger bias on the origin of the hadronic jet being reconstructed as a τ candidate. For the lower threshold trigger the relative abundance of quarks and gluons in the hard scattering that produce hadronic jets in the tag-and-probe jet pair is not biased above $20-30$ GeV. For the higher threshold trigger the tag jet needs to have a p_{T} of at least 95 GeV, which means that in the region where the largest discrepancy can be seen in Fig. 6.5, around $50-100$ GeV, the probe jet needs to be significantly lower in p_{T} than the tag jet. This happens most often when the probe jet is mismeasured or split by a hard gluon emission. In both of these cases the probability for the probe jet to have been initiated by a hard gluon is increased. Gluon-initiated jets tend to have a higher particle multiplicity and more spread-out showers in the detector. Since the τ-identification algorithms primarily select τ candidates on the basis of the particle multiplicity and the spread of signals in the calorimeter and tracking detectors, the misidentification probability for gluon jets is much lower than for quark jets. This effect can be observed over the entire range of p_{T} below the higher trigger threshold. On the right hand side of Fig. 6.5 one can see a comparison of the measured f_{ID} with the combination of all triggers during early and late data-taking periods of 2010, where for the early period the L1 10 GeV trigger was the statistically dominant one, and for the late data-taking period the EF 95 GeV trigger was statistically dominant. It can be seen that the effect of a trigger bias is reduced in the combination of all available triggers to less than a 10% relative effect. As will be shown in the following sections this is well below the level of uncertainty reached in this study.

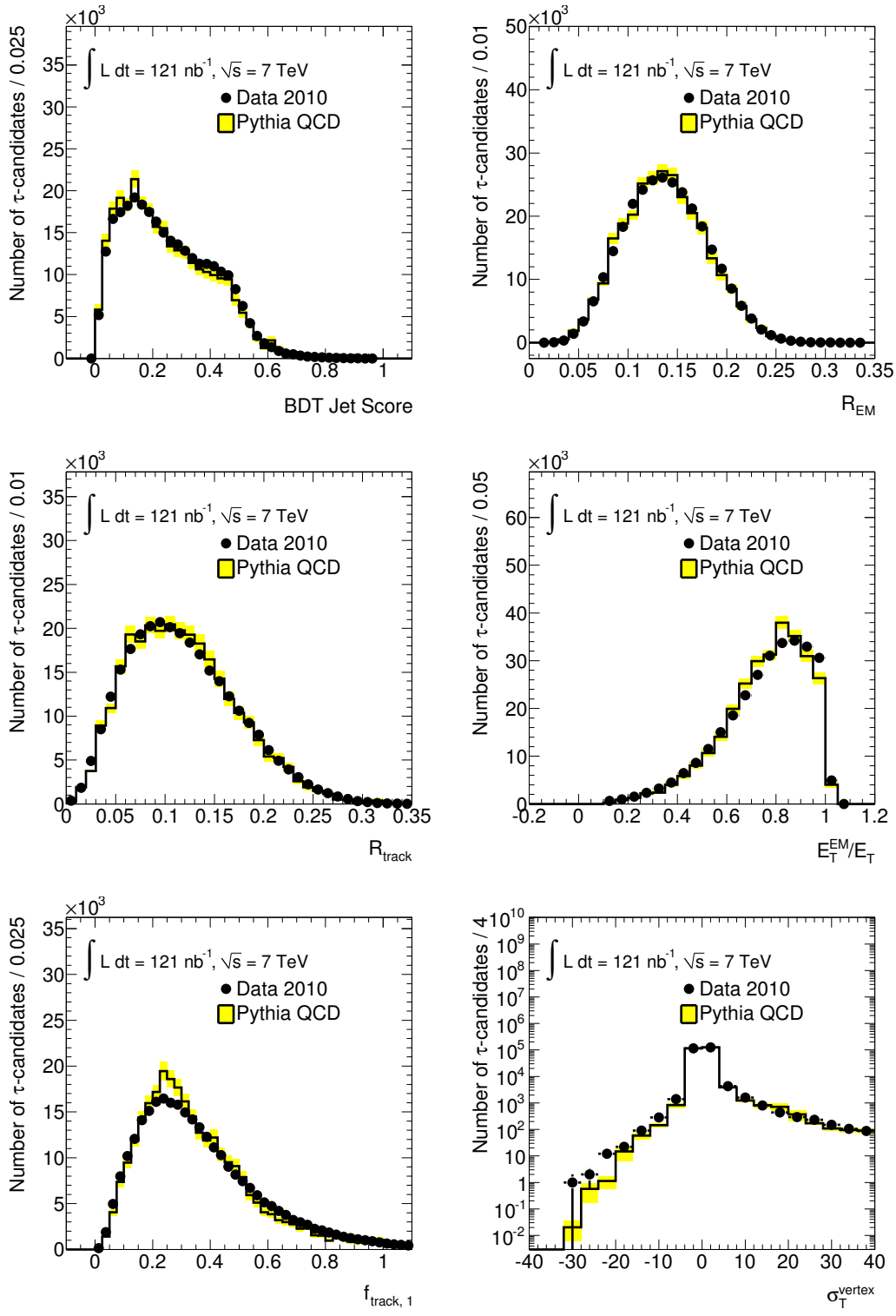


Figure 6.3: For early 2010 ATLAS data selected with a L1 jet trigger threshold of 10 GeV: Identification variables for the reconstructed τ candidates matched to the probe jet: BDT jet score (a), electromagnetic radius R_{EM} (b), track average distance R_{track} (c), longitudinal fraction of the electromagnetic energy $E_{\mathrm{T}}^{\mathrm{EM}}/E_{\mathrm{T}}$ (d), fraction of the leading track p_{T} to E_{T}, $f_{\mathrm{track,1}}$ (e) and flight path significance of the tracks $\sigma_{\mathrm{T}}^{\mathrm{vertex}}$ (f). The integrated luminosity quoted is the integrated luminosity recorded in the early 2010 data-taking period multiplied by the prescale factor for this trigger.

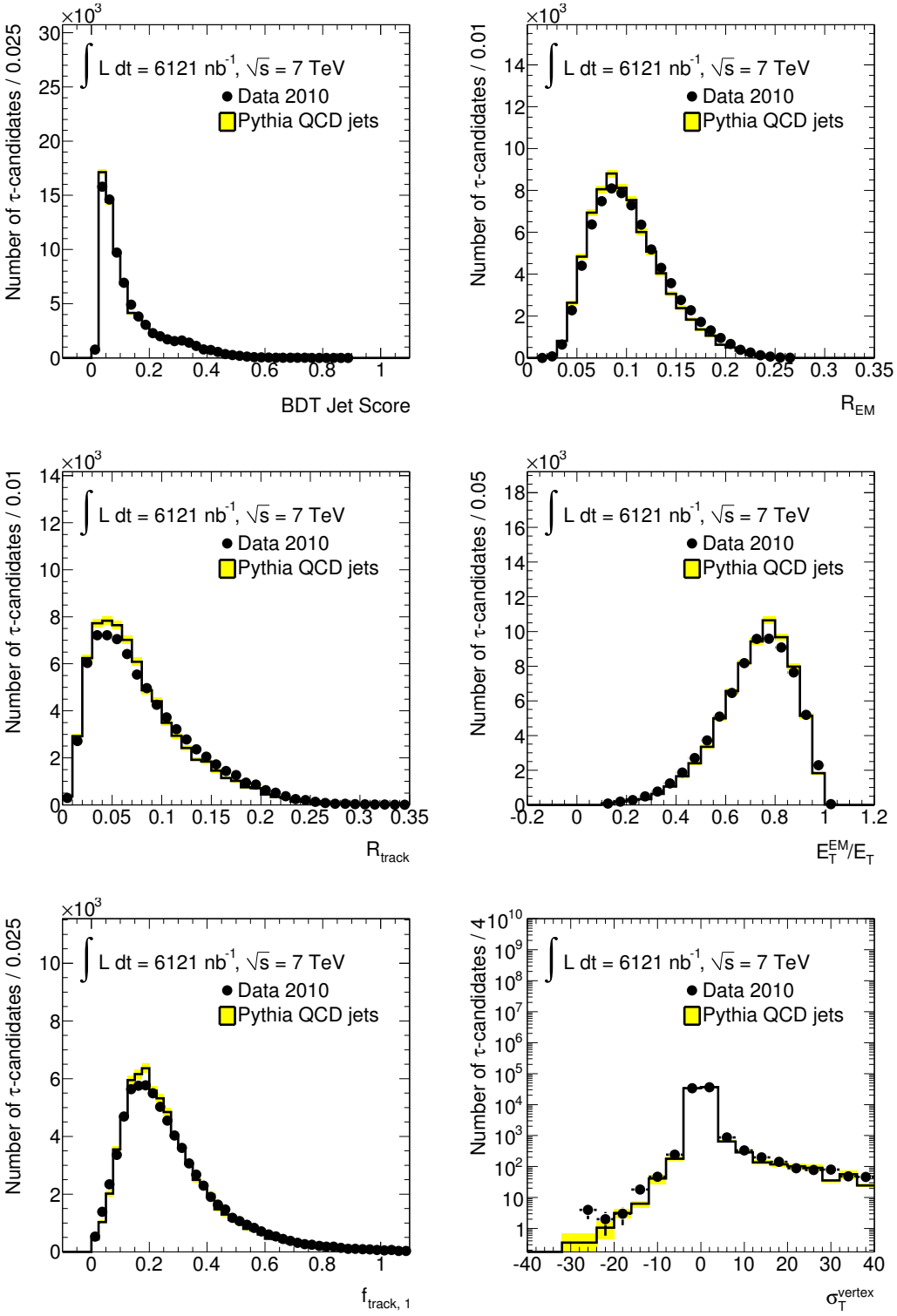


Figure 6.4: For late 2010 ATLAS data selected with a EF jet trigger threshold of 95 GeV: Identification variables for the reconstructed τ candidates matched to the probe jet: BDT jet score (a), electromagnetic radius R_{EM} (b), track average distance R_{track} (c), longitudinal fraction of the electromagnetic energy $E_{\mathrm{T}}^{\mathrm{EM}}/E_{\mathrm{T}}$ (d), fraction of the leading track p_{T} to E_{T}, $f_{\mathrm{track,1}}$ (e) and flight path significance of the tracks $\sigma_{\mathrm{T}}^{\mathrm{vertex}}$ (f). The integrated luminosity quoted is the integrated luminosity recorded in the late 2010 data-taking period multiplied by the prescale factor for this trigger.

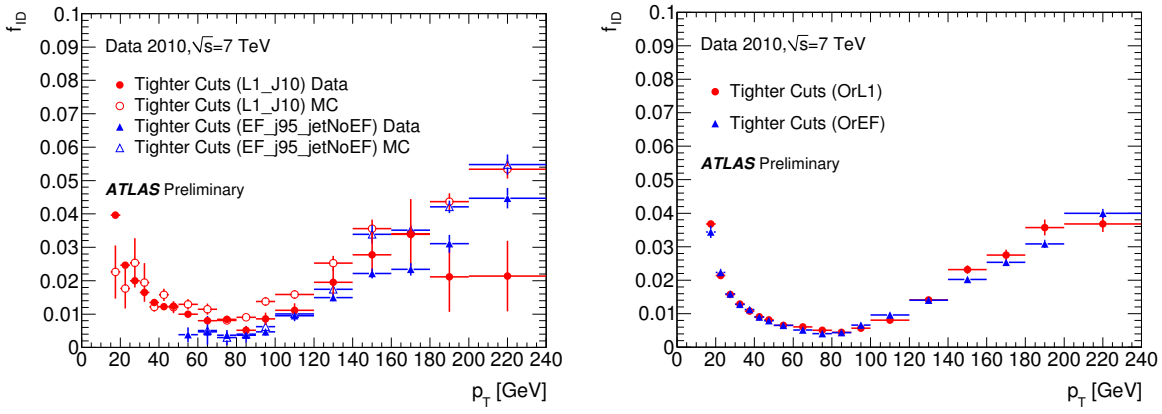


Figure 6.5: Comparison of the misidentification probability, as a function of τ candidate p_{T}, for data and simulation, for the tighter cut-based working point [103]. On the left-hand side plot the solid and open round markers correspond to the lowest threshold (10 GeV) L1 trigger in data and Monte Carlo, respectively. The solid and open triangles correspond to the highest threshold (95 GeV) EF trigger in data and simulation, respectively. The result of combining either all L1 triggers or all EF triggers in data is shown on the right-hand side. The uncertainties shown are statistical only.

6.2.3 Results

The misidentification probability measured with the di-jet event selection in the 2010 dataset is shown in Fig. 6.6 as a function of p_{T} for all three commonly used identification algorithms. Since most physics analyses involving τ leptons require the τ candidates to have either one or three reconstructed tracks associated with them, f_{ID} is only shown for these two cases, for which it differs significantly. The distribution of f_{ID} is also split between events with two or less reconstructed vertices and events with three or more reconstructed vertices, because of the dependence of f_{ID} on the amount of activity around the jet, which comes from additional low-energy collisions. One can clearly see this dependence in f_{ID} when comparing the left and right column in Fig. 6.6. Additionally the dependence of f_{ID} on the pseudorapidity η of the τ candidate has been studied. It was found that there is no strong dependence on η. Hence it is justified not to bin f_{ID} in η to preserve a higher number of candidates in each bin of p_{T} and track multiplicity.

The uncertainties shown as shaded areas around the measured values denote the combination of statistical uncertainty of the data and the systematic uncertainties evaluated as described in Sec. 6.2.4. Overall the misidentification probability was found to be within $0.1 - 10\%$, depending on the p_{T} and number of reconstructed tracks of the τ candidate and the number of reconstructed primary vertices in the event.

6.2.4 Systematic Uncertainties

The following possible sources of systematic error on f_{ID} were studied and found to be negligible:

- The effect of varying the matching criterion of the probe jets to the reconstructed τ candidates (default: $\Delta R < 0.2$) was studied.
- The contamination of the probe jet sample with real τ leptons was investigated.

The number of additional pile-up interactions in the event, i.e. the number of primary vertices, was found to affect f_{ID} significantly as shown in Fig. 6.7. A higher number of reconstructed vertices leads to

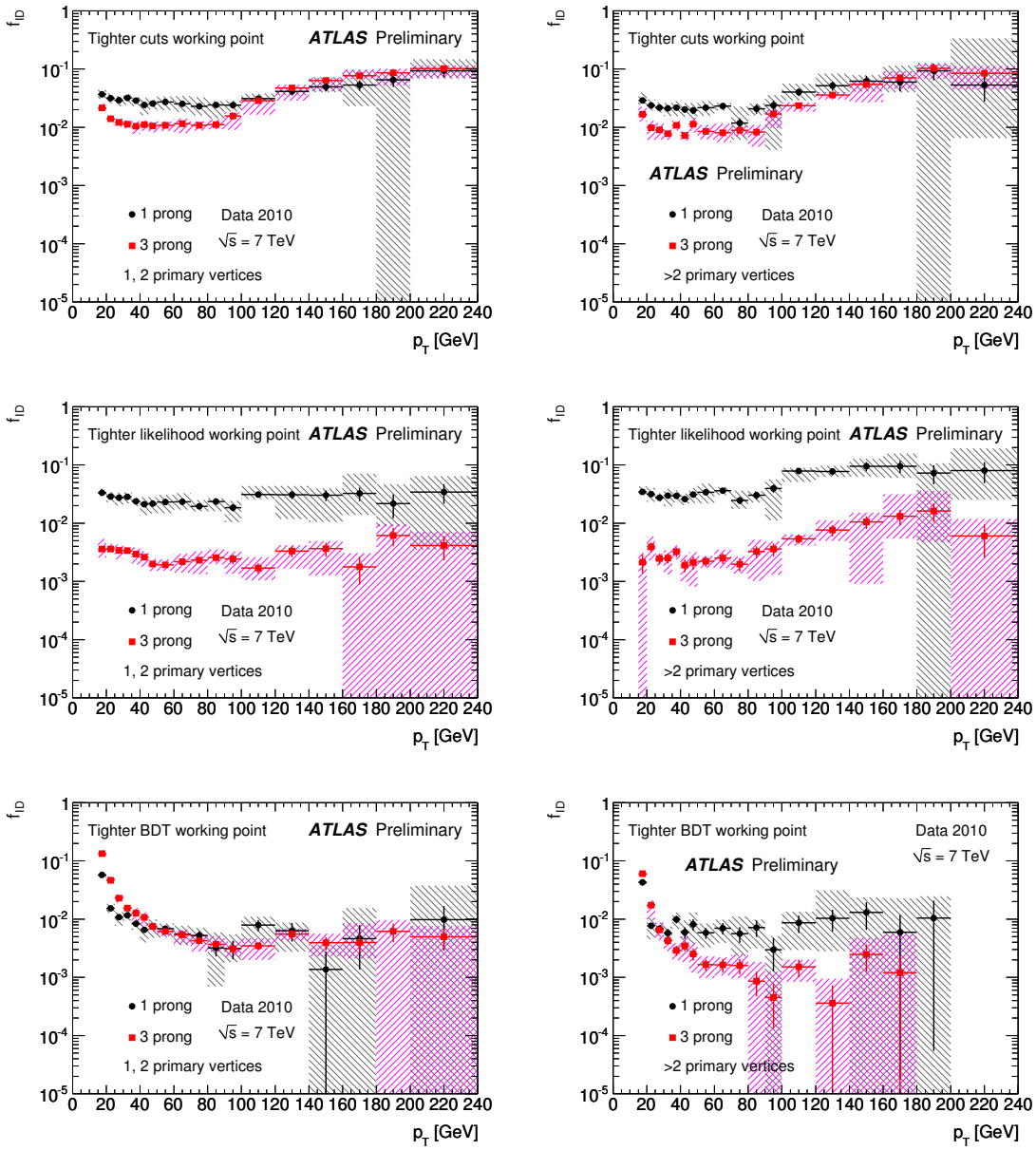


Figure 6.6: Misidentification probabilities f_{ID} for tau identification measured at the tigher working point in the full 2010 ATLAS dataset of 35pb^{-1} for the cut-based (a and b), projective likelihood (c and d) and BDT$_j$ (e and f) classifiers as a function of p_{T} [103]. Because of the dependence of f_{ID} on the amount of pile-up in the event, f_{ID} is shown for low numbers of reconstructed vertices on the left and higher numbers of vertices on the right.

a reduced misidentification probability, since fewer τ candidates fulfill the necessary isolation criteria, which are implicit in the τ-identification algorithms. Instead of assigning a systematic uncertainty due to this important effect, it was decided to measure f_{ID} in bins of the number of reconstructed primary vertices. Additionally the possible systematic variation of f_{ID} due to the selection criteria for the di-jet event selection were investigated. The list of selection criteria that were found to have a noticeable effect on f_{ID} is given below, together with a description of how the uncertainty due to each selection criterion

was estimated.

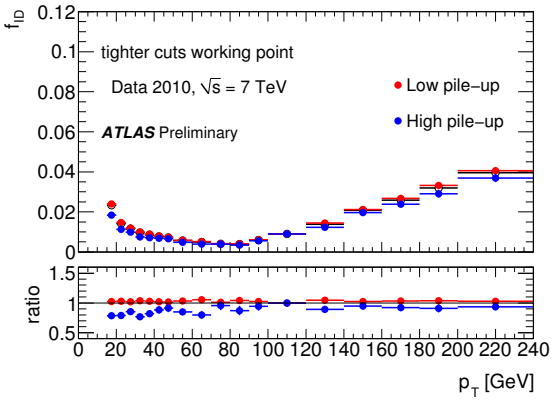


Figure 6.7: Distribution of f_{ID} for the full 2010 ATLAS dataset of 35pb^{-1} separated into events with high and low number of reconstructed primary vertices to illustrate the effect of the presence of pile-up on the probability to misidentify hadronic jets as τ leptons [103]. The red markers denote f_{ID} for events with one or two reconstructed primary vertices. The blue markers show f_{ID} for events with more than two reconstructed primary vertices.

1. The choice of the value used in the p_{T}- and ϕ-balancing criterion of the tag and the probe jet can have an influence on the misidentification rate. There are two main reasons for this: differences in the response between the two jets in the tag-and-probe pair on the one hand and the presence of additional jets in the event on the other hand. These are correlated both with the origin of the jet, i.e. quark or gluon, and the isolation of the probe jet. The event sample is therefore split into sub-samples satisfying $(\pi - 0.3) < \Delta\phi < (\pi - 0.1)$ and $\Delta\phi \geq (\pi - 0.1)$, and $|\Delta p_{\mathrm{T}}| < 0.44 \cdot p_{\mathrm{T}}^{\mathrm{max}}/2$ and $0.44 \cdot p_{\mathrm{T}}^{\mathrm{max}}/2 < |\Delta p_{\mathrm{T}}| < p_{\mathrm{T}}^{\mathrm{max}}/2$. The separation thresholds for these samples was chosen such that the two sub-samples are roughly of equal size. For both the p_{T}- and ϕ-balancing criterion the deviation of f_{ID} between the sub-samples and the full sample was taken as a separate systematic uncertainty evaluated in each p_{T}- and track-multiplicity bin.

2. The number of tracks in the tag jet can influence the probe jet. Since the average number of charged particles within a jet is higher for gluon-initiated jets, requiring a large number of tracks within the tag jet tends to select gluon-initiated jets on the tag side. This affects the relative abundance of quark- and gluon-initiated jets on the probe side, since di-quark final states are suppressed. To estimate this effect, the event sample is split into one where the tag jet has more than seven tracks, and one with less than seven tracks, leading to two samples of roughly the same size. The misidentification probability is determined separately for the two samples and the observed difference to the mean of both - the default value measured on the full sample - is included in the systematic uncertainty.

6.3 Measurement of f_{ID} in different Event Topologies

In addition to the study described in Sec. 6.2 the misidentification probability for hadronic jets was also measured in two further studies, based on different event topologies. Since these studies were not directly performed as part of this thesis, they are only briefly discussed here and a comparison of the measured values of f_{ID} with the ones measured with the di-jet selection on simulated samples is

presented together with an interpretation of the observed differences. The important difference in the probe-jet selection in all three studies is the origin of the probe jets. Depending on the process in which the probe jet is produced the relative abundance of quark- and gluon-initiated jets varies significantly. As discussed in the previous sections, this has a large effect on the measured f_{ID}, since the τ-identification algorithms are trained to select narrow, well-collimated jets. Because quark jets tend to have a lower particle multiplicity and a lower propability for large angle-emissions of gluons, the measured f_{ID} is expected to be higher if the sample of jets is quark dominated (see Fig. 6.8).

The two studies to be compared with the di-jet selection are:

1. $Z(\rightarrow \ell\ell)$ + jets topology: In this study a Z boson decaying into two same-flavour, light leptons (either electrons or muons) is used as a tag for the event. Because of the very clear signal of two isolated, well-identified electrons or muons with an invariant mass compatible with the Z-boson mass, one can select a sample of these events with a very high purity. Any additional hadronic jets in these events can then be used to probe f_{ID}. The expected fraction of quark-initiated jets in the probe jet sample is significantly higher compared to the sample of probe jets from di-jet events (see Fig. 6.9) and therefore the misidentification probability measured in this topology is also higher. The only drawback of this selection is the low cross section of Z-boson production compared with QCD-multi-jet production. This study therefore suffered from very low statistical precision in 2010 ATLAS data, limiting the precision of the comparison of f_{ID} with the other channels.

2. γ+jet topology: Using the same tag-and-probe strategy as in the di-jet event selection, one can exchange the tag jet with an isolated photon that is balanced with the the probe jet in p_{T} and ϕ. In this selection the fraction of quark-initiated jets is the highest (well above 90%) of all three topologies considered. While the cross section of γ+jet production is much larger than for Z-boson production, it is still several orders of magnitude lower than for QCD-multi-jet production. While this does not prohibit a statistically meaningful comparison with the measurement on the di-jet event sample in the 2010 data, it means that, when selecting a sample of γ+jet events in data, there is still a significant contamination, on the order of $\approx 20\%$, from di-jet events, where one jet is misidentified as a photon. Since no accurate estimate of this contamination in the selected

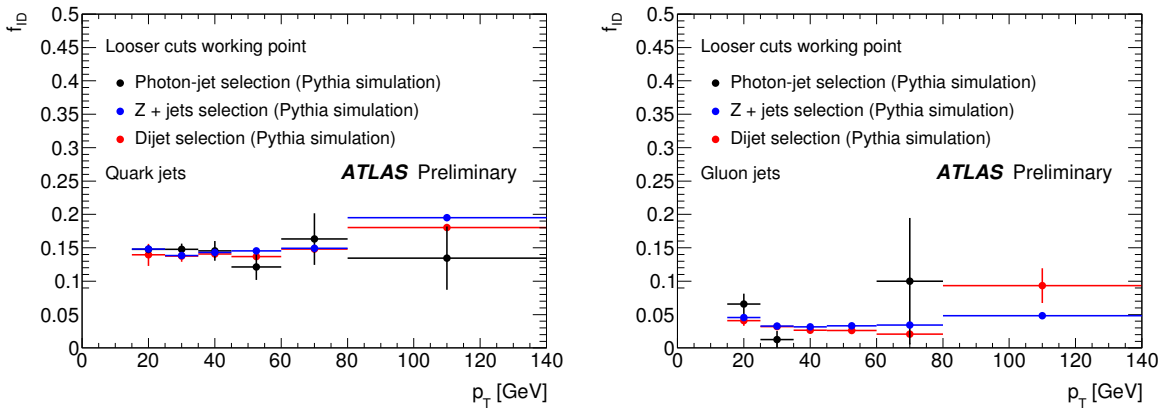


Figure 6.8: Misidentification probability as a function of τ-lepton p_{T} for quark-initiated (left) and gluon-initiated (right) probe jets for the looser cut-based working point, as determined from Monte Carlo simulation for three different event topologies [103]. The errors shown are statistical only.

data sample was available, a precise estimate of the expected relative abundances of quark- and gluon-initiated jets was not possible for this topology in data.

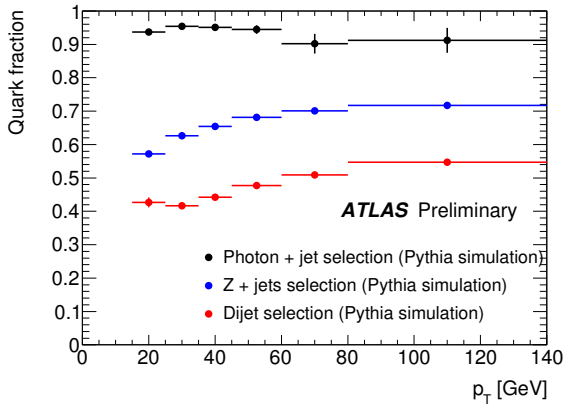


Figure 6.9: Fraction of quark-initiated probe jets as a function of the p_{T} of the probe jet as determined from Monte Carlo simulation for three different event topologies [103]. The errors shown are statistical only.

Given a measurement of f_{ID} and a reliable estimate of the ratio of quark and gluon jets from simulation for all three of these studies, it should be possible to determine f_{ID} as a function of the fraction of quark- or gluon-initiated jets in any given data sample. This is suggested by the good agreement of f_{ID} measured in all three different topologies for simulated samples when selecting pure quark or pure gluon jets as well as by the comparison of the f_{ID} distributions as they were measured in each of the three topologies in 2010 ATLAS data. However due to the reasons stated in the description of the different topologies, it was not possible to conclusively test this proposition with the available results collected from ATLAS data in 2010.

6.4 Update of the Analysis to 1fb^{-1}

As mentioned in the introduction to this chapter, the τ-identification algorithms were re-optimised for 2011 data taking. The most important reasons for this were the increase in instantaneous luminosity at the LHC, leading to higher numbers of pile-up interactions per bunch crossing, and the better understanding of some of the more complicated, but also more powerful, identification variables. Therefore the measurement of f_{ID} presented in Sec. 6.2 can not be applied to estimate the multi-jet background in physics analyses performed on 2011 ATLAS data, like the one presented in Chapter 7. Therefore the measurement of f_{ID} in both the di-jet and the $Z(\rightarrow \ell\ell)$+jets topologies were performed again on the data taken with the ATLAS Experiment between January and July of 2011 corresponding to an integrated luminosity of 1.04 fb^{-1}. The modifications to the analyses performed in 2010 and the results of these measurements are described in the following sections.

6.5 Measuring f_{ID} in 2011 ATLAS Data in Di-Jet Topologies

6.5.1 Dataset and Selection - 2011 Di-Jet Topology

The updated measurement in 2011 could be performed with a single jet trigger with a threshold of $p_{\mathrm{T}} > 30$ GeV, which was sufficient to collect enough events to measure f_{ID} up to a transverse momentum

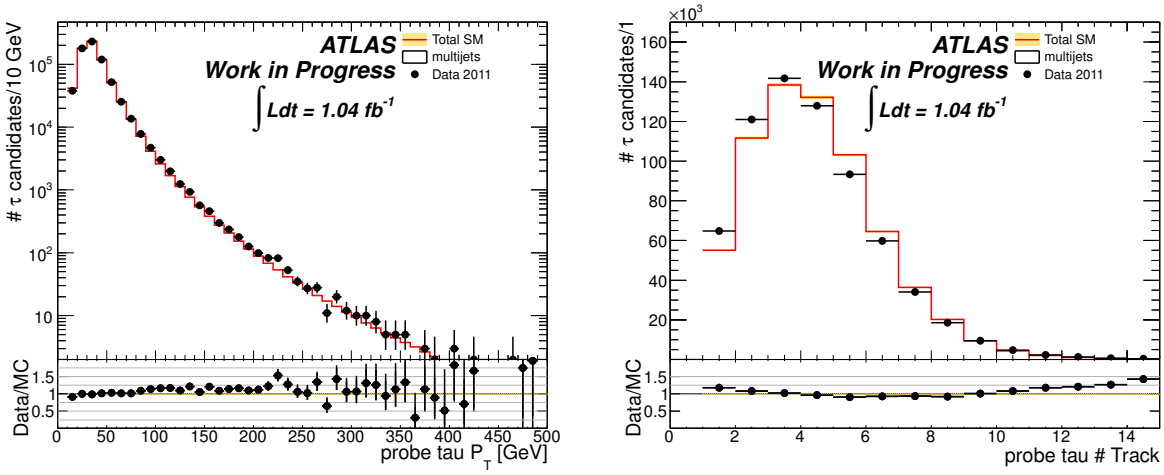


Figure 6.10: The p_{T} distribution (left) and the number of tracks (right) of the τ candidates associated with probe jets for the events selected with the di-jet event selection in the 2011 ATLAS dataset. The integrated luminosity quoted is the integrated luminosity recorded in the early 2011 data-taking period. The solid markers are the data, the open histogram shows the prediction from Monte Carlo simulation and the statistical uncertainty on this prediction is represented by the shaded area.

of about 100 GeV. One other important modification to the tag-and-probe jet pair selection was made. In the 2010 measurement it was necessary to accept any tag-and-probe jet pair that fulfilled the criteria described in Sec. 6.1.1 to increase the available number of probe jets. As a result there were possibly more than one tag-and-probe pair in each event and jets that were not part of clear di-jet but rather of multi-jet topologies entered the probe jet sample. In the 2011 measurement the available amount of data allowed for a more stringent di-jet selection, meaning that only the leading di-jet pair fulfilling all the requirements was considered in each event. As a final refinement to further reject three-jet topologies the p_{T}-balancing criterion was tightened to $|\Delta p_{\mathrm{T}}| < 0.3 \cdot p_{\mathrm{T}}^{\mathrm{max}}$. The p_{T} and number of reconstructed tracks distributions for the selected data samples are shown in Fig. 6.10.

Good agreement between data and simulation can again be observed for the p_{T} spectrum, while the distribution of the number of reconstructed tracks associated with the τ candidates shows the same shift observed in the comparison of 2010 data with simulation. Across most of the p_{T} range from 15 to 100 GeV there are more than ten thousand probe-τ candidates available, allowing for a measurement of f_{ID} in p_{T} bins of 10 GeV. The distributions of the identification variables and the BDT$_j$ classifier output are shown for the 2011 dataset in Fig. 6.11. The distributions of these variables in data are also described reasonably well by the simulation. Deviations are seen in some of the track-based variables, which can be understood from a mismodelling of the reconstructed track multiplicity. The effect of this can also be observed in the discriminant output where it leads to a significant overestimation of the number of τ candidates with BDT$_j$ scores higher than 0.8.

6.5.2 Results - 2011 Di-Jet Topology

The results of measuring f_{ID} in the 2011 data sample are shown in Fig. 6.12 as a function of p_{T} for the BDT$_j$ classifier at the working point found to be most suitable when optimising the event selection described in Chapter 7. For this working point tight selection criteria are applied to one-prong candidates, while medium criteria are applied to three-prong candidates. Again the dependence of f_{ID} on the

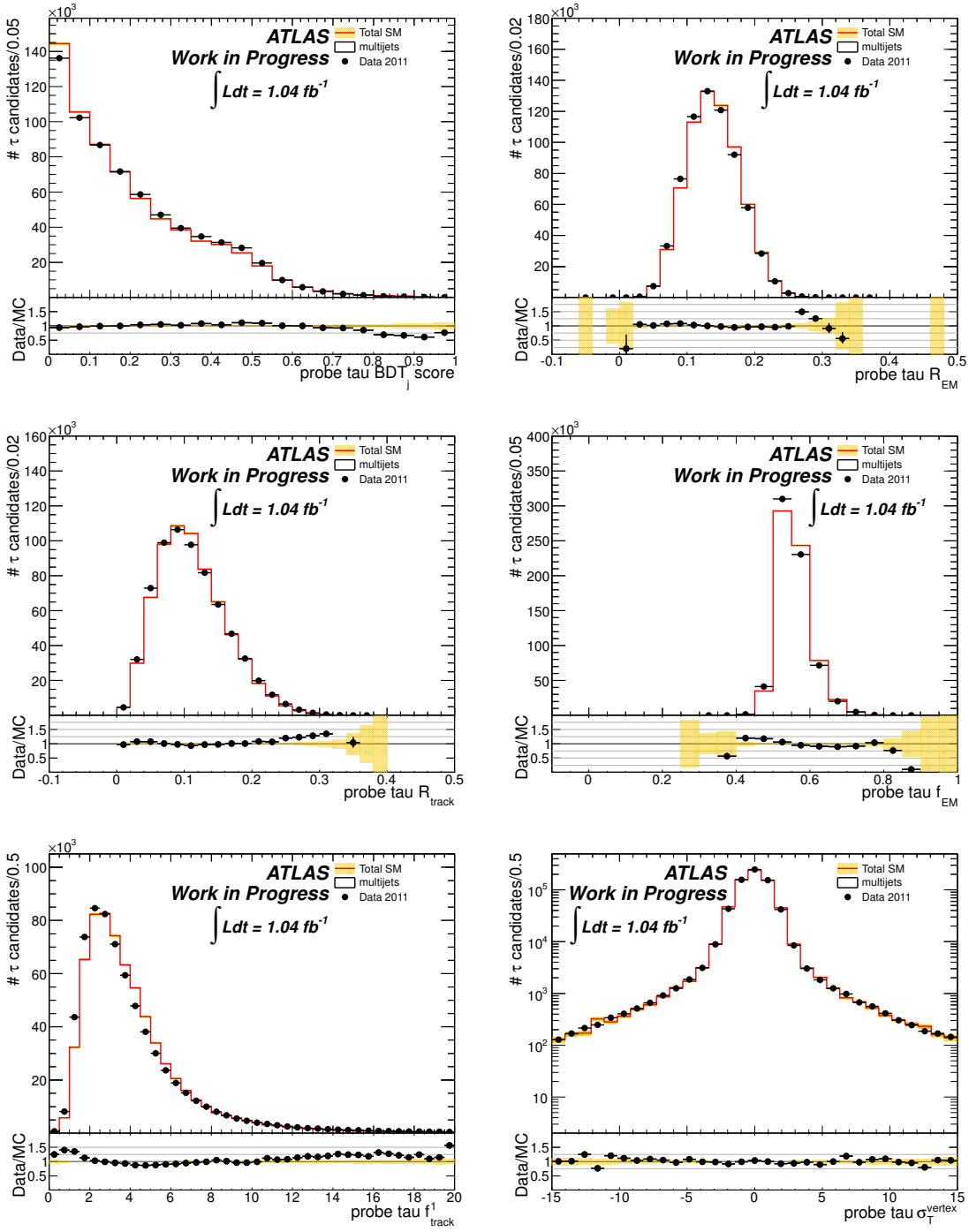


Figure 6.11: Identification variables for the reconstructed τ candidates matched to the probe jets in 2011 ATLAS data amounting to 1.04fb^{-1}: BDT jet score (a), electromagnetic radius R_{EM} (b), average track distance R_{track} (c), longitudinal fraction of the electromagnetic energy $E_{\mathrm{T}}^{\mathrm{EM}}/E_{\mathrm{T}}$ (d), fraction of the leading track p_{T} to E_{T}, $f_{\mathrm{track,1}}$ (e) and flight path significance $\sigma_{\mathrm{T}}^{\mathrm{vertex}}$ (f).

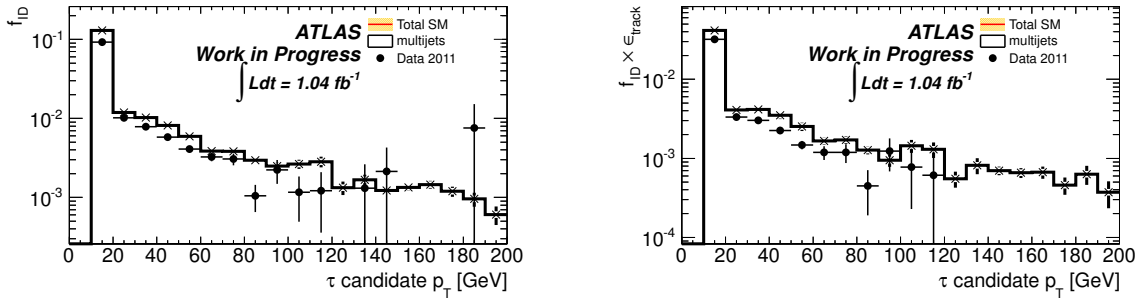


Figure 6.12: Distribution of the misidentification probability f_{ID} measured in a 2011 ATLAS dataset of 1.04fb^{-1} for the BDT$_j$ classifier as a function of p_{T} (left). Because the analysis in Chapter 7 uses the product of f_{ID} with the efficiency for requiring one or three reconstructed tracks, this product is also shown on the right.

variables η, number of reconstructed tracks and number of reconstructed vertices has been studied. As in the measurement for 2010, it was found that f_{ID} does not significantly depend on η. Because of the strong dependence of f_{ID} on the number of reconstructed vertices in the event, the re-optimisation of the τ-identification algorithms was performed in bins of number of reconstructed vertices. Therefore this dependence is much reduced and it is justified to measure f_{ID} integrated over the full range of number of reconstructed vertices in the event.

Fig. 6.12 also shows the product of f_{ID} with the efficiency for requiring the probe-τ candidates to be reconstructed with one or three tracks. This combination will be used in some of the background estimates of the SUSY analysis presented in Chapter 8 and is therefore given here for reference.

6.6 Measuring f_{ID} in 2011 ATLAS Data in $Z(\rightarrow \mu\mu)$ + jets Topologies

As was discussed in Sec. 6.3, f_{ID} depends strongly on the topology of the events, most significantly on the relative abundance of quark-initiated and gluon-initiated jets. As noted in that section, no process-independent measurement of f_{ID} as a function of the fraction of quark- and gluon-initiated jets in a sample is available at present. Therefore the misidentification probability measured in Sec. 6.5 is applicable only for QCD-multi-jet topologies and not for other processes that have a significantly different ratio of the number of quark and gluon jets. Because the analysis presented in Chapter 7 has to deal with a significant background contribution from W- and Z-boson production, it is necessary to measure f_{ID} also for hadronic jets being produced in association with vector bosons. In 2010 such a measurement was performed in the $Z(\rightarrow \ell\ell)$ + jets topology, using both the decays of the Z boson to electrons and to muons to select a sample of events as described in Sec. 6.3. This measurement was performed again in the context of this thesis on the 2011 dataset collected with the ATLAS Experiment. 1.04fb^{-1} of data have been collected and analysed. In contrast to the study described in 6.3 this study was performed only in the muon channel, meaning the decay $Z \rightarrow \mu\mu$ and with different analysis object definitions as described in Chapter 4, which were updated to 2011 ATLAS data analysis prescriptions published by the different performance working groups of the ATLAS collaboration. Especially the definition of the muon objects was significantly altered since the measurement in 2010 was performed.

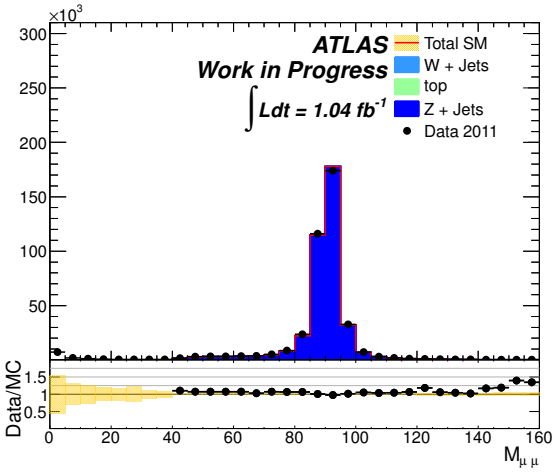


Figure 6.13: Invariant mass distribution of di-muon pairs in 2011 ATLAS data amounting to 1.04fb^{-1}.

6.6.1 Dataset and Selection - 2011 $Z(\rightarrow \mu\mu)$ + jets Topology

After requiring the data to be selected by a single muon trigger with a p_{T} threshold of 18 GeV, exactly two identified muon objects of opposite charge are required. From the sum of the four-momenta of these two objects the invariant mass is calculated and required to lie within the range 75 – 105 GeV, i.e. to be consistent with the Z-boson mass. The distribution of invariant mass values is shown in Fig. 6.13 for the data collected with ATLAS in 2011 and for simulated $Z \rightarrow \mu\mu$ decays normalised to the cross section taken from [59–61]. As can be seen, the agreement between data and simulation is good in the selected mass window, while some discrepancies are only observed far outside of this mass window.

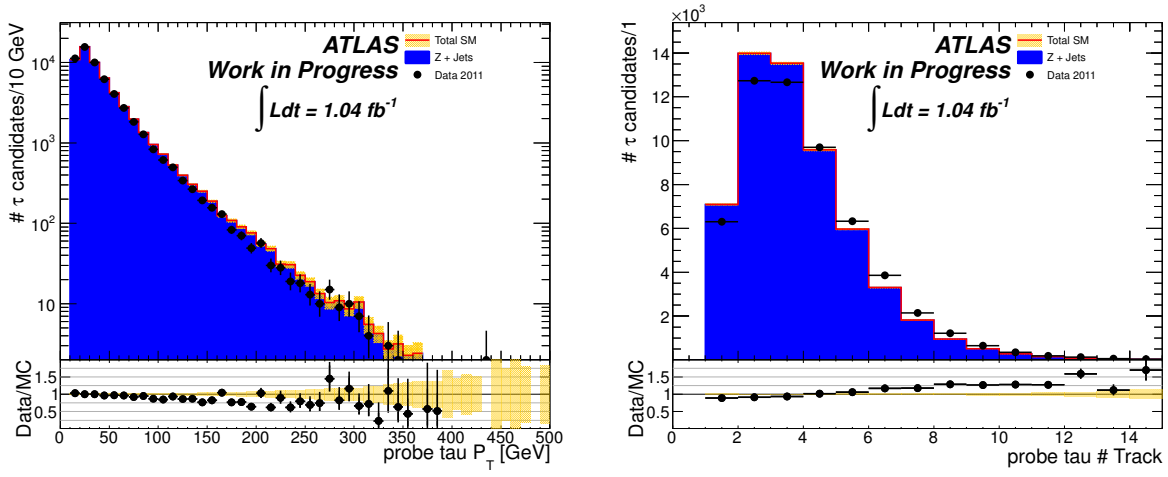


Figure 6.14: The p_{T} distribution (left-hand side) and the number of tracks (right-hand side) of the τ candidates associated to probe jets for the events selected with the $Z(\rightarrow \mu\mu)$+jets event selection in the 2011 ATLAS dataset. The integrated luminosity quoted is the luminosity recorded in the early 2011 data-taking period. The solid markers are the data, the blue histogram shows the prediction from Monte Carlo and the statistical uncertainty on this prediction is represented by the shaded area.

In this data sample all hadronic jets that are reconstructed as τ candidates with $p_{\mathrm{T}} > 15$ GeV are

considered as a viable probe for the τ-identification algorithms.

6.6.2 Results - 2011 $Z(\rightarrow \mu\mu)$ + jets Topology

The p_{T} and number of reconstructed tracks distributions for the selected probe-τ candidates is shown in Fig. 6.14. While the agreement in the p_{T} distribution is good, the track multiplicity is again not well described by the simulation. As was the case for the measurement in the di-jet topology for the 2011 dataset, this disagreement also translates into some disagreements in the distributions of identification variables (see Fig. 6.16), especially for the track-based variables. These discrepancies and the resulting disagreement in the distribution of the BDT$_j$ classifier is even more pronounced than in the measurement in the di-jet topologies. This results in a more significant disagreement between f_{ID} measured in data and simulation for the BDT$_j$ classifier, as shown in Fig. 6.15. This more pronounced disagreement in f_{ID} for the simulated events generated with `ALPGEN` is a well known phenomenon from other ATLAS physics analyses involving τ leptons in 2011 [106], [98] and is thought to be caused by mismodelling of the relative abundance of hard quarks and gluons in the simulation. This large difference in the misidentification probability for hadronic jets being produced in association with vector bosons will also be encountered in the studies of these background presented in Chapter 8.

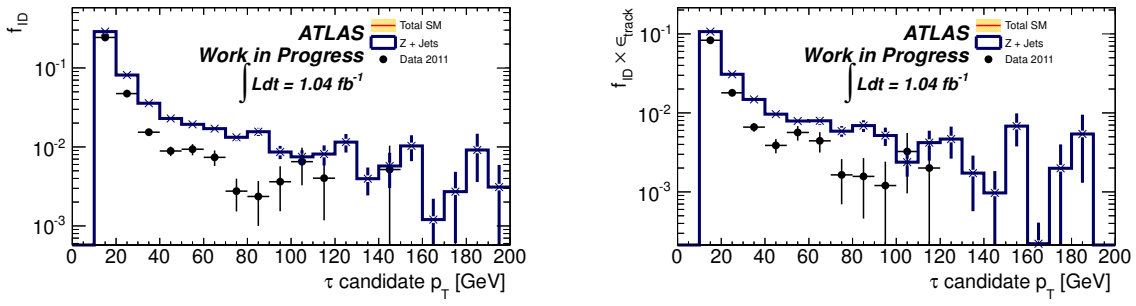


Figure 6.15: Misidentification probability f_{ID} measured in the 2011 ATLAS dataset of 1.04fb^{-1} for the BDT$_j$ classifier as a function of τ-candidate p_{T} (left) measured in the $Z(\rightarrow \mu\mu)$ + jets topology. Because the analysis in Chapter 7 uses the product of f_{ID} with the efficiency for requiring one or three reconstructed tracks, this product is also shown on the right.

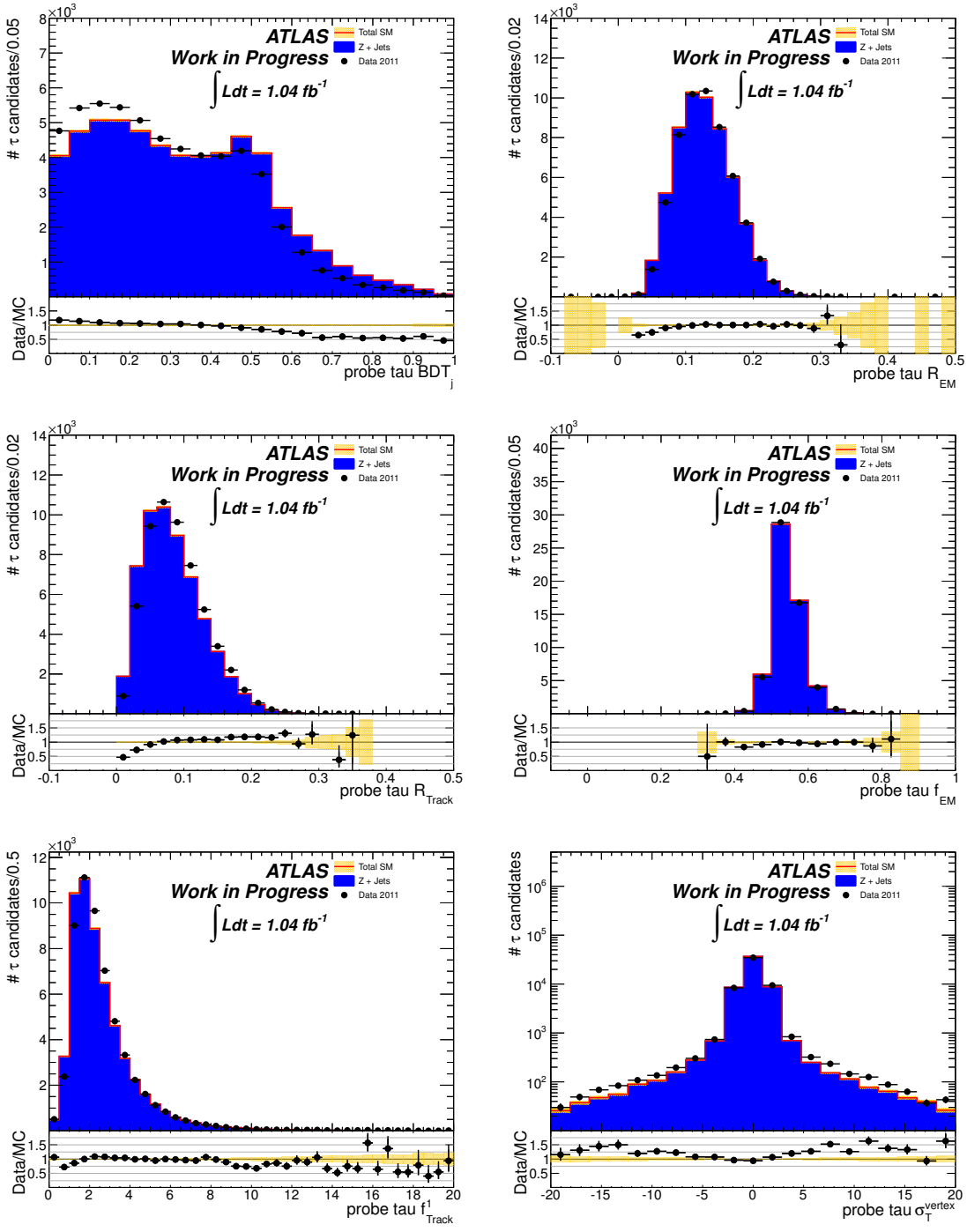


Figure 6.16: Identification variables in 2011 ATLAS data amounting to 1.04fb^{-1} for the reconstructed τ candidates matched to the probe jet: BDT jet score (a), electromagnetic radius R_{EM} (b), track average distance R_{track} (c), longitudinal fraction of the electromagnetic energy $E_{\mathrm{T}}^{\mathrm{EM}}/E_{\mathrm{T}}$ (d), fraction of the leading track p_{T} to E_{T}, $f_{\mathrm{track,1}}$ (e) and flight path significance $\sigma_{\mathrm{T}}^{\mathrm{vertex}}$ (f).

Chapter 7

SUSY Searches With Hadronic τ-Lepton Decays

The application of the methods for τ-lepton identification described in Chapter 5 and results gathered in Chapter 6 in the context of a physics analysis in a τ-lepton final state is presented in this chapter. While most of the known Standard-Model processes with τ leptons in the final state like $Z(\rightarrow \tau\tau)$ [98] and $W(\rightarrow \tau\nu)$ [99] and $t\bar{t}$ [107] have been observed at the ATLAS Experiment, τ leptons also provide an interesting probe for physics beyond the Standard Model. In many supersymmetric models the coupling to the fermions of the third generation is enhanced, making searches in τ-lepton final states especially promising. With this in mind, a search for Supersymmetry in τ lepton final states was developed and optimised for sensitivity across a large range of SUSY scenarios. The data used for this study are presented in Sec. 7.1 and Sec. 7.2 describes the event selection and the dominant contributions from Standard Model processes to the selected event samples. The basic requirements defining the analysis channel are presented in Sec. 7.2.3. Sec. 7.2.4 and Sec. 7.2.5 then present further event selection criteria applied to reject QCD multi-jet production and electroweak backgrounds, respectively. The techniques used to estimate these background processes are presented in Chapter 8 and their uncertainties, determined in the given dataset, are presented in Chapter 9. The results of the search and their interpretation in the studied SUSY signal scenarios are summarized in Chapter 10.

7.1 Data Sample and Trigger Selection

In this analysis pp collision data are used, which were collected at $\sqrt{s} = 7$ TeV between March and August of 2011. The data-taking runs[1], together with their unique run numbers and grouped according to the time period in which they were recorded, are summarised in Table 7.1. They correspond to a total integrated luminosity of 2.05 fb^{-1}.

Data are selected according to a so-called "good runs list (GRL)" selection to ensure the basic quality criteria needed for performance of the reconstruction and identification of the particles used in the analysis. These criteria are common to all $E_{\mathrm{T}}^{\mathrm{miss}}$-based SUSY searches performed at the ATLAS Experiment. After requiring GRL criteria, the total integrated luminosity amounts to 2.05 fb^{-1}, with the uncertainty on this luminosity estimated to be 3.4% [82].

Pile-up Re-weighting

As in the studies presented in the previous chapters, the simulated samples used for this analysis only roughly describe the amount of pile-up interactions found in 2011 ATLAS data. Therefore a reweighting method similar to the one described in Chapter 6 needs to be applied to the simulated samples to improve the description of pile-up in data. Unlike the studies in Chapter 6, where the distributions of the number of reconstructed vertices in data and simulation were used to derive the reweighting factors for each

[1] A data-taking run is one period of several hours in which collisions from one fill of the LHC machine are recorded.

Period	Run numbers	Runs	$\int \mathcal{L}dt$ before GRL [pb^{-1}]	$\int \mathcal{L}dt$ after GRL [pb^{-1}]
B	177986–178109	7	17.38	11.39
D	179710–180481	23	184.77	154.08
E	180614–180776	5	52.20	42.56
F	182013–182519	16	157.30	122.69
G	182726–183462	28	571.94	464.30
H	183544–184169	13	285.82	240.26
I	185353–186493	27	399.72	304.71
J	186516–186755	9	233.01	212.18
K	186873–187815	21	660.54	499.90
All		149	2562.7	2052.0

Table 7.1: Data-taking periods used in the analysis with corresponding integrated luminosities.

event, for this study the number of primary interactions per bunch crossing μ was used. This has become common practice at the ATLAS Experiment, because it has the advantage - compared to the reweighting using the number of reconstructed vetices - of not being dependent on the modelling of the vertex reconstruction efficiency in the simulation. The quantity μ is exactly known in the simulation and can be computed in data using the integrated luminosity measured over short periods of time [108].

Trigger

The trigger used to select events for this analysis is based on one high-p_{T} ($>$ 75 GeV) jet and large ($>$ 45 GeV) $E_{\mathrm{T}}^{\mathrm{miss}}$ at the electromagnetic energy scale.

It is important to note that the simulation does not reproduce the efficiency of this type of trigger well, therefore only the selection criteria for calibrated jet p_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ are applied to the simulated samples, namely $p_{\mathrm{T}}^{\mathrm{jet1}} > 130$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV. As can be seen in Fig. 7.1, this trigger can be assumed to be fully efficient for these selection criteria in data and the uncertainty on this assumption is considered small enough to be neglected. These assumptions are based on in-depth studies performed in the context of [7], where the efficiencies of the jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ parts of the trigger were measured separately. This was achieved by measuring the efficiency as a function of $p_{\mathrm{T}}^{\mathrm{jet1}}$ after requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV and measuring the efficiency as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ after requiring $p_{\mathrm{T}}^{\mathrm{jet1}} > 130$ GeV.

7.2 Event Selection and Background Suppression

This section describes the event selection designed to select SUSY-like events by a signature of several high-p_{T} jets, $E_{\mathrm{T}}^{\mathrm{miss}}$ and at least one identified hadronic τ-lepton decay. The objects used in this selection are defined in Chapter 4. The overall approach when designing this event selection has been to use a minimal set of cuts in order to minimise the dependence on a particular SUSY model. Wherever possible, the cuts are common to other $E_{\mathrm{T}}^{\mathrm{miss}}$-based SUSY searches at the ATLAS Experiment. Where appropriate, cuts have been optimised to maximise the significance. As an estimator of this significance, the Asimov approximation [109] is defined as:

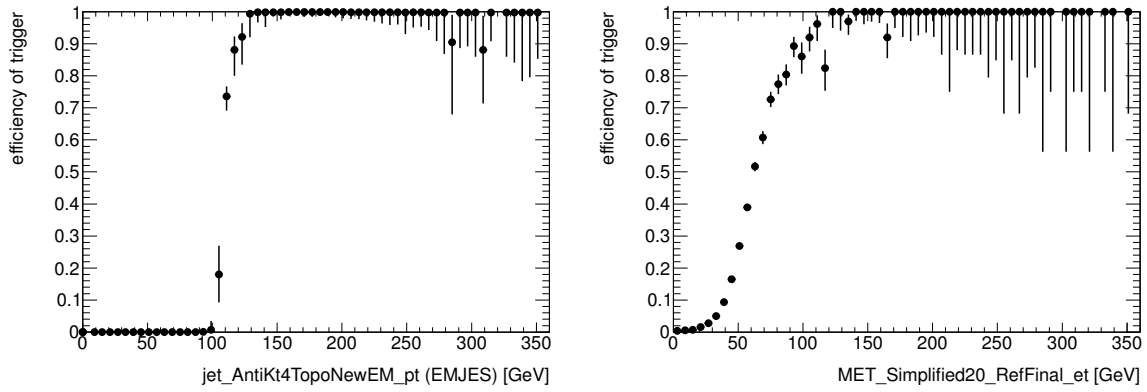


Figure 7.1: Projections of the trigger efficiencies using data from period D. Left: Projection onto the $p_{\mathrm{T}}^{\mathrm{jet1}}$ axis after requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV. Right: Projection onto the $E_{\mathrm{T}}^{\mathrm{miss}}$ axis after requiring $p_{\mathrm{T}}^{\mathrm{jet1}} > 130$ GeV.

$$z_A \equiv \sqrt{2\left[\left(N_{\mathrm{Sig}} + N_{\mathrm{BG}}\right)\ln\left(1 + \frac{N_{\mathrm{Sig}}}{N_{\mathrm{BG}}}\right) - N_{\mathrm{Sig}}\right]}.$$

The quantity z_A reduces to the familiar formula $N_{\mathrm{Sig}}/\sqrt{N_{\mathrm{BG}}}$ in the large-statistics limit, but tends to describe the Poisson fluctuations for low event yields better. Here, N_{Sig} is the number of events expected from signal processes and N_{BG} is the sum of expected events from all SM background processes. The optimisation was performed using the theory cross sections for signal an Standard Model backgrounds from Sec. 2.4.1, not including any of the data-driven background estimates from Chapter 8. It has been checked, however, that including these estimates in the optimisation does not substantially affect the result.

7.2.1 Dominant Background Processes

While this analysis is in many respects similar to other SUSY searches requiring high-p_{T} jets and $E_{\mathrm{T}}^{\mathrm{miss}}$, the one defining criterion of this event selection is the requirement of at least one identified hadronic τ-lepton decay. Fig. 7.2 shows the number of selected hadronic τ candidates for the two benchmark GMSB signal points defined in Sec. 2.4.2 and the dominant Standard-Model backgrounds as well as the signal significance of a cut on this variable. One immediately sees the advantage of requiring one or more hadronic τ-lepton decays for GMSB models or other SUSY models with similar topologies. The Standard Model processes that contribute to the selected event sample are:

Top quark production: In events with $t\bar{t}$ pairs or single top quarks both real τ leptons and neutrinos can be produced that satisfy the event selection requirements. In addition a large number of high-p_{T} jets are produced, especially in cases where one of the top quarks in $t\bar{t}$ events decays hadronically.

***W* + jets:** The decay of a W boson can also produce a real τ lepton and a neutrino. If the W boson is produced in association with high-p_{T} jets, it can contribute as a background to this search.

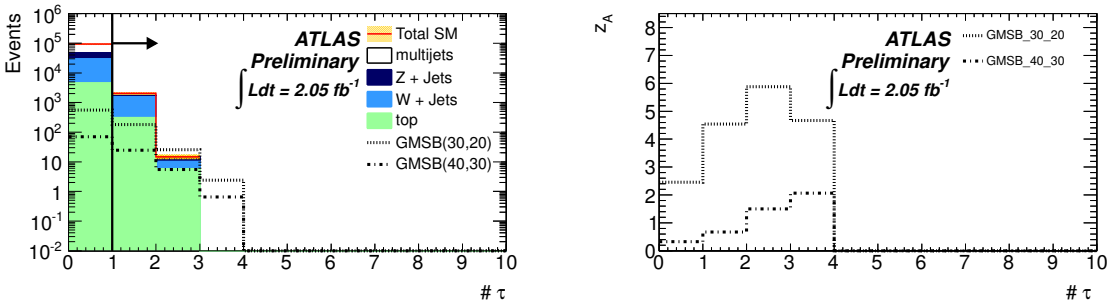


Figure 7.2: Distribution of the number of selected τ candidates after jet and $E_{\mathrm{T}}^{\mathrm{miss}}$requirements (left), with only luminosity scaling applied to Monte Carlo samples. The corresponding significance for the two example signal points over the full Standard Model background is shown on the right as a function of the cut value on the number of selected τ candidates. The error band shows sum of the statistical errors of the SM contributions.

***Z* + jets:** The main contribution from this process comes from the decay $Z \to \nu\nu$. If this decay is accompanied by several high-p_{T} jets, the neutrinos from the decay are strongly boosted, causing large $E_{\mathrm{T}}^{\mathrm{miss}}$. Since no real τ leptons are produced, the identified τ candidate must come from a hadronic jet.

QCD multi-jet production: This process can only contribute due to serious mismeasurements of the hadronic jets. Both the hard $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement and the τ lepton identification reduce the contribution from this background, but they are offset by the large cross section for QCD multi-jet production.

Other Backgrounds: Other processes that might contribute to the background for this search were also investigated, including Drell-Yan, diboson and b-quark-associated vector boson production. All of them contribute significantly less than one event to the sample selected in the signal region, based on Monte Carlo estimates, and are therefore neglected in this thesis.

7.2.2 Event Selection Overview

The selection is divided into three stages: event preselection and cleaning, rejection of QCD multi-jet and other backgrounds where the $E_{\mathrm{T}}^{\mathrm{miss}}$ comes mainly from detector effects, and final kinematic cuts to further separate SUSY signals from SM backgrounds. A veto on identified electrons or muons is applied as part of the preselection. This veto allows for a future combination of these results with other searches for SUSY in similar topologies with light leptons. The complete event selection can be summarised as follows:

- event preselection and cleaning,
- light-lepton veto,
- one high-p_{T} jet ($p_{\mathrm{T}}^{\mathrm{jet1}} > 130$ GeV) and large $E_{\mathrm{T}}^{\mathrm{miss}}$ ($E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV),
- a second low-p_{T} jet ($p_{\mathrm{T}}^{\mathrm{jet2}} > 30$ GeV),
- at least one τ candidate (identified as described in Sec. 5),
- large $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$,

- large $\Delta\phi > 0.3$ rad., where $\Delta\phi$ denotes the separation in azimuthal angle between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the two leading jets,
- large transverse mass $m_{\mathrm{T}} > 110$ GeV between the leading τ candidate and $E_{\mathrm{T}}^{\mathrm{miss}}$, which is defined as

$$m_{\mathrm{T}} = \sqrt{m_{\tau}^2 + 2 \cdot p_{\mathrm{T}}^{\tau} \cdot E_{\mathrm{T}}^{\mathrm{miss}} \cdot \left(1 - \cos\Delta\phi(\tau, E_{\mathrm{T}}^{\mathrm{miss}})\right)} \tag{7.1}$$

  and
- large $m_{\mathrm{eff}} > 600$ GeV, where the effective mass m_{eff} of an event is defined as:

$$m_{\mathrm{eff}} = p_{\mathrm{T}}^{\tau} + \sum p_{\mathrm{T}}^{\mathrm{jet}} + E_{\mathrm{T}}^{\mathrm{miss}}, \tag{7.2}$$

  where the sum runs over the two leading jets in the event. The combination of all these event selection criteria defines the "signal region (SR)".

7.2.3 Jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ Requirements

In many SUSY scenarios, the decay chains of SUSY particles are expected to produce a large number of high-p_{T} jets and large $E_{\mathrm{T}}^{\mathrm{miss}}$ from the LSPs, which escape the detector unseen. Therefore events are required to contain at least one jet with $p_{\mathrm{T}} > 130$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV, after applying the event cleaning and general data quality requirements described above. The thresholds of the leading jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ requirements are motivated by the plateau of the trigger efficiency (see Sec. 7.1), while requiring another jet with $p_{\mathrm{T}} > 30$ GeV has been shown to improve the signal significance, as can be seen in Fig. 7.3a.

7.2.4 QCD Multi-Jet Background Suppresion

Most of the QCD multi-jet events that pass the above selection show significant $E_{\mathrm{T}}^{\mathrm{miss}}$ only because of detector effects, which are caused by one or more mismeasured jets. Since these effects are difficult to model in simulation, the goal of the selection requirements described in this section is not to improve the overall signal significance of the selection, but to ensure that the selected sample after applying these requirements is mostly free of QCD multi-jet events and other backgrounds with "fake" $E_{\mathrm{T}}^{\mathrm{miss}}$. Therefore the optimisation curves for all these variables show that the overall discovery significance does not increase with a requirement on these variables, (see Fig. 7.4) except when considering the significance with respect to the QCD multi-jet background only (see Fig. 7.5).

Multi-Jet events where fake $E_{\mathrm{T}}^{\mathrm{miss}}$ can be associated with a mismeasurement of one of the two leading jets can be effectively removed by requiring an angular separation $\Delta\phi > 0.3$ rad. between each of the two leading jets and the $E_{\mathrm{T}}^{\mathrm{miss}}$ direction in the transverse plane. Distributions of the angular separation of the two leading jets from the $E_{\mathrm{T}}^{\mathrm{miss}}$ direction (denoted as $\Delta\phi(\mathrm{jet}_1, E_{\mathrm{T}}^{\mathrm{miss}})$ and $\Delta\phi(\mathrm{jet}_2, E_{\mathrm{T}}^{\mathrm{miss}})$) are shown in Fig. 7.4. For multi-jet events where the angular correlation between the leading jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ is lost, the ratio $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$ between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the effective mass of the event, m_{eff}, has proven to give a good separation between processes with "real" $E_{\mathrm{T}}^{\mathrm{miss}}$ from LSPs or neutrinos and events where the $E_{\mathrm{T}}^{\mathrm{miss}}$ is caused by detector effects. The distribution of the ratio of $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{eff} can be seen in Fig. 7.4. A selection criterion of $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$ has proven to remove a large fraction of the QCD multi-jet background without decreasing the significance of the selection.

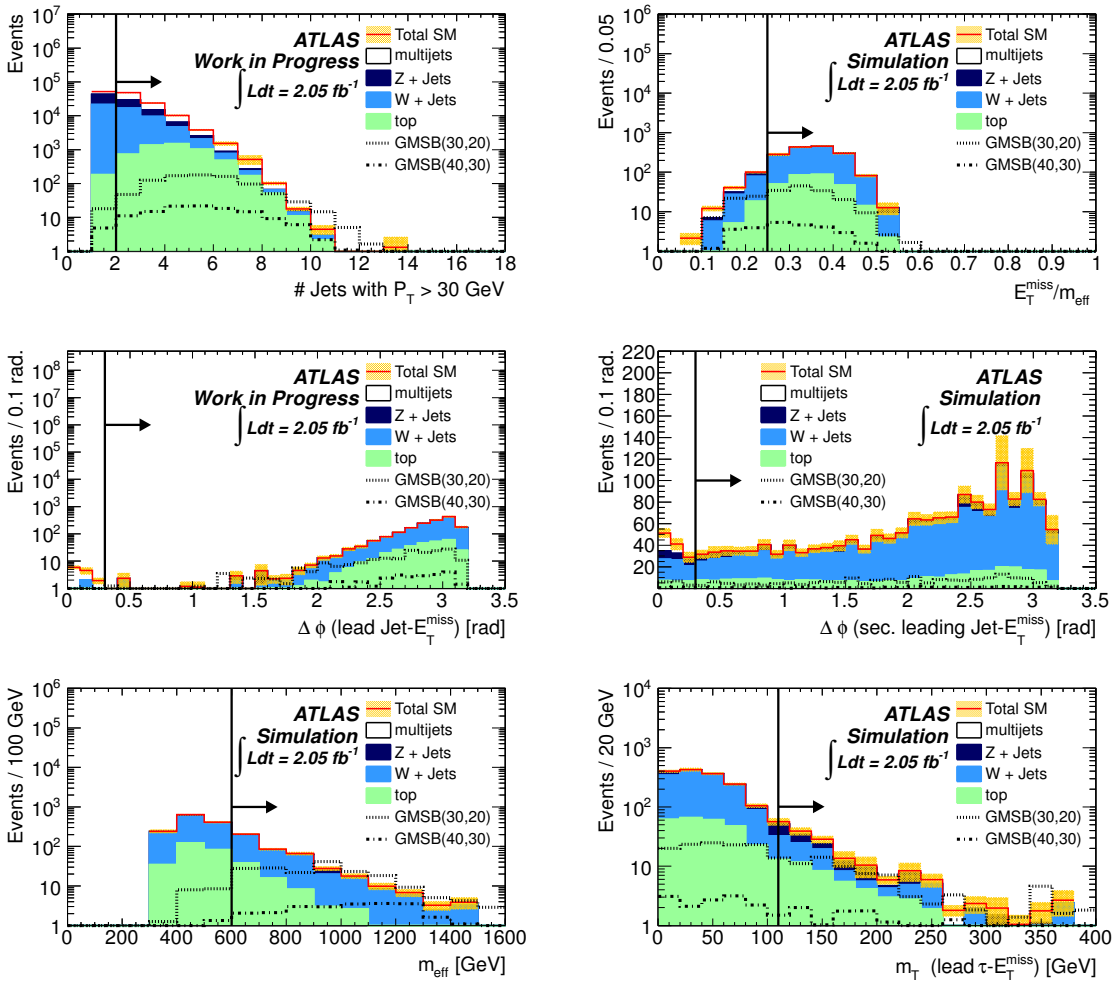


Figure 7.3: Event selection variables after baseline selection. Shown are: (a) the multiplicity of jets with p_{T} above 30 GeV, (b) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$, (c) $\Delta\phi$ between the leading jet and $E_{\mathrm{T}}^{\mathrm{miss}}$, (d) $\Delta\phi$ between the second-leading jet and $E_{\mathrm{T}}^{\mathrm{miss}}$, (e) m_{eff}, and (f) the transverse mass m_{T} of the leading τ candidate. All distributions are from Monte Carlo samples with only luminosity scaling applied. The error band shows sum of the statistical errors of the SM contributions. The black lines and arrows show the optimized selection criteria.

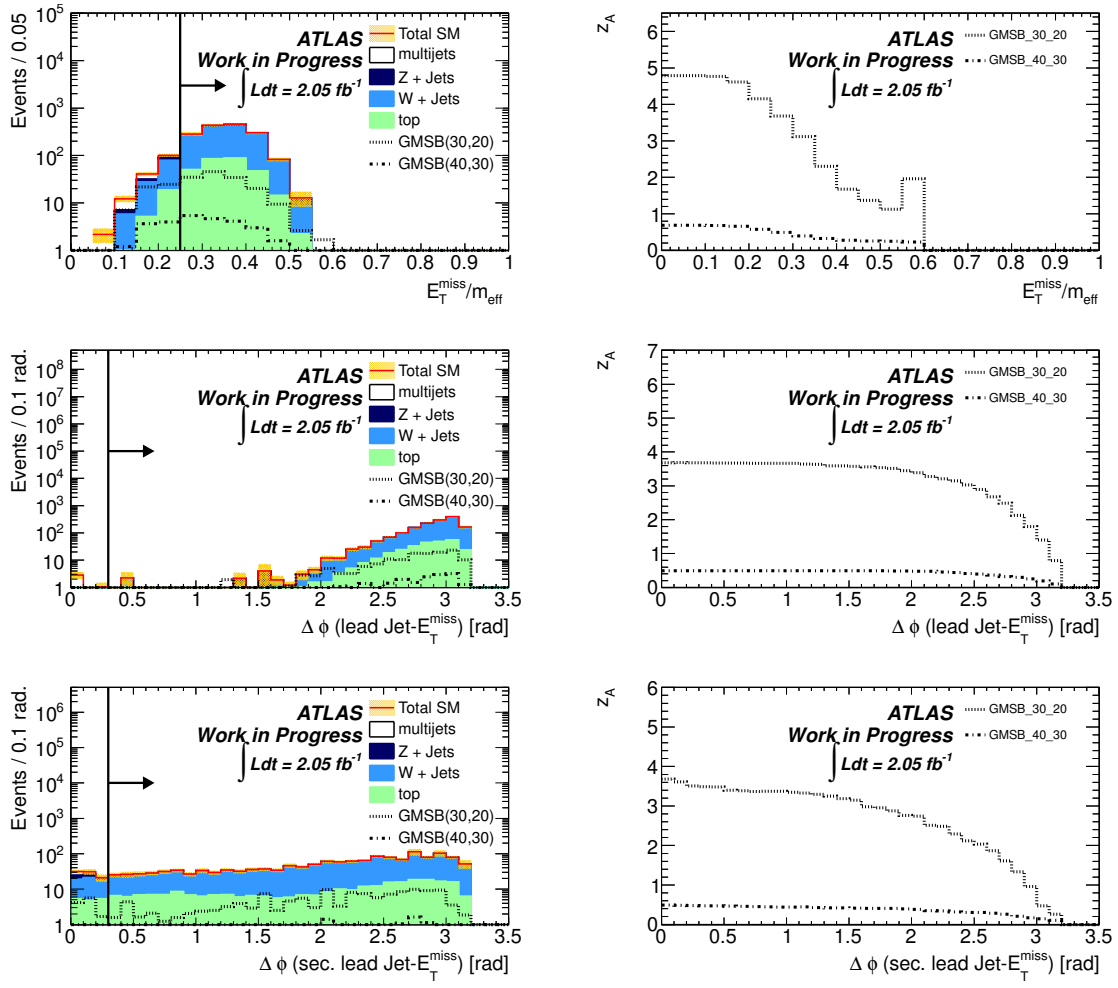


Figure 7.4: Distributions of event selection variables: (a) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$, (b) $\Delta\phi$ between the leading jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ and (c) $\Delta\phi$ between the second-leading jet and $E_{\mathrm{T}}^{\mathrm{miss}}$. The corresponding significance for the two example signal points over the full Standard Model background is shown in the figure on the right as a function of the cut value on each of these variables. Each distribution and its corresponding significance curve is shown just before the cut on this variable is applied. All distributions are from Monte Carlo samples with only luminosity scaling applied. The error band shows sum of the statistical errors of the SM contributions.

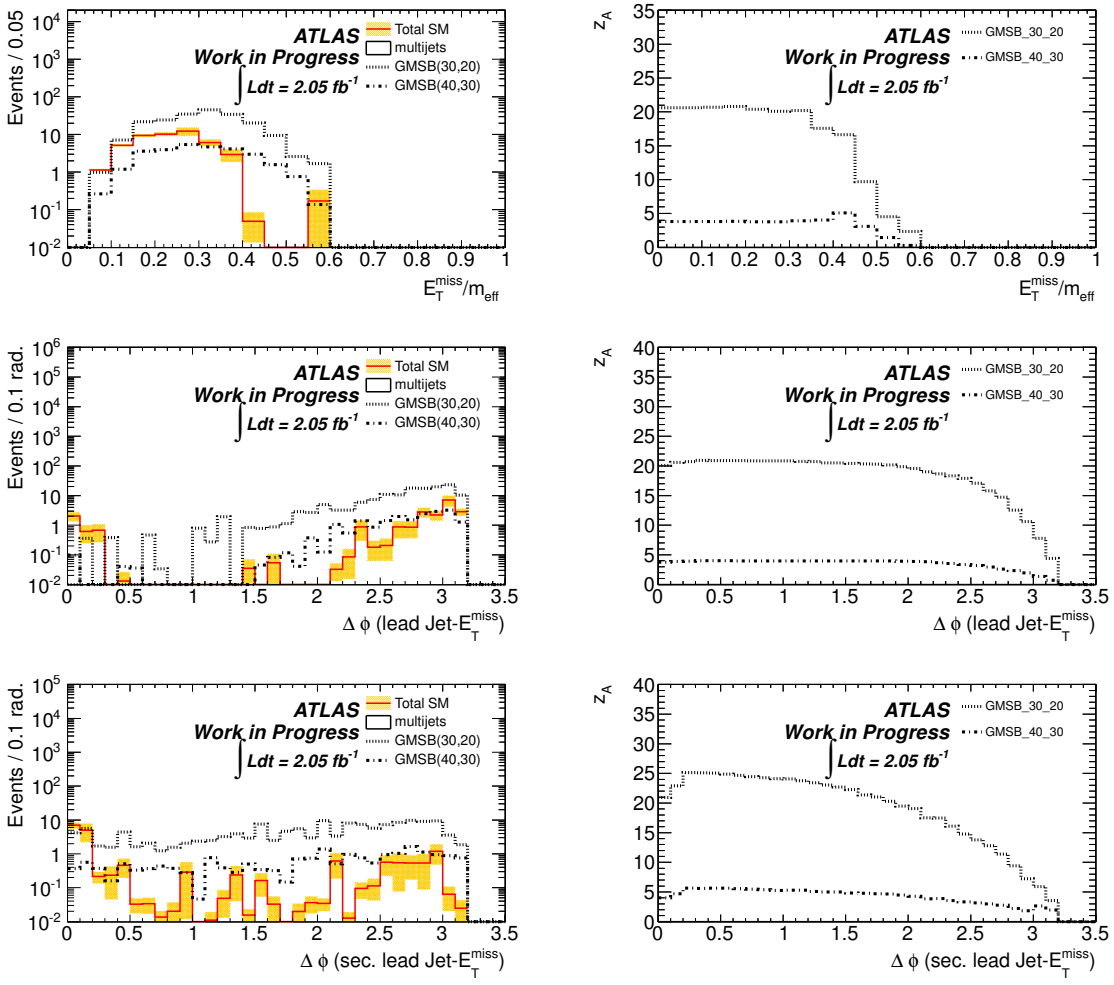


Figure 7.5: Optimisation of event selection variables against QCD only: (a) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$, (b) $\Delta\phi$ between the leading jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ and (c) $\Delta\phi$ between the second-leading jet and $E_{\mathrm{T}}^{\mathrm{miss}}$. The corresponding significance for the two example signal points, if only the background contribution from QCD multi-jet events is considered, is shown on the right as a function of the cut value on each of these variables. Each distribution and its corresponding significance curve is shown just before the cut on this variable is applied. All distributions are from Monte Carlo samples with only luminosity scaling applied. The error band shows sum of the statistical errors of the SM contributions.

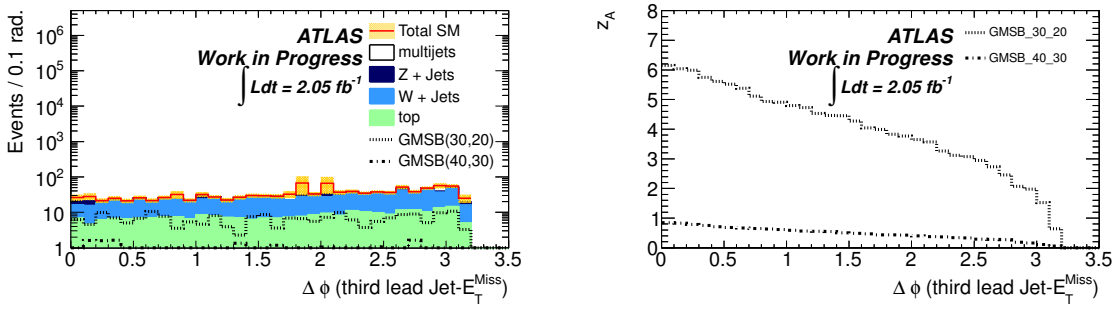


Figure 7.6: Distribution (left) of $\Delta\phi$ between the third-leading jet and $E_{\mathrm{T}}^{\mathrm{miss}}$, for events with at least three jets. The corresponding significance for the two example signal points, if only the background contribution from QCD multi-jet events is considered, is shown on the right as a function of the cut value on this variable. A cut on this variable is not seen to improve the significance.

To further ensure that the selected sample is basically free of contributions from QCD multi-jet events, the potential usefulness of requiring angular separation between the third-leading jet, if there are at least three jets in the event, and $E_{\mathrm{T}}^{\mathrm{miss}}$ was studied. The distribution of this variable and the significance of a cut on it are shown in Fig. 7.6. One can see that including a requirement on the third jet would noticeably decrease the significance for both of the benchmark points considered. The contribution from QCD multi-jet production is therefore controlled by the previous requirements on $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$, $\Delta\phi(\mathrm{jet}_1, E_{\mathrm{T}}^{\mathrm{miss}})$ and $\Delta\phi(\mathrm{jet}_2, E_{\mathrm{T}}^{\mathrm{miss}})$.

7.2.5 Suppression of Background Processes with real $E_{\mathrm{T}}^{\mathrm{miss}}$

After suppressing events with fake $E_{\mathrm{T}}^{\mathrm{miss}}$, the dominant remaining backgrounds contain real $E_{\mathrm{T}}^{\mathrm{miss}}$, mostly from leptonic W decays, where the W boson is either produced directly or where it comes from the decay of a top quark. To identify these events, the transverse mass m_{T} of the leading selected τ candidate and $E_{\mathrm{T}}^{\mathrm{miss}}$ is used. In cases where the τ candidate comes from a leptonic W decay (either from a real τ lepton or a misidentified electron or muon), this variable must satisfy $m_{\mathrm{T}} < m_W$ up to resolution effects. Fig. 7.7 shows distributions of m_{T} after the other selection cuts have been applied. The requirement on this variable, as indicated by the arrow in the figure, is $m_{\mathrm{T}} > 110$ GeV. Another class of background events, which is only slightly reduced by the requirement on m_{T}, comes from the process Z+jets, where the Z boson decays to a pair of neutrinos (hereafter denoted as $Z(\rightarrow \nu\nu)$+jets). In these events the two neutrinos from the Z-boson decay are highly boosted due to the high p_{T} threshold for the leading jet against which the Z boson recoils, leading to large values of $E_{\mathrm{T}}^{\mathrm{miss}}$. To further improve the signal sensitivity with respect to all these background, the final selection step is a requirement on the effective mass: $m_{\mathrm{eff}} > 600$ GeV. The m_{eff} distribution and optimisation curve are shown in Fig. 7.7 (c) and (d), respectively.

7.2.6 Expected Event Yields

An overview of the event selection obtained using simulated samples is shown in Table 7.2, where the simulated samples are normalised to an integrated luminosity of 2.05fb^{-1}. One expects 20.6 SM background events after applying all event selection criteria. In contrast, the two SUSY benchark points GMSB(30,20) and GMSB(40,30), which were used for optimising the selection, yield 47.8 and 9.1

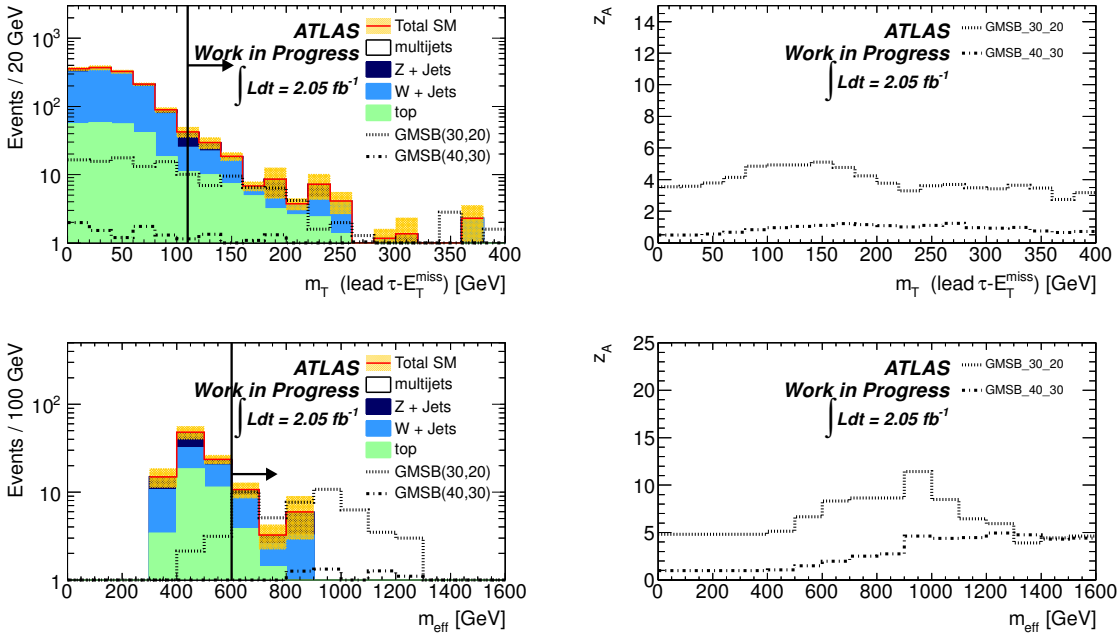


Figure 7.7: (a) Transverse mass m_{T} of the leading τ candidate and $E_{\mathrm{T}}^{\mathrm{miss}}$ and (b) the significance as a function of the cut on this variable. (c) shows effective mass m_{eff} of the leading τ candidate, the two leading jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ and (d) the significance as a function of the cut on this variable. Both are shown just before the cut on this variable is made in the event selection. All distributions are from Monte Carlo samples with only luminosity scaling applied. The error band shows the sum of statistical errors on the SM backgrounds only.

events, respectively. This corresponds to significances of 10.8 and 3.6 using the Asimov approximation. The notation GMSB(40,30) stands for the GMSB model with $\Lambda = 40$ TeV and $\tan\beta = 30$ and analogously for GMSB(30,20).

	GMSB(30,20)	GMSB(40,30)	top	W	Z	QCD
before cuts	4821 ± 51	927.4 ± 9.7	414250 ± 420	(6452.6 ± 2.6)×10^4	(8542.1 ± 3.3)×10^3	(4684.9 ± 4.0)×10^6
event cleaing	4800 ± 51	921.9 ± 9.7	412990 ± 310	(6269.0 ± 2.5)×10^4	(8343.0 ± 3.3)×10^3	(4682.5 ± 4.0)×10^6
$n_e = 0$	3300 ± 42	573.0 ± 7.7	344900 ± 300	(5189.4 ± 2.3)×10^4	(6759.6 ± 3.1)×10^3	(4671.2 ± 4.0)×10^6
$n_\mu = 0$	2388 ± 36	377.6 ± 6.2	283500 ± 290	(4050.5 ± 2.0)×10^4	(5139.0 ± 3.0)×10^3	(4670.7 ± 4.0)×10^6
$p_{\mathrm{T}}^{\mathrm{jet1}} > 130$ GeV	1204 ± 25	164.1 ± 4.1	68820 ± 150	163760 ± 700	82420 ± 510	(2921.0 ± 7.4)×10^5
$E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV	901 ± 22	126.1 ± 3.6	5508 ± 20	50440 ± 480	41430 ± 390	(42.4 ± 1.7)×10^3
$p_{\mathrm{T}}^{\mathrm{jet2}} > 30$ GeV	761 ± 20	101.3 ± 3.3	4819 ± 18	26120 ± 230	18350 ± 260	(30.1 ± 1.4)×10^3
$n_\tau > 0$	203 ± 10	28.7 ± 1.7	311.9 ± 4.1	1220 ± 42	145 ± 15	47.5 ± 3.8
$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$	148.1 ± 8.7	19.6 ± 1.4	287.7 ± 3.9	1131 ± 42	128 ± 14	21.5 ± 3.5
$\Delta\phi$ cuts	136.3 ± 8.4	18.2 ± 1.3	271.1 ± 3.8	1089 ± 41	114 ± 13	6.0 ± 1.3
$m_{\mathrm{T}} > 110$ GeV	53.1 ± 5.2	9.9 ± 1.0	39.5 ± 1.5	38.7 ± 4.5	27.4 ± 8.7	1.4 ± 0.6
$m_{\mathrm{eff}} > 600$ GeV	47.8 ± 5.0	9.1 ± 1.0	6.6 ± 0.6	9.0 ± 2.4	4.9 ± 3.0	0.1 ± 0.1

Table 7.2: Overview of the event selection using simulated samples. The simulated samples are normalised to 2.05fb^{-1} and do not include normalisation factors from control regions. "Before cuts" refers to the expectations from MC samples with no selection applied. For comparison, the event yields at the end of the event selection from the data-driven background estimates can be seen in Table 8.6 in Chapter 8.

Chapter 8

Background Estimation Techniques

This chapter describes the methods and control regions used to estimate the dominant background contributions to the search for Supersymmetry in τ lepton final states. One important contribution from W boson production, with the W decaying into a real τ lepton and a τ neutrino, is discussed in Sec. 8.1. Another important contribution comes from events where a real W or Z boson is present but the identified τ candidate comes from the misidentification of a hadronic jet produced in association with the vector boson. In both the events with real- and fake-τ candidates the W can either be promptly produced, or come from a top quark decay in $t\bar{t}$ or single top quark events. Another important background comes from $Z(\rightarrow \nu\nu)$ + jets production. Because these events only contain fake τ candidates, the number of simulated events after applying the full selection to this background is very small. A scaling technique is used to increase the availabe number of events, which is described in Sec. 8.2. The estimation of the total contribution from this and the other backgrounds with fake-τ candidates is then described in Sec. 8.3. These background contributions with real- or fake-τ candidates are estimated by determining the normalisation from control regions in the data and then deriving a transfer factor from simulation to extrapolate back to the signal region. An alternative, completely data-driven method to estimate the $Z(\rightarrow \nu\nu)$ + jets contribution is discussed in Sec. 8.3.1. The smallest, non-negligible contribution comes from QCD multi-jet production and is described in Sec. 8.4. For this background, two complementary approaches were investigated, where one is a completely data-driven background estimate and the other is based on deriving the normalisation from data and applying a combination of transfer factors determined from data and simulation. Several of the background estimation techniques presented here, among them the estimation of the W, top and QCD multi-jet background, were developed in close collaboration with Felix Bührer in the context of his diploma thesis [110].

8.1 Estimation of the Background with real τ leptons from W + jets and Top

As explained in Sec. 7.2.5, there is a significant contribution to the total Standard Model background from W + jets and top quark production, where a W boson decays into a real lepton. No matter whether this lepton is a real τ lepton that decays hadronically or the identified τ candidate comes from a light lepton, the transverse mass of the τ candidates and $E_{\mathrm{T}}^{\mathrm{miss}}$ is bounded by the mass of the W boson. Therefore a control region for such events can be obtained by inverting the m_{T} requirement and dropping the requirement on m_{eff}. The distribution of m_{T} for simulated W and top quark events is shown in Fig. 8.1 separately for the two cases where the identified τ candidate is matched to a true hadronic τ-lepton decay in the simulation (real) or not (fake). Since most events with real-τ candidates have m_{T} below 70 GeV, while there is still a significant contribution from events with fake-τ candidates reconstructed from hadronic jets in the region 70 GeV $< m_{\mathrm{T}} <$ 110 GeV, the requirement for the true-τ enriched control region is set to $m_{\mathrm{T}} <$ 70 GeV. This control region is hereafter denoted as the "true-τ enhanced control

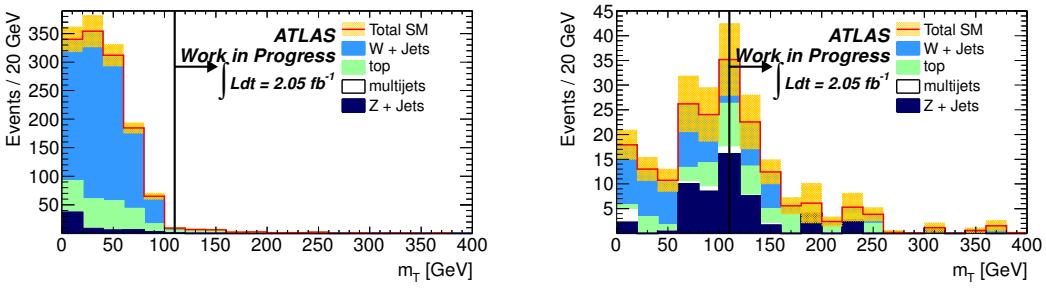


Figure 8.1: m_{T} distributions for events with truth-matched τ candidates (left) and non-truth-matched τ candidates (right).

region ($\mathrm{CR_{true}}$)". The number of events in $\mathrm{CR_{true}}$ for data and all backgrounds considered are shown in Table 8.1.

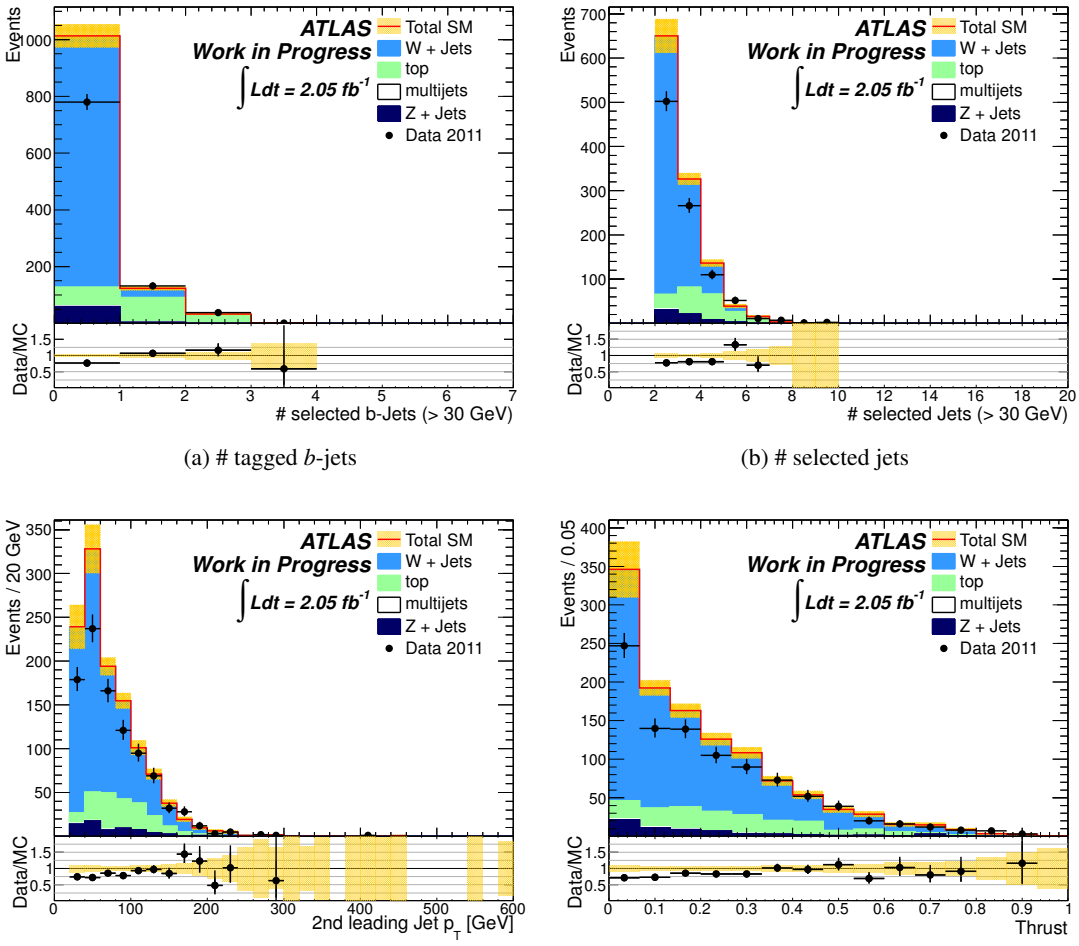
	True τ	Fake τ	Total
Top	180.8 ± 8.1	5.6 ± 0.3	186.4 ± 8.4
W + jets	882.2 ± 38.0	36.8 ± 1.6	919.0 ± 40.0
Z + jets	54.1 ± 5.8	8.1 ± 0.9	62.2 ± 6.7
Multi-Jet	0 ± 0	1.8 ± 1.8	1.8 ± 1.8
Total	111.7 ± 39.7	52.3 ± 2.6	1169.4 ± 41.5
Data			951

Table 8.1: Numbers of observed and expected events in the true-τ-dominated W/top control region $\mathrm{CR_{true}}$, defined as $m_{\mathrm{T}} < 70$ GeV. The numbers shown for W + jets and top are from Monte Carlo simulation and do not include the correction factors derived from this control region.

It can be seen that the by far dominant contribution is direct W + jets production with the next largest contribution being $t\bar{t}$ and single top quark production, which are summed in the contribution labelled "top". Only small contributions come from Z + jets and QCD multi-jet production. The true-τ purity is 97% for top, 96% for W + jets and 87% for Z + jets. In QCD multi-jet production the fraction of events with real τ leptons is negligible.

8.1.1 Separating the Contributions from W + jets and Top

The goal of defining $\mathrm{CR_{true}}$ is to derive the normalisation of the background with real hadronic τ-lepton decays from data. Since $\mathrm{CR_{true}}$ is not a pure control region for any one background process, but is a mixture of W + jets and top quark events, with a small contamination from QCD and Z + jets events, it is desirable to separate out the W + jets and top quark contributions and derive separate scale factors for the simulated samples for these two processes. The small contributions from Z + jets and multi-jet events in this control region are estimated as follows: Since the Z + jets process is expected to be reasonably well described in this control region, as it is dominated by events with real τ leptons, for which the efficiency is known to be well modelled (see Sec. 5.3.1), this contribution is taken from simulation directly. The even smaller contribution from QCD is estimated using simulated events reweighted as described in Appendix A. A data-driven cross check for this estimate was also developed, but since the QCD multi-



(a) # tagged b-jets

(b) # selected jets

(c) p_{T} of the second leading jet

(d) thrust

Figure 8.2: Input variables for the W/top separation BDT, shown in the true-τ enhanced $\mathrm{CR}_{\mathrm{true}}$. The error band shows the combined statistical uncertainty on the SM backgrounds.

jet contribution is so small in $\mathrm{CR}_{\mathrm{true}}$ this additional study is not detailed here. It is, however, important to note that this data-driven cross check yields consistent results with using reweighted, simulated QCD multi-jet events.

The normalisation of W + jets and top quark in $\mathrm{CR}_{\mathrm{true}}$ can be determined using the the number of tagged b-jets in the event (see Fig. 8.2). This can be achieved by fitting the data distribution in a binned, maximum-likelihood fit using templates for the different processes derived from simulation. The normalisation of the Z+jets and multi-jet contribution are fixed as described above and the normalisation of the W+jets and top quark contributions are allowed to vary to fit the data. The implementation chosen for the fit is the so called `TFractionFitter` developed by Barlow and Beeston [111].

Since the number of b-tagged jets per event comes with its own set of systematic uncertainties, namely the uncertainties on the b-jet identification efficiency and on the light jet misidentification probability, it is desirable to be not solely dependent on this observable. As a compromise, a multivariate classifier (a boosted decision tree) was trained to separate W + jets and top. The following four variables are used as

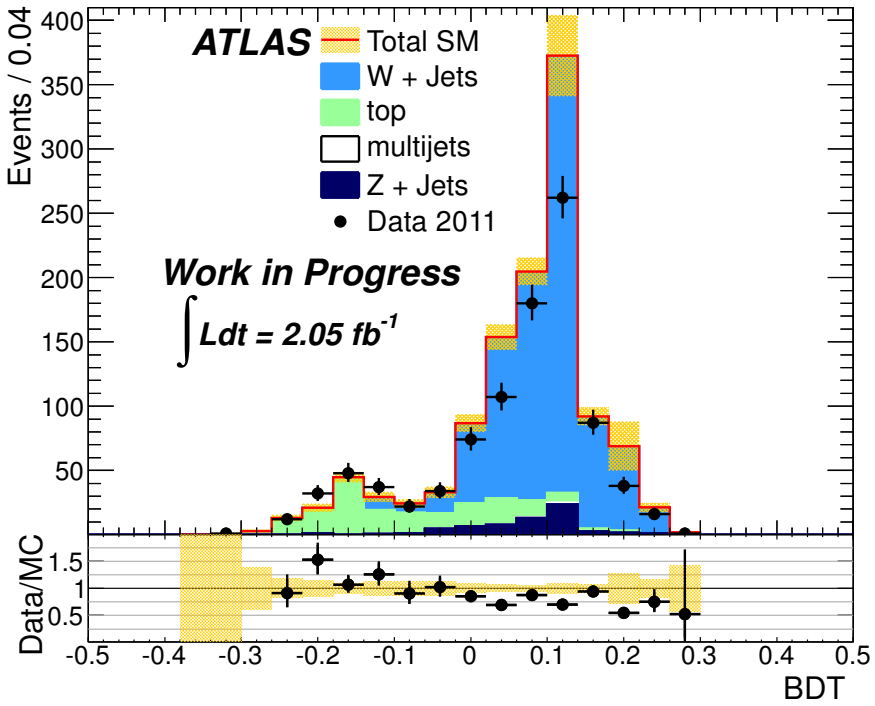


Figure 8.3: BDT output computed from the input variables in Fig. 8.2, for data and MC simulation in CR_{true}. The error band denotes the total statistical uncertainty on the SM background.

input to the BDT:

- the number of tagged b-jets with $p_T > 30$ GeV in the event,
- the number of jets with $p_T > 30$ GeV,
- the p_T of the second leading jet,
- the transverse thrust T of the event, which is defined as: $T = max_{|n|=1} \frac{\sum_i \vec{n} \cdot \vec{p}_i}{|\sum_i \vec{p}_i|}$, where the thrust axis is determined by the normal vector $\vec{n}$ for which a maximum of T is attained and $\vec{p}_i$ are the momenta of all jets in the event.

Distributions of these input variables are shown in Fig. 8.2. It can already be seen that the relative abundance of the two processes seems not to be modelled well with simulated events normalised to the theoretical cross sections. The resulting classifier (Fig. 8.3) also shows clear discrepancies. In Table 8.2 the numbers of W + jets and top quark events expected from simulation are compared to the results of fitting the normalisation of the two processes to the data BDT distribution.

From the "nominal" fit results, where the BDT classifier uses the full list of variables noted above, one can derive scale factors to correct the normalisation of the W and top contributions:

$$\begin{aligned} f_W &= 0.71 \pm 0.03, \\ f_{\text{top}} &= 1.22 \pm 0.13 \, . \end{aligned}$$

	Nominal fit	Fit without b-tagging	MC prediction
W+jets	650 ± 58	661 ± 57	919 ± 40
Top	228 ± 31	218 ± 42	186.4 ± 6.4

Table 8.2: Results of two fits in CR_{true} compared to MC predictions, once using a BDT classifier with b-tagging information and once without b-tagging information. The two fits are consistent, but the errors are larger for the fit without b-tagging information. The χ^2/ndf is 10.8/8 (7.0/8) for the fit with (without) b-tagging information, and the linear correlation between the two scale factors is -29.5% (-57.3%).

The table also containes the results of fitting the background templates to data for a different BDT classifier, which does not include the number of identified b-jets as an input variable. While the uncertainties obtained from a fit to this classifier are larger, the good agreement with the fit result obtained with b-tagging information gives confidence that no systematic bias is introduced by the b-tagging information. To verify the scale factors derived here, the input distributions for the BDT and the classifier itself are again shown in Fig. 8.4 and Fig. 8.5, respectively, with the scale factors for W + jets and top applied. Improved agreement can be observed for all distributions, such that one can say that the simulation agrees very well with the data within the uncertainties given. These uncertainties include the statistical uncertainties on the data and simulation as well as the statistical uncertainties on the scale factors derived from the fit.

Using the scale factors obtained from this fit, the number of background events with real hadronic τ lepton decays in the signal region can be estimated[1] as:

$$\hat{N}_i^{\mathrm{SR}} = f_i \times N_i^{\mathrm{MC,SR}} , \tag{8.1}$$

which yields in the final signal region $m_{\mathrm{eff}} > 600$ GeV:

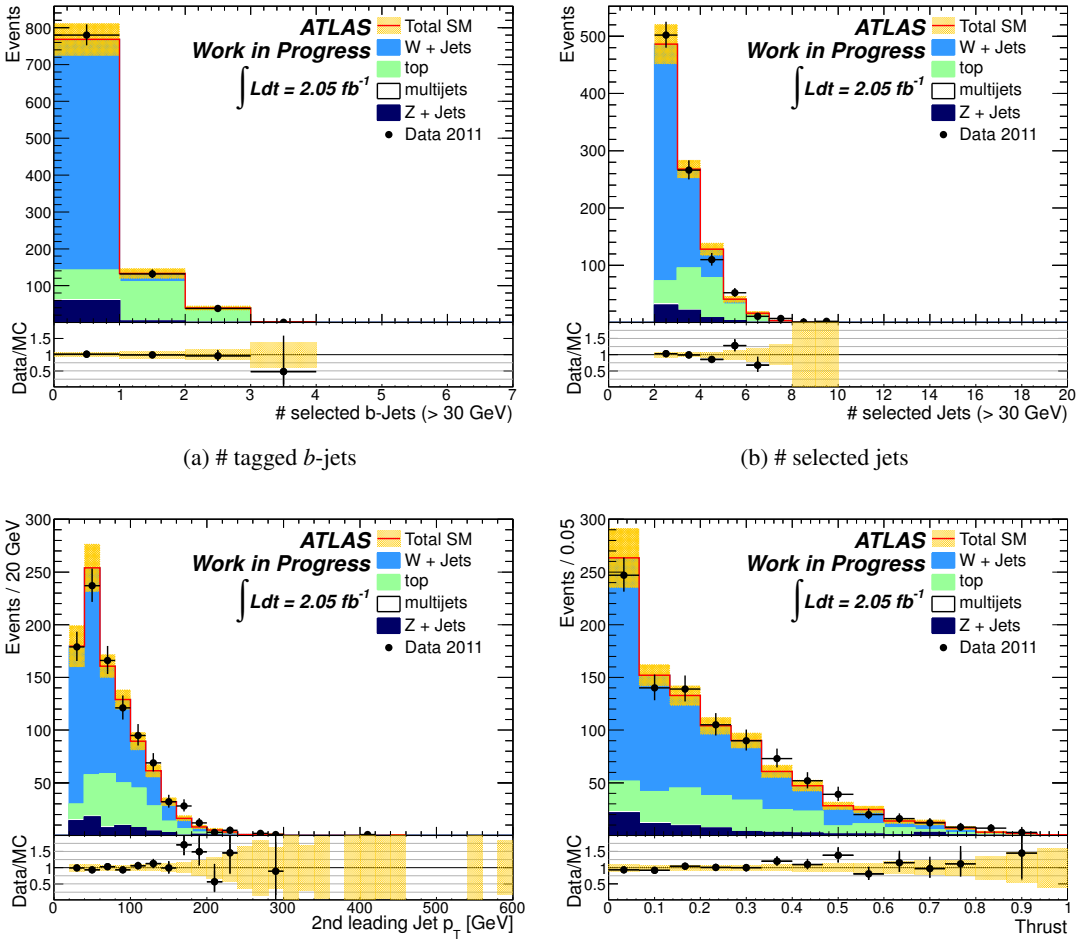
$$\begin{aligned} \hat{N}_W^{m_{\mathrm{eff}}>600\ \mathrm{GeV}} &= 3.0 \pm 1.1, \\ \hat{N}_{\mathrm{top}}^{m_{\mathrm{eff}}>600\ \mathrm{GeV}} &= 2.9 \pm 1.3 . \end{aligned}$$

The quoted errors also include the Monte Carlo statistical uncertainties in the signal region. The systematic uncertainties of extrapolating from CR_{true} to the signal region will be discussed in Chapter 9. As a final check, the W and top backgrounds in the signal region can be estimated without fitting the separate contributions, i.e., assuming $f_W = f_{\mathrm{top}} = 1$. In this case, the two values $\hat{N}_i$ move by about 1σ (W increasing, top decreasing), but the sum of the two backgrounds changes by less than 2%, so any potential unknown problems with this method do not affect the sum of background events with true τ leptons in the signal region.

8.2 $Z(\rightarrow \nu\nu)$ + jets Event Weighting using Misidentification Probabilities

Similar to the QCD multi-jet background, the number of hadronic τ-lepton decays in $Z(\rightarrow \nu\nu)$ + jets events is negligible. This means that when using simulated $Z(\rightarrow \nu\nu)$ + jets events the available number

[1] the $\hat{N}_i^{SR}$ symbol denotes a statistical estimator of the number of background events of type i in the signal region



(a) # tagged b-jets

(b) # selected jets

(c) p_{T} of the second leading jet

(d) thrust

Figure 8.4: Input variables for the W/top separation BDT after applying the obtained scale factors. The errors shown include Monte Carlo statistics and the statistical uncertainties on the scale factors.

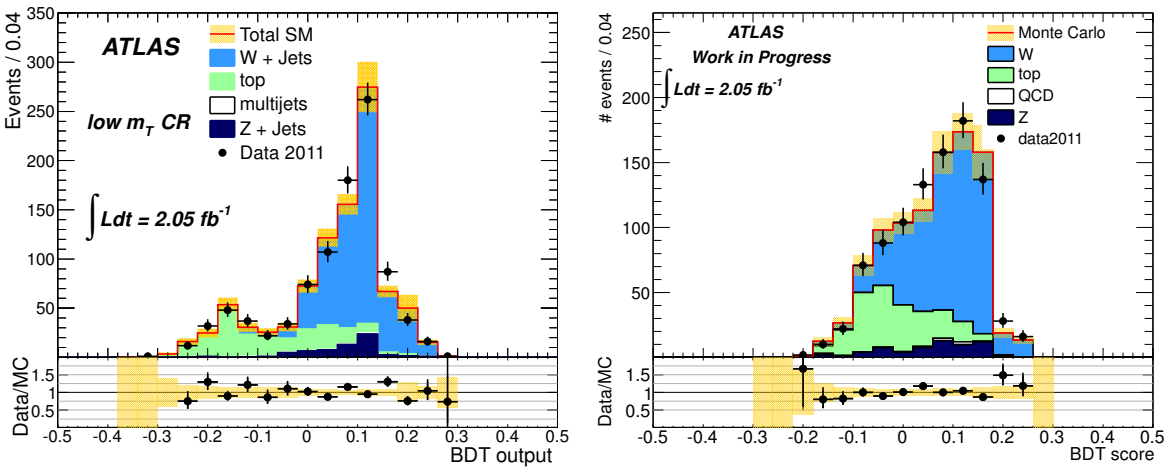


Figure 8.5: Results of two fits to the BDT distribution from data in CR_{true}. Left: fit to the nominal BDT distribution with b-tagging information included. Right: fit to the BDT distribution without b-tagging information. The results obtained from these two fits are consistent, but the BDT without b-tagging is less discriminating.

of events is reduced greatly when requiring an identified τ candidate. Unlike for the QCD background, it is expected that the normalisation to the theoretical cross section and the overall description of observables for this background is sufficient to base the background estimate purely on simulation. The only observable for which the simulation differs significantly from the data is the misidentification probability for the hadronic jets, produced together with the Z boson, to be identified as hadronic τ-lepton decays. This discrepancy between the vaules of f_{ID} measured in data and in simulation, which was observed in the study described in Sec. 6.6.2, is shown in Fig. 8.6. To correct for this mismodelling in the simulation and to improve the available number of simulated events at the end of the selection, the simulated $Z(\rightarrow \nu\nu)$ + jets sample is reweighted using the misidentification probability f_{ID} of the hadronic jets obtained from $Z(\rightarrow \mu\mu)$ + jets events in data. The weights, which are applied to the simulated events, are calculated in the following way: the same preselection used for the determination of the misidentification probability, f_{ID}, in Sec. 6.6.2 is applied to all reconstructed τ candidates in the event. For this list of n candidates, with their respective values of $f_{\mathrm{ID,i}}$, all possible combinations of BDT_j pass or fail decisions are considered, and the absolute probability w for each combination is calculated as

$$w = \prod_{i=1}^{n} [f_{\mathrm{ID,i}}]_{\text{if candidate i passes}} , [1 - f_{\mathrm{ID,i}}]_{\text{if candidate i fails}} \cdot \tag{8.2}$$

From all possible combinations in which at least one candidate passes the identification criteria, one is chosen randomly, according to their relative probability, and its absolute probability w is then applied as an event weight. This method retains all events in which at least one reconstructed τ candidate is found which fullfils the preselection criteria and ensures that the efficiency for the τ-identification step is correctly modelled. This procedure is the same as for the QCD multi-jet simulation explained in detail in Appendix A. The only difference is that by default the measurement of f_{ID} from $Z(\rightarrow \mu\mu)$ + jets events detailed in Sec. 6.6 is used to determine the event weights. This measurement of f_{ID} is well-suited to describe the fake-τ candidates in $Z(\rightarrow \nu\nu)$ + jets events, since the kinematics of the jets and the relative abundance of quark- and gluon-initiated jets is expected to be the same for $Z(\rightarrow \mu\mu)$ + jets and $Z(\rightarrow \nu\nu)$ + jets events.

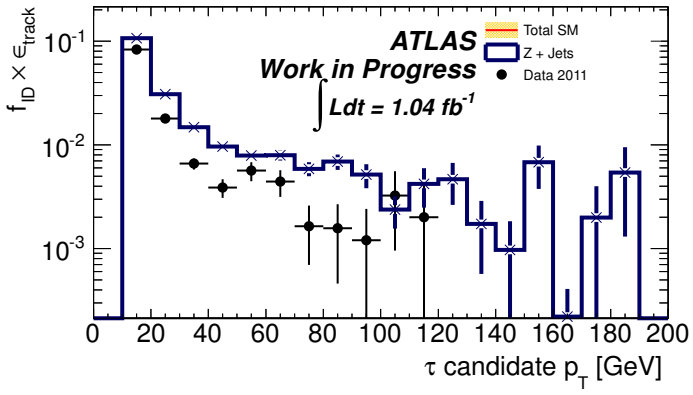


Figure 8.6: Misidentification probability f_{ID} measured in Z + jets events for the BDT$_j$ working point used in this analysis. The measured values of f_{ID} are shown down to 15 GeV, but only τ candidates with $pt > 20$ GeV are considered in the analysis.

	True τ	Fake τ	Total
Top	53.3 ± 7.5	37.8 ± 5.8	91.1 ± 9.4
W + jets	80.5 ± 6.9	33.3 ± 4.1	113.8 ± 8.0
Z + jets	5.1 ± 1.6	41.5 ± 10.8	46.6 ± 10.9
Multi-Jet	0 ± 0	2.9 ± 1.0	2.9 ± 1.0
Total	139 ± 10	116 ± 13	254 ± 17
Data			197

Table 8.3: Numbers of observed and expected events in the fake-τ-enhanced control region. The numbers of expected W + jets and top-quark events have been corrected by the correction factors measured in the true-τ-dominated region. The fake-τ correction factor obtained from data is 0.50 ± 0.08.

8.3 Estimation of the Background with fake τ Leptons from W + jets, $Z(\rightarrow \nu\nu)$ + jets and Top

Using the event weighting and scale factors detailed above, the agreement between data and simulation is again verified for all events passing the requirement $m_{\mathrm{T}} > 70$ GeV but failing either of the two signal region requirements of $m_{\mathrm{T}} > 110$ GeV and $m_{\mathrm{eff}} > 600$ GeV. The important difference between this additional control region (hereafter denoted as $\mathrm{CR}_{\mathrm{fake}}$) and $\mathrm{CR}_{\mathrm{true}}$ is the relative abundance of events with real- and fake-τ candidates. While $\mathrm{CR}_{\mathrm{true}}$ is composed of more than 90% events with a real hadronic τ lepton decay, $\mathrm{CR}_{\mathrm{fake}}$ only contains about 55% events with real τ leptons. This is much closer to the value in the signal region, where the events with real τ leptons make up only about 40% of the event yield. The number of events, separated by true and fake contributions, and the yield of data events in $\mathrm{CR}_{\mathrm{fake}}$ are detailed in Table 8.3. Here, the W and top contributions have already been corrected for the scale factors derived in $\mathrm{CR}_{\mathrm{true}}$. To derive a scale factor for the total fake contribution one can subtract all of the expected true-τ contributions and compare to the total expected fake-τ contribution, giving:

$$f_{\mathrm{fake}} = \frac{N^{\mathrm{CR}_{\mathrm{fake}}} - \hat{N}^{\mathrm{CR}_{\mathrm{fake}}}_{\mathrm{true}}}{\hat{N}^{\mathrm{CR}_{\mathrm{fake}}}_{\mathrm{fake}}} = \frac{197 - (139 \pm 10)}{(116 \pm 13)} = 0.50 \pm 0.08\ .$$

Here, the uncertainty includes the statistical uncertainties of the data and the simulated samples in $\mathrm{CR}_{\mathrm{fake}}$ and the uncertainties on the scale factors derived in $\mathrm{CR}_{\mathrm{true}}$. Note that no attempt is made to derive separate fake-τ scale factors for W, Z or top events, but just a single scale factor for all fake-τ candidates in non-QCD events is defined. With this scale factor, the fake-τ background in the signal region is estimated by scaling the combined W, top, and Z fake-τ MC event yield by f_{fake} (taking the relative abundances of their contributions from simulation). Before any scaling, one obtains:

$$\begin{aligned}
\hat{N}^{m_{\mathrm{eff}}>600\ \mathrm{GeV}}_{\mathrm{W,fake}} &= 4.7 \pm 1.8 \\
\hat{N}^{m_{\mathrm{eff}}>600\ \mathrm{GeV}}_{\mathrm{top,fake}} &= 4.0 \pm 1.2 \\
\hat{N}^{m_{\mathrm{eff}}>600\ \mathrm{GeV}}_{\mathrm{Z,fake}} &= 4.9 \pm 0.5
\end{aligned}$$

where the errors given are only the statistical uncertainties of the simulated samples. The total fake background would therefore be 13.6 ± 2.2 events. Scaling the W and top contributions by the scale factors determined in $\mathrm{CR}_{\mathrm{true}}$ and all contributions by f_{fake} yields:

$$\hat{N}^{m_{\mathrm{eff}}>600\ \mathrm{GeV}}_{\mathrm{fake}} = 6.6 \pm 1.9 \tag{8.3}$$

for the signal region. To be conservative an additional systematic uncertainty from the mixture of the various backgrounds in the Monte Carlo simulation is included in the total uncertainty on this estimate. This additional uncertainty is estimated by varying the relative abundances of W, top, and Z events up and down by 50% and the QCD multi-jet background by 100% and propagating these changes to the final background estimate.

To verify that the shapes of important observables are well described, Fig. 8.7 shows several kinematic variables in $\mathrm{CR}_{\mathrm{fake}}$ after scaling the fake-τ contribution. Good agreement between the rescaled simulation and the data is seen.

The truth composition of fake-τ candidates in the simulation was also studied. It was found that the relative abundances of misidentified candidates due to electrons, muons, light quarks, heavy quarks, and gluons are all found to be similar between $\mathrm{CR}_{\mathrm{fake}}$ and the signal region, both for one- and three-prong τ candidates. This agreement adds further confidence that the fake-τ contributions to the background in the signal region can be modelled using $\mathrm{CR}_{\mathrm{fake}}$.

8.3.1 Alternate Estimation of the $Z(\to \nu\nu)$ + jets Background from $Z(\to \mu\mu)$ + jets Events in Data

In order to better understand the $Z(\to \nu\nu)$ + jets background, which for the final results is taken directly from simulation, an attempt is made to produce a data-driven cross-check. This is done by selecting a control region enriched in $Z(\to \mu\mu)$ + jets events. In these events the selected muons are removed and their transverse momenta are added vectorially to $E_{\mathrm{T}}^{\mathrm{miss}}$ to mimic the decay of a Z boson to neutrinos. The event sample thus obtained should provide a good model of $Z(\to \nu\nu)$ + jets events, including the additional jets in the event. Two of the jets must satisfy the event selection requirements and a third must be misidentified as a hadronic τ lepton decay.

Events are selected with a single-muon trigger with a p_{T} threshold of 18 GeV and are required to contain two oppositely charged muon candidates (according to the definition in Chapter 4). The distribution of the invariant mass of the two selected muon candidates, $M_{\mu\mu}$, is shown in Fig. 8.8. As can be seen from this figure, the selected event sample in data seems to be well described by pure $Z(\to \mu\mu)$ + jets Monte Carlo. The most probable sources of contamination from other Standard Model processes are

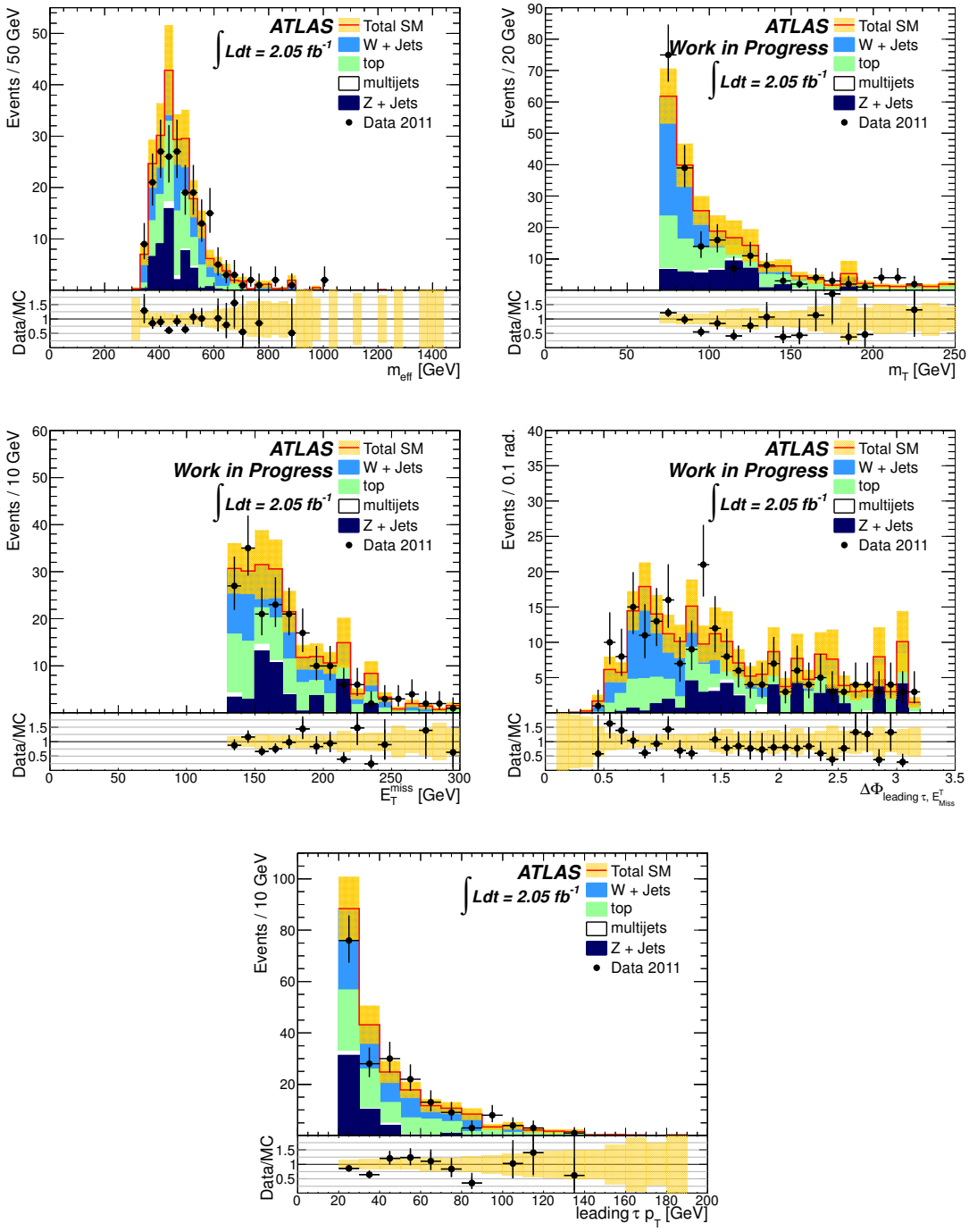


Figure 8.7: Kinematic distributions in CR_{fake} after applying both the W/top scale factors from CR_{true} and the combined fake-τ scale factor from CR_{fake}. The errors shown include Monte Carlo statistics and the statistical uncertainties on the W/top scale factors from the fit.

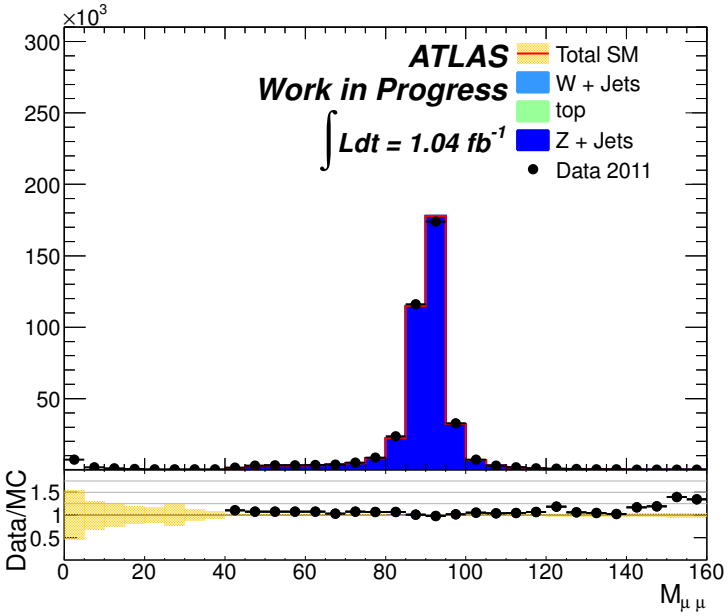


Figure 8.8: Dimuon invariant mass distribution for the $Z \to \mu^+\mu^-$ control region. The distribution is shown after selecting exactly two muons in the event with opposite charge. A requirement on the invariant mass of $75 < M_{\mu\mu} < 105$ GeV is applied to select a clean sample of $Z \to \mu^+\mu^-$ events.

negligible in a window around the Z-boson mass of $75 < M_{\mu\mu} < 105$ GeV. This mass window is selected to further reduce the remaining background processes, yielding a very pure sample of $Z(\to \mu\mu)$+jets events in data. One can then apply the standard event selection of the SUSY search to the remainder of events, treating the muons as if they had not been reconstructed and adding their transverse momenta to $E_{\mathrm{T}}^{\mathrm{miss}}$ vectorially. The number of observed events in 1.04 fb^{-1} of data and the number expected from Monte Carlo simulation are shown in Table 8.4.

With the current data sample, very few events pass the default event selection. After requiring two jets (with p_{T} thresholds of 130 and 30 GeV) and $E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV, the sample size has been reduced by a factor of 600. After requiring that the third jet be identified as a τ candidate, the sample is reduced by another two orders of magnitude, as would be expected for a sample where no real hadronic τ lepton decays are present and the identified τ candidates must be reconstructed from hadronic jets. The remaining sample contains 5 events and is too small to make any quantitative statements about the misidentification probability f_{ID}. None of the preceeding selection criteria can be relaxed since the kinematics of the jets need to match the nominal selection in order to avoid biasing f_{ID}. After applying the full selection, no events pass in data, and therefore no estimate of the $Z(\to \nu\nu)$ + jets background contribution in the final signal region is possible. While there are too few $Z(\to \mu\mu)$ + jets events to take advantage of this method at the moment, this method has promise for the near future. With the full 2011 data sample of 5 fb^{-1}, one would expect 25 events to pass the τ-identification step, which would make the statistical error on this number comparable to the total uncertainty on the MC-based estimate for the $Z(\to \nu\nu)$+jets contribution (see Chapter 9). With larger data samples, or complementing this study with a $Z \to e^+e^-$ control sample as well, this method should be able to improve upon the current estimates.

Even though this estimate is not competitive in terms of precision with the estimate derived from the simulation at the moment, it is still useful to attempt a comparison at earlier stages of the selection. In order to compare this selection to event yields for the nominal selection, one needs to apply a scale

	Data	Z	top	W
beforeCuts	75778407	$(4734.8 \pm 1.8)\times10^3$	159240 ± 160	$(3251.1 \pm 1.3)\times10^4$
event cleaning	40666337	868830 ± 520	25595 ± 63	$(5176.4 \pm 5.1)\times10^3$
$n_\mu = 2$	410592	405330 ± 360	932 ± 12	40.6 ± 9.8
opposite charge	410372	405330 ± 360	922 ± 12	26.8 ± 6.4
$75 > M_{\mu\mu} > 105$ GeV	362371	371650 ± 340	198.8 ± 5.8	5.2 ± 1.3
$p_{\mathrm{T}}^{\mathrm{jet1}} > 130$ GeV	2298	2664 ± 29	33.3 ± 2.4	$0 \pm$
$E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV	1337	1614 ± 23	10.9 ± 1.4	$0 \pm$
$p_{\mathrm{T}}^{\mathrm{jet2}} > 30$ GeV	577	686 ± 15	8.2 ± 1.3	$0 \pm$
$n_\tau > 0$	5	7.0 ± 1.5	0.3 ± 0.2	$0 \pm$
$E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$	4	6.7 ± 1.5	0.3 ± 0.2	$0 \pm$
$\Delta\phi$ cuts	4	6.5 ± 1.5	0.3 ± 0.2	$0 \pm$
$m_{\mathrm{T}} > 110$ GeV	1	2.0 ± 0.8	$0 \pm$	$0 \pm$
$m_{\mathrm{eff}} > 600$ GeV	0	1.2 ± 0.7	$0 \pm$	$0 \pm$

Table 8.4: Number of observed events in 1.04 fb^{-1} of data and numbers expected from Monte Carlo simulation for the $Z(\rightarrow \nu\nu)$ + jets background study. After reconstructing a good $Z(\rightarrow \mu\mu)$ + jets candidate, very few events pass the remaining event selection. Especially the requirements on the two leading jets and on τ identification reduce the event yields dramatically.

factor of about 2.0 for the luminosity (2.05 fb^{-1} in the nominal selection versus 1.04 fb^{-1} in this study), a factor of 6 for the branching ratio of $Z(\rightarrow \nu\nu)$ + jets (all three flavours) versus $Z(\rightarrow \mu\mu)$ + jets, and a factor of 2.6 to correct for the dimuon event selection efficiency (as determined from simulated $Z(\rightarrow \nu\nu)$ + jets events). Together, these give a scale factor of 31.2. If one applies this factor to the 5 selected $Z(\rightarrow \mu\mu)$ + jets events in data after the τ-identification step, this yields an estimate of 156 ± 70 selected $Z(\rightarrow \nu\nu)$ + jets events at the same stage in the nominal selection, where the relative error here is only $\sqrt{5}/5$. This is consistent with the expectation from $Z(\rightarrow \nu\nu)$ + jets simulation in the nominal selection: 134.3 ± 5.9 events.

8.4 QCD Multi-Jet Background Estimation

Two complementary approaches to estimate the event yield for the QCD multi-jet background are presented in this section. Because the number of real τ leptons produced in QCD multi-jet events is negligible, the observed background yield is very much dependent on the probability f_{ID} for a hadronic jet to be misidentified as a hadronic τ lepton decay, as discussed in Chapter 6. This background cannot be reliably predicted from simulation alone because the production cross section has a large uncertainty and because most of the missing transverse energy originates from instrumental effects. For these reasons, it is preferable to determine the QCD background from data rather than relying on simulation.

QCD-enhanced control regions are defined by inverting one or more of the QCD-suppression cuts discussed in Sec. 7.2.4. These control regions can be used to normalise the QCD predictions based on simulation and to estimate the τ misidentification probability f_{ID} for jets. In order to not rely only on estimates of f_{ID} for QCD jets from simulation, the dedicated measurements of the misidentification probability for hadronic jets to be identified as hadronic τ-lepton decays from Chapter 6 are applied to the simulated samples in the one approach which still relies partly on simulation. Because this background contributes less than one event after full event selection the two approaches used to estimate

it are only summarised here. The full details are given in Appendix A and B.

- First Approach: Simulated QCD events are used, in which all jets that are reconstructed as τ candidates are considered, regardless of whether they passed identification or not. One or more of these candidates are selected as identified τ candidates at random, taking the misidentification probability for each candidate, measured as function of τ-candidate p_{T} as detailed in Chapter 6, into account. Just like in the reweighting method for the $Z(\rightarrow \nu\nu)$ + jets background, described in Sec. 8.2, the events are weighted for each τ candidate in the event, with f_{ID} if the candidate was selected and with $1 - f_{\mathrm{ID}}$ if it was not. This allows more events to pass the event selection (although with weights less than 1) and therefore leads to much smaller statistical fluctuations than unweighted Monte Carlo samples. The absolute normalisation of this sample is taken from a QCD control region in data.

- Second Approach: A selection of events in data is used in which the τ candidate is chosen randomly from the jets in the event. In this approach the full kinematic event selection used in the main analysis can be applied to data events, in particular, the m_{T} cut which suppresses a large part of the true-τ lepton background. At the same time, the random selection yields a sample with a wide variety of fake-τ candidates without bias towards large BDT_j scores. To estimate the number of QCD events in the signal region, the data are divided into four regions based on two discriminants: the BDT_j score of the τ candidate and the differences in azimuthal angle between the missing transverse energy vector and the leading and the second-leading jet $\Delta\phi(\mathrm{jet}_1, E_{\mathrm{T}}^{\mathrm{miss}})$ and $\Delta\phi(\mathrm{jet}_2, E_{\mathrm{T}}^{\mathrm{miss}})$. The definition of the four resulting regions in the data is illustrated in Table 8.5. The two regions with low angular separation between the two leading jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ are completely dominated by QCD multi-jet events. In this sample the ratio of events with (region B) and without (region A) an identified τ candidate can be determined. This ratio is used to scale the data sample that passes $\Delta\phi(\mathrm{jet}_1, E_{\mathrm{T}}^{\mathrm{miss}})$ and $\Delta\phi(\mathrm{jet}_2, E_{\mathrm{T}}^{\mathrm{miss}})$ requirements but without an identified τ candidate (region C), which is again dominated by QCD multi-jet events. With this method, hereafter denoted as the "random-τ method", an estimate of the QCD multi-jet contribution in the signal region is obtained in a data-driven way, which only relies on simulation to predict the small contributions from other SM processes that need to be subtracted from the data in region A, B and C.

	fail τID	pass τID
	$N_{trk} < 6, \lvert Q\rvert < 3$	$N_{trk} = 1$ or $3, \lvert Q\rvert = 1$
Small $\Delta\phi$ (inverted cuts)	Control region A	Control region B
Large $\Delta\phi$ (nominal cuts)	Control region C	Signal region D

Table 8.5: Definition of signal and control regions for the random-τ background estimation method.

The two methods produce consistent results and the final background estimate of 0.5 ± 0.6 events from QCD multi-jet production is based on the random-τ method. The event weighting method is used as a crosscheck and to smooth out statistical fluctations when estimating systematic uncertainties and to model the distribution of QCD events in the figures.

Cut	top	W + jets	Z + jets	QCD	$\sum_{SM}$
$m_{eff} > 600$ GeV	5.6 ± 0.8	4.7 ± 1.3	2.4 ± 0.6	0.5 ± 0.6	13.2 ± 1.7

Table 8.6: Numbers of events estimated for Standard Model backgrounds with an integrated luminosity of 2.05 fb^{-1}. The uncertainties quoted for the W, Z and top samples are the combination of the statistical uncertainties of the simulated samples and of the scale factors determined in CR_{true} and CR_{fake}. The quoted uncertainty on the contribution from QCD multi-jet events is the full uncertainty of the data-driven method as explained in Appendix A.

8.5 Summary of Background Estimates

The numbers of events estimated for Standard Model backgrounds in the signal region with an integrated luminosity of 2.05 fb^{-1} are summarized in Table 8.6. The estimates for the W+jets and top backgrounds are based on simulated samples with the scale factor from CR_{true} applied for all events and the scale factor from CR_{fake} applied in addition, if the identified τ candidate in the event is not truth-matched. The expected background yields for Z + jets are based on simulated events reweighted with the efficiency of the τ identification as determined from data and scaled with the fake-τ scale factor from CR_{fake}. The QCD multi-jet contribution is estimated with the random-τ method described in Sec. 8.4. The total background estimate is significantly reduced compared to the estimate derived purely from simulation (see Table 7.2). This is mainly due to the low fake-τ scale factor from CR_{fake} of 0.5, since a large part of the background events is expected to contain only fake-τ candidates. The top sample has a real τ purity of about 35%, while the W + jets sample has a real τ purity of about 50% in the signal region; the Z + jets and QCD samples are completely dominated by fake-τ candidates. The expected yields for GMSB(30,20) and GMSB(40,30) signal scenarios would be 47.8 and 9.1 events, respectively.

Chapter 9

Systematic Uncertainties

After deriving all the necessary methods to estimate the background contributions in the final signal region, it is necessary to estimate the uncertainties on these background estimates. In this chapter both the systematic uncertainties on the background estimates, as well as the uncertainties on the observables used in the event selection are described. The list of quantities used as inputs for the analysis, which are subject to uncertainties, includes:

- jet, electron, muon and τ energy scale
- jet, electron, muon and τ energy resolution
- $E_{\mathrm{T}}^{\mathrm{miss}}$
- amount of pile-up
- trigger efficiencies
- electron, muon and τ identification efficiency
- the rejection against background provided by electron, muon and τ identification
- signal and background cross sections from Monte Carlo predictions
- luminosity

The following sections give detailed descriptions of each individual systematic uncertainty.

9.1 Jet Energy Scale

The energy measurement of hadronic jets is not infinitely precise. Therefore the uncertainty on the jet energy scale (JES) was determined by the Jet/Etmiss performance group of the ATLAS collaboration in MC studies where nominal results are compared to samples with varied hadronic shower and physics models, alternative detector configurations and by a data vs. MC comparison of the jet response as function of η [112]. Additional corrections are taken into account for close-by jets as described in [113]. These corrections are applied to all jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$, matching the acceptance for jets in computing $E_{\mathrm{T}}^{\mathrm{miss}}$.

The JES uncertainty is based on data and Monte Carlo studies for the 2010 dataset. Hence additional systematic Jet Energy Scale uncertainties are considered for the changed pileup conditions in the 2011 data taking. Those numbers are taken from initial studies on data taken in 2011. Their p_{T}- and η-dependent values range from 2% to 7% and are added in quadrature to the values obtained for the basic and close-by jet uncertainties [114].

Changes in the jet energies are propagated to $E_{\mathrm{T}}^{\mathrm{miss}}$ by the following relation:

$$E_{x,y}^{\mathrm{miss,new}} = E_{x,y}^{\mathrm{miss,old}} + \sum_{\mathrm{jets}} p_{x,y}^{\mathrm{old}} - \sum_{\mathrm{jets}} p_{x,y}^{\mathrm{new}} \,. \tag{9.1}$$

After modifying the jet transverse momenta and propagating these changes to $E_{\mathrm{T}}^{\mathrm{miss}}$, the whole event selection is redone on the modified event leading to a different selection efficiency. The correction is applied once in the positive direction, scaling all jet momenta up simultaneously, and once in the negative direction, scaling all momenta down. The JES uncertainty in the central calorimeter region ($|\eta| < 0.8$) is lower than 2.5% for jets with transverse momenta in the range $60 < p_{\mathrm{T}} < 800\,\mathrm{GeV}$, and less than 4.6% for the full p_{T} range above 20 GeV. In the endcap ($0.8 < |\eta| < 2.8$) and forward ($2.8 < |\eta| < 4.5$) regions, the uncertainty for jets with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ is below 4% and 6% respectively. The JES uncertainty is the largest for jets in the range $20 < p_{\mathrm{T}} < 30\,\mathrm{GeV}$ in the most forward region $3.2 < |\eta| < 4.5$, where it amounts to 14% [112].

These corrections are applied before the jet-τ overlap removal step, so that this systematic uncertainty is also applied to the jet objects corresponding to identified τ candidates. This is appropriate since the definition of $E_{\mathrm{T}}^{\mathrm{miss}}$ used in this analysis treats hadronic τ candidates as jets.

The systematic uncertainty due to these variations is determined by the difference of the expected number of events in the signal region with and without varying the jet energy scale.

9.2 Jet Energy Resolution

The jet energy resolution is only simulated with finite precision by the GEANT4 detector simulation. The agreement between the jet energy resolution in data and simulation has been studied using the spread of the p_{T} imbalance in dijet events [115] and with different in-situ techniques [116]. Deviations have been found to be on the order of 10%. The deviation has been evaluated in bins of jet p_{T} and η. All jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$ are smeared randomly and in an uncorrelated way with a Gaussian function of mean 1 and a standard deviation computed from these estimated resolution values.

The modified jet energies are propagated to $E_{\mathrm{T}}^{\mathrm{miss}}$ following Eq. 9.1. These corrections are applied before the jet-τ overlap removal, so that these uncertainties are applied to the τ candidates as well.

Again, the whole event selection procedure is redone using smeared values resulting in a different number of expected SM background events. The difference between event numbers with and without jet energy smearing is evaluated as systematic uncertainty on the background expectation.

9.3 Electron and Muon Energy/Momentum Scale and Resolution

The electron energy and muon momentum scales and resolutions are evaluated using the published set of uncertainties from [101] and [117], respectively. For both lepton types the momentum of the leptons in the simulation is smeared to match the position and width of the Z-boson peak observed in data. These uncertainties only enter in the event selection via the electron and muon vetos and the effects are found to be negligible ($< 10^{-3}$ relative difference) for all simulated samples.

9.4 Missing Transverse Energy

The missing transverse energy is affected by uncertainties in the energy measurement of objects in the ATLAS detector. The variations of the jet and lepton energies described above are therefore propag-

ated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation as described in the respective sections. Those contributions are the main uncertainties on the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement.

Additionally, activity associated with calorimeter energies outside identified physical objects are taken into account [118]. These corrections are applied to all calorimeter clusters with $|\eta| < 4.5$.

9.5 Tau Energy Scale

The systematic uncertainty due to the energy scale of hadronic τ-lepton decays, as detailed in [91], has been considered. The uncertainty is dependent on p_{T}, η and the number of reconstructed tracks associated with the τ candidate and ranges from 3.5% to 9.5%.

The systematic effect of the τ energy scale is evaluated in the same way as described for the jet energy scale. Scaling of the energies of the τ candidates is not propagated to the missing transverse energy since the definition of $E_{\mathrm{T}}^{\mathrm{miss}}$ used for this analysis does not include a term for τ candidates. Because all selected τ candidates are seeded by jets (see Chapter 5), the influence of varying energy scales on the $E_{\mathrm{T}}^{\mathrm{miss}}$ computation is already covered by the jet energy scale variations.

9.6 Pile-up Influence

By reweighting the simulation to reproduce the pile-up conditions seen in data the overall number of events (i.e. the sum of all event weights) changes slightly, depending on the MC dataset. This effect has been studied and found to be in general smaller than 0.5% .

9.7 Electron and Muon Identification Efficiencies

The muon reconstruction efficiency uncertainty has been obtained using a tag-and-probe method with a sample of $Z \to \mu\mu$ events from 2010 data [119], and the corrections applied to the simulated samples in this analysis represent an update with 2011 data. Similar studies were performed on a sample $Z \to ee$ events from 2010 and 2011 data to determine the electron reconstruction efficiency and its uncertainties. The systematic uncertainties on these efficiencies only enter in the SUSY analysis via the electron and muon vetos, and the effects are found to be negligible for all MC background samples.

9.8 Tau Efficiency and Misidentification Probability

Systematic uncertainties on the τ identification efficiency and on f_{ID} for the 2011 dataset have already been described in detail in Chapters 5 and 6. In general, they depend on the τ-identification algorithm, the kinematics of the τ candidates, and the number of associated tracks. These uncertainties vary from 4.5% to 12%. To evaluate the resulting systematic effects on the analysis, the simulated samples are reweighted, to increase or decrease the τ efficiency by the measured uncertainties. When performing this efficiency reweighting, only truth-matched τ candidates are considered. It is important to note that this uncertainty only applies to signal MC samples; for background samples, the normalisation of events with real τ leptons is determined directly from data.

For the misidentification probabilities for hadronic jets, the measurements presented in Chapter 6 are applied to the QCD-multi-jet and $Z(\to \nu\nu)$ + jets samples. For the other backgrounds and the signal MC samples the uncertainty measured in the fake-τ-dominated control region ($\mathrm{CR}_{\mathrm{fake}}$) in data is used. As with the τ efficiency, the systematic effects are evaluated by weighting events up or down by the

measured uncertainty. However, in the case of the fake-τ uncertainty, the leading τ candidate in the event is required to be *not* truth matched.

9.9 Trigger Efficiency

The effect of the trigger efficiency with respect to the offline event selection has been studied, see Sec. 7.1. As stated there, the trigger selection applied in this analysis can be considered fully efficient above the jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ thresholds applied in the offline event selection and the uncertainty on this efficiency is considered small enough to be neglected.

9.10 Luminosity

The total luminosity of the data runs used is determined by the use of the Good Runs List (GRL) to select only data fulfilling all quality requirements for a given analysis. Its uncertainty depends on the total uncertainty of the Luminosity measurement in ATLAS which is studied by the Luminosity Working Group. For the 2011 data this uncertainty is determined to be 3.4%[82].

9.11 Theory and Generator Uncertainties

For electroweak backgrounds, the same methods for the theory and generator uncertainties as in a search for SUSY in final states with jets, $E_{\mathrm{T}}^{\mathrm{miss}}$ and b-jets [10] are applied, which was published concurrently with this analysis. For W + jets MC samples, the generator uncertainties are estimated by varying the cross section of each sample up and down by 25% times the number of associated hard partons. A similar procedure is applied to Z + jets samples, but the results are found to be negligible. For top samples, the signal-region acceptance is estimated at truth level for a variety of samples, changing the shower-matching scales, the factorisation scale, and the α_S reweighting scale.

For signal MC samples, parton distribution function, α_S, and renormalisation scale uncertainties are included.

9.12 Systematic Uncertainties on QCD Multi-jet Background

Since the final QCD multi-jet background estimation is based on the random-τ method, the statistical uncertainties in the control regions in data are included in the uncertainty on this background estimate. Because the random-τ method relies on subtraction of the W and top yields, the statistical and systematic uncertainties on the normalisation of these backgrounds in each region of the random-τ method is added to the statistical uncertainty of the data. The subtraction of simulated W and top events with fake-τ leptons is in fact the dominant uncertainty on the final estimate of the QCD-multi-jet background.

9.13 Systematic Uncertainties on W, top, and Z Backgrounds

The uncertainties on the W, top, and Z backgrounds are estimated by the variation of the transfer factors, the ratio of the number of events between the signal region and the control regions. This is done because many of the systematic uncertaintes are correlated between the two regions and, since the control regions are used to estimate the backgrounds, the resulting systematic effects partially cancel. For

Systematic Uncertainty [%]	W	top
b-tag efficiency up	-3.7	10.5
b-tag efficiency down	4.0	-11.8
Jet energy resolution	0.1	0.4
Jet energy scale up	-1.4	5.7
Jet Energy scale down	-1.2	-2.8
Tau energy scale up	-0.5	-0.2
Tau energy scale down	0.4	-1.7
cluster energy scale up	0.5	-1.3
cluster energy scale down	-0.2	1.8

Table 9.1: Relative deviations on the scale factors measured for the W and top backgrounds in $\mathrm{CR_{true}}$ with different systematic variations.

the true-τ contribution to the W and top backgrounds, the changes in the ratio $N^{\mathrm{SR}}/N^{\mathrm{CR_{true}}}$ for Monte Carlo simulation are taken as the systematic uncertainty; for the combined fake-τ contribution, the ratio $N^{\mathrm{SR}}/N^{\mathrm{CR_{fake}}}$ is taken instead.

Systematic uncertainties on the W/top fit in $\mathrm{CR_{true}}$ are also evaluated by varying the jet energy scale and resolution, the τ energy scale, and the b-tagging efficiency up and down for the Monte Carlo samples and deriving new fit templates for the W/top-separation BDT. The fit to data is then repeated with the modified templates and the resulting differences in the scale factors are taken as a systematic uncertainty. These are summarised in Table 9.1.

When combining these uncertainties, the uncertainties due to the jet and τ energy scales are treated as fully correlated since many of the underlying sources of uncertainty (such as uncertainties in the detector material model) are common between the two. This is a conservative choice as these two variations shift the transfer factors in the same direction and treating them as correlated yields a larger uncertainty than assuming them to be uncorrelated.

An additional uncertainty on the fake-τ background determination, as discussed in Sec. 8.3, is included.

9.14 Summary of Systematic Uncertainties

Table 9.2 summarises the systematic uncertainties after the final m_{eff} cut for all SM backgrounds and for two GMSB signal points.

Systematic Uncertainty [%]	GMSB(30,20)	GMSB(40,30)	top	W	Z	QCD
Jet Energy Scale	5.2	2.1	5.5	3.9	3.7	–
Jet Energy Resolution	10.0	4.5	6.2	9.3	6.1	–
τ Energy Scale	1.6	3.4	3.8	6.0	9.8	–
τ Efficiency	11.6	11.9	0	0	0	–
Cluster energy scale	0.9	0.2	1.0	1.9	2.6	–
W/top fit stat.	–	–	9.0	4.8	–	–
W/top fit syst.	–	–	10.9	3.5	–	–
CR_{fake} stat.	0.5	0.2	9.7	8.0	15.4	–
CR_{fake} extrapolation to SR	–	–	1.9	1.9	1.9	–
random-τ method	–	–	–	–	–	120
Total Theory	8.6	10.2	6.3	4.7	–	–
MC stat.	7.0	7.1	9.2	26.7	21.2	–
Total	20.1	18.6	23.4	32.0	30.3	120

Table 9.2: Systematic uncertainties. All numbers are given in percent after the last cut step for two example GMSB signal points and for the relevant Standard Model backgrounds. The total systematic line includes the JES-TES correlation and treats the true- and fake-τ contributions separately in W and top and so is not simply the sum in quadrature of the individual contributions.

Chapter 10

Results and Exclusion Limits

The numbers of events observed in data and the number of expected background events are summarised in Table 10.1 for the last two steps in the selection (m_{T} and m_{eff} req.). The expected background yields are based on control region studies described in Sec. 8. The W + jets and top estimates include the contributions of both real and fake τ leptons. The W + jets sample has a τ purity of about 35% and the top sample of 50% for these final two cuts. The Z + jets and QCD samples are completely dominated by fake-τ candidates. With 11 events observed in data and 13.2 ± 4.2 expected from SM processes, no significant excess of events beyond the Standard Model expectation is found. The expected yields for GMSB(30,20) and GMSB(40,30) signal scenarios would be 53.1 and 9.9 events after the m_{T} cut, respectively, and 47.8 and 9.1 events after the requirement on m_{eff}.

Because the background estimate for events with fake-τ candidates depends on a scale factor that is significantly different from one, an additional cross-check is performed to validate the extrapolation from the fake-τ-dominated control region, $\mathrm{CR}_{\mathrm{fake}}$, to the signal region, even though no deviation from the background estimates is observed. Fig. 10.1 shows distributions of the τ-candidate p_{T} and the number of tracks, for both $\mathrm{CR}_{\mathrm{fake}}$ and the signal region. Since the τ-misidentification probability depends strongly on both of these variables, the fact that the two sets of distributions are similar and that the data in both regions are well described by the simulation suggests that the extrapolation is valid. Several kinematic distributions are shown in Fig. 10.2 after the m_{T} requirement. Good agreement between data and expected backgrounds is seen. Additional distributions can be found in Appendix C.

As no excess is seen beyond the expected Standard Model backgrounds, exclusion limits on the studied GMSB signal scenarios are calculated. As discussed in Sec. 2.2.4 limits are set in the $\Lambda - \tan\beta$ plane, where the other parameters are constrained to the following values: $M_{\mathrm{mess}} = 250\,\mathrm{TeV}$, $N_5 = 3$, $\mathrm{sign}(\mu) = +$ and $C_{\mathrm{grav}} = 1$. The event selection used in this analysis is especially sensitive to these signal scenarios, because of the large branching ratio to τ leptons for large parts of the parameter space. Studies of the sensitivity of this analysis to mSUGRA/CMSSM scenarios, also for large values of $\tan\beta$, have shown that the fully hadronic inclusive searches performed at ATLAS [7, 8] have a higher

Cut	Data	top	W + jets	Z + jets	QCD	$\sum_{\mathrm{SM}}$
$m_{\mathrm{T}} > 110$ GeV	64	32.4 ± 7.4	20.3 ± 4.2	13.8 ± 4.9	2.9 ± 2.9	69.4 ± 24.3
$m_{\mathrm{eff}} > 600$ GeV	11	5.6 ± 1.4	4.7 ± 1.5	2.4 ± 0.7	0.5 ± 0.6	13.2 ± 4.2

Table 10.1: Numbers of events observed in data and estimated Standard Model backgrounds with an integrated luminosity of 2.05 fb^{-1}. The expected background yields for Z + jets are based on simulated events, while those for the other backgrounds are based on control region studies described in Sec. 8. The W + jets and top estimates include the contributions of both real and fake τ leptons, where the top sample has a real τ purity of about 35%, while the W + jets sample has a real τ purity of about 50% for the two cuts given here; the Z + jets and QCD samples are completely dominated by fake τ candidates.

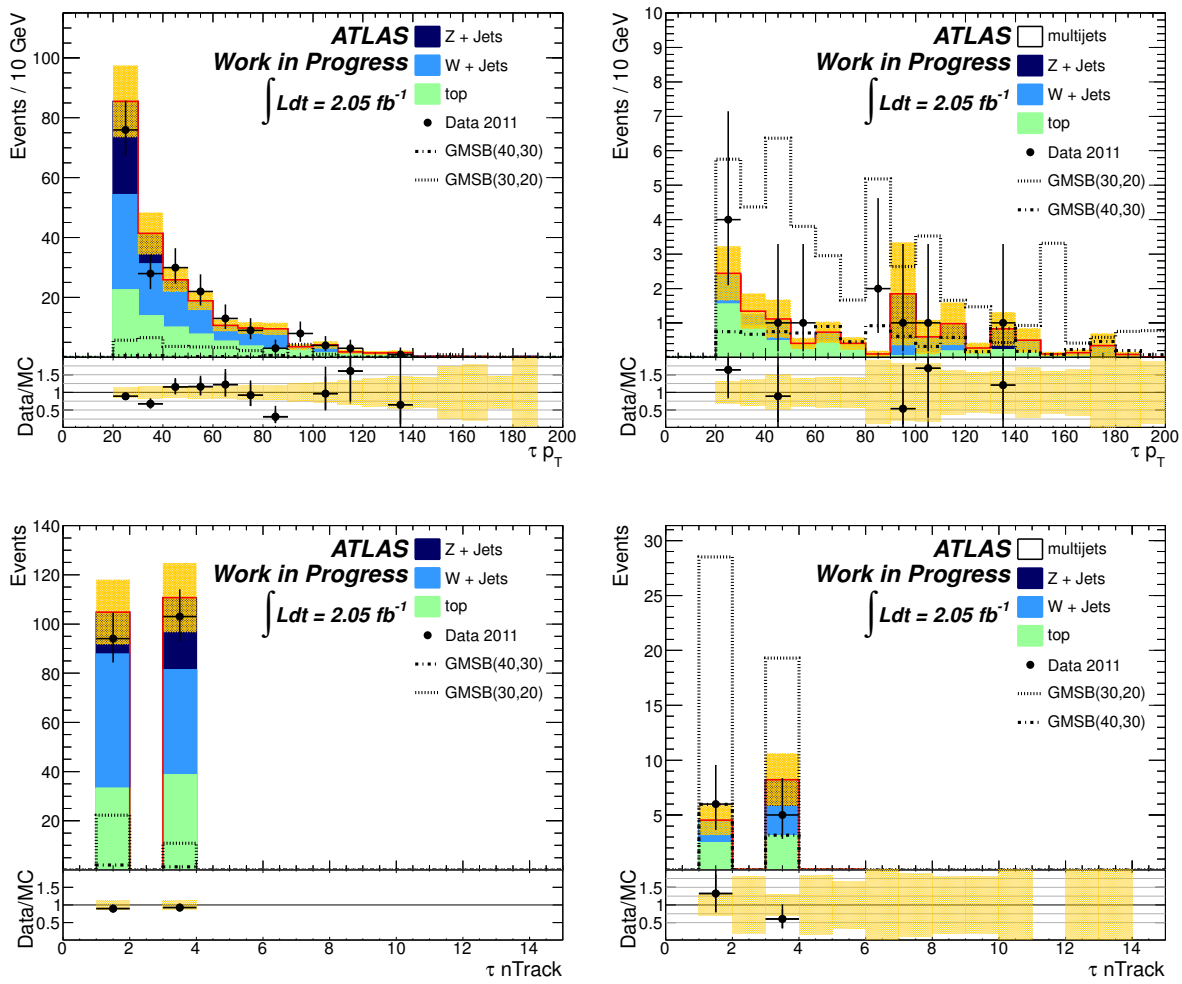


Figure 10.1: Distributions of the τ candidate p_{T} and number of tracks for $\mathrm{CR}_{\mathrm{fake}}$ (left) and the signal region (right).

sensitivity. Therefore no interpretation of the search with hadronic τ-lepton decays was performed in mSUGRA/CMSSM scenarios. The method used to compute exclusion limits on the studied GMSB models is based on a profile log-likelihood ratio test. The likelihood function is written as a product of a Poisson distribution and a probability density function:

$$L(n_S|\mu, b, \theta) = P(n_S|\lambda_S(\mu, b, \theta)) \times P_{\text{Syst}}(\theta^0, \theta), \tag{10.1}$$

where n_S is the number of observed events in data and λ_S the number of expected events. The number of expected events depends on the parameter μ called "signal strength" of the SUSY model. The case $\mu = 0$ corresponds to a model with no signal contribution (background only), while $\mu = 1$ corresponds to the sum of the expected number of SUSY events and background. The other parameters are the background normalisation factors b and the nuisance parameters θ that parametrise the systematic uncertainties. The probability density function $P_{\text{Syst}}(\theta^0, \theta)$ includes the systematic uncertainties, where θ^0 are the nominal values around which θ are varied to maximise the likelihood.

The complete method is described in detail in [121]. It is also used by the ATLAS collaboration for other SUSY publications [7]. Due to the recommendations of the ATLAS statistics forum the CL_s method [122] is used as baseline instead of the CL_{s+b} method that uses the p-value p_{s+b}. The CL_s method artificially introduces a penalty factor of $(1-p_b)^{-1}$ that encodes the difference of the background-only p-value from unity. This has the advantage of limiting the exclusion power of an analysis where the background-only p-value is close to one, which corresponds to an analysis with low signal sensitivity.

The observed and expected 95% confidence level (C.L.) exclusion limits are computed and shown in Fig. 10.3. The regions of GMSB parameter space with different NLSPs are labelled by the name of the NLSP and separated by light grey lines. It can be clearly seen that the exclusion range of this analysis is larger in the $\tilde{\tau}_1$ NLSP region due to an enhanced branching fraction to τ leptons. Fig. 10.5 shows a comparison with other ATLAS searches that were also interpreted in GMSB signal scenarios. The exclusion range is shown to be comparable with the inclusive di-tau [12] search and significantly larger than for the search with light leptons [33] for high $\tan\beta$. This is due to the increased branching ratio to τ leptons for high values of $\tan\beta$, which offsets the higher selection efficiency for light leptons compared to τ leptons, but also in part due to the larger dataset analysed in the searches with hadronic τ-lepton decays. Searches for these signal scenarios were also performed with LEP data, where the most stringent limits come from the OPAL experiment [120]. The exclusion ranges of the OPAL searches are shown as colored bands in Fig. 10.3 and Fig. 10.5. It can be seen that the results presented here significantly improve on their sensitivity. In the signal models introduced in Sec. 2.4.2, the production of supersymmetric particles can be excluded up to $\Lambda = 30$ TeV, independent of $\tan\beta$, and up to $\Lambda = 43$ TeV for large values of $\tan\beta$. In addition, an upper limit of 8.5 on the number of events observed due to non-SM sources is also derived at the 95% confidence level. This limit corresponds to an upper limit on the visible cross section of 4.0 fb. The visible cross section is defined as:

$$\sigma_{vis} = \sigma_{prod} \times f_\tau \times a \times \epsilon, \tag{10.2}$$

where σ_{prod} is production cross section, f_τ is the branching fraction to at least one τ lepton, a is the acceptance, and ϵ is the efficiency using the event selection described above. The acceptance and efficiency are shown across the GMSB $\Lambda - \tan\beta$ plane in Fig. 10.4. The acceptance includes all fiducial and kinematic selection requirements and is defined at generator level. The efficiency is the final selection efficiency at reconstruction level, defined relative to the acceptance at generator level. Note that the efficiency may be larger than 100%, particularly when there are true, low-p_{T}, light leptons present. The inefficiency to reconstruct these leptons increases the total efficiency since light leptons are used in veto mode.

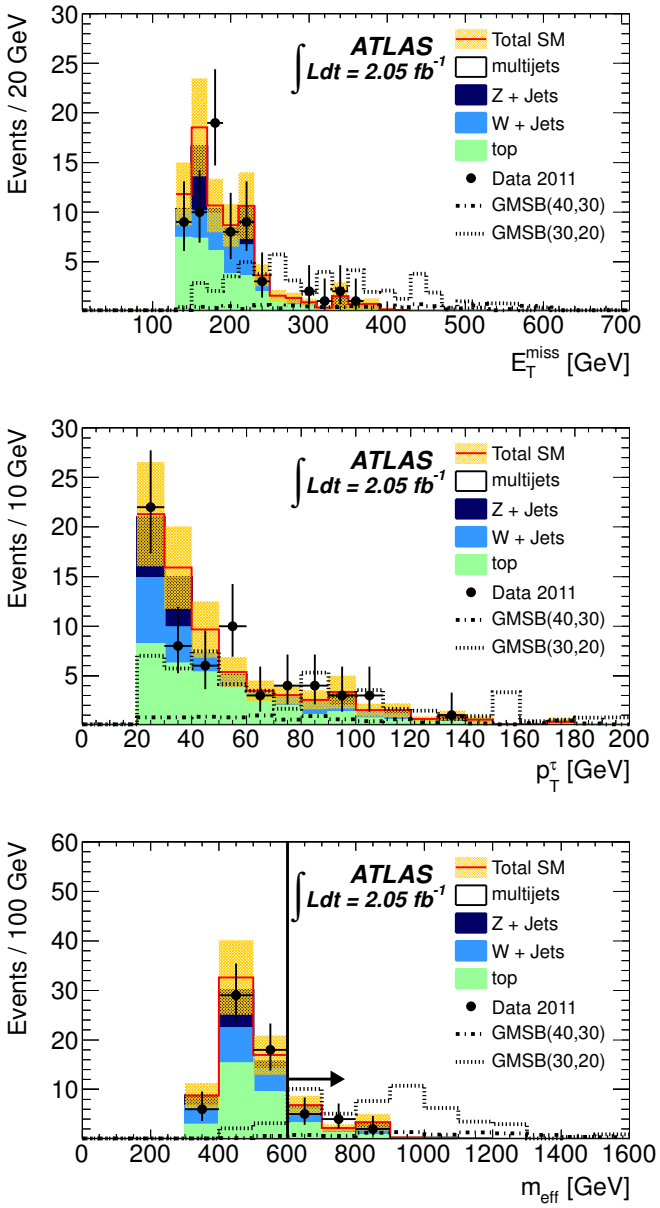


Figure 10.2: Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$, p_{T}^{τ}, and m_{eff} for data with all selection requirements except for that on m_{eff}, along with the corresponding estimated backgrounds. Backgrounds are taken from simulation and normalised using control regions in data. The solid (red) line with shaded (yellow) error band corresponds to the total SM prediction. The data are shown as solid markers. The error band indicates the size of the total (statistical and systematic) uncertainty.

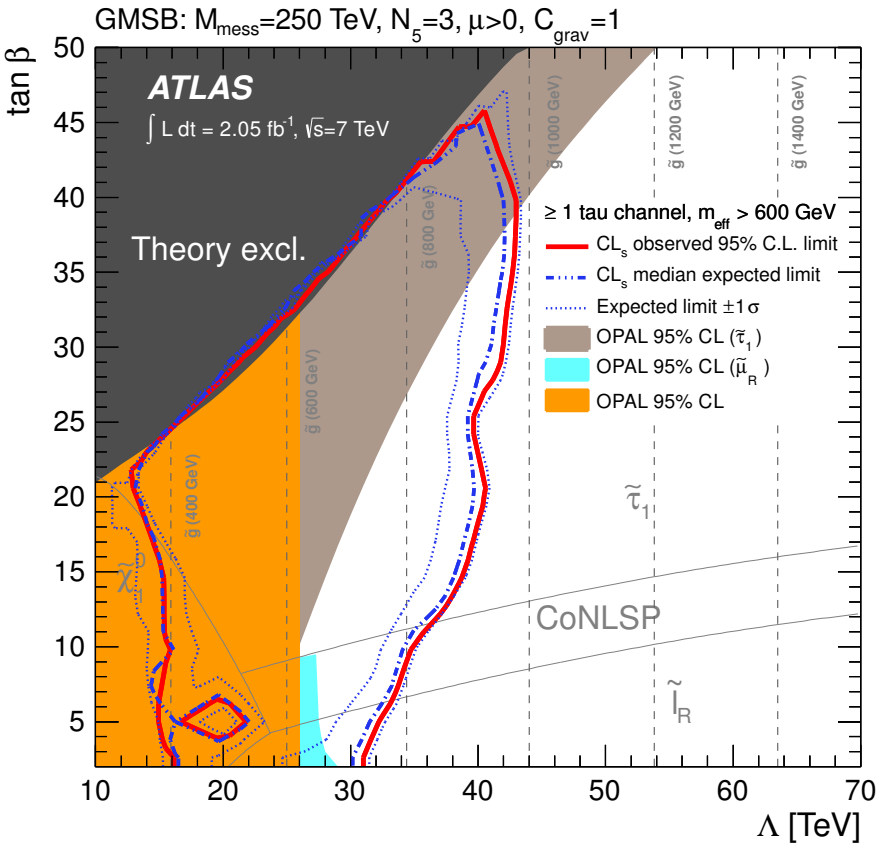


Figure 10.3: Expected and observed 95% C.L. exclusion limits in the $M_{\text{mess}} = 250$ TeV, $N_5 = 3$, $\mu > 0$, $C_{\text{grav}} = 1$ slice of GMSB, together with the most stringent previous limits from OPAL [120]. The identity of the NLSP is indicated, with CoNLSP the region where the $\tilde{\tau}$ and $\tilde{\ell}$ are nearly degenerate.

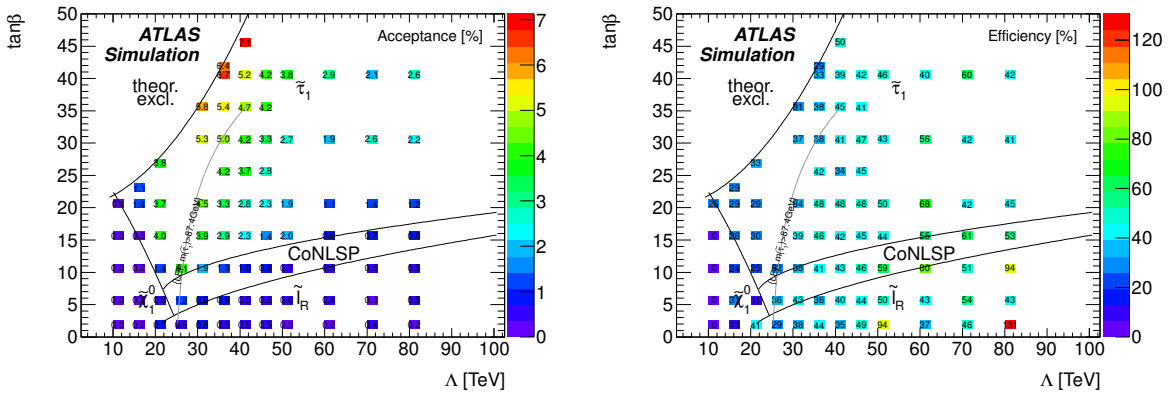


Figure 10.4: Acceptance (left) and efficiency (right) across the GMSB grid. The acceptance includes all fiducial and kinematic selection cuts and is defined at generator level. The efficiency is the final selection efficiency at reconstruction level, defined relative to the acceptance at generator level. Note that efficiency may be larger than 100%, particularly when there are true, low-p_{T}, light leptons present. The inefficiency to reconstruct these leptons increases the total efficiency since light leptons are used in veto mode.

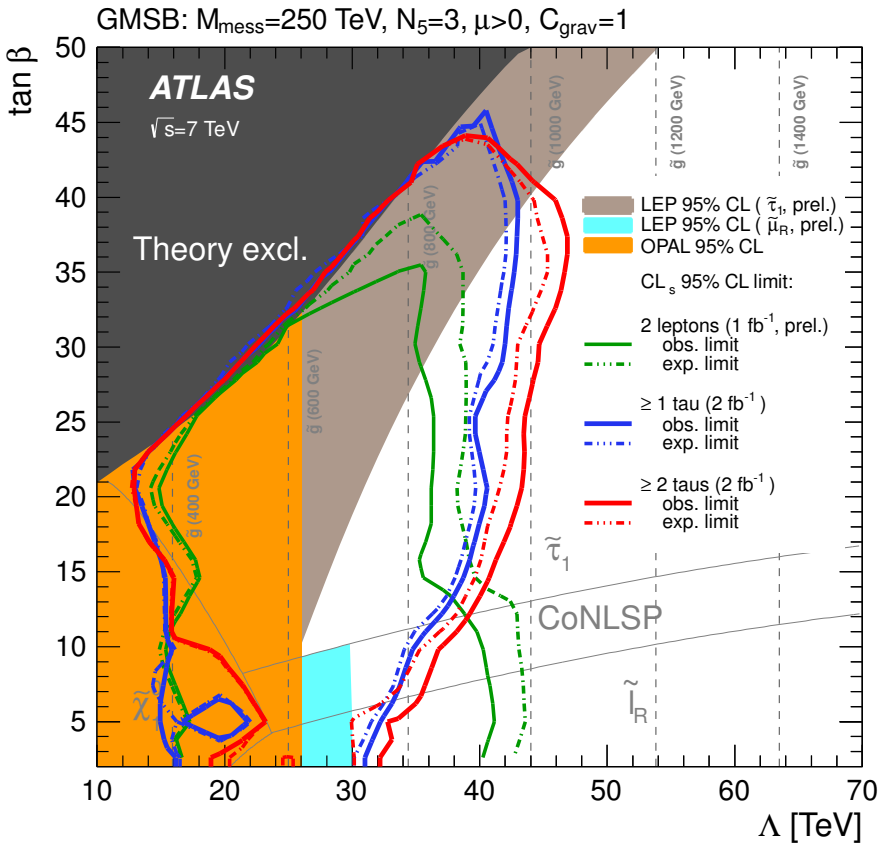


Figure 10.5: Comparison of expected and observed 95% C.L. exclusion limits in the $M_{\text{mess}} = 250$ TeV, $N_5 = 3$, $\mu > 0$, $C_{\text{grav}} = 1$ slice of GMSB for ATLAS searches with at least one or two τ leptons as well as a search with at least two light leptons in the final state. The most stringent previous limits from OPAL [120] are shown as colored areas. The identity of the NLSP is indicated, with CoNLSP the region where the $\tilde{\tau}$ and $\tilde{\ell}$ are nearly degenerate.

Chapter 11

Summary

At the ATLAS Experiment measurements with τ leptons contribute to measuring known Standard Model processes at the terascale as well as offering an interesting probe for physics beyond the Standard Model, although the separation of hadronic τ-lepton decays from jet and electron backgrounds poses a considerable challenge. This thesis presents studies of τ-lepton identification at the ATLAS Experiment as well as a search for physics beyond the Standard Model in τ-lepton final states.

To search for new physics in final states with hadronic τ-lepton decays a reliable and efficient reconstruction algorithm for hadronic τ decays is needed to separate real τ decays and backgrounds from quark- or gluon-initiated jets and electrons. Therefore the first part of this thesis documents work on the development and validation of methods to reconstruct and identify hadronic τ lepton decays. Two existing τ-reconstruction algorithms were further developed and integrated into a single algorithm that reconstructs hadronic τ-lepton decays using calorimter jets as seeds. This algorithm has since become the standard method of identifying hadronic τ lepton decays at the ATLAS Experiment. The suppression of jet background, one of the crucial aspects of τ reconstruction and the SUSY analysis in this thesis, was studied and the probabilities of misidentifying quark- or gluon-initiated jets as hadronic τ-lepton decays were measured in both the 2010 and 2011 ATLAS data at a centre-of-mass energy of $\sqrt{s} = 7$ TeV, using samples of di-jet events.

In the second part of this thesis a search for events with highly energetic jets, missing transverse energy and at least one τ lepton is presented. The τ lepton is reconstructed in the hadronic decay mode. The data used for this search amounts to $2.05\,\text{fb}^{-1}$ of proton-proton collisions recorded between March and August of 2011 at the ATLAS Experiment. This is the first published search at the LHC that includes events with exactly one hadronic τ-lepton decay. The Standard Model background processes that contribute most to the selected event sample are W- and Z-boson production in association with jets as well as top quark production. Eleven data events remain after the full event selection, which is in good agreement with the expectation from Standard Model background processes of 13.2 ± 4.2 events. While a search in τ-lepton final states is applicable to many scenarios for physics beyond the Standard Model, specific R-parity conserving Supersymmetry signal scenarios were investigated in this thesis. Like many other scenarios for new physics, these scenarios predict an excess of events with high-p_{T} jets and missing transverse momentum. Gauge Mediated SUSY Breaking (GMSB) scenarios are furthermore especially promising for searches in τ-lepton final states, because of the large number of leptons expected in the SUSY decay chains and an enhanced production of third-generation fermions over a large part of the paramter space. Exclusion limits are derived at 95% confidence level (C.L.) in the context of GMSB as a function of the GMSB parameters Λ and $\tan\beta$, for fixed values of the other GMSB parameters: $M_{\text{mess}} = 250\,\text{TeV}$, $N_5 = 3$, $\text{sign}(\mu) = +$ and $C_{\text{grav}} = 1$. A model-independent, upper limit on the visible cross section of 4.0 fb for non-Standard Model processes is derived at 95% C.L., where the visible cross section is defined as the product of production cross section of the non-Standard Model process, its branching fraction to at least one τ lepton, acceptance, and efficiency of the event selection applied. The exclusion range achieved in this thesis is similar to that from other ATLAS results

in the di-tau and di-lepton final states derived from 2011 datasets of similar size. The limits presented in this thesis also significantly extend previous limits placed by other experiments, among which those published by the OPAL experiment were the most stringent.

Chapter 12

Outlook

The methods for τ-lepton identification presented in this thesis are continually applied to new measurements of Standard Model processes and the search for SUSY or new particles decaying to τ leptons. Especially the discovery of a new bosonic particle by both the ATLAS [5] and CMS [6] collaborations in the summer of 2012 gives new urgency to searches for the decay of a Standard Model Higgs boson to τ leptons. For these searches and other measurements with τ leptons at the ATLAS experiment the uncertainties on the energy scale and identification efficiency of the τ-reconstruction algorithm presented in this thesis are being studied with ever larger samples of $Z \to \tau\tau$ and $W \to \tau\nu$ events in ATLAS data. With the increase in the available number of data events containing real τ leptons and refinements made to the τ-efficiency measurements presented in Chapter 5, the uncertainties on the τ-lepton identification can be decreased significantly below the levels of uncertainty assumed for the SUSY analysis in this thesis. The studies of the probability to misidentify jets as hadronic τ-lepton decays (f_{ID}), presented in Chapter 6, will also have to be performed again on the 2012 dataset, due to changes in the beam and data-taking conditions as well as improvements made to the τ-identification algorithms. It can be hoped that, by measuring the misidentification probability in several samples of jets with significantly different abundances of quark- and gluon-initiated jets, a process-independent parametrization of f_{ID} will be possible in the future.

Since the publication of the results from this thesis, both the ATLAS and CMS collaborations have published further results from searches for events with large missing transverse energy, jets and at least one τ lepton on the full 2011 LHC dataset at a centre-of-mass energy of $\sqrt{s} = 7$ TeV [123, 124]. In addition to the increase in the amount of available data, the ATLAS search includes a combination of the fully hadronic one- and two-tau channels with a search for events with one τ lepton and one electron or muon. This channel complements the sensitivity of these searches especially at low and intermediate values of $\tan\beta$. None of the further searches in τ-lepton final states with the $\sqrt{s} = 7$ TeV LHC data have observed an excess over Standard Model background expectation, but the limits on GMSB models have been extended.

Another increase in sensitivity to GMSB models can be expected when these search strategies are applied to the even larger 2012 LHC datasets of proton-proton collisions at the higher centre-of-mass energy of $\sqrt{s} = 8$ TeV. Further use can also be made of the results derived from both datasets, by interpreting them in other scenarios for physics beyond the Standard Model, especially those that predict an enhanced production of third-generation fermions.

Appendix A

QCD Multi-jet Background Estimate using f_{ID} Measured in Data

In this approach, instead of using the τ-identification decisions from the simulation, f_{ID} as measured in QCD di-jet topologies in data (see Sec. 6.5) is used as an event weight. Because the f_{ID} measurement as presented in the previous chapter only gives the probability for a τ candidate reconstructed from a hadronic jet to pass the BDT_j identification algorithm, the p_{T} and η acceptance requirements, the requirements for at least one reconstructed track and for the charge of the candidate to be $|q| == 1$ and the $e^{\pm}$-veto are all applied to the reconstructed τ candidates as a preselection. The f_{ID}-weighting, which only represents the efficiency for the BDT_j decision and the requirement for the candidates to have either one or three reconstructed tracks, can only be applied to candidates that pass this preselection. As explained in Sec. 6.5, f_{ID} measured in 2011 ATLAS data is only binned in τ-candidate p_{T}. While f_{ID} shows some dependence on $|\eta|$ and the number of prongs associated with the τ candidate, this dependence is integrated out to improve the statistical precision up to high values of p_{T} of the τ candidates. This is justified as the fake rate is nearly flat in $|\eta|$ (except for the calorimeter crack region), so binning in $|\eta|$ wound not significantly improve the precision of this result. Integrating over the number of prongs is also justified, as it is explicitly checked that the simulation provides a good description of the ratio of 1- to 3-prong τ candidates (see, for example, Fig. 10.1), so this simplification is expected to be safe to apply as well. The misidentification probabilities used in this analysis, including an acceptance cut $p_{\mathrm{T}} > 15$ GeV and multiplied with the efficiency for the requirement on $nTrack == 1 \,\|\, nTrack == 3$, are shown in Fig. A.1. The discrepancy between data and simulation seen in the first bin is one reason why this analysis is restricted to τ candidates with $p_{\mathrm{T}} > 20$ GeV.

The weights, which are applied to the simulated events, are calculated in the following way: the same preselection used for the determination of the misidentification probability is applied to all reconstructed τ candidates in the event. For this list of n candidates, with their respective fake rates $f_{\mathrm{ID,i}}$, all possible combinations of BDT_j pass or fail decisions are considered, and the absolute probability w for each combination is calculated as

$$w = \prod_{i=1}^{n} [f_{\mathrm{ID,i}}]_{\text{if candidate i passes}} \,, [1 - f_{\mathrm{ID,i}}]_{\text{if candidate i fails}} \,. \tag{A.1}$$

From all possible combinations in which at least one candidate passes the identification criteria, one is chosen randomly, according to their relative probability, and its absolute probability w is then applied as an event weight. While this method improves the available QCD statistics by the inverse of the average misidentification probability per fake τ candidate, this only ensures that the efficiency for the τ-identification step is correctly modelled. The overall normalisation, which is uncertain due to the large errors on the production cross section and possible mismodelling of the efficiency of the jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ requirements for the QCD multi-jet background still needs to be determined from data. To achieve this, the part of the event selection that is designed to reject QCD events is inverted: an event is accepted

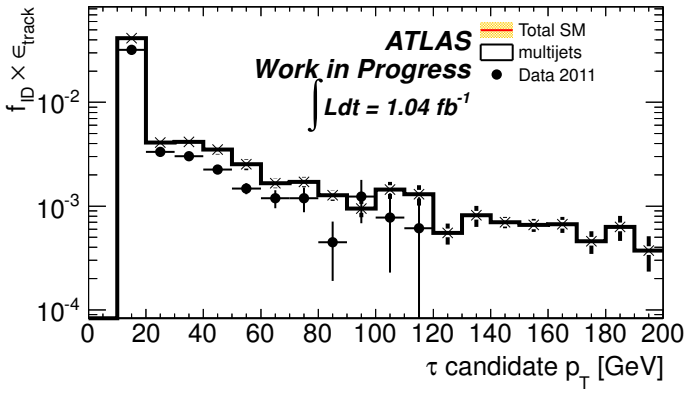


Figure A.1: Fake rates from QCD di-jet events for the BDT$_j$ working point used in this analysis. The bin covering 10–20 GeV only contains entries above the τ acceptance cut of 15 GeV. These τ candidates are not included in the analysis, which requires $p_{\mathrm{T}} > 20$ GeV.

	Data	top	W	Z	QCD
N^{CR}	319	33.8 ± 3.9	130.6± 8.2	31.3± 6.6	41.5±3.6

Table A.1: Number of observed and expected events in the control region for the QCD event-weighting method. The uncertainties given for Z and QCD are statistical only. For the W and top contributions, the uncertainties include both the Monte Carlo statistical and the statistical uncertainties on the scale factors derived in their respective control regions.

if it fails any of the nominal $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$, $\Delta\phi(\mathrm{jet}_1, E_{\mathrm{T}}^{\mathrm{miss}})$ and $\Delta\phi(\mathrm{jet}_2, E_{\mathrm{T}}^{\mathrm{miss}})$ cuts. A comparison of the data obtained by this selection and the reweighted QCD Monte Carlo sample is shown in Fig. A.2. For this comparison, the electroweak processes are already corrected for their normalisation as explained in Sec. 8.1 and Sec. 8.3. The shapes agree within the limited statistics of the MC sample, but the normalisation is clearly underestimated in the simulation. The numbers of events in both data and simulation are summarised in Table A.1. From these numbers, after subtracting off the non-QCD contributions, a QCD normalisation factor of 3.15 ± 0.95 is derived, where the error includes statistical uncertainties and uncertainties on the W and top subtraction discussed in Sec. 8.1 and Sec. 8.3. Fig. A.3 shows the same distributions as Fig. A.2 after the QCD scaling has been applied, improving the overall agreement.

Applying this scaling and the event weighting to the simulated QCD sample after the final signal selections yields

$$N_{\mathrm{weighed}}^{\mathrm{QCD}}\Big|_{m_{\mathrm{T}}>110\ \mathrm{GeV}} = 4.3 \pm 2.1 \tag{A.2}$$

and

$$N_{\mathrm{weighed}}^{\mathrm{QCD}}\Big|_{m_{\mathrm{T}}>110\ \mathrm{GeV},\ m_{\mathrm{eff}}>600\ \mathrm{GeV}} = 0.3 \pm 0.1\ . \tag{A.3}$$

As an additional crosscheck, this study is repeated using a different set of misidentification probabilities, derived from $Z(\rightarrow \mu\mu)$ + jets events instead of QCD di-jet events. These measured f_{ID} values, as presented in Sec. 6.6, are based on a very different mixture of quark and gluon jets, on which f_{ID}

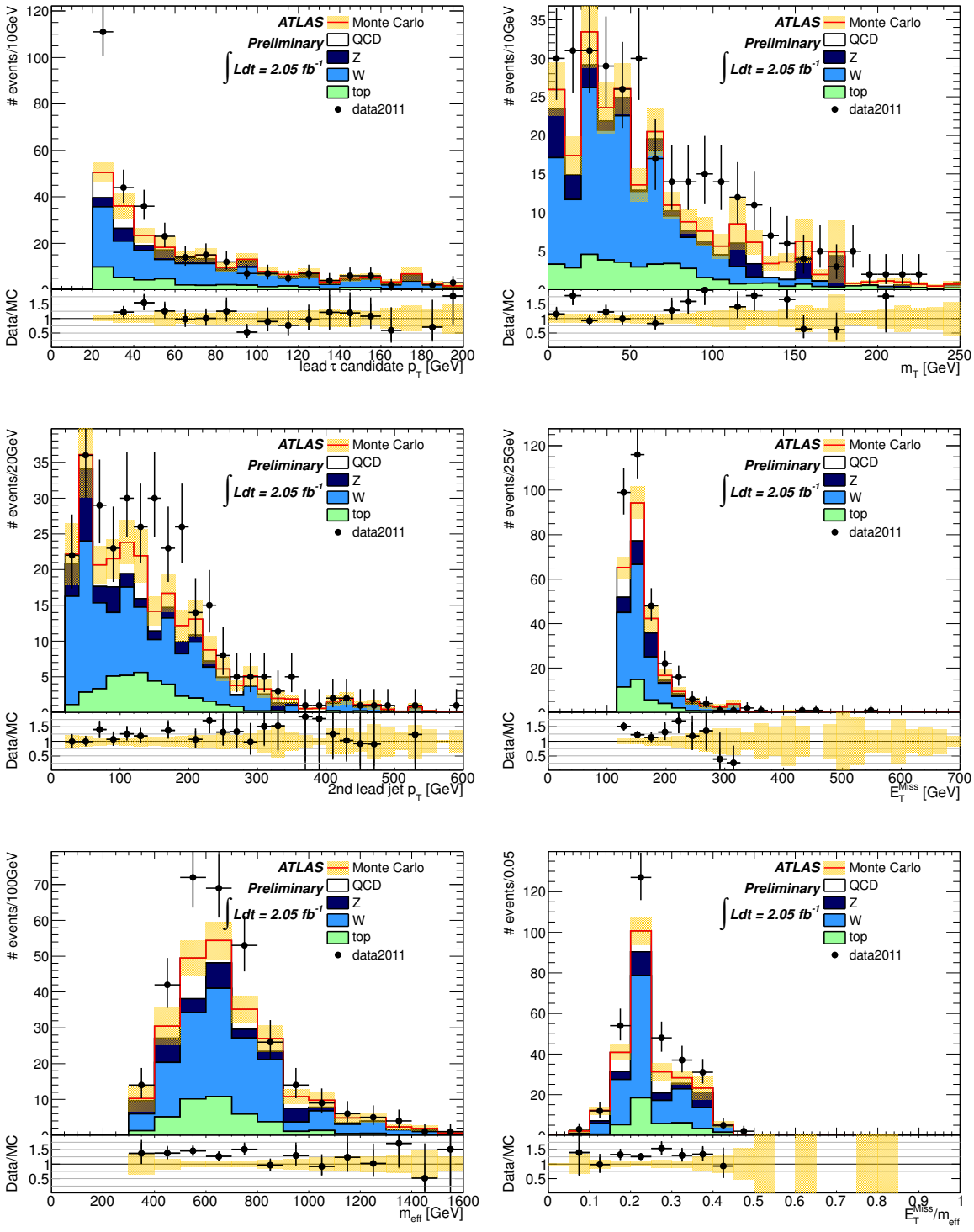


Figure A.2: Comparison of data and simulation in QCD control region for event-weighting method. Errors shown are statistical only.

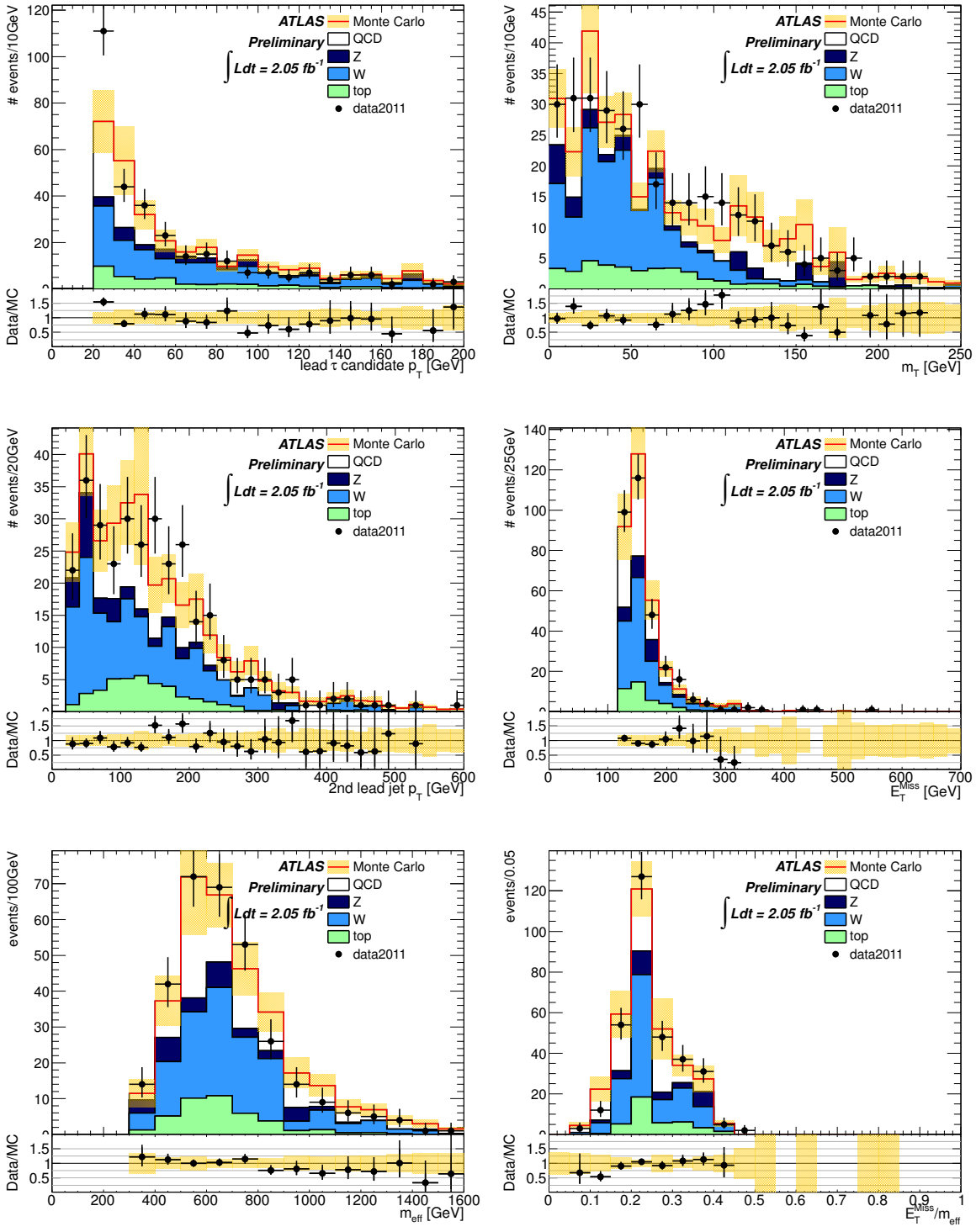


Figure A.3: Comparison of data and simulation in QCD control region for the event-weighting method. The QCD contribution is scaled such that the sum of QCD and other Standard Model processes is equal to the observed number of events in data. Errors shown are statistical only.

is known to depend strongly. Comparing the difference in predicted events in the signal region, due to using this other set of f_{ID} values, can therefore set a limit on possible variation in the QCD background estimate, due to the dependence of f_{ID} on the quark-gluon mixture in the selected sample of events. With these f_{ID} values one obtains, if the procedure above is repeated and extrapolated to the final selection criteria:

$$N'^{\mathrm{QCD}}_{\mathrm{weighed}}\Big|_{m_{\mathrm{T}}>110\ \mathrm{GeV}} = 4.5 \pm 1.9 \tag{A.4}$$

and

$$N'^{\mathrm{QCD}}_{\mathrm{weighed}}\Big|_{m_{\mathrm{T}}>110\ \mathrm{GeV},\ m_{\mathrm{eff}}>600\ \mathrm{GeV}} = 0.3 \pm 0.1\ . \tag{A.5}$$

Compared with Equations A.2 and A.3, respectively, one sees very good agreement.

The p_{T}^{τ} distribution in Fig. A.3 shows some disagreement between data and MC, either a deficit of data in the first bin or an excess of data in the second (because the overall normalisation is fixed to data, these two possibilities are indistinguishable). To check that this is consistent with a statistical fluctuation, the background estimation in the signal region is repeated while reweighting the p_{T}^{τ} spectrum of MC in the signal region by the ratio of distributions seen in Fig. A.3. This results in a 16% shift in the expected event yield, consistent with the expected precision from the limited number of simulated QCD events. This does not affect the agreement with the random-τ method for estimating the QCD multi-jet background, which is discussed in the next section. Since the primary QCD estimation is the random-τ method and this does not rely on the p_{T}^{τ} distribution in simulation, the final QCD estimate is safe and no systematic uncertainty is added due to the modelling of this distribution.

Appendix B

QCD Multi-jet Background Estimation using the Random-τ Method

An alternative approach to estimate the number of QCD events is given by applying the so-called random-τ method in a QCD-enriched control region. To this end, the τ candidate used for all steps of the event selection involving τ candidates is selected at random. To ensure a sensible choice of τ candidate, which has a higher chance of passing the τ-identification criteria afterwards, a very loose preselection defined by $N_{\text{track}} < 6$ and $|Q| < 3$ is applied.

After applying the event selection up to the QCD rejection cuts, the m_{T} and a looser $E_{\text{T}}^{\text{miss}}/m_{\text{eff}} > 0.1$ are applied. This leaves the requirements on $\Delta\phi(\text{jet}_1, E_{\text{T}}^{\text{miss}})$, $\Delta\phi(\text{jet}_2, E_{\text{T}}^{\text{miss}})$ and τ identification, which are used to separate the resulting data sample into four disjoint regions. The final cuts on N_{track}, $|Q|$, and BDT_j are applied to the randomly selected candidate to divide the sample, with Regions B and D passing 'tight' τ ID criteria and Regions A and C with 'loose-but-not-tight' criteria. Regions B and D (or A and C) are then divided from each other by the requirement that events in D (C) pass both $\Delta\phi(\text{jet}_1, E_{\text{T}}^{\text{miss}})$ and $\Delta\phi(\text{jet}_2, E_{\text{T}}^{\text{miss}})$ requirements, while events in A (B) fail at least one of these requirements. The definition of the four regions is summarised in Table B.1. Region D is the *signal region*, i.e., it corresponds to the nominal selection used in the analysis. Regions A, B, and C are fake-τ dominated *control regions*.

The remainder of the event selection follows the main signal selection as closely as possible, including the main kinematic cuts. With the currently available data, though, the sample sizes in the control regions, in particular region B, are still quite small. To deal with this, the final m_{eff} requirement is removed, and - as stated previously - the $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}$ requirement is relaxed in all four regions from the nominal cut $E_{\text{T}}^{\text{miss}}/m_{\text{eff}} > 0.25$ to $E_{\text{T}}^{\text{miss}}/m_{\text{eff}} > 0.1$. The final background estimate will be scaled to extrapolate from this $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}$ cut into the nominal signal region. While regions A and B are clearly dominated by QCD events, regions C and D have a sizeable contribution of events from W, $t\bar{t}$ production and $Z(\rightarrow \nu\nu)$ + jets. To obtain a real QCD estimate from this method, care must be taken to estimate these contributions. Because of the contribution from $Z(\rightarrow \nu\nu)$ + jets in region C, which can only be taken purely from the Monte Carlo estimate, an additional requirement on $E_{\text{T}}^{\text{miss}}/m_{\text{eff}} < 0.25$ is applied in region C, to decrease the contamination from W, top, and especially $Z(\rightarrow \nu\nu)$ + jets.

Table B.2 gives the numbers of data and simulated background events and Figs. B.1 and B.2 show the m_{eff} and $E_{\text{T}}^{\text{miss}}/m_{\text{eff}}$ distributions for all four regions A, B, C, and D. Except for a global normalisation

	fail τID	pass τID
	$N_{trk} < 6, \lvert Q\rvert < 3$	$N_{trk} = 1$ or $3, \lvert Q\rvert = 1$
Small $\Delta\phi$ (inverted cuts)	Control region A	Control region B
Large $\Delta\phi$ (nominal cuts)	Control region C	Signal region D

Table B.1: Definition of signal and control regions for the random-τ background estimation method.

	data	W + Jets	top	Z + Jets	QCD
Region A	11368	187.4±25.4	177.8±25.5	168.7±34.7	9602±886
Region B	36	1.9±1.7	1.2±0.9	3.9±5.0	108±91
Region C	1160	103.6±44.2	83.9±30.6	66.4±40.2	340±35
Region D	42	10.8±4.1	12.7±3.2	11.3±7.8	4.4±4.1

Table B.2: Numbers of data and simulated events in the four regions of the QCD background estimation method. The Numbers for W + Jets and $t\bar{t}$ are corrected with the scale factors obtained in the $W/t\bar{t}$ enriched control region. Errors shown include Monte Carlo statistical uncertainties and statistical uncertainties on the scale factors on W and top.

difference, a good agreement in the shape of the distributions for data and simulation is observed. In Fig. B.2, the large Z rejection in region C can be seen.

The numbers of QCD events in regions A–C are then estimated by subtracting the estimated W, top, and Z contributions. The W and top contributions are estimated with Monte Carlo samples with additional scale factors derived from a separate control sample, as explained in Sec. 8.1 and Sec. 8.3. The Z contribution is taken purely from Monte Carlo estimates.

The number of QCD events expected in the signal region D is then estimated as

$$N_D^{\mathrm{QCD}}\Big|_{0.1<E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}<0.25} = \frac{N_B^{\mathrm{corrected}}}{N_A^{\mathrm{corrected}}} N_C^{\mathrm{corrected}}\Big|_{0.1<E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}<0.25} = 2.4 \pm 0.6\ . \tag{B.1}$$

The error on $N_D^{\mathrm{QCD}}\Big|_{E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}>0.1}$ takes into account the statistical uncertainties on the data events in the different regions, the statistical uncertainties on the Monte Carlo events of the samples that are subtracted, and the uncertainties on the scale factors derived for W and $t\bar{t}$. It has been checked that, even if we assumed a very conservative 100% uncertainty on the predicted number of events from W,Z and top production, the result would not change significantly.

To obtain a final estimate of the number of QCD events in the signal region for the nominal selection, this N_D^{QCD} needs to be scaled by the ratio of the event yields for $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.25$ and $0.25 > E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.1$. This scaling factor is taken from data in the QCD-dominated region A. The resulting QCD background estimates are

$$N_D^{\mathrm{QCD}}\Big|_{E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}>0.25} = 1.6 \pm 0.4 \tag{B.2}$$

for the number of events after the m_{T} cut, and

$$N_D^{\mathrm{QCD}}\Big|_{E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}>0.25 \& m_{\mathrm{eff}}>600\ \mathrm{GeV}} = 0.47 \pm 0.11 \tag{B.3}$$

in the high m_{eff} signal region. These results are consistent with those obtained by weighting the simulated QCD-multi-jet events by f_{ID} and normalising them in a control region, seen in Equations A.2 and A.3, respectively.

One can also consider uncertainties due to the extrapolation factor for the $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$ cut. Since this factor is taken from region A in data, any correlation between $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$ and τ identification would shift the results. To check this, one can also estimate the QCD background by taking the extrapolation

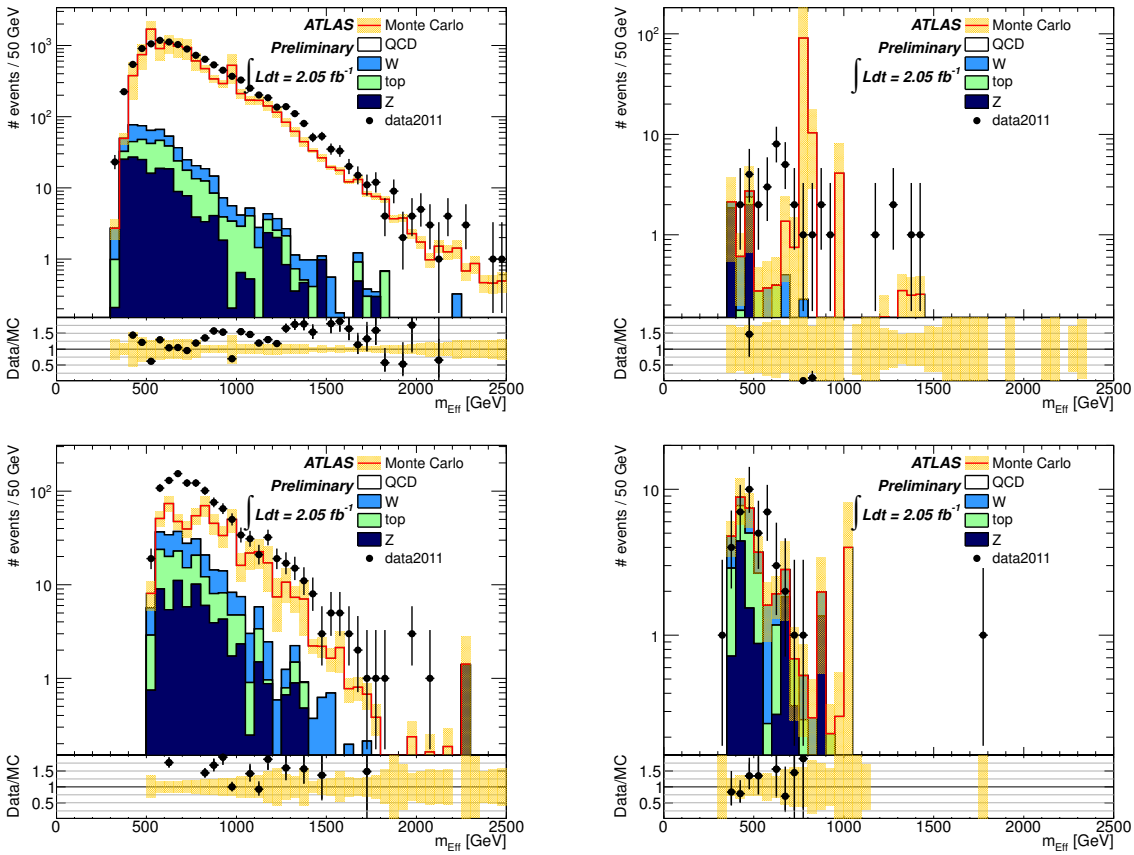


Figure B.1: Effective mass distribution for all four regions of the QCD background estimation method. The W and top contributions are scaled using the scale factors obtained in Sec. 8.1 and Sec. 8.3. Upper left: control region A, upper right: control region B, lower left: control region C, lower right: signal region D. Errors shown include Monte Carlo statistics and the statistical uncertainties on the scale factors on W and top.

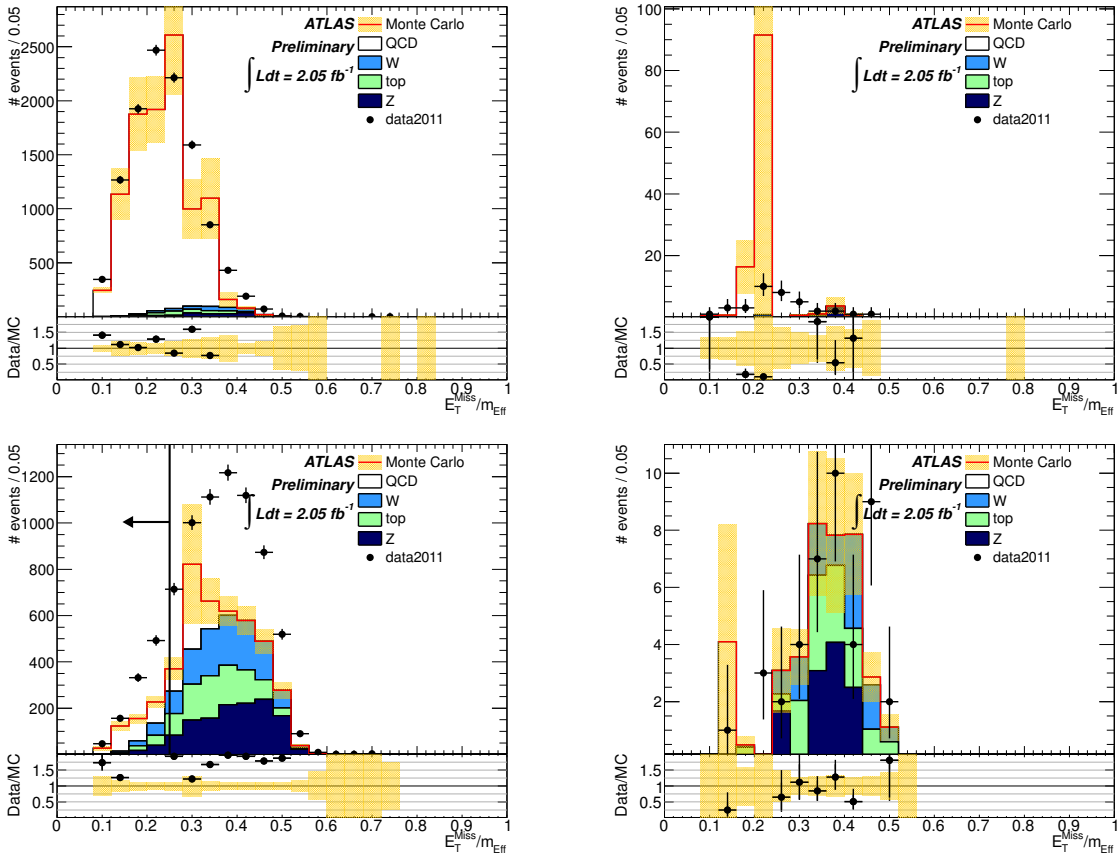


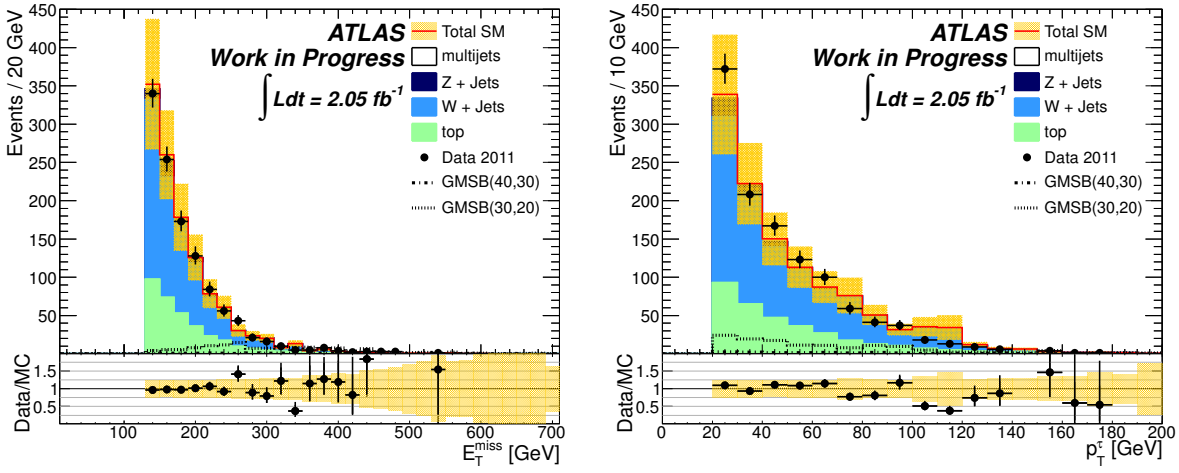
Figure B.2: $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$ distributions for all four regions of the QCD background estimation method. The W and top contributions are scaled using the scale factors obtained in Sec. 8.1 and Sec. 8.3. Upper left: control region A, upper right: control region B, lower left: control region C, lower right: signal region D. Errors shown include Monte Carlo statistics and the statistical uncertainties on the scale factors on W and top. In region C, the additional requirement $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} < 0.25$ is also indicated. This cut is intended to reduce the contamination from electroweak processes, particularly $Z \to \nu\nu$.

factor from region B in data. The scale factor obtained from Region A is 0.194 ± 0.009, while that from Region B is 0.4 ± 0.3. Region B is extremely statistically limited, but, since it requires the τ candidates to pass identification criteria, is a check of any possible correlations. Using the factor from region B, one estimates 1.0 ± 0.7 events in the final signal region. The difference between these two estimates can be taken as an additional systematic ucnertainty on the final estimate in the signal region,

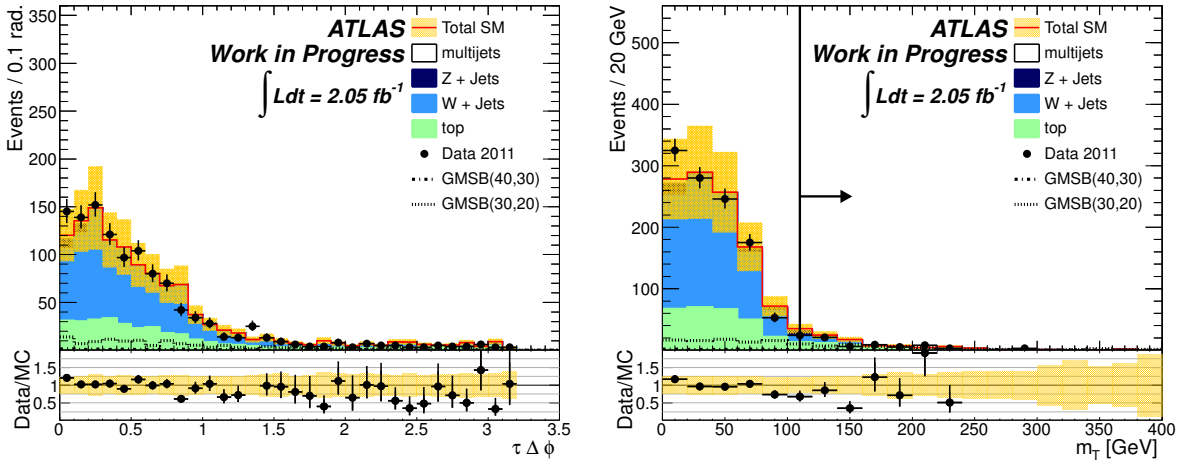
$$N_D^{\mathrm{QCD}}\Big|_{E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}>0.25 \& m_{\mathrm{eff}}>600\ \mathrm{GeV}} = 0.5 \pm 0.6\ . \tag{B.4}$$

Appendix C

Additional Distributions at the Late Stages of the Event Selection

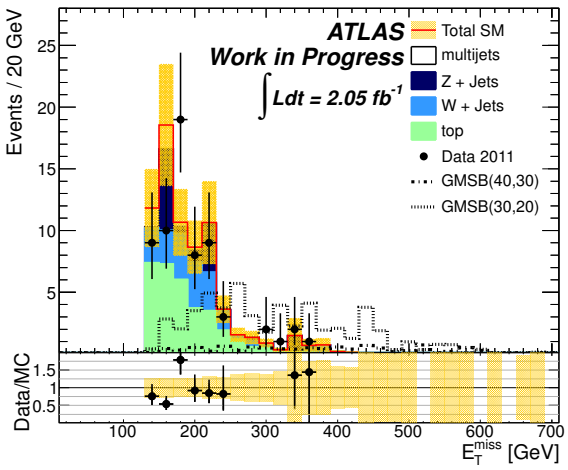


(a) $E_{\mathrm{T}}^{\mathrm{miss}}$ (b) leading p_{T}^{τ}

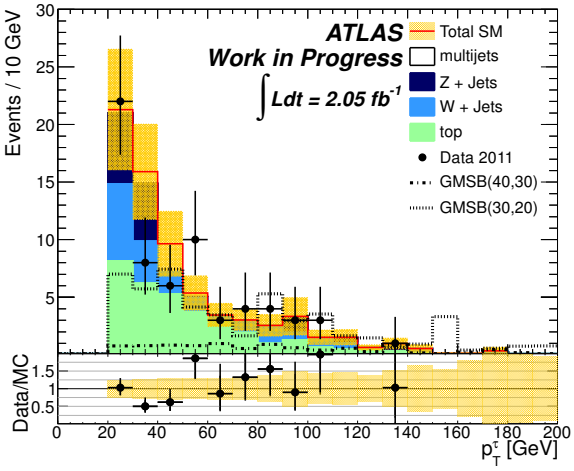


(c) $\Delta\phi(\tau\,, E_{\mathrm{T}}^{\mathrm{miss}})$ (d) m_{T}^{τ}

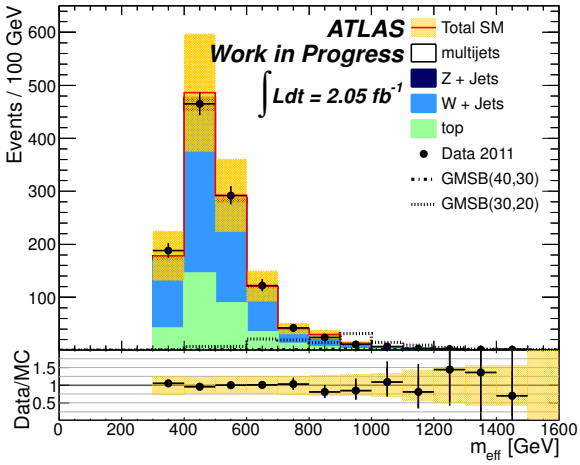
Figure C.1: Distributions before cutting on the transverse mass,(a) $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) leading τ momentum, (c) the azimuthal angle $\Delta\phi(\tau\,, E_{\mathrm{T}}^{\mathrm{miss}})$ and (d) the transverse mass m_{T}^{τ}. The yellow band denotes the combination of systematic and statistical uncertainties for the sum of the standard model backgrounds. The error bars on the data denote only their poissonian statistical uncertainties. Fig. (d) also indicates the second-to-last selection cut, $m_{\mathrm{T}}^{\tau} > 110$ GeV.



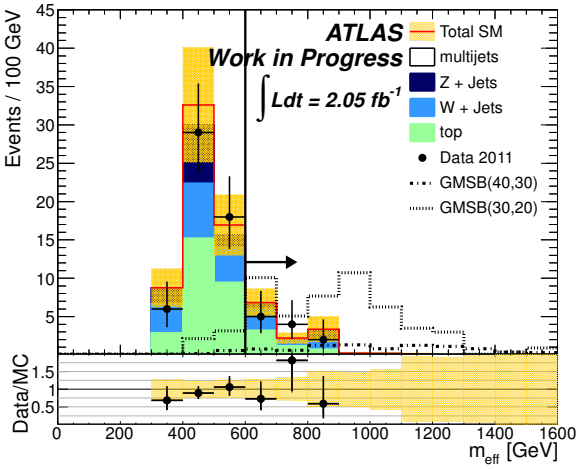
(a) $E_{\mathrm{T}}^{\mathrm{miss}}$



(b) leading p_{T}^{τ}



(c) m_{eff} before m_{T}^{τ} cut



(d) m_{eff} after m_{T}^{τ} cut

Figure C.2: Distributions of (a) $E_{\mathrm{T}}^{\mathrm{miss}}$ and (b) leading p_{T}^{τ} after the transverse mass requirement, and distributions of m_{eff} (c) before and (d) after the m_{T}^{τ} requirement. The yellow band denotes the combination of systematic and statistical uncertainties for the sum of the standard model backgrounds. The error bars on the data denote only their poissonian statistical uncertainties. Fig. (d) also indicates the final selection cut, $m_{\mathrm{eff}} > 600$ GeV.

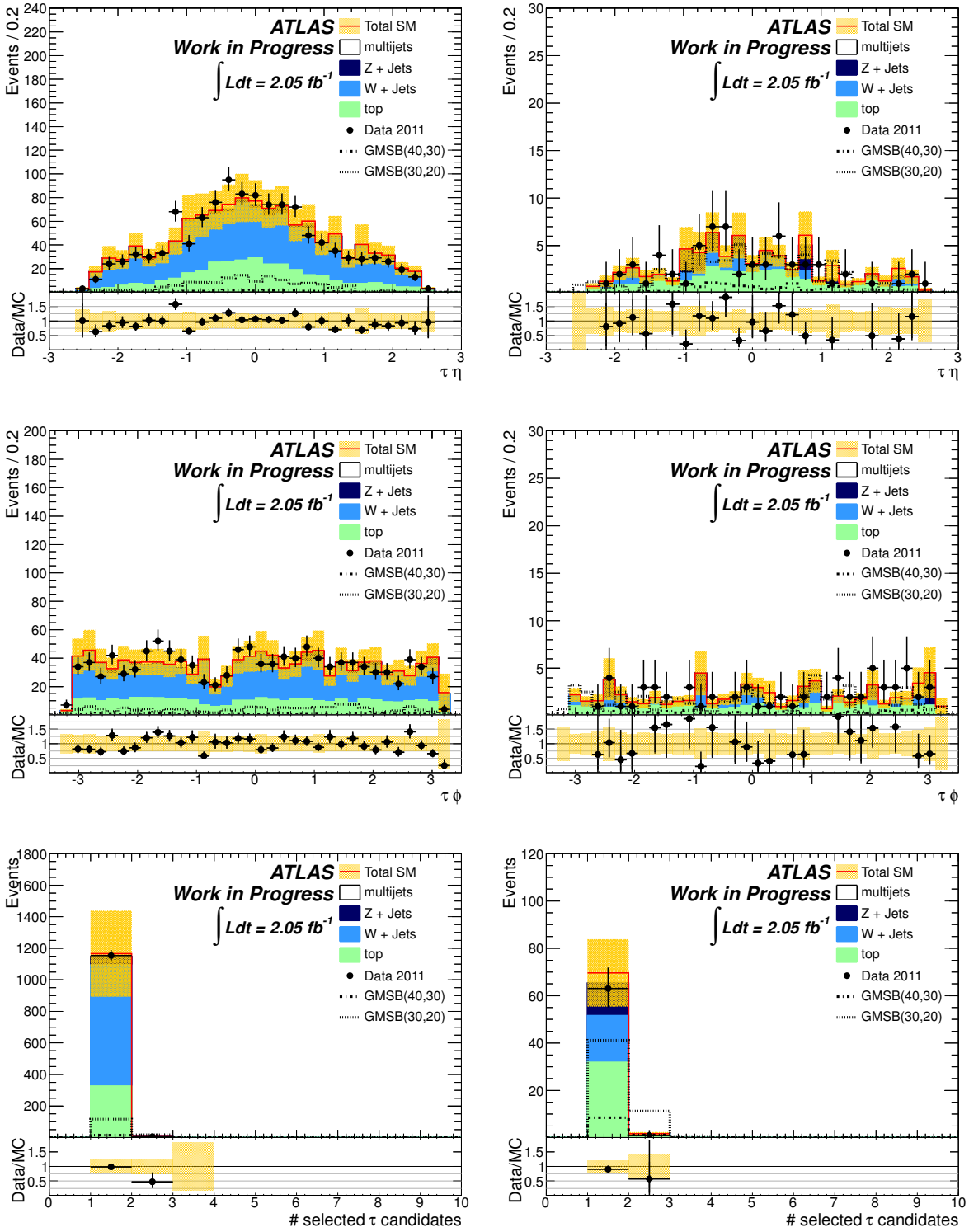
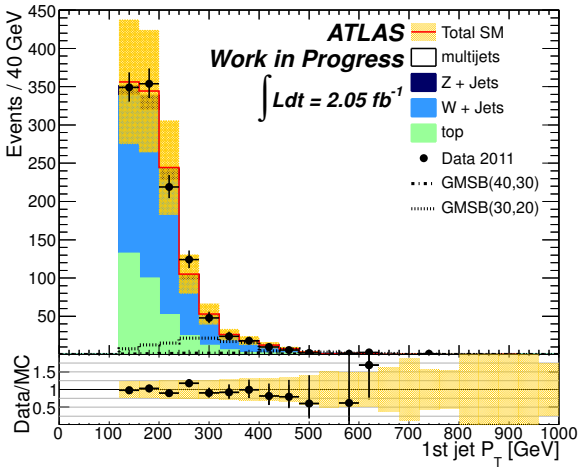
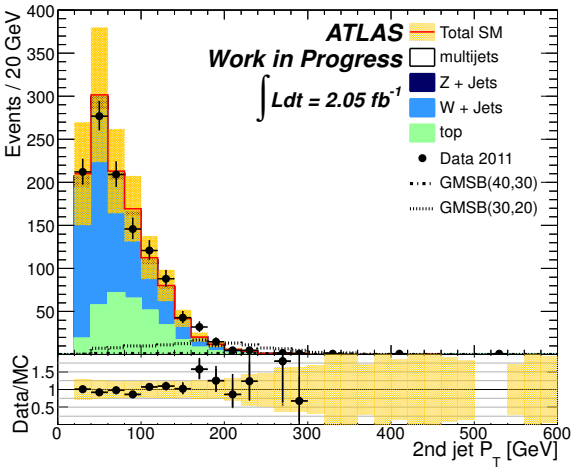


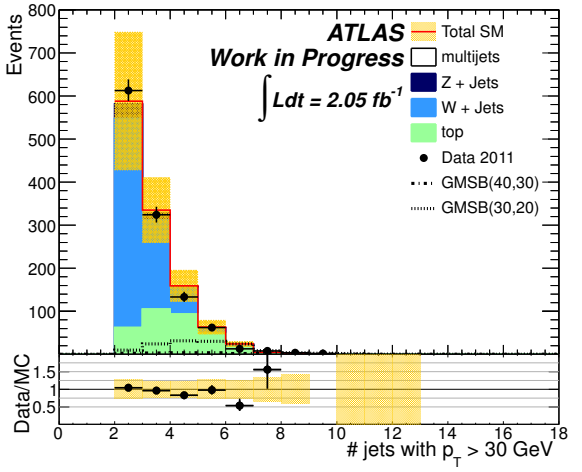
Figure C.3: η and ϕ distributions of the leading selected τ candidate and the multiplicity of selected τ candidates, shown before (left column) and after (right) the m_{T}^{τ} requirement. The yellow band denotes the combination of systematic and statistical uncertainties for the sum of the standard model backgrounds. The error bars on the data denote only their poissonian statistical uncertainties.



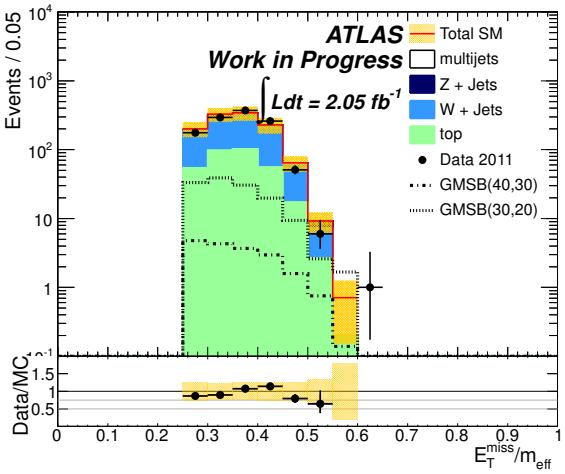
(a) $p_{\mathrm{T}}^{\mathrm{jet1}}$



(b) $p_{\mathrm{T}}^{\mathrm{jet2}}$

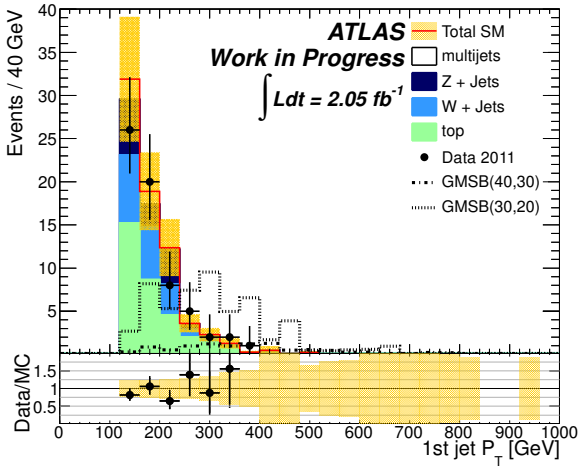


(c) jet multiplicity

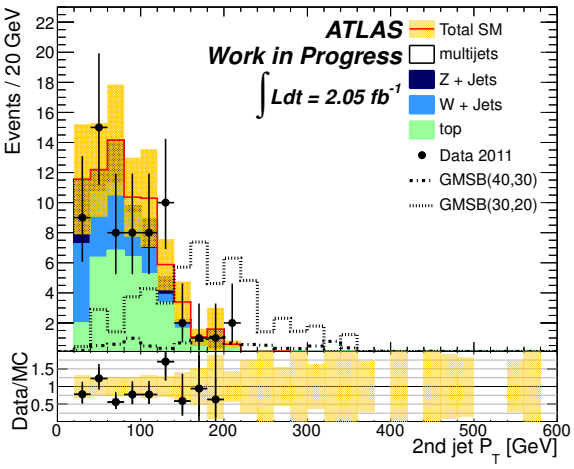


(d) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$

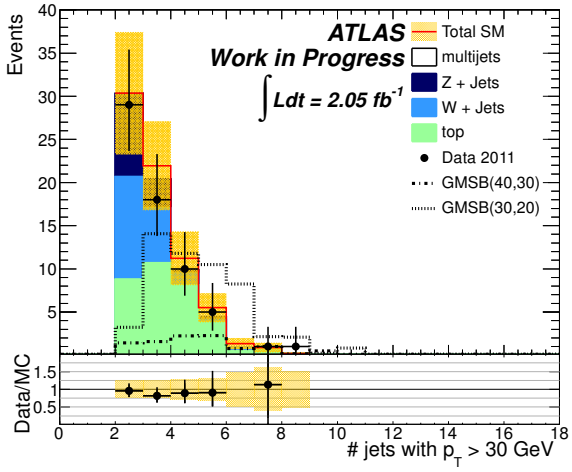
Figure C.4: jet kinematic distributions and multiplicity before m_{T}^{τ} requirement.The yellow band denotes the combination of systematic and statistical uncertainties for the sum of the standard model backgrounds. The error bars on the data denote only their poissonian statistical uncertainties.



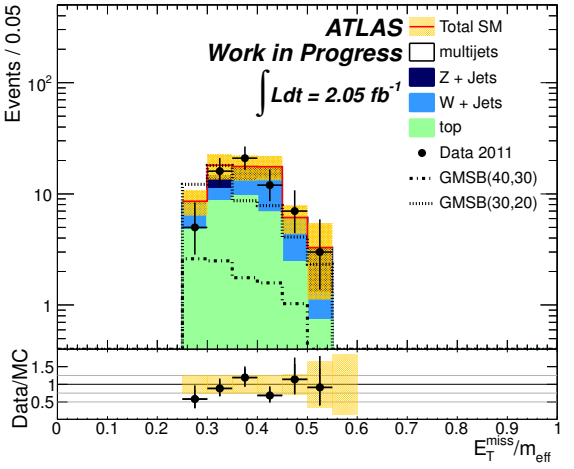
(a) $p_{\mathrm{T}}^{\mathrm{jet1}}$



(b) $p_{\mathrm{T}}^{\mathrm{jet2}}$



(c) jet multiplicity



(d) $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}}$

Figure C.5: jet kinematic distributions and multiplicity after m_{T}^{τ} requirement.The yellow band denotes the combination of systematic and statistical uncertainties for the sum of the standard model backgrounds. The error bars on the data denote only their poissonian statistical uncertainties.

Bibliography

[1] P. W. Higgs. 'Broken Symmetries and the Masses of Gauge Bosons'. In: *Phys. Rev. Lett.* 13 (1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[2] S. L. Glashow. 'Partial Symmetries of Weak Interactions'. In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[3] A. Salam and J.C. Ward. 'Electromagnetic and weak interactions'. In: *Phys. Lett.* 13 (1964), pp. 168–171. DOI: 10.1016/0031-9163(64)90711-5.

[4] S. Weinberg. 'A Model of Leptons'. In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[5] ATLAS Collaboration. 'Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC'. In: *Phys.Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv:1207.7214 [hep-ex].

[6] CMS Collaboration. 'Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC'. In: *Phys.Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv:1207.7235 [hep-ex].

[7] 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'. In: *Phys.Lett.* B710 (2012), pp. 67–85. DOI: 10.1016/j.physletb.2012.02.051. arXiv:1109.6572 [hep-ex].

[8] ATLAS Collaboration. 'Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using 4.7fb^{-1} of $\sqrt{s}$ = 7 TeV proton-proton collision data'. In: (2012). arXiv:1208.0949 [hep-ex].

[9] ATLAS Collaboration. 'Search for supersymmetry in final states with jets, missing transverse momentum and one isolated lepton in $\sqrt{s}$ = 7 TeV pp collisions using 1 fb^{-1} of ATLAS data'. In: *Phys.Rev.* D85 (2012), p. 012006. DOI: 10.1103/PhysRevD.85.012006. arXiv:1109.6606 [hep-ex].

[10] ATLAS Collaboration. 'Search for supersymmetry in pp collisions at $\sqrt{s}$ = 7 TeV in final states with missing transverse momentum and b-jets with the ATLAS detector'. In: *Phys.Rev.* D85 (2012), p. 112006. DOI: 10.1103/PhysRevD.85.112006. arXiv:1203.6193 [hep-ex].

[11] ATLAS Collaboration. 'Search for supersymmetry with jets, missing transverse momentum and at least one hadronically decaying tau lepton in proton-proton collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector'. In: *Phys. Lett.* B714 (2012), pp. 197–214. arXiv:1204.3852 [hep-ex].

[12] ATLAS Collaboration. 'Search for events with large missing transverse momentum, jets, and at least two tau leptons in 7 TeV proton-proton collision data with the ATLAS detector'. In: *Phys.Lett.* B714 (2012), pp. 180–196. DOI: 10.1016/j.physletb.2012.06.055. arXiv:1203.6580 [hep-ex].

[13] CMS Collaboration. *Search for Physics Beyond the Standard Model in Events with Opposite-sign Tau Pairs and Missing Energy*. Tech. rep. CMS PAS SUS-11-007. Geneva: CERN, 2011.

[14] A. Djouadi. 'The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model'. In: *Phys.Rept.* 457 (2008), pp. 1–216. DOI: 10.1016/j.physrep.2007.10.004. arXiv:hep-ph/0503172 [hep-ph].

[15] H. Fritzsch, M. Gell-Mann and H. Leutwyler. 'Advantages of the Color Octet Gluon Picture'. In: *Phys.Lett.* B47 (1973), pp. 365–368. DOI: 10.1016/0370-2693(73)90625-4.

[16] D.J. Gross and F. Wilczek. 'Asymptotically Free Gauge Theories. 1'. In: *Phys.Rev.* D8 (1973), pp. 3633–3652. DOI: 10.1103/PhysRevD.8.3633.

[17] H. David Politzer. 'Asymptotic Freedom: An Approach to Strong Interactions'. In: *Phys.Rept.* 14 (1974), pp. 129–180. DOI: 10.1016/0370-1573(74)90014-3.

[18] K. Nakamura et al. (Particle Data Group). In: *J. Phys.* G 37 (2010), p. 075021.

[19] M. Kobayashi and T. Maskawa. 'CP Violation in the Renormalizable Theory of Weak Interaction'. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[20] R. Haag, J.T. Lopuszanski and M. Sohnius. 'All Possible Generators of Supersymmetries of the S Matrix'. In: *Nucl. Phys.* B88 (1975), p. 257. DOI: 10.1016/0550-3213(75)90279-5.

[21] S.R. Coleman and J. Mandula. 'All Possible Symmetries of the S Matrix'. In: *Phys. Rev.* 159 (1967), pp. 1251–1256. DOI: 10.1103/PhysRev.159.1251.

[22] A. Djouadi. 'The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model'. In: *Phys.Rept.* 459 (2008), pp. 1–241. DOI: 10.1016/j.physrep.2007.10.005. arXiv:hep-ph/0503173 [hep-ph].

[23] C. Anders. 'Private Communication'. In: (2012).

[24] Stephen P. Martin. 'A Supersymmetry Primer'. In: (1997). arXiv:hep-ph/9709356.

[25] L. Alvarez-Gaume, M. Claudson and M.B. Wise. 'Low-energy supersymmetry'. In: *Nucl. Phys.* B207 (1982), p. 96.

[26] M. Dine, W. Fischler and M. Srednicki. 'Supersymmetric technicolor'. In: *Nucl. Phys.* B189 (1981), p. 575. DOI: 10.1016/0550-3213(81)90582-4.

[27] S. Dimopoulos and S. Raby. 'Supercolor'. In: *Nucl. Phys.* B192 (1981), p. 353. DOI: 10.1016/0550-3213(81)90430-2.

[28] C.R. Nappi and B.A. Ovrut. 'Supersymmetric extension of the SU(3)×SU(2)×U(1) Model'. In: *Phys. Lett.* B113 (1982), p. 175. DOI: 10.1016/0370-2693(82)90418-X.

[29] M. Dine and A.E. Nelson. 'Dynamical supersymmetry breaking at low-energies'. In: *Phys. Rev.* D48 (1993), p. 1277. eprint: hep-ph/9303230.

[30] M. Dine, A.E. Nelson and Y. Shirman. 'Low-energy dynamical supersymmetry breaking simplified'. In: *Phys. Rev.* D51 (1995), p. 1362. eprint: hep-ph/9408384.

[31] M. Dine et al. 'New tools for low-energy dynamical supersymmetry breaking'. In: *Phys. Rev.* D53 (1996), p. 2658. eprint: hep-ph/9507378.

[32] LEPSUSYWG, ALEPH, DELPHI, L3 and OPAL Experiments. *Combined LEP GMSB stau/smuon/selectron results, 189-208 GeV*. Tech. rep. LEPSUSYWG/02-09.2. 2002.

[33] ATLAS Collaboration. *Constraining the gauge-mediated Supersymmetry breaking model in final states with two leptons, jets and missing transverse momentum with the ATLAS experiment at $\sqrt{s}$ = 7 TeV*. Tech. rep. ATLAS-CONF-2011-156. Geneva: CERN, 2011.

[34] E. Boos et al. 'Generic user process interface for event generators'. In: (2001). arXiv:hep-ph/0109068.

[35] G. Altarelli and G. Parisi. 'Asymptotic Freedom in Parton Language'. In: *Nucl.Phys.* B126 (1977), p. 298. DOI: 10.1016/0550-3213(77)90384-4.

[36] V.N. Gribov and L.N. Lipatov. 'Deep inelastic e p scattering in perturbation theory'. In: *Sov.J.Nucl.Phys.* 15 (1972), pp. 438–450.

[37] Y.L. Dokshitzer. 'Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.' In: *Sov.Phys.JETP* 46 (1977), pp. 641–653.

[38] S. Alekhin et al. 'The 3-, 4-, and 5-flavor NNLO Parton from Deep-Inelastic- Scattering Data and at Hadron Colliders'. In: *Phys. Rev.* D81 (2010), p. 014032. DOI: 10.1103/PhysRevD.81.014032. arXiv:0908.2766 [hep-ph].

[39] P.M. Nadolsky et al. 'Implications of CTEQ global analysis for collider observables'. In: *Phys. Rev.* D78 (2008), p. 013004. DOI: 10.1103/PhysRevD.78.013004. arXiv:0802.0007 [hep-ph].

[40] H. Lai et al. 'New parton distributions for collider physics'. In: *Phys. Rev.* D82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv:1007.2241 [hep-ph].

[41] S. Habib. 'Combined measurement and QCD analysis of the inclusive e+- p scattering cross sections at HERA'. In: *PoS* DIS2010 (2010), p. 035.

[42] V. Radescu. 'Hera Precision Measurements and Impact for LHC Predictions'. In: (2011). arXiv:1107.4193 [hep-ex].

[43] A. D. Martin et al. 'Parton distributions for the LHC'. In: *Eur. Phys. J.* C63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv:0901.0002 [hep-ph].

[44] R.D. Ball et al. 'A first unbiased global NLO determination of parton distributions and their uncertainties'. In: *Nucl. Phys.* B838 (2010), pp. 136–206. DOI: 10.1016/j.nuclphysb.2010.05.008. arXiv:1002.4407 [hep-ph].

[45] R.D. Ball et al. 'Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology'. In: *Nucl. Phys.* B849 (2011), pp. 296–363. DOI: 10.1016/j.nuclphysb.2011.03.021. arXiv:1101.1300 [hep-ph].

[46] T. Sjostrand, S. Mrenna and P. Skands. 'PYTHIA 6.4 Physics and Manual'. In: *JHEP* 0605 (2006), p. 026.

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

[48] B. Andersson et al. 'Parton Fragmentation and String Dynamics'. In: *Phys. Rept.* 97 (1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[49] M. A. Dobbs et al. 'Les Houches guidebook to Monte Carlo generators for hadron collider physics'. In: (2004). arXiv:hep-ph/0403045.

[50] J. M. Butterworth, J. R. Forshaw and M. H. Seymour. 'Multiparton interactions in photoproduction at HERA'. In: *Z. Phys.* C 72 (1996), 637.

[51] S. Agostinelli et al. 'GEANT4: A Simulation toolkit'. In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[52] ATLAS Collaboration. 'The ATLAS Simulation Infrastructure'. In: *Eur. Phys. J.* C70 (2010), pp. 823–874. DOI: 10.1140/epjc/s10052-010-1429-9.

[53] P. Calafiura et al. 'The athena control framework in production, new developments and lessons learned'. In: *Proceedings Computing in High-Energy Physics (CHEP '04)* (2005), pp. 456–458.

[54] ATLAS Collaboration. *First tuning of HERWIG/JIMMY to ATLAS data*. Tech. rep. ATL-PHYS-PUB-2010-014. Geneva: CERN, 2010.

[55] ATLAS Collaboration. *Charged particle multiplicities in p p interactions at* $\sqrt{s} = 0.9$ *and 7 TeV in a diffractive limited phase-space measured with the ATLAS detector at the LHC and new PYTHIA6 tune*. Tech. rep. ATLAS-CONF-2010-031. Geneva: CERN, 2010.

[56] ATLAS Collaboration. 'Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC'. In: *New J. Phys.* 13 (2011), p. 053033. DOI: 13/5/053033.

[57] M. L. Mangano et al. 'ALPGEN a generator for hard multiparton processes in hadronic collisions'. In: *JHEP* 07 (2003), p. 001. eprint: hep-ph/0206293.

[58] J. Pumplin et al. 'New generation of parton distributions with uncertainties from global QCD analysis'. In: *JHEP* 0207 (2002), p. 012. arXiv:hep-ph/0201195 [hep-ph].

[59] R. Hamberg, W.L. van Neerven and T. Matsuura. 'A Complete calculation of the order $\alpha - s^2$ correction to the Drell-Yan K factor'. In: *Nucl.Phys.* B359 (1991), pp. 343–405. DOI: 10.1016/0550-3213(91)90064-5.

[60] K. Melnikov and F. Petriello. 'The W boson production cross section at the LHC through $O(\alpha_s^2)$'. In: *Phys.Rev.Lett.* 96 (2006), p. 231803. DOI: 10.1103/PhysRevLett.96.231803. arXiv:hep-ph/0603182 [hep-ph].

[61] K. Melnikov and F. Petriello. 'Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$'. In: *Phys.Rev.* D74 (2006), p. 114017. DOI: 10.1103/PhysRevD.74.114017. arXiv:hep-ph/0609070 [hep-ph].

[62] S. Frixione and B.R. Webber. 'Matching NLO QCD computations and parton shower simulations'. In: *JHEP* 0206 (2002), p. 029. arXiv:hep-ph/0204244 [hep-ph].

[63] R. Bonciani et al. 'NLL resummation of the heavy quark hadroproduction cross-section'. In: *Nucl.Phys.* B529 (1998), pp. 424–450. DOI: 10.1016/j.nuclphysb.2008.06.006, 10.1016/S0550-3213(98)00335-6. arXiv:hep-ph/9801375 [hep-ph].

[64] M. Beneke et al. 'Threshold expansion of the gg(qq-bar) to QQ-bar + X cross section at $O(\alpha_s^4)$'. In: *Phys.Lett.* B690 (2010), pp. 483–490. DOI: 10.1016/j.physletb.2010.05.038. arXiv:0911.5166 [hep-ph].

[65] S. Moch and P. Uwer. 'Theoretical status and prospects for top-quark pair production at hadron colliders'. In: *Phys.Rev.* D78 (2008), p. 034003. DOI: 10.1103/PhysRevD.78.034003. arXiv:0804.1476 [hep-ph].

[66] J.M. Campbell and R.K. Ellis. 'An Update on vector boson pair production at hadron colliders'. In: *Phys.Rev.* D60 (1999), p. 113006. DOI: 10.1103/PhysRevD.60.113006. arXiv:hep-ph/9905386 [hep-ph].

[67] J.M. Campbell, R.K. Ellis and C. Williams. 'Vector boson pair production at the LHC'. In: *JHEP* 1107 (2011), p. 018. DOI: 10.1007/JHEP07(2011)018. arXiv:1105.0020 [hep-ph].

[68] S. Jadach et al. 'The tau decay library TAUOLA: Version 2.4'. In: *Comput.Phys.Commun.* 76 (1993), pp. 361–380. DOI: 10.1016/0010-4655(93)90061-G.

[69] P. Golonka et al. 'The Tauola photos F environment for the TAUOLA and PHOTOS packages: Release. 2.' In: *Comput.Phys.Commun.* 174 (2006), pp. 818–835. DOI: 10.1016/j.cpc.2005.12.018. arXiv:hep-ph/0312240 [hep-ph].

[70] E. Barberio and Z. Was. 'PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0'. In: *Comput.Phys.Commun.* 79 (1994), pp. 291–308. DOI: 10.1016/0010-4655(94)90074-4.

[71] M. Bahr et al. 'Herwig++ Physics and Manual'. In: *Eur.Phys.J.* C58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv:0803.0883 [hep-ph].

[72] F.E. Paige et al. *ISAJET 7.69: A Monte Carlo event generator for pp, anti-p p, and e+e- reactions*. 2003. arXiv:hep-ph/0312045 [hep-ph].

[73] A. Sherstnev and R.S. Thorne. 'Parton Distributions for LO Generators'. In: *Eur.Phys.J.* C55 (2008), pp. 553–575. DOI: 10.1140/epjc/s10052-008-0610-x. arXiv:0711.2473 [hep-ph].

[74] W. Beenakker et al. 'Squark and gluino production at hadron colliders'. In: *Nucl.Phys.* B492 (1997), pp. 51–103. DOI: 10.1016/S0550-3213(97)00084-9. arXiv:hep-ph/9610490 [hep-ph].

[75] W. Beenakker et al. 'Stop production at hadron colliders'. In: *Nucl.Phys.* B515 (1998), pp. 3–14. DOI: 10.1016/S0550-3213(98)00014-5. arXiv:hep-ph/9710451 [hep-ph].

[76] W. Beenakker et al. 'The Production of charginos / neutralinos and sleptons at hadron colliders'. In: *Phys.Rev.Lett.* 83 (1999), pp. 3780–3783. DOI: 10.1103/PhysRevLett.100.029901,10.1103/PhysRevLett.83.3780. arXiv:hep-ph/9906298 [hep-ph].

[77] T. Plehn. 'Measuring the MSSM Lagrangean'. In: *Czech.J.Phys.* 55 (2005), B213–B220. arXiv:hep-ph/0410063 [hep-ph].

[78] *https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults*.

[79] ATLAS Collaboration. 'The ATLAS Experiment at the CERN Large Hadron Collider'. In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[80] V. Cindro et al. 'The ATLAS beam conditions monitor'. In: *JINST* 3 (2008), P02004. DOI: 10.1088/1748-0221/3/02/P02004.

[81] H. Stenzel. *Luminosity calibration from elastic scattering*. Tech. rep. ATL-LUM-PUB-2007-001. Geneva: CERN, 2007.

[82] ATLAS Collaboration. *Updated Luminosity Determination in pp Collisions at root(s)=7 TeV using the ATLAS Detector*. Tech. rep. ATLAS-CONF-2011-011. Geneva: CERN, 2011.

[83] ATLAS Collaboration. *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*. Tech. rep. Geneva: CERN, 2009. arXiv:0901.0512 [hep-ex].

[84] W. Lampl et al. *Calorimeter clustering algorithms: Description and performance*. Tech. rep. ATL-LARG-PUB-2008-002. Geneva: CERN, 2008.

[85] T. Barillari et al. *Local hadronic calibration*. Tech. rep. ATL-LARG-PUB-2009-001-2. Geneva: CERN, 2009.

[86] M. Cacciari and G.P. Salam. 'Dispelling the N^3 myth for the k_t jet-finder'. In: *Phys. Lett.* B641 (2006), pp. 57–61. DOI: 10.1016/j.physletb.2006.08.037. arXiv:hep-ph/0512210.

[87] M. Cacciari, G.P. Salam and G. Soyez. 'The anti–k_t jet clustering algorithm'. In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv:0802.1189 [hep-ph].

[88] G.P. Salam. 'Towards Jetography'. In: *Eur. Phys. J.* C67 (2010), pp. 637–686. DOI: 10.1140/epjc/s10052-010-1314-6. arXiv:0906.1833 [hep-ph].

[89] G. Piacquadio and C. Weiser. 'A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS'. In: *J.Phys.Conf.Ser.* 119 (2008), p. 032032.

[90] ATLAS Collaboration. *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*. Tech. rep. ATLAS-CONF-2011-102. Geneva: CERN, 2011.

[91] ATLAS Collaboration. *Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS*. Tech. rep. ATLAS-CONF-2011-152. Geneva: CERN, 2011.

[92] ATLAS Collaboration. *Commissioning of the ATLAS Tau-Lepton Reconstruction Using 900 GeV Minimum-Bias Data*. Tech. rep. ATLAS-CONF-2010-012. Geneva: CERN, 2010.

[93] ATLAS Collaboration. *Reconstruction of hadronic tau candidates in QCD events at ATLAS with 7 TeV proton-proton collisions*. Tech. rep. ATLAS-CONF-2010-059. Geneva: CERN, 2010.

[94] ATLAS Collaboration. *Tau Reconstruction and Identification Performance in ATLAS*. Tech. rep. ATLAS-CONF-2010-086. Geneva: CERN, 2010.

[95] A. Hoecker et al. 'TMVA: Toolkit for Multivariate Data Analysis'. In: *PoS* ACAT (2007), p. 040. arXiv:physics/0703039.

[96] Y. Freund and R.E. Shapire. 'Experiments with a New Boosting Algorithm'. In: *Proceedings 13th International Conference on Machine Learning*. 1996.

[97] Y. Freund and R.E. Schapire. *A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting*. 1995.

[98] ATLAS Collaboration. 'Measurement of the Z to tau tau Cross Section with the ATLAS Detector'. In: *Phys.Rev.* D84 (2011), p. 112006. DOI: 10.1103/PhysRevD.84.112006. arXiv:1108.2016 [hep-ex].

[99] ATLAS Collaboration. 'Measurement of the W to tau nu Cross Section in pp Collisions at $\sqrt{s}$ = 7 TeV with the ATLAS experiment'. In: *Phys.Lett.* B706 (2012), pp. 276–294. DOI: 10.1016/j.physletb.2011.11.057. arXiv:1108.4101 [hep-ex].

[100] ATLAS Collaboration. *Measurement of hadronic tau decay identification efficiency using W to tau nu events*. Tech. rep. ATLAS-CONF-2011-093. Geneva: CERN, 2011.

[101] ATLAS Collaboration. *Expected electron performance in the ATLAS experiment*. Tech. rep. ATL-PHYS-PUB-2011-006. Geneva: CERN, 2011.

[102] ATLAS Collaboration. *Dimuon composition in ATLAS at 7 TeV*. Tech. rep. ATLAS-CONF-2011-003. Geneva: CERN, 2011.

[103] ATLAS Collaboration. *Measurement of the Mis-identification Probability of Tau Leptons from Hadronic Jets and from Electrons*. Tech. rep. ATLAS-CONF-2011-113. Geneva: CERN, 2011.

[104] ATLAS Collaboration. *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of sqrts = 7 TeV*. Tech. rep. ATLAS-CONF-2010-038. Geneva: CERN, 2010.

[105] ATLAS Collaboration. *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons*. Tech. rep. ATLAS-CONF-2011-077. Geneva: CERN, 2011.

[106] ATLAS Collaboration. *Search for neutral MSSM Higgs bosons decaying to tau+tau- pairs in proton-proton collisions at sqrt(s) = 7 TeV with the ATLAS detector*. Tech. rep. ATLAS-CONF-2011-132. Geneva: CERN, 2011.

[107] ATLAS Collaboration. *Measurement of the $t\bar{t}$ production cross section in the final state with a hadronically decaying tau lepton and jets using the ATLAS detector*. Tech. rep. ATLAS-CONF-2012-032. Geneva: CERN, 2012.

[108] ATLAS Collaboration. 'Luminosity Determination in pp Collisions at $\sqrt{s}$=7 TeV Using the ATLAS Detector at the LHC'. In: *Eur.Phys.J.* C71 (2011), p. 1630. DOI: 10.1140/epjc/s10052-011-1630-5. arXiv:1101.2185 [hep-ex].

[109] E. Gross G. Cowan K. Cranmer and O. Vitells. 'Asymptotic formulae for likelihood-based tests of new physics'. In: *Eur. Phys. J.* C71 (2011), p. 1554. eprint: arXiv:1007.1727.

[110] F. Buehrer. 'Estimation of the W+jets, top and QCD backgrounds in the search for supersymmetry with jets, missing transverse momentum and tau leptons'. In: (2011).

[111] R. Barlow and C. Beeston. 'Fitting using finite Monte Carlo samples'. In: *Comput.Phys.Commun.* 77 (1993), pp. 219–228. DOI: 10.1016/0010-4655(93)90005-W.

[112] ATLAS Collaboration. *Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s}$=7 TeV in ATLAS 2010 data*. Tech. rep. ATLAS-CONF-2011-032. Geneva: CERN, 2011.

[113] ATLAS Collaboration. *Close-by Jet Effects on Jet Energy Scale Calibration in pp Collisions at $\sqrt{s}$=7 TeV with the ATLAS Detector*. Tech. rep. ATLAS-CONF-2011-062. Geneva: CERN, 2011.

[114] ATLAS Collaboration. *In-situ jet energy scale and jet shape corrections for multiple interactions in the first ATLAS data at the LHC*. Tech. rep. ATLAS-CONF-2011-030. Geneva: CERN, 2011.

[115] 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*. Tech. rep. ATLAS-CONF-2010-054. Geneva: CERN, 2010.

[116] G Romeo et al. *Jet Energy Resolution from In-situ Techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s}$ = 7 TeV*. Tech. rep. ATL-COM-PHYS-2011-240. Geneva: CERN, 2011.

[117] ATLAS Collaboration. *Muon Momentum Resolution in First Pass Reconstruction of pp Collision Data Recorded by ATLAS in 2010*. Tech. rep. ATLAS-CONF-2011-046. Geneva: CERN, 2011.

[118] ATLAS Collaboration. *Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV*. Tech. rep. ATLAS-CONF-2012-101. Geneva: CERN, 2012.

[119] ATLAS Collaboration. *Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector*. Tech. rep. ATLAS-CONF-2011-063. Geneva: CERN, 2011.

[120] OPAL Collaboration, G. Abbiendi *et al.*. In: *Eur. Phys. J.* C 46 (2006), p. 307.

[121] ATLAS Collaboration. *Setting exclusion limits in ATLAS supersymmetry searches with a likelihood ratio based method*. Tech. rep. ATL-COM-PHYS-2011-004. Geneva: CERN, 2011.

[122] A. Read. 'Presentation of search results: the CLs technique'. In: *Journal of Physics G: Nucl. Part. Phys.* 28 (2002), p. 2693.

[123] ATLAS Collaboration. 'Search for Supersymmetry in Events with Large Missing Transverse Momentum, Jets, and at Least One Tau Lepton in 7 TeV Proton-Proton Collision Data with the ATLAS Detector'. In: (2012). arXiv:1210.1314 [hep-ex].

[124] CMS Collaboration. *Search for supersymmetry in all-hadronic events with tau leptons*. Tech. rep. CMS PAS SUS-12-004. Geneva: CERN, 2012.

List of Figures

List of Tables

Acknowledgements

Nach fünf Jahren (und ein paar Monaten) ist die letzte Hürde nun genommen, das Machwerk eingereicht und verteidigt und entgegen meiner ewigen Proteste und Nörgeleien habe ich mich entschieden eine weitere Stelle in der Forschung anzutreten. Kurzum also ein guter Zeitpunkt zurückzuschauen und sich klarzumachen (und dann auch schriftlich zuzugeben) dass man Nichts von Alledem geschafft hätte ohne die Unterstützung vieler Freunde bei und abseits der Arbeit.

Der erste Dank gebührt sicher meinem Doktorvater Prof. Dr. Jochen Dingfelder für seine unerschöpfliche Geduld, Enthusiasmus und gute Laune. Jochen, wir wissen immer noch nicht wie du dass machst. Ich hätte mir keinen besseren Doktorvater wünschen können.

Bedanken möchte ich mich auch bei Prof. Dr. Markus Schumacher dafür dass er mich im letztem Jahr meiner Doktorarbeit in seine Gruppe aufgenommen hat und mir die Gelegenheit gegeben hat mit so vielen wirklich tollen Kollegen zusammen an der Publikation zum ersten "Anzeichen" eines Signals bei der Suche nach dem Zerfall des Higgs Bosons in Tauleptonen zu arbeiten. Ich bin mir sehr bewusst welches grosse Glück ich hatte unter zwei Gruppenleitern arbeiten zu dürfen, die man sich nicht nur fachlich sondern auch menschlich als Beispiel nehmen sollte.

Mein Dank gilt auch Prof. Dr. Klaus Desch, Prof. Dr. Herbert Dreiner und Prof. Dr. Armin Cremers dass sie sich die Zeit genommen haben auf meine Doktorarbeit und deren Verteidigung ein Dreiviertel Jahr zu warten und dann trotzdem bei der Verteidigung nicht allzu viele fiese Fragen gestellt haben.

Ausserdem möchte ich Prof. Dr. Hans-Christian Schultz-Coulon und Prof. Dr. Karl Jakobs für die freundliche Aufnahme in ihre Gruppen danken.

Besonderer Dank gilt meinen direkten Betreuern bei den verschiedenen Projekten meiner Doktorarbeit; Dr. Stan Lai, Dr. Sylvie Brunet, Dr. Philip Bechtle und Dr. Michael Mazur. Ohne ihre geduldige Anleitung und produktive Kritik wären die Analysen, die in dieser Doktorarbeit veröffentlicht sind, nicht zur Publikationsreife gekommen.

Genauso wichtig war mir auch die Zusammenarbeit und Freundschaft meiner direkten Gruppenmitglieder aus der Gruppe Dingfelder; Veit Scharf, Felix Bührer, Liv Wiik, Julian Glatzer und Christoph Anders sowie aus der Gruppe Schumacher; Dr. Michael Böhler, Dirk Sammel, Florian Kiss, Holger v. Radziewski, Christian Schillo und Julian Maluck. Dabei habe ich natürlich nicht die lieben Kollegen aus den anderen Gruppen in Heidelberg, Freiburg und Bonn vergessen. Euch Allen ein dickes Dankeschön für den guten Zusammenhalt und (manchmal) gute Laune im Büro und auch für das Ignorieren des schlimmsten Schwachsinnes den ich ab und zu von mir gegeben habe. Besonders erwähnt werden muss hier Dr. Matthias Werner, der immer schon vorher einen Grossteil meiner unqualifizierten Kommentare angehört und gefiltert hat, bevor ich die anderen damit belästigen konnte.

Da ich sicherlich eine grosse Zahl von Kollegen vergessen habe, möchte ich mich hiermit auch bei den

vielen Kollegen, die mir im Laufe meiner Arbeit bei ATLAS begegnet sind, für die gute Zusammenarbeit bedanken.

Zum Schluss möchte ich noch meiner Familie und meinen Freunden ausserhalb der Arbeit für den Rückhalt und das Verständnis danken, speziell meinen lieben Eltern die mich immer respektiert und unterstützt haben. Ich bin sehr froh dass ich für ein paar Jahre fast völlig in meine Arbeit abtauchen konnte und nun da ich wieder aufgetaucht bin noch so viele Menschen da sind die mir das nicht übel genommen haben. Genauso macht es mich aber auch traurig das zwei besondere Menschen das nicht mehr miterleben können: Justin und Brigitte, ihr fehlt mir.