Measurement of the mass of the top quark using the transverse decay length and lepton transverse momentum techniques

Dissertation

Zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften der Fakultät Physik an der Technischen Universität Dortmund

vorgelegt von

Christian Jung

Lehrstuhl für Experimentelle Physik IV

Fakultät Physik

May 2, 2014

Contents

Chapter 1

Introduction

Giant machines have been built to gain insight into the smallest pieces of nature. This is driven by the fact that to resolve smaller sizes, more energy is needed. The real driving force however is mankind's passion for understanding the reality it finds itself in. In this thesis but a small contribution is made to this struggle. It is a measurement of one of the fundamental parameters of the theory called the "Standard Model of Particle Physics". This parameter is the mass of the particle known as the "top quark". A better knowledge of the parameters of the standard model might open the door to a better understanding of nature or simply help making better prediction using that model.

After this introduction, in chapter 2 an overview of the standard model of particle physics is given. It is the theory which describes with astonishing precision the structure and interactions of all known matter. The emphasis on this chapter is put on the top quark. With a mass of roughly a gold atom, it is the heaviest fundamental particle known to date. Due to this high mass it is suspected to play an important role in the consistency of the standard model and searches for physics beyond it. These implications are stronger the higher the precision on the mass of the top quark is, motivating the analysis presented in this thesis.

Any measurement wants the best possible knowledge of the detectors and techniques it is using. The experiment used for this thesis is the ATLAS experiment. A description of the detector and its capabilities are given in chapter 3.

The measurement of the mass of the top quark uses Monte-Carlo simulation techniques to a very large extent. The simulation of a proton-proton interaction producing a pair of top quarks is a complicated process. A short description of the basic concepts behind these simulation techniques can be found in chapter 4.

The discussion of the current status of measurements of the mass of the top quark and techniques for achieving that goal is presented in chapter 5. Focus is put on the template method which will also be used in this thesis. Additionally that chapter describes the limiting uncertainties on measuring the mass of the top quark. One of them is the "jet energy scale uncertainty", an uncertainty on the precision of measuring the energy of jets.

The description of the measurement done in this thesis starts in chapter 6. It explains the variables used, the transverse decay length of bottom hadrons and the transverse momentum of the lepton, both stemming from top quark decays. These variables have been designed to be as insensitive as possible to the knowledge of the jet energy scale, minimizing its impact on the measurement.

Finally, chapter 7 presents the actual measurement. It presents the selection of objects and events used in the analysis. It also describes the application of the techniques described in chapter 6 to the official ATLAS simulation. The largest part of the chapter is devoted to the discussion of the systematic uncertainties affecting the measurement. After that the full method is applied to the data taken by the ATLAS detector and the mass of the

top quark is extracted.

After that the work is summed up in chapter 8. It puts the result of this thesis into context and also discusses potential future steps.

1.1 Units and Conventions

This thesis uses natural units: This means that, for simplicity's sake, the values of the speed of light c and Plancks constant $\hbar$ have been set to unity: $c = \hbar = 1$. As a consequence, mass, m, momentum, p, and energy, E, all share the unit of energy. In high-energy physics energy is typically quantified in electron volts, eV. One electron volt is the energy an electron acquires when it is accelerated by a voltage of 1 V. This leads to $1\,\mathrm{eV} \approx 1.602 \cdot 10^{-19}\,\mathrm{J}$.

A few additional definitions are being used in this thesis: The coordinate system is cartesic and right-handed. Its origin is located right at the center of the ATLAS detector with the positive x-direction pointing towards the center of the LHC ring. The positive y-direction points upwards, letting the z-direction be parallel to the beam pipe. From these basic positions the following derived quantities can be derived and are more commonly used:

- The transverse momentum, $p_T = \sqrt{p_x^2 + p_y^2}$, measuring the momentum perpendicular to the z-direction.
- The azimuthal angle, $\varphi = \arctan\left(p_x/p_y\right)$, indicating the arc between the y-axis and the particle.
- The polar angle, $\theta = \operatorname{arccot}\left(\frac{p_z}{p_T}\right)$.
- The pseudorapidity, $\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$, being a measure of how parallel to the beampipe the object travelled.

The transverse momentum, the azimuthal angle and the pseudorapidity together fully describe the three-vector of a particle. Together with the energy of it, the full four-vector of it can be build. These quantities are used instead of the cartesic coordinates because they take into account the rotational symmetry of the detector more natively.

Chapter 2

The Standard Model of Particle Physics

The standard model of particle physics is a theory of particles and their interactions. It describes with very high precision all known particles and forces except gravity. It was majorly developed by Glashow [1] and later Weinberg [2] and Salam [3] in the 1960s. Since then it has been extended by several new discoveries but the basic principles have withstood experimental tests since then.

In this chapter an overview of the standard models constituents and interactions is given. First the building blocks of all known matter is described. Afterwards the different forces, their unification and the Higgs mechanism are discussed.

2.1 Building Blocks of Matter

All matter we experience in daily life consists out of three elementary particles: electrons, up-quarks and down-quarks. However, there exists more particles which to current knowledge are thought to be elementary. In this section an overview of the known building blocks of matter is given. It is loosely oriented at the historical order of discovery. Incidentally this is caused by the increasing energy experiments where able to achive and thus produce them in interactions. It is categorised by the type of particle: molecules and atoms, quarks and leptons.

2.1.1 Molecules and Atoms

The body of an average human with a weight of $m = 70\,\mathrm{kg}$ consists of roughly 10^{28} atoms. Most of these atoms are bound in molecules like water and various carbohydrates. The most abundant molecule in a human body is in fact water, making up 50% of all molecules [4].

A water molecule is made of three atoms: Two hydrogen atoms and one oxygen atom. The difference between the various atoms is the number of proton and neutrons in its core. The number of protons Z determines the type of the atom, i.e. $Z = 1$ is a hydrogen atom or $Z = 12$ a carbon atom.

Since protons are positively charged a core made only out of protons would be unstable as the electromagnetic force would break it apart. In fact the core is held together by the nuclear force which is related to the strong force. Its properties require there to be an additional numbers of neutrons N in the core to make it stable. The sum of the number of protons and neutrons in a core is called mass number $A = Z + N$.

Still, only certain combinations of Z and N are stable. The free neutron decays with a halftime of about ten minutes into a proton, an electron and a neutrino [5], a decay via

the weak nuclear force. In a core with similar number of protons and neutrons this decay is not possible energetically, however in a core with many more neutrons than protons it is and a neutron from this core decays, shifting the core towards more equal N and Z. A similar effect happens in a core with many more protons than neutrons.

2.1.2 Quarks

Looking closer into these decays it is not the protons and neutrons themselves that decay.

Protons and neutrons are made out of up- and down-quarks: A proton consists out of two up quarks and one down quark while the neutron consists out of one up and two down quarks.

A quark is an elementary particle discovered in deep-inelastic scattering experiments at the Stanford Linear Accelerator Complex (SLAC) in 1968 [6, 7]. They were earlier proposed by Gell-Mann in 1964 [8, 9]. Further experiments found the properties of these quarks: The up quark has a charge of $Q_u = +2/3\,\mathrm{e}$ with e being the elementary charge and a mass of $m_u = 2.3^{+0.7}_{-0.5}\,\mathrm{MeV}$ [5] in the $\overline{\mathrm{MS}}$ scheme which will be of particular interest later. The down quarks charge and mass were estimated to be $Q_d = -1/3\,\mathrm{e}$ and $m_d = 4.8^{+0.7}_{-0.3}\,\mathrm{MeV}$[5]. The spin of both of these quarks was found to be $1/2$.

The model proposed by Gell-Mann included a third quark, called the strange quark. It was introduced to explain the plethora of additional particles discovered in the the years from 1930 to 1950. Example for such particles are pions and kaons. Its properties share some striking similarities with the ones from the down quark. It has the same charge $Q_s = -1/3\,\mathrm{e}$ and same spin, the only difference seems to be its mass which in the $\overline{\mathrm{MS}}$ scheme is $m_s = 95 \pm 5\,\mathrm{MeV}$ [5].

One interesting property of quarks is that they do not exist in isolation [10]. All known quarks have been observed only in bound states of multiple quarks. This originates in the properties of the strong force which will be discussed in the next section.

With the discovery of the J/ψ particle in 1974 an additional quark was discovered: The charm quark. Like the up quark the charm quark has a charge of $Q_c = +2/3\,\mathrm{e}$ but its mass is several orders of magnitude larger: $m_c = 1.28 \pm 0.3\,\mathrm{GeV}$ [5].

These striking similarities between up and charm quark and respectively between down and strange quark in that they share all properties but their mass hint to an underlying symmetry. This lead to the characterisation of generations of matter. The up- and down quark form the first generation while the charm- and strange quark form the second generation of quark matter.

Using work from Cabibbo [11], Kobayashi and Maskawa predicted a third generation of quarks [12]. This prediction was confirmed by experiment with the discovery of the bottom quark in 1977 by the E288 experiment. The bottom quark has similar properties to the down- and strange quarks in that its properties like the charge ($Q_b = -1/3\,\mathrm{e}$) are the same. Only its mass is different, which was estimated to be $m_b = 4.18 \pm 0.03\,\mathrm{GeV}$ [5].

To complete this third generation of an additional quark is expected. It was named the "top quark" and a search over several decades for it was made until its discovery at the Tevatron collider in 1995 by the D0 and CDF experiments [13, 14]. As predicted the top quark shares all its properties except its mass with the up quark. Its spin was found to be $1/2$ and its charge $Q_t = +2/3\,\mathrm{e}$. In the $\overline{\mathrm{MS}}$ scheme its mass was estimated to be $m^{\overline{\mathrm{MS}}}_{\mathrm{Top}} = 160^{+5}_{-4}\,\mathrm{GeV}$ [5]. As the mass of the top quark is of big interest for this thesis further discussion of the mass of the top quark will be given in chapter 5.

So far no additional generations of quarks has been found. If there are further generation of quarks in existence their mass must be larger than $m_{t'} > 400 - 680\,\mathrm{GeV}$ [5], the limits depending on the assumptions made in the search.

2.1.3 Mesons and Baryons

As pointed out in the previous section, no free quark has been observed yet. Instead they only exist in bound states of a quark and an anti-quark or of either three quarks or three anti-quarks. A state consisting out of a quark and an anti-quark is called "Meson", the one with three quarks "Baryon". The proton and the neutron are examples of baryons, the previously mentioned pions and kaons are examples of mesons. However there exist many more mesons and baryons, with their properties depending on the content of quarks. Basically all combinations of quarks have been found to build mesons, the top quark being the one notably exception which will be discussed later.

As all quarks have a spin of $1/2$ all mesons have to have either spin 0 or spin 1 meaning that all mesons are bosons.

2.1.4 Leptons

In addition to quarks there exists another type of elementary particle building up all known matter, the leptons. In fact the first elementary particles to be discovered is a lepton: the electron. The electron also is the best studied elementary particle up to date, its properties are measured to very high precision: Its mass is $m_e = (0.510998928 \pm 0.000000011)\,\mathrm{MeV}$ [5] and its charge was measured to be $Q_e = (1.602176565 \pm 0.000000035) \times 10^{-19}\,\mathrm{C} \equiv 1e$ [5]. As the quarks it is a fermion with spin of $1/2$.

One production mechanism of electrons it the β-decay of unstable nuclei. Studying these decays it was observed that they produce an additional particle which is not electrically charged. Because of this it was called the neutrino or electron neutrino [15]. Its mass is so small that it was first estimated to be consisted with 0. To fulfil the conservation of angular momentum in the decay of the nuclei its spin also must be $1/2$. Electron and electron neutrino together are - similar to the up quark and down quark - called the first generation of leptonic matter.

Similar to the quarks two additional generations of leptons exist: The muon is a heavier version of the electron it has the same charge and spin but its mass is about 200 times larger than the one of the electron: $m_\mu = (105.6583715 \pm 0.0000035)\,\mathrm{MeV}$ [5]. The corresponding neutrino is called muon neutrino. The third generations charged lepton is the tau lepton and the tau neutrino. The tau leptons mass is again larger than the one of the muon: $m_\tau = (1776.82 \pm 0.16)\,\mathrm{MeV}$ [5].

With the observation of neutrino oscillations [16] - one neutrino changing its flavour to another one - the assumption of massless neutrinos turned out to be wrong. However, the absolute scale of the neutrino masses is still unknown. From studying the decay of tritium atoms it is known that the mass of the electron neutrino must be less than $m_{\nu_e} < 2\,\mathrm{eV}$ [5]. Neutrino oscillation experiments can measure the difference between the neutrino masses but not the absolute scale. They find that the difference of neutrino masses[1] is of the order of $10 - 50\,\mathrm{meV}$ [5].

2.2 On Interactions

2.2.1 The Electromagnetic Force

The electromagnetic force is – together with gravity – the first force discovered by humans. The first observations of it are most likely lightning or the magnetic effect of some minerals. During the ancient, medieval and renaissance eras, the electronic and magnetic effects were researched, but it was until the industrial age that a theory of electromagnetism describing all the observed effects was developed: Based on the work by Gauss, Coulomb and Faraday

[1] Technically these experiments measure the difference between neutrino mass eigenstates.

the british scientist Maxwell developed a set of equations which can describe all effects of classical electromagnetism [17]. In differential form they read:

$$\overrightarrow{\nabla} \cdot \overrightarrow{E} = \frac{\rho}{\epsilon_0} \tag{2.1}$$

$$\overrightarrow{\nabla} \times \overrightarrow{E} = -\frac{\partial \overrightarrow{B}}{\partial_t} \tag{2.2}$$

$$\overrightarrow{\nabla} \cdot \overrightarrow{B} = 0 \tag{2.3}$$

$$\overrightarrow{\nabla} \times \overrightarrow{B} = \mu_0 \overrightarrow{I} + \mu_0 \epsilon_0 \frac{\partial \overrightarrow{E}}{\partial_t} \tag{2.4}$$

$\overrightarrow{E}$ and $\overrightarrow{B}$ are the field strength of the electric and magnetic fields, ρ and $\overrightarrow{I}$ are the charge density and the electric current.

With the advent of quantum mechanics in the 20th century and the discovery that electromagnetic waves also have particle properties Planck and Einstein introduced a particle for these waves. This particle was later named photon. This later lead to the concept of the exchange particle which can be depicted the following way: A particle emits a photon and recoils against it. Later that photon hits another particle and this particle gets pushed away. The behaviour is similar to two particles repelling each other. The analogy breaks down as there is also an attractive force which would be a negative recoil in this picture. The photon has a mass of 0 and carries no electric charge. It is a boson with a spin of 1.

In the quantum field theory picture all interaction is described by a Lagrangian

$$\mathcal{L}_{QED} = i\overline{\Psi} \not{D} \Psi - m\overline{\Psi}\Psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{2.5}$$

Ψ denotes the fermion field of the interacting particle and m its mass. F is the field strength tensor and $\not{D}$ the gauge covariant derivative. It is needed to make the whole equation gauge invariant under the gauge transformation of the electromagnetic force. The beauty of this gauge covariant derivative is that it contains the interaction term:

$$\not{D} = \partial_\mu - iQA_\mu \tag{2.6}$$

with A_μ being the photon field. This means by just requiring the gauge invariance of the Lagrangian one obtains the interaction term of the corresponding interaction. This alone is the concept of the theory of gauge interactions. The strength of the coupling is giving by the constant in front of the photon field - in case of the electromagnetic force this means it couples to all particles having an non-zero electric charge.

2.2.2 The Weak Force

The weak force is the force which is the reason for the β decay of nuclei.

The charge of the weak force is called weak isospin I_3. All quarks and leptons have a weak isospin of either $+1/2$ or $-1/2$. The structure is that each generation has a particle with a weak isospin of $+1/2$ and one with $-1/2$, i.e. the up quark has an isospin of $+1/2$ and the bottom quark has one of $-1/2$. All generations therefore can be organised in a doublet of the form $\binom{u}{d}$ where the upper particle has an isospin of 1 more than the lower one.

The weak force manifests itself in two different types: One transforms a particle with negative isospin to one with positive, like the weak decay of a down quark in a neutron to an up quark, transforming the neutron to a proton. The resulting electron and electron neutrino are again part of a weak isospin doublet. The weak isospin is a conserved quantity

in all interactions. This means that the exchange particle of this manifestation has to carry weak and electric charge. The corresponding particle is the $W^{\pm}$ boson. Unlike the photon it is not massless, in fact its mass is with $m_W = 80.385 \pm 0.015\,\mathrm{GeV}$ [5] quite large. This also explains the relative weakness of this force in comparison to the electromagnetic force - a muon with its comparatively low mass can not directly produce a W boson. The decay has to involve a virtual W boson which in the language of quantum field theory suppresses the interaction probability. The spin of the W boson is 1.

The other manifestation is in the form of neutral currents, e.g. meaning that an electron scatters of an electron neutrino or pair-production of a weakly charged particle and its antiparticle. This requires another exchange particle to be present in the weak force. It is called the Z boson. It does carry neither electrical charge nor weak isospin. Its mass is $m_Z = 91.1876 \pm 0.0021\,\mathrm{GeV}$ [5].

The Lagrangian for the weak force is the following:

$$\mathcal{L}_W = i\overline{\Psi}\not{D}\Psi - \frac{1}{4}W^{a\mu\nu}W^a_{\mu\nu} \tag{2.7}$$

The W^a denote the three boson fields in the weak force and $\not{D}$ the gauge covariant derivative for the weak force.

One interesting feature of the weak force is that it couples differently to particles and anti-particles: It only interacts with right-handed particles and left-handed antiparticles [18]. The handedness of a particle is described by the projection of its spin to its direction of movement: If they point into the same direction it is called right-handed, if they point in opposite direction it is called left-handed.

The weak force was found to not couple to the quarks as they were described in section 2.1.2. Instead there is a set of separate eigenstates, called weak eigenstates. The eigenstates described in the previous section are called the mass eigenstates. By definition the weak eigenstates for up, charm and top quark are the same as their mass eigenstates. The transformation for down, strange and bottom quark is described by a matrix proposed by Kobayashi and Maskawa when they predicted the existence of the bottom and top quarks [12]. It has the following elements and absolute values [19]:

$$|V_{CKM}| = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97425^{+0.00022}_{-0.00014} & 0.22543^{+0.00060}_{-0.00094} & 0.003490^{+0.000209}_{-0.000094} \\ 0.22529^{+0.00059}_{-0.00095} & 0.97344^{+0.00022}_{-0.00018} & 0.04077^{+0.00130}_{-0.00048} \\ 0.00851^{+0.00047}_{-0.00023} & 0.04002^{+0.00129}_{-0.00048} & 0.999142^{+0.000043}_{-0.000025} \end{pmatrix}$$

It is a complex unitary matrix. It is diagonal-dominant meaning that mostly the weak eigenstates are the same as the mass eigenstates. However there is a small admixture of the other mass eigenstates which makes changes of quarks between the generations possible, e.g. a bottom quark can decay into a charm quark. The CKM-Matrix has a complex phase which is not shown here. It leads to the possibility of CP violation in the standard model.

A similar matrix exists for the three generation of leptons. Here the weak and mass eigenstates are the same for the charged leptons and there exists a similar matrix for the neutrino eigenstates, the PMNS-Matrix [20, 21]. For the neutrinos it is not possible to observe the mass eigenstates directly - they only interact via the weak interaction. This leads to the feature of neutrino oscillations - one flavor of neutrino changing its type to another one.

2.2.3 The Strong Force

The strong force couples to another quantity of particles that is called colour. Unlike charge and weak isospin where there is one type of charge, colour has three. They are called red, blue and green. From the so far discussed particles, only quarks carry colour

charge. This has the consequence that there are in fact more quarks than expected from interactions of the previous forces: There is one type of quark for each colour, e.g. a red up quark, a blue up quark, a green up-quark and the same for all the other quarks. This makes the interaction strength of the other forces seem stronger to quarks than it actually is. The Lagrangian of the strong force reads:

$$\mathcal{L} = i\overline{\Psi}\not{D}\Psi - \frac{1}{4}G^{a\mu\nu}G^{a}_{\mu\nu} \tag{2.8}$$

As for the other colours there are an exchange particles for the strong force. In the above equation their fields are denoted G^a. They are called gluons and carry a colour charge as well. It is of the form colour/anticolor (e.g. red/anti-green). This again leads to the possibility of self-interaction of gluons. In total there are eight different combinations of that form which leads to the existence of eight different gluons.

The interaction strength of the strong force can become so large that the theory becomes non-pertubative. In these cases other approaches for describing the interactions need to be used.

2.3 The Standard Model of Particle Physics

The Standard Model of Particle Physics is a model to describe all known particles and their interactions. It contains in total 24 elementary particles: 18 quarks (six for each colour) and 6 leptons. In addition there are the exchange particles of the three forces of the standard model: Electromagnetic Force, Weak Force and Strong Force. The Standard Model unifies the electromagnetic force and the weak force together to the Electroweak force: In the Electroweak force there are four bosons, W_1, W_2, W_3 and B. The three W bosons still couple to the weak isospin, however the B boson couples to the hypercharge Y. These charges are related to each other in the following way:

$$Y = 2(Q - I_3) \tag{2.9}$$

The Z boson and the photon are linear combinations of the W_3 and the B boson in the way that the photon does only couple to charged particles.

2.3.1 The Higgs Mechanism and Boson

One final piece exists in the standard model: The Higgs boson. It is a consequence of the Higgs field which is added to the standard model to describe the behaviour of mass terms. A simple mass current in the Lagrangian $m\overline{\Psi}\Psi$ is not renormalizable. However its coupling with a scalar boson field $\Phi\overline{\Psi}\Psi$ is. Such a coupling is introduced with the Higgs field. The masses of all particles are therefore generated by their interaction with the Higgs field and are proportional to their interaction strength.

The interaction of the Higgs field with itself has an additional consequence: The existence of another boson, the Higgs boson [22, 23]. This has been on of the major predictions of the standard model. A boson with similar properties as the standard models Higgs boson was discovered in 2012 by the ATLAS and CMS collaborations at the LHC [24, 25]. First estimations of the mass are in the 125 GeV region [24, 25].

2.4 The Top Quark

With its very high mass the top quark plays a special role in the standard model. It is the only known particle which can decay into a real W boson, as it has a higher mass than the latter. This leads to the consequence that the average lifetime of a top quark is

$\tau_t = 0.5 \times 10^{-24}$ s [26]. This has the interesting consequence that the top quark decays before it forms mesons or baryons with other quarks – for this very short time the top acts as a free quark. Since the CKM-Matrix element V_{tb} is very close to unity the top quark is expected to almost always decay into a bottom quark and a W boson $t \to bW$.

The very high mass also makes the top quarks mass a very important parameter in the standard-model. There exist predictions for the Higgs mass depending mostly on the W boson mass and the top mass. An example is shown in Figure 2.1. The green areas denote the measured values for the masses W boson and top quark. Where they intersect, the allowed region for the mass of the Higgs boson by these measurements exists. The dashed diagonal lines are the allowed diagonals for a given mass of the Higgs boson. The blue- and grey-shaded areas are best-fit values for the Higgs boson using various input parameters. Excellent agreement is observed, thus the mass of the top quark is a important parameter for the consistency of the standard model.

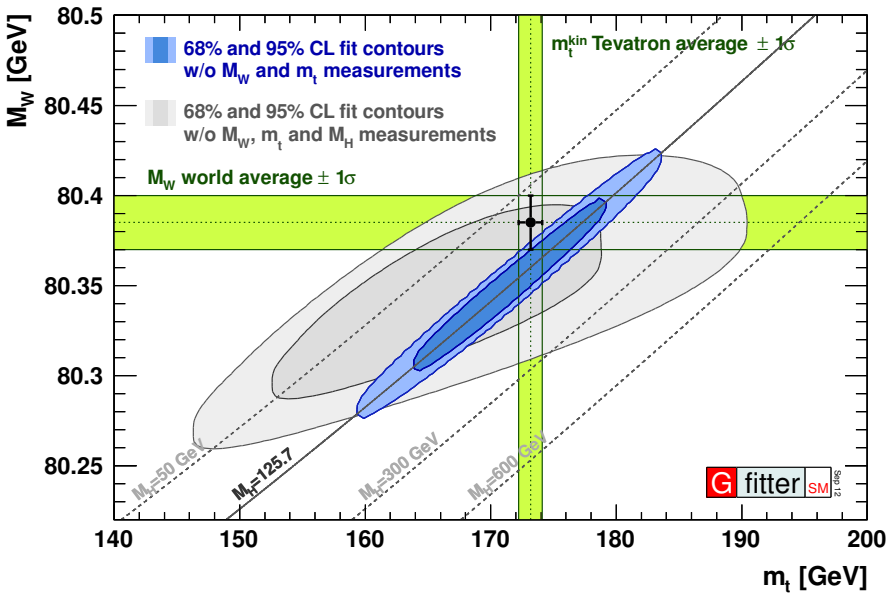

Figure 2.1: Prediction of the Higgs mass based on W boson mass and top quark mass measurements together with the measurement of the Higgs boson mass assuming that it is the standard models Higgs boson. The correlation between top and W boson mass is neglected in this figure [27].

Chapter 3

The ATLAS Experiment

3.1 CERN and LHC

Located at the Swiss-french border near Geneva, CERN is the worlds largest center for research in particle physics. It houses several experiments, present and past. At CERN several thousand scientists and engineers from a multitude of nations work together on answering the fundamental questions of the universe. The discovery of the W- and Z-bosons in 1983 and the discovery of a boson with consistent properties as the standard models Higgs boson in 2012 belong to its long list of discoveries.

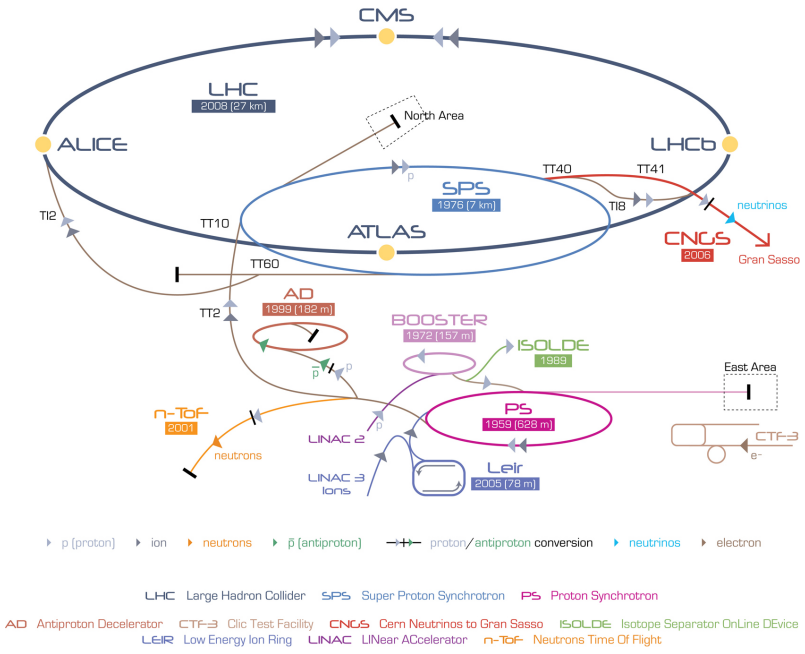

Figure 3.1: An overview of the accelerator complex at CERN. It includes the whole chain from the extraction of protons in LINAC2 its way through the BOOSTER, PS and SPS until they finally arrive in the LHC. It also includes several additional accelerator and storage rings and other experiments, i.e. CNGS, a neutrino beam travelling from CERN to the LNGS facility in Italy [28].

Figure 3.1 shows a sketch of the of the accelerator complex at CERN. It includes several accelerators and experiments, like the anti-proton decelerator. It also shows the Large Hadron Collider and its four big experiments.

3.1.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the most powerful particle accelerator built until today. In it protons have been accelerated up to $\sqrt{s} = 8\,\text{TeV}$. In the future it will be possible to increase that energy up to $\sqrt{s} = 14\,\text{TeV}$. It houses four big and two small experiments:

- ALICE (A Large Ion Colliding Experiment) mainly studies the collisions of lead ions which also can be accelerated in the LHC
- CMS (Central Muon Solenoid) is one of the two general-purpose experiments.
- LHCb was built to do precision measurements of the decays and properties of B-Hadrons
- ATLAS (A Toroidal LHC Apparatus) is the other general-purpose experiment. Like CMS the main goals are the discovery and study of the Higgs Boson, precision measurements of top quark properties and the general search for new phenomena
- TOTEM (Total Elastic and Diffractive Cross-Section Measurement) is one of the smaller experiments. It wants to measure the total cross-section of proton-proton interactions
- LHCf studies ultra-forward parton showers to enhance the performance of cosmic-ray simulation programs.

The LHC is designed to deliver a luminosity of $L = 10^{34}\,{}^1/_{\text{cm}^2\text{s}}$ at $\sqrt{s} = 14\,\text{TeV}$. In its $\sqrt{s} = 7\,\text{TeV}$ runs in 2010 and 2011 and its $\sqrt{s} = 8\,\text{TeV}$ run in 2012 the peak luminosity to ATLAS was around $L_{peak} \approx 8 \cdot 10^{33}\,{}^1/_{\text{cm}^2\text{s}}$ (cp. Figure 3.2).

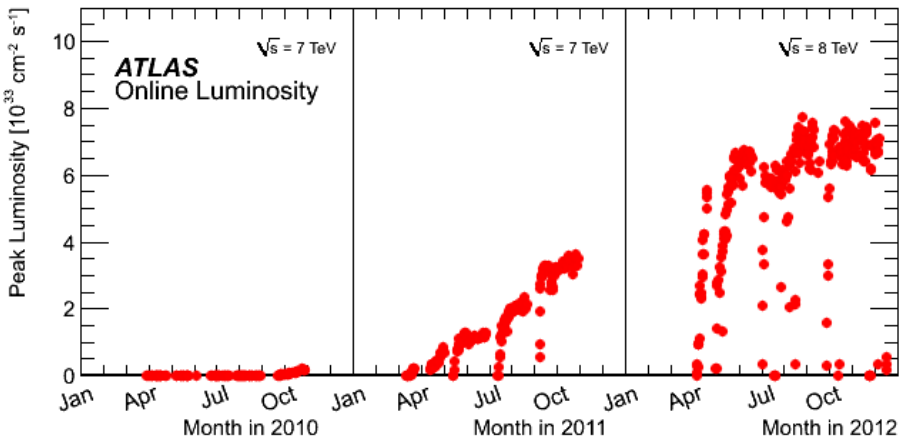

Figure 3.2: The peak luminosity delivered to the ATLAS experiment versus the time of the runs in 2010-2012. The luminosity was steadily increased in 2011 and levelled out early in the 2012 run [29].

This high luminosity leads to the effect that multiple interactions happen during each crossing of two bunches. This effect is called "pileup". A measure for it is the average number of interactions per bunch-crossing, $\langle\mu\rangle$. Figure 3.3 shows the distribution of $\langle\mu\rangle$ in the 2011 and 2012 runs of the LHC.

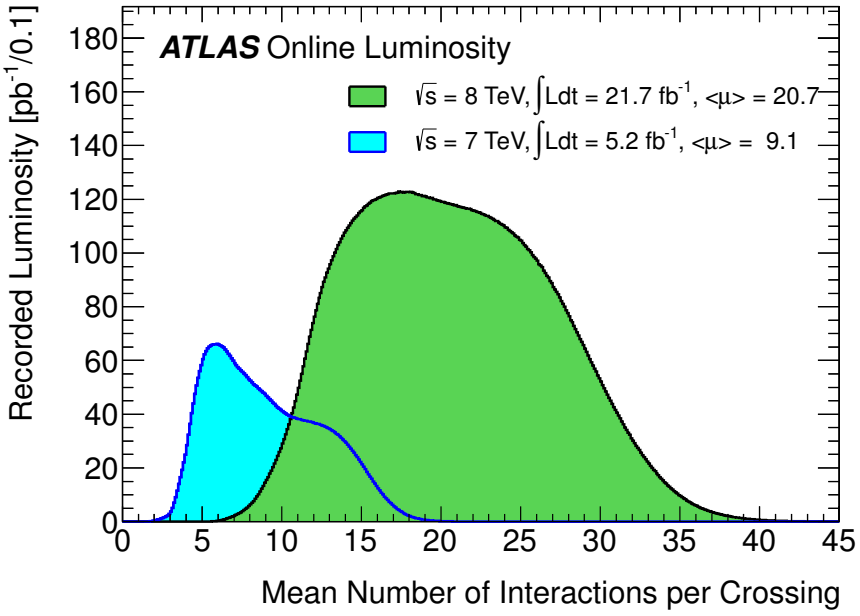

Figure 3.3: The mean number of interactions per bunch-crossing as observed by the ATLAS experiment for the 2011 and 2012 runs. [29].

3.2 The ATLAS Detector

The ATLAS detector is one of the four large experiments operating at the LHC. It is the largest particle physics detector ever built at a particle accelerator. Its dimensions are 44 m in length and 22 m in height. In total it has approximately one hundred million readout channels. It has been built and is operated by an international collaboration of around four thousand scientists.

The main goals of the ATLAS experiment are the discovery and study of the standard model Higgs boson, precision measurement of properties of the standard model, particularly the top quark, and a broad search for physics beyond the standard model.

Figure 3.4 shows a sketch of the detector. It contains roughly out of the inner detector, calorimeters, the muon system and a system of magnetic fields. Additionally each of these components plays a part in the trigger decision. A description of all of these parts will be given in the following sections.

3.2.1 Inner Detector

The inner detector of ATLAS consists out of a series of tracking systems inside a solenoidal magnetic field with a field strength of $B_{ID} = 2\,\mathrm{T}$ [30]. The tracking detectors detect the passing of electromagnetically charged particles. This allows the reconstruction of tracks by combining the hits in the several detectors. Charged particles are bent by the magnetic field which makes their momentum and sign of charge measurable by the curvature the reconstructed track.

The pixel detector

By far the most readout channels of the ATLAS detector are concentrated in the pixel detector. Having its first layer, the b-layer, at only 50.5 mm radial distance from the interaction point the amount of cells needed to obtain sufficient resolution needs to be very high. A typical pixel has a size of $50\,\mu\mathrm{m} \times 400\,\mu\mathrm{m}$ in the φ-z-plane. The pixel

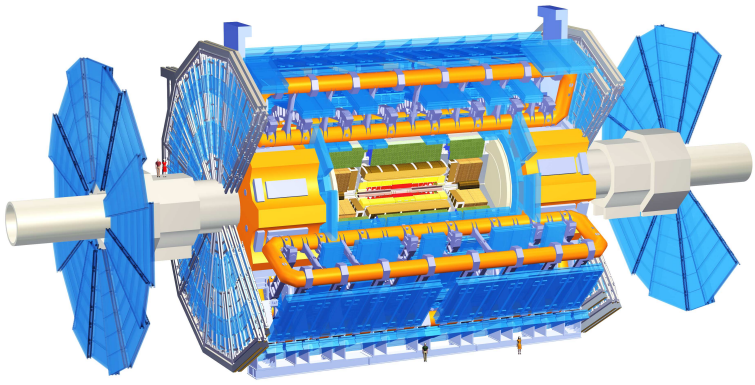

Figure 3.4: A sketch of the ATLAS detector. From the innermost layer to the outermost it consists out of the PIXEL detector, the semiconductor tracker (SCT), the transition radiation tracker (TRT), followed by the electromagnetic calorimeter (LAr) and the hadronic calorimeter (TILE). The outermost part is the muon system consisting out of thin gap chambers (TGC), monitored drift tubes (MDT), resistive plate chambers (RPC) and cathode strip chambers (CSC). Between the inner detector (PIXEL, SCT, TRT) and the electromagnetic calorimeter is a solenoid magnet. The muon system is inside a toroidal magnet field generated by the large coils visible in the sketch [30].

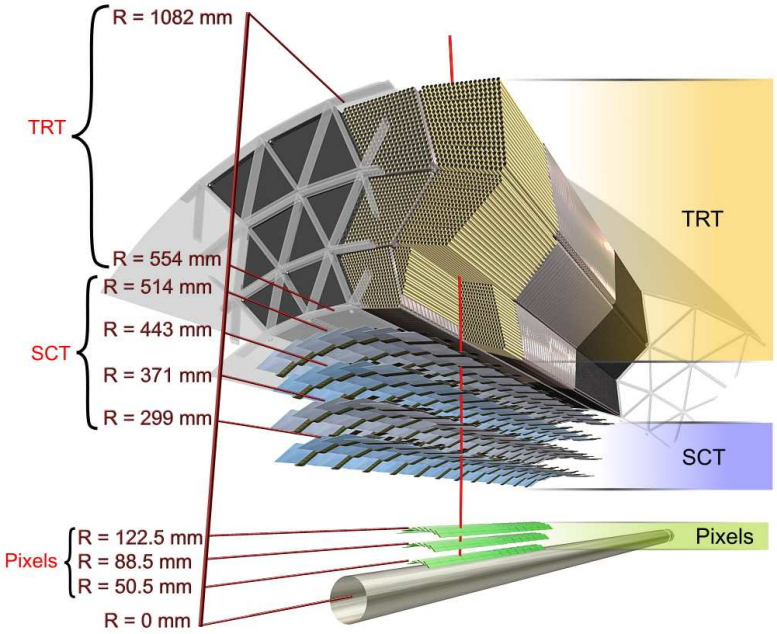

Figure 3.5: The inner detector of the ATLAS experiment. At $R = 50.5\,\text{mm}$, $R = 88.5\,\text{mm}$ and $R = 122.5\,\text{mm}$ are the three layers of the PIXEL sub-detector. The SCT has layers at $R = 299\,\text{mm}$, $R = 371\,\text{mm}$, $R = 443\,\text{mm}$ and $R = 514\,\text{mm}$. The volume of the TRT is between $R = 554\,\text{mm}$ and $R = 1082\,\text{mm}$ [30].

detector is a silicon-based semiconductor detector. A passing charged particle will create electron-hole pairs which form a current if sufficient voltage is applied to the sensor. The presence of this current can then be interpreted as a hit in that pixel. The size of the current is proportional to the energy loss of the charged particle in the sensor. As a typical track will pass multiple pixels a resolution of 10 μm can be achieved in the φ-direction [30].

Apart from the b-layer there are two more pixel layers, layer 1 at $R = 88.5\,\text{mm}$ and layer 2 at $R = 122.5\,\text{mm}$. Each of the layers spans 400.5 mm in both directions in z. Additionally there are two sets of three disks perpendicular to the beam on each side of the other layers. Their positions in z are 495 mm, 580 mm and 650 mm. In total the pixel detector consists out of 1744 modules, 286 in the b-layer, 494 in layer 1, 676 in layer to and 48 in each of the 6 disks. Each module has a size of $19\,\text{mm} \times 63\,\text{mm}$ and consists out of 46080 read-out channels.

The semiconductor tracker

The semiconductor tracker is another silicon-based tracking detector. Other then the pixel detector it does not contain out of single pixels but out of 12 cm long strips. Two of these strips form one of the in total 4088 modules. To retain resolution perpendicular to the strip the two modules on a strip are inclined at an angle of 2.29°. Thus, a resolution of 580 μm in the z-direction is achieved. The size of the strip in the R-φ-direction is 80 μm resulting in a resolution of 17 μm .

There are four barrels of SCT modules, they exist at $R = 299\,\text{mm}$, $R = 371\,\text{mm}$, $R = 443\,\text{mm}$ and $R = 514\,\text{mm}$. Additionally there are – similar to the pixel – end-caps at various directions in z [30].

The transition radiation tracker

The transition radiation tracker uses a different technology to detect passing charged particles. It consists out of straws filled with an admixture of xenon and carbon dioxide. A passing charged particle will emit transition radiation at the boundary. That radiation is then ionising the gas thereby creating free electrons. Electrons are collected at an anode inside the straw. The current generated by these electrons can be used to generate a hit in that straw. Additionally, as the number of electrons generated is proportional to the γ-factor the TRT can be used to identify the type of the particle.

A single straw consists out of a 35 μm thick layer of films: the base is a 25 μm thick polyimide film. Inside that film is – after a thin 0.2 μm Al layer a 6 μm thick film of graphite-polyimide. On the other side of that film is a 5 μm thick layer made out of polyurethane. The whole TRT spans the area between $R = 554\,\text{mm}$ and $R = 1082\,\text{mm}$. In total it has 73 layers of straws in the barrel and additional 160 planes in the end-cap [30].

3.2.2 The Calorimeters

Calorimeters are used to absorb all of the energy of passing particles. This is done by putting large amounts of absorber materials in the direction of flight of these particles. This will result in the generation of particle showers inside the calorimeter. The energy of these showers then can be measured by measuring the energy the shower particles deposit inside the calorimeter.

The ATLAS experiments employs two different calorimeters – an electromagnetic one and a hadronic one. Both of them are sampling calorimeters, i.e. they consist out of a material where the particles shower and another one which actually measures the deposited energy. The electromagnetic calorimeter uses lead as the showering material and liquid

argon as the absorber. The hadronic calorimeter uses different technologies in the barrel and the end-cap.

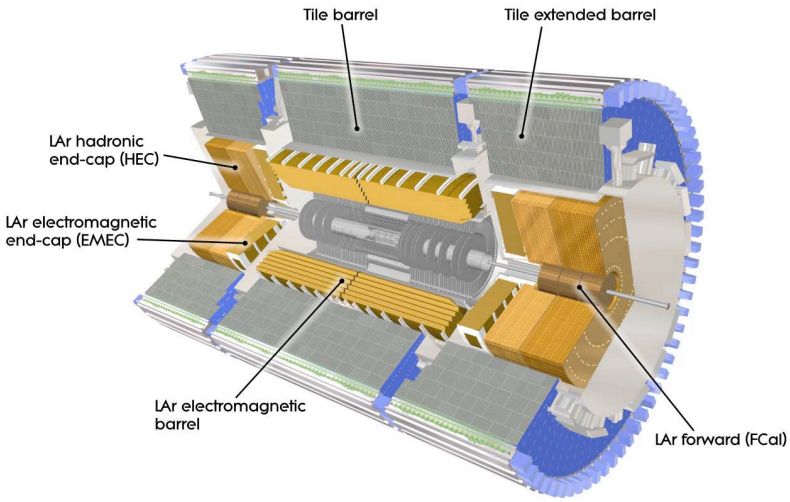

Figure 3.6: A sketch of the calorimeters of the ATLAS detector. It consists out of the LAr electromagnetic calorimeter and the Tile hadronic calorimeter. The LAr calorimeter consists out of a barrel and two forward calorimeters while the Tile consists out of barrel and an extended barrel. Additionally the LAr calorimeter has a electromagnetic and a hadronic end-cap, resting inside the Tile extended barrel [30].

The electromagnetic calorimeter

The electromagnetic calorimeter of ATLAS uses liquid argon as the active material. Hence, it is referred to as the LAr calorimeter. Passing electromagnetically interacting particles, particularly electrons and photons, interact with the lead absorber and generate a shower of secondary electromagnetic particles. The energy of these particles is then measured inside the active material, and thus the energy of the original particle can be measured. Depending on the pseudorapidity of the particle it consists of 25-45 radiation lengths of material.

The LAr calorimeter barrel is built in three layers with various cell sizes: The first one ranges from $X_0 = 1.7 - 4.3$, the second from there until $X_0 = 20.3$. After that there is a trigger tower with a radiation length of $X_0 = 2$. The LAr cells are constructed in an accordion-like structure. Additionally they are structured into cells of various sizes: In layer 1 they are strips of $147.3\,\text{mm} \times 4.69\,\text{mm}$. Layer 2 consists out of square cells and layer three out of towers [30].

Additionally there is an end-cap for the electromagnetic calorimeter. It employs similar geometry as the barrel but spans a larger amount of pseudorapidity. In between the barrel and the end-cap there is the so-called “crack-region” where no calorimeter material exists. It spans from $1.37 < |\eta| < 1.52$.

The hadronic calorimeters

The hadronic calorimeter of ATLAS consists out of three parts: The Tile calorimeter, which makes up the bulk of the hadronic calorimeters, the LAr hadronic end-cap and the LAr forward calorimeter. Each of the three parts overlaps with at least one of the others to eliminate gaps in the coverage.

The tile calorimeter is a sampling calorimeter. Other than the LAr calorimeters it uses steel as the absorbers and scintillating tiles as the active material. It consists out

of a barrel spanning up to a pseudorapidity of $|\eta| < 1.0$ and two extended barrels in the $0.8 < |\eta| < 1.7$ region. Both types of barrels are segmented into three layers with various radiation lengths, in total 7.4 independent of the barrel type [30].

The LAr hadronic end-cap – as the electromagnetic calorimeter – uses liquid argon as the active material. Copper is used as the showering material. The hadronic end-cap spans from $|\eta| = 1.5$ until $|\eta| = 3.2$ and is built out of two wheels per side [30].

Finally there is a hadronic LAr calorimeter, consisting out of several absorbers: One layer of copper and two layers of tungsten, again using liquid argon as the active material. It has approximately 10 radiation lengths and covers a pseudorapidity of up to $|\eta| = 4.5$, starting at $|\eta| = 3.1$ [30].

3.2.3 Muon System

The ATLAS muon spectrometer is another series of tracking stations. Located behind the calorimeters muons are basically the only charged particles which will travel until there. The muon system consists out of four different components. Two have finer granularity and are used for the actual reconstruction, while the other two are less granular but much faster and are mainly used in the trigger system.

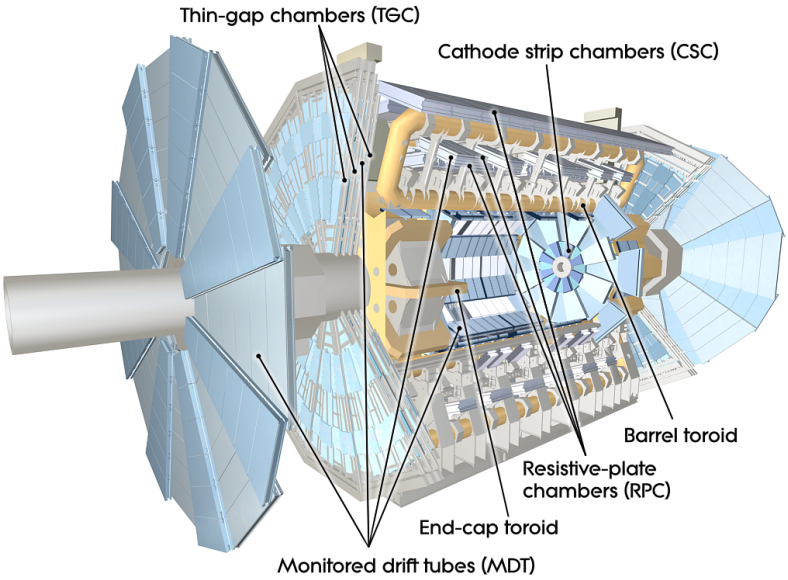

Figure 3.7: A sketch of the ATLAS muon system. The special toroid magnets are depicted in yellow, while the various parts of the system are shown in blue.[30].

The monitored drift tubes make up the bulk of the precision tracking channels. A passing charged particle will interact with the argon-carbon-dioxide gas inside the chamber. The produced electrons drift towards a wire inside the chamber, generating a current. In total there are 1150 chambers with a total of 354,000 channels. They are arranged in three layers at $R = 5\,\text{m}$, $R = 7.5\,\text{m}$ and $R = 10\,\text{m}$. They cover a η-region up to 2.7 (2.0 in the first layer). Monitored drift tubes are also used in the big wheels. They provide a resolution varying from $50\,\mu\text{m}$ to $300\,\mu\text{m}$ depending on the impact parameter of the track w.r.t. to the wire [30].

In the small wheels a different technology is used: They consist out of cathode strip chambers. They are multiwire proportional chambers. They reach a resolution of around $60\,\mu\text{m}$. Each passing track passes through at least three chambers providing sufficient amount of information for momentum and charge reconstruction [30].

The detectors used for the trigger system are called resistive plate chambers and thin-gap chambers. The former are mainly used in the barrel region while the latter in the wheels. The resistive plate chambers consist out of two parallel plates with a gas filled inside them. A passing particle will generate secondary particles which then can again be measured as a current at the plates. The thin-gap chambers are multiwire proportional chambers like the CSCs. The difference between the RPCs/TGCs with respect to the MDTs/CSCs is that the size of the chambers is much smaller. Therefore the time until the signal reaches the electrodes is much shorter. This allows them to be fast enough for the triggering decision.

The muon system is filled with an air-cored toroidal magnet field with a varying field strength. This will still bend a muon with momentum of more then 1 TeV on a trajectory sufficient for momentum and charge reconstruction.

3.2.4 The Trigger

Without a trigger system the amount of data the ATLAS detector would have to store is gigantic. The unfiltered data rate would be in the order of $^{\mathrm{PByte}}/_{\mathrm{s}}$ [31]. In 2007 the worlds total ability to store data was estimated to be around 290 EByte [32]. If all data of ATLAS would be stored the worlds present ability to store it would run out after about 80 days[1]. Even with the rapid increase in computing technology an efficient system to reduce the data rate is therefore imperative. The most likely interactions when to protons collide have been thoroughly studied at previous experiments. Thus, they can safely be discarded. The trigger system of the ATLAS detector is designed to achieve just that: It rapidly decides whether an event contains a interesting signature and otherwise discards it.

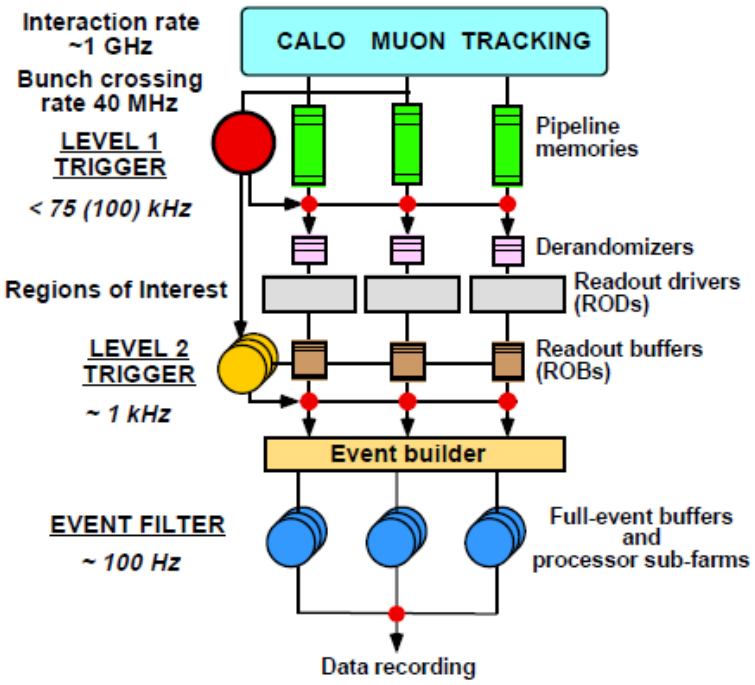

Figure 3.8: Schematic view of the ATLAS trigger system[30].

The ATLAS trigger chain has three big components: level 1 (L1), level 2 (L2) and the event filter. In each step the amount of information used and the processing time

[1]This of course neglects the increase in storage capability since then.

increase. While L1 must make a decision within 2.5 μs, the event filter may take up to several seconds to decide whether an event is kept or not [30].

The level 1 trigger is seeded by the calorimeters and the trigger-specific parts of the muon system. The electromagnetic and hadronic calorimeters look for the deposition of large amounts of energy in a small, but less granular, region. Several thresholds are programmable for the amount of energy, allowing it to be tuned to the instantaneous luminosity. Triggers for jets, electrons, photons and tau leptons are in place. The second layer of RPCs is used to seed muon triggers in the barrel. From there a line to the interaction point is defined and a hit in the first RPC layer is searched. The deviation of that point from the line is a measure for the transverse momentum of the muon. For high transverse momentum muons additionally the third layer of RPC is used. A similar technique is used in the endcap, but here the third plane of TGC is the pivot plane.

If an event is selected by a level 1 trigger the information is passed to level 2. This information includes a 'region of interest', the physical location of the object which fired the trigger. The information of all sub-detectors is stored in so-called read-out-buffers. First, the level 2 trigger verifies the level 1 trigger object in the region of interest, using more sophisticated information. For example, in the muon system a simple circular fit is performed to measure the transverse momentum of the object. If the verification is successful additional information of other sub-detectors is used. At this level also information of the inner detector can be used, allowing the use of triggers identifying jets originating from bottom-quarks. The event filter is the last step of the trigger chain. It fully reconstructs the event allowing for a large variety of possible selections.

3.2.5 The Analysis Chain

An event which passes the even filter trigger requirement is then stored on disk. The computing chain starts at "Tier0", a large computing farm at CERN. At this stage the data is in byte-stream form. At Tier0 the byte stream is converted into the "RAW" format, which is basically a compressed byte stream. On Tier0 the data is then reconstructed within 24 to 48 h after the run. The reconstructed data is stored as "Event Summary Data" (ESD). This format contains the basic information of all sub detectors, enough to potentially rerun the reconstruction. Another output is the "Analysis Object Data" (AOD). They only contain reconstructed objects like tracks, electrons or jets and are intended to have all the necessary information for any analysis. For more specific analysis needs the D3PD format exists. It contains a tree in the ROOT-format [33] including only the information the performance or physics group in question needs. Specific analysis teams usually define even smaller ROOT-trees for their specific analyses. These usually apply the final calibrations and perform a very basic event selection.

3.2.6 Data Quality

The detector will not at all times perform in perfect conditions. Even though the operation team of the detector is working at all times, hardware or software failures will happen from time to time. Also at the beginning or end of each run parts of the detector may still be ramping their high-voltage before they actually take data. The data from these times may not be usable for data analysis. This depends on the needs of the analysis and the severity of the defect. In ATLAS, each subsystem and reconstruction group flags all the data, adding so-called "defects" to them. A defect may be that the high-voltage of the PIXEL detector is off for the first minutes of data-taking or that there is a hot-spot somewhere in the calorimeter. The granularity in time for these defects is called a "luminosity-block" and is typically one minute long. Each physics analysis creates a list of defects they deem as intolerable for their analysis. The result is a new list, called "good runs list" (GRL)

which contains a list of run numbers and luminosity blocks which are usable for their analysis.

3.2.7 Simulation

Many physics processes are simulated to estimate their contributions to various analysis. For this, a Monte-Carlo generator amongst other software is used. The details of this will be discussed in chapter 4. The detector response of the simulated events also needs to be simulated. A common tool for this is the Geant4-package [34]. It simulates the interaction of matter with detector material. It is used in the simulation of the response of the ATLAS detector to simulated events. Two modes of operation are used in this simulation:

- Full Simulation. This does simulate the whole particle interactions with detector elements in the optimal way of the toolkit. The simulation of the interaction in the calorimeters uses a large amount of time.
- Fast Simulation. This replaces the full calorimeter shower interaction with presimulated showers. It is considerably faster then the full simulation but slightly less accurate.

The output of the Geant4-simulation is very similar to the RAW data format of the real detector. It also contains information about the truth input of the generator. It can be fed into the default analysis chain as if it was taken by the real detector.

3.3 Reconstruction of Physics Objects

The information of the detector needs to be interpreted as physical objects: A muon will leave hits in the inner detector, deposit very little energy in the calorimeters and then leave hits in the muon chambers. However, the information one obtains from the detector is backwards: One has a couple of hits in the inner detector and in the muon spectrometer in vaguely the same region of the detector and a small deposition in the calorimeter. This information now needs to be combined to reconstruct the original muon. This must be done while there are several additional physics objects leaving their signatures in the detector. Even more challenging, there are objects which will show a very similar signature as a muon but are in fact something else. An example would be a charged kaon which will leave hits in the tracker and deposit energy in the calorimeter. There it may decay into an actual muon which will leave hits in the muon system. Any reconstruction algorithm must be able to discriminate as much as possible between the desired physics object and the background sources while retaining a high reconstruction efficiency.

3.3.1 Tracking

Tracking is the reconstruction of the traces of charged particles. This is used for the reconstruction of the actual interaction points, as all particles from a single interaction must originate from the point of that interaction. Tracking is also used in the identification of higher-level objects, for instance electrons and photons have similar signatures except that electrons are charged and therefore leave a track while photons do not.

The actual tracking starts from the hits in the pixel and SCT. These are combined into 'clusters' inside a single layer. These clusters exist because a particle will not always pass a pixel/SCT plane perpendicularly. The tracking uses these clusters as seeds for tracks and extends them to the TRT information [35]. After all tracks have been found they are combined to find the primary vertex of the event. Because of pileup there may be more than one primary vertex and by association of physics objects with the signal primary vertex background from pileup can be suppressed.

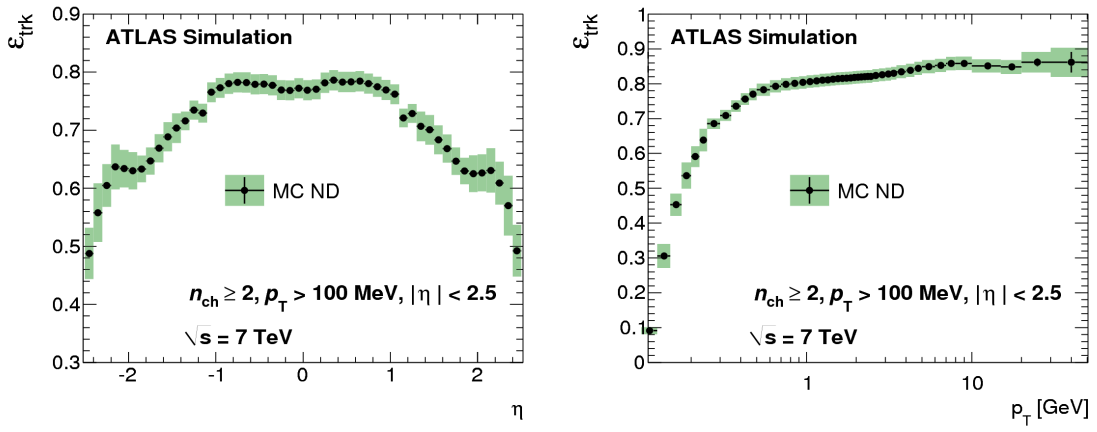

Figure 3.9: The track reconstruction efficiency as a function of track η and p_T. The more central a track is and the higher its momentum is, the higher its reconstruction efficiency is [35].

A very important quantity is the reconstruction efficiency of tracks. This efficiency may depend on the pseudorapidity of the track, as the particle will pass through different parts of the detector. Also it will depend on the transverse momentum of the track, if it is too low the particle will not escape the magnetic field. The track reconstruction efficiency using simulated data is shown in Figure 3.9. A track with $p_T > 1\,\text{GeV}$ has, in general, a reconstruction efficiency of larger than 80%. Also, the more central it is, the higher its reconstruction efficiency is.

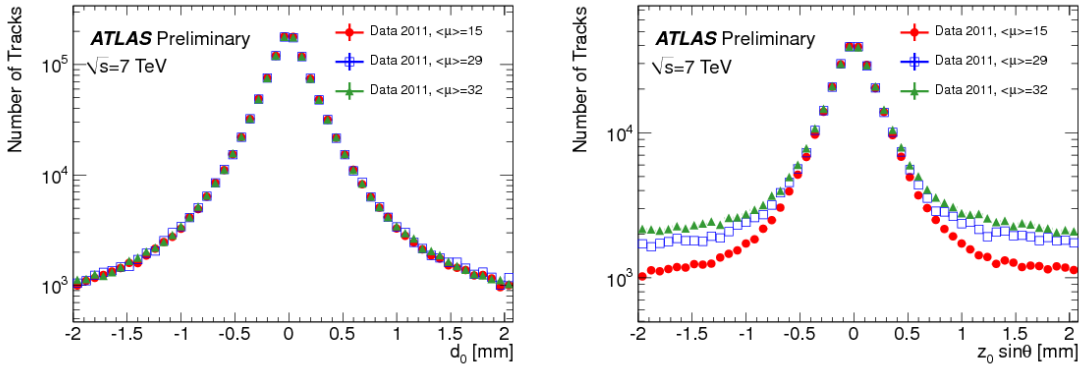

Figure 3.10: The impact parameter distributions for various average number of interactions $\langle\mu\rangle$ measured by the ATLAS collaboration in 2011 high-luminosity data. left: transverse impact parameter, right: longitudinal impact parameter [36].

Figure 3.10 shows the impact parameter distributions of tracks reconstructed with the ATLAS inner detector. An impact parameter is the minimum distance between the track and its primary vertex, the width of the distribution is a measure for the tracking resolution. The figure shows the impact parameter distributions for various number of average proton-proton-interactions per bunch-crossing $\langle\mu\rangle$ which is a measure for the amount of pileup. The transverse impact parameter distribution is very stable against the amount of pileup which is as desired. The longitudinal impact parameter distribution becomes slightly broader with increasing pile-up, the effect is small though [35].

Another important part of tracking is the efficiency of primary vertex reconstruction, i.e. the probability to find a primary vertex if it exists. The efficiency as a function of number of tracks in the beamspot is shown in Figure 3.11. In general the efficiency is very

high, being basically unity for more than three tracks [35].

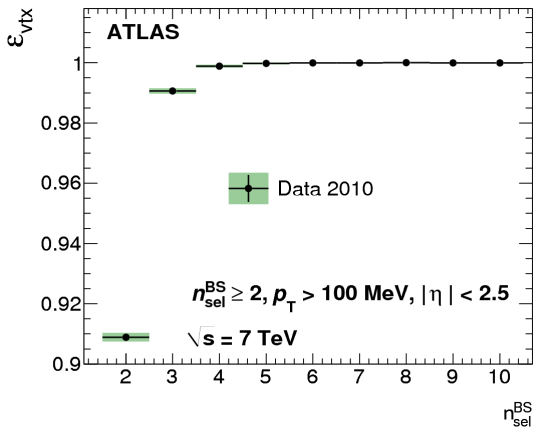

Figure 3.11: The vertex reconstruction efficiency of the ATLAS detector as a function of number of tracks in the beamspot [35].

3.3.2 Electrons

Electrons are a higher-order object than tracks. While tracks may be any charged particle, for the identification of electrons, more information needs to be combined. The reconstruction of an electron starts with a seed in the electromagnetic calorimeter. From there a matching track in the inner detector is tried to be matched. If no track is matched, or the matched tracks matches the profile of a photon conversion, the candidate is discarded. The resolution of the pseudorapidity and azimuthal angle of the inner detector track is better for the inner detector track, therefore that information is used for these two variables, while the energy resolution of the calorimeter is better and is used there.

Additionally reconstructed electrons are classified into three categories: loose, medium and tight [37]:

- Loose electrons need to satisfy a loose matching cut between the inner detector track and the calorimeter shower, as well as a few simple shower-shape cuts
- Medium electrons, on top of the loose cuts, tighten the cuts on the inner detector track and add more sophisticated shower-shape cuts
- Tight electrons add a tighter matching cut by explicitly requiring a hit in the innermost pixel layer, and also cutting on information from the TRT

The electron reconstruction efficiency depends on the applied set of cuts. It can be easily measured using $Z \rightarrow ee$ events: One requires at least one tight electron in the event. This electron is called the "tag" electron. Then a second electron is searched for, which, when combined with the tag electron has an invariant mass close to the one of the Z boson. This electron is called the "probe" electron. The background is very small, if such a match can be found, therefore one can easily measure the efficiencies for the various set of cuts for the probe electron. For low energy electrons a similar tag-and-probe-measurement can be done, using instead the J/ψ resonance instead of the Z resonance. Results of these measurements, for various values of pseudorapidity and transverse energy are shown in Figure 3.12 for medium electrons and in Figure 3.13 for tight electrons. The reconstruction efficiency for tight electrons is clearly lower than for medium electrons. Tight electrons have a better rejection of backgrounds, which mainly stems from misidentified jets and

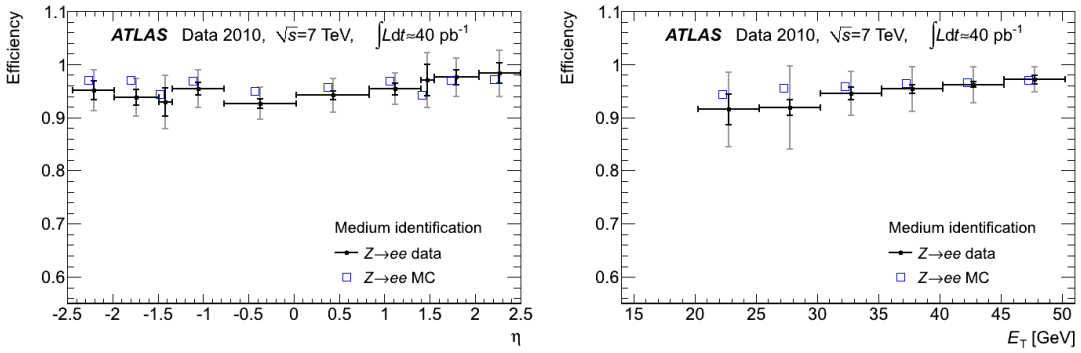

Figure 3.12: The reconstruction efficiency for medium electrons for data and simulation in various bins of pseudorapidity (left) and transverse energy (right) [37].

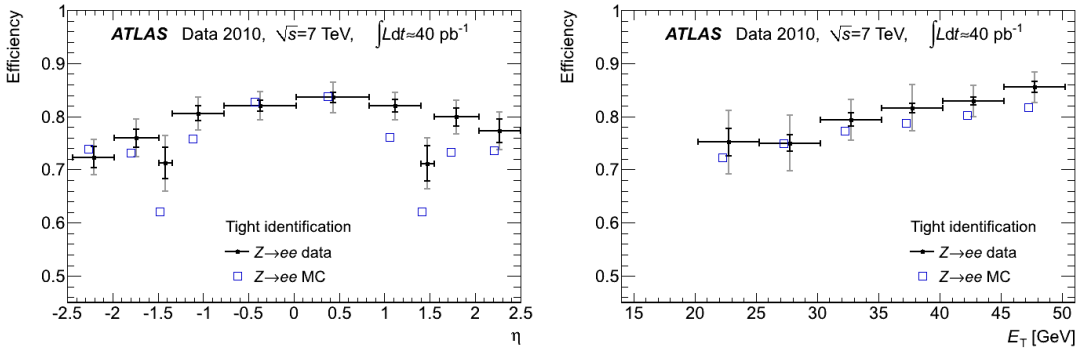

Figure 3.13: The reconstruction efficiency for tight electrons for data and simulation in various bins of pseudorapidity (left) and transverse energy (right) [37].

from photons. In general, the jet rejection of the tight cuts is about a factor of 40 times better than the medium cuts [31].

Another important quantity is the energy resolution for electrons. The energy resolution depends mostly on the resolution of the calorimeter. It ranges from $\sim 1\%$ for high p_T and low η up to $\sim 5\%$ at low p_T and high η [31].

3.3.3 Photons

The photon reconstruction starts in a similar way as the electron reconstruction. A seed from the electromagnetic calorimeter is extrapolated to the inner detector. If it does not match to a track there, or the matched track passes as a photon conversion, the object becomes a photon candidate. On top of that there is a similar selection as for the "tight" electrons. The cuts rejecting photons for electrons have been inverted to instead select photons. The performance of the photon reconstruction is similar as for the electron reconstruction as well as in resolution [31, 38].

3.3.4 Muons

In general, the signature of the muon is a track in the muon spectrometer. Basically any other particle – with the exception of neutrinos which will not leave a track – will be stopped by either the electromagnetic or the hadronic calorimeter. The muon reconstruction therefore begins by the reconstruction of a track in the muon spectrometer. This track is then either taken standalone or combined with a track in the inner detector. The respective muons are therefore called "standalone" or "combined" muon. There are two analysis chains of combining the inner detector and the muon spectrometer tracks. One does a statistical combination of the two tracks while the other one uses the full covariance matrix to build a combined track. Additionally for the rejection of background from hadrons there are additional cuts, like calorimeter isolation or hit requirements in the inner detector – similar to the electrons – which are applied [39].

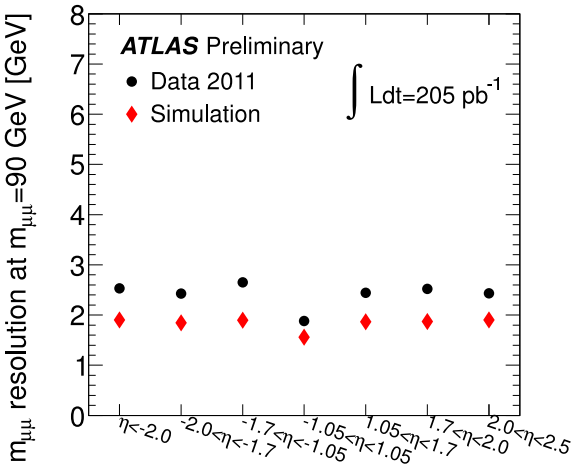

Figure 3.14: Combined dimuon mass resolution for various ranges of pseudorapidity in data and simulation. The resolution is slightly worse for the actual data [39].

Figure 3.14 shows the mass resolution of combined dimuons in 2011 data. The invariant mass was cut to be at a window around the Z-boson mass. In general the resolution in the data is around 2.5%, while the resolution in simulation is a bit better. The difference is caused by the residual misalignment of the detector and the simulation needs to be corrected to match the one in data [39].

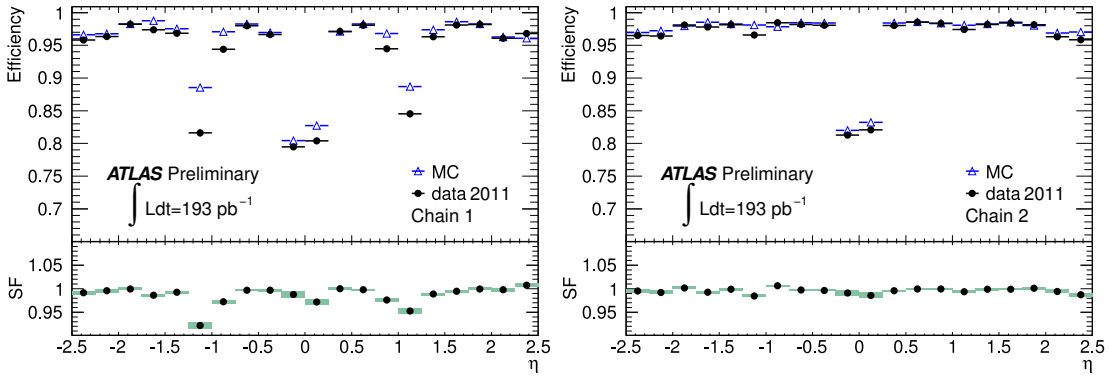

Figure 3.15: Reconstruction efficiency measured by the ATLAS collaboration in 2011 data for Chain1 (left) and Chain2 (right), depending on the pseudorapidity of the muon [39].

The reconstruction efficiency of muons is displayed in Figure 3.15. In general, the reconstruction efficiency of muons is very high, close to unity. There are gaps in the reconstruction efficiency at $\eta = 0$ for both chains as there is a gap for detector services in this region. The simulation describes the muon reconstruction efficiency very well [39].

3.3.5 Jets

Jets are – other than the other discussed objects – not single high energy particles, but instead a spray of particles. Usually they originate from a single high-energy-particle, a quark or gluon, which due to the nature of QCD creates a plethora of secondary particles. Depending on the initial momentum of the original particle, the spray will have a wider or lower opening angle. The created particles deposit their energy in the calorimeters. A jet algorithm starts by taking all clusters of energy in the calorimeter. The particular algorithm used is called "anti-k_t" [40] jet algorithm.

The anti-k_t algorithm works in the following way. It calculates the k_t distance for each pair of particles:

$$d_{ij} = \min\left(k_{ti}^2, k_{tj}^2\right) \Delta R_{ij}^2 / R^2 \tag{3.1}$$

k_{ti} denotes the transverse momentum of the particle i, and ΔR is the distance in the rapidity-phi-plane. R is related to the jet-radius and is a parameter of the algorithm. Typical values in ATLAS are $R = 0.4$ or $R = 0.6$. Additionally, for each particle the so-called beam distance, d_{iB}, is calculated:

$$d_{iB} = \frac{1}{k_{ti}^2} \tag{3.2}$$

When this procedure has finished for each pair of particles, all d_{ij} and d_{iB} are sorted into a list and the minimum element is considered. If it is a d_{ij}, the particles i and j are merged and all distances are recalculated. If it is instead a d_{iB}, the particle i is considered a final jet and removed from the considered particles. This procedure is repeated until there are only final jets left in the event.

The calibration of jet energies is challenging as they are the most complex objects in the detector. Energy might be lost due to some particles escaping the calorimeter. Another source of uncertainty is particles never reaching the calorimeter due or their too low momenta. The calibration of jet energies usually involves the balance of momentum of a jet and a easier to measure particle, like photons or leptons. Two main sources of uncertainties remain: The "jet energy scale" (JES), the scale of jet energy measurements, and the "jet energy resolution" (JER), the resolution. Figure 3.16 shows the jet energy scale uncertainty for various values of p_T and η. Depending on the values of these variables,

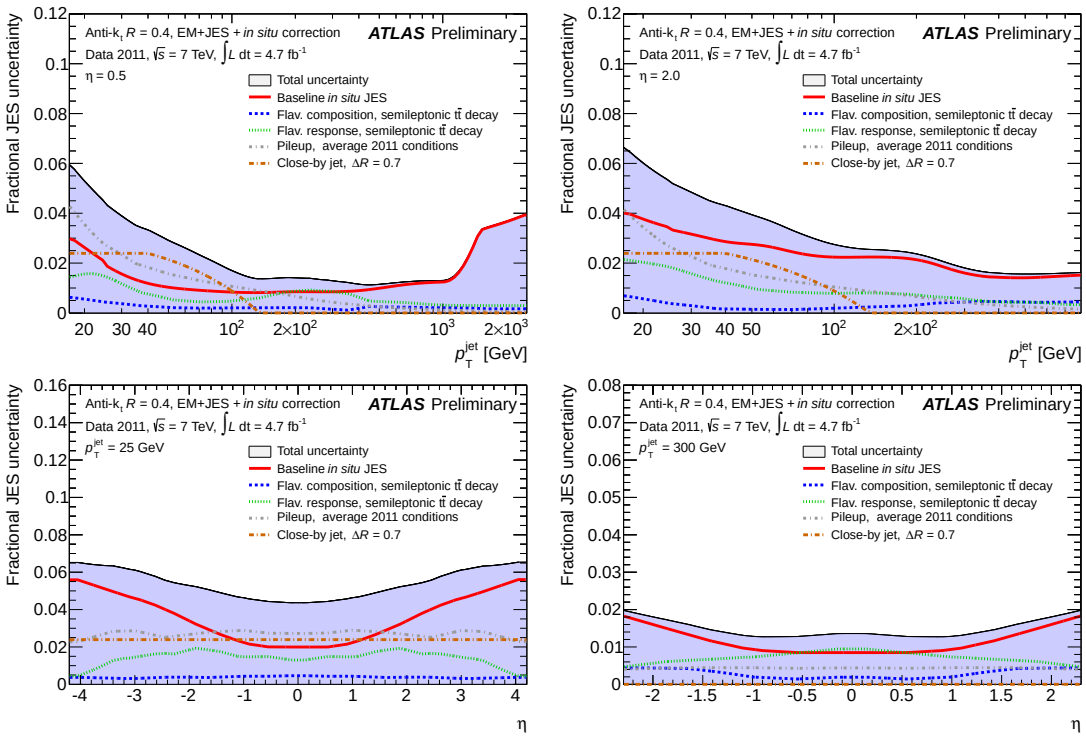

Figure 3.16: The components of the jet energy scale uncertainty. Top: As a function of the transverse momentum of the jet for $\eta = 0.5$ (left), and for $\eta = 2.0$ (right). Bottom: As a function of the pseudorapidity for $p_T^{jet} = 25\,\text{GeV}$ (left), and for $p_T^{jet} = 25\,\text{GeV}$ (right) [41].

the jet energy scale uncertainty is between 2% and 6% [41]. The jet energy resolution uncertainty for various calibration schemes is shown in Figure 3.17. Depending on the transverse momentum on the jet, the resolution can be up to 20% [42].

The jet energy scale uncertainty shown neglects one component which only applies to bottom quark jets. Bottom quarks can decay into a lepton-neutrino pair. Neither of these particles deposit much (or any) of their energy in the hadronic calorimeter. This effect needs to be separately addressed and leads to an additional term in the jet energy scale uncertainty when dealing with jets originating from bottom quarks.

3.3.6 Missing Transverse Energy

Neutrinos interact very rarely with the detector material. Therefore they will leave the detector almost undetected and the information about their kinematics is lost. However one can exploit the conservation of energy to gain information about the neutrino: The colliding partons have very small momentum transverse to the beam axis. The sum of transverse momenta of all the final state objects from a parton-parton collision is therefore very close to 0. Using that, information about the transverse momentum of a neutrino can be extracted. If more than one neutrino is present, only the combined momentum can be extracted. This reconstruction method would also be able to find new particles which do not interact with the detector more than a neutrino would. As one reconstructs the energy which is missing for the conservation of energy, the object is called “Missing Transverse Energy”.

The reconstruction of missing energy takes all reconstructed physics objects. Additionally it considers objects like soft jets or unclustered energy in the calorimeter. It then builds the negative sum of all objects, taking the direction information into account. The

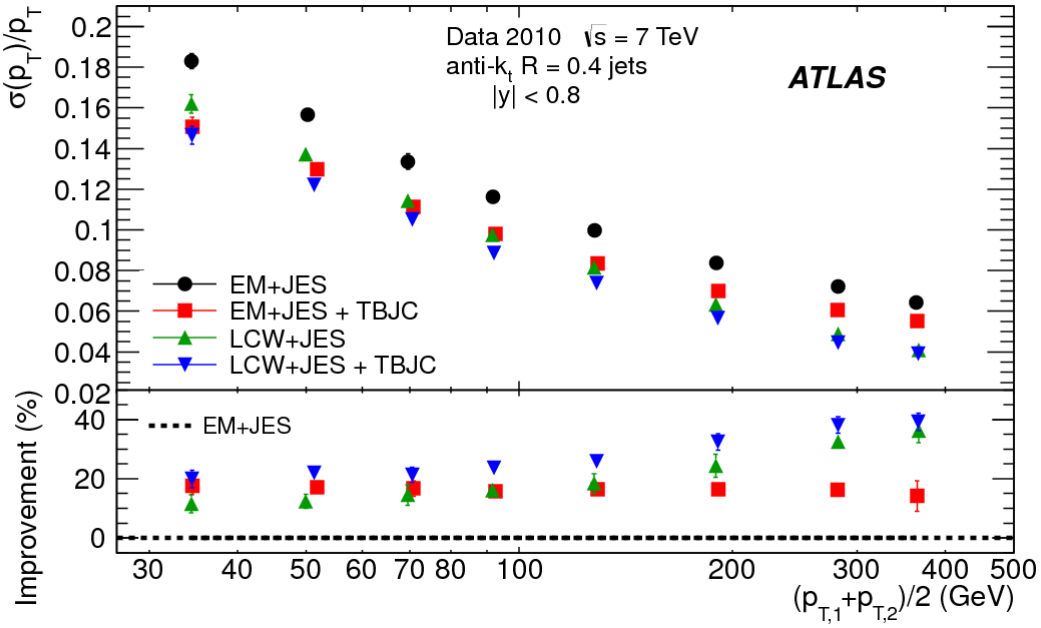

Figure 3.17: The jet energy resolution as measured by the ATLAS calibration in 2010 data for various jet calibration schemes [42].

information is very dependent on the resolution of all the input objects. A good quantity to measure the performance of the missing energy is the missing energy resolution.

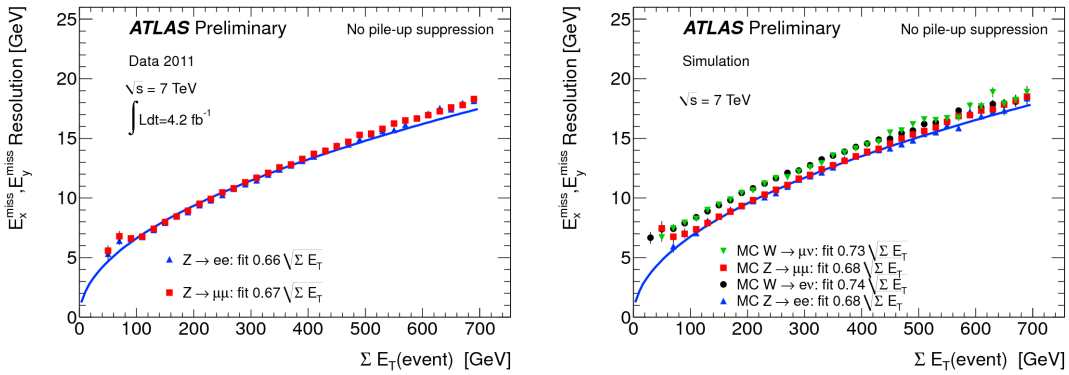

Figure 3.18: The missing energy resolution in data (left) and various simulations (right) against the total amount of transverse energy in the event. The fit is a function of the form $k\sqrt{\sum E_T}$. k denotes the free parameter in the fit. Its fit values are printed in the Figures [43].

Figure 3.18 shows the resolution of missing energy for various values of total transverse energy in the event. In general the resolution gets worse for more total energy. The values have been fit with a root function for a smoother description. The fit results show that the resolution of missing energy is well-described for $Z \to \mu\mu$ and $Z \to ee$ events [43].

3.3.7 Flavor Tagging

Flavour tagging is the technique of identifying jets which originate from bottom quarks. Because of the structure of the CKM-Matrix (compare chapter 2), a bottom quark would basically all the time decay into a top quark – a decay which is forbidden because of the

higher mass of the top quark. Thus the suppressed, but allowed decay to a charm quark is the most likely decay of a bottom quark. This suppression leads to a significant lifetime of hadrons containing at least one bottom quark: They are able to travel up to several centimetres in the detector. In the reconstruction this leads to tracks originating not from the primary vertex, but instead from a secondary vertex. The properties of these tracks can be exploited to identify these jets. This is a powerful technique to separate signatures which are similar except for the presence of bottom quark jets, like the W+jets signature and the $t\bar{t}$ signature.

In praxis more information than the presence of a secondary vertex is used: The impact parameter of the tracks from bottom quarks is larger than the one from lighter quarks. The reconstructed mass of the tracks at the secondary vertex or the amount of charged particles in the jet are other examples of variables which are sensitive to the flavor of the originating particle. Several algorithms have been developed to identify bottom quark jets. The most performant one for 2011 data is called "MV1". It is a neuronal network combining the output of several other algorithms: "SV1" and "JetFitter" [44] which use the properties of secondary vertices and the IP3D [44] which uses the impact parameters of tracks.

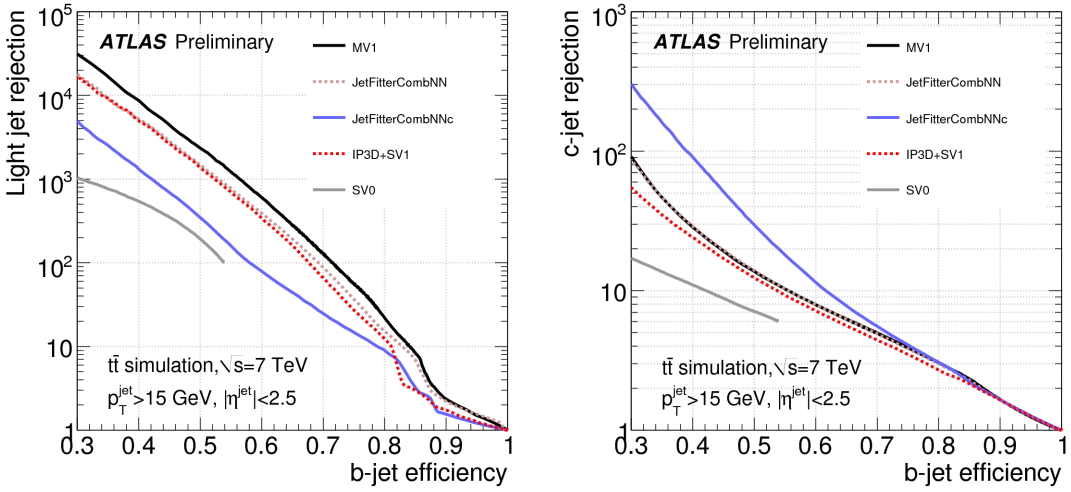

Figure 3.19: The light-jet (left) and charm-jet (right) rejection against the b-jet tagging efficiency for various algorithms. The used sample is a $t\bar{t}$-sample [45].

Figure 3.19 shows the performance of some of these algorithms. It is quantified as the rejection of light or charm quark jets at a given bottom quark jets tagging efficiency. Typically a fixed efficiency is chosen which corresponds to a cut on the output distributions of the algorithm. By changing this cut, looser or tighter selections can be chosen.

The efficiency and fake rate of these algorithms have been measured in data [46]. Two approaches have been applied by the ATLAS calibration: One is to exploit the fact, that muons from a bottom hadron decay will have a significant transverse momentum with respect to the jet axis. This transverse momentum is called p_{T}^{rel}. Its distribution is displayed in Figure 3.20. This transverse momentum stems from the high mass of the bottom quark.

Another important fact is that the p_{T}^{rel} variable is uncorrelated to the bottom quark tagging algorithms described earlier. Therefore a template fit to the number of jets before and after the tagging requirement will yield the total number of bottom quark jets, N_b^{total} as well as the number of tagged bottom quark jets, N_b^{tagged}, in the sample. The efficiency,

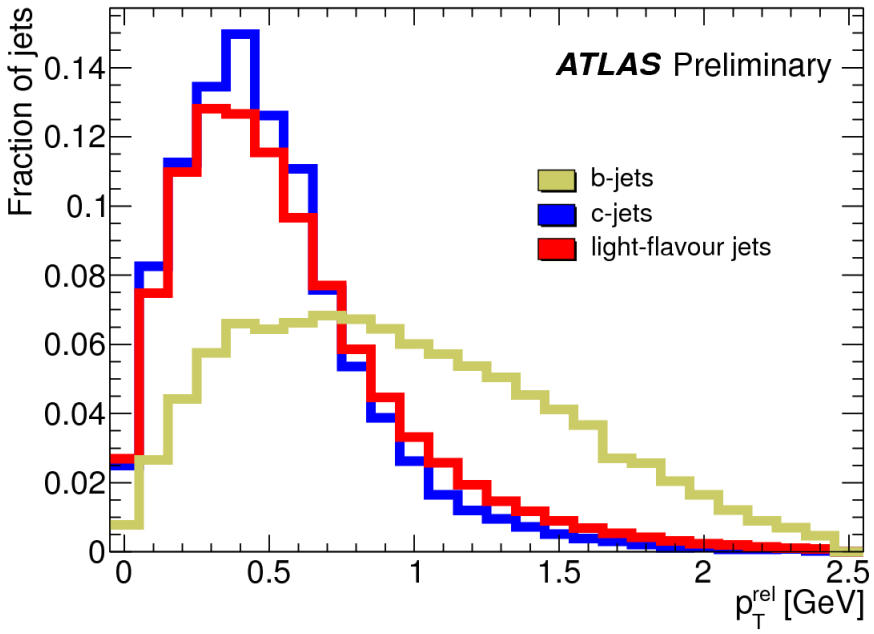

Figure 3.20: The distribution of the p_{T}^{rel} variable for various flavors of jets [46].

ϵ_b can then simply be extracted in the following way:

$$\epsilon_b = \frac{N_b^{tagged}}{N_b^{total}} \tag{3.3}$$

The measurement has been applied to data containing dijet events. In order to increase the number of bottom quark jets in the sample a tag-and-probe approach has been used: One jet has been tagged by a tight cut value of a bottom tagging algorithm. The other jet, if it contains a muon, is used for the analysis. Since the bottom quark tagging efficiency depends on the transverse momentum of the jet, the measurement has been done in several bins of it. Figure 3.21 shows the fit results for one of the bins of jet transverse momentum. Only jets with $40\,\mathrm{GeV} < p_T^{jet} < 50\,\mathrm{GeV}$ have been used in this figure.

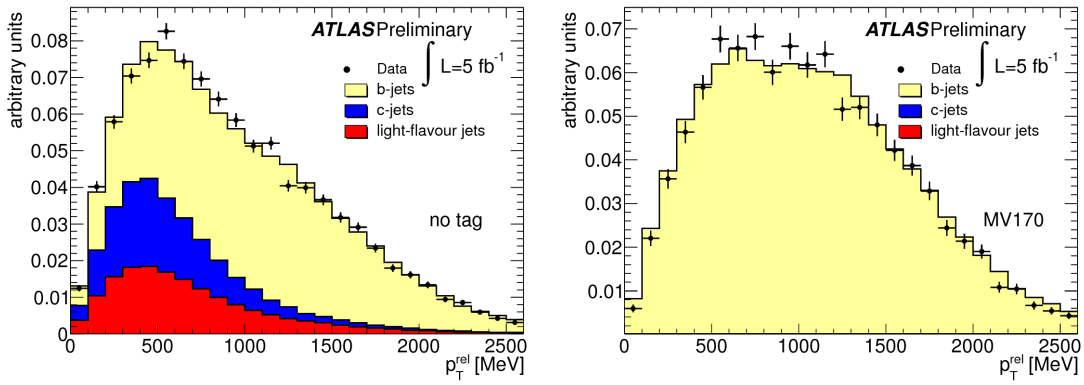

Figure 3.21: Example fits to the data for the inclusive sample (left) and the sample tagged by the MV1 tagger at 70% nominal efficiency (right). The transverse momentum of the jets has been between 40 GeV and 50 GeV [45].

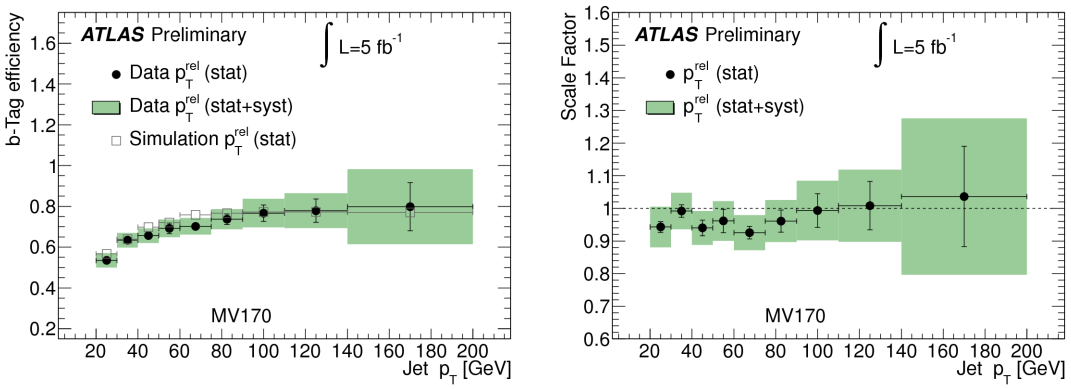

Figure 3.22: The results of the p_{T}^{rel} analysis: The fitted efficiency for the data and the simulation tagging efficiency (left) and the scale factor needed to correct the simulation (right) [45].

Figure 3.22 shows the result of the analysis. On the left the tagging efficiency of the simulation and the measured efficiency in data for the various jet transverse momentum bins is displayed. On the right, the ratio of the two, $\kappa_b = \frac{\epsilon_b^{data}}{\epsilon_b^{MC}}$, called "scale factor" is shown. It is needed to correct the simulation to match the performance in the data. In general the scale factor is a bit lower than one, indicating that the algorithms perform slightly worse in the data than in the simulation.

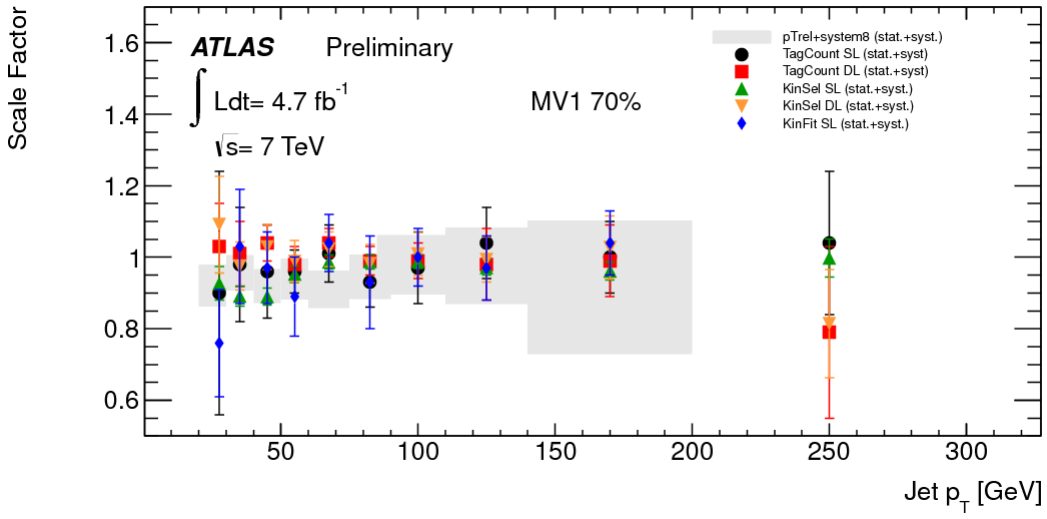

Figure 3.23: The scale factors measured by several $t\bar{t}$ based calibration methods as well as measured with the p_{T}^{rel} based measurement. While the $t\bar{t}$ scale factors tend to be a bit higher than the p_{T}^{rel} based ones, good agreement is observed within uncertainties [47]

One important note for the p_{T}^{rel} analysis is that it can only measure the tagging efficiency for the subset of bottom quark jets where a decay into a muon happened[46]. The largest systematic uncertainty of the measurement is the extrapolation uncertainty to the inclusive sample.

Other measurements of the bottom quark tagging efficiency do not have that limitation. For instance, the decay of the top quarks can be used to measure the bottom jet tagging efficiency. The dileptonic $t\bar{t}$ has a very clean signature even when not requiring bottom jet tagging. Therefore this signature can be used to measure the tagging efficiency by

counting how often the jets in the final state are tagged. This, and more sophisticated techniques, have been applied to this decay channel and the semileptonic $t\bar{t}$ decay channel. The results are displayed in Figure 3.23. The figure also shows the combination of the p_{T}^{rel} measurement combined with another measurement which exploits the p_{T}^{rel} variable in a different way, called "System8" [45]. In general, good agreement within uncertainties is observed between all measurements of the bottom jet tagging efficiency.

Additionally the fake rate needs to be measured. This must be done separately for jets originating from charm quark and light quarks as the performance of the bottom jet tagging algorithms differs between them.

For charm jets the D^* meson from the charm decay has been fully reconstructed to select a very pure sample of charm jets. The number of charm jets before and after applying the bottom jet tagging algorithm is extracted in a fit. The results of this analysis are shown in Figure 3.24.

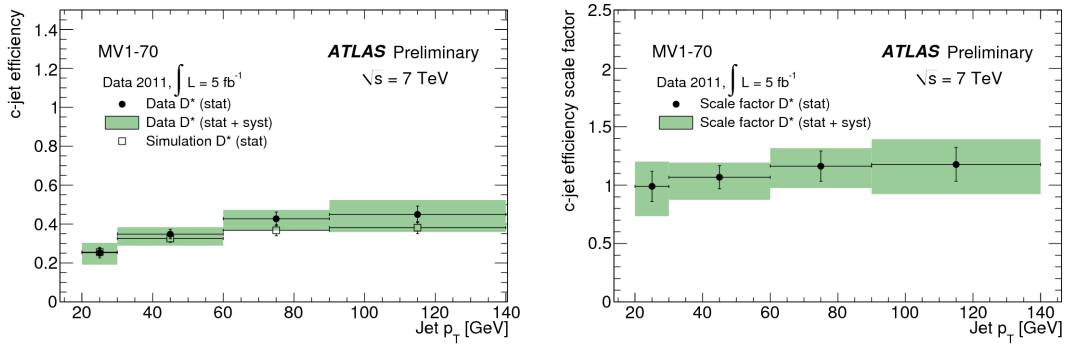

Figure 3.24: The measured efficiency (left) and scaling factor (right) for jets originating from charm quarks in bins of transverse momentum of the jet [48].

The light jet fake rate has been measured by two methods. One does fits to the invariant mass at the vertex, similar to the p_{T}^{rel} analysis. The other measurement exploits that the properties of light jets are almost symmetrical with respect to the primary vertex. Therefore by inverting e.g. the impact parameters of the tracks, one can measure the negative tag rate, and extrapolate that to the positive tag rate. The result of the latter measurement is shown in 3.25.

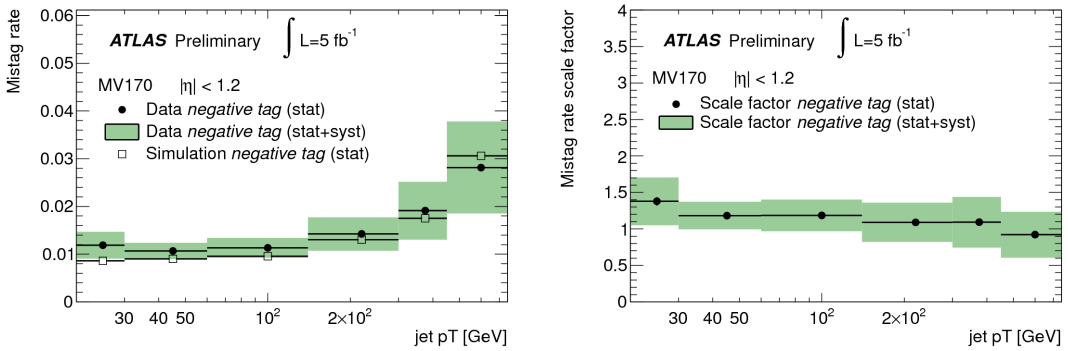

Figure 3.25: The measured efficiency (left) and scaling factor (right) for jets originating from light quarks in bins of transverse momentum of the jet. Only jets with $|\eta| < 1.2$ have been considered in this Figure [49].

Chapter 4

Simulation of Top-Quark Production and Decay

4.1 Simulation of a Hadron-Hadron Interaction

An interaction at the LHC is not an easy thing to describe theoretically. Since the colliding protons are not elementary the actual interactions are done by partons which will not carry the full energy and momentum of the proton. Since the protons consist of more than one of these partons there may also be additional hard interactions. The fact that most of the processes are governed by the strong force does not make the task easier. Instead, there are additional hurdles like the flow of the colour charge that need to be taken into account. Also, whilst the hard process most of the time can be described pertubatively the hadronization and the following parton shower cannot.

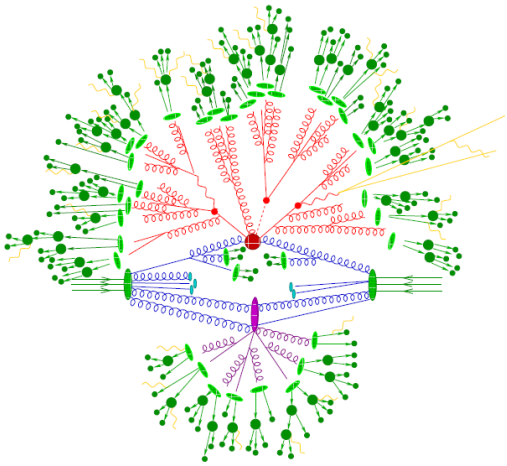

Figure 4.1: A sketch of a hadron-hadron interaction. The signal process is depicted in red, the additional hard interactions in pink. The parton shower is shown by the green blobs and line while the yellow lines represent additional radiations [50].

However, there exists a theorem that states that each of these terms can be factored out and described separately: The factorisation theorem [51]. All current simulation tools for hadron-hadron interactions use this to simulate the interactions. It is usually split up into the following categories:

Parton Distribution Functions (PDFs) The parton density functions give the probability of finding a parton (gluon or quark) in a hadron parametrized in the energy of the hadron and the fraction of momentum the parton carries.

Hard Process The actual signal interaction. The hard process only considers the matrix element of the parton-parton interaction to a given final state.

Hadronization and Parton Shower These cover the formation of hadrons from the particles in the final state and their decay. They may also simulate the additional interactions and final state radiations.

Each of these steps will now be described in more detail.

4.1.1 On Perturbation Theory

The creation of a top-antitop-quark-pair is dominantly a process governed by the strong force. The interaction Lagrangian can be written as a perturbation series in order of α_s:

$$\mathcal{L} = \alpha_s A_1 + \alpha_s^2 A_2 + \alpha_s^3 A_3 + \alpha_s^4 A_4 + O\left(\alpha_s^5\right) \tag{4.1}$$

It is obvious that this perturbation series only converges for $\alpha_s < 1$. As the value of α_s depends on the interaction energy this is only given for certain interaction energies.

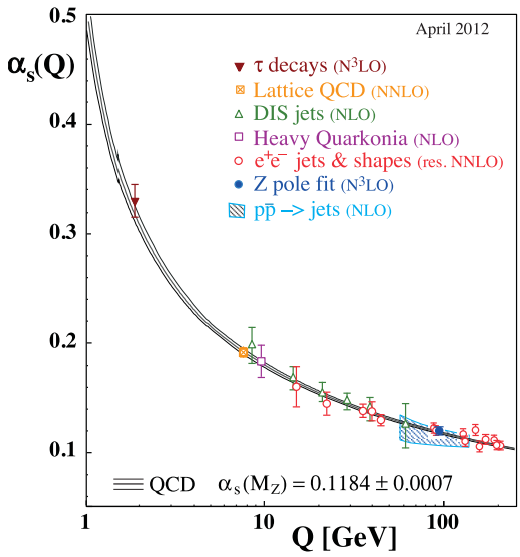

Figure 4.2: The running coupling of the strong force. For lower interaction energy Q the value of the interaction increases and will eventually reach value larger than one [5].

Figure 4.2 shows the dependence of the strong coupling constant α_s on the interaction strength Q. For higher interaction energy the coupling strength decreases while for lower interaction energies the interaction strength will reach values larger than one. This effectively means that in these regions the approach of perturbation theory is not viable, instead other descriptions need to be found.

4.1.2 Parton Distribution Functions

As stated in chapter 2 a proton is not elementary. It consists out of two up-quarks and a down quark. Additionally there is a "sea" of additional quarks and gluons. With increasing energy of the proton the amount of sea quarks and gluons increases. A single constituent of the proton therefore will never carry the full energy of the proton. To correctly describe an interaction involving at least one proton the probability to find a constituent with a given energy E_p must therefore be known. This is done by parton distribution functions. Based on a set of measurements of the momentum fractions of hadrons they give the probability to find a parton of a given type and momentum fraction x [52].

The parton distribution functions were measured to good precision at several fixed-target experiments and at the HERA collider where a beam of electrons with an energy of 27.5 GeV was brought into collision with a proton beam accelerated to 920 GeV [53, 54]. These measurements must then be extrapolated to the energy of the collision in question.

Several groups attempt to fit the input data: The CTEQ groups parametrizes the input parton distribution functions with the following functional form [55]:

$$xf\left(x,\mu\right) = A_0 x^{A_1}(1-x)^{A_2} e^{A_3 x}\left(1+e^{A_4}x\right)^{A_5} \tag{4.2}$$

for each of the different flavors. μ is the scale at which the PDF is to be evaluated. In total the CTEQ group quotes 20 different parameters. The uncertainties on them are transformed to a set of independent eigenvectors to be able to vary them in an uncorrelated way. Additional input parameters are the masses of the quarks in question and the value of the strong interaction strength α_s. The CTEQ group provides them in several different versions, i.e. the CTEQ6M set uses the quark masses in the $\overline{MS}$ scheme. Also there exist versions where the strength of α_s is also extracted from the fit. The most recent result of the CTEQ collaboration is the CT10 PDF set. It contains in total 52 eigenvector sets to be evaluated. The default values of the PDF at $\mu = 85$ GeV is shown in Figure 4.3.

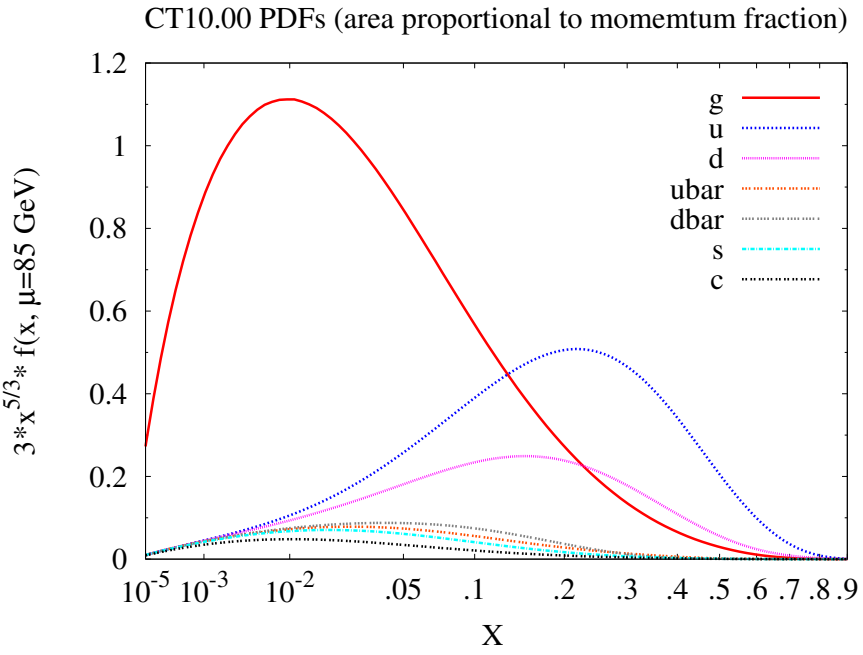

Figure 4.3: The parton distribution functions as provided by the CTEQ collaboration at a μ of 85 GeV[56].

In addition there are other groups providing PDF sets. One example of them is the MSTW collaboration. They differ in the CTEQ group mainly in choice of the parametrization of the input PDF [57]:

$$xf\left(x,\mu\right) = A(1-x)^{\eta}(1+\epsilon x^{0.5}+\gamma x)x^{\delta} \tag{4.3}$$

This also leads to a different number of eigendirections which parametrise the uncertainties.

4.1.3 Matrix Element

The matrix element is the calculation of the cross-section for the actual interaction, say two gluons fusing to a single gluon which then form a pair of top quarks. These interactions

happen at scales where the strong coupling constant can be described by a perturbation series. It is therefore possible to calculate the value of the matrix element up to a certain order of α_s and neglect the other orders as their contributions will be small.

Figure 4.4 shows the two lowest order Feynman diagrams contributing to the strong creation of $t\bar{t}$ pairs. As there are two strong interaction vertices it is of the order of α_s^2. Figure 4.5 shows some examples for higher order diagrams of $t\bar{t}$ pair creation. All of them show additional strong vertices meaning that they are suppressed by additional factors of α_s.

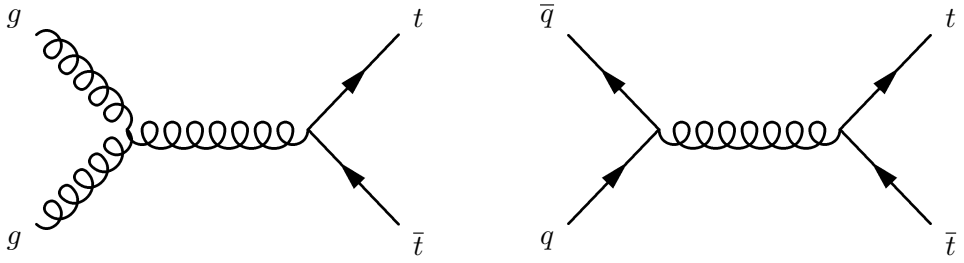

Figure 4.4: Leading order processes for the generation of $t\bar{t}$ events

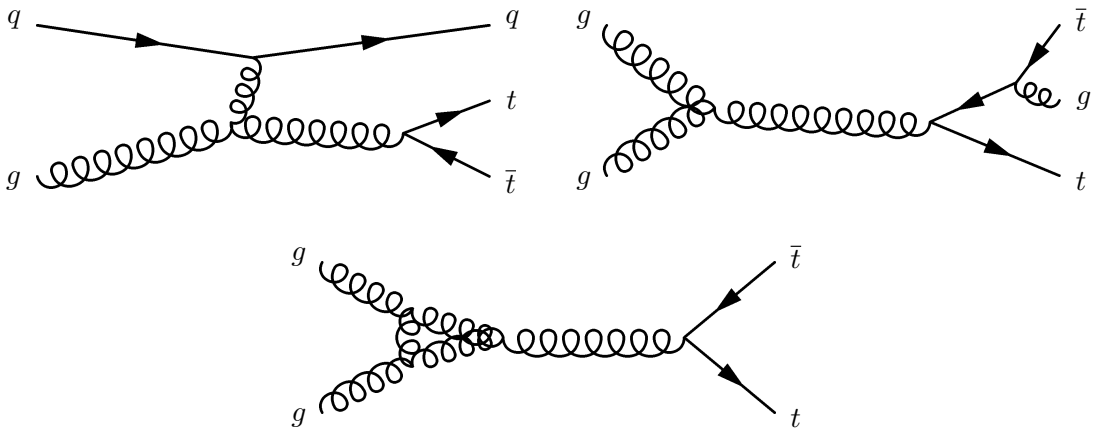

Figure 4.5: Example higher order processes for the generation of $t\bar{t}$ events

Two ways of calculating the matrix elements are in use in the high energy physics community:

- Tree level element generators which calculate the matrix elements of the style in Figure 4.4. Additionally they may include additional radiations (like the middle of Figure 4.5)
- So-called NLO generators which generate the matrix element up to one loop (for instance the right one in Figure 4.5)

The advantage of the tree-level is that they are easier to calculate up to high number of final state objects while the NLO one generally gives better results.

4.1.4 Parton Shower

The final step of the simulation of the $t\bar{t}$ interaction is the parton shower. It is the emission of soft gluons from final state objects. Depending on the energy and angle of the emission there exist divergences in QCD (called soft and collinear) which can not be calculated with perturbation theory. At this stage the parton shower approximation is used.

It considers all emissions down to a certain scale value q_0 which is called the cutoff. Emissions below the cutoff are not considered by the parton shower approximation. The probability for an emission can be calculated to be

$$P(\Phi_{rad})\mathrm{d}\Phi_{rad} = \frac{\alpha_s q}{\pi}\frac{\mathrm{d}q}{q}P(z)\frac{d\phi}{2\pi} \tag{4.4}$$

with $P(z)$ being the so-called splitting function [51].

It follows for the probability of no emission below the cutoff between two scales $q_1, q_2, q_2 > q_1$:

$$\Delta_s(q_1, q_2) = \exp\left(-\int_{q_2}^{q_1}\frac{\alpha_s q}{\pi}\frac{\mathrm{d}q}{q}\int_{z_0}^{1}P(z)\mathrm{d}z\right) \tag{4.5}$$

This is the so-called Sudakov form factor and plays a very important role in the technical implementation of the parton shower approximation: Each parton in the final state can emit additional partons at a scale q_1 which is chosen by generating a random number R_1 for $\Delta_s(Q, q_1)$ (Q being the initial scale) and then solving for q_1. This process is repeated for a new scale q_2 which has to be lower than q_1 until one reaches the cutoff. The result is a shower of partons with properties near the chosen cutoff.

The choice of scale variable is usually either the angle of emission (the first emission has the biggest emission angle) or energy (the first emission has the highest energy).

Usually the fragmentation and the simulation of additional interactions is also done during this step of simulation even though they are not directly related to the parton shower approximation.

4.2 Simulation Techniques for Top Interactions

4.2.1 Tree Level: AcerMC and ALPGEN

Tree level event generator have the advantage that they can generate many events in rather short time. They lack the high precision of the next- to-leading-order event generators but their data-description can still be excellent.

AcerMC provides the matrix element for several processes involving top quarks either in pair production or in single top production [58].

Alpgen is specialised for generating heavy quarks or vector bosons in association with many additional partons in the matrix element [59].

4.2.2 Next-to Leading Order Matrix Element: POWHEG and MC@NLO

The main challenge when designing next-to-leading-order event generators is the matching with the parton shower: There is an overlap in that the emission of an additional parton is simulated in both steps. The present next-to-leading-order matrix element generators use different approaches to resolve this overlap.

The Mc@nlo [60] framework extends the matrix-element calculation by an additional term with negative sign. This term is constructed in a way to generate the overlapping events. With the negative sign these events are subtracted from the original matrix element which results in a matrix element which does not overlap with the parton shower anymore. However the Mc@nlo framework does require the exact knowledge of the form of the following parton shower. Interfacing it with a new parton shower generator requires significant change to the calculation.

A different approach is used by the Powheg method [61]. Here the idea is to generate the hardest emission of an additional parton in the matrix element and then let the parton

shower take over for the following emissions. This only requires a general way of telling the parton shower what has been done in the matrix element. A general interface for this exists in the Les Houches Interface [62] and the commonly used parton shower algorithms support it. An additional advantage of the POWHEG method is that there is no negative term in the matrix element which would otherwise lead to negatively weighted events which are hard to interpret physically.

4.2.3 Parton Shower Generators: Herwig and Pythia

Two packages are commonly used for handling the simulation of the parton shower: HERWIG [63] and PYTHIA [64].

The main difference between the two is that HERWIG is an angular-ordered shower, meaning that the scale variable chosen for the evaluation is the emission angle, while PYTHIA is a p_T-ordered shower, emitting the hardest radiation first.

Chapter 5

Techniques for measuring the Mass of the Top Quark

5.1 The Template Method

The basic idea behind the template method is to find a distribution which is sensitive to the observable one is interested in and then fit that distribution with an analytic function. The fit result will then depend on the observable and can be used to extract the observable. In praxis one uses simulation with varied values of the observables to study the dependence of the distribution and uses that to calibrate the function one uses to fit the distribution.

For example one can consider the following model: The observable in question would be the mean μ of a Gaussian distribution:

$$f(x) = \frac{N}{\sqrt{2\pi}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{5.1}$$

Due to experimental effects one measures the mean shifted by $+0.2$ of its true value. Additionally the width of the Gaussian is proportional to the inverse mean: $\sigma \sim {}^1/_\mu$ and the normalisation depends exponentially on it: $N \sim \exp(-\mu)$.

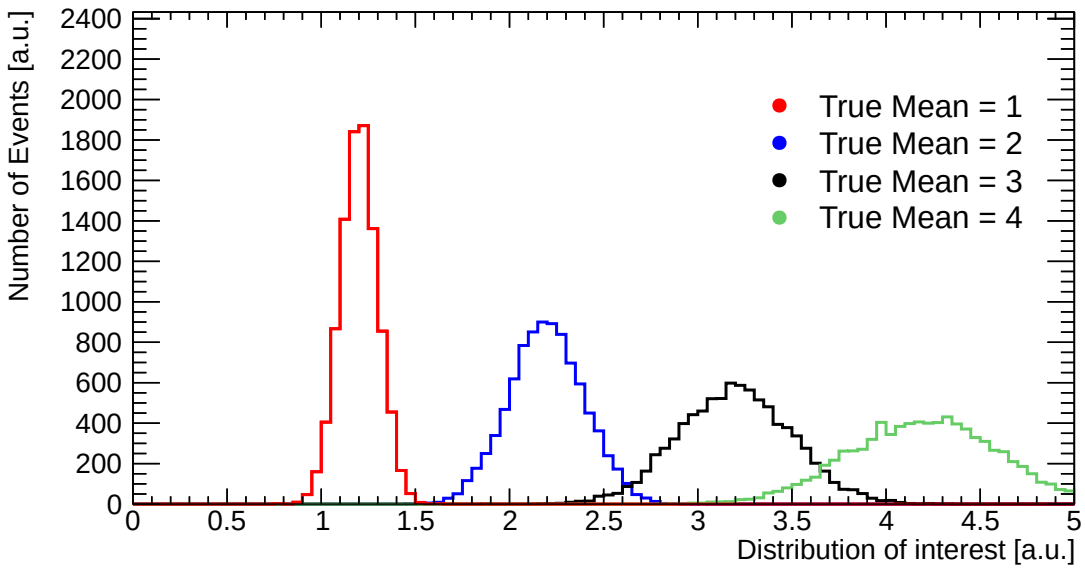

Figure 5.1: Distributions for various values of μ for a toy Gaussian simulation

The template method would now generate templates of that distribution for various values of μ. Figure 5.1 shows the distributions of these templates. The shift in the distribution towards higher values is clearly visible as well as the additional dependence of norm and width of the Gaussian on its mean. Each of these distributions would now be fit with a Gaussian and its parameters correlated with the true mean. These resulting graphs for each of these three are shown in Figure 5.2.

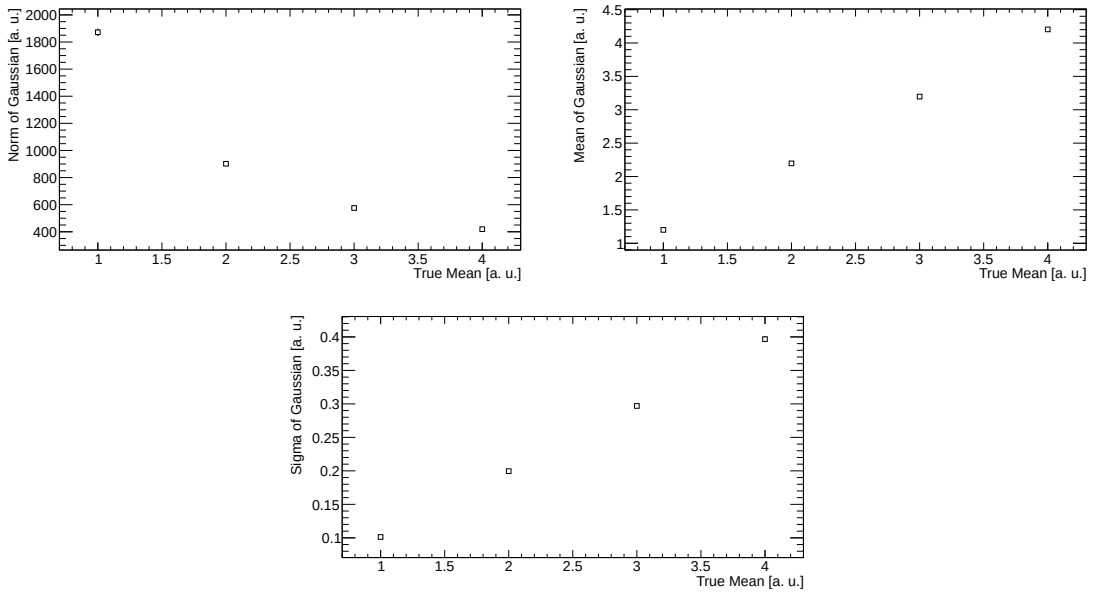

Figure 5.2: Calibration curves for the norm (top left), mean (top right) and sigma (bottom) of the toy Gaussian models for the calibration of the underlying mean

These graphs are then fit with a parametrization (usually a linear function) which can then be used to predict the observable distribution for any value of μ. This allows the performance of a likelihood fit with the data distribution which does only depend on the assumed μ. The resulting value and uncertainty from this likelihood fit is the result of the measurement in data.

The advantages of the template method are clear. While it would be possible to fit a Gaussian distribution to data and correct the resulting mean by the experimental effects the template method uses the additional information of the other parameters as well. An additional feature of the template method is the easy addition of more sensitive variables to the likelihood fit. This is done by calibrating their behaviour as well and then simply adding them to the likelihood expression. In the same manner more than one observable can be extracted from the data. However, the template method requires a good description of the data by the simulation, but this is true for almost all of the techniques presented in this thesis.

5.2 The Calibration Curve Approach

Another approach similar to the template method approach is the calibration curve approach. It uses a variable which is sensitive to the property to be measured. However instead of trying to find a full parametrization of it it uses more simple quantities of the distribution, e.g. its mean value as a single parameter.

The dependence of this parameter is parametrized versus the property. This parametrization is called calibration curve. Since it is a single parameter, simply calculating that parameter is enough to get an estimate of the property: The calibration curve can directly

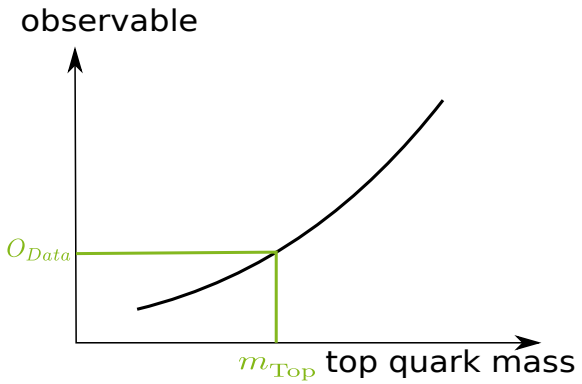

Figure 5.3: A sketch showing the calibration curve technique. The black line is the calibration curve, parametrizing the behaviour of the observable against the top quark mass. The green colour shows the approach for extracting the mass: The observable is taken from the data and the calibration curve used to extract the top quark mass.

be used to get the value and its uncertainty.

While using less information than a full template analysis the calibration curve technique can be used in situations where an analytic description of the observable is hard to find, thus allowing also the use of such observables.

5.3 The Matrix Method

The matrix method was designed to measure the top quark in the presence of limited statistics. For each event it calculates the probability $P_{t\bar{t}}$ that the event is a top quark pair production event in the given selection and an assumed top mass:

$$P(x, m_{\mathrm{Top}}) = A(x)\frac{1}{\sigma(m_{\mathrm{Top}})}\int \mathrm{d}\sigma(y, m_{\mathrm{Top}})\mathrm{d}q_1\mathrm{d}q_2 f(q_1)f(q_2) \times W(y, x) \qquad (5.2)$$

In this formula x represent the measured objects in the detector (i.e. jets or leptons). $A(x)$ is the acceptance of these events which is independent of the top quarks mass.

$W(x, y)$ are transfer functions which transfer the measured objects to the partonic (i.e. quarks) objects. The integral calculates the partonic cross-section depending on the top quark mass. Typically, these integrals are solved using Monte-Carlo integration techniques. Depending on the number of final state objects this integral can become very CPU-intense to solve.

Additionally the probability for the event being a background event needs to be calculated as well. A maximum likelihood fit is then performed, thus finding the value of the top quark mass which gives the overall maximum probability.

The big advantage of the matrix method is that well-measured events will contribute more to the final likelihood than badly measured events. This leads to better statistical accuracy than the template method. However, in situations with many final state objects the calculation of the cross-sections will use very large amounts of computing time.

5.4 Direct Measurement of the Mass of the Top Quark from the its Cross-Section

Any method which requires simulation to measure the top quark mass faces an additional issue: The term of mass is not well-defined for a strongly-interacting particle as the top

quark. Higher order corrections and their divergences of QCD impact the value of the top quark mass as it is implemented in the simulation.

There are two big definitions of the top quark mass which are stable against the divergences of the mass of the top quark: The $\overline{MS}$-mass and the pole mass. In the $\overline{MS}$-scheme the infinities are absorbed into the value of the mass, while the pole mass is the point where the cross-section is maximal for the production of the top quark.

One possible way to measure the mass of the top quark in an unambiguous way is to exploit the dependence of the top cross-section on the top mass: The higher the top mass the lower the top cross-section is. This theoretical dependence is known with very high precision so if one measures the top cross-section the top mass can be inferred. Also this dependence can easily be parametrized in the $\overline{MS}$ mass thereby directly extracting a stable top mass.

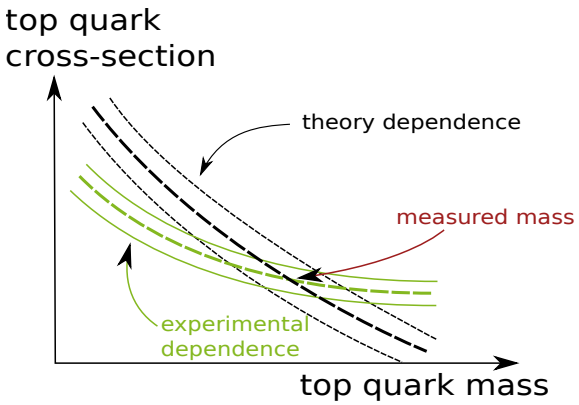

Figure 5.4: A sketch explaining the technique for measuring the top mass from the top cross-section. The black lines show the theory dependence of the top cross-section on the top quark mass (dashed) and its uncertainty (solid), the green lines the experimental dependence. The measured mass is where the dashed lines intersect, the uncertainty the area spanned by the solid lines around it.

Any real measurement of the top cross-section will also have a slight dependence of the top quark mass: It will use properties of the top events which can change with the top mass. This must also be parametrized and will give another curve of m_{Top} vs. σ_{Top}. This is sketched in Figure 5.4. The two parametrizations intersect at one point which gives the extracted top mass. Both lines also have their uncertainties (theory and experimental ones) which will span an area. This area gives the uncertainty on the extracted top mass using this technique.

5.5 Previous Measurements of the Mass of the Top Quark

5.5.1 Reconstruction of the Top Quark

Before one can attempt to measure the mass of the top quark one must identify a mode of production and observables to be used for the measurement.

Semileptonic Channel

Typically measurements of the mass of the top quark are done in the channel where one top quark decays hadronically and one decays leptonically. This allows to reconstruct the invariant mass of top and anti-top quark from their respective decay products.

This is sketched in Figure 5.5: The final state consists out of four jets, two of them originating from b-hadrons, a charged lepton and a neutrino. Events containing τ-leptons

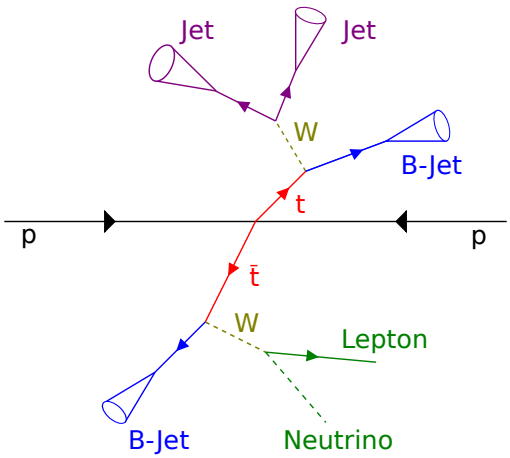

Figure 5.5: A sketch of top quark pair production and subsequent semileptonic decay. Both produced top quarks decay into a B-Jet and a W boson. One of the two W bosons decays into a lepton and neutrino, while the other one decays into two light jets.

are usually discarded for this decay channel due to their complicated decay modes.

This channel has the need of reconstruction of the top quark, an assignment of jets to the top quarks must be made. Typically this uses the invariant mass of the W to assign two of the four jets to the hadronic top quark. After this there are several options to find the hadronic b-jet: A simple example is maximising the p_T of the top quark candidate. On can also use the known W mass for constraining the longitudinal momentum of the neutrino. Then one can assign the b-jets to top quarks by using the one of the two possibilities which gives the smallest difference in mass of both quarks.

Dileptonic Channel

The full reconstruction of the invariant mass in the dileptonic channel is not possible. This is due to the ambiguity of the two neutrinos in this state: It consists out of two b-jets, two charged leptons and two neutrinos. From the latter one can only measure the sum of transverse momentum. The two three-vectors have in total six free parameters. Two of them can be constrained by requiring that the sum of the x- and y- components have to be the observed missing energy values in x- and y-direction. Two more can be constrained by using the mass of the W boson. This makes it impossible to exactly reconstruct the invariant mass of the two top quarks and one has to find other variables to measure the top quark mass.

All-hadronic Channel

In the all-hadronic channel the final state consists out of six jets, two of them b-jets. The reconstruction of the top quarks in the all-hadronic channel is done in a similar way as in the semileptonic channel. One can use the well known mass of the W to assign 2 pairs of jets to the the W bosons and then use some other algorithm to assign the b-jets.

5.5.2 1D-Template Method

Using the invariant mass of the two top quark candidates in the semileptonic channel one can directly apply a template method to extract the top quark mass. Another example of

a 1D-Template method is the measurement of the top quark mass in the dilepton decay modes. As in this mode the invariant mass of the top quarks can not be fully reconstructed, different observables need to be used: For instance one could associate leptons to b-jets and then use the invariant mass of the two objects. Another popular choice is called neutrino weighting. Here, one tries to recover the neutrino momenta by scanning over the possible phase-space of neutrino momenta and calculating a likelihood. The combination with the highest likelihood is taken and the invariant mass of the top quarks is built.

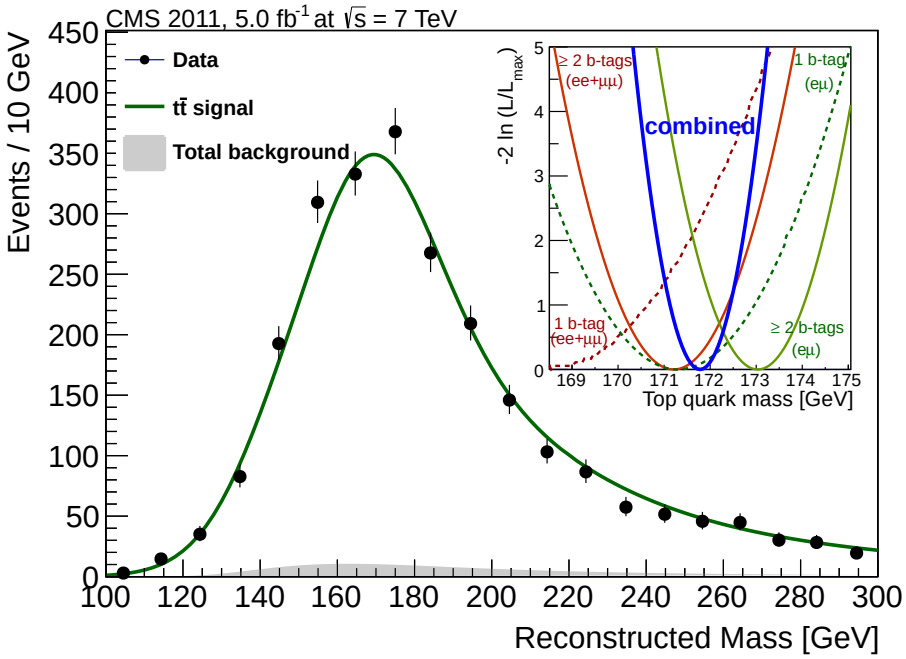

Figure 5.6: Reconstructed invariant mass of the dileptonic top quark candidates in data of the CMS collaboration together with the template fit to extract the top quark mass. The background is shown in the gray filled area. The figure in the top right shows the likelihood curves for various selections depending on the type of the leptons and the number of b-tags as well as their combination [65].

This approach was used by the CMS collaborationL Their reconstructed events are shown in Figure 5.6. This figure also shows the one-dimensional template fit and the likelihood for various top quark masses. They quote: $m_{\mathrm{Top}} = (172.5 \pm 0.4_{\mathrm{stat.}} \pm 1.5_{\mathrm{syst.}})\,\mathrm{GeV}$ [65]. An overview of the systematic uncertainties estimated is shown in Table 5.1. The dominant uncertainties are stemming from the calibration of the jet energies. Sizeable contributions are made by theory uncertainties like the choice of renormalization and factorisation scales.

5.5.3 2D-Template Method

The typical template method in use is a two-dimensional template method. It expands the number of observables by another one, typically the invariant mass of the W boson decaying into jets. The mass of the W boson has been measured to very high precision [66] so it can be used to in-situ calibrate the jet energy measurement. This gained knowledge can be propagated to the invariant mass of the hadronic top quark(s). The jet calibration is reduced to one global parameter, called "Jet Scale Factor" and is also determined in data. The advantage of this procedure is that the jet scale factor absorbs large parts of the uncertainties on the jet calibration thus reducing their effect on the top quark mass. However, this procedure increases the statistical uncertainty on the measurement as more

Table 5.1: An overview of the systematic uncertainties as quoted by the CMS collaboration of their measurement of the top quark mass in the dilepton channel using a neutrino weighting technique [65].

Source	Δm_{Top}[GeV]
jet energy scale	$^{+0.90}_{-0.97}$
b-jet energy scale	$^{+0.76}_{-0.66}$
jet energy resolution	± 0.14
lepton energy scale	± 0.14
unclustered missing energy	± 0.12
b-tagging efficiency	± 0.05
mistag rate	± 0.08
fit calibration	± 0.40
background normalisation	± 0.05
matching scale	± 0.19
renormalization an factorisation scale	± 0.55
pileup	± 0.11
PDFs	± 0.09
underlying event	± 0.26
colour reconnection	± 0.13
Monte-Carlo generator	± 0.04
total	± 1.48

information is extracted from the same amount of data.

An example for a 2D-Template measurement is the measurement of the top quark mass from CDF done using $8.7\,\mathrm{fb}^{-1}$ data [67]. They simultaneously fit the invariant mass distributions of the top as well as the invariant mass of the W boson. To enhance the statistical accuracy of the measurement they not only use the best assignment of jets to partons for reconstruction of the top quark mass but also the second-best, denoted $m_{\mathrm{Top}}(2)$. They extract the top quark mass to be $(m_{\mathrm{Top}} = 172.85 \pm 0.71_{\mathrm{stat.}} \pm 0.85_{\mathrm{syst.}})\,\mathrm{GeV}$ [67]. Example fits to the three observable in events containing at least one b-tag are shown in Figure 5.7. Their systematic uncertainties on the measurement are overviewed in Table 5.2.

5.5.4 3D-Template Method

ATLAS has measured the mass of the top quark using a three-dimensional template analysis. The additional dimension is used to additionally constrain the additional term in the jet energy scale, called the b-jet energy scale. It stems from the fact that jets originating from b-quarks are harder to measure. This is due to the fact, that some amount of b-quarks decay into a lepton and a neutrino. To also constrain this variable a third observable, sensitive to the b-jet energy scale is added: It is the ratio of the transverse momenta of the two light- and the two b-jets: $R_{lb} = \frac{p_T^{l1}+p_T^{l2}}{p_T^{b1}+p_T^{b2}}$. This allows the introduction of a b-jet scale factor and thus a constraint on the b-jet energy scale. They find that the top mass is $m_{\mathrm{Top}} = (172.31 \pm 0.75_{\mathrm{stat.\ +\ JSF\ +\ bJSF}} \pm 1.34_{\mathrm{syst.}})\,\mathrm{GeV}$ [68]. The fits to the three observable is shown in Figure 5.8 and the systematic uncertainties break down to the ones shown in Table 5.3. In this analysis one of the biggest uncertainty is the calibration of the b-tagging efficiency. This stems from a special sensitivity the R_{lb} variable has to the calibration of the b-tagging algorithms. Other big contributions arise from the residual jet energy scale and the amount of initial and final state radiation.

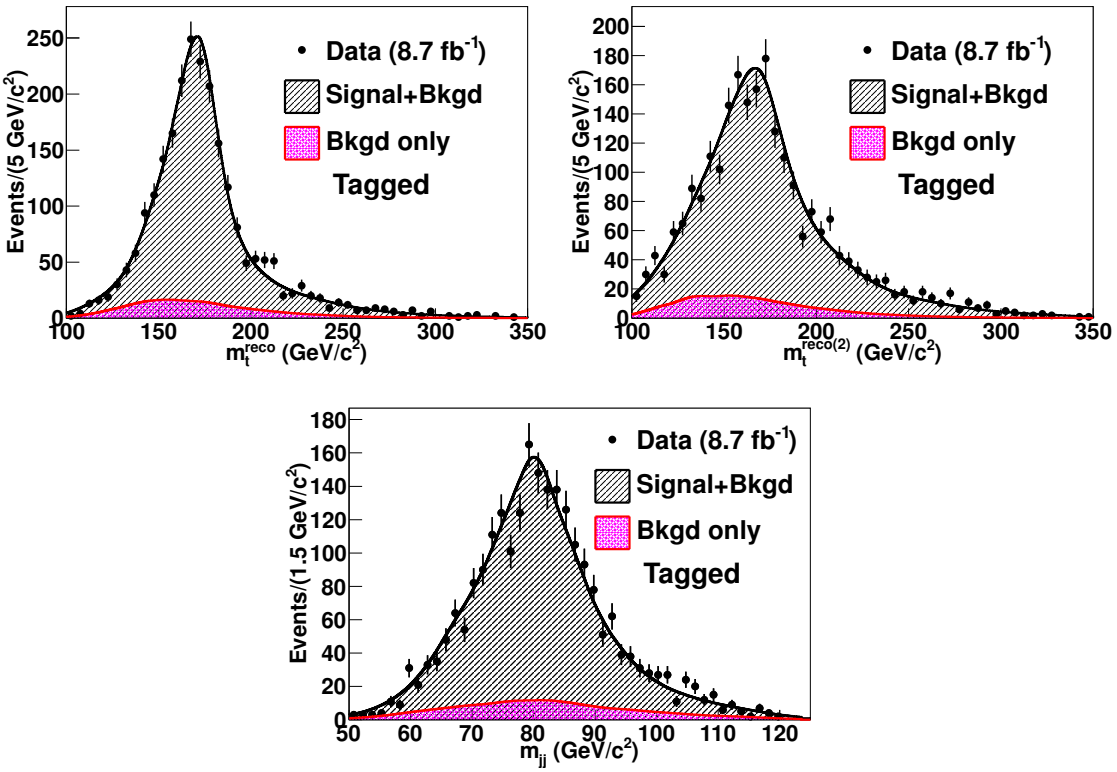

Figure 5.7: Invariant mass distributions of the best (top left) and second-best (top right) assignment of jets to top quark candidates and the invariant mass of the two jets assigned to hadronically decaying W boson (bottom). The template fit to each of the distributions is also shown [67].

Table 5.2: Systematic uncertainties on the top quark mass as determined by the CDF collaboration. [67].

Source	Δm_{Top}[GeV]
residual jet energy scale	± 0.52
signal modelling	± 0.57
b-jet energy scale	± 0.18
b-tagging efficiency	± 0.03
ISR/FSR	± 0.06
PDFs	± 0.08
gluon fusion fraction	± 0.03
lepton energy scale	± 0.03
background shape	± 0.20
multiple hadron interaction	± 0.07
colour reconnection	± 0.21
mc statistics	± 0.05
total	± 0.85

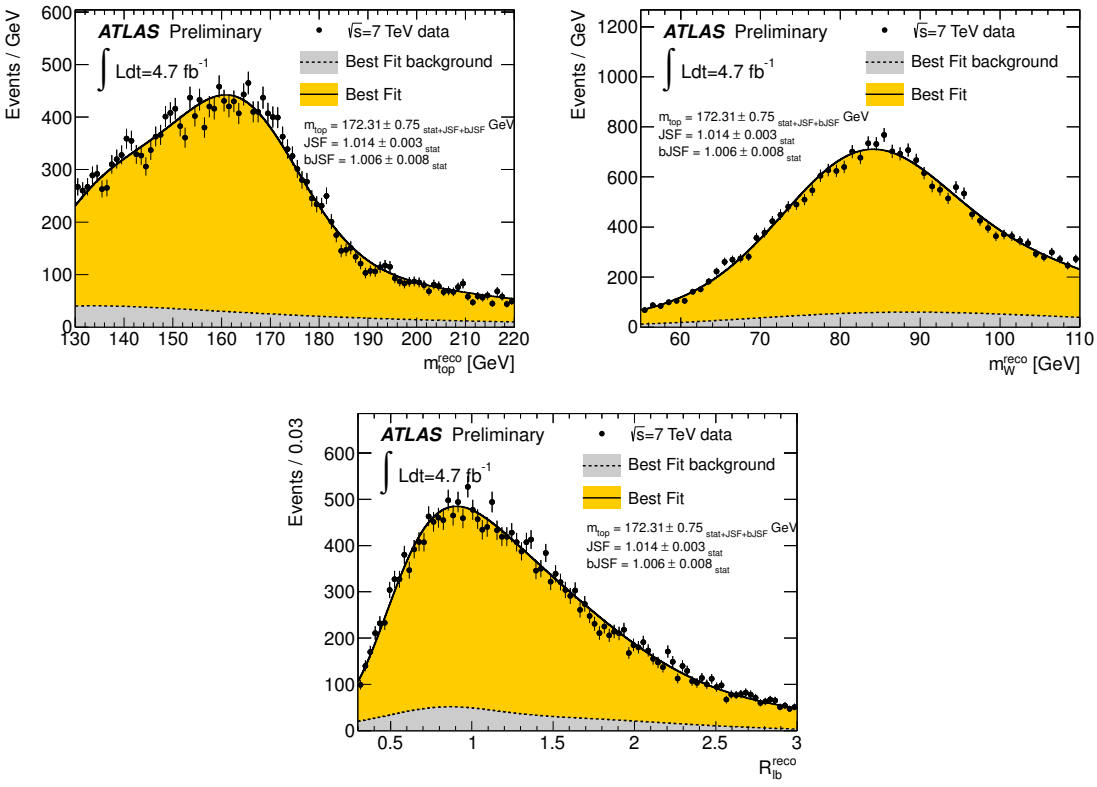

Figure 5.8: The distribution of the invariant mass of the top quark candidate (top left), the hadronically decaying W candidate (top right) and the R_{lb} variable (bottom) as found by the ATLAS collaboration in $4.7\,\mathrm{fb}^{-1}$ of data. The template fits to each of the distributions is shown as well [68].

Table 5.3: Systematic uncertainties on the top quark mass as estimated by ATLAS for their 3D template fit [68].

Source	Δm_{Top}[GeV]
jet energy scale factor	±0.27
b-jet energy scale factor	±0.67
method calibration	±0.13
signal mc generator	±0.19
hadronization	±0.27
underlying event	±0.12
colour reconnection	±0.32
ISR/FSR	±0.45
PDFs	±0.17
single top normalisation	±0.00
W+jets background	±0.03
QCD multijet background	±0.10
jet energy scale	±0.79
b-jet energy scale	±0.08
jet energy resolution	±0.22
jet reconstruction efficiency	±0.05
b-tagging efficiency and mistag rate	±0.81
lepton energy scale	±0.04
missing transverse momentum	±0.03
pile-up	±0.03
total	±1.35

5.5.5 Matrix Method

So far the matrix method has only been applied by the CDF and D0 collaborations at the Tevatron collider. This is due to the lower center of mass energy at the Tevatron. There, top quarks are mainly produced at rest which substantially simplifies the calculations needed to be done for solving the matrix elements.

An example is the application of the matrix method is the measurement done by D0 using $2.6\,\mathrm{fb}^{-1}$ of data. Similar as in the 2D-Template method they extend the matrix method by another observable to constrain the jet energy scale. Equation (5.2) can be expanded to measure more than one quantity. P does then not only depend on the top quark mass but also on a jet scale factor. Doing that D0 finds that $m_{\mathrm{Top}} = (176.01 \pm 1.01_{\mathrm{stat.}} \pm 0.79_{\mathrm{JES}} \pm 1.02_{\mathrm{syst.}})\,\mathrm{GeV}$ [69]. The observed likelihoods for the top quark mass and the jet scale factor are shown in Figure 5.9. A breakdown of the systematic uncertainties is displayed in Table 5.4. Dominant uncertainties are the jet energy scale and modelling related uncertainties like the hadronization and underlying event. The result also demonstrates the power of the matrix method — while using only a third of the data compared to the 2D-template measurement the statistical uncertainty is only 50% larger.

5.6 Dominating Uncertainties on present Measurements of the Mass of the Top Quark

The measurements presented in the previous section show similar dominating uncertainties: They are limited by the jet energy scale and the modelling of the top-antitop in the

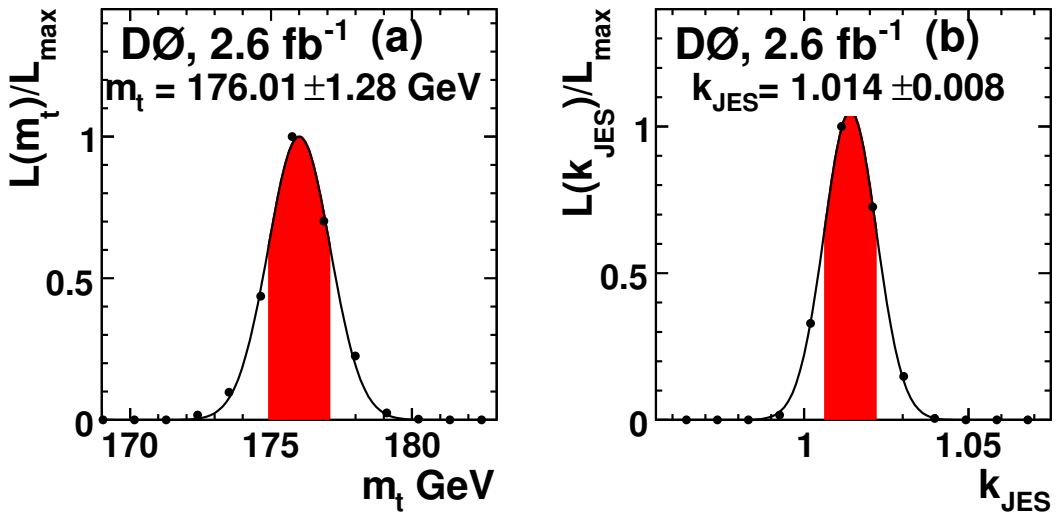

Figure 5.9: Projections of the observed data likelihoods for the top quark mass (left) and the jet scale factor (right) in the D0 application of the matrix method to measure the top quark mass [69].

Table 5.4: An overview of the systematic uncertainties as quoted by the D0 collaboration for their application of the matrix method to 2.6 fb^{-1} of data [69].

Source	Δm_{Top}[GeV]
higher-order effects	±0.25
ISR/FSR	±0.26
hadronization an underlying event	±0.58
colour reconnection	±0.28
multiple $p\bar{p}$ interactions	±0.07
modelling of background	±0.16
W+jets heavy-flavour scale factor	±0.07
modelling of b-jets	±0.09
choice of PDF	±0.24
residual jet energy scale	±0.21
data-mc jet response difference	±0.28
b-tagging efficiency	±0.08
trigger efficiency	±0.01
lepton momentum scale	±0.17
jet energy resolution	±0.32
jet id efficiency	±0.26
multijet contamination	±0.14
signal fraction	±0.10
mc calibration	±0.20
total	±1.02

simulation. These analyses are typical in that almost all measurements of the top quark mass have them as their limiting uncertainties.

The reasons for the uncertainties originating from the jet energy scale are simple: All described methods directly use the energies of the reconstructed jets in building their observables (in case of the template methods) or for the likelihood. As clearly visible in the one-dimensional template analysis by CMS (cp. Table 5.1) in that case the jet energy scale easily dominates the total uncertainty on the top quark mass measurement. This also explains the introduction of the two- and three-dimensional analyses by the various experiments. However, these multidimensional analyses will not completely marginalise the jet energy scale uncertainty.

The cause of the modelling-related uncertainties depend to a large extend on the individual analysis. As an example the uncertainty ATLAS calls hadronization in Table 5.3 is the difference between using Pythia and Herwig as the showering algorithm. The effect is caused by the dependence of the jet energies on the hadronization model. Normally that would be included in the jet energy scale uncertainty, but as that measurement does an in-situ calibration an additional effect on the top quark mass is introduced. Another example is the matrix method, as it uses the matrix element of the $t\bar{t}$ process naturally has a strong dependence on the choice of the parton distribution function (cp. Equation 5.2). These uncertainties can be controlled by using sensitive observables for the parameter in question and use data to tune the simulation.

Chapter 6

Development of a Template Method for the Transverse Decay Length and Momentum of the Lepton

The analysis of this thesis is a measurement of the mass of the top quark. The observables used have been designed to be as independent as possible on the jet energy scale. This will be explained in detail in this chapter. They are the transverse decay length of B Hadrons, L_{xy}, and the transverse momentum of the lepton from the top decay, $p_{\mathrm{T}}^{\mathrm{Lep}}$.

6.1 Observables with a small Dependence on the Jet Energy Scale

6.1.1 The Decay Length

In a top-antitop event there are typically two jets originating from bottom quarks. As discussed earlier, because of the structure of the CKM-Matrix, a hadron containing a bottom quark will travel up to several centimetres in the detector, a distance which can be measured very well. The reconstruction of this distance does not require any information from the calorimeter, therefore the variable will have a very small dependence on the jet energy scale. This distance is also called the "decay length" of the bottom hadron. Typically only the components perpendicular to the beam direction are considered for this variable due to the better tracking resolution. This variable is then called the "transverse decay length", L_{xy}. A sketch explaining the decay length variable is shown in Figure 6.1.

The decay length of bottom hadrons in top quark events depends on the mass of the top quark. This is due to the higher lorentz-factor, γ_B, the hadron receives if the mass of the top quark is larger. The lorentz-factor is defined in the following way:

$$\gamma_B = \frac{E_B}{m_B} \tag{6.1}$$

E_B and m_B denote the energy and mass of the bottom hadron. E_B can be further expressed by using the information of the two-body decay of the top quark into the bottom quark and the W boson:

$$p_b \cdot p_t = \frac{1}{2}\left(m_t^2 - m_W^2 + m_b^2\right) \tag{6.2}$$

$$= 2\left(E_t E_b - |\vec{p}_t||\vec{p}_b|\cos(\theta_{tb})\right) \tag{6.3}$$

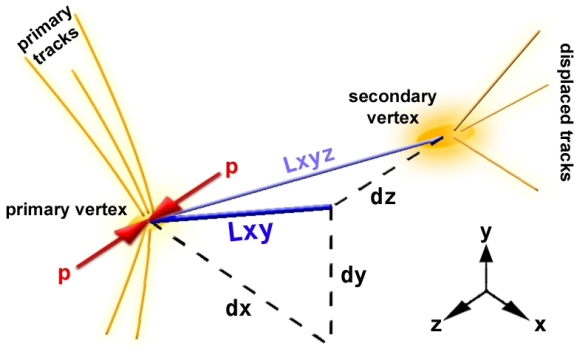

Figure 6.1: A sketch of the decay length variable. The protons interact at the primary vertex. The bottom hadron travels the distance to the secondary vertex. The total decay length is the three-dimensional distance, L_{xyz}, while for the analysis only the transverse decay length, L_{xy} is used [70].

m_t, m_W and m_b are the respective masses of the top quark, W boson and bottom quark, E_t and E_b the energies of the top and bottom quarks, $\vec{p}_t$ and $\vec{p}_b$ the three-vectors of them and θ_{tb} the angle between these vectors.

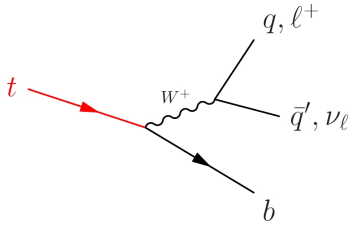

Figure 6.2: Feynman diagram of the top quark decay into a bottom quark and a W boson. The W boson may further decay into a lepton-neutrino- or a quark-anti-quark-pair.

This formula can be simplified by exploiting the following assumptions:

1. The mass of the top quark is much larger than the mass of the bottom quark, leading top $|\vec{p}_b| \approx E_b$. With $m_{\mathrm{Top}} \approx 175\,\mathrm{GeV}$ and $m_{\mathrm{Bottom}} \approx 5\,\mathrm{GeV}$ this condition is easily fulfilled.

2. The angle between the top and the bottom quark is relatively large, meaning $\cos(\theta_{tb}) \approx 1$. This has been studied in previous work and found to be true [71].

3. The momentum of the top quark is relatively slow, meaning that $E_t \approx m_{\mathrm{Top}} + \frac{1}{2}\frac{|\vec{p}_{\mathrm{Top}}|^2}{m_{\mathrm{Top}}}$. This is valid for top momenta up to roughly 200 GeV (compare Figure 6.3). At a center of mass energy of 7 TeV this is given for most of the produced top quarks. Also, this assumption is only made for simplicity purposes, since the dependence on the mass of the top quark is not diminished by its invalidity.

Using all these approximations, after a straightforward calculation one obtains:

$$\gamma_B = \frac{E_B}{m_B} \approx \frac{1}{2}\frac{m_{\mathrm{Top}} + |\vec{p}_{\mathrm{top}}|}{m_B}\left(1 - \left(\frac{m_W}{m_{\mathrm{Top}}}\right) + \left(\frac{m_B}{m_{\mathrm{Top}}}\right)\right) \tag{6.4}$$

The lorentz factor of the bottom hadron therefore is directly dependent on the mass of the top quark. The higher the mass of the top quark is, the higher the resulting lorentz factor will be. Due to special relativity this will directly lead to a larger travel distance of the bottom hadron in the detector frame of reference. Thus, by measuring the decay length it is possible to extract information about the mass of the top quark.

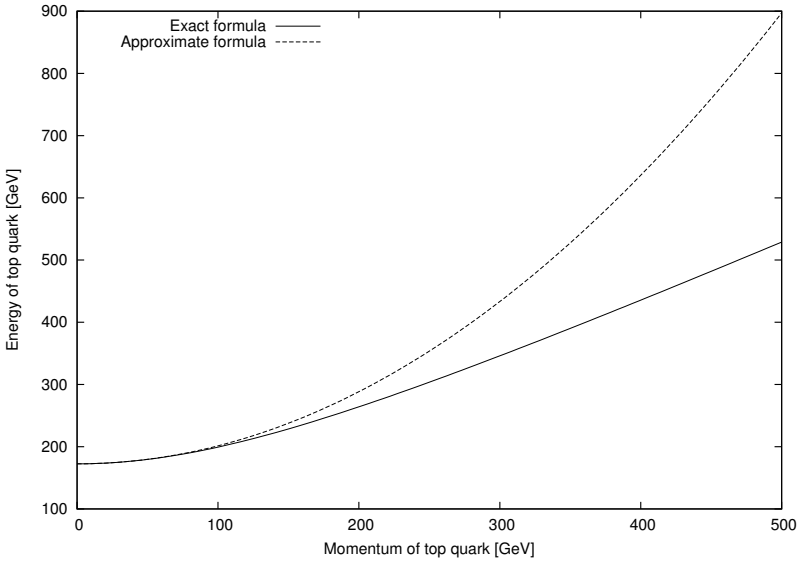

Figure 6.3: Approximate formula of assumption 3 in the text against the exact value. The assumed mass of the top quark has been 172.5 GeV. The approximation holds up to top quark momenta of about 200 GeV.

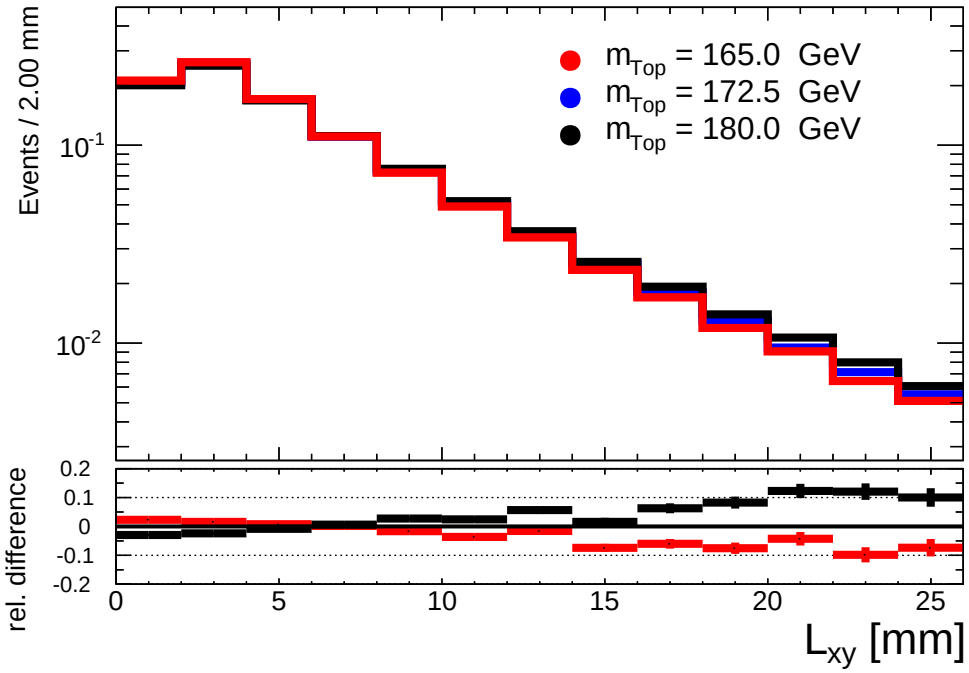

Figure 6.4: The decay length distribution in simulation for various assumed top quark masses. The used Monte-Carlo generator for all these samples is POWHEG, interfaced with PYTHIA.

This is further shown in Figure 6.4. It shows the distribution of decay lengths in ATLAS simulation. The used Monte-Carlo generator was POWHEG [61]. The dependence of the decay length distribution is clearly visible: The higher the mass of the top quark, the higher the average decay length will be.

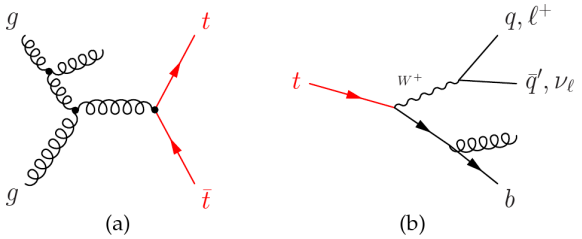

Figure 6.5: Feynman diagram of the $t\bar{t}$ initial (a) and final (b) state radiation processes.

The decay length distribution also has other dependencies. As seen in Equation (6.4), a different momentum of the top quark will yield different observed decay lengths. A good modelling of the momentum of the top quark is therefore imperative for the application of this method. Additionally, the variable is very sensitive to QCD final state radiation: If the bottom quark radiates a gluon before decaying its kinematics will be changed. In general, the more radiation occurs, the shorter the average decay length will be [72]. This is shown in Figure 6.6. This sensitivity is so strong, that the decay length would not be usable for a top quark mass measurement: If one would naively calculate the systematic effects of final state radiation on the transverse decay length the uncertainty would be about $\Delta^{\mathrm{FSR}} m_{\mathrm{Top}} \approx 5\,\mathrm{GeV}$.

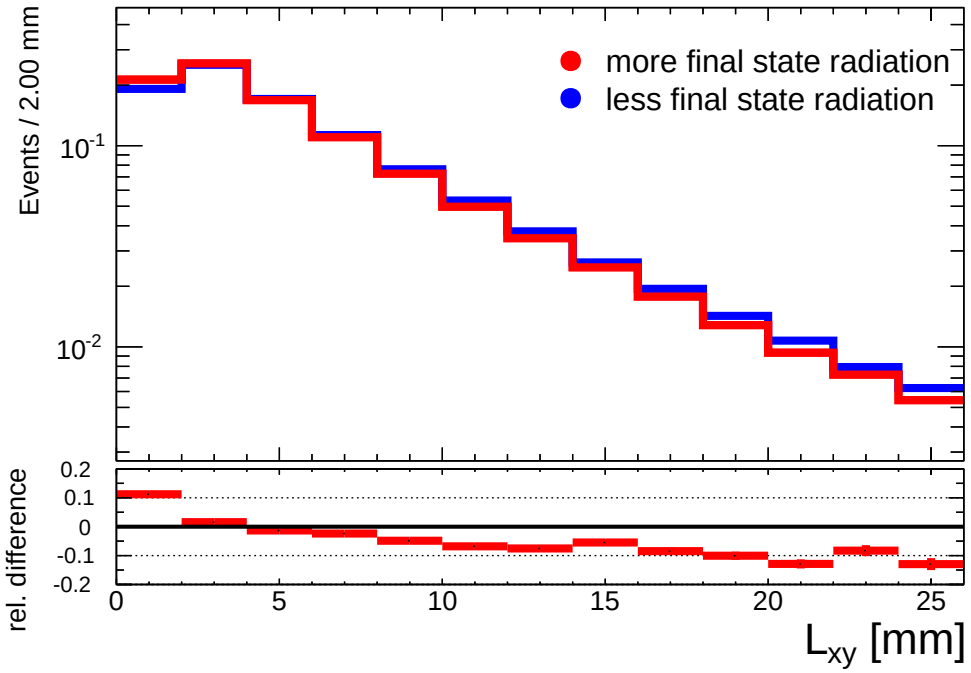

Figure 6.6: The decay length distribution in simulation for various assumed amounts of QCD final state radiation. The used Monte-Carlo generator for all these samples is ACERMC, interfaced with PYTHIA. The top quark mass for both samples was assumed to be 172.5 GeV.

However that sensitivity can be reduce by treating that systematic effect as a so-called "nuisance parameter", similarly to the JSF parameters mentioned in chapter 5. Thus the

data is used to also fit the effects of that systematic uncertainty. This will increase the statistical uncertainty on the measurement as more information is extracted but greatly reduces the uncertainty on the effects of final state radiation.

6.1.2 The transverse Momentum of the Lepton

Another variable which uses only very little calorimeter information is the transverse momentum of the lepton(s), $p_{\mathrm{T}}^{\mathrm{Lep}}$ from the $t\bar{t}$ decay. The reasoning behind the dependence of the transverse momentum is similar to the one of the transverse decay length: The higher the mass of the top quark, the higher the energy which is distributed on its decay products.

The transverse momentum has some advantages: As has been described in chapter 3, electrons and muons can be calibrated with very high accuracy due to the very precisely known mass of the Z boson. The energy scales of leptons should therefore impact the uncertainty on the mass of the top quark much less than the jet energy scale would if one would use a variable like the transverse momentum of the leading jet for the mass measurement.

As with the transverse decay length, the transverse momentum of the lepton does not depend on the top quarks mass as strongly as the invariant mass of its decay products would. A large amount of data is needed to acquire a similar precision as the analyses described in the previous chapter. Fortunately the LHC is able to provide this amount of data.

Figure 6.7 shows the dependence of the lepton transverse momentum distribution on the mass of the top quark. Only the contributions from electrons and muons have been considered in the creation of the figure. A good dependence of the distribution on the mass of the top quark is visible. Again, the higher the mass of the top quark, the higher the average momentum of the lepton will be.

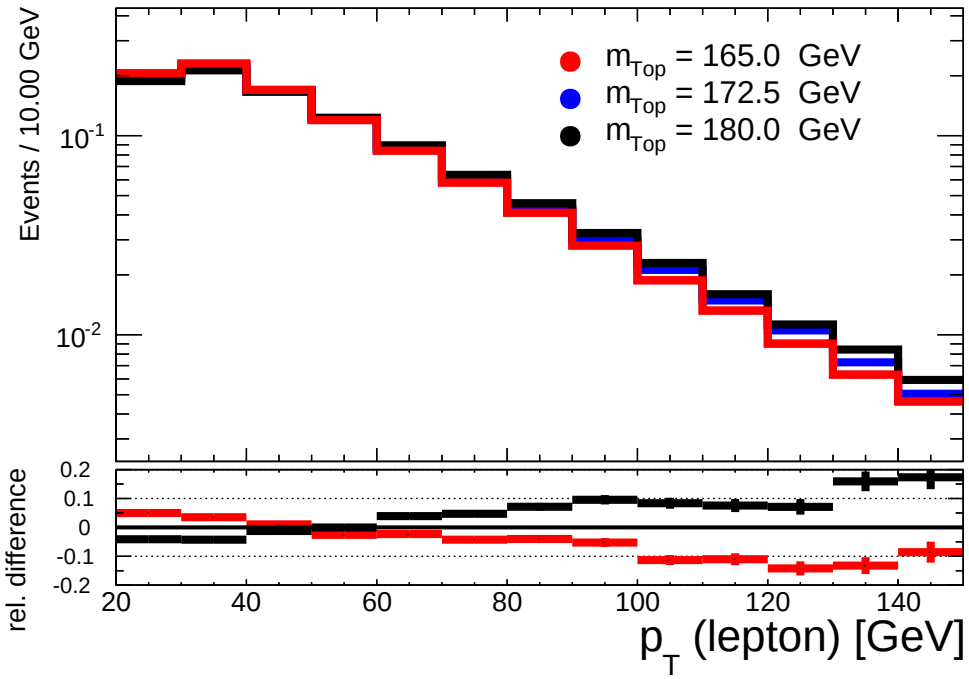

Figure 6.7: The lepton transverse momentum distribution in simulation for various assumed top quark masses. The used Monte-Carlo generator for all these samples is POWHEG, interfaced with PYTHIA.

The transverse momentum of the lepton has a similar dependence as the transverse decay length on the initial momentum of its mother top quark. As will be shown later this may introduce dependencies on modelling parameters, like the parton shower, which seem to be contra-intuitive. However, the dependence on the final state radiation modelling is very low. As there is little to none final state radiation from the mother top quark, and neither W boson nor the lepton are subject to QCD final state radiation this is expected. Figure 6.8 shows this effect. In comparison to Figure 6.6 the effect is very small.

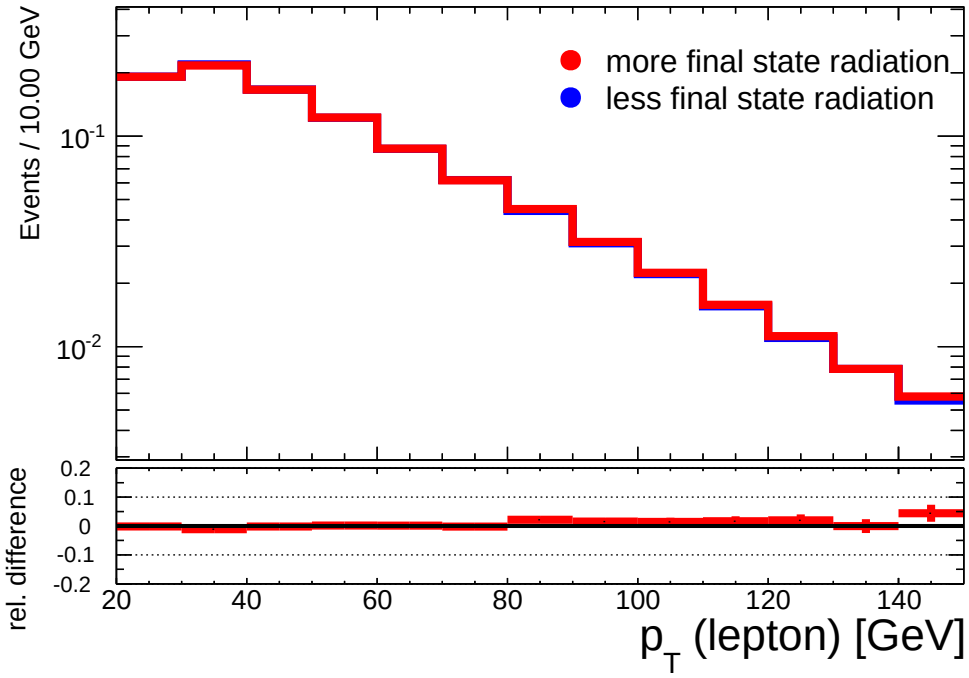

Figure 6.8: The lepton transverse momentum distribution in simulation for various assumed amounts of QCD final state radiation. The used Monte-Carlo generator for all these samples is AcerMC, interfaced with Pythia. The top quark mass for both samples was assumed to be 172.5 GeV.

6.1.3 Residual Jet Energy Dependence of the Parameters

In principle, both the transverse decay length and the transverse momentum of the lepton can be measured without using energy information of the hadronic calorimeter. In praxis, a small dependence on the jet energy scale remains. It originates in the impacts of the event selection which any analysis needs to apply in order to select $t\bar{t}$ events.

For the transverse decay length the effect comes from the correlation of the transverse momentum of the jet with the decay length of the bottom hadron it is associated with. This correlation is shown in Figure 6.9. Clearly, the higher the transverse momentum of the jet, the higher the transverse decay length will be. When there is a cut on the minimum transverse momentum of the jets in the event selection, varying the jet energy scale causes events to pass (not pass) the cut which did not (did) before. This changes the sample of used jets and, via the described correlation, changes the decay length and therefore introduces a dependence on the jet energy scale into this method.

This effect is displayed in Figure 6.10. It shows the default leading jet transverse momentum spectrum together with an upward- and a downward-variation of the jet energy scale. The total number of events changes from 18054 (default) to 17969 (downwards) and 18111 (upwards).

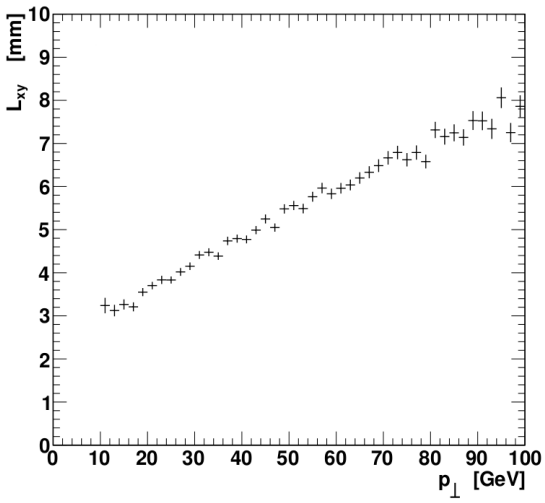

Figure 6.9: Correlation between the transverse momentum of the jet and the transverse decay length [71].

A similar effect exists for the transverse momentum of the lepton in the event. The effect is of a higher order than for the transverse decay length: Both the transverse momentum of the lepton and the transverse momenta of the jets are used in the calculation of the missing energy. A cut on the missing energy therefore impacts the transverse momentum of the lepton. Both these effects are rather small in impact and do not spoil the initial goal of having a very insensitive measurement of the mass of the top quark to the jet energy scale. They are important to mention though as they explain the small effects which will be exposed later.

6.2 Final State Radiation as a Nuisance Parameter

As mentioned in section 6.1.1, in order to use the information of the transverse decay length for the analysis, a special treatment of the effects of final state radiation is needed. It will be treated as a nuisance parameter. This means that the fit wich will be described in the next section will be extended by another observable called the FSR strength, S_F.

In the simulation, the final state radiation is modelled by the Pythia hadronization software. The strength of final state radiation is a parameter of the simulation which governs the strength of the QCD coupling when calculating the probability of another emission in the parton shower. This is governed by the `PARP(72)` parameter in the Pythia showering algorithm [64]. This parameter is the value of Λ_{QCD} in that calculation. Thus, a higher value of `PARP(72)` will yield more emissions in the parton shower. There exist two more parameters which have been changed in the ATLAS samples for ISR/FSR evaluation. They change the strength of initial state radiation: `PARP(64)` is a scale multiplier to Λ_{QCD} for parton emission in the initial state. Contrary to `PARP(72)` a larger value of it means less parton emission. The third parameter, `PARP(67)` is a starting scale multiplier for emission in the initial state.

From the values of these parameters, the FSR strength has been defined to be $S_F = +1$ for the current ATLAS upwards variation and $S_F = -1$ for the downwards variation. For the construction of the calibration also templates with $S_F = \pm 0.5$ have been constructed. The exact values are shown in Table 6.1. The samples used in the ATLAS variation also have variation on the initial state radiation. Their effect on the analysis is small, especially because they have been constrained by a separate analysis [73]. For completeness sake

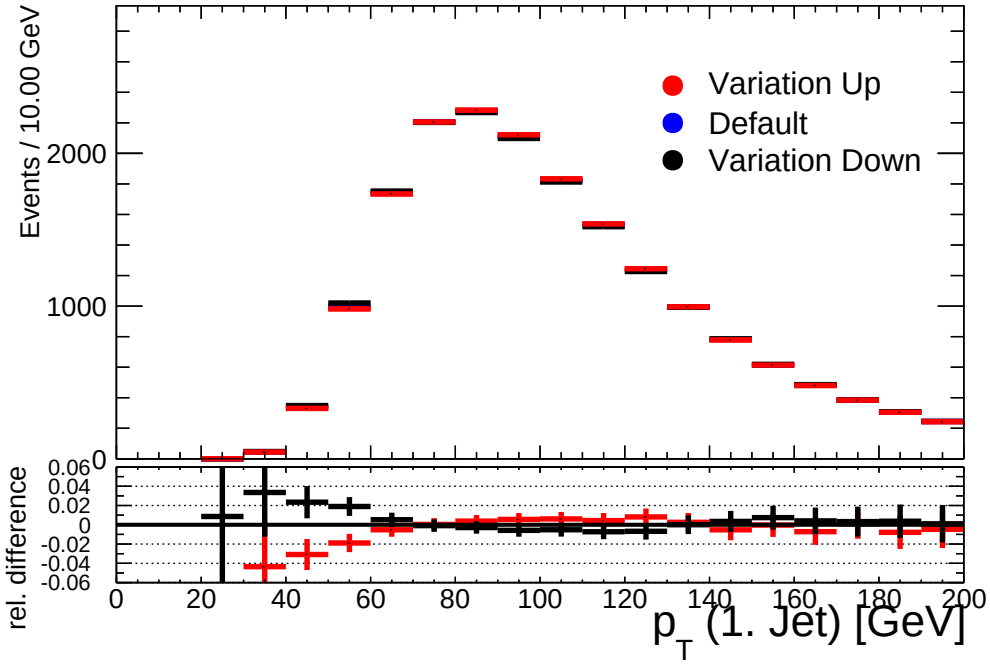

Figure 6.10: Display of migration effects when doing a variation of the jet energy scale. The number of events as well as the shape of the jet transverse momentum spectrum changes.

these variations are also shown in Table 6.1. In total there are four samples for the parametrization of the final state radiation, with FSR strengths of +1, $+\frac{1}{2}$, $-\frac{1}{2}$ and -1.

		assigned FSR strength				
Parameter	Description	1	0.5	0	-0.5	-1
`PARP(67)`	starting scale multiplier for ISR	1.40	1.20	1	0.80	0.60
`PARP(64)`	scale multiplier for Λ_{QCD} in ISR	0.90	0.95	1	2.25	3.50
`PARP(72)`	value of Λ_{QCD} in FSR	0.37	0.315	0.26	0.185	0.11

Table 6.1: Parameter variations for the amount of radiation. The two parameters which affect the scale parameter of the QCD, Λ_{QCD} act in the following way. The multiplier `PARP(64)` decreases the value of the strong coupling constant, α_s as it is multiplied to the scale. Thus, a higher value for it reduces the amount of ISR showers. Contrary, `PARP(72)` increases the amount of FSR as it is the actual value used for the scale.

6.3 A Template Fit for the transverse Momentum of the Lepton and the transverse Decay Length

6.3.1 Motivation

Earlier attempts of using the transverse decay length [70] or the transverse momentum of the lepton [74] have used the mean value of the respective distributions. This attempt has some serious flaws, as the mean value of the distribution is suspect to strong fluctuations as will be shown in the following.

Figure 6.11 shows two times the same distribution of 10000 random numbers, drawn

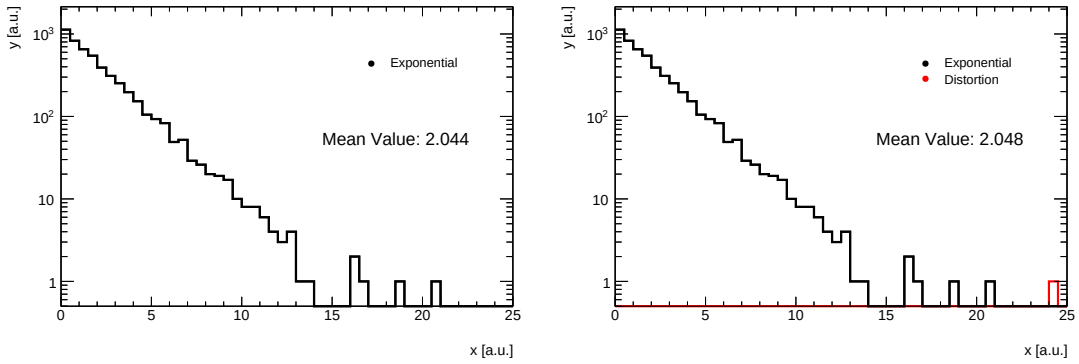

Figure 6.11: Distribution of random numbers, drawn from an exponential function. Left: without distortion, Right: with an additional distortion. The mean value of the distribution changes by 0.2%

from the following exponential function:

$$f(x) = \exp(-0.5x) \tag{6.5}$$

The mean value of the left distribution is 2.044, very close to the actual mean of 2.010. For the right distribution, an additional "event" with value of 24 has been added as a distortion. This changes the mean value to 2.048. The change is about 0.2%, which seems small at first but if compared to the achieved precision of the mass of the top quark is quite large. This effect gets in fact blown up when translating the mean value of the observed distribution to an actual mass value.

This effect can be offset a bit by using a more stable location parameter for a distribution like the median of the distribution. Still, reducing a whole distribution to a single parameter has some disadvantages as is displayed in Figure 6.12. It shows three distributions, a Gaussian, an exponential and a breit-wigner function, all with the same mean value of 2.010. Also, one is limited to a calibration curve approach for estimating the mass of the top quark. All in all, the use of a template method to describe both the transverse decay length and the lepton transverse momentum is highly desired and shall be developed in the next section.

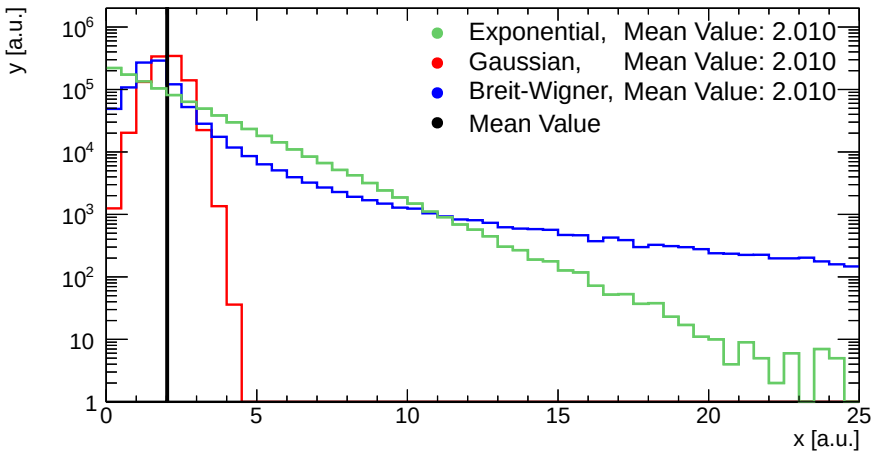

Figure 6.12: Distribution of random numbers, drawn from an exponential, a Gaussian and a breit wigner function. All functions have been constructed to have the same mean value.

6.3.2 Signal Parametrization

For any template method, analytic parametrization of the input distributions need to be found. In [75], the convolution of a fermi function with a gamma function was proposed as a proper description for the lepton transverse momentum distribution. The gamma function is supposed to model the falling end of the spectrum at high values of lepton transverse momentum and the fermi function the sharp decline at low values. Using this work as a suggestion, it was found that for both, the transverse momentum of the lepton and the transverse decay length, the sum of a fermi function and a gamma function does describe the distributions as well. The functional form, with the free parameters $p_i, i \in N_0$ is the following:

$$f(x) = \frac{1}{\exp\left(\left(x + p_1\right)/p_0\right)/p_2 + 1} + \frac{\left(\frac{x-p_4}{p_5}\right)^{p_3-1} \exp\left(-\frac{x-p_4}{p_5}\right)}{p_5 \Gamma(p_3)} \tag{6.6}$$

Both functions have three free parameters. To note is, that the parameter called p_2 is not part of a normal fermi function. It is necessary for the description of the transverse decay length distribution. In fact, for the lepton transverse momentum analysis this parameter has been fixed, as it does not depend on either the mass of the top quark. The same is true for parameters 0 and 3. For similar reasons, parameters 1, 2 and 3 are also fixed for the transverse decay length parametrization. Table 6.2 shows the values these parameters were fixed to. These fixed values of these parameters govern the general shape of the distribution and do depend slightly on the event selection (e.g. the minimum cut on the lepton transverse momentum). They need to be estimated in a separate step.

parameter	L_{xy} fixed value	$p_{\mathrm{T}}^{\mathrm{Lep}}$ fixed value
p_0	n/a	$(31.3 \pm 0.1)\,\mathrm{GeV}$
p_1	$(199 \pm 1)\,\mathrm{mm}$	n/a
p_2	-0.30 ± 0.01	1
p_3	0.45 ± 0.01	22.7 ± 0.1
p_4	n/a	n/a
p_5	n/a	n/a

Table 6.2: Values for the fixed parameters of Equation 6.6.

The resulting functions do describe the distribution of transverse decay length and lepton momentum very well. Figure 6.13 shows example fits to simulation samples. The fit range had to be limited to be between 0.5 mm and 25 mm for the transverse decay length distribution. Decay lengths with less than half a milimeter are hard to reconstruct, therefore they had to be excluded. At the range of more than 25 mm the influence of the beam pipe starts to affect the shape of the distribution. This is very well modelled by the simulation, but hard to take into account in an analytic formula. The fit range of the lepton transverse momentum has been restricted from 28 GeV to 150 GeV. The lower bound is the minimum possible value for an accurate description of the distribution. Lower values are too close to the cut value of 25 GeV for electrons. The upper cut value was chosen because of a known modelling issue of the simulation for leptons in that energy range.

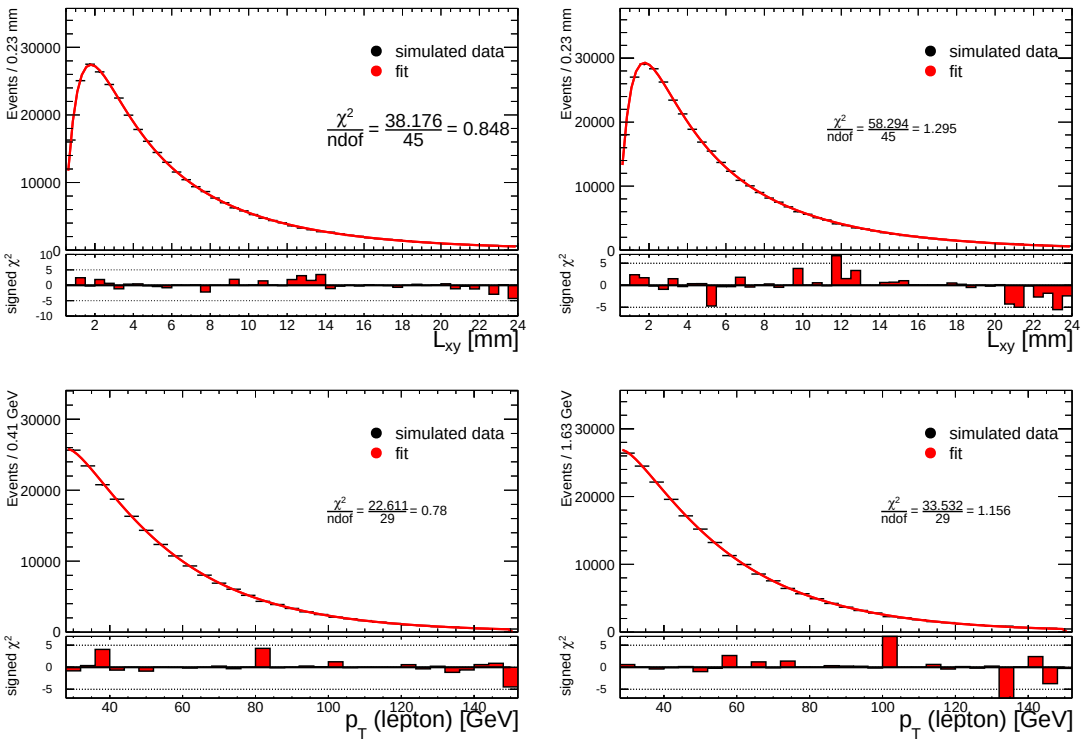

Figure 6.13: Example fits of the transverse decay length (top) and lepton transverse momentum (bottom). These fits have been done to the 'default' POWHEG+PYTHIA sample with an assumed top quark mass of 172.5 GeV (left), and to the ACERMC+PYTHIA sample with 50% more final state radiation and the same assumed top quark mass (right).

6.3.3 Removing Correlation between the Parameters

The fit parameters of the parametrization in Equation 6.6 are correlated. This can be expressed as the correlation matrix of the parameters:

$$V_{L_{\mathrm{xy}}} = \begin{pmatrix} 1. & 0.68 & 0.09 \\ 0.68 & 1. & 0.06 \\ 0.09 & 0.06 & 1. \end{pmatrix} \tag{6.7}$$

for the transverse decay length, and

$$V_{p_{\mathrm{T}}^{\mathrm{Lep}}} = \begin{pmatrix} 1.0 & 0.01 & 0.19 \\ 0.01 & 1.0 & 0.58 \\ 0.19 & 0.58 & 1.0 \end{pmatrix} \tag{6.8}$$

for the transverse momentum of the lepton. This correlation in fact spoils the sensitivity of the method. If parameters have a high correlation, both of them can be changed in a way that the fit quality is unchanged. The correlation should therefore be removed. The simplest way is to transform the parameters to a parameter space where all parameter are uncorrelated. For that the information of the covariance matrix can be used. The used technique is called the "Cholesky decomposition" [76]. It is a way to construct a transformation matrix to the linear independent parameter space. The final transformation matrix gives new fit parameters a_i which relate to the original parameters p_i in the following way:

$$\begin{pmatrix} p_0 \\ p_4 \\ p_5 \end{pmatrix} = \begin{pmatrix} 0.015\,\mathrm{mm} & 0 & 0 \\ -0.021\,\mathrm{mm} & 0.0143\,\mathrm{mm} & 0 \\ 0.002\,\mathrm{mm} & < 0.0001\,\mathrm{mm} & 0.0019\,\mathrm{mm} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}, \tag{6.9}$$

where the a_i denote the actually calibrated parameters. The matrix for the lepton transverse momentum parameters is:

$$\begin{pmatrix} p_1 \\ p_4 \\ p_5 \end{pmatrix} = \begin{pmatrix} 0.8342\,\mathrm{GeV} & 0 & 0 \\ 0.0005\,\mathrm{GeV} & 0.0055\,\mathrm{GeV} & 0 \\ 0.0472\,\mathrm{GeV} & -0.0876\,\mathrm{GeV} & 0.0447\,\mathrm{GeV} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \tag{6.10}$$

These parameters a_i are the final parameters which let the templates depend on the mass of the top quark and the amount of QCD final state radiation. Simulation samples with various values for both observables are fit and the dependence of each parameter of them is fit with a function. This is called calibration will be further described in the next chapter when the actual analysis is discussed. The calibration can then be used to generate templates which do only depend on the mass of the top quark and the amount of final state radiation. This can than be used to maximise a likelihood between the observed data and these two observables to extract values and uncertainties for them. The algorithm used to perform the minimization is called Migrad. It is part of the Minuit[77] package. The actual implementation in this theses is called pyMinuit [78].

6.3.4 Correlation between the Resulting Variables

Figure 6.14 shows the correlation between the estimated top quark mass and the estimated FSR strength. A correlation between the two quantities is visible – for a higher FSR strength the estimated mass is lower. This is due to the fact that the transverse decay length also has some sensitivity to the top quark mass.

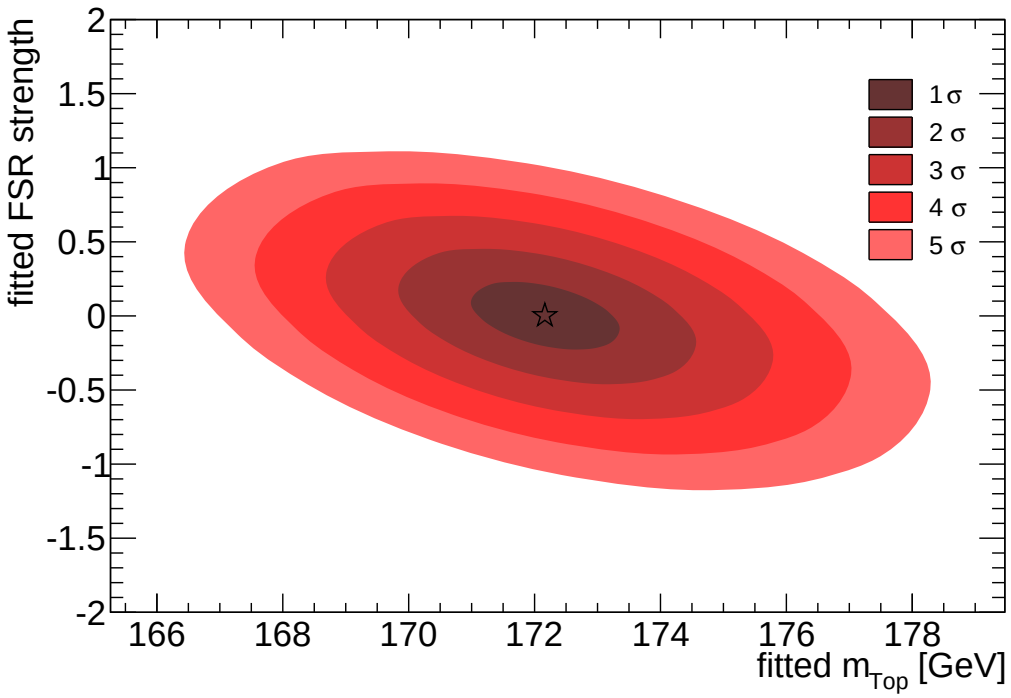

Figure 6.14: The correlation between the estimated top mass and the estimated FSR strength in the sample with an assumed top mass of 172.5 GeV. It is anti correlated with a correlation coefficient of -0.43. The fit has been done using a likelihood fit to the lepton p_T and transverse decay length distributions.

6.3.5 Background Treatment

Any analysis will have a fraction of background events in their data sample. This background needs to be taken into account in some way in the parametrization. One possible approach is to also find a template description for the background. The parameters of these templates will not depend on the mass of the top quark or the amount of final state radiation. In fact, they could be fixed by fitting only to the estimated background distributions. The fit would have an additional parameter which governs the fraction of background events, but it could also be fixed to the background estimation. Another options is to use the estimated shape directly as histogram templates. This approach is used here. In data the estimated background shape is subtracted from the data to get a 'signal-only' fit.

Chapter 7

Measurement of the Mass of the Top Quark with the Transverse Decay Length and Momentum of the Lepton

This chapter describes the actual measurement performed. It begins by describing the selection cuts used for identifying the physics objects. After that, the event selection cuts for selecting signal and control regions are described. After that the calibration of the templates is discussed. The the considered systematic uncertainties and their causes are described. Finally the method is applied to data and the mass of the top quark is measured.

This analysis uses the entire $\sqrt{s} = 7\,\text{TeV}$ dataset collected by the ATLAS detector in 2011. In total it amounts to an integrated luminosity of $\int L\text{d}t = 4.7\,\text{fb}^{-1}$.

7.1 Object and Event Selection

7.1.1 Object Definitions

This analysis makes use of jets, electrons, muons, missing energy and flavour tagging. The selection criteria applied for the reconstruction of each of them will now be described in more detail.

Jet definition

Jets have been reconstructed with the AntiKt algorithm ($R = 0.4$). They have been calibrated using the EM+JES calibration scheme. To reject electrons faking a jet, the closest jet to a selected electron (see below) is removed if it is within $\Delta R < 0.2$ of it. Pileup jets are rejected by applying a cut on the momentum-weighted fraction of tracks which are associated to this jet and the primary vertex. If that fraction, called "jet vertex fraction" is larger than 75% the jet is accepted.

Electron definition

Electrons need to fulfil the "tight++" electron identification criteria (cp. section 3.3.2). Their absolute pseudorapidity must be between $0 < |\eta| < 2.47$ with the window of $1.37 < |\eta| < 1.52$ excluded. Additionally, any electron which is within a $\Delta R < 0.4$ to any jet with $p_T > 25\,\text{GeV}$ is removed. Electrons with hadronic calorimeter activity in their flight direction are also rejected. These cuts are applied on the total transverse momentum

and energy in the hadronic calorimeter in a cone with size of $\Delta R < 0.3$ for the transverse momentum and $\Delta R < 0.2$ for the energy. The exact cuts depend on the selection efficiency for isolated electrons. The cuts have been chosen to have a signal efficiency of 90%.

Muon definition

Muons have to be combined muons with tight selection criterion. Algorithm chain 2 has been used (see section 3.3.4). Their absolute pseudorapidity has to be less than 2.5. To select muons from the correct primary interaction, a cut of $z_0 < 2.0\,\text{mm}$ on the longitudinal impact parameter is in place. Additionally, there are hit requirements for the various tracking layers: Muons need to have a hit in the b-layer, at least one hit in the other two pixel layers and at least 6 hits in the SCT. The cut requirements on the TRT depend on the pseudorapidity: For $|\eta| < 1.9$ there have to be at least 5 hits. Also, the fraction of outliers must be less than 10%. If the absolute pseudorapidity is higher than that value, the cut on the fraction of outliers is only applied if the track has at least 5 TRT hits associated to it. If the muon candidate passes a dead sensor in one of the semiconductor detectors, a hit is assumed for all of these requirements. To reject fake leptons from jets, muons within a cone of $\Delta R < 0.4$ of any jet are removed from the event, and they need to be isolated with similar cone variables as for the electrons. The exact cut values are:

- The calorimeter p_T in a cone with size of $R < 0.3$ around the muons is $p_T^{\text{cone}} < 2.5\,\text{GeV}$
- The calorimeter E in a cone with size of $R < 0.2$ around the muons is $E_T^{\text{cone}} < 4\,\text{GeV}$

Missing Energy definition

The missing energy, E_T^{miss}, is, as described in section 3.3.6 the negative sum of all other objects found in the event. The missing transverse energy depends slightly on the selection criteria for these objects as stated above. The used definition of missing transverse energy is consistent with the above definition of objects. In addition to the objects discussed above, unclustered energy in calorimeter cells is also included in the calculation. When the four-vectors of any object are changed by a systematic uncertainty the changes are propagated to the missing transverse energy as well.

Flavour tagging

The MV1 b-tagging algorithm is used for the identification of bottom quark jets. For this analysis the operating point with 70% nominal efficiency is used. It corresponds to a cut value of $w_{\text{MV1}} > 0.60173$ for a jet to be considered tagged. As described in chapter 3 the MV1 algorithm is a neuronal net combining the output of several other flavor tagging algorithms, two of which use secondary vertex information (JetFitter and SV1) and one using track impact parameter information (IP3D).

7.1.2 Common Preselection

All signal and control regions share a common preselection. It reads the following:

1. A primary vertex with at least four tracks associated to it is present in the event
2. The event has been triggered by an appropriate single lepton trigger:
   (a) For electrons this is, depending on the run period the trigger called "EF_e20_medium" (periods B-J), "EF_e22_medium" (period K) or either "EF_e22vh_medium1" or "EF_e45_medium1" trigger (periods L and M);-

(b) For muons two different triggers were in place depending on the run period: It is either the trigger "EF_mu18" (periods B through I) or the trigger "EF_mu18_medium" (periods J to M).

This is also summarised in Table 7.1

3. There is exactly one isolated lepton (electron or muon) with transverse momentum of at least 25 GeV if it is an electron or at least 20 GeV if it is a muon

4. This lepton was the one which fired the trigger

5. Quality cuts on the event: There is no electron which is overlapping with a muon or vice versa; No noise in the LAr calorimeter

6. There are at least three jets with transverse momentum of more than 25 GeV in the event.

Table 7.1: Used muon and electron triggers in the various periods of data-taking.

Data Period	Muon Trigger	Electron Trigger
B-I	EF_mu18	EF_e20_medium
J	EF_mu18_medium	EF_e20_medium
K	EF_mu18_medium	EF_e22_medium
L-M	EF_mu18_medium	EF_e22vh_medium1

Additionally, there is a requirement which is only present in the data: The event must not be rejected by the data quality flags. A GoodRunList (see section 3.2.6) is applied. All detector elements needed to be in good condition for an event to be accepted.

This common preselection allows to select regions where either signal or the two larger background components (W+jets or QCD fakes) are dominant.

7.1.3 Background Estimation

The amount of background needs to be estimated. This is especially important for the larger backgrounds. The smaller backgrounds, Z+jets production, single top production and diboson production, do not impact the analysis very much. Their yield has been estimated by the theoretical cross-section for the respective cross-sections. The cut efficiency and the shape for these processes has been taken from MC simulation. Table 7.2 shows an overview of the various processes and their estimated cross-sections.

Table 7.2: Used cross-sections for the smaller backgrounds.

Process	cross-section [pb]
Single Top	84 [79, 80, 81]
Z+Jets	3256.64[82]
Diboson	4.69 [82]

For the other two considered backgrounds, W+jets production and QCD fakes an accurate description of both the yield and the shape is necessary.

7.1.4 QCD Fakes Estimation

A fake lepton is called an object which is either no lepton (e.g. a jet) or a lepton from the cascade of decays in a jet which passes the lepton selection criterion. The probability for this happening is in general very low ($\ll 1\%$), but the cross-section for QCD multijet production cross-section is several orders of magnitude larger than the $t\bar{t}$ cross-section. A small amount of QCD multijet events with fake leptons will pass the event selection.

While the QCD multijet process can be simulated with reasonable description of the data, it is unrealistic to use Monte-Carlo to describe the QCD fakes. Due to the very low fake efficiency one would have to simulate a very large number of events to obtain sufficient statistics of simulated events for an accurate background shape.

Another way is possible by using a data-driven technique to extract the fake rate and shape. The technique is called the "matrix method". It makes use of two groups of events, one called "loose" events, the other one called "tight" events. Tight events must always also be loose events for the method to work as presented here. The criterion itself does not have very strong requirements: The efficiency for the fake events, ε_f, must be very different from the efficiency for real events, ε_r. The total number of fake events which pass the the tight selection, N_f^t can then be calculated from the total number of tight events, N^t, the total number of loose events N^f and the efficiencies for loose and tight events in the following way:

$$N_f^t = \frac{\varepsilon_f}{\varepsilon_r - \varepsilon_f}(N_f^l \varepsilon_r - N^t) \tag{7.1}$$

For QCD fakes one is interested in the probability that a QCD multijet event fakes a tight lepton. Therefore, lepton criteria are good choices for the selection criterion. In this analysis three different criterions are used, two for muons and one for electrons:

- Both muon analysis use the isolation requirement of the muon as the criterion. A fake muon from QCD multijet production is likely produced inside or close to a jet, therefore its fake efficiency is rather low. For a $t\bar{t}$ lepton this is not the case, it is produced from an isolated W boson. The tight selection is identical to the one described earlier this selection. The fake selection drops the following two cuts:

  - The calorimeter p_T in a cone with size of $R < 0.3$ around the muons is $p_T^{\text{cone}} < 2.5\,\text{GeV}$
  - The calorimeter E in a cone with size of $R < 0.2$ around the muons is $E_T^{\text{cone}} < 4\,\text{GeV}$

  The two methods differ in the way of parametrization the efficiencies as functions of the event topology.

- The electron matrix method tight definition is the same as the electron selection earlier in this section. The loose selection drops the b-layer hit requirement for the track associated to the electron and the isolation criteria.

The formula given in Equation (7.1) can be generalised to obtain shapes of distributions. Instead of calculating the total number of events in the selection, one can separate the events into several bins of the distribution one is interested in. Then one applies Equation (7.1) for each of these bins and obtains the number of tight fake events for that bin. The obtained histogram is the distribution of fake events. The uncertainty on the QCD fakes estimation is dominated by the uncertainties on the measurement of the fake and real efficiency. It is $\pm 50\%$ on the number of fake events which pass the tight selection.

7.1.5 W+Jets Estimation

The production of W+jets has a similar signature as the $t\bar{t}$ process: It involves an isolated lepton, missing energy and additional jets. The amount of bottom quark jets is in general lower though, so flavour tagging is a viable way of reducing that background. Still the amount of W+jets events which pass the final selection can be up to 10-20%. An accurate description of it is therefore necessary. The theoretical cross-section for the $W+n$ partons process is hard to calculate for a large number of partons, n. However, the ratio of positively charged to negatively charged W bosons is much easier to calculate and is therefore better modelled by the simulation. The ratio is not 0.5 as one would naively expect. This is due to the initial state at the LHC which is not charge-symmetric. In fact, the ratio $R = {}^{N_+}/_{N_-}$ is larger than one. If the ratio of simulation R_M and data R_D is the same, the total number of W+jets events, $D = D_+ + D_-$ can be estimated from R_M, and the difference of $D_+ - D_-$ in the following way:

$$
\begin{aligned}
D &= D_+ + D_- && (7.2)\\
&= D_+ + D_- \left(\frac{D_+ - D_-}{D_+ - D_-}\right) && (7.3)\\
&= (D_+ - D_-)\left(\frac{D_+ + D_-}{D_+ - D_-}\right) && (7.4)\\
&= (D_+ - D_-)\left(\frac{{}^{D_+}/_{D_-} + 1}{{}^{D_+}/_{D_-} - 1}\right) && (7.5)\\
&= (D_+ - D_-)\left(\frac{R_D + 1}{R_D - 1}\right) && (7.6)\\
&= (D_+ - D_-)\left(\frac{R_M + 1}{R_M - 1}\right). && (7.7)
\end{aligned}
$$

The last line uses that R_M equals R_D. This method assumes that all other contributions to the data, especially the $t\bar{t}$ process is charge-symmetric. For the single top quark and diboson contribution a small asymmetry exists. Their simulated contributions are subtracted from the data before applying Equation (7.7). Thus the remainder which yields D_+ and D_- only consists out of W+jets events.

The shape of the W+jets process is taken from the simulation. Control regions have been used to verify that the modelling of the simulation is adequate. One aspect still needs to be corrected for. The production rates of different heavy flavour in association with a W boson may vary from the simulation fractions. For that several groups of W+ jets production are defined: W+light flavour, $W+bb$ (two bottom quarks), $W+cc$ (two charm quarks), and $W+c$ (one charm quark). Correction scale factors have been measured to change the relative fractions of these four components, but not the overall yield.

7.1.6 Signal and Control Regions

Starting from the common base selection described in Section 7.1.2, two control and one signal region have been defined. Both control regions are separated from the signal region by the cut on the number of jets in the event and the flavour tagging requirement: The signal region requires four or more jets while the control regions both require exactly three jets in the event. Both control regions cut on no flavour tagged jet in the event, while in the signal region one or more jets must be tagged. All three regions also have cuts on the transverse mass of the W boson, m_T^W[1], and the missing energy. Both variables are very powerful in discriminating between QCD fakes and events with real W bosons. The QCD fake region requires $E_{\mathrm{T}}^{\mathrm{miss}} < 30\,\mathrm{GeV}$ and $m_T^W < 30\,\mathrm{GeV}$, the W+jets region and the signal

[1] $m_T^W = E_{\mathrm{T}}^{\mathrm{miss}} p_T^{\mathrm{lep}} \left(1 - \cos\left(\Delta\phi_{\mathrm{lep}}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)$

region $E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$ and $m_W^T > 30\,\mathrm{GeV}$. Finally, the signal region rejects events where any two jets are closer to each other than $\Delta R = 0.8$. Table 7.3 shows an overview over the cuts which define the various regions.

Table 7.3: Additional cut defining the QCD Fakes and W+jets control regions and the signal region.

Cut	QCD Fakes region	W+jets region	Signal region
Number of jets	$N_{\mathrm{jets}} = 3$	$N_{\mathrm{jets}} = 3$	$N_{\mathrm{jets}} >= 4$
Number of b-tagged jets	$N_{\mathrm{tags}} = 0$	$N_{\mathrm{tags}} = 0$	$N_{\mathrm{tags}} >= 1$
Missing Energy	$E_{\mathrm{T}}^{\mathrm{miss}} < 30\,\mathrm{GeV}$	$E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$	$E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$
Transverse W mass	$m_W^T < 30\,\mathrm{GeV}$	$m_W^T > 30\,\mathrm{GeV}$	$m_W^T > 30\,\mathrm{GeV}$
Miscellaneous	n/a	n/a	$\min(\Delta R_{\mathrm{jet,jet}}) > 0.8$

Figures 7.1 shows the transverse momentum of the lepton and the transverse momentum of the jet with the highest transverse momentum (also called the "leading jet") in data together with various signal and background components. The used selection in this plot is the QCD fakes control region selection. The figure also shows the result of the fake lepton contribution as described in Section 7.1.4. For the muon channel there are little other contributions then the QCD fakes estimate, in the electron channel there is a sizeable contribution from the amount of Z+jets production. The shown uncertainties on these plots are the uncertainty on the fake lepton contribution. These stem from the uncertainties on the measurements on the real and fake efficiencies. Within these uncertainties, both distributions are well described by the estimation. The two methods of the muon channel have been averaged in the subfigures in the muon channel. All in all the description of the data is very good showing that the QCD fake estimate is working well.

Figure 7.2 shows the same distributions as Figure 7.1 but for the W+jets control region. Clearly, this control region is dominate by W+jets production. The shown uncertainty includes the uncertainty on the estimate on the W+ jets yield as described in Section 7.1.5. The description of the data by the simulation, combined with the yield estimate is very good. However, in the lepton transverse momentum spectrum there is some disagreement between the data and the simulation for events where the transverse momentum of the lepton is larger than 150 GeV. This disagreement is the reason why the fit range was chosen to stop at that value.

Table 7.4 shows the expected amount of $t\bar{t}$ signal and the various background processes as well as the observed numbers of events in data in the signal regions. The used cross-section of top-antitop quark pair production is $\sigma_{t\bar{t}} = 177^{+10}_{-11}\,\mathrm{pb}$. It has been calculated for a mass of the top quark of $m_{\mathrm{Top}} = 172.5\,\mathrm{GeV}$. The matrix element has been calculated at next-to-next-to-leading order (NNLO) with next-to-next-to-leading logarithmic resummation [40, 83, 84, 85, 86, 87]. This region is clearly dominated by the amount of the $t\bar{t}$ process. Dominant background is the W+jets process, followed by the single top process. All in all the remaining background fraction is 19%.

Figures 7.3 through 7.5 show various distributions in the signal region. Figure 7.3 shows the leading jets transverse momentum and pseudorapidity. Both distributions are well-modelled by the simulation. In Figure 7.4 distributions of the missing transverse energy and the transverse mass of the W boson are shown. These two distributions are sensitive to various background components (especially) QCD fakes and the fact that they are well-modelled shows again that the background has been correctly estimated.

Finally in Figure 7.5 distributions of the number of tagged events by the used flavour tagging algorithm and the minimum distance between any two jets in the event are shown. Both distributions – as the transverse mass of the W boson and the missing transverse energy – are used in the event selection and should be well-modelled by the simulation.

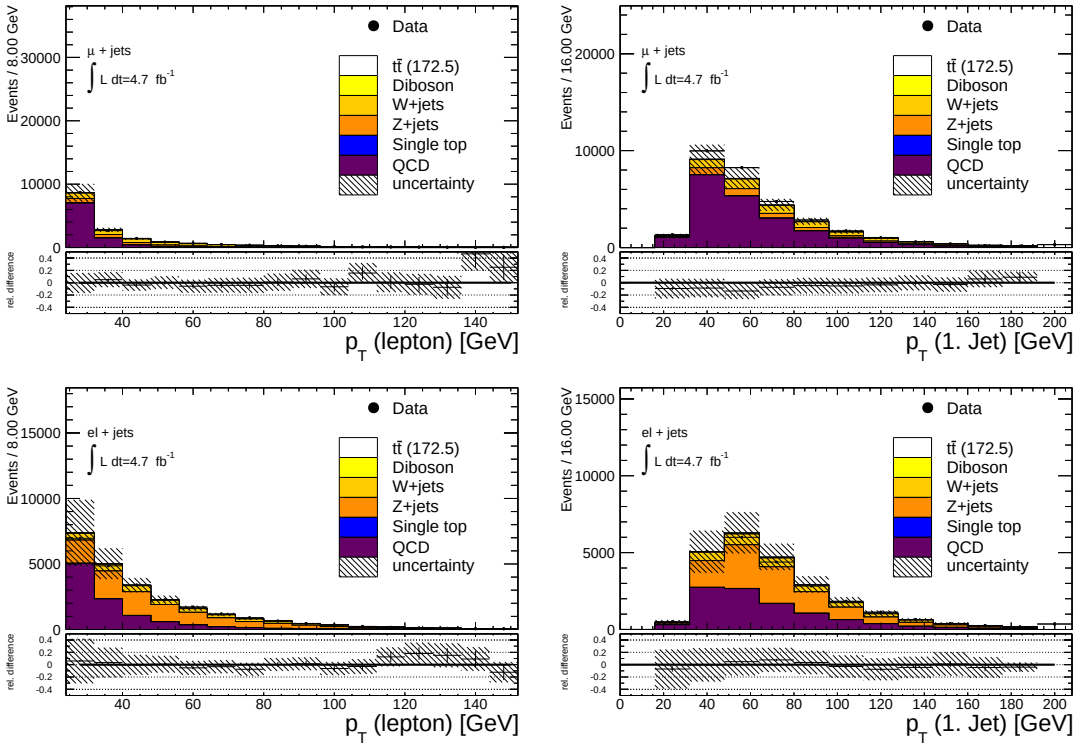

Figure 7.1: The transverse momenta of the lepton (left) and the leading jet (right) in the QCD fakes region. Top: If the lepton is a muon, Bottom: if it is an electron. The shown uncertainty includes the uncertainty on the fake estimate.

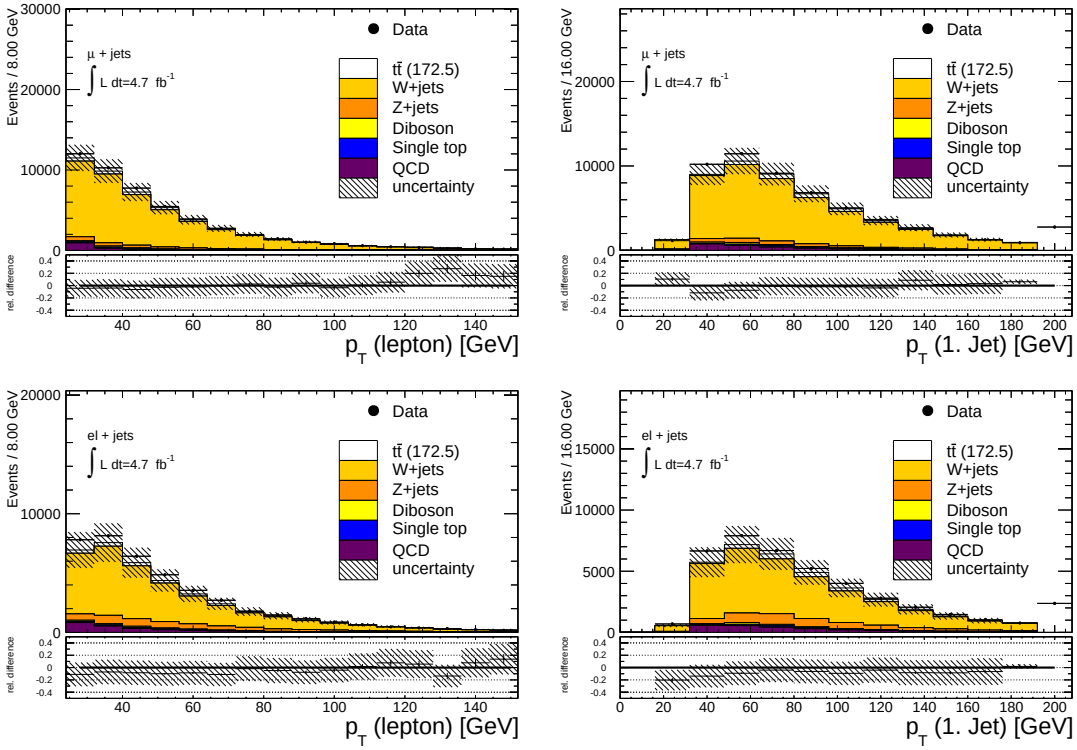

Figure 7.2: The transverse momenta of the lepton (left) and the leading jet (right) in the $W + jets$ region. Top: If the lepton is a muon, Bottom: if it is an electron. The shown uncertainty includes the uncertainty on the W+jets yield estimate.

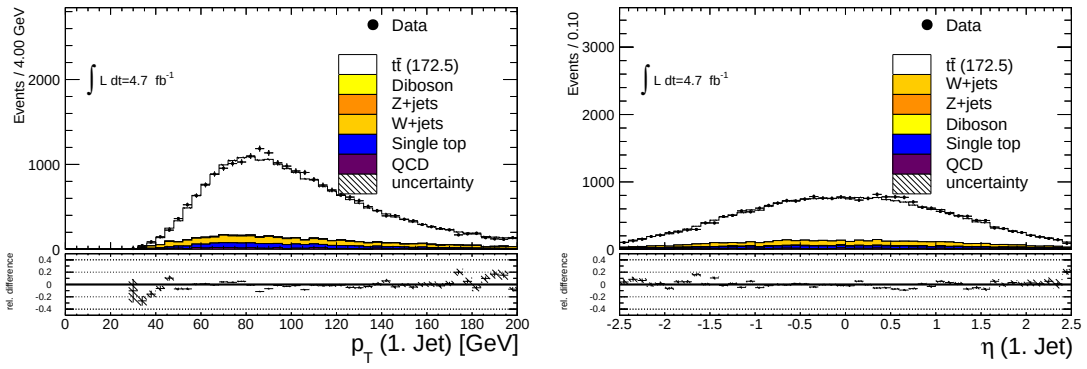

Figure 7.3: The transverse momenta (left) and the pseudorapidity (right) of the leading jet in the signal region.

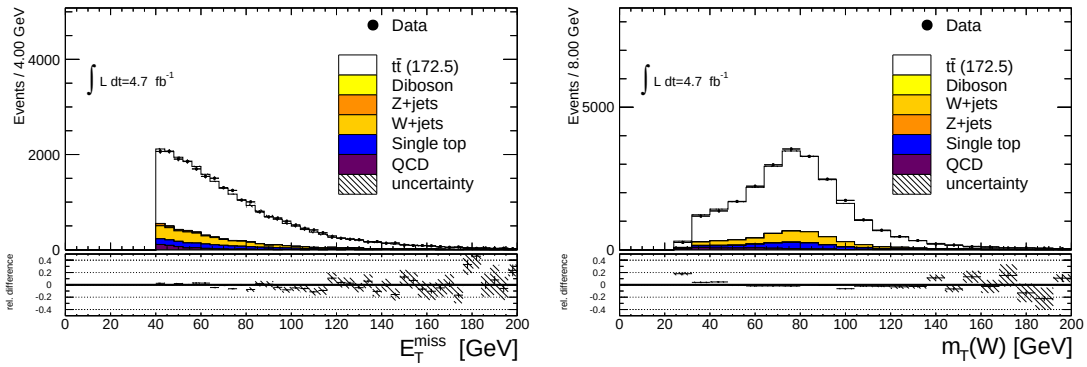

Figure 7.4: The transverse momenta (left) and the pseudorapidity (right) of the leading jet in the signal region.

Table 7.4: Number of expected events for the signal and various background processes as well as obtained number of events in data.

Process	Electron Channel	Muon Channel
QCD	140.2 ± 70.1	182.9 ± 91.5
Single Top	567.6 ± 7.0	799.7 ± 8.2
Diboson	18.3 ± 0.8	25.5 ± 1.0
W+jets	956.5 ± 124.4	1567.0 ± 219.4
Z+jets	128.1 ± 6.1	70.2 ± 5.2
$t\bar{t}$ (172.5)	8141.8 ± 19.5	$11\,246.6 \pm 22.9$
Total MC	9952.4 ± 144.4	$13\,891.8 \pm 239.0$
Data	$10\,017.0 \pm 100.1$	$14\,146.0 \pm 118.9$

This is true for both distributions. The number of b-tagged jets is slightly mismodelled but this is covered by the uncertainty on the flavour tagging calibration.

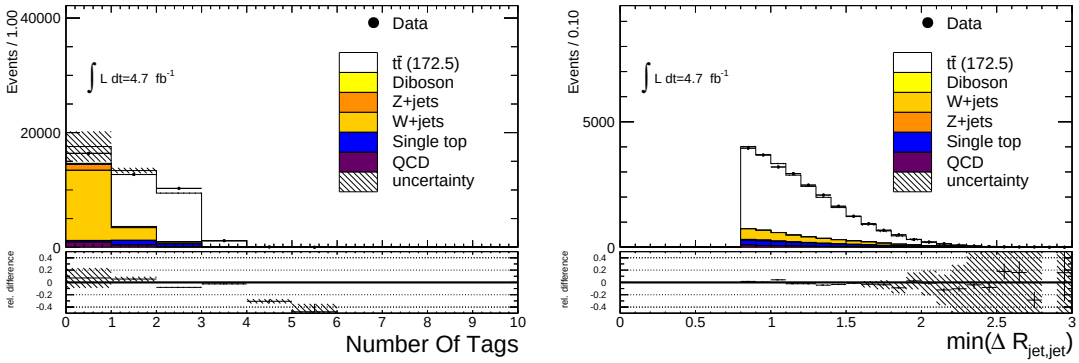

Figure 7.5: Number of b-tagged jets (left) and the minimum distance between any two jets in the event (right) in the signal region.

7.2 The Calibration Curves

The parametrization described in section 6.3.2 needs to be calibrated. That means, that the dependence of the fit parameters on the mass of the top quark and the strength of QCD final state radiation is fit. These fits for the various parameters a_i are shown in Figures 7.6 and 7.7.

The values of the a_i must be transformed with the matrices shown in Equations (6.9) and (6.10) to be the parameters shown in Equation (6.6). The sensitivity of the parameters on the mass of the top quark and the amount of final state radiation can be best judged by the relative size of the error bars on the various mass and final state radiation values: As the amount of events is the same for each point between the various a_i, a smaller shown error bar means a stronger dependence on that a_i than on another one with larger shown error bars. As expected, the parameters of the transverse decay length distribution depend on both, the mass of the top quark and the amount of final state radiation. The parameters for the transverse momentum of the lepton do not depend as strongly on the final state radiation strength as on the mass of the top quark. This allows a proper extraction of both of the observables.

Another thing to note is that for parameters 1 and 2 of the transverse decay length, the slope of the calibration is positive for both, the mass of the top quark and the amount of final state radiation. For parameter 0 this is not the case. This means that by adding the information of the transverse decay length to the fit the accuracy on the mass of the top quark is increased compared by using only the transverse momentum of the lepton.

7.2.1 Extraction of the Mass of the Top Quark

The following procedure has been applied to extract the mass of the top quark: For an assumed pair of $\left(m'_{\text{Top}}, S'_F\right)$ the values of the top mass calibration curves $M\left(m_{\text{Top}}\right)$ have been used to get an initial set of template parameter values

$$X_i = M\left(m'_{\text{Top}}\right), \tag{7.8}$$

where X_i is the i-th template parameter. The calibration for the assumed value of the final state radiation is then added, assuming that the S_F calibration and the top quark mass calibration are independent of each other. By construction $S_F = 0$ denotes the nominal

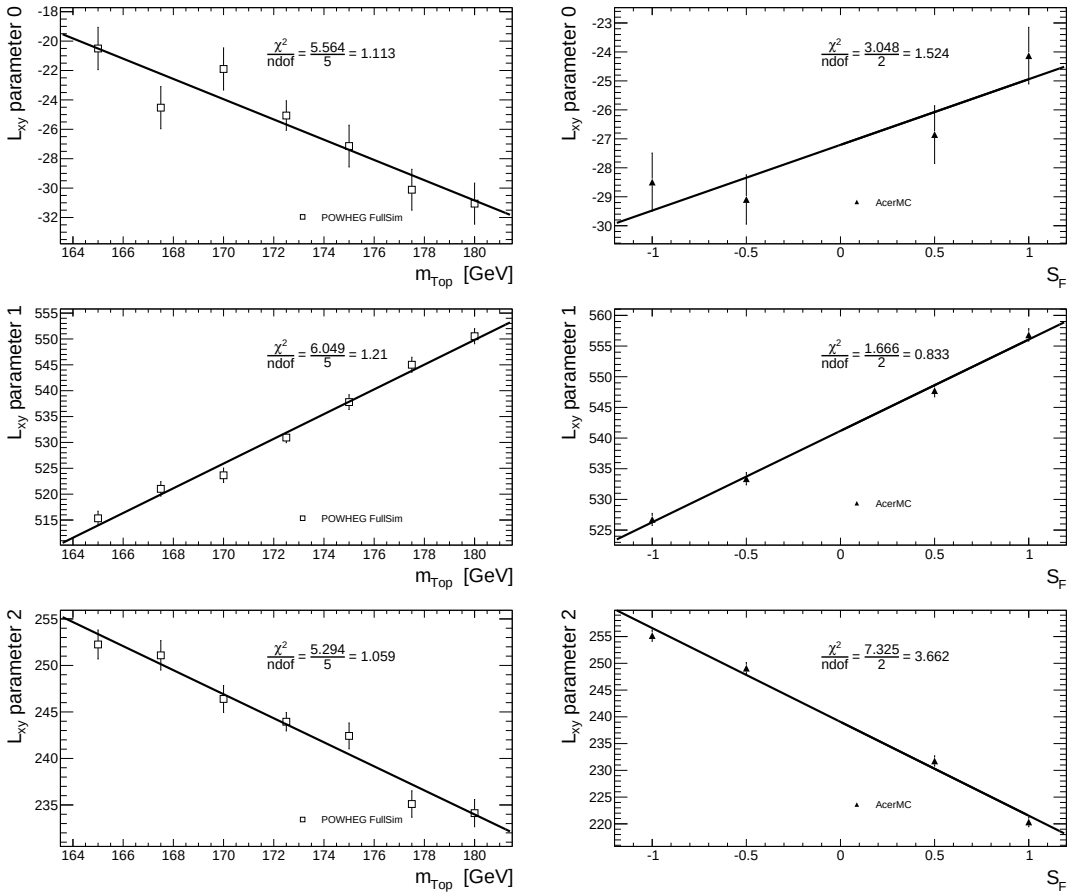

Figure 7.6: The calibration fits for the various parameters for the L_{xy}-template. Left: against the mass of the top quark, Right: against the amount of final state radiation. The parameters have no dimension by construction. They are applied to the parametrization of Equation (6.6) after applying the transformation matrix in Equation (6.9).

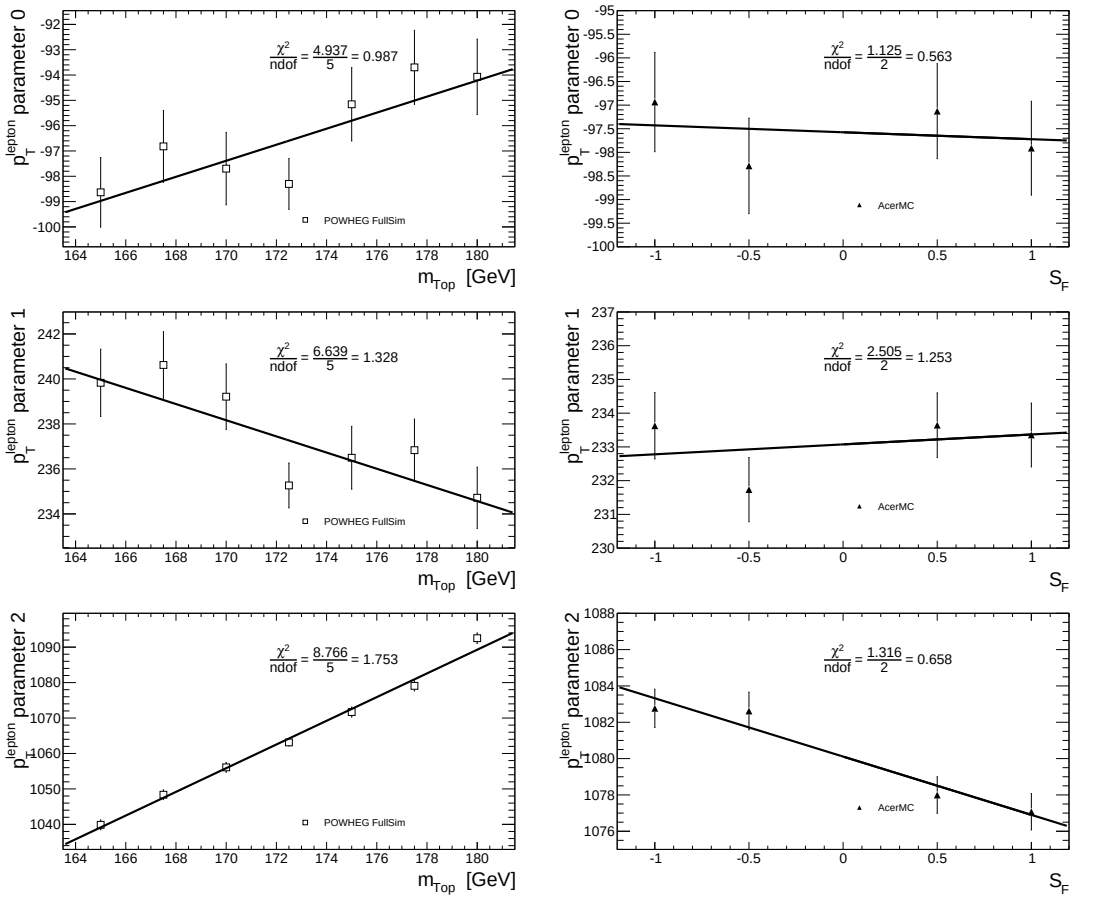

Figure 7.7: The calibration fits for the various parameters for the $p_{\mathrm{T}}^{\mathrm{Lep}}$-template. Left: against the mass of the top quark, Right: against the amount of final state radiation. The parameters have no dimension by construction. They are applied to the parametrization of Equation (6.6) after applying the transformation matrix in Equation (6.9).

amount of final state radiation. Thus, an offset for a given $S_F^{'}$ value can be calculated in the following way using the FSR strength calibration:

$$\Delta X_i = F\left(S_F^{'}\right) - F(0) \tag{7.9}$$

Combining this information yields the parameter values for a pair of $\left(m_{\mathrm{Top}}^{'}, S_F^{'}\right)$:

$$X_i^{\mathrm{final}}\left(m_{\mathrm{Top}}^{'}, S_F^{'}\right) = X_i + \Delta X_i = M\left(m_{\mathrm{Top}}^{'}\right) + F\left(S_F^{'}\right) - F(0) \tag{7.10}$$

These values are then used to construct the templates for the lepton p_T and decay length distributions and a binned likelihood is performed to find the values for (m_{Top}, S_F) which maximises this likelihood with respect to the data.

7.2.2 Constructed Templates

Figures 7.8 and 7.9 demonstrate how the templates change when varying the input mass of the top quark and amount of final state radiation. The template for the transverse decay length – as expected – changes with both input values. The parametrization of the transverse momentum of the lepton does also show the expected behaviour. It depends strongly on the mass of the top quark but is almost independent on the value of the final state radiation parametrization.

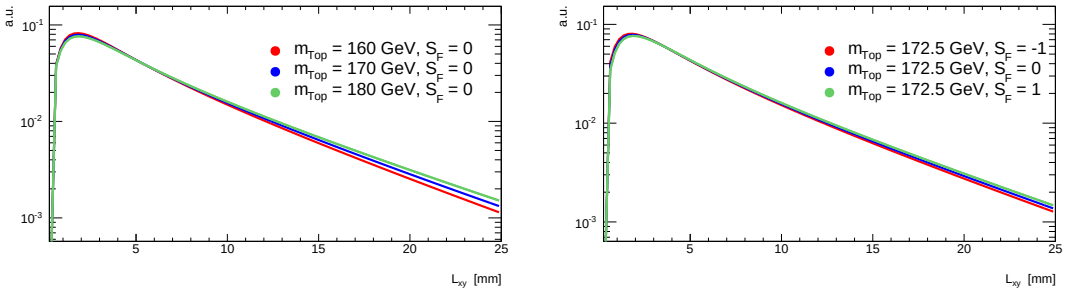

Figure 7.8: Change of the template parametrization of the transverse decay length distribution with changing mass of the top quark and constant final state radiation strength (left) and vice-versa (right). The templates have been normalized to unity.

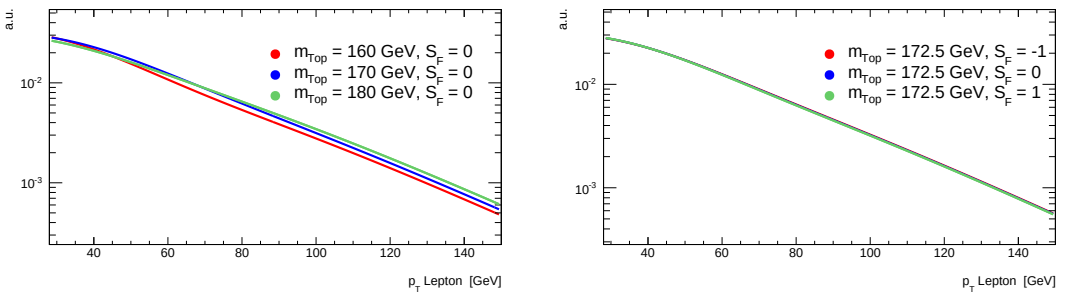

Figure 7.9: Change of the template parametrization of the transverse momentum of the lepton distribution with changing mass of the top quark and constant final state radiation strength (left) and vice-versa (right). The templates have been normalized to unity.

7.3 Expected Sensitivity and Fit Bias

The correctness of the calibration and the technique itself must be verified. This is done by doing pseudoexperiments. From a given transverse decay length and lepton transverse momentum distribution N random events are drawn each, where N is a random number from a Poisson distribution with mean value μ:

$$P(N) = \frac{\mu^N \exp(-\mu)}{N!}. \tag{7.11}$$

μ is set to the expected number of events in data. This lets the number of events in the distribution fluctuate around the expected value μ to mimic the effect the statistical fluctuation of the data events. After the proper number of events are drawn, both distributions are fit with the technique described in the previous sections. This fit yields values for the mass of the top quark and the strength of final state radiation as well as their errors. This procedure is repeated 2000 times.

The obtained values for the mass of the top quark and the amount of final state radiation can than be histogrammed. The results are shown in Figure 7.10 for an assumed mass of the top quark of 175.0 GeV. The assumed strength of final state radiation is 0. Another result of the pseudoexperiments is the expected statistical uncertainty. This is shown in Figure 7.11 for the same set of pseudoexperiments as the ones in Figure 7.10. From the distribution of these values several quantities can be calculated which give information about the fit procedures correctness and power.

One important fact is that, in total, more events are generated to be used in the pseudoexperiments than there are in the generation of the underlying distribution. This leads to the fact that a single event of the underlying distribution will be used in more than one pseudoexperiment. Because of this, the uncertainty on the various values which will be used in the following are not simply the mean error of their respective distributions. Instead the root mean square, modified by the Barlow correction factor, R_B, needs to be used[88]. For pseudoevents drawn from the same distribution, where an event may be used more than once in the generation it reads:

$$R_B = \sqrt{\frac{1}{N_{\text{pseudo}}} + \frac{n}{m}}. \tag{7.12}$$

In this equation N_{pseudo} denotes the number of pseudoexperiments generated, n the average number of generated events in each pseudoexperiment and m the total number of events used in the underlying distribution. Clearly, for a small number of pseudoexperiments this number depends on N_{pseudo}, but for a large number if converges against $\frac{n}{m}$ as no more information can be gained from the original distribution. For the following series of plots N_{pseudo} is 2000, $n = \mu = 25000$ (for an assumed mass of the top quark of 175 GeV) and $m = 320265$, R_b is 0.28.

7.3.1 Bias of the Method

A bias in the method means that the fit results would differ from the true result in any way. For instance, the result could be shifted by a constant value into any direction. To identify these values, the difference between the assumed and the fitted value for each quantity, for the mass of the top quark:

$$\Delta m_{\text{Top}} = m_{\text{Top}}^{\text{assumed}} - m_{\text{Top}}^{i} \tag{7.13}$$

$m_{\text{Top}}^{\text{assumed}}$ is the assumed mass in the sample which has been used for the generation of the pseudodata, and m_{Top}^{i} is the result of pseudoexperiment i. This is done for a various

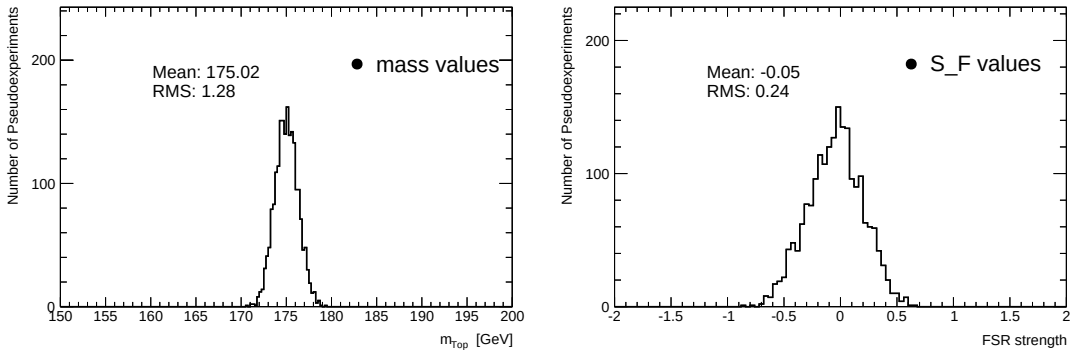

Figure 7.10: Distribution of fitted top mass values (left) and the FSR strength values (right) in 2000 pseudoexperiments. The basis for the pseudodata has been the sample with $m_{\mathrm{Top}} = 175.0\,\mathrm{GeV}$ and $S_F = 0$.

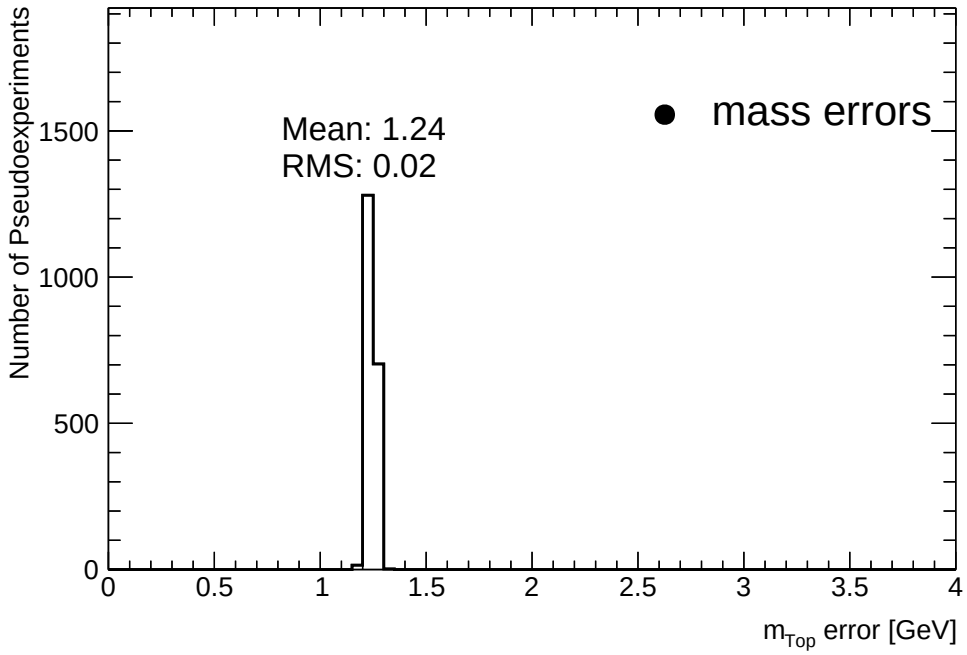

Figure 7.11: Distribution of fitted top mass errors in 2000 pseudoexperiments. The basis for the pseudodata has been the sample with $m_{\mathrm{Top}} = 175.0\,\mathrm{GeV}$ and $S_F = 0$.

number of assumed values. The result of these pseudoexperiment is put into a graph against the assumed value of the parameter. If the fit is unbiased, the values will fluctuate around 0. If a bias would be present, either a constant offset or some other systematic trend would be visible in the figure.

For this analysis the result of this validity check are shown in Figure 7.12. Both, the difference in the mass of the top quark and the strength of final state radiation fluctuate around 0. Also, the graph has been fit with a constant function. The black line is the result and the red shaded area the uncertainty. The exact values are:

$$\Delta m_{\mathrm{Top}} = 0.01 \pm 0.12\,\mathrm{GeV} \tag{7.14}$$

$$\Delta S_F = 0.00 \pm 0.02 \tag{7.15}$$

The fit is therefore free of any bias.

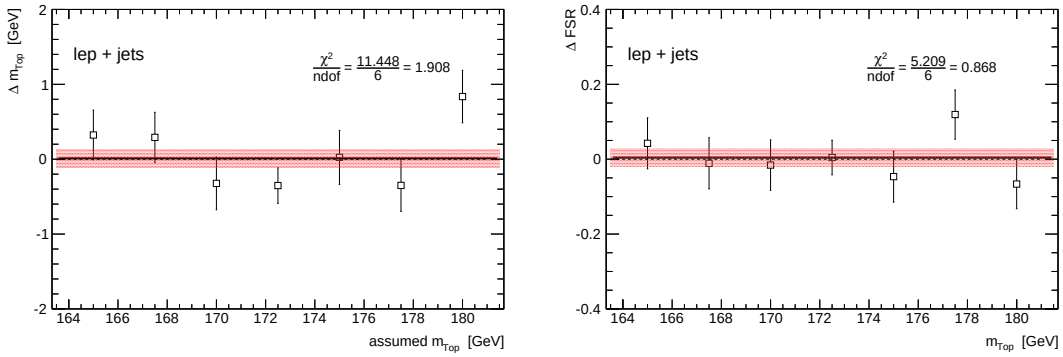

Figure 7.12: Bias of the technique for estimating the top mass (left) and the FSR strength (right) as a function of the assumed mass of the top quark.

7.3.2 Expected Statistical Uncertainty

The fit also gives information about the expected error on the mass of the top quark. As the fit is constructed in the way that the number of events fluctuates around the number of expected data events, the distribution of the error on the mass of the top quark should, on average, be the expected error on the respective observable. The spread of this distribution gives the precision on this value. Also, as the cross-section of the $t\bar{t}$ -process depends on the mass on the top quark, the statistical error will change with changing mass of the top quark. This effect is demonstrated in Figure 7.13. The expected uncertainty on the mass of the top quark slightly increases with the assumed mass of the top quark. The expected error at a mass of the top quark of 175 GeV is:

$$m_{\mathrm{Top}}^{\mathrm{error}} = 1.2\,\mathrm{GeV} \tag{7.16}$$

7.3.3 Pull Width

Another important quantity is the distribution of pulls. The pull is a quantity which is the bias of an individual pseudoexperiment divided by the error of that pseudoexperiment:

$$P_i = \frac{m_{\mathrm{Top}}^{\mathrm{assumed}} - m_{\mathrm{Top}}^{i}}{m_{\mathrm{Top}}^{\mathrm{error},i}} \tag{7.17}$$

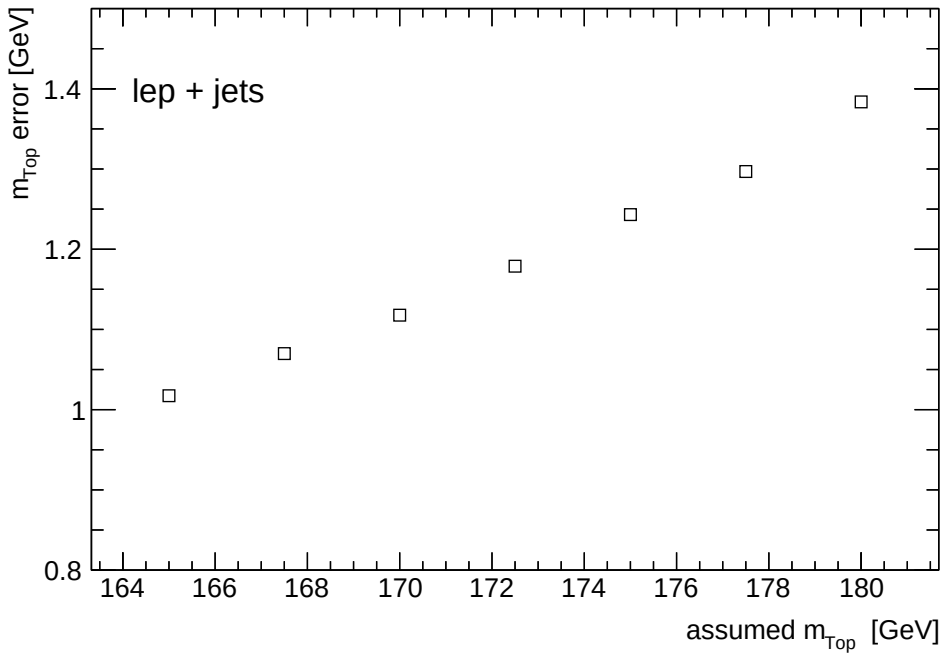

Figure 7.13: Expected error of the technique for estimating the top mass as a function of the assumed mass of the top quark.

As with the bias distribution if the fit is unbiased the mean value of this distribution should be 0. The pull distributions for same set of pseudoexperiments as in Figure 7.10 is shown in Figure 7.14. No new information is gained by the mean value of these distributions. The important quantity of this distribution is the RMS of it. If the error on the fit result has been estimated correctly the RMS of this distribution should be equal to 1. This must also be true independently of the assumed mass of the top quark. This is shown for various masses of the top quark in Figure 7.15. A fit with a straight line yields:

$$P_{m_{\mathrm{Top}}} = 0.99 \pm 0.01 \tag{7.18}$$

The error on the fit is therefore correctly estimated.

7.4 Systematic Uncertainties

There are several effects which may influence the measurement. They can be roughly separated into two categories:

- Reconstruction related uncertainties: These are uncertainties on the calibration or reconstruction efficiencies of the physics objects used in the analysis
- Modelling related uncertainties: The simulation has various assumptions which are motivated from other measurements. This leaves some uncertainty on the exact values of the simulation parameters.

The possible effects on the distributions needs to be evaluated and possible shifts taken into account when calculating a final uncertainty.

The procedure to evaluate the effects of these systematic uncertainties in this analysis is the following. Say, $m_{\mathrm{Top}}^{\mathrm{default}}$ is the result of 2000 pseudo-experiments using the nominal

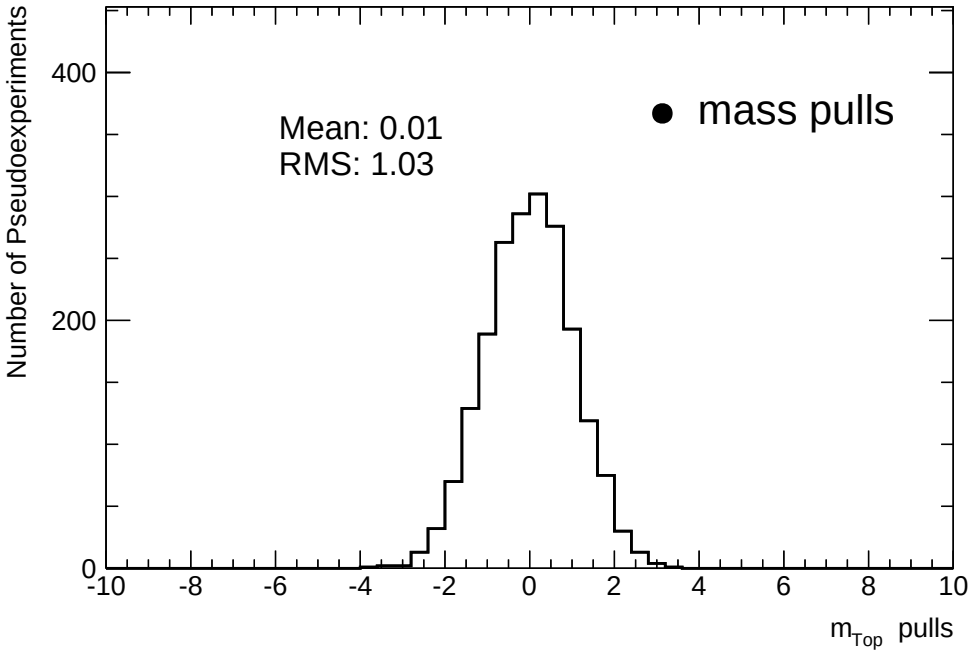

Figure 7.14: Distribution of fitted top mass pull values in 2000 pseudoexperiments. The basis for the pseudodata has been the sample with $m_{\mathrm{Top}} = 175.0\,\mathrm{GeV}$ and $S_F = 0$.

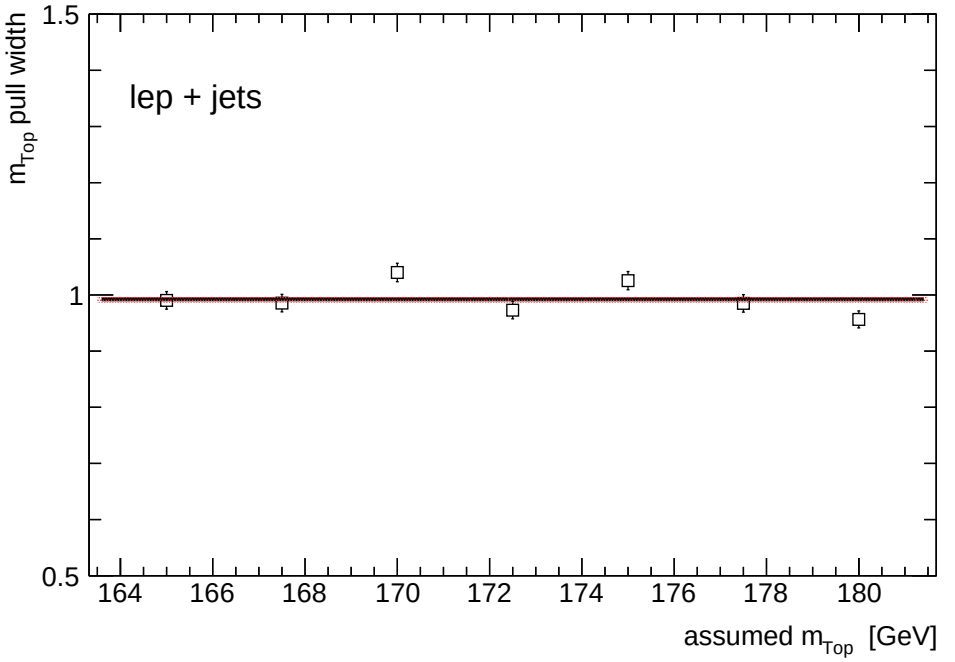

Figure 7.15: Pull width of the technique to estimate the top mass as a function of the assumed mass of the top quark.

sample. A systematic uncertainty would then be evaluated by changing one of the parameters in question and then using that modified sample as the basis for another 2000 pseudo-experiments, yielding a $m_{\text{Top}}^{\text{shifted}}$. The systematic uncertainty for that parameter would then be:

$$\Delta^{\text{systematic}} m_{\text{Top}} = |m_{\text{Top}}^{\text{default}} - m_{\text{Top}}^{\text{shifted}}| \tag{7.19}$$

Some systematic shifts provide an upward and downward variation. In the presence of this, individually 2000 pseudoexperiments are done for the upward and downward variation, resulting in a $m_{\text{Top}}^{\text{upward}}$ and $m_{\text{Top}}^{\text{downward}}$. The uncertainty in that case is evaluated in this way:

$$\Delta^{\text{systematic}} m_{\text{Top}} = \frac{|m_{\text{Top}}^{\text{upward}} - m_{\text{Top}}^{\text{downward}}|}{2} \tag{7.20}$$

For all systematic variations their effect on the estimated strength of final state radiation is also estimated by the fit. A systematic variation may impact S_F in a different way as it impacts the mass of the top quark. This typically indicates different sensitivities of the transverse momentum of the lepton and the transverse decay length to that systematic variation. Thus, for all systematic variations its change to S_F is also quoted. The above considerations (Equations (7.19) and (7.20)) are analogously applied for the resulting strength of final state radiation.

7.4.1 Reconstruction Related Uncertainties

The calibration of all physics objects as well as their reconstruction efficiency measurements are all measurements themselves and have associated uncertainties. When performing the default analysis the default results of these measurements, which have been explained in more detail in chapter 3, have been applied. The uncertainties on these measurements need to be propagated to the signal distributions.

Jet Energy Scale

The effect of the jet energy scale on this measurement has already been explained in detail in section 6.1.3. The jet energy scale has various components of uncertainties, some of which are displayed in Figure 3.16. Within the ATLAS experiment, two possible treatments of the jet energy scale uncertainty are considered:

- For measurements where the jet energy scale is not a dominant systematic uncertainty, the quadratic sum of all components can be used for a single upwards and a single downwards shift.
- If the jet energy scale is a dominant systematic uncertainty it is recommended to vary groups of the individual components and then take the quadratic sum of the individual uncertainties.

Even though the jet energy scale is not a dominant uncertainty on this analysis by construction, in order to be compatible with other ATLAS measurements of the mass of the top quark, the latter approach has been used. In detail, the various components have been split into the following groups:

- Statistical uncertainty.
- Detector knowledge uncertainty.
- Modelling related uncertainties.
- The b-jet energy scale uncertainty.

- Pileup related uncertainties.
- Uncertainties on the intercalibration of various regions of pseudorapidity.
- A component of a high-transverse momentum single particle forming a jet .
- The effects of close-by jets.
- Uncertainties on the flavor composition and response.
- Other (called "Mixed" uncertainties).

In total there are 21 components which need to be taken into account. All of the variations have an upwards and a downwards variation, thus Equation (7.20) has been used for the calculation of the uncertainties. The resulting uncertainties are summarised in Table 7.5. As constructed, the jet energy scale uncertainty on this measurement is very small. In total the jet energy scale uncertainty is $\Delta^{\text{JES}} m_{\text{Top}} = 0.30\,\text{GeV}$ on the mass of the top quark and $\Delta^{\text{JES}} S_F = 0.09$ on the strength of the final state radiation. This compares very well with the measurements presented in chapter 5.

Table 7.5: Breakdown of the jet energy scale systematic uncertainties of the top quark mass and FSR strength determination.

	Combined		Electron channel		Muon channel	
Systematic	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F
B-Jet Energy Scale	0.04	0.05	0.01	0.06	0.06	0.05
Detector Component 1	0.08	0.03	0.06	0.02	0.10	0.03
Detector Component 2	0.01	0.01	0.00	0.01	0.02	0.01
Mixed Component 1	0.02	0.01	0.01	0.01	0.03	0.01
Mixed Component 2	0.04	0.01	0.07	0.01	0.02	0.01
Modelling Component 1	0.11	0.05	0.13	0.05	0.10	0.04
Modelling Component 2	0.07	0.00	0.08	0.00	0.06	0.00
Modelling Component 3	0.06	0.00	0.06	0.00	0.05	0.00
Modelling Component 4	0.01	0.00	0.01	0.00	0.02	0.01
Statistical Component 1	0.10	0.02	0.14	0.02	0.07	0.02
Statistical Component 2	0.02	0.00	0.02	0.00	0.02	0.00
Statistical Component 3	0.03	0.01	0.02	0.01	0.04	0.01
η Intercalibration – Theory	0.09	0.05	0.06	0.05	0.11	0.04
η Intercalibration – Statistical	0.04	0.01	0.06	0.01	0.02	0.01
Pileup Offset μ Component	0.01	0.01	0.01	0.00	0.00	0.02
Pileup Offset N_{PV} Component	0.01	0.00	0.01	0.00	0.02	0.01
Relative Non-Closure Component	0.00	0.00	0.00	0.00	0.00	0.00
High-p_T Single Particle Component	0.00	0.00	0.00	0.00	0.00	0.00
Close-by Jets Component	0.19	0.00	0.21	0.00	0.20	0.01
Flavor Composition Component	0.02	0.00	0.03	0.00	0.01	0.01
Flavor Response Component	0.04	0.00	0.01	0.00	0.07	0.01
Total	0.30	0.09	0.33	0.10	0.31	0.09

Jet Energy Resolution and Missing Energy

In section 6.1.3 it has been claimed that the impact of the jet energy resolution on the transverse momentum of the lepton stems mostly from its impact on the missing energy. This is shown in Figure 7.16. The figure shows the impact of the jet energy resolution on the missing energy, together with the impact of one jet energy scale component. The jet energy resolutions impact is much stronger than the impact of the jet energy scale component. This impacts the lepton transverse momentum as shown in Figure 7.17. When

changing the jet energy resolution the lepton transverse momentum is greatly changed when cutting on the missing transverse energy. When omitting the cut the uncertainty becomes negligible. Not cutting on the missing energy is not feasible though, as it limits the impact of the QCD fakes background. Additionally the missing energy calibration has two uncertainty components. Together with the impact of the jet energy resolution they are displayed in Table 7.6. The effect of the jet energy resolution dominates these uncertainties, with an uncertainty of $\Delta^{\mathrm{JER}} m_{\mathrm{Top}} = 0.71\,\mathrm{GeV}$ on the mass of the top quark and $\Delta^{\mathrm{JER}} S_F = 0.03$ on the strength of final state radiation. It is one of the largest uncertainties on the measurement.

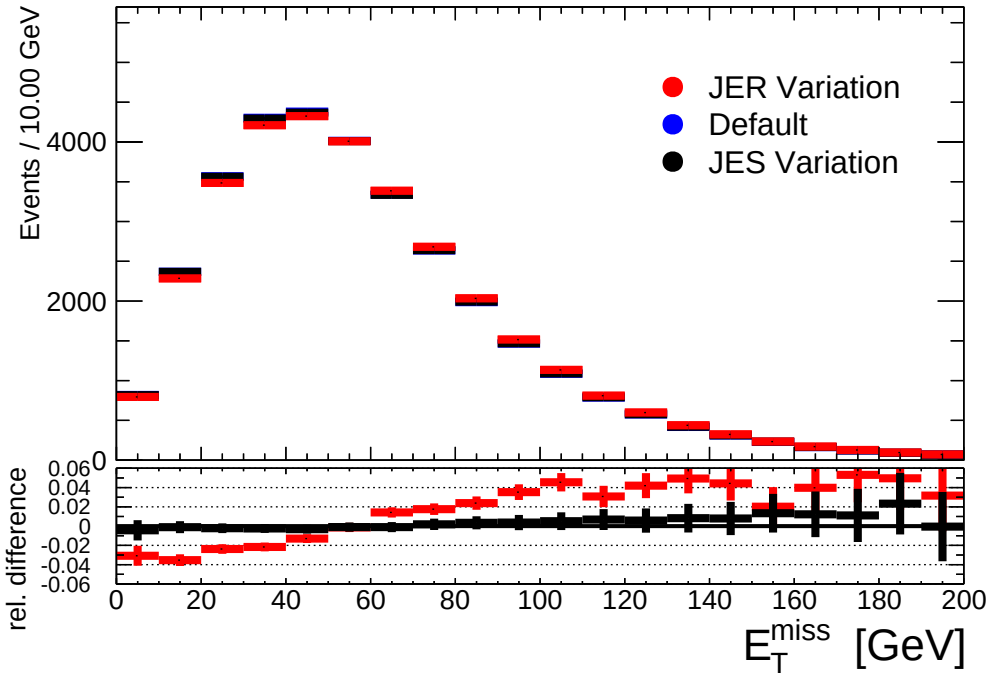

Figure 7.16: Distribution of the missing energy for the default selection ("Default"), the jet energy resolution systematic ("JER") and one direction of the flavor composition component of the jet energy scale uncertainty ("JES").

Table 7.6: Breakdown of the missing energy related systematic uncertainties of the top quark mass and FSR strength determination.

	Combined		Electron channel		Muon channel	
Systematic	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F
MET Cell Term	0.11	0.01	0.08	0.02	0.12	0.00
MET Pileup Term	0.05	0.00	0.02	0.01	0.07	0.00
Jet Energy Resolution	0.71	0.03	0.65	0.03	0.70	0.03
Total	0.72	0.03	0.66	0.04	0.72	0.03

Flavour Tagging Uncertainties

The calibration of the flavour tagging algorithm has been extensively described in section 3.3.7. As with the jet energy scale uncertainties they have been split up into various components. Other than with the jet energy scale uncertainty their uncertainty has not been grouped into categories. Instead, the full covariance matrix between the various bins

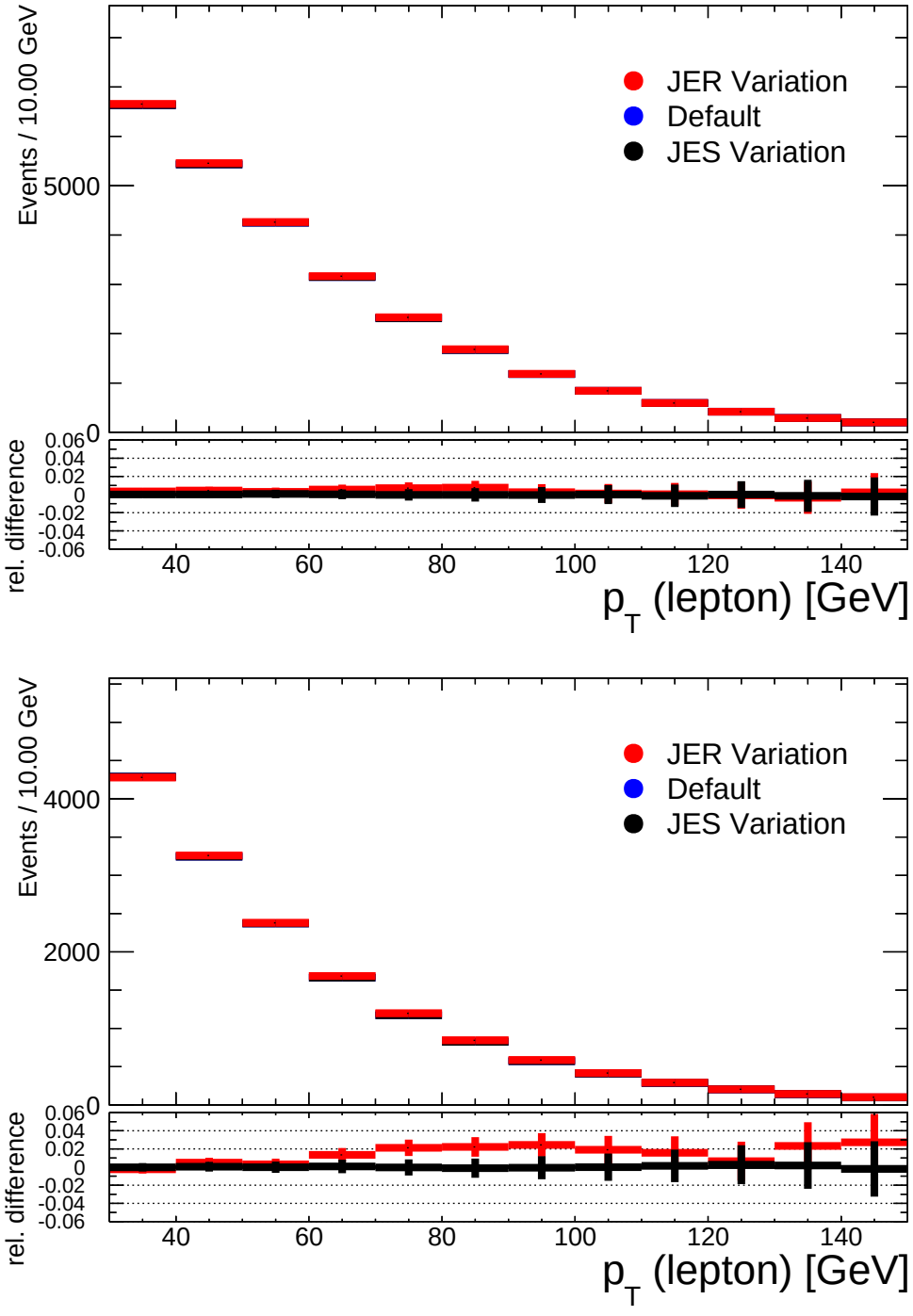

Figure 7.17: Distribution of the lepton transverse momentum for the default selection variation ("Default"), the jet energy resolution systematic ("JER") and one direction of the flavor composition component of the jet energy scale uncertainty ("JES"). Top: Without cutting of the value of the missing energy, Bottom : With the default cut on the missing energy.

of jet transverse momentum has been built and then decomposed into eigenvectors. As it is a 10×10 matrix there are 10 eigendirections. For each eigendirection there is a component of uncertainties to be evaluated, each with an upward and a downward fluctuation. The resulting uncertainties are summarised in Table 7.7. In total the uncertainty on the mass of the top quark is $\Delta^{\text{flavour tagging}} m_{\text{Top}} = 0.21\,\text{GeV}$ and on the FSR strength $\Delta^{\text{flavour tagging}} S_F = 0.07$. Additionally also the measurement of the charm-jet and light-jet fake tagging rates uncertainties have been propagated to the final result. They yield in a small uncertainty.

Table 7.7: Breakdown of the b-tagging efficiency systematic uncertainties of the top quark mass and FSR strength determination.

	Combined		Electron channel		Muon channel	
Systematic	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F
B-Tagging Efficiency 1	0.02	0.01	0.03	0.01	0.00	0.01
B-Tagging Efficiency 2	0.02	0.00	0.05	0.00	0.01	0.00
B-Tagging Efficiency 3	0.01	0.01	0.00	0.00	0.02	0.00
B-Tagging Efficiency 4	0.02	0.00	0.04	0.01	0.03	0.01
B-Tagging Efficiency 5	0.02	0.02	0.05	0.02	0.04	0.03
B-Tagging Efficiency 6	0.09	0.01	0.10	0.01	0.06	0.02
B-Tagging Efficiency 7	0.06	0.00	0.10	0.01	0.06	0.01
B-Tagging Efficiency 8	0.06	0.04	0.09	0.03	0.05	0.04
B-Tagging Efficiency 9	0.04	0.02	0.02	0.01	0.03	0.01
B-Tagging Efficiency 10	0.16	0.05	0.16	0.05	0.15	0.06
Total	0.21	0.07	0.21	0.07	0.19	0.08

Lepton Energy Scales and Reconstruction Efficiencies

The uncertainties on the reconstruction efficiency and energy resolution of leptons are not expected to have a strong impact on the measurement as they are known to a very precise level. However, they directly impact the lepton transverse momentum and therefore impact the measurement. The reconstruction efficiencies for the leptons have been varied within their uncertainties and as expected have a small impact on the measurement. The exact results are shown in Table 7.8. The resulting uncertainty of $\Delta^{\text{lepton}} m_{\text{Top}} = 0.36\,\text{GeV}$ and $\Delta^{\text{lepton}} S_F = 0.04$ is about the same size as the uncertainty of the jet energy scale, demonstrating how the quantities of leptons can be measured with much higher precision than the quantities of jets.

Table 7.8: Breakdown of the lepton related systematic uncertainties of the top quark mass and FSR strength determination.

	Combined		Electron channel		Muon channel	
Systematic	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F
Electron Energy Resolution	0.00	0.00	0.00	0.01	0.00	0.00
Electron Energy Scale	0.31	0.03	0.67	0.07	0.00	0.00
Lepton ID Efficiency	0.08	0.01	0.08	0.01	0.00	0.00
Lepton Reconstruction Efficiency	0.02	0.00	0.01	0.00	0.01	0.00
Lepton Trigger Efficiency	0.03	0.00	0.01	0.00	0.02	0.00
Muon Energy Resolution	0.02	0.00	0.00	0.00	0.02	0.01
Muon Energy Scale	0.14	0.01	0.00	0.00	0.27	0.02
Total	0.36	0.04	0.67	0.07	0.27	0.03

Other Reconstruction Related Uncertainties

Two additional uncertainties related to the reconstruction have been considered. One is the reconstruction efficiency for jets. It has been published by the collaboration [89]. The systematic is applied by randomly removing jets from event according to the jet reconstruction inefficiency, which is about 1%. The other uncertainty is the cut efficiency on the jet vertex fraction cut. Scale factors have been measured and applied and their uncertainty is evaluated. Both effects lead to very small systematic uncertainties ($\Delta m_{\text{Top}} < 50\,\text{MeV}/\Delta S_F < 0.01$).

7.4.2 Simulation Related Uncertainties

There are several sources of simulation related uncertainties. Most of them originate from the fact, that the parameters of the simulation have been tuned to match various measurements from either ATLAS or other experiments. Naturally, these measurements also have uncertainties. The parameters of the simulation can only be tuned to a best-fit central value and a band of uncertainty around it. If a measurement depends on the effects of the change of parameters, these uncertainties need to be considered for a proper estimation of its own uncertainties. A description of the considered simulation related uncertainties will be given in the following.

Monte-Carlo Generator Uncertainties

The default Monte-Carlo generator for this analysis is Powheg [61]. Other Monte-Carlo generators, like Mc@nlo [60] or Alpgen [59] describe the data equally well within the uncertainties of the data. As such, no generator can be strictly preferred over the other. Thus, if the different generators predict different shapes for either observable, none of them can be deemed the "correct" one.

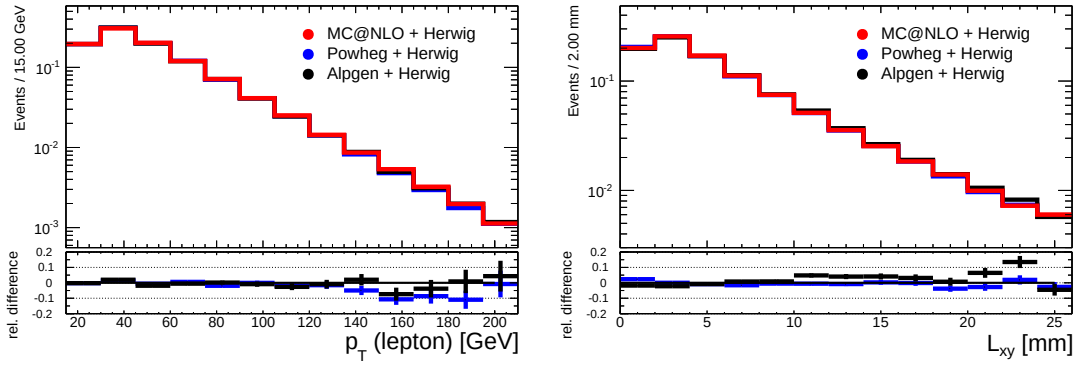

Figure 7.18: The transverse momentum of the lepton (left) and the transverse decay length (right) for the Powheg, Mc@nlo and Alpgen Monte-Carlo generators. The selection is a loose semileptonic $t\bar{t}$ selection to increase the statistical accuracy.

Figure 7.18 shows the transverse momentum of the lepton and the transverse decay length for the three mentioned Monte-Carlo generators. For better comparison, all of them have been showered with Herwig and simulated using fast simulation. The event selection requirements regarding the missing transverse energy and the transverse mass of the W boson have been dropped for this figure to enhance the statistical accuracy. The three Monte-Carlo generators indeed predict different shapes for the observables. This can be traced back to the prediction of different momenta of the top quark. As mentioned in chapter 6, the transverse momentum of the top quark also impacts both observables. This is also shown in Figure 7.19. It shows the truth spectrum of the transverse momentum of

the top quark and the truth momentum of the lepton for all three generators. The same trend as in Figure 7.18 can be seen. Also, the transverse momentum of the top quark shows the same effects. This explains the effects seen on the shapes. The full difference of the predictions from the various generators is taken as a systematic uncertainty.

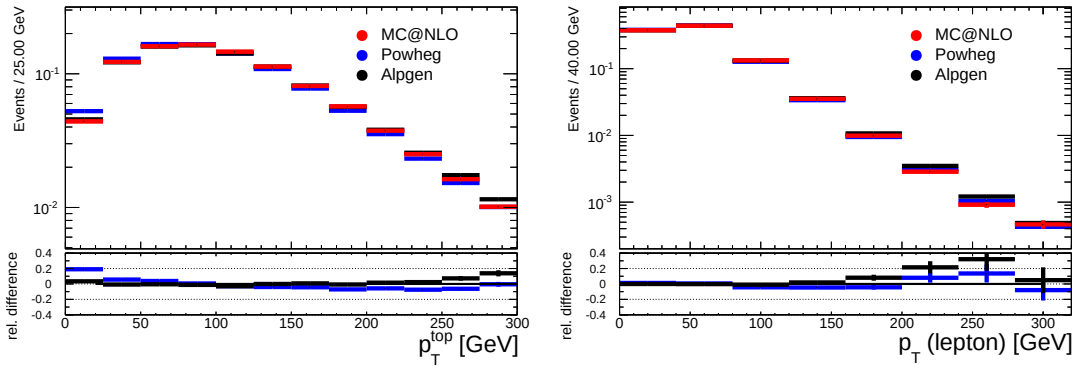

Figure 7.19: The truth transverse momentum of the top quark (left) and the lepton (right) for the Powheg, Mc@nlo and Alpgen Monte-Carlo generators. The selection is a very loose semileptonic $t\bar{t}$ selection to increase the statistical accuracy.

Hadronization

Similarly as there is no Monte-Carlo generator to prefer over the other, none of the two considered hadronization models, Pythia and Herwig can be a priori preferred over the other. The impact of the hadronization model on the transverse decay length is easy to understand. Both hadronization models implement different variants of fragmentation models. This governs the momentum fraction the bottom hadron receives from the original bottom quark. This translates into a different transverse decay length when using either fragmentation model.

The effect of the hadronization on the transverse momentum of the lepton is harder to understand. In first order the hadronization only impacts the momenta of quark and gluon jets, not the properties of leptons. The explanation is again the effect the hadronization has on the transverse momentum of the top quark. The changes the parton shower does to the momenta of the initial state objects are propagated to the momentum of the top quark. Thus, different parton shower models may predict different transverse momenta of the top quarks and thus of the transverse momentum of the lepton in the event. Figure 7.20 shows that this is the case for this analysis and impacts the analysis. The effect of this can be lowered by cutting on the minimum distance between the jets in the event. This removes events with a large amount of initial state radiation thus limiting the effects.

Parton Distribution Functions

As mentioned in chapter 4, the Parton Distribution Functions are a vital ingredient for the simulation. Typically they are a global fit to a large amount of measurements by various experiments. Several groups have performed global fits to the data and the resulting parameters have been extrapolated to the LHC energy. For this measurement several PDF sets have been considered. Besides the default Ct10 [56], the central values of the Mstw [90], Nnpdf [91] and Herapdf [92] groups have been calculated. All of these groups also give a list of variations which represent the uncertainties. These have also been varied. The maximum difference is the one of the central values between Nnpdf and Ct10 with $\Delta^{PDF} m_{\mathrm{Top}} = 0.58\,\mathrm{GeV}$ and $\Delta S_F = 0.04$.

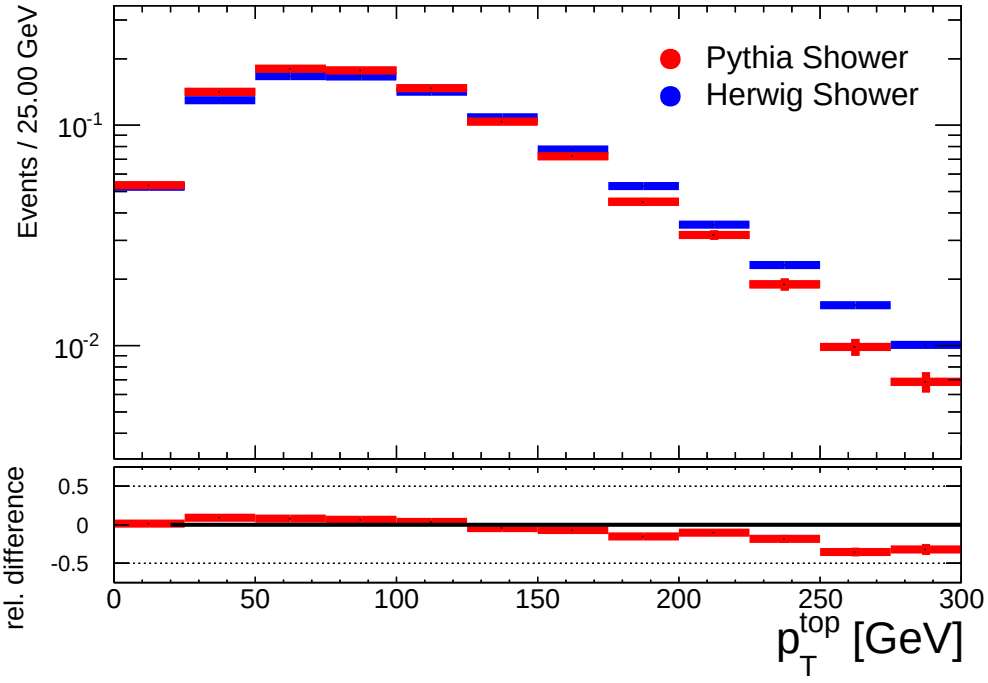

Figure 7.20: The truth transverse momentum of the top quark for the HERWIG and PYTHIA parton shower simulations. The selection is a very loose semileptonic $t\bar{t}$ selection to increase the statistical accuracy.

Other Modelling Uncertainties

There are other sources of modelling related uncertainties: The renormalization and factorisation scales have been varied within their uncertainties. A small effect on this measurement was found. Also the effects of colour reconnection, another ad-hoc source of modelling which had to be introduced into the simulation to properly model the data has measured parameters with uncertainties. Finally the initial and final state radiation has been varied. As these are the same variations which have been done when calibrating the templates, the effect should be very close to zero. The result is that this is the case. Also, the effect on the FSR strength should be equal to one as this is the way the quantity has been defined.

All resulting uncertainties from the modelling have been summarised in Table 7.9. Leading effects are the uncertainties from hadronization, the Monte-Carlo generator and the Parton Distributon Functions.

Table 7.9: Breakdown of the modelling related systematic uncertainties of the top quark mass and FSR strength determination.

	Combined		Electron channel		Muon channel	
Systematic	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F
Colour Reconnection	0.07	0.04	0.12	0.02	0.31	0.07
Hadronization	1.14	0.36	1.23	0.42	1.00	0.30
MC Generator	1.62	0.76	1.86	0.85	1.85	0.68
muR/muF variation	0.02	0.04	0.15	0.01	0.17	0.08
ISR/FSR	0.15	n/a	0.26	n/a	0.02	n/a
PDF, inter NNPDF/CT10	0.58	0.04	0.62	0.01	0.63	0.08

7.4.3 Background

Additionally to the so far mentioned uncertainties there are a couple of other systematic influences on the measurement. These are all related to the treatment of the background. The estimation of the background has uncertainties. They lead to a small uncertainty in case of the W+jets estimation and a medium uncertainty for the QCD estimation. Additionally there are shape uncertainties which lead to negligible effects. Finally a special treatment for the single top background is neede. regarding the single top background. The mass of the top quark does change the shape of the single top background. However, the current treatment of the background does not allow to take this into account. An additional systematic, has been introduced for this negligence. The analysis has been repeated by subtracting the difference in shape between the default single top background with $m_{\text{Top}} = 172.5\,\text{GeV}$ and two samples with $\pm 7.5\,\text{GeV}$. This gives an upward and a downward variation which has been treated in the usual way. Table 7.10 shows the exact values for these systematics.

Table 7.10: Breakdown of the background related systematic uncertainties of the top quark mass and FSR strength determination.

	Combined		Electron channel		Muon channel	
Systematic	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F
QCD Normalisation	0.27	0.00	0.51	0.01	0.16	0.00
Single Top Mass	0.33	0.05	0.28	0.03	0.39	0.05
W+HF Fractions	0.00	0.00	0.00	0.00	0.00	0.00
W+Jets Normalisation	0.01	0.10	0.32	0.12	0.04	0.10

7.5 Result

The observed distributions of the transverse decay length and the transverse momentum of the lepton together with the expected predicted simulation distributions is shown in Figure 7.21. Good agreement between data and simulation is observed, even though the assumed mass of the top quark in the $t\bar{t}$ contribution is 172.5 GeV and the assumed FSR strength is 0.

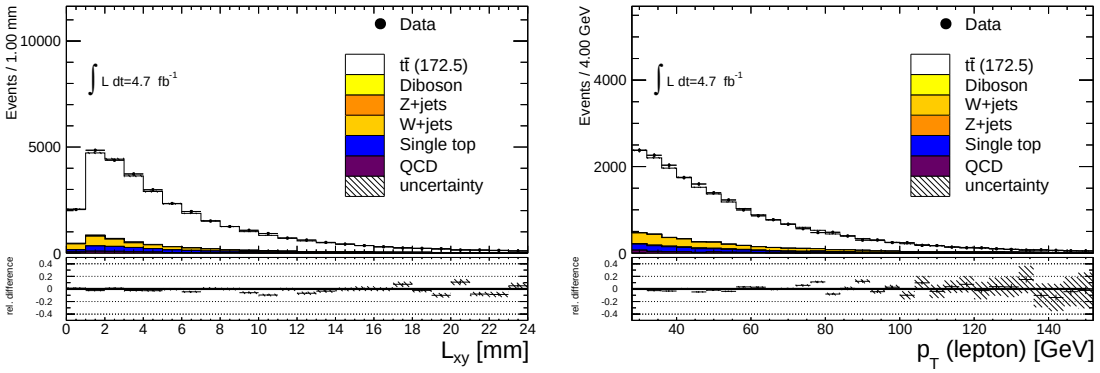

Figure 7.21: The observed data distributions for the transverse decay length (left) and the transverse momentum of the lepton (right). The filled areas are the various background estimations. Good agreement between the data and the simulation is observed. The assumed mass of the top quark in this figure is 172.5 GeV and the assumed FSR strength 0.

The template fitting procedure has been applied to the background-subtracted data. The distributions for both observables after subtracting the background is shown in Figure 7.22. This Figure also shows the templates with the best-fit values. The χ^2 per degree of freedom values are 1.031 for the transverse decay length and 0.991 for the transverse momentum of the lepton, indicating a good description of the data by the fit. The resulting mass of the top quark is:

$$m_{\text{Top}} = 170.4 \pm 1.1\,\text{GeV} \tag{7.21}$$

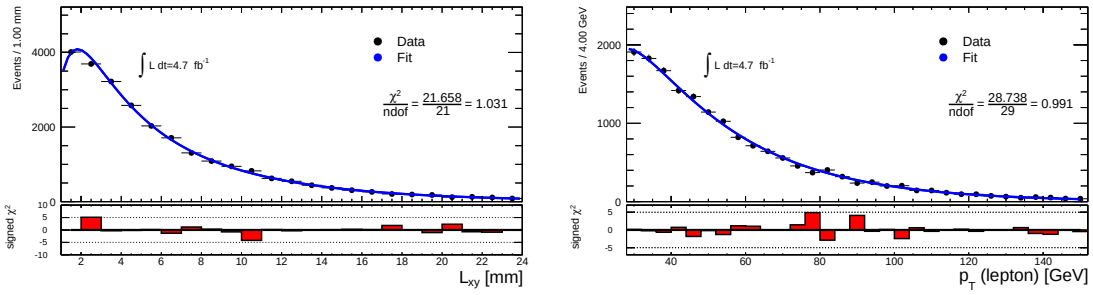

Figure 7.22: The observed background-subtracted data distributions for the transverse decay length (left) and the transverse momentum of the lepton (right). The black line is the parametrization with the best fitting values. The χ^2/dof for both distributions is close to 1.

Together with the systematic uncertainties, of which a total overview is shown in Table 7.11, this yields the following result:

$$m_{\text{Top}} = 170.4 \pm 1.1_{\text{stat.}} \pm 2.3_{\text{syst.}}\,\text{GeV} \tag{7.22}$$

The observed mass of the top quark is in very good agreement with the previous measurements described in 5. The measurement is dominated by systematic uncertainties, especially by the estimation of the Monte-Carlo generator uncertainty. The result for S_F is 0.42 ± 0.22, indicating more radiation in the data than in the default simulation.

Table 7.11: Systematic uncertainties of the top quark mass and FSR strength determination.

	Combined		Electron channel		Muon channel	
Systematic	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F	Δm_{Top}[GeV]	ΔS_F
B-Tagging Efficiency	0.21	0.07	0.28	0.07	0.19	0.08
C-Tagging Mistag Rate	0.00	0.00	0.00	0.00	0.00	0.00
Jet Vertex Fraction	0.00	0.00	0.00	0.00	0.00	0.00
Light-Tagging Mistag Rate	0.00	0.00	0.00	0.00	0.00	0.00
MET Cell Term	0.11	0.01	0.08	0.02	0.12	0.00
MET Pileup Term	0.05	0.00	0.02	0.01	0.07	0.00
QCD Normalisation	0.27	0.00	0.51	0.01	0.16	0.00
Single Top Mass	0.33	0.05	0.28	0.03	0.39	0.05
W+HF Fractions	0.00	0.00	0.00	0.00	0.00	0.00
W+Jets Norm	0.01	0.10	0.32	0.12	0.04	0.10
Jet Energy Resolution	0.71	0.03	0.65	0.03	0.70	0.03
Jet Reconstruction Efficiency	0.02	0.00	0.02	0.00	0.02	0.00
Jet Energy Scale	0.30	0.09	0.33	0.10	0.31	0.09
Electron Energy Resolution	0.00	0.00	0.00	0.01	0.00	0.00
Electron Energy Scale	0.31	0.03	0.67	0.07	0.00	0.00
Lepton ID Efficiency	0.08	0.01	0.08	0.01	0.00	0.00
Lepton Reconstruction Efficiency	0.02	0.00	0.01	0.00	0.01	0.00
Lepton Trigger Efficiency	0.03	0.00	0.01	0.00	0.02	0.00
Muon Energy Resolution	0.02	0.00	0.00	0.00	0.02	0.01
Muon Energy Scale	0.14	0.01	0.00	0.00	0.27	0.02
Colour Reconnection	0.07	0.04	0.12	0.02	0.31	0.07
Hadronization	1.14	0.36	1.23	0.42	1.00	0.30
MC Generator	1.62	0.76	1.86	0.85	1.85	0.68
muR/muF variation	0.02	0.04	0.15	0.01	0.17	0.08
ISR/FSR	0.15	n/a	0.26	n/a	0.02	n/a
PDF, inter NNPDF/CT10	0.58	0.04	0.62	0.01	0.63	0.08
Total	2.32	0.86	2.63	0.97	2.43	0.77

Chapter 8

Summary and Conclusion

A measurement of the mass of the top quark using the transverse momentum of the lepton and decay length of the B-Hadron has been presented. The result is

$$m_{\text{Top}} = (170.4 \pm 1.1_{\text{stat.}} \pm 2.3_{\text{syst.}})\ \text{GeV} \tag{8.1}$$

This is compatible with previous measurements of the mass of the top quark, done by either the ATLAS collaboration or other experiments. An overview of the result on the mass of the top quark including this result and various others is shown in Figure 8.1. The total uncertainty on the result of this analysis, $\Delta^{\text{total}} m_{\text{Top}} = 2.6\,\text{GeV}$ is larger than results by other measurements. However, with an jet energy scale uncertainty of only $\Delta^{\text{JES}} m_{\text{Top}} = 0.3\,\text{GeV}$ it has one of the smallest uncertainties caused by this source. In a combination of results this will help reducing the total uncertainty on the mass of the top quark. The value of 0.42 on the strength on final state radiation indicates that the simulation underestimates the strength of final state radiation. There is currently work ongoing aiming to publish the results found in this thesis in the context of an official ATLAS publication.

Additionally the uncertainties can be compared with those one would obtain by using only one of the two variables. If one considers only the transverse decay length, a statistical error of $\Delta m_{\text{Top}}^{\text{stat.}} = 1.7\,\text{GeV}$ and a systematic uncertainty of $\Delta m_{\text{Top}}^{\text{syst.}} = 7.8\,\text{GeV}$ is obtained, dominated by the uncertainty on initial and final state radiation. The statistical uncertainty obtained by using the transverse momentum of the lepton is with $\Delta m_{\text{Top}}^{\text{stat.}} = 1.4\,\text{GeV}$ a bit lower than the one obtained by the transverse decay length alone but still larger than the one of the presented measurement. The systematic uncertainty obtained is $\Delta m_{\text{Top}}^{\text{syst.}} = 2.7\,\text{GeV}$. Combining the two variables is therefore worthwhile compared with using only the transverse momentum of the lepton alone.

The dominant uncertainties on the measurement are caused by imperfect knowledge of the simulation parameters, especially the choice of Monte-Carlo generator. Other large components are the choice of a hadronization model. With the advent of more accurate Monte-Carlo simulation programs like aMc@nlo or Sherpa 2.0, and hadronization simulations like Herwig++ and Pythia 8 as well as more accurate tunes these uncertainties may decrease significantly in the future.

One large uncertainty on the measurement is also the statistical uncertainty. It can be reduced significantly be applying the technique to the $\sqrt{s} = 8\,\text{TeV}$ 2012 data set of ATLAS. It does not only consist out of more than four times as much integrated luminosity ($20.7\,\text{fb}^{-1}$) but the higher center-of-mass energy increases the cross-section from 177.3 pb to 252.9 pb. In total this gives a data set which is six times larger than the one used in this analysis. This will not only reduce the statistical uncertainty but will also allow further optimisation of the event and object selection for additional reduction of the systematic uncertainties.

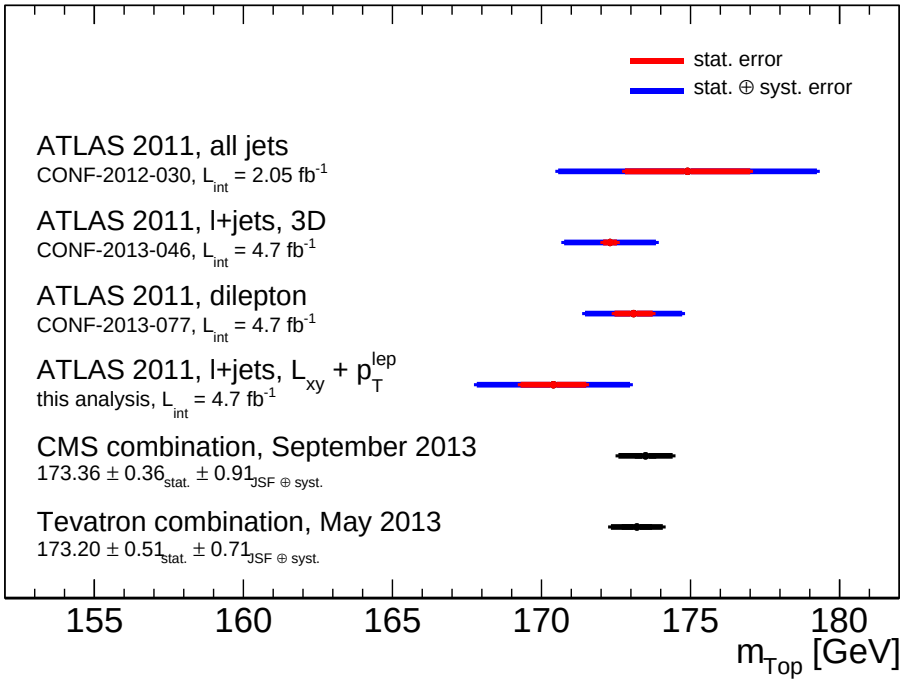

Figure 8.1: Comparison of the top quark mass measurement of this thesis with various other measurements of the top quark mass [93, 94].

Acknowledgements

There are many people who accompanied me on my way of crafting this thesis and I'd like to thank for doing that. First there is of course my supervisor, PD Dr. Reiner Klingenberg. He provided many valuable insights into the world of particle physics and always managed to find time for the concerns of not only me but also the rest of the chair. Another person who added a lot of knowledge to this analysis is Prof. Dr. I. Brock. Without him and his group the lepton transverse momentum might have never made it into this thesis. I also want to thank him, Dr. Ulf Berges and Prof. Dr. J. Stolze for agreeing to participate in the defence examination.

I also want to thank the head of the chair of Experimentelle Physik IV in Dortmund, Prof. Dr. Claus Gößling for giving me the opportunity to do my Ph.D. work. He did let me spend most of my Ph.D. time at CERN which is a great place to work on this kind of analysis.

Then there are the people who I spent most of my time with while being at CERN. Thanks Florian, Kerim, Michele, Kate, Muhammad, Paul, Sascha, Kathrin and 肖肖. You made the long working days in the office much more enjoyable then they would have been without you. In particular I want to thank 肖肖 for all the physics and non-physics related discussions about everything, from b-tagging calibration to the historical aspects of the current diplomatic relations between China and Japan. And of course that the way is the goal.

Also in Dortmund at the chair of E4 there are so many people that I'll probably forget half of them. But I'll try anyways. In no particular order there are: Markus, Michael, Thomas, Ingo, Ingo, Mona, Jan, Silke, Silke, Julia, Andrea, Tobias, Tobias, Tobias, Bane, Jenny, Dominik, Christian, Till, Robert, Andreas and Bettina. In particular order I'd like to thank Hendrik for making such a good team with me and Joerg for teaching me the ways of analysis in ATLAS. Also Oli taught me so many things that he deserves a special thanks.

More people here in Dortmund are my friends from the group of Dorph[1] as well as some other friends, especially Phillip (They will never figure this one out), Simon and Gregor (この時を！). Thank you all for all the good times we spent together.

I especially want to thank my family for all the patience and support they gave me.

There are some random people I want to thank: I would like to thank Martin Luther for making my existence possible. I would like to thank ZP Theart, Herman Li, Sam Totman, Frédéric Leclercq, Dave Mackintosh and Vadim Pruzhanov for being awesome and Tim Schafer for introducing them to me. Finally, there is one thing left to say: This document may contain nuts[2].

[1] DORtmunder PHysiker, yes, very creative, I know

[2] T. Pratchett, I. Stewart, J. Cohen, *The Science of Discworld II: The Globe*, Elbury (2002), ISBN-10: 0091882737

Bibliography

[1] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588.

[2] S. Weinberg, *A Model of Leptons*, Phys .Rev. Lett. **19** (1967) 1264–1266.

[3] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168–171.

[4] B. R. Watson PE, Watson ID, *Total body water volumes for adult males and females estimated from simple anthropometric measurements.*, Am J Clin Nutr. **Jan; 33(1)** (1980) 27–39.

[5] J. Beringer et al., *Review of Particle Physics*, Phys. Rev. D **86** (2012) . `http://pdg.lbl.gov`.

[6] E. D. Bloom, D. H. Coward, H. DeStaebler, J. Drees, G. Miller, L. W. Mo, R. E. Taylor, M. Breidenbach, J. I. Friedman, G. C. Hartmann, and H. W. Kendall, *High-Energy Inelastic $e-p$ Scattering at 6° and 10°*, Phys. Rev. Lett. **23** (Oct, 1969) 930–934. `http://link.aps.org/doi/10.1103/PhysRevLett.23.930`.

[7] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. DeStaebler, J. Drees, L. W. Mo, and R. E. Taylor, *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, Phys. Rev. Lett. **23** (Oct, 1969) 935–939. `http://link.aps.org/doi/10.1103/PhysRevLett.23.935`.

[8] M. Gell-Mann, *A schematic model of baryons and mesons*, Physics Letters **8** (1964) no. 3, 214 – 215. `http://www.sciencedirect.com/science/article/pii/S0031916364920013`.

[9] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Part I*, .

[10] K. G. Wilson, *Confinement of Quarks*, Phys. Rev. **D10** (1974) 2445–2459.

[11] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10, 12** (1963) 531–533.

[12] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49 No. 2** (1973) 652–657.

[13] D0 Collaboration Collaboration, S. Abachi et al., *Search for High Mass Top Quark Production in $p\overline{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **74** (Mar, 1995) 2422–2426. `http://link.aps.org/doi/10.1103/PhysRevLett.74.2422`.

[14] CDF Collaboration, F. Abe et al., *Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (Apr, 1995) 2626–2631. `http://link.aps.org/doi/10.1103/PhysRevLett.74.2626`.

[15] L. M. Brown, *The idea of the neutrino*, Phys. Today **31N9** (1978) 23–28.

[16] M. Gonzalez-Garcia and M. Maltoni, *Phenomenology with massive neutrinos*, Physics Reports **460** (2008) 1 – 129. http://www.sciencedirect.com/science/article/pii/S0370157308000148.

[17] J. C. Maxwell, *A Dynamical Theory of the Electromagnetic Field*, Philosophical Transactions of the Royal Society of London **155** (1865) 459–512, http://rstl.royalsocietypublishing.org/content/155/459.full.pdf+html. http://rstl.royalsocietypublishing.org/content/155/459.short.

[18] R. Feynman and M. Gell-Mann, *Theory of Fermi interaction*, Phys. Rev. **109** (1958) 193–198.

[19] CKMfitter Group Collaboration, J. Charles et al., *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur.Phys.J. **C41** (2005) 1–131, arXiv:hep-ph/0406184 [hep-ph].

[20] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Progress of Theoretical Physics **28** (Nov., 1962) 870–880.

[21] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, Soviet Journal of Experimental and Theoretical Physics **26** (May, 1968) 984.

[22] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323.

[23] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132–133.

[24] G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012) no. 1, 1 – 29. http://www.sciencedirect.com/science/article/pii/S037026931200857X.

[25] S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) no. 1, 30 – 61. http://www.sciencedirect.com/science/article/pii/S0370269312008581.

[26] A. Quadt, *Top quark physics at hadron colliders*, Eur. Phys. J. **C48** (2006) 835–1000.

[27] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, et al., *The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC*, Eur. Phys. J. **C72** (2012) 2205, arXiv:1209.2716 [hep-ph].

[28] C. Lefevre, *The CERN accelerator complex. Complexe des accelerateurs du CERN (2006)*, . http://cdsweb.cern.ch/record/979035/?ln=en.

[29] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73** (2013) 2518, arXiv:1302.4393 [hep-ex].

[30] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[31] ATLAS Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment — Detector, Trigger and Physics*, arXiv.org:0901.0512. CERN-OPEN-2008-020.

[32] M. Hilbert and P. Lopez, *The World's Technological Capacity to Store, Communicate, and Compute Information*, Science **332** (2011) no. 6025, 60–65, http://www.sciencemag.org/content/332/6025/60.full.pdf. http://www.sciencemag.org/content/332/6025/60.abstract.

[33] I. Antcheva et al., {*ROOT*} *A C++ framework for petabyte data storage, statistical analysis and visualization*, Computer Physics Communications **180** (2009) no. 12, 2499 – 2512. http://www.sciencedirect.com/science/article/pii/S0010465509002550.

[34] S. Agostinelli et al., *Geant4 a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) no. 3, 250 – 303. http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[35] ATLAS Collaboration, G. Aad et al., *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, New J. Phys. **13** (2011) 053033, arXiv:1012.5104 [hep-ex].

[36] *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, Tech. Rep. ATLAS-CONF-2012-042, CERN, Geneva, Mar, 2012.

[37] ATLAS Collaboration, G. Aad et al., *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. **C72** (2012) 1909, arXiv:1110.3174 [hep-ex].

[38] *Measurements of the photon identification efficiency with the ATLAS detector using 4.9 fb-1 of pp collision data collected in 2011*, Tech. Rep. ATLAS-CONF-2012-123, CERN, Geneva, Aug, 2012.

[39] ATLAS Collaboration, G. Aad et al., *Combined Muon Performance Public Results*, Https://twiki.cern.ch/twiki/bin/view/atlaspublic/muonperformancepublicplots#anchorplots2011. Accessed: 2012-07-31.

[40] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ph].

[41] *Jet energy scale and its systematic uncertainty in proton-proton collisions at sqrt(s)=7 TeV with ATLAS 2011 data*, Tech. Rep. ATLAS-CONF-2013-004, CERN, Geneva, Jan, 2013.

[42] ATLAS Collaboration, G. Aad et al., *Jet energy resolution in proton-proton collisions at* $\sqrt{s} = 7$ *TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73** (2013) 2306, arXiv:1210.6210 [hep-ex].

[43] *Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2012-101, CERN, Geneva, Jul, 2012.

[44] *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011.

[45] *Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb-1 of Data from the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-043, CERN, Geneva, Mar, 2012.

[46] *Calibrating the b-Tag Efficiency and Mistag Rate in 35 pb-1 of Data with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2011-089, CERN, Geneva, Jun, 2011.

[47] *Measuring the b-tag efficiency in a top-pair sample with 4.7 fb-1 of data from the ATLAS detector*, Tech. Rep. ATLAS-CONF-2012-097, CERN, Geneva, Jul, 2012.

[48] *b-jet tagging calibration on c-jets containing $D*+$ mesons*, Tech. Rep. ATLAS-CONF-2012-039, CERN, Geneva, Mar, 2012.

[49] *Measurement of the Mistag Rate with 5 fb−1 of Data Collected by the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-040, CERN, Geneva, Mar, 2012.

[50] T. Gleisberg, S. Hoeche, F. Krauss, A. Schalicke, S. Schumann, et al., *SHERPA 1. alpha: A Proof of concept version*, JHEP **0402** (2004) 056, arXiv:hep-ph/0311263 [hep-ph].

[51] P. Nason and B. Webber, *Next-to-Leading-Order Event Generators*, Ann. Rev. Nucl. Part. Sci. **62** (2012) 187–213, arXiv:1202.1251 [hep-ph].

[52] M. Gluck, E. Reya, and A. Vogt, *Dynamical parton distributions revisited*, Eur. Phys. J. **C5** (1998) 461–470, arXiv:hep-ph/9806404 [hep-ph].

[53] H1 Collaboration Collaboration, S. Aid et al., *The Gluon density of the proton at low x from a QCD analysis of F2*, Phys. Lett. **B354** (1995) 494–505, arXiv:hep-ex/9506001 [hep-ex].

[54] ZEUS Collaboration Collaboration, M. Derrick et al., *Measurement of the F2 structure function in deep inelastic e+ p scattering using 1994 data from the ZEUS detector at HERA*, Z. Phys. **C72** (1996) 399–424, arXiv:hep-ex/9607002 [hep-ex].

[55] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[56] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[57] A. Martin, R. Roberts, W. Stirling, and R. Thorne, *Uncertainties of predictions from parton distributions. 1: Experimental errors*, Eur. Phys. J. **C28** (2003) 455–473, arXiv:hep-ph/0211080 [hep-ph].

[58] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, Comput. Phys. Commun. **184** (2013) 919–985, arXiv:hep-ph/0405247 [hep-ph].

[59] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **0307** (2003) 001, arXiv:hep-ph/0206293 [hep-ph].

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

[61] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **0711** (2007) 070, arXiv:0709.2092 [hep-ph].

[62] E. Boos, M. Dobbs, W. Giele, I. Hinchliffe, J. Huston, et al., *Generic user process interface for event generators*, arXiv:hep-ph/0109068 [hep-ph].

[63] G. M. S. M. K. O. P. R. M. S. G. Corcella, I.G. Knowles and B. Webber, *HERWIG 6.5*, JHEP **0101** (2001) 010.

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

[65] CMS Collaboration Collaboration, S. Chatrchyan et al., *Measurement of the top-quark mass in $t\bar{t}$ events with dilepton final states in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C72** (2012) 2202, arXiv:1209.2393 [hep-ex].

[66] CDF Collaboration, D0 Collaboration Collaboration, T. E. W. Group, *2012 Update of the Combination of CDF and D0 Results for the Mass of the W Boson*, arXiv:1204.0042 [hep-ex].

[67] CDF Collaboration, T. Aaltonen et al., *Precision Top-Quark Mass Measurements at CDF*, Phys. Rev. Lett. **109** (2012) 152003, arXiv:1207.6758 [hep-ex].

[68] *Measurement of the Top Quark Mass from $\sqrt{s} = 7$ TeV ATLAS Data using a 3-dimensional Template Fit*, Tech. Rep. ATLAS-CONF-2013-046, CERN, Geneva, May, 2013.

[69] D0 Collaboration, V. M. Abazov et al., *Precise measurement of the top-quark mass from lepton+jets events at D0*, Phys. Rev. **D84** (2011) 032004, arXiv:1105.6287 [hep-ex].

[70] J. Walbersloh, *Investigations on the decay length method for the measurement of the top quark mass at the atlas experiment*. PhD thesis, TU Dortmund University, Germany, August, 2009. https://hdl.handle.net/2003/26461.

[71] C. A. Jung, *Weiterentwicklung der Zerfallslaengenmethode zur Messung der Masse des Top-Quarks*, April, 2010.

[72] H. Esch, *Systematische Studien zur Gluonabstrahlung in der Top-Antitop-Paarproduktion am ATLAS-Experiment Untersuchungen im Rahmen der Bestimmung des Wirkungsquerschnittes und der Masse des Top-Quarks*, Master's thesis, TU Dortmund University, Germany.

[73] ATLAS Collaboration, G. Aad et al., *Measurement of $t\bar{t}$ production with a veto on additional central jet activity in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector*, Eur. Phys. J. **C72** (2012) 2043, arXiv:1203.5015 [hep-ex].

[74] J. Stillings, *Top Quark Mass Measurement using Lepton Transverse Momenta at ATLAS*, Master's thesis, Bonn University, Germany, November, 2009.

[75] CDF, *Measurement of the top quark mass from the lepton transverse momentum in the tt to dilepton chanel at the Tevatron*, tech. rep., 2007.

[76] D. Dereniowski and M. Kubale, *Cholesky Factorization of Matrices in Parallel and Ranking of Graphs*. 2004.

[77] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, .

[78] *pyminuit: Minuit numerical function minimization in Python*, Https://code.google.com/p/pyminuit2/, 2007. `https://code.google.com/p/pyminuit2/`.

[79] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. **D81** (2010) 054028, `arXiv:1001.5034 [hep-ph]`.

[80] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys. Rev. **D82** (2010) 054018, `arXiv:1005.4451 [hep-ph]`.

[81] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. **D83** (2011) 091503, `arXiv:1103.2792 [hep-ph]`.

[82] ATLAS Collaboration, G. Aad et al., *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at* $\sqrt{s} = 7$ *TeV*, Eur. Phys. J. **C71** (2011) 1577, `arXiv:1012.1792 [hep-ex]`.

[83] P. Baernreuther, M. Czakon, and A. Mitov, *Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to* $q\bar{q} \to t\bar{t} + X$, Phys. Rev. Lett. **109** (2012) 132001, `arXiv:1204.5201 [hep-ph]`.

[84] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, `arXiv:1112.5675 [hep-ph]`.

[85] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, JHEP **1301** (2013) 080, `arXiv:1210.6832 [hep-ph]`.

[86] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **1212** (2012) 054, `arXiv:1207.0236 [hep-ph]`.

[87] M. Czakon, P. Fiedler, and A. Mitov, *The total top quark pair production cross-section at hadron colliders through* $O(alpha_S^4)$, `arXiv:1303.6254 [hep-ph]`.

[88] R. J. Barlow, *Application of the Bootstrap resampling technique to Particle Physics experiments*, MAN/HEP **4** (1999) .

[89] ATLAS Collaboration 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*, .

[90] A. Sherstnev and R. S. Thorne, *Parton Distributions for LO Generators*, Eur. Phys. J. **C55** (2008) 553–575.

[91] S. Forte, L. Garrido, J. I. Latorre, and A. Piccione, *Neural network parametrization of deep inelastic structure functions*, JHEP **0205** (2002) 062, `arXiv:hep-ph/0204232 [hep-ph]`.

[92] F. Aaron et al., *Combined measurement and QCD analysis of the inclusive e-p scattering cross sections at HERA*, Journal of High Energy Physics **2010** (2010) no. 1, 1–63. `http://dx.doi.org/10.1007/JHEP01%282010%29109`.

[93] CMS Collaboration Collaboration, *Top quark mass combinations using the 2010, 2011 and 2012 data*, Tech. Rep. CMS-PAS-TOP-13-002, CERN, Geneva, 2013.

[94] Tevatron Electroweak Working Group, CDF Collaboration, D0 Collaboration Collaboration, M. Muether and CDF, *Combination of CDF and DO results on the mass of the top quark using up to 8.7 fb-1 at the Tevatron*, `arXiv:1305.3929 [hep-ex]`.

[95] M. Haugh, *The Monte Carlo Framework, Examples from Financa and Generating Correlated Random Variables*, IEOR **E4703** (2004) .

[96] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, third ed., 1999.

[97] E. Jones, T. Oliphant, P. Peterson, et al., *SciPy: Open source scientific tools for Python*, Http://www.scipy.org/, 2001. `http://www.scipy.org/`.

Appendix A

Mathematical Appendix

A.1 The Cholesky Decomposition

The Cholesky Decomposition is a procedure to decompose a hermitian matrix M into a lower and an upper triangular matrix:

$$M = LU \tag{A.1}$$

Special about is, that the upper triangular matrix is the transposition of the lower triangular matrix:

$$L = U^T \tag{A.2}$$

The important feature of this analysis is that if M is the covariance matrix of a system with correlated variables, a vector of uncorrelated variables $\vec{u}$ can be transformed to a vector with the correlation of the system $\vec{c}$ simply by multiplying it with L [95]:

$$\vec{c} = L\vec{u} \tag{A.3}$$

The decomposition works in the following way: Let m_{ij} and l_{ij} denote the element in the i-th row and the j-th column of M and L. One starts with $j = 1$ and calculates

$$l_{jj} = \sqrt{m_{jj} - \sum_{k=1}^{j-1} l_{jk}^2}. \tag{A.4}$$

For $j = 1$ this is simply $\sqrt{m_{jj}}$. This is repeated for $j = 2$. In the next step, for all $j > i$, the first off-diagonal element is calculated:

$$l_{ij} = \frac{1}{l_{jj}} \left(m_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk} \right). \tag{A.5}$$

For $j < i$, l_{ij} is 0 as the algorithm calculates a triangular matrix. j is increased again and Equations (A.4) and (A.5) are applied until all elements of L have been computed. If the matrix is not hermitian, the square root in Equation (A.4) will be negative or the division in (A.5) will be a division by zero. The algorithm can therefore also be used to check if a matrix is hermitian.

An example for the algorithm: For

$$M = \begin{pmatrix} 2 & 1 \\ 2 & 4 \end{pmatrix}, \tag{A.6}$$

L is:

$$L = \begin{pmatrix} \sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{3.5} \end{pmatrix}. \tag{A.7}$$

LL^T is equal to M which can be verified easily.

The implementation used is from LAPACK[96], interfaced through the python extension scipy/numpy [97].

A.2 Fermi- and Gamma-Distributions

A.2.1 The Fermi Distribution

The fermi distribution originates in statistical mechanics. It describes the average occupation number n of a system of fermions with temperature T:

$$n(E) = \frac{1}{\exp\left(\frac{E-\mu}{kT}\right) + 1} \tag{A.8}$$

k is Boltzmann's konstant and μ the chemical potential of the system. The distribution depends on E, the energy of the state in question. μ and T are free parameters of the system. The shape of the distribution mainly depends of the ratio of μ and T. Figure A.1 shows several fermi distributions for different values of that ratio.

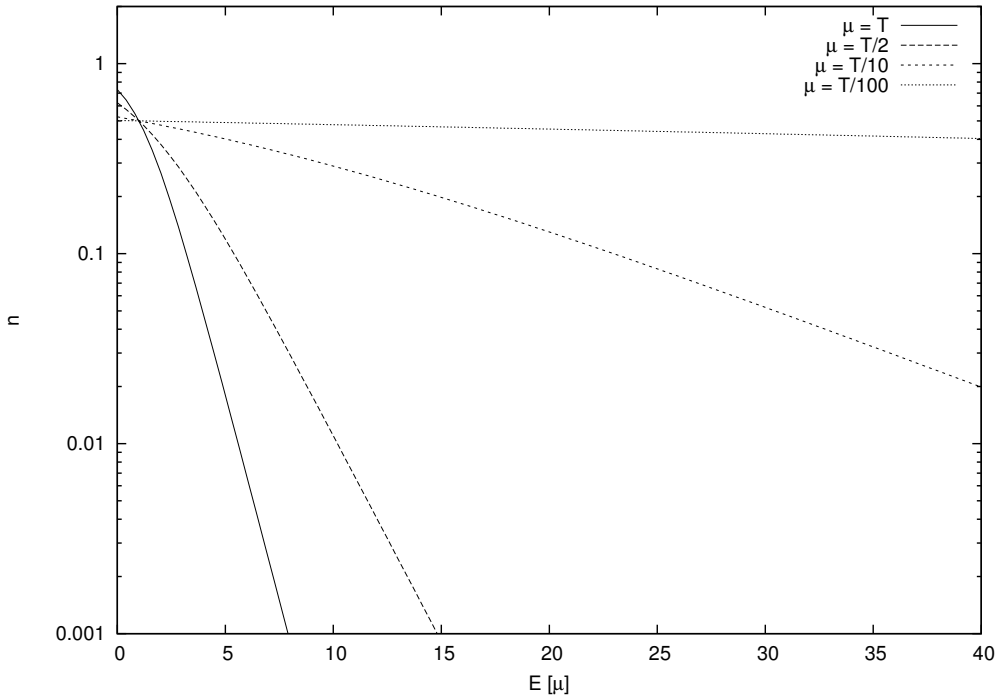

Figure A.1: Example for several fermi distributions for different ratios of μ and T.

A.2.2 The Gamma Distribution

The gamma distribution is a mathematical function:

$$G(x) = \frac{\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} \exp\left(-\frac{x-\gamma}{\beta}\right)}{\beta\Gamma(\alpha)} \tag{A.9}$$

α, β and γ are free parameters. The classic version of the gamma distribution does not include the parameter γ, it is a constant offset to x. Figure A.2 shows different versions of the gamma function for various values of the three free parameters.

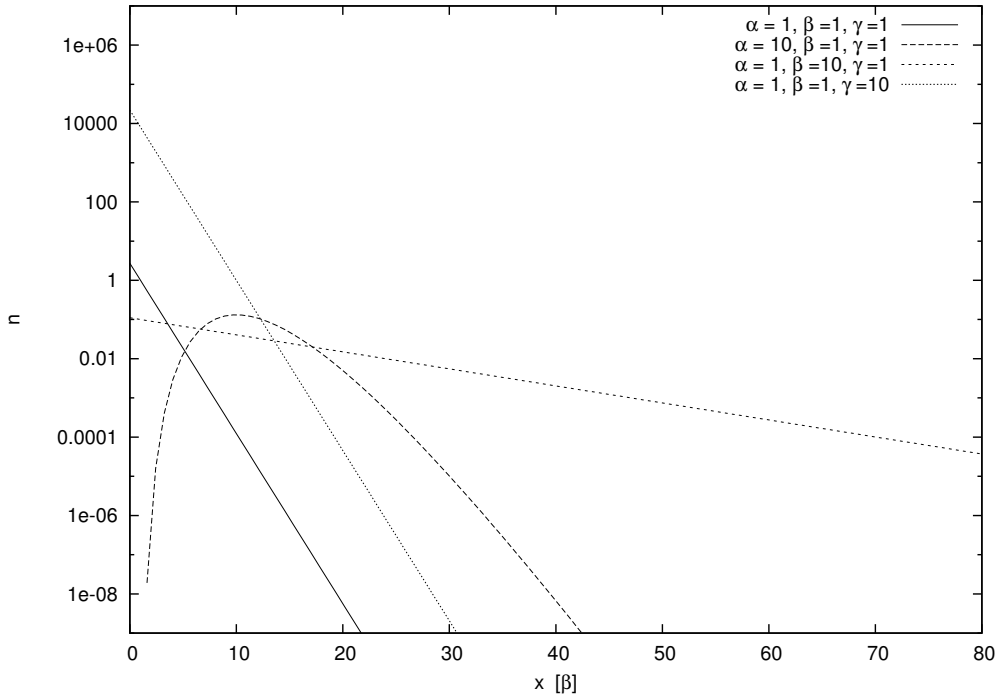

Figure A.2: Example for several fermi distributions for different values of α, β and γ.